UNIVERSITY COLLEGE LONDON

FACULTY OF MATHEMATICAL & PHYSICAL SCIENCES

Department of Mathematics

THE EDGE-REINFORCED BRANCHING RANDOM WALK ON THE TRIANGLE AND GENERALISED BALLS AND BINS WITH POSITIVE FEEDBACK

GIORDANO GIAMBARTOLOMEI

Supervisor:

Prof. Nadia Sidorova

Thesis submitted for the degree of

Doctor of Philosophy

Academic Year 2022

Declaration

I, Giordano Giambartolomei, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Giordano Giambartolomei

Abstract

Edge-reinforced random walks are processes with reinforcement, on which the effect of branching has not been investigated. Our discrete time model starts with a particle, which branches at a constant rate, at the vertex of a triangle initialised with edge crossing numbers. The offspring particles, independently of each other, traverse the incident edges at random, with probabilities proportional to the edge crossing numbers, correspondingly updated at each traversal. Then the process repeats. We show the convergence of the proportions of edge crossings to a random variable using dynamical systems techniques, and prove that two events have positive probability: when none of the edges is crossed negligibly, and when exactly one is. We show that all edges are crossed infinitely many times and conjecture that no two edges can be negligibly crossed.
This conjecture stems from connections between this model and balls and bins, where balls are added to bins at random, following certain rules. There is positive feedback when the probability of incoming balls choosing a bin with m balls is proportional to a power of m, bigger than 1; no feedback when the power is 1. In a time-dependent version, the number of balls added at discrete times varies, yielding different regimes of growth. Generalising results known for two bins to any number of bins, we investigate the proportion of balls in each bin, depending on feedback and regime of growth. We focus on the events of monopoly (eventually one of the bins will receive all incoming balls) and dominance (one of the bins gets all but a negligible number of balls). When there is no feedback, neither monopoly nor dominance occur. When feedback is introduced, several regimes are identified, at which dominance and monopoly occur. While at certain regimes monopoly does not occur, we conjecture dominance to always occur.

Impact Statement

Edge-reinforced random walks on graphs are well studied processes with reinforcement, but the branching of the particles performing the walk has not yet been investigated. The introduction of this model in the first part of the present work, opens up a fresh line of investigation not only interesting for its sheer originality; but because technically, it pushes the boundaries of how dynamical systems techniques are implemented into probabilistic work. For example, stochastic approximation is well known to be successful in unravelling the asymptotics of several discrete processes with reinforcement, by borrowing from continuous dynamical systems techniques that are coupled with martingales. However, in a branching random walk, the regime of growth of the particles makes stochastic approximation and other established approaches not applicable. In this study a new dynamical approach is developed, which does not dodge the discrete setting of the problem, and takes full advantage of martingale theory and of the particles' branching rate. This line of investigation can grow in many ways, from considering other graphs than the triangle, like done for standard edge-reinforced random walks; to adding a feedback to the model, like done in the balls and bins model, generalised and studied in the second part of the work.

In terms of potential applications, besides the theoretical development of a new stochastic-analytic dynamical approach, it might be worth taking into consideration the time-dependent balls and bins model with positive feedback, generalised to an arbitrary finite number of bins. This model has deep connections with the branching random walk, and some of the martingale techniques that proved successful on the former, helped approaching the latter. Generalising the balls and bins model can find applications in areas that have traditionally benefited from various types of such models, such as computer science and economics. In these disciplines the probabilistic understanding of network evolution is crucial. Generalising balls and bins and studying branching random walks helps the development of a toolbox useful to characterise probabilistically the onset of preferential attachment in the evolving dynamics on networks.

Aknowledgements

I would like to express my gratitude to my supervisor, Nadia Sidorova, who gave me the opportunity of becoming a mathematician by working on a beautiful and original project in probability. If today I am having so much fun doing mathematics, I owe it to the creativity and skills I developed while working on it. This would not have been possible without the expertise and experience she kindly shared over the years.

I am sincerely grateful to my BSc supervisor, Nicola Arcozzi, for his unfailing support in my academic journey. Moreover, I owe to him my first taste of research: ever since, doing mathematics has been the only option that makes sense to me.

It is my pleasure to thank also my secondary supervisor, Leonid Parnovski, for his always timely academic service and support.

I am also thankful to the UCL Department of Mathematics, who generously supported me with adequate equipment, to make sure my research and teaching would not be hindered by the pandemic, and funds, when it was finally possible to travel abroad to conferences again.

I would like to express my wholehearted appreciation to Eva. I am very fortunate to have someone by my side as capable and trustworthy as she is, who selflessly supports me with many daily contingencies.

I conclude by thanking my family back in Italy for their understanding, which enables me to enjoy research to my heart's content.

Contents

1 Introduction 1

1.1 Reinforced random walks . . . 1
1.2 Edge-reinforced branching random walks . . . 5
1.3 Balls and bins with positive feedback . . . 11
1.4 Generalised balls and bins with positive feedback . . . 13
1.5 The connection between ERBRW on graphs and generalised BB . . . 15
1.6 Note for the reader . . . 18
1.7 Notation . . . 18

I Edge-reinforced branching random walk on the triangle 20

2 Introduction 22

2.1 Iterative equations of the model . . . 22
2.2 Reduction to a perturbed dynamical system . . . 27
2.3 Outline of contents . . . 28
2.4 Note for the reader . . . 33

3 The dynamical system with $\rho = 1$ 34

3.1 Preliminaries . . . 34
3.2 Fixed points and potential function . . . 37
3.3 Convergence bounded away from the boundary . . . 40
3.4 Convergence to the boundary with $\ell = 0$ . . . 43
3.5 Convergence to the boundary with $\ell > 0$ . . . 60
3.5.1 Convergence of boundary orbits . . . 60
3.5.2 Structure of the set of accumulation points . . . 65
3.5.3 Convergence of regular orbits . . . 70
3.6 Convergence of the dynamical system . . . 84

4 The ERBRW stochastic process 85

4.1 Preliminaries . . . 85
4.2 Fixed points and potential . . . 87
4.3 Convergence on $\mathscr{B}$ . . . 90
4.4 Convergence on $\mathscr{E}_0$ . . . 93
4.5 Convergence on $\mathscr{E}_>$ . . . 117
4.5.1 Convergence with boundary initial conditions . . . 119
4.5.2 Structure of the set of accumulation points . . . 124
4.5.3 Convergence with regular initial conditions . . . 130
4.6 Convergence of the stochastic process . . . 149
4.7 Typical asymptotic behaviours of the ERBRW . . . 150
4.8 No monopoly . . . 165

4.9 No dominance conjecture . . . 167

II Generalised balls and bins with positive feedback 173

5 Introduction 175
5.1 Iterative equations of the model . . . 175
5.2 A useful heuristics for *dominance* . . . 177
5.3 Outline of contents . . . 178
5.4 Note for the reader . . . 181
5.5 Notation . . . 181

6 No monopoly 182

7 No dominance 185

8 Getting away from the equilibrium 186
8.1 Preliminaries . . . 186
8.2 Infinitesimal deviations from the equilibria in presence of feedback . . . 189

9 Dominance and monopoly 218
9.1 Preliminaries . . . 218
9.2 Dominance in presence of feedback . . . 226
9.3 Monopoly in presence of feedback . . . 239

III Supplements to Part I 240

A The dynamical system with $\rho < 1$ 242
A.1 Preliminaries . . . 242
A.2 Fixed points and potential function . . . 244
A.3 Convergence bounded away from the boundary . . . 247
A.4 Convergence to the boundary with $\ell = 0$ . . . 247
A.5 Convergence to the boundary with $\ell > 0$ . . . 253
A.5.1 Convergence of boundary orbits . . . 253
A.5.2 Structure of the set of accumulation points . . . 256
A.5.3 Convergence of regular orbits . . . 259
A.6 Convergence of the dynamical system . . . 272

B Construction of $\vartheta(\mu)$ and $\theta(\mu)$ 273

IV Supplements to Part II 277

C Supplements to Chapter 6 279

D Supplements to Chapter 7 282

E Supplements to Chapter 8 284

F Supplements to Chapter 9 291

List of Figures

1.1 ERBRW on the triangle . . . 7
1.2 Standard simplex (gray) and its medial triangle (red) . . . 8
1.3 Convergence of Θ_n (green) and π_n (blue) to equilibrium points (Θ, π_Θ) (red) . . . 10
1.4 Convergence of Θ_n (green) to Θ (red) and divergence of π_n (blue) about π_Θ (gray) . . . 10
1.5 Two bins and ERBRWs . . . 16
1.6 Three bins and ERBRWs . . . 16
1.7 Convergence of Θ_n (green) to a vertex v (red) of the simplex (dominance) and divergence of π_n (blue) about π_v (gray) . . . 17

2.1 ERBRW on the triangle with loops . . . 23
2.2 Convergence of p_n (green) and q_n (blue) to internal equilibrium points (p, q_p) (in red) . . . 29
2.3 Convergence of p_n (green) to a boundary limit p (red) and asymptotic oscillation of q_n (blue) about (q_p) (red) along the eigendirection $e_{-1}(p)$ (violet) . . . 31

3.1 $\mathbb{S}_1^2 \cap \Pi_0$ . . . 40
3.2 Internal equilibria . . . 41
3.3 $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$ as the Cartesian product of the dark gray and the red regions . . . 44
3.4 $R_{\varepsilon,\frac{\delta}{8}}$ in green . . . 55
3.5 $\mathcal{K}^{\ell}_{\varepsilon',\delta}$: in gray the coordinates for p, in red the eigencordinates for q, with $q_* = q_{p_*} + \frac{\ell}{2}e_{-1}(p_*)$, $\hat{q}_* = q_{p_*} - \frac{\ell}{2}e_{-1}(p_*)$ . . . 71

5.1 $\psi^{(1)}$ (red) and $\psi^{(1)}_n$ (blue) and cobwebbing . . . 179

9.1 P_{n^*} (red) and S_{n^*} (blue) . . . 225
9.2 Δ_d and Δ^*_{d-1}, where $\hat{\Theta}_{d-2} = (\Theta_1, \dots, \Theta_{d-2})$. . . . 231

B.1 Nonnegligibility of the positive solution to (B.2) and (B.3) . . . 276
B.2 Finding $\theta(\mu)$ given $\rho(\mu) = 1/3$ for $\mu = 3/2$ . . . 276

Chapter 1

Introduction

This first chapter is a general introduction to the two main problems addressed in this work, following the same order in which they will be studied. We start by introducing the reader to the area the first problem belongs to, by reviewing some of the literature on reinforced random walks, and then describe the subject of our study, the *edge-reinforced branching random walk* on the triangle, including the corresponding main results obtained. We next introduce the reader to the area the second problem belongs to, by similarly reviewing first the literature on balls and bins with positive feedback, and subsequently describing the *generalised balls and bins model with positive feedback*, including the main results obtained.

We conclude this general introduction with a description of the connections between the two problems; a note for the reader about how to use the Appendix; a brief commentary on the notation followed.

Each of the two problems will have its own separate technical introduction in the corresponding part of the work dedicated to it, so as to keep this first chapter as free of technicalities as possible, allowing for a more immediate understanding.

1.1 Reinforced random walks

Consider a connected locally finite graph and fix a vertex as starting point at time zero. Select a neighbour to move to, uniformly at random. At time one move to the neighbour selected. At each discrete time n, repeat the selection process, from the currently occupied vertex, and move to it at time $n+1$. The sequence of points thus visited is a *random walk* on the graph.

Random walks on graphs have been intensively studied for more than a century, as they arise in many models in mathematics and physics, on top of having important algorithmic applications. The properties of the graph affect the behaviour of the random walk and can be a crucial factor on the types of investigations that can succeed. The classical theory of random walks typically focused on simple (no multiple edges, no loops) but infinite graphs, like lattices in various dimensions, and the type of investigation was qualitative: for example whether the random walk returns to its starting point almost surely, whether it returns infinitely often, its limiting distribution etc. In more recent work, questions of more quantitative character have also been studied, even on more general, but finite graphs: how long it takes to return to the starting point, how long it takes to visit all vertices, how fast the random walker's distribution tends to a limit distribution etc.

Random walks on graphs are Markov chains, and Markov chains are random walks on directed weighted graphs. Therefore the study of random walks on graphs does not differ much from the study of Markov chains. For simplicity let us consider only random walks on simple and undirected graphs (some of the results we will talk about do extend to cases when the graph is not simple: sometimes parallel edges are allowed, sometimes even loops; some results also apply to directed graphs). In this case, the neighbour is selected uniformly at random in the following sense: if the graph is unweighted, this means that each neighbour has a probability to be selected equal to the reciprocal of the degree of the current vertex; if the graph is weighted, each neighbour has a probability to be selected equal to the weight of the edge the random walker will have to traverse to reach it, divided by the total weight of all edges incident to the currently occupied vertex. The former is clearly a particular case of the latter, if we choose all weights equal to 1, hence all graphs considered will be weighted (clearly, if parallel edges are allowed, one can go in the opposite direction too, since they can be used to simulate weights in unweighted graphs).

In 1986 P. Diaconis introduced *reinforcement* in this model. To be more specific, *linear edge-reinforcement*: given a weighted graph (recall that in this literature review we assume all graphs to be connected, locally finite and simple), a starting point, and the usual scheme of random selection of a neighbour, it also happens that upon each traversal of the edge leading to the selected neighbour, the weight of the edge increases by 1. This means that in the future, it will be more likely that the traversed edge is going to be crossed than it was in the past. The random walker is *self-interacting*: it remembers where it has been, and prefers to cross familiar edges. Thus the new process, in general, is not Markovian anymore, as it depends on the whole history. In [9, 13] this model was studied on a finite graph, with all edges' initial weights set to 1 (*initially fair*): almost sure recurrence (the random walker returns to the stating vertex infinitely often: note that by a standard Borel-Cantelli argument, this is equivalent to visiting each vertex infinitely often) and the convergence of the normalised edge occupation vector to a random limit, having density continuous with respect to the Lebesgue measure on the simplex, were shown. The formula of the density was also derived explicitly, but this remained unpublished, until about 10 years later, when M. Keane and S. Rolles rederived it in [20]. As to infinite graphs, P. Diaconis posed the question, of whether on $\mathbb{Z}^d$ the linear edge-reinforced random walk (LERRW) almost surely returns to the starting point infinitely often (by random walk on $\mathbb{Z}^d$ it is meant a random walk on the d-dimensional lattice graph).

In [35] R. Pemantle, under P. Diaconis' supervision, introduced the vertex-reinforced random walk on a graph, where the weights accumulate on the vertices rather than on the edges. In the most simple version of this model, instead of a standard weighted graph, we consider a vertex-weighted graph and pick a starting point: since the weights are not on the edges, but on the vertices, the reinforcement scheme will be to increase the weight of a vertex every time it gets occupied by the random walker. From the current vertex, the random walker selects the next vertex at random among its neighbours. Each neighbour has a probability to be selected equal to its own weight divided by the total weight of the neighbours.

In more general reinforced random walks the new weight may depend on many factors: the linear scheme (with respect to the number of edge crossings in the case of edge-reinforcement, with respect to the number of visits to the vertices in the case of vertex-reinforcement) of adding 1 (or any fixed constant) when updating the weights is not the only one that has been studied. Several generalisations of P. Diaconis' model exploiting different reinforcement schemes received attention over the years: for example in [11] the ERRW of *matrix type* is introduced, where the constant added to the current weight of the traversed edge depends on the edge and how many times that edge has already been traversed; similar generalisations have been studied for vertex-reinforced random walks (see [38]). The field of reinforced processes overall has grown significantly since 1986 (see [39] and [28] for a general survey), but in this review we will focus solely on ERRWs.

There is no question that ‘random walks on graphs is one of those notions that tend to pop up everywhere once you start looking for them’ [26]. Even while shuffling a deck of cards, for instance, you can construct a graph whose vertices are all permutations of the deck, and any two of them are adjacent if they come by one shuffle move. Then repeated shuffle moves correspond to a random walk on this graph (see [12]). Additionally with ERRWs on graphs we can model self-interaction when exploring an unknown environment: for instance in [33] LERRWs have been taken as a simple model for the motion of myxobacteria; these bacteria produce a slime trail and prefer to glide on the slime produced earlier.

Linear reinforcement is special, since it produces a *partially exchangeable* process: if two finite paths are such that every edge in each of them is crossed the same number of times, then they have the same probability to be the beginning of a LERRW. The order in which the edges are visited does not matter. In [42, 29] it was proved, thanks to partial exchangeability, that a LERRW is a mixture of Markov chains both on finite and infinite graphs respectively. What this means, in layman’s terms, is that there is a measure in the space of Markov chains such that our process first picks a Markov chain using this measure (called the *mixing measure*) and then does the random walk as per the chain picked. With more recent terminology, we would say that it is a *random walk in a random environment*. Partial exchangeability, being a mixture of Markov chains and recurrence are notions interconnected to each other.

The question P. Diaconis posed on $\mathbb{Z}^d$ provoked a substantial amount of study on almost sure recurrence/transience criteria (by transience it is meant that the random walker returns to the starting point finitely often almost surely) for LERRWs on infinite graphs. The first of such results was in [36], where a phase transition between almost sure recurrence and almost sure transience was identified on infinite binary trees, thanks to the construction of a random environment of Pólya urns for the LERRW: consider the reinforcement parameter being an arbitrary constant $c > 1$ instead of 1; it was shown that there is a constant $c_0 \approx 4.29$ such that if $c < c_0$ there is almost sure transience, if $c > c_0$ there is almost sure recurrence. While on acyclic graphs (such as trees) recurrence is easier to understand, for graphs with cycles such as $\mathbb{Z}^d$ for $d \geq 2$ getting results took longer: a first relatively general one arrived with [43], where it was shown that if the initial weights are sufficiently large (the reinforcement parameter is, instead, kept to 1) then for any finite tree G, the LERRW on $\mathbb{Z} \times G$ is almost surely recurrent. Eventually in [44] recurrence would be successfully shown in any dimension d for $\mathbb{Z}^d$ with small enough initial weights (thus following for any graph having bounded degrees as well). Notably this proof relied on a connection with certain quantum models, called the hyperbolic σ-model, that allowed to exploit partial exchangeability to its full potential. Another proof appeared in [1] immediately after, which does not rely on the quantum model nor on explicit calculation (see [21] for an account including a light touch introduction to the quantum model used).

Since in LERRWs the transition probabilities are proportional to the weights, and we can thus expect the edge crossings to grow proportionally with time, with some dependence on the initial weights; we would also expect that, with stronger reinforcement, *preferential attachment* arises, meaning that some (random) edge will be disproportionally visited as time passes: at a large time with probability close to 1 the walk visits all edges but one a very small number of times. *Superlinear* reinforcement schemes (also called *strong reinforcement*) yield this type of asymptotic behaviour indeed: they are defined such that the transition probabilities are proportional to a function f of the edge crossings, defined on the positive integers and taking values on the positive reals. The main line of investigation is, in this case, finding the conditions for which such a reinforcement gets the random walk stuck at some edge. In [11] it was shown that if the function f grows fast enough, such that

$$\sum_{n=1}^{\infty} \frac{1}{f(n)} < \infty, \qquad \text{H}$$

then the walk on $\mathbb{Z}$ almost surely gets eventually trapped in a single edge; if the series diverges, it is recurrent almost surely (it visits all vertices infinitely often). In the argument the lack of exchangeability due to a stronger reinforcement scheme imposes a change of toolbox: no more random environments, but martingale techniques combined with Rubin's construction, which consists of an exponential embedding of the process which we will describe more in detail in Section 1.3. We will call strong edge-reinforced random walk (SERRW) those for which Condition H is satisfied. When performing the walk on $\mathbb{Z}^d$, with $d \geq 2$, the analysis gets more difficult. The bipartite nature of the d-dimensional lattice requires considering $\sum_{n=1}^{\infty} 1/f(2n)$ and $\sum_{n=1}^{\infty} 1/f(2n+1)$ separately; in [46] it was first shown that if both sums converge (thus Condition H holds), on $\mathbb{Z}^d$ the process is almost surely trapped on a single edge. Although the argument extends to any bipartite graph of bounded degree, it would not work even on a single triangle, thus Sellke states as a conjecture that Condition H ensures that the process is almost surely trapped on a single (random) edge on the triangle. Moreover, it was not possible to prove that if both series diverge (and thus Condition H does not hold), recurrence would follow on $\mathbb{Z}^d$: it was only possible to prove that almost surely the range was infinite and each coordinate, separately, would vanish infinitely often. In [23] the first significant progress towards settling Sellke's conjecture was made, by showing that if we restrict $f(n) = n^\alpha$ where $\alpha > 1$, then for any graph of bounded degree the SERRW is almost surely trapped on a single edge. Here the fact that the argument used in [46] generalises to any graph of bounded degree without odd cycles was used. More in general, through the adaptation of Rubin's construction used in [11, 46], it was possible to prove, via graph-based techniques and martingale arguments, that the SERRW on any graph of bounded degree is either almost surely trapped on a single edge or on an odd cycle. In light of this, the focus of Sellke's conjecture on the triangle was crucial: solving the problem on the triangle, and more in general on odd cycles, would yield the more general claim that if Condition H holds, on any graph of bounded degree the SERRW is almost surely trapped on a single edge. Hence we will refer to this claim as Sellke's conjecture instead. In [24] martingales techniques combined with stochastic approximations techniques yielded the sought result on odd cycles for any nondecreasing weight function f satisfying Condition H. Due to the case analysis required, depending on whether the graph contains odd cycles or not [25], this is essentially as close as it got to a full proof of Sellke's conjecture for the following decade: with a nondecreasing (the actual general condition is more technical and we omit it) weight function f satisfying Condition H, the SERRW is almost surely trapped on a single edge of any graph of bounded degree. Through a alternative approach, using the order statistics on the number of edges, in [10] Sellke's conjecture was finally shown (with Condition H needing minor refinements). We observe that the argument extends to graphs with loops.

1.2 Edge-reinforced branching random walks

Let us briefly go back to random walks on graphs and superimpose branching to the neighbour selecting scheme: assume that at time zero the random walker spontaneously dies by producing a certain number μ of descendants, which carry on hopping at random, independently of each other, to new sites; once on the new site, each descendant dies, generating other $\mu > 1$ descendants, which will follow the same scheme, iteratively. Such model is known as a *branching random walk* (BRW) on graphs. For simplicity we described only a *pure birth* version of the model, but in BRWs the random walkers can also be subject to spontaneous extinction (death with no descendants); moreover both the reproduction and extinction rate can be dependent on the sites, or be random according to some given distribution (in this case the model is known as BRW *in a random environment*). From a general standpoint a BRW can be seen as a spatial generalisation of the Galton-Watson process, and it has been extensively studied. Among the applications found, as mentioned in [7], modern models of disease propagation incorporate spatial interaction by allowing a pathogen to be passed on only to the neighbours of an infected host [34]; a virus can multiply at a host cell and then infect any of the neighbouring ones at random [45]; the total number of infected cells therefore corresponds to the number of distinct sites visited by a BRW [15]. We will not delve into the vast literature existing on BRWs, as it would be too dispersive for our scope. The basic question that one answers studying branching processes is whether it survives, which means that with positive probability at any time there is someone alive; while we saw that the classical question for random walks is whether the walker returns (with positive probability or, equivalently, with probability one) infinitely many times to some fixed site. For BRWs the first question asks whether there is global survival, that is, with positive probability at any time there is someone alive somewhere; while the second question deals with local survival, that is, with positive probability the process returns infinitely many times to some fixed site. We refer the interested reader to [6] for a survey concerning these aspects.

Conceptually, as a generalisation of random walks on graphs, BRWs are obtained analogously to how we will obtain our model as a generalisation of ERRWs; moreover, BRWs make a compelling case for our study. In fact, while BRWs are a well-established area of research, to the best of our knowledge there has not even been any attempt to similarly generalise reinforced random walks, before the present work. Thus, as in the construction of the BRW, we will consider an ERRW to which we superimpose the branching of the random walkers (for simplicity referred to as *particles* from now on). In other words we consider particles, which split into $\mu > 1$ offspring particles, while performing a random walk (on a graph) with an edge-reinforcement scheme. We call this process an *edge-reinforced branching random walk* (ERBRW) on a graph, and we will focus specifically on the triangle, with linear reinforcement. Since this model evolves with pure birth, we will be concerned only with the study of localisation.

More formally, define the ERBRW on the triangle (as per the scheme in Figure 1.1) in the following way: let $T_n^{(1)}$, $T_n^{(2)}$, $T_n^{(3)}$ denote the number of edge crossings, $N_n^{(1)}$, $N_n^{(2)}$, $N_n^{(3)}$ the number of particles at the vertices, at time n. At time 0, we start with one particle at any of the vertices. This hypothesis is not essential, but we adopt it for simplicity: our analysis would not change if we started from an arbitrary distribution of particles at the vertices, since there will always be a positive probability that these particles gather all at the same vertex after some time. This particle branches with deterministic constant factor $\mu > 1$, and then the μ offspring particles choose, independently of each other, the incident edge to traverse, according to a linear reinforcement scheme. Before describing the scheme in detail, we make some comments on two key assumptions.

- We will allow μ to be a noninteger, in which case the interpretation of the random walkers as particles needs to be replaced by unit and fractional masses (the unit masses are seen as normal particles, the fractional masses as smaller particles). To keep this first description of the model as immediate as possible, we will not treat the details of a nonintegral branching factor, which will be dealt with in the more technical description given in the introduction to Part I. Moreover, in the course of this work we will see that there is no substantial loss of generality in adopting only a particle-like point of view.

- The offspring particles are assumed *not lazy*: they cannot remain at the vertex where they already are. Hence we also exclude the presence of *loops* based at the vertices of the triangle (it is customary, in the reinforced random walks literature, to only consider simple graphs). Simulations suggest that adding loops would not yield a qualitatively different dynamics for an undirected graph. We leave further details on the triangle with loops to the introduction to Part I.

After the initial particle has branched, the offspring particles, under the assumptions aforementioned, will travel, independently of each other, through any of the two incident edges at random, with probability of choosing either one of them proportional to the positive number of edge crossings initially deterministically assigned $T_0^{(1)}$, $T_0^{(2)}$, $T_0^{(3)}$. Once the particles reach the new vertices, the edge crossings of the traversed edges are updated, and this process repeats. Each particle's crossing increases the edge crossings by one. The total number of particles at time $n \in \mathbb{N}_0$ is then

$$N_n^{(1)} + N_n^{(2)} + N_n^{(3)} =: \sigma_n = \mu^n,$$

and if we let the $\tau_0 := T_0^{(1)} + T_0^{(2)} + T_0^{(3)}$, the total number of crossings up to time n is then

$$T_n^{(1)} + T_n^{(2)} + T_n^{(3)} =: \tau_n = \tau_0 + \sum_{i=1}^{n} \sigma_i.$$

For all $i \in \{1, 2, 3\}$, let

$$\Theta_n^{(i)} := \frac{T_n^{(i)}}{\tau_n}$$

and

$$\pi_n^{(i)} := \frac{N_n^{(i)}}{\sigma_n}$$

be the corresponding proportions of edge crossings and particles at the vertices at time n; note that as vectors

$$\Theta_n, \pi_n \in \{(x, y, z) \in [0,1]^3 : \ x + y + z = 1\}.$$

For every $(x, y) \in [0,1]^2 \cap \{0 < x + y \leq 1\}$ let

$$\phi(x, y) := \frac{x^\alpha}{x^\alpha + y^\alpha}.$$

We define the ERBRW on the triangle with and without feedback as follows.

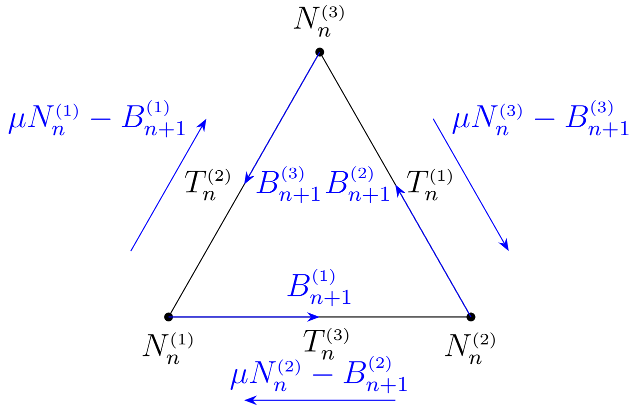


Figure 1.1: ERBRW on the triangle

- ERBRW with *no feedback*: $\alpha = 1$. Let $i \neq j \neq k \in \{1, 2, 3\}$, with k referring to the occupied vertex and i, j referring to the incident edges. Given the past up to time n, after branching, each particle at the kth vertex, independently of the others at random, traverses the ith or jth edge at time $n+1$ independently, with probability proportional to the corresponding number of edge crossings $T_n^{(i)}$, $T_n^{(j)}$, that is with probabilities $\phi(\Theta_n^{(i)}, \Theta_n^{(j)})$ and $\phi(\Theta_n^{(j)}, \Theta_n^{(i)})$ respectively, with $\alpha = 1$. The number of particles traversing one of the incident edges from the kth vertex at time $n+1$, according to the diagram in Figure 1.1, are binomial random variables denoted as $B_{n+1}^{(k)}$, .

- ERBRW *with feedback*: $\alpha > 1$. Using the same notation as in the previous case, given the past up to time n, after branching, each particle at the kth vertex, independently of the others at random, traverses the ith or jth edge at time $n + 1$ independently, with probability proportional to the corresponding number of edge crossings raised to the power of $\alpha > 1$, $(T_n^{(i)})^\alpha$, $(T_n^{(j)})^\alpha$, that is with probabilities $\phi(\Theta_n^{(i)}, \Theta_n^{(j)})$ and $\phi(\Theta_n^{(j)}, \Theta_n^{(i)})$ respectively, with $\alpha > 1$. This corresponds to the branching analogue of the SERRW studied in [23].

The case $\alpha > 1$ will not be directly studied in this dissertation, which is concerned with $\alpha = 1$. The terminology relating to feedback arises from the literature on balls and bins models, which have connections with ERBRWs on graphs. The choice of starting the study of the ERBRW on the triangle graph is motivated by the symmetries of this graph. The equations governing the model inherit this helpful symmetry, along with the benefits of a low number of degrees of freedom.

In Part I we show three main qualitative localisation results for the ERBRW *with no feedback*. Denote, for simplicity of exposition, the standard simplex in three dimensions as

$$\Sigma := \{(x, y, z) \in [0, 1]^3 : x + y + z = 1\}.$$

We will say *almost surely*, meaning *with probability* 1, while talking about events, *negligibility* means *with probability* 0. The first result is the following.

Theorem 1.1. *There is an almost surely Σ-valued bounded random variable Θ such that almost surely $\Theta_n \longrightarrow \Theta$ as $n \longrightarrow \infty$.*

In the context of the ERBRW on the triangle define *dominance* as the event $\mathcal{D}$ in which the edge crossings along two of the edges become negligible as the time n grows. Here *negligible* has the meaning of the corresponding proportions *vanishing* as n grows.

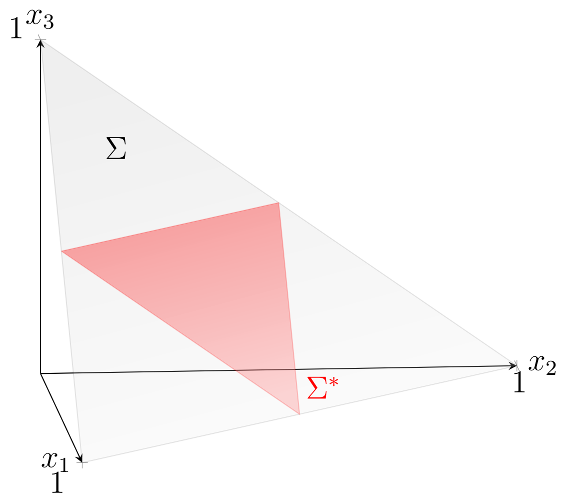


Figure 1.2: Standard simplex (gray) and its medial triangle (red)

Definition 1.2. Dominance *is the event*

$$\mathcal{D} := \left\{ \exists i \in \{1,2,3\} : \lim_{n \longrightarrow \infty} \Theta_n^{(i)} = 1 \right\}.$$

Denote the event on which $\{\Theta_n\}$ approaches an edge of the simplex Σ while being bounded away from the set of its vertices V as $\mathscr{E} := \mathscr{S} \cap \mathcal{D}^c$, where $\mathscr{S} := \{\Theta_n \longrightarrow \partial\Sigma\}$. Having defined

$$\pi_{\Theta_n} := \frac{\mathbf{1} - \Theta_n}{2},$$

we anticipate that, in the dynamical interpretation of the random process $\{(\Theta_n, \pi_n)\}$, which will be explored in Chapter 2, the distance of π_n from π_{Θ_n} is crucial in gauging whether the process is close to equilibrium or not. Denoting as $\|\cdot\|_1$ the ℓ^1 norm on $\mathbb{R}^3$ (it is computed as the sum of the absolute values of the columns, and will be referred to as 1-norm for simplicity); we will show in Lemma 4.5 that there is a nonnegative bounded random variable ℓ such that almost surely

$$\|\pi_n - \pi_{\Theta_n}\|_1 \longrightarrow \ell.$$

Denote

$$\begin{aligned}
\mathscr{B} &:= \{\Theta_n \text{ is bounded away from } \partial\Sigma\}, \\
\mathscr{E}_> &:= \mathscr{E} \cap \{\ell > 0\} = \{\Theta_n \text{ approaches } \partial\Sigma \setminus V\} \cap \{\pi_n \text{ is bounded away from } \pi_{\Theta_n}\}, \\
\mathcal{D}_> &:= \mathcal{D} \cap \{\ell > 0\} = \{\Theta_n \longrightarrow v \in V\} \cap \{\pi_n \text{ is bounded away from } \pi_v\}.
\end{aligned}$$

We now state our second result.

Theorem 1.3. *The following hold:*

i) $\mathbb{P}(\mathscr{B}) > 0$;

ii) $\mathbb{P}(\mathscr{E}_>) > 0$.

Let Σ^* be the portion of Σ delimited by its medial triangle, boundary excluded (that is, the interior of the triangle formed by connecting the midpoints of the edges of the simplex, as per Figure 1.2). Not only Theorem 1.3 reveals which asymptotic scenarios for Θ_n are nonnegligible, but it has a straightforward corollary, which describes the asymptotics of π_n in those scenarios.

Corollary 1.4. *The following hold:*

i) $\mathbb{P}(\pi_n \text{ converges in } \Sigma^*) > 0$*;*

ii) $\mathbb{P}(\pi_n \text{ diverges in } \Sigma) > 0$*;*

iii) $\mathbb{P}(\{\pi_n \text{ converges in } \Sigma^*\} \setminus \mathscr{B}) = \mathbb{P}(\mathscr{B} \setminus \{\pi_n \text{ converges in } \Sigma^*\}) = 0$*;*

iv) $\mathbb{P}(\{\pi_n \text{ diverges in } \Sigma\} \cap \mathscr{E}_>) > 0$ *and* $\mathbb{P}(\{\pi_n \text{ diverges in } \Sigma\} \setminus (\mathscr{E}_> \cup \mathcal{D}_>)) = \mathbb{P}((\mathscr{E}_> \cup \mathcal{D}_>) \setminus \{\pi_n \text{ diverges in } \Sigma\}) = 0$*.*

The content of Theorem 1.3 and Corollary 1.4 becomes intuitive if we observe the two typical outcomes of Python simulations for the ERBRW, shown in Figures 1.3 and 1.4 (the violet dot is the centre of the simplex, about which one reflects any Θ to get π_Θ, upon halving of the reflected point). The two figures depict the two asymptotics, shown to be nonnegligible.

- Convergence of $\{\Theta_n\}$ in the interior of the simplex with convergence of $\{\pi_n\}$ in the interior of the medial triangle of the simplex. Note that in addition Corollary 1.4 (iii) states that this is the only possible scenario when either $\{\Theta_n\}$ converges in the interior of the simplex or $\{\pi_n\}$ converges in the interior of the medial triangle.

- Convergence of $\{\Theta_n\}$ to the boundary of the simplex (bounded away from the vertices) with divergence of $\{\pi_n\}$. Note that the second part of Corollary 1.4 (iv) states that we cannot rule out that when $\{\pi_n\}$ diverges, $\{\Theta_n\}$ might tend to a vertex (dominance). However, this scenario does not appear in the simulations, which is one of the reasons justifying Conjecture 1.7, which in turn would imply that convergence of $\{\Theta_n\}$ to the boundary of the simplex (bounded away from the vertices) is the only possible scenario when $\{\pi_n\}$ diverges. Note that this does not rule out convergence of both $\{\pi_n\}$ to the boundary of the medial triangle and $\{\Theta_n\}$ to the boundary of the simplex. This behaviour, however, does not appear in the simulations, suggesting that the corresponding event may be negligible.

For the ERBRW on the triangle we also define *monopoly* as the event $\mathcal{M}$ on which all but finitely many crossings happen along exactly one edge, or equivalently, eventually all particles stop crossing two edges.

Definition 1.5. Monopoly *is the event*

$$\mathcal{M} := \{\exists!\, i \in \{1,2,3\} :\ T_n^{(i)} \longrightarrow \infty\}.$$

Note that $\mathcal{M} \subseteq \mathcal{D}$. The third and last result is the following.

Theorem 1.6. *Almost surely, for all* $i \in \{1,2,3\}$*,* $T_n^{(i)} \longrightarrow \infty$ *as* $n \longrightarrow \infty$*. In particular,* $\mathbb{P}(\mathcal{M}) = 0$.

Conjecture 1.7. *Let* $\alpha = 1$*. Then* $\mathbb{P}(\mathcal{D}) = 0$*.*

We leave further results for the introduction to Part I.

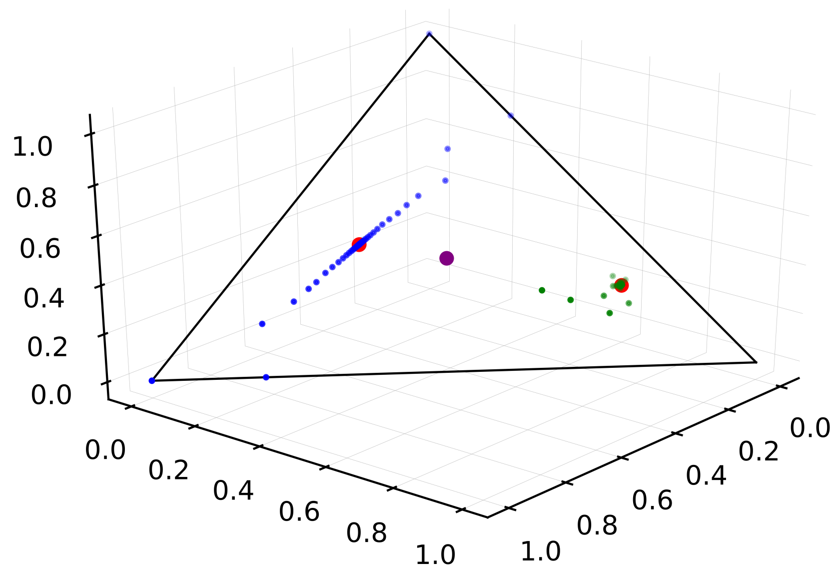


Figure 1.3: Convergence of Θ_n (green) and π_n (blue) to equilibrium points (Θ, π_Θ) (red)

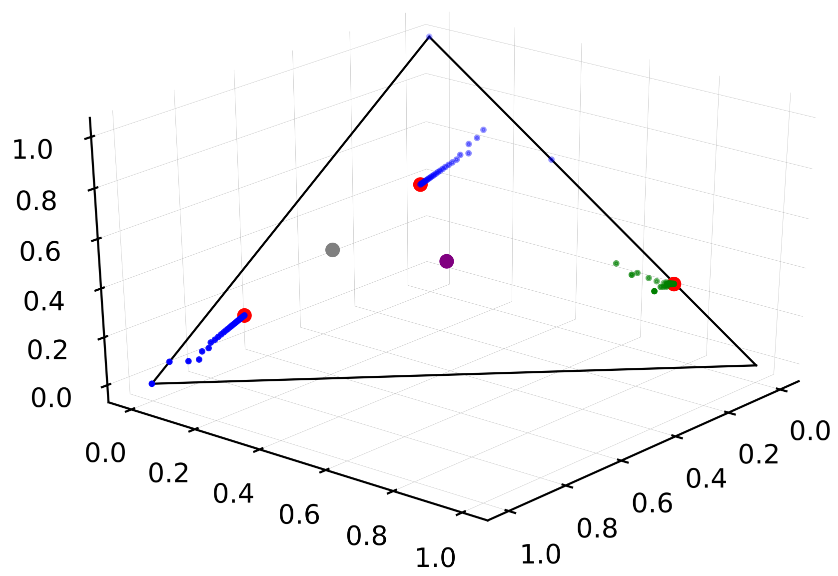


Figure 1.4: Convergence of Θ_n (green) to Θ (red) and divergence of π_n (blue) about π_Θ (gray)

1.3 Balls and bins with positive feedback

Balls and bins (BB) is a probabilistic model where balls are added to bins at random, following certain rules, concerning the number of balls, bins and the probabilistic laws governing how a ball falls in a specific bin. The classical setting is with two bins and the arrival of one ball at each discrete time. The most fundamental BB model with reinforcement is the Pólya urn with two colours. This urn model is in fact equivalent, upon identifying the colours (say red and white) with the bins, to throwing balls one at a time, at random, in two bins, with the probability of each ball landing in one bin proportional to the number of balls already in the bin. Adding *feedback* to this model means that the probability of a new ball arriving choosing a bin to land into, is proportional to a power of $\alpha \neq 1$ of the balls already in the bin (or more in general a function f of the number of balls). BB with feedback were introduced in the mathematical literature by [14]: the case $\alpha > 1$ is called *positive feedback* and the case $\alpha < 1$ *negative feedback*. Economists had already used these models [2], since positive feedback models the tendency, often observed in economic competition, of fast growth for the company that obtains first a nonnegligible initial advantage among a small number of companies, to the point of reaching monopoly or almost; negative feedback models instead a situation in which the competitor that has an advantage has difficulty keeping it, for example due to inefficiencies [47].

The convergence of the proportion of balls of each colour in a Pólya urn process to a beta-distributed random variable, as the number of allocated balls tends to infinity, is a well-known fact [18], which settles the case $\alpha = 1$ for the classical BB. In [14, §2] it was shown that for $\alpha > 1$ the proportion of balls in each of the two bins converges to either 0 or 1, as the number of allocated balls tends to infinity, a scenario called *dominance*. In [19], in the context of modelling neuron growth, a stronger result was shown: that if $\alpha > 1$, almost surely one of the (possibly more than two) bins gets all but finitely many balls, as the number of allocated balls tends to infinity. A scenario that we will call *monopoly*. In [32] monopoly was studied more quantitatively for two bins, showing that 'it takes a long time before a clear leader emerges, but once it does, it is likely to stick' [32, §1]. This was achieved through a technique known as *Rubin's construction*, which consists of embedding the BB process into a continuous time process, built from exponentially distributed random variables, related to the arrival times of the balls at each bin. It is the independence and explicit distribution of these arrival times, which allows studying more quantitative aspects of monopoly. As a matter of fact, thanks to Rubin's construction, in [31] the probability of monopoly of a bin, given its initial number of balls, has also been successfully approximated by the normal cumulative distribution function for large initial total number of balls. When $\alpha < 1$, in [14, §3] it was shown that the number of balls in the bins tends to a near-equal state, no matter the initial number of balls in the bins.

Generalisations of this model with more than two bins have been studied in [14, §4], obtaining similar results as in the two bins scenario by standard union bounds. In [8] preferential attachment has been studied in a model displaying a number of bins which grows as the balls arrive, one at a time. A second type of generalisations can also be obtained by considering different feedback functions for each bin, introducing asymmetry in the model. Very recently in [30] monopoly has been studied for a two bins asymmetric model, that is with feedbacks α_1 and α_2, generalising, among other results, the normal approximation obtained for the symmetric case in [31]. Additional generalisations can be obtained with various types of constraints on the allocated number of balls, and more can be said about negative feedback, but in this work we would like to focus on a third type of generalisation: time-dependent BB, for which the number of balls added varies with time.

In all the BB models considered so far, the bins always received one ball at a time. To explore time-dependence, let us go back to the Pólya urn once more. In [11, Appendix] a proof of Rubin's construction is provided, by solving the following time-dependent scheme for the Pólya urn: instead of adding the classical single ball of the same colour as the drawn one, σ_n white balls are added to the urn the nth time a white ball is drawn, and $\hat{\sigma}_n$ red balls are added the nth time a red ball is drawn, where $\{\sigma_n\}$ and $\{\hat{\sigma}_n\}$ are two deterministic integer-valued nonnegative sequences. Rubin's construction can be seen as the exponential embedding of this time-dependent model. Let $\{\tau_n\}$ be the partial sums of $\{\sigma_n\}$ and let $\{\hat{\tau}_n\}$ be the partial sums of $\{\hat{\sigma}_n\}$. Thanks to the embedding, monopoly, which is the event in which a colour is drawn all but finitely many times, can be studied through the convergence or divergence of $\sum_{n=0}^{\infty} 1/\tau_n$ and $\sum_{n=0}^{\infty} 1/\hat{\tau}_n$. If both series converge, monopoly is almost sure (with either colour having a positive probability of being the one eventually always drawn); if both diverge there is no monopoly; if one diverges and one converges, then monopoly of the colour corresponding to the convergent series is almost sure. In conclusion, it all depends on how fast the number of balls added for each colour grows, as this is the main mechanism providing the advantage necessary to reach a monopolistic regime in this model. The Rubin's construction is tailored to this type of time-dependence, where the time-evolution of the number of balls thrown is coordinated with the arrival times of the colours drawn.

In [37] a more general time-dependent version of the Pólya urn was studied, which, rephrased in terms of BB, would be equivalent to having two bins, and at time n the number of balls that arrives is σ_n and, in a bulk, all σ_n balls go in either of the bins, with probability of choosing a bin proportional to the number of balls in it. This type of time-dependency is already beyond Rubin's construction's reach, because by using only one sequence $\{\sigma_n\}$ for both bins, some of the coordination between the time-dependency and the arrival time at the bin chosen is lost. Martingale techniques come to the rescue. Let τ_n be the total number of balls in the bins at time n. Then dominance is almost sure if and only if $\sum_{n=0}^{\infty} \sigma_{n+1}^2/\tau_n^2 = \infty$. Monopoly is harder to analyse without Rubin's construction. In [49] it was shown that if $\sum_{n=0}^{\infty} \sigma_{n+1}^2/\tau_n^2 < \infty$, dominance is nonnegligible if and only if $\sum_{n=0}^{\infty} 1/\tau_n < \infty$. Moreover the study of the phase transition between no dominance and nonnegligible dominance was found to be closely related to the phase transition between no monopoly and nonnegligible monopoly.

In [48] N. Sidorova studied BB with two bins and positive feedback not only under time-dependence, but with each of the σ_n balls arriving at time n choosing *independently*, rather than in a bulk, to go in either bin; this changes significantly the evolution of the model. For example, in the model with balls added in a bulk, if σ_n grows fast enough, dominance is ensured without the need of any feedback; when the balls choose the bins independently, dominance never occurs if there is no feedback, no matter how fast σ_n grows. In the study of BB we will conduct, we generalise the latter model to more than two bins, with particular emphasis on dominance and positive feedback. Before commenting further on the results obtained in [48], we describe this generalisation more in detail.

1.4 Generalised balls and bins with positive feedback

Let $d \geq 2$ be an arbitrary number of bins and let $\{\sigma_n\}$ be an integer-valued positive sequence, representing the number of added balls at time $n \in \mathbb{N}$. Denote by $\tau_0 = T_0^{(1)} + \ldots + T_0^{(d)}$ the initial total number of balls in the d bins, where $T_0^{(i)}$ denotes the initial deterministic positive number of balls in the ith bin. For each $n \in \mathbb{N}$, let again

$$T_n^{(1)} + \ldots + T_n^{(d)} = \tau_n =: \tau_0 + \sum_{i=1}^{n} \sigma_i$$

be the total number of balls in the bins at time n. Minimal regularity conditions will be added to $\{\sigma_n\}$, when considered in full generality, but for the purpose of this introduction it will be enough to focus on $\sigma_n = \mu^n$, which is the relevant case for the connections of BB with the ERBRW. Denote by

$$\Theta_n^{(i)} := \frac{T_n^{(i)}}{\tau_n}$$

the proportion of balls in the ith bin at time n. Clearly they are valued in the standard simplex in d dimensions

$$\Delta^{d-1} := \{(x_1, \ldots, x_d) \in [0,1]^d : \; x_1 + \ldots + x_d = 1\}.$$

Note that, when there is no ambiguity, we will often switch from denoting the components with upper indices to denoting them with lower indices, when a time index is not involved. Let, for every integer $1 \leq i \leq d$ and $x \in \Delta^{d-1}$,

$$\psi^{(i)}(x) := \frac{x_i^\alpha}{\sum_{j=1}^{d} x_j^\alpha}.$$

In a model with no feedback, $\alpha = 1$; with feedback, $\alpha > 1$. Given the past up to time n, each of the σ_{n+1} balls thrown at time $n+1$ will fall, independently of each other at random, in the ith bin with probability proportional to the number of balls already in it, $T_n^{(i)}$, and the number of balls already in it raised to the power of $\alpha > 1$, $(T_n^{(i)})^\alpha$, respectively; that is with probability $\psi^{(i)}(\Theta_n)$. For each integer $1 \leq i \leq d$ consider the number of balls going in the ith bin at time $n+1$: these random variables jointly define a multinomial random vector of components denoted as $B_{n+1}^{(i)}$ (more technical details will be given in the introduction to Part II).

In the BB model with $d \geq 2$ we define, following the literature, *dominance* as the event $\mathcal{D}$ on which the number of balls in all but one of the d bins is negligible, as the number of allocated balls grows to infinity.

Definition 1.8. Dominance *is the event*

$$\mathcal{D} := \left\{ \exists\, i \in \{1, \ldots, d\} : \lim_{n \to \infty} \Theta_n^{(i)} = 1 \right\}.$$

We also define *monopoly* as the event $\mathcal{M}$ on which eventually all balls are added to only one of the d bins.

Definition 1.9. Monopoly *is the event*

$$\mathcal{M} := \left\{ \exists\, i \in \{1, \ldots, d\} : B_n^{(i)} = \sigma_n, \text{ ev.} \right\}.$$

The notation *ev.* stands for *eventually*, and means that there is a large enough time index $N \in \mathbb{N}$, such that for all $n \geq N$, $B_n^{(i)} = \sigma_n$. Note that $\mathcal{M} \subseteq \mathcal{D}$.

N. Sidorova showed that, for two bins:

- if there is no feedback, dominance is negligible ([48, Theorem 1.1]);
- if there is positive feedback, dominance is almost sure ([48, Theorem 1.2]).

No particularly strong restrictions are necessary on $\{\sigma_n\}$: it is assumed either bounded, or divergent to infinity. In Part II we obtain, as main results, the extension of these two theorems to $d > 2$ bins. The first result holds for $d > 2$ bins with the same level of generality as for two bins.

Theorem 1.10. *Let $\alpha = 1$. Then Θ_n converges almost surely to a bounded random variable Θ and $\mathbb{P}(\mathcal{D}) = 0$.*

The second result, due to technicalities in the argument arising for $d > 2$ bins, requires more restrictions on $\{\sigma_n\}$ than for two bins, and they are set through the following quantities:

- $\rho_n := \frac{\sigma_{n+1}}{\tau_n}$: we assume that $\{\rho_n\}$ is either bounded or diverges to infinity;
- $\theta_n := \frac{\log \tau_n}{\alpha^n}$: we assume that $\{\sigma_n\}$ is such that $\theta_n \longrightarrow \theta \in [0, \infty]$;
- $\lambda := \limsup_{n\to\infty} \frac{\sigma_{n+1}\sigma_{n-1}^{\alpha}}{\sigma_n^{\alpha+1}}$.

Theorem 1.11. *Let $\alpha > 1$. Then:*

- *if ρ_n is bounded, $\mathbb{P}(\mathcal{D}) = 1$;*
- *if $\rho_n \longrightarrow \infty$, $\theta = 0$ and $\lambda < 1$, $\mathbb{P}(\mathcal{D}) = 1$.*

We do not believe that the additional restrictions $\theta = 0$ and $\lambda < 1$ are necessary, as they arise from technical aspects of the argument.

Conjecture 1.12. *Let $\alpha > 1$. Then if $\rho_n \longrightarrow \infty$, $\mathbb{P}(\mathcal{D}) = 1$.*

As to monopoly, a time-dependent analysis is far more involved with the regularity of $\{\sigma_n\}$ (captured by λ) and finer details of its rate of growth (captured by ρ_n and θ_n). We will give a more detailed description of our results on monopoly (which we consider in some sense secondary to those about dominance) and more examples about the various regimes of growth and the parameters λ and θ, in the introduction to Part II. For the sake of a good understanding of the connections between BB and ERBRWs, the reader may feel less overwhelmed if, at a first reading, the focus is kept on the case in which ρ_n is bounded. In the rest of this section we will in fact discuss several regimes of growth, but note that for $\sigma_n = \mu^n$, ρ_n is bounded since it converges to a constant; then the regime of growth to which the ERBRW belongs is when ρ_n is bounded. We interpret this regime as *slow growth*, as opposed to, for example, $\sigma_n = \mu^{\mu^n}$, which is such that $\rho_n \longrightarrow \infty$, a regime referred to as *fast growth*. Given that the main focus of this introduction is the ERBRW and its connection with BB, fast growth is a less interesting case at a first reading.

With this in mind, N. Sidorova shows that, for two bins:

- if there is no feedback, monopoly is negligible ([48, Lemma 2.2]);

- if there is feedback, $\rho_n \longrightarrow \infty$, $\theta = 0$ and $\lambda < 1$, monopoly is negligible (part of [48, Theorem 1.4]).

. In [48, Theorem 1.4] she also shows that, for two bins, with positive feedback:

- if ρ_n is bounded, monopoly is almost sure;
- if $\rho_n \longrightarrow \infty$, $\theta = 0$ and $\lambda < 1$, monopoly is almost sure.

In Part II we also obtain the extension of these results to $d > 2$ bins. We start in reverse order, since the generalisation of the last two results follows directly from Theorem 1.11, and in fact it can be noted that it states that monopoly is almost sure in all the regimes in which we showed that dominance is almost sure.

Corollary 1.13. *Let $\alpha > 1$. Then:*

- *if ρ_n is bounded, $\mathbb{P}(\mathcal{M}) = 1$;*
- *if $\rho_n \longrightarrow \infty$, $\theta = 0$ and $\lambda < 1$, $\mathbb{P}(\mathcal{M}) = 1$.*

We conclude by stating the generalisation of the results about negligible monopoly.

Theorem 1.14. *Let $\alpha = 1$. Then $\mathbb{P}(\mathcal{M}) = 0$.*

Theorem 1.15. *Let $\alpha > 1$ and $\theta = \infty$. Then $\mathbb{P}(\mathcal{M}) = 0$.*

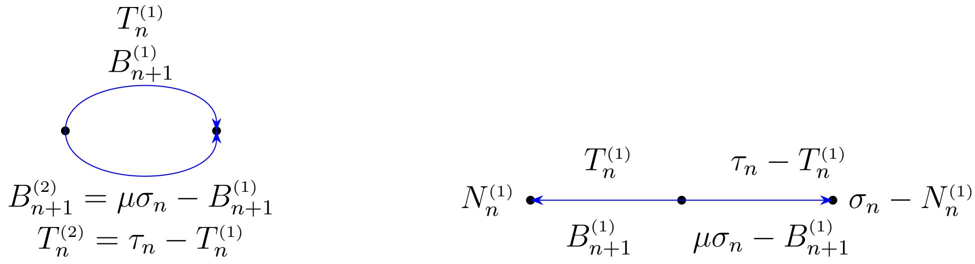
Theorem 1.16. *Let $\alpha > 1$, $\rho_n \longrightarrow \infty$, $\theta = 0$ and $\lambda > 1$. Then $\mathbb{P}(\mathcal{M}) = 0$.*

As anticipated, in the next section, we will only focus on the cases in which ρ_n bounded, when referring to the results for BB.

1.5 The connection between ERBRW on graphs and generalised BB

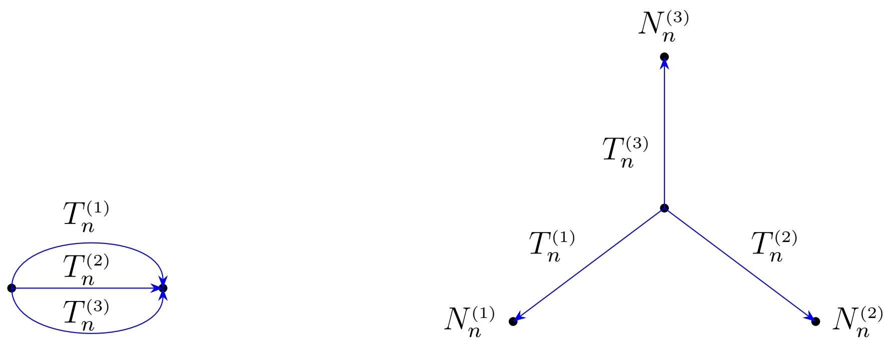
Note that the event of dominance for BB with $d = 3$ in Definition 1.8 coincides with the event of dominance for the ERBRW on the triangle in Definition 1.2, and corresponds to the random walk getting stuck on an edge, a situation previously discussed when reviewing SERRW. Similarly the event of monopoly for BB with $d = 3$ in Definition 1.9 can be rephrased so as to coincide with the event of monopoly for the ERBRW on the triangle in Definition 1.5, and corresponds to the random walk eventually stopping crossing all but one edge. In both models, we are particularly interested in finding when dominance and monopoly are negligible and when they are almost sure, depending on the feedback. We now focus on $\{\Theta_n\}$, so as to see clearly the connections between the two models. This connection will not come as a surprise, if one considers that the BB model is a generalisation of the Pólya urn, which is, in some sense, the building block of all processes with reinforcement.

Let us begin with $d = 2$ in BB, and then consider the ERBRW between two nodes and a double edge, defined along the lines of the ERBRW on the triangle (see Figure 1.5a; equivalently, if one prefers to also keep, in analogy with the definition on the triangle, two nontrivial $N_n^{(i)}$, we can think of it as an ERBRW on the binary tree of height one in Figure 1.5b, but one must perform the branching and count the beginning of the discrete times every other move, rather than at every move: starting from the root in the middle at time 0, the particles branch and choose the incident edge to cross; once on a leaf, the next time unit starts, and we update the $T_n^{(i)}$ and



(a) ERBRW on the double edge (b) ERBRW on the binary tree

Figure 1.5: Two bins and ERBRWs



(a) ERBRW on the triple edge (b) ERBRW on the ternary tree

Figure 1.6: Three bins and ERBRWs

$N_n^{(i)}$; then they must return to the root node, without increasing the edge crossing count; there they branch again, and repeat: the successive time unit will start once they reach again the leaf, so that in the recorded $N_n^{(i)}$ we never see when the vertices get empty). In both cases the vector of proportions Θ_n can be studied, by symmetry, by just focusing on $\Theta_n^{(1)}$ since $\Theta_n^{(2)} = 1 - \Theta_n^{(1)}$. Both constructions yield the same stochastic process. We already mentioned that in [48] the sequence of balls thrown in the bins is very general, and clearly it includes the case $\sigma_n = \mu^n$. The aforementioned results of no dominance and no monopoly in absence of feedback and dominance and monopoly with positive feedback apply directly to the ERBRW on the double edge, with and without feedback.

We can proceed similarly and consider an ERBRW between two nodes and a triple edge with and without feedback (see Figure 1.6a; equivalently, we can think of it as an ERBRW on the ternary tree of height one in Figure 1.6b, with the same precautions as those adopted on the binary tree), and note that it can similarly be identified with a time-dependent BB model with $d = 3$, with or without feedback. Hence the need for the study of a generalised (that is, $d > 2$) time-dependent BB model, to further the understanding of these types of ERBRWs. By Proposition 4.48 and Theorem 1.14 (no dominance nor monopoly in absence of feedback) and Theorem 1.11 and Corollary 1.13 (dominance and monopoly with positive feedback, in particular for $\sigma_n = \mu^n$), imply that the study of dominance and monopoly for the ERBRW on multiple edges between two nodes (or equivalently, on d-ary trees of height one) is completed, with and without feedback.

This connection between the two models can be somehow exploited when approaching the ERBRW on graphs that do not offer an immediate identification with BB. The main focus of the present work will be the study of ERBRW on the triangle with no feedback. Simulations, both with and without feedback, suggest that the

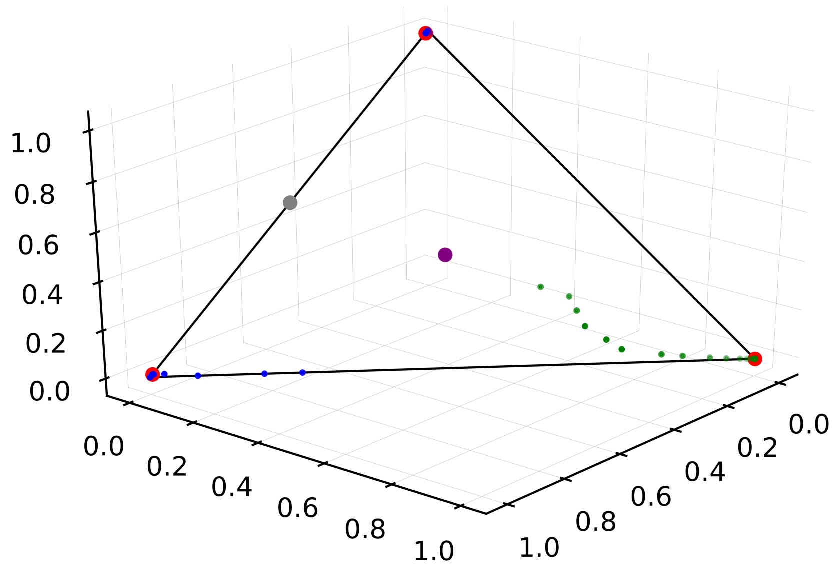


Figure 1.7: Convergence of Θ_n (green) to a vertex v (red) of the simplex (dominance) and divergence of π_n (blue) about π_v (gray)

results obtained for BB with $d = 3$ and $\sigma_n = \mu^n$ agree with the behaviour displayed by the random walk on the triangle. This further supports Conjecture 1.7 (which is stated for the ERBRW with no feedback) and gives rise to a further conjecture for ERBRW with positive feedback (which we will not treat in this work), well supported by simulations (see Figure 1.7, a simulation for $\alpha = 2$).

Conjecture 1.17. *Let $\alpha > 1$. Then, for the ERBRW on the triangle, $\mathbb{P}(\mathcal{D}) = 1$.*

When studying ERBRWs it will be necessary to complement the general methodology developed in the study of BB with a much more dynamical approach. Discrete dynamical systems have an important role in the heuristics of the arguments for BB, but martingale theory, like in the case of SERRWs, definitely plays a much more significant and decisive role, when carrying out those arguments. For example proving Theorem 1.10 is fairly easy, due to $\{\Theta_n\}$ being a martingale in BB. When studying the ERBRW, this is no longer true, and our main goal will be recovering the almost sure convergence of $\{\Theta_n\}$ through an original implementation and development of discrete *dynamical systems* techniques in a random setting. The application of dynamical systems to reinforced processes is a well established method: a survey on some of these methods can be found in [39]; among them, *stochastic approximation* is well-known for successfully dealing with random perturbations of a dynamical system, usually arising as martingale differences. A survey of stochastic approximation techniques can be found in [4]. This theory traditionally relies on ODE methods, taking advantage of slowly decaying martingale differences perturbations. The branching we introduce in the ERRW model, however, makes the martingale increments decay much faster than what is required to get a good approximation in the continuum via ODE methods. Thus we develop alternative dynamical techniques, which rely on fast decreasing martingale perturbations.

1.6 Note for the reader

In the first part of the work we will study the ERBRW, as we consider generalised BB (which we actually studied first) as instrumental to it: the study of BB, along with the simulations, prompted all the conjectures proposed in this work, and set the high-level structure of the methodology followed in analysing the ERBRW. Nonetheless Parts I and II can be considered and read independently. The simulations aforementioned both for the ERBRW and the corresponding dynamical system, are all run with $\mu = 2$, for 60 iterates.

Due to the overall level of technicality and length, we designed an appendix, where we moved work that might be considered not essential at a first reading (Part III), or not original (Part IV), but in some parts, which will be flagged, it will be relied upon. Chapter A in Part III is original work that offers a window on the inner workings of our arguments, and might also be relevant for future developments: when necessary the reader will be referred to it; Chapter B in Part III contains a short heuristic supplement we will not rely upon directly; Part IV contains some technical results that come as a straightforward variation of those in [48], and have been included for self-containedness, due to the arguments significantly relying on them.

1.7 Notation

We will adopt the standard probabilistic notation *ev.* for *eventually* and *i.o.* for *infinitely often*. More formally, given a sequence of events $\{E_n\}$ in a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, we define the events

$$\{E_n,\ \text{ev.}\} := \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} E_n = \{\omega \in \Omega,\ \exists m(\omega) \in \mathbb{N},\ \omega \in E_n\ \forall n \geq m(\omega)\}$$

and

$$\{E_n,\ \text{i.o.}\} := \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} E_n = \{\omega \in \Omega,\ \forall m \in \mathbb{N}\ \exists n(\omega) \geq m,\ \omega \in E_{n(\omega)}\}.$$

Given real-valued random variables X, Y, we will often (especially for random times) adopt the probabilistic notation $X \wedge Y := \min\{X, Y\}$ and $X \vee Y := \max\{X, Y\}$.

The complementary of an event $E \subset \Omega$ will always be denoted as E^c. Standard asymptotic notations such as $\mathcal{O}$, o, Ω, $\asymp$ and $\sim$ will often be adapted to the probabilistic setting in the following way: $\mathcal{O}_\omega$, o_ω, Ω_ω, $\asymp_\omega$ and $\sim_\omega$. Throughout this work such notation always means that, on the event considered, the constants involved in the standard definition are random: they apply for almost all ω in the event considered, often with pointwise dependence on ω. For example if $T_n^{(1)} = \mathcal{O}_\omega(\mu^n)$ on Ω, this means that for almost every $\omega \in \Omega$, there is a constant $C = C(\omega)$ such that

$$\limsup_{n \longrightarrow \infty} \left| \frac{T_n^{(1)}(\omega)}{\mu^n} \right| \leq C(\omega).$$

We will use this notation with some flexibility. In fact the same notation may be used if, in particular cases, the constant applies uniformly, but only for almost all ω in any event considered, or if it applies for every ω in the event but not uniformly. It will be clear from the context and our comments. When we switch to standard notation, it means that the constant applies uniformly and for all ω on the event considered.

We will not be overly formal about vector notation. When we use row notation for inline formulas, we mean the transpose of a column vector, as we will omit the transpose symbol for simplicity. Thus in inline formulas all row vectors are intended as column vectors. Also, when using boldface for asymptotic notations and numbers, we mean that they denote vectors. For example in a 3-dimensional setting $\mathbf{\frac{1}{3}}$ means $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, and $\Theta_n = \boldsymbol{o_\omega(1)}$ means that Θ_n is an almost surely vanishing 3-dimensional random vector, that is for every $\varepsilon > 0$, for almost every ω there is $N(\omega) \in \mathbb{N}$ such that for all $n \geq N(\omega)$, $\|\Theta_n(\omega)\| < \varepsilon$, for some vector norm $\|\cdot\|$ which will, depending on the context, either be the 1-norm or the Euclidean norm.

We leave further comments of more specific notational character to the introductions to Parts I and II.

Part I

Edge-reinforced branching random walk on the triangle

Chapter 2

Introduction

This introductory chapter provides all the technical details regarding the ERBRW on the triangle: we first describe the process in terms of a system of iterative equations, and then perform some manipulations so as to reduce the stochastic iterative scheme to that of a randomly perturbed deterministic iteration map, which we refer to as *randomly perturbed dynamical system*.

This is the general set-up that preludes to the analysis carried out in this part of the work, concerning the ERBRW on the triangle. We conclude this chapter with a high-level description of the arguments leading to our main results and a note for the reader.

2.1 Iterative equations of the model

Let us begin with recalling the probabilistic model of the ERBRW on the triangle outlined in Chapter 1, so as to add the necessary adjustments concerning the extension to a possibly nonintegral branching factor $\mu > 1$ and some discussions regarding the nonlaziness assumption.

At time 0 we start with one particle at any of the vertices, which branches with deterministic constant factor μ. If μ is not an integer, the interpretation of the random walkers as particles needs to be replaced by unit and fractional masses. When a nonintegral branching occurs, the resulting (possibly nonintegral) total mass is seen as composed of particle-like unit masses (making up the integer part of the total mass) and an additional fractional mass, which has a particle-like behaviour too. This will not change the analysis significantly: we identify the unit masses with particles, while the extra fractional mass is seen as a smaller particle. Although we will always adopt, as the main point of view in the exposition, the analogy with particles, we will address explicitly masses only in the isolated instances where the analysis slightly differs.

By the next time unit, the offspring particles (unit and small masses alike) will have travelled independently through any of the incident edges at random, with probability of choosing either one of them proportional to the positive number of edge crossings initially deterministically assigned $T_0^{(1)}$, $T_0^{(2)}$, $T_0^{(3)}$ (more in general raised to a power of $\alpha > 1$ if the feedback is positive, but we will not study positive feedback in this part). The offspring particles cannot remain at the same vertex in the next time unit, thus we forbid loops. Although the assumption that the graph is simple is customary in the reinforced random walk literature, we would like to note that the model can be extended to include loops (or even parallel edges, but in the comments that follow we briefly only address loops). The natural way loops are counted for an edge-reinforced model on an undirected graph, slightly differs from the graph theoretical convention

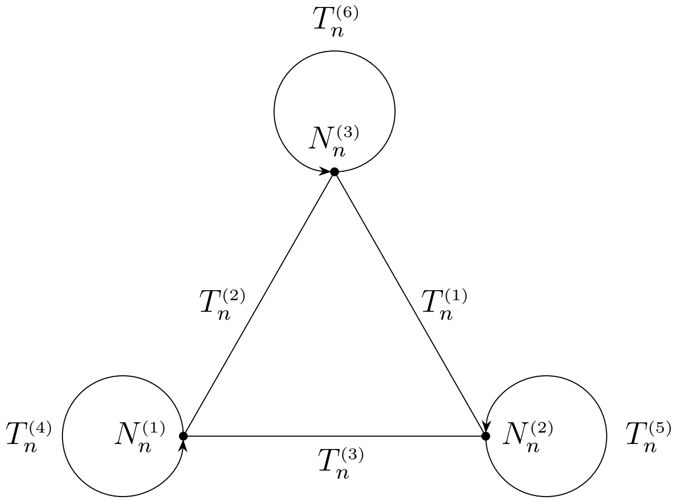


Figure 2.1: ERBRW on the triangle with loops

(where it is customary to count loops twice, as they come with two incidences on the same vertex). For edge-reinforced random walks it is more natural to count loops only once, as the probabilities change with the edge crossings. Thus we can think of a loop as a directed arc, having the same head and tail (see Figure 2.1). We will use the word *edge* solely for the triangle's undirected edges. According to this convention, loops are at a disadvantage in a triangle. At each vertex, a loop is competing against two incident edges, which the particles can choose to traverse from both ends. Let $i \neq j \neq k \in \{1, 2, 3\}$ with k referring to the occupied vertex, i, j referring to the incident edges and $k + 3$ referring to the loop based at k. Given the past up to time n, after branching, at time $n + 1$ each particle (unit or fractional mass alike) at the kth vertex, independently of the others at random, traverses the ith or jth edge with probability $\psi(\Theta_n^{(i)}, \Theta_n^{(j)}, \Theta_n^{(k+3)})$ and $\psi(\Theta_n^{(j)}, \Theta_n^{(i)}, \Theta_n^{(k+3)})$ respectively, while it traverses the loop with probability $\psi(\Theta_n^{(k+3)}, \Theta_n^{(i)}, \Theta_n^{(j)})$, where

$$\psi(x, y, z) = \psi(x, z, y) = \frac{x^\alpha}{x^\alpha + y^\alpha + z^\alpha}.$$

This scheme puts the loops at disadvantage because the edges, once the particles have branched enough to be scattered among all vertices, can get crossed from both ends.

The resulting dynamics is conjectured to be the following: the loops end up being traversed negligibly many times, as they attract fewer and fewer particles. At the same time, as the loops get progressively neglected by the particles, the dynamics on the triangle evolves the same as that of a model with no loops. Quantitatively, loops will simply slow down the model's evolution towards its asymptotic behaviour with no loops: in the beginning, through appropriate weighting of the loops, particles can obviously be likely to be attracted to them. However, qualitatively, as the number of particles grows and starts populating all the vertices due to random fluctuations, the edges of the triangle, crossed from both ends, get disproportionally more competitive against the loops. This informal heuristics, confirmed by simulations, suggests that a model with loops is not much more interesting than a model without loops. Once it is shown that all loops are negligibly crossed almost surely, it should be possible to recover Theorems 1.1 and 1.3 and Corollary 1.4, by suitably adapting the arguments; Conjecture 1.7 and Conjecture 1.17 still seem to apply. Clearly it is possible to add loops following a different convention, that is allowing them to be traversed from both ends, thus counting twice, but this would yield a completely different model, which is out of our scope.

Going back to the offspring particles on the triangle with no loops, once they branch and traverse either of the edges, they reach the new vertices, the edge crossings of the chosen edge are updated, and the process repeats. Each particle's crossing increases the edge crossings by one, so unit masses have the same effect, while factional masses increase the edge crossings by the corresponding fraction of mass, leading to possibly nonintegral edge crossings $T_n^{(i)}$ in the case of nonintegral μ. The total number of particles and mass (for nonintegral μ) at time n is clearly still

$$N_n^{(1)} + N_n^{(2)} + N_n^{(3)} =: \sigma_n = \mu^n,$$

and the total number of crossings up to time n is still

$$T_n^{(1)} + T_n^{(2)} + T_n^{(3)} =: \tau_n = \tau_0 + \sum_{i=1}^{n} \sigma_i.$$

Thus $\Theta_n^{(i)}$ and $\pi_n^{(i)}$ still represent the corresponding proportions of edge crossings and particles (mass) at the vertices at time n. Recall that we exclusively work with no feedback, thus

$$\phi(x, y) := \frac{x}{x+y}$$

will be the function providing us with the probabilities each particle (mass) has, to cross the incident edges. Let $i \neq j \neq k \in \{1, 2, 3\}$ with k referring to the occupied vertex and i, j referring to the incident edges. Given the past up to time n, after branching, at time $n+1$ each particle (unit or fractional mass alike) at the kth vertex, independently of the others at random, traverses the ith or jth edge with probability $\phi(\Theta_n^{(i)}, \Theta_n^{(j)})$ and $\phi(\Theta_n^{(j)}, \Theta_n^{(i)})$ respectively. Since the number of particles traversing one of the incident edges from the kth vertex at time $n + 1$(according to the diagram in Figure 1.1) is a binomial random variables $B_{n+1}^{(k)}$, in the case of nonintegral μ the binomial $B_{n+1}^{(k)}$ will denote the total amount of integral mass traversing one of the incident edges from the kth vertex at time $n + 1$, and will therefore not be sufficient to denote the total mass traversing the edge. It will be necessary to add to the binomial, the fractional mass traversing the corresponding edge, by exploiting a Bernoulli random variable $I_{n+1}^{(k)}$ suitably rescaled to the fraction. We now formalise this through the model's equations.

For integral μ, $\{B_{n+1}^{(i)}\}$ are binomial random variables distributed, conditionally on the past, as follows:

$$B_{n+1}^{(1)} \sim \text{Bin}(\mu N_n^{(1)}, \phi(\Theta_n^{(3)}, \Theta_n^{(2)})) \tag{2.1}$$
$$B_{n+1}^{(2)} \sim \text{Bin}(\mu N_n^{(2)}, \phi(\Theta_n^{(1)}, \Theta_n^{(3)})) \tag{2.2}$$
$$B_{n+1}^{(3)} \sim \text{Bin}(\mu N_n^{(3)}, \phi(\Theta_n^{(2)}.\Theta_n^{(1)})), \tag{2.3}$$

and are otherwise independent of each other and the past. Denote, as customary, the supporting probability space $(\Omega, \mathscr{F}, \mathbb{P})$ endowed with the filtration $\{\mathscr{F}_n\}_{n\in\mathbb{N}_0}$, where $\mathscr{F}_0 = \{\emptyset, \Omega\}$, $\mathscr{F}_n := \sigma(B_1, \ldots, B_n)$ for all $n \in \mathbb{N}$ and $\mathscr{F}_\infty := \sigma\left(\bigcup_{n=0}^{\infty} \mathscr{F}_n\right)$. We will denote by $\mathbb{P}_{\mathscr{F}_n}$, $\mathbb{E}_{\mathscr{F}_n}$, $\text{Var}_{\mathscr{F}_n}$ and $\text{Cov}_{\mathscr{F}_n}$ the conditional probability, expectation, variance and covariance respectively.

For nonintegral μ, $\{B_{n+1}^{(i)}\}$ are binomial random variables and $\{I_{n+1}^{(i)}\}$ are Bernoulli random variables distributed, conditionally on the past, as follows:

$$B^{(1)}_{n+1} \sim \mathrm{Bin}(\lfloor \mu N^{(1)}_n \rfloor, \phi(\Theta^{(3)}_n, \Theta^{(2)}_n)) \tag{2.4}$$
$$B^{(2)}_{n+1} \sim \mathrm{Bin}(\lfloor \mu N^{(2)}_n \rfloor, \phi(\Theta^{(1)}_n, \Theta^{(3)}_n)) \tag{2.5}$$
$$B^{(3)}_{n+1} \sim \mathrm{Bin}(\lfloor \mu N^{(3)}_n \rfloor, \phi(\Theta^{(2)}_n, \Theta^{(1)}_n)), \tag{2.6}$$
$$I^{(1)}_{n+1} \sim \mathrm{Ber}(\phi(\Theta^{(3)}_n, \Theta^{(2)}_n)) \tag{2.7}$$
$$I^{(2)}_{n+1} \sim \mathrm{Ber}(\phi(\Theta^{(1)}_n, \Theta^{(3)}_n)) \tag{2.8}$$
$$I^{(3)}_{n+1} \sim \mathrm{Ber}(\phi(\Theta^{(2)}_n, \Theta^{(1)}_n)), \tag{2.9}$$

otherwise independent of each other and the past. Clearly the filtration will be $\mathscr{F}_n := \sigma(B_1, I_1, \ldots, B_n, I_n)$ instead. As aforementioned, we will add $\{I^{(i)}_{n+1}\}$ to the binomials, upon rescaling by the factors $\{\mu N^{(1)}_n\}$, $\{\mu N^{(2)}_n\}$, $\{\mu N^{(3)}_n\}$ respectively, where $\{\cdot\}$ denotes the fractional part. This will lead, for nonintegral μ, to the definition of the random variables

$$\tilde{B}^{(1)}_{n+1} := B^{(1)}_{n+1} + \{\mu N^{(1)}_n\} I^{(1)}_{n+1} \tag{2.10}$$
$$\tilde{B}^{(2)}_{n+1} := B^{(2)}_{n+1} + \{\mu N^{(2)}_n\} I^{(2)}_{n+1}, \tag{2.11}$$
$$\tilde{B}^{(3)}_{n+1} := B^{(3)}_{n+1} + \{\mu N^{(3)}_n\} I^{(3)}_{n+1}. \tag{2.12}$$

The iterative description of our model, reflected in Figure 1.1, turns into a system of difference equations (for nonintegral μ, we simply replace $B^{(i)}_{n+1}$ with $\tilde{B}^{(i)}_{n+1}$).

For integral μ we obtain

$$T^{(1)}_{n+1} = T^{(1)}_n + \mu N^{(3)}_n + B^{(2)}_{n+1} - B^{(3)}_{n+1} \tag{2.13}$$
$$T^{(2)}_{n+1} = T^{(2)}_n + \mu N^{(1)}_n + B^{(3)}_{n+1} - B^{(1)}_{n+1} \tag{2.14}$$
$$T^{(3)}_{n+1} = T^{(3)}_n + \mu N^{(2)}_n + B^{(1)}_{n+1} - B^{(2)}_{n+1} \tag{2.15}$$
$$N^{(1)}_{n+1} = \mu N^{(2)}_n + B^{(3)}_{n+1} - B^{(2)}_{n+1} \tag{2.16}$$
$$N^{(2)}_{n+1} = \mu N^{(3)}_n + B^{(1)}_{n+1} - B^{(3)}_{n+1} \tag{2.17}$$
$$N^{(3)}_{n+1} = \mu N^{(1)}_n + B^{(2)}_{n+1} - B^{(1)}_{n+1}. \tag{2.18}$$

while for nonintegral μ we obtain the same equations, but with every term $B^{(k)}_{n+1}$ replaced by $\tilde{B}^{(k)}_{n+1}$. For simplicity of exposition, it is best to simply work with only one notation, with the understanding that when μ is not an integer everywhere a binomial $B^{(k)}_{n+1}$ appears, it should be replaced by $\tilde{B}^{(k)}_{n+1}$. This will be beneficial since there will be very few places, in the analysis, where the random variables $\tilde{B}^{(k)}_{n+1}$ will have to be dealt with separately, with a specific approach.

These equations can be trivially manipulated to yield the corresponding equations for the proportions of edge crossings

$$\Theta^{(1)}_{n+1} = \frac{\tau_n}{\tau_{n+1}} \Theta^{(1)}_n + \frac{\sigma_{n+1}}{\tau_{n+1}} \pi^{(3)}_n + \frac{1}{\tau_{n+1}} (B^{(2)}_{n+1} - B^{(3)}_{n+1})$$
$$\Theta^{(2)}_{n+1} = \frac{\tau_n}{\tau_{n+1}} \Theta^{(2)}_n + \frac{\sigma_{n+1}}{\tau_{n+1}} \pi^{(1)}_n + \frac{1}{\tau_{n+1}} (B^{(3)}_{n+1} - B^{(1)}_{n+1})$$
$$\Theta^{(3)}_{n+1} = \frac{\tau_n}{\tau_{n+1}} \Theta^{(3)}_n + \frac{\sigma_{n+1}}{\tau_{n+1}} \pi^{(2)}_n + \frac{1}{\tau_{n+1}} (B^{(1)}_{n+1} - B^{(2)}_{n+1}),$$

and for the proportions of particles (mass) at the vertices

$$\pi_{n+1}^{(1)} = \pi_n^{(2)} + \frac{1}{\sigma_{n+1}}(B_{n+1}^{(3)} - B_{n+1}^{(2)})$$
$$\pi_{n+1}^{(2)} = \pi_n^{(3)} + \frac{1}{\sigma_{n+1}}(B_{n+1}^{(1)} - B_{n+1}^{(3)})$$
$$\pi_{n+1}^{(3)} = \pi_n^{(1)} + \frac{1}{\sigma_{n+1}}(B_{n+1}^{(2)} - B_{n+1}^{(1)}).$$

Note that

$$\tau_n = \tau_0 + \sum_{i=1}^{n} \sigma_i = \tau_0 + \sum_{i=1}^{n} \mu^i = \tau_0 + \frac{1-\mu^{n+1}}{1-\mu} - 1 = \tau_0 + \frac{\mu}{\mu-1}(\mu^n - 1).$$

Next we extract the martingale parts by adding and subtracting the conditional expectations, which happen to be the same for both cases, integral and nonintegral μ. In fact, conditionally on the corresponding $\mathscr{F}_n$, both $B_{n+1}^{(i)}$ and $\tilde{B}_{n+1}^{(i)}$ have the same expectation, due to integer part and fractional part adding up to the original total mass, and the probability parameter of the $I_{n+1}^{(i)}$ being the same as that of the binomial. Thus we have another set of equations holding for both integral and nonintegral μ, with the usual understanding that one must replace the $B_{n+1}^{(i)}$ with the $\tilde{B}_{n+1}^{(i)}$:

$$\Theta_{n+1}^{(1)} = \Theta_n^{(1)} + \frac{\sigma_{n+1}}{\tau_{n+1}}(-\Theta_n^{(1)} + \pi_n^{(2)}\phi(\Theta_n^{(1)}, \Theta_n^{(3)}) + \pi_n^{(3)}\phi(\Theta_n^{(1)}, \Theta_n^{(2)})) + S_{n+1}^{(1)} \tag{2.19}$$
$$\Theta_{n+1}^{(2)} = \Theta_n^{(2)} + \frac{\sigma_{n+1}}{\tau_{n+1}}(-\Theta_n^{(2)} + \pi_n^{(3)}\phi(\Theta_n^{(2)}, \Theta_n^{(1)}) + \pi_n^{(1)}\phi(\Theta_n^{(2)}, \Theta_n^{(3)})) + S_{n+1}^{(2)} \tag{2.20}$$
$$\Theta_{n+1}^{(3)} = \Theta_n^{(3)} + \frac{\sigma_{n+1}}{\tau_{n+1}}(-\Theta_n^{(3)} + \pi_n^{(1)}\phi(\Theta_n^{(3)}, \Theta_n^{(2)}) + \pi_n^{(2)}\phi(\Theta_n^{(3)}, \Theta_n^{(1)})) + S_{n+1}^{(3)} \tag{2.21}$$
$$\pi_{n+1}^{(1)} = \pi_n^{(3)}\phi(\Theta_n^{(2)}, \Theta_n^{(1)}) + \pi_n^{(2)}\phi(\Theta_n^{(3)}, \Theta_n^{(1)}) + R_{n+1}^{(1)} \tag{2.22}$$
$$\pi_{n+1}^{(2)} = \pi_n^{(1)}\phi(\Theta_n^{(3)}, \Theta_n^{(2)}) + \pi_n^{(3)}\phi(\Theta_n^{(1)}, \Theta_n^{(2)}) + R_{n+1}^{(2)} \tag{2.23}$$
$$\pi_{n+1}^{(3)} = \pi_n^{(2)}\phi(\Theta_n^{(1)}, \Theta_n^{(3)}) + \pi_n^{(1)}\phi(\Theta_n^{(2)}, \Theta_n^{(3)}) + R_{n+1}^{(3)} \tag{2.24}$$

where for each $n \in \mathbb{N}_0$ we have defined

$$S_{n+1}^{(1)} := \frac{1}{\tau_{n+1}}(B_{n+1}^{(2)} - \mu N_n^{(2)}\phi(\Theta_n^{(1)}, \Theta_n^{(3)})) + \frac{1}{\tau_{n+1}}(\mu N_n^{(3)} - B_{n+1}^{(3)} - \mu N_n^{(3)}\phi(\Theta_n^{(1)}, \Theta_n^{(2)})) \tag{2.25}$$

$$S_{n+1}^{(2)} := \frac{1}{\tau_{n+1}}(B_{n+1}^{(3)} - \mu N_n^{(3)}\phi(\Theta_n^{(2)}, \Theta_n^{(1)})) + \frac{1}{\tau_{n+1}}(\mu N_n^{(1)} - B_{n+1}^{(1)} - \mu N_n^{(1)}\phi(\Theta_n^{(2)}, \Theta_n^{(3)})) \tag{2.26}$$

$$S_{n+1}^{(3)} := \frac{1}{\tau_{n+1}}(B_{n+1}^{(1)} - \mu N_n^{(1)}\phi(\Theta_n^{(3)}, \Theta_n^{(2)})) + \frac{1}{\tau_{n+1}}(\mu N_n^{(2)} - B_{n+1}^{(2)} - \mu N_n^{(2)}\phi(\Theta_n^{(3)}, \Theta_n^{(1)})) \tag{2.27}$$

$$R_{n+1}^{(1)} := \frac{1}{\sigma_{n+1}}(B_{n+1}^{(3)} - \mu N_n^{(3)}\phi(\Theta_n^{(2)}, \Theta_n^{(1)})) + \frac{1}{\sigma_{n+1}}(\mu N_n^{(2)} - B_{n+1}^{(2)} - \mu N_n^{(2)}\phi(\Theta_n^{(3)}, \Theta_n^{(1)})) \tag{2.28}$$

$$R_{n+1}^{(2)} := \frac{1}{\sigma_{n+1}}(B_{n+1}^{(1)} - \mu N_n^{(1)}\phi(\Theta_n^{(3)}, \Theta_n^{(2)})) + \frac{1}{\sigma_{n+1}}(\mu N_n^{(3)} - B_{n+1}^{(3)} - \mu N_n^{(3)}\phi(\Theta_n^{(1)}, \Theta_n^{(2)})) \tag{2.29}$$

$$R_{n+1}^{(3)} := \frac{1}{\sigma_{n+1}}(B_{n+1}^{(2)} - \mu N_n^{(2)}\phi(\Theta_n^{(1)}, \Theta_n^{(3)})) + \frac{1}{\sigma_{n+1}}(\mu N_n^{(1)} - B_{n+1}^{(1)} - \mu N_n^{(1)}\phi(\Theta_n^{(2)}, \Theta_n^{(3)})) \tag{2.30}$$

2.2 Reduction to a perturbed dynamical system

Recall that $\Sigma := \{p \in \mathbb{R}^3 : p_1+p_2+p_3 = 1, p_1,p_2,p_3 \geq 0\}$. Let the set of vertices of Σ be $V := \{v_1, v_2, v_3\}$ (where v_i is the ith element of the canonical basis of $\mathbb{R}^3$) and the set of edges (with endpoints removed) of Σ be $E := \{E_1, E_2, E_3\}$ (where E_i denotes the edge opposite to v_i). Let $\partial\Sigma$ denote the boundary of Σ and $\mathring{\Sigma}$ the interior of Σ. Let

$$M_p := \begin{pmatrix} 0 & \phi(p_3,p_1) & \phi(p_2,p_1) \\ \phi(p_3,p_2) & 0 & \phi(p_1,p_2) \\ \phi(p_2,p_3) & \phi(p_1,p_3) & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{p_3}{p_1+p_3} & \frac{p_2}{p_1+p_2} \\ \frac{p_3}{p_2+p_3} & 0 & \frac{p_1}{p_1+p_2} \\ \frac{p_2}{p_2+p_3} & \frac{p_1}{p_1+p_3} & 0 \end{pmatrix}.$$

For each $\Theta \in \Sigma_0 := \Sigma \setminus V$ the matrix M_Θ is well defined, (2.19) to (2.24) can be rewritten as

$$\pi_{n+1} = M_{\Theta_n}\pi_n + R_{n+1} \tag{2.31}$$

$$\Theta_{n+1} = \Theta_n + \rho_{n+1}(\mathbf{1} - \Theta_n - \pi_n - \pi_{n+1}) = \Theta_n + \rho_{n+1}(\mathbf{1} - \Theta_n - \pi_n - M_{\Theta_n}\pi_n) + S_{n+1} \tag{2.32}$$

where

$$\rho_{n+1} := \frac{\sigma_{n+1}}{\tau_{n+1}},$$

having used the fact that $S^{(i)}_{n+1} = -\rho_{n+1}R^{(i)}_{n+1}$ for all $i \in \{1,2,3\}$. Note that since for all $i \in \{1,2,3\}$, $T^{(i)}_0 > 0$, $\Theta_0 \notin \partial\Sigma$; also, by (2.13) to (2.15) for all $i \in \{1,2,3\}$, $T^{(i)}_{n+1} \geq T^{(1)}_n$, hence for all $n \in \mathbb{N}_0$, $\Theta_n \notin \partial\Sigma$.

Due to the denominators τ_{n+1} and σ_{n+1}, which grow geometrically fast, the martingale differences will be shown to be negligible to some degree, so by dropping them from (2.31) and (2.32), one will be left with the predictable component of the process. It is a dynamical system that will most likely drive the asymptotic behaviour of the stochastic process. This observation is at the core of our method. We study the predictable component of the process as a dynamical system in its own right, and therefore we change notation from (Θ, π) to (p, q). This yields the discrete nonlinear, nonautonomous dynamical system

$$q_{n+1} = M_{p_n}q_n \tag{2.33}$$

$$p_{n+1} = (1-\rho_n)p_n + \rho_{n+1}(\mathbf{1} - q_n - q_{n+1}) \tag{2.34}$$

well defined on $\Sigma_0 \times \Sigma$. Note that

$$\rho_n = \frac{\mu^n}{\tau_0 - \frac{\mu}{\mu-1} + \mu^n\frac{\mu}{\mu-1}} \longrightarrow \rho := \frac{\mu-1}{\mu}$$

as $n \longrightarrow \infty$. Replacing ρ_n by its limit ρ in (2.33) and (2.34) will simplify the analysis of the dynamical system, so one can study

$$q_{n+1} = M_{p_n}q_n \tag{2.35}$$

$$p_{n+1} = (1-\rho)p_n + \rho(\mathbf{1} - q_n - q_{n+1}) \tag{2.36}$$

instead, which is well defined on $\Sigma_0 \times \Sigma$. Note that (2.36) is a convex combination of the *past* (p_n) and the *update* $(1 - q_n - q_{n+1})$. To further simplify this system, we shall let $\rho = 1$ in (2.36), so as to suppress the past component of (2.35) and (2.36), which yields

$$q_{n+1} = M_{p_n}q_n \tag{2.37}$$

$$p_{n+1} = \mathbf{1} - q_n - q_{n+1}. \tag{2.38}$$

This iteration is well defined on Σ_0^2 and, for every $i \in \{1,2,3\}$, it is well defined also for $p_0 \in E_i$ with $q_0 \in \Sigma_0 \cup \{v_i\}$. This represents the starting point of our analysis. Our main goal is to prove almost sure convergence of the process, via gaining knowledge of the asymptotic behaviour of the sample paths. However, the facts proved to obtain almost sure convergence will be powerful enough to yield further results, involving preferential attachment. This knowledge is indeed very helpful in ruling out monopoly, in determining which asymptotics are nonnegligible, and is likely to be the key in proving the conjectured negligibility of dominance.

2.3 Outline of contents

The convergence argument for the stochastic process is better understood by starting from an overview of the two main results concerning the dynamical systems aforementioned. The first result is its convergence.

Theorem 2.1. *For any orbit, $\{p_n\}$ converges to some $p_* \in \Sigma$, depending on the initial condition.*

Let $q_p := (\mathbf{1} - p)/2$, that is, the halved reflection about $\frac{\mathbf{1}}{\mathbf{3}}$ (the centre of the simplex). In Remark 3.13 we show that $\|q_n - q_{p_n}\|_1$ converges to a limit ℓ, depending on the initial conditions. For all $p \in \partial\Sigma \setminus V$ denote $e_{-1}(p)$, the eigenvector of M_p corresponding to the eigenvalue -1 (in Lemma 3.19 we show all the properties of M_p). Recall that Σ^* denotes the portion of Σ delimited by its medial triangle (boundary excluded), thus its closure $\overline{\Sigma}^*$ is the medial triangle. We now state our second main result, following from the convergence of $\{p_n\}$.

Corollary 2.2. *For any orbit, $\{q_n\}$ either converges in $\overline{\Sigma}^*$ or is asymptotic to the 2-cycle*

$$\left\{ q_{p_*} \pm \frac{\ell}{2} e_{-1}(p_*) \right\},$$

where $p_* := \lim_{n\longrightarrow\infty} p_n$ and $\ell := \lim_{n\longrightarrow\infty} \|q_n - q_{p_n}\|_1$, depending on the initial condition.

The dynamical system does offer significant challenges, and a large portion of this part will be dedicated to it. Since the same methods that work for the system obeying (2.37) and (2.38), with technical adjustments, work also for the one obeying (2.35) and (2.36) (which is studied in Chapter A, producing analogous results) and consequently for the more general (2.33) and (2.34); for simplicity we will focus this outline on (2.37) and (2.38), which is studied in Chapter 3.

We start by identifying the equilibrium points of the system: $\{(p, q_p) : \ p \in \Sigma\}$. We distinguish these equilibria between internal ones (when $p \notin \partial\Sigma$) and boundary ones (when $p \in \partial\Sigma$). The standard stability analysis is not fruitful, due to the number of dimensions of the system and the density of the equilibrium points. Hence we introduce a nonnegative potential $V(p,q) := \|q - q_p\|_1$. In Section 3.2 it is shown that this potential yields, loosely speaking, a gradient-like dynamics, since it is nonincreasing along the orbit of the dynamical system. However, the iteration map of the system is not defined on a compact set and the equilibria are dense: therefore most standard topological dynamical results involving gradient-like systems do not apply to this specific system. Nonetheless, admitting a potential is still a valuable property of the system, and it will allow us to build tools to show convergence of the system. Since the potential is a monotone nonincreasing function on the orbits, it has a limit

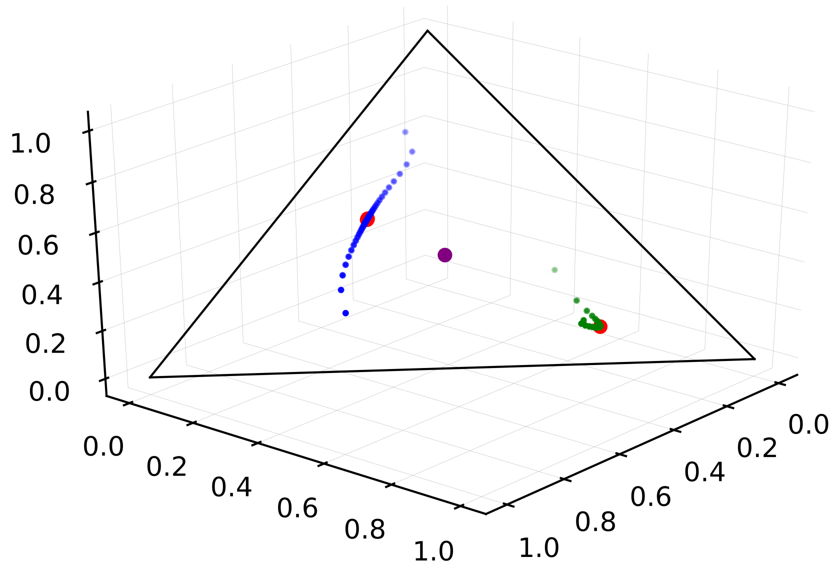


Figure 2.2: Convergence of p_n (green) and q_n (blue) to internal equilibrium points (p, q_p) (in red)

$\ell \geq 0$. In Proposition 3.15 it is shown that if $\{p_n\}$ is bounded away from $\partial\Sigma$, then the potential decays geometrically along the orbits, and convergence of $\{(p_n, q_n)\}$ to one of the internal equilibrium points follows. In Figure 2.2 it is possible to get an intuitive understanding of this fact by observing the output of a Python simulation ($\mu = 2$).

The methodology just described is at the core of our approach. For this system, finding suitable initial conditions, from which a specific asymptotic behaviour follows, is infeasible. Rather, we prescribe some generic asymptotic behaviour for $\{p_n\}$ (for instance being bounded away from $\partial\Sigma$) and prove the convergence as a result. Hence we devise a case analysis of mutually exclusive asymptotic behaviours, and prove convergence of $\{p_n\}$ for each of them. This cannot be done without understanding $\{q_n\}$'s asymptotics. Proposition 3.15 takes care of the case, when $\{p_n\}$ is bounded away from $\partial\Sigma$ (and $\ell = 0$). There are two other mutually exclusive cases: $\{p_n\}$ approaches the boundary with either $\ell = 0$ or $\ell > 0$. In each of these cases we distinguish between two subcases: convergence to any of the vertices of Σ; existence of a subsequence of $\{p_n\}$, bounded away from the vertices of Σ. We also make the following conjecture, well supported by numerical evidence.

Conjecture 2.3. *For any orbit, $\{p_n\}$ never converges to a vertex of the simplex.*

This conjecture is not required in order to show convergence of the stochastic process $\{\Theta_n\}$. However, it is a crucial part of a tentative argument for Conjecture 1.7, which states that $\{\Theta_n\}$ almost never converges to the vertices of the simplex.

In Section 3.4 it is shown that if $\ell = 0$ and $\{p_n\}$ approaches the boundary, with a subsequence $\{p_{n_j}\}$ bounded away from the vertices, the system will approach a boundary equilibrium point based at an edge (recall that the edges E_i have been defined with the endpoints removed), from which $\{(p_n, q_n)\}$ can be shown to converge to a (possibly different) boundary equilibrium point based on the same edge, that is

(p, q_p) with p in the edge (thus, not a vertex). This is the conclusion reached with Theorem 3.27. The build up to this theorem is a list of technical lemmas that deal with the representation of the dynamical system via eigencoordinates (with respect to the eigenvectors of M_p, lying on the linear space, to which the simplex is parallel) and the handling of the error terms arising from the representation. The change of coordinates captures the oscillatory nature of the dynamical behaviour of $\{q_n\}$, which contributes to the convergence of $\{p_n\}$ (by looking at Figure 2.2, it is rather evident that oscillations are also involved when converging to an internal equilibrium; however, in that case the proof does not need to rely on them). To this argument, the assumption of the existence of a subsequence of $\{p_n\}$ bounded away from the vertices is essential: it would not work, by solely relying on an initial condition, close enough to some suitable boundary equilibrium. Equivalently, it is not known whether this asymptotic boundary behaviour is actually displayed by the system or not. In the many simulations we performed, it never appeared. If it does happen for some initial conditions, the basin of attraction has to necessarily be meagre (this will follow from Section 3.5). On the other hand, Proposition 3.17 shows that it is possible, for the case in which $\{p_n\}$ is bounded away from the boundary, to identify an open neighbourhood of initial conditions in Σ close enough to an internal equilibrium, such that the system converges to a (possibly different) internal equilibrium.

This discussion brings us to the last question to be answered about the dynamical system: is the dynamical system also convergent when $\ell > 0$, $\{p_n\}$ approaches the boundary and there is a subsequence of $\{p_n\}$ bounded away from the vertices? If yes, is this last hypothesis necessary, or we can identify an open neighbourhood of initial conditions for the convergence? This is dealt with in Section 3.5, and it is the case in which the system displays the richest asymptotics, that is convergence of $\{p_n\}$ in an edge, but divergence of $\{q_n\}$ (more precisely, it is asymptotically 2-periodic). This is the content of Theorem 3.48, and the build up to this theorem requires a new toolbox: a complete description of the set of accumulation points of the orbits, explored through the asymptotics of the system under *boundary initial conditions* (that is with $p_0 \in \partial\Sigma \setminus V$), which yield *boundary orbits* (orbits such that eventually $p_n \in \partial\Sigma$). The study of boundary orbits is conducted in Section 3.5.1 and the results are the following: we have convergence of $\{p_n\}$ within the edge, on which the initial condition is, away from the vertices; we have either convergence of $\{q_n\}$ in $\partial\Sigma^*$ or asymptotic 2-periodicity of $\{q_n\}$. In Section 3.5.2 this result is used to derive the description of the set of accumulation points of *regular orbits* (that is, such that eventually $p_n \in \mathring{\Sigma}$) approaching the boundary. This set is not very informative as a whole, but when fixing the specific value of $\ell > 0$ for the orbit considered, it narrows down to only two possible configurations: for any p chosen on an edge, q_n approaches, oscillating, two points, denoted as q and $\hat{q}$, which lie on either side of q_p in the direction of the eigenvector of M_p corresponding to the eigenvalue -1 (denoted as $e_{-1}(p)$, see Figure 2.3 for a Python simulation with $\mu = 2$).

Although the eigenvectors are the key to understanding the system, the oscillations of $\{q_n\}$ being bounded away from $\{q_{p_n}\}$ makes a change to eigencoordinates not fruitful, when $\ell > 0$. In Lemma 3.46 the asymptotic oscillations are shown to give rise to a geometric decay of the component of $\{p_n\}$ that vanishes along the subsequence aforementioned, and this couples with a geometric upper bound on the increments of one of the two other components, as shown in Lemma 3.47; all of which yields convergence of $\{p_n\}$ as per Theorem 3.48, where we feed one estimate into the other, through a suitably engineered sequence of hitting times. The set-up of this argument relies only on the initial conditions belonging to a suitably small open neighbourhood of an oscillatory limit configuration. The existence of such a neighbourhood, starting

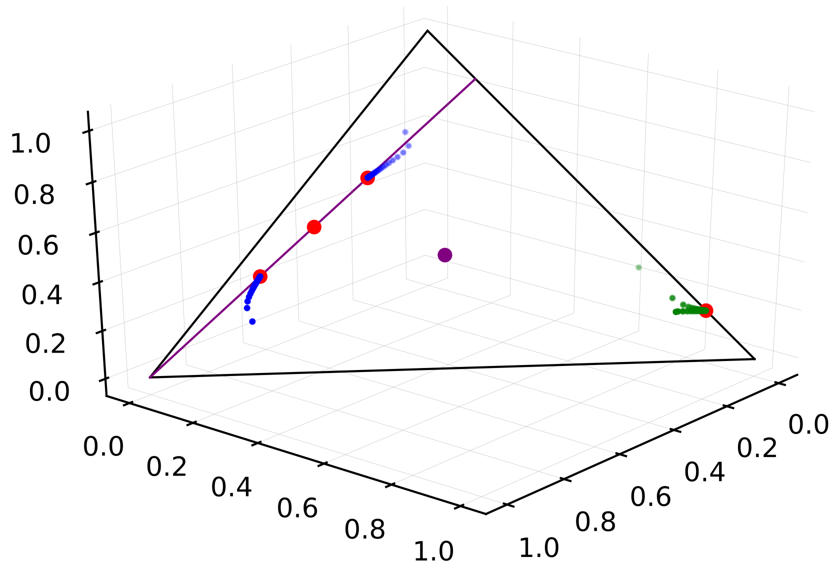


Figure 2.3: Convergence of p_n (green) to a boundary limit p (red) and asymptotic oscillation of q_n (blue) about (q_p) (red) along the eigendirection $e_{-1}(p)$ (violet)

within which the system tends to a (possibly different) limit configuration of the same type on the same edge, shows that, as anticipated, it is not possible, for the limit points in the boundary case $\ell = 0$, to have a fat basin of attraction. Because of the density of the two sets of accumulation points (the one in the boundary case $\ell > 0$ and the one in the boundary case $\ell = 0$), a contradiction would follow, if this were true, as the two neighbourhoods would intersect. This suggests that the set of initial conditions for the boundary case having $\ell = 0$ should be at least meagre, possibly negligible.

The work in Chapter 3 can be generalised for the dynamical system described by (2.35) and (2.36): this is done in Chapter A: it has its own challenges, but they are of technical nature. Generalising the results in Chapter A to the general case described by (2.33) and (2.34) is trivial and can directly be dealt with while studying the stochastic process's sample paths, which is done in Chapter 4. We see the stochastic process obeying (2.31) and (2.32) with *regular initial conditions* (that is $T_0^{(i)} > 0$ for all $i \in \{1, 2, 3\}$) as a randomly perturbed dynamical system, so that its sample paths can be analysed through the work made on the dynamical system, by taming the random perturbations coming from the martingale increments R_{n+1} and S_{n+1}. The geometric growth of the number of particles in the system ensures that a geometrically decaying upper bound eventually holds for these perturbation terms (Lemma 4.1). This allows enough control, to be able to proceed pointwise in $\omega \in \Omega$, with Ω partitioned suitably into events that match the case analysis of the deterministic system (recall the events defined in Chapter 1, right after Definition 1.2). In this fashion we prove Theorem 1.1, stating that $\{\Theta_n\}$ converges almost surely to a random variable Θ. Although it is not the main objective of our work, as it is a far easier case to analyse, it is worth noting that in Section 4.5.1, in analogy with the results obtained for the dynamical system, we derive convergence of $\{\Theta_n\}$ also for *boundary initial conditions*, that is when there is exactly one $i \in \{1,2,3\}$ such that $T_0^{(i)} = 0$: this is the content of Theorem 4.26. In this case the stochastic process is more rigid. We show that $\{\Theta_n\}$ converges almost surely to a point Θ in the edge to which the initial condition Θ_0 belongs, and that whether $\{\pi_n\}$ diverges or converges is almost surely determined by the initial condition π_0 (in particular, if there is a component of $\pi_0 - \pi_{\Theta_0}$ along $e_{-1}(\Theta_0)$, $\{\pi_n\}$ almost surely diverges; otherwise it almost surely converges to π_Θ). Overall Chapter 4 is a further generalisation of the deterministic case previously discussed, and it is the most technical section, due to how riddled some estimates get when perturbed. Theoretically speaking, it is a self-contained chapter. However, we preferred an exposition that relies on the chapters dedicated to the dynamical system, which have thus been included. This sheds more light on the methodology we followed, and benefits the reader's understanding, even though it comes at the price of a longer presentation.

In Section 4.7 we show that the two asymptotic behaviours of convergence of $\{(\Theta_n, \pi_n)\}$ to internal equilibria and convergence of $\{\Theta_n\}$ only, to the boundary (away from the vertices) are both nonnegligible events. Informally, we refer to them as *typical*. In Figures 1.3 and 1.4 one can observe, by comparison, how well the process's typical sample paths align, after a few noticeable initial fluctuations, with the typical orbits of the dynamical system, which can be observed in Figures 2.2 and 2.3. The approach to this result is algorithmic. An iterative sequence of moves, each having positive probability, is devised to approach arbitrarily either of the two limit configurations. We show this by relying on two particular cases: $\left\{\frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{1}}{\mathbf{3}}\right\}$ and $\{((0, {}^1\!/\!_2, {}^1\!/\!_2), ({}^3\!/\!_4, {}^1\!/\!_8, {}^1\!/\!_8)), ((0, {}^1\!/\!_2, {}^1\!/\!_2), ({}^1\!/\!_4, {}^3\!/\!_8, {}^3\!/\!_8))\}$, but it could have been any other configurations of the same type. Theorem 1.3 shows, via a probabilistic argument exploiting moderate deviations of the binomials from their mean, that it is possible to get close enough to these configurations at the same time that the negligibility of the martingale increments kicks in, thus allowing the workings of Proposition 4.8 and Theorem 4.42 (the random analogue of Proposition 3.17 and Theorem 3.48) to drive the system to the corresponding type of asymptotic behaviour.

In Section 4.8 the properties of the sample paths investigated while studying the convergence of the stochastic process are combined with a martingale argument (initially devised for BB) exploiting the predictable quadratic variation, so as to show, in Theorem 1.6, that almost surely all edges get infinitely many crossings from the particles. This means that it almost never happens that one or two edges get only finitely many crossings of particles. Therefore monopoly does not occur.

To conclude, in Section 4.9 we show some progress made towards proving Conjecture 1.7 (negligibility of dominance). The conjecture rests on several grounds: firstly, simulations support it; next, the connection with BB and Theorem 1.10, stating that no feedback implies negligible dominance for any number of bins $d \geq 2$ (in particular for $d = 3$); lastly, the quantitative estimate of Proposition 4.48, obtained through a bootstrap argument involving martingale theory, coupled with nonautonomous linear dynamical systems results stemming from the works of Perron, Frobenius and Poincaré. This estimate in particular suggests that the stochastic process would follow very closely the deterministic dynamical system, when near the vertices. Since the simulations we performed strongly suggest that the dynamical system does not tend to the vertices, proving Conjecture 1.7 seems likely to require a deeper study of the dynamical system near the vertices, so as to show Conjecture 2.3 first. This is a task, for which we still have to develop the necessary toolbox and it also justifies keeping the study of the dynamical system conducted in Chapter 3 as the core of the first part of this dissertation. There is in fact still work to be done on the dynamical aspects of the problem.

Due to the challenges offered by the case $\alpha = 1$, we have not yet attempted the study of the ERBRW with positive feedback. However, both simulations and the success in showing Theorem 1.11 for any number of bins $d \geq 2$, which applies to several regimes of growth, including $\sigma_n = \mu^n$, give us confidence that dominance is almost sure when feedback is added to the ERBRW (Conjecture 1.17).

2.4 Note for the reader

In Section 4.5.1 we focus on the study of the model with boundary initial conditions, which is much easier. If the reader has no interest in it or does not need the intuition it provides on the workings of the model, this section can be skipped as it has no consequences for the study of regular initial conditions, which is the main goal of this work.

At times, in Section 4.5.2 we will rely on some results in Chapter A: by no means this requires that the whole of Chapter A be read. The results quoted will suffice, and it is very likely that, having gone through Chapter 3, the reader will be able to fill in the gaps without even reading the arguments. At other times the results from Chapter A are mentioned merely for comparison.

Chapter 3

The dynamical system with $\rho = 1$

In this chapter the convergence of $\{p_n\}$ for the dynamical system described by (2.37) and (2.38) is shown. For one-step iterations arguments, a less cumbersome notation will sometimes be used, in order to omit the time index, and (2.37) and (2.38) will often be written as

$$\hat{q} = M_p q$$
$$\hat{p} = \mathbf{1} - q - \hat{q},$$

where we recall that

$$M_p := \begin{pmatrix} 0 & \frac{p_3}{p_1+p_3} & \frac{p_2}{p_1+p_2} \\ \frac{p_3}{p_2+p_3} & 0 & \frac{p_1}{p_1+p_2} \\ \frac{p_2}{p_2+p_3} & \frac{p_1}{p_1+p_3} & 0 \end{pmatrix}.$$

3.1 Preliminaries

In this section we motivate some preliminary reductions, which can be made in order to formally simplify the study of the dynamical system. Since for the stochastic process described by (2.31) and (2.32), $\Theta_n \in \mathring{\Sigma}$ for all $n \in \mathbb{N}_0$, one could think that our interest is limited to initial conditions such that $p_0 \notin \partial\Sigma$. This seems also a natural choice for the deterministic dynamical system, being the iteration matrix in (2.37) not well defined on $V \subset \partial\Sigma$. However, in general we will consider $p_0 \in \Sigma_0$, especially in Section 3.5.1; the reason is that, as we will see, $p_0 \in \partial\Sigma \setminus V$ yields a *boundary orbit*.

Definition 3.1. *A* boundary orbit *is an orbit of the dynamical system* $\{(p_n, q_n)\}$*, such that eventually* $p_n \in \partial\Sigma \setminus V$*.*

The importance of boundary orbits arises from being helpful in studying *regular orbits* that approach the boundary.

Definition 3.2. *A* regular orbit *is an orbit of the dynamical system* $\{(p_n, q_n)\}$*, such that eventually* $p_n \in \mathring{\Sigma}$*.*

Regular orbits are the more challenging to study and the more informative about the stochastic process, and thus the main object of our interest. The following remark shows what can go wrong when allowing both p_0 and q_0 in the boundary of the simplex without any further restrictions.

Remark 3.3. *Let* $p_0 \in E_i$ *and* $q_0 = v_j$ *for some* $i \neq j \in \{1, 2, 3\}$*. Then* $p_1 = v_k$*, with* $i \neq j \neq k \in \{1, 2, 3\}$*, which is inadmissible for the iteration scheme described by* (2.37) *and* (2.38)*.*

Proof. Without loss of generality, by symmetry, assume $i = 1$ and $j = 2$. Let $p_0 = (0, p_0^{(2)}, 1 - p_0^{(2)})$ and denote $a = \phi(p_0^{(2)}, p_0^{(3)})$. As

$$q_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1-a & 0 & 0 \\ a & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

it follows that

$$p_1 = \mathbf{1} - v_1 - v_2 = v_3,$$

thus (2.37) is inconsistent, since M_{v_3} is not well-defined. □

This justifies why, although for the stochastic process we have that $\pi_0 \in V$, we cannot assume an arbitrary $q_0 \in V$ for the dynamical system, if we let p_0 be on the boundary of the simplex. There is only one case, which does not lead to an inconsistent iteration, and the resulting orbit is quite trivial, as it can be seen from the following remark.

Remark 3.4. *Let $p_0 \in E_i$ and $q_0 = v_i$ for some $i \in \{1, 2, 3\}$. Then $p_n = p_0$ for all $n \in \mathbb{N}$ and q_n is 2-periodic.*

Proof. Without loss of generality, by symmetry, assume $i = 1$. Let $p_0 = (0, p_0^{(2)}, 1 - p_0^{(2)})$ and denote $a = p_0^{(2)}$. As

$$q_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1-a & 0 & 0 \\ a & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1-a \\ a \end{pmatrix},$$

it follows that

$$p_1 = \mathbf{1} - v_1 - \begin{pmatrix} 0 \\ 1-a \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ a \\ 1-a \end{pmatrix} = p_0$$

and

$$q_2 = \begin{pmatrix} 0 & 1 & 1 \\ 1-a & 0 & 0 \\ a & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1-a \\ a \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = q_0.$$

By induction this shows that q_n is 2-periodic, with p_n fixed. □

The same happens, with one-step delay, to any initial condition $p_0 \notin \partial\Sigma$ and $q_0 \in V$. The vertex pushes p_1 to the correct edge, so as to start the 2-cycle, as the following remark shows.

Remark 3.5. *Let $p_0 \notin \partial\Sigma$ and $q_0 \in V$. Then $p_n = p_1$ for all $n \in \mathbb{N}$ and q_n is 2-periodic.*

Proof. By symmetry, without loss of generality, this can be shown by performing the computation for this specific starting point $q_0 = (1, 0, 0)$. Since $q_1 = (0, \phi(p_0^{(3)}, p_0^{(2)}), \phi(p_0^{(2)}, p_0^{(3)})) = (0, 1-a, a)$, where we defined $a := \phi(p_0^{(2)}, p_0^{(3)})$, and since $p_0 \notin \partial\Sigma$, we have that $0 < a < 1$. It follows that $p_1 = \mathbf{1} - q_1 - q_0 = (0, a, 1-a) \in \partial\Sigma$. Then

$$M_{p_1} = \begin{pmatrix} 0 & 1 & 1 \\ 1-a & 0 & 0 \\ a & 0 & 0 \end{pmatrix}$$

and as a result $q_2 = q_0$. Thus $p_2 = \mathbf{1} - q_2 - q_1 = \mathbf{1} - q_0 - q_1 = p_1$ and $q_3 = q_2$. The orbit $\{(p_n, q_n)\}$ is an eventual 2-cycle, since except for (p_0, q_0), it oscillates between the values (p_1, q_1) and (p_1, q_0) previously determined. □

Since $q_0 \in V$ yields only boundary orbits, whose asymptotic behaviour is easily computed in any case except for the nonadmissible one, we can assume $q_0 \notin V$ and avoid unnecessary technicalities in various hypotheses of our claims, when dealing with regular orbits. Furthermore, it is easy to show that $p_0 \in \partial\Sigma \setminus V$ always yields boundary orbits, when q_0 is admissible, so when dealing with regular orbits we can assume $p_0 \notin \partial\Sigma$.

Remark 3.6. *Let $p_0 \in E_i$ for some $i \in \{1,2,3\}$ and $q_0 \neq v_j$ for every $j \in \{1,2,3\} \setminus \{i\}$. Then $p_n^{(i)} = 0$ for all $n \in \mathbb{N}$.*

Proof. By symmetry, without loss of generality, assume that $i = 1$, that is $p_0 \in E_1$, which is equivalent to $p_0^{(1)} = 0$ and $0 < p_0^{(2)} < 1$. Then since $q_0 \neq v_1$,

$$p_1 = \mathbf{1} - q_0 - \begin{pmatrix} 0 & 1 & 1 \\ 1 - p_0^{(2)} & 0 & 0 \\ p_0^{(2)} & 0 & 0 \end{pmatrix} q_0.$$

From the form taken by M_{p_0}, it follows immediately that $p_1^{(1)} = 0$. Therefore, by induction, it follows that for all $n \in \mathbb{N}_0$, $p_n^{(1)} = 0$, that is we are in presence of a boundary orbit on the edge E_1. □

When $p_0 \notin \partial\Sigma$, we can also assume $q_0 \notin \partial\Sigma$, since the case with q_0 on an edge reduces to the case where q_0 is in the interior.

Remark 3.7. *Let $p_0 \notin \partial\Sigma$ and $q_0 \in E_i$ for some $i \in \{1,2,3\}$. Then $p_1 \notin \partial\Sigma$ and $q_1 \notin \partial\Sigma$.*

Proof. By symmetry, without loss of generality, we can show the claim via explicit calculation for $q_0 = (0, a, 1-a)$, with $0 < a < 1$. Since $p_0 \notin \partial\Sigma$, it follows that $q_1 \notin \partial\Sigma$ because

$$q_1 = \begin{pmatrix} a\phi(p_0^{(3)}, p_0^{(1)}) + (1-a)\phi(p_0^{(2)}, p_0^{(1)}) \\ (1-a)\phi(p_0^{(1)}, p_0^{(2)}) \\ a\phi(p_0^{(1)}, p_0^{(3)}) \end{pmatrix}$$

and $\phi(p_0^{(i)}, p_0^{(j)}) = 0$ if and only if $p_0^{(i)} = 0$ (which is not allowed). Hence

$$p_1 = \mathbf{1} - q_1 - q_0 = \begin{pmatrix} 1 - a\phi(p_0^{(3)}, p_0^{(1)}) - (1-a)\phi(p_0^{(2)}, p_0^{(1)}) \\ (1-a)\phi(p_0^{(2)}, p_0^{(1)}) \\ a\phi(p_0^{(3)}, p_0^{(1)}) \end{pmatrix} \notin \partial\Sigma,$$

since $\phi(p_0^{(i)}, p_0^{(j)}) = 1$ if and only if $p_0^{(j)} = 0$ (which is not allowed) and therefore the convex combination subtracted to 1 in the first component is subunitary (while the other terms are nonzero). □

As a result the case when $p_0 \notin \partial\Sigma$ and $q_0 \in \partial\Sigma \setminus V$ has been reduced to that of an orbit not starting at the boundary, since one can relabel (p_1, q_1) as (p_0, q_0) and then apply the following remark.

Remark 3.8. *If $p_0 \notin \partial\Sigma$ and $q_0 \notin \partial\Sigma$, then $p_n \notin \partial\Sigma$ and $q_n \notin \partial\Sigma$ for all $n \in \mathbb{N}$.*

Proof. It is enough to show that if for some $n \in \mathbb{N}_0$, $p_n \notin \partial\Sigma$ and $q_n \notin \partial\Sigma$, then $p_{n+1} \notin \partial\Sigma$ and $q_{n+1} \notin \partial\Sigma$. First note that

$$q_{n+1} = \begin{pmatrix} \phi(p_n^{(3)}, p_n^{(1)})q_n^{(2)} + \phi(p_n^{(2)}, p_n^{(1)})q_n^{(3)} \\ \phi(p_n^{(3)}, p_n^{(2)})q_n^{(1)} + \phi(p_n^{(1)}, p_n^{(2)})q_n^{(3)} \\ \phi(p_n^{(2)}, p_n^{(3)})q_n^{(1)} + \phi(p_n^{(1)}, p_n^{(3)})q_n^{(2)} \end{pmatrix} \notin \partial\Sigma,$$

which follows from $\phi(p_n^{(i)}, p_n^{(j)}) = 0$ if and only if $p_n^{(i)} = 0$ (which is not the case) and $q_n^{(i)} \neq 0$ for all $i \in \{1, 2, 3\}$. Denote by J a 3 by 3 matrix of ones and define

$$\overline{M}_{p_n} := J - I - M_{p_n} = \begin{pmatrix} 0 & \phi(p_n^{(1)}, p_n^{(3)}) & \phi(p_n^{(1)}, p_n^{(2)}) \\ \phi(p_n^{(2)}, p_n^{(3)}) & 0 & \phi(p_n^{(2)}, p_n^{(1)}) \\ \phi(p_n^{(3)}, p_n^{(2)}) & \phi(p_n^{(3)}, p_n^{(1)}) & 0 \end{pmatrix}.$$

Then

$$p_{n+1} = \mathbf{1} - (M_{p_n} + I)q_n = Jq_n - (M_{p_n} + I)q_n = \overline{M}_{p_n} q_n.$$

Since $\overline{M}_{p_n}$ has the same entries as M_{p_n}, but swapped within the columns, $p_{n+1} \notin \partial\Sigma$ for the same reasoning applied to q_{n+1}. The claim now follows by induction, from the initial conditions given. ☐

Unless otherwise stated, all orbits will be considered having *regular initial conditions.*

Definition 3.9. *We call* regular initial conditions*, those yielding* regular orbits*, and* boundary initial conditions*, those yielding* boundary orbits*.*

Remark 3.10. *As a result of these introductory remarks, not only $p_0 \notin \partial\Sigma$ and $q_0 \notin \partial\Sigma$ are regular initial conditions, but studying the system for such initial conditions is equivalent to studying it for all regular initial conditions.*

We conclude with a general property which will be exploited later on.

Remark 3.11. *If $p_{n+1} - p_n \longrightarrow \mathbf{0}$ as $n \longrightarrow \infty$ and p_n does not converge to any of the vertices, then there is a subsequence $\{p_{n_j}\}_{j\in\mathbb{N}}$ bounded away from V.*

Proof. By contradiction, if there is no such subsequence, since p_n does not converge to any of the vertices (by hypothesis) but any of its subsequences approaches the set of vertices V (by contradiction), we can extract two disjoint subsequences $\{p_{n_k}\}_{k\in\mathbb{N}}$ and $\{p_{n_l}\}_{l\in\mathbb{N}}$ from $\{p_n\}_{n\in\mathbb{N}}$, such that

$$\{p_n\}_{n\in\mathbb{N}} = \{p_{n_k}\}_{k\in\mathbb{N}} \cup \{p_{n_l}\}_{l\in\mathbb{N}},$$

$p_{n_k} \longrightarrow v_i$ for some $i \in \{1, 2, 3\}$ (by boundedness) and $p_{n_l} \longrightarrow V \setminus \{v_i\}$. Since p_n is either p_{n_k} for some k, or p_{n_l} for some l, there are infinitely many k and l_k such that $p_{n_{l_k}} = p_{n_k+1}$. For any ε fixed, by the hypothesis $p_{n+1} - p_n \longrightarrow \mathbf{0}$, for all k large enough $\|p_{n_{l_k}} - p_{n_k}\|_1 < \varepsilon$. But the 1-distance between $V \setminus \{v_i\}$ (which p_{n_l} approaches) and v_i (which p_{n_k} approaches) is 2. Since ε is arbitrary, we have a contradiction. ☐

3.2 Fixed points and potential function

From (2.37) and (2.38) it is immediate to derive the fixed point equations, which are

$$q = M_p q$$
$$p = \mathbf{1} - 2q,$$

and yield a dense set of equilibrium points $\{(p, q_p) : p \in \Sigma_0\}$, where

$$q_p := \frac{\mathbf{1} - p}{2}.$$

Verifying that $M_p q_p = q_p$ is a straightforward computation (see Lemma 3.19 (a)). The whole analysis of this system's asymptotic behaviour will revolve around the fixed points, which are of two types: internal equilibria ($p \notin \partial\Sigma$) and boundary equilibria ($p \in \partial\Sigma$). Under the assumptions made, orbits never get stuck at equilibria, as we show in the following remark.

Remark 3.12. *If $p_0 \in \Sigma_0$ and $q_{p_0} \neq q_0 \in \Sigma_0$, then for all $p \in \Sigma_0$, $(p_n, q_n) \neq (p, q_p)$ for all $n \in \mathbb{N}$.*

Proof. If $q_{n+1} = q_{p_{n+1}}$ for some $n \in \mathbb{N}_0$, then by rearranging (2.38) we obtain

$$q_n = \mathbf{1} - p_{n+1} - q_{n+1} = \mathbf{1} - p_{n+1} - q_{p_{n+1}} = q_{p_{n+1}}.$$

Since then $q_{p_{n+1}} = M_{p_n} q_{p_{n+1}}$, it follows that $q_{p_{n+1}}$ is an eigenvector for the eigenvalue 1 of M_{p_n}. Also, it is normalised with respect to the 1-norm and it is a nonnegative vector (all of its components are nonnegative). Since 1 is a simple eigenvalue for the matrix M_p for any $p \in \Sigma_0$, the corresponding eigenspace is the eigenline with direction q_p (see Lemma 3.19 for the computations relative to these elementary facts about the matrix). This implies that $p_n = p_{n+1}$. More precisely, rewriting as $\{tq_{p_n},\ t \in \mathbb{R}\}$ the eigenspace spanned by q_{p_n}, $q_{p_{n+1}} = M_{p_n} q_{p_{n+1}}$ if and only if $q_{p_{n+1}} = tq_{p_n}$, with $t \neq 0$, as $\|q_{p_{n+1}}\|_1 = \|q_{p_n}\|_1 = 1$. Taking norms yields $1 = \|q_{p_{n+1}}\|_1 = |t|\|q_{p_n}\|_1 = |t|$. Since $q^{(i)}_{p_{n+1}} \geq 0$ and $q^{(i)}_{p_n} \geq 0$ for all $i \in \{1, 2, 3\}$, we have that $t > 0$ and therefore $t = 1$, which results in $q_{p_{n+1}} = q_{p_n}$, which is equivalent to $p_n = p_{n+1}$.

Iterating this argument backwards implies that (p_0, q_0) is an equilibrium configuration (precisely, $(p_{n+1}, q_{p_{n+1}})$), a contradiction. □

Let $V(p, q) := \|q - q_p\|_1$. Since by (2.38) it holds that

$$q_{\hat{p}} := \frac{\mathbf{1} - \hat{p}}{2} = \frac{\hat{q} + q}{2}, \tag{3.1}$$

one can rewrite

$$\hat{q} - q_{\hat{p}} = \hat{q} - \frac{\hat{q} + q}{2} = \frac{\hat{q} - q}{2} = \frac{\hat{q} - q_p - (q - q_p)}{2} = \frac{M_p(q - q_p) - (q - q_p)}{2}.$$

Hence

$$\hat{q} - q_{\hat{p}} = \frac{M_p - I}{2}(q - q_p). \tag{3.2}$$

Denote $L_p := \frac{M_p - I}{2}$.

Remark 3.13. *Since $\|L_p\|_1 = 1$ for all $p \in \Sigma_0$, taking the norm on both sides of (3.2) yields $V(\hat{p}, \hat{q}) \leq V(p, q)$, and therefore the continuous nonnegative function $V(p, q)$ is nonincreasing along the orbits of the dynamical system, and defines a* Lyapunov potential function *for this system as a result. Moreover, since $V(p, q)$ is nonnegative and nonincreasing, it immediately follows that there is $0 \leq \ell \in \mathbb{R}$, dependent on the initial conditions, such that $V(p_n, q_n) \longrightarrow \ell$ as $n \longrightarrow \infty$.*

Loosely speaking, dynamical systems admitting a potential are often referred to as *gradient-like*. Note that in Chapter 1 we have already anticipated the existence of such limit also for the stochastic process, and we denoted it, for simplicity, with the same letter ℓ used in this section, for the deterministic dynamical system. Clearly, in that case ℓ is a random variable, while in this case it is deterministic: the fact that we are using the same notation does not mean that, if the initial conditions $(p_0, q_0) = (\Theta_0, \pi_0)$, then the limit ℓ in the two cases is the same; it simply means that both quantities have analogous roles in the convergence arguments for the dynamical system and the stochastic process, and this is regardless of initial conditions, as long as they are regular.

Lemma 3.14. *For every $p \notin \partial\Sigma$ and $q \neq q_p$,*

$$V(\hat{p}, \hat{q}) < V(p, q).$$

Therefore, the potential is eventually strictly decreasing along the regular orbits of the dynamical system.

Proof. If $q = q_p$, then $\hat{q} = q = q_p$ and $\hat{p} = p$. Therefore, $V(\hat{p}, \hat{q}) = V(p, q) = 0$. Assume $q \neq q_p$. By (3.2), $V(\hat{p}, \hat{q}) < V(p, q)$ if and only if

$$\left\| L_p \frac{q - q_p}{\|q - q_p\|_1} \right\|_1 < 1.$$

To show this we will consider the action of L_p on the intersection of the 1-norm unit sphere $\mathbb{S}^2_1$ with the plane Π_0 of Cartesian equation $x + y + z = 0$, where

$$v := \frac{q - q_p}{\|q - q_p\|_1}$$

lies. In particular it is enough to show that for every $p \notin \partial\Sigma$ and $v \in \mathbb{S}^2_1 \cap \Pi_0$ one has $\|L_p v\|_1 < 1$. Note that (see Figure 3.1) the intersection aforementioned is a regular hexagon with set of vertices

$$\left\{ \left(0, \frac{1}{2}, -\frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}, 0\right), \left(-\frac{1}{2}, 0, \frac{1}{2}\right), \left(0, -\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, -\frac{1}{2}, 0\right), \left(\frac{1}{2}, 0, -\frac{1}{2}\right) \right\}.$$

As a consequence of the linearity of L_p, the claim is proved if one can show that each of these vertices is shrunk strictly inside the hexagonal portion of the plane. By symmetry, without loss of generality, this can be shown by performing the calculation explicitly for the vertex $v = (1/2, 0, -1/2)$. First compute

$$L_p v = \frac{1}{4} \begin{pmatrix} -1 - \frac{p_2}{p_1 + p_2} \\ \frac{p_3}{p_3 + p_2} - \frac{p_1}{p_1 + p_2} \\ 1 + \frac{p_2}{p_2 + p_3} \end{pmatrix}$$

and then consider that since $p \notin \partial\Sigma$,

$$\begin{aligned} \|L_p v\|_1 &= \frac{1}{4} \left(2 + \frac{p_2}{p_1 + p_2} + \frac{p_2}{p_2 + p_3} + \left| \frac{p_3}{p_3 + p_2} - \frac{p_1}{p_1 + p_2} \right| \right) \\ &= \begin{cases} \frac{1}{2} \left(1 + \frac{p_2}{p_1 + p_2} \right) < 1, & \frac{p_3}{p_3 + p_2} \geq \frac{p_1}{p_1 + p_2} \\ \frac{1}{2} \left(1 + \frac{p_2}{p_2 + p_3} \right) < 1, & \frac{p_3}{p_3 + p_2} < \frac{p_1}{p_1 + p_2} \end{cases} \end{aligned} \tag{3.3}$$

because of the fact that none of the coordinates of $p \notin \partial\Sigma$ is zero. Note that since by Remark 3.8 $p_n \notin \partial\Sigma$ for all $n \in \mathbb{N}_0$ and by Remark 3.12 $q_n \neq q_{p_n}$, we can conclude that the potential is strictly decreasing along the orbits (that is, V is a strict Lyapunov potential function for the dynamical system). □

Overall there are three mutually exclusive asymptotic scenarios, in each of which it is possible to prove that $\{p_n\}$ converges (whereas $\{q_n\}$ may or may not):

- $\{p_n\}$ is bounded away from the boundary (Section 3.3);
- $\ell = 0$ and $\{p_n\}$ is not bounded away from the boundary (Section 3.4);
- $\ell > 0$ and $\{p_n\}$ is not bounded away from the boundary (Section 3.5).

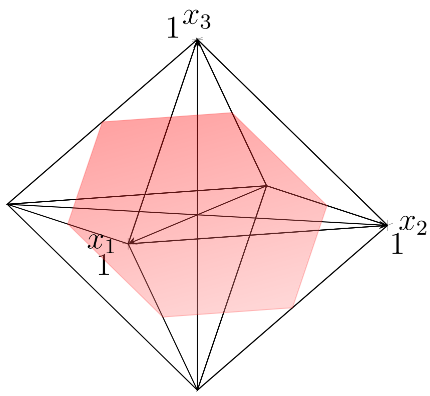


Figure 3.1: $\mathbb{S}_1^2 \cap \Pi_0$

3.3 Convergence bounded away from the boundary

The main goal of this section is showing convergence of the dynamical system when $\{p_n\}$ is known to be bounded away from the boundary of the simplex.

Proposition 3.15. *If $\{p_n\}$ is bounded away from $\partial\Sigma$, there is a constant $0 < \not{c} < 1$, dependent on the initial conditions, such that*

$$V(p_{n+1}, q_{n+1}) < \not{c}\, V(p_n, q_n).$$

Hence

$$\ell := \lim_{n \longrightarrow \infty} V(p_n, q_n) = 0,$$

and the dynamical system converges to an internal equilibrium.

Proof. By p_n *bounded away from* $\partial\Sigma$ it is meant that for all n, $p_n \in \Sigma_\varepsilon$, for some $\varepsilon > 0$ small enough, where the compact $\Sigma_\varepsilon \subset \Sigma$ is defined as $\Sigma_\varepsilon := \{x \in \Sigma : \ x_i \geq \varepsilon, \ \forall\, i \in \{1, 2, 3\}\}$. Note that the definition of Σ_ε is one of those cases, mentioned in the introduction, in which the notation for the components switches from upper to lower index, in absence of time index.

If $\{p_n\}$ is bounded away from the boundary then the functions $\phi_{i,j}(p_n) := \phi(p_n^{(i)}, p_n^{(j)})$ will be bounded away from 1 (since the value 1 and 0 are attained only at boundary points, as we saw in deriving (3.3)), so the constant $\not{c}$, applying uniformly on Σ_ε, can be found by upper-bounding (3.3), rather than with 1, with

$$\not{c} := \max_{i \neq j \in \{1,2,3\}} \max_{p \in \Sigma_\varepsilon} \phi(p_i, p_j)$$

which is well defined for every $\varepsilon > 0$ small enough, by the continuity of ϕ and the compactness of Σ_ε. From the geometric decaying upper bound on $V(p_n, q_n)$ it follows that

$$\ell := \lim_{n \longrightarrow \infty} V(p_n, q_n) = 0.$$

To show convergence of the dynamical system to an internal equilibrium, first of all observe that

$$\hat{p} - p = 2q_p - q - \hat{q} = M_p(q_p - q) + q_p - q$$

and therefore, since $\|M_p\|_1 = 1$, it follows that

$$\|\hat{p} - p\|_1 \leq 2V(p, q). \tag{3.4}$$

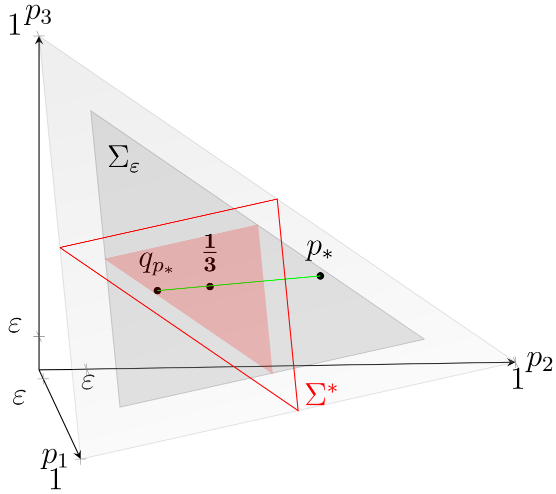


Figure 3.2: Internal equilibria

The geometric decay of $V(p_n, q_n)$ ensured by $\{p_n\}$ being bounded away from $\partial\Sigma$ implies that

$$\sum_{n=0}^{\infty} V(p_n, q_n) < \infty.$$

Then

$$\sum_{n=0}^{\infty} \|p_{n+1} - p_n\|_1 < \infty$$

and the convergence of $\{p_n\}$ follows immediately. Since $\ell = 0$, the convergence of p_n to some limit p_* bounded away from the boundary, implies the convergence of q_n to q_{p_*} bounded away from the boundary (as they belong to Σ^*, more precisely to the transformation of Σ_ε through the homothety $(\mathbf{1} - p)/2$, see Figure 3.2). □

Remark 3.16. *Clearly the same would hold if* $\{p_n\}$ *could only eventually (that is for all* n *large enough only) be bounded away from the boundary. The only difference would be that the constant* $0 < c < 1$ *such that* $V(p_{n+1}, q_{n+1}) < cV(p_n, q_n)$ *holds only eventually, which is enough to yield convergence of* $\{(p_n, q_n)\}$. *However, note that by Remarks 3.4 and 3.6 we cannot have such case, where* $p_0 \in \partial\Sigma$ *and at some later time* n, $p_n \notin \partial\Sigma$.

The following proposition and corollary show that internal equilibria are stable, implying that convergence bounded away from the boundary can also be ensured with suitable initial conditions, close enough to one of the internal equilibria. From a technical standpoint, these last two results will be mainly relied upon in Section 3.4, specifically Lemma 3.24. More in general, they provide a deeper understanding of the dynamical system.

Denote by $B(p, r) \subset \mathbb{R}^3$ the ball centered at p having radius r in the distance dist generated by the 1-norm. Define $U((p, q), r, r') := B(p, r) \times B(q, r')$ and $U((p, q), r) := U((p, q), r, r)$.

Proposition 3.17. *For every* $p \notin \partial\Sigma$ *and a small enough* $0 < \varepsilon' < \mathrm{dist}(p, \partial\Sigma)$, *there is a* $\delta' > 0$ *suitably smaller than* ε' *such that, if* $(p_0, q_0) \in U((p, q_p), \delta')$ *then* $(p_n, q_n) \in U((p, q_p), \varepsilon', {}^{\varepsilon'}\!/2)$ *for all* $n \in \mathbb{N}$.

Proof. Given p and ε' as in the statement, let $0 < \delta' < \varepsilon'$ (it will be further restricted if necessary). By Proposition 3.15 and (3.2) it is known that since $B(p, \varepsilon')$ is bounded away from $\partial\Sigma$, for all $(p', q') \in U((p, q_p), \varepsilon')$, $\|L_{p'}(q - q_{p'})\|_1 \leq \mathcal{c}\|q - q_{p'}\|_1$ for some $0 < \mathcal{c} < 1$. Further restrict $\delta' < (1 - \mathcal{c})\varepsilon'/4$, and consider $(p_0, q_0) \in U((p, q_p), \delta')$. The following claim will be shown by induction: for all $n \in \mathbb{N}$, $p_n \in B(p, \varepsilon')$ and $q_n \in B(q_p, \varepsilon'/2)$. Consider that

$$\|q_0 - q_{p_0}\|_1 \leq \|q_0 - q_p\|_1 + \|q_p - q_{p_0}\|_1 = \|q_0 - q_p\|_1 + \frac{\|p - p_0\|_1}{2} < \frac{3}{2}\delta'.$$

Therefore

$$\|q_1 - q_{p_1}\|_1 = \|L_{p_0}(q_0 - q_{p_0})\|_1 \leq \mathcal{c}\|q_0 - q_{p_0}\|_1 = \frac{3}{2}\mathcal{c}\delta'.$$

Recall that for all n, $\|p_{n+1} - p_n\|_1 \leq 2\|q_n - q_{p_n}\|_1$ by (3.4), hence $\|p_1 - p_0\|_1 < 3\delta'$. As a consequence, noting that $q_{p_1} - q_p = (p - p_1)/2$,

$$\begin{aligned}
&\|p_1 - p\|_1 \leq \|p_1 - p_0\|_1 + \|p_0 - p\|_1 < 4\delta' < (1 - \mathcal{c})\varepsilon' < \varepsilon' \\
&\|q_1 - q_p\|_1 \leq \|q_1 - q_{p_1}\|_1 + \|q_{p_1} - q_p\|_1 \leq \frac{3}{2}\mathcal{c}\delta' + \frac{\|p_1 - p\|_1}{2} \leq \frac{3}{2}\mathcal{c}\delta' + 2\delta' < 2\delta'(1 + \mathcal{c}) \\
&\qquad < (1 - \mathcal{c})(1 + \mathcal{c})\frac{\varepsilon'}{2} = (1 - \mathcal{c}^2)\frac{\varepsilon'}{2} < \frac{\varepsilon'}{2}.
\end{aligned}$$

This immediately implies that we can use again the geometric decay of the potential:

$$\|q_2 - q_{p_2}\|_1 = \|L_{p_1}(q_1 - q_{p_1})\|_1 \leq \mathcal{c}\|q_1 - q_{p_1}\|_1 \leq \mathcal{c}^2\|q_0 - q_{p_0}\|_1 < \frac{3}{2}\mathcal{c}^2\delta'.$$

For $n = 1$, it has been shown that if $(p_0, q_0) \in U((p, q_p), \delta')$, where $\delta' < (1 - \mathcal{c})\varepsilon'/4$, with $\mathcal{c}$ being the subunitary constant uniformly holding on $B(p, \varepsilon')$, then $\|q_1 - q_{p_1}\|_1 < 3\mathcal{c}\delta'/2$, $\|p_1 - p\|_1 < 4\delta'$, $p_1 \in B(p, \varepsilon')$, $\|q_1 - q_p\|_1 \leq 3\mathcal{c}\delta'/2 + 2\delta' < 2\delta'(1 + \mathcal{c})$, $q_1 \in B(q_p, \varepsilon'/2)$. Assume as induction hypothesis that

$$\|q_n - q_{p_n}\|_1 < \frac{3}{2}\mathcal{c}^n\delta',$$

that

$$\|p_n - p\|_1 < 4\delta'\sum_{i=0}^{n-1}\mathcal{c}^i,$$

so that $p_n \in B(p, \varepsilon')$, and

$$\|q_n - q_p\|_1 < 2\delta'\sum_{i=0}^{n}\mathcal{c}^i,$$

so that $q_n \in B(q_p, \varepsilon'/2)$, and consider p_{n+1}. Since

$$\begin{aligned}
\|p_{n+1} - p\|_1 \leq \|p_{n+1} - p_n\|_1 + \|p_n - p\|_1 \leq 2\|q_n - q_{p_n}\|_1 + \|p_n - p\|_1 < \\
3\mathcal{c}^n\delta' + 4\delta'\sum_{i=0}^{n-1}\mathcal{c}^i < 4\delta'\sum_{i=0}^{n}\mathcal{c}^i,
\end{aligned}$$

this shows that $p_{n+1} \in B(p, \varepsilon')$ since

$$4\delta'\sum_{i=0}^{n}\mathcal{c}^i < 4\delta'\sum_{i=0}^{\infty}\mathcal{c}^i = 4\frac{\delta'}{1 - \mathcal{c}} < \varepsilon'$$

by hypothesis, and therefore it also follows that

$$\|q_{n+1} - q_{p_{n+1}}\|_1 < c\|q_n - q_{p_n}\|_1 < \frac{3}{2}c^{n+1}\delta'.$$

Since

$$\begin{aligned}\|q_{n+1} - q_p\|_1 \leq \|q_{n+1} - q_{p_{n+1}}\|_1 + \|q_{p_{n+1}} - q_p\|_1 &\leq \frac{3}{2}c^{n+1}\delta' + \frac{\|p_{n+1} - p\|_1}{2} \\ &\leq \frac{3}{2}c^{n+1}\delta' + 2\delta' \sum_{i=0}^{n} c^i < 2\delta' \sum_{i=0}^{n+1} c^i,\end{aligned}$$

this shows that $q_{n+1} \in B(q_p, \varepsilon'/2)$ since

$$2\delta' \sum_{i=0}^{n+1} c^i < 2\delta' \sum_{i=0}^{\infty} c^i = 2\frac{\delta'}{1-c} < \frac{\varepsilon'}{2}$$

by hypothesis. □

By Proposition 3.17, once the system is confined in such a neighbourhood of some internal equilibrium (p, q_p) bounded away from the boundary, the final convergence claim in Proposition 3.15 can be enacted, immediately yielding the following.

Corollary 3.18. *If (p_0, q_0) is close enough to an internal equilibrium (p, q_p), the system converges to a (possibly different) internal equilibrium.*

Proposition 3.17 is not strong enough to ensure that the limit is the same (p, q_p), by the density of the set $\{(p, q_p) : p \in \Sigma\} \subset \Sigma^2$.

3.4 Convergence to the boundary with $\ell = 0$

The main goal of this section is to show that if $\{p_n\}$ approaches the boundary and the limit of the potential $\ell := \lim_{n \longrightarrow \infty} V(p_n, q_n) = 0$, the dynamical system converges.

First of all note that by the remarks made at the opening of this chapter, $\ell = 0$ is not possible for any type of boundary orbit having q_n entering a 2-cycle. Since all other boundary cases are trivial, we can work under regular initial conditions as usual. In this case, assuming that there is no convergence to the vertex (if it were, there would be nothing to prove, since p_n would tend to a vertex $v \in V$ and q_n to q_v due to $\ell = 0$), by Remark 3.11 (which applies by (3.4) and $\ell = 0$) there will be a subsequence $\{p_{n_j}\}_{j \in \mathbb{N}}$ bounded away from the vertices and, by assumption, not bounded away from the boundary. Extracting a convergent subsubsequence $\{p_{n_{j_l}}\}_{l \in \mathbb{N}}$ by boundedness of $\{p_n\}$, relabelling it with $\{n_k\}$, we can assume that there is a subsequence $\{p_{n_k}\}_{k \in \mathbb{N}}$ bounded away from the vertices and convergent to a point in an edge. By symmetry, without loss of generality, assume this subsequence to converge to $p_* \in E_1$, with $\delta/2 \leq p_*^{(2)} \leq 1 - \delta/2$ for some $\delta > 0$ small enough fixed. Then $p_{n_k}^{(1)} \longrightarrow 0$ as $k \longrightarrow \infty$ and $p_{n_k}^{(2)}$ is eventually bounded away from 0 and 1. In this section we will look at $q_n - q_{p_n} \in \Pi_0$ through a change to eigencoordinates derived from the two eigenvectors of the matrix M_{p_n} that span Π_0. Since $\|q_n - q_{p_n}\|_1 \longrightarrow 0$, there will be a small enough $\delta > 0$ fixed and an $\varepsilon > 0$ arbitrarily small, dependent on δ, such that for some large enough $K \in \mathbb{N}$, for all $k > K$,

$$(p_{n_k}, q_{n_k}) \in \mathscr{K}^*_{\varepsilon, \frac{\delta}{2}} := \left\{(p, q) \in \Sigma^2 : \ 0 < p_1 \leq \varepsilon, \ \frac{\delta}{2} \leq p_2 \leq 1 - \frac{\delta}{2}, 0 < |\alpha|, |\beta| \leq \varepsilon\right\}$$

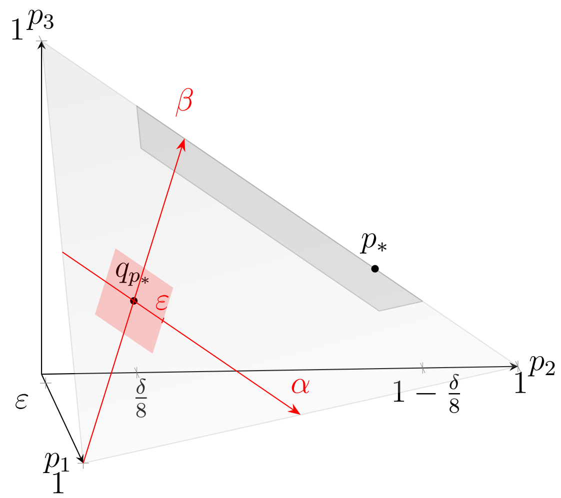


Figure 3.3: $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$ as the Cartesian product of the dark gray and the red regions

because both eigencoordinates of $q_n - q_{p_n}$, which are denoted as α_n, β_n, will vanish as $n \longrightarrow \infty$, and therefore eventually $|\alpha_n| \in \{0 < \alpha \leq \varepsilon,\ \alpha \in \mathbb{R}\}$ and $|\beta_n| \in \{0 < \beta \leq \varepsilon,\ \beta \in \mathbb{R}\}$ for any arbitrary $\varepsilon > 0$ (we will define precisely the eigensystem for these eigencoordinates in Lemma 3.19). We will also make use of a similarly defined set $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$ and we will also need the set

$$\mathscr{K}_{\varepsilon,\frac{\delta}{8}} := \left\{ p \in \Sigma :\ 0 < p_1 \leq \varepsilon,\ \frac{\delta}{8} \leq p_2 \leq 1 - \frac{\delta}{8} \right\}.$$

When it does not create confusion, ε or $^{\delta}/_{8}$ or both, will be dropped from the notation referring to these two sets. We will also adopt the following notation.

- $f(p) = \mathcal{O}(g(p_1))$ if for ε small enough $f(p)/g(p_1)$ is well defined and bounded in $\mathscr{K}_{\varepsilon,\frac{\delta}{8}}$.
- $f(p, \alpha, \beta) = \mathcal{O}(g_1(p_1, \alpha, \beta), \ldots, g_k(p_1, \alpha, \beta))$ if for a sufficiently small ε

  $$\frac{f(p_1, p_2, 1 - p_1 - p_2, \alpha, \beta)}{|g_1(p_1, \alpha, \beta)| + \ldots + |g_k(p_1, \alpha, \beta)|}$$

  is well-defined and bounded on $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$.
- $r := (p_1, \alpha, \beta)$.

Note that from the given definition, assuming well-definedness whenever necessary, the usual rules hold.

- $\mathcal{O}(f(r)) \pm \mathcal{O}(g(r)) = \mathcal{O}(f(r), g(r))$.
- $\mathcal{O}(f(r))\mathcal{O}(g(r)) = \mathcal{O}(f(r)g(r))$.
- if $f(r) = \mathcal{O}(g(r))$ with $g(r) = \mathcal{O}(h(r))$, then $f(r) = \mathcal{O}(h(r))$.
- if $f(r) = \mathcal{O}(g(r), h(r))$ with with $g(r) = \mathcal{O}(m(r))$, then $f(r) = \mathcal{O}(m(r), h(r))$, and so on.

The following lemma shows all the elementary properties of the matrix M_p that will be useful.

Lemma 3.19. *Let $p = (p_1, p_2, p_3) \in \Sigma_0$.*

a) M_p has eigenvalue 1 with eigenvector q_p;

b) M_p has two further real eigenvalues $-1 \leq \lambda_{-1}(p) \leq \lambda_0(p) \leq 0$;

c) M_p is invariant on the plane Π_0 in $\mathbb{R}^3$ having equation $x + y + z = 0$;

d) If $p \neq \frac{\mathbf{1}}{\mathbf{3}}$ then $\lambda_{-1}(p) \neq \lambda_0(p)$ and the corresponding eigenvectors $e_{-1}(p)$, $e_0(p)$ lie in Π_0;

e) If $p = \frac{\mathbf{1}}{\mathbf{3}}$ then $\lambda_{-1}(p) = \lambda_0(p) = \lambda := -\frac{1}{2}$, having geometric multiplicity 2 and eigenspace Π_0;

f) If $p \in \partial\Sigma \setminus V$ then $\lambda_{-1}(p) = -1$ and $\lambda_0(p) = 0$. In particular if $p_1 = 0$ then the corresponding eigenvectors can be chosen to be $e_{-1}(p) = (-1, 1 - p_2, p_2)$ and $e_0(p) = (0, 1, -1)$;

g) $\lambda_0(p) = -2p_1 + \mathcal{O}(p_1^2)$ and $\lambda_{-1}(p) = -1 + 2p_1 + \mathcal{O}(p_1^2)$ as $p_1 \longrightarrow 0$;

h) $e_{-1}(p)$ and $e_0(p)$ can be chosen to depend smoothly on p and having the norm bounded away from zero on $\mathscr{K}_{\varepsilon,\frac{\delta}{8}}$ for ε small enough.

Proof.

a) Each column of M_p adds up to 1, and therefore its transpose M_p' has rows adding up to 1, which means that M_p' has eigenvalue 1 with eigenvector $\mathbf{1}$. Thus M_p has an eigenvalue of 1, since a matrix and its transpose have the same spectrum. By inspection it is easy to guess that q_p is the eigenvector of the eigenvalue 1 for M_p. This is equivalent to calculating $M_p q_p = q_p$ (*a posteriori* the calculation makes finding first the eigenvalue of 1 redundant). By symmetry, it is enough to show the calculation for the first component only, without loss of generality:

$$\begin{aligned}(M_p q_p)^{(1)} &= \frac{p_3}{p_1 + p_3}\frac{1 - p_2}{2} + \frac{p_2}{p_1 + p_2}\frac{1 - p_3}{2} = \frac{p_3}{1 - p_2}\frac{1 - p_2}{2} + \frac{p_2}{1 - p_3}\frac{1 - p_3}{2} \\ &= \frac{p_3}{2} + \frac{p_2}{2} = \frac{1 - p_1}{2} = q_p^{(1)}.\end{aligned}$$

b) Note that the *trace* $\operatorname{tr}(M_p) = 0$ and the *determinant*

$$\det(M_p) = 2\frac{p_1 p_2 p_3}{(p_1 + p_3)(p_1 + p_2)(p_2 + p_3)},$$

thus from (a) one can conclude that $\lambda_1 + \lambda_2 = -1$ and that

$$\lambda_1 \lambda_2 = \frac{2p_1 p_2 p_3}{(p_1 + p_3)(p_1 + p_2)(p_2 + p_3)}.$$

From (a) it is also already known that the characteristic polynomial $p_{M_p}(\lambda)$ is a monic cubic having a factor $\lambda - 1$, hence it is of the form $(\lambda - 1)q_{M_p}(\lambda)$, where $q_{M_p}(\lambda)$ is a monic (otherwise the product would not be monic) quadratic, having roots λ_1 and λ_2, and therefore

$$q_{M_p}(\lambda) = \lambda^2 + \lambda + \frac{2p_1 p_2 p_3}{(p_1 + p_3)(p_1 + p_2)(p_2 + p_3)}.$$

As a result the discriminant of q_{M_p} is

$$1 - \frac{8p_1 p_2 p_3}{(p_1 + p_3)(p_1 + p_2)(p_2 + p_3)},$$

hence the remaining two eigenvalues will be real if and only if it is shown that for all $p_1,\ p_2,\ p_3 \geq 0$ such that $p_1 + p_2 + p_3 = 1$,

$$\frac{p_1 p_2 p_3}{(p_1 + p_3)(p_1 + p_2)(p_2 + p_3)} \leq \frac{1}{8}.$$

Since then none of the p_i can be zero, under the standard simplex' constraint this inequality is equivalent to

$$\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} \geq 9. \tag{3.5}$$

This can be seen by noting that on the simplex

$$\frac{p_1 p_2 p_3}{(p_1 + p_3)(p_1 + p_2)(p_2 + p_3)} = \frac{p_1 p_2 p_3}{(1 - p_1)(1 - p_2)(1 - p_3)}$$
$$= \frac{1}{\left(\frac{1}{p_1} - 1\right)\left(\frac{1}{p_2} - 1\right)\left(\frac{1}{p_3} - 1\right)},$$

and then taking the reciprocal yields

$$\left(\frac{1}{p_1} - 1\right)\left(\frac{1}{p_2} - 1\right)\left(\frac{1}{p_3} - 1\right) \geq 8.$$

Multiplying out and rearranging yields

$$\frac{1}{p_1 p_2 p_3} - \frac{1}{p_1 p_2} - \frac{1}{p_2 p_3} - \frac{1}{p_1 p_3} + \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} \geq 9.$$

Under the simplex' constraint

$$\frac{1}{p_1 p_2 p_3} - \frac{1}{p_1 p_2} - \frac{1}{p_2 p_3} - \frac{1}{p_1 p_3} = \frac{1 - p_1 - p_2 - p_3}{p_1 p_2 p_3} = 0,$$

hence the claim. Applying the Lagrange multipliers method (denoting the multiplier with m) to the left-hand side of (3.5) under the simplex constraint, yields $-p_i^{-2} = m$ for all $i \in \{1, 2, 3\}$, since the gradient of the constraint is $\mathbf{1}$. Hence $p_1 = p_2 = p_3 = 1/\sqrt{-m}$, which yields $p_i = 1/3$ for all $i \in \{1, 2, 3\}$ under the simplex constraints. This is a minimum (attaining a value of 9) since the function $p_1^{-1} + p_2^{-1} + p_3^{-1}$ has positive definite Hessian

$$\begin{pmatrix} \frac{2}{p_1^3} & 0 & 0 \\ 0 & \frac{2}{p_2^3} & 0 \\ 0 & 0 & \frac{2}{p_3^3} \end{pmatrix}$$

for all p inside the simplex, and therefore it is convex on the interior of the standard simplex. To conclude there is only one such point at which the inequality becomes equality, the equilibrium of the simplex $\frac{\mathbf{1}}{\mathbf{3}}$. On the boundary $\partial\Sigma_0$ this does not happen since the original expression has the numerator $p_1 p_2 p_3 = 0$, hence the expression is zero, which is obviously less than $1/8$. Hence the eigenvalues λ_1 and λ_2 are always real and distinct for all $p \neq \frac{\mathbf{1}}{\mathbf{3}}$, while they are equal to $-1/2$ at $\frac{\mathbf{1}}{\mathbf{3}}$. This proves then the first part of (d) and (e). Also, since $\lambda_1 + \lambda_2 = -1$ and $\lambda_1 \lambda_2 \geq 0$, it follows that they are both nonpositive and in particular $\lambda_1, \lambda_2 \in [-1, 0]$. For the inequality see (g).

c) Since M'_p has eigenvalue 1 with eigenvector $\mathbf{1}$, the span of the eigenvectors corresponding to λ_1 and λ_2 is the orthogonal complement of the span of $\mathbf{1}$, that is Π_0. Hence M_p is invariant on Π_0. This also proves the second part of (d) and (e).

d) Already partially proved in (b) and (c).

e) Since

$$M_{\frac{1}{3}} = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

the eigenspace equation

$$M_{\frac{\mathbf{1}}{\mathbf{3}}} v = -\frac{1}{2} v$$

is equivalent to the equation $v_1 + v_2 + v_3 = 0$, hence the geometric multiplicity of $-1/2$ is 2.

f) Assume $p \in E_1$, then $p_1 = 0$, thus

$$M_p = \begin{pmatrix} 0 & 1 & 1 \\ 1 - p_2 & 0 & 0 \\ p_2 & 0 & 0 \end{pmatrix}.$$

From (b) it is also known that $\lambda_1 + \lambda_2 = -1$ and $\lambda_1 \lambda_2 = 0$, which implies that $\lambda_{-1}(p) := \lambda_1 = -1$ and $\lambda_0(p) := \lambda_2 = 0$. The eigenspace equations are then:

- $M_p v = 0$, which yields $v_1 = 0$ and $v_2 = -v_3$, which yields the eigenvector of $\lambda_0(p)$, $e_0(p) = (0, 1, -1)$;
- $M_p v = -v$, which yields $(1-a)v_1 = -v_2$ and $a v_1 = -v_3$ (the third equation $v_1 + v_2 + v_3 = 0$ is obtained from those two by addition hence it is omitted). Fixing $v_1 = -1$ yields the eigenvector of $\lambda_{-1}(p)$, $e_{-1}(p) = (-1, 1 - p_2, p_2)$.

g) We derive the smallest eigenvalue first, by directly solving the characteristic polynomial, that is the quadratic $q_{M_p}(\lambda)$, for $p_1 \in (0, \varepsilon]$ with ε arbitrarily small, and $p_2 \in [\delta/8, 1 - \delta/8]$ with $\delta > 0$ small enough fixed. From

$$q_{M_p}(\lambda) = \lambda^2 + \lambda + \frac{2p_1 p_2 (1 - p_1 - p_2)}{(1 - p_1)(1 - p_2)(p_1 + p_2)}$$

and expanding in Taylor series the factor $(1 - p_1)^{-1} = 1 + \mathcal{O}(p_1)$, it follows that

$$\frac{2p_1 p_2 (1 - p_1 - p_2)}{(1 - p_1)(1 - p_2)(p_1 + p_2)} = \frac{2p_1 p_2}{(1 - p_2)(p_1 + p_2)} - \frac{2p_1 p_2^2}{(1 - p_2)(p_1 + p_2)} + \mathcal{O}(p_1^2)$$
$$= 2p_1 p_2 \left(\frac{1}{(1 - p_2)(p_1 + p_2)} - \frac{p_2}{(1 - p_2)(p_1 + p_2)} \right) + \mathcal{O}(p_1^2) = \frac{2p_1 p_2}{p_1 + p_2} + \mathcal{O}(p_1^2)$$
$$= \frac{2p_1 (p_2 + p_1 - p_1)}{p_1 + p_2} + \mathcal{O}(p_1^2) = 2p_1 + \mathcal{O}(p_1^2),$$

and therefore the eigenvalues are solutions of $\lambda^2 + \lambda + 2p_1 + \mathcal{O}(p_1^2)$, yielding

$$\lambda_0(p) = \frac{-1 + \sqrt{1 - 8p_1 + \mathcal{O}(p_1^2)}}{2}.$$

Expanding in Taylor series the square root yields

$$\lambda_0(p) = \frac{-1 + 1 - 4p_1 + \mathcal{O}(p_1^2)}{2} = -2p_1 + \mathcal{O}(p_1^2).$$

Since $\lambda_{-1}(p) = -1 - \lambda_0(p) = -1 - 2p_1 + \mathcal{O}(p_1^2)$, the result follows. Note that the calculation also shows that $\lambda_0(p) > \lambda_{-1}(p)$ away from $\frac{\mathbf{1}}{\mathbf{3}}$.

h) By a standard application of the *Implicit Function Theorem*, it is possible to show that simple eigenvalues and eigenvectors of M_p (under any smooth normalisation condition given) depend smoothly on the parameter $p \in \mathscr{K}_{\varepsilon,\frac{\delta}{8}}$, because the term $\phi(p_i, p_j)$ in the matrix varies smoothly for $p \in \Sigma_0$, and p is bounded away from $\mathbf{\frac{1}{3}}$ (since $p \in \mathscr{K}_{\varepsilon,\frac{\delta}{8}}$, it follows that $p_1 \in (0, \varepsilon]$ for ε sufficiently small to be determined and δ fixed small enough). For an ε small enough, we show the claim for each eigenpair, by considering all p close enough to an arbitrary $p^* \in E_1 \cap \overline{\mathscr{K}}$, where from now on we denote $\mathscr{K} := \mathscr{K}_{\varepsilon,\frac{\delta}{8}}$. The eigenpair considered will be either the one with eigenvalue approaching 0 (denoted with index $i = 1$) or the one approaching -1 (denoted with index $i = 2$) on the edge (which is E_1 as always, without loss of generality) as $p \longrightarrow p^*$, since there is nothing to prove for the eigenpair with eigenvalue of 1. This relabelling of the eigenpairs is only adopted in this argument, for the sake of consistency with the expression of their norm, as it will soon be clear. Thus for $i \in \{1, 2\}$, the eigenpair will be denoted $(\lambda(p), e(p)) := (\lambda_i(p), e_i(p))$, with

$$\lambda_i(p) = \begin{cases} \lambda_0(p), & i = 1 \\ \lambda_{-1}(p), & i = 2 \end{cases},$$

and

$$e_i(p) = \begin{cases} e_0(p), & i = 1 \\ e_{-1}(p), & i = 2 \end{cases}.$$

Since i is fixed, and the proof is the same for both indices, upon the relabelling performed, the index will be omitted in the notation of λ and e from now on. Consider that

$$M_{p^*} e(p^*) = \lambda(p^*) e(p^*)$$

with $\|e(p^*)\|_2$ bounded away from 0 uniformly on $E_1 \cap \overline{\mathscr{K}}$. In fact on the edge, we adopt an indexing consistent with the one established for $p \in \mathscr{K}$, that is with $i = 1$ denoting the eigenpair relative to 0, $i = 2$ that relative to -1. On the boundary the representation chosen in (f) always yields

$$\|e_0(p^*)\|_2^2 \equiv 2 \equiv 2[1 - p_1^* + (p_1^*)^2]$$

for $i = 1$ (trivially due to $p_1^* = 0$; it will be clear in a moment that the reason for this trivial expression is simply to write one proof that works for both $i \in \{1, 2\}$ without changes) and

$$\|e_{-1}(p^*)\|_2^2 = 2[1 - p_2^* + (p_2^*)^2] > 2\frac{\delta}{8}\left(1 + \frac{\delta}{8}\right)$$

for $i = 2$. Then we need to solve for

$$\begin{aligned} &M_p e(p) = \lambda(p) e(p) \\ &e(p^*)' e(p) = e(p^*) \cdot e(p) = 2[1 - p_i^* + (p_i^*)^2], \end{aligned}$$

where $'$ in the normalisation condition denotes the transpose, having written the inner product via matrix multiplication. Denote

$$F(e, \lambda, p) := \begin{pmatrix} M_p e - \lambda e \\ e(p^*) \cdot e - 2[1 - p_i^* + (p_i^*)^2] \end{pmatrix}.$$

The problem of showing smoothness can be rephrased as finding a smooth function $(e(p), \lambda(p))$ such that $F(e(p), \lambda(p), p) = \mathbf{0}$. F is smooth, so we calculate the Jacobian

$$\frac{\partial F}{\partial(e,\lambda)}(e,\lambda,p) = \begin{pmatrix} M_p - \lambda I & -e \\ e(p^*)' & 0 \end{pmatrix}$$

and denote

$$J_{p^*} := \frac{\partial F}{\partial(e,\lambda)}(e,\lambda,p)\bigg|_{p=p^*} = \begin{pmatrix} M_{p^*} - \lambda(p^*)I & -e(p^*) \\ e(p^*)' & 0 \end{pmatrix}.$$

We show that J_{p^*} is invertible, by showing that its kernel is null. This follows by contradiction. Let the kernel's equation $J_{p^*}W = \mathbf{0}$ hold for a vector $\mathbf{0} \neq W = (E, \Lambda) \in \mathbb{R}^{3+1}$ (there is no risk of ambiguity with E, the set of edges for the simplex in this context). The equation is equivalent to

$$\begin{aligned} &M_{p^*}E - \lambda(p^*)E - \Lambda e(p^*) = \mathbf{0} \\ &e(p^*) \cdot E = 0. \end{aligned}$$

If $E \neq \mathbf{0}$, the second condition implies orthogonality, and thus linear independence, of E and $e(p^*)$. Thus we can rewrite M_{p^*} in a new basis $\mathscr{B}$ (there is no risk of ambiguity with $\mathscr{B}$, the event when the stochastic process is bounded away from the boundary of the simplex) whose first two elements are $e(p^*)$ and E, in this order. Since $M_{p^*}e(p^*) = \lambda(p^*)e(p^*)$ and $M_{p^*}E = \Lambda e(p^*) + \lambda(p^*)E$,

$$[M_{p^*}]_{\mathscr{B}} = \begin{pmatrix} \lambda(p^*) & \Lambda & * & \dots & * \\ 0 & \lambda(p^*) & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & \vdots & & \vdots & \\ 0 & 0 & * & \dots & * \end{pmatrix}$$

and therefore, from the determinant formula of block-upper-triangular matrices, it follows that $\lambda(p^*)$ is at least a double root of the characteristic polynomial of M_{p^*}. However it is known that it is a simple eigenvalue, hence a contradiction. So $E = \mathbf{0}$. But then $\Lambda e(p^*) = \mathbf{0}$, which implies $\Lambda = 0$. Hence $W = \mathbf{0}$ and therefore J_{p^*} is invertible. The invertibility of J_{p^*} implies, by the *Implicit Function Theorem*, that the equation $F(e, \lambda, p) = \mathbf{0}$ has a unique smooth solution $(e(p), \lambda(p))$ defined on an open neighbourhood of p_* small enough. Note that since $\|e(p^*)\|_1 \geq \|e(p^*)\|_2$, using the *Cauchy-Schwarz inequality* in the normalisation condition of the problem yields

$$2(1 - p_i^* + (p_i^*)^2) = e(p^*) \cdot e(p) \leq \|e(p^*)\|_2 \|e(p)\|_2 \leq \|e(p^*)\|_1 \|e(p)\|_1 = 2\|e(p)\|_1,$$

which bounds away from zero the 1-norm of the eigenvector $e(p)$ varying smoothly near p^*. We conclude by noting that since $p^* \in E_1 \cap \overline{\mathscr{K}}$ is arbitrary, the smooth solution is unique and the boundedness away from 0 of its norm is uniform, the local solution extends to the whole of $\mathscr{K}_{\varepsilon,\frac{\delta}{8}}$ for some ε small enough, due to the set being simply connected and relatively compact.

□

We now describe in more detail the eigencoordinates, which will be used. For each $p \in \Sigma_0$ fix $e_{-1}(p)$ and $e_0(p)$ as usual. From Lemma 3.19 (e, d) we have the unique representation of $q - q_p = \alpha e_0(p) + \beta e_{-1}(p)$, with $(\alpha, \beta) \in \mathbb{R}^2$ denoting the new eigencoordinates. Upon one iterate with M_p, denote $\hat{q} - q_{\hat{p}} = \hat{\alpha} e_0(\hat{p}) + \hat{\beta} e_{-1}(\hat{p})$.

Lemma 3.20.

$$\hat{\alpha} = \alpha\frac{\lambda_0(p)-1}{2}\frac{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(\hat{p}) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p}) \end{vmatrix}} + \beta\frac{\lambda_{-1}(p)-1}{2}\frac{\begin{vmatrix} e_{-1}^{(i)}(p) & e_{-1}^{(i)}(\hat{p}) \\ e_{-1}^{(j)}(p) & e_{-1}^{(j)}(\hat{p}) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p}) \end{vmatrix}} \tag{3.6}$$

$$\hat{\beta} = \alpha\frac{\lambda_0(p)-1}{2}\frac{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_0^{(i)}(p) \\ e_0^{(j)}(\hat{p}) & e_0^{(j)}(p) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p}) \end{vmatrix}} + \beta\frac{\lambda_{-1}(p)-1}{2}\frac{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(p) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p}) \end{vmatrix}} \tag{3.7}$$

Proof. From (3.2), a system of three linear equations in two variables follows,

$$\hat{\alpha}e_0(\hat{p}) + \hat{\beta}e_{-1}(\hat{p}) = \alpha\frac{\lambda_0(p)-1}{2}e_0(p) + \beta\frac{\lambda_{-1}(p)-1}{2}e_{-1}(p)$$

with $(\hat{\alpha}, \hat{\beta}) \in \mathbb{R}^2$. The system can be solved by picking any two of the three equations. Since the vectors, which will give the coefficients to the matrix of this linear system of two equations, are $e_0(\hat{p})$ and $e_{-1}(\hat{p})$, and since they are linearly independent, the matrix will have nonzero determinant, being the absolute value of its determinant always equal to the absolute value of one of the coordinates of the vector product $e_0(\hat{p}) \times e_{-1}(\hat{p})$. Since all of the coordinates of the vector product are identical and therefore nonzero (because of linear independence and nondegeneracy of the eigenvectors, which lie in Π_0, which has normal vector $\mathbf{1}$, to which the vector product will be parallel) and since they can be computed, in absolute value, from selecting rows $i \neq j \in \{1,2,3\}$ from the matrix $(e_0(\hat{p})|e_{-1}(\hat{p}))$; we have that, for any $i \neq j \in \{1,2,3\}$ chosen, the linear system can be reduced to

$$\begin{pmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p}) \end{pmatrix}\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \alpha\frac{\lambda_0(p)-1}{2}e_0^{(i)}(p) + \beta\frac{\lambda_{-1}(p)-1}{2}e_{-1}^{(i)}(p) \\ \alpha\frac{\lambda_0(p)-1}{2}e_0^{(j)}(p) + \beta\frac{\lambda_{-1}(p)-1}{2}e_{-1}^{(j)}(p) \end{pmatrix}.$$

This system has unique solution, which can be calculated by Cramer's rule,

$$\hat{\alpha} = \frac{\begin{vmatrix} \alpha\frac{\lambda_0(p)-1}{2}e_0^{(i)}(p) + \beta\frac{\lambda_{-1}(p)-1}{2}e_{-1}^{(i)}(p) & e_{-1}^{(i)}(\hat{p}) \\ \alpha\frac{\lambda_0(p)-1}{2}e_0^{(j)}(p) + \beta\frac{\lambda_{-1}(p)-1}{2}e_{-1}^{(j)}(p) & e_{-1}^{(j)}(\hat{p}) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p}) \end{vmatrix}}$$

$$\hat{\beta} = \frac{\begin{vmatrix} e_0^{(i)}(\hat{p}) & \alpha\frac{\lambda_0(p)-1}{2}e_0^{(i)}(p) + \beta\frac{\lambda_{-1}(p)-1}{2}e_{-1}^{(i)}(p) \\ e_0^{(j)}(\hat{p}) & \alpha\frac{\lambda_0(p)-1}{2}e_0^{(j)}(p) + \beta\frac{\lambda_{-1}(p)-1}{2}e_{-1}^{(j)}(p) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p}) \end{vmatrix}},$$

which yields the iteration given in (3.6) and (3.7). The ratios of the determinants do not depend on the choice of $i \neq j$ because of the aforementioned fact that all coordinates of the vector product are equal; also, the same change of sign will appear, if anywhere, both at the numerator and the denominator, and will therefore ultimately cancel out; moreover, all vectors involved lie in Π_0, so for $k \neq i \neq j$, $e_0(\hat{p})^{(k)} = -e_0(\hat{p})^{(i)} - e_0(\hat{p})^{(j)}$, $e_0(p)^{(k)} = -e_0(p)^{(i)} - e_0(p)^{(j)}$, $e_{-1}(\hat{p})^{(k)} = -e_{-1}(\hat{p})^{(i)} - e_{-1}(\hat{p})^{(j)}$, $e_{-1}(p)^{(k)} = -e_{-1}(p)^{(i)} - e_{-1}(p)^{(j)}$, meaning that the system is overdetermined by an equation, which is linear combination of the other two. $\square$

We proceed to expand in Taylor series about $(p_1, \alpha, \beta) = (0, 0, 0)$ the iteration for the eigencoordinates achieved in Lemma 3.20 and the standard iteration for p. We start with p. Since it belongs to the simplex, $p_3 = 1 - p_1 - p_2$, so for simplicity we perform the analysis only on the first two components.

Lemma 3.21.

$$\hat{p}_1 = p_1 + \rho_1(r) \tag{3.8}$$
$$\hat{p}_2 = p_2 - 2(1-p_2)p_1\beta + \rho_2(r), \tag{3.9}$$

where $\rho_1(r) = \mathcal{O}(\beta p_1, \alpha p_1)$ *and* $\rho_2(r) = \mathcal{O}(\alpha, \beta p_1^2)$.

Proof. Since

$$\hat{p} - p = \mathbf{1} - p - \hat{q} - q = 2q_p - M_p q - q = -(M_p + I)(q - q_p),$$

we have that

$$\hat{p} = p - \alpha(1 + \lambda_0(p))e_0(p) - \beta(1 + \lambda_{-1}(p))e_{-1}(p), \tag{3.10}$$

from which, reading off the first two components and applying Lemma 3.19 (f, g, h), it follows that

$$\begin{aligned}
\hat{p}_1 &= p_1 - \alpha(1 + \lambda_0(p))e_0^{(1)}(p) - \beta(1 + \lambda_{-1}(p))e_{-1}^{(1)}(p) \\
&= p_1 - \alpha(1 - 2p_1 + \mathcal{O}(p_1^2))\mathcal{O}(p_1) - \beta(2p_1 + \mathcal{O}(p_1^2))(-1 + \mathcal{O}(p_1)) \\
&= p_1 + \mathcal{O}(\alpha p_1, \beta p_1) \\
\hat{p}_2 &= p_2 - \alpha(1 + \lambda_0(p))e_0^{(2)}(p) - \beta(1 + \lambda_{-1}(p))e_{-1}^{(2)}(p) \\
&= p_2 - \alpha(1 - 2p_1 + \mathcal{O}(p_1^2))(1 + \mathcal{O}(p_1)) - \beta(2p_1 + \mathcal{O}(p_1^2))(1 - p_2 + \mathcal{O}(p_1)) \\
&= p_2 - 2(1-p_2)p_1\beta + \mathcal{O}(\alpha, \beta p_1^2),
\end{aligned}$$

having used, in the second step of both equations, the smoothness of the eigenvectors to linearise as p approaches the edge E_1 and the relative compactness of $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$ to estimate uniformly the Jacobian terms appearing in

$$\begin{aligned}
e_0(p) &= e_0((0, p_2, 1-p_2) + (p_1, 0, -p_1)) = e_0((0, p_2, 1-p_2)) + \boldsymbol{\mathcal{O}(\|(p_1, 0, -p_1)\|_1)} \\
&= (0, 1, -1) + \boldsymbol{\mathcal{O}(p_1)}, \\
e_{-1}(p) &= e_{-1}((0, p_2, 1-p_2) + (p_1, 0, -p_1)) = e_{-1}((0, p_2, 1-p_2)) + \boldsymbol{\mathcal{O}(\|(p_1, 0, -p_1)\|_1)} \\
&= (-1, 1-p_2, p_2) + \boldsymbol{\mathcal{O}(p_1)}.
\end{aligned}$$

□

Lemma 3.22.

$$\hat{\alpha} = -\frac{\alpha}{2}(1 + \rho_3(r)) + \rho_4(r) \tag{3.11}$$
$$\hat{\beta} = -\beta(1 - p_1) + \rho_5(r), \tag{3.12}$$

where $\rho_3(r) = \mathcal{O}(\alpha, p_1)$, $\rho_4(r) = \mathcal{O}(\beta\alpha, \beta^2 p_1)$, $\rho_5(r) = \mathcal{O}(\alpha^2, \alpha\beta, \beta^2 p_1, \beta p_1^2)$.

Proof. By Lemma 3.21 it follows that $\hat{p} = p + \boldsymbol{\mathcal{O}(\alpha, \beta p_1)}$, since trivially both $\rho_1(r)$ and $\rho_2(r)$ are $\mathcal{O}(\alpha, \beta p_1)$, so

$$\begin{aligned}
\hat{p}_1 &= p_1 + \mathcal{O}(\alpha, \beta p_1) \\
\hat{p}_2 &= p_2 + \mathcal{O}(\beta p_1) + \mathcal{O}(\alpha, \beta p_1) = \mathcal{O}(\alpha, \beta p_1) \\
\hat{p}_3 &= 1 - \hat{p}_1 - \hat{p}_2 = 1 - p_1 - p_2 + \mathcal{O}(\alpha, \beta p_1) = p_3 + \mathcal{O}(\alpha, \beta p_1).
\end{aligned}$$

We can plug these estimates, along with that of Lemma 3.19 (g), in (3.6) and (3.7) obtained in Lemma 3.20; more specifically we can plug them in the terms next to α and β in (3.6) and (3.7). This yields, due to smoothness of the eigenvectors' components and relative compactness of $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$, the following estimates for those terms involved in (3.6):

$$\frac{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1)\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) \\ e_0^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1)\end{vmatrix}} = \frac{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(p)\end{vmatrix}+\mathcal{O}(\alpha,\beta p_1)}{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(p)\end{vmatrix}+\mathcal{O}(\alpha,\beta p_1)};$$

$$\frac{\begin{vmatrix} e_{-1}^{(i)}(p) & e_{-1}^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) \\ e_{-1}^{(j)}(p) & e_{-1}^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1)\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) \\ e_0^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1)\end{vmatrix}} = \frac{\begin{vmatrix} e_{-1}^{(i)}(p) & e_{-1}^{(i)}(p) \\ e_{-1}^{(j)}(p) & e_{-1}^{(j)}(p)\end{vmatrix}+\mathcal{O}(\alpha,\beta p_1)}{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(p)\end{vmatrix}+\mathcal{O}(\alpha,\beta p_1)}.$$

Hence (3.6) becomes

$$\begin{aligned}
\hat{\alpha} &= \alpha\left(-\frac{1}{2}+\mathcal{O}(p_1)\right)(1+\mathcal{O}(\alpha,\beta p_1)) \\
&+\beta\left(-1+\mathcal{O}(p_1)\right)\frac{0+\mathcal{O}(\alpha,\beta p_1)}{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(p)\end{vmatrix}+\mathcal{O}(\alpha,\beta p_1)} \\
&= \alpha\left(-\frac{1}{2}+\mathcal{O}(p_1)\right)(1+\mathcal{O}(\alpha,\beta p_1))+\beta\left(-1+\mathcal{O}(p_1)\right)\mathcal{O}(\alpha,\beta p_1) \\
&= \alpha\left(-\frac{1}{2}+\mathcal{O}(\alpha,p_1)\right)+\beta\mathcal{O}(\alpha,\beta p_1) = -\frac{\alpha}{2}(1+\mathcal{O}(\alpha,p_1))+\mathcal{O}(\beta\alpha,\beta^2 p_1).
\end{aligned}$$

Doing the same with the corresponding terms in (3.7) yields the following estimates:

$$\frac{\begin{vmatrix} e_0^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_0^{(i)}(p) \\ e_0^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_0^{(j)}(p)\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) \\ e_0^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1)\end{vmatrix}} = \frac{\begin{vmatrix} e_0^{(i)}(p) & e_0^{(i)}(p) \\ e_0^{(j)}(p) & e_0^{(j)}(p)\end{vmatrix}+\mathcal{O}(\alpha,\beta p_1)}{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(p)\end{vmatrix}+\mathcal{O}(\alpha,\beta p_1)};$$

$$\frac{\begin{vmatrix} e_0^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(j)}(p)\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(i)}(p)+\mathcal{O}(\alpha,\beta p_1) \\ e_0^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1) & e_{-1}^{(j)}(p)+\mathcal{O}(\alpha,\beta p_1)\end{vmatrix}} = \frac{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(p)\end{vmatrix}+\mathcal{O}(\alpha,\beta p_1)}{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(p)\end{vmatrix}+\mathcal{O}(\alpha,\beta p_1)}.$$

Hence (3.7) becomes

$$\begin{aligned}
\hat{\beta} &= \alpha\left(-\frac{1}{2}+\mathcal{O}(p_1)\right)(0+\mathcal{O}(\alpha,\beta p_1))+\beta\left(-1+p_1+\mathcal{O}(p_1^2)\right)(1+\mathcal{O}(\alpha,\beta p_1)) \\
&= \alpha\mathcal{O}(\alpha,\beta p_1)+\beta(-1+p_1+\mathcal{O}(p_1^2)+\mathcal{O}(\alpha,\beta p_1)) = \alpha\mathcal{O}(\alpha,\beta p_1) \\
&+\beta(-1+p_1+\mathcal{O}(\alpha,\beta p_1,p_1^2)) = \mathcal{O}(\alpha^2,\alpha\beta p_1)-\beta(1-p_1)+\mathcal{O}(\alpha\beta,\beta^2 p_1,\beta p_1^2) \\
&= -\beta(1-p_1)+\mathcal{O}(\alpha^2,\alpha\beta,\beta^2 p_1,\beta p_1^2).
\end{aligned}$$

Note that often we have implicitly used the fact that the determinant in the denominators is bounded away from zero. □

Lemma 3.23. *Let the constant* $\theta := 1/16$. *There is* $c > 0$ *such that for all sufficiently small* $\varepsilon > 0$*, on the closure* $\overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$ *it holds that*

$$\begin{aligned} |\rho_1(r)| &< \theta p_1 \\ |\rho_2(r)| &< c|\alpha| + p_1|\beta| \\ |\rho_3(r)| &< \theta \\ |\rho_4(r)| &< \theta|\alpha| + \theta p_1|\beta| \\ |\rho_5(r)| &< \theta|\alpha| + \theta p_1|\beta|. \end{aligned}$$

Proof. Start with any given $\varepsilon > 0$ suitably small, so as to have well-definedness of all quantities involved, to possibly be further reduced, and $\delta > 0$ fixed, small enough to have all quantities involved well-defined too. To simplify the notation, since we will exclusively work on $\overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$, we will denote $\mathscr{K}^* := \mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$.

Starting with ρ_1, by Lemma 3.21 it holds that on $\overline{\mathscr{K}}^*$ there is a constant $c_1 > 0$, such that $|\rho_1(r)| \leq c_1 p_1(|\alpha| + |\beta|)$. If now one further restricts $\varepsilon < \theta/(2c_1)$, it follows that $c_1 p_1(|\alpha| + |\beta|) < \theta p_1$, since $c_1(|\alpha| + |\beta|) \leq 2\varepsilon c_1 < 1$, thus yielding the desired estimate. Note that further restricting ε is consistent with c_1, because the same constant upper bound applies, being δ fixed, when $\overline{\mathscr{K}}^*$ gets smaller, as ε gets possibly reduced in the future steps.

Moving on to ρ_2, by Lemma 3.21 it holds that on $\overline{\mathscr{K}}^*$ there is a constant $c_2 > 0$, such that $|\rho_2(r)| \leq c_2(|\alpha| + |\beta|p_1^2)$. Define $c := c_2$ and further restrict $\varepsilon < 1/c$ if necessary. Then $c_2(|\alpha| + |\beta|p_1^2) < c|\alpha| + |\beta|p_1$, since $cp_1 \leq c\varepsilon < 1$. Similarly to the previous step, this bound is consistent with further reducing ε in future steps if necessary.

As to $\rho_3(r)$, by Lemma 3.22 we know that on $\overline{\mathscr{K}}^*$ there is a constant $c_3 > 0$, such that $|\rho_3(r)| \leq c_3(|\alpha| + p_1)$. Similarly to what done previously, further restrict $\varepsilon < \theta/(2c_3)$ if necessary, then it follows that $c_3(|\alpha| + p_1) \leq 2c_3\varepsilon < \theta$, yielding $|\rho_3(r)| \leq \theta$.

For $\rho_4(r)$, by Lemma 3.22 it is known that on $\overline{\mathscr{K}}^*$ there is a constant $c_4 > 0$, such that $|\rho_4(r)| \leq c_4(|\beta||\alpha| + \beta^2 p_1) = c_4|\beta|(|\alpha| + |\beta|p_1)$. Similarly to what done in previous steps, further restrict $\varepsilon < \theta/c_4$ if necessary, then it follows that $c_4|\beta| \leq c_4\varepsilon < \theta$, yielding $|\rho_4(r)| < \theta(|\alpha| + |\beta|p_1)$.

Lastly $\rho_5(r)$. By Lemma 3.22 it holds that on $\overline{\mathscr{K}}^*$ there is a constant $c_5 > 0$, such that $|\rho_5(r)| \leq c_5(\alpha^2 + |\beta||\alpha| + \beta^2 p_1 + |\beta|p_1^2) = c_5(|\alpha| + |\beta|)|\alpha| + c_5(|\beta| + p_1)p_1|\beta|$. As always, further restrict $\varepsilon < \theta/(2c_5)$ if necessary. Then it follows that $c_5(|\alpha| + |\beta|) \leq 2c_5\varepsilon < \theta$ and $c_5(p_1 + |\beta|) \leq 2c_5\varepsilon < \theta$, yielding $|\rho_5(r)| < \theta|\alpha| + \theta|\beta|p_1)$.

All in all, starting from a given ε defining constants c_1, c, c_3, c_4, c_5, possibly further restricted such that

$$\varepsilon < \min\left\{\frac{\theta}{2c_1}, \frac{1}{c}, \frac{\theta}{2c_3}, \frac{\theta}{c_4}, \frac{\theta}{2c_5}\right\},$$

all obtained five estimates will hold on $\overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$. □

For further arguments it will be necessary to add the requirement to ε that, given δ, c, θ,

$$\varepsilon < \min\left\{\theta, \frac{\delta(1-2\theta)}{16(3+c)}\right\}$$

and that ε be so small, that $\overline{\mathscr{K}}_{2\varepsilon,\frac{\delta}{8}}$ does not intersects E_2 nor E_3, and every point in $\overline{\mathscr{K}}_{2\varepsilon,\frac{\delta}{8}}$ is closer to E_1 than to E_2 and E_3 (in 1-norm).

Consider now the orbit $\{(p_n, q_n)\}$ as described at the beginning of this section, and recall that the subsequence $\{p_{n_k}\}_{k\in\mathbb{N}}$ was, by construction, bounded away from the vertices and convergent to $p_* \in E_1$. For each n one will have $r_n := (p_n^{(1)}, p_n^{(2)}, \alpha_n, \beta_n)$ (note that the lower index has been moved to an upper one to allow the index n for the orbit's time, as usual). For every $k \geq K$, define the times

$$\tau_k := \inf \left\{ n > n_k : \ \Theta_n^{(2)} \notin \left[\frac{\delta}{8}, 1 - \frac{\delta}{8} \right] \right\} \in \mathbb{N} \cup \infty.$$

We already saw that there will be an arbitrarily large K such that for $m := n_k$, for any $k \geq K$, $p_m^{(2)} \in [\delta/2, 1 - \delta/2]$ and $|\alpha_n|, |\beta_n| < \varepsilon$ for all $n \geq m$ (that is, the subsequence of the orbit is in $\mathscr{K}^*_{\varepsilon, \frac{\delta}{2}}$). It is left to show that, by choosing a suitable $k \geq K$ large enough, and letting $m = n_k$, we can have $p_n^{(1)} < \varepsilon$ for all $m \leq n < \tau_k$, on top of the previous conditions. In order to do this, we will add one more requirement on K: since $\ell = 0$, $\|p_{n+1} - p_n\|_1 \longrightarrow 0$, we can choose K large enough, such that for all $k \geq K$,

$$|p_{n_k}^{(2)} - p_*^{(2)}| < \frac{1}{2} \min \left\{ p_*^{(2)} - \frac{\delta}{8}, 1 - \frac{\delta}{8} - p_*^{(2)} \right\}$$

and for any $n \geq n_K$,

$$\|p_{n+1} - p_n\|_1 < \frac{1}{2} \min \left\{ p_*^{(2)} - \frac{\delta}{8}, 1 - \frac{\delta}{8} - p_*^{(2)} \right\}.$$

This assumption ensures that for all $k \geq K$, $\tau_{n_k} > n_k + 1$ (so that there is always some $n_k < n < \tau_k$), since

$$p_{n_k+1}^{(2)} \leq |p_{n_k+1}^{(2)} - p_{n_k}^{(2)}| + |p_{n_k}^{(2)} - p_*^{(2)}| + p_*^{(2)} < \min \left\{ p_*^{(2)} - \frac{\delta}{8}, 1 - \frac{\delta}{8} - p_*^{(2)} \right\} + p_*^{(2)} \leq 1 - \frac{\delta}{8}$$

and

$$\begin{aligned} p_{n_k+1}^{(2)} &\geq p_*^{(2)} - |p_{n_k+1}^{(2)} - p_{n_k}^{(2)} + p_{n_k}^{(2)} - p_*^{(2)}| \geq p_*^{(2)} - \left(|p_{n_k+1}^{(2)} - p_{n_k}^{(2)}| + |p_{n_k}^{(2)} - p_*^{(2)}| \right) \\ &> p_*^{(2)} - \min \left\{ p_*^{(2)} - \frac{\delta}{8}, 1 - \frac{\delta}{8} - p_*^{(2)} \right\} \geq \frac{\delta}{8}. \end{aligned}$$

Lemma 3.24. *There exists a $k \geq K$ large enough such that, having defined $m := n_k$, for all $m \leq n < \tau_k$, $p_n^{(1)} \leq \varepsilon$.*

Proof. For $n = n_k$ it is trivial. For $n > n_k$ one needs to distinguish between two cases and proceed by contradiction. The core of the argument is the same for both cases, only the preparation slightly differs.

If $\tau_k \in \mathbb{N}$ for all $k \geq K$, suppose by contradiction that there is a subsequence $\{k_r\}$ (with $k_r \geq K$ for all $r \in \mathbb{N}$) such that for every r, for some $n_{k_r} < n < \tau_{k_r}$, $p_n^{(1)} > \varepsilon$. This implies that we can construct a subsequence $\{p_{n_r}\}$ for which $p_{n_r}^{(1)} > \varepsilon$ and $p_{n_r}^{(2)} \in [\delta/8, 1 - \delta/8]$. From this subsequence, a subsubsequence $\{p_{n_{r_l}}\}$ can be extracted - denote it $\{p_{n_l}\}$ for simplicity - such that $p_{n_l - 1}^{(1)} \leq \varepsilon$ and $\varepsilon < p_{n_l}^{(1)} < 2\varepsilon$. This is true because, looking back at the original sequence, at least $p_{n_{k_{r_l}}}^{(1)} \leq \varepsilon$, so the sequence exits $(0, \varepsilon]$ after having been inside the interval for at least one time, and one can choose n_l to be the first time of exit from $(0, \varepsilon]$ for every l. Furthermore, the vanishing potential implies $\|p_n - p_{n-1}\|_1 \longrightarrow 0$ by (3.4), so for all l large enough, $|p_{n_l}^{(1)} - p_{n_l - 1}^{(1)}| \leq \|p_{n_l} - p_{n_l - 1}\|_1 < \varepsilon$, hence

$$\varepsilon < p_{n_l}^{(1)} \leq |p_{n_l}^{(1)} - p_{n_l - 1}^{(1)}| + p_{n_l - 1}^{(1)} < 2\varepsilon.$$

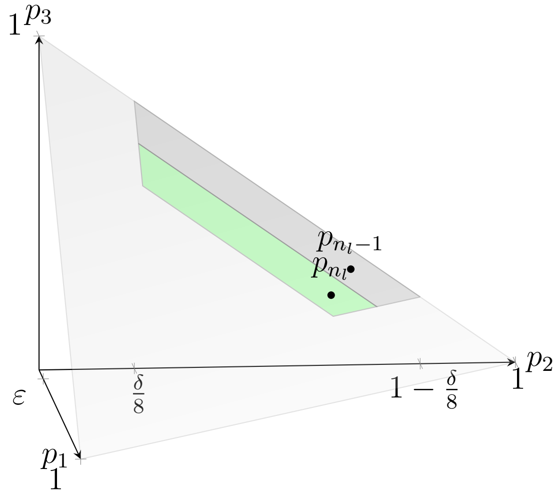


Figure 3.4: $R_{\varepsilon,\frac{\delta}{8}}$ in green

At the same time, by construction $\delta/8 < p^{(2)}_{n_l} < 1-\delta/8$, since $n_{k_{r_l}} < n_l < \tau_{k_{r_l}}$. Consider now the set $\{(p,q_p) : \ p \in \overline{R}\}$ where $R = R_{\varepsilon,\frac{\delta}{8}} := \mathscr{K}_{2\varepsilon,\frac{\delta}{8}} \setminus \mathscr{K}_{\varepsilon,\frac{\delta}{8}}$ (see Figure 3.4). By construction of ε, every $p \in \overline{R}$ does not lie on $\partial\Sigma$ and it is closer to E_1 than to the other two edges. Note that $p_{n_l} \in \mathring{R}$. For every p, fixing a small enough $\varepsilon' < \varepsilon = \mathrm{dist}(\overline{R}, E_1)/2$ (recall that we work in 1-norm), letting $\not c$ the positive subunitary constant such that $\|L_p(q-q_p)\|_1 \leq \not c\|q-q_p\|_1$ for all $p \in \overline{R}$ (following the same construction and definitions as in Lemma 3.14, it is known that this constant holds uniformly on the whole compact set $\Sigma_{\varepsilon'} := \{p \in \Sigma : \ p_i \geq \varepsilon', \ \forall\ i \in \{1,2,3\}\}$, so in particular it is uniform on $\overline{R} \subset \Sigma_{\varepsilon'}$ by construction of ε); one can set $\delta' := (1-\not c)\varepsilon'/8$. Then since $\|q_n - q_{p_n}\|_1 \longrightarrow 0$, for some large enough $\bar{l}$, $(p_{n_{\bar{l}}}, q_{n_{\bar{l}}}) \in U((p_{n_{\bar{l}}}, q_{p_{n_{\bar{l}}}}), \delta')$, hence by Proposition 3.17 it is known that for all $n \geq n_{\bar{l}}$, $p_n \in B(p_{n_{\bar{l}}}, \varepsilon')$. Indeed Proposition 3.17 applies, since by construction $\varepsilon' < \mathrm{dist}(\overline{R}, \partial\Sigma)$, and therefore for every $p \in \overline{R}$, $\varepsilon' < \mathrm{dist}(p, \partial\Sigma)$. But then eventually $\mathrm{dist}(p_n, E_1) \geq \mathrm{dist}(B(p_{n_{\bar{l}}}, \varepsilon'), E_1) \geq \mathrm{dist}(\overline{R}, E_1) - \varepsilon' > 2\varepsilon - \varepsilon = \varepsilon$. This is in contradiction with $p^{(1)}_{n_k} \longrightarrow 0$.

If for some $\bar{k} \geq K$, $\tau_{\bar{k}} = \infty$, then for all $k \geq \bar{k}$, $\tau_k = \infty$. Suppose again, by the same construction as above, that, by contradiction, there is a subsequence $\{p_{n_r}\}$ for which $p^{(1)}_{n_r} > \varepsilon$, with $n_r > n_{k_r}$ and $k_r \geq \bar{k}$ for all $r \in \mathbb{N}$. Note that this time it automatically holds that $p^{(2)}_{n_r} \in [\delta/8, 1-\delta/8]$ (since all $\tau_{k_r} = \infty$ for all $k_r \geq \bar{k}$, $p^{(2)}_n \in [\delta/8, 1-\delta/8]$ for all $n \geq n_{\bar{k}}$, in particular the last condition imposed on K is not necessary for this case). This case only differs in that it does not require any degree of control on $p^{(2)}_n$, and it is now clear that we can just proceed as in the previous case. Another subsequence $\{p_{n_{r_l}}\}$ can be extracted - denote it simply $\{p_{n_l}\}$ as before - such that $p^{(1)}_{n_l-1} \leq \varepsilon$ and $p^{(1)}_{n_l} > \varepsilon\}$. Hence $\varepsilon < p^{(1)}_{n_l} < 2\varepsilon$, along with $\delta/8 < p^{(2)}_{n_l} < 1-\delta/8$. From here on the previous argument takes care of things. □

In the following all the proofs are made with respect to the large enough $m = n_k$, with $k \geq K$, existing by Lemma 3.24, and therefore the corresponding τ_k will be simply denoted as τ.

Lemma 3.25. *For all $m \leq n \leq \tau$,*

$$|\alpha_n| \leq \max\left\{\left(\frac{3}{4}\right)^{n-m}|\alpha_m|, p^{(1)}_n|\beta_n|\right\}.$$

Proof. Proceed by induction. If $n = m$, the statement $|\alpha_m| \leq \max\{|\alpha_m|, p_m^{(1)}|\beta_m|\}$ is trivially true by definition of m and $\overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{2}}$. If $n = m + 1 < \tau$, recall that by (3.11) in Lemma 3.22 and Lemma 3.23 on $\overline{\mathscr{K}}^* := \overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$ it holds that

$$|\alpha_{m+1}| \leq \frac{|\alpha_m|}{2}(1 + |\rho_3(r_m)|) + |\rho_4(r_m)| \leq \frac{|\alpha_m|}{2}(1+\theta) + \theta|\alpha_m| + \theta p_m^{(1)}|\beta_m|,$$

which applies by definition of m. Then it follows that

$$|\alpha_{m+1}| \leq |\alpha_m|\frac{1+3\theta}{2} + \theta p_m^{(1)}|\beta_m|.$$

If $|\alpha_m| \geq p_m^{(1)}|\beta_m|$, then

$$|\alpha_{m+1}| \leq |\alpha_m|\frac{1+3\theta}{2} + \theta|\alpha_m| = \frac{1+5\theta}{2}|\alpha_m| \leq \frac{3}{4}|\alpha_m|,$$

as $\theta < {}^1/{}_{10}$ and ${}^1/{}_{10}$ is the value, at which the equation $(1+5x)/2 = {}^3/{}_4$ holds (with the left-hand side being increasing). If instead $|\alpha_m| < p_m^{(1)}|\beta_m|$, then it holds that

$$|\alpha_{m+1}| \leq p_m^{(1)}|\beta_m|\frac{1+3\theta}{2} + \theta p_m^{(1)}|\beta_m| = \frac{1+5\theta}{2}p_m^{(1)}|\beta_m|. \tag{3.13}$$

By definition of m and by (3.8) in Lemma 3.21, and by Lemma 3.23,

$$p_{m+1}^{(1)} \geq p_m^{(1)} - |\rho_1(r_m)| \geq p_m^{(1)} - \theta p_m^{(1)} = (1-\theta)p_m^{(1)}, \tag{3.14}$$

since at time m the orbit is in $\overline{\mathscr{K}}^*$, so the lemmas apply. Note also that, since by hypothesis $|\alpha_m| < p_m^{(1)}|\beta_m| < |\beta_m|$ and since by assumption $\varepsilon \leq \theta$, by definition of m, it follows that $p_m^{(1)} < \varepsilon \leq \theta$. This yields, by applying (3.12) and Lemma 3.23, that

$$|\beta_{m+1}| \geq |\beta_m|(1-p_m^{(1)}) - |\rho_5(r_m)| > |\beta_m|(1-p_m^{(1)}) - \theta|\alpha_m| - \theta p_m^{(1)}|\beta_m| \geq |\beta_m|(1-\theta) - 2\theta|\beta_m|,$$

which yields

$$|\beta_{m+1}| \geq |\beta_m|(1-3\theta). \tag{3.15}$$

Then plugging the bounds in (3.14) and (3.15) into (3.13) yields

$$\alpha_{m+1} \leq \frac{1+5\theta}{2}\frac{p_{m+1}^{(1)}|\beta_{m+1}|}{(1-\theta)(1-3\theta)} < p_{m+1}^{(1)}|\beta_{m+1}|,$$

since $(1+5x)/[2(1-x)(1-3x)] = 1$ holds only at $(13-\sqrt{145})/12 > (13-\sqrt{147})/12 = (13-7\sqrt{3})/12 > [13-7(1+{}^3/{}_4)]/12 =: \theta$ for $x < {}^1/{}_3$, and the function $(1+5x)/[2(1-x)(1-3x)]$ is monotone increasing on that interval.

Assume now the hypothesis for any $m+1 \leq n < \tau$. This time, when carrying on with the inductive step, it will not be possible to appeal to the definition of m, but it will be necessary to rely on Lemma 3.24, which ensures that $(p_n, q_n) \in \overline{\mathscr{K}}^*$ and therefore makes it possible for the same lemmas we just used to apply, in the corresponding parts of the induction step. First we have

$$|\alpha_{n+1}| \leq |\alpha_n|\frac{1+3\theta}{2} + \theta p_n^{(1)}|\beta_n|$$

by (3.11) in Lemma 3.22 and Lemma 3.23. Next, if $({}^3/{}_4)^{n-m}|\alpha_m| \geq p_n^{(1)}|\beta_n|$, the induction hypothesis reads

$$|\alpha_n| \leq \left(\frac{3}{4}\right)^{n-m}|\alpha_m| = \max\left\{\left(\frac{3}{4}\right)^{n-m}|\alpha_m|, p_n|\beta_n|\right\}$$

and then

$$\begin{aligned}|\alpha_{n+1}| &\le \left(\frac{3}{4}\right)^{n-m}\frac{1+3\theta}{2}|\alpha_m| + \left(\frac{3}{4}\right)^{n-m}\theta|\alpha_m| = \left(\frac{3}{4}\right)^{n-m}\frac{1+5\theta}{2}|\alpha_m|\\ &\le \left(\frac{3}{4}\right)^{n+1-m}|\alpha_m|.\end{aligned}$$

If instead $(3/4)^{n-m}|\alpha_m| < p_n^{(1)}|\beta_n|$, then by the induction hypothesis it holds that

$$|\alpha_{n+1}| \le \frac{1+3\theta}{2}\max\left\{\left(\frac{3}{4}\right)^{n-m}|\alpha_m|, p_n^{(1)}|\beta_n|\right\} + \theta p_m^{(1)}|\beta_m| = \frac{1+5\theta}{2}p_m^{(1)}|\beta_m|. \quad (3.16)$$

By (3.8) and Lemma 3.23

$$p_{n+1}^{(1)} \ge (1-\theta)p_n^{(1)}. \quad (3.17)$$

Note also that since in this case $(3/4)^{n-m}|\alpha_m| < p_n^{(1)}|\beta_n|$, the induction hypothesis reads $|\alpha_n| \le p_n^{(1)}|\beta_n|$ and by construction $\varepsilon \le \theta$; it follows by Lemma 3.24 that $p_n^{(1)} < \varepsilon \le \theta$. Both facts yield $|\alpha_n| \le \theta|\beta_n|$, and therefore by applying (3.12) and Lemma 3.23, we have that

$$|\beta_{n+1}| > |\beta_n|(1-3\theta). \quad (3.18)$$

Then plugging the bounds (3.17) and (3.18) into (3.16) yields

$$\alpha_{n+1} \le \frac{1+5\theta}{2}\frac{p_{n+1}^{(1)}|\beta_{n+1}|}{(1-\theta)(1-3\theta)} < p_{n+1}^{(1)}|\beta_{n+1}|,$$

which proves that

$$|\alpha_{n+1}| \le \max\left\{\left(\frac{3}{4}\right)^{n+1-m}|\alpha_m|, p_{n+1}^{(1)}|\beta_{n+1}|\right\}.$$

□

Let

$$\sigma := \inf\left\{n \ge m : \left(\frac{3}{4}\right)^{n-m}|\alpha_m| \le p_n^{(1)}|\beta_n|\right\} \in \mathbb{N}\cup\infty.$$

Lemma 3.26.

a) If $\tau < \infty$, then $\sigma < \tau$ and $p_\sigma^{(2)} \in [\delta/4, 1-\delta/4]$;

b) If $\sigma < \infty$, then for all $\sigma \le n \le \tau$,

$$\left(\frac{3}{4}\right)^{n-m}|\alpha_m| \le p_n^{(1)}|\beta_n|.$$

Proof.

a) For all $m \le n \le \sigma \wedge \tau$, by definition of σ, $p_n^{(1)}|\beta_n| < (3/4)^{n-m}|\alpha_m|$ and as a result by Lemma 3.25, $\alpha_n \le (3/4)^{n-m}|\alpha_m|$. By (3.9), Lemma 3.23 and Lemma 3.24

(this last one used, as in the previous Lemma 3.25, to ensure that the bounds on the errors in Lemma 3.23 hold) it follows that

$$\begin{aligned}|p^{(2)}_{\sigma\wedge\tau} - p^{(2)}_m| &\le \sum_{n=m}^{\sigma\wedge\tau-1} |p^{(2)}_{n+1} - p^{(2)}_n| = \sum_{n=m}^{\sigma\wedge\tau-1} |2(1-p^{(2)}_n)p^{(1)}_n\beta_n - \rho_2(r_n)| \\ &< 2\sum_{n=m}^{\sigma\wedge\tau-1} p^{(1)}_n|\beta_n| + \sum_{n=m}^{\sigma\wedge\tau-1} c|\alpha_n| + p^{(1)}_n|\beta_n| = 3\sum_{n=m}^{\sigma\wedge\tau-1} p^{(1)}_n|\beta_n| \\ &+ c\sum_{n=m}^{\sigma\wedge\tau-1}|\alpha_n| \le (3+c)|\alpha_m| \sum_{n=m}^{\sigma\wedge\tau-1}\left(\frac{3}{4}\right)^{n-m} \le (3+c)|\alpha_m|\sum_{i=0}^{\infty}\left(\frac{3}{4}\right)^{i} \\ &= 4(3+c)|\alpha_m| < 4(3+c)\varepsilon < 4(3+c)\frac{\delta(1-2\theta)}{16(3+c)} < \frac{\delta}{4}.\end{aligned}$$

Since $p^{(2)}_m \in [\delta/2, 1-\delta/2]$, having travelled a distance less than $\delta/4$, it follows that $p^{(2)}_{\sigma\wedge\tau} \in [\delta/4, 1-\delta/4] \subset [\delta/8, 1-\delta/8]$, hence $\sigma\wedge\tau \neq \tau$, otherwise by definition of τ, $p^{(2)}_{\sigma\wedge\tau}$ would not belong to $[\delta/8, 1-\delta/8]$. Hence $\sigma\wedge\tau = \sigma$, that is $\sigma < \tau$, and in particular $p^{(2)}_\sigma \in [\delta/4, 1-\delta/4]$.

b) Note that by definition of σ, the case $n = \sigma$ is trivially true. If $\sigma < \infty$, $\sigma < \tau$, because if $\tau = \infty$, $\sigma < \tau = \infty$; whereas by part (a) if $\tau < \infty$, then $\sigma < \tau$. Hence one can only assume the claim to be true up to some $\sigma < n < \tau$. Recall that the steps in Lemma 3.25 apply again and produce (3.17), that is $p^{(1)}_{n+1} \ge (1-\theta)p^{(1)}_n$, and (3.18), that is $|\beta_{n+1}| \ge (1-3\theta)|\beta_n|$. This last equation is still true for n, since the induction hypothesis $(3/4)^{n-m}|\alpha_m| \le p^{(1)}_n|\beta_n|$ and Lemma 3.25 imply that $|\alpha_n| \le p^{(1)}_n|\beta_n|$, which is the key to the relevant estimate in order to get the result. Putting these facts altogether yields that

$$\begin{aligned}p^{(1)}_{n+1}|\beta_{n+1}| &\ge (1-\theta)(1-3\theta)p^{(1)}_n|\beta_n| \ge (1-\theta)(1-3\theta)\left(\frac{3}{4}\right)^{n-m}|\alpha_m| \\ &> \left(\frac{3}{4}\right)^{n+1-m}|\alpha_m|,\end{aligned}$$

where the induction hypothesis has been used in the second last inequality, and the last inequality follows from $(1-x)(1-3x)$ being decreasing at the left of $2/3$ (the vertex of the parabola) and hitting $3/4$ at $(4-\sqrt{13})/6 > \theta$, since $\sqrt{13} < 361/100$ because $130000 < 130321 = 361^2$, and therefore $(4-\sqrt{13})/6 > (4-361/100)/6 = 13/200 > 1/16$, since the last inequality is equivalent to $208 > 200$.

□

Theorem 3.27. $\tau = \infty$ *and the dynamical system converges.*

Proof. Suppose first that $\sigma = \infty$. By Lemma 3.26 (a), if $\tau < \infty$, $\sigma < \tau < \infty$ against the hypothesis, hence $\tau = \infty$. By (3.9) and Lemmas 3.23 to 3.25 and the definition of σ it follows that

$$\begin{aligned}\sum_{n=m}^{\infty}|p^{(2)}_{n+1} - p^{(2)}_n| &= \sum_{n=m}^{\infty}|2(1-p^{(2)}_n)p^{(1)}_n\beta_n + \rho_2(r_n)| \le 3\sum_{n=m}^{\infty}p^{(1)}_n|\beta_n| + c\sum_{n=m}^{\infty}|\alpha_n| \\ &\le (3+c)|\alpha_m|\sum_{n=m}^{\infty}\left(\frac{3}{4}\right)^{n-m} \le 4(3+c)\varepsilon < \frac{\delta}{4},\end{aligned}$$

so $p^{(2)}_n$ converges within $[\delta/4, 1-\delta/4] \subset [\delta/8, 1-\delta/8]$. The existence of a subsequence $p_{n_j} \longrightarrow p_* \in E_1 \cap \overline{\mathscr{K}}_{\varepsilon,\frac{\delta}{8}}$, ensures that ε can be chosen arbitrarily small, and by

Lemma 3.24 this means that $p_n^{(1)}$ can be chosen to stay arbitrarily small while $p_n^{(2)}$ converges within $[\delta/8, 1-\delta/8]$. All of this implies that p_n is eventually inside $\overline{\mathscr{K}}_{\varepsilon,\frac{\delta}{8}}$ with ε arbitrarily small, while $p_n^{(2)}$ converges. Since ε is arbitrary, $p_n^{(1)} \longrightarrow 0$ and therefore $p_n \longrightarrow p_* \in E_1 \cap \overline{\mathscr{K}}_{\varepsilon,\frac{\delta}{8}}$.

Suppose now that $\sigma < \infty$. By Lemma 3.26 (b), $\sigma < \tau$. By (3.12) in Lemma 3.22 for any $\sigma \leq k < \tau$

$$\beta_{k+1} + (-1)^{k-\sigma}\beta_\sigma = \sum_{n=\sigma}^{k}(-1)^{k-n}(\beta_{n+1}+\beta_n) = \sum_{n=\sigma}^{k}(-1)^{k-n}(\beta_n p_n^{(1)} + \rho_5(r_n)).$$

By the definition of σ, Lemmas 3.23 to 3.25 and Lemma 3.26 (b) for all $\sigma \leq n \leq k$,

$$|\rho_5(r_n)| < \theta|\alpha_n| + \theta p_n^{(1)}|\beta_n| \leq 2\theta p_n^{(1)}|\beta_n|,$$

since $|\alpha_n| \leq p_n^{(1)}|\beta_n| = \max\{(3/4)^{n-m}|\alpha_m|, p_n^{(1)}|\beta_n|\}$. It follows that $\{\beta_n\}_{n=\sigma}^{k}$ has an alternating sign (the sign of two consecutive terms flips, whenever the previous term is nonzero, that is; zero terms do not provide contribution to the sum we are trying to estimate with this argument, which is $\sum_{n=\sigma}^{k} p_n^{(1)}|\beta_n|$) since

$$\begin{aligned}\beta_{n+1}\beta_n &= (-1+p_n^{(1)})\beta_n^2 + \rho_5(r_n)\beta_n \leq (-1+p_n^{(1)})\beta_n^2 + |\rho_5(r_n)||\beta_n| \\ &\leq (-1+p_n^{(1)}+2\theta p_n^{(1)})\beta_n^2 < [-1+\theta(1+2\theta)]\beta_n^2 < 0\end{aligned}$$

by Lemma 3.24, and by the fact that $\varepsilon \leq \theta < 1/(1+2\theta)$, since for $0 \leq x < 1/2$ the function $x(1+2x) < 1$ and $\theta \in (0, 1/2)$. Since for all $k \geq m$, $|\beta_k| < \varepsilon$, and the sign alternates as aforementioned,

$$\begin{aligned}2\varepsilon > |\beta_{k+1} + (-1)^{k-\sigma}\beta_\sigma| &= \left|\sum_{n=\sigma}^{k}(-1)^{k-n}(\beta_n p_n^{(1)} + \rho_5(r_n))\right| = \left|\sum_{n=\sigma}^{k}(-1)^{k-n}\mathrm{sign}(\beta_n)|\beta_n|p_n^{(1)}\right. \\ &\left.+\sum_{n=\sigma}^{k}(-1)^{k-n}\rho_5(r_n)\right| = \left|\sum_{n=\sigma}^{k}\mathrm{sign}(\beta_k)|\beta_n|p_n^{(1)} + \sum_{n=\sigma}^{k}(-1)^{k-n}\rho_5(r_n)\right| \geq \sum_{n=\sigma}^{k}|\beta_n|p_n^{(1)} \\ &-\left|\sum_{n=\sigma}^{k}(-1)^{k-n}\rho_5(r_n)\right| \geq \sum_{n=\sigma}^{k}|\beta_n|p_n^{(1)} - \sum_{n=\sigma}^{k}|\rho_5(r_n)| \geq (1-2\theta)\sum_{n=\sigma}^{k}|\beta_n|p_n^{(1)} \geq 0.\end{aligned}$$

In conclusion it has been shown that

$$\sum_{n=\sigma}^{k} p_n^{(1)}|\beta_n| < \frac{2\varepsilon}{1-2\theta}. \tag{3.19}$$

The main argument for this case can now start, by showing that $\tau = \infty$. Suppose, by contradiction, $\tau < \infty$. Then we can use (3.19) with $k = \tau - 1$, and therefore by (3.9) in Lemma 3.21, Lemmas 3.23 to 3.26 we obtain, by the same estimate as in the previous case ($\sigma = \infty$; more precisely Lemma 3.26 (a) ensures $\sigma < \tau$, while Lemma 3.26 (b) ensures $(3/4)^{n-m}|\alpha_m| \leq p_n^{(1)}|\beta_n|$) and the definition of σ, that

$$\begin{aligned}|p_\sigma^{(2)} - p_\tau^{(2)}| &\leq \sum_{n=\sigma}^{\tau-1}|p_{n+1}^{(2)} - p_n^{(2)}| \leq 3\sum_{n=\sigma}^{\tau-1}p_n^{(1)}|\beta_n| + c\sum_{n=\sigma}^{\tau-1}|\alpha_n| \leq 3\sum_{n=\sigma}^{\tau-1}p_n^{(1)}|\beta_n| \\ &+ c\sum_{n=\sigma}^{\tau-1}\max\left\{\left(\frac{3}{4}\right)^{n-m}|\alpha_m|, p_n^{(1)}|\beta_n|\right\} = 3\sum_{n=\sigma}^{\tau-1}p_n^{(1)}|\beta_n| + c\sum_{n=\sigma}^{\tau-1}p_n^{(1)}|\beta_n| \\ &= (3+c)\sum_{n=\sigma}^{\tau-1}p_n^{(1)}|\beta_n| \leq \frac{(3+c)2\varepsilon}{1-2\theta} < \frac{\delta}{8}.\end{aligned}$$

But if $\tau < \infty$ by Lemma 3.26 (a), $p^{(2)}_\sigma \in [\delta/4, 1 - \delta/4]$, and since it has just been proved that $p^{(2)}_\tau$ has travelled a distance less than $\delta/8$, this yields that $p^{(2)}_\tau \in [\delta/8, 1-\delta/8]$ in contradiction with the definition of τ. Hence $\tau = \infty$. Then again, by (3.9) in Lemma 3.21, Lemmas 3.23 to 3.25 and Lemma 3.26 (b) it follows that

$$\begin{aligned}\sum_{n=\sigma}^{\infty} |p^{(2)}_{n+1} - p^{(2)}_n| &= \sum_{n=\sigma}^{\infty} |2(1-p^{(2)}_n)p^{(1)}_n\beta_n + \rho_2(r_n)| \le 3\sum_{n=\sigma}^{\infty} p^{(1)}_n|\beta_n| + c\sum_{n=\sigma}^{\infty}|\alpha_n| \\ &\le 3\sum_{n=\sigma}^{\infty} p^{(1)}_n|\beta_n| + c\sum_{n=\sigma}^{\infty}|\alpha_n| \le 3\sum_{n=\sigma}^{\infty} p^{(1)}_n|\beta_n| \\ &+ c\sum_{n=\sigma}^{\infty}\max\left\{\left(\frac{3}{4}\right)^{n-m}|\alpha_m|, p^{(1)}_n|\beta_n|\right\} = 3\sum_{n=\sigma}^{\infty} p^{(1)}_n|\beta_n| + c\sum_{n=\sigma}^{\infty} p^{(1)}_n|\beta_n| \\ &= (3+c)\sum_{n=\sigma}^{\infty} p^{(1)}_n|\beta_n| \le \frac{2\varepsilon(3+c)}{1-2\theta} \le \frac{\delta}{8} < \infty.\end{aligned}$$

The last step follows from (3.19) holding uniformly for any $k \ge \sigma$, since now $\tau = \infty$. This yields the convergence of $p^{(2)}_n$ within $[\delta/8, 1-\delta/8]$, and by the same reasoning as in the previous case ($\sigma = \infty$), exploiting ε being arbitrarily small, this results again in the convergence of p_n inside the edge.

In both cases, convergence of $\{p_n\}$ to $p_* \in E_1$ implies convergence of $\{q_n\}$ to q_{p_*}, since $\ell = 0$, and therefore convergence of the whole orbit. □

Remark 3.28. *Repeating this argument for $p_* \in E_i$ with $i \in \{2,3\}$, by exploiting the symmetry of the model, defining σ and τ accordingly in terms of the corresponding coordinates and showing an analogous version of Theorem 3.27 for $i \in \{2,3\}$, yields convergence of any orbit approaching the boundary with $\ell = 0$.*

3.5 Convergence to the boundary with $\ell > 0$

The main goal of this section is to show that if $\{p_n\}$ approaches the boundary and the limit of the potential $\ell := \lim_{n\to\infty} V(p_n, q_n) > 0$, the dynamical system converges. Compared with the case studied in Section 3.4, a key feature of the case with $\ell > 0$ is that the orbit does not admit a subsequence $\{p_{n_j}\}_{j\in\mathbb{N}}$ bounded away from the boundary, as otherwise we would have $\ell = 0$, due to the geometric decay of the decreasing potential along this subsequence, ensured by Proposition 3.15. This will simplify certain arguments, compared to those in Section 3.4. In particular, results the likes of Lemma 3.24 will not be needed. On the other hand, the lack of information about the set of accumulation points for the orbits (which is quite trivial when $\ell = 0$: they are the equilibria, including boundary ones) will complicate the analysis, which requires an explicit study of the set of accumulation points of the orbit.

3.5.1 Convergence of boundary orbits

Even though we seek to understand regular orbits, the description of their set of accumulation points, if they approach the boundary, comes from the study of boundary orbits, which can be reduced to the case of $p_n \in E_i$ for some $i \in \{1,2,3\}$, for all $n \in \mathbb{N}_0$, with $q_0 \in \Sigma_0 := \Sigma \setminus V$. The requirement on q_0 follows by Remarks 3.3 and 3.4: recall that if $p_0 \in E_i$ and $q_0 = v_k$ for $k \ne i$, the iteration is inconsistent, while if $q_0 = v_i$, then $p_n = p_0$ and $\{q_n\}$ is 2-periodic.

Lemma 3.29. *Let $p_0 \in E_i$ for some $i \in \{1,2,3\}$. Then there exists $p_* \in E_i$ such that $p_n \longrightarrow p_*$ as $n \longrightarrow \infty$.*

Proof. By symmetry, without loss of generality, assume $i = 1$, that is $p_0 \in E_1$, or equivalently $p_0^{(1)} = 0$ and $0 < p_0^{(2)} < 1$. Then

$$p_1 = \mathbf{1} - q_0 - \begin{pmatrix} 0 & 1 & 1 \\ 1 - p_0^{(2)} & 0 & 0 \\ p_0^{(2)} & 0 & 0 \end{pmatrix} q_0.$$

It follows immediately from the form taken by M_{p_0} that $p_1^{(1)} = 0$ and therefore, by induction, that for all $n \in \mathbb{N}_0$, $p_n^{(1)} = 0$. Thus we are in presence of a boundary orbit. We can also see that

$$p_1^{(2)} = 1 - q_0^{(2)} - (1 - p_0^{(2)})q_0^{(1)} > 0,$$

since $0 < p_0^{(2)} < 1$, and therefore, by induction, $0 < p_n^{(2)} < 1$ for all $n \in \mathbb{N}$. We can conclude that for all $n \in \mathbb{N}_0$,

$$M_{p_n} = \begin{pmatrix} 0 & 1 & 1 \\ 1 - p_n^{(2)} & 0 & 0 \\ p_n^{(2)} & 0 & 0 \end{pmatrix}.$$

Given these facts, we show convergence by estimating $p_{n+1}^{(2)} - p_n^{(2)}$. Note that

$$\begin{aligned} p_{n+1}^{(2)} - p_n^{(2)} &= 1 - (1 - p_n^{(2)})q_n^{(1)} - q_n^{(2)} - p_n^{(2)} = (1 - p_n^{(2)})(1 - q_n^{(1)}) - q_n^{(2)} \\ &= (1 - p_n^{(2)})(1 - q_n^{(1)}) - (1 - p_{n-1}^{(2)})q_{n-1}^{(1)}. \end{aligned}$$

Since $p_n^{(1)} = 0$, we have that $q_{n-1}^{(1)} = 1 - q_n^{(1)}$ for all $n \in \mathbb{N}$; therefore, factoring this quantity out, yields that

$$p_{n+1}^{(2)} - p_n^{(2)} = -(1 - q_n^{(1)})(p_n^{(2)} - p_{n-1}^{(2)}). \tag{3.20}$$

Note that (3.20) implies that

$$|p_{n+1}^{(2)} - p_n^{(2)}| \le |p_n^{(2)} - p_{n-1}^{(2)}|,$$

so the increments of $p_n^{(2)}$ are monotone decreasing with an alternating sign (whenever the sign does not alternate, we are in the trivial case of the increments being eventually identically zero, which is a trivial case of convergence). Note also that the argument so far makes no hypotheses on $q_0 \in \Sigma_0$. Iterating once more yields

$$p_{n+1}^{(2)} - p_n^{(2)} = (1 - q_n^{(1)})(1 - q_{n-1}^{(1)})(p_{n-1}^{(2)} - p_{n-2}^{(2)}) = q_n^{(1)}(1 - q_n^{(1)})(p_{n-1}^{(2)} - p_{n-2}^{(2)}),$$

which implies that

$$|p_{n+1}^{(2)} - p_n^{(2)}| \le \frac{1}{4}|p_{n-1}^{(2)} - p_{n-2}^{(2)}|$$

and thus $|p_{n+1}^{(2)} - p_n^{(2)}| \longrightarrow 0$. By the *Leibniz test* of convergence

$$\sum_{n=0}^{\infty}(p_{n+1}^{(2)} - p_n^{(2)}) < \infty,$$

hence $p_n^{(2)}$ converges to some $p_*^{(2)}$. It is left to show that $0 < p_*^{(2)} < 1$. Since $p_n \in E_1$ for all n and since $p_k^{(2)} - p_{k-1}^{(2)}$ alternates the sign while vanishing monotonically, by the formula

$$p_*^{(2)} - p_0^{(2)} = p_*^{(2)} - p_n^{(2)} + \sum_{k=1}^{n}(p_k^{(2)} - p_{k-1}^{(2)}) = o(1) + \sum_{k=1}^{n}(p_k^{(2)} - p_{k-1}^{(2)}), \tag{3.21}$$

which holds since $|p_*^{(2)} - p_n^{(2)}| \longrightarrow 0$ as $n \longrightarrow \infty$, the claim follows by exhaustion of the following three cases.

- If $p_1^{(2)} = p_0^{(2)}$, by (3.20) all increments are identically zero, so $p_n^{(2)} = p_0^{(2)}$ for all n and therefore $p_n = p_0$ for all n, which is the trivial case of convergence to $p_* = p_0 \in E_1$ previously mentioned.
- If $p_1^{(2)} > p_0^{(2)}$, due to the monotonicity and alternating of the signs,

$$0 \leq \sum_{k=1}^{n} p_k^{(2)} - p_{k-1}^{(2)} \leq p_1^{(2)} - p_0^{(2)},$$

  and since n can be chosen arbitrarily large in (3.21), this implies that

$$0 \leq p_*^{(2)} - p_0^{(2)} \leq p_1^{(2)} - p_0^{(2)},$$

  hence $p_* \in E_1$.
- If $p_1^{(2)} < p_0^{(2)}$, due to the monotonicity and alternating of the signs,

$$p_1^{(2)} - p_0^{(2)} \leq \sum_{k=1}^{n} p_k^{(2)} - p_{k-1}^{(2)} \leq 0,$$

  and since n can be chosen arbitrarily large in (3.21), this implies that

$$p_1^{(2)} - p_0^{(2)} \leq p_*^{(2)} - p_0^{(2)} \leq 0,$$

  hence $p_* \in E_1$.

□

Fundamental to the key geometric ideas behind Corollary 3.30 is that for all $p \in E_i$, $e_0(p)$ is constant and always parallel to the edge E_i (see Lemma 3.19 (f)). Recall that as $e_1(p)$ we can take q_p (see Lemma 3.19 (a)).

Corollary 3.30. *If $p_0 \in E_i$, then for all $q_0 \in \Sigma_0$ there exists $\lim_{n\longrightarrow\infty} p_n =: p_* \in E_i$ and $\beta \geq 0$ (dependent on the initial conditions) such that the set of accumulation points of the boundary orbit is $\{(p_*, q_{p_*} \pm \beta e_{-1}(p_*))\}$. Moreover, letting $\ell := \lim_{n\longrightarrow\infty} V(p_n, q_n)$ yields $\beta = \ell/2$.*

Proof. By symmetry, without loss of generality set $i = 1$. Note that

$$\Phi(p,q) := \begin{pmatrix} \mathbf{1} - q - M_p q \\ M_p q \end{pmatrix}$$

is continuous on $E_1 \times \Sigma_0$. By Lemma 3.29 as $n \longrightarrow \infty$,

$$(I + M_{p_n})q_n = \mathbf{1} - p_{n+1} \longrightarrow \mathbf{1} - p_* = 2q_{p_*}$$

so

$$\frac{I + M_{p_n}}{2} q_n \longrightarrow q_{p_*}$$

and therefore

$$\frac{I + M_{p_*}}{2}(q_n - q_{p_*}) \longrightarrow \mathbf{0},$$

since

$$\left(\frac{I + M_{p_n}}{2} - \frac{I + M_{p_*}}{2}\right) q_n + \frac{I + M_{p_*}}{2} q_n = \frac{I + M_{p_n}}{2} q_n \longrightarrow q_{p_*} = \frac{I + M_{p_*}}{2} q_{p_*},$$

and by continuity

$$\frac{I + M_{p_n}}{2} - \frac{I + M_{p_*}}{2} \longrightarrow \mathbf{0}.$$

This means that $q_n - q_{p_*}$ either becomes aligned with $e_{-1}(p_*)$ or it vanishes as $n \longrightarrow \infty$ (in which case $\beta = 0$). Note that $\|q_n - q_{p_*}\|_1$ converges by the monotonicity of the potential, which, although not strictly, still applies to boundary orbits by Remark 3.13. Since $V(p_n, q_n) \longrightarrow \ell \geq 0$, $\|q_n - q_{p_*}\|_1 \longrightarrow \ell$, which determines β, because $\|e_{-1}(p_*)\|_1 = 2$ for all $p_* \in E_1$. If $\ell = 0$, $\beta = 0$. Indeed, if $\ell > 0$, $\|q_n - q_{p_*}\|_1 = \|\beta_n e_{-1}(p_*) + o(\mathbf{1})\|_1 \longrightarrow \ell > 0$. If $\beta_n \longrightarrow 0$, we would have a contradiction with $\ell > 0$, so $\beta_n \not\longrightarrow 0$. If $|\beta_n|$ does not converge, by boundedness there would be two convergent subsequences $|\beta_{n_j}| \longrightarrow \tilde{\beta}$ and $|\beta_{n_k}| \longrightarrow \bar{\beta}$, with $\tilde{\beta} \neq \bar{\beta}$. But $\|\beta_{n_j} e_{-1}(p_*) + o(\mathbf{1})\|_1 \longrightarrow \|\tilde{\beta} e_{-1}(p_*)\|_1 = 2\tilde{\beta}$ and $\|\beta_{n_k} e_{-1}(p_*) + o(\mathbf{1})\|_1 \longrightarrow \|\bar{\beta} e_{-1}(p_*)\|_1 = 2\bar{\beta}$. Both limit values must be equal to ℓ, hence $\tilde{\beta} = \bar{\beta} = \ell/2$. Hence $\{|\beta_n|\}$ converges to $\beta = \ell/2 > 0$, and thus its accumulation points are $\{\pm\beta\}$. This yields the claim. □

We can now exploit Lemma 3.29 to show that for a boundary orbit, $\{q_n\}$ is asymptotically 2-periodic whenever $q_0 - q_{p_0}$ is not aligned with $e_0(p_0)$.

Proposition 3.31. *If $p_0 \in E_i$, then for all $q_0 \in \Sigma_0$ there exists $\lim_{n\longrightarrow\infty} p_n = p_* \in E_i$ and $\beta \geq 0$ (dependent on the initial conditions) such that the boundary orbit approaches the 2-cycle $\{(p_*, q_{p_*} \pm \beta e_{-1}(p_*))\}$, with $\beta = |\beta_0|$, where $q_0 = q_{p_0} + \alpha_0 e_0(p_0) + \beta_0 e_{-1}(p_0)$.*

Proof. By Corollary 3.30 it is known that the set of accumulation points is as per the claim, we only need to show that the orbit alternates between the two accumulation points. We will adopt the boundary eigenvectors as per Lemma 3.19 and, by symmetry, we will assume $p_* \in E_1$, without loss of generality. Recall that $e_0(p_n) = e_0 := (0, 1, -1)$ for all $n \in \mathbb{N}_0$ and $e_{-1}(p_n) = (-1, 1 - p_n^{(2)}, p_n^{(2)})$, so that

$$e_{-1}(p_{n+1}) - e_{-1}(p_n) = (p_{n+1}^{(2)} - p_n^{(2)}) e_0.$$

Then since $q_{p_n} = e_1(p_n)$ and $q_n - q_{p_n} = \alpha_n e_0 + \beta_n e_{-1}(p_n)$ for all n, for a boundary orbit (3.2) reads as

$$\alpha_{n+1} e_0 + \beta_{n+1} e_{-1}(p_{n+1}) = -\frac{\alpha_n}{2} e_0 - \beta_n e_{-1}(p_n),$$

which yields

$$\begin{aligned} -\beta_{n+1} &= \beta_n \\ \alpha_{n+1} + (1 - p_{n+1}^{(2)})\beta_{n+1} &= -\frac{\alpha_n}{2} - (1 - p_n^{(2)})\beta_n \\ -\alpha_{n+1} + p_{n+1}^{(2)}\beta_{n+1} &= \frac{\alpha_n}{2} - p_n^{(2)}\beta_n. \end{aligned}$$

The first equation plugged into the second makes the latter a scalar multiple of the third equation, so the system is consistent and overdetermined, and we solve it by keeping only the first and the third equation, and use the first equation to simplify the third, obtaining

$$\alpha_{n+1} = -\frac{\alpha_n}{2} + \beta_n (p_n^{(2)} - p_{n+1}^{(2)}) \tag{3.22}$$

$$\beta_{n+1} = -\beta_n. \tag{3.23}$$

Using (3.1) we can also rewrite $p_{n+1} - p_n$ in eigencoordinates as

$$\begin{aligned} p_{n+1} - p_n &= -2(q_{p_{n+1}} - q_{p_n}) = -2\left(\frac{q_{n+1}+q_n}{2} - q_{p_n}\right) = -((M_{p_n}+I)q_n - 2q_{p_n}) \\ &= -((M_{p_n}+I)q_n - (M_{p_n}+I)q_{p_n}) = -(M_{p_n}+I)(q_n - q_{p_n}) = -\alpha_n e_0. \end{aligned}$$

Due to $(M_{p_n} + I)e_{-1}(p_n) = \mathbf{0}$, we obtain $p^{(2)}_{n+1} - p^{(2)}_n = -\alpha_n$, which turns (3.22) and (3.23) into the system

$$\alpha_{n+1} = \left(\beta_n - \frac{1}{2}\right)\alpha_n \tag{3.24}$$

$$\beta_{n+1} = -\beta_n. \tag{3.25}$$

By (3.25), β_n oscillates between β_0 and $-\beta_0$. If we can show that $\alpha_n \longrightarrow 0$, we have the asymptotic 2-periodicity claimed. We will actually show the stronger fact, that the asymptotic 2-cycle is approached with a geometric decay of the eigencoordinate α_n. Since $\|e_{-1}(p)\|_2 = \sqrt{2}\sqrt{1 - p^{(2)}_n + (p^{(2)}_n)^2}$ and the parabola $3/4 \le x^2 - x + 1 < 1$ for $0 < x < 1$ (the minimum is at $1/2$),

$$\sqrt{\frac{3}{2}} \le \|e_{-1}(p)\|_2 \le \sqrt{2},$$

$$|\beta_n|\sqrt{\frac{3}{2}} \le |\beta_n|\|e_{-1}(p_n)\| = \|\beta_n e_{-1}(p_n)\|_2 \le \frac{1}{\sqrt{2}}$$

where the last inequality on the right comes from the geometry of the simplex: since $e_0(p_n)$ is parallel to E_1, if $\|\beta_n e_{-1}(p_n)\|_2 > 1/\sqrt{2}$, then $q_n \notin \Sigma$. As a brief explanation of this elementary geometric fact, note that for boundary orbits with $p_n \in E_1$, q_{p_n} is on the line, parallel to E_1, joining $(1/2, 0, 1/2)$ to $(1/2, 1/2, 0)$; this line passes through the midpoint of $(1,0,0) + e_{-1}(p_n)$, hence $\|\beta_n e_{-1}(p_n)\|_2 < \sqrt{2}/2$, where $\sqrt{2}$ is the Euclidean length of the edge of the simplex (half of the edge being the largest, which the projection $\|\beta_n e_{-1}(p_n)\|_2$ can be, due to the geometric properties of the simplex). Therefore $|\beta_n| < 1/\sqrt{3}$. Thus we use this estimate in the two-steps iteration obtained for α_n, by iterating (3.24) once more and plugging (3.25) in it. This yields that

$$\begin{aligned} \alpha_{n+1} &= \left(\beta_n - \frac{1}{2}\right)\left(\beta_{n-1} - \frac{1}{2}\right)\alpha_{n-1} = -\left(\beta_{n-1} + \frac{1}{2}\right)\left(\beta_{n-1} - \frac{1}{2}\right)\alpha_{n-1} \\ &= -\left(\beta^2_{n-1} - \frac{1}{4}\right)\alpha_{n-1}. \end{aligned}$$

It follows that

$$|\alpha_{n+1}| < \left|\beta^2_{n-1} - \frac{1}{4}\right| |\alpha_{n-1}| < \frac{1}{12}|\alpha_{n-1}|. \tag{3.26}$$

Denote $M := \max\{|\alpha_0|, |\alpha_1|\}$, then it is easy to see that by induction,

$$|\alpha_{2k}| < \left(\frac{1}{12}\right)^k |\alpha_0|$$

and

$$|\alpha_{2k+1}| < \left(\frac{1}{12}\right)^k |\alpha_1|,$$

so a common geometrically decaying upper bound can be found as

$$|\alpha_n| < \left(\frac{1}{12}\right)^{\lfloor \frac{n}{2} \rfloor} M,$$

hence

$$|\alpha_n| < M\left(\frac{1}{2\sqrt{3}}\right)^n, \tag{3.27}$$

yielding a geometrically decaying upper bound on the first eigencoordinate. That $\beta = \beta_0$, is obvious by the alternating β_n coordinate. In particular, $\beta = 0$ if and only if $q_0 = q_{p_0} + \alpha_0 e_0(p_0)$ for some real α_0 (hence, these are the initial conditions of boundary orbits having $\ell = 0$). □

Remark 3.32. *The geometric upper bound on the decay of $|\alpha_n|$ for a boundary orbit, expressed in (3.27), is uniform, since $M := \max\{|\alpha_0|, |\alpha_1|\}$ holds uniformly by the boundedness of the simplex. Hence there exists a uniform constant $\tilde{M}$ such that, for any boundary orbit,*

$$|\alpha_n| < \tilde{M}\left(\frac{1}{2\sqrt{3}}\right)^n \tag{3.28}$$

The following is immediate from Proposition 3.31. Recall that Σ^* denotes the medial triangle in Σ (boundary excluded).

Corollary 3.33. *If $p_0 \in E_i$ and $q_0 = q_{p_0} + \alpha_0 e_0(p_0) + \beta_0 e_{-1}(p_0) \in \Sigma_0$, $p_n \longrightarrow p_* \in E_i$, while $\{q_n\}$ either converges to $q_{p_*} \in \partial\Sigma^*$ if $\beta_0 = 0$ or approaches the 2-cycle $\{(p_*, q_{p_*} \pm \beta e_{-1}(p_*))\}$, with $\beta = |\beta_0|$.*

3.5.2 Structure of the set of accumulation points

In this section we study the set of accumulation points of regular orbits approaching the boundary, and therefore assume $p_0 \notin \partial\Sigma$. It will be useful to describe the dynamical system in terms of its iteration map $\Phi(p, q)$, which is continuous on Σ_0^2:

$$\begin{pmatrix} p_{n+1} \\ q_{n+1} \end{pmatrix} = \Phi(p_n, q_n) = \begin{pmatrix} \Phi_p(p_n, q_n) \\ \Phi_q(p_n, q_n) \end{pmatrix} := \begin{pmatrix} \mathbf{1} - q_n - M_{p_n} q_n \\ M_{p_n} q_n \end{pmatrix}$$

Boundary orbits are the main tool to finding the structure of the set of accumulation points of regular orbits characterised by $\ell > 0$. Not only, but for a truly informative characterisation, particular attention must be paid to the asymptotics of a regular orbit approaching the set of vertices V. To this end, it is useful to define, for any $i \neq j \neq k$, $e_{-1}(v_i) := v_j - v_k$. Even though the matrix M_{v_i} is not well defined, $e_{-1}(v_i)$ artificially defined as such, will play the role of the corresponding eigenvector $e_{-1}(p)$ (well defined on the edges), in capturing the direction of the asymptotic oscillations of $\{q_n\}$, for an orbit such that $\{p_n\}$ approaches the vertices. We will see that if $\{p_n\}$ approaches the vertices, the orbit has, along a subsequence, an asymptotics analogous to that of boundary orbits.

Lemma 3.34. *Let $\{(p_n, q_n)\}$ be an orbit, such that $\{p_n\}$ is not bounded away from a vertex, that is such that there is $\{n_k\}_{k\in\mathbb{N}}$, with $p_{n_k} \longrightarrow v_i$ for some $i \in \{1, 2, 3\}$. Then the set of accumulation points of $\{(p_{n_k}, q_{n_k})\}$ and $\{(p_{n_k+1}, q_{n_k+1})\}$ is a subset of $\{(v_i, q_{v_i} \pm \frac{\ell}{2} e_{-1}(v_i))\}$. Moreover, if $\{n_k\}$ is such that also $\{q_{n_k}\}$ converges, $\{q_{n_k-1}\}$, $\{q_{n_k}\}$, and $\{q_{n_k+1}\}$ asymptotically oscillate between $q_* = q_{v_i} \pm \frac{\ell}{2} e_{-1}(v_i)$ and $\hat{q}_* = q_{v_i} \mp \frac{\ell}{2} e_{-1}(v_i)$, while $\{p_{n_k}\}$ and $\{p_{n_k+1}\}$ all tend to v_i, that is $(p_{n_k}, q_{n_k}) \longrightarrow (v_i, q_*)$ and $(p_{n_k+1}, q_{n_k+1}) \longrightarrow (v_i, \hat{q}_*)$.*

Proof. By symmetry assume $i = 2$, without loss of generality. Let an arbitrary convergent subsubsequence $(p_{n_{k_l}}, q_{n_{k_l}}) \longrightarrow (v_2, q_*)$ as $l \longrightarrow \infty$. We do not know if there actually are convergent subsequences $\{p_{n_{k_l}}\}$ that tend to the vertices (we will later conjecture that there are not), but we will show that if there are, they must have the same type of accumulation points as those of boundary orbits, by determining the form of q_*, for now left unknown. Relabel as $\{n_r\}$ the subsubsequence $\{n_{k_l}\}$. Note that

$$q_{n_r} + q_{n_r - 1} = \mathbf{1} - p_{n_r} \longrightarrow (1, 0, 1),$$

which implies that

$$q^{(2)}_{n_r} \longrightarrow 0 = q^{(2)}_*,$$

while for $i \in \{1, 3\}$,

$$q^{(i)}_{n_r} + q^{(i)}_{n_r - 1} \longrightarrow 1$$

as $r \longrightarrow \infty$. Since $V(p_{n_r}, q_{n_r}) \longrightarrow V(v_2, q_*) = \ell$ one can fully determine q_*. Indeed $q_{v_2} = (1/2, 0, 1/2)$ and

$$\ell = V(v_2, q_*) = |q^{(1)}_* - q^{(1)}_{v_2}| + |q^{(2)}_* - q^{(2)}_{v_2}| + |q^{(1)}_* - q^{(1)}_{v_2}| = 2|q^{(1)}_* - q^{(1)}_{v_2}| = 2|q^{(3)}_* - q^{(3)}_{v_2}|,$$

thus

$$\left| q^{(3)}_* - \frac{1}{2} \right| = \frac{\ell}{2}.$$

Hence either $q^{(3)}_* = (1+\ell)/2$ or $q^{(3)}_* = (1-\ell)/2$, that is either $q_* = ((1+\ell)/2, 0, (1-\ell)/2)$ or $q_* = ((1-\ell)/2, 0, (1+\ell)/2)$. Without loss of generality, assume the latter scenario. Since it has been shown that

$$q^{(i)}_{n_r} + q^{(i)}_{n_r - 1} \longrightarrow \begin{cases} 0, & i = 2 \\ 1, & i \neq 2, \end{cases}$$

trivially $q_{n_r - 1} \longrightarrow ((1 + \ell)/2, 0, (1 - \ell)/2) =: \hat{q}_*$, the complementary form of q_*. Slightly less trivially, the same holds for $q_{n_r + 1}$ (this part, technically not necessary at this point, will be of use when starting from an arbitrary convergent subsubsequence of $\{(p_{n_k+1}, q_{n_k+1})\}$, but we anticipate it now, so as to avoid repeating the start of the argument twice). As $r \longrightarrow \infty$,

$$q^{(1)}_{n_r+1} = \frac{p^{(3)}_{n_r}}{p^{(1)}_{n_r} + p^{(3)}_{n_r}} q^{(2)}_{n_r} + \frac{p^{(2)}_{n_r}}{p^{(1)}_{n_r} + p^{(2)}_{n_r}} q^{(3)}_{n_r} \longrightarrow \frac{1+\ell}{2},$$

because even though its limit does not exist,

$$0 \le \frac{p^{(3)}_{n_r}}{p^{(1)}_{n_r} + p^{(3)}_{n_r}} \le 1,$$

and multiplied by $q^{(2)}_{n_r} \longrightarrow 0$, the term vanishes, while

$$\frac{p^{(2)}_{n_r}}{p^{(1)}_{n_r} + p^{(2)}_{n_r}} \longrightarrow 1$$

and $q^{(3)}_{n_r} \longrightarrow (1 + \ell)/2$. As $r \longrightarrow \infty$,

$$q^{(2)}_{n_r+1} = \frac{p^{(3)}_{n_r}}{p^{(2)}_{n_r} + p^{(3)}_{n_r}} q^{(1)}_{n_r} + \frac{p^{(1)}_{n_r}}{p^{(1)}_{n_r} + p^{(2)}_{n_r}} q^{(3)}_{n_r} \longrightarrow 0,$$

since both terms next to $q^{(1)}_{n_r}$ and $q^{(3)}_{n_r}$ vanish. Therefore, $q_{n_r+1} \longrightarrow \hat{q}_*$, and as a result, also $p_{n_r+1} \longrightarrow v_2$ and $q_{p_{n_r+1}} \longrightarrow q_{v_2} = (1/2, 0, 1/2)$ as $r \longrightarrow \infty$.

We only showed the form of the possible accumulation points for $\{(p_{n_k}, q_{n_k})\}$, but that it is the same for $\{(p_{n_k+1}, q_{n_k+1})\}$ follows by simply noting that, if we started with an arbitrary convergent subsubsequence denoted as $(p_{n_r+1}, q_{n_r+1}) \longrightarrow (p_*, \hat{q}_*)$, with p_* and $\hat{q}_*$ to be determined; knowing by hypothesis that $p_{n_r} \longrightarrow v_2$, would imply, through (2.34), that $q_{n_r} \longrightarrow \mathbf{1} - v_2 - \hat{q}_* =: q_*$. Thus we obtain $(p_{n_r}, q_{n_r}) \longrightarrow (v_2, q_*)$ and we can proceed, through the same argument as above, with showing that the form of q_* and $\hat{q}_*$ is as per the claim, thus obtaining that $p_* = v_2$. The second part of the claim trivially follows by taking $n_r = n_k$ in the argument above. □

Next we show that regular orbits such that $\{p_n\}$ approaches the boundary with $\ell > 0$, behave asymptotically like boundary orbits. Recall that if $\ell > 0$, an orbit cannot have any subsequence bounded away from the boundary $\partial\Sigma$.

Proposition 3.35. *Let $\{(p_n, q_n)\}$ be an orbit such that $\ell > 0$. The set of accumulation points of the orbit is a subset of*

$$\{(p, q_p \pm \beta e_{-1}(p)) : \ p \in \partial\Sigma, \ \beta > 0, \ V(p, q_p \pm \beta e_{-1}(p)) = \ell\} \, .$$

Proof. Recall that if the orbit is not bounded away from the boundary and $\ell > 0$, then for every $\{n_k\}$, $p_{n_k} \longrightarrow \partial\Sigma$, since there is no subsequence bounded away from the boundary. By boundedness, consider $(p_{n_k}, q_{n_k}) \longrightarrow (p_0^*, q_0^*)$ as $k \longrightarrow \infty$. If $p_0^* \in V$ by Lemma 3.34, one gets the type of limit points claimed, with $\beta = \ell/2$. Hence let us assume $p_0^* \notin V$. Then $p_0^* \in E_i$ for some $i \in \{1, 2, 3\}$. By symmetry, without loss of generality, assume $i = 1$. We need to prove that q_0^* is of the form $q_{p_0^*} \pm \beta e_{-1}(p_0^*)$, where $\beta > 0$ is such that $V(p_0^*, q_{p_0^*} \pm \beta e_{-1}(p_0^*)) = \ell$. Consider then $\{(p_{n_k-1}, q_{n_k-1})\}$, which has a convergent subsequence $(p_{n_{k_s}-1}, q_{n_{k_s}-1}) \longrightarrow (p_1^*, q_1^*)$ as $s \longrightarrow \infty$. For simplicity, denote $n_r := n_{k_s} - 1$. The key idea is that also $p_1^* \notin V$, as by Lemma 3.34, if $p_1^* = v_i$, for some $i \in \{1, 2, 3\}$, $(p_{n_r}, q_{n_r}) \longrightarrow (v_i, q_1^*)$ as $r \longrightarrow \infty$, where $q_1^* - q_{v_i}$ is in the eigenspace spanned by $e_{-1}(v_i)$; then it must follow that $(p_{n_r+1}, q_{n_r+1}) \longrightarrow (v_i, \hat{q}_1^*)$, where $\hat{q}_1^*$ denotes the point complementary to q_1^*, as described in the proof of Lemma 3.34. The contradiction is the following: since $n_r + 1 = n_{k_s}$, on one hand we concluded that $p_{n_{k_s}} \longrightarrow v_i \in V$, whereas it was assumed $p_{n_k} \longrightarrow p_0^* \notin V$.

We obtained that p_1^* is not a vertex. Thus using the fact that Φ is continuous at (p_1^*, q_1^*) yields

$$\begin{aligned}\Phi(p_1^*, q_1^*) &= \Phi\left(\lim_{r\to\infty} (p_{n_r}, q_{n_r})\right) = \lim_{r\to\infty} \Phi\,(p_{n_r}, q_{n_r}) = \lim_{r\to\infty} (p_{n_r+1}, q_{n_r+1}) = \lim_{r\to\infty} (p_{n_{k_s}}, q_{n_{k_s}}) \\ &= (p_0^*, q_0^*),\end{aligned}$$

which also means that $p_1^* \in E_1$, being on the boundary. This argument can be iterated and applied to $(p_{n_r}, q_{n_r}) \longrightarrow (p_1^*, q_1^*)$ as $r \longrightarrow \infty$; that is, we consider a convergent subsequence of $\{(p_{n_r-1}, q_{n_r-1})\}$, $(p_{n_{r_j}-1}, q_{n_{r_j}-1}) \longrightarrow (p_2^*, q_2^*)$ as $j \longrightarrow \infty$, which can be relabelled as $n_l := n_{r_j} - 1$ and then, via the same argument, prove that $\Phi(p_2^*, q_2^*) = (p_1^*, q_1^*)$. This yields $\Phi^2(p_2^*, q_2^*) = (p_0^*, q_0^*)$. The iteration of this procedure through subsequences m times, produces the finite segment of a boundary orbit, as $\Phi^m(p_m^*, q_m^*) = (p_0^*, q_0^*)$, and $p_m^* \in E_1$ for any $m \in \mathbb{N}$.

Finally we show that q_0^* is on the eigenspace spanned by $e_{-1}(p_0^*)$, which will be done by using the same eigencoordinates as in the proof of Proposition 3.31. Recall that, as a consequence of the notation set at the beginning of this section, via the iteration map $\Phi = (\Phi_p, \Phi_q)$, for all $0 \leq k \leq m$,

$$\Phi^k = \begin{pmatrix} (\Phi^k)_p \\ (\Phi^k)_q \end{pmatrix},$$

hence

$$(\Phi^m)_q(p^*_m, q^*_m) - q_{(\Phi^m)_p(p^*_m,q^*_m)} = q^*_0 - q_{p^*_0} = \alpha_0 e_0 + \beta_0 e_{-1}(p^*_0).$$

We can choose m arbitrarily large, with $p^*_m \in E_1$ and $q^*_m \in \Sigma$. Thus by Remark 3.32, which ensures a uniform geometric upper bound on the decay of the α-eigencoordinate, iterating m times from $q^*_m - q_{p^*_m}$ to $q^*_0 - q_{p^*_0}$, will imply that $\alpha_0 = 0$. Indeed by Remark 3.32, after m iterates, $|\alpha_0| < \tilde{M}(2\sqrt{3})^{-m}$ for every m, and m grows as we iterate the previous argument, so that this upper bound vanishes uniformly, implying that $\alpha_0 = 0$ as m is arbitrary. The claim now follows trivially. □

Note that if $\{(p_n, q_n)\}$ is such that $\ell = 0$, Proposition 3.35 follows trivially by taking $\beta = 0$ and $\ell = 0$, that is the accumulation points are in $\{(p, q_p) : p \in \Sigma\}$.

Remark 3.36. *Consider an orbit such that $\ell > 0$ and, for some $\{n_k\}_{k\in\mathbb{N}}$, $p_{n_k} \longrightarrow p_* \in E_i$ for some $i \in \{1, 2, 3\}$. The set of accumulation points of $\{(p_{n_k}, q_{n_k})\}$ and $\{(p_{n_k+1}, q_{n_k+1})\}$ is a subset of $\{(p_*, q_{p_*} \pm \beta e_{-1}(p_*)) : \beta > 0,\ V(p_*, q_{p_*} \pm \beta e_{-1}(p_*)) = \ell\}$. Moreover if $\{n_k\}$ is such that also $\{q_{n_k}\}$ converges, $\{q_{n_k}\}$ and its shift $\{q_{n_k+1}\}$ asymptotically oscillate between $q_* = q_{p_*} \pm \beta e_{-1}(p_*)$ and $\hat{q}_* = q_{p_*} \mp \beta e_{-1}(p_*)$, that is if $(p_{n_k}, q_{n_k}) \longrightarrow (p_*, q_*)$, then $(p_{n_k+1}, q_{n_k+1}) \longrightarrow (p_*, \hat{q}_*)$ as $k \longrightarrow \infty$.*

Proof. We will exploit the continuity of Φ on $E_i \times \Sigma$. Starting with $p_{n_k} \longrightarrow p_* \in E_i$, we must extract any convergent subsubsequence $\{(p_{n_{k_l}}, q_{n_{k_l}})\}_{l\in\mathbb{N}}$ and check its limit. Relabel it with $\{n_r\}$ for simplicity. For the q-component's shift we have, by Proposition 3.35, that $q_* = q_{p_*} \pm \beta e_{-1}(p_*)$. Since

$$q_{p_*} \mp \beta e_{-1}(p_*) = \hat{q}_* := M_{p_*} q_* = M_{p_*}(q_{p_*} \pm \beta e_{-1}(p_*)),$$

by (2.33); the continuity and the hypothesis $p_{n_r} \longrightarrow p_* \in E_i$ imply that if $q_{n_r} \longrightarrow q_*$, then $q_{n_r+1} = M_{p_{n_r}} q_{n_r} \longrightarrow \hat{q}_*$. Hence if we start with an arbitrary subsubsequence $\{(p_{n_r+1}, q_{n_r+1})\}_{r\in\mathbb{N}}$, convergent to some $(p, \hat{q}_*)$ to be determined under the hypothesis given, having $p_{n_r} \longrightarrow p_* \in E_i$, by (2.34) it follows that $q_{n_r} \longrightarrow \mathbf{1} - p_* - q =: q_*$. We can repeat the argument above by applying Proposition 3.35 and (2.33) to show the claimed form of q_* and $\hat{q}_*$. For the p-component's shift, since $q_* + \hat{q}_* = 2q_{p_*}$, if $p_{n_r} \longrightarrow p_*$, having already shown that $q_{n_r} \longrightarrow q_*$ and $q_{n_r+1} \longrightarrow \hat{q}_*$, by (2.34) it follows that

$$p_{n_r+1} = \mathbf{1} - q_{n_r+1} - q_{n_r} \longrightarrow \mathbf{1} - \hat{q}_* - q_* = \mathbf{1} - 2q_{p_*} = p_*.$$

Hence if we start with an arbitrary subsubsequence $\{(p_{n_r+1}, q_{n_r+1})\}_{r\in\mathbb{N}}$ convergent to some $(p, \hat{q}_*)$, with only p left to be determined, it follows that $p = p_*$. The rest of the claim is trivial by taking $n_r = n_k$. □

Corollary 3.37. *Let $\{(p_n, q_n)\}$ be an orbit. Then $p_{n+1} - p_n \longrightarrow \mathbf{0}$ as $n \longrightarrow \infty$.*

Proof. If $\ell = 0$ the claim follows from (3.4) in Proposition 3.15. Assume $\ell > 0$ and denote $d_n := p_{n+1} - p_n$. The claim is equivalent to showing that $d_n \longrightarrow \mathbf{0}$. Since d_n is bounded, if every convergent subsequence d_{n_k} converges to $\mathbf{0}$, then d_n converges to $\mathbf{0}$. Consider then a convergent subsequence $d_{n_k} \longrightarrow d$. We will now show that by Lemma 3.34 and Remark 3.36, $d = \mathbf{0}$. There are in fact two cases, depending on whether $\{(p_{n_k}, q_{n_k})\}$ converges or not.

- If $\{(p_{n_k}, q_{n_k})\}$ converges, it could be that $\{p_{n_k}\}$ tends to a vertex or to a point inside an edge. If it is a vertex, $p_{n_k} \longrightarrow v_i \in V$ for some $i \in \{1, 2, 3\}$ by Lemma 3.34 and $p_{n_k+1} \longrightarrow v_i$ too, as $k \longrightarrow \infty$, and therefore $d_{n_k} \longrightarrow \mathbf{0}$. If it is not a vertex, $p_{n_k} \longrightarrow p_* \in E_i$ for some $i \in \{1, 2, 3\}$, then by Remark 3.36, $p_{n_k+1} \longrightarrow p_*$ too. Hence $d_{n_k} \longrightarrow \mathbf{0}$. In any case, $d = \mathbf{0}$.

- If $\{(p_{n_k}, q_{n_k})\}$ does not converge, by boundedness one can pick a subsubsequence $\{(p_{n_{k_r}}, q_{n_{k_r}})\}$ that does converge. Since $d_{n_k} \longrightarrow d$, also $d_{n_{k_r}} \longrightarrow d$ as $r \longrightarrow \infty$. However, the previous argument applies to $d_{n_{k_r}}$, since $(p_{n_{k_r}}, q_{n_{k_r}})$ converges, thus falling back in the previous case. Thus $d_{n_{k_r}}$ vanishes as $r \longrightarrow \infty$. It follows that $d = \mathbf{0}$ by the uniqueness of the limit.

□

Remark 3.38. *Let $\{(p_n, q_n)\}$ be an orbit with $\{p_n\}$ not convergent to any of the vertices. By Corollary 3.37 and Remark 3.11, there is a subsequence of $\{p_n\}$ bounded away from V.*

The following claim is trivially true if $\ell = 0$.

Corollary 3.39. *Let $\{(p_n, q_n)\}$ be an orbit such that $\ell > 0$, that is $V(p_n, q_n) = \|\alpha_n e_0(p_n) + \beta_n e_{-1}(p_n)\|_1 \longrightarrow \ell > 0$. Then $\alpha_n \longrightarrow 0$ and $|\beta_n| \longrightarrow \ell/2$ as $n \longrightarrow \infty$.*

Proof. For the first part of the statement, following as always the notation of Lemma 3.19, consider that in eigencoordinates (3.10) holds, which we rearrange as

$$p_{n+1} - p_n = -\alpha_n(1 + \lambda_0(p_n))e_0(p_n) - \beta_n(1 + \lambda_{-1}(p_n))e_{-1}(p_n),$$

and consider that as $p_n \longrightarrow \partial\Sigma$, $1 + \lambda_{-1}(p_n) \longrightarrow 0$ and $1 + \lambda_0(p_n) \longrightarrow 1$. As a direct result of Corollary 3.37 and Lemma 3.19 (h), one sees that $\alpha_n \longrightarrow 0$ as $n \longrightarrow \infty$ (this is trivially true also if $\ell = 0$).

As to the second part of the statement, since $\{|\beta_n|\}$ is bounded, consider any convergent subsequence $|\beta_{n_j}| \longrightarrow \ell'$. Assume by contradiction that $\ell' \neq \ell/2$. Since $\{p_{n_j}\}$ is bounded, extract a convergent subsubsequence $\{p_{n_{j_l}}\}$. Relabel it with $\{n_k\}$ for simplicity. Since the potential limit along this orbit is $\ell > 0$, we have that $p_{n_k} \longrightarrow p_* \in \partial\Sigma$ by Proposition 3.15, under the assumption that $|\beta_{n_k}| \longrightarrow \ell' \neq \ell/2$. If $p_* \in V$, then by Lemma 3.34, $|\beta_{n_k}| \longrightarrow \ell/2$, and therefore $p_* \in E_i$ for some $i \in \{1, 2, 3\}$. By symmetry, without loss of generality, assume $i = 1$. Then by the smoothness of the eigenvectors proved in Lemma 3.19 (h), it is known that $e_0(p_{n_k}) \longrightarrow (0, 1, -1) = e_0(p_*)$, with $e_0(p_{n_k}) - e_0(p_*) = \boldsymbol{\mathcal{O}(p_{n_k} - p_*)}$ and $e_{-1}(p_{n_k}) \longrightarrow (-1, 1 - p_*^{(2)}, p_*^{(2)}) = e_{-1}(p_*)$, with $e_{-1}(p_{n_k}) - e_{-1}(p_*) = \boldsymbol{\mathcal{O}(p_{n_k} - p_*)}$. Thus

$$\begin{aligned}
&\|\alpha_{n_k} e_0(p_{n_k}) + \beta_{n_k} e_{-1}(p_{n_k})\|_1 = \\
&\|\alpha_{n_k}[e_0(p_{n_k}) - e_0(p_*)] + \beta_{n_k}[e_{-1}(p_{n_k}) - e_{-1}(p_*)] + \alpha_{n_k} e_0(p_*) + \beta_{n_k} e_{-1}(p_*)\|_1 = \\
&\|\boldsymbol{\mathcal{O}(p_{n_k} - p_*)} + \alpha_{n_k} e_0(p_*) + \beta_{n_k} e_{-1}(p_*)\|_1 = \|\boldsymbol{o(1)} + \beta_{n_k} e_{-1}(p_*)\|_1,
\end{aligned}$$

since $\alpha_n \longrightarrow 0$ by the first part of this argument. Thus

$$V(p_{n_k}, q_{n_k}) = \|\alpha_{n_k} e_0(p_{n_k}) + \beta_{n_k} e_{-1}(p_{n_k})\|_1 = \|\boldsymbol{o(1)} + \beta_{n_k} e_{-1}(p_*)\|_1.$$

If, as assumed, $|\beta_{n_k}| \longrightarrow \ell' \neq \ell/2$, this would imply that

$$0 \leftarrow V(p_{n_k}, q_{n_k}) - \ell = \|\boldsymbol{o(1)} + \beta_{n_k} e_{-1}(p_*)\|_1 - \ell = |\beta_{n_k}|\|\boldsymbol{o(1)} + e_{-1}(p_*)\|_1 - \ell \rightarrow 2\ell' - \ell \neq 0,$$

because $\ell' \neq 0$ (otherwise $V(p_{n_k}, q_{n_k}) \longrightarrow 0$, as $k \longrightarrow \infty$, against the hypothesis). Thus we showed that $\ell' = \ell/2$, implying that $|\beta_{n_j}| \longrightarrow \ell/2$. Since $\{|\beta_{n_j}|\}$ is an arbitrary convergent subsequence of the bounded sequence $\{|\beta_n|\}$, we must have that $|\beta_n| \longrightarrow \ell/2$. □

Remark 3.40. *For any orbit, $\ell < 1$, since by (3.1), for all $n \in \mathbb{N}$,*

$$V(p_n, q_n) := \|q_n - q_{p_n}\|_1 = \frac{\|q_n - q_{n-1}\|_1}{2} \leq \frac{\|q_n\|_1 + \|q_{n-1}\|_1}{2} \leq 1,$$

and by Lemma 3.14, for all $n \in \mathbb{N}$,

$$V(p_{n+1}, q_{n+1}) = \|q_{n+1} - q_{p_{n+1}}\|_1 < \|q_n - q_{p_n}\|_1 = V(p_n, q_n).$$

3.5.3 Convergence of regular orbits

Let us start with an approach similar to that of Section 3.4, by defining a set on the boundary such that, if eventually $\{(p_n, q_n)\}$ enters it, convergence of $\{p_n\}$ follows. The additional complication will be controlling $\{q_n\}$, whose oscillations will drive the convergence of $\{p_n\}$, which in turn helps with controlling the said oscillations.

If an orbit $\{(p_n, q_n)\}$ having $\ell > 0$ is such that $p_n \longrightarrow v_i$ for some $i \in \{1, 2, 3\}$, there is nothing to prove in terms of convergence of $\{p_n\}$. Moreover, by Lemma 3.34, taking $n_k = n$, either $q_n \longrightarrow q_{v_i} - \frac{\ell}{2}e_{-1}(v_i)$ for n even and $q_n \longrightarrow q_{v_i} + \frac{\ell}{2}e_{-1}(v_i)$ for n odd or *vice versa*, that is, asymptotic 2-periodicity of $\{q_n\}$ follows. If an orbit with $\ell > 0$ is such that p_n does not converge to any vertex, by Remark 3.38 and the boundedness of p_n, there will be a subsequence $\{p_{n_j}\}_{j\in\mathbb{N}}$ bounded away from the vertices. Extracting a convergent subsubsequence $\{p_{n_{j_l}}\}_{l\in\mathbb{N}}$ by boundedness of p_n, relabel it with $\{n_k\}$, we can assume that there is a subsequence $\{p_{n_k}\}$ bounded away from the vertices, and there is $p_* \in E_i$ for some $i \in \{1, 2, 3\}$, such that $p_{n_k} \longrightarrow p_*$, as $k \longrightarrow \infty$. Because of the structure of the accumulation points of $\{(p_n, q_n)\}$ proved in Proposition 3.35, the properties of the shift of $\{(p_{n_k}, q_{n_k})\}$ shown in Remark 3.36, and Corollary 3.39, by the geometry of the simplex and Remark 3.40, it will be possible to fix such a subsequence so that $q_{n_k} \longrightarrow q_* := q_{p_*} + \frac{\ell}{2}e_{-1}(p_*)$ as $k \longrightarrow \infty$ too. As always, by symmetry, without loss of generality we will assume $i = 1$ in all arguments that will follow. Let us start with fixing $\delta > 0$ small enough (it will be necessary to further reduce it later on) so that $\delta < p_*^{(2)} < 1 - \delta$, $\delta < q_*^{(1)} < 1/2 - \delta$, $\delta < q_*^{(2)} < 1 - \delta$ and $q_*^{(3)} > \delta$. There will be an ε' small enough and K large enough such that, having defined $m := n_K$, if $p_m^{(1)}, |\alpha_m|, ||\beta_m| - \ell/2| \leq \varepsilon'$, then $\delta \leq p_m^{(2)} \leq 1 - \delta$, $\delta \leq q_m^{(1)} \leq 1/2 - \delta$, $\delta \leq q_m^{(2)} \leq 1 - \delta$ and $q_m^{(3)} \leq \delta$ (see Figure 3.5 for a graphical intuitive representation) and $|\alpha_n|, ||\beta_n| - \ell/2| \leq \varepsilon'$ for all $n \geq m$ (α_n and β_n refer to the eigencoordinates of $q_n - q_{p_n}$ as usual). Also, since $\|q - q_p\|_1 \leq 2$ (due to the diameter of the simplex in 1-norm) by Lemma 3.19 (g, h) there is a constant $B > 1$ large enough such that $|\beta| < B$, and then by (3.10) for any δ fixed small enough, there is a $\varepsilon < \delta$ small enough (to be further restricted) such that,

$$\begin{aligned}\|\hat{p} - p\|_1 &\leq 3|\alpha| + B\frac{\|e_{-1}(p)\|_1}{2}(1 + \lambda_{-1}(p)) \leq 3|\alpha| + Bp_1\|e_{-1}(p)\|_1(1 + \mathcal{O}(p_1)) \\ &\leq 3B(|\alpha| + p_1)\end{aligned}$$

for all $p_1 < \varepsilon$. Define $\varepsilon' < \varepsilon/(12B)$ (ε' will be further restricted). Having defined

$$\mathcal{K}^{\ell}_{\varepsilon',\delta} := \left\{(p, q) \in \Sigma^2 : \ 0 < p^{(1)}, \ |\alpha|, \ \left||\beta| - \frac{\ell}{2}\right| \leq \varepsilon', \ \delta \leq p^{(2)} \leq 1 - \delta\right\},$$

and similarly $\mathcal{K}^{\ell}_{\varepsilon,\delta'}$, where $\delta' := \delta/2$, we can conclude that, by construction of m, $(p_m, q_m) \in \mathcal{K}^{\ell}_{\varepsilon',\delta}$ with $\delta < q_m^{(1)} < 1/2 - \delta$ and

$$\|p_{m+1} - p_m\|_1 < 6B\varepsilon' < \frac{\varepsilon}{2}, \tag{3.29}$$

thus ensuring,

$$p_{m+1}^{(1)} \leq p_m^{(1)} + \|p_{m+1} - p_m\|_1 < \varepsilon' + \frac{\varepsilon}{2} < \varepsilon. \tag{3.30}$$

$$p_{m+1}^{(2)} \leq p_m^{(2)} + \|p_{m+1} - p_m\|_1 < 1 - \delta + \frac{\varepsilon}{2} < 1 - \delta', \tag{3.31}$$

$$p_{m+1}^{(2)} \geq p_m^{(2)} - \|p_{m+1} - p_m\|_1 > \delta - \frac{\varepsilon}{2} > \delta'. \tag{3.32}$$

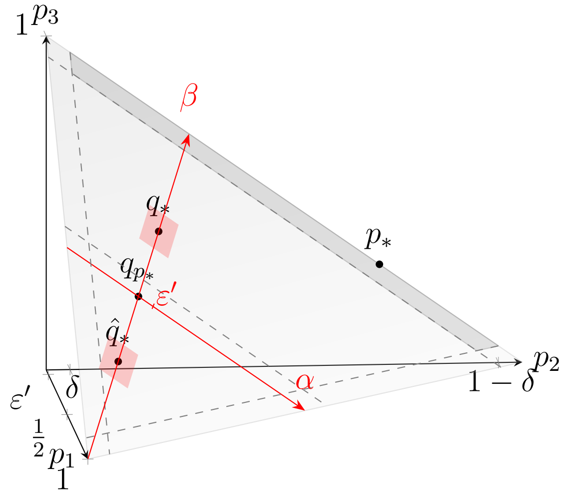


Figure 3.5: $\mathcal{K}^{\ell}_{\varepsilon',\delta}$: in gray the coordinates for p, in red the eigencordinates for q, with $q_* = q_{p_*} + \frac{\ell}{2} e_{-1}(p_*)$, $\hat{q}_* = q_{p_*} - \frac{\ell}{2} e_{-1}(p_*)$

When there is no ambiguity we will often simplify the notation as $\mathcal{K} := \mathcal{K}^{\ell}_{\varepsilon',\delta} \subset \mathcal{K}^* := \mathcal{K}^{\ell}_{\varepsilon,\delta'}$. Before expressing all the conditions on how small δ, ε' and ε need to be, we will establish a few iterative formulas.

Remark 3.41. *For all $n \geq 0$, $p^{(1)}_{n+1} \leq 2p^{(1)}_n$.*

Proof. Since for all $n \geq 0$,

$$p^{(1)}_{n+1} = p^{(1)}_n \left(\frac{q^{(2)}_n}{1 - p^{(2)}_n} + \frac{q^{(3)}_n}{p^{(1)}_n + p^{(2)}_n} \right) = p^{(1)}_n \left(\frac{q^{(2)}_n}{q^{(2)}_{n-1} + q^{(2)}_n} + \frac{q^{(3)}_n}{q^{(3)}_{n-1} + q^{(3)}_n} \right),$$

the claim follows. $\square$

Remark 3.42. *For any $n \geq 0$,*

$$p^{(1)}_{n+2} = p^{(1)}_{n+1}[(1 - q^{(1)}_{n+1})\vartheta_{n+1} - p^{(1)}_{n+1}\vartheta_{n+1} + p^{(1)}_n \vartheta'_{n+1}],$$

where

$$\begin{aligned} \vartheta_{n+1} &:= \frac{1}{(1 - p^{(1)}_n)} \left(\frac{1 - p^{(2)}_n}{1 - p^{(2)}_{n+1}} + \frac{p^{(2)}_n}{p^{(1)}_{n+1} + p^{(2)}_{n+1}} \right) \\ &= \frac{1}{1 - p^{(1)}_m} \left(2 + \frac{p^{(2)}_{m+1} - p^{(2)}_m}{1 - p^{(2)}_{m+1}} + \frac{p^{(2)}_m - p^{(2)}_{m+1}}{p^{(1)}_{m+1} + p^{(2)}_{m+1}} - \frac{p^{(1)}_{m+1}}{p^{(1)}_{m+1} + p^{(2)}_{m+1}} \right), \quad (3.33) \\ \vartheta'_{n+1} &:= -\frac{q^{(1)}_n}{(1 - p^{(2)}_{n+1})(1 - p^{(1)}_n)} + \frac{q^{(2)}_n}{(1 - p^{(2)}_{n+1})(1 - p^{(3)}_n)} + \frac{q^{(3)}_n}{(p^{(1)}_{n+1} + p^{(2)}_{n+1})(1 - p^{(2)}_n)}. \quad (3.34) \end{aligned}$$

Proof. For any $n \geq 0$,

$$\begin{aligned} p^{(1)}_{n+2} &= p^{(1)}_{n+1} \left(\frac{q^{(2)}_{n+1}}{1 - p^{(2)}_{n+1}} + \frac{q^{(3)}_{n+1}}{p^{(1)}_{n+1} + p^{(2)}_{n+1}} \right) \\ &= p^{(1)}_{n+1} \left(\frac{\frac{1 - p^{(1)}_n - p^{(2)}_n}{1 - p^{(1)}_n} q^{(1)}_n + \frac{p^{(1)}_n}{p^{(1)}_n + p^{(2)}_n} q^{(3)}_n}{1 - p^{(2)}_{n+1}} + \frac{\frac{p^{(2)}_n}{1 - p^{(1)}_n} q^{(1)}_n + \frac{p^{(1)}_n}{1 - p^{(2)}_n} q^{(2)}_n}{p^{(1)}_{n+1} + p^{(2)}_{n+1}} \right) \end{aligned}$$

and therefore, rearranging the factor in the bracket as

$$
\begin{aligned}
&\frac{1-p_n^{(2)}}{(1-p_n^{(1)})(1-p_{n+1}^{(2)})}q_n^{(1)} - \frac{p_n^{(1)}}{(1-p_n^{(1)})(1-p_{n+1}^{(2)})}q_n^{(1)} + \frac{p_n^{(1)}}{(p_n^{(1)}+p_n^{(2)})(1-p_{n+1}^{(2)})}q_n^{(3)} \\
&+ \frac{p_n^{(2)}}{(1-p_n^{(1)})(p_{n+1}^{(1)}+p_{n+1}^{(2)})}q_n^{(1)} + \frac{p_n^{(1)}}{(1-p_n^{(2)})(p_{n+1}^{(1)}+p_{n+1}^{(2)})}q_n^{(2)} = \frac{1-p_n^{(2)}}{(1-p_n^{(1)})(1-p_{n+1}^{(2)})}q_n^{(1)} \\
&+ \frac{p_n^{(2)}}{(1-p_n^{(1)})(p_{n+1}^{(1)}+p_{n+1}^{(2)})}q_n^{(1)} - \frac{p_n^{(1)}}{(1-p_n^{(1)})(1-p_{n+1}^{(2)})}q_n^{(1)} + \frac{p_n^{(1)}}{(1-p_n^{(2)})(p_{n+1}^{(1)}+p_{n+1}^{(2)})}q_n^{(2)} \\
&+ \frac{p_n^{(1)}}{(p_n^{(1)}+p_n^{(2)})(1-p_{n+1}^{(2)})}q_n^{(3)} = q_n^{(1)}\vartheta_{n+1} + p_n^{(1)}\vartheta'_{n+1}
\end{aligned}
$$

and, by rearranging the first component of (2.38), using $q_n^{(1)} = 1 - q_{n+1}^{(1)} - p_{n+1}^{(1)}$, the claim follows. □

Remark 3.43. *For any* $n \geq 0$*,*

$$
p_{n+2}^{(2)} - p_{n+1}^{(2)} = -q_n^{(1)}\left(p_{n+1}^{(2)} - p_n^{(2)} + \eta_{n+1} - \eta'_{n+1} - \eta_n + \eta''_{n+1} + \eta'''_{n+1}\right),
$$

where

$$
\eta_n := \frac{p_n^{(2)}p_n^{(1)}}{1-p_n^{(1)}} \tag{3.35}
$$

$$
\eta'_{n+1} := \frac{q_n^{(2)}}{q_n^{(1)}}\frac{p_{n+1}^{(2)}}{p_{n+1}^{(1)}+p_{n+1}^{(2)}}\frac{p_n^{(1)}}{1-p_n^{(2)}} \tag{3.36}
$$

$$
\eta''_{n+1} := p_n^{(2)}\frac{p_{n+1}^{(1)}}{p_{n+1}^{(1)}+p_{n+1}^{(2)}} \tag{3.37}
$$

$$
\eta'''_{n+1} = p_n^{(2)}\frac{p_{n+1}^{(1)}}{p_{n+1}^{(1)}+p_{n+1}^{(2)}}\frac{p_n^{(1)}}{1-p_n^{(1)}}. \tag{3.38}
$$

Proof.

$$
\begin{aligned}
p_{n+2}^{(2)} - p_{n+1}^{(2)} &= \frac{p_{n+1}^{(2)}}{1-p_{n+1}^{(1)}}q_{n+1}^{(1)} + \frac{p_{n+1}^{(2)}}{p_{n+1}^{(1)}+p_{n+1}^{(2)}}q_{n+1}^{(3)} - p_{n+1}^{(2)} = p_{n+1}^{(2)}\left(\frac{q_{n+1}^{(1)}}{1-p_{n+1}^{(1)}} - 1\right) \\
&+ \frac{p_{n+1}^{(2)}}{p_{n+1}^{(1)}+p_{n+1}^{(2)}}\left(q_n^{(1)}\frac{p_n^{(2)}}{1-p_n^{(1)}} + q_n^{(2)}\frac{p_n^{(1)}}{1-p_n^{(2)}}\right) = p_{n+1}^{(2)}\Bigg[-\frac{q_n^{(1)}}{1-p_{n+1}^{(1)}} \\
&+ \frac{p_n^{(1)}}{(p_{n+1}^{(1)}+p_{n+1}^{(2)})(1-p_n^{(2)})}q_n^{(2)}\Bigg] + p_n^{(2)}\frac{q_n^{(1)}}{1-p_n^{(1)}}\frac{p_{n+1}^{(2)}}{p_{n+1}^{(1)}+p_{n+1}^{(2)}} = -q_n^{(1)} \\
&\left[\frac{p_{n+1}^{(2)}}{1-p_{n+1}^{(1)}} - \frac{q_n^{(2)}}{q_n^{(1)}}\frac{p_{n+1}^{(2)}}{p_{n+1}^{(1)}+p_{n+1}^{(2)}}\frac{p_n^{(1)}}{1-p_n^{(2)}} - p_n^{(2)}\frac{p_{n+1}^{(2)}}{(p_{n+1}^{(1)}+p_{n+1}^{(2)})(1-p_n^{(1)})}\right] \\
&= -q_n^{(1)}\Bigg[p_{n+1}^{(2)} + \frac{p_{n+1}^{(2)}p_{n+1}^{(1)}}{1-p_{n+1}^{(1)}} - \frac{q_n^{(2)}}{q_n^{(1)}}\frac{p_{n+1}^{(2)}}{p_{n+1}^{(1)}+p_{n+1}^{(2)}}\frac{p_n^{(1)}}{1-p_n^{(2)}} \\
&- p_n^{(2)}\left(1 - \frac{p_{n+1}^{(1)}}{p_{n+1}^{(1)}+p_{n+1}^{(2)}}\right)\left(1 + \frac{p_n^{(1)}}{1-p_n^{(1)}}\right)\Bigg].
\end{aligned}
$$

Thus the claim follows. □

The fixed small enough parameter δ will, in addition, be required to satisfy $\delta < 1/45$. Define

$$
\gamma = \gamma(\delta') := \sqrt{1 - 4(\delta')^2 + 144(\delta')^3}.
$$

Note that the constant γ is positive subunitary, since $\gamma' := \gamma^2 = 1-4(\delta')^2+144(\delta')^3$ is too, because the polynomial $-4x^2+144x^3$ is negative monotone decreasing on $(0, 1/90)$ with minimum at $1/54 > 1/90$ of value $-1/2187$. Require

$$\varepsilon < \min\left\{(\delta')^5, \frac{\delta'(1-\gamma^2)}{4}\right\}.$$

Define

$$D := 4 + \frac{1}{\delta'}\left(2 + \frac{1}{\delta'}\right)$$

and let Γ be a constant such that

$$0 < \Gamma < \frac{\frac{c}{1-\delta'} - 1}{D},$$

where $0 < 1-\delta' < c < 1$ is another constant such that $c(1-\delta')^{-1} > 1$, which is consistent since $0 < 1-\delta' < 1$ (no confusion can arise with the constant c introduced in Section 3.4). Define $\lambda := \max\{\gamma, c\}$ and further restrict

$$\varepsilon' < \min\left\{\frac{\varepsilon}{12B}, \frac{\varepsilon}{8}\frac{1-\gamma^2}{D}, \frac{\delta\Gamma\lambda(1-\lambda)}{2(2+\Gamma)}\right\}.$$

This has ultimately determined the size of $\mathcal{K}$ (the smaller set we need, to kick start the arguments in the lemmas that will follow) while $\mathcal{K}^*$ (the larger set, on which all constants defined so far exist and apply uniformly, as the orbit travels through it) had been already previously fixed to determine the constants necessary to define $\mathcal{K}$.

Lemma 3.44. *Let $\gamma' := \gamma^2$, assume that $p_{m+l} \leq 2(\gamma')^{\lfloor \frac{l}{2} \rfloor} p_m^{(1)}$ for all $0 \leq l \leq 2k-b$, where $b \in \{0,1\}$. For all $b \leq j \leq 2k$,*

$$\delta' < q_{m+2k-j}^{(1)} < \frac{1}{2} - \delta'$$

if j is even and

$$\frac{1}{2} + \delta' < q_{m+2k-j}^{(1)} < 1 - \delta'$$

if j is odd.

Proof. Iterate the first component of (2.38) after rearranging it as

$$q_{m+l}^{(1)} = 1 - p_{m+l}^{(1)} - q_{m+l-1}^{(1)}.$$

It yields

$$q_{m+l}^{(1)} = \begin{cases} q_m^{(1)} + \sum_{j=1}^{l} (-1)^{j+1} p_{m+j}^{(1)} & l \text{ even} \\ 1 - q_m^{(1)} + \sum_{j=1}^{l} (-1)^{j} p_{m+j}^{(1)} & l \text{ odd.} \end{cases} \tag{3.39}$$

Recall that by construction

$$\varepsilon < \frac{\delta'(1-\gamma')}{4}.$$

Since $q_m^{(1)} < 1/2 - \delta$,

$$\begin{aligned} q_{m+2k-j}^{(1)} = q_m^{(1)} + \sum_{l=1}^{2k-j} (-1)^{l+1} p_{m+l}^{(1)} < \frac{1}{2} - \delta + \sum_{l=1}^{2k-j} p_{m+l}^{(1)} < \frac{1}{2} - \delta + 4p_m^{(1)} \sum_{l=0}^{\lfloor \frac{2k-j}{2} \rfloor} (\gamma')^l \\ < \frac{1}{2} - \delta + 4\frac{\varepsilon}{1-\gamma'} < \frac{1}{2} - \delta' \end{aligned} \tag{3.40}$$

for all even $b \le j \le 2k$ (with the bound for $j = 2k$ holding also with δ, by adopting empty sum convention) and

$$q^{(1)}_{m+2k-j} = 1 - q^{(1)}_m + \sum_{l=1}^{2k-j}(-1)^l p^{(1)}_{m+l} > \frac{1}{2} + \delta - \sum_{l=1}^{2k-j} p^{(1)}_{m+l} > \frac{1}{2} + \delta - 4\varepsilon' \sum_{l=0}^{\lfloor \frac{2k-j}{2} \rfloor} (\gamma')^l$$
$$> \frac{1}{2} + \delta - 4\frac{\varepsilon}{1-\gamma'} > \frac{1}{2} + \delta' \tag{3.41}$$

for all odd $b \le j \le 2k$. Similarly, since $\delta < q^{(1)}_m < 1 - \delta$,

$$q^{(1)}_{m+2k-j} = q^{(1)}_m + \sum_{l=1}^{2k-j}(-1)^{l+1} p^{(1)}_{m+l} < \frac{1}{2} - \delta + \sum_{l=1}^{2k-j} p^{(1)}_{m+l} > \delta - 4p^{(1)}_m \sum_{l=0}^{\lfloor \frac{2k-j}{2} \rfloor} (\gamma')^l$$
$$> \delta - 4\frac{\varepsilon}{1-\gamma'} > \delta' \tag{3.42}$$

for all even $b \le j \le 2k$ and

$$q^{(1)}_{m+2k-j} = 1 - q^{(1)}_m + \sum_{l=1}^{2k-j}(-1)^l p^{(1)}_{m+l} < 1 - \delta + \sum_{l=1}^{2k-j} p^{(1)}_{m+l} < 1 - \delta + 4\varepsilon' \sum_{l=0}^{\lfloor \frac{2k-j}{2} \rfloor} (\gamma')^l$$
$$< 1 - \delta + 4\frac{\varepsilon}{1-\gamma'} < 1 - \delta' \tag{3.43}$$

for all odd $b \le j \le 2k$. □

Lemma 3.45. *Let $\gamma' := \gamma^2$, assume that $p_{m+l} \le 2(\gamma')^{\lfloor \frac{l}{2} \rfloor} p^{(1)}_m$ for all $0 \le l \le 2k - b$, where $b \in \{0, 1\}$, and $\delta' < p^{(2)}_{m+l} < 1 - \delta'$. For all $b \le j \le 2k - 1$,*

$$|p^{(2)}_{m+2k-j} - p_{m+2k-j-1}| < \varepsilon$$

Proof. Recall that
$$\varepsilon' < \frac{\varepsilon}{8D}(1-\gamma').$$
Iterate Remark 3.43, setting $n = m + 2(k-1) - j$, down to time m, it yields

$$|p^{(2)}_{m+2k-j} - p^{(2)}_{m+2k-j-1}| \le |p^{(2)}_{m+1} - p^{(2)}_m| + \sum_{l=1}^{2k-j-1} E_{m+l} < \frac{\varepsilon}{2} + \sum_{l=1}^{2k-2} E_{m+l},$$

where $E_{m+l} := \eta_{m+l} + \eta'_{m+l} + \eta_{m+l-1} + \eta''_{m+l} + \eta'''_{m+l}$. Note that the hypotheses allow to apply Lemma 3.44, thus $\delta' < q^{(1)}_{m+l-1} < 1 - \delta'$ for all $1 \le l \le 2k - j - 1$. This implies that, by using Remark 3.41 and the assumptions, for all $1 \le l \le 2k - j - 1$,

$$\eta_{m+l-1} < p^{(1)}_{m+l-1} < 2(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m \tag{3.44}$$
$$\eta_{m+l} < 2p^{(1)}_{m+l-1} < 4(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m \tag{3.45}$$
$$\eta'_{m+l} < \frac{1}{(\delta')^2} p^{(1)}_{m+l-1} < 2\frac{1}{(\delta')^2}(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m \tag{3.46}$$
$$\eta''_{m+l} < \frac{2}{\delta'} p^{(1)}_{m+l-1} < \frac{4}{\delta'}(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m \tag{3.47}$$
$$\eta'''_{m+l} < p^{(1)}_{m+l-1} < 2(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m. \tag{3.48}$$

Hence
$$E_{m+l} < 2D(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m, \tag{3.49}$$

which yields a bound uniform in k on the increments of the $p^{(2)}$-component:

$$|p^{(2)}_{m+2k-1} - p^{(2)}_{m+2k-2}| < \frac{\varepsilon}{2} + 2D\varepsilon' \sum_{l=1}^{\infty} (\gamma')^{\lfloor \frac{l-1}{2} \rfloor} = \frac{\varepsilon}{2} + 4D\frac{\varepsilon'}{1-\gamma'} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon. \quad (3.50)$$

□

Define the hitting time

$$\sigma := \inf\left\{ n \geq m : \ p^{(2)}_n \notin \left[\frac{\delta}{2}, 1 - \frac{\delta}{2}\right]\right\} \in \mathbb{N} \cup \infty.$$

Note that $\sigma > m+1$ by construction of m.

Lemma 3.46. *For all $m \leq n \leq \sigma$, $p^{(1)}_n \leq 2\gamma^{n-m-1} p^{(1)}_m$.*

Proof. Let $\gamma^2 = \gamma'(\delta') := 1 - 4(\delta')^2 + 341(\delta')^3$, we will first show that for every $k \geq 0$ such that $m + 2k < \sigma$,

$$p^{(1)}_{m+2k} \leq 2(\gamma')^k p^{(1)}_m \quad (3.51)$$

$$p^{(1)}_{m+2k+1} \leq 2(\gamma')^k p^{(1)}_m \quad (3.52)$$

and if $\sigma = m + 2\bar{k}$ for some $\bar{k} > 0$, we leave (3.52) for $k = \bar{k}$ out. Recall that by construction $\delta' < {}^{1}/_{90}$,

$$\varepsilon < \delta' \frac{1-\gamma'}{4}$$

and

$$\varepsilon' < \frac{\varepsilon}{8}\frac{1-\gamma'}{D}.$$

Using Remark 3.41 with $n = m$ yields the first odd case for $k = 0$ (the even one being trivial):

$$p^{(1)}_{m+1} \leq 2p^{(1)}_m.$$

Let $n = m$ in (3.33) and (3.34) and apply the hypothesis made in (3.30) to (3.32), and $|p^{(2)}_{m+1} - p^{(2)}_m| < {}^{\varepsilon}/_{2}$, which follows from (3.29). Then the following estimates follow:

$$|\vartheta_{m+1}| = \frac{1}{1-p^{(1)}_m}\left(\frac{1-p^{(2)}_m}{1-p^{(2)}_{m+1}} + \frac{p^{(2)}_m}{p^{(1)}_{m+1} + p^{(2)}_{m+1}}\right) \leq \frac{2}{\delta'(1-\varepsilon)} \quad (3.53)$$

$$|\vartheta_{m+1}| \leq 2 + 2\frac{p^{(1)}_m}{1-p^{(1)}_m} + \frac{|p^{(2)}_{m+1} - p^{(2)}_m|}{(1-p^{(1)}_m)(1-p^{(2)}_{m+1})} + \frac{|p^{(2)}_m - p^{(2)}_{m+1}|}{(1-p^{(1)}_m)(p^{(1)}_{m+1} + p^{(2)}_{m+1})}$$
$$+ \frac{p^{(1)}_{m+1}}{(1-p^{(1)}_m)(p^{(1)}_{m+1} + p^{(2)}_{m+1})} \leq 2 + 2\frac{p^{(1)}_m}{1-p^{(1)}_m} + 3\frac{\varepsilon}{\delta'(1-\varepsilon)} \leq 2 + \frac{\varepsilon}{1-\varepsilon}\left(2 + \frac{3}{\delta'}\right) \quad (3.54)$$

$$|\vartheta'_{m+1}| \leq \frac{1}{(1-p^{(2)}_{m+1})(1-p^{(1)}_m)} + \frac{1}{(1-p^{(2)}_{m+1})(p^{(1)}_m + p^{(2)}_m)} + \frac{1}{(p^{(1)}_{m+1} + p^{(2)}_{m+1})(1-p^{(2)}_m)}$$
$$\leq \frac{1}{\delta'}\left(\frac{1}{1-\varepsilon} + \frac{2}{\delta'+\varepsilon}\right). \quad (3.55)$$

Plug the estimates (3.53) to (3.55) into Remark 3.42 applied to $n = m$, it yields

$$p^{(1)}_{m+2} \leq p^{(1)}_{m+1}\left[2(1-q^{(1)}_{m+1}) + \frac{\varepsilon}{1-\varepsilon}\left(2 + \frac{3}{\delta'}\right) + p^{(1)}_{m+1}\frac{2}{\delta'(1-\varepsilon)} + p^{(1)}_m \frac{1}{\delta'}\right.$$
$$\left.\left(\frac{1}{1-\varepsilon} + \frac{2}{\delta'+\varepsilon}\right)\right] \leq p^{(1)}_{m+1}\left[2(1-q^{(1)}_{m+1}) + \varepsilon\left(\frac{2}{1-\varepsilon} + \frac{2}{\delta'+\varepsilon} + \frac{6}{\delta'(1-\varepsilon)}\right)\right].$$

Since by construction $\varepsilon < (\delta')^5 < {}^1\!/_{90^5}$ and recalling that since $q^{(1)}_m < {}^1\!/_2 - \delta$, we have

$$q^{(1)}_{m+1} = 1 - p^{(1)}_{m+1} - q^{(1)}_m > 1 - \varepsilon - \frac{1}{2} + \delta > \frac{1}{2} + \delta',$$

which yields

$$\begin{aligned} p^{(1)}_{m+2} &\le 2p^{(1)}_m \left[2(1-q^{(1)}_{m+1}) + \varepsilon\left(\frac{4}{\delta'} + \frac{6}{(\delta')^2}\right)\right] < 2p^{(1)}_m \left[1 - 2\delta' + 10\frac{\varepsilon}{(\delta')^2}\right] \\ &< 2p^{(1)}_m \left[1 - 2\delta' + 10(\delta')^3\right]. \end{aligned}$$

Note that $\gamma' = 1 - 4(\delta')^2 + 144(\delta')^3 > 1 - 2\delta' + 10(\delta')^3$ for all positive δ', because the difference $1 - 4x^2 + 144x^3 - (1 - 2x + 10x^3) = 2(x - 2x^2 + 67x^3)$ is nonnegative increasing on the positive reals, having derivative $2(1 - 4x + 201x^2) > 0$, due to negative discriminant of the parabola. Thus we have shown that $p^{(1)}_{m+2} \le 2\gamma' p^{(1)}_m$. Note that the constant γ' holds uniformly for $p_m \in \mathcal{K}$, $p_{m+1} \in \mathcal{K}^*$ by the steps above. This has been ensured by $p_m \in \mathcal{K}$ and the initial tuning of the constants involved in the definition of the set $\mathcal{K}$. The case $k = 1$ is not concluded yet though, as if $\sigma > m + 2$, we will have to do one more step. This will allow us to see clearly how this argument iterates and the pattern that arises, especially how the constant γ' arises. By the geometric decay proved so far,

$$p^{(1)}_{m+2} < 2\gamma' p^{(1)}_m < 2\varepsilon' < \varepsilon, \tag{3.56}$$

and by the definition of σ, $\delta' < p^{(2)}_{m+2} < 1 - \delta'$; thus the same estimates in (3.53) and (3.55) apply to ϑ_{m+2} and ϑ'_{m+2}, with the due shift of time indices. However, (3.54) does not apply automatically, since nothing guarantees that the same bound applies on the shifted increments of the $p^{(2)}$-component. Let us first assume that indeed, also $|p^{(2)}_{m+2} - p^{(2)}_{m+1}| < \varepsilon$ holds and therefore, that also (3.54) applies with the due shift of indices. Then plugging them into Remark 3.42, applied to $n = m + 1$, yields

$$\begin{aligned} p^{(1)}_{m+3} &\le p^{(1)}_{m+2}\left[2(1 - q^{(1)}_{m+2}) + 10\frac{\varepsilon}{(\delta')^2}\right] = p^{(1)}_{m+2}\left[2(p^{(1)}_{m+2} + q^{(1)}_{m+1}) + 10\frac{\varepsilon}{(\delta')^2}\right] \\ &\le p^{(1)}_{m+1}\left[2(1 - q^{(1)}_{m+1}) + 10\frac{\varepsilon}{(\delta')^2}\right]\left[2q^{(1)}_{m+1} + \varepsilon\left(2 + \frac{10}{(\delta')^2}\right)\right] \\ &= p^{(1)}_{m+1}\left[4q^{(1)}_{m+1}(1 - q^{(1)}_{m+1}) + \varepsilon\left(4(1 - q^{(1)}_{m+1}) + \frac{20}{(\delta')^2}\right) + \varepsilon^2\frac{10}{(\delta')^2}\left(2 + \frac{10}{(\delta')^2}\right)\right] \\ &< 2p^{(1)}_m\left[1 - 4\delta^2 + \varepsilon\frac{24}{(\delta')^2} + \varepsilon^2\frac{120}{(\delta')^4}\right] < 2p^{(1)}_m\left[1 - 4(\delta')^2 + 24(\delta')^3 + 120(\delta')^6\right] \\ &< 2p^{(1)}_m\left[1 - 4(\delta')^2 + 144(\delta')^3\right], \end{aligned}$$

hence $p^{(1)}_{m+3} \le 2\gamma' p^{(1)}_m$, which completes the claim for $k = 1$ (and $\sigma > m + 2$). We now show that the upper bound on the $p^{(2)}$-component indeed keeps applying uniformly, by using Remark 3.43 for $n = m$, which yields the upper bound

$$|p^{(2)}_{m+2} - p^{(2)}_{m+1}| < |p^{(2)}_{m+1} - p^{(2)}_m| + E_{m+1}, \tag{3.57}$$

where $E_{m+1} := \eta_{m+1} + \eta'_{m+1} + \eta_m + \eta''_{m+1} + \eta'''_{m+1}$. Since $p^{(1)}_{m+2} < 2\gamma' p^{(1)}_m < 2\varepsilon' < \varepsilon$, by the definition of σ, which ensures $\delta' < p^{(2)}_{m+2} < 1 - \delta'$, and by exploiting Remark 3.41

with $n = m$, we can estimate

$$\eta_m = \frac{p_m^{(2)} p_m^{(1)}}{p_m^{(2)} + p_m^{(3)}} < p_m^{(1)} \tag{3.58}$$

$$\eta_{m+1} = \frac{p_{m+1}^{(2)} p_{m+1}^{(1)}}{p_{m+1}^{(2)} + p_{m+1}^{(3)}} < 2p_m^{(1)} \tag{3.59}$$

$$\eta'_{m+1} := \frac{q_m^{(2)}}{q_m^{(1)}} \frac{p_{m+1}^{(2)}}{p_{m+1}^{(1)} + p_{m+1}^{(2)}} \frac{p_m^{(1)}}{1 - p_m^{(2)}} < \frac{1}{(\delta')^2} p_m^{(1)} \tag{3.60}$$

$$\eta''_{m+1} := p_m^{(2)} \frac{p_{m+1}^{(1)}}{p_{m+1}^{(1)} + p_{m+1}^{(2)}} < \frac{2}{\delta'} p_m^{(1)} \tag{3.61}$$

$$\eta'''_{m+1} = p_m^{(2)} \frac{p_{m+1}^{(1)}}{p_{m+1}^{(1)} + p_{m+1}^{(2)}} \frac{p_m^{(1)}}{p_m^{(2)} + p_m^{(3)}} < p_m^{(1)}. \tag{3.62}$$

These estimates yield

$$E_{m+1} \le D p_m^{(1)}.$$

Since by construction

$$\varepsilon' < \frac{\varepsilon}{8} \frac{1-\gamma'}{D},$$

$$E_{m+1} < \frac{\varepsilon}{8}(1-\gamma') < \frac{\varepsilon}{8},$$

and as a result of $|p_{m+1}^{(2)} - p_m^{(2)}| < \varepsilon/2$, which yields

$$|p_{m+2}^{(2)} - p_{m+1}^{(2)}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{8} < \varepsilon.$$

Apart from this first instances in the base cases, this estimate will be less immediate in the later steps and we will rely on Lemma 3.45.

To summarise what has been proved in this two-steps argument, there is a constant γ' holding uniformly on $\mathcal{K}^*$ for both cases, $\sigma = m+2$ and $\sigma > m+2$. In the first case $p_{m+1}^{(1)} < 2(\gamma')^0 p_m^{(1)}$ (case $k = 0$) and $p_{m+2}^{(1)} < 2\gamma' p_m^{(1)}$ (half case $k = 1$); in the second case both $p_{m+1}^{(1)} < 2(\gamma')^0 p_m^{(1)}$ (case $k = 0$), and $p_{m+2}^{(1)} < 2\gamma' p_m^{(1)}$ and $p_{m+3}^{(1)} < 2\gamma' p_m^{(1)}$ (full case $k = 1$). Before proceeding with iterating this two-steps geometric decay, note that the estimate on $q_n^{(1)}$'s oscillations above and below $1/2$ has to iterate at each step. For example, for the next step it will hold because

$$q_{m+2}^{(1)} = 1 - p_{m+2}^{(1)} - q_{m+1}^{(1)} < 1 - \left(\frac{1}{2} + \delta'\right) = \frac{1}{2} - \delta'.$$

Apart from these first few steps, this condition will not be so immediate to verify, because geometric terms will start adding up, and we will rely on Lemma 3.44. Assume that $m + 3 < n < \sigma$ for some n and let us prove the claim for $n + 1$. There are two cases to consider: the even step $n = m + 2k - 1$ to $n + 1 = m + 2k$ first and the odd step $n = m + 2k$ to $n + 1 = m + 2k + 1$ afterwards, for all $k \in \mathbb{N}$ such that n is in the mentioned range.

- In the even step one has the induction hypothesis that for all $1 \le j \le 2k$,

$$p_{m+2k-j}^{(1)} < 2(\gamma')^{\lfloor \frac{2k-j}{2} \rfloor} p_m^{(1)} \tag{3.63}$$

  and it needs to be shown (3.51), that is $p_{m+2k}^{(1)} < 2(\gamma')^k p_m^{(1)}$. As to the oscillations of $q_n^{(1)}$, they are δ'-bounded away from $1/2$ in the correct order, thanks to the induction hypothesis and Lemma 3.44 applied with $b = 1$:

$$q_{m+2k-j}^{(1)} < \frac{1}{2} - \delta'$$

for all even $1 \le j \le 2k$ (with bounds for $j = 2k$ holding also with δ) and

$$q^{(1)}_{m+2k-j} > \frac{1}{2} + \delta'$$

for all odd $1 \le j \le 2k$. All that remains to be shown, is that

$$p^{(1)}_{m+2k} < \gamma' p^{(1)}_{m+2k-2},$$

by using

$$q^{(1)}_{m+2k-1} > \frac{1}{2} + \delta'.$$

Since by the geometric decay ensured by (3.63),

$$p^{(1)}_{m+2k-1} < 2(\gamma')^{k-1} p^{(1)}_m$$

which ensures (3.56) up to $p^{(1)}_{m+2k-1}$, and since $\sigma > m + 2k - 1$ implies $\delta' < p^{(2)}_{m+2k-1} < 1 - \delta'$; the estimates in (3.53) to (3.55) apply also to ϑ_{m+2k-1} and ϑ'_{m+2k-1} (with the due shift of time indices) since by Lemma 3.45 with $b = 1$ it holds that $|p^{(2)}_{m+2k-1} - p^{(2)}_{m+2k-1}| < \varepsilon$. We will assume this for now and show it after the conclusion of this case. By plugging the aforementioned estimates into Remark 3.42 applied to $n = m + 2k - 2$ yields the same estimate as that obtained for $p^{(1)}_{m+2}$:

$$\begin{aligned} p^{(1)}_{m+2k} &\le p^{(1)}_{m+2k-1} \left[2(1 - q^{(1)}_{m+2k-1}) + \varepsilon \left(\frac{4}{\delta'} + \frac{6}{(\delta')^2} \right) \right] < p^{(1)}_{m+2k-1}[1 - 2\delta' + 10(\delta')^3] \\ &< \gamma' p^{(1)}_{m+2k-1} \end{aligned}$$

resulting into (3.51) by (3.63).

- In the odd step one has (3.63) holding for all $0 \le j \le 2k$, and it needs to be proved (3.52), that is $p^{(1)}_{m+2k+1} < 2(\gamma')^k p^{(1)}_m$. For the oscillations of $q^{(1)}$ we proceed similarly but, again, with a different range for j, by exploiting Lemma 3.44 with $b = 0$:

$$q^{(1)}_{m+2k-j} < \frac{1}{2} - \delta'$$

for all even $0 \le j \le 2k$ (with bounds for $j = 2k$ holding also with δ) and

$$q^{(1)}_{m+2k-j} > \frac{1}{2} + \delta'$$

for all odd $0 \le j \le 2k$. All that remains to show is that

$$p^{(1)}_{m+2k+1} < \gamma' p^{(1)}_{m+2k-1}$$

by using

$$q^{(1)}_{m+2k-1} > \frac{1}{2} + \delta'.$$

Since by the geometric decay ensured by (3.63) in the new range of indices,

$$p^{(1)}_{m+2k-1} < 2(\gamma')^{k-1} p^{(1)}_m,$$

ensuring (3.56) up to $p^{(1)}_{m+2k-1}$, and $\sigma > m + 2k$ implies $\delta' < p^{(2)}_{m+2k} < 1 - \delta'$, the estimates in (3.53) to (3.55) apply also to ϑ_{m+2k} and ϑ'_{m+2k} (with the due shift of time indices) since by Lemma 3.45 with $b = 0$, we have $|p^{(2)}_{m+2k} - p^{(2)}_{m+2k-1}| < \varepsilon$;, also the previous step's estimates for ϑ_{m+2k-1} and ϑ'_{m+2k-1} keep applying, and they are vital, since in this step the bound needed is yielded by iterating the previous step into the current one, producing a two-step estimate, necessary because a one step estimate would not yield a subunitary constant, due to $q^{(1)}_{m+2k} < 1/2 - \delta'$, and thus $2(1-q^{(1)}_{m+2k}) > 1$. Therefore, by plugging the extended estimates into Remark 3.42 applied to $n = m+2k-1$, and also using the old estimates from the previous even step, yields the same estimate as that obtained for $p^{(1)}_{m+3}$:

$$p^{(1)}_{m+2k+1} \leq p^{(1)}_{m+2k-1}\left[2(1-q^{(1)}_{m+2k-1}) + 10\frac{\varepsilon}{(\delta')^2}\right]\left[2q_{m+2k-1}\varepsilon\left(2+\frac{10}{(\delta')^2}\right)\right]$$
$$< p^{(1)}_{m+2k-1}[1-4(\delta')^2 + 144(\delta')^3] < \gamma' p^{(1)}_{m+2k-1},$$

resulting into (3.52) by (3.63) in the new range of indices.

Having proved the two-steps claim, we can easily derive the main claim by simply setting $\gamma := \sqrt{\gamma'}$, so as to express the two-steps geometric decaying upper bound as a one-step geometric decaying one. It has been shown that for all integers $1 \leq l \leq \sigma - m$,

$$p^{(1)}_{m+l} < 2(\gamma')^{\lfloor \frac{l}{2} \rfloor} p^{(1)}_m.$$

Since

$$\left\lfloor \frac{l}{2} \right\rfloor \geq \frac{l-1}{2},$$

it follows that

$$p^{(1)}_{m+l} < 2\sqrt{\gamma'}^{\,l-1} p^{(1)}_m,$$

hence for the uniform constant γ, we have that for all $m < n \leq \sigma$,

$$p^{(1)}_n < 2\gamma^{n-m-1} p^{(1)}_m.$$

□

For any $\tau \geq m$, define

$$\zeta := \inf\left\{n > \tau : \frac{|p^{(2)}_{n+1} - p^{(2)}_n|}{p^{(1)}_n} < \frac{1}{\Gamma}\right\}.$$

Lemma 3.47. *Suppose there exists $m \leq \tau < \sigma$ such that*

$$\frac{p^{(1)}_\tau}{|p^{(2)}_{\tau+1} - p^{(2)}_\tau|} \leq \Gamma.$$

Then for all $\tau \leq n \leq \zeta \wedge \sigma$,

$$|p^{(2)}_{n+1} - p^{(2)}_n| < c^{n-m}|p^{(2)}_{m+1} - p^{(2)}_m|.$$

Proof. We show the claim for $\tau = m$, as it extends trivially. Clearly the claim is trivially true for $n = m$. If $\zeta = m+1$, $\zeta \wedge \sigma = m+1$, and we need to show it for $n = m+1$. It is known that $p^{(1)}_m \leq \Gamma|p^{(2)}_{m+1} - p^{(2)}_m|$ by hypothesis and $p^{(1)}_{m+1} < 2p^{(1)}_m$ by

Remark 3.41 with $n = m$, along with the hypotheses (3.30) to (3.32) by construction of m. Then (3.58) to (3.62) apply, and it follows that

$$\begin{aligned}
\eta_m &< p^{(1)}_m < \Gamma|p^{(2)}_{m+1} - p^{(2)}_m| \\
\eta_{m+1} &< 2p^{(1)}_m < 2\Gamma|p^{(2)}_{m+1} - p^{(2)}_m| \\
\eta'_{m+1} &< \frac{1}{(\delta')^2} p^{(1)}_m < \frac{\Gamma}{(\delta')^2}|p^{(2)}_{m+1} - p^{(2)}_m| \\
\eta''_{m+1} &< \frac{2}{\delta'} p^{(1)}_m < \frac{2}{\delta'}\Gamma|p^{(2)}_{m+1} - p^{(2)}_m| \\
\eta'''_{m+1} &< p^{(1)}_m < \Gamma|p^{(2)}_{m+1} - p^{(2)}_m|.
\end{aligned}$$

Plugging these estimate into Remark 3.43 with $n = m$ yields

$$|p^{(2)}_{m+2} - p^{(2)}_{m+1}| \leq q^{(1)}_m \left[1 + \Gamma D\right] |p^{(2)}_{m+1} - p^{(2)}_m|,$$

since by construction

$$1 + \Gamma D < \frac{c}{1 - \delta'},$$

$$q^{(1)}_m \left[1 + \Gamma D\right] \leq \frac{q^{(1)}_m c}{1 - \delta'} \leq c$$

and it follows that

$$|p^{(2)}_{m+2} - p^{(2)}_{m+1}| \leq c|p^{(2)}_{m+1} - p^{(2)}_m|,$$

which completes the claim, and can be also regarded as the first step of the induction argument for $\zeta \wedge \sigma > m + 1$.

Assuming now $\zeta > m + 1$, $\zeta \wedge \sigma > m + 1$, thus the claim will be shown for all $m + 1 < n \leq \zeta \wedge \sigma$, by induction. Assuming, for any $m \leq n < \zeta \wedge \sigma$, $|p^{(2)}_{n+1} - p^{(2)}_n| < c^{n-m}|p^{(2)}_{m+1} - p^{(2)}_m|$, we need to show that $|p^{(2)}_{n+2} - p^{(2)}_{n+1}| < c^{n-m+1}|p^{(2)}_{m+1} - p^{(2)}_m|$. Here it will be crucial to remember that in Lemma 3.46 we ensured the boundedness of $q^{(1)}$ away from the boundary of the simplex, in order to obtain the validity of the constants involved, in the coming steps. This means that again, in parallel with the iteration of $p^{(2)}$'s increments' geometrically decaying upper bound, one will have to control $q^{(1)}$, in order to iterate the estimate on η'_n. Recall that $\delta' < \delta < q^{(1)}_m < 1 - \delta < 1 - \delta'$. Since for all $m < n \leq \zeta \wedge \sigma$, by Lemma 3.46 it holds that $p^{(1)}_n < 2\gamma^{n-m-1}\varepsilon$, or equivalently $p^{(1)}_{m+k} < 2\gamma^{k-1}\varepsilon$ for all $k \in \mathbb{N}$ such that $n = m + k$ is within the bounds above; by (3.39), for all suitable k

$$q^{(1)}_{m+k} \geq \begin{cases} q^{(1)}_m - \sum_{j=1}^k p^{(1)}_{m+j} \geq \delta - 2\varepsilon \sum_{j=1}^k \gamma^{j-1} > \delta - 2\frac{\varepsilon}{1-\gamma} > \delta' & k \text{ even}, \\ 1 - q^{(1)}_m - \sum_{j=1}^k p^{(1)}_{m+j} \geq \delta - 2\varepsilon \sum_{j=1}^k \gamma^{j-1} \geq \delta - 2\frac{\varepsilon}{1-\gamma} > \delta' & k \text{ odd}, \end{cases}$$

since $\varepsilon < \delta'(1-\gamma)/2$, as by construction $\varepsilon < \delta'(1-\gamma^2)/4$ and $(1-\gamma^2)/4 \leq (1-\gamma)/2$, being this inequality equivalent to $0 \leq (\gamma - 1)^2$. This ensures that upper bounding the reciprocals of $q^{(1)}$ appearing in η'_n with δ' can carry on during the induction step. As to lower bounding it with $1 - \delta'$, one can proceed analogously:

$$q^{(1)}_{m+k} \leq \begin{cases} q^{(1)}_m + \sum_{j=1}^k p^{(1)}_{m+j} \leq 1 - \delta + 2\varepsilon \sum_{j=1}^k \gamma^{j-1} < 1 - \delta + 2\frac{\varepsilon}{1-\gamma} < 1 - \delta' \\ k \text{ even}, \\ 1 - q^{(1)}_m + \sum_{j=1}^k p^{(1)}_{m+j} \leq 1 - \delta + 2\varepsilon \sum_{j=1}^k \gamma^{j-1} < 1 - \delta + 2\frac{\varepsilon}{1-\gamma} < 1 - \delta' \\ k \text{ odd}. \end{cases}$$

Recall that the inductive hypothesis is that, changing from n to $m + k$ indexing, for some $k \geq 0$ such that $m + k + 1 < \zeta \wedge \sigma$, $|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| < c^k|p^{(2)}_{m+1} - p^{(2)}_m|$ and it needs to be shown that $|p^{(2)}_{m+k+2} - p^{(2)}_{m+k+1}| < c^{k+1}|p^{(2)}_{m+1} - p^{(2)}_{m}|$, and it will be done by showing that $|p^{(2)}_{m+k+2} - p^{(2)}_{m+k+1}| < c|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|$. Since by the definition of σ it still holds that $\delta' \leq p^{(2)}_{m+k+1} \leq 1-\delta'$, by the definition of ζ it still holds that $p^{(1)}_{m+k} \leq \Gamma|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|$ and the geometric decay of the first component ensures the bounds on q_{m+k} and $q^{(1)}_{m+k+1}$ as above. Then it follows that

$$\begin{aligned}
\eta_{m+k} &< p^{(1)}_{m+k} < \Gamma|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\eta_{m+k+1} &< 2p^{(1)}_{m+k} < 2\Gamma|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\eta'_{m+k+1} &< \frac{1}{(\delta')^2}p^{(1)}_{m+k} < \frac{1}{(\delta')^2}\Gamma|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\eta''_{m+k+1} &< \frac{2}{\delta'}p^{(1)}_{m+k} < \frac{2}{\delta'}\Gamma|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\eta'''_{m+k+1} &< p^{(1)}_{m+k} < \Gamma|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|.
\end{aligned}$$

Hence

$$|p^{(2)}_{m+k+2} - p^{(2)}_{m+k+1}| \leq q^{(1)}_{m+k}\,(1+\Gamma D)\,|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|,$$

since by construction

$$1+\Gamma D < \frac{c}{1-\delta'},$$

$$q^{(1)}_{m+k}(1+\Gamma D) < q^{(1)}_{m+k}\frac{c}{1-\delta'} = c\frac{q^{(1)}_{m+k}}{1-\delta'} \leq c,$$

and it follows that

$$|p^{(2)}_{m+k+2} - p^{(2)}_{m+k+1}| \leq c|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|.$$

By the induction hypothesis this implies

$$|p^{(2)}_{m+k+2} - p^{(2)}_{m+k+1}| \leq cc^k|p^{(2)}_{m+1} - p^{(2)}_{m}| = c^{k+1}|p^{(2)}_{m+1} - p^{(2)}_{m}|.$$

For $m < \tau < \sigma$ the proof is very similar: $(p_\tau, q_\tau) \in \mathcal{K}^*$ by Lemma 3.46 and definition of m and σ. Thus all the estimate for the base case apply with τ instead of m, and the inductive step is the same. □

Theorem 3.48. *$\{p_n\}$ converges to $p_* \in E_1$.*

Proof. Define $\zeta_0 := m$, and a doubly sequence of hitting times $\{\zeta_i\}$, $\{\tau_i\}$ for all $i \in \mathbb{N}$ along with the usual hitting time σ:

$$\begin{aligned}
\tau_i &:= \inf\left\{ n \geq \zeta_{i-1} : \frac{|p^{(2)}_{n+1} - p^{(2)}_{n}|}{p^{(1)}_{n}} \geq \frac{1}{\Gamma} \right\} \in \mathbb{N}\cup\infty \\
\zeta_i &:= \inf\left\{ n \geq \tau_i : \frac{|p^{(2)}_{n+1} - p^{(2)}_{n}|}{p^{(1)}_{n}} < \frac{1}{\Gamma} \right\} \in \mathbb{N}\cup\infty.
\end{aligned}$$

Note that for all $i \in \mathbb{N}$ such that $\zeta_{i-1} < \infty$, $\tau_i > \zeta_{i-1}$, and for all $i \in \mathbb{N}$ such that $\tau_i < \infty$, $\zeta_i > \tau_i$. We prove by contradiction that $\sigma = \infty$. Assume that $\sigma < \infty$, then by convention $\infty \wedge \sigma = \sigma$, thus by empty sum convention we have that the sum

$$|p^{(2)}_{\sigma} - p^{(2)}_{m}| \leq \sum_{i=0}^{\infty}\left(\sum_{n=\zeta_i\wedge\sigma}^{\tau_{i+1}\wedge\sigma-1} |p^{(2)}_{n+1} - p^{(2)}_{n}| + \sum_{n=\tau_{i+1}\wedge\sigma}^{\zeta_{i+1}\wedge\sigma-1} |p^{(2)}_{n+1} - p^{(2)}_{n}| \right)$$

is a finite sum, as either $\zeta_{\bar{i}} \leq \sigma < \tau_{\bar{i}+1}$ or $\tau_{\bar{i}} \leq \sigma < \zeta_{\bar{i}}$ for some $\bar{i} \geq 0$.

In the case $\zeta_{\bar{i}} \le \sigma < \tau_{\bar{i}+1}$, for all $\zeta_{\bar{i}} \le n \le \sigma$, $|p^{(2)}_{n+1} - p^{(2)}_n| < \Gamma^{-1} p^{(1)}_n$. By Lemma 3.46 the geometric decaying upper bound on the first component carries on at least until $p^{(1)}_\sigma$, which means that for all $\zeta_{\bar{i}} \le n < \sigma$,

$$|p^{(2)}_{n+1} - p^{(2)}_n| < \Gamma^{-1} p^{(1)}_n < 2\Gamma^{-1}\lambda^{n-m-1} p^{(1)}_m < \frac{2}{\Gamma}\lambda^{n-m-1}\varepsilon' < \left(1 + \frac{2}{\Gamma}\right)\lambda^{n-m-1}\varepsilon'$$

and the same argument applies for all $\zeta_i \le n < \tau_{i+1}$ for all $i < \bar{i}$, if there are any. For all $\tau_{\bar{i}} \le n < \zeta_{\bar{i}}$, if there are any, a different argument is needed (and similarly for all $\tau_i \le n < \zeta_i$, for $i < \bar{i}$, if there are any). In fact in this case, rearranging the condition in the hitting time's definition, for all $\tau_{\bar{i}} \le n < \zeta_{\bar{i}}$ one has that

$$\frac{p^{(1)}_n}{|p^{(2)}_{n+1} - p^{(2)}_n|} \le \Gamma,$$

which is the type of condition in Lemma 3.47. This condition, for $n = \tau_{\bar{i}}$, yields that we can apply Lemma 3.47 started at $\tau = \tau_{\bar{i}}$, implying that

$$|p^{(2)}_{n+1} - p^{(2)}_n| < \lambda^{n-\tau_{\bar{i}-1}+1} |p^{(2)}_{\tau_{\bar{i}-1}} - p^{(2)}_{\tau_{\bar{i}-1}-1}|$$

for all $\tau_{\bar{i}} \le n \le \zeta_{\bar{i}} = \zeta_{\bar{i}} \wedge \sigma$. Observe that $|p^{(2)}_{\tau_{\bar{i}}} - p^{(2)}_{\tau_{\bar{i}}-1}|$ falls in the range treated earlier, hence

$$|p^{(2)}_{\tau_{\bar{i}}} - p^{(2)}_{\tau_{\bar{i}}-1}| < \left(1 + \frac{2}{\Gamma}\right)\lambda^{\tau_{\bar{i}}-1-m-1}\varepsilon',$$

and therefore (note that it is for this very step that $1 + {}^2\!/\Gamma$ has been introduced instead of keeping working with just ${}^2\!/\Gamma$)

$$\begin{aligned}|p^{(2)}_{n+1} - p^{(2)}_n| &< \lambda^{n-\tau_{\bar{i}}+1}|p^{(2)}_{\tau_{\bar{i}}} - p^{(2)}_{\tau_{\bar{i}}-1}| < \lambda^{n-\tau_{\bar{i}}+1}\left(1 + \frac{2}{\Gamma}\right)\lambda^{\tau_{\bar{i}}-1-m-1}\varepsilon' \\ &= \left(1 + \frac{2}{\Gamma}\right)\lambda^{n-m-1}\varepsilon'.\end{aligned}$$

In the case $\tau_{\bar{i}} \le \sigma < \zeta_{\bar{i}}$, we proceed similarly, but the other way around: for all $\tau_{\bar{i}} \le n < \sigma = \zeta_{\bar{i}} \wedge \sigma$ and $\tau_i \le n \le \zeta_i = \zeta_i \wedge \sigma$, for all $i < \bar{i}$, if there are any, $|p^{(2)}_{n+1} - p^{(2)}_n| < (1 + {}^2\!/\Gamma)\,\lambda^{n-m-1}\varepsilon'$ by Lemma 3.47; while, if there are any, for all $\zeta_{i-1} \le n < \tau_i$, for all $i \le \bar{i}$, $|p^{(2)}_{n+1} - p^{(2)}_n| < (1 + {}^2\!/\Gamma)\,\lambda^{n-m-1}\varepsilon'$ by Lemma 3.46.

In both cases the conclusion is always that for every $m \le n < \sigma$,

$$|p^{(2)}_{n+1} - p^{(2)}_n| < \left(1 + \frac{2}{\Gamma}\right)\lambda^{n-m-1}\varepsilon'.$$

Therefore, by construction,

$$\begin{aligned}|p^{(2)}_\sigma - p^{(2)}_m| &\le \sum_{n=m}^{\sigma-1}|p^{(2)}_{n+1} - p^{(2)}_n| < \left(1 + \frac{2}{\Gamma}\right)\varepsilon'\sum_{n=m}^{\sigma-1}\lambda^{n-m-1} = \left(1 + \frac{2}{\Gamma}\right)\frac{\varepsilon'}{\lambda}\sum_{i=0}^{\sigma-m-1}\lambda^i \\ &< \left(1 + \frac{2}{\Gamma}\right)\frac{\varepsilon'}{\lambda}\sum_{i=0}^{\infty}\lambda^i = \varepsilon'\frac{2+\Gamma}{\Gamma\lambda(1-\lambda)} < \frac{\delta}{2}.\end{aligned}$$

The contradiction is that, since $p^{(2)}_m \in [\delta, 1-\delta]$ then having $p^{(2)}_\sigma$ travelled less than ${}^\delta\!/2$ away from $p^{(2)}_m$, we have that

$$p^{(2)}_\sigma \in \left[\frac{\delta}{2}, 1 - \frac{\delta}{2}\right],$$

in contradiction with the very own definition of σ. Since $\sigma = \infty$, $p_n^{(1)} \longrightarrow 0$ as $n \longrightarrow \infty$. As to $\{p_n^{(2)}\}$, one can repeat the two cases argument for any n (essentially by replacing σ with n and n with k for previous time indices, when necessary), getting in any case the geometric estimate

$$|p_{n+1}^{(2)} - p_n^{(2)}| < \left(1 + \frac{2}{\Gamma}\right) \lambda^{n-m-1} \varepsilon',$$

which yields

$$\sum_{n=m}^{\infty} |p_{n+1}^{(2)} - p_n^{(2)}| < \infty.$$

Therefore $p_n^{(2)} \longrightarrow p_*^{(2)} \in [\delta/2, 1 - \delta/2]$, yielding convergence of $p_n \longrightarrow p_* \in E_1$. □

Corollary 3.49. *As $\{p_n\}$ converges to $p_* \in E_1$, $\{q_n\}$ is asymptotically 2-periodic to $\{q_{p_*} \pm \frac{\ell}{2} e_{-1}(p_*)\}$.*

Proof. The asymptotic 2-periodicity of the $\{q_n\}$ follows directly from Theorem 3.48, Corollary 3.39 and Remark 3.36. Indeed, by Theorem 3.48 and Corollary 3.39, it is known that p_{m+l} converges to $p_* \in E_1$, with $\alpha_{m+l} \longrightarrow 0$ and $|\beta_{m+l}| \longrightarrow \ell/2$ as $l \longrightarrow \infty$. Thus

$$q_{m+l} - q_{p_{m+l}} = \alpha_{m+l} e_0(p_{m+l}) + \beta_{m+l} e_{-1}(p_{m+l}) = o(1) + \beta_{m+l} e_{-1}(p_*),$$

and therefore by Lemma 3.19 (h) we have that

$$q_{m+l} = q_{p_*} + \beta_{m+l} e_{-1}(p_*) + o(1)$$

with $|\beta_{m+l}| \longrightarrow \ell/2$. By Remark 3.36 it is also known that if $q_{m+2k} \longrightarrow q_{p_*} - \frac{\ell}{2} e_{-1}(p_*)$ as $k \longrightarrow \infty$, then $q_{m+2k+1} \longrightarrow q_{p_*} + \frac{\ell}{2} e_{-1}(p_*)$ as $k \longrightarrow \infty$. This is the only option as in the argument of Lemma 3.46 it has been shown how the even shifts of the $q^{(1)}$-component stay below $1/2$ and the odd ones stay above (and this carries on for all k now that by Theorem 3.48 it is known that $\sigma = \infty$). Hence the asymptotic 2-periodicity of q_{m+k} as $k \longrightarrow \infty$ follows. □

Note that the arguments, up to Theorem 3.48, rely entirely on the conditions related to hitting the set $\mathcal{K}$ suitably small, hence they can be entirely rephrased in terms of initial conditions alone, eliminating the assumption of $\ell > 0$ and nonconvergence to the vertices. In particular, the value of ℓ is crucial, only in defining the shape and size of the set $\mathcal{K}$ and $\mathcal{K}^*$, and knowing $\ell > 0$ is only required to ensure that $(p_m, q_m) \in \mathcal{K}$ for some m large enough. Besides, ℓ is only necessary in Corollary 3.49. But one could alternatively start with initial condition $(p_0, q_0) \in \mathcal{K}$ so defined and proceed with the same arguments. The only thing that would change is that in Corollary 3.49, when trying to show the asymptotic 2-periodicity of $\{q_n\}$, we would not know that $|\beta_n| \longrightarrow \ell/2$, but only that $\alpha_n \longrightarrow 0$ and that $\{q_n\}$ diverges, since in the argument of Lemma 3.44 it has been shown how the $\{q_n^{(1)}\}$ alternates between values below $1/2$, bounded away from 0 and $1/2$, and above $1/2$, bounded away from 1 and $1/2$ (and this carries on for all k now that by Theorem 3.48 it is known that $\sigma = \infty$). Hence by setting $m = 0$ in the arguments of Section 3.5.3 up to Theorem 3.48 leads to the following remark.

Remark 3.50. *Let $(p_0, q_0) \in \mathcal{K}^{\ell}_{\varepsilon',\delta}$ for any fixed admissible value $\ell > 0$. Then by Theorem 3.48, there exists $p_* \in E_1$, such that $p_n \longrightarrow p_*$ and $\{q_n\}$ diverges.*

The same is not true for the argument in Section 3.4, where determining $\ell = 0$ in advance is essential for the system not to jump back, infinitely often, in a part of the simplex bounded away from the boundary. In Section 3.4 the asymptotic assumptions, as already mentioned, are more essential than to the argument shown in this section, and cannot be reinterpreted in terms of initial conditions alone.

Remark 3.51. *If $p_* \in E_i$ for $i \in \{2,3\}$ one can proceed by exploiting the symmetry of the model, define σ, ζ_i and τ_i accordingly in terms of the corresponding coordinates, and show an analogous version of Theorem 3.48 and Remark 3.50 for $i \in \{2,3\}$ as well, thus yielding convergence of $\{p_n\}$ to some $p_* \in \partial\Sigma \setminus V$ and asymptotic 2-periodicity of $\{q_n(\omega)\}$ to $\{q_{p_*} \pm \frac{\ell}{2} e_{-1}(p_*)\}$ for any orbit having $\ell > 0$ and a subsequence bounded away from the vertices.*

3.6 Convergence of the dynamical system

In this section we put together all the convergence results gathered so far, so as to show, firstly, the convergence of $\{p_n\}$, secondly, that $\{q_n\}$ may or may not converge.

Proof of Theorem 2.1. Let $p_0 \notin \partial\Sigma$. By Lemma 3.14 the limit ℓ of the potential function exists. If $\{p_n\}$ is bounded away from the boundary, it converges by Proposition 3.15. If $\ell = 0$ and $\{p_n\}$ is not bounded away from the boundary, it converges by Remark 3.28. If $\ell > 0$ and $\{p_n\}$ is not bounded away from the boundary, it converges by Remark 3.51. By mutual exclusion the only case left is convergence to a vertex. Let $p_0 \in \partial\Sigma \setminus V$ and $q_0 \in \Sigma_0$. Then $\{p_n\}$ converges by Lemma 3.29. Let $p_0 \in E_i$ and $q_0 = v_i$. Then $\{p_n\}$ converges by Remark 3.4. There are no cases left, as both $p_0 \in V$ and $p_0 \in E_i$ with $q_0 \in V \setminus \{v_i\}$ are not admissible, as discussed in the preliminaries to this chapter. □

Proof of Corollary 2.2. Let $p_0 \notin \partial\Sigma$. By Lemma 3.14 the limit ℓ of the potential function exists. By Theorem 2.1 if $\ell = 0$, the convergence to $\overline{\Sigma}^*$ is trivial. If $\ell > 0$, the convergence to the limit 2-cycle follows either by Remark 3.51 if $p_* \in V$, or by the introductory remarks to Section 3.5.3. Let $p_0 \in \partial\Sigma \setminus V$ and $q_0 \in \Sigma_0$. Then $\{q_n\}$ either converges in $\partial\Sigma^* \subset \overline{\Sigma}^*$ or is asymptotic to a 2-cycle by Corollary 3.33. Let $p_0 \in E_i$ and $q_0 = v_i$. Then $\{q_n\}$ is 2-periodic by Remark 3.4, and thus trivially asymptotic to a 2-cycle. There are no cases left, as both $p_0 \in V$ and $p_0 \in E_i$ with $q_0 \in V \setminus \{v_i\}$ are not admissible, as discussed in the preliminaries to this chapter. □

Chapter 4

The ERBRW stochastic process

In this section we turn to the stochastic process described by (2.31) and (2.32) seen as a randomly perturbed dynamical system, taking full advantage of the tools developed by studying the deterministic dynamical system (as explained in the introduction, we will not rely directly on the results of Chapters A and 3, but on the methods established to derive them). For one-step iterations arguments, a less cumbersome notation will sometimes be used, in order to omit the time index, and (2.31) and (2.32) will often be written as

$$\hat{\pi} = M_\Theta \pi + \hat{R}$$
$$\hat{\Theta} = (1-\hat{\rho})\Theta + \hat{\rho}(\mathbf{1} - \pi - \hat{\pi}),$$

where we recall that

$$M_\Theta := \begin{pmatrix} 0 & \frac{\Theta_3}{\Theta_1+\Theta_3} & \frac{\Theta_2}{\Theta_1+\Theta_2} \\ \frac{\Theta_3}{\Theta_2+\Theta_3} & 0 & \frac{\Theta_1}{\Theta_1+\Theta_2} \\ \frac{\Theta_2}{\Theta_2+\Theta_3} & \frac{\Theta_1}{\Theta_1+\Theta_3} & 0 \end{pmatrix}.$$

The main goal of this chapter is to prove the almost sure convergence of the $\{\Theta_n\}$ to a random variable Θ, that is Theorem 1.1, based on the fast decay of the random perturbation coming from the martingale increments $\{R_n^{(i)}\}_i$. As per the construction of the model, we always assume regular initial conditions, that is $\Theta_0 \notin \partial\Sigma$. Nonetheless, in Section 4.5.1 we will also study the model given boundary initial conditions, referring to $\Theta_0 \in E_i$ for some $i \in \{1,2,3\}$, since it is simpler and provides some intuition about the model with regular initial conditions.

4.1 Preliminaries

One of the main tools that relates the stochastic process' asymptotics to the dynamical system's, is the fast decay of the perturbation terms $\{R_n^{(i)}\}_i$ extracted in the Doob's decomposition.

Lemma 4.1. *For any $1 < \nu < \sqrt{\mu}$ fixed, we have that almost surely, eventually $|R_n^{(i)}| \le \nu^{-n}$ for all $i \in \{1,2,3\}$.*

Proof. Without loss of generality, the claim will be shown for $i = 1$. Recall that by (2.28)

$$R_{n+1}^{(1)} := \frac{1}{\sigma_{n+1}} \left[\left(B_{n+1}^{(3)} - \mathbb{E}_{\mathscr{F}_n} B_{n+1}^{(3)}\right) - \left(B_{n+1}^{(2)} - \mathbb{E}_{\mathscr{F}_n} B_{n+1}^{(2)}\right) \right],$$

where conditionally on $\mathscr{F}_n$

$$B^{(3)}_{n+1} \sim \operatorname{Bin}\left(\mu N^{(3)}_n, \frac{\Theta^{(2)}_n}{\Theta^{(1)}_n + \Theta^{(2)}_n}\right)$$
$$B^{(2)}_{n+1}, \sim \operatorname{Bin}\left(\mu N^{(2)}_n, \frac{\Theta^{(1)}_n}{\Theta^{(1)}_n + \Theta^{(3)}_n}\right).$$

If μ is nonintegral, we would have to replace $B^{(3)}_{n+1}$ and $B^{(2)}_{n+1}$ with $\tilde{B}^{(3)}_{n+1}$ and $\tilde{B}^{(2)}_{n+1}$, obtaining the corresponding expression for $R^{(1)}_{n+1}$, where we recall that

$$\tilde{B}^{(3)}_{n+1} \sim \operatorname{Bin}\left(\lfloor \mu N^{(3)}_n \rfloor, \frac{\Theta^{(2)}_n}{\Theta^{(1)}_n + \Theta^{(2)}_n}\right) + \{\mu N^{(3)}_n\} \operatorname{Ber}\left(\frac{\Theta^{(2)}_n}{\Theta^{(1)}_n + \Theta^{(2)}_n}\right)$$
$$\tilde{B}^{(2)}_{n+1} \sim \operatorname{Bin}\left(\lfloor \mu N^{(2)}_n \rfloor, \frac{\Theta^{(1)}_n}{\Theta^{(1)}_n + \Theta^{(3)}_n}\right) + \{\mu N^{(2)}_n\} \operatorname{Ber}\left(\frac{\Theta^{(1)}_n}{\Theta^{(1)}_n + \Theta^{(3)}_n}\right).$$

In both cases (for nonintegral μ it is crucial that, by the specification of the model, the Bernoulli random variables are independent from the binomials conditionally on $\mathscr{F}_n$) by the *Cauchy-Schwarz inequality*

$$\begin{aligned}
\mathbb{E}_{\mathscr{F}_n}\left[R^{(1)}_{n+1}\right]^2 &= \frac{1}{\sigma^2_{n+1}}\left[\operatorname{Var}_{\mathscr{F}_n} B^{(3)}_{n+1} + \operatorname{Var}_{\mathscr{F}_n} B^{(2)}_{n+1} - 2\operatorname{Cov}_{\mathscr{F}_n}\left(B^{(3)}_{n+1}, B^{(2)}_{n+1}\right)\right] \\
&\le \frac{1}{\sigma^2_{n+1}}\left[\operatorname{Var}_{\mathscr{F}_n} B^{(3)}_{n+1} + \operatorname{Var}_{\mathscr{F}_n} B^{(2)}_{n+1} + 2\sqrt{\operatorname{Var}_{\mathscr{F}_n} B^{(3)}_{n+1} \operatorname{Var}_{\mathscr{F}_n} B^{(2)}_{n+1}}\right] \\
&\le \frac{1}{\sigma^2_{n+1}}\left(\frac{\sigma_{n+1}}{4} + \frac{\sigma_{n+1}}{4} + 2\sqrt{\left(\frac{\sigma_{n+1}}{4}\right)^2}\right) = \frac{1}{\sigma_{n+1}} = \frac{1}{\mu^{n+1}},
\end{aligned}$$

since for all $i \in \{1,2,3\}$, $\mu N^{(i)}_n \le \mu\sigma_n = \sigma_{n+1}$. With $\tilde{B}^{(3)}_{n+1}$ and $\tilde{B}^{(2)}_{n+1}$ the same bound would follow, by simply observing that for $i \in \{2,3\}$ the conditional variances of the independent Bernoulli elements are multiplied by the factors $\{\mu N^{(i)}_n\}^2 \le \{\mu N^{(i)}_n\}$, from which it follows that $\operatorname{Var}_{\mathscr{F}_n} \tilde{B}^{(3)}_{n+1} \le \operatorname{Var}_{\mathscr{F}_n} B^{(3)}_{n+1}$ and $\operatorname{Var}_{\mathscr{F}_n} \tilde{B}^{(2)}_{n+1} \le \operatorname{Var}_{\mathscr{F}_n} B^{(2)}_{n+1}$.

We conclude that

$$\mathbb{E}_{\mathscr{F}_{n-1}}\left[R^{(1)}_n\right]^2 \le \mu^{-n},$$

where $\mu > 1$. By the *conditional Markov's inequality* it follows that for any constant $1 < \nu < \sqrt{\mu}$, having set $\gamma := \nu^2/\mu < 1$,

$$\mathbb{P}_{\mathscr{F}_{n-1}}\left(|R^{(1)}_n| \ge \nu^{-n}\right) \le \nu^{2n}\mathbb{E}_{\mathscr{F}_{n-1}}\left[R^{(1)}_n\right]^2 \le \gamma^n,$$

which implies, by *Lévy's extension of Borel-Cantelli Lemma*, that

$$\mathbb{P}\left(|R^{(1)}_n| \ge \nu^{-n}, \text{ i.o.}\right) = 0,$$

as $\sum_n \gamma^n < \infty$. Recall that the probabilistic notation *i.o.* means *infinitely often* and *ev.* means *eventually.* Thus $\mathbb{P}\left(|R^{(1)}_n| < \nu^{-n}, \text{ ev.}\right) = 1$. ☐

Lemma 4.1 is the key to generalise the results obtained in Chapters A and 3. Since $\|R_n\|_1 = \sum_i |R^{(i)}_n|$, one can restate Lemma 4.1 as $\mathbb{P}\left(\|R_n\|_1 < 3\nu^{-n}, \text{ ev.}\right) = 1$, or equivalently that for almost every $\omega \in \Omega$, there is a random time $m = m(\omega) \in \mathbb{N}$ such that $\|R_n\|_1 < 3\nu^{-n}$ for all $n \ge m$. We stress that there will be no need to take the earliest such time, in the approach we will adopt.

There is a deterministic perturbation, coming from having a time-dependent ρ_n instead of its limit ρ, when adapting the arguments established for the deterministic dynamical system. This perturbation is geometric too.

Remark 4.2.

$$\rho_n - \rho = \mathcal{O}\left(\frac{1}{\mu^n}\right).$$

Proof.

$$\rho_n - \rho := \frac{\sigma_n}{\tau_n} - \frac{\mu-1}{\mu} = \frac{\mu^n}{\tau_0 + \frac{\mu}{\mu-1}(\mu^n - 1)} - \frac{\mu-1}{\mu} = \frac{1}{\frac{\tau_0}{\mu^n} + \frac{\mu}{\mu-1} - \frac{\mu}{\mu^n(\mu-1)}} - \frac{\mu-1}{\mu}$$

$$= \frac{\mu-1}{\mu}\left(\frac{1}{1+\frac{1}{\mu^n}\left(\tau_0\frac{\mu-1}{\mu}-1\right)} - 1\right) = \frac{\mu-1}{\mu}\frac{\frac{1}{\mu^n}\left(1-\tau_0\frac{\mu-1}{\mu}\right)}{1+\frac{1}{\mu^n}\left(\tau_0\frac{\mu-1}{\mu}-1\right)}.$$

□

Finally, it is easy to show the result corresponding to Remark 3.11 within the random setting. Recall that $\mathcal{D} = \{\omega \in \Omega, \exists\, v \in V,\ \Theta_n \longrightarrow v\}$ is the event of *dominance*, and note that we will often omit the dependence on ω from $\Theta_n(\omega)$ and $\pi_n(\omega)$ for simplicity, when the randomness is obvious and plays no particular role in the argument. For clarity, we show this transition in the conclusion of the following remark, which could be rewritten without any reference to ω, except for the first line.

Remark 4.3. *For any $\omega \in \mathcal{D}^c$ such that $\Theta_{n+1} - \Theta_n \longrightarrow \mathbf{0}$ as $n \longrightarrow \infty$, there is a subsequence $\{\Theta_{n_j}\}_{j\in\mathbb{N}}$ bounded away from V.*

Proof. By contradiction, if for some $\omega \in \mathcal{D}^c$ there is no such subsequence, since $\{\Theta_n(\omega)\}$ does not converge to any of the vertices (by hypothesis) but any of its subsequences approaches the set of vertices V (by contradiction), we can extract two disjoint subsequences $\{\Theta_{n_k}(\omega)\}_{k\in\mathbb{N}}$ and $\{\Theta_{n_l}(\omega)\}_{l\in\mathbb{N}}$ from $\{\Theta_n(\omega)\}$ such that

$$\{\Theta_n(\omega)\} = \{\Theta_{n_k}(\omega)\} \cup \{\Theta_{n_l}(\omega)\},$$

$\Theta_{n_k}(\omega) \longrightarrow v_i$ for some $i \in \{1,2,3\}$ (by boundedness) and $\Theta_{n_l}(\omega) \longrightarrow V\setminus\{v_i\}$. Since $\Theta_n(\omega)$ is either $\Theta_{n_k}(\omega)$ for some k, or $\Theta_{n_l}(\omega)$ for some l, for the ω fixed, there is a subsubsequence $\{\Theta_{n_{l_k}}\}_{k\in\mathbb{N}}$ such that for infinitely many k, $\Theta_{n_{l_k}} = \Theta_{n_k+1}$. For any ε fixed, by the hypothesis $\Theta_{n+1} - \Theta_n \longrightarrow \mathbf{0}$, for all k large enough $\|\Theta_{n_{l_k}} - \Theta_{n_k}\|_1 < \varepsilon$. But the 1-distance between $V\setminus\{v_i\}$ (which Θ_{n_l} approaches) and v_i (which Θ_{n_k} approaches) is 2. Since ε is arbitrary, we have a contradiction. □

4.2 Fixed points and potential

The equilibrium points for the deterministic dynamical system will be still very useful for the study of the stochastic process. Therefore, we let

$$\pi_\Theta := \frac{\mathbf{1} - \Theta}{2},$$

and we will refer to them, with abuse of language, as fixed points of the stochastic process, since they satisfy $\pi_\Theta = M_\Theta \pi_\Theta$. The same potential function, which translates into the new notation as $V(\Theta, \pi) := \|\pi - \pi_\Theta\|_1$, is also of use, but monotonicity is lost due to the random perturbations, although its convergence property can be recovered. For this reason, with abuse of language, we will still refer to it as a potential. Note that in the new notation, (A.1) and (3.1) read as

$$\pi_{\hat{\Theta}} = \frac{\mathbf{1} - \hat{\Theta}}{2} = (1-\hat{\rho})\frac{\mathbf{1} - \Theta}{2} + \frac{\hat{\rho}}{2}(\pi + \hat{\pi}) = (1-\hat{\rho})\pi_\Theta + \frac{\hat{\rho}}{2}(\pi + \hat{\pi}). \tag{4.1}$$

Thus (A.2) and (3.2) are perturbed, since

$$\hat{\pi}-\pi_{\hat{\Theta}} = \left(1-\frac{\hat{\rho}}{2}\right)\hat{\pi}-\frac{\hat{\rho}}{2}\pi-\left(1-\frac{\hat{\rho}}{2}\right)\pi_{\Theta}+\frac{\hat{\rho}}{2}\pi_{\Theta} = \left(1-\frac{\hat{\rho}}{2}\right)M_{\Theta}\pi+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}$$
$$-\left(1-\frac{\hat{\rho}}{2}\right)\pi_{\Theta}-\frac{\hat{\rho}}{2}(\pi-\pi_{\Theta}) = \left(1-\frac{\hat{\rho}}{2}\right)(M_{\Theta}\pi-\pi_{\Theta})+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}$$
$$-\frac{\hat{\rho}}{2}(\pi-\pi_{\Theta}) = \left(1-\frac{\hat{\rho}}{2}\right)M_{\Theta}(\pi-\pi_{\Theta})-\frac{\hat{\rho}}{2}(\pi-\pi_{\Theta})+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}.$$

Hence

$$\hat{\pi}-\pi_{\hat{\Theta}} = L_{\Theta}(\pi-\pi_{\Theta})+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}, \tag{4.2}$$

where

$$L_{\Theta} := \left(1-\frac{\hat{\rho}}{2}\right)M_{\Theta}-\frac{\hat{\rho}}{2}I,$$

and $\|L_{\Theta}\|_1 \le 1$ as in Sections A.2 and 3.2. Denote, for any given $\omega \in \Omega$ and the corresponding sample path $\{(\Theta_n(\omega), \pi_n(\omega))\}$ (often denoted simply as (Θ_n, π_n)), $v_n := \pi_n - \pi_{\Theta_n}$. Then $V(\Theta_n, \pi_n) = \|v_n\|_1$. Note that the iterative scheme (4.2), which in the new notation reads as

$$v_{n+1} = L_{\Theta_n}v_n + \left(1-\frac{\rho_{n+1}}{2}\right)R_{n+1},$$

started at any given time $\bar{m}$, can be compared with the iterative scheme $\bar{v}_{n+1} = L_{\Theta_n}\bar{v}_n$ started at $\bar{v}_{\bar{m}} := v_{\bar{m}}$. Denote, for any $n \ge \bar{m}$, $\overline{V}(\Theta_n, \pi_n) := \|\bar{v}_n\|_1$.

Remark 4.4. *For any given $1 < \nu < \sqrt{\mu}$, for almost every $\omega \in \Omega$, let $m = m(\omega) \in \mathbb{N}$ such that $\|R_k\|_1 < 3\nu^{-k}$ for all $k \ge m$. Then for every $\bar{m} \ge m$, uniformly in $n \ge \bar{m}$,*

$$|V(\Theta_{n+1}, \pi_{n+1}) - \overline{V}(\Theta_{n+1}, \pi_{n+1})| \le \frac{3}{\nu(\nu-1)}\frac{1}{\nu^{\bar{m}}}.$$

Proof. The existence of $m \in \mathbb{N}$ for almost every such ω and ν fixed is ensured by Lemma 4.1. Define, for any $k \le n$,

$$P_{n,k} := \prod_{i=0}^{n-k} L_{\Theta_{n-i}} = L_{\Theta_n}\dots L_{\Theta_k}.$$

Having that $\bar{v}_{n+1} = P_{n,\bar{m}}v_{\bar{m}}$, it holds that

$$v_{n+1} = \bar{v}_{n+1} + \sum_{k=\bar{m}+1}^{n+1}\left(1-\frac{\rho_{k+1}}{2}\right)P_{n,k}R_k.$$

Since $\|L_{\Theta_i}\|_1 \le 1$, by submultiplicativity of the matrix norm $\|P_{n,k}\|_1 \le 1$, thus

$$\|v_{n+1}-\bar{v}_{n+1}\|_1 \le \sum_{k=\bar{m}+1}^{n+1}\left(1-\frac{\rho_{k+1}}{2}\right)\|P_{n,k}\|_1\|R_k\|_1 < 3\sum_{k=\bar{m}+1}^{n+1}\nu^{-k} = 3\frac{\nu^{-\bar{m}-1}-\nu^{-n-2}}{1-\nu^{-1}}$$
$$= \frac{3}{\nu-1}\frac{1}{\nu^{\bar{m}}}\left(1-\frac{1}{\nu^{n-\bar{m}+1}}\right) < \frac{3}{\nu-1}\frac{1}{\nu^{\bar{m}}}$$

and therefore, uniformly in $n \ge \bar{m}$,

$$|V(\Theta_{n+1}, \pi_{n+1}) - \overline{V}(\Theta_{n+1}, \pi_{n+1})| := |\|v_{n+1}\|_1 - \|\bar{v}_{n+1}\|_1| \le \|v_{n+1}-\bar{v}_{n+1}\|_1$$
$$\le \frac{3}{\nu-1}\frac{1}{\nu^{\bar{m}}}.$$

□

Lemma 4.5. *There is a random variable $\ell : \Omega \longrightarrow [0,2]$ such that almost surely, $V(\Theta_n, \pi_n) \longrightarrow \ell$ as $n \longrightarrow \infty$.*

Proof. Fix any given $1 < \nu < \sqrt{\mu}$. For almost every fixed $\omega \in \Omega$, for which there is a finite time $m = m(\omega)$ such that for all $k > m$, $\|R_k(\omega)\|_1 < 3\nu^{-k}$ (by Lemma 4.1), suppose (by contradiction) that $V(\Theta_n, \pi_n) = \|v_n\|_1$ does not converge. In other words, suppose that for some $\delta > 0$, there are subsequences $\{n_j\}$ and $\{n_k\}$ such that

$$|\|v_{n_j+1}\|_1 - \|v_{n_k+1}\|_1| > \delta. \tag{4.3}$$

Let

$$\bar{m} \geq \max\left\{m, \frac{\log \frac{9}{\delta(\nu-1)}}{\log \nu}\right\}.$$

By Remark 4.4, uniformly in $n \geq \bar{m}$

$$|V(\Theta_{n+1}, \pi_{n+1}) - \overline{V}(\Theta_{n+1}, \pi_{n+1})| \leq \frac{3}{\nu - 1}\frac{1}{\nu^{\bar{m}}} < \frac{\delta}{3},$$

where the last inequality follows by construction of $\bar{m}$. Thus it is possible to choose J, K large enough, such that for all $j \geq J$ and $k \geq K$, $n_j, n_k \geq \bar{m}$, that is $|\|v_{n_k+1}\|_1 - \|\bar{v}_{n_k+1}\|_1| < \delta/3$ and $|\|v_{n_j+1}\|_1 - \|\bar{v}_{n_j+1}\|_1| < \delta/3$. By Remark A.6 (which guarantees, upon translating it into this section's notation, that $\|\bar{v}_n\|_1$ is nonnegative nonincreasing, and therefore implies that it is a convergent, and thus Cauchy, sequence) it is possible to choose - by taking the maximum - J, K to be also large enough to ensure that $n_j, n_k \geq N$ for all $j \geq J$ and $k \geq K$, where N is such that for all $n, n' \geq N$, $|\|\bar{v}_n\|_1 - \|\bar{v}_{n'}\|_1| < \delta/3$. Therefore, $|\|\bar{v}_{n_j+1}\|_1 - \|\bar{v}_{n_k+1}\|_1| < \delta/3$ for all $j \geq J$ and $k \geq K$. We have reached a contradiction with (4.3) for almost every ω considered, since we have shown that eventually

$$|V(\Theta_{n_j+1}, \pi_{n_j+1}) - V(\Theta_{n_k+1}, \pi_{n_k+1})| := |\|v_{n_j+1}\|_1 - \|v_{n_k+1}\|_1| \leq |\|v_{n_k+1}\|_1 - \|\bar{v}_{n_k+1}\|_1|$$
$$+ |\|v_{n_j+1}\|_1 - \|\bar{v}_{n_j+1}\|_1| + |\|\bar{v}_{n_j+1}\|_1 - \|\bar{v}_{n_k+1}\|_1| < 3\frac{\delta}{3} = \delta.$$

In conclusion, for almost every $\omega \in \Omega$, $\ell_*(\omega) := \lim_{n\longrightarrow\infty} V(\Theta_n, \pi_n)$ exists and $0 \leq \ell_* \leq 2$ (since 2 is the diameter of Σ in the 1-norm). We can thus define $\ell(\omega) := \limsup_{n\longrightarrow\infty} V(\Theta_n(\omega), \pi_n(\omega))$ for all $\omega \in \Omega$, which will be bounded between 0 and 2 as well, and $\mathscr{F}_\infty$-measurable, thus resulting to be a well defined random variable on $(\Omega, \mathscr{F}, \mathbb{P})$. □

As already mentioned in Chapter 1, we will partition the sample space Ω into three main events: $\mathcal{D} := \{\exists v \in V :\ \Theta_n \longrightarrow v\}$, $\mathscr{E} := \mathscr{S} \cap \mathcal{D}^c$ where $\mathscr{S} := \{\Theta_n \longrightarrow \partial\Sigma\}$, and $\mathscr{B} := \{\Theta_n \text{ bounded away from } \partial\Sigma\}$. Since $\mathscr{B} \subseteq \mathcal{D}^c$ and since $\mathscr{S} \cup \mathscr{B} = \Omega$, because Θ_n either approaches the boundary of the simplex or is bounded away from it, it follows that $\mathscr{E} \cup \mathscr{B} = (\mathscr{S} \cap \mathcal{D}^c) \cup \mathscr{B} = (\mathscr{S} \cup \mathscr{B}) \cap (\mathcal{D}^c \cup \mathscr{B}) = \Omega \cap \mathcal{D}^c = \mathcal{D}^c$. Thus the following remark holds.

Remark 4.6. *$\mathcal{D}$, $\mathscr{E}$, $\mathscr{B}$, form a partitioning of Ω.*

We can also consider a different partitioning of Ω given by the two events $\{\ell = 0\}$ and $\{\ell > 0\}$. Finally consider the following events:

- $\mathscr{B}_0 := \mathscr{B} \cap \{\ell = 0\}$, $\mathscr{B}_> := \mathscr{B} \cap \{\ell > 0\}$;
- $\mathscr{E}_0 := \mathscr{E} \cap \{\ell = 0\}$, $\mathscr{E}_> := \mathscr{E} \cap \{\ell > 0\}$;

- $\mathcal{D}_0 := \mathcal{D} \cap \{\ell = 0\}$, $\mathcal{D}_> := \mathcal{D} \cap \{\ell > 0\}$.

We will work on each event separately, showing that for almost every ω in each event $\{\Theta_n(\omega)\}$ converges, yielding almost sure convergence of $\{\Theta_n\}$ to a random variable Θ.

4.3 Convergence on $\mathscr{B}$

The main goal of this section is showing that the stochastic process $\{(\Theta_n, \pi_n)\}$ converges almost surely on $\mathscr{B}$.

Proposition 4.7. *$\mathbb{P}(\mathscr{B} \setminus \mathscr{B}_0) = 0$ and thus, for almost every $\omega \in \mathscr{B}$, the sample path $\{(\Theta_n(\omega), \pi_n(\omega))\}$ converges to an internal equilibrium.*

Proof. We prove first the almost sure convergence of Θ_n. Consider that

$$\begin{aligned}\Theta_{n+1} - \Theta_n &= \rho_{n+1}(\mathbf{1} - \Theta_n - \pi_{n+1} - \pi_n) = \rho_{n+1}(2\pi_{\Theta_n} - \pi_{n+1} - \pi_n) \\ &= -\rho_{n+1}(M_{\Theta_n} + I)v_n - \rho_{n+1}R_{n+1}\end{aligned}$$

and therefore

$$\|\Theta_{n+1} - \Theta_n\|_1 \leq 2\|v_n\|_1 + \|R_{n+1}\|_1. \tag{4.4}$$

We will use (4.4), which holds in general, to show that for almost every $\omega \in \mathscr{B}$,

$$\sum_k \|\Theta_{k+1} - \Theta_k\|_1 < \infty,$$

which implies directly that since

$$\Theta_{n+1} = \Theta_m + \sum_{k=m}^{n} (\Theta_{k+1} - \Theta_k),$$

Θ_n converges as $n \longrightarrow \infty$, where $m = m(\omega) \in \mathbb{N}$ exists by Lemma 4.1, and is a random time such that, given a fixed $1 < \nu < \sqrt{\mu}$ and for almost every ω, for all $k \geq m$, $\|R_k\|_1 \leq 3\nu^{-k}$. By (4.4)

$$\sum_{k=m}^{n} \|\Theta_{k+1} - \Theta_k\|_1 \leq 2 \sum_{k=m}^{n} \|v_k\|_1 + \sum_{k=m+1}^{n+1} \|R_k\|_1$$

and $\sum_k \|R_k\|_1 < \infty$ due to Lemma 4.1, so we only need to prove that $\sum_k \|v_k\|_1 < \infty$. To this end, note that for any fixed $\omega \in B$, there is $\varepsilon = \varepsilon(\omega)$ such that $\Theta_n \in \Sigma_\varepsilon$ for all $n \in \mathbb{N}$, where Σ_ε is defined as in Lemma 3.14. Consider that, upon translating it into this section's notation, the estimate of (A.3) in Lemma A.8 still applies to L_{Θ_n}, having the time-dependent $0 < \rho_{n+1} < 1$ instead of ρ, by Remark 4.2. Then, translating Lemma A.8 in the notation we set for this section, we have that $\|L_{\Theta_n} v_n\|_1 < \mathscr{c}\|v_n\|_1$ for some $0 \leq \mathscr{c}(\omega) < 1$. Hence by taking norms in (4.2) and by the triangle inequality, one gets

$$\|v_{n+1}\|_1 < \mathscr{c}\|v_n\|_1 + \|R_{n+1}\|_1. \tag{4.5}$$

For all $n \geq m$, (4.5) reads

$$\|v_{n+1}\|_1 < \mathscr{c}\|v_n\|_1 + 3\nu^{-n-1}. \tag{4.6}$$

Let $\lambda := \max\{c, \nu^{-1}\}$ and iterate the bound of (4.6) down to time m, starting at any $n > m$, yielding

$$\|v_{n+1}\|_1 \leq c^{n-m+1}\|v_m\|_1 + 3\sum_{k=m+1}^{n+1} c^{n-k+1}\nu^{-k} \leq \lambda^{n-m+1}\|v_m\|_1 + 3\sum_{k=m+1}^{n+1} \lambda^{n-k+1}\lambda^k$$
$$= \lambda^{n-m+1}\|v_m\|_1 + 3(n-m)\lambda^{n+1}.$$

By the boundedness of v_m, following from the boundedness of Σ, one sees that on $\mathscr{B}$, $\|v_k\|_1 = \mathcal{O}_\omega(k\lambda^k)$, hence $\sum_k \|v_k\|_1 < \infty$. This shows that:

- $\|v_k\|_1$ vanishes for almost all $\omega \in \mathscr{B}$, and hence that $\mathbb{P}(\mathscr{B} \setminus \mathscr{B}_0) = 0$, because by construction on $\mathscr{B} \setminus \mathscr{B}_0$ we have that $\|v_k\|_1$ does not vanish;
- Θ_n converges to some $\Theta \notin \partial\Sigma$ (indeed by definition of $\mathscr{B}$ the limit cannot be on the boundary), thus yielding convergence of π_n to $\pi_\Theta \in \Sigma^*$ (recall that Σ^* is the interior of the medial triangle of the simplex), since $\|v_n\|_1 = \|\pi_n - \pi_{\Theta_n}\|_1 \longrightarrow 0$ as $n \longrightarrow \infty$.

□

The argument of Proposition 3.17 is robust under geometric perturbations, and can be adjusted to the random setting as follows. As we continue working in 1-norm, in the following $U((\Theta,\pi),r,r') := B(\Theta,r) \times B(\pi,r')$ and $\mathrm{dist}(\cdot,\cdot)$ are with respect to the 1-norm, and $U((\Theta,\pi),r) := U((\Theta,\pi),r,r)$, having denoted by $B(\Theta,r)$ the ball centred at Θ of radius r, with respect to the distance generated by the 1-norm.

Proposition 4.8. *For every $1 < \nu < \sqrt{\mu}$, $\Theta \notin \partial\Sigma$ and a small enough $0 < \varepsilon' < \mathrm{dist}(\Theta, \partial\Sigma)$ given, there is a $\delta' > 0$ small enough such that, if for $\omega \in \Omega$ fixed there is a finite $m = m(\omega)$ large enough, such that $\|R_{n+1}\|_1 < 3\nu^{-n-1}$ for all $n \geq m$, $3\nu^{-m} < \delta'$ and $(\Theta_m(\omega), \pi_m(\omega)) \in U((\Theta,\pi_\Theta),\delta')$; then*

$$(\Theta_n(\omega), \pi_n(\omega)) \in U\left((\Theta,\pi_\Theta), \varepsilon', \frac{\varepsilon'}{2}\right)$$

for all $n > m$.

Proof. Given Θ and ε' as in the statement, let $0 < \delta' < \varepsilon'$ (it will be further restricted if necessary); by (4.5) in Proposition 4.7 it is known that since $B(\Theta,\varepsilon')$ is bounded away from $\partial\Sigma$, for all $(\Theta',\pi') \in U((\Theta,\pi_\Theta),\varepsilon')$, $\|L_{\Theta'}(\pi - \pi_{\Theta'})\|_1 \leq c\|\pi - \pi_{\Theta'}\|_1 + \|\hat{R}\|_1$ for some $0 < c < 1$. Denote by $\lambda := \max\{c, \nu^{-1}\}$. Further restrict, given $1 < \nu < \sqrt{\mu}$,

$$\delta' < \frac{\varepsilon'}{\frac{3}{1-\lambda^2} + \frac{\nu}{\nu-1}}.$$

For almost every fixed $\omega \in \Omega$, the time $m = m(\omega) \in \mathbb{N}$ is well defined by Lemma 4.1 and the monotonicity of ν^{-n}. If $(\Theta_m(\omega), \pi_m(\omega)) \in U((\Theta,\pi_\Theta),\delta')$ we will show by induction that for all $n \geq m+1$, $\Theta_n \in B(\Theta,\varepsilon')$ and $\pi_n \in B(\pi_\Theta, \varepsilon'/2)$.

Consider that

$$\|\pi_m - \pi_{\Theta_m}\|_1 \leq \|\pi_m - \pi_\Theta\|_1 + \|\pi_\Theta - \pi_{\Theta_m}\|_1 = \|\pi_m - \pi_\Theta\|_1 + \frac{\|\Theta - \Theta_m\|_1}{2} < \frac{3}{2}\delta'.$$

Then since $3\nu^{-m} < \delta'$

$$\|\pi_{m+1} - \pi_{\Theta_{m+1}}\|_1 \leq \|L_{\Theta_m}(\pi_m - \pi_{\Theta_m})\|_1 + \|R_{m+1}\|_1 \leq c\|\pi_m - \pi_{\Theta_m}\|_1 + 3\nu^{-m-1}$$
$$< \frac{3}{2}c\delta' + \delta'\nu^{-1} \leq \frac{3\delta'}{2}2\lambda^1.$$

Recall that for all n, $\|\Theta_{n+1}-\Theta_n\|_1 \leq 2\|\pi_n - \pi_{\Theta_n}\|_1 + \|R_{n+1}\|_1$ by (4.4), hence $\|\Theta_{m+1} - \Theta_m\|_1 < 3\delta' + 3\nu^{-m-1}$. As a consequence, noting that $\pi_{\Theta_{m+1}} - \pi_\Theta = (\Theta - \Theta_{m+1})/2$,

$$\|\Theta_{m+1}-\Theta\|_1 \leq \|\Theta_{m+1}-\Theta_m\|_1+\|\Theta_m-\Theta\|_1 < 3\delta'+3\nu^{-m-1}+\delta' < 3\delta'(1\lambda^0)+\delta'(1+\nu^{-1})$$

which implies in particular by construction of δ' that

$$\|\Theta_{m+1} - \Theta\|_1 < \delta'(3+1+\nu^{-1}) < \delta'\left(\frac{3}{(1-\lambda)^2} + \frac{\nu}{\nu-1}\right) < \varepsilon',$$

and

$$\begin{aligned}\|\pi_{m+1} - \pi_\Theta\|_1 \leq \|\pi_{m+1} - \pi_{\Theta_{m+1}}\|_1 + \|\pi_{\Theta_{m+1}} - \pi_\Theta\|_1 = \|\pi_{m+1} - \pi_{\Theta_{m+1}}\|_1 \\ + \frac{\|\Theta_{m+1} - \Theta\|_1}{2} \leq \frac{3}{2}\delta' 2\lambda^1 + \frac{3}{2}\delta' 1\lambda^0 + \frac{\delta'}{2}(1+\nu^{-1}),\end{aligned}$$

which implies in particular that

$$\|\pi_{m+1} - \pi_\Theta\|_1 \leq \frac{\delta'}{2}\left(3(1\lambda^0 + 2\lambda^1) + (1+\nu^{-1})\right) < \frac{\delta'}{2}\left(\frac{3}{(1-\lambda)^2} + \frac{\nu}{\nu-1}\right) < \frac{\varepsilon'}{2}.$$

It immediately follows, by exploiting again the geometric decay by the constant factor c, still valid by the previous estimates, that

$$\begin{aligned}\|\pi_{m+2} - \pi_{\Theta_{m+2}}\|_1 &\leq \|L_{\Theta_{m+1}}(\pi_{m+1} - \pi_{\Theta_{m+1}})\|_1 + \|R_{m+2}\|_1 \leq c\|\pi_{m+1} - \pi_{\Theta_{m+1}}\|_1 \\ &+ \|R_{m+2}\|_1 \leq c^2\|\pi_m - \pi_{\Theta_m}\|_1 + c\|R_{m+1}\|_1 + \|R_{m+2}\|_1 \\ &< \frac{3}{2}c^2\delta' + 3c\nu^{-m-1} + 3\nu^{-m-2} < \frac{3}{2}c^2\delta' + \delta' c\nu^{-1} + \delta'\nu^{-2} < \frac{3\delta'}{2}3\lambda^2.\end{aligned}$$

To sum up what proved for $k=1$, it has been shown that if $(\Theta_m, \pi_m) \in U((\Theta, \pi_\Theta), \delta')$, with c being the subunitary constant uniformly holding on $B(\Theta, \varepsilon')$, then

$$\|\pi_{m+1} - \pi_{\Theta_{m+1}}\|_1 < \frac{3}{2}\delta'(2\lambda^1),$$

$$\|\Theta_{m+1} - \Theta\|_1 < 3\delta'(1\lambda^0) + \delta'(1+\nu^{-1}),$$

so that $\Theta_{m+1} \in B(\Theta, \varepsilon')$, and

$$\|\pi_{m+1} - \pi_\Theta\|_1 \leq \frac{3}{2}\delta'(2\lambda^1 + 1\lambda^0) + \frac{\delta'}{2}(1+\nu^{-1}),$$

so that $\pi_{m+1} \in B(\pi_\Theta, \varepsilon'/2)$.

Assume as induction hypothesis that

$$\|\pi_{m+k} - \pi_{\Theta_{m+k}}\|_1 < \frac{3}{2}\delta'(k+1)\lambda^k,$$

that

$$\|\Theta_{m+k} - \Theta\|_1 < 3\delta'\sum_{i=0}^{k} i\lambda^{i-1} + \delta'\sum_{i=0}^{k}\nu^{-i},$$

so that $\Theta_{m+k} \in B(\Theta, \varepsilon')$, and

$$\|\pi_{m+k} - \pi_\Theta\|_1 < \frac{3}{2}\delta'\sum_{i=0}^{k+1} i\lambda^{i-1} + \frac{\delta'}{2}\sum_{i=0}^{k}\nu^{-i}$$

so that $\pi_{m+k} \in B(\Theta, \varepsilon'/2)$, and consider Θ_{m+k+1}. Since

$$\begin{aligned}\|\Theta_{m+k+1} - \Theta\|_1 \leq \|\Theta_{m+k+1} - \Theta_{m+k}\|_1 + \|\Theta_{m+k} - \Theta\|_1 \leq 2\|\pi_{m+k} - \pi_{\Theta_{m+k}}\|_1 \\ + \|R_{m+k+1}\|_1 + \|\Theta_{m+k} - \Theta\|_1 < 3\delta'(k+1)\lambda^k + \delta'\nu^{-k-1} \\ + 3\delta' \sum_{i=0}^{k} i\lambda^{i-1} + \delta' \sum_{i=0}^{k} \nu^{-i} = 3\delta' \sum_{i=0}^{k+1} i\lambda^{i-1} + \delta' \sum_{i=0}^{k+1} \nu^{-i},\end{aligned}$$

this shows that $\Theta_{m+k+1} \in B(\Theta, \varepsilon')$, since

$$\|\Theta_{m+k+1} - \Theta\|_1 < \delta' \left(3\sum_{i=0}^{\infty} i\lambda^{i-1} + \sum_{i=0}^{\infty} \nu^{-i} \right) = \delta' \left(\frac{3}{(1-\lambda)^2} + \frac{\nu}{\nu - 1} \right) < \varepsilon'$$

by construction of δ', and therefore it also holds that

$$\begin{aligned}\|\pi_{m+k+1} - \pi_{\Theta_{m+k+1}}\|_1 < \|L_{\Theta_{m+k}}(\pi_{m+k} - \pi_{\Theta_{m+k}})\|_1 + \|R_{m+k+1}\|_1 < c\|\pi_{m+k} - \pi_{\Theta_{m+k}}\|_1 \\ + \delta'\nu^{-k-1} < \frac{3}{2}c\delta'(k+1)\lambda^k + \delta'\nu^{-k-1} \leq \frac{3}{2}\delta'(k+1)\lambda^{k+1} + \delta'\lambda^{k+1} \\ < \frac{3}{2}\delta'(k+2)\lambda^{k+1}.\end{aligned}$$

Since

$$\begin{aligned}\|\pi_{m+k+1} - \pi_\Theta\|_1 \leq \|\pi_{m+k+1} - \pi_{\Theta_{m+k+1}}\|_1 + \|\pi_{\Theta_{m+k+1}} - \pi_\Theta\|_1 \leq \frac{3}{2}\delta'(k+2)\lambda^{k+1} \\ + \frac{1}{2}\|\Theta_{m+k+1} - \Theta\|_1 < \frac{3}{2}\delta'(k+2)\lambda^{k+1} + \frac{3}{2}\delta' \sum_{i=0}^{k+1} i\lambda^{i-1} + \frac{\delta'}{2} \sum_{i=0}^{k+1} \nu^{-i} \\ = \frac{3}{2}\delta' \sum_{i=0}^{k+2} i\lambda^{i-1} + \frac{\delta'}{2} \sum_{i=0}^{k+1} \nu^{-i},\end{aligned}$$

this shows that $\pi_{m+k+1} \in B(\pi_\Theta, \varepsilon'/2)$, since

$$\|\pi_{m+k+1} - \pi_\Theta\|_1 < \frac{\delta'}{2} \left(3\sum_{i=0}^{\infty} i\lambda^{i-1} + \sum_{i=0}^{\infty} \nu^{-i} \right) = \frac{\delta'}{2} \left(\frac{3}{(1-\lambda)^2} + \frac{\nu}{\nu - 1} \right) < \frac{\varepsilon'}{2}$$

by hypothesis. □

Corollary 4.9. *For almost all $\omega \in \Omega$, by Lemma 4.1 there is a random time $m = m(\omega) \in \mathbb{N}$ large enough such that if $(\Theta_m(\omega), \pi_m(\omega))$ is close enough to an internal equilibrium (Θ, π_Θ), by Propositions 4.7 and 4.8 the stochastic process converges to a random internal equilibrium.*

4.4 Convergence on $\mathscr{E}_0$

The main goal of this section is showing that the stochastic process $\{(\Theta_n, \pi_n)\}$ converges almost surely on $\mathscr{E}_0$.

For almost every $\omega \in \mathscr{E}_0$ fixed, $\|\Theta_{n+1} - \Theta_n\|_1 \longrightarrow 0$ as $n \longrightarrow \infty$, by (4.4) (which holds in general, not just on $\mathscr{B}$, where it has been used so far) and Lemma 4.1. By Remark 4.3 it follows that for almost every $\omega \in \mathscr{E}_0$ there is a subsequence

$\{\Theta_{n_j}\}_{j\in\mathbb{N}}$ bounded away from the vertices. By boundedness, this implies the existence of a subsubsequence $\{\Theta_{n_{j_l}}\}$ (relabelled with n_k for simplicity) such that $\Theta_{n_k}(\omega) \longrightarrow \Theta_*(\omega) \in E_i$ for some $i \in \{1,2,3\}$, for almost every $\omega \in \mathscr{E}_0$. Define

$$\mathscr{E}_0^{(i)} := \{\omega \in \mathscr{E}_0 : \; \exists\{n_k\}_{k\in\mathbb{N}},\ \Theta_{n_k}(\omega) \longrightarrow \Theta_*(\omega) \in E_i \text{ as } k \longrightarrow \infty\}.$$

Then we have that

$$\mathbb{P}\left(\mathscr{E}_0 \setminus \bigcup_{i=1}^{3} \mathscr{E}_0^{(i)}\right) = 0.$$

By symmetry, we will show the argument of convergence on $\mathscr{E}_0^{(1)}$ without loss of generality. For almost every $\omega \in \mathscr{E}_0^{(1)}$, $\Theta_{n_k}^{(1)} \longrightarrow 0$ as $k \longrightarrow \infty$, while $\{\Theta_{n_k}^{(2)}\}_{k\in\mathbb{N}}$ is bounded away from 0 and 1. Since on $\mathscr{E}_0^{(1)}$, $\|\pi_n - \pi_{\Theta_n}\|_1 \longrightarrow 0$, and for almost every such ω by Lemma 4.1 there is an $\bar{m} = \bar{m}(\omega)$ such that for any fixed $1 < \nu < \sqrt{\mu}$, $3\nu^{-\bar{m}} < \varepsilon$ and $|R_n^{(i)}| < \nu^{-n}$ for all $i \in \{1,2,3\}$ and $n \geq \bar{m}$; there will be, in conclusion, a large enough K such that, for any $k \geq K = K(\omega)$, $n_k \geq \bar{m}$ and for any sufficiently small enough $\delta > 0$, $\delta = \delta(\omega)$, and an arbitrarily small $\varepsilon > 0$, dependent on δ, $(\Theta_{n_k}, \pi_{n_k}, R_{n_k+1})$ belongs to

$$\mathscr{K}^*_{\varepsilon,\frac{\delta}{2}} := \left\{(\Theta, \pi, \hat{R}) \in \Sigma^2 \times \Pi_0 : \; 0 < \Theta_1 \leq \varepsilon,\ \frac{\delta}{2} \leq \Theta_2 \leq 1 - \frac{\delta}{2},\ 0 \leq |\alpha|,\ |\beta|,\ \|\hat{R}\|_1 \leq \varepsilon\right\}$$

where, for each $\Theta \in \Sigma_0$, the usual notation for the eigenvectors spanning Π_0 from Lemma 3.19 has been adopted, leading to the representation $\pi - \pi_\Theta = \alpha e_0(\Theta) + \beta e_{-1}(\Theta)$, so that α_n and β_n are eigencoordinates for $\pi_n - \pi_{\Theta_n}$. Since $\pi_n(\omega) - \pi_{\Theta_n(\omega)} \longrightarrow \mathbf{0}$ and the norm of the linearly independent eigenvectors is bounded away from zero, for any $\varepsilon > 0$ eventually $|\alpha_n| \in \{0 < \alpha \leq \varepsilon,\ \alpha \in \mathbb{R}\}$ and $|\beta_n| \in \{0 < \beta \leq \varepsilon,\ \beta \in \mathbb{R}\}$. Define a similar set $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$ and denote

$$\mathscr{K}_{\varepsilon,\frac{\delta}{8}} := \left\{(\Theta, \hat{R}) \in \Sigma \times \Pi_0 : \; 0 < \Theta_1 \leq \varepsilon,\ \frac{\delta}{8} \leq \Theta_2 \leq 1 - \frac{\delta}{8},\ 0 \leq \|\hat{R}\|_1 \leq \varepsilon\right\}.$$

Note that ignoring the $\hat{R}$ coordinate in the Cartesian product, the intuitive picture for these sets is related to Figure 3.3.

We will adopt the same modified $\mathcal{O}$-notation as in Section 3.4, that is $f(\Theta) = \mathcal{O}(g(\Theta_1))$ if for $\varepsilon > 0$ small enough $f(\Theta)/g(\Theta_1)$ is well-defined and bounded on $\mathscr{K}_{\varepsilon,\frac{\delta}{8}}$. We will also adapt it to the new set of variables when necessary: $f(\Theta, \alpha, \beta, \hat{R}) = \mathcal{O}(g_1(\Theta_1, \alpha, \beta, \|\hat{R}\|_1), \ldots, g_k(\Theta_1, \alpha, \beta, \|\hat{R}\|_1))$ if, for sufficiently small $\varepsilon > 0$,

$$\frac{f(\Theta, \alpha, \beta, \hat{R})}{|g_1(\Theta_1, \alpha, \beta, \|\hat{R}\|_1)| + \ldots + |g_k(\Theta_1, \alpha, \beta, \|\hat{R}\|_1)|}$$

is well-defined and bounded on $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$. Denote $r_n := (\Theta_n^{(1)}, \Theta_n^{(2)}, \alpha_n, \beta_n, \|R_{n+1}\|_1)$.

Lemma 4.10.

$$\hat{\alpha} = \alpha\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_0(\Theta)-\frac{\hat{\rho}}{2}\right]\frac{\begin{vmatrix} e_0^{(i)}(\Theta) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\Theta) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}} + \beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_{-1}(\Theta)-\frac{\hat{\rho}}{2}\right]\frac{\begin{vmatrix} e_{-1}^{(i)}(\Theta) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_{-1}^{(j)}(\Theta) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}} + \left(1-\frac{\hat{\rho}}{2}\right)\frac{\begin{vmatrix} \hat{R}^{(i)} & e_{-1}^{(i)}(\hat{\Theta}) \\ \hat{R}^{(j)} & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}} \tag{4.7}$$

$$\hat{\beta} = \alpha\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_0(\Theta)-\frac{\hat{\rho}}{2}\right]\frac{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_0^{(i)}(\Theta) \\ e_0^{(j)}(\hat{\Theta}) & e_0^{(j)}(\Theta)\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}} + \beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_{-1}(\Theta)-\frac{\hat{\rho}}{2}\right]\frac{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\Theta) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\Theta)\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}} + \left(1-\frac{\hat{\rho}}{2}\right)\frac{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & \hat{R}^{(i)} \\ e_0^{(j)}(\hat{\Theta}) & \hat{R}^{(j)}\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}} \tag{4.8}$$

Proof. By (4.2) we obtain a system of three linear equations in two variables $(\hat{\alpha},\hat{\beta}) \in \mathbb{R}^2$,

$$\hat{\alpha}e_0(\hat{\Theta})+\hat{\beta}e_{-1}(\hat{\Theta}) = \alpha\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_0(\Theta)-\frac{\hat{\rho}}{2}\right]e_0(\Theta)+\beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_{-1}(\Theta)-\frac{\hat{\rho}}{2}\right]e_{-1}(\Theta) + \left(1-\frac{\hat{\rho}}{2}\right)\hat{R},$$

which can therefore be solved by picking any two of the three equations as in Lemma 3.20. For any $i \neq j$ chosen, the linear system

$$\begin{pmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta})\end{pmatrix}\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix} = \begin{pmatrix} \alpha\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_0(\Theta)-\frac{\hat{\rho}}{2}\right]e_0^{(i)}(\Theta)+\beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_{-1}(\Theta)-\frac{\hat{\rho}}{2}\right]e_{-1}^{(i)}(\Theta)+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}^{(i)} \\ \alpha\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_0(\Theta)-\frac{\hat{\rho}}{2}\right]e_0^{(j)}(\Theta)+\beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_{-1}(\Theta)-\frac{\hat{\rho}}{2}\right]e_{-1}^{(j)}(\Theta)+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}^{(j)}\end{pmatrix}$$

is solved similarly to how it has been done in Lemma 3.20, via Cramer's rule,

$$\hat{\alpha} = \frac{\begin{vmatrix} \alpha\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_0(\Theta)-\frac{\hat{\rho}}{2}\right]e_0^{(i)}(\Theta)+\beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_{-1}(\Theta)-\frac{\hat{\rho}}{2}\right]e_{-1}^{(i)}(\Theta)+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}^{(i)} & e_{-1}^{(i)}(\hat{\Theta}) \\ \alpha\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_0(\Theta)-\frac{\hat{\rho}}{2}\right]e_0^{(j)}(\Theta)+\beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_{-1}(\Theta)-\frac{\hat{\rho}}{2}\right]e_{-1}^{(j)}(\Theta)+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}^{(j)} & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}}$$

$$\hat{\beta} = \frac{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & \alpha\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_0(\Theta)-\frac{\hat{\rho}}{2}\right]e_0^{(i)}(\Theta)+\beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_{-1}(\Theta)-\frac{\hat{\rho}}{2}\right]e_{-1}^{(i)}(\Theta)+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}^{(i)} \\ e_0^{(j)}(\hat{\Theta}) & \alpha\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_0(\Theta)-\frac{\hat{\rho}}{2}\right]e_0^{(j)}(\Theta)+\beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\lambda_{-1}(\Theta)-\frac{\hat{\rho}}{2}\right]e_{-1}^{(j)}(\Theta)+\left(1-\frac{\hat{\rho}}{2}\right)\hat{R}^{(j)}\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta})\end{vmatrix}},$$

yielding the claim. The ratios of the determinants do not depend on the choice of $i \neq j$, because $\hat{R} \in \Pi_0$, and therefore for $k \neq i \neq j$, $\hat{R}_k = -\hat{R}_i - \hat{R}_j$. This means that the concluding remarks of Lemma 3.20 also apply to this system. □

Lemma 4.11.

$$\hat{\Theta}_1 = \Theta_1 + \rho_1(r) \tag{4.9}$$
$$\hat{\Theta}_2 = \Theta_2 - 2\hat{\rho}(1-\Theta_2)\Theta_1\beta + \rho_2(r), \tag{4.10}$$

where $\rho_1(r) = \mathcal{O}(\beta\Theta_1, \alpha\Theta_1, \|\hat{R}\|_1)$ *and* $\rho_2(r) = \mathcal{O}(\alpha, \beta\Theta_1^2, \|\hat{R}\|_1)$.

Proof. Since

$$\hat{\Theta} - \Theta = \hat{\rho}(\mathbf{1} - \Theta - \hat{\pi} - \pi) = \hat{\rho}(2\pi_\Theta - M_\Theta\pi - \hat{R} - \pi) = -\hat{\rho}(M_\Theta + I)(\pi - \pi_\Theta) - \hat{\rho}\hat{R},$$

one has that

$$\hat{\Theta} = \Theta - \hat{\rho}\alpha(1+\lambda_0(\Theta))e_0(\Theta) - \hat{\rho}\beta(1+\lambda_{-1}(\Theta))e_{-1}(\Theta) - \hat{\rho}\hat{R}, \tag{4.11}$$

from which, reading off the first two components and applying Lemma 3.19 (f, g, h), it follows that

$$\begin{aligned}
\hat{\Theta}_1 &= \Theta_1 - \hat{\rho}\alpha(1 - 2\Theta_1 + \mathcal{O}(\Theta_1^2))\mathcal{O}(\Theta_1) - \hat{\rho}\beta(2\Theta_1 + \mathcal{O}(\Theta_1^2))(-1 + \mathcal{O}(\Theta_1)) - \hat{\rho}\hat{R}_1 \\
&= \Theta_1 + \mathcal{O}(\alpha\Theta_1, \beta\Theta_1) - \hat{\rho}\hat{R}_1 \\
\hat{\Theta}_2 &= \Theta_2 - \hat{\rho}\alpha(1 - 2\Theta_1 + \mathcal{O}(\Theta_1^2))(1 + \mathcal{O}(\Theta_1)) - \hat{\rho}\beta(2\Theta_1 + \mathcal{O}(\Theta_1^2))(1 - \Theta_2 + \mathcal{O}(\Theta_1)) \\
&\quad - \hat{\rho}\hat{R}_2 = \Theta_2 - 2\hat{\rho}(1-\Theta_2)\Theta_1\beta + \mathcal{O}(\alpha, \beta\Theta_1^2) - \hat{\rho}\hat{R}_2,
\end{aligned}$$

having used the smoothness of the eigenvectors to linearise as Θ approaches the edge E_1, and the relative compactness of $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$ to estimate uniformly the Jacobian term as in the concluding remark of Lemma 3.21. □

Lemma 4.12.

$$\hat{\alpha} = -\frac{\hat{\rho}}{2}\alpha(1 + \rho_3(r)) + \rho_4(r) \tag{4.12}$$
$$\hat{\beta} = -\beta\left[1 - (2-\hat{\rho})\,\Theta_1\right] + \rho_5(r), \tag{4.13}$$

where $\rho_3(r) = \mathcal{O}(\alpha, \Theta_1, \|\hat{R}\|_1)$, $\rho_4(r) = \mathcal{O}(\beta\alpha, \beta^2\Theta_1, \|\hat{R}\|_1)$, $\rho_5(r) = \mathcal{O}(\alpha^2, \alpha\beta, \beta^2\Theta_1, \beta\Theta_1^2, \|\hat{R}\|_1)$.

Proof. By Lemma 4.11 it follows that $\hat{\Theta} = \Theta + \boldsymbol{\mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)}$, because $\hat{\Theta}_3 = 1 - \hat{\Theta}_1 - \hat{\Theta}_2 = 1 - \Theta_1 - \Theta_2 + \mathcal{O}(\alpha, \beta\Theta_1) + \hat{\rho}(\hat{R}_1 + \hat{R}_2) = \Theta_3 + \mathcal{O}(\alpha, \beta\Theta_1) - \hat{\rho}\hat{R}_3$. We plug this estimate, along with that of Lemma 3.19 (g), in the terms next to α and β, in (4.7) and (4.8). This yields, due to smoothness of the eigenvectors' components and relative compactness of $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$, the following estimates for those terms involved in (4.7):

$$\frac{\begin{vmatrix} e_0^{(i)}(\Theta) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\Theta) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix}} = \frac{\begin{vmatrix} e_0^{(i)}(\Theta) & e_{-1}^{(i)}(\Theta) \\ e_0^{(j)}(\Theta) & e_{-1}^{(j)}(\Theta) \end{vmatrix} + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)}{\begin{vmatrix} e_0^{(i)}(\Theta) & e_{-1}^{(i)}(\Theta) \\ e_0^{(j)}(\Theta) & e_{-1}^{(j)}(\Theta) \end{vmatrix} + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)}$$

$$\frac{\begin{vmatrix} e_{-1}^{(i)}(\Theta) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_{-1}^{(j)}(\Theta) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix}} = \frac{\begin{vmatrix} e_{-1}^{(i)}(\Theta) & e_{-1}^{(i)}(\Theta) \\ e_{-1}^{(j)}(\Theta) & e_{-1}^{(j)}(\Theta) \end{vmatrix} + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)}{\begin{vmatrix} e_0^{(i)}(\Theta) & e_{-1}^{(i)}(\Theta) \\ e_0^{(j)}(\Theta) & e_{-1}^{(j)}(\Theta) \end{vmatrix} + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)}$$

$$\left| \frac{\begin{vmatrix} \hat{R}^{(i)} & e_{-1}^{(i)}(\hat{\Theta}) \\ \hat{R}^{(j)} & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix}} \right| \leq \frac{\|\hat{R}\|_1 \|e_{-1}(\hat{\Theta})\|_1}{\left\| \begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix} \right\|} = \mathcal{O}(\|\hat{R}\|_1),$$

where the last estimate follows by *Hadamard's inequality* (which holds for the Euclidean norm), the fact that the Euclidean norm is always smaller than the 1-norm, the boundedness away from zero of the determinant in the denominator, the fact that $\|e_{-1}(\hat{\Theta})\|_1$ approaches 2 as $\hat{\Theta}$ approaches the boundary and Lemma 3.19 (h). Hence (4.7) becomes

$$\begin{aligned}
\hat{\alpha} &= \alpha\left(-\frac{\hat{\rho}}{2} + \mathcal{O}(\Theta_1)\right)(1 + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)) \\
&\quad + \beta\left(-\left(1 - \frac{\hat{\rho}}{2}\right) - \frac{\hat{\rho}}{2} + \mathcal{O}(\Theta_1)\right)\mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1) + \mathcal{O}(\|\hat{R}\|_1) \\
&= \alpha\left(-\frac{\hat{\rho}}{2} + \mathcal{O}(\Theta_1)\right)(1 + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)) + \beta\left(-1 + \mathcal{O}(\Theta_1)\right)\mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1) \\
&\quad + \mathcal{O}(\|\hat{R}\|_1) = \alpha\left(-\frac{\hat{\rho}}{2} + \mathcal{O}(\alpha, \Theta_1, \|\hat{R}\|_1)\right) + \beta\mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1) + \mathcal{O}(\|\hat{R}\|_1) \\
&= -\frac{\hat{\rho}}{2}\alpha(1 + \mathcal{O}(\alpha, \Theta_1, \|\hat{R}\|_1)) + \mathcal{O}(\beta\alpha, \beta^2\Theta_1, \|\hat{R}\|_1).
\end{aligned}$$

Doing the same with the corresponding terms in (4.8) yields

$$\frac{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_0^{(i)}(\Theta) \\ e_0^{(j)}(\hat{\Theta}) & e_0^{(j)}(\Theta) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix}} = \frac{\begin{vmatrix} e_0^{(i)}(\Theta) & e_0^{(i)}(\Theta) \\ e_0^{(j)}(\Theta) & e_0^{(j)}(\Theta) \end{vmatrix} + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)}{\begin{vmatrix} e_0^{(i)}(\Theta) & e_{-1}^{(i)}(\Theta) \\ e_0^{(j)}(\Theta) & e_{-1}^{(j)}(\Theta) \end{vmatrix} + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)}$$

$$\frac{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\Theta) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\Theta) \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix}} = \frac{\begin{vmatrix} e_0^{(i)}(\Theta) & e_{-1}^{(i)}(\Theta) \\ e_0^{(j)}(\Theta) & e_{-1}^{(j)}(\Theta) \end{vmatrix} + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)}{\begin{vmatrix} e_0^{(i)}(\Theta) & e_{-1}^{(i)}(\Theta) \\ e_0^{(j)}(\Theta) & e_{-1}^{(j)}(\Theta) \end{vmatrix} + \mathcal{O}(\alpha, \beta\Theta_1, \|\hat{R}\|_1)}$$

$$\left| \frac{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & \hat{R}^{(i)} \\ e_0^{(j)}(\hat{\Theta}) & \hat{R}^{(j)} \end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix}} \right| \leq \frac{\|e_0(\hat{\Theta})\|_1 \|\hat{R}\|_1}{\left\| \begin{vmatrix} e_0^{(i)}(\hat{\Theta}) & e_{-1}^{(i)}(\hat{\Theta}) \\ e_0^{(j)}(\hat{\Theta}) & e_{-1}^{(j)}(\hat{\Theta}) \end{vmatrix} \right\|} = \mathcal{O}(\|\hat{R}\|_1),$$

hence equation (4.8) becomes

$$\begin{aligned}
\hat{\beta} &= \alpha\left(-\frac{\hat{\rho}}{2}+\mathcal{O}(\Theta_1)\right)\mathcal{O}(\alpha,\beta\Theta_1,\|\hat{R}\|_1)\\
&+\beta\left[\left(1-\frac{\hat{\rho}}{2}\right)\left(-1+2\Theta_1+\mathcal{O}(\Theta_1^2)\right)-\frac{\hat{\rho}}{2}\right](1+\mathcal{O}(\alpha,\beta\Theta_1,\|\hat{R}\|_1))+\mathcal{O}(\|\hat{R}\|_1)\\
&=\mathcal{O}(\alpha^2,\alpha\beta\Theta_1,\alpha\|\hat{R}\|_1)+\beta\left[-1+(2-\hat{\rho})\,\Theta_1+\mathcal{O}(\Theta_1^2)+\mathcal{O}(\alpha,\beta\Theta_1,\|\hat{R}\|_1)\right]\\
&+\mathcal{O}(\|\hat{R}\|_1)=\mathcal{O}(\alpha^2,\alpha\beta\Theta_1,\alpha\|\hat{R}\|_1)-\beta\left[1-(2-\hat{\rho})\,\Theta_1\right]\\
&+\mathcal{O}(\alpha\beta,\beta^2\Theta_1,\beta\Theta_1^2,\beta\|\hat{R}\|_1)+\mathcal{O}(\|\hat{R}\|_1)=-\beta\left[1-(2-\hat{\rho})\,\Theta_1\right]\\
&+\mathcal{O}(\alpha^2,\alpha\beta,\beta^2\Theta_1,\beta\Theta_1^2,\|\hat{R}\|_1),
\end{aligned}$$

having again used the determinant in the denominators being bounded away from zero. □

The constant $1<\nu<\sqrt{\mu}$ has yet to be fixed in its range of validity: its choice will depend on $\mu>1$, entailing some additional technicalities. Let

$$\nu=\nu(\mu):=\begin{cases}\frac{7}{5}, & \mu\geq 2\\ \mu^{\frac{1}{2}-\vartheta}, & 1<\mu<2,\end{cases}$$

then $\vartheta=\vartheta(\mu)$ can be determined such that $0<\vartheta<1/2$ and such that the conditions listed in the next lemma hold. To give some intuition, note that if $\mu\geq 2$ the choice of ν is consistent with the fact that, in Lemmas 4.16 and 4.17 and Theorem 4.18, for $\mu\geq 2$ it will be required that ν is greater than $4/3$. This is the easiest case, as it is possible to identify a value, $4/3$, with respect to which all the estimates, which we will derive, hold uniformly for all $\mu\geq 2$. Now, $7/5$ satisfies both being greater than $4/3$ (as $21>20$) and smaller than $\sqrt{2}$ (since $50>49$ is equivalent to $5\sqrt{2}>7$). Since $\sqrt{\mu}\geq\sqrt{2}$ in this case, $1<4/3<7/5<\sqrt{2}\leq\sqrt{\mu}$, $\nu=7/5$ is a consistent choice. As to the case $1<\mu<2$, the choice will be less explicit, as it depends on μ through the function $\vartheta(\mu)$, which, in Chapter B, we determine *constructively* to be possible to be fixed as the constant $1/12$, on the grounds of slightly more relaxed conditions than those seen for $\mu\geq 2$, which are still sufficient for the arguments of Lemmas 4.16 and 4.17 and Theorem 4.18 nonetheless. In proving the following lemma we avoid the constructive approach, for the sake of brevity, since it is possible to verify the claim through mere computations. These are beneficial for the reader, as they allow to familiarise with the quantities involved, which at this stage might feel somehow detached from the context. However they will be the key to the arguments aforementioned.

Lemma 4.13. *Let* $\eta=\eta(\mu):=\min\{\rho,1-\rho\}$ *and*

$$\nu=\nu(\mu):=\begin{cases}\frac{7}{5}, & \mu\geq 2\\ \mu^{\frac{5}{12}}, & 1<\mu<2.\end{cases}$$

Let $0<\theta<1/2$ *be*

$$\theta:=\begin{cases}\frac{1}{16}, & \mu\geq 2\\ \frac{2+\frac{1}{\sqrt[3]{\mu}}}{3-\frac{1}{\mu}}-1 & 1<\mu<2.\end{cases}$$

Let

$$a=a(\theta,\mu):=\frac{\rho+\eta+(4+\rho+\eta)\theta}{2}$$

and

$$b = b(\theta,\mu) := \frac{a(\theta,\mu)}{d(\theta,\mu)},$$

where

$$d = d(\theta,\mu) := \left[1 - \left(1 - \frac{1}{\sqrt[3]{\mu}}\right)\theta\right]\left[1 - 2\frac{2-\rho+\eta+2\theta}{3(1+\sqrt[3]{\mu})}(\sqrt[3]{\mu}-1)\theta\right].$$

Then for all $\mu > 1$, $0 < a < b < 1$ and $d > a > \frac{1}{\nu}$.

Proof. Note that

$$\eta = \eta(\mu) := \min\{\rho, 1-\rho\} = \begin{cases} 1-\rho, & \mu \geq 2 \\ \rho, & 1 < \mu < 2, \end{cases}$$

since for all $\mu > 1$, $\mu \geq 2$ is equivalent to $1 - \rho \leq \rho$, as $\rho = (\mu - 1)/\mu$, thus implying also that $\mu < 2$ is equivalent to $\rho < 1 - \rho$. Note that $0 < \theta < 1/2$: this is trivial for $\mu \geq 2$, while for $1 < \mu < 2$ it follows from

$$\theta < \frac{2+1}{3-1} - 1 = \frac{1}{2}$$

and

$$\theta > \frac{2 + \frac{1}{\sqrt[3]{2}}}{3 - \frac{1}{2}} - 1 = \frac{4}{5} + \frac{2}{5\sqrt[3]{2}} - 1 > \frac{4}{5} + \frac{1}{5} - 1 = 0.$$

This implies the positivity of the two factors appearing in d.

- Since $\mu > 1$, $\left(1 - 1/\sqrt[3]{\mu}\right)\theta < \theta < 1/2$, thus the first factor is positive.
- Since

  $$2 - \rho + \eta = 2 - \rho + \min\{\rho, 1-\rho\} \leq 2,$$

  we have that

  $$2\frac{2-\rho+\eta+2\theta}{3(1+\sqrt[3]{\mu})}(\sqrt[3]{\mu}-1)\theta < \frac{2}{3}2\theta(1+\theta)\frac{\sqrt[3]{\mu}-1}{1+\sqrt[3]{\mu}} < \frac{\sqrt[3]{\mu}-1}{1+\sqrt[3]{\mu}} < 1$$

  due to the parabola $2x(1+x)$ being increasing in $(0, 1/2)$ and valued 0 at 0 and $3/2$ at $1/2$; thus the second factor is positive too.

The constant $\theta = \theta(\mu)$ has been constructed along with $\vartheta(\mu)$, such that, for all $\mu \geq 2$, $0 < a < 3/4$, $0 < b < 1$ and $3/4 > 1/\nu$; while for all $1 < \mu < 2$ it only satisfies the less restrictive conditions $0 < a < b < 1$ and $a > 1/\nu$. That this is true for all $\mu \geq 2$, is easily seen. In this case, if $\theta < (13 - \sqrt{145})/12 < 1/12$,

$$a \leq \frac{1+5\theta}{2} < \frac{3}{4}$$

and, since $1 - 1/\sqrt[3]{\mu} < 1$ and since we have just verified that

$$2\frac{2-\rho+\eta+2\theta}{3(1+\sqrt[3]{\mu})}(\sqrt[3]{\mu}-1)\theta < \frac{4}{3}\theta(1+\theta) < \frac{8}{3}\theta < 3\theta,$$

$d > (1-\theta)(1-3\theta)$, thus implying

$$b \leq \frac{1+5\theta}{2(1-\theta)(1-3\theta)} < 1.$$

The estimate on a follows because at $1/10 > (13-\sqrt{145})/12$, it holds that $(1+5x)/2 = 3/4$ and the slope of the line on the left-hand side of this equation is positive. The estimate on b follows since the function $(1+5x)/(2(1-x)(1-3x)) = 1$ at $(13-\sqrt{145})/12 > 1/13$ (which is verified by rearranging and squaring both sides), and on $(0, 1/3)$, the function is monotone increasing due to its derivative being $[3(3-2x-5x^2)]/[2(1-x)^2(1-3x)^2]$ and the concave quadratic $3-2x-5x^2$ having roots at -1 and $3/5$. Thus $\theta = 1/16$ satisfies the requirements. The value of θ chosen draws also motivation from another requirement that will be crucial in Lemma 4.17: it ensures that $d > (1-\theta)(1-3\theta) > 3/4$, because the convex parabola $(1-x)(1-3x)$ is subunitary, decreasing for $0 < x < 2/3$ (note that the vertex of the parabola is exactly at $2/3$) and hits $3/4$ at $(4-\sqrt{13})/6 > 1/16$ (seen by rearranging and squaring both sides).

We now show that all the conditions are met also for $1 < \mu < 2$. First of all we show that $a > 1/\nu$. Since in this case $\eta(\mu) = \rho = 1 - 1/\mu$, $2+\rho = 3 - 1/\mu$, by construction

$$a = \rho + (2+\rho)\left(\frac{2+\frac{1}{\sqrt[3]{\mu}}}{3-\frac{1}{\mu}} - 1\right) = -2 + \left(3 - \frac{1}{\mu}\right)\frac{2+\frac{1}{\sqrt[3]{\mu}}}{3-\frac{1}{\mu}} = \frac{1}{\sqrt[3]{\mu}} > \frac{1}{\mu^{\frac{5}{12}}} =: \frac{1}{\nu}.$$

Secondly, we show that $0 < a < b < 1$. Note, from the previous line, that $a = 1/\sqrt[3]{\mu}$; clearly $0 < a < 1$, since $1 < \mu < 2$; as to $0 < b < 1$, it is equivalent to $0 < a/d < 1$ and since $a = 1/\sqrt[3]{\mu}$, we can just show that $0 < 1/d < \sqrt[3]{\mu}$, in order to yield the claim. We already verified the positivity, while the upper bound is verified by showing equivalently that

$$\frac{3}{2}\frac{1+\sqrt[3]{\mu}}{\sqrt[3]{\mu}-1} > \frac{\theta\left(2-\rho+\eta+2\theta\right)}{1-\frac{1}{\sqrt[3]{\mu}\left[1-\left(1-\frac{1}{\sqrt[3]{\mu}}\right)\theta\right]}},$$

which holds, since the denominator of the right-hand side is positive. Recalling that $2-\rho+\eta = 2$ and that all fractions have positive numerators and denominators, due to $\theta < 1/2$ and $\mu > 1$,

$$\frac{\theta\left(2-\rho+\eta+2\theta\right)}{1-\frac{1}{\sqrt[3]{\mu}\left[1-\left(1-\frac{1}{\sqrt[3]{\mu}}\right)\theta\right]}} = \frac{2\theta(1+\theta)}{1-\frac{1}{\sqrt[3]{\mu}\left[1-\left(1-\frac{1}{\sqrt[3]{\mu}}\right)\theta\right]}} < \frac{3}{2}\frac{1}{1-\frac{1}{\sqrt[3]{\mu}\left[1-\frac{1}{2}\left(1-\frac{1}{\sqrt[3]{\mu}}\right)\right]}} = \frac{3}{2}\frac{1}{1-\frac{2}{\sqrt[3]{\mu}+1}}$$
$$= \frac{3}{2}\frac{1+\sqrt[3]{\mu}}{\sqrt[3]{\mu}-1}.$$

Note that $d > a$ holds automatically since it follows from $b = a/d < 1$. This will be crucial for Lemma 4.17. □

The index K will be required to be large enough, to ensure that $\rho-\eta < \rho_{n+1} < \rho+\eta$ for all $n \geq n_K$.

Lemma 4.14. *Let the positive constant*

$$\tilde{c} := \frac{2}{3}\left(\sqrt{2}-1\right).$$

There exists $c > 0$ and $M \geq 1$ such that for all sufficiently small ε, on the closure

$\overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$ *it holds that*

$$\begin{aligned}
|\rho_1(r)| &< \left(1-\frac{1}{\sqrt[3]{\mu}}\right)\theta\Theta_1 + M\|\hat{R}\|_1\\
|\rho_2(r)| &< c|\alpha| + \Theta_1|\beta| + M\|\hat{R}\|_1\\
|\rho_3(r)| &< \tilde{c}\theta + M\|\hat{R}\|_1\\
|\rho_4(r)| &< \tilde{c}\theta(|\alpha| + \Theta_1|\beta|) + M\|\hat{R}\|_1\\
|\rho_5(r)| &< \frac{\theta}{2\nu}(|\alpha| + \Theta_1|\beta|) + M\|\hat{R}\|_1.
\end{aligned}$$

Proof. Let $\delta > 0$ fixed small enough, and $\varepsilon > 0$ arbitrarily small (possibly dependent on δ and to be further reduced) so as to ensure well-definedness of all quantities involved. We will work exclusively on $\overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$, so denote $\overline{\mathscr{K}}^* := \overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$.

Starting with ρ_1, by Lemma 4.11 it holds that on $\overline{\mathscr{K}}^*$ there is some $c_1 > 0$, such that $|\rho_1(r)| \leq c_1\Theta_1(|\alpha| + |\beta|) + c_1\|\hat{R}\|_1$. If now we further restrict

$$\varepsilon < \left(1-\frac{1}{\sqrt[3]{\mu}}\right)\frac{\theta}{2c_1},$$

it follows that

$$c_1\Theta_1(|\alpha| + |\beta|) + c_1\|\hat{R}\|_1 < \left(1-\frac{1}{\sqrt[3]{\mu}}\right)\theta\Theta_1 + c_1\|\hat{R}\|_1,$$

since

$$c_1(|\alpha| + |\beta|) \leq 2\varepsilon c_1 < \left(1-\frac{1}{\sqrt[3]{\mu}}\right)\theta,$$

yielding the desired estimate. Further restricting ε is consistent with c_1, as the same constant upper bound applies on the new $\overline{\mathscr{K}}^*$, as it shrinks, being δ fixed, as explained in Lemma 3.23 (this mention will be implicit in subsequent steps).

Moving on to ρ_2, by Lemma 4.11 it holds that on $\overline{\mathscr{K}}^*$ there is some $c_2 > 0$, such that $|\rho_2(r)| \leq c_2(|\alpha| + |\beta|\Theta_1^2) + c_2\|\hat{R}\|_1$. Let $c := c_2$ and further restrict $\varepsilon < {}^1\!/_c$. Then $c_2(|\alpha| + |\beta|\Theta_1^2) + c_2\|\hat{R}\|_1 < c(|\alpha| + \|\hat{R}\|_1) + |\beta|\Theta_1$, since $c\Theta_1 \leq c\varepsilon < 1$.

As to $\rho_3(r)$, by Lemma 4.12 it holds that on $\overline{\mathscr{K}}^*$ there is some $c_3 > 0$, such that $|\rho_3(r)| \leq c_3(|\alpha| + \Theta_1) + c_3\|\hat{R}\|_1$. We add the restriction $\varepsilon < \tilde{c}\theta/(2c_3)$, then it follows that $c_3(|\alpha| + \Theta_1) \leq 2c_3\varepsilon < \tilde{c}\theta$, yielding $|\rho_3(r)| \leq \tilde{c}\theta + c_3\|\hat{R}\|_1$.

For $\rho_4(r)$, by Lemma 4.12 it holds that on $\overline{\mathscr{K}}^*$ there is some $c_4 > 0$, such that $|\rho_4(r)| \leq c_4(|\beta||\alpha| + \beta^2\Theta_1) + c_4\|\hat{R}\|_1 = c_4|\beta|(|\alpha| + |\beta|\Theta_1) + c_4\|\hat{R}\|_1$. Further restricting $\varepsilon < \tilde{c}\theta/c_4$, it follows that $c_4|\beta| \leq c_4\varepsilon < \tilde{c}\theta$, yielding $|\rho_4(r)| < \tilde{c}\theta(|\alpha| + |\beta|\Theta_1) + c_4\|\hat{R}\|_1$.

Lastly $\rho_5(r)$. By Lemma 4.12 it holds that on $\overline{\mathscr{K}}^*$ there is some $c_5 > 0$, such that $|\rho_5(r)| \leq c_5(\alpha^2 + |\beta||\alpha| + \beta^2\Theta_1 + |\beta|\Theta_1^2) + c_5\|\hat{R}\|_1 = c_5(|\alpha| + |\beta|)|\alpha| + c_5(|\beta| + \Theta_1)\Theta_1|\beta| + c_5\|\hat{R}\|_1$. Further restricting $\varepsilon < \theta/(4c_5\nu)$, it follows that $c_5(|\alpha| + |\beta|) \leq 2c_5\varepsilon < \theta/(2\nu)$ and $c_5(\Theta_1 + |\beta|) \leq 2c_5\varepsilon < \theta/(2\nu)$, yielding

$$|\rho_5(r)| < \frac{\theta}{2\nu}(|\alpha| + |\beta|\Theta_1) + c_5\|\hat{R}\|_1.$$

All in all, from a given initial ε defining constants c_1, c, c_3, c_4, c_5, we further restrict it so that

$$\varepsilon \leq \min\left\{\left(1-\frac{1}{\sqrt[3]{\mu}}\right)\frac{\theta}{2c_1}, \frac{1}{c}, \frac{\tilde{c}\theta}{2c_3}, \frac{\tilde{c}\theta}{c_4}, \frac{\theta}{4c_5\nu}\right\},$$

and denote $M := \max\{1,\ c_1,\ c,\ c_3,\ c_4,\ c_5\}$. Then all previous five estimates will hold on the newly constructed $\overline{\mathscr{K}}^*$. □

It will be necessary, for further arguments, to add restrictions on ε, given δ, ν, c, $\tilde{c}$, c_i, M, θ. To sum up all those added so far and the ones that will be needed from this point on:

$$\varepsilon < \min\left\{\frac{1}{18}, \frac{2\theta}{3}\frac{\sqrt[3]{\mu}-1}{1+\sqrt[3]{\mu}}, \left(1-\frac{1}{\sqrt[3]{\mu}}\right)\frac{\theta}{2c_1}, \frac{1}{c}, \frac{\tilde{c}\theta}{2c_3}, \frac{\tilde{c}\theta}{c_4}, \frac{\theta}{2c_5}, \frac{\delta}{A}, \frac{\delta}{B}\right\},$$

where

$$A := 4\left[(3+c)\left(\frac{1}{1-a}+\frac{d\nu}{(1-d)(d\nu-1)^2}\right)+\frac{1}{\nu-1}\left(1+\frac{3c\nu}{\nu-1}\right)\right],$$
$$B := 8\left[(3+c)\left(\iota\frac{2+\frac{\nu}{(\nu-1)^2}}{1-\theta}+4\frac{\nu}{(\nu-1)^2}\right)+\frac{1}{\nu-1}\left(1+\frac{3\nu}{\nu-1}\right)\right]$$

and $\iota := \bar{k}+1 \in \mathbb{N}$, where

$$\bar{k} := \max\left\{k \in \mathbb{N}: \ (a\nu)^k < \frac{6}{\theta\nu}(k+1)\right\},$$

and is finite due to $a\nu > 1$ by Lemma 4.13. This definition ensures that for all $k \geq \iota$

$$(a\nu)^k \geq \frac{6(k+1)}{\theta\nu}.$$

Finally, we will add the implicit condition, given the fixed $\delta = \delta(\omega)$, that ε be small enough to allow $\overline{\mathscr{K}}_{2\varepsilon,\frac{\delta}{8}}$ to not intersect E_2 and E_3, and every point in $\overline{\mathscr{K}}_{2\varepsilon,\frac{\delta}{8}}$ to be closer to E_1 than to E_2 and E_3. This construction of ε is consistent with all the constants already defined, since further reducing ε at any step necessary, keeps the new set within the one constructed out of the previous more relaxed ε, and therefore the constants keep holding uniformly.

Consider now, for any fixed $\omega \in \mathscr{E}_0^{(1)}$, the random times

$$\tau_k := \inf\left\{n > n_k: \ \Theta_n^{(2)} \notin \left[\frac{\delta}{8}, 1-\frac{\delta}{8}\right]\right\} \in \mathbb{N} \cup \infty.$$

We already saw that there is an arbitrarily large $K = K(\omega)$ such that for $m = n_k$, for any $k \geq K$, $\Theta_m^{(2)} \in [\delta/2, 1-\delta/2]$, $|\alpha_n|, |\beta_n| < \varepsilon$ for all $n \geq m$ and $\|R_{n+1}\|_1 < 3\nu^{-n-1} < \varepsilon$ for all $n \geq m$ (that is the subsequence of the orbit is in $\mathscr{K}^*_{\varepsilon,\frac{\delta}{2}}$). It is left to show that we can satisfy, by choosing a suitable k large enough, letting $m = n_k$, also that $\Theta_n^{(1)} < \varepsilon$ for all $m \leq n < \tau_k$. In order to do this we will put on K one more requirement: since on $\{\ell = 0\}$, $\|\Theta_{n+1} - \Theta_n\|_1 \longrightarrow 0$, we can choose $K = K(\omega)$ large enough, such that for all $k \geq K$,

$$|\Theta_{n_k}^{(2)} - \Theta_*^{(2)}| < \frac{1}{2}\min\left\{\Theta_*^{(2)} - \frac{\delta}{8}, 1-\frac{\delta}{8}-\Theta_*^{(2)}\right\}$$

and for any $n \geq n_K$,

$$\|\Theta_{n+1} - \Theta_n\|_1 < \frac{1}{2}\min\left\{\Theta_*^{(2)} - \frac{\delta}{8}, 1-\frac{\delta}{8}-\Theta_*^{(2)}\right\}.$$

This assumption ensures that for all $k \geq K$, $\tau_{n_k} > n_k + 1$ (so that there is always some $n_k < n < \tau_k$), since

$$\Theta_{n_k+1}^{(2)} \leq |\Theta_{n_k+1}^{(2)} - \Theta_{n_k}^{(2)}| + |\Theta_{n_k}^{(2)} - \Theta_*^{(2)}| + \Theta_*^{(2)} < \min\left\{\Theta_*^{(2)} - \frac{\delta}{8}, 1-\frac{\delta}{8}-\Theta_*^{(2)}\right\} + \Theta_*^{(2)} \leq 1-\frac{\delta}{8}$$

and

$$\Theta^{(2)}_{n_k+1} \geq \Theta^{(2)}_* - |\Theta^{(2)}_{n_k+1} - \Theta^{(2)}_{n_k} + \Theta^{(2)}_{n_k} - \Theta^{(2)}_*| \geq \Theta^{(2)}_* - \left(|\Theta^{(2)}_{n_k+1} - \Theta^{(2)}_{n_k}| + |\Theta^{(2)}_{n_k} - \Theta^{(2)}_*|\right)$$
$$> \Theta^{(2)}_* - \min\left\{\Theta^{(2)}_* - \frac{\delta}{8}, 1 - \frac{\delta}{8} - \Theta^{(2)}_*\right\} \geq \frac{\delta}{8}.$$

Additionally, since with all parameters set as such, the subsequence of the orbit stays in the set $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$, the constant M always holds, and therefore we can take the index K large enough, such that

$$3\frac{M}{\nu^{n_K}} < \varepsilon.$$

Lemma 4.15. *For almost every $\omega \in \mathscr{E}^{(1)}_0$, there exists $k = k(\omega) \geq K$ large enough such that, letting $m = m(\omega) := n_k$, for all $m \leq n < \tau_k$, $\Theta^{(1)}_n \leq \varepsilon$.*

Proof. For $n = n_k$ it is trivial. For $n > n_k$ one needs to distinguish between two cases and proceed by contradiction. The core of the argument is the same for both events, only the preparation slightly differs.

If $\omega \in \mathscr{E}^{(1)}_0$ is such that $\tau_k \in \mathbb{N}$ for all $k \geq K$, suppose by contradiction that there is a subsequence $\{k_r\}_{r\in\mathbb{N}}$ (with $k_r \geq K$) such that for all r, for some $n_{k_r} < n < \tau_{k_r}$, $\Theta^{(1)}_n > \varepsilon$. This implies that there is a subsequence $\{\Theta_{n_r}\}$ for which $\Theta^{(1)}_{n_r} > \varepsilon$ and $\Theta^{(2)}_{n_r} \in [\delta/8, 1-\delta/8]$. From this subsequence, for almost every such ω, a subsubsequence $\{\Theta_{n_{r_s}}\}_{s\in\mathbb{N}}$ can be extracted - denote it $\{\Theta_{n_l}\}$ for simplicity - such that $\Theta^{(1)}_{n_l-1} \leq \varepsilon$ and $\varepsilon < \Theta^{(1)}_{n_l} < 2\varepsilon$. This is true because for every ω considered, at least $\Theta^{(1)}_{n_{k_{r_s}}} \leq \varepsilon$, so the sequence exits $(0, \varepsilon]$ after having been inside the interval at least one time, and one can choose n_l as the first time of exit from $(0, \varepsilon]$. Furthermore, for every ω considered the potential vanishes, thus for almost every ω considered $\|\Theta_n - \Theta_{n-1}\|_1 \longrightarrow 0$ by (4.4) and Lemma 4.1. Thus for all l large enough, $|\Theta^{(1)}_{n_l} - \Theta^{(1)}_{n_l-1}| \leq \|\Theta_{n_l} - \Theta_{n_l-1}\|_1 < \varepsilon$, and therefore for almost every ω considered,

$$\varepsilon < \Theta^{(1)}_{n_l} \leq |\Theta^{(1)}_{n_l} - \Theta^{(1)}_{n_l-1}| + \Theta^{(1)}_{n_l-1} < 2\varepsilon.$$

At the same time, by construction $\delta/8 < \Theta^{(2)}_{n_l} < 1-\delta/8$, since $n_{k_{r_s}} < n_l < \tau_{k_{r_s}}$. Consider now the set $\{(\Theta, \pi_\Theta) :\ \Theta \in \overline{R}\}$ where $R = R_{\varepsilon,\frac{\delta}{8}} := \mathscr{K}_{2\varepsilon,\frac{\delta}{8}} \setminus \mathscr{K}_{\varepsilon,\frac{\delta}{8}}$. By construction of ε, every $\Theta \in \overline{R}$ does not lie on $\partial\Sigma$ and it is closer to E_1 than to the other two edges. Note that $\Theta_{n_l} \in \mathring{R}$. For every Θ, fixing a small enough $\varepsilon' < \varepsilon = \mathrm{dist}(\overline{R}, E_1)/2$ (since we are working as usual in 1-norm), letting c the positive subunitary constant such that $\|L_\Theta(\pi - \pi_\Theta)\|_1 \leq c\|\pi - \pi_\Theta\|_1$ for all $\Theta \in \overline{R}$ (following the same construction as in Lemma 4.5, it is known that this constant holds uniformly on the whole compact $\Sigma_{\varepsilon'}$, so in particular it is uniform on $\overline{R} \subset \Sigma_{\varepsilon'}$); one can set, by recalling that $\lambda := \max\{c, \nu^{-1}\}$,

$$\delta' := \frac{\varepsilon'}{\frac{3}{1-\lambda^2} + \frac{\nu}{\nu-1}}.$$

Then since $\|\pi_n - \pi_{\Theta_n}\|_1 \longrightarrow 0$, for some large enough $\bar{l}$, $(\Theta_{n_{\bar{l}}}, \pi_{n_{\bar{l}}}) \in U((\Theta_{n_{\bar{l}}}, \pi_{\Theta_{n_{\bar{l}}}}), \delta')$, hence by Proposition 4.8 it is known that for almost every ω considered, for all $n \geq n_{\bar{l}}$, $\Theta_n \in B(\Theta_{n_{\bar{l}}}, \varepsilon')$. Indeed the proposition applies since by construction $\varepsilon' < \mathrm{dist}(\overline{R}, \partial\Sigma)$ and therefore for every $\Theta \in \overline{R}$, $\varepsilon' < \mathrm{dist}(\Theta, \partial\Sigma)$. But then eventually $\mathrm{dist}(\Theta_n, E_1) \geq \mathrm{dist}(B(\Theta_{n_{\bar{l}}}, \varepsilon'), E_1) \geq \mathrm{dist}(\overline{R}, E_1) - \varepsilon' > 2\varepsilon - \varepsilon = \varepsilon$, which is in contradiction, for almost every ω considered, with $\Theta^{(1)}_{n_k} \longrightarrow 0$.

If $\omega \in \mathscr{E}^{(1)}_0$ is such that for some $\bar{k} \geq K$, $\tau_{\bar{k}} = \infty$, then for all $k \geq \bar{k}$, $\tau_k = \infty$. Suppose again, by the same construction as above, that, by contradiction, there is a subsequence $\{\Theta_{n_r}\}$, for which $\Theta^{(1)}_{n_r} > \varepsilon$, $n_r > n_{k_r}$ for all $r \in \mathbb{N}$, $k_r \geq \bar{k}$. Note that this time it automatically holds that $\Theta^{(2)}_{n_r} \in [\delta/8, 1 - \delta/8]$ (since all $\tau_{k_r} = \infty$ for all $k_r \geq \bar{k}$, $\Theta^{(2)}_n \in [\delta/8, 1 - \delta/8]$ for all $n \geq n_{\bar{k}}$, in particular the last condition imposed on K is not necessary for this case). This case does not require any degree of control on the $\Theta^{(2)}$-component, and we can just proceed as in the previous case. Another subsubsequence $\{\Theta_{n_l}\}_{l\in\mathbb{N}}$ can be extracted such that $\Theta^{(1)}_{n_l-1} \leq \varepsilon$ and $\varepsilon < \Theta^{(1)}_{n_l} < 2\varepsilon\}$ along with $\delta/8 < \Theta^{(2)}_{n_l} < 1 - \delta/8$. It is clear enough that the previous argument can be repeated *verbatim*. □

In the following all the proofs are made with respect to the large enough $m = n_k$, with $k \geq K$, existing by Lemma 4.15, and therefore the corresponding τ_k will be simply denoted as τ.

Lemma 4.16. *For almost every $\omega \in \mathscr{E}^{(1)}_0$, for all $m \leq n \leq \tau$,*

$$|\alpha_n| \leq \max\left\{a^{n-m}\varepsilon, \Theta^{(1)}_n|\beta_n|\right\} + 9M\frac{n-m}{\nu^n}.$$

Proof. Proceed by induction. If $n = m$, the statement $|\alpha_m| \leq \max\{\varepsilon, \Theta^{(1)}_m|\beta_m|\} = \varepsilon$ is trivially true for all $\mu > 1$ by the definition of the time m and of $\mathscr{K}^*_{\varepsilon,\frac{\delta}{2}}$.

If $n = m+1 < \tau$, recall that for almost every $\omega \in \mathscr{E}^{(1)}_0$, being $(\Theta_m, \pi_m, R_{m+1}) \in \overline{\mathscr{K}}^* := \overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$, by (4.12) and Lemma 4.14 it holds that

$$|\alpha_{m+1}| \leq \rho_{m+1}\frac{|\alpha_m|}{2}(1 + |\rho_3(r_m)|) + |\rho_4(r_m)| \leq \frac{\rho+\eta}{2}|\alpha_m|\left(1 + \theta + \frac{3M}{\nu^{m+1}}\right) + \theta|\alpha_m| + \theta\Theta^{(1)}_m|\beta_m| + \frac{3M}{\nu^{m+1}},$$

which applies by definition of m. Then from $3M\nu^{-m} < \varepsilon$ and $|\alpha_m| \leq \varepsilon$ it follows that

$$|\alpha_{m+1}| \leq \varepsilon\left(\frac{\rho+\eta+(2+\rho+\eta)\theta}{2}\right) + \theta\Theta^{(1)}_m|\beta_m| + \left(1+\frac{\varepsilon}{2}\right)\frac{3M}{\nu^{m+1}}.$$

If $\varepsilon \geq \Theta^{(1)}_m|\beta_m|$, then

$$|\alpha_{m+1}| \leq \varepsilon\left(\frac{\rho+\eta+(2+\rho+\eta)\theta}{2}\right) + \theta\varepsilon + \left(1+\frac{\varepsilon}{2}\right)\frac{3M}{\nu^{m+1}} < a\varepsilon + \frac{9M}{\nu^{m+1}},$$

since $1 + \varepsilon/2 < 3$. In principle, one should now consider the scenario when instead $\varepsilon < \Theta^{(1)}_m|\beta_m|$ and proceed with showing the claim again, however, due to the definition of the time m and of the set $\mathscr{K}^*$, we know that $\Theta^{(1)}_m|\beta_m| \leq \varepsilon^2 < \varepsilon$. So, and this holds specifically for the times m and $m+1$, there is no need to do so. It will be necessary, however, when performing the induction step past the $(m+1)$st time.

Assume now the hypothesis for any $m+1 \leq n < \tau$. This time in carrying out the inductive step, it will not be possible to appeal only to the definition of m, but it will be necessary to rely on Lemma 4.15 as well, which ensures that $(\Theta_n, \pi_n, R_{n+1}) \in \overline{\mathscr{K}}^*$ for almost every $\omega \in \mathscr{E}^{(1)}_0$, and therefore ensures that the same lemmas aforementioned apply in the corresponding steps of the remaining part of the argument. The constants involved in the error terms will require different estimates, depending on whether $\mu \geq 2$ or $1 < \mu < 2$. Recall that in Lemma 4.14 we denoted

$$\tilde{c} := \frac{2}{3}\left(\sqrt{2} - 1\right) < 1.$$

If $\mu \geq 2$, for almost every $\omega \in \mathscr{E}_0^{(1)}$, we have that

$$|\alpha_{n+1}| \leq \frac{\rho+\eta}{2}|\alpha_n| \left(1 + \tilde{c}\theta + \frac{3M}{\nu^{n+1}}\right) + \tilde{c}\theta|\alpha_n| + \tilde{c}\theta\Theta_n^{(1)}|\beta_n| + \frac{3M}{\nu^{n+1}}$$
$$< |\alpha_n|\frac{\rho+\eta+(2+\rho+\eta)\theta}{2} + \theta\Theta_n^{(1)}|\beta_n| + \left(1+\frac{\varepsilon}{2}\right)\frac{3M}{\nu^{n+1}}$$

by (4.12) and Lemma 4.14. By induction hypothesis, for almost every $\omega \in \mathscr{E}_0^{(1)}$, it follows that

$$|\alpha_n| < \max\left\{a^{n-m}\varepsilon, \Theta_n^{(1)}|\beta_n|\right\} + 9M\frac{n-m}{\nu^n}.$$

If ω is such that $a^{n-m}\varepsilon \geq \Theta_n^{(1)}|\beta_n|$,

$$|\alpha_n| < a^{n-m}\varepsilon + 9M\frac{n-m}{\nu^n},$$

and then

$$|\alpha_{n+1}| < a^{n-m}\frac{\rho+\eta+(2+\rho+\eta)\theta}{2}\varepsilon + 9M\frac{1+3\theta}{2}\frac{n-m}{\nu^n} + a^{n-m}\theta\varepsilon$$
$$+ \left(1+\frac{\varepsilon}{2}\right)\frac{3M}{\nu^{n+1}} = a^{n-m}\frac{\rho+\eta+(4+\rho+\eta)\theta}{2}\varepsilon + \frac{1+3\theta}{2}\frac{9M(n-m)}{\nu^n}$$
$$+ \left(1+\frac{\varepsilon}{2}\right)\frac{3M}{\nu^{n+1}} \leq a^{n+1-m}\varepsilon + \frac{9M(n-m)}{\nu^{n+1}} + \left(1+\frac{\varepsilon}{2}\right)\frac{3M}{\nu^{n+1}}$$
$$< a^{n+1-m}\varepsilon + \frac{9M(n-m+1)}{\nu^{n+1}},$$

since by construction $(1+3\theta)/2 < 5/7 = 1/\nu$ (as $(1+3/16)/2 = 19/32 < 20/32 = 5/8$) and $1+\varepsilon/2 < 3$. If instead ω is such that $a^{n-m}\varepsilon < \Theta_n^{(1)}|\beta_n|$, then the induction hypothesis becomes

$$|\alpha_n| < \Theta_n^{(1)}|\beta_n| + 9M\frac{n-m}{\nu^n},$$

and therefore

$$|\alpha_{n+1}| \leq \frac{\rho+\eta}{2}\Theta_n^{(1)}|\beta_n|\left(1+\tilde{c}\theta+\frac{3M}{\nu^{n+1}}\right) + 9M\frac{\rho+\eta}{2}\left(1+\tilde{c}\theta+\frac{3M}{\nu^{n+1}}\right)\frac{n-m}{\nu^n}$$
$$+ 2\tilde{c}\theta\Theta_n^{(1)}|\beta_n| + \tilde{c}\theta 9M\frac{n-m}{\nu^n} + \frac{3M}{\nu^{n+1}} \leq \frac{\rho+\eta}{2}\Theta_n^{(1)}|\beta_n|(1+\theta) + 2\theta\Theta_n^{(1)}|\beta_n|$$
$$+ \frac{\varepsilon^2}{2}\frac{3M}{\nu^{n+1}} + \frac{9\nu}{2}\left(1+3\tilde{c}\theta+\frac{\varepsilon}{\nu}\right)M\frac{n-m}{\nu^{n+1}} + \frac{3M}{\nu^{n+1}} \leq a\Theta_n^{(1)}|\beta_n|$$
$$+ \frac{27}{4}(1+2\theta)M\frac{n-m}{\nu^{n+1}} + \left(1+\frac{\varepsilon^2}{2}\right)\frac{3M}{\nu^{n+1}}.$$

The last inequality follows since $\varepsilon < 1/18 < \theta$ and $3\tilde{c} < 1$ implies

$$1 + 3\tilde{c}\theta + \frac{\varepsilon}{\nu} < 1 + \theta + \frac{5}{9}\theta = 1 + \frac{12}{7}\theta < 1 + 2\theta,$$

and $9\nu/2 = 63/10 < 6.75 = 27/4$. Thus $\theta = 1/16$ yields

$$\frac{27}{4}(1+2\theta) < 8,$$

implying

$$|\alpha_{n+1}| \leq a\Theta_m^{(1)}|\beta_m| + \frac{8M(n-m)}{\nu^{n+1}} + \left(1+\frac{\varepsilon^2}{2}\right)\frac{3M}{\nu^{n+1}}. \tag{4.14}$$

By (4.9) and Lemma 4.14, for almost every $\omega \in \mathscr{E}_0^{(1)}$,

$$\Theta_n^{(1)} \le \frac{\Theta_{n+1}^{(1)} + \frac{3M}{\nu^{n+1}}}{1 - \left(1 - \frac{1}{\sqrt[3]{\mu}}\right)\theta}. \tag{4.15}$$

Note also that for almost every such ω, the induction hypothesis implies that $|\alpha_n| \le \varepsilon|\beta_n| + 9M(n-m)/\nu^n$ by Lemma 4.15. This yields, by applying (4.13) and Lemma 4.14, that

$$\begin{aligned}
|\beta_{n+1}| &\ge |\beta_n|\,(1-(2-\rho+\eta)\Theta_n^{(1)}) - |\rho_5(r_n)| > |\beta_n|\,(1-(2-\rho+\eta)\varepsilon) \\
&\quad - \frac{\theta}{2\nu}|\alpha_n| - \frac{\theta}{2\nu}\Theta_n^{(1)}|\beta_n| - \frac{3M}{\nu^{n+1}} > |\beta_n|\,(1-(2-\rho+\eta)\varepsilon) - 2\varepsilon\theta|\beta_n| \\
&\quad - \frac{\theta}{2}\frac{9M(n-m)}{\nu^{n+1}} - \frac{3M}{\nu^{n+1}} \ge |\beta_n|\left[1 - 2\frac{2-\rho+\eta+2\theta}{3(1+\sqrt[3]{\mu})}(\sqrt[3]{\mu}-1)\theta\right] - \frac{\theta}{2}\frac{9M(n-m)}{\nu^{n+1}} \\
&\quad - \frac{3M}{\nu^{n+1}},
\end{aligned}$$

hence

$$|\beta_n| < \frac{|\beta_{n+1}| + \frac{\theta}{2}\frac{9M(n-m)}{\nu^{n+1}} + \frac{3M}{\nu^{n+1}}}{1 - 2\frac{2-\rho+\eta+2\theta}{3(1+\sqrt[3]{\mu})}(\sqrt[3]{\mu}-1)\theta}. \tag{4.16}$$

Thus plugging the bounds in (4.15) and (4.16) into (4.14) yields that

$$\begin{aligned}
\alpha_{n+1} &\le b\left(\Theta_{n+1}^{(1)} + \frac{3M}{\nu^{n+1}}\right)\left(|\beta_{n+1}| + \frac{9M(n-m)}{\nu^{n+1}} + \frac{3M}{\nu^{n+1}}\right) + \frac{8M(n-m)}{\nu^{n+1}} \\
&\quad + \left(1 + \frac{\varepsilon^2}{2}\right)\frac{3M}{\nu^{n+1}} < \Theta_{n+1}^{(1)}|\beta_{n+1}| + \left(8 + 9\Theta_{n+1}^{(1)} + 27\frac{M}{\nu^{n+1}}\right)\frac{M(n-m)}{\nu^{n+1}} \\
&\quad + \left(1 + \frac{\varepsilon^2}{2} + \Theta_{n+1}^{(1)} + |\beta_{n+1}| + \frac{3M}{\nu^{n+1}}\right)\frac{3M}{\nu^{n+1}} < \Theta_{n+1}^{(1)}|\beta_{n+1}| + \frac{9M(n-m)}{\nu^{n+1}} + \frac{9M}{\nu^{n+1}} \\
&= \Theta_{n+1}^{(1)}|\beta_{n+1}| + \frac{9M(n-m+1)}{\nu^{n+1}},
\end{aligned} \tag{4.17}$$

where the last inequality follows from

$$8 + 9\Theta_{n+1}^{(1)} + 27\frac{M}{\nu^{n+1}} < 8 + 9\varepsilon + \frac{9\varepsilon}{\nu^{n-m+1}} \le 8 + 9\varepsilon\left(1 + \frac{1}{\nu}\right) < 8 + 18\varepsilon < 9$$

due to $\varepsilon < 1/18$, and

$$1 + \frac{\varepsilon^2}{2} + \Theta_{n+1}^{(1)} + |\beta_{n+1}| + \frac{3M}{\nu^{n+1}} < 1 + 3\varepsilon + \frac{\varepsilon}{\nu^{n-m+1}} \le 1 + \varepsilon\left(3 + \frac{1}{\nu}\right) < 1 + 4\varepsilon < 3$$

due to $\varepsilon < 1/2$. In conclusion, we showed that

$$|\alpha_{n+1}| \le \max\left\{a^{n+1-m}\varepsilon, \Theta_{n+1}^{(1)}|\beta_{n+1}|\right\} + 9M\frac{n+1-m}{\nu^{n+1}}.$$

If $1 < \mu < 2$, we start again with

$$|\alpha_{n+1}| \le \frac{\rho+\eta}{2}|\alpha_n|\left(1 + \tilde{c}\theta + \frac{3M}{\nu^{n+1}}\right) + \tilde{c}\theta|\alpha_n| + \tilde{c}\theta\Theta_n^{(1)}|\beta_n| + \frac{3M}{\nu^{n+1}}$$

by (4.12) and Lemma 4.14, but keep $\tilde{c}$ in the estimate, as θ is possibly not as small as in the case $\mu \geq 2$. Next, by the usual induction hypothesis holding for almost every $\omega \in \mathscr{E}_0^{(1)}$, if ω is such that $a^{n-m}\varepsilon \geq \Theta_n^{(1)}|\beta_n|$,

$$|\alpha_n| < a^{n-m}\varepsilon + 9M\frac{n-m}{\nu^n}$$

and then we have that

$$\begin{aligned}|\alpha_{n+1}| &\leq a^{n-m}\frac{\rho+\eta}{2}\varepsilon\left(1+\tilde{c}\theta+\frac{3M}{\nu^{n+1}}\right)+9M\frac{\rho+\eta}{2}\left(1+\tilde{c}\theta+\frac{3M}{\nu^{n+1}}\right)\frac{n-m}{\nu^n}\\ &+2\tilde{c}\theta a^{n-m}\varepsilon+9\tilde{c}\theta M\frac{n-m}{\nu^n}+\frac{3M}{\nu^{n+1}}<a^{n-m}\frac{\rho+\eta+(4+\rho+\eta)\theta}{2}\varepsilon\\ &+\frac{9}{2}\left(1+3\tilde{c}\theta+\frac{\varepsilon}{\nu}\right)\frac{M(n-m)}{\nu^n}+\left(1+\frac{\varepsilon}{2}\right)\frac{3M}{\nu^{n+1}}\leq a^{n+1-m}\varepsilon\\ &+\frac{9M(n-m)}{\nu^{n+1}}+\frac{9M}{\nu^{n+1}}=a^{n+1-m}\varepsilon+\frac{9M(n-m+1)}{\nu^{n+1}},\end{aligned}$$

where the last inequality follows since by construction $1+\varepsilon/2 < 3$ and, having $\varepsilon < 1/18$ for $1<\mu<2$ by construction as well,

$$\frac{1+3\tilde{c}\theta+\frac{\varepsilon}{\nu}}{2} < \frac{1}{\nu}.$$

This claim is easily verified by recalling the definitions of all constants involved for this range of μ. It is equivalent to $\nu + 3\tilde{c}\theta\nu + \varepsilon < 2$, which is equivalent to

$$\mu^{\frac{5}{12}} + (\sqrt{2}-1)\left(\frac{2+\frac{1}{\sqrt[3]{\mu}}}{3-\frac{1}{\mu}}-1\right)+\varepsilon < 2.$$

Hence we only need to verify that

$$\varepsilon < 2-\mu^{\frac{5}{12}} - (\sqrt{2}-1)\left(\frac{2+\frac{1}{\sqrt[3]{\mu}}}{3-\frac{1}{\mu}}-1\right).$$

We can simply find a lower bound for the right-hand side, which exceeds $1/18 > \varepsilon$. Then noting that

$$\begin{aligned}2-\mu^{\frac{5}{12}}-(\sqrt{2}-1)\left(\frac{2+\frac{1}{\sqrt[3]{\mu}}}{3-\frac{1}{\mu}}-1\right) &> 2-2^{\frac{5}{12}}-\frac{\sqrt{2}-1}{2} > 2-\sqrt{2}-\frac{\sqrt{2}-1}{2} =\\ &\frac{5-3\sqrt{2}}{2} > \frac{5-\frac{9}{2}}{2}=\frac{1}{4}>\frac{1}{18},\end{aligned}$$

the claim follows.

If instead ω is such that $a^{n-m}\varepsilon < \Theta_n^{(1)}|\beta_n|$, then the induction hypothesis yields

$$|\alpha_n| < \Theta_n^{(1)}|\beta_n| + 9M\frac{n-m}{\nu^n}$$

for almost every such ω, and therefore

$$\begin{aligned}|\alpha_{n+1}| &\leq \frac{\rho+\eta}{2}\Theta_n^{(1)}|\beta_n|\left(1+\tilde{c}\theta+\frac{3M}{\nu^{n+1}}\right)+9M\frac{\rho+\eta}{2}\left(1+\tilde{c}\theta+\frac{3M}{\nu^{n+1}}\right)\frac{n-m}{\nu^n}\\ &+2\tilde{c}\theta\Theta_n^{(1)}|\beta_n|+\tilde{c}\theta 9M\frac{n-m}{\nu^n}+\frac{3M}{\nu^{n+1}}\leq\frac{\rho+\eta}{2}\Theta_n^{(1)}|\beta_n|(1+\theta)+2\theta\Theta_n^{(1)}|\beta_n|\\ &+\frac{\varepsilon^2}{2}\frac{3M}{\nu^{n+1}}+\frac{9\nu}{2}\left(1+3\tilde{c}\theta+\frac{\varepsilon}{\nu}\right)M\frac{n-m}{\nu^{n+1}}+\frac{3M}{\nu^{n+1}},\end{aligned}$$

yielding (4.14), by factoring out all the $\Theta_n^{(1)}|\beta_n|$ and estimating $\tilde{c} < 1$ only in the corresponding term. This follows since for $1 < \mu < 2$,

$$\frac{9\nu}{2}\left(1 + 3\tilde{c}\theta + \frac{\varepsilon}{\nu}\right) = \frac{9}{2}\mu^{\frac{5}{12}}\left[1 + 2(\sqrt{2}-1)\left(\frac{2+\frac{1}{\sqrt[3]{\mu}}}{3-\frac{1}{\mu}} - 1\right)\right] + \frac{9\varepsilon}{2},$$

which is, by the condition aforementioned, strictly less than 8 as $9\varepsilon/2 < 1/4$ and the convex function added to it achieves its maximum on $[1,2]$ at 2 with value

$$\frac{9}{2}2^{\frac{5}{12}}\left[1 + 2(\sqrt{2}-1)\left(\frac{2+\frac{1}{\sqrt[3]{2}}}{3-\frac{1}{2}} - 1\right)\right] = \frac{9}{2^{\frac{7}{12}}}\left(\frac{9-2\sqrt{2}}{5}\right) < 8 - \frac{1}{4} = \frac{31}{4}.$$

The last inequality holds true, since it is equivalent to

$$\left(\frac{9-2\sqrt{2}}{5}\right)^{12} < \left(\frac{31}{36}\right)^{12} 2^7$$

and while $(9-2\sqrt{2})/5 < (9 - 14/5)/5 = 31/25$ (whose power of 12 is easily verified to be less than 14), $31/36 > 5/6$, (whose power of 12 multiplied by 2^7 is easily verified to be larger than 14). The convexity of

$$f(x) = x^{\frac{5}{12}}\left[1 + 2(\sqrt{2}-1)\left(\frac{2+\frac{1}{\sqrt[3]{x}}}{3-\frac{1}{x}} - 1\right)\right] = x^{\frac{5}{12}}\left[3 - 2\sqrt{2} + 2(\sqrt{2}-1)\left(\frac{2+\frac{1}{\sqrt[3]{x}}}{3-\frac{1}{x}}\right)\right]$$

(and thus of $9f(x)/2$) on $[1,2]$ follows easily from computing $f''(x)$ to be

$$\frac{5}{12x^{\frac{7}{12}}}\left[3 - 2\sqrt{2} + 2(\sqrt{2}-1)\left(\frac{2+\frac{1}{\sqrt[3]{x}}}{3-\frac{1}{x}}\right)\right] - 2(\sqrt{2}-1)x^{\frac{5}{12}}\frac{\frac{1}{3x^{\frac{4}{3}}}\left(3-\frac{1}{x}\right) + \frac{1}{x^2}\left(2+\frac{1}{\sqrt[3]{x}}\right)}{\left(3-\frac{1}{x}\right)^2}.$$

Setting $f''(x) > 0$ and rearranging by first multiplying both sides by $x^{\frac{7}{12}}$ and then multiplying out the factor of x thus produced in the subtracted term yields, by finally moving the subtracted term to the right-hand side and dividing both sides by $2(\sqrt{2}-1)$, equivalently,

$$\frac{5}{12}\left(\frac{3-2\sqrt{2}}{2(\sqrt{2}-1)} + \frac{2+\frac{1}{\sqrt[3]{x}}}{3-\frac{1}{x}}\right) > \frac{\frac{1}{\sqrt[3]{x}} + \frac{2}{3x\sqrt[3]{x}} + \frac{2}{x}}{\left(3-\frac{1}{x}\right)^2}.$$

At last, multiplying both sides by the quadratic denominator yields

$$\frac{5}{12}\frac{3-2\sqrt{2}}{2(\sqrt{2}-1)}\left(3-\frac{1}{x}\right)^2 + 6 - \frac{2}{x} + \frac{3}{\sqrt[3]{x}} - \frac{1}{x\sqrt[3]{x}},$$

which rearranged gives

$$6 + \frac{5}{12}\frac{3-2\sqrt{2}}{2(\sqrt{2}-1)}\left(3-\frac{1}{x}\right)^2 > -\frac{2}{\sqrt[3]{x}} + \frac{4}{x} + \frac{5}{3x\sqrt[3]{x}}.$$

Finally, noting that the left-hand side is greater than 6, while the right-hand side is less than 6, we can easily conclude that $f''(x) > 0$ is verified for all $1 \leq x \leq 2$.

The fact that the convex function $9f/2$ has maximum at $x = 2$ follows directly from convexity and the fact that

$$\frac{9}{2^{\frac{7}{12}}}\left[1+2(\sqrt{2}-1)\left(\frac{2+\frac{1}{\sqrt[3]{2}}}{3-\frac{1}{2}}-1\right)\right] = \frac{9}{2}f(2) > \frac{9}{2}f(1) = \frac{9}{\sqrt{2}}.$$

This is seen by rearranging the inequality into

$$1+\frac{4}{5}(\sqrt{2}-1)\left(\frac{3}{4}+\frac{1}{\sqrt[3]{2}}\right) > \sqrt[12]{2}$$

and noting that, since $\sqrt[3]{2} < 4$, the left-hand side is greater than $1 + 4(\sqrt{2} - 1)/5$, which is in turn, given that $\sqrt{2} > {}^{7}/_{5}$, greater than ${}^{33}/_{25}$, which, raised to the power of 12, is trivially greater than 2. Having recovered (4.14), from now on one can conclude the same as for $\mu \geq 2$: by (4.9) and Lemma 4.14 we have (4.15) and again, since under this case the induction hypothesis implies $|\alpha_n| \leq \varepsilon|\beta_n| + 9M(n-m)\nu^{-n}$ by Lemma 4.15, we can apply (4.13) and Lemma 4.14, to get (4.16); plugging the bounds (4.15) and (4.16) into (4.14) yields (4.17) by the same estimates as for the case $\mu \geq 2$. □

Let

$$D := \frac{d\nu\varepsilon}{(d\nu-1)^2},$$

which is well-defined since by Lemma 4.13, $d > a > {}^{1}/_{\nu}$. Let also

$$D_{m,n} := \sum_{j=m}^{n} d^{n-j}\frac{j-m}{\nu^j}.$$

For all $\omega \in \mathscr{E}_0^{(1)}$, define the stopping time

$$\sigma := \inf\left\{n \geq m:\ a^{n-m}\varepsilon + Dd^{n-m} \leq \Theta_n^{(1)}|\beta_n| + 9M\varepsilon D_{m,n}\right\} \in \mathbb{N}\cup\infty.$$

Lemma 4.17.

a) For almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\tau < \infty\}$, $\sigma < \tau$ and $\Theta_\sigma^{(2)} \in [{}^{\delta}/_{4}, 1-{}^{\delta}/_{4}]$;

b) For almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\}$, for all $\sigma \leq n \leq \tau$

$$a^{n-m}\varepsilon + Dd^{n-m} \leq \Theta_n^{(1)}|\beta_n| + 9M\varepsilon D_{m,n}.$$

Proof.

a) For every $\omega \in \mathscr{E}_0^{(1)} \cap \{\tau < \infty\}$, if ω is such that $\sigma = m$ the claim is trivial, since by definition of m, for almost every such ω, $\Theta_\sigma^{(2)} \in [{}^{\delta}/_{2}, 1-{}^{\delta}/_{2}] \subset [{}^{\delta}/_{4}, 1-{}^{\delta}/_{4}] \subset [{}^{\delta}/_{8}, 1-{}^{\delta}/_{8}]$. Also $\tau > m = \sigma$ by definition. If ω is such that $\sigma > m$, for all $m \leq n < \sigma\wedge\tau$, by definition of σ,

$$\Theta_n^{(1)}|\beta_n| < a^{n-m}\varepsilon + Dd^{n-m} - 9M\varepsilon D_{m,n}.$$

Note that

$$D_{m,n} = \frac{d^{n-m}}{\nu^m}\sum_{j=m}^{n}\frac{j-m}{(d\nu)^{j-m}} < \frac{d^{n-m}}{\nu^m}\frac{d\nu}{(d\nu-1)^2},$$

since $d\nu > 1$ implies

$$\sum_{k=0}^{\infty} \frac{k}{(d\nu)^k} = \frac{1}{d\nu} \sum_{k=1}^{\infty} \frac{k}{(d\nu)^{k-1}} = \frac{1}{d\nu\left(1-\frac{1}{d\nu}\right)^2} = \frac{d\nu}{(d\nu-1)^2},$$

and therefore

$$9M\varepsilon D_{m,n} < \frac{9M\varepsilon d^{n-m}}{\nu^m} \frac{d\nu}{(d\nu-1)^2} < d^{n-m} \frac{3d\nu\varepsilon^2}{(d\nu-1)^2} < Dd^{n-m},$$

since $3\varepsilon^2 < \varepsilon$ as $\varepsilon < 1/3$. This implies that $Dd^{n-m} - 9M\varepsilon D_{m,n} > 0$, and therefore

$$\Theta_n^{(1)}|\beta_n| < a^{n-m}\varepsilon + Dd^{n-m}$$

and

$$\max\left\{a^{n-m}\varepsilon, \Theta_n^{(1)}|\beta_n|\right\} < \max\left\{a^{n-m}\varepsilon, a^{n-m}\varepsilon + Dd^{n-m}\right\} = a^{n-m}\varepsilon + Dd^{n-m}.$$

As a result, by Lemma 4.16, for almost every such ω,

$$|\alpha_n| \le a^{n-m}\varepsilon + Dd^{n-m} + 9M\frac{n-m}{\nu^n},$$

which implies that

$$|\alpha_n| < a^{n-m}\varepsilon + Dd^{n-m} + 3\varepsilon\frac{n-m}{\nu^{n-m}}.$$

By (4.10), Lemmas 4.14 and 4.15 (Lemma 4.15 is used as in Lemma 4.16, to ensure that the bounds on the errors hold even after time m) for almost every such ω it follows that

$$\begin{aligned}
|\Theta_{\sigma\wedge\tau}^{(2)} - \Theta_m^{(2)}| &\le \sum_{n=m}^{\sigma\wedge\tau-1} |\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| = \sum_{n=m}^{\sigma\wedge\tau-1} |2\rho_{n+1}(1-\Theta_n^{(2)})\Theta_n^{(1)}\beta_n - \rho_2(r_n)| \\
&< 2\sum_{n=m}^{\sigma\wedge\tau-1} \Theta_n^{(1)}|\beta_n| + \sum_{n=m}^{\sigma\wedge\tau-1} c|\alpha_n| + \Theta_n^{(1)}|\beta_n| + \frac{3M}{\nu^{n+1}} \\
&\le 3\sum_{n=m}^{\sigma\wedge\tau-1} \Theta_n^{(1)}|\beta_n| + c\sum_{n=m}^{\sigma\wedge\tau-1} |\alpha_n| + \varepsilon\sum_{n=m}^{\sigma\wedge\tau-1} \frac{1}{\nu^{n-m+1}} \\
&< (3+c)\varepsilon\sum_{n=m}^{\sigma\wedge\tau-1} a^{n-m} + (3+c)D\sum_{n=m}^{\sigma\wedge\tau-1} d^{n-m} + 3c\varepsilon\sum_{n=m}^{\sigma\wedge\tau-1} \frac{n-m}{\nu^{n-m}} \\
&+ \varepsilon\sum_{n=m}^{\sigma\wedge\tau-1} \frac{1}{\nu^{n-m+1}} \le (3+c)\varepsilon\sum_{i=0}^{\infty} a^i + (3+c)D\sum_{i=0}^{\infty} d^i + 3c\varepsilon\sum_{i=0}^{\infty} \frac{i}{\nu^i} \\
&+ \varepsilon\sum_{i=1}^{\infty} \frac{1}{\nu^i} = \varepsilon\left[\frac{3+c}{1-a} + \frac{3+c}{1-d}\frac{d\nu}{(d\nu-1)^2} + 3c\frac{\frac{1}{\nu}}{\left(1-\frac{1}{\nu}\right)^2} + \frac{\frac{1}{\nu}}{1-\frac{1}{\nu}}\right] \\
&= \varepsilon\frac{A}{4} < \frac{\delta}{4},
\end{aligned}$$

since by construction $\varepsilon < \delta/A$. Since $\Theta_m^{(2)} \in [\delta/2, 1-\delta/2]$ and for almost every $\omega \in \mathscr{E}_0^{(1)}$ the travelled distance has been less than $\delta/4$, it follows that for almost every such ω, $\Theta_{\sigma\wedge\tau}^{(2)} \in [\delta/4, 1-\delta/4] \subset [\delta/8, 1-\delta/8]$. This implies that for almost every $\omega \in \mathscr{E}_0^{(1)}$, $\sigma\wedge\tau \neq \tau$ (by definition of τ, if $\sigma\wedge\tau = \tau$, $\Theta_{\sigma\wedge\tau}^{(2)} \notin [\delta/8, 1-\delta/8]$, so if there were a nonnegligible event in $\mathscr{E}_0^{(1)}$ such that $\sigma\wedge\tau = \tau$, we would reach an almost sure contradiction on $\mathscr{E}_0^{(1)}$). Hence for almost every $\omega \in \mathscr{E}_0^{(1)}$, $\sigma\wedge\tau = \sigma$, that is $\sigma < \tau$, and in particular $\Theta_\sigma^{(2)} \in [\delta/4, 1-\delta/4]$.

b) Note first that for almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\}$, $\sigma(\omega) < \tau(\omega)$, because for every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\} \cap \{\tau = \infty\}$, trivially $\sigma(\omega) < \tau(\omega) = \infty$; whereas by part (a) for almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\tau < \infty\}$, $\sigma < \tau$, thus implying that for almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\} \cap \{\tau < \infty\}$, $\sigma(\omega) < \tau(\omega)$. Hence the remark follows. We can thus consider, for almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\}$, $\sigma \le n \le \tau$. Note that by definition of σ, the case $n = \sigma \ge m$ is trivially true. Assume the claim to be true up to some $\sigma < n < \tau$. The steps in Lemma 4.16 yielding (4.15) and (4.16) can be reproduced now under the induction hypothesis

$$a^{n-m}\varepsilon + Dd^{n-m} \le \Theta_n^{(1)}|\beta_n| + 9M\varepsilon D_{m,n},$$

which implies, due to $Dd^{n-m} - 9M\varepsilon D_{m,n} > 0$, that $a^{n-m}\varepsilon < \Theta_n^{(1)}|\beta_n|$ and therefore, being $\max\{a^{n-m}\varepsilon, \Theta_n^{(1)}|\beta_n|\} = \Theta_n^{(1)}|\beta_n|$, by Lemma 4.16 we have that

$$|\alpha_n| \le \Theta_n^{(1)}|\beta_n| + 9M\frac{n-m}{\nu^n}.$$

This ensures that (4.15) and (4.16) can be derived again as in Lemma 4.16. By using $\theta < 1/2$, $9\theta/2 < 9/4 < 3$ in (4.16) and by (4.15), we have that, rearranging,

$$\begin{aligned}&\left(\Theta_{n+1}^{(1)} + \frac{3M}{\nu^{n+1}}\right)\left(|\beta_{n+1}| + 3M\frac{n-m+1}{\nu^{n+1}}\right) \ge d\Theta_n^{(1)}|\beta_n| \\ &\ge d\left(a^{n-m}\varepsilon + Dd^{n-m} - 9M\varepsilon D_{m,n}\right) > a^{n+1-m}\varepsilon + Dd^{n-m+1} - 9M\varepsilon d D_{m,n},\end{aligned}$$

where the induction hypothesis has been used in the second last inequality, and the last inequality follows from Lemma 4.13, in particular from $d > a$. On the other hand note that

$$\begin{aligned}&\left(\Theta_{n+1}^{(1)} + \frac{3M}{\nu^{n+1}}\right)\left(|\beta_{n+1}| + 3M\frac{n-m+1}{\nu^{n+1}}\right) < \Theta_{n+1}^{(1)}|\beta_{n+1}| + 3M\varepsilon\frac{(n-m+1)}{\nu^{n+1}} \\ &\left(1 + \frac{1}{n-m+1} + \frac{1}{\nu^{n-m+1}}\right) < \Theta_{n+1}^{(1)}|\beta_{n+1}| + 9\varepsilon M\frac{(n-m+1)}{\nu^{n+1}}.\end{aligned}$$

Hence putting the two results together and rearranging, yields the inequality

$$a^{n+1-m}\varepsilon + Dd^{n-m+1} < \Theta_{n+1}^{(1)}|\beta_{n+1}| + 9\varepsilon M\left(dD_{m,n} + \frac{(n-m+1)}{\nu^{n+1}}\right).$$

Note that

$$dD_{m,n} + \frac{(n-m+1)}{\nu^{n+1}} = \sum_{j=m}^{n} d^{n-j+1}\frac{j-m}{\nu^j} + \frac{(n-m+1)}{\nu^{n+1}} = D_{m,n+1}$$

and thus the induction step is complete, as we have shown that

$$a^{n+1-m}\varepsilon + Dd^{n-m+1} < \Theta_{n+1}^{(1)}|\beta_{n+1}| + 9\varepsilon M D_{m,n+1}.$$

□

Theorem 4.18. *For almost all $\omega \in \mathscr{E}_0^{(1)}$, $\tau = \infty$ and the sample path $\{(\Theta_n(\omega), \pi_n(\omega))\}$ converges to $(\Theta_*(\omega), \pi_{\Theta_*(\omega)})$.*

Proof. We start with partitioning $\mathscr{E}_0^{(1)} = (\mathscr{E}_0^{(1)} \cap \{\sigma = \infty\}) \cup (\mathscr{E}_0^{(1)} \cap \{\sigma < \infty\})$, where σ is the stopping time defined just before the previous lemma.

Consider first $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma = \infty\}$. By Lemma 4.17 (a), for almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\tau < \infty\}$, $\sigma(\omega) < \tau(\omega) < \infty$, hence $\mathscr{E}_0^{(1)} \cap \{\sigma = \infty\} \cap \{\tau < \infty\}$ is negligible, and we can focus only on $\omega \in \mathscr{E}_0 \cap \{\sigma = \infty\} \cap \{\tau = \infty\}$. By Lemmas 4.11 and 4.14 to 4.16 and the definition of σ, we can estimate as in Lemma 4.17 (a) for all $n \geq m$, yielding

$$\begin{aligned}\sum_{n=m}^{\infty} |\Theta^{(2)}_{n+1} - \Theta^{(2)}_n| &= \sum_{n=m}^{\infty} |2\rho_{n+1}(1-\Theta^{(2)}_n)\Theta^{(1)}_n\beta_n + \rho_2(r_n)| \\ &< 3\sum_{n=m}^{\infty} \Theta^{(1)}_n|\beta_n| + c\sum_{n=m}^{\infty}|\alpha_n| + \varepsilon\sum_{n=m}^{\infty}\frac{1}{\nu^{n-m+1}} < \frac{\delta}{4}.\end{aligned}$$

Therefore for almost every such ω, $\Theta^{(2)}_n$ converges within $[\delta/4, 1-\delta/4] \subset [\delta/8, 1-\delta/8]$. We know that for almost every ω, there is a subsequence $\Theta_{n_k}(\omega) \longrightarrow \Theta_*(\omega) \in E_1 \cap \overline{\mathscr{K}}_{\varepsilon,\frac{\delta}{8}}$, and also that $\varepsilon > 0$ can be chosen arbitrarily small, which, by Lemma 4.15 applied with $\tau = \infty$, means that for almost every such ω, $\Theta^{(1)}_n(\omega)$ is eventually arbitrarily small. Therefore, for almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma = \infty\} \cap \{\tau = \infty\}$ fixed, eventually $\Theta_n(\omega)$ does not exit any set $\overline{\mathscr{K}}_{\varepsilon,\frac{\delta}{8}}$ with ε is arbitrarily small and $\delta = \delta(\omega)$ fixed small enough, while $\Theta^{(2)}_n(\omega) \longrightarrow \Theta^{(2)}_*(\omega) \in [\delta/8, 1-\delta/8]$. Thus $\Theta^{(1)}_n(\omega) \longrightarrow 0$, finally implying that $\Theta_n(\omega) \longrightarrow \Theta_*(\omega) \in E_1 \cap \overline{\mathscr{K}}_{\varepsilon,\frac{\delta}{8}}$. Then the convergence of $\pi_n(\omega) \longrightarrow \pi_{\Theta_*(\omega)}$ for almost every such ω trivially follows from $\ell(\omega) = 0$ for all $\omega \in \mathscr{E}_0$.

Consider now $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\}$. By Lemma 4.17 (b), for almost every such ω, $\sigma < \tau$. Note that by (4.13), $\beta_{n+1} + \beta_n = (2-\rho_{n+1})\Theta^{(1)}_n\beta_n + \rho_5(r_n)$, and that for any integers $q \geq p$,

$$\sum_{n=p}^{q}(-1)^{q-n}(\beta_{n+1}+\beta_n) = \beta_{q+1} + (-1)^{q-p}\beta_p.$$

Thus for any $\sigma \leq k < \tau$,

$$\beta_{k+1}+(-1)^{k-\sigma}\beta_\sigma = \sum_{n=\sigma}^{k}(-1)^{k-n}(\beta_{n+1}+\beta_n) = \sum_{n=\sigma}^{k}(-1)^{k-n}\left[(2-\rho_{n+1})\,\beta_n\Theta^{(1)}_n + \rho_5(r_n)\right]. \tag{4.18}$$

By definition of σ, Lemmas 4.14 to 4.16 and Lemma 4.17 (b) for all $\sigma \leq n \leq k$, for almost every such ω,

$$|\rho_5(r_n)| < \frac{\theta}{2\nu}|\alpha_n| + \frac{\theta}{2\nu}\Theta^{(1)}_n|\beta_n| + \frac{3M}{\nu^{n+1}} \leq \frac{\theta}{\nu}\Theta^{(1)}_n|\beta_n| + 3M\frac{n-m+1}{\nu^{n+1}}.$$

Indeed, being $9M\varepsilon D_{m,n} - Dd^{n-m} < 0$, it holds that

$$\begin{aligned}|\alpha_n| &\leq \max\left\{a^{n-m}\varepsilon, \Theta^{(1)}_n|\beta_n|\right\} + 9M\frac{n-m}{\nu^n} \\ &\leq \max\{\Theta^{(1)}_n|\beta_n| + 9M\varepsilon D_{m,n} - Dd^{n-m}, \Theta^{(1)}_n|\beta_n|\} + 9M\frac{n-m}{\nu^n} \\ &= \Theta^{(1)}_n|\beta_n| + 9M\frac{n-m}{\nu^n},\end{aligned}$$

then $9\theta/(2\nu) < 3/\nu$ implies the claim. As $\nu > 1$, we can conclude that

$$|\rho_5(r_n)| < \theta\Theta^{(1)}_n|\beta_n| + 3M\frac{n-m+1}{\nu^{n+1}}, \tag{4.19}$$

We now use (4.19) to show that $\{\beta_n\}_{n=\sigma}^{k}$ has either alternating sign or an almost geometric decay, and that the transitions between the two regimes are such that it will be possible to bound, uniformly in k, the term $\sum_{n=\sigma}^{k} \Theta_n^{(1)} |\beta_n|$ appearing in (4.18). First note that by (4.13) and (4.19) and Lemma 4.15, for almost every ω considered,

$$\beta_{n+1}\beta_n = \left(-1 + (2-\rho_{n+1})\,\Theta_n^{(1)}\right)\beta_n^2 + \beta_n\rho_5(r_n) \le \left(-1 + 2\Theta_n^{(1)}\right)\beta_n^2 + |\beta_n||\rho_5(r_n)| \le \left(-1 + 2\Theta_n^{(1)}\right)\beta_n^2 + \theta\Theta_n^{(1)}\beta_n^2 + 3M|\beta_n|\frac{n-m+1}{\nu^{n+1}} < \left[-1 + \frac{4}{3}\theta + \theta^2\right]\beta_n^2 + 3M|\beta_n|\frac{n-m+1}{\nu^{n+1}},$$

as $\Theta_n^{(1)} < \varepsilon < 2\theta/3$ by construction. If

$$3M\frac{n-m+1}{\nu^{n+1}} \le \frac{\theta}{6}|\beta_n|,$$

then it follows that (as long as $\beta_n \neq 0$, because in this case the term would not even count in the contributions to the series we are trying to estimate, so we would just skip to the next nonzero one)

$$\beta_{n+1}\beta_n < \left[-1 + \left(\frac{4}{3} + \frac{1}{6}\right)\theta + \theta^2\right]\beta_n^2 = \left(-1 + \frac{3}{2}\theta + \theta^2\right)\beta_n^2 < 0,$$

since for all $0 < x < 1/2$ the parabola $-1 + 3x/2 + x^2$ is increasing, valued -1 at 0 and 0 at $1/2$. By Lemma 4.13 this fact implies that for all choices of θ, depending on $\mu > 1$, the factor is strictly negative. Which means that the sign of the β_n alternates at this time. If

$$3M\frac{n-m+1}{\nu^{n+1}} > \frac{\theta}{6}|\beta_n|,$$

it is possible to estimate

$$|\beta_n| < \frac{18M}{\theta}\frac{n-m+1}{\nu^{n+1}} < 4\frac{n-m+1}{\nu^{n-m+1}},$$

which produces an almost geometric upper bound, implying, were it to continue, the convergence of the series we are seeking to estimate. We then need to control the number of transitions between one regime and the other. Recall that we are working for all $\sigma \le n \le k < \tau$, and define two new stopping times sequences for all $\sigma \le n < \tau$ as follows: for all

$$\omega \in N_{\le} := \left\{3M\frac{\sigma-m+1}{\nu^{\sigma+1}} \le \frac{\theta}{6}|\beta_\sigma|\right\} \cap \{\sigma < \infty\} \cap \mathscr{E}_0^{(1)},$$

denote $\eta_0 := \sigma$ and for all $i \in \mathbb{N}$ define

$$\eta_i := \inf\left\{n \ge \sigma_i :\ 3M\frac{n-m+1}{\nu^{n+1}} \le \frac{\theta}{6}|\beta_n|\right\}$$

and

$$\sigma_i := \inf\left\{n \ge \eta_{i-1} :\ 3M\frac{n-m+1}{\nu^{n+1}} > \frac{\theta}{6}|\beta_n|\right\};$$

for all

$$\omega \in N_{>} := \left\{3M\frac{\sigma-m+1}{\nu^{\sigma+1}} > \frac{\theta}{6}|\beta_\sigma|\right\} \cap \{\sigma < \infty\} \cap \mathscr{E}_0^{(1)},$$

for all $i \in \mathbb{N}$ denote $\sigma_0 := \sigma$ and define

$$\eta_i := \inf\left\{n \ge \sigma_{i-1} :\ 3M\frac{n-m+1}{\nu^{n+1}} \le \frac{\theta}{6}|\beta_n|\right\}$$

and

$$\sigma_i := \inf\left\{n \geq \eta_i :\ 3M\frac{n-m+1}{\nu^{n+1}} > \frac{\theta}{6}|\beta_n|\right\}.$$

The crucial fact is then the following: that the condition

$$|\beta_n| < \frac{18M}{\theta}\frac{n-m+1}{\nu^{n+1}}$$

defining the σ_i random times can be combined with Lemma 4.17 (b), which ensures that for almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\}$, $|\beta_n| \geq \Theta_n^{(1)}|\beta_n| \geq a^{n-m}\varepsilon + Dd^{n-m} - 9M\varepsilon D_{m,n}$, for all $\sigma \leq n \leq k < \tau$, for all k. For the range of n considered then this yields

$$a^{n-m}\varepsilon + Dd^{n-m} - 9M\varepsilon D_{m,n} < \frac{18M}{\theta}\frac{n-m+1}{\nu^{n+1}}.$$

Recall that by construction $Dd^{n-m} - 9M\varepsilon D_{m,n} > 0$, thus implying

$$a^{n-m}\varepsilon < \frac{18M}{\theta}\frac{n-m+1}{\nu^{n+1}} < \frac{6\varepsilon}{\theta}\frac{n-m+1}{\nu^{n-m+1}} = \frac{6\varepsilon}{\theta\nu}\frac{n-m+1}{\nu^{n-m}}.$$

If we rearrange this, we get

$$(a\nu)^{n-m} < \frac{6(n-m+1)}{\theta\nu}.$$

This condition cannot be satisfied infinitely often, as by Lemma 4.13, $a\nu > 1$ and we would have an exponential sequence upper-bounded by a linear one. Let

$$\bar{n} := \min\left\{m \leq n \in \mathbb{N} :\ \forall\, j \geq n,\ (a\nu)^{j-m} \geq \frac{6(j-m+1)}{\theta\nu}\right\}.$$

Note that $\bar{n} - m = \iota$. For almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\}$, there will be a uniformly bounded random time $\bar{i} = \bar{i}(\omega) < \iota$ such that: on $N_{\leq}$, $\iota = \eta_{\bar{i}} < \infty$ and $\sigma_{\bar{i}+1} = \infty$ and for all $i \geq \bar{i}+1$ $\eta_i = \sigma_i = \infty$; on $N_>$, $\iota = \eta_{\bar{i}} < \infty$ and $\sigma_{\bar{i}} = \infty$ and for all $i \geq \bar{i}+1$ $\eta_i = \sigma_i = \infty$. On both events, the number of transitions between oscillatory and almost geometric decay is bounded by ι, which is deterministic. This observation allows us to work on $\mathscr{E}_0^{(1)} \cap \{\sigma < \infty\} = N_{\leq} \cup N_>$, and we will show that for almost every ω in this partitioning, $\tau(\omega) = \infty$. It is enough to show this in full detail on $N_>$ as the argument is similar on $N_{\leq}$, with the corresponding ordering of the η_i and σ_i random times.

Since for almost every $\omega \in N_>$, for all $k \geq m$, $|\beta_k| < \varepsilon$, and the sign alternates as aforementioned, for all $\eta_i \leq n \leq k \wedge (\sigma_i - 1)$, for all $1 \leq i \leq \bar{i}$, by (4.18) it follows that

$$\begin{aligned}
&|\beta_{k\wedge(\sigma_i-1)+1} + (-1)^{k\wedge(\sigma_i-1)-\eta_i}\beta_{\eta_i}| = \left|\sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} (-1)^{k\wedge(\sigma_i-1)-n}[(2-\rho_{n+1})\Theta_n^{(1)}\beta_n + \rho_5(r_n)]\right| \\
&= \left|\sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} (-1)^{k\wedge(\sigma_i-1)-n}\mathrm{sign}(\beta_n)|\beta_n|(2-\rho_{n+1})\Theta_n^{(1)} + \sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} (-1)^{k\wedge(\sigma_i-1)-n}\rho_5(r_n)\right| \\
&= \left|\sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} \mathrm{sign}(\beta_{k\wedge(\sigma_i-1)})|\beta_n|\Theta_n^{(1)}(2-\rho_{n+1}) + \sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} (-1)^{k\wedge(\sigma_i-1)-n}\rho_5(r_n)\right| \\
&\geq \sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} (2-\rho_{n+1})|\beta_n|\Theta_n^{(1)} - \left|\sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} (-1)^{k\wedge(\sigma_i-1)-n}\rho_5(r_n)\right| \\
&\geq \sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} |\beta_n|\Theta_n^{(1)} - \sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} |\rho_5(r_n)| \geq (1-\theta)\sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} |\beta_n|\Theta_n^{(1)} - 3M\sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} \frac{n-m+1}{\nu^{n+1}} \\
&\geq (1-\theta)\sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} |\beta_n|\Theta_n^{(1)} - \varepsilon\sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} \frac{n-m+1}{\nu^{n-m+1}},
\end{aligned}$$

where the second last inequality follows by (4.19). Since

$$|\beta_{k\wedge(\sigma_i-1)+1} + (-1)^{k\wedge(\sigma_i-1)-\eta_i}\beta_{\eta_i}| < 2\varepsilon,$$

by rearranging we can conclude that

$$\sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} \Theta_n^{(1)}|\beta_n| < \frac{2 + \sum_{n=\eta_i}^{k\wedge(\sigma_i-1)} \frac{n-m+1}{\nu^{n-m+1}}}{1-\theta}\varepsilon. \tag{4.20}$$

It is known that we have finitely many sums of this type, and the number of them is uniformly upper-bounded by ι. As to the other sums, the estimate is simpler, since for all $\eta_{i-1} \leq n \leq k \wedge (\eta_i - 1)$, for all $1 \leq i \leq \bar{i}$,

$$\sum_{n=\sigma_{i-1}}^{k\wedge(\eta_i-1)} \Theta_n^{(1)}|\beta_n| < 4\varepsilon \sum_{n=\sigma_{i-1}}^{k\wedge(\eta_i-1)} \frac{n-m+1}{\nu^{n-m+1}}. \tag{4.21}$$

In conclusion, for any $\sigma \leq k < \tau$ we have an upper bound uniform both in k and in ω, achieved by splitting the whole series into these segments, and adding up the two estimates' contributions:

$$\sum_{n=\sigma}^{k} \Theta_n^{(1)}|\beta_n| < \iota\varepsilon\frac{2 + \sum_{n=m}^{\infty} \frac{n-m+1}{\nu^{n-m+1}}}{1-\theta} + 4\varepsilon\sum_{n=m}^{\infty} \frac{n-m+1}{\nu^{n-m+1}} = \varepsilon\left(\iota\frac{2 + \frac{\nu}{(\nu-1)^2}}{1-2\theta} + 4\frac{\nu}{(\nu-1)^2}\right).$$

As mentioned earlier, for almost every $\omega \in N_{\leq}$, for all $\eta_{i-1} \leq n \leq k \wedge (\sigma_i - 1)$, for all $1 \leq i \leq \bar{i}$, one can proceed with estimating in the same way, to get (4.20), but with the appropriate indexing of the random times, that is

$$\sum_{n=\eta_{i-1}}^{k\wedge(\sigma_i-1)} \Theta_n^{(1)}|\beta_n| < \frac{2 + \sum_{n=\eta_{i-1}}^{k\wedge(\sigma_i-1)} \frac{n-m+1}{\nu^{n-m+1}}}{1-\theta}\varepsilon.$$

One derives similarly also (4.21) for all $\eta_i \le n \le k \wedge (\sigma_i - 1)$, for all $1 \le i \le \bar{i}$:

$$\sum_{n=\sigma_i}^{k\wedge(\eta_i-1)} \Theta_n^{(1)}|\beta_n| < 4\varepsilon \sum_{n=\sigma_i}^{k\wedge(\eta_i-1)} \frac{n-m+1}{\nu^{n-m+1}}.$$

This shows that the bound

$$\sum_{n=\sigma}^{k} \Theta_n^{(1)}|\beta_n| < \varepsilon\left(\iota\frac{2+\frac{\nu}{(\nu-1)^2}}{1-\theta} + 4\frac{\nu}{(\nu-1)^2}\right) \tag{4.22}$$

applies uniformly in $m \le k < \tau \in \mathbb{N} \cup \infty$ and for almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\}$.

Thanks to these results we now show that $\mathscr{E}_0^{(1)} \cap \{\sigma < \infty\} \cap \{\tau < \infty\}$ is negligible. For almost every fixed $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\} \cap \{\tau < \infty\}$ one can use (4.22) with $k = \tau - 1$, and by Lemmas 4.11 and 4.14 to 4.17, one has the following (recall that $\sigma < \tau$ is ensured by Lemma 4.17 (a), while $a^{n-m}\varepsilon < \Theta_n^{(1)}|\beta_n|$ is shown as in Lemma 4.17 (b)):

$$\begin{aligned}
|\Theta_\sigma^{(2)} - \Theta_\tau^{(2)}| &\le \sum_{n=\sigma}^{\tau-1} |\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| \le 3\sum_{n=\sigma}^{\tau-1} \Theta_n^{(1)}|\beta_n| + c\sum_{n=\sigma}^{\tau-1} |\alpha_n| + \varepsilon\sum_{n=\sigma}^{\tau-1} \frac{1}{\nu^{n-m+1}} \\
&\le 3\sum_{n=\sigma}^{\tau-1} \Theta_n^{(1)}|\beta_n| + c\sum_{n=\sigma}^{\tau-1} \max\left\{a^{n-m}\varepsilon, \Theta_n^{(1)}|\beta_n|\right\} + 3\varepsilon\frac{n-m}{\nu^{n-m}} \\
&+ \varepsilon\sum_{n=\sigma}^{\tau-1} \frac{1}{\nu^{n-m+1}} < (3+c)\sum_{n=\sigma}^{\tau-1} \Theta_n^{(1)}|\beta_n| + 3\varepsilon\sum_{n=m}^{\infty} \frac{n-m}{\nu^{n-m}} + \varepsilon\sum_{n=m}^{\infty} \frac{1}{\nu^{n-m+1}} \\
&\le (3+c)\varepsilon\left(\iota\frac{2+\frac{\nu}{(\nu-1)^2}}{1-\theta} + 4\frac{\nu}{(\nu-1)^2}\right) + 3\varepsilon\frac{\nu}{(\nu-1)^2} + \varepsilon\frac{1}{\nu-1} \\
&= \varepsilon\frac{B}{8} < \frac{\delta}{8}.
\end{aligned}$$

But since $\omega \in \mathscr{E}_0^{(1)} \cap \{\tau < \infty\}$, by Lemma 4.17 (a), $\Theta_\sigma^{(2)} \in [\delta/4, 1-\delta/4]$ for almost every such ω, and since it has just been proved that $\Theta_\tau^{(2)}$ has travelled a distance less than $\delta/8$ for almost all $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\} \cap \{\tau < \infty\}$, $\Theta_\tau^{(2)} \in [\delta/8, 1-\delta/8]$ for almost every such ω, that is an almost sure contradiction with the definition of τ. Hence the considered event must be negligible. Thus we can consider just $\mathscr{E}_0 \cap \{\sigma < \infty\} \cap \{\tau = \infty\}$. We can perform the same exact estimates as above (except invoking Lemma 4.17 (a) of course): thanks to Lemmas 4.11 and 4.14 to 4.16, and Lemma 4.17 (b), it holds that for almost every $\omega \in \mathscr{E}_0 \cap \{\sigma < \infty\} \cap \{\tau = \infty\}$,

$$\sum_{n=\sigma}^{\infty} |\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| < (3+c)\sum_{n=\sigma}^{\infty} \Theta_n^{(1)}|\beta_n| + 3\varepsilon\sum_{n=m}^{\infty} \frac{n-m}{\nu^{n-m}} + \varepsilon\sum_{n=m}^{\infty} \frac{1}{\nu^{n-m+1}} \le \varepsilon\frac{B}{8} < \frac{\delta}{8}.$$

The last step follows from (4.22) holding uniformly for any $k \ge \sigma$. This finally yields the convergence of $\Theta_n^{(2)}$ in $[\delta/8, 1-\delta/8]$ for almost every $\omega \in \mathscr{E}_0^{(1)}$ by definition of τ. Exploiting again ε being arbitrarily small, as we did on $\mathscr{E}_0^{(1)} \cap \{\sigma = \infty\}$, we conclude that $\Theta_n^{(1)}$ vanishes. Thus similarly, for almost every $\omega \in \mathscr{E}_0^{(1)} \cap \{\sigma < \infty\}$, the convergence of Θ_n to some $\Theta_* \in E_1$ implies the convergence of π_n to π_{Θ_*}. □

Remark 4.19. *On $\mathscr{E}_0^{(i)}$ with $i \in \{2,3\}$ one can proceed by exploiting the symmetry of the model, define σ and τ accordingly in terms of the corresponding coordinates, and show an analogous version of Theorem 4.18 for $i \in \{2,3\}$ as well, thus yielding convergence of sample paths for almost every $\omega \in \mathscr{E}_0$.*

4.5 Convergence on $\mathscr{E}_>$

The main goal of this section is to show that $\{\Theta_n\}$ converges almost surely on $\mathscr{E}_>$. Let us start with an introductory remark corresponding to the one opening Section 3.5, but adapted to the random setting.

Remark 4.20. *For almost every $\omega \in \{\ell > 0\}$, there is no subsequence $\{\Theta_{n_k}\}_{k\in\mathbb{N}}$ bounded away from $\partial\Sigma$. Thus $\Theta_n \longrightarrow \partial\Sigma$ almost surely on $\{\ell > 0\}$.*

Proof. Consider that for almost every $\omega \in \{\ell > 0\}$, $\|v_n\|_1 \longrightarrow \ell$ by Lemma 4.5, and by Lemma 4.1 we also have that $\|R_{n+1}\|_1 \longrightarrow 0$ for almost all $\omega \in \{\ell > 0\}$, as $n \longrightarrow \infty$. Assume by contradiction that there is a nonnegligible event $A \subsetneq \{\ell > 0\}$ on which a subsequence $\{\Theta_{n_k}\}_{k\in\mathbb{N}}$ is bounded away from the boundary. For almost all $\omega \in A$, at the times $\{n_k\}$, we would be able to apply (4.5) from Proposition 4.7, with some constant $0 < \mathscr{c}(\omega) < 1$, yielding

$$\|v_{n_k+1}\|_1 < \mathscr{c}\|v_{n_k}\|_1 + \|R_{n_k+1}\|_1.$$

Taking the limit as $k \longrightarrow \infty$ on both sides yields $\ell \leq \mathscr{c}\ell$, which implies that $\mathscr{c} = 1$ for almost all $\omega \in A$, since $\ell > 0$. Thus a contradiction has been reached. $\square$

Define the event

$$\mathscr{V} := \left\{\exists\{n_k\}_{k\in\mathbb{N}},\ \exists v \in V : \lim_{k\longrightarrow\infty} \Theta_{n_k} = v\right\} \supseteq \mathcal{D}.$$

Note that $\mathscr{V} = \bigcup_{i=1}^{3} \mathscr{V}^{(i)}$, where for every $i \in \{1, 2, 3\}$ we define

$$\mathscr{V}^{(i)} := \left\{\exists\{n_k\}_{k\in\mathbb{N}} : \lim_{k\longrightarrow\infty} \Theta_{n_k} = v_i\right\}.$$

The following lemma corresponds to Lemma 3.34. For reasons that will soon be clear, we anticipate its proof in this part of the section.

Lemma 4.21. *For almost every fixed $\omega \in \mathscr{V}$ and any subsequence $\{n_k\}$, such that $\Theta_{n_k}(\omega) \longrightarrow v \in V$, the set of accumulation points of $\{(\Theta_{n_k-1}, \pi_{n_k-1})\}$, $\{(\Theta_{n_k}, \pi_{n_k})\}$ and $\{(\Theta_{n_k+1}, \pi_{n_k+1})\}$ is a subset of $\left\{\left(v, \pi_v \pm \frac{\ell}{2} e_{-1}(v)\right)\right\}$. Moreover, if the considered ω and $\{n_k\}$ are such that also $\{\pi_{n_k}\}$ converges, $\{\pi_{n_k-1}\}$, $\{\pi_{n_k}\}$ and $\{\pi_{n_k+1}\}$ asymptotically oscillate between $\pi_* = \pi_v \pm \frac{\ell}{2} e_{-1}(v)$ and $\hat{\pi}_* = \pi_v \mp \frac{\ell}{2} e_{-1}(v)$, while $\{\Theta_{n_k-1}\}$,$\{\Theta_{n_k}\}$,$\{\Theta_{n_k+1}\}$ all tend to v: that is $(\Theta_{n_k-1}, \pi_{n_k-1}) \longrightarrow (v, \hat{\pi}_*)$, $(\Theta_{n_k}, \pi_{n_k}) \longrightarrow (v, \pi_*)$ and $(\Theta_{n_k+1}, \pi_{n_k+1}) \longrightarrow (v, \hat{\pi}_*)$ as $k \longrightarrow \infty$.*

Proof. By symmetry, without loss of generality, assume that $\omega \in \mathscr{V}^{(2)}$ and let $\{n_{k_l}\}_{l\in\mathbb{N}}$ be such that $(\Theta_{n_{k_l}}, \pi_{n_{k_l}}) \longrightarrow (v_2, \pi_*)$ as $l \longrightarrow \infty$. For simplicity relabel as n_r the subsubsequence $\{n_{k_l}\}$. The first part of the argument is the same as in Lemma A.26, with p,q replaced by Θ,π and the ρ coefficient being now time-dependent. Briefly, note that, by (2.32), as $r \longrightarrow \infty$,

$$(1 - \rho_{n_r})\Theta^{(1)}_{n_r-1} + \rho_{n_r}(1 - \pi^{(1)}_{n_r-1} - \pi^{(1)}_{n_r}) \longrightarrow 0 \tag{4.23}$$

$$(1 - \rho_{n_r})\Theta^{(2)}_{n_r-1} + \rho_{n_r}(1 - \pi^{(2)}_{n_r-1} - \pi^{(2)}_{n_r}) \longrightarrow 1 \tag{4.24}$$

$$(1 - \rho_{n_r})\Theta^{(3)}_{n_r-1} + \rho_{n_r}(1 - \pi^{(3)}_{n_r-1} - \pi^{(3)}_{n_r}) \longrightarrow 0, \tag{4.25}$$

and therefore (4.24) implies that $\pi^{(2)}_{n_r-1} + \pi^{(2)}_{n_r} \longrightarrow 0$, since $0 \leq \pi^{(2)}_{n_r-1} \leq 1$. Hence, $\pi^{(2)}_{n_r} \longrightarrow 0 = \pi^{(2)}_*$ and $\Theta^{(2)}_{n_r-1} \longrightarrow 1$, which also implies that for $i \in \{1, 3\}$, $\Theta^{(i)}_{n_r-1} \longrightarrow 0$, that is $\Theta_{n_r-1} \longrightarrow v_2$ too. (4.23) and (4.25) directly imply also that for $i \in \{1, 3\}$,

$\pi^{(i)}_{n_r}+\pi^{(i)}_{n_r-1} \longrightarrow 1$ as $r \longrightarrow \infty$, then for all $i \in \{1,3\}$, $1-\pi^{(i)}_{n_r}-\pi^{(i)}_{n_r-1} \longrightarrow 0$ as $r \longrightarrow \infty$. Since by Lemma 4.5 for almost all $\omega \in \mathscr{V}^{(2)}$, $V(\Theta_{n_r}, \pi_{n_r}) \longrightarrow V(v_2, \pi_*) = \ell(\omega)$, one can determine π_*. Indeed the definition of the potential has not changed from Lemma 3.34 so, by the same argument from the corresponding part, we have that either $\pi_* = ((1+\ell)/2, 0, (1-\ell)/2)$ or $\pi_* = ((1-\ell)/2, 0, (1+\ell)/2)$. Without loss of generality, assume the latter scenario. Then since it has been shown that

$$\pi^{(i)}_{n_r} + \pi^{(i)}_{n_r-1} \longrightarrow \begin{cases} 0, & i = 2 \\ 1, & i \neq 2 \end{cases},$$

quite trivially $\pi_{n_r-1} \longrightarrow ((1+\ell)/2, 0, (1-\ell)/2) =: \hat{\pi}_*$, the complementary form of π_*. The same holds for π_{n_r+1}.

The argument now slightly differs from the one in Lemmas A.26 and 3.34, due to the presence of perturbation terms, which, however, vanish almost surely by Lemma 4.1 and therefore do not significantly change the proof of the last fact. As $r \longrightarrow \infty$, $R^{(1)}_{n_r+1} \longrightarrow 0$ almost surely, so by (2.31),

$$\pi^{(1)}_{n_r+1} = \frac{\Theta^{(3)}_{n_r}}{\Theta^{(1)}_{n_r} + \Theta^{(3)}_{n_r}} \pi^{(2)}_{n_r} + \frac{\Theta^{(2)}_{n_r}}{\Theta^{(1)}_{n_r} + \Theta^{(2)}_{n_r}} \pi^{(3)}_{n_r} + R^{(1)}_{n_r+1} \longrightarrow \frac{1+\ell}{2},$$

because

$$0 \leq \frac{\Theta^{(3)}_{n_r}}{\Theta^{(1)}_{n_r} + \Theta^{(3)}_{n_r}} \leq 1,$$

and $\pi^{(2)}_{n_r} \longrightarrow 0$ for almost every $\omega \in \mathscr{V}^{(2)}$, while

$$\frac{\Theta^{(2)}_{n_r}}{\Theta^{(1)}_{n_r} + \Theta^{(2)}_{n_r}} \longrightarrow 1,$$

and $\pi^{(3)}_{n_r} \longrightarrow (1+\ell)/2$ for almost every $\omega \in \mathscr{V}^{(2)}$. By (2.31),

$$\pi^{(2)}_{n_r+1} = \frac{\Theta^{(3)}_{n_r}}{\Theta^{(2)}_{n_r} + \Theta^{(3)}_{n_r}} \pi^{(1)}_{n_r} + \frac{\Theta^{(1)}_{n_r}}{\Theta^{(1)}_{n_r} + \Theta^{(2)}_{n_r}} \pi^{(3)}_{n_r} + R^{(2)}_{n_r+1} \longrightarrow 0$$

as both terms next to $\pi^{(1)}_{n_r}$ and $\pi^{(3)}_{n_r}$, along with $R^{(2)}_{n_r}$, vanish for almost every $\omega \in \mathscr{V}^{(2)}$. Therefore $\pi_{n_r+1} \longrightarrow \hat{\pi}_*$ and $\pi_{n_r} + \pi_{n_r+1} \longrightarrow (1,0,1)$ for almost every $\omega \in \mathscr{V}^{(2)}$. As a result

$$\Theta_{n_r+1} = (1-\rho_{n_r+1})\Theta_{n_r} + \rho_{n_r+1}[\mathbf{1} - (\pi_{n_r} + \pi_{n_r+1})] \longrightarrow (1-\rho)v_2 + \rho(\mathbf{1} - (1,0,1)) = v_2$$

and $\pi_{\Theta_{n_r+1}} \longrightarrow \pi_{v_2} = (1/2, 0, 1/2)$ as $r \longrightarrow \infty$, for almost every $\omega \in \mathscr{V}^{(2)}$. As in Lemma A.26, we now start from an arbitrary convergent subsubsequence of $\{\Theta_{n_k+1}(\omega), \pi_{n_k+1}(\omega)\}$ and of $\{\Theta_{n_k-1}, \pi_{n_k-1}\}$. We will denote them $\{\Theta_{n_r+1}, \pi_{n_r+1}\}$ and $\{\Theta_{n_r-1}, \pi_{n_r-1}\}$, and their limit will be, in each case separately, denoted as $(\Theta_*, \hat{\pi}_*)$, to be determined. The underlying hypothesis is that for the ω considered $\Theta_{n_r}(\omega) \longrightarrow v$. In the case of the convergent forward shift subsubsequence, by (2.32) we can see that

$$\pi_{n_r} = \frac{(1-\rho_{n_r+1})\Theta_{n_r} - \Theta_{n_r+1}}{\rho_{n_r+1}} + \mathbf{1} - \pi_{n_r+1} \longrightarrow \frac{(1-\rho)v - \Theta_*}{\rho} + \mathbf{1} - \hat{\pi}_* =: \pi_*,$$

so we have that $(\Theta_{n_r}, \pi_{n_r}) \longrightarrow (v, \pi_*)$ and we obtain, through the same argument as the one shown above, that the forms of π_* and $\hat{\pi}_*$ are the claimed ones, and that $\Theta_* = v$ for almost every ω considered. We proceed similarly in the case of the convergent backward shift subsubsequence. By (2.32), we can see that

$$\pi_{n_r} = \frac{(1-\rho_{n_r})\Theta_{n_r-1} - \Theta_{n_r}}{\rho_{n_r}} + \mathbf{1} - \pi_{n_r} \longrightarrow \frac{(1-\rho_{n_r})\Theta_* - v}{\rho_{n_r}} + \mathbf{1} - \hat{\pi}_* =: \pi_*,$$

so $(\Theta_{n_r}, \pi_{n_r}) \longrightarrow (v, \pi_*)$, and we can repeat the same strategy just discussed in the previous case. The second part of the claim trivially follows by taking $n_r = n_k$ in the argument above. □

4.5.1 Convergence with boundary initial conditions

In this section the probabilistic model is forced to admit boundary initial conditions, to briefly study its corresponding asymptotic behaviour. Although the main objective of our analysis is the model having $\Theta_0 \notin \partial\Sigma$ (regular initial conditions), it is interesting to study, like we did for the dynamical system in Section 3.5, also the evolution of the stochastic process with boundary initial conditions. The asymptotic behaviours we will observe give precious intuition regarding the regular case. At the same time, the system is much more rigid, with boundary initial conditions, hence the results are easier to obtain.

Lemma 4.22. *Let $\Theta_0 \in E_i$ for some $i \in \{1, 2, 3\}$. Then for almost all $\omega \in \Omega$ there exists $\Theta_*(\omega) \in \overline{E}_i$, such that $\Theta_n(\omega) \longrightarrow \Theta_*(\omega)$ as $n \longrightarrow \infty$.*

Proof. Without loss of generality, by symmetry, assume $i = 1$, that is $\Theta_0 \in E_1$, or equivalently $\Theta_0^{(1)} = 0$ and $0 < \Theta_0^{(2)} < 1$. It follows immediately, from the form taken by M_{Θ_0}, similarly to the initial part of the proof of Lemma 3.29, that

$$\pi_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 - \Theta_0^{(2)} & 0 & 0 \\ \Theta_0^{(2)} & 0 & 0 \end{pmatrix} \pi_0 + R_1 = \begin{pmatrix} \pi_0^{(2)} + \pi_0^{(3)} + R_1^{(1)} \\ (1 - \Theta_0^{(2)})\pi_0^{(1)} + R_1^{(2)} \\ \Theta_0^{(2)}\pi_0^{(1)} + R_1^{(3)} \end{pmatrix}.$$

However, since $B_1^{(2)} \sim \text{Bin}(\mu N_0^{(2)}, 0)$ and $B_1^{(3)} \sim \text{Bin}(\mu N_0^{(3)}, 1)$, as $\Theta_0^{(1)} = 0$, $B_1^{(2)} = 0$ and $B_1^{(3)} = \mu N_0^{(3)}$ almost surely. Then (2.28) implies that almost surely

$$R_1^{(1)} = \frac{B_1^{(3)} - \mu N_0^{(3)}}{\sigma_1} + \frac{\mu N_0^{(2)} - B_1^{(2)} - \mu N_0^{(2)}}{\sigma_1} = 0.$$

It follows that almost surely

$$\Theta_1^{(1)} = (1 - \rho_1)\Theta_0^{(1)} + \rho_1(1 - \pi_0^{(1)} - \pi_0^{(2)} - \pi_0^{(3)}) = 0$$

and therefore, by induction, one has that almost surely $\Theta_n^{(1)} = 0$ for all $n \in \mathbb{N}$. Note also that almost surely, as a consequence of $R_n^{(1)} = 0$ for all $n \in \mathbb{N}$ and $R_n \in \Pi_0 := \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$, we have that $R_n^{(3)} = -R_n^{(2)}$ almost surely. Due to this fact, M_{Θ_n} has the same form as M_{Θ_0} for all n, almost surely, and it follows that

$$\pi_{n+1} = \begin{pmatrix} 0 & 1 & 1 \\ 1 - \Theta_n^{(2)} & 0 & 0 \\ \Theta_n^{(2)} & 0 & 0 \end{pmatrix} \pi_n + R_{n+1} = \begin{pmatrix} \pi_n^{(2)} + \pi_n^{(3)} \\ (1 - \Theta_n^{(2)})\pi_n^{(1)} + R_{n+1}^{(2)} \\ \Theta_n^{(2)}\pi_n^{(1)} - R_{n+1}^{(2)} \end{pmatrix}.$$

Therefore, by (2.32)

$$\begin{aligned} \Theta_{n+1}^{(2)} - \Theta_n^{(2)} &= \rho_{n+1}[1 - \Theta_n^{(2)} - (1 - \Theta_n^{(2)})\pi_n^{(1)} - R_{n+1}^{(2)} - \pi_n^{(2)}] \\ &= \rho_{n+1}\left[(1 - \Theta_n^{(2)})(1 - \pi_n^{(1)}) - \pi_n^{(2)} - R_{n+1}^{(2)}\right] = \rho_{n+1}[(1 - \Theta_n^{(2)})(1 - \pi_n^{(1)}) \\ &- (1 - \Theta_{n-1}^{(2)})\pi_{n-1}^{(1)} - R_n^{(2)} - R_{n+1}^{(2)}]. \end{aligned}$$

Since almost surely $\Theta_n^{(1)} = 0$ for all n, by (2.32) we also have that almost surely $\pi_{n-1}^{(1)} = 1 - \pi_n^{(1)}$ for all $n \in \mathbb{N}$, and thus that

$$\Theta_{n+1}^{(2)} - \Theta_n^{(2)} = -\rho_{n+1}(1 - \pi_n^{(1)})(\Theta_n^{(2)} - \Theta_{n-1}^{(2)} + r_n), \tag{4.26}$$

where $r_n := R_n^{(2)} + R_{n+1}^{(2)}$. Note that by Remark 4.2, there is a deterministic $\tilde{m}$ large enough, such that for all $n \geq \tilde{m}$,

$$\rho_{n+1} < \rho + \frac{1 - \rho}{2} = \tilde{\rho} := \frac{\rho + 1}{2} < 1,$$

hence $\rho_{n+1}(1-\pi_n^{(1)}) < \tilde{\rho}$ for all $n \geq \tilde{m}$.

We show that $\Theta_n^{(2)}$ converges almost surely, by showing that for almost all $\omega \in \Omega$,

$$\sum_{k=0}^{\infty} |\Theta_{k+1}^{(2)} - \Theta_k^{(2)}| < \infty.$$

Fix $1 < \nu < \sqrt{\mu}$, by Lemma 4.1 for almost all $\omega \in \Omega$, there is $\bar{m} = \bar{m}(\omega)$ such that for all $n \geq \bar{m}$, $|R_n^{(2)}| \leq \nu^{-n}$. Let $m = m(\omega) := \max\{\tilde{m}, \bar{m}(\omega)\}$. For all $n \geq m$ we iterate (4.26) down to m, and it follows that

$$\begin{aligned}\Theta_{n+1}^{(2)} - \Theta_n^{(2)} &= (-1)^{n-m}\tilde{\rho}^{n-m} \prod_{k=m+1}^{n} (1-\pi_k^{(1)})(\Theta_{m+1}^{(2)} - \Theta_m^{(2)}) \\ &\quad + \sum_{k=m+1}^{n} (-1)^{n-k+1}\tilde{\rho}^{n-k+1} \prod_{j=k+1}^{n} (1-\pi_k^{(1)}) r_k,\end{aligned}$$

and since for almost all ω for all $n \geq m$, $|r_n| \leq |R_{n+1}^{(2)}| + |R_n^{(2)}| \leq 2\nu^{-n}$, if we define $\lambda := \max\{\tilde{\rho}, \nu^{-1}\}$, the following estimate holds,

$$\begin{aligned}|\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| &\leq \tilde{\rho}^{n-m}|\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| + \sum_{k=m+1}^{n} \tilde{\rho}^{n-k+1}|r_k| \leq \tilde{\rho}^{n-m}|\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| \\ &\quad + 2\sum_{k=m+1}^{n} \frac{\tilde{\rho}^{n-k+1}}{\nu^k} \leq \lambda^{n-m}|\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| + 2\sum_{k=m+1}^{n} \lambda^{n-k+1}\lambda^k \\ &= \lambda^{n-m}|\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| + 2(n-m)\lambda^{n+1} \\ &= n\lambda^n \left(2\lambda\left(1-\frac{m}{n}\right) + \frac{|\Theta_{m+1}^{(2)} - \Theta_m^{(2)}|}{n\lambda^m}\right).\end{aligned}$$

Thus $\Theta_{n+1}^{(2)} - \Theta_n^{(2)} = \mathcal{O}_\omega(n\lambda^n)$, which shows the claim of convergence of the series of absolute increments. Since for all n, $\Theta_n^{(1)} = 0$ almost surely, the almost sure convergence of $\Theta_n^{(2)}$ implies the almost sure convergence of Θ_n to some Θ_* in the closure of E_1. □

Note that Lemma 4.22 already implies almost sure convergence of the stochastic process with boundary initial conditions. However, due to the perturbations, we lost the alternating series estimate of Lemma 3.29, and as a result we cannot exclude the vertices on E_i from being possible candidates for the limit of a boundary sample path. For this reason, we needed to prove Lemma 4.21 earlier than we did with Lemma 3.34 in Section 3.5. We can now exploit it to make progress in understanding the sample paths of $\{\pi_n\}$.

Corollary 4.23. *If $\Theta_0 \in E_i$, then for all $\pi_0 \in \Sigma$ and almost every $\omega \in \Omega$, the set of accumulation points of the sample paths is $\{(\Theta_*, \pi_{\Theta_*} \pm \beta e_{-1}(\Theta_*))\}$ for some $\Theta_* \in \overline{E_i}$ and $\beta \geq 0$. Moreover, $\beta = \ell(\omega)/2$.*

Proof. If $\omega \in \mathcal{D}$, $\Theta_n \longrightarrow \Theta_* \in V$, and the claim follows from Lemma 4.21. Assume $\omega \in \mathcal{D}^c$, and without loss of generality, by symmetry, let $i = 1$. In this case by Lemma 4.22 the limit $\Theta_* \in E_1$, and we can exploit continuity of the nonautonomous iteration map

$$\Phi_{n+1}(\Theta, \pi) := \begin{pmatrix} (1-\rho_{n+1})\Theta + \rho_{n+1}(\mathbf{1} - \pi - M_\Theta \pi) \\ M_\Theta \pi \end{pmatrix} + \begin{pmatrix} -\rho_{n+1} R_{n+1} \\ R_{n+1} \end{pmatrix}$$

on $E_1 \times \Sigma$. Using $\mathbf{1} - \Theta_n = 2\pi_{\Theta_n}$ and rearranging (2.32) we obtain

$$(I + M_{\Theta_n})\pi_n = 2\pi_{\Theta_n} - \left(\frac{1}{\rho_{n+1}}(\Theta_{n+1} - \Theta_n) + R_{n+1} \right).$$

By Lemmas 4.1 and 4.22, almost surely $\Theta_{n+1} - \Theta_n, R_{n+1} \longrightarrow \mathbf{0}$ as $n \longrightarrow \infty$, thus implying that $(I + M_{\Theta_n})\pi_n \longrightarrow 2\pi_{\Theta_*}$ almost surely. It then follows that

$$\frac{I + M_{\Theta_n}}{2}\pi_n \longrightarrow \pi_{\Theta_*}$$

and therefore

$$\frac{I + M_{\Theta_*}}{2}(\pi_n - \pi_{\Theta_*}) \longrightarrow \mathbf{0}$$

almost surely, as

$$\frac{I + M_{\Theta_n}}{2}\pi_n - \pi_{\Theta_*} = \left[\frac{I + M_{\Theta_n}}{2} - \frac{I + M_{\Theta_*}}{2} \right] \pi_n + \frac{I + M_{\Theta_*}}{2}(\pi_n - \pi_{\Theta_*})$$

and almost surely

$$\frac{I + M_{\Theta_n}}{2}\pi_n - \pi_{\Theta_*} \longrightarrow \mathbf{0},$$

while

$$\left[\frac{I + M_{\Theta_n}}{2} - \frac{I + M_{\Theta_*}}{2} \right] \longrightarrow \mathbf{0}$$

by continuity. Thus we conclude that, like on $\mathcal{D}$, for almost all $\omega \in \mathcal{D}^c$, $\pi_n - \pi_{\Theta_*}$ either becomes aligned with $e_{-1}(\Theta_*)$ as $n \longrightarrow \infty$ or vanishes (in which case $\beta = 0$). Lemma 4.5 ensures, even though without monotonicity, the existence of the limit ℓ for $V(\Theta_n, \pi_n)$ also for boundary sample paths (indeed, note that the proof of Lemma 4.5 does not make any hypothesis such as Θ_0 being in the interior of the simplex). Therefore one can proceed to determine β as in Corollary 3.30. It suffices to change the notation in p, q to Θ, π; then the argument is exactly the same, and there is no need to repeat it here. ☐

Proposition 4.24. *If $\Theta_0 \in E_i$, then for all $\pi_0 \in \Sigma$ and for almost all $\omega \in \Omega$, there exists $\Theta_* \in \overline{E}_i$ and $\beta \geq 0$ such that the boundary sample path approaches the 2-cycle $\{(\Theta_*, \pi_{\Theta_*} \pm \beta e_{-1}(\Theta_*))\}$, with $\beta = |\beta_0|$, where $\pi_0 = \pi_{\Theta_0} + \alpha_0 e_0(\Theta_0) + \beta_0 e_{-1}(\Theta_0)$.*

Proof. We only need to show the cycling between the two accumulation points, since Proposition 4.7 ensures that the set of accumulation points is as per the claim. Unlike in Proposition 3.31, we need to consider the vertex as a possible accumulation point. However, whether Θ_* is a vertex or not, with the usual choice of eigenvectors introduced in Proposition 3.31 and Lemma 3.34, which extends to the vertices as well, as far as e_{-1} is concerned, we have

$$e_{-1}(\Theta_{n+1}) - e_{-1}(\Theta_n) = (\Theta^{(2)}_{n+1} - \Theta^{(2)}_n)e_0.$$

Then for a boundary sample path (4.2) reads

$$\alpha_{n+1}e_0 + \beta_{n+1}e_{-1}(\Theta_{n+1}) = -\frac{\rho_{n+1}}{2}\alpha_n e_0 - \beta_n e_{-1}(\Theta_n) + \left(1 - \frac{\rho_{n+1}}{2}\right) R_{n+1}, \qquad (4.27)$$

which yields, being $R^{(1)}_{n+1} = 0$ and $R^{(3)}_{n+1} = -R^{(2)}_{n+1}$ for all n as shown in Lemma 4.22,

$$\begin{aligned} -\beta_{n+1} &= \beta_n \\ \alpha_{n+1} + (1 - \Theta^{(2)}_n)\beta_{n+1} &= -\frac{\rho_{n+1}}{2}\alpha_n - \beta_n(1 - \Theta^{(2)}_n) + \left(1 - \frac{\rho_{n+1}}{2}\right) R^{(2)}_{n+1} \\ -\alpha_{n+1} + \Theta^{(2)}_n\beta_{n+1} &= \frac{\rho_{n+1}}{2}\alpha_n - \beta_n\Theta^{(2)}_n - \left(1 - \frac{\rho_{n+1}}{2}\right) R^{(2)}_{n+1} \end{aligned}$$

for almost all $\omega \in \Omega$. The first equation plugged into the second makes the latter a scalar multiple of the third, so, as in Proposition 3.31, we keep only the first and the third equation. Simplifying the third equation by means of the first yields

$$\alpha_{n+1} = -\frac{\rho_{n+1}}{2}\alpha_n + \beta_n(\Theta_n^{(2)} - \Theta_{n+1}^{(2)}) + R_{n+1}^{(2)} \tag{4.28}$$

$$\beta_{n+1} = -\beta_n. \tag{4.29}$$

We can also rewrite $\Theta_{n+1} - \Theta_n$ in eigencoordinates by rearranging (2.32) and using $\pi_n - \pi_{\Theta_n} = \alpha_n e_0(\Theta_n) + \beta_n e_{-1}(\Theta_n)$. It yields

$$\begin{aligned}\Theta_{n+1} - \Theta_n &= \rho_{n+1}(2\pi_{\Theta_n} - \pi_{n+1} - \pi_n) = -\rho_{n+1}\left[(M_{\Theta_n} + I)(\pi_n - \pi_{\Theta_n}) + R_{n+1}\right] \\ &= -\rho_{n+1}(\alpha_n e_0 + R_{n+1}),\end{aligned}$$

which implies that $\Theta_{n+1}^{(2)} - \Theta_n^{(2)} = -\rho_{n+1}\left(\alpha_n + R_{n+1}^{(2)}\right)$, which turns (4.28) and (4.29) into the following system, almost surely,

$$\alpha_{n+1} = \rho_{n+1}\left(\beta_n - \frac{1}{2}\right)\alpha_n + \left[1 + \rho_{n+1}\left(\beta_n - \frac{1}{2}\right)\right]R_{n+1}^{(2)} \tag{4.30}$$

$$\beta_{n+1} = -\beta_n. \tag{4.31}$$

These equations hold almost surely, so by (4.31) it is clear that almost surely β_n oscillates between β_0 and $-\beta_0$. The asymptotic 2-periodicity will follow from showing that α_n vanishes almost surely. The same estimates of Corollary 3.30 leading to $|\beta_n| < 1/\sqrt{3}$ apply and allow us to prove this claim. We iterate (4.30) once more and plug it into (4.31), yielding

$$\begin{aligned}\alpha_{n+1} &= \rho_{n+1}\rho_n\left(\beta_n - \frac{1}{2}\right)\left(\beta_{n-1} - \frac{1}{2}\right)\alpha_{n-1} + \rho_{n+1}\left(\beta_n - \frac{1}{2}\right)\left[1 + \rho_n\left(\beta_{n-1} - \frac{1}{2}\right)\right]R_n^{(2)} \\ &\quad + \left[1 + \rho_{n+1}\left(\beta_n - \frac{1}{2}\right)\right]R_{n+1}^{(2)} = -\rho_{n+1}\rho_n\left(\beta_{n-1} + \frac{1}{2}\right)\left(\beta_{n-1} - \frac{1}{2}\right)\alpha_{n-1} \\ &\quad - \rho_{n+1}\left(\beta_{n-1} + \frac{1}{2}\right)\left[1 + \rho_n\left(\beta_{n-1} - \frac{1}{2}\right)\right]R_n^{(2)} - \left[1 + \rho_{n+1}\left(\beta_{n-1} + \frac{1}{2}\right)\right]R_{n+1}^{(2)} \\ &\quad = -\rho_{n+1}\rho_n\left(\beta_{n-1}^2 - \frac{1}{4}\right)\alpha_{n-1} + r_{n+1},\end{aligned}$$

where

$$r_{n+1} := -\rho_{n+1}\left(\beta_{n-1} + \frac{1}{2}\right)\left[1 + \rho_n\left(\beta_{n-1} - \frac{1}{2}\right)\right]R_n^{(2)} - \left[1 + \rho_{n+1}\left(\beta_{n-1} + \frac{1}{2}\right)\right]R_{n+1}^{(2)}.$$

Note that for all $n \geq m$, since

$$\left|\beta_{n-1} \pm \frac{1}{2}\right| < \frac{1}{\sqrt{3}} + \frac{1}{2} < \frac{2}{\sqrt{3}} < \frac{4}{3},$$

$$|r_{n+1}| < \left|\beta_{n-1} + \frac{1}{2}\right|\left[1 + \left|\beta_{n-1} - \frac{1}{2}\right|\right]|R_n^{(2)}| + \left[1 + \left|\beta_{n-1} + \frac{1}{2}\right|\right]|R_{n+1}^{(2)}| < \left(1 + \frac{4}{3}\right)^2 \frac{1}{\nu^n}.$$

This leads to

$$|\alpha_{n+1}| < \tilde{\rho}^2\left|\beta_{n-1}^2 - \frac{1}{4}\right||\alpha_{n-1}| + |r_{n+1}| < \frac{1}{12}|\alpha_{n-1}| + \frac{6}{\nu^n}.$$

Recall that, in Lemma 4.22, we showed that:

- By Remark 4.2, there is a deterministic $\tilde{m}$ large enough, such that $\rho_{n+1} < \tilde{\rho} < 1$ for all $n \geq \tilde{m}$ and for all fixed $1 < \nu < \sqrt{\mu}$;
- By Lemma 4.1 for almost all $\omega \in \Omega$, there is $\bar{m} = \bar{m}(\omega)$, such that for all $n \geq \bar{m}$, $|R_n^{(2)}| \leq \nu^{-n}$.

Thus let $m = m(\omega) \geq \max\{\tilde{m}, \bar{m}(\omega)\}$. For all $n > m$,

$$|\alpha_{n+1}| < \frac{1}{12}|\alpha_{n-1}| + \frac{6}{\nu^n}. \tag{4.32}$$

Set $\lambda := \max\{1/12, \nu^{-1}\}$. Then for any $k \in \mathbb{N}$, if $n+1 = m+2k$, iterating (4.32) down to m yields

$$\begin{aligned}|\alpha_{m+2k}| &< \frac{1}{12^k}|\alpha_m| + 6\sum_{j=0}^{k-1}\frac{1}{12^j}\frac{1}{\nu^{m+2k-1-2j}} < \lambda^k|\alpha_m| + 6\lambda^{m+2k-1}\sum_{j=0}^{k-1}\lambda^{-j}\\ &= \lambda^k|\alpha_m| + 6\lambda^{m+2k-1}\frac{1-\lambda^{-k}}{1-\lambda^{-1}} = \lambda^k|\alpha_m| + 6\frac{\lambda^{m+2k-1}}{\lambda^{k-1}}\frac{1-\lambda^k}{1-\lambda}\\ &< \lambda^k|\alpha_m| + 6\frac{\lambda^{m+k}}{1-\lambda} = \lambda^k\left(|\alpha_m| + 6\frac{\lambda^m}{1-\lambda}\right).\end{aligned}$$

Similarly, if $n+1 = m+2k+1$, iterating (4.32) down to $m+1$ yields

$$\begin{aligned}|\alpha_{m+2k+1}| &< \frac{1}{12^k}|\alpha_{m+1}| + 6\sum_{j=0}^{k-1}\frac{1}{12^j}\frac{1}{\nu^{m+2k-2j}} = \frac{1}{12^k}|\alpha_{m+1}| + 6\sum_{j=0}^{k-1}\frac{1}{12^j}\frac{1}{\nu^{m+1+2k-1-2j}}\\ &< \lambda^k\left(|\alpha_{m+1}| + 6\frac{\lambda^{m+1}}{1-\lambda}\right) < \lambda^k\left(|\alpha_{m+1}| + 6\frac{\lambda^m}{1-\lambda}\right).\end{aligned}$$

Let $M := \max\{|\alpha_m|, |\alpha_{m+1}|\}$. Then for all $n > m$,

$$|\alpha_{n+1}| < \left(M + 6\frac{\lambda^m}{1-\lambda}\right)\lambda^{\lfloor\frac{n+1-m}{2}\rfloor} < \frac{1}{\sqrt{\lambda}^m}\left(M + 6\frac{\lambda^m}{1-\lambda}\right)\sqrt{\lambda}^{n+1} = \mathcal{O}_\omega\left(\lambda^{\lfloor\frac{n+1}{2}\rfloor}\right).$$

Thus for all $n > m$,

$$|\alpha_n| < \frac{1}{\sqrt{\lambda}^m}\left(M + 6\frac{\lambda^m}{1-\lambda}\right)\sqrt{\lambda}^n. \tag{4.33}$$

This yields a geometrically decaying upper bound on the first eigencoordinate. That $\beta = |\beta_0|$ almost surely, is obvious by the alternating β_n coordinate. In particular, almost surely, $\beta = 0$ if and only if $\pi_0 = \pi_{\Theta_0} + \alpha_0 e_0(\Theta_0)$ for some real α_0. □

Remark 4.25. *The geometric upper bound on the decay of $|\alpha_n|$ for a boundary sample path is uniform on the simplex, by the same reasoning as in Remark 3.32, hence there exist a constant $\tilde{M}$, dependent on $\omega \in \Omega$, such that for almost all ω,*

$$|\alpha_n| < \tilde{M}\sqrt{\lambda}^n. \tag{4.34}$$

Equivalently, $\alpha_n = \mathcal{O}_\omega\left(\sqrt{\lambda}^n\right)$.

Theorem 4.26. *If $\Theta_0 \in E_i$ for any $i \in \{1,2,3\}$, $\{\Theta_n\}$ converges almost surely to an almost surely $\overline{E}_i$-valued bounded random variable Θ.*

Proof. For almost every $\omega \in \Omega$, $\{\Theta_n(\omega)\}$ converges to $\Theta_*(\omega) \in \overline{E}_i$ by Lemma 4.22, thus by defining $\Theta^{(j)}(\omega) := \limsup_{n\longrightarrow\infty} \Theta_n^{(j)}(\omega)$ for all $j \in \{1,2,3\}$, we obtain a random variable Θ to which $\{\Theta_n\}$ converges almost surely, which is well defined on $(\Omega, \mathscr{F}, \{\mathscr{F}_n\}, \mathbb{P})$, since it is $\mathscr{F}_\infty$-measurable, and is almost surely $\overline{E}_i$-valued, since by construction $\mathbb{P}(\omega :\ \Theta(\omega) = \Theta_*(\omega)) = 1$. □

The following is an immediate consequence of Theorem 4.26 and Corollary 4.23. Recall that Σ^* denotes the medial triangle in Σ (boundary excluded).

Corollary 4.27. *Let $\Theta_0 \in E_i$ for any $i \in \{1,2,3\}$ and $\pi_0 = \pi_{\Theta_0} + \alpha_0 e_0(\Theta_0) + \beta_0 e_{-1}(\Theta_0)$. Then $\{\pi_n\}$ either almost surely diverges if $|\beta_0| > 0$, or almost surely converges to π_Θ (which is almost surely $\partial\Sigma^*$-valued) if $\beta_0 = 0$, where Θ denotes the almost sure limit of $\{\Theta_n\}$.*

4.5.2 Structure of the set of accumulation points

In this section we go back to regular initial conditions. Unlike in Section 3.5.2, we will not make direct use of the results in Section 4.5.1, since in Proposition 4.29 we will be able to rely directly on the dynamical behaviour of deterministic boundary orbits to understand sample paths approaching the boundary. It is useful to describe the stochastic process as a dynamical system having nonautonomous iteration map Φ_{n+1}, including a random perturbation part, which is vanishing by Lemma 4.1. The iteration map has almost sure limit Φ, which is continuous on $\Sigma_0 \times \Sigma$:

$$
\begin{aligned}
\begin{pmatrix}\Theta_{n+1}\\ \pi_{n+1}\end{pmatrix} = \Phi_{n+1}(\Theta_n, \pi_n) &= \begin{pmatrix}(\Phi_{n+1})_\Theta(\Theta_n,\pi_n)\\ (\Phi_{n+1})_\pi(\Theta_n,\pi_n)\end{pmatrix}\\
&:= \begin{pmatrix}(1-\rho_{n+1})\Theta_n + \rho_{n+1}(\mathbf{1} - \pi_n - M_{\Theta_n}\pi_n)\\ M_{\Theta_n}\pi_n\end{pmatrix} + \begin{pmatrix}-\rho_{n+1}R_{n+1}\\ R_{n+1}\end{pmatrix},
\end{aligned}
$$

where

$$
\begin{aligned}
\Phi_{n+1}(\Theta,\pi) &:= \begin{pmatrix}(1-\rho_{n+1})\Theta + \rho_{n+1}(\mathbf{1} - \pi - M_\Theta\pi)\\ M_\Theta\pi\end{pmatrix} + \begin{pmatrix}-\rho_{n+1}R_{n+1}\\ R_{n+1}\end{pmatrix}\\
&\longrightarrow \begin{pmatrix}(1-\rho)\Theta + \rho(\mathbf{1} - \pi - M_\Theta\pi)\\ M_\Theta\pi\end{pmatrix} =: \Phi(\Theta,\pi).
\end{aligned}
$$

Remark 4.28. *For almost every $\omega \in \Omega$ fixed, the limit map Φ is a uniform limit for Φ_{n+1} with respect to the variables Θ and π and the 1-norm.*

Proof. By Remark 4.2, the boundedness of Σ and Lemma 4.1, we have that

$$
\sup_{\Sigma_0\times\Sigma} \|\Phi\left(\Theta,\pi\right) - \Phi_{n+1}\left(\Theta,\pi\right)\|_1 \le 2\left(|\rho-\rho_{n+1}| + \|R_{n+1}\|_1\right) = \mathcal{O}_\omega\left(\frac{1}{\mu^n} + \frac{1}{\nu^n}\right).
$$

The limit is possibly pointwise in ω, since the random time, at which the geometric upper bound starts being active as per Lemma 4.1, has not been shown uniformly bounded. □

Proposition 4.29. *For almost every $\omega \in \{\ell > 0\}$, the set of accumulation points of the sample path $\{(\Theta_n(\omega), \pi_n(\omega))\}$ is a subset of*

$$
\{(\Theta, \pi_\theta \pm \beta e_{-1}(\Theta)) :\ \Theta \in \partial\Sigma,\ V(\Theta, \pi_\Theta \pm \beta e_{-1}(\Theta)) = \ell(\omega)\}\,.
$$

Proof. First recall that by Remark 4.20 for almost all $\omega \in \{\ell > 0\}$, for every $\{n_k\}$, $\Theta_{n_k} \longrightarrow \partial\Sigma$. Consider any convergent $(\Theta_{n_k}, \pi_{n_k}) \longrightarrow (\Theta_0^*, \pi_0^*)$ as $k \longrightarrow \infty$, for some $\Theta_0^* \in \partial\Sigma$, $\pi_0^* \in \Sigma$. If ω is such that $\Theta_0^* \in V$, then $\omega \in \mathscr{V}$, so by Lemma 4.21, for almost every such ω,

$$\pi_0^* = \pi_{\Theta_0^*} \pm \frac{\ell(\omega)}{2} e_{-1}(\Theta_0^*);$$

since then, due to $\|e_{-1}(\Theta_0^*)\|_1 = 2$,

$$\|\pi_0^* - \pi_{\Theta_0^*}\|_1 = \left\| \frac{\ell(\omega)}{2} e_{-1}(\Theta_0^*) \right\|_1 = \ell(\omega),$$

the claim follows. If ω is such that $\Theta_0^* \notin V$, then $\Theta_0^*(\omega) \in E_i$ for some $i \in \{1, 2, 3\}$. We need to show that for almost every such ω, π_0^* is of the form $\pi_{\Theta_0^*} \pm \beta e_{-1}(\Theta_0^*)$, where $\beta > 0$ is such that $V(\Theta_0^*, \pi_{\Theta_0^*} \pm \beta e_{-1}(\Theta_0^*)) = \ell(\omega)$. For every such ω, consider then $\{(\Theta_{n_k-1}, \pi_{n_k-1})\}_{k\in\mathbb{N}}$, which, by boundedness, has a convergent subsequence $(\Theta_{n_{k_s}-1}, \pi_{n_{k_s}-1}) \longrightarrow (\Theta_1^*, \pi_1^*)$ as $s \longrightarrow \infty$. For simplicity denote $n_r := n_{k_s} - 1$. We claim that for almost every such ω, $\Theta_1^*(\omega) \notin V$. In fact if such ω's constitute a nonnegligible event on which the subsequential limit $\Theta_1^* \in V$, for almost every such ω in the nonnegligible event, for some $v \in V$, $(\Theta_{n_r}, \pi_{n_r}) \longrightarrow (v, \pi_1^*)$ as $r \longrightarrow \infty$, where $\pi_1^* - \pi_v$ is in the eigenspace spanned by $e_{-1}(v)$; then by Lemma 4.21, it must follow that for almost every such ω in this nonnegligible event, $(\Theta_{n_r+1}, \pi_{n_r+1}) \longrightarrow (v, \hat{\pi}_1^*)$, where $\hat{\pi}_1^*$ denotes the point complementary to π_1^*, as per the notation in Lemma 4.21. The contradiction for almost all such ω's is thus reached, since $n_r + 1 = n_{k_s}$, so we concluded that with positive probability $\Theta_{n_{k_s}} \longrightarrow v \in V$, while it was assumed that $\Theta_{n_k} \longrightarrow \Theta_0^* \notin V$ for all considered ω.

Since for almost every ω considered, $\Theta_1^*(\omega)$ is not a vertex, the limit map Φ is continuous at (Θ_1^*, π_1^*), therefore for almost every such ω,

$$\begin{aligned} \Phi(\Theta_1^*, \pi_1^*) &= \Phi\left(\lim_{r\to\infty} (\Theta_{n_r}, \pi_{n_r})\right) = \lim_{r\to\infty} \Phi(\Theta_{n_r}, \pi_{n_r}) = \lim_{r\to\infty} \Phi_{n_r+1}(\Theta_{n_r}, \pi_{n_r}) \\ &= \lim_{r\to\infty} (\Theta_{n_r+1}, \pi_{n_r+1}) = \lim_{s\to\infty} (\Theta_{n_{k_s}}, \pi_{n_{k_s}}) = (\Theta_0^*, \pi_0^*), \end{aligned}$$

where the crucial part is the third equality, which is the main difference from Propositions A.27 and 3.35, and it follows by Remark 4.28. Indeed, as $r \longrightarrow \infty$, by Lemma 4.1 and Remark 4.2, for some $1 < \nu < \sqrt{\mu}$ fixed, for almost every ω,

$$\|\Phi(\Theta_{n_r}, \pi_{n_r}) - \Phi_{n_r+1}(\Theta_{n_r}, \pi_{n_r})\|_1 = \mathcal{O}_\omega\left(\frac{1}{\mu^{n_r}} + \frac{1}{\nu^{n_r}}\right)$$

and since both $\Phi(\Theta_{n_r}, \pi_{n_r})$, and $\Phi_{n_r+1}(\Theta_{n_r}, \pi_{n_r})$ have a well defined limit, then both limits must be the same, for almost every such ω. In conclusion, for almost every such ω, we can treat the sequence of limit points yielded by iterating this procedure (exactly as done in Propositions A.27 and 3.35), as a deterministic finite segment of boundary orbit, since the limit map Φ is the same iteration map as Φ_ρ, used in Section A.5: $\Phi^m(\Theta_m^*, \pi_m^*) = (\Theta_0^*, \pi_0^*)$, and $\Theta_m^* \in E_1$ for any m. We can see that π_0^* is on the eigenspace spanned by $e_{-1}(\Theta_0^*)$ for almost every such ω in eigencoordinates, exactly as in the conclusion of Propositions A.27 and 3.35, by formally changing the p, q notation into the Θ, π one. It is worth noting that we did not need to use the results about boundary sample paths, in particular Proposition 4.24 and Remark 4.25. It was enough to rely on the limit deterministic map's driving force, to show that $|\alpha_0| = 0$, where $\pi_0^* - \pi_{\Theta_0^*} = \alpha_0 e_0 + \beta_0 e_{-1}(\Theta_0^*)$. Working pointwise in ω, this leaves only two choices for β_0 on the eigenline: the values $\pm\beta$, such that $V(\Theta_0^*, \pi_{\Theta_0^*} \pm \beta e_{-1}(\Theta_0^*)) = \ell(\omega)$. ∎

Note that for almost every $\omega \in \{\ell = 0\}$, Proposition 4.29 follows trivially with $\beta = 0$, that is $\{(\Theta, \pi_\Theta) : \Theta \in \Sigma\}$.

Remark 4.30. *For every fixed $\omega \in \mathscr{E}_>$, consider any $\{n_k\}_{k\in\mathbb{N}}$ such that $\{\Theta_{n_k}\}$ converges, that is $\Theta_{n_k}(\omega) \longrightarrow \Theta_*(\omega) \in E_i$ for some $i \in \{1,2,3\}$. For almost every such ω, the set of accumulation points of $\{(\Theta_{n_k}, \pi_{n_k})\}$ and $\{(\Theta_{n_k+1}, \pi_{n_k+1})\}$ is a subset of $\{(\Theta_*, \pi_{\Theta_*} \pm \beta e_{-1}(\Theta_*)) :\ \beta = \beta(\omega) > 0,\ V(\Theta_*, \pi_{\Theta_*} \pm \beta e_{-1}(\Theta_*)) = \ell(\omega)\}$. Moreover, for almost every such ω for which also $\{\pi_{n_k}\}$ converges, $\{\pi_{n_k}\}$ and its shift $\{\pi_{n_k+1}\}$ asymptotically oscillate between $\pi_* = \pi_{\Theta_*} \pm \beta e_{-1}(\Theta_*)$ and $\hat{\pi}_* = \pi_{\Theta_*} \mp \beta e_{-1}(\Theta_*)$, that is, if the considered ω is such that $(\Theta_{n_k}, \pi_{n_k}) \longrightarrow (\Theta_*, \pi_*)$, then for almost every such ω, $(\Theta_{n_k+1}, \pi_{n_k+1}) \longrightarrow (\Theta_*, \hat{\pi}_*)$ as $k \longrightarrow \infty$.*

Proof. We exploit the continuity of the limit map Φ on $E_i \times \Sigma$. Starting with $\Theta_{n_k} \longrightarrow \Theta_* \in E_i$, to find the accumulation points we must extract any convergent $\{(\Theta_{n_{k_l}}, \pi_{n_{k_l}})\}_{l\in\mathbb{N}}$ and check its limit. To simplify the notation we relabel it as $\{(\Theta_{n_r}, \pi_{n_r})\}$. For the π-component's shifts we have that by Proposition 4.29, for almost every $\omega \in \mathscr{E}_>$, $\pi_* = \pi_{\Theta_*} \pm \beta e_{-1}(\Theta_*)$. Since

$$\pi_{\Theta_*} \mp \beta e_{-1}(\Theta_*) = \hat{\pi}_* := M_{\Theta_*}\pi_* = M_{\Theta_*}(\pi_{\Theta_*} \pm \beta e_{-1}(\Theta_*)),$$

(2.31), Lemma 4.1, and Remark 4.28 imply that, for almost every ω such that $\Theta_{n_r} \longrightarrow \Theta_* \in E_i$ and $\pi_{n_r} \longrightarrow \pi_*$, we have that $\pi_{n_r+1} = M_{\Theta_{n_r}}\pi_{n_r} + R_{n_r+1} \longrightarrow \hat{\pi}_*$. Therefore, like in Remark A.28, we start with an arbitrary subsubsequence $\{(\Theta_{n_r+1}(\omega), \pi_{n_r+1}(\omega))\}_{r\in\mathbb{N}}$ convergent to some $(\Theta, \hat{\pi}_*)$, both to be determined under the hypothesis given. Having that $\Theta_{n_r}(\omega) \longrightarrow \Theta_*(\omega) \in E_i$, by (2.32) we know that

$$\pi_{n_r} = \frac{(1-\rho_{n_r+1})\Theta_{n_r} - \Theta_{n_r+1}}{\rho_{n_r+1}} + \mathbf{1} - \pi_{n_r+1} \longrightarrow \frac{(1-\rho)\Theta_* - \Theta}{\rho} + \mathbf{1} - \hat{\pi}_* =: \pi_*.$$

We can repeat the argument above by applying Proposition 4.29 and (2.31), so as to show the claimed form of π_* and $\hat{\pi}_*$ for almost every ω considered.

For the Θ-component's shift, since $\pi_* + \hat{\pi}_* = 2\pi_{\Theta_*}$, and having already showed that for almost every ω considered $\pi_{n_r} \longrightarrow \pi_*$ and $\pi_{n_r+1} \longrightarrow \hat{\pi}_*$; if $\Theta_{n_r} \longrightarrow \Theta_*$, then by (2.32) it follows that

$$\begin{aligned}\Theta_{n_r+1} &= (1-\rho_{n_r+1})\Theta_{n_r} + \rho_{n_r+1}(\mathbf{1} - \pi_{n_r+1} - \pi_{n_r}) \longrightarrow (1-\rho)\Theta_* + \rho[\mathbf{1} - (\pi_* + \hat{\pi}_*)] \\ &= \Theta_*,\end{aligned}$$

as $\mathbf{1} - (\pi_* + \hat{\pi}_*) = \mathbf{1} - 2\pi_{\Theta_*} = \Theta_*$. Hence if we start with an arbitrary subsubsequence $\{(\Theta_{n_r+1}(\omega), \pi_{n_r+1}(\omega))\}_{r\in\mathbb{N}}$ convergent to some $(\Theta, \hat{\pi}_*)$ with only Θ left to be determined, it follows that $\Theta = \Theta_*$ for almost every ω considered. The rest of the claim is trivial by taking $n_r = n_k$. □

Note that Remark 4.30 trivially holds also for $\omega \in \mathscr{E}_0$, with $\ell = 0$, $\pi_* = \hat{\pi}_* = \pi_{\Theta_*}$, $\beta = 0$. The next corollary can be proved as its deterministic counterparts Corollary 3.37, the only difference being that some steps hold only almost surely.

Corollary 4.31. *For almost all $\omega \in \{\ell > 0\}$, $\Theta_{n+1} - \Theta_n \longrightarrow \mathbf{0}$ as $n \longrightarrow \infty$.*

Proof. Denote $d_n := \Theta_{n+1} - \Theta_n$. The claim is equivalent to showing that $d_n \longrightarrow \mathbf{0}$ for almost every such ω. Since d_n is bounded, if every convergent subsequence d_{n_k} converges to $\mathbf{0}$, then d_n converges to $\mathbf{0}$. Consider then a convergent subsequence $d_{n_k} \longrightarrow d$. We will show that by Lemma 4.21 and Remark 4.30, $d = \mathbf{0}$. There are in fact two cases, depending on whether $\{(\Theta_{n_k}, \pi_{n_k})\}$ converges or not.

- If ω is such that $\{(\Theta_{n_k}, \pi_{n_k})\}$ converges, it could be that Θ_{n_k} tends to a vertex or inside an edge. If it is a vertex, $\Theta_{n_k} \longrightarrow v \in V$ by Lemma 4.21 as $k \longrightarrow \infty$, $\Theta_{n_k+1} \longrightarrow v$ too for almost every such ω and therefore $d_{n_k} \longrightarrow \mathbf{0}$ for almost every such ω. If it is not a vertex $\Theta_{n_k} \longrightarrow \Theta_* \in E_i$ for some $i \in \{1, 2, 3\}$, then by Remark 4.30 $\Theta_{n_k+1} \longrightarrow \Theta_*$ too for almost every such ω. Hence $d_{n_k} \longrightarrow \mathbf{0}$. In any case, $d = \mathbf{0}$.

- If ω is such that $\{(\Theta_{n_k}, \pi_{n_k})\}$ does not converge, by boundedness again one can pick a subsubsequence $(\Theta_{n_{k_r}}, \pi_{n_{k_r}})$ that does so. Since $d_{n_k} \longrightarrow d$, also $d_{n_{k_r}} \longrightarrow d$ as $r \longrightarrow \infty$. However the previous argument applies to $d_{n_{k_r}}$ since $(\Theta_{n_{k_r}}, \pi_{n_{k_r}})$ converges, thus falling into the previous case, and therefore $d_{n_{k_r}}$ vanishes for almost every such ω as $r \longrightarrow \infty$. Then for almost every such ω, $\mathbf{0} = d = \lim_{k\longrightarrow\infty} d_{n_k} = \lim_{r\longrightarrow\infty} d_{n_{k_r}}$.

□

Note that for almost every $\omega \in \{\ell = 0\}$, Corollary 4.31 follows trivially from (4.4) and Lemma 4.1.

Remark 4.32. *For almost all $\omega \in \mathscr{E}_>$, by Remark 4.3 and Corollary 4.31 there exists a subsequence of $\{\Theta_n\}$ bounded away from V.*

The following claim holds trivially also on $\{\ell = 0\}$.

Corollary 4.33. *For almost every $\omega \in \{\ell > 0\}$, as $n \longrightarrow \infty$, $\alpha_n \longrightarrow 0$ and $|\beta_n| \longrightarrow \ell/2$.*

Proof. We first show that α_n vanishes. Considering that in eigencoordinates (4.11) holds, which can be rearranged as

$$\Theta_{n+1} - \Theta_n = -\rho_{n+1}(\alpha_n(1 + \lambda_0(\Theta_n))e_0(\Theta_n) + \beta_n(1 + \lambda_{-1}(\Theta_n))e_{-1}(\Theta_n) + R_{n+1}),$$

and considering that for almost every $\omega \in \{\ell > 0\}$ by Remark 4.20, $\Theta_n \longrightarrow \partial\Sigma$ and, as a result, by Lemma 3.19 (f), $1 + \lambda_{-1}(\Theta_n) \longrightarrow 0$ and $1 + \lambda_0(\Theta_n) \longrightarrow 1$, and by Lemma 4.1, $R_{n+1} \longrightarrow \mathbf{0}$; by Corollary 4.31 and Lemma 3.19 (h) it follows that for almost every such ω, $\alpha_n \longrightarrow 0$ as $n \longrightarrow \infty$.

We now show that $|\beta_n(\omega)| \longrightarrow \ell(\omega)/2$ for almost every $\omega \in \{\ell > 0\}$. Since $|\beta_n|$ is bounded, consider any convergent subsequence $|\beta_{n_j}(\omega)| \longrightarrow \ell'(\omega)$. Assume by contradiction that $\ell'(\omega) \neq \ell(\omega)/2$ for all ω belonging to a nonnegligible event in $\{\ell > 0\}$. Since Θ_{n_j} is bounded, extract a convergent subsubsequence $\{\Theta_{n_{j_l}}\}_{l\in\mathbb{N}}$. Relabel it with $\{n_k\}$ for simplicity. Since for almost every ω considered the potential limit along this orbit is $\ell(\omega) > 0$, we have $\Theta_{n_k}(\omega) \longrightarrow \Theta_*(\omega) \in \partial\Sigma$ by Remark 4.3, and $|\beta_{n_k}(\omega)| \longrightarrow \ell'(\omega) \neq \ell(\omega)/2$ by assumption. Recall that by Remark 4.20, $\mathbb{P}(\{\ell > 0\} \setminus (\mathscr{E}_> \cup \mathcal{D}_>)) = 0$. So we only need to consider two events: the one for which ω is such that $\Theta_*(\omega) \in V$ and the one for which $\Theta_*(\omega)$ belongs to $\partial\Sigma \setminus V$. If ω is such that $\Theta_*(\omega) \in V$, $|\beta_{n_k}(\omega)| \longrightarrow \ell(\omega)/2$ for almost every such ω by Lemma 4.21. Since $\{n_k\}$ is a subsequence of $\{n_j\}$, and $|\beta_{n_j}(\omega)| \longrightarrow \ell'(\omega)$, this leaves only a negligible event, on which we could escape a contradiction with the assumption $\ell'(\omega) \neq \ell(\omega)/2$. If ω is such that $\Theta_*(\omega)$ belongs to E_i for some $i \in \{1, 2, 3\}$ (recall that $\partial\Sigma \setminus V$ is partitioned into E_1, E_2, E_3), by symmetry, without loss of generality, consider any such ω belonging to the event for which $i = 1$. We show the argument explicitly only on this event, as on the events, for which $i = 2$ and $i = 3$, the argument is the same, upon interchanging the coordinates according to the symmetry of the model. For ω such that $i = 1$ fixed, by the smoothness of the eigenvectors proved in Lemma 3.19 (h), it is known that $e_0(\Theta_{n_k}) \longrightarrow (0,1,-1) = e_0(\Theta_*)$ with $e_0(\Theta_{n_k}) - e_0(\Theta_*) = \mathcal{O}(\boldsymbol{\Theta_{n_k} - \Theta_*})$, and $e_{-1}(\Theta_{n_k}) \longrightarrow (-1, 1-\Theta_*^{(2)}, \Theta_*^{(2)}) =: e_{-1}(\Theta_*)$ with $e_{-1}(\Theta_{n_k}) - e_{-1}(\Theta_*) = \mathcal{O}(\boldsymbol{\Theta_{n_k} - \Theta_*})$. Thus

$$\begin{aligned}&\|\alpha_{n_k} e_0(\Theta_{n_k}) + \beta_{n_k} e_{-1}(\Theta_{n_k})\|_1 = \|\alpha_{n_k}[e_0(\Theta_{n_k}) - e_0(\Theta_*)] + \beta_{n_k}[e_{-1}(\Theta_{n_k}) - e_{-1}(\Theta_*)] \\ &\quad + \alpha_{n_k} e_0(\Theta_*) + \beta_{n_k} e_{-1}(\Theta_*)\|_1 = \|\mathcal{O}(\boldsymbol{\Theta_{n_k} - \Theta_*}) + \alpha_{n_k} e_0(\Theta_*) + \beta_{n_k} e_{-1}(\Theta_*)\|_1 \\ &\quad = \|o_{\boldsymbol{\omega}}(\mathbf{1}) + \beta_{n_k} e_{-1}(\Theta_*)\|_1,\end{aligned}$$

since $\alpha_n \longrightarrow 0$ for almost every ω considered. Thus

$$V(\Theta_{n_k}, \pi_{n_k}) = \|o_{\boldsymbol{\omega}}(\mathbf{1}) + \beta_{n_k} e_{-1}(\Theta_*)\|_1.$$

Since ω was assumed such that $|\beta_{n_j}(\omega)| \longrightarrow \ell'(\omega) \neq \ell(\omega)/2$, being $\{n_k\}$ a subsequence of $\{n_j\}$, also $|\beta_{n_k}(\omega)| \longrightarrow \ell'(\omega)$, and this would imply that

$$0 \leftarrow V(\Theta_{n_k}, \pi_{n_k}) - \ell = \|o_{\boldsymbol{\omega}}(\mathbf{1}) + \beta_{n_k} e_{-1}(\Theta_*)\|_1 - \ell = |\beta_{n_k}| \|o_{\boldsymbol{\omega}}(\mathbf{1}) + e_{-1}(\Theta_*)\|_1 - \ell \to 2\ell' - \ell$$

because $\ell' \neq 0$ (otherwise $V(\Theta_{n_k}, \pi_{n_k}) \longrightarrow 0$, as $k \longrightarrow \infty$, but $\omega \in \{\ell > 0\}$). Thus for almost every ω considered, $2\ell'(\omega) - \ell(\omega) = 0$, or equivalently $\ell'(\omega) = \ell(\omega)/2$, which leaves only a negligible event on which we could escape a contradiction with the assumption $\ell'(\omega) \neq \ell(\omega)/2$. We have thus shown that for almost every $\omega \in \{\ell > 0\}$, $\ell'(\omega) = \ell(\omega)/2$, where $\ell'(\omega)$ is the limit of an arbitrary convergent subsequence $\{|\beta_{n_j}|\}$. Since for almost every $\omega \in \{\ell > 0\}$, any convergent subsequence $\{|\beta_{n_j}(\omega)|\}$ of the bounded sequence $\{|\beta_n(\omega)|\}$ tends to $\ell(\omega)/2$, it follows that for almost every such ω, $|\beta_n(\omega)| \longrightarrow \ell(\omega)/2$. □

For almost every $\omega \in \mathscr{E}_>$ fixed, by Remark 4.32 it follows that there is a subsequence $\{\Theta_{n_j}\}_{j\in\mathbb{N}}$ bounded away from the vertices. By boundedness, this implies the existence of a subsubsequence $\{\Theta_{n_{j_l}}\}_{l\in\mathbb{N}}$ (relabelled with n_k for simplicity) such that $\Theta_{n_k}(\omega) \longrightarrow \Theta_*(\omega) \in E_i$ for some $i \in \{1,2,3\}$, for almost every $\omega \in \mathscr{E}_>$. Define

$$\mathscr{E}_>^{(i)} := \{\omega \in \mathscr{E}_> : \exists \{n_k\}_{k\in\mathbb{N}},\ \Theta_{n_k}(\omega) \longrightarrow \Theta_*(\omega) \in E_i \text{ as } k \longrightarrow \infty\}.$$

Thus

$$\mathbb{P}\left(\mathscr{E}_> \setminus \bigcup_{i=1}^{3} \mathscr{E}_>^{(i)}\right) = 0.$$

By symmetry, similarly to what done on $\mathscr{E}_0$, without loss of generality, we will show arguments that apply almost surely on $\mathscr{E}_>$ only on $\mathscr{E}_0^{(1)}$. The following is the first such example.

Remark 4.34. *For almost every $\omega \in \mathscr{E}_>$, $\ell < 1$.*

Proof. Without loss of generality, let $\omega \in \mathscr{E}_>^{(1)}$, so that there is $\{n_k\}$ such that $\Theta_{n_k}(\omega) \longrightarrow \Theta_*(\omega) \in E_1$. By Remark 4.30 and Corollary 4.33, for almost every such ω, the limit points of $\{\pi_{n_k}\}$ and $\{\pi_{n_k+1}\}$ are, without loss of generality, $\pi_* = \pi_{\Theta_*} - \beta e_{-1}(\Theta_*)$ and $\hat{\pi}_* = \pi_{\Theta_*} + \beta e_{-1}(\Theta_*)$, with $\beta = \ell(\omega)/2$. Note that $\pi_{\Theta_*}^{(1)} = 1/2$ and that $\pi_{\Theta_{n_k+1}}^{(1)} = (1 - \Theta_{n_k+1}^{(1)})/2 < 1/2$ for all $k \in \mathbb{N}$, since $\Theta_{n_k+1}^{(1)} > 0$ by definition of the model. Note also that for all k large enough, due to the nonzero angle that the eigendirection of $e_{-1}(\Theta_*)$ forms with E_2 and E_3, and due to $\beta > 0$ and the elementary geometry of the simplex (see Figure 3.5), we have that $\pi_{n_k}^{(1)} - \pi_{\Theta_{n_k}}^{(1)} > 0$, $\pi_{n_k}^{(2)} - \pi_{\Theta_{n_k}}^{(2)} < 0$, $-(\pi_{n_k}^{(1)} - \pi_{\Theta_{n_k}}^{(1)} + \pi_{n_k}^{(2)} - \pi_{\Theta_{n_k}}^{(2)}) = \pi_{n_k}^{(3)} - \pi_{\Theta_{n_k}}^{(3)} < 0$, and $\pi_{n_k+1}^{(1)} - \pi_{\Theta_{n_k+1}}^{(1)} < 0$,

$\pi^{(2)}_{n_k+1} - \pi^{(2)}_{\Theta_{n_k+1}} > 0$, $-(\pi^{(1)}_{n_k+1} - \pi^{(1)}_{\Theta_{n_k+1}} + \pi^{(2)}_{n_k+1} - \pi^{(2)}_{\Theta_{n_k+1}}) = \pi^{(3)}_{n_k+1} - \pi^{(3)}_{\Theta_{n_k+1}} > 0$. Therefore,

$$\begin{aligned}\|\pi_{n_k+1} - \pi_{\Theta_{n_k+1}}\|_1 = \sum_{i\in\{1,2,3\}} |\pi^{(i)}_{n_k+1} - \pi^{(i)}_{\Theta_{n_k+1}}| = -(\pi^{(1)}_{n_k+1} - \pi^{(1)}_{\Theta_{n_k+1}}) + \pi^{(2)}_{n_k+1} - \pi^{(2)}_{\Theta_{n_k+1}} \\ - (\pi^{(1)}_{n_k+1} - \pi^{(1)}_{\Theta_{n_k+1}} + \pi^{(2)}_{n_k+1} - \pi^{(2)}_{\Theta_{n_k+1}}) = 2(\pi^{(1)}_{\Theta_{n_k+1}} - \pi^{(1)}_{n_k+1}) < 2 \cdot \frac{1}{2} = 1.\end{aligned}$$

By Lemma 4.1 and Corollary 4.33, for almost every such ω, we can fix $1 < \nu < \sqrt{\mu}$ and let $m = n_{\bar{k}}$, with $\bar{k}$ large enough, such that for all $n \geq m$, $\|R_{n+1}\|_1 < 3\nu^{-n-1}$ and $\pi_n \neq \pi_{\Theta_n}$; then for all $n \geq \bar{m} \geq m$,

$$|V(\Theta_{n+1}, \pi_{n+1}) - \overline{V}(\Theta_{n+1}, \pi_{n+1})| \leq \frac{3}{\nu - 1} \frac{1}{\nu^{\bar{m}}}$$

by Remark 4.4. Lemma A.8, rephrased in Θ, π notation instead of p, q notation and with a time-dependent parameter ρ_{n+1} instead of ρ (it is trivial to verify that these are only formal changes and do not alter the substance of the argument of the lemma), applies to $\overline{V}(\Theta_{n+1}, \pi_{n+1})$. Indeed $\Theta_n \notin \partial\Sigma$ for all n, according to the model. Thus we can conclude that $\overline{V}(\Theta_n, \pi_n)$ is strictly decreasing for all $n \geq m$, and is subunitary at time m (recall that at time m, we define $\overline{V}(\Theta_m, \pi_m) := V(\Theta_m, \pi_m)$, as per definition in Remark 4.4). We show that this set-up implies that for almost every considered ω, $V(\Theta_n(\omega), \pi_n(\omega))$ can only converge to subunitary values of $\ell(\omega)$.

Assume by contradiction that $\ell(\omega) \geq 1$ for all ω belonging to a nonnegligible subset of $\mathscr{E}^{(1)}_{>}$. For every such ω in the nonnegligible event, define

$$0 < \delta < \frac{6}{\nu - 1}$$

small enough, such that $V(\Theta_m, \pi_m) \leq 1 - \delta$. The parameter δ exists by the previous part of this argument. By monotonicity, for every such ω, $\overline{V}(\Theta_{n+1}, \pi_{n+1}) < 1 - \delta$, for all $n \geq m$. There exists $m' > m$ large enough such that, for almost every such ω, for all $n \geq m'$, $|V(\Theta_{n+1}, \pi_{n+1}) - \overline{V}(\Theta_{n+1}, \pi_{n+1})| > \delta/2$, since if $\ell \geq 1$, there will be m' such that for all $n \geq m'$, $V(\Theta_{n+1}, \pi_{n+1}) > 1 - \delta/2$ for almost every ω considered. Let

$$\bar{m} := \max\left\{ m', \frac{\log \frac{6}{(\nu-1)\delta}}{\log \nu} \right\}.$$

By construction of $\bar{m}$, for all $n \geq \bar{m}$,

$$|V(\Theta_{n+1}, \pi_{n+1}) - \overline{V}(\Theta_{n+1}, \pi_{n+1})| \leq \frac{3}{\nu - 1} \frac{1}{\nu^{\bar{m}}} \leq \frac{\delta}{2},$$

since

$$\bar{m} \geq \frac{\log \frac{6}{(\nu-1)\delta}}{\log \nu},$$

and therefore

$$\nu^{\bar{m}} \geq \nu^{\frac{\log \frac{6}{(\nu-1)\delta}}{\log \nu}} = \left(e^{\log \nu}\right)^{\frac{\log \frac{6}{(\nu-1)\delta}}{\log \nu}} = \frac{6}{(\nu - 1)\delta}.$$

However $\bar{m} \geq m'$, so the inequality should be reversed and strict, which leaves us with a contradiction, which we could only escape on a negligible event. Hence the subset, assumed nonnegligible, must actually be negligible. □

4.5.3 Convergence with regular initial conditions

We proceed with considering, without loss of generality, only $\omega \in \mathscr{E}_{>}^{(1)}$, for which, following the same line of reasoning at the start of Section 3.5.3, there exists $\{n_k\}_{k\in\mathbb{N}}$ such that $\{\Theta_{n_k}(\omega)\}$ is bounded away from the vertices, and for some $\Theta_*(\omega) \in E_1$, $\Theta_{n_k}(\omega) \longrightarrow \Theta_*(\omega)$ as $k \longrightarrow \infty$. Because of the almost sure structure of the accumulation points of the sample paths $\{(\Theta_n(\omega), \pi_n(\omega))\}$ proved in Proposition 4.29, the properties of the shift of $\{(\Theta_{n_k}(\omega), \pi_{n_k}(\omega))\}$ shown in Remark 4.30, and Corollary 4.33, by the geometry of the simplex and Remark 4.34, it will be possible to fix such a subsequence, so that

$$\pi_{n_k}(\omega) \longrightarrow \pi_*(\omega) := \pi_{\Theta_*(\omega)} + \frac{\ell}{2} e_{-1}(\Theta_*(\omega))$$

as $k \longrightarrow \infty$. Let us fix $\delta = \delta(\omega) > 0$ small enough, so that $\delta < \Theta_*^{(2)}(\omega) < 1-\delta$, $\delta < \pi_*^{(1)}(\omega) < {}^1\!/\!_2 - \delta$, $\pi_*^{(3)}(\omega) > \delta$. There will be an $\varepsilon' = \varepsilon'(\omega)$ small enough and $K = K(\omega)$ large enough such that, having defined $m := n_K$, if $\Theta_m^{(1)}(\omega), |\alpha_m(\omega)|, ||\beta_m(\omega)| - {}^{\ell(\omega)}\!/\!_2| \leq \varepsilon'$, then $\delta \leq \Theta_m^{(2)}(\omega) \leq 1-\delta$, $\delta \leq \pi_m^{(1)}(\omega) \leq {}^1\!/\!_2 - \delta$, $\delta \leq \pi_m^{(2)}(\omega) \leq 1-\delta$, $\pi_m^{(3)}(\omega) \geq \delta$ and $|\alpha_n|, ||\beta_n| - {}^\ell\!/\!_2| \leq \varepsilon'$ for all $n \geq m$ for almost every ω. Recall that α_n and β_n refer to the eigencoordinates of $\pi_n - \pi_{\Theta_n}$. Moreover, since $\|\pi - \pi_\Theta\|_1 \leq 2$ (due to the diameter of the simplex in 1-norm) by Lemma 3.19 (g, h) there is a constant $B > 1$ large enough such that $|\beta| < B$, and then by (4.11) for any δ fixed small enough, there is a $\varepsilon < \delta$ small enough (to be further restricted) such that,

$$\begin{aligned}\|\hat{\Theta} - \Theta\|_1 &\leq 3|\alpha| + B\frac{\|e_{-1}(\Theta)\|_1}{2}(1 + \lambda_{-1}(\Theta)) + \|\hat{R}\|_1 \\ &\leq 3|\alpha| + B\Theta_1\|e_{-1}(\Theta)\|_1(1 + \mathcal{O}(\Theta_1)) + \|\hat{R}\|_1 \leq 3B(|\alpha| + \Theta_1) + \|\hat{R}\|_1\end{aligned}$$

for all $\Theta_1 < \varepsilon$. Recalling that $\tilde{\rho} := (1+\rho)/2$, define

$$R := 1 + \tilde{\rho}\left(\frac{2}{\delta} - 1\right)$$

and

$$\varepsilon' < \min\left\{\frac{\varepsilon}{R}, \frac{\varepsilon}{2(6B+1)}\right\}$$

(ε' is to be further restricted too). Finally define

$$\mathcal{K}^\ell_{\varepsilon',\delta} := \left\{(\Theta, \pi) \in \Sigma^2 : \ 0 < \Theta^{(1)},\ |\alpha|,\ \left||\beta| - \frac{\ell}{2}\right| \leq \varepsilon',\ \delta \leq \Theta^{(2)} \leq 1-\delta\right\},$$

and similarly $\mathcal{K}^\ell_{\varepsilon,\delta'}$. Note that there is a large enough $\bar{m} = \bar{m}(\omega) \in \mathbb{N}$ such that $3\nu^{-\bar{m}} < \varepsilon'$ and for all $n \geq \bar{m}$, for all $i \in \{1,2,3\}$, $|R_{n+1}^{(i)}| < \nu^{-n-1}$ by Lemma 4.1, $0 < {}^\rho\!/\!_2 < \rho_{n+1} < \tilde{\rho} < 1$, ${}^{\rho_{n+1}}\!/\!_{\rho_n} < 3/2$ (as it tends to 1) and $\rho_{n+1} - \rho_n < \varepsilon$ (as it converges). Then by additionally requiring K to be large enough such that $m(\omega) := n_{K(\omega)} \geq \bar{m}(\omega)$, we ensure that for almost every $\omega \in \mathscr{E}_{>}^{(1)}$, $(\Theta_m(\omega), \pi_m(\omega)) \in \mathcal{K}^\ell_{\varepsilon',\delta}$ with $\pi_m^{(1)} < {}^1\!/\!_2 - \delta$ and for all $n \geq m$, for all $i \in \{1,2,3\}$, $|R_{n+1}^{(i)}| < \nu^{-n-1} < {}^{\varepsilon'}\!/\!_3$. By construction

$$\|\Theta_{m+1} - \Theta_m\|_1 < (6B+1)\varepsilon' < \frac{\varepsilon}{2}, \tag{4.35}$$

thus ensuring, having defined $\delta' := {}^\delta\!/\!_2$, that

$$\Theta_{m+1}^{(1)} \leq \Theta_m^{(1)} + \|\Theta_{m+1} - \Theta_m\|_1 < \varepsilon' + \frac{\varepsilon}{2} < \varepsilon, \tag{4.36}$$

$$\Theta_{m+1}^{(2)} \leq \Theta_m^{(2)} + \|\Theta_{m+1} - \Theta_m\|_1 < 1 - \delta + \frac{\varepsilon}{2} < 1 - \delta', \tag{4.37}$$

$$\Theta_{m+1}^{(2)} \geq \Theta_m^{(2)} - \|\Theta_{m+1} - \Theta_m\|_1 > \delta - \frac{\varepsilon}{2} > \delta'. \tag{4.38}$$

When there is no ambiguity, we will often simplify the notation as $\mathcal{K} := \mathcal{K}^{\ell}_{\varepsilon',\delta} \subset \mathcal{K}^* := \mathcal{K}^{\ell}_{\varepsilon,\delta'}$ and omit ω as customary. Through these parameters, these sets are clearly defined pointwise for almost every $\omega \in \mathscr{E}^{(1)}_{>}$. The intuitive picture to always keep in mind is Figure 3.5.

On $\mathscr{E}^{(1)}_{>}$ we define the stopping time

$$\sigma := \inf \left\{ n > m+1 : \ \Theta^{(2)}_n \notin \left[\frac{\delta}{2}, 1 - \frac{\delta}{2}\right] \right\} \in \mathbb{N} \cup \infty.$$

We derive some iterative formulas and bounds.

Remark 4.35. *For almost every $\omega \in \mathscr{E}^{(1)}_{>}$ and $m \leq n < \sigma$,*

$$\Theta^{(1)}_{n+1} < R\Theta^{(1)}_n + \frac{1}{\nu^{n+1}}.$$

Proof. Since for any $m \leq n < \sigma$,

$$1 - \pi^{(1)}_{n+1} - \pi^{(1)}_n = \Theta^{(1)}_n \left(\frac{\pi^{(2)}_n}{1 - \Theta^{(2)}_n} + \frac{\pi^{(3)}_n}{1 - \Theta^{(3)}_n} \right) - R^{(1)}_{n+1} < \frac{2}{\delta}(1 - \pi^{(1)}_n) - R^{(1)}_{n+1} \leq \frac{2}{\delta} - R^{(1)}_{n+1},$$

hence, being $\delta < 1/2$,

$$\Theta^{(1)}_{n+1} \leq (1 - \rho_{n+1})\Theta^{(1)}_n + \frac{2\rho_{n+1}}{\delta}\Theta^{(1)}_n + \rho_{n+1}|R^{(1)}_{n+1}| = \left[1 + \rho_{n+1}\left(\frac{2}{\delta} - 1\right)\right]\Theta^{(1)}_n + \rho_{n+1}|R^{(1)}_{n+1}|.$$

Hence the claim follows by definition of m. □

The next two remarks establish two more iterative formulas.

Remark 4.36. *For any $n \geq 0$,*

$$\Theta^{(1)}_{n+2} = \Theta^{(1)}_{n+1} \left\{ (1 - \rho_{n+2}) + \rho_{n+2} \left[(1 - \pi^{(1)}_{n+1})\vartheta_{n+1} - \Theta^{(1)}_{n+1}\rho_{n+1}\vartheta_{n+1} + \Theta^{(1)}_n \vartheta''_{n+1} + r_{n+1} \right] \right\} - \rho_{n+2} R^{(1)}_{n+2},$$

where ϑ_{n+1} and ϑ'_{n+1} are defined as in (3.33) and (3.34) in Remark 3.42 (with the usual formal change from p, q notation to Θ, π notation),

$$\vartheta''_{n+1} := \frac{1 - \rho_{n+1}}{\rho_{n+1}} \vartheta_{n+1} + \vartheta'_{n+1} \tag{4.39}$$

and

$$r_{n+1} := -R^{(1)}_{n+1}\vartheta_{n+1} + \frac{R^{(2)}_{n+1}}{1 - \Theta^{(2)}_{n+1}} + \frac{R^{(3)}_{n+1}}{\Theta^{(1)}_{n+1} + \Theta^{(2)}_{n+1}}. \tag{4.40}$$

Proof. For any $n \geq 0$,

$$\begin{aligned}\Theta^{(1)}_{n+2} &= (1 - \rho_{n+2})\Theta^{(1)}_{n+1} + \rho_{n+2}\Theta^{(1)}_{n+1} \left(\frac{\pi^{(2)}_{n+1}}{1 - \Theta^{(2)}_{n+1}} + \frac{\pi^{(3)}_{n+1}}{\Theta^{(1)}_{n+1} + \Theta^{(2)}_{n+1}} \right) - \rho_{n+2} R^{(1)}_{n+2} \\ &= \Theta^{(1)}_{n+1} \left\{ (1 - \rho_{n+2}) + \rho_{n+2} \left(\pi^{(1)}_n \vartheta_{n+1} + \Theta^{(1)}_n \vartheta'_{n+1} + \frac{R^{(2)}_{n+1}}{1 - \Theta^{(2)}_{n+1}} + \frac{R^{(3)}_{n+1}}{\Theta^{(1)}_{n+1} + \Theta^{(2)}_{n+1}} \right) \right\} \\ &\quad - \rho_{n+2} R^{(1)}_{n+2},\end{aligned}$$

by the second step (in which we rearranged the factor in the brackets) in the proof of Remark 3.42, and since by (2.32)

$$\pi^{(1)}_n = (1 - \pi^{(1)}_{n+1}) + \frac{1 - \rho_{n+1}}{\rho_{n+1}} \Theta^{(1)}_n - \frac{1}{\rho_{n+1}} \Theta^{(1)}_{n+1} - R^{(1)}_{n+1},$$

the claim follows. □

Remark 4.37. *For any $n \geq 0$,*

$$\Theta^{(2)}_{n+2} - \Theta^{(2)}_{n+1} = -\rho_{n+2}\pi^{(1)}_n \left(\Theta^{(2)}_{n+1} - \Theta^{(2)}_n + \xi_{n+1} + \xi'_{n+1} - \eta'_{n+1} - \eta_n + \eta''_{n+1} + \eta'''_{n+1} + r'_{n+1}\right)$$

where η_n, η'_{n+1}, η''_{n+1} and η'''_{n+1} are defined as (3.35) to (3.38) in Remark 3.43 (with the usual formal change from p, q notation to Θ, π notation) and

$$\xi_{n+1} := \Theta^{(1)}_{n+1} \frac{\Theta^{(2)}_{n+1}}{1 - \Theta^{(1)}_{n+1}} \left(1 + \frac{\frac{1}{\rho_{n+1}} - 1}{\pi^{(1)}_n}\right)$$

$$\xi'_{n+1} := \Theta^{(1)}_n \Theta^{(2)}_{n+1} \frac{1 - \frac{1}{\rho_{n+1}}}{\pi^{(1)}_n (1 - \Theta^{(1)}_{n+1})}$$

$$r'_{n+1} := -\frac{R^{(3)}_{n+1}}{\pi^{(1)}_n (\Theta^{(1)}_{n+1} + \Theta^{(2)}_{n+1})} + \frac{R^{(2)}_{n+2}}{\pi^{(1)}_n \Theta^{(2)}_{n+1}}.$$

Proof. For all $n \geq 0$,

$$\Theta^{(2)}_{n+2} - \Theta^{(2)}_{n+1} = \rho_{n+2}\left(\frac{\Theta^{(2)}_{n+1}}{1-\Theta^{(1)}_{n+1}}\pi^{(1)}_{n+1} + \frac{\Theta^{(2)}_{n+1}}{\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1}}\pi^{(3)}_{n+1} - R^{(2)}_{n+2} - \Theta^{(2)}_{n+1}\right)$$
$$= \rho_{n+2}\Bigg[\Theta^{(2)}_{n+1}\left(\frac{\pi^{(1)}_{n+1}}{1-\Theta^{(1)}_{n+1}} - 1 - \frac{R^{(2)}_{n+2}}{\Theta^{(2)}_{n+1}}\right) + \frac{\Theta^{(2)}_{n+1}}{\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1}}\Bigg(\pi^{(1)}_n\frac{\Theta^{(2)}_n}{1-\Theta^{(1)}_n} + \pi^{(2)}_n\frac{\Theta^{(1)}_n}{1-\Theta^{(2)}_n}$$
$$+ R^{(3)}_{n+1}\Bigg)\Bigg] = \rho_{n+2}\Bigg\{\Theta^{(2)}_{n+1}\Bigg[\frac{\pi^{(1)}_{n+1}}{1-\Theta^{(1)}_{n+1}} - 1 + \frac{\Theta^{(1)}_n}{(\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1})(1-\Theta^{(2)}_n)}\pi^{(2)}_n + \frac{R^{(3)}_{n+1}}{\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1}}$$
$$- \frac{R^{(2)}_{n+2}}{\Theta^{(2)}_{n+1}}\Bigg] + \Theta^{(2)}_n\frac{\pi^{(1)}_n}{1-\Theta^{(1)}_n}\frac{\Theta^{(2)}_{n+1}}{\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1}}\Bigg\} = \rho_{n+2}\Bigg\{\Theta^{(2)}_{n+1}\Bigg[-\pi^{(1)}_n - \Theta^{(1)}_{n+1}\frac{\pi^{(1)}_n + \frac{1}{\rho_{n+1}} - 1}{1-\Theta^{(1)}_{n+1}}$$
$$- \Theta^{(1)}_n\frac{1-\frac{1}{\rho_{n+1}}}{1-\Theta^{(1)}_{n+1}} + \frac{\Theta^{(1)}_n}{(\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1})(1-\Theta^{(2)}_n)}\pi^{(2)}_n + \frac{R^{(3)}_{n+1}}{\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1}} - \frac{R^{(2)}_{n+2}}{\Theta^{(2)}_{n+1}}\Bigg]$$
$$+ \Theta^{(2)}_n\pi^{(1)}_n\left(1+\frac{\Theta^{(1)}_n}{1-\Theta^{(1)}_n}\right)\left(1-\frac{\Theta^{(1)}_{n+1}}{\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1}}\right)\Bigg\} = -\rho_{n+2}\pi^{(1)}_n\Bigg[\Theta^{(2)}_{n+1} - \Theta^{(2)}_n$$
$$+ \Theta^{(1)}_{n+1}\frac{\Theta^{(2)}_{n+1}}{1-\Theta^{(1)}_{n+1}}\left(1+\frac{\frac{1}{\rho_{n+1}}-1}{\pi^{(1)}_n}\right) + \Theta^{(1)}_n\Theta^{(2)}_{n+1}\frac{1-\frac{1}{\rho_{n+1}}}{\pi^{(1)}_n(1-\Theta^{(1)}_{n+1})}$$
$$- \Theta^{(1)}_n\frac{\Theta^{(2)}_{n+1}}{(\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1})(1-\Theta^{(2)}_n)}\frac{\pi^{(2)}_n}{\pi^{(1)}_n} - \Theta^{(2)}_n\frac{\Theta^{(1)}_n}{1-\Theta^{(1)}_n} + \Theta^{(1)}_{n+1}\frac{\Theta^{(2)}_n}{\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1}}$$
$$+ \Theta^{(2)}_n\frac{\Theta^{(1)}_n}{1-\Theta^{(1)}_n}\frac{\Theta^{(1)}_{n+1}}{\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1}} - \frac{R^{(3)}_{n+1}}{\pi^{(1)}_n(\Theta^{(1)}_{n+1}+\Theta^{(2)}_{n+1})} + \frac{R^{(2)}_{n+2}}{\pi^{(1)}_n\Theta^{(2)}_{n+1}}\Bigg],$$

since by (2.32),

$$1 - \pi^{(1)}_{n+1} = \pi^{(1)}_n + \left(1 - \frac{1}{\rho_{n+1}}\right)\Theta^{(1)}_n + \frac{\Theta^{(1)}_{n+1}}{\rho_{n+1}},$$

and therefore

$$\frac{\pi^{(1)}_{n+1}}{1-\Theta^{(1)}_{n+1}} - 1 = -\frac{1-\pi^{(1)}_{n+1}}{1-\Theta^{(1)}_{n+1}} + \frac{\Theta^{(1)}_{n+1}}{1-\Theta_{n+1^{(1)}}} = -\frac{\pi^{(1)}_n}{1-\Theta^{(1)}_{n+1}} - \frac{1-\frac{1}{\rho_{n+1}}}{1-\Theta^{(1)}_{n+1}}\Theta^{(1)}_n$$
$$- \frac{\Theta^{(1)}_{n+1}}{\rho_{n+1}(1-\Theta^{(1)}_{n+1})} + \frac{\Theta^{(1)}_{n+1}}{1-\Theta^{(1)}_{n+1}} = -\pi^{(1)}_n - \Theta^{(1)}_{n+1}\frac{\pi^{(1)}_n + \frac{1}{\rho_{n+1}} - 1}{1-\Theta^{(1)}_{n+1}} - \Theta^{(1)}_n\frac{1-\frac{1}{\rho_{n+1}}}{1-\Theta^{(1)}_{n+1}}.$$

Thus the claim follows. ∎

Fix a small enough

$$\delta < \min\left\{\frac{4}{105}\rho^3, 2\sqrt[3]{\frac{1-\tilde{\rho}}{3+\frac{8}{\rho}}}\right\}$$

and define

$$\gamma = \gamma(\delta') \coloneqq \sqrt{\max\left\{1-\rho^2(\delta')^2+\frac{35}{\rho}(\delta')^3, \frac{1}{\nu}\right\}}.$$

The constant γ is positive subunitary, since $\gamma' \coloneqq \gamma^2 = \max\{1-\rho^2(\delta')^2+35(\delta')^3/\rho, \nu^{-1}\}$ is. Indeed $0 < \nu^{-1} < 1$ and $0 < 1-\rho^2(\delta')^2+35(\delta')^3/\rho < 1$ since $-\rho x^2+35x^3/\rho$ is negative monotone decreasing on $(0, 2\rho^3/105)$ and it takes minimum at $2\rho^3/105$ of value $-212\rho^8/105^3 > -{}^{212}/_{105^3}$. Recall that

$$R \coloneqq 1+\tilde{\rho}\left(\frac{2}{\delta}-1\right).$$

Further require

$$\varepsilon < \min\left\{(\delta')^5, \frac{\rho\delta'(1-\gamma^2)}{8\left(R+\frac{1}{1-\gamma^2}\right)}, \frac{\rho\delta'}{4\left(1+\frac{R+\frac{1}{1-\gamma^2}}{1-\gamma}\right)}\right\}.$$

Define also

$$D \coloneqq 2\left[R\left(R+\frac{1}{1-\gamma^2}\right)+1\right]\left(1+\frac{1}{\delta'\rho}\right)+\left(R+\frac{1}{1-\gamma^2}\right)\left[2+\frac{1}{\delta'}\left(\frac{2}{\rho}-1\right)+\frac{1}{(\delta')^2}\right] + \frac{2}{(\delta')^2}$$

$$\tilde{D} = 4+\frac{1}{\delta'}\left(\frac{4}{\rho}-1\right)+\frac{3}{(\delta')^2}$$

and let Γ be a constant such that

$$0 < \Gamma < \frac{\delta'}{\tilde{D}(1-\delta')},$$

$\lambda \coloneqq \max\{\gamma, \tilde{\rho}\}$ and

$$\tilde{R} \coloneqq R^2+\frac{R}{1-\gamma^2}+1.$$

Finally further restrict

$$\varepsilon' < \min\left\{\frac{\varepsilon}{R+\frac{3}{1-\gamma^2}}, \frac{\varepsilon}{2(6B+1)}, \frac{\varepsilon(1-\gamma^2)}{4D}, \frac{\delta\Gamma\lambda(1-\lambda)}{2(\tilde{R}+\Gamma)}\right\}.$$

This has ultimately determined the size of $\mathcal{K}$ (the smaller set needed to kick start the arguments in the lemmas that will follow) while $\mathcal{K}^*$ (the larger set, on which all constants defined so far exist, and apply uniformly as the orbit travels through it) had been already previously fixed, to determine the constants necessary to define $\mathcal{K}$.

Lemma 4.38. *For all $\omega \in \mathscr{E}^{(1)}_{>}$ such that for all $0 \le l \le 2k-b$,*

$$\Theta^{(1)}_{m+l} \le (\gamma')^{\lfloor \frac{l}{2} \rfloor}\left(\Theta^{(1)}_m + \frac{3}{\nu^m(1-\gamma')}\right),$$

where $b \in \{0,1\}$*, it holds that for all* $b \le j \le 2k$*,*

$$\delta' < \pi^{(1)}_{m+2k-j} < \frac{1}{2} - \delta'$$

if j *is even and*

$$\frac{1}{2} + \delta' < \pi^{(1)}_{m+2k-j} < 1 - \delta'$$

if j *is odd.*

Proof. Iterate the first component of (2.32) after rearranging it as

$$\pi^{(1)}_{m+l} = 1 - \pi^{(1)}_{m+l-1} + \frac{1-\rho_{m+l}}{\rho_{m+l}}\Theta^{(1)}_{m+l-1} - \frac{1}{\rho_{m+l}}\Theta^{(1)}_{m+l}.$$

It yields

$$\pi^{(1)}_{m+l} = \\ = \begin{cases} \pi^{(1)}_m - \frac{1-\rho_{m+1}}{\rho_{m+1}}\Theta^{(1)}_m + \sum_{j=1}^{l-1}(-1)^{j+1}\left(\frac{1}{\rho_{m+j}} + \frac{1-\rho_{m+j+1}}{\rho_{m+j+1}}\right)\Theta^{(1)}_{m+j} - \frac{1}{\rho_{m+l}}\Theta^{(1)}_{m+l} & l \text{ even} \\ 1 - \pi^{(1)}_m + \frac{1-\rho_{m+1}}{\rho_{m+1}}\Theta^{(1)}_m + \sum_{j=1}^{l-1}(-1)^{j}\left(\frac{1}{\rho_{m+j}} + \frac{1-\rho_{m+j+1}}{\rho_{m+j+1}}\right)\Theta^{(1)}_{m+j} - \frac{1}{\rho_{m+l}}\Theta^{(1)}_{m+l} & l \text{ odd.} \end{cases} \tag{4.41}$$

Recall that by construction

$$\varepsilon < \frac{\rho\delta'(1-\gamma')}{8\left(R + \frac{1}{1-\gamma'}\right)}.$$

Then since $\pi^{(1)}_m < 1/2 - \delta$,

$$\begin{aligned} \pi^{(1)}_{m+2k-j} &= \pi^{(1)}_m - \frac{1-\rho_{m+1}}{\rho_{m+1}}\Theta^{(1)}_m + \sum_{l=1}^{2k-j-1}(-1)^{l+1}\left(\frac{1}{\rho_{m+l}} + \frac{1-\rho_{m+l+1}}{\rho_{m+l+1}}\right)\Theta^{(1)}_{m+l} \\ &- \frac{1}{\rho_{m+2k-j}}\Theta^{(1)}_{m+2k-j} < \frac{1}{2} - \delta + \frac{4}{\rho}\sum_{l=0}^{2k-j}\Theta^{(1)}_{m+l} < \frac{1}{2} - \delta + \frac{8}{\rho}\left(R + \frac{1}{1-\gamma'}\right)\varepsilon\sum_{l=0}^{\lfloor\frac{2k-j}{2}\rfloor}(\gamma')^l \\ &< \frac{1}{2} - \delta + \frac{8}{\rho}\left(R + \frac{1}{1-\gamma'}\right)\frac{\varepsilon}{1-\gamma'} < \frac{1}{2} - \delta' \end{aligned} \tag{4.42}$$

for all even $b \le j \le 2k$ (with the bound for $j = 2k$ holding also with δ, by adopting empty sum convention) and

$$\begin{aligned} \pi^{(1)}_{m+2k-j} &= 1 - \pi^{(1)}_m + \frac{1-\rho_{m+1}}{\rho_{m+1}}\Theta^{(1)}_m + \sum_{l=1}^{2k-j-1}(-1)^{l}\left(\frac{1}{\rho_{m+l}} + \frac{1-\rho_{m+l+1}}{\rho_{m+l+1}}\right)\Theta^{(1)}_{m+l} \\ &- \frac{1}{\rho_{m+2k-j}}\Theta^{(1)}_{m+2k-j} > \frac{1}{2} + \delta - \frac{4}{\rho}\sum_{l=0}^{2k-j}\Theta^{(1)}_{m+l} > \frac{1}{2} + \delta - \frac{8}{\rho}\left(R + \frac{1}{1-\gamma'}\right)\varepsilon\sum_{l=0}^{\lfloor\frac{2k-j}{2}\rfloor}(\gamma')^l \\ &> \frac{1}{2} + \delta - \frac{8}{\rho}\left(R + \frac{1}{1-\gamma'}\right)\frac{\varepsilon}{1-\gamma'} > \frac{1}{2} + \delta' \end{aligned} \tag{4.43}$$

for all odd $b \le j \le 2k$. Similarly, since $\delta < \pi^{(1)}_m < 1-\delta$,

$$\pi^{(1)}_{m+2k-j} = \pi^{(1)}_m - \frac{1-\rho_{m+1}}{\rho_{m+1}}\Theta^{(1)}_m + \sum_{l=1}^{2k-j-1}(-1)^{l+1}\left(\frac{1}{\rho_{m+l}} + \frac{1-\rho_{m+l+1}}{\rho_{m+l+1}}\right)\Theta^{(1)}_{m+l}$$
$$- \frac{1}{\rho_{m+2k-j}}\Theta^{(1)}_{m+2k-j} > \delta - \frac{4}{\rho}\sum_{l=0}^{2k-j}\Theta^{(1)}_{m+l} > \delta - \frac{8}{\rho}\left(R + \frac{1}{1-\gamma'}\right)\varepsilon\sum_{l=0}^{\lfloor\frac{2k-j}{2}\rfloor}(\gamma')^l$$
$$> \delta - \frac{8}{\rho}\left(R + \frac{1}{1-\gamma'}\right)\frac{\varepsilon}{1-\gamma'} > \delta' \tag{4.44}$$

for all even $b \le j \le 2k$ and

$$\pi^{(1)}_{m+2k-j} = 1 - \pi^{(1)}_m + \frac{1-\rho_{m+1}}{\rho_{m+1}}\Theta^{(1)}_m + \sum_{l=1}^{2k-j-1}(-1)^{l}\left(\frac{1}{\rho_{m+l}} + \frac{1-\rho_{m+l+1}}{\rho_{m+l+1}}\right)\Theta^{(1)}_{m+l}$$
$$- \frac{1}{\rho_{m+2k-j}}\Theta^{(1)}_{m+2k-j} < 1-\delta + \frac{4}{\rho}\sum_{l=0}^{2k-j}\Theta^{(1)}_{m+l} < 1-\delta + \frac{8}{\rho}\left(R + \frac{1}{1-\gamma'}\right)\varepsilon\sum_{l=0}^{\lfloor\frac{2k-j}{2}\rfloor}(\gamma')^l$$
$$< 1-\delta + \frac{8}{\rho}\left(R + \frac{1}{1-\gamma'}\right)\frac{\varepsilon}{1-\gamma'} < 1-\delta' \tag{4.45}$$

for all odd $b \le j \le 2k$. □

Lemma 4.39. *For all $\omega \in \mathscr{E}^{(1)}_>$ such that for all $0 \le l \le 2k-b$,*

$$\Theta^{(1)}_{m+l} \le (\gamma')^{\lfloor\frac{l}{2}\rfloor}\left(\Theta^{(1)}_m + \frac{3}{\nu^m(1-\gamma')}\right),$$

where $b \in \{0,1\}$, and $\delta' < \Theta^{(2)}_{m+l} < 1-\delta'$, it holds that for all $b \le j \le 2k-1$,

$$|\Theta^{(2)}_{m+2k-j} - \Theta^{(2)}_{m+2k-j-1}| < \varepsilon.$$

Proof. Recall that

$$\varepsilon' < \frac{\varepsilon(1-\gamma')}{4D}.$$

Iterate Remark 4.37 setting $n = m + 2(k-1) - j$ down to time m, it yields

$$|\Theta^{(2)}_{m+2k-j} - \Theta^{(2)}_{m+2k-j-1}| \le |\Theta^{(2)}_{m+1} - \Theta^{(2)}_m| + \sum_{l=1}^{2k-j-1}E_{m+l} < \frac{\varepsilon}{2} + \sum_{l=1}^{2k-2}E_{m+l},$$

where $E_{m+l} := \xi_{m+l} + \xi'_{m+l} + \eta_{m+l} + \eta'_{m+l} + \eta_{m+l-1} + \eta''_{m+l} + \eta'''_{m+l} + |r'_{m+l}|$. Note that the hypotheses allow to apply Lemma 4.38, thus $\delta' < \pi^{(1)}_{m+l-1} < 1-\delta'$ for all $1 \le l \le 2k-j-1$. This implies that, using Remark 4.35 and the assumptions, for all $1 \le l \le 2k-j-1$,

$$\eta_{m+l-1} < \Theta^{(1)}_{m+l-1} < (\gamma')^{\lfloor\frac{l-1}{2}\rfloor}\left(R + \frac{1}{1-\gamma'}\right)\varepsilon' \tag{4.46}$$

$$\eta_{m+l} < R\Theta^{(1)}_{m+l-1} + \frac{1}{\nu^{m+l}} < \left[R\left(R + \frac{1}{1-\gamma'}\right) + (\gamma')^{l-\lfloor\frac{l-1}{2}\rfloor}\right](\gamma')^{\lfloor\frac{l-1}{2}\rfloor}\varepsilon' \tag{4.47}$$

$$\eta'_{m+l} < \frac{\Theta^{(1)}_{m+l-1}}{(\delta')^2} < \frac{1}{(\delta')^2}(\gamma')^{\lfloor\frac{l-1}{2}\rfloor}\left(R + \frac{1}{1-\gamma'}\right)\varepsilon' \tag{4.48}$$

$$\eta''_{m+l} < \frac{1}{\delta'}\left(R\Theta^{(1)}_{m+l-1} + \frac{1}{\nu^{m+l}}\right) < \frac{1}{\delta'}\left[R\left(R + \frac{1}{1-\gamma'}\right) + (\gamma')^{l-\lfloor\frac{l-1}{2}\rfloor}\right](\gamma')^{\lfloor\frac{l-1}{2}\rfloor}\varepsilon' \tag{4.49}$$

$$\eta'''_{m+l} < \Theta^{(1)}_{m+l-1} < (\gamma')^{\lfloor\frac{l-1}{2}\rfloor}\left(R + \frac{1}{1-\gamma'}\right)\varepsilon', \tag{4.50}$$

and similarly we can estimate the three error terms

$$\xi_{m+l} < \left[1+\frac{1}{\delta'}\left(\frac{2}{\rho}-1\right)\right]\left(R\Theta^{(1)}_{m+l-1}+\frac{1}{\nu^{m+l}}\right)$$
$$< \left[1+\frac{1}{\delta'}\left(\frac{2}{\rho}-1\right)\right]\left[R\left(R+\frac{1}{1-\gamma'}\right)+(\gamma')^{l-\lfloor\frac{l-1}{2}\rfloor}\right](\gamma')^{\lfloor\frac{l-1}{2}\rfloor}\varepsilon' \tag{4.51}$$
$$\xi'_{m+l} < \frac{1}{\delta'}\left(\frac{2}{\rho}-1\right)\left(R+\frac{1}{1-\gamma'}\right)(\gamma')^{\lfloor\frac{l-1}{2}\rfloor}\varepsilon' \tag{4.52}$$
$$|r'_{m+l}| \leq \frac{|R^{(3)}_{m+l}|}{\pi^{(1)}_{m+l-1}(\Theta^{(1)}_{m+l}+\Theta^{(2)}_{m+l})}+\frac{|R^{(2)}_{m+l+1}|}{\pi^{(1)}_{m+l-1}\Theta^{(2)}_{m+l}} < \frac{2}{(\delta')^2\nu^{m+l}} < \frac{2(\gamma')^l}{(\delta')^2}\varepsilon'. \tag{4.53}$$

Hence

$$E_{m+l} \leq D(\gamma')^{\lfloor\frac{l-1}{2}\rfloor}\varepsilon', \tag{4.54}$$

which yields a bound uniform in k on the increments of the $\Theta^{(2)}$-component,

$$|\Theta^{(2)}_{m+2k-1}-\Theta^{(2)}_{m+2k-2}| < \frac{\varepsilon}{2}+D\varepsilon'\sum_{l=1}^{\infty}(\gamma')^{\lfloor\frac{l-1}{2}\rfloor} = \frac{\varepsilon}{2}+2D\frac{\varepsilon'}{1-\gamma'} < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} < \varepsilon. \tag{4.55}$$

□

Lemma 4.40. *For almost every $\omega \in \mathscr{E}^{(1)}_{>}$, for all $m+1 \leq n \leq \sigma$,*

$$\Theta^{(1)}_n \leq \gamma^{n-m-1}\left(R\Theta^{(1)}_m+\frac{3}{\nu^m(1-\gamma^2)}\right).$$

Proof. Recall that

$$\gamma' := \gamma^2 = \max\left\{1-\rho^2\frac{(\delta')^2}{\rho^2}+\frac{35}{\rho}(\delta')^3, \frac{1}{\nu}\right\}.$$

We will first show, by adopting empty sum convention, that for almost all $\omega \in \mathscr{E}^{(1)}_{>}$, for every $k \geq 0$ such that $m+2k \leq \sigma$,

$$\Theta^{(1)}_{m+2k} \leq (\gamma')^k R\Theta^{(1)}_m+\frac{1}{\nu^m}\left(3\sum_{l=k+1}^{2k-1}(\gamma')^l+2(\gamma')^{2k}\right) \tag{4.56}$$
$$\Theta^{(1)}_{m+2k+1} \leq (\gamma')^k R\Theta^{(1)}_m+\frac{1}{\nu^m}\left(3\sum_{l=k+1}^{2k}(\gamma')^l+(\gamma')^{2k+1}\right) \tag{4.57}$$

and, if $m+2k=\sigma$, we only show the claim up to (4.56). Recall that by construction

$$\delta' < \min\left\{\frac{2}{105}\rho^3, \sqrt[3]{\frac{1-\tilde{\rho}}{3+\frac{8}{\rho}}}\right\}$$

and

$$\varepsilon < \min\left\{(\delta')^5, \frac{\rho\delta'(1-\gamma')}{8\left(R+\frac{1}{1-\gamma'}\right)}\right\}.$$

Recall also that

$$\varepsilon' < \min\left\{\frac{\varepsilon}{R+\frac{3}{1-\gamma'}}, \frac{\varepsilon(1-\gamma')}{4D}\right\}$$

and $3\nu^{-m} < \varepsilon'$. Choosing $n = m$ in Remark 4.35 yields (4.57) for $k = 0$ (for $k = 0$, (4.56) is trivial, since $R > 1$): for almost every $\omega \in \mathscr{E}^{(1)}_{>}$,

$$\Theta^{(1)}_{m+1} < R\Theta^{(1)}_m + \frac{1}{\nu^{m+1}} < R\Theta^{(1)}_m + \frac{\gamma'}{\nu^m}.$$

Let $n = m$ in (3.33) and (3.34) (after changing formally to Θ, π notation) and apply the hypotheses made in (4.36) to (4.38), and $|\Theta^{(2)}_{m+1} - \Theta^{(2)}_m| < \varepsilon/2$, which follows from (4.35), and is crucial to (4.59), the same way it was to (3.54). Then the estimates (3.53) to (3.55) follow, as in the corresponding step at the beginning of Lemmas A.38 and 3.46, for almost every $\omega \in \mathscr{E}^{(1)}_{>}$:

$$|\vartheta_{m+1}| \leq \frac{2}{\delta'(1-\varepsilon)}, \tag{4.58}$$

$$|\vartheta_{m+1}| \leq 2 + \frac{\varepsilon}{1-\varepsilon}\left(2 + \frac{3}{\delta'}\right), \tag{4.59}$$

$$|\vartheta'_{m+1}| \leq \frac{1}{\delta'}\left(\frac{1}{1-\varepsilon} + \frac{2}{\delta'+\varepsilon}\right). \tag{4.60}$$

Plugging (4.58) to (4.60) into (4.39) applied to $n = m$, yields

$$|\vartheta''_{m+1}| \leq \left(\frac{2}{\rho} - 1\right)|\vartheta_{m+1}| + |\vartheta'_{m+1}| \leq \left(\frac{2}{\rho} - 1\right)\frac{2}{\delta'(1-\varepsilon)} + \frac{1}{\delta'}\left(\frac{1}{1-\varepsilon} + \frac{2}{\delta'+\varepsilon}\right)$$

and therefore, for almost every $\omega \in \mathscr{E}^{(1)}_{>}$,

$$|\vartheta''_{m+1}| < \frac{1}{\delta'}\left[\left(\frac{4}{\rho} - 1\right)\frac{1}{1-\varepsilon} + \frac{2}{\delta'}\right]. \tag{4.61}$$

By plugging (4.58) into (4.40) applied to $n = m$, yields

$$|r_{m+1}| \leq \frac{1}{\nu^{m+1}}\left(|\vartheta_{m+1}| + \frac{2}{\delta'}\right) \leq \frac{1}{\nu^{m+1}}\left(\frac{2}{\delta'(1-\varepsilon)} + \frac{2}{\delta'}\right),$$

thus yielding, for almost every $\omega \in \mathscr{E}^{(1)}_{>}$, that

$$|r_{m+1}| < \frac{4}{\delta'(1-\varepsilon)\nu^{m+1}}. \tag{4.62}$$

Plug the estimates (4.58) to (4.62) into Remark 4.36 applied to $n = m$, it yields

$$\begin{aligned}
\Theta^{(1)}_{m+2} &\leq \Theta^{(1)}_{m+1}\Bigg\{(1-\rho_{m+2}) + \rho_{m+2}\Bigg[2(1-\pi^{(1)}_{m+1}) + \frac{\varepsilon}{1-\varepsilon}\left(2+\frac{3}{\delta'}\right) + \frac{\Theta^{(1)}_{m+1}}{\rho_{m+1}}\frac{2}{\delta'(1-\varepsilon)} \\
&+ \frac{\Theta^{(1)}_m}{\delta'}\left[\left(\frac{4}{\rho}-1\right)\frac{1}{1-\varepsilon} + \frac{2}{\delta'}\right] + |r_{m+1}|\Bigg]\Bigg\} + |R^{(1)}_{m+2}| \leq \Theta^{(1)}_{m+1}\Bigg\{(1-\rho_{m+2}) \\
&+ \rho_{m+2}\Bigg[2(1-\pi^{(1)}_{m+1}) + \varepsilon\Bigg\{\frac{1}{1-\varepsilon}\left(2+\frac{3}{\delta'}\right) + \frac{2}{\rho}\frac{2}{\delta'(1-\varepsilon)} + \frac{1}{\delta'}\left[\left(\frac{4}{\rho}-1\right)\frac{1}{1-\varepsilon} + \frac{2}{\delta'}\right]\Bigg\} \\
&+ \frac{4}{\delta'(1-\varepsilon)\nu^{m+1}}\Bigg]\Bigg\} + \frac{1}{\nu^{m+2}} < \Theta^{(1)}_{m+1}\Bigg\{(1-\rho_{m+2}) + \rho_{m+2}\Bigg[2(1-\pi^{(1)}_{m+1}) \\
&+ 2\varepsilon\left(\frac{1}{1-\varepsilon} + \left(\frac{4}{\rho}+1\right)\frac{1}{\delta'(1-\varepsilon)} + \frac{1}{(\delta')^2}\right) + \frac{4}{\delta'(1-\varepsilon)}\varepsilon'\Bigg]\Bigg\} + \frac{1}{\nu^{m+2}} \\
&\leq \Theta^{(1)}_{m+1}\left\{(1-\rho_{m+2}) + 2\rho_{m+2}(1-\pi^{(1)}_{m+1}) + 2\varepsilon\left(\frac{1}{1-\varepsilon} + 2\left(1+\frac{2}{\rho}\right)\frac{1}{\delta'(1-\varepsilon)} + \frac{1}{(\delta')^2}\right)\right\} \\
&+ \frac{1}{\nu^{m+2}}
\end{aligned}$$

for almost every $\omega \in \mathscr{E}_{>}^{(1)}$. In the last inequality we used $\varepsilon' < \varepsilon/2$. Recall that $\varepsilon < (\delta')^5 < 2^5/105^5$. Then $\pi_m^{(1)} < 1/2 - \delta$ implies that for almost every $\omega \in \mathscr{E}_{>}^{(1)}$,

$$\pi_{m+1}^{(1)} = 1 - \pi_m^{(1)} + \frac{1-\rho_{m+1}}{\rho_{m+1}}\Theta_m^{(1)} - \frac{1}{\rho_{m+1}}\Theta_{m+1}^{(1)} > \frac{1}{2} + \delta - \frac{2}{\rho}\Theta_m^{(1)} - \frac{2}{\rho}\left(R\Theta_m^{(1)} + \frac{1}{\nu^{m+1}}\right)$$
$$> \frac{1}{2} + \delta - \frac{4}{\rho}(R+1)\,\varepsilon > \frac{1}{2} + \delta'$$

by construction of ε. Since

$$(1-\rho_{m+2}) + 2\rho_{m+2}(1-\pi_{m+1}^{(1)}) + 2\varepsilon\left(\frac{1}{1-\varepsilon} + 2\left(1+\frac{2}{\rho}\right)\frac{1}{\delta'(1-\varepsilon)} + \frac{1}{(\delta')^2}\right)$$
$$< 1 - \rho_{m+2} + 2\rho_{m+2}(1-\pi_{m+1}^{(1)}) + 2\varepsilon\left(\frac{1}{\delta'} + \left(3+\frac{4}{\rho}\right)\frac{1}{(\delta')^2}\right) < 1 - 2\rho_{m+2}\delta'$$
$$+ 4\left(1+\frac{1}{\rho}\right)\frac{\varepsilon}{(\delta')^2} < 1 - \rho\delta' + \frac{8}{\rho}(\delta')^3,$$

by (4.56) applied to $k=1$, for almost every $\omega \in \mathscr{E}_{>}^{(1)}$ we have that

$$\Theta_{m+2}^{(1)} < \Theta_{m+1}^{(1)}\left(1 - \rho\delta' + \frac{8}{\rho}(\delta')^3\right) + \frac{1}{\nu^{m+2}} \leq \gamma'\Theta_{m+1}^{(1)} + \frac{1}{\nu^{m+2}} < R\gamma'\Theta_m^{(1)} + \frac{\gamma'}{\nu^{m+1}}$$
$$+ \frac{1}{\nu^{m+2}} < R\gamma'\Theta_m^{(1)} + \frac{2}{\nu^m}(\gamma')^2.$$

Note that for any $\delta' > 0$, $\gamma' \geq 1 - \rho^2(\delta')^2 + 35(\delta')^3/\rho > 1 - \rho\delta' + 8(\delta')^3/\rho$ since it is equivalent to $\rho - \rho^2\delta' + 27(\delta')^2/\rho > 0$, and $\rho - \rho^2 x + 27x^2/\rho > 0$ is a convex parabola with symmetry axis parallel to the y-axis taking value $\rho > 0$ at 0 and having negative discriminant $\rho^4 - 108/\rho$ (since $\rho < 1$, the discriminant is less than -107). If $\omega \in \mathscr{E}_{>}^{(1)}$ is such that $\sigma > m+2$, the case $k=1$ is not yet concluded. By the geometric decaying upper bound proved so far and the construction of ε', for almost every ω considered

$$\Theta_{m+2}^{(1)} < \left(R + \frac{3}{1-\gamma'}\right)\varepsilon' < \varepsilon.$$

By the definition of σ and the fact that $\delta' < p_{m+2}^{(2)} < 1-\delta'$, also the same estimates in (4.58) and (4.60) to (4.62) apply, for almost every such ω, to ϑ_{m+2}, ϑ'_{m+2}, ϑ''_{m+2} and r_{m+2}, with the due shift of time indices. However, (4.59) does not apply automatically since nothing guarantees that the same bound applies on the shifted increments of the $\Theta^{(2)}$-component. Let us first assume that indeed it also holds that $|\Theta_{m+2}^{(2)} - \Theta_{m+1}^{(2)}| < \varepsilon$ for almost every ω considered, and therefore that also (4.59) applies for almost every ω considered, with the due shift of indices. Plugging these into Remark 4.36 applied to $n = m+1$, and recalling that for all $n \geq m$, $\rho_{n+1}/\rho_n < 3/2$ and $\rho_{n+1} - \rho_n < \varepsilon$, yields that

$$\Theta^{(1)}_{m+3} \le \Theta^{(1)}_{m+2}\left[1-\rho_{m+3}+2\rho_{m+3}(1-\pi^{(1)}_{m+2})+4\left(1+\frac{1}{\rho}\right)\frac{\varepsilon}{(\delta')^2}\right]+\rho_{m+3}|R^{(1)}_{m+3}|$$
$$< \Theta^{(1)}_{m+2}\left[1-\rho_{m+3}+2\rho_{m+3}\left(\pi^{(1)}_{m+1}-\frac{1-\rho_{m+2}}{\rho_{m+2}}\Theta^{(1)}_{m+1}+\frac{1}{\rho_{m+2}}\Theta^{(1)}_{m+2}\right)+\frac{8}{\rho}\frac{\varepsilon}{(\delta')^2}\right]+\frac{1}{\nu^{m+3}}$$
$$\le \left\{\Theta^{(1)}_{m+1}\left[1-\rho_{m+2}+2\rho_{m+2}(1-\pi^{(1)}_{m+1})+\frac{8}{\rho}\frac{\varepsilon}{(\delta')^2}\right]+\frac{1}{\nu^{m+2}}\right\}\left[1-\rho_{m+3}+2\rho_{m+3}\pi^{(1)}_{m+1}\right.$$
$$\left.+3\Theta^{(1)}_{m+2}+\frac{8}{\rho}\frac{\varepsilon}{(\delta')^2}\right]+\frac{1}{\nu^{m+3}} < \left\{\Theta^{(1)}_{m+1}\left[1-\rho_{m+2}+2\rho_{m+2}(1-\pi^{(1)}_{m+1})+\frac{8}{\rho}\frac{\varepsilon}{(\delta')^2}\right]\right.$$
$$\left.+\frac{1}{\nu^{m+2}}\right\}\left[1-\rho_{m+3}+2\rho_{m+3}\pi^{(1)}_{m+1}+\varepsilon\left(3+\frac{8}{\rho(\delta')^2}\right)\right]+\frac{1}{\nu^{m+3}} < \Theta^{(1)}_{m+1}\left[1-\rho_{m+2}\right.$$
$$\left.+2\rho_{m+2}(1-\pi^{(1)}_{m+1})+\frac{8}{\rho}\frac{1\varepsilon}{(\delta')^2}\right]\left[1-\rho_{m+3}+2\rho_{m+3}\pi^{(1)}_{m+1}+\varepsilon\left(3+\frac{8}{\rho(\delta')^2}\right)\right]$$
$$+\frac{1}{\nu^{m+2}}\left[1-\rho_{m+3}+2\rho_{m+3}\pi^{(1)}_{m+1}+(\delta')^5\left(3+\frac{8}{\rho(\delta')^2}\right)\right]+\frac{1}{\nu^{m+3}}$$
$$< \Theta^{(1)}_{m+1}\left[1-\rho_{m+2}+2\rho_{m+2}(1-\pi^{(1)}_{m+1})+\frac{8}{\rho}\frac{\varepsilon}{(\delta')^2}\right]\left[3-\rho_{m+3}-2\delta'+\varepsilon\left(3+\frac{8}{\rho(\delta')^2}\right)\right]$$
$$+\frac{2}{\nu^{m+2}}+\frac{1}{\nu^{m+3}} < \Theta^{(1)}_{m+1}\left[1-\rho_{m+2}\rho_{m+3}+4\rho_{m+2}\rho_{m+3}\pi^{(1)}_{m+1}(1-\pi^{(1)}_{m+1})\right.$$
$$\left.+(\rho_{m+3}-\rho_{m+2})(2\pi^{(1)}_{m+1}-1)+\varepsilon\left(3+\frac{8}{\rho(\delta')^2}\right)+\frac{24}{\rho}\frac{\varepsilon}{(\delta')^2}+\varepsilon^2\frac{8}{\rho(\delta')^2}\left(3+\frac{8}{\rho(\delta')^2}\right)\right]$$
$$+\frac{2}{\nu^{m+2}}+\frac{1}{\nu^{m+3}}$$

for almost every ω considered, since the condition

$$\delta < 2\sqrt[3]{\frac{1-\tilde{\rho}}{3+\frac{8}{\rho}}}$$

ensures

$$1-\rho_{m+3}+2\rho_{m+3}\pi^{(1)}_{m+1}+(\delta')^5\left(3+\frac{8}{\rho(\delta')^2}\right) < 1+\tilde{\rho}+\left(3+\frac{8}{\rho}\right)\delta'^3 < 2.$$

Noting that for almost every ω considered

$$1-\rho_{m+2}\rho_{m+3}+4\rho_{m+2}\rho_{m+3}\pi^{(1)}_{m+1}(1-\pi^{(1)}_{m+1})+(\rho_{m+3}-\rho_{m+2})(2\pi^{(1)}_{m+1}-1)$$
$$+\varepsilon\left(3+\frac{8}{\rho(\delta')^2}\right)+\frac{24}{\rho}\frac{\varepsilon}{(\delta')^2}+\varepsilon^2\frac{8}{\rho(\delta')^2}\left(3+\frac{8}{\rho(\delta')^2}\right) \le 1-\rho^2(\delta')^2+(\delta')^5$$
$$+3(\delta')^5+\frac{32}{\rho}(\delta')^3+\frac{24}{\rho}(\delta')^8+(\delta')^6\frac{64}{\rho^2} < 1-\rho^2(\delta')^2+(\delta')^3\left[\frac{32}{\rho}+4(\delta')^2\right.$$
$$\left.+\frac{64}{\rho^2}(\delta')^3+\frac{24}{\rho}(\delta')^5\right] \le 1-\rho^2(\delta')^2+\frac{35}{\rho}(\delta')^3,$$

we can conclude that (4.57) holds for almost every $\omega \in \mathscr{E}^{(1)}_{>}$ considered, for $k=1$:

$$\Theta^{(1)}_{m+3} < \Theta^{(1)}_{m+1}\left(1-\rho^2(\delta')^2+\frac{35}{\rho}(\delta')^3\right)+\frac{2}{\nu^{m+2}}+\frac{1}{\nu^{m+3}}$$
$$< \gamma'\Theta^{(1)}_{m+1}+\frac{2}{\nu^{m+2}}+\frac{1}{\nu^{m+3}} < \gamma' R\Theta^{(1)}_{m}+\frac{1}{\nu^{m}}\left(3(\gamma')^2+(\gamma')^3\right).$$

We now show that the upper bound on the $\Theta^{(2)}$-component keeps applying for almost every ω considered, by using Remark 4.37 applied to $n = m$, yielding the upper bound corresponding to Lemma 3.45 by defining $E_{m+1} := \xi_{m+1} + \xi'_{m+1} + \eta_{m+1} + \eta'_{m+1} + \eta_m + \eta''_{m+1} + \eta'''_{m+1} + |r'_{m+1}|$, that is

$$|\Theta^{(2)}_{m+2} - \Theta^{(2)}_{m+1}| \le |\Theta^{(2)}_{m+1} - \Theta^{(2)}_{m}| + E_{m+1}.$$

Since, as previously mentioned, for almost every ω considered, $\Theta^{(1)}_{m+2} < \varepsilon$, by the definition of σ, which ensures that also $\delta' < \Theta^{(2)}_{m+1} < 1-\delta'$, and by exploiting Remark 4.35 applied to $n = m$, we can estimate

$$\eta_m = \frac{\Theta^{(2)}_m\Theta^{(1)}_m}{\Theta^{(2)}_m + \Theta^{(3)}_m} < \Theta^{(1)}_m \tag{4.63}$$

$$\eta_{m+1} = \frac{\Theta^{(2)}_{m+1}\Theta^{(1)}_{m+1}}{\Theta^{(2)}_{m+1} + \Theta^{(3)}_{m+1}} < R\Theta^{(1)}_m + \frac{1}{\nu^{m+1}} \tag{4.64}$$

$$\eta'_{m+1} := \frac{\pi^{(2)}_m}{\pi^{(1)}_m}\frac{\Theta^{(2)}_{m+1}}{\Theta^{(1)}_{m+1} + \Theta^{(2)}_{m+1}}\frac{\Theta^{(1)}_m}{1 - \Theta^{(2)}_m} < \frac{1}{(\delta')^2}\Theta^{(1)}_m \tag{4.65}$$

$$\eta''_{m+1} := \Theta^{(2)}_m\frac{\Theta^{(1)}_{m+1}}{\Theta^{(1)}_{m+1} + \Theta^{(2)}_{m+1}} < \frac{1-\delta'}{\delta'}\left(R\Theta^{(1)}_m + \frac{1}{\nu^{m+1}}\right) < \frac{1}{\delta'}\left(R\Theta^{(1)}_m + \frac{1}{\nu^{m+1}}\right) \tag{4.66}$$

$$\eta'''_{m+1} = \Theta^{(2)}_m\frac{\Theta^{(1)}_{m+1}}{\Theta^{(1)}_{m+1} + \Theta^{(2)}_{m+1}}\frac{\Theta^{(1)}_m}{\Theta^{(2)}_m + \Theta^{(3)}_m} < \Theta^{(1)}_m \tag{4.67}$$

$$\xi_{m+1} = \frac{\Theta^{(2)}_{m+1}}{\Theta^{(2)}_{m+1} + \Theta^{(3)}_{m+1}}\left(1 + \frac{\frac{1}{\rho_{m+1}} - 1}{\pi^{(1)}_m}\right)\Theta^{(1)}_{m+1} < \left[1 + \frac{1}{\delta'}\left(\frac{2}{\rho} - 1\right)\right]\left(R\Theta^{(1)}_m + \frac{1}{\nu^{m+1}}\right) \tag{4.68}$$

$$\xi'_{m+1} = \frac{\Theta^{(2)}_{m+1}}{\Theta^{(2)}_{m+1} + \Theta^{(3)}_{m+1}}\frac{\frac{1}{\rho_{m+1}} - 1}{\pi^{(1)}_m}\Theta^{(1)}_m < \frac{1}{\delta'}\left(\frac{2}{\rho} - 1\right)\Theta^{(1)}_m \tag{4.69}$$

$$|r'_{m+1}| \le \frac{|R^{(3)}_{m+1}|}{\pi^{(1)}_m(\Theta^{(1)}_{m+1} + \Theta^{(2)}_{m+1})} + \frac{|R^{(2)}_{m+2}|}{\pi^{(1)}_m\Theta^{(2)}_{m+1}} \le \frac{2}{(\delta')^2\nu^{m+1}} \tag{4.70}$$

for almost every ω considered. Then by recalling that $\Theta^{(1)}_m, \nu^{-m} < \varepsilon'$ and $R > 1$, we have that for almost every ω considered

$$E_{m+1} < \left[2 + \frac{1}{\delta'}\left(\frac{2}{\rho} - 1\right) + \frac{3}{(\delta')^2} + (R+1)\left(1 + \frac{2}{\delta'\rho}\right)\right]\varepsilon' < D\varepsilon'.$$

Therefore, by construction of ε' we get, for almost every $\omega \in \mathscr{E}^{(1)}_{>}$ considered, that

$$E_{m+1} \le \frac{\varepsilon}{4}.$$

Since $|\Theta^{(2)}_{m+1} - \Theta^{(2)}_m| < \varepsilon/2$, this yields that for almost every ω considered

$$|\Theta^{(2)}_{m+2} - \Theta^{(2)}_{m+1}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} < \varepsilon.$$

Apart from the base cases, this estimate will be less immediate in further steps and we will rely on Lemma 4.39.

To summarise what we proved in this two steps argument: there is a constant $\gamma' = \gamma'(\delta', \nu)$ holding uniformly on $\mathcal{K}^*$ almost surely on both events in $\mathscr{E}^{(1)}_{>}$: $\{\sigma = m+2\}$ and $\{\sigma > m+2\}$. For almost every ω in the first event, $\Theta^{(1)}_{m+1} < (\gamma')^0 R\Theta^{(1)}_m + \gamma'\nu^{-m}$

(case $k=0$), and $\Theta^{(1)}_{m+2} < \gamma' R\Theta^{(1)}_m + 2(\gamma')^2\nu^{-m})$ (half case $k=1$); for almost every ω in the second event both $\Theta^{(1)}_{m+1} < (\gamma')^0 R\Theta^{(1)}_m + \gamma'\nu^{-m})$ (case $k=0$), and $\Theta^{(1)}_{m+2} < \gamma' R\Theta^{(1)}_m + 2(\gamma')^2\nu^{-m})$ and $\Theta^{(1)}_{m+3} < \gamma' R\Theta^{(1)}_m + \nu^{-m}(3(\gamma')^2 + (\gamma')^3)$ (full case $k=1$). Note, before proceeding, that the estimate on $\pi^{(1)}_n$'s oscillations above and below $1/2$ has to iterate at each step. For example, for the next step it will hold for almost every $\omega \in \mathscr{E}^{(1)}_>$ considered, because for all $j \in \mathbb{N}$,

$$\frac{1}{\rho_{m+j}}, \ \frac{1-\rho_{m+j}}{\rho_{m+j}}, \ \frac{1}{\rho_{m+j}} + \frac{1-\rho_{m+j+1}}{\rho_{m+j+1}} < \frac{4}{\rho},$$

hence

$$\begin{aligned}
\pi^{(1)}_{m+2} &= 1 - \pi^{(1)}_{m+1} + \frac{1-\rho_{m+2}}{\rho_{m+2}}\Theta^{(1)}_{m+1} - \frac{1}{\rho_{m+2}}\Theta^{(1)}_{m+2} = \pi^{(1)}_m - \frac{1-\rho_{m+1}}{\rho_{m+1}}\Theta^{(1)}_m + \frac{1}{\rho_{m+1}}\Theta^{(1)}_{m+1} \\
&\quad + \frac{1-\rho_{m+2}}{\rho_{m+2}}\Theta^{(1)}_{m+1} - \frac{1}{\rho_{m+2}}\Theta^{(1)}_{m+2} < \frac{1}{2} - \delta + \frac{1-\rho_{m+1}}{\rho_{m+1}}\Theta^{(1)}_m + \left(\frac{1}{\rho_{m+1}} + \frac{1-\rho_{m+2}}{\rho_{m+2}}\right) \\
&\quad \Theta^{(1)}_{m+1} + \frac{1}{\rho_{m+2}}\Theta^{(1)}_{m+2} < \frac{1}{2} - \delta + \frac{4}{\rho}\left[2\left(R\Theta^{(1)}_m + \frac{1}{\nu^m}\right) + \gamma'\left(R\Theta^{(1)}_m + \frac{1}{\nu^m}\right)\right] \\
&\quad < \frac{1}{2} - \delta + \frac{8}{\rho}(R+1)\varepsilon(1+\gamma') < \frac{1}{2} - \delta'
\end{aligned}$$

by construction of ε. If $\omega \in \mathscr{E}^{(1)}_>$ is such that $m+3 < n < \sigma$, we show the claim for $n+1$. There are two steps to perform, each requiring a separate argument, depending on ω: the even step from $n = m+2k-1$ to $n+1 = m+2k$ and the odd step from $n = m+2k$ to $n+1 = m+2k+1$, for all $k \in \mathbb{N}$ such that n is in the mentioned range.

- In the even step, putting (4.56) and (4.57) together, one has the induction hypothesis that for almost every $\omega \in \mathscr{E}^{(1)}_>$ considered, for all $1 \le j \le 2k-1$,

$$\Theta^{(1)}_{m+2k-j} < \begin{cases} (\gamma')^{\lfloor \frac{2k-j}{2} \rfloor} R\Theta^{(1)}_m + \frac{1}{\nu^m}\left(3\sum_{l=\lfloor \frac{2k-j}{2}\rfloor+1}^{2k-j-1}(\gamma')^l + 2(\gamma')^{2k-j}\right), & j \text{ even} \\ (\gamma')^{\lfloor \frac{2k-j}{2} \rfloor} R\Theta^{(1)}_m + \frac{1}{\nu^m}\left(3\sum_{l=\lfloor \frac{2k-j}{2}\rfloor+1}^{2k-j-1}(\gamma')^l + (\gamma')^{2k-j}\right), & j \text{ odd} \end{cases} \tag{4.71}$$

  so that for almost every ω considered

$$\Theta^{(1)}_{m+2k-j} < (\gamma')^{\lfloor \frac{2k-j}{2} \rfloor}\left(R\Theta^{(1)}_m + \frac{3}{\nu^m}\sum_{l=1}^{2k-j}(\gamma')^l\right) < (\gamma')^{\lfloor \frac{2k-j}{2} \rfloor}\left(R + \frac{1}{1-\gamma'}\right)\varepsilon', \tag{4.72}$$

  and (4.56) needs to be shown. As to the oscillations of $\pi^{(1)}$, they are δ'-bounded away from $1/2$ in the correct order, for almost every ω considered, thanks to (4.72), which allows the use of Lemma 4.38 with $b=1$. For almost every ω considered,

$$\pi^{(1)}_{m+2k-j} < \frac{1}{2} - \delta'$$

  for all even $1 \le j \le 2k$ (with the bound for $j = 2k$ holding also with δ by adopting empty sum convention) and

$$\pi^{(1)}_{m+2k-j} > \frac{1}{2} + \delta'$$

  for all odd $1 \le j \le 2k$ (with bound for $j = 2k$ holding also with δ by adopting empty sum convention). All that remains to be shown is that

$\Theta^{(1)}_{m+2k} < \gamma'\Theta^{(1)}_{m+2k-1} + \tilde{\rho}\nu^{-m-2k}$, by using $\pi^{(1)}_{m+2k-1} > 1/2 + \delta'$ for almost every ω considered. Since (4.71) holds, implying

$$\Theta^{(1)}_{m+2k-1} < (\gamma')^{k-1} R\Theta^{(1)}_{m} + \frac{1}{\nu^m}\left(3\sum_{l=k}^{2(k-1)} (\gamma')^l + (\gamma')^{2k-1}\right)$$

for almost every $\omega \in \mathscr{E}^{(1)}_{>}$ considered, and since ω is such that $\sigma > m + 2k - 1$, which implies that $\delta' < \Theta^{(2)}_{m+2k-1} < 1-\delta'$; the estimates in (4.58) to (4.61) apply also to ϑ_{m+2k-1}, ϑ'_{m+2k-1} and ϑ''_{m+2k-1} (with the due shift of time indices) for almost every ω considered, because by Lemma 4.39 with $b = 1$ we know that $|\Theta^{(2)}_{m+2k-1} - \Theta^{(2)}_{m+2k-1}| < \varepsilon$, for almost every $\omega \in \mathscr{E}^{(1)}_{>}$ considered. Plugging the aforementioned estimates into Remark 4.36 applied to $n = m + 2k - 2$, yields the same estimate as that obtained for $\Theta^{(1)}_{m+2}$ earlier, that is,

$$\begin{aligned}
\Theta^{(1)}_{m+2k} &< \Theta^{(1)}_{m+2k-1}\Bigg\{(1-\rho_{m+2k}) + 2\rho_{m+2k}(1-\pi^{(1)}_{m+2k-1}) + 2\varepsilon\Bigg(\frac{1}{1-\varepsilon} \\
&\quad + \frac{2}{\delta'(1-\varepsilon)}\left(1+\frac{2}{\rho}\right) + \frac{1}{(\delta')^2}\Bigg)\Bigg\} + \frac{1}{\nu^{m+2k}} < \Theta^{(1)}_{m+2k-1}\left(1-\rho\delta' + \frac{8}{\rho}(\delta')^3\right) \\
&\quad + \frac{1}{\nu^{m+2k}} < \gamma'\Theta^{(1)}_{m+2k-1} + \frac{(\gamma')^{2k}}{\nu^m} < \gamma'(\gamma')^{k-1}R\Theta^{(1)}_{m} \\
&\quad + \frac{\gamma'}{\nu^m}\left(3\sum_{l=k}^{2(k-1)}(\gamma')^l + (\gamma')^{2k-1}\right) + \frac{(\gamma')^{2k}}{\nu^m} \\
&= (\gamma')^k R\Theta^{(1)}_{m} + \frac{1}{\nu^m}\left(3\sum_{l=k+1}^{2k-1}(\gamma')^l + 2(\gamma')^{2k}\right)
\end{aligned}$$

for almost every $\omega \in \mathscr{E}^{(1)}_{>}$ considered, which is (4.56).

- In the odd step one has the induction hypothesis (4.72) for all $0 \le j \le 2k-1$, since we have just shown, for almost every $\omega \in \mathscr{E}^{(1)}_{>}$ considered, (4.56) for the even step, so that the new range is extended by one index, and we need to show (4.57). For the oscillations of $\pi^{(1)}$, we proceed similarly to the even step but with a different range for j, by exploiting Lemma 4.38 applied with $b = 0$. For almost every $\omega \in \mathscr{E}^{(1)}_{>}$ considered

$$\pi^{(1)}_{m+2k-j} < \frac{1}{2} - \delta'$$

for all even $0 \le j \le 2k$ (with the bound for $j = 2k$ holding also with δ by adopting empty sum convention) and

$$\pi^{(1)}_{m+2k-j} > \frac{1}{2} + \delta'$$

for all odd $0 \le j \le 2k$ (with bound for $j = 2k$ holding also with δ, by adopting empty sum convention). All that has to be shown explicitly is that for almost every $\omega \in \mathscr{E}^{(1)}_{>}$ considered

$$\theta^{(1)}_{m+2k+1} < \gamma'\theta^{(1)}_{m+2k-1} + \frac{2}{\nu^{m+2k}} + \frac{1}{\nu^{m+2k+1}},$$

by using

$$\pi^{(1)}_{m+2k-1} > \frac{1}{2} + \delta'.$$

Since, like in the even step, (4.71) holds, implying

$$\Theta^{(1)}_{m+2k-1} < (\gamma')^{k-1} R\Theta^{(1)}_m + \frac{1}{\nu^m}\left(3\sum_{l=k}^{2(k-1)}(\gamma')^l + (\gamma')^{2k-1}\right)$$

for almost every $\omega \in \mathscr{E}^{(1)}_>$ considered, and since ω is such that $\sigma > m+2k$, implying that $\delta' < \Theta^{(2)}_{m+2k} < 1-\delta'$; the estimates in (4.58) to (4.61) apply up to ϑ_{m+2k}, ϑ'_{m+2k} and ϑ''_{m+2k} (with the due shift of time indices), since Lemma 4.39 with $b=0$ ensures that for almost every ω considered $|\Theta^{(2)}_{m+2k} - \Theta^{(2)}_{m+2k}| < \varepsilon$ as well; we will use explicitly also the previous step's estimates for ϑ_{m+2k-1} and ϑ'_{m+2k-1}. They are vital, since in this step the bound needed is yielded by iterating the previous even step into the current odd one, producing a two-step estimate, because a one-step estimate would not yield a subunitary constant due to $\pi^{(1)}_{m+2k} < 1/2 - \delta'$, which would imply $2(1-\pi^{(1)}_{m+2k}) > 1$. Therefore, by plugging these estimates into Remark 4.36 applied to $n = m+2k-1$, and also using the estimate from the previous even step, yields the estimate corresponding to the one obtained for $\Theta^{(1)}_{m+3}$:

$$\begin{aligned}\Theta^{(1)}_{m+2k+1} &< \Theta^{(1)}_{m+2k-1}\left[1-\rho_{m+2k} + 2\rho_{m+2k}(1-\pi^{(1)}_{m+2k-1}) + \frac{8}{\rho}\frac{\varepsilon}{(\delta')^2}\right] \\ &\left[1-\rho_{m+2k+1} + 2\rho_{m+2k+1}\pi^{(1)}_{m+2k-1} + \varepsilon\left(3+\frac{8}{\rho(\delta')^2}\right)\right] \\ &+ \frac{1}{\nu^{m+2k}}\left[1-\rho_{m+2k+1} + 2\rho_{m+2k+1}\pi^{(1)}_{m+2k-1} + (\delta')^5\left(3+\frac{8}{\rho(\delta')^2}\right)\right] \\ &+ \frac{1}{\nu^{m+2k+1}} < \gamma'\Theta^{(1)}_{m+2k-1} + \frac{2}{\nu^{m+2k}} + \frac{1}{\nu^{m+2k+1}} < \gamma'(\gamma')^{k-1}R\Theta^{(1)}_m + \frac{1}{\nu^m}\gamma' \\ &\left(3\sum_{l=k}^{2(k-1)}(\gamma')^l + (\gamma')^{2k-1}\right) + 2\frac{1}{\nu^{m+2k}} + \frac{1}{\nu^{m+2k+1}} \\ &< (\gamma')^k R\Theta^{(1)}_m + \frac{1}{\nu^m}\left(3\sum_{l=k+1}^{2k-1}(\gamma')^l + (\gamma')^{2k+1}\right),\end{aligned}$$

which is, for almost every $\omega \in \mathscr{E}^{(1)}_>$ considered, the sought estimate: (4.57).

We conclude that, by factoring out $(\gamma')^k$, (4.56) and (4.57) yield

$$\begin{aligned}\Theta^{(1)}_{m+2k} &< (\gamma')^k\left(R\Theta_m + \frac{3}{\nu^m(1-\gamma')}\right) \\ \Theta^{(1)}_{m+2k+1} &< (\gamma')^k\left(R\Theta_m + \frac{3}{\nu^m(1-\gamma')}\right).\end{aligned}$$

Recall that $\gamma := \sqrt{\gamma'}$. We have shown that for almost every $\omega \in \mathscr{E}^{(1)}_>$, for all integers $1 \le l \le \sigma - m$,

$$\Theta^{(1)}_{m+l} < (\gamma')^{\lfloor \frac{l}{2} \rfloor}\left(R\Theta^{(1)}_m + \frac{3}{\nu^m(1-\gamma')}\right).$$

Since

$$\left\lfloor \frac{l}{2} \right\rfloor \ge \frac{l-1}{2},$$

it follows that

$$(\gamma')^{\lfloor \frac{l}{2} \rfloor} < \sqrt{\gamma'}^{\,l-1}.$$

Then the two-steps geometric decaying upper bound can be expressed as a one-step geometric decaying upper bound. It has been shown that for almost every $\omega \in \mathscr{E}_>^{(1)}$, for all integers $1 \leq l \leq \sigma - m$,

$$\Theta_{m+l}^{(1)} < \gamma^{l-1} \left(R\Theta_m^{(1)} + \frac{3}{\nu^m(1-\gamma^2)} \right).$$

Hence for almost every $\omega \in \mathscr{E}_>^{(1)}$, for all $m \leq n \leq \sigma$,

$$\Theta_n^{(1)} < \gamma^{n-m-1} \left(R\Theta_m^{(1)} + \frac{3}{\nu^m(1-\gamma^2)} \right).$$

□

For every $\omega \in \mathscr{E}_>^{(1)}$ fixed, for any $\tau \geq m$ define a hitting time

$$\zeta := \inf \left\{ n > \tau : \frac{|\Theta_{n+1}^{(2)} - \Theta_n^{(2)}|}{R\Theta_n^{(1)} + \nu^{-n-1}} < \frac{1}{\Gamma} \right\}.$$

Lemma 4.41. *For almost every $\omega \in \mathscr{E}_>^{(1)}$ such that there exists a time $m(\omega) \leq \tau(\omega) < \sigma(\omega)$, such that*

$$\frac{R\Theta_{\tau(\omega)}^{(1)}(\omega) + \nu^{-\tau-1}}{|\Theta_{\tau(\omega)+1}^{(2)}(\omega) - \Theta_{\tau(\omega)}^{(2)}(\omega)|} \leq \Gamma,$$

it holds that for all $\tau(\omega) \leq n \leq \zeta(\omega) \wedge \sigma(\omega)$,

$$|\Theta_{n+1}^{(2)}(\omega) - \Theta_n^{(2)}(\omega)| < \tilde{\rho}^{n-m} |\Theta_{m+1}^{(2)}(\omega) - \Theta_m^{(2)}(\omega)|.$$

Proof. We show the claim for almost every $\omega \in \mathscr{E}_>^{(1)} \cap \{\tau = m\}$ first. For every ω considered, $\zeta = m+1$, thus $\zeta \wedge \sigma = m+1$; for almost every ω considered it is also known that:

- $\Theta_{m+1}^{(1)} < R\Theta_m^{(1)} + \nu^{-m-1}$ by Remark 4.35 with $n = m$ holds;
- the condition $R\Theta_m^{(1)} + \nu^{-m-1} \leq \Gamma |\Theta_{m+1}^{(2)} - \Theta_m^{(2)}|$ is satisfied;
- the hypotheses made in (4.36) to (4.38) are satisfied.

Since trivially both $\Theta_m^{(1)}, \nu^{-m-1} < R\Theta_m^{(1)} + \nu^{-m-1}$ as $R > 1$, by (4.63) to (4.67) it follows, for almost every ω considered, that

$$\begin{aligned}
\eta_m &< \Theta_m^{(1)} < \Gamma |\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| \\
\eta_{m+1} &< R\Theta_m^{(1)} + \frac{1}{\nu^{m+1}} < \Gamma |\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| \\
\eta'_{m+1} &< \frac{1}{(\delta')^2} \Theta_m^{(1)} < \frac{\Gamma}{(\delta')^2} |\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| \\
\eta''_{m+1} &< \frac{1}{\delta'} \left(R\Theta_m^{(1)} + \frac{1}{\nu^{m+1}} \right) < \frac{1}{\delta'} \Gamma |\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| \\
\eta'''_{m+1} &< \Theta_m^{(1)} < \Gamma |\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| \\
\xi_{m+1} &< \left[1 + \frac{1}{\delta'} \left(\frac{2}{\rho} - 1 \right) \right] \left(R\Theta_m^{(1)} + \frac{1}{\nu^{m+1}} \right) < \left[1 + \frac{1}{\delta'} \left(\frac{2}{\rho} - 1 \right) \right] \Gamma |\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| \\
\xi'_{m+1} &< \frac{1}{\delta'} \left(\frac{2}{\rho} - 1 \right) \Theta_m^{(1)} < \frac{1}{\delta'} \left(\frac{2}{\rho} - 1 \right) \Gamma |\Theta_{m+1}^{(2)} - \Theta_m^{(2)}| \\
|r'_{m+1}| &< \frac{2}{(\delta')^2 \nu^{m+1}} < \frac{2}{(\delta')^2} \Gamma |\Theta_{m+1}^{(2)} - \Theta_m^{(2)}|
\end{aligned}$$

Plugging these estimates into Remark 4.37 applied to $n = m$ yields, for almost every ω considered,

$$|\Theta^{(2)}_{m+2} - \Theta^{(2)}_{m+1}| < \tilde{\rho}\pi^{(1)}_m \left(1 + \Gamma\tilde{D}\right) |\Theta^{(2)}_{m+1} - \Theta^{(2)}_m|.$$

Since by construction $\Gamma < \delta'/[\tilde{D}(1-\delta')]$,

$$\pi^{(1)}_m(1 + \Gamma\tilde{D}) < \pi^{(1)}_m \left(1 + \frac{\delta'}{1-\delta'}\right) = \frac{\pi^{(1)}_m}{1-\delta'} \leq 1,$$

and therefore it follows that

$$|\Theta^{(2)}_{m+2} - \Theta^{(2)}_{m+1}| \leq \tilde{\rho}|\Theta^{(2)}_{m+1} - \Theta^{(2)}_m|$$

for almost every ω considered, and the first step of the induction is complete.

If ω is such that $\zeta(\omega) > m+1$, then $\zeta(\omega) \wedge \sigma(\omega) > m+1$. We have already seen, in the previous step, that $\delta' < \delta < \pi^{(1)}_m < 1-\delta < 1-\delta'$ is crucial for the argument, so we will have to make sure it iterates for almost every ω considered. Recall that

$$\varepsilon < \frac{\rho\delta'}{4\left(1 + \frac{R+\frac{1}{1-\gamma^2}}{1-\gamma}\right)}.$$

Since for all $m < n \leq \zeta(\omega) \wedge \sigma(\omega)$, by Lemma 4.40, it holds that for almost every such ω, $\Theta^{(1)}_n < \gamma^{n-m-1}\left(R\Theta^{(1)}_m + 3\nu^{-m}/(1-\gamma)\right)$, or equivalently

$$\Theta^{(1)}_{m+k} < \gamma^{k-1}\left(R\Theta^{(1)}_m + \frac{3}{\nu^m(1-\gamma^2)}\right) < \gamma^{k-1}\left(R + \frac{1}{1-\gamma^2}\right)\varepsilon'$$

for all $k \in \mathbb{N}$ such that $n = m+k$ is within the bounds above; by (4.41), for almost every ω considered, for all such k's

$$\pi^{(1)}_{m+k} \geq \begin{cases} \pi^{(1)}_m - \frac{4}{\rho}\sum_{j=0}^k \Theta^{(1)}_{m+j} \geq \delta - \frac{4}{\rho}\varepsilon\left(1 + \left(R + \frac{1}{1-\gamma^2}\right)\sum_{j=0}^k \gamma^j\right) & \\ > \delta - \frac{4}{\rho}\varepsilon\left(1 + \frac{R+\frac{1}{1-\gamma^2}}{1-\gamma}\right) > \delta', & k \text{ even} \\ 1 - \pi^{(1)}_m - \frac{4}{\rho}\sum_{j=0}^k \Theta^{(1)}_{m+j} \geq \delta - \frac{4}{\rho}\varepsilon\left(1 + \left(R + \frac{1}{1-\gamma^2}\right)\sum_{j=0}^k \gamma^j\right) & \\ \geq \delta - \frac{4}{\rho}\varepsilon\left(1 + \frac{R+\frac{1}{1-\gamma^2}}{1-\gamma}\right) > \delta', & k \text{ odd.} \end{cases}$$

by construction of ε. This ensures that when estimating η'_n, ξ_n, ξ'_n and r'_n with the constants, which upper bound the reciprocals of the $\pi^{(1)}$-component, can carry out during the induction step, for almost every ω considered. As to the constant, which lower bound the reciprocals involving the $\pi^{(1)}$-component, one can proceed analogously for almost every ω considered:

$$\pi^{(1)}_{m+k} \leq \begin{cases} \pi^{(1)}_m + \frac{4}{\rho}\sum_{j=0}^k \Theta^{(1)}_{m+j} \leq 1 - \delta + \frac{4}{\rho}\varepsilon\left(1 + \left(R + \frac{1}{1-\gamma^2}\right)\sum_{j=0}^k \gamma^j\right) & \\ < 1 - \delta + \frac{4}{\rho}\varepsilon\left(1 + \frac{R+\frac{1}{1-\gamma^2}}{1-\gamma}\right) < 1 - \delta', & k \text{ even} \\ 1 - q^{(1)}_m + \frac{2-\rho}{\rho}\sum_{j=0}^k p^{(1)}_{m+j} \leq 1 - \delta + \frac{4}{\rho}\varepsilon\left(1 + \left(R + \frac{1}{1-\gamma^2}\right)\sum_{j=0}^k \gamma^j\right) & \\ < 1 - \delta + \frac{4}{\rho}\varepsilon\left(1 + \frac{R+\frac{1}{1-\gamma^2}}{1-\gamma}\right) < 1 - \delta', & k \text{ odd.} \end{cases}$$

The inductive hypothesis is then that, for almost every ω considered, for some $k \geq 0$ such that $m(\omega)+k+1 < \zeta(\omega) \wedge \sigma(\omega)$, $|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}| < \tilde{\rho}^k|\Theta^{(2)}_{m+1} - \Theta^{(2)}_m|$, and it needs to be shown that $|\Theta^{(2)}_{m+k+2} - \Theta^{(2)}_{m+k+1}| < \tilde{\rho}^{k+1}|\Theta^{(2)}_{m+1} - \Theta^{(2)}_{m}|$ for almost every such ω, which will be done by showing that $|\Theta^{(2)}_{m+k+2} - \Theta^{(2)}_{m+k+1}| < \tilde{\rho}|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}|$ for almost every such ω. Since by the definition of σ, it still holds that $\delta' \leq \Theta^{(2)}_{m+k+1} \leq 1 - \delta'$, by the definition of ζ it still holds that $R\Theta^{(1)}_{m+k} + \nu^{m+k+1} \leq \Gamma|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}|$, and the geometric decaying upper bound of Lemma 4.40 ensures the bounds on $\pi^{(1)}_{m+k}$ and $\pi^{(1)}_{m+k+1}$ for almost every ω considered; it follows that for almost every ω considered,

$$\begin{aligned}
\eta_{m+k} &< \Theta^{(1)}_{m+k} < \Gamma|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}| \\
\eta_{m+k+1} &< R\Theta^{(1)}_{m+k} + \frac{1}{\nu^{m+k+1}} < \Gamma|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}| \\
\eta'_{m+k+1} &< \frac{1}{(\delta')^2}\Theta^{(1)}_{m+k} < \frac{\Gamma}{(\delta')^2}|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}| \\
\eta''_{m+k+1} &< \frac{1}{\delta'}\left(R\Theta^{(1)}_{m+k} + \frac{1}{\nu^{m+k+1}}\right) < \frac{1}{\delta'}\Gamma|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}| \\
\eta'''_{m+k+1} &< \Theta^{(1)}_{m+k} < \Gamma|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}| \\
\xi_{m+k+1} &< \left[1 + \frac{1}{\delta'}\left(\frac{2}{\rho} - 1\right)\right]\left(R\Theta^{(1)}_{m+k} + \frac{1}{\nu^{m+k+1}}\right) \\
&< \left[1 + \frac{1}{\delta'}\left(\frac{2}{\rho} - 1\right)\right]\Gamma|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}| \\
\xi'_{m+k+1} &< \frac{1}{\delta'}\left(\frac{2}{\rho} - 1\right)\Theta^{(1)}_{m+k} < \frac{1}{\delta'}\left(\frac{2}{\rho} - 1\right)\Gamma|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}| \\
|r'_{m+k+1}| &< \frac{2}{(\delta')^2\nu^{m+k+1}} < \frac{2}{(\delta')^2}\Gamma|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}|.
\end{aligned}$$

Plugging these estimates into Remark 4.37 applied to $n = m + k$ yields, for almost every ω considered,

$$|\Theta^{(2)}_{m+k+2} - \Theta^{(2)}_{m+k+1}| < \tilde{\rho}\pi^{(1)}_{m+k}\left(1 + \Gamma\tilde{D}\right)|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}|.$$

Since by construction $\Gamma < \delta'/[\tilde{D}(1 - \delta')]$,

$$\pi^{(1)}_{m+k}(1 + \Gamma\tilde{D}) < \pi^{(1)}_{m+k}\left(1 + \frac{\delta'}{1 - \delta'}\right) = \frac{\pi^{(1)}_{m+k}}{1 - \delta'} \leq 1$$

for almost every ω considered, and therefore it follows that

$$|\Theta^{(2)}_{m+k+2} - \Theta^{(2)}_{m+k+1}| \leq \tilde{\rho}|\Theta^{(2)}_{m+k+1} - \Theta^{(2)}_{m+k}|$$

for almost every ω considered, and the induction is complete.

If $\omega \in \mathscr{E}^{(1)}_{>} \cap \{m < \tau < \sigma\}$, one proceeds analogously: since $(\Theta_\tau, \pi_\tau) \in \mathcal{K}^*$ by Lemma 4.40 and definition of m and σ, all the estimates of the base case just completed apply with τ instead of m, and the inductive step stays the same. □

Theorem 4.42. *For almost every $\omega \in \mathscr{E}^{(1)}_{>}$, $\Theta_n(\omega) \longrightarrow \Theta_*(\omega) \in E_1$.*

Proof. On $\mathscr{E}^{(1)}_{>}$, define $\zeta_0 := m$, and a doubly sequence of (possibly infinite) hitting times $\{\zeta_i\}$, $\{\tau_i\}$ for all $i \in \mathbb{N}$, and the usual stopping time σ:

$$\begin{aligned}
\tau_i &:= \inf\left\{n \geq \zeta_{i-1} : \frac{|\Theta^{(2)}_{n+1} - \Theta^{(2)}_{n}|}{R\Theta^{(1)}_{n} + \nu^{-n-1}} \geq \frac{1}{\Gamma}\right\} \in \mathbb{N} \cup \infty \\
\zeta_i &:= \inf\left\{n \geq \tau_i : \frac{|\Theta^{(2)}_{n+1} - \Theta^{(2)}_{n}|}{R\Theta^{(1)}_{n} + \nu^{-n-1}} < \frac{1}{\Gamma}\right\} \in \mathbb{N} \cup \infty.
\end{aligned}$$

Note that for all $i \in \mathbb{N}$ and considered ω, such that $\zeta_{i-1}(\omega) < \infty$, $\tau_i(\omega) > \zeta_{i-1}(\omega)$, and for all $i \in \mathbb{N}$ and considered ω, such that $\tau_i(\omega) < \infty$, $\zeta_i(\omega) > \tau_i(\omega)$. We prove first that $\sigma = \infty$ for almost every $\omega \in \mathscr{E}_>^{(1)}$. To this end, we work first on $\mathscr{E}_>^{(1)} \cap \{\sigma < \infty\}$, and show that it is a negligible event, by reaching a contradiction for almost every ω in it. For any $\omega \in \mathscr{E}_>^{(1)} \cap \{\sigma < \infty\}$, $\infty \wedge \sigma = \sigma$ by convention, thus by also using empty sum convention, we have that the sum

$$|\Theta_\sigma^{(2)} - \Theta_m^{(2)}| \leq \sum_{i=0}^{\infty} \left(\sum_{n=\zeta_i \wedge \sigma}^{\tau_{i+1} \wedge \sigma - 1} |\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| + \sum_{n=\tau_{i+1} \wedge \sigma}^{\zeta_{i+1} \wedge \sigma - 1} |\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| \right)$$

is a finite sum, since ω is such that there is a $\bar{i}(\omega) \in \mathbb{N}$, for which either $\zeta_{\bar{i}} \leq \sigma < \tau_{\bar{i}+1}$ or $\tau_{\bar{i}} \leq \sigma < \zeta_{\bar{i}}$. Therefore, these two events partition $\mathscr{E}_>^{(1)} \cap \{\sigma < \infty\}$.

In the part of the event where $\zeta_{\bar{i}} \leq \sigma < \tau_{\bar{i}+1}$, for all $\zeta_{\bar{i}} \leq n \leq \sigma$, $|\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| < \Gamma^{-1}(R\Theta_n^{(1)} + \nu^{-n-1})$. By Lemma 4.40, the almost sure geometric decay of the $\Theta^{(1)}$-component carries on at least until $\Theta_\sigma^{(1)}$, which means that for almost all ω considered, for all $\zeta_{\bar{i}} \leq n < \sigma$,

$$\begin{aligned}
|\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| &< \Gamma^{-1} \left[R\Theta_n^{(1)} + \frac{1}{\nu^{n+1}} \right] < \Gamma^{-1} \Bigg[R\gamma^{n-m-1} \left(R\Theta_m^{(1)} + \frac{3}{1-\gamma^2} \nu^{-m} \right) \\
&\quad + \frac{1}{\nu^{n-m-1}} \frac{1}{\nu^m + 2} \Bigg] < \Gamma^{-1} \lambda^{n-m-1} \left[R \left(R\Theta_m^{(1)} + \frac{\nu^{-m}}{1-\gamma^2} \right) + \frac{1}{\nu^m} \right] \\
&< \Gamma^{-1} \lambda^{n-m-1} \left[R^2 \Theta_m^{(1)} + \left(\frac{R}{1-\gamma^2} + 1 \right) \nu^{-m} \right] < \frac{\tilde{R}}{\Gamma} \lambda^{n-m-1} \varepsilon' \\
&< \left(1 + \frac{\tilde{R}}{\Gamma} \right) \lambda^{n-m-1} \varepsilon',
\end{aligned}$$

and the same argument applies for all $\zeta_i \leq n < \tau_{i+1}$ for all $i < \bar{i}$, if there are any. On the other hand, for all $\tau_{\bar{i}} \leq n < \zeta_{\bar{i}}$, a different argument is needed (and similarly for all $\tau_i \leq n < \zeta_i$ for $i < \bar{i}$, if there are any). In fact, for the ω considered, rearranging the condition in the hitting times definition, for all $\tau_{\bar{i}} \leq n < \zeta_{\bar{i}}$ one has that

$$\frac{\Theta_n^{(1)}}{|\Theta_{n+1}^{(2)} - \Theta_n^{(2)}|} \leq \Gamma,$$

which is the type of condition relating to Lemma 4.41. This condition, for $n = \tau_i$, yields that we can apply Lemma 4.41 for almost all ω considered, implying that

$$|\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| < \lambda^{n-\tau_{\bar{i}}+1} |\Theta_{\tau_{\bar{i}}}^{(2)} - \Theta_{\tau_{\bar{i}}-1}^{(2)}|$$

for all $\tau_{\bar{i}} \leq n \leq \zeta_{\bar{i}} = \zeta_{\bar{i}} \wedge \sigma$. Observe that $|\Theta_{\tau_{\bar{i}}}^{(2)} - \Theta_{\tau_{\bar{i}}-1}^{(2)}|$ falls in the range treated earlier, hence

$$|\Theta_{\tau_{\bar{i}}}^{(2)} - \Theta_{\tau_{\bar{i}}-1}^{(2)}| < \left(1 + \frac{\tilde{R}}{\Gamma} \right) \lambda^{\tau_{\bar{i}}-1-m} \varepsilon',$$

and therefore (note that it is for this very step that the addition $1 + \tilde{R}/\Gamma$ has been introduced) for almost every ω considered

$$\begin{aligned}
|\Theta_{n+1}^{(2)} - \Theta_n^{(2)}| &< \lambda^{n-\tau_{\bar{i}}+1} |\Theta_{\tau_{\bar{i}}}^{(2)} - \Theta_{\tau_{\bar{i}}-1}^{(2)}| < \lambda^{n-\tau_{\bar{i}}+1} \left(1 + \frac{\tilde{R}}{\Gamma} \right) \lambda^{\tau_{\bar{i}}-1-m-1} \varepsilon' \\
&= \left(1 + \frac{\tilde{R}}{\Gamma} \right) \lambda^{n-m-1} \varepsilon'.
\end{aligned}$$

In the part of the event where $\tau_{\bar{i}} \leq \sigma < \zeta_{\bar{i}}$, one proceeds the other way around, for almost every ω considered: for all $\tau_{\bar{i}} \leq n < \sigma = \zeta_{\bar{i}} \wedge \sigma$ and $\tau_i \leq n \leq \zeta_i = \zeta_i \wedge \sigma$ for all $i < \bar{i}$ (if there are any), $|\Theta^{(2)}_{n+1} - \Theta^{(2)}_n| < (1 + \tilde{R}/\Gamma)\, \lambda^{n-m-1}\varepsilon'$ by Lemma 4.41; while for all $\zeta_{\bar{i}-1} \leq n < \tau_{\bar{i}}$ for all $i \leq \bar{i}$, $|\Theta^{(2)}_{n+1} - \Theta^{(2)}_n| < (1 + \tilde{R}/\Gamma)\, \lambda^{n-m-1}\varepsilon'$ by Lemma 4.40.

In conclusion, for almost every $\omega \in \mathscr{E}^{(1)}_{>} \cap \{\sigma < \infty\}$, for every $m \leq n < \sigma$,

$$|\Theta^{(2)}_{n+1} - \Theta^{(2)}_n| < \left(1 + \frac{\tilde{R}}{\Gamma}\right) \lambda^{n-m-1}\varepsilon'.$$

Therefore, since by construction

$$\varepsilon' < \frac{\delta\Gamma\lambda(1-\lambda)}{2(\tilde{R} + \Gamma)},$$

$$|\Theta^{(2)}_\sigma - \Theta^{(2)}_m| \leq \sum_{n=m}^{\sigma-1} |\Theta^{(2)}_{n+1} - \Theta^{(2)}_n| < \left(1 + \frac{\tilde{R}}{\Gamma}\right) \varepsilon' \sum_{n=m}^{\sigma-1} \lambda^{n-m-1} = \left(1 + \frac{\tilde{R}}{\Gamma}\right) \frac{\varepsilon'}{\lambda} \sum_{i=0}^{\sigma-m-1} \lambda^i$$
$$< \left(1 + \frac{\tilde{R}}{C}\right) \frac{\varepsilon'}{\lambda} \sum_{i=0}^{\infty} \lambda^i = \varepsilon' \frac{\tilde{R} + \Gamma}{\Gamma\lambda(1-\lambda)} < \frac{\delta}{2}.$$

For almost every $\omega \in \mathscr{E}^{(1)}_{>} \cap \{\sigma < \infty\}$, one has reached a contradiction, since for almost every ω in the event, $\Theta^{(2)}_m \in [\delta, 1-\delta]$, and having $\Theta^{(2)}_\sigma$ travelled less than $\delta/2$ away from $\Theta^{(2)}_m$, for almost every such ω, we have that

$$\Theta^{(2)}_\sigma \in \left[\frac{\delta}{2}, 1 - \frac{\delta}{2}\right],$$

against the very definition of σ. The contradiction can only be escaped on a negligible event. Thus $\mathbb{P}(\mathscr{E}^{(1)}_{>} \cap \{\sigma < \infty\}) = 0$.

Since $\sigma(\omega) = \infty$ for almost every $\omega \in \mathscr{E}^{(1)}_{>}$, $\Theta^{(1)}_n \longrightarrow 0$ for almost every such ω by Lemma 4.40. As to $\Theta^{(2)}_n$, one can apply the strategy already used in the argument by contradiction, but for any n (replace σ with n and n with k for previous time indices, when necessary), getting - with respect to any possible realisation of the doubly sequence of random times, which defines some event in $\mathscr{E}^{(1)}_{>}$ - the geometric estimate

$$|\Theta^{(2)}_{n+1} - \Theta^{(2)}_n| < \left(1 + \frac{\tilde{R}}{\Gamma}\right) \lambda^{n-m-1}\varepsilon'$$

for almost every ω considered. This yields

$$\sum_{n=m}^{\infty} |\Theta^{(2)}_{n+1} - \Theta^{(2)}_n| < \infty.$$

Therefore $\Theta^{(2)}_n \longrightarrow \Theta^{(2)}_* \in [\delta/2, 1 - \delta/2]$ for almost every ω considered, yielding overall convergence of $\Theta_n \longrightarrow \Theta_* \in E_1$ for almost every $\omega \in \mathscr{E}^{(1)}_{>}$. □

Corollary 4.43. *For almost every $\omega \in \mathscr{E}^{(1)}_{>}$, as $\{\Theta_n(\omega)\}$ converges to some $\Theta_*(\omega) \in E_1$, $\{\pi_n(\omega)\}$ is asymptotically 2-periodic to $\{\pi_{\Theta_*} \pm \frac{\ell}{2} e_{-1}(\Theta_*)\}$.*

Proof. Recall that for almost every $\omega \in \mathscr{E}^{(1)}_{>}$, by definition of $m(\omega)$, $\pi^{(1)}_m(\omega) < 1/2 - \delta$. By Theorem 4.42 and Corollary 4.33, it is known that for almost every $\omega \in \mathscr{E}^{(1)}_{>}$,

$\{\Theta_{m+r}(\omega)\}$ converges to some $\Theta_*(\omega) \in E_1$, with $\alpha_{m+r} \longrightarrow 0$ and $|\beta_{m+r}| \longrightarrow \ell/2$ as $r \longrightarrow \infty$ for almost every such ω, where

$$\pi_{m+r} - \pi_{\Theta_{m+r}} = \alpha_{m+r} e_0(\Theta_{m+r}) + \beta_{m+r} e_{-1}(\Theta_{m+r}) = o_\omega(1) + \beta_{m+r} e_{-1}(\Theta_*).$$

Hence by Lemma 3.19 (h), we have that

$$\pi_{m+r} = \pi_{\Theta_*} + \beta_{m+r} e_{-1}(\Theta_*) + o_\omega(1)$$

with $|\beta_{m+r}| \longrightarrow \ell/2$. By Remark 4.30 it is also known that if ω is such that $\pi_{m+2k} \longrightarrow \pi_{\Theta_*} - \frac{\ell}{2} e_{-1}(\Theta_*)$ as $k \longrightarrow \infty$, then for almost every such ω, $\pi_{m+2k+1} \longrightarrow \pi_{\Theta_*} + \frac{\ell}{2} e_{-1}(\Theta_*)$ as $k \longrightarrow \infty$. This is the only option for the sample path considered, as in the argument of Lemma 4.40 (precisely in the inductive step) it has been shown how the even shifts of the $\pi^{(1)}$-component, starting from $m(\omega)$, stay below $1/2$, and the odd ones stay above (and this now trivially carries on for all k, since $\sigma = \infty$ for almost all such ω, by the argument in Theorem 4.42). Hence we obtain the asymptotic 2-periodicity of π_{m+r} as $r \longrightarrow \infty$, for almost every $\omega \in \mathscr{E}_>^{(1)}$. ☐

Remark 4.44. *On $\mathscr{E}_>^{(i)}$ with $i \in \{2,3\}$ one can proceed by exploiting the symmetry of the model, define σ, ζ_i and τ_i accordingly in terms of the corresponding coordinates, and show an analogous version of Theorem 4.42 for $i \in \{2,3\}$ as well, thus yielding convergence of $\{\Theta_n(\omega)\}$ to some $\Theta_*(\omega) \in \partial\Sigma \setminus V$ and asymptotic 2-periodicity of $\{\pi_n(\omega)\}$ to $\{\pi_{\Theta_*} \pm \frac{\ell}{2} e_{-1}(\Theta_*)\}$, for almost every $\omega \in \mathscr{E}_>$.*

4.6 Convergence of the stochastic process

In this section we put together all the convergence results gathered so far, so as to show firstly, the almost sure convergence of $\{\Theta_n\}$, secondly, that $\{\pi_n\}$ may or may not converge, depending on the event considered.

Proof of Theorem 1.1. For every $\omega \in \mathcal{D}$, Θ_n converges to one of the vertices by definition of $\mathcal{D}$, that is $\Theta_*(\omega) = v_i \in V$ for some $i \in \{1,2,3\}$. For almost every $\omega \in B$, $\Theta_n(\omega)$ converges within the interior of the simplex by Proposition 4.7, so $\Theta_*(\omega) \in \mathring{\Sigma}$. For almost every $\omega \in \mathscr{E}_0$, $\Theta_n(\omega)$ converges within an edge of the simplex by Remark 4.19, so $\Theta_*(\omega) \in \partial\Sigma \setminus V$. For almost every $\omega \in \mathscr{E}_>$, Θ_n converges within an edge of the simplex by Remark 4.44, so again $\Theta_*(\omega) \in \partial\Sigma \setminus V$. These sample paths' limits, denoted as $\Theta_*(\omega) := \lim_{n\longrightarrow\infty} \Theta_n(\omega)$, are defined almost everywhere, since $\mathcal{D}$, $\mathscr{B}$, $\mathscr{E}_0$, $\mathscr{E}_>$, almost surely partition Ω. Thus we can define, componentwise for all $\omega \in \Omega$, the random variables $\Theta^{(i)}(\omega) := \limsup_{n\longrightarrow\infty} \Theta_n^{(i)}(\omega)$ for all $i \in \{1,2,3\}$. Then by construction, the stochastic process Θ_n converges almost surely to the random variable Θ, which is well-defined in $(\Omega, \mathscr{F}, \{\mathscr{F}_n\}, \mathbb{P})$ by the standard theory, since Θ is $\mathscr{F}_\infty$-measurable and almost surely Σ-valued, since by construction $\mathbb{P}(\omega : \ \Theta(\omega) = \Theta_*(\omega)) = 1$. ☐

Recall that Σ^* denotes the portion of Σ delimited by its medial triangle (boundary excluded).

Corollary 4.45. *Almost surely $\{\pi_n\}$ does not converge to any vertex of Σ.*

Proof. For almost all $\omega \in \{\ell = 0\}$, $\{\pi_n\}$ converges in $\overline{\Sigma}^*$ (which is bounded away from V), since for every $\Theta \in \Sigma$, $\pi_\Theta \in \overline{\Sigma}^*$, and for all such ω, $\pi_n - \pi_{\Theta_n} \longrightarrow \mathbf{0}$, with Θ_n converging for almost every ω by Theorem 1.1. For almost every $\omega \in \mathcal{D}_>$, $\{\pi_n\}$ diverges by Lemma 4.21. For almost every $\omega \in \mathscr{B}$, $\ell = 0$ by Proposition 4.7, so it is only left to consider $\omega \in \mathscr{E}_>$. For almost every such ω, $\{\pi_n\}$ diverges by Remark 4.44. ☐

4.7 Typical asymptotic behaviours of the ERBRW

The main result of this section is Theorem 1.3, through which it is shown that two typical asymptotics of the model are nonnegligible: convergence to an internal equilibrium (Θ, π_Θ), and convergence of $\{\Theta_n\}$ to the boundary (with $\{\pi_n\}$ diverging).

We first start by showing, in two lemmas, that with positive probability, in a large enough (finite) number of moves, the process can get arbitrarily close to an internal equilibrium configuration and to an oscillatory boundary configuration. This can be done by following a suitable algorithm (whose steps have positive probability), which forces the system as close to either of the configurations, as many steps are run. For ease,

$$C_0 := \left\{ \left(\frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{1}}{\mathbf{3}} \right) \right\}$$

will be chosen for the internal equilibrium configuration and

$$C_> := \left\{ \left(\left(0, \frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{4}, \frac{3}{8}, \frac{3}{8}\right) \right), \left(\left(0, \frac{1}{2}, \frac{1}{2}\right), \left(\frac{3}{4}, \frac{1}{8}, \frac{1}{8}\right) \right) \right\}$$

will be chosen for the oscillatory configuration at the boundary. Consider firstly, keeping Figure 1.1 in mind, that with positive probability all particles can be, at some finite time, at vertex 1 (this argument is obviously symmetric, but for simplicity we base it at vertex 1). Indeed, since for all $i \in \{1, 2, 3\}$, $T_0^{(i)} > 0$, if the initial particle is not there at time zero, there is a positive probability $\phi(\Theta_0^{(i)}, \Theta_0^{(j)})^\mu$ (with i, j depending on which vertex one starts with) that after branching, all particles go to vertex 1. More generally, at any time m, when particles are more spread out, there is a positive probability that they all go to vertex 1: first, conditionally on $\mathscr{F}_m$, there is a positive probability $\phi(\Theta_m^{(3)}, \Theta_m^{(2)})^{\mu N_m^{(1)}} \phi(\Theta_m^{(1)}, \Theta_m^{(3)})^{\mu N_m^{(2)}} \phi(\Theta_m^{(1)}, \Theta_m^{(2)})^{\mu N_m^{(3)}}$ that at time $m+1$, after branching, all the particles in vertex 1 go to vertex 2, those in vertex 2 go to vertex 3 and those in vertex 3 go to vertex 2, merging with those that previously were in vertex 1. Similarly, once the particles are all in vertex 2 or vertex 3, conditionally on this event, there is a positive probability (based on the updated parameters entering the binomials' probability parameters) that, after branching, all particles go to vertex 1 from both vertex 2 and vertex 3. Since vertex 1 had no particles at time $m+1$, at time $m+2$ all particles are in vertex 1. It is clear, after this first discussion, that using the binomials' probability parameters and conditioning on previous events, the particles can make any prescribed move (in accordance to the model) across the triangle, after each branching, with positive probability, so we will omit, for simplicity, the specific probability and conditioning performed at each step. Since with positive probability at some time the system can be forced to have all particles in vertex 1, we will assume this, without loss of generality, to be the starting point for the algorithm that takes us to any configuration sought (let us call this *Step 0*). Without loss of generality, we can also assume, due to previous considerations, that $m \geq 0$, the starting time of the algorithm, is chosen large enough so that the divisions in *Step 1-2* produce all nonzero quotients (this will simply make the intuition behind the algorithm easier, it is not a necessary assumption). Do not confuse this deterministic m with the random time $m(\omega)$ in the previous sections: from now on m will solely denote a deterministic starting time for our algorithm. The role of $m(\omega)$ will later on be taken up by N. Finally, as customary, for simplicity of exposition, we will always adopt the particle-like point of view, rather than the mass-like one. This will come with no loss of generality, since all Euclidean division involved in the argument naturally extend to divisions with nonintegral dividend, by allowing for a nonintegral remainder.

Lemma 4.46. *With positive probability, at a time large enough the stochastic process is arbitrarily close to the configuration C_0.*

Proof.

- *Step 1* Starting with all particles at vertex 1 at some time m, with positive probability each of these moves can be performed iteratively after they branch and become $\mu N_m^{(1)}$ particles: divide these particles in three lots (possibly with remainder), that is $\mu N_m^{(1)} = 3q_m + r_m$, with $r_m \in \{0,1,2\}$ and $q_m \geq 0$, and send $2q_m$ to vertex 3, $q_m + r_m$ to vertex 2. After this is done, $N_{m+1}^{(1)} = 0$, $N_{m+1}^{(2)} = q_m + r_m$, $N_{m+1}^{(3)} = 2q_m$.

- *Step 2* Then they branch, so that in vertex 1 there is nothing, in vertex 2 there are $\mu q_m + \mu r_m$ particles and in vertex 3, $2\mu q_m$. Divide now the μr_m particles into three lots again: $\mu r_m = 3q'_{m+1} + r_{m+1}$, with $r_{m+1} \in \{0,1,2\}$ and $q'_{m+1} \geq 0$. Denote $q_{m+1} = \mu q_m$. From vertex 2 send $q_{m+1} + q'_{m+1} + r_{m+1}$ to vertex 3, and $2q'_{m+1}$ to vertex 1. *While these particles are travelling along the edges* (this clause means simply that we always refer to the number of particles before the movements had begun: even though we have to state the moves sequentially, according to the model they happen all simultaneously), from vertex 3 send q_{m+1} particles to vertex 1 and q_{m+1} to vertex 2. Then $N_{m+2}^{(1)} = q_{m+1} + 2q'_{m+1}$, $N_{m+2}^{(2)} = q_{m+1}$ and $N_{m+2}^{(3)} = q_{m+1} + q'_{m+1} + r_{m+1}$.

- *Step 3* The particles now branch again, so we end up with $\mu q_{m+1} + 2\mu q'_{m+1}$ particles in vertex 1, μq_{m+1} in vertex 2 and $\mu q_{m+1} + \mu q'_{m+1} + \mu r_{m+1}$ in vertex 3. Divide μr_{m+1} into three lots again: $\mu r_{m+1} = 3q'_{m+2} + r_{m+2}$ with $r_{m+2} \in \{0,1,2\}$, $q'_{m+2} \geq 0$. From vertex 2 send all the μq_{m+1} particles to vertex 3, where we send also $\mu q'_{m+1}$ from vertex 1. While these particles are travelling along the edges, from vertex 1 send all the remaining $\mu q_{m+1} + \mu q'_{m+1}$ to vertex 2 and from vertex 3 send $q'_{m+2} + r_{m+2}$ to vertex 2 as well. While these particles are travelling along the edges, the $\mu q_{m+1} + \mu q'_{m+1} + 2q'_{m+2}$ particles still remaining in vertex 3 go to vertex 1. Denote $q_{m+2} = \mu q_{m+1} + \mu q'_{m+1}$. Then, with all edge crossing updated accordingly, $N_{m+3}^{(1)} = q_{m+2} + 2q'_{m+2}$, $N_{m+3}^{(2)} = q_{m+2} + q'_{m+2} + r_{m+2}$ and $N_{m+3}^{(3)} = q_{m+2}$. This is the same partitioning of the imbalance as the previous one, but mirrored with respect to the second and the third vertex. Hence one only needs to show one more step, with the mirrored moves, and then cycle between these last two sets of moves to keep the imbalance between the vertices uniformly bounded.

- *Step 4* The particles now branch, so there are $\mu q_{m+2} + 2\mu q'_{m+2}$ particles in vertex 1, $\mu q_{m+2} + \mu q'_{m+2} + \mu r_{m+2}$ in vertex 2 and μq_{m+2} in vertex 3. Divide μr_{m+2} into three lots again: $\mu r_{m+2} = 3q'_{m+3} + r_{m+3}$ with $r_{m+3} \in \{0,1,2\}$, $q'_{m+3} \geq 0$. We now mirror the moves from *Step 3*: from vertex 1 send $\mu q'_{m+2}$ particles to vertex 2, where we also send all the μq_{m+2} particles from vertex 3. While these particles are travelling along the edges, from vertex 2 send the $q'_{m+3} + r_{m+3}$ particles just divided to vertex 3, where we also send all the remaining $\mu q_{m+2} + \mu q'_{m+2}$ from vertex 1. While these particles are travelling along the edges, send all the remaining $\mu q_{m+2} + \mu q'_{m+2} + 2q'_{m+3}$ particles from vertex 2 to vertex 1. Denote $q_{m+3} = \mu q_{m+2} + \mu q'_{m+2}$. Thus, with all edge crossings updated accordingly, we have $N_{m+4}^{(1)} = q_{m+3} + 2q'_{m+3}$, $N_{m+4}^{(2)} = q_{m+3}$ and $N_{m+4}^{(3)} = q_{m+3} + q'_{m+3} + r_{m+3}$. Go back to *Step 3* (obviously replacing the previous time index with the current one and repeat).

We now state more rigorously the cycle, implementing also a stopping condition for an arbitrarily small $\varepsilon > 0$ fixed.

After *Step1-2* are executed, let the time variable $n = m+1$ and repeat the following steps:

- *Step 3* The particles branch, so we end up with $\mu q_n + 2\mu q'_n$ particles in vertex 1, μq_n in vertex 2 and $\mu q_n + \mu q'_n + \mu r_n$ in vertex 3. Divide μr_n into three lots: $\mu r_n = 3q'_{n+1} + r_{n+1}$ with $r_{n+1} \in \{0, 1, 2\}$, $q'_{n+1} \geq 0$. From vertex 2 send all the μq_n particles to vertex 3, where we send also $\mu q'_n$ from vertex 1. From vertex 1 send all the remaining $\mu q_n + \mu q'_n$ to vertex 2, from vertex 3 send $q'_{n+1} + r_{n+1}$ to vertex 2 as well, while the remaining $\mu q_n + \mu q'_n + 2q'_{n+1}$ go to vertex 1. Denote $q_{n+1} = \mu q_n + \mu q'_n$. Then, with all edge crossing updated accordingly, $N^{(1)}_{n+2} = q_{n+1} + 2q'_{n+1}$, $N^{(2)}_{n+2} = q_{n+1} + q'_{n+1} + r_{n+1}$ and $N^{(3)}_{n+2} = q_{n+1}$. If $\|\pi_{n+2} - \frac{\mathbf{1}}{\mathbf{3}}\|_1 < \varepsilon$ and $\|\Theta_{n+2} - \frac{\mathbf{1}}{\mathbf{3}}\|_1 < \varepsilon$: **stop**. Otherwise: go to *Step 4*.
- *Step 4* The particles now branch, so there are $\mu q_{n+1} + 2\mu q'_{n+1}$ particles in vertex 1, $\mu q_{n+1} + \mu q'_{n+1} + \mu r_{n+1}$ in vertex 2 and μq_{n+1} in vertex 3. Divide μr_{n+1} into three lots again: $\mu r_{n+1} = 3q'_{n+2} + r_{n+2}$ with $r_{n+2} \in \{0, 1, 2\}$, $q'_{n+2} \geq 0$. We now mirror the moves from *Step 3*: from vertex 1 send $\mu q'_{n+1}$ particles to vertex 2, were we also send all the μq_{n+1} particles from vertex 3. From vertex 2 send the $q'_{n+2} + r_{n+2}$ particles just divided to vertex 3, where we also send all the remaining $\mu q_{n+1} + \mu q'_{n+1}$ from vertex 1. Lastly send all the remaining $\mu q_{n+1} + \mu q'_{n+1} + 2q'_{n+2}$ particles from vertex 2 to vertex 1. Denote $q_{n+2} = \mu q_{n+1} + \mu q'_{n+1}$. Thus, with all edge crossings updated accordingly, we have $N^{(1)}_{n+3} = q_{n+2} + 2q'_{n+2}$, $N^{(2)}_{n+3} = q_{n+2}$ and $N^{(3)}_{n+3} = q_{n+2} + q'_{n+2} + r_{n+2}$. If $\|\pi_{n+3} - \frac{\mathbf{1}}{\mathbf{3}}\|_1 < \varepsilon$ and $\|\Theta_{n+3} - \frac{\mathbf{1}}{\mathbf{3}}\|_1 < \varepsilon$: **stop**. Otherwise: $n \leftarrow n+2$ and go to *Step 3*.

It is clear from this algorithm that the largest imbalance between the $\{N^{(i)}_{n+1}\}$ is eventually less than $3q'_n + r_n = \mu r_{n-1} \leq 2\mu$ thanks to redistributing the newly generated quotients step after step. The largest imbalance between the increments of the $\{T^{(i)}_{n+1}\}$, when updating them, is $2\mu^2$. Indeed within the cycle, when redistributing the terms, either we send $\mu q_n + q'_{n+1} + r_{n+1}$ particles through edge 1, $\mu q_n + 2\mu q'_n + 2q'_{n+1}$ particles through edge 2 and $\mu q_n + \mu q'_n$ particles through edge 3 in *Step3* (in this case the largest imbalance is bounded by $3\mu q'_n + 3q'_{n+1} + r_{n+1} = 3\mu q'_n + \mu r_n = \mu^2 r_{n-1} \leq 2\mu^2$); or we send corresponding quantities in the successive step, but in a mirrored fashion: $\mu q_{n+1} + q'_{n+2} + r_{n+2}$ particles through edge 1, $\mu q_{n+1} + \mu q'_{n+1}$ particles through edge 2 and $\mu q_{n+1} + 2\mu q'_{n+1} + 2q'_{n+2}$ particles through edge 3 (in this case the largest imbalance is bounded by $3\mu q'_{n+1} + 3q'_{n+2} + r_{n+2} = 3\mu q'_{n+1} + \mu r_{n+1} = \mu^2 r_n \leq 2\mu^2$). Hence

$$\begin{aligned}\left\|\pi_n - \frac{\mathbf{1}}{\mathbf{3}}\right\|_1 &= \sum_i \left|\pi^{(i)}_n - \frac{1}{3}\right| = \sum_i \left|\frac{3N^{(i)}_n - \sigma_n}{3\sigma_n}\right| = \sum_i \left|\frac{3N^{(i)}_n - \sum_j N^{(j)}_n}{3\sigma_n}\right| \\ &\leq \frac{1}{3\sigma_n}\sum_{i,j} |N^{(i)}_n - N^{(j)}_n| = \frac{2}{3\sigma_n}\sum_{i>j} |N^{(i)}_n - N^{(j)}_n| \leq \frac{12}{3\sigma_n}\mu = \frac{2\mu}{\sigma_n} \longrightarrow 0\end{aligned}$$

and similarly, no matter what $T^{(i)}_m$ one starts with, eventually the largest imbalance between the increments $\Delta T^{(i)}_n := T^{(i)}_{n+1} - T^{(i)}_n \leq 2\mu^2$, and therefore

$$\begin{aligned}\left\|\Theta_n - \frac{\mathbf{1}}{\mathbf{3}}\right\|_1 &= \sum_i \left|\Theta^{(i)}_n - \frac{1}{3}\right| = \sum_i \left|\frac{3T^{(i)}_n - \tau_n}{3\tau_n}\right| = \sum_i \left|\frac{3T^{(i)}_n - \sum_j T^{(j)}_n}{3\tau_n}\right| \\ &\leq \frac{1}{3\tau_n}\sum_{i,j} |T^{(i)}_n - T^{(j)}_n| = \frac{2}{3\tau_n}\sum_{i>j} |T^{(i)}_n - T^{(j)}_n|\end{aligned}$$

$$= \frac{2}{3\tau_n}\sum_{i>j}\left|T_m^{(i)} - T_m^{(j)} + \sum_{k=m}^{n-1}\Delta T_k^{(i)} - \Delta T_k^{(j)}\right| \leq \frac{2}{3\tau_n}\sum_{i>j}|T_m^{(i)} - T_m^{(j)}|$$
$$+ \frac{2}{3\tau_n}\sum_{i>j}\left|\sum_{k=m}^{n-1}\Delta T_k^{(i)} - \Delta T_k^{(j)}\right| \leq \frac{2}{3\tau_n}\sum_{i>j}|T_m^{(i)} - T_m^{(j)}|$$
$$+ \frac{2}{3\tau_n}\sum_{k=m}^{n-1}\sum_{i>j}|\Delta T_k^{(i)} - \Delta T_k^{(j)}| \leq \frac{2}{3\tau_n}\sum_{i>j}|T_m^{(i)} - T_m^{(j)}| + \frac{12}{3\tau_n}\sum_{k=m}^{n-1}\mu^2$$
$$\leq \frac{2}{3\tau_n}\sum_{i>j}|T_m^{(i)} - T_m^{(j)}| + \frac{4\mu^2(n-m)}{\tau_n} \longrightarrow 0$$

as m is fixed and $\tau_n \sim \mu^{n+1}/(\mu-1)$. These considerations prove that the algorithm eventually stops and we are therefore arbitrarily close to the equilibrium sought. Also, it shows that, given the initial conditions $T_m^{(i)}$, there is a deterministic $N > m$, such that ever after both Θ_n and π_n) are ε-close to C_0 as long as the index is large enough and $n = N$ satisfies

$$2\frac{\mu}{\sigma_n} < \varepsilon$$
$$\frac{2}{3\tau_n}\sum_{i>j}|T_m^{(i)} - T_m^{(j)}| + \frac{4\mu^2(n-m)}{\tau_n} < \varepsilon,$$

with $\sigma_n = \mu^n$ and $\tau_n = \tau_0 + \mu(\mu^n - 1)/(\mu - 1)$ by definition. □

We now describe the algorithm to approach the configuration $C_>$.

Lemma 4.47. *With positive probability, at a time large enough the stochastic process is arbitrarily close to $C_>$.*

Proof.

- *Step 1* Starting with all particles at vertex 1 at some time m, after they branch and become $\mu N_m^{(1)}$ particles, divide them in eight lots (possibly with remainder), that is $\mu N_m^{(1)} = 8q_m + r_m$, with $0 \leq r_m \leq 7$ and $q_m \geq 0$, and send $8q_m$ to vertex 3, $8q_m + r_m$ to vertex 2. After this $N_{m+1}^{(1)} = 0$, $N_{m+1}^{(2)} = 8q_m + r_m$, $N_{m+1}^{(3)} = 8q_m$.

- *Step 2* Then they branch, so that in vertex 1 there is nothing, in vertex 2 there are $8\mu q_m + \mu r_m$ particles and in vertex 3, $8\mu q_m$. Divide now the μr_m particles in eight lots: $\mu r_m = 8q'_{m+1} + r_{m+1}$, with $0 \leq r_{m+1} \leq 7$ and $q'_{m+1} \geq 0$. Denote $q_{m+1} = \mu q_m$. Send to vertex 1, $q_{m+1} + 2q'_{m+1}$ particles from vertex 2 and q_{m+1} from vertex 3. While these particles are travelling along the edges, from vertex 3 send the remaining $3q_{m+1}$ particles to vertex 2 while from vertex 2 the remaining $3q_{m+1} + 6q'_{m+1}$ particles are sent to vertex 3. Then $N_{m+2}^{(1)} = 2q_{m+1} + 2q'_{m+1}$, $N_{m+2}^{(2)} = 3q_{m+1}$ and $N_{m+2}^{(3)} = 3q_{m+1} + 6q'_{m+1} + r_{m+1}$.

- *Step 3* The particles now branch again, so we end up with $2\mu q_{m+1} + 2\mu q'_{m+1}$ particles in vertex 1, $3\mu q_{m+1}$ in vertex 2 and $3\mu q_{m+1} + 6\mu q'_{m+1} + \mu r_{m+1}$ in vertex 3. Divide μr_{m+1} into eight lots again: $\mu r_{m+1} = 8q'_{m+2} + r_{m+2}$ with $0 \leq r_{m+2} \leq 7$, $q'_{m+2} \geq 0$. From vertex 2 send all the $3\mu q_{m+1}$ particles to vertex 1, where we send also $3\mu q_{m+1} + 6\mu q'_{m+1} + 6q'_{m+2}$ from vertex 3. While these particles are travelling along the edges, from vertex 1 send $\mu q_{m+1} + \mu q'_{m+1}$ to vertex 3 and $\mu q_{m+1} + \mu q'_{m+1}$ to vertex 2. While these particles are travelling along the edges, the $2q'_{m+2} + r_{m+2}$ particles still remaining in vertex 3 go to vertex 2. Denote $q_{m+2} = \mu q_{m+1} + \mu q'_{m+1}$. Then, with all the edge crossings updated accordingly, $N^{(1)}_{m+3} = 6q_{m+2} + 6q'_{m+2}$, $N^{(2)}_{m+3} = q_{m+2} + 2q'_{m+2} + r_{m+2}$ and $N^{(3)}_{m+3} = q_{m+2}$. This is the same partitioning of the imbalance as the previous one, but mirrored with respect to the second and the third vertex. Hence one only needs to do one more step, with the mirrored moves, and then cycle between these last two sets of moves to keep the imbalance between the vertices uniformly bounded. Note that only remainder terms were send through edge 1 (which is the edge whose edge crossings proportion will have to be proved vanishing).

- *Step 4* The particles now branch, so there are $6\mu q_{m+2} + 6\mu q'_{m+2}$ particles in vertex 1, $\mu q_{m+2} + 2\mu q'_{m+2} + \mu r_{m+2}$ in vertex 2 and μq_{m+2} in vertex 3. Divide μr_{m+2} into eight lots again: $\mu r_{m+2} = 8q'_{m+3} + r_{m+3}$ with $0 \le r_{m+3} \le 7$, $q'_{m+3} \ge 0$. We now mirror the moves from *Step 3* to produce an oscillation, while sending only remainder terms through edge 1: from vertex 1 send $3\mu q_{m+2} + 3\mu q'_{m+2}$ particles to vertex 2 and $3\mu q_{m+2} + 3\mu q'_{m+2}$ to vertex 3. While these particles are travelling along the edges, send all the μq_{m+2} particles in vertex 3 and $\mu q_{m+2} + 2\mu q'_{m+2}$ of the particles in vertex 2, to vertex 1. While these particles are travelling along the edges, from vertex 2 send the remaining $6q'_{m+3} + r_{m+3}$ particles just divided to vertex 3. Denote $q_{m+3} = \mu q_{m+2} + \mu q'_{m+2}$. Thus, with all the edge crossings updated accordingly, we have $N^{(1)}_{m+4} = 2q_{m+3} + 2q'_{m+3}$, $N^{(2)}_{m+4} = 3q_{m+3}$ and $N^{(3)}_{m+4} = 3q_{m+3} + 6q'_{m+3} + r_{m+3}$. Go back to *Step 3* (replacing the previus time index with the current one and repeat).

We now state more rigorously the cycle, implementing also a stopping condition for an arbitrarily small $\varepsilon > 0$ fixed.

After *Step1-2* are executed let the time variable $n = m+1$ and repeat the following steps:

- *Step 3* The particles branch, so we end up with $2\mu q_n + 2\mu q'_n$ particles in vertex 1, $3\mu q_n$ in vertex 2 and $3\mu q_n + 6\mu q'_n + \mu r_n$ in vertex 3. Divide μr_n into eight lots again: $\mu r_n = 8q'_{n+1} + r_{n+1}$ with $0 \le r_{n+1} \le 7$, $q'_{n+1} \ge 0$. From vertex 2 send all the $3\mu q_n$ particles to vertex 1, where we send also $3\mu q_n + 6\mu q'_n + 6q'_{n+1}$ from vertex 3. While these particles are travelling along the edges, from vertex 1 send $\mu q_n + \mu q'_n$ to vertex 3 and $\mu q_n + \mu q'_n$ to vertex 2. While these particles are travelling along the edges, the $2q'_{n+1} + r_{n+1}$ particles still remaining in vertex 3 go to vertex 2. Denote $q_{n+1} = \mu q_n + \mu q'_n$. Then, with all edge crossing updated accordingly, $N^{(1)}_{n+2} = 6q_{n+1} + 6q'_{n+1}$, $N^{(2)}_{n+2} = q_{n+1} + 2q'_{n+1} + r_{n+1}$ and $N^{(3)}_{n+2} = q_{n+1}$. Go to *Step 4*.

- *Step 4* The particles now branch, so there are $6\mu q_{n+1} + 6\mu q'_{n+1}$ particles in vertex 1, $\mu q_{n+1} + 2\mu q'_{n+1} + \mu r_{n+1}$ in vertex 2 and μq_{n+1} in vertex 3. Divide μr_{n+1} into eight lots again: $\mu r_{n+1} = 8q'_{n+2} + r_{n+2}$ with $0 \le r_{n+2} \le 7$, $q'_{n+2} \ge 0$. We now mirror the moves from *Step 3* to produce an oscillation, while sending only remainder terms through edge 1: from vertex 1 send $3\mu q_{n+1} + 3\mu q'_{n+1}$ particles to vertex 2 and $3\mu q_{n+1} + 3\mu q'_{n+1}$ to vertex 3. While these particles are travelling along the edges, send all the μq_{n+1} particles in vertex 3 and $\mu q_{n+1} + 2\mu q'_{n+1}$ of the particles in vertex 2 to vertex 1. While these particles are travelling along the edges, from vertex 2 send the remaining $6q'_{n+2} + r_{n+2}$ particles just divided to vertex 3. Denote $q_{n+2} = \mu q_{n+1} + \mu q'_{n+1}$. Thus, with all the edge crossings updated accordingly, we have $N^{(1)}_{n+3} = 2q_{n+2} + 2q'_{n+2}$, $N^{(2)}_{n+3} = 3q_{n+2}$ and $N^{(3)}_{n+3} = 3q_{n+2} + 6q'_{n+2} + r_{n+2}$. If $\|\pi_{n+2} - (3/4, 1/8, 1/8)\|_1 < \varepsilon$ and $\|\Theta_{n+2} - (0, 1/2, 1/2)\|_1 < \varepsilon$

and $\|\pi_{n+3} - (1/4, 3/8, 3/8)\|_1 < \varepsilon$ and $\|\Theta_{n+3} - (0, 1/2, 1/2)\|_1 < \varepsilon$: **stop**. Otherwise: $n \leftarrow n+2$ and go to *Step 3*.

The algorithm eventually stops because it is such that $\|\pi_{m+2k+1} - (3/4, 1/8, 1/8)\|_1 \longrightarrow 0$ and $\|\pi_{m+2k} - (1/4, 3/8, 3/8)\|_1 \longrightarrow 0$, while $\|\Theta_{m+k} - (0, 1/2, 1/2)\|_1 \longrightarrow 0$ if let run without stopping condition (meaning, as $k \longrightarrow \infty$). This can be seen as follows:

$$\left\|\pi_{m+2k+1} - \left(\frac{3}{4}, \frac{1}{8}, \frac{1}{8}\right)\right\|_1 = \left|\pi^{(1)}_{m+2k+1} - \frac{3}{4}\right| + \sum_{i\in\{2,3\}} \left|\pi^{(i)}_{m+2k+1} - \frac{1}{8}\right| = \left|\frac{N^{(1)}_{m+2k+1}}{\sigma_{m+2k+1}} - \frac{3}{4}\right|$$
$$+ \sum_{i\in\{2,3\}} \left|\frac{N^{(i)}_{m+2k+1}}{\sigma_{m+2k+1}} - \frac{1}{8}\right| = \left|\frac{N^{(1)}_{m+2k+1}}{8q_{m+2k} + 8q'_{m+2k} + r_{m+2k}} - \frac{3}{4}\right|$$
$$+ \sum_{i\in\{2,3\}} \left|\frac{N^{(i)}_{m+2k+1}}{8q_{m+2k} + 8q'_{m+2k} + r_{m+2k}} - \frac{1}{8}\right| = \left|\frac{6q_{m+2k} + 6q'_{m+2k}}{8q_{m+2k} + 8q'_{m+2k} + r_{m+2k}} - \frac{3}{4}\right|$$
$$+ \left|\frac{q_{m+2k} + 2q'_{m+2k} + r_{m+2k}}{8q_{m+2k} + 8q'_{m+2k} + r_{m+2k}} - \frac{1}{8}\right| + \left|\frac{q_{m+2k}}{8q_{m+2k} + 8q'_{m+2k} + r_{m+2k}} - \frac{1}{8}\right|$$

$$= \left|\frac{24q_{m+2k} + 24q'_{m+2k} - 24q_{m+2k} - 24q'_{m+2k} - 3r_{m+2k}}{4\sigma_{m+2k+1}}\right|$$
$$+ \left|\frac{8q_{m+2k} + 16q'_{m+2k} + 8r_{m+2k} - 8q_{m+2k} - 8q'_{m+2k} - r_{m+2k}}{8\sigma_{m+2k+1}}\right|$$
$$+ \left|\frac{8q_{m+2k} - 8q_{m+2k} - 8q'_{m+2k} - r_{m+2k}}{8\sigma_{m+2k+1}}\right| = \frac{3r_{m+2k}}{4\sigma_{m+2k+1}} + \frac{8q'_{m+2k} + 7r_{m+2k}}{8\sigma_{m+2k+1}}$$
$$+ \frac{8q_{m+2k} +' r_{m+2k}}{8\sigma_{m+2k+1}} = \frac{3r_{m+2k}}{4\sigma_{m+2k+1}} + \frac{\mu r_{m+2k-1} + 6r_{m+2k}}{8\sigma_{m+2k+1}} + \frac{\mu r_{m+2k-1}}{8\sigma_{m+2k+1}}$$
$$\leq \frac{21}{4\sigma_{m+2k+1}} + \frac{7\mu + 42}{8\sigma_{m+2k+1}} + \frac{7\mu}{8\sigma_{m+2k+1}} \longrightarrow 0$$

and similarly

$$\left\|\pi_{m+2k} - \left(\frac{1}{4}, \frac{3}{8}, \frac{3}{8}\right)\right\|_1 = \left|\frac{2q_{m+2k-1} + 2q'_{m+2k-1}}{8q_{m+2k-1} + 8q'_{m+2k-1} + r_{m+2k-1}} - \frac{1}{4}\right|$$
$$+ \left|\frac{3q_{m+2k-1}}{8q_{m+2k-1} + 8q'_{m+2k-1} + r_{m+2k-1}} - \frac{3}{8}\right| + \left|\frac{3q_{m+2k-1} + 6q'_{m+2k-1} + r_{m+2k-1}}{8q_{m+2k-1} + 8q'_{m+2k-1} + r_{m+2k-1}} - \frac{3}{8}\right|$$
$$= \left|\frac{8q_{m+2k-1} + 8q'_{m+2k-1} - 8q_{m+2k-1} - 8q'_{m+2k-1} - r_{m+2k-1}}{4\sigma_{m+2k}}\right|$$
$$+ \left|\frac{24q_{m+2k-1} - 24q_{m+2k-1} - 24q'_{m+2k-1} - 3r_{m+2k-1}}{8\sigma_{m+2k}}\right|$$
$$+ \left|\frac{24q_{m+2k-1} + 48q'_{m+2k-1} + 8r_{m+2k-1} - 24q_{m+2k-1} - 24q'_{m+2k-1} - 3r_{m+2k-1}}{8\sigma_{m+2k}}\right|$$

$$= \frac{r_{m+2k-1}}{4\sigma_{m+2k}} + \frac{24q'_{m+2k-1} + 3r_{m+2k-1}}{8\sigma_{m+2k}} + \frac{24q'_{m+2k-1} + 5r_{m+2k-1}}{8\sigma_{m+2k}} = \frac{r_{m+2k-1}}{4\sigma_{m+2k-1}}$$
$$+ \frac{3\mu r_{m+2k-2}}{8\sigma_{m+2k}} + \frac{3\mu r_{m+2k-2} + 2r_{m+2k-1}}{8\sigma_{m+2k}} \leq \frac{7}{4\sigma_{m+2k}} + \frac{21\mu}{8\sigma_{m+2k}} + \frac{21\mu + 14}{8\sigma_{m+2k}} \longrightarrow 0.$$

Lastly it is easy to see that $\Theta^{(1)}_n \longrightarrow 0$, since for $n = m+2k$, edge 1 is traversed only by $2q'_{m+2k} + r_{m+2k} < 8q'_{m+2k} + r_{m+2k} = \mu r_{m+2k-1} \leq 7\mu$ particles, so $\Delta T^{(1)}_n < 7\mu$; for

$n = m+2k+1$, edge 1 is traversed only by $6q'_{m+2k+1} + r_{m+2k+1} < 8q'_{m+2k+1} + r_{m+2k} = \mu r_{m+2k} \leq 7\mu$ particles, so $\Delta T_n^{(1)} < 7\mu$ again. Hence eventually $\Delta T_n^{(1)} < 7\mu$, and therefore

$$\Theta_n^{(1)} = \frac{T_{m+1}^{(1)} + \sum_{k=m+1}^{n-1} \Delta T_k^{(1)}}{\tau_n} \leq \frac{T_{m+1}^{(1)}}{\tau_n} + \frac{\sum_{k=m+1}^{n-1} \Delta T_k^{(1)}}{\tau_n} \leq \frac{T_{m+1}^{(1)}}{\tau_n} + \frac{7\mu(n-m)}{\tau_n} \longrightarrow 0$$

due to the geometric asymptotics of τ_n and m being fixed. Now, having $\Theta_n^{(1)} \longrightarrow 0$, to show that the other two components tend to $1/2$, it is enough to show that $|\Theta_n^{(2)} - \Theta_n^{(3)}| \longrightarrow 0$, as all components add up to 1. Since for $n = m+2k$, edge 2 is traversed by $3\mu q_{m+2k} + 6\mu q'_{m+2k} + 6q'_{m+2k+1} + \mu q_{m+2k} + \mu q'_{m+2k} = 4\mu q_{m+2k} + 7\mu q'_{m+2k} + 6q'_{m+2k+1}$ particles and edge 3 by $3\mu q_{m+2k} + \mu q_{m+2k} + \mu q'_{m+2k} = 4\mu q_{m+2k} + \mu q'_{m+2k}$ particles, it follows that

$$|\Delta T_n^{(2)} - \Delta T_n^{(3)}| = 6\mu q'_{m+2k} + 6q'_{m+2k+1} \leq \mu^2 r_{m+2k-1} + \mu r_{m+2k} \leq 7(\mu^2 + \mu);$$

for $n = m+2k+1$, edge 2 is traversed by $3\mu q_{m+2k+1} + 3\mu q'_{m+2k+1} + \mu q_{m+2k+1} = 4\mu q_{m+2k+1} + 3\mu q'_{m+2k+1}$ particles and edge 3 by $3\mu q_{m+2k+1} + 3\mu q'_{m+2k+1} + \mu q_{m+2k+1} + 2\mu q'_{m+2k+1} = 4\mu q_{m+2k+1} + 5\mu q'_{m+2k+1}$ particles, so it follows that

$$|\Delta T_n^{(2)} - \Delta T_n^{(3)}| = 2\mu q'_{m+2k+1} \leq \mu^2 r_{m+2k} \leq 7\mu^2 < 7(\mu^2 + \mu).$$

In conclusion

$$|\Theta_n^{(2)} - \Theta_n^{(3)}| = \left| \frac{T_{m+1}^{(2)} - T_{m+1}^{(3)} + \sum_{k=m+1}^{n-1} \Delta T_k^{(2)} - \Delta T_k^{(3)}}{\tau_n} \right| \leq \frac{|T_{m+1}^{(2)} - T_{m+1}^{(3)}|}{\tau_n}$$
$$+ \frac{\sum_{k=m+1}^{n-1} |\Delta T_k^{(2)} - \Delta T_k^{(3)}|}{\tau_n} \leq \frac{|T_{m+1}^{(2)} - T_{m+1}^{(3)}|}{\tau_n} + \frac{7(\mu^2+\mu)(n-m)}{\tau_n} \longrightarrow 0.$$

Similarly to the conclusion of Lemma 4.46, this also shows that, given the initial conditions $T_m^{(i)}$, there is a deterministic $N > m$, such that ever after both Θ_n and π_n are ε-close to $C_>$. □

Since the algorithms in Lemmas 4.46 and 4.47 have been proved to force the system arbitrarily close to either of the two configurations in finite time; it is now enough to show that the asymptotic behaviour expected of the stochastic process follows almost surely in each of the two events, defined via the corresponding algorithm. This fact will be at the base of proving Theorem 1.3, which claims the following:

i) $\mathbb{P}(\mathscr{B}) > 0$;

ii) $\mathbb{P}(\mathscr{E}_>) > 0$.

Let us denote by $\mathscr{A}_\varepsilon^N(C_0)$ and $\mathscr{A}_\varepsilon^N(C_>)$ the events defined by performing N steps of either of the algorithms introduced in Lemmas 4.46 and 4.47 respectively, to force the system to be at least ε-close to either C_0 or $C_>$, with N large enough, such that $\varepsilon > 0$ is so small, that we can proceed either with the argument of Proposition 4.8 or with that of Theorem 4.42, if we were given that, for all $n \geq N$, for all $i \in \{1,2,3\}$, $|R_{n+1}^{(i)}| < \nu^{-n-1}$ (that is, the deterministic N takes on the role of the random $m(\omega)$ in Chapter 4). Clearly we are meant to proceed with Proposition 4.8 if the limiting configuration is C_0: in this case ε must meet the conditions of δ' in Proposition 4.8, with respect to a suitable ε' fixed so that $B(\mathbf{\frac{1}{3}}, \varepsilon')$ does not intersect $\partial\Sigma$. We proceed with Theorem 4.42 if the limiting configuration is $C_>$: in this case ε must meet all the conditions of ε' in Section 4.5.3, with respect to a suitable δ fixed, so that all the conditions put on δ in Section 4.5.3 apply with respect to $\Theta_* = (0, 1/2, 1/2)$, $\pi_* = (1/4, 3/8, 3/8)$, $\hat{\pi}_* = (3/4, 1/8, 1/8)$. Indeed Proposition 4.8 and Theorem 4.42 have as initial condition, that the system is close to either of the two types of configurations respectively, at the same time that the geometric decaying upper bound on the martingale increments starts holding. In the context of those theorems the time constraint on when the martingale increments start being geometrically upper bounded is avoided by proceeding pointwise in ω and relying on Lemma 4.1. But in the present context, that approach is not viable, as we are trying to show that the two typical asymptotic behaviours are nonnegligible. Rather, we have just the initial condition, to which the system is driven as close as long is the running time of the algorithm. However the driving force of the geometric decay is actually the branching of the particles, that is the factor $1/\sigma_n$ involved in the formula of $R^{(i)}_{n+1}$. This is the fundamental fact behind our confidence that, with positive probability, we can force the system until a time large enough is reached, at which the geometrically decaying upper bound has already started applying. As a last remark, before proceeding with the proof, one should note that the strategy just described would not work for the nonoscillatory boundary limiting point, say, for example, $((0, 1/2, 1/2), (1/2, 1/4, 1/4))$. Devising an algorithm to force the system near this point is not the issue: the proof of the convergence to such type of limit points relies more fundamentally than the others on the asymptotic definition of the event considered, namely, it requires knowledge that $\omega \in \{\ell = 0\}$ where ℓ is the limit of the potential. For this reason we will not prove that the system converges to nonoscillatory boundary limit points with positive probability. Moreover, on top of such theoretical grounds, the simulations do not show this behaviour at all, suggesting that it may be a negligible event for the stochastic process after all. To keep the notation as simple as possible, we will not repeat the following argument twice, once for $\mathscr{A}^N_\varepsilon(C_0)$ and once for $\mathscr{A}^N_\varepsilon(C_>)$, since, except for the conclusion, it is the same; hence we abuse the notation and simply write $\mathscr{A}^N_\varepsilon$ denoting either of the two.

Proof of Theorem 1.3.

Step 1. Define

$$\mathscr{A}_N := \bigcap_{n \geq N} A_n,$$

where

$$A_n := \bigcap_{i=1}^{3} A^{(i)}_n \cap \overline{A}^{(i)}_n$$

and, denoting the complementary binomials as $\overline{B}^{(i)}_{n+1} := \mu N^{(i)}_n - B^{(i)}_{n+1}$ for $i \in \{1, 2, 3\}$,

$$A^{(1)}_n = \left\{ \frac{|B^{(1)}_{n+1} - \mu N^{(1)}_n \phi(\Theta^{(3)}_n, \Theta^{(2)}_n)|}{\sigma_{n+1}} < \frac{1}{2\nu^{n+1}} \right\},$$
$$A^{(2)}_n = \left\{ \frac{|B^{(2)}_{n+1} - \mu N^{(2)}_n \phi(\Theta^{(1)}_n, \Theta^{(3)}_n)|}{\sigma_{n+1}} < \frac{1}{2\nu^{n+1}} \right\},$$
$$A^{(3)}_n = \left\{ \frac{|B^{(3)}_{n+1} - \mu N^{(1)}_n \phi(\Theta^{(2)}_n, \Theta^{(1)}_n)|}{\sigma_{n+1}} < \frac{1}{2\nu^{n+1}} \right\},$$

$$\overline{A}_n^{(1)} = \left\{ \frac{|\overline{B}_{n+1}^{(1)} - \mu N_n^{(1)} \phi(\Theta_n^{(2)}, \Theta_n^{(3)})|}{\sigma_{n+1}} < \frac{1}{2\nu^{n+1}} \right\},$$
$$\overline{A}_n^{(2)} = \left\{ \frac{|\overline{B}_{n+1}^{(2)} - \mu N_n^{(2)} \phi(\Theta_n^{(3)}, \Theta_n^{(1)})|}{\sigma_{n+1}} < \frac{1}{2\nu^{n+1}} \right\},$$
$$\overline{A}_n^{(3)} = \left\{ \frac{|\overline{B}_{n+1}^{(3)} - \mu N_n^{(3)} \phi(\Theta_n^{(1)}, \Theta_n^{(2)})|}{\sigma_{n+1}} < \frac{1}{2\nu^{n+1}} \right\}.$$

For nonintegral μ, the argument will have all $B_{n+1}^{(i)}$ replaced by $\tilde{B}_{n+1}^{(i)}$ in the corresponding positions. Since by Lemmas 4.46 and 4.47, $\mathbb{P}(\mathscr{A}_\varepsilon^N) > 0$, we can show that $\mathbb{P}(\mathscr{A}_\varepsilon^N \cap \mathscr{A}_N) = \mathbb{P}(\mathscr{A}_N|\mathscr{A}_\varepsilon^N)\mathbb{P}(\mathscr{A}_\varepsilon^N) > 0$, since on this event, the arguments of Proposition 4.8 and Theorem 4.42 can be started, each with respect to the corresponding limiting configuration considered implicitly in the notation $\mathscr{A}_\varepsilon^N$ (C_0 or $C_>$). This boils down to showing that $\mathbb{P}(\mathscr{A}_N|\mathscr{A}_\varepsilon^N) > 0$. Denote by $\mathscr{A}_N^c$ the complement of $\mathscr{A}_N$. By the monotonicity of the (conditional) probability measure,

$$\mathbb{P}(\mathscr{A}_N|\mathscr{A}_\varepsilon^N) = \mathbb{P}\left(\bigcap_{n\geq N} A_n \Big| \mathscr{A}_\varepsilon^N\right) = \lim_{M\longrightarrow\infty} \mathbb{P}\left(\bigcap_{n=N}^{M} A_n \Big| \mathscr{A}_\varepsilon^N\right). \tag{4.73}$$

Since by simply prescribing values for the binomial, as shown in the algorithmic part, it trivially holds that $\mathbb{P}\left(A_N \cap \mathscr{A}_\varepsilon^N\right) = \mathbb{P}\left(A_N|\mathscr{A}_\varepsilon^N\right)\mathbb{P}\left(\mathscr{A}_\varepsilon^N\right) > 0$, one can write

$$\mathbb{P}\left(\bigcap_{n=N}^{M} A_n \Big| \mathscr{A}_\varepsilon^N\right) = \mathbb{P}\left(\bigcap_{n=N}^{M} A_n \Big| \mathscr{A}_\varepsilon^N\right) = \frac{\mathbb{P}\left(\bigcap_{n=N+1}^{M} A_n \cap (A_N \cap \mathscr{A}_\varepsilon^N)\right)}{\mathbb{P}\left(A_N \cap \mathscr{A}_\varepsilon^N\right)} \frac{\mathbb{P}\left(A_N \cap \mathscr{A}_\varepsilon^N\right)}{\mathbb{P}\left(\mathscr{A}_\varepsilon^N\right)}$$
$$= \mathbb{P}\left(\bigcap_{n=N+1}^{M} A_n \Big| A_N \cap \mathscr{A}_\varepsilon^N\right) \mathbb{P}\left(A_N \Big| \mathscr{A}_\varepsilon^N\right).$$

As long as we consider finitely many moves, preforming sequentially one move at a time, one can always ensure by induction that, for every $n \leq k \leq M$,

$$\mathbb{P}\left(\bigcap_{n=N}^{k} A_n \cap \mathscr{A}_\varepsilon^N\right) = \mathbb{P}\left(A_k \Big| \bigcap_{n=N}^{k-1} A_n \cap \mathscr{A}_\varepsilon^N\right) \mathbb{P}\left(\bigcap_{n=N}^{k-1} A_n \cap \mathscr{A}_\varepsilon^N\right) > 0.$$

Therefore, if we denote for all $k \geq N$,

$$A_k^N := \bigcap_{n=N}^{k} A_n \cap \mathscr{A}_\varepsilon^N,$$

it we can see by induction that iterating yields

$$\mathbb{P}\left(\bigcap_{n=N}^{M} A_n \Big| \mathscr{A}_\varepsilon^N\right) = \prod_{k=N+1}^{M} \mathbb{P}\left(A_k \Big| A_{k-1}^N\right) \mathbb{P}\left(A_N \Big| \mathscr{A}_\varepsilon^N\right).$$

If we also denote $A_{N-1}^N := \mathscr{A}_\varepsilon^N$, we can express it in a more compact way:

$$\mathbb{P}\left(\bigcap_{n=N}^{M} A_n \Big| \mathscr{A}_\varepsilon^N\right) = \prod_{k=N}^{M} \mathbb{P}\left(A_k \Big| A_{k-1}^N\right) = \prod_{k=N}^{M} \left[1 - \mathbb{P}\left(A_k^c \Big| A_{k-1}^N\right)\right].$$

Step 2. In the last identity of *Step 1*, we expressed the probability using the complementary event, because it is easier to match the standard notation used for infinite products. Since we need to show that

$$0 < \lim_{M\longrightarrow\infty} \prod_{k=N}^{M} \left[1 - \mathbb{P}\left(A_k^c \middle| A_{k-1}^N\right)\right] = \prod_{k=N}^{\infty} \left[1 - \mathbb{P}\left(A_k^c \middle| A_{k-1}^N\right)\right], \tag{4.74}$$

one can simply apply the standard theory of infinite products of the form $\prod_{k=N}^{\infty}(1-p_k)$ with $0 < p_k < 1$, and it is a well known elementary fact that $\prod_{k=N}^{\infty}(1-p_k) = 0$ if and only if $\sum_{k=N}^{\infty} p_k = \infty$. In conclusion it is enough to show that

$$\sum_{k=N}^{\infty} \mathbb{P}\left(A_k^c \middle| A_{k-1}^N\right) < \infty, \tag{4.75}$$

in order to have $\mathbb{P}(\mathscr{A}_N|\mathscr{A}_\varepsilon^N) > 0$, which is the claim sought. We will do so by showing the geometric decay of the probabilities $\mathbb{P}\left(A_k^c \middle| A_{k-1}^N\right)$. First of all note that $A_k^c = \bigcup_{i=1}^{3} A_k^{(i)c} \cup \overline{A}_k^{(i)c}$, so

$$\mathbb{P}\left(A_k^c \middle| A_{k-1}^N\right) \leq \sum_{i=1}^{3} \mathbb{P}\left(A_k^{(i)c} \middle| A_{k-1}^N\right) + \mathbb{P}\left(\overline{A}_k^{(i)c} \middle| A_{k-1}^N\right) = \sum_{i=1}^{3} \mathbb{E}\left[\mathbb{P}_{\mathscr{F}_k}\left(A_k^{(i)c}\right) \middle| A_{k-1}^N\right] \\ + \mathbb{E}\left[\mathbb{P}_{\mathscr{F}_k}\left(\overline{A}_k^{(i)c}\right) \middle| A_{k-1}^N\right]. \tag{4.76}$$

In this step we show that the identity follows by the tower property. In fact note that the events A_{k-1}^N are $\mathscr{F}_k$-measurable, so $\mathscr{G}_k := \sigma\left(A_{k-1}^N\right) \subseteq \mathscr{F}_k$; let $\mathbb{I}_k$ be any of the indicator functions $\mathbb{1}_{A_k^{(i)c}}$ or $\mathbb{1}_{\overline{A}_k^{(i)c}}$, where $i \in \{1, 2, 3\}$; the identity then follows if we show that

$$\mathbb{E}\left(\mathbb{I}_k | A_{k-1}^N\right) = \mathbb{E}\left(\mathbb{E}_{\mathscr{F}_k} I_k | A_{k-1}^N\right). \tag{4.77}$$

This follows from the tower property applied to $\mathbb{E}(\mathbb{I}_k|\mathscr{F}_k|\mathscr{G}_k)$, which ensures that

$$\mathbb{E}(\mathbb{I}_k|\mathscr{F}_k|\mathscr{G}_k) = \mathbb{E}_{\mathscr{G}_k}(\mathbb{I}_k) = \mathbb{1}_{A_{k-1}^N}\mathbb{E}\left[\mathbb{I}_k|A_{k-1}^N\right] + \mathbb{1}_{(A_{k-1}^N)^c}\mathbb{E}\left[\mathbb{I}_k|(A_{k-1}^N)^c\right].$$

On the other hand

$$\mathbb{E}(\mathbb{I}_k|\mathscr{F}_k|\mathscr{G}_k) = \mathbb{E}_{\mathscr{G}_k}(\mathbb{E}_{\mathscr{F}_k}\mathbb{I}_k) = \mathbb{1}_{A_{k-1}^N}\mathbb{E}(\mathbb{E}_{\mathscr{F}_k}\mathbb{I}_k|A_{k-1}^N) + \mathbb{1}_{(A_{k-1}^N)^c}\mathbb{E}(\mathbb{E}_{\mathscr{F}_k}\mathbb{I}_k|(A_{k-1}^N)^c).$$

Choosing $\omega \in A_{k-1}^N$ turns the identity

$$\mathbb{1}_{A_{k-1}^N}\mathbb{E}\left[\mathbb{I}_k|A_{k-1}^N\right] + \mathbb{1}_{(A_{k-1}^N)^c}\mathbb{E}\left[\mathbb{I}_k|(A_{k-1}^N)^c\right] = \mathbb{1}_{A_{k-1}^N}\mathbb{E}\left[\mathbb{E}_{\mathscr{F}_k}\mathbb{I}_k|A_{k-1}^N\right] \\ + \mathbb{1}_{(A_{k-1}^N)^c}\mathbb{E}\left[\mathbb{E}_{\mathscr{F}_k}\mathbb{I}_k|(A_{k-1}^N)^c\right]$$

into (4.77).

Step 3. As a result of (4.77) in *Step 2*, it will be enough to bound geometrically all the terms $\mathbb{P}_{\mathscr{F}_k}\left(A_k^{(i)c}\right)$ and $\mathbb{P}_{\mathscr{F}_k}\left(\overline{A}_k^{(i)c}\right)$. The argument will not differ for any of these terms, since they are all of the same form, for what the argument is concerned. Hence we proceed with

$$\mathbb{P}_{\mathscr{F}_k}\left(A_k^{(1)c}\right) = \mathbb{P}_{\mathscr{F}_k}\left(\frac{|B_{k+1}^{(1)} - \mu N_k^{(1)}\phi(\Theta_k^{(3)}, \Theta_k^{(2)})|}{\sigma_{k+1}} \geq \frac{1}{2\nu^{k+1}}\right).$$

It is possible to further simplify this expression. For brevity, denote $P_k^{(i,j)} := \phi(\Theta_k^{(i)}, \Theta_k^{(j)})$. Since

$$\mathbb{P}_{\mathscr{F}_k}\left(\frac{|B_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(3,2)}|}{\sigma_{k+1}} \geq \frac{1}{2\nu^{k+1}}\right) = \mathbb{P}_{\mathscr{F}_k}\left(B_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(3,2)} \geq \frac{\sigma_{k+1}}{2\nu^{k+1}}\right)$$
$$+ \mathbb{P}_{\mathscr{F}_k}\left(B_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(3,2)} \leq -\frac{\sigma_{k+1}}{2\nu^{k+1}}\right) = \mathbb{P}_{\mathscr{F}_k}\left(B_{k+1}^{(1)} \geq \mu N_k^{(1)} P_k^{(3,2)} + \frac{\sigma_{k+1}}{2\nu^{k+1}}\right)$$
$$+ \mathbb{P}_{\mathscr{F}_k}\left(\mu N_k^{(1)} - (B_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(3,2)}) \geq \mu N_k^{(1)} + \frac{\sigma_{k+1}}{2\nu^{k+1}}\right)$$
$$= \mathbb{P}_{\mathscr{F}_k}\left(B_{k+1}^{(1)} \geq \mu N_k^{(1)} P_k^{(3,2)} + \frac{\sigma_{k+1}}{2\nu^{k+1}}\right) + \mathbb{P}_{\mathscr{F}_k}\left(\overline{B}_{k+1}^{(1)} \geq \mu N_k^{(1)} P_k^{(2,3)} + \frac{\sigma_{k+1}}{2\nu^{k+1}}\right)$$
$$= \mathbb{P}_{\mathscr{F}_k}\left(\frac{B_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(3,2)}}{\sigma_{k+1}} \geq \frac{1}{2\nu^{k+1}}\right) + \mathbb{P}_{\mathscr{F}_k}\left(\frac{\overline{B}_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(2,3)}}{\sigma_{k+1}} \geq \frac{1}{2\nu^{k+1}}\right)$$

and the two terms obtained can be dealt with similarly, it is enough to show the argument for the geometric decay of

$$\mathbb{P}_{\mathscr{F}_k}\left(\frac{B_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(3,2)}}{\sigma_{k+1}} \geq \frac{1}{2\nu^{k+1}}\right).$$

The main strategy in this step will be to find a Chernoff bound at a moderate deviations regime from the conditional mean. Rewrite the term as

$$\mathbb{P}_{\mathscr{F}_k}\left(\frac{X_{k+1}^{(1)}}{\sqrt{\sigma_{k+1}}} \geq \frac{\sqrt{\sigma_{k+1}}}{2\nu^{k+1}}\right),$$

where $X_{k+1}^{(1)} := B_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(3,2)}$ and therefore $\mathbb{E}_{\mathscr{F}_k} X_{k+1}^{(1)} = 0$ and $\operatorname{Var}_{\mathscr{F}_k} X_{k+1}^{(1)} = \mu N_k^{(1)} P_k^{(3,2)} P_k^{(2,3)}$. Clearly for nonintegral μ the only difference is that, having denoted $\tilde{X}_{k+1}^{(1)} := \tilde{B}_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(3,2)}$, $\operatorname{Var}_{\mathscr{F}_k} \tilde{X}_{k+1}^{(1)} = (\lfloor \mu N_k^{(1)} \rfloor + \{\mu N_k^{(1)}\}^2) P_k^{(3,2)} P_k^{(2,3)} \leq \mu N_k^{(1)} P_k^{(3,2)} P_k^{(2,3)}$. Recall that $1 < \nu < \sqrt{\mu}$, thus

$$\frac{\sqrt{\sigma_{k+1}}}{\nu^{k+1}} = \eta^{k+1},$$

with $\eta := \sqrt{\mu}/\nu > 1$. Therefore for any $\lambda \in \mathbb{R}$,

$$\mathbb{P}_{\mathscr{F}_k}\left(\frac{X_{k+1}^{(1)}}{\sqrt{\sigma_{k+1}}} \geq \frac{\eta^{k+1}}{2}\right) = \mathbb{P}_{\mathscr{F}_k}\left(\exp\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} X_{k+1}^{(1)}\right) \geq \exp\left(\frac{\lambda}{2}\eta^{k+1}\right)\right)$$
$$\leq \exp\left(-\frac{\lambda}{2}\eta^{k+1}\right) \mathbb{E}_{\mathscr{F}_k}\left(\exp\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} X_{k+1}^{(1)}\right)\right).$$

Denote the conditional generating function term as

$$\Phi_{\mathscr{F}_k}\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}}\right) := \mathbb{E}_{\mathscr{F}_k}\left(\exp\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} X_{k+1}^{(1)}\right)\right).$$

For nonintegral μ, having to use $\tilde{X}_{k+1}^{(1)}$ instead of $X_{k+1}^{(1)}$, we will denote it as $\tilde{\Phi}_{\mathscr{F}_k}$. Note that $\lambda/\sqrt{\sigma_{k+1}} \longrightarrow 0$ as $k \longrightarrow \infty$. The main idea here is that by the elementary properties of moment generating functions and the mean and variance of $X_{k+1}^{(1)}$, it is known that

$$\Phi_{\mathscr{F}_k}(0) = 1,\ \Phi_{\mathscr{F}_k}(0)' = 0,\ \Phi_{\mathscr{F}_k}(0)'' = \mu N_k^{(1)} P_k^{(3,2)} P_k^{(2,3)}.$$

Hence, as $k \longrightarrow \infty$, we should be able to exploit

$$\Phi_{\mathscr{F}_k}\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}}\right) \approx 1 + \frac{1}{2}\mu N_k^{(1)} P_k^{(3,2)} P_k^{(2,3)} \left(\frac{\lambda}{\sqrt{\sigma_{k+1}}}\right)^2 = 1 + \frac{\pi_k^{(1)}}{2} P_k^{(3,2)} P_k^{(2,3)} \lambda^2 = \mathcal{O}(1),$$

to conclude that this factor is indeed harmless to the overall bound (the conclusion of this heuristic argument for nonintegral μ yields similarly $\tilde{\Phi}_{\mathscr{F}_k}(\lambda/\sqrt{\sigma_{k+1}}) = \mathcal{O}(1)$, due to the estimate on the variance aforementioned). More rigorously, we can compute directly the conditional moment generating function:

$$\Phi_{\mathscr{F}_k}\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}}\right) = \left(P_k^{(3,2)} \exp\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}}\right) + P_k^{(2,3)}\right)^{\mu N_k^{(1)}} \exp\left(-\frac{\lambda}{\sqrt{\sigma_{k+1}}}\mu N_k^{(1)} P_k^{(3,2)}\right)$$
$$= \exp\left\{\mu N_k^{(1)} \log\left[P_k^{(3,2)} \exp\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}}\right) + P_k^{(2,3)}\right] - \frac{\lambda}{\sqrt{\sigma_{k+1}}}\mu N_k^{(1)} P_k^{(3,2)}\right\}.$$

For nonintegral μ, one can exploit the conditional independence of the Bernoulli from the binomial term, which yields

$$\tilde{\Phi}_{\mathscr{F}_k}\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}}\right) = \exp\left\{\lfloor \mu N_k^{(1)} \rfloor \log\left[P_k^{(3,2)} \exp\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}}\right) + P_k^{(2,3)}\right]\right.$$
$$\left. + \log\left[P_k^{(3,2)} \exp\left(\frac{\lambda\{\mu N_k^{(1)}\}}{\sqrt{\sigma_{k+1}}}\right) + P_k^{(2,3)}\right] - \frac{\lambda}{\sqrt{\sigma_{k+1}}}\mu N_k^{(1)} P_k^{(3,2)}\right\}.$$

Expanding in Taylor series to the second order at the origin, the exponential firstly and the logarithm secondly, yields

$$\log\left(P_k^{(3,2)} \exp\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}}\right) + P_k^{(2,3)}\right) = \log\left[P_k^{(3,2)} + P_k^{(3,2)}\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} + \frac{\lambda^2}{2\sigma_{k+1}}\right.\right.$$
$$\left.\left. + \mathcal{O}\left(\frac{\lambda^3}{\sigma_{k+1}^{\frac{3}{2}}}\right)\right) + P_k^{(2,3)}\right] = \log\left[1 + P_k^{(3,2)}\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} + \frac{\lambda^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3}{\sigma_{k+1}^{\frac{3}{2}}}\right)\right)\right]$$
$$= P_k^{(3,2)}\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} + \frac{\lambda^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3}{\sigma_{k+1}^{\frac{3}{2}}}\right)\right) - \frac{1}{2}\left(P_k^{(3,2)}\right)^2\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} + \frac{\lambda^2}{2\sigma_{k+1}}\right.$$
$$\left. + \mathcal{O}\left(\frac{\lambda^3}{\sigma_{k+1}^{\frac{3}{2}}}\right)\right)^2 + \mathcal{O}\left(\left(P_k^{(3,2)}\right)^3\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} + \frac{\lambda^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3}{\sigma_{k+1}^{\frac{3}{2}}}\right)\right)^3\right)$$
$$= P_k^{(3,2)}\frac{\lambda}{\sqrt{\sigma_{k+1}}} + P_k^{(3,2)}\frac{\lambda^2}{2\sigma_{k+1}} - \left(P_k^{(3,2)}\right)^2\frac{\lambda^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3}{\sigma_{k+1}^{\frac{3}{2}}}\right)$$
$$= P_k^{(3,2)}\frac{\lambda}{\sqrt{\sigma_{k+1}}} + \left[P_k^{(3,2)} - \left(P_k^{(3,2)}\right)^2\right]\frac{\lambda^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3}{\sigma_{k+1}^{\frac{3}{2}}}\right)$$
$$= P_k^{(3,2)}\frac{\lambda}{\sqrt{\sigma_{k+1}}} + P_k^{(3,2)}P_k^{(2,3)}\frac{\lambda^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3}{\sigma_{k+1}^{\frac{3}{2}}}\right),$$

and similarly, since $\{\mu N_k^{(1)}\} \leq 1$,

$$\log \left[P_k^{(3,2)} \exp\left(\frac{\lambda \{\mu N_k^{(1)}\}}{\sqrt{\sigma_{k+1}}} \right) + P_k^{(2,3)} \right] =$$

$$P_k^{(3,2)} \frac{\lambda \{\mu N_k^{(1)}\}}{\sqrt{\sigma_{k+1}}} + P_k^{(3,2)} P_k^{(2,3)} \frac{\lambda^2 \{\mu N_k^{(1)}\}^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3}{\sigma_{k+1}^{\frac{3}{2}}} \right).$$

Plugging this expansions in yields, uniformly on the probability space, that

$$\Phi_{\mathscr{F}_k}\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} \right) = \exp\left[\mu N_k^{(1)} P_k^{(3,2)} \frac{\lambda}{\sqrt{\sigma_{k+1}}} + \mu N_k^{(1)} P_k^{(3,2)} P_k^{(2,3)} \frac{\lambda^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3 \mu N_k^{(1)}}{\sigma_{k+1}^{\frac{3}{2}}} \right) \right.$$

$$\left. - \frac{\lambda}{\sqrt{\sigma_{k+1}}} \mu N_k^{(1)} P_k^{(3,2)} \right] = \exp\left[\mu N_k^{(1)} P_k^{(3,2)} P_k^{(2,3)} \frac{\lambda^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3 \mu N_k^{(1)}}{\sigma_{k+1}^{\frac{3}{2}}} \right) \right]$$

$$= \exp\left[\pi_k^{(1)} P_k^{(3,2)} P_k^{(2,3)} \frac{\lambda^2}{2} + \mathcal{O}\left(\frac{\lambda^3 \pi_k^{(1)}}{\sqrt{\sigma_{k+1}}} \right) \right] \leq \exp\left[\frac{\lambda^2}{8} + \mathcal{O}\left(\frac{\lambda^3}{\sqrt{\sigma_{k+1}}} \right) \right] = e + o(1)$$

and

$$\tilde{\Phi}_{\mathscr{F}_k}\left(\frac{\lambda}{\sqrt{\sigma_{k+1}}} \right) = \exp\left[\lfloor \mu N_k^{(1)} \rfloor P_k^{(3,2)} \frac{\lambda}{\sqrt{\sigma_{k+1}}} + \lfloor \mu N_k^{(1)} \rfloor P_k^{(3,2)} P_k^{(2,3)} \frac{\lambda^2}{2\sigma_{k+1}} + P_k^{(3,2)} \frac{\lambda \{\mu N_k^{(1)}\}}{\sqrt{\sigma_{k+1}}} \right.$$

$$\left. + P_k^{(3,2)} P_k^{(2,3)} \frac{\lambda^2 \{\mu N_k^{(1)}\}^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3 \mu N_k^{(1)}}{\sigma_{k+1}^{\frac{3}{2}}} \right) - \frac{\lambda}{\sqrt{\sigma_{k+1}}} \mu N_k^{(1)} P_k^{(3,2)} \right] =$$

$$\exp\left[(\lfloor \mu N_k^{(1)} \rfloor + \{\mu N_k^{(1)}\}^2) P_k^{(3,2)} P_k^{(2,3)} \frac{\lambda^2}{2\sigma_{k+1}} + \mathcal{O}\left(\frac{\lambda^3 \mu N_k^{(1)}}{\sigma_{k+1}^{\frac{3}{2}}} \right) \right]$$

$$\leq \exp\left[\pi_k^{(1)} P_k^{(3,2)} P_k^{(2,3)} \frac{\lambda^2}{2} + \mathcal{O}\left(\frac{\lambda^3 \pi_k^{(1)}}{\sqrt{\sigma_{k+1}}} \right) \right] \leq \exp\left[\frac{\lambda^2}{8} + \mathcal{O}\left(\frac{\lambda^3}{\sqrt{\sigma_{k+1}}} \right) \right] = e + o(1),$$

having chosen $\lambda = 2\sqrt{2}$. Hence for both integral and nonintegral μ (we adopt as main notation that of the integral case, as usual)

$$\mathbb{P}_{\mathscr{F}_k}\left(\frac{B_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(3,2)}}{\sigma_{k+1}} \geq \frac{1}{2\nu^{k+1}} \right) = \mathcal{O}\left(\exp\left(-\sqrt{2}\eta^{k+1} \right) \right).$$

In these estimates only the following fact have been used: $0 \leq \pi_k^{(i)} \leq 1$, $0 \leq P_k^{(i,j)} \leq 1$, $0 \leq P_k^{(i,j)} P_k^{(j,i)} \leq 1/4$ and all asymptotic behaviours are deterministically driven by $\sigma_{k+1} = \mu^{k+1}$. Thus the same estimates can be performed on the complementary binomial:

$$\mathbb{P}_{\mathscr{F}_k}\left(\frac{\overline{B}_{k+1}^{(1)} - \mu N_k^{(1)} P_k^{(2,3)}}{\sigma_{k+1}} \geq \frac{1}{2\nu^{k+1}} \right) = \mathcal{O}\left(\exp\left(-\sqrt{2}\eta^{k+1} \right) \right).$$

Therefore

$$\mathbb{P}_{\mathscr{F}_k}\left(A_k^{(1)c} \right) = 2\mathcal{O}\left(\exp\left(-\sqrt{2}\eta^{k+1} \right) \right) = \mathcal{O}\left(\exp\left(-\sqrt{2}\eta^{k+1} \right) \right).$$

The same argument, extending similarly to all other binomials, yields that for all $i \in \{1, 2, 3\}$,

$$\mathbb{P}_{\mathscr{F}_k}\left(A_k^{(i)c} \right) = \mathcal{O}\left(\exp\left(-\sqrt{2}\eta^{k+1} \right) \right), \ \mathbb{P}_{\mathscr{F}_k}\left(\overline{A}_k^{(i)c} \right) = \mathcal{O}\left(\exp\left(-\sqrt{2}\eta^{k+1} \right) \right).$$

Step 4. Plugging the geometric estimates obtained in *Step 3* in (4.76) yields

$$\begin{aligned}\mathbb{P}\left(A_k^c \middle| A_{k-1}^N\right) &\leq \sum_{i=1}^{3} \mathbb{E}\left[\mathbb{P}_{\mathscr{F}_k}\left(A_k^{(i)c}\right) \middle| A_{k-1}^N\right] + \mathbb{E}\left[\mathbb{P}_{\mathscr{F}_k}\left(\overline{A}_k^{(i)c}\right) \middle| A_{k-1}^N\right] \\ &= 6\mathbb{E}\left[\mathcal{O}\left(e^{-\sqrt{2}\eta^{k+1}}\right) \middle| A_{k-1}^N\right] = \mathcal{O}\left(e^{-\sqrt{2}\eta^{k+1}}\right),\end{aligned}$$

due to the uniformity of the $\mathcal{O}$-constant on the probability space. As $\eta > 1$, we obtain (4.75) and thus (4.74). Hence using (4.73) we have obtained $\mathbb{P}(\mathscr{A}_N | \mathscr{A}_\varepsilon^N) > 0$.

Step 5. In *Step 1* we argued that this implies that $\mathbb{P}(\mathscr{A}_\varepsilon^N \cap \mathscr{A}_N) > 0$, so if we now note that

$$\{|R_{n+1}^{(i)}| < \nu^{-n-1},\ \forall i \in \{1,2,3\}\} \supseteq A_n,$$

it follows that

$$\bigcap_{n \geq N} \{|R_{n+1}^{(i)}| < \nu^{-n-1},\ \forall i \in \{1,2,3\}\} \supseteq \bigcap_{n \geq N} A_n = \mathscr{A}_N,$$

and therefore

$$\mathbb{P}\left(|R_{n+1}^{(i)}| < \nu^{-n-1},\ \forall i \in \{1,2,3\},\ \forall n \geq N | \mathscr{A}_\varepsilon^N\right) \geq \mathbb{P}(\mathscr{A}_N | \mathscr{A}_\varepsilon^N) > 0.$$

In conclusion

$$\mathbb{P}\left(\{|R_{n+1}^{(i)}| < \nu^{-n-1},\ \forall i \in \{1,2,3\},\ \forall n \geq N\} \cap \mathscr{A}_\varepsilon^N\right) > 0,$$

so there is a nonnegligible event on which we can repeat:

i) The argument of Proposition 4.8, when we take $\mathscr{A}_\varepsilon^N = \mathscr{A}_\varepsilon^N(C_0)$, yielding nonnegligible convergence to an internal equilibrium point.

ii) The argument of Theorem 4.42, when we take $\mathscr{A}_\varepsilon^N = \mathscr{A}_\varepsilon^N(C_>)$, yielding nonnegligible convergence to the boundary for $\{\Theta_n\}$, while it stays bounded away from the vertices.

□

By Theorem 1.3, $\{\Theta_n\}$ converges with positive probability both in $\mathring{\Sigma}$ and $\partial\Sigma$. Equivalently, both of the following events have positive probability: all three edges of the triangle are asymptotically nonnegligibly crossed by the particles and exactly one edge is asymptotically negligibly crossed.

We now prove Corollary 1.4, in which the following is claimed:

i) $\mathbb{P}(\pi_n \text{ converges in } \Sigma^*) > 0$.

ii) $\mathbb{P}(\pi_n \text{ diverges in } \Sigma) > 0$.

iii) $\mathbb{P}(\{\pi_n \text{ converges in } \Sigma^*\} \setminus \mathscr{B}) = \mathbb{P}(\mathscr{B} \setminus \{\pi_n \text{ converges in } \Sigma^*\}) = 0$.

iv) $\mathbb{P}(\{\pi_n \text{ diverges in } \Sigma\} \cap \mathscr{E}_>) > 0$ and $\mathbb{P}(\{\pi_n \text{ diverges in } \Sigma\} \setminus (\mathscr{E}_> \cup \mathcal{D}_>)) = \mathbb{P}((\mathscr{E}_> \cup \mathcal{D}_>) \setminus \{\pi_n \text{ diverges in } \Sigma\}) = 0$.

Proof of Corollary 1.4.

i) By the conclusion of the argument in Theorem 1.3 (i), when using Proposition 4.8 at C_0, we can set up the constants ε and δ such that $\{\pi_n(\omega)\}$ converges within a neighbourhood of $\mathbf{\frac{1}{3}}$ small enough to be contained in Σ^*, for almost every ω in the nonnegligible event

$$\{|R^{(i)}_{n+1}| < \nu^{-n-1},\ \forall i \in \{1,2,3\},\ \forall n \geq N\} \cap \mathscr{A}^N_\varepsilon(C_0).$$

ii) By the conclusion of the argument in Theorem 1.3 (ii), when using Theorem 4.42 at $C_>$, even though we do not know the value of ℓ, similarly to how argued in Remark 3.50, we can set up the constants ε and δ such that $\{\pi_n(\omega)\}$ diverges, due to the almost sure oscillations of $\pi^{(1)}_n(\omega)$ following from Lemma 4.38 (which ensures that almost surely $\pi^{(1)}_n$ alternates between values less than $1/2$, bounded away from 0 and $1/2$, and values larger than $1/2$, bounded away from $1/2$ and 1), for almost every ω in the nonnegligible event

$$\{|R^{(i)}_{n+1}| < \nu^{-n-1},\ \forall i \in \{1,2,3\},\ \forall n \geq N\} \cap \mathscr{A}^N_\varepsilon(C_>).$$

iii) Assume that $\mathbb{P}(N) > 0$, where $N \coloneqq \{\pi_n \text{ converges in } \Sigma^*\} \setminus \mathscr{B}$, by contradiction. By Remark 4.6, $N \subseteq \mathscr{E} \cup \mathcal{D}$. Thus every $\omega \in N$ is such that either $\omega \in \mathscr{E}$ or $\omega \in \mathcal{D}$. Then we have the following cases.

- $\omega \in \mathscr{E}_0$, which leads to an almost sure contradiction, since by Theorem 4.18 and Remark 4.19, almost every such ω is such that $(\Theta_n(\omega), \pi_n(\omega)) \longrightarrow (\Theta_*(\omega), \pi_{\Theta_*(\omega)})$, with $\Theta_*(\omega) \in \partial\Sigma$, implying that $\pi_{\Theta_*(\omega)} \in \partial\Sigma^*$, against the hypothesis that $\{\pi_n(\omega)\}$ converges in Σ^* (recall that this set is defined with its boundary excluded).
- $\omega \in \mathscr{E}_>$, which leads to an almost sure contradiction, since by Corollary 4.43 and Remark 4.44, for almost every such ω, we have asymptotic 2-periodicity of $\{\pi_n(\omega)\}$, that is it has two distinct limit points, and therefore it diverges.
- $\omega \in \mathcal{D}$, which leads to an almost sure contradiction, since by Lemma 4.21, for almost every ω, $\{\pi_n(\omega)\}$ only admits limit points in $\partial\Sigma$.

It follows that N is a negligible event. Moreover, $\mathbb{P}(\mathscr{B} \setminus \{\pi_n \text{ converges in } \Sigma^*\}) = 0$ directly by Proposition 4.7, where, among other things, we showed that for almost every $\omega \in \mathscr{B}$, $\{\pi_n(\omega)\}$ converges in Σ^*.

iv) $\mathbb{P}(\{\pi_n \text{ diverges in } \Sigma\} \cap \mathscr{E}_>) > 0$ trivially follows from Theorem 1.3 (ii), Corollary 1.4 (ii) and Corollary 4.43 and Remark 4.44 (the last two ensuring that for almost every $\omega \in \mathscr{E}_>$, we have asymptotic 2-periodicity of $\{\pi_n(\omega)\}$). In order to show that $\mathbb{P}(\{\pi_n \text{ diverges in } \Sigma\} \setminus (\mathscr{E}_> \cup \mathcal{D}_>)) = 0$, assume by contradiction that $\mathbb{P}(N) > 0$, where $N \coloneqq \{\pi_n \text{ diverges in } \Sigma\} \setminus (\mathscr{E}_> \cup \mathcal{D}_>)$. By Remark 4.6, $N \subseteq \mathscr{E}_0 \cup \mathcal{D}_0 \cup \mathscr{B}$. Thus every $\omega \in N$ is such that either $\omega \in \mathscr{E}_0 \cup \mathcal{D}_0$ or $\omega \in \mathscr{B}$. Then we have the following cases.

- $\omega \in \mathscr{E}_0 \cup \mathcal{D}_0$, which leads to an almost sure contradiction since by Theorem 4.18 and Remark 4.19 and definition of $\mathcal{D}_0$, for almost every such ω, $(\Theta_n(\omega), \pi_n(\omega)) \longrightarrow (\Theta_*(\omega), \pi_{\Theta_*(\omega)})$, with $\Theta_*(\omega) \in \partial\Sigma$. Thus $\{\pi_n(\omega)\}$ converges.
- $\omega \in \mathscr{B}$, which leads to an almost sure contradiction, since by Proposition 4.7, for almost every such ω, $(\Theta_n(\omega), \pi_n(\omega)) \longrightarrow (\Theta_*(\omega), \pi_{\Theta_*(\omega)})$, with $\Theta_*(\omega) \in \mathring{\Sigma}$. Thus $\{\pi_n(\omega)\}$ converges.

Moreover, $\mathbb{P}((\mathscr{E}_> \cup \mathcal{D}_>) \setminus \{\pi_n \text{ diverges in } \Sigma\}) = 0$, since:

- by Corollary 4.43 and Remark 4.44, for almost every $\omega \in \mathscr{E}_>$, we have asymptotic 2-periodicity of $\{\pi_n(\omega)\}$, and therefore divergence in Σ;
- by Lemma 4.21, for almost every $\omega \in \mathcal{D}_>$, $\{\pi_n(\omega)\}$ diverges in Σ.

□

4.8 No monopoly

In this section we show the negligibility of monopoly, which is, as per Definition 1.5, the event $\mathcal{M}$ on which all but finitely many crossings happen along exactly one edge.

Proof of Theorem 1.6. If we can show the stronger claim, that all three edges are crossed infinitely many times, then monopoly does not occur. Via martingale arguments we focus on showing that for all $i \in \{1,2,3\}$, $T_n^{(i)} \longrightarrow \infty$ as $n \longrightarrow \infty$. By symmetry, without loss of generality, assume the index considered to be $i = 1$.

Denote $\overline{B}_{n+1}^{(3)} = \mu N_n^{(3)} - B_{n+1}^{(3)} \sim \operatorname{Bin}(\mu N_n^{(3)}, P_n^{(1,2)})$ and recall $B_{n+1}^{(2)} \sim \operatorname{Bin}(\mu N_n^{(2)}, P_n^{(1,3)})$ conditionally on $\mathscr{F}_n$, where $P_n^{(1,2)} := \Theta_n^{(1)}/(\Theta_n^{(1)} + \Theta_n^{(2)})$ and $P_n^{(1,3)} := \Theta_n^{(1)}/(\Theta_n^{(1)} + \Theta_n^{(3)})$. Then

$$
\begin{aligned}
T_{n+1}^{(1)} &= T_n^{(1)} + B_{n+1}^{(2)} - \mu N_n^{(2)} P_n^{(1,3)} + \overline{B}_{n+1}^{(3)} - \mu N_n^{(3)} P_n^{(1,2)} + \mu N_n^{(2)} P_n^{(1,3)} + \mu N_n^{(3)} P_n^{(1,2)} \\
&= T_0^{(1)} + \sum_{k=1}^{n+1} B_k^{(2)} - \mu N_{k-1}^{(2)} P_{k-1}^{(1,3)} + \sum_{k=1}^{n+1} \overline{B}_k^{(3)} - \mu N_{k-1}^{(3)} P_{k-1}^{(1,2)} + \sum_{k=0}^{n} \mu N_k^{(2)} P_k^{(1,3)} \\
&\quad + \sum_{k=0}^{n} \mu N_k^{(3)} P_k^{(1,2)} =: T_0^{(1)} + M_{n+1}^{(2)} + \overline{M}_{n+1}^{(3)} + Y_{n+1}^{(2)} + \overline{Y}_{n+1}^{(3)}.
\end{aligned}
$$

For nonintegral μ we will have $\overline{\tilde{B}}_{n+1}^{(3)} = \mu N_n^{(3)} - \tilde{B}_{n+1}^{(3)} \sim \operatorname{Bin}(\lfloor \mu N_n^{(3)} \rfloor, P_n^{(1,2)}) + \{\mu N_n^{(3)}\} \operatorname{Ber}(P_n^{(1,2)})$ and $\tilde{B}_{n+1}^{(2)} \sim \operatorname{Bin}(\lfloor \mu N_n^{(2)} \rfloor, P_n^{(1,3)}) + \{\mu N_n^{(2)}\} \operatorname{Ber}(P_n^{(1,3)})$ conditionally on $\mathscr{F}_n$, thus getting the corresponding decomposition

$$
T_{n+1}^{(1)} = T_0^{(1)} + \tilde{M}_{n+1}^{(2)} + \overline{\tilde{M}}_{n+1}^{(3)} + Y_{n+1}^{(2)} + \overline{Y}_{n+1}^{(3)},
$$

where $\tilde{M}_{n+1}^{(2)} := \sum_{k=1}^{n+1} \tilde{B}_k^{(2)} - \mu N_{k-1}^{(2)} P_{k-1}^{(1,3)}$ and $\overline{\tilde{M}}_{n+1}^{(3)} := \sum_{k=1}^{n+1} \overline{\tilde{B}}_k^{(3)} - \mu N_{k-1}^{(3)} P_{k-1}^{(1,2)}$.

The first step to prove that $T_n^{(1)} \longrightarrow \infty$ almost surely is to show that $Y_n^{(2)} + \overline{Y}_n^{(3)} \longrightarrow \infty$ almost surely. Note that both $P_k^{(1,2)}, P_k^{(1,3)} \geq \Theta_k^{(1)}$ and $T_k^{(1)} \geq 1$ for all $k \geq 0$, so

$$
Y_n^{(2)} + \overline{Y}_n^{(3)} \geq \sum_{k=0}^{n} \mu (N_k^{(2)} + N_k^{(3)}) \Theta_k^{(1)} = \sum_{k=0}^{n} \frac{\sigma_{k+1}}{\tau_k} \left(\frac{N_k^{(2)}}{\sigma_k} + \frac{N_k^{(3)}}{\sigma_k} \right) T_k^{(1)} \geq \sum_{k=0}^{n} \frac{\sigma_{k+1}}{\tau_k} (\pi_k^{(2)} + \pi_k^{(3)}).
$$

Since $\mu > 1$,

$$
\frac{\sigma_{k+1}}{\tau_k} = \frac{\mu^{k+1}}{\tau_0 + \mu \frac{\mu^k - 1}{\mu - 1}} \longrightarrow \mu - 1 > 0,
$$

hence by limit comparison,

$$
\sum_{k=0}^{n} \frac{\sigma_{k+1}}{\tau_k} (\pi_k^{(2)} + \pi_k^{(3)}) = \infty
$$

if

$$
\sum_{k=0}^{n} (1 - \pi_k^{(1)}) = \infty.
$$

Therefore, to show that almost surely $Y_n^{(2)} + \overline{Y}_n^{(3)} \longrightarrow \infty$, it is enough to ensure that almost surely $\pi_k^{(1)} \not\longrightarrow 1$, meaning that almost surely there is a subsequence $\pi_{k_j}^{(1)}$ bounded away from 1: this follows by Corollary 4.45. Finally, by showing that

$$\{Y_n^{(2)} + \overline{Y}_n^{(3)} \longrightarrow \infty\} \subset \{T_n^{(1)} \longrightarrow \infty\}, \tag{4.78}$$

it will follow that almost surely $T_n^{(1)} \longrightarrow \infty$. By repeating a similar argument for the other two components, it will follow that $\mathbb{P}(\mathcal{M}) = 0$. To show the inclusion, assume $Y_n^{(2)} + \overline{Y}_n^{(3)} \longrightarrow \infty$ and consider that

$$\langle M^{(2)} \rangle_n = \sum_{k=0}^{n} \mu N_k^{(2)} P_k^{(1,3)}(1 - P_k^{(1,3)}) \leq Y_n^{(2)}, \quad \left\langle \overline{M}^{(3)} \right\rangle_n = \sum_{k=0}^{n} \mu N_k^{(3)} P_k^{(1,2)}(1 - P_k^{(1,2)}) \leq \overline{Y}_n^{(3)}$$

with all the possible cases being the following: $\langle M^{(2)} \rangle_n$ and $\left\langle \overline{M}^{(3)} \right\rangle_n$ both converge, one converges and the other diverges, or both diverge. Before proceeding, note that

$$\begin{aligned}
\left\langle \tilde{M}^{(2)} \right\rangle_n &= \sum_{k=0}^{n} (\lfloor \mu N_k^{(2)} \rfloor + \{\mu N_k^{(2)}\}^2) P_k^{(1,3)}(1 - P_k^{(1,3)}) \leq \langle M^{(2)} \rangle_n \\
\left\langle \overline{\tilde{M}}^{(3)} \right\rangle_n &= \sum_{k=0}^{n} (\lfloor \mu N_k^{(3)} \rfloor + \{\mu N_k^{(3)}\}^2) P_k^{(1,2)}(1 - P_k^{(1,2)}) \leq \left\langle \overline{M}^{(3)} \right\rangle_n,
\end{aligned}$$

and therefore, the same argument we are about to show holds for nonintegral μ. Thus we explicitly treat only the integral case, from now on.

The standard theory of the angle bracket process applies. Recall that:

- if $\langle M^{(i)} \rangle_n$ converges, then $M_n^{(i)}$ converges almost surely by [50, §12.13];
- if $\langle M^{(i)} \rangle_n$ diverges, then $M_n^{(i)} / \langle M^{(i)} \rangle_n$ vanishes by [50, §12.14].

Therefore we have the following cases.

Case 1. If both $\langle M^{(2)} \rangle_n$ and $\left\langle \overline{M}^{(3)} \right\rangle_n$ converge, then both martingales $M_n^{(2)}$ and $\overline{M}_n^{(3)}$ converges, yielding $T_n^{(1)} = T_0^{(1)} + M_n^{(2)} + \overline{M}_n^{(3)} + Y_n^{(2)} + \overline{Y}_n^{(3)} \longrightarrow \infty$.

Case 2. If only one of them converges, assume, by symmetry, without loss of generality, that $\left\langle \overline{M}^{(3)} \right\rangle_n$ is the convergent one. Then $\overline{M}_n^{(3)}$ converge. On the other hand $\langle M^{(2)} \rangle_n \longrightarrow \infty$, which implies $M_n^{(2)} / \langle M^{(2)} \rangle_n \longrightarrow 0$, hence $M_n^{(2)} / Y_n^{(2)} \longrightarrow 0$ by the inequality aforementioned, and also $M_n^{(2)} / (Y_n^{(2)} + \overline{Y}_n^{(3)}) \longrightarrow 0$. In conclusion, on this event

$$\begin{aligned}
T_n^{(1)} &= \left(\frac{T_0^{(1)}}{Y_n^{(2)} + \overline{Y}_n^{(3)}} + \frac{M_n^{(2)}}{Y_n^{(2)} + \overline{Y}_n^{(3)}} + \frac{\overline{M}_n^{(3)}}{Y_n^{(2)} + \overline{Y}_n^{(3)}} + 1 \right) (Y_n^{(2)} + \overline{Y}_n^{(3)}) \\
&= (1 + o_\omega(1)) \left(Y_n^{(2)} + \overline{Y}_n^{(3)} \right) \longrightarrow \infty,
\end{aligned}$$

because trivially, since $\overline{M}_n^{(3)}$ converges, $\overline{M}_n^{(3)} / (Y_n^{(2)} + \overline{Y}_n^{(3)}) \longrightarrow 0$.

Case 3. If both $\langle M^{(2)} \rangle_n$ and $\left\langle \overline{M}^{(3)} \right\rangle_n$ diverge, by a similar factorisation as in the previous case, and exploiting the same argument just used for $M_n^{(2)}$, also for $\overline{M}_n^{(3)}$, one reaches the same conclusion.

Hence (4.78) has been proved and the claim follows. The argument extends by symmetry to all components of T_n, by noting that

$$\langle M^{(i)} \rangle_n = \left\langle \overline{M}^{(i)} \right\rangle_n,$$

and that for each component, the cases are all the same, with the predictable parts diverging almost surely, as π_n never converges to any of the vertices by Corollary 4.45. □

4.9 No dominance conjecture

In this section we discuss Conjecture 1.7, asserting that $\mathbb{P}(\mathcal{D}) = 0$, where we defined the event of dominance $\mathcal{D}$ as the event, on which the edge crossings along two of the edges become eventually negligible (see Definition 1.2). The conjecture is based on three grounds:

- It holds for the generalised balls and bins model (Theorem 1.10), which has a special connection with this model.
- It is numerically supported by simulations of (2.31) and (2.32). In particular, simulating the dynamical systems (2.33) and (2.34) and (2.35) and (2.36) results in only two asymptotic behaviours, the typical ones corresponding to Section 4.7, exemplified by Figures 2.2 and 2.3 (hence we conjecture also that $\{p_n\}$ in (2.33) and (2.34) and (2.35) and (2.36) does not converge to the vertices).
- The martingale increments $\{R_n^{(i)}\}_i$ not only satisfy Lemma 4.1, but if the system approaches the corner, they can be shown to become negligible, compared to the terms in what we could call the *deterministic* part of (2.31) and (2.32). This suggests that it should be possible, for almost all $\omega \in \mathcal{D}$, to show a contradiction by exploiting the asymptotic behaviour of the dynamical system, which is conjectured to not converge to the vertices.

In the following proposition we make some progress, by proving the negligibility of the martingale increments near the vertices (for simplicity, we will only treat the case of integral μ, but the result follows also for nonintegral μ). Partition the event of dominance as $\mathcal{D} = \bigcup_{i=1}^3 \mathcal{D}^{(i)} = \bigcup_{i=1}^3 \mathcal{D}_0^{(i)} \cup \mathcal{D}_>^{(i)}$, where $\mathcal{D}^{(i)} := \{\Theta_n \longrightarrow v_i\}$, and the indices 0 and $>$ denote the usual intersection with the events, on which the values of ℓ are, respectively, 0 and greater than 0.

Proposition 4.48. *For every $i, j, k \in \{1, 2, 3\}$, $i \neq j \neq k$, for almost every $\omega \in \mathcal{D}_0^{(i)}$, $R_{n+1}^{(i)} = o_\omega(\Theta_n^{(j)})$ and $R_{n+1}^{(i)} = o_\omega(\Theta_n^{(k)})$.*

Proof. By symmetry, without loss of generality, it will suffice to show the argument for $i = 2$. In Theorem 1.6 it has been shown how, as $n \longrightarrow \infty$, $T_n^{(1)} = (1 + o_\omega(1))(Y_n^{(2)} + \overline{Y}_n^{(3)})$, where $Y_n^{(2)} := \sum_{k=0}^n \mu N_k^{(2)} P_k^{(1,3)}$, $\overline{Y}_n^{(3)} = \sum_{k=0}^n \mu N_k^{(3)} P_k^{(1,2)}$. There the initial index m was zero, but it can be taken to be arbitrary in general, and the $(1 + o_\omega(1))$ factor holds as $n \longrightarrow \infty$, for m fixed. Similarly $T_n^{(3)} = (1 + o_\omega(1))(\overline{Y}_n^{(2)} + Y_n^{(1)})$, where $\overline{Y}_n^{(2)} := \sum_{k=0}^n \mu N_k^{(2)} P_k^{(3,1)}$, $Y_n^{(1)} = \sum_{k=0}^n \mu N_k^{(1)} P_k^{(3,2)}$. Theorem 1.6 shows that $T_n^{(i)} \longrightarrow \infty$ for all $i \in \{1, 2, 3\}$, but more can be said, given this estimate, and the knowledge of the almost sure limit points of the stochastic process. Due to the asymptotic nature of the argument, it will be crucial to observe that the deterministic coefficients that will appear in the summations

$$\frac{\sigma_{k+1}}{\tau_k} = \frac{\mu^{k+1}}{\tau_0 + \frac{\mu^{K+1}}{\mu-1} - \frac{\mu}{\mu-1}} = \frac{\mu-1}{1 + \frac{\tau_0(\mu-1)}{\mu^{k+1}} - \frac{1}{\mu^k}} = (\mu-1)\left(1 + \mathcal{O}\left(\frac{1}{\mu^k}\right)\right) \tag{4.79}$$

and

$$\frac{\sigma_{k+1}}{\tau_{k-1}} = \frac{\mu^{k+1}}{\tau_0 + \frac{\mu^{k}}{\mu-1} - \frac{\mu}{\mu-1}} = \frac{\mu(\mu-1)}{1 + \frac{\tau_0(\mu-1)}{\mu^{k}} - \frac{1}{\mu^{k-1}}} = \mu(\mu-1)\left(1 + \mathcal{O}\left(\frac{1}{\mu^k}\right)\right). \tag{4.80}$$

We will rely on *Perron-Frobenius Theorems*, with the argument shown only for the first component, since for the third it is analogous. Start with

$$\begin{aligned} T^{(1)}_{n+1} &= (1 + o_\omega(1))\left(\sum_{k=0}^{n} \mu N^{(2)}_k P^{(1,3)}_k + \mu N^{(3)}_k P^{(1,2)}_k\right) \\ &= (1 + o_\omega(1))\left(\sum_{k=m}^{n} \mu N^{(2)}_k P^{(1,3)}_k + \mu N^{(3)}_k P^{(1,2)}_k\right) \\ &= (1 + o_\omega(1))\sum_{k=m}^{n} \frac{\sigma_{k+1}}{\tau_k}\left(\frac{\pi^{(2)}_k}{\Theta^{(1)}_k + \Theta^{(3)}_k} + \frac{\pi^{(3)}_k}{\Theta^{(2)}_k + \Theta^{(1)}_k}\right) T^{(1)}_k \geq \sum_{k=m}^{n} \zeta_k T^{(1)}_k, \end{aligned}$$

where

$$\frac{\sigma_{k+1}}{\tau_k}(1 + o_\omega(1))\left(\frac{\pi^{(2)}_k}{\Theta^{(1)}_k + \Theta^{(3)}_k} + \frac{\pi^{(3)}_k}{\Theta^{(2)}_k + \Theta^{(1)}_k}\right) > \frac{\mu-1}{2} - \varepsilon_k =: \zeta_k$$

for some $\varepsilon_k(\omega) = o_\omega(1)$ slow enough (compared to the sample path taken for almost every ω on the event, pointwise), since on $\mathcal{D}^{(2)}_0$ the coefficient $\pi^{(3)}_k$ tends to $1/2$ and $\Theta^{(2)}_k + \Theta^{(1)}_k \longrightarrow 1$, but it is not known whether $\pi^{(2)}_k/(\Theta^{(1)}_k + \Theta^{(3)}_k)$ is even bounded, hence the inequality sign. The term, which $T^{(1)}_{n+1}$ majorises, $\sum_{k=m}^{n} \zeta_k T^{(1)}_k$, is generated by the following dynamical system. Take $T_m = T^{(1)}_m > 1$ as initial value, and set $S_m = 0$, then the summation is the first component of the orbit of the 2-dimensional dynamical system

$$\begin{aligned} T_{n+1} &= S_n + \zeta_n T_n \\ S_{n+1} &= S_n + \zeta_n T_n \end{aligned}$$

for all $n \geq m$. In matrix form:

$$\begin{pmatrix} T_{n+1} \\ S_{n+1} \end{pmatrix} = \begin{pmatrix} \zeta_n & 1 \\ \zeta_n & 1 \end{pmatrix}\begin{pmatrix} T_n \\ S_n \end{pmatrix}.$$

Denote the iteration matrix A_n and let $X_n = (T_n, S_n)$. What is needed, is the asymptotics of $X_{n+1} = A_n X_n$, specifically of its first component, keeping in mind that in our case, X_n is a positive vector (and each of its components are unbounded). It will now be shown that this system the hypotheses of the various strengthened forms of the *weak Perron-Frobenius Theorem* (see [40] for some background, and in particular Theorem 1.3, which will be used). We will derive two mutually exclusive asymptotics for it, and then apply the information about the specific orbit we have on $\mathcal{D}^{(2)}_0$, so as to narrow down the slowest possible asymptotic regime of growth for T_{n+1} (since $T^{(1)}_{n+1}$ is lower bounded by the dynamics of T_{n+1}, the asymptotics for the latter will yield a lower bound asymptotics for the former). First of all, note that the matrix $A_n \longrightarrow A$, where, denoting $\zeta := \mu - 1$,

$$A = \begin{pmatrix} \zeta & 1 \\ \zeta & 1 \end{pmatrix}.$$

The characteristic polynomial of A_n and A are $p_n(\lambda) = \lambda[\lambda - (1 + \zeta_n)]$ and $p(\lambda) = \lambda[\lambda - (1+\zeta)]$, yielding two simple eigenvalues $\lambda_+ = \zeta + 1 = (\mu+1)/2 > 1$ and $\lambda_- = 0$ respectively. Similarly for A_n, and by continuity of simple eigenvalues it is known that $\lambda_n^+ \longrightarrow \lambda^+$, while $\lambda_n^- = 0$ for all n. The corresponding eigenvectors of A can be chosen to be $e_+ = (1, 1)$ (this is the Perron eigenvector) and $e_- = (1, -\zeta)$ (note that coming from the eigenspace equation $y = -\zeta x$, this eigenvector is never nonnegative, being $\zeta = \mu - 1 > 0$). In [40, Theorem 1.3] it is shown that for a nonnegative matrix $A_n \longrightarrow A$ as such (that is with distinct eigenvalues), and a nonnegative orbit X_n generated, like ours, by the iteration scheme, either X_n is eventually zero (which is not our case, being unbounded on the considered event); or the limit (which exist by [40, Theorem 1.2]) of $\sqrt[n]{\|X_n\|_1}$ (any norm in general, but in this case it is simpler to work in 1-norm) is an eigenvalue of A, with a nonnegative eigenvector, which only leaves the option

$$\lim_{n \longrightarrow \infty} \sqrt[n]{\|X_n\|_1} = \frac{\mu+1}{2},$$

due to the zero eigenvalue not having a nonnegative eigenvector. If one notes that $S_n = T_n$, then this says that

$$\frac{\mu+1}{2} = \lim_{n \longrightarrow \infty} \sqrt[n]{2|T_n|} = \lim_{n \longrightarrow \infty} \sqrt[n]{T_n}.$$

This does not allow for a regime as slow as $\sqrt{\mu}^n$, since for all $\mu > 1$, $\mu + 1 > 2\sqrt{\mu}$. Hence for any $\mu > 1$ fixed, there is some positive $p = p(\mu)$ smaller than $1/2$, such that $T_n = \Omega_\omega(\mu^{(\frac{1}{2}+p)n})$. Indeed if for every $0 < p < 1/2$ given, there is a subsequence $\{n_j\}$, such that $T_{n_j} < \mu^{(\frac{1}{2}+p)n_j}$, then choose any

$$\bar{p} < \frac{\log \frac{\mu+1}{2}}{\log \mu} - \frac{1}{2}.$$

This choice is feasible, since the term on the right-hand side is positive for all $\mu > 1$, because this term being positive is equivalent to $\mu + 1 > 2\sqrt{\mu}$, via taking logarithms. It should be noted that $\bar{p}$ has been defined so that $\mu^{\frac{1}{2}+\bar{p}} < (\mu+1)/2$. It follows that

$$\frac{\mu+1}{2} = \lim_{n \longrightarrow \infty} \sqrt[n]{T_n} \le \lim_{j \longrightarrow \infty} \sqrt[n_j]{\mu^{(\frac{1}{2}+\bar{p})n_j}} = \mu^{\frac{1}{2}+\bar{p}} < \frac{\mu+1}{2},$$

which is a contradiction. Thus the claim, that $T_n^{(1)} = \Omega_\omega(\mu^{(\frac{1}{2}+p)n})$ has been proved (since $T_n^{(1)} \ge T_n$).

Before concluding, we would like to stress that this stronger estimate on $\mathcal{D}_0^{(2)}$, that for every $\mu > 1$ there is a $p = p(\mu)$ small enough (choose it at least smaller than $1/2$) such that $\mu^{(\frac{1}{2}+p)n} = o_\omega(T_n^{(1)})$ (and the same of course holding for $T_n^{(3)}$), can be bootstrapped. We will show how the bootstrapping works in the conclusion to this section, after this proof.

The claim that $R_{n+1}^{(2)} = o_\omega(\Theta_n^{(1)})$ and $R_{n+1}^{(2)} = o_\omega(\Theta_n^{(3)})$ now easily follows. This is only shown for the index $i = 1$, as the method is similar for $i = 3$. By Lemma 4.1, $R_{n+1}^{(2)} = \mathcal{O}_\omega(\nu^{-n-1}) = \mathcal{O}_\omega(\nu^{-n})$ for any $1 < \nu < \sqrt{\mu}$, while $\Theta_n^{(1)} = T_n^{(1)}/\tau_n \sim (\mu - 1)T_n^{(1)}/\mu^{n+1}$ and since $\mu^{(1/2+p)n} = o_\omega(T_n^{(1)})$, it follows that

$$\frac{1}{\mu^{\left(\frac{1}{2}-p\right)n}} = \frac{\mu^{\left(\frac{1}{2}+p\right)n}}{\mu^n} = o_\omega\left(\frac{T_n^{(1)}}{\mu^n}\right) = o_\omega\left(\frac{\mu}{\mu-1}\Theta_n^{(1)}\right) = o_\omega\left(\Theta_n^{(1)}\right).$$

Hence by choosing $\nu = \mu^{\frac{1-p}{2}}$, noting that $\nu > \mu^{1/2-p}$, it follows that

$$R_{n+1}^{(2)} = \mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{(1-p)}{2}n}}\right) = o_\omega\left(\frac{1}{\mu^{\left(\frac{1}{2}-p\right)n}}\right) = o_\omega\left(\Theta_n^{(1)}\right). \tag{4.81}$$

It should now be clear the reason for strengthening the initial estimate by this additional power of p: the separation given by the square root regime alone does not allow for room in between the possible regimes upper-bounding $R^{(2)}_{n+1}$ and the ones lower-bounding $\Theta^{(1)}_n$ and $\Theta^{(3)}_n$. $\square$

We conclude with discussing how further bootstrapping could be used to improve the estimates (4.81), and possibly make the negligibility obtained in Proposition 4.48 stronger, if needed, when analysing the stochastic process as a perturbed dynamical system on $\mathcal{D}^{(2)}_0$.Since we showed that almost surely, eventually, the martingale increment perturbations, appearing in the expression for $\pi^{(2)}_n$, become negligible with respect to the main terms involving $\Theta^{(1)}_n$ and $\Theta^{(3)}_n$, it is possible to conclude that on $\mathcal{D}^{(2)}_0$, for every $k \geq m$,

$$\begin{aligned}\pi^{(2)}_k &= \pi^{(1)}_{k-1}\frac{\Theta^{(3)}_{k-1}}{\Theta^{(2)}_{k-1}+\Theta^{(3)}_{k-1}} + \pi^{(3)}_{k-1}\frac{\Theta^{(1)}_{k-1}}{\Theta^{(2)}_{k-1}+\Theta^{(1)}_{k-1}} + R^{(2)}_k \\ &= \left(\frac{\pi^{(1)}_{k-1}}{\Theta^{(2)}_{k-1}+\Theta^{(3)}_{k-1}}\Theta^{(3)}_{k-1} + \frac{\pi^{(3)}_{k-1}}{\Theta^{(2)}_{k-1}+\Theta^{(1)}_{k-1}}\Theta^{(1)}_{k-1}\right)(1+r^{(2)}_k),\end{aligned}$$

where

$$r^{(2)}_k = \frac{R^{(2)}_k}{\frac{\pi^{(1)}_{k-1}}{\Theta^{(2)}_{k-1}+\Theta^{(3)}_{k-1}}\Theta^{(3)}_{k-1} + \frac{\pi^{(3)}_{k-1}}{\Theta^{(2)}_{k-1}+\Theta^{(1)}_{k-1}}\Theta^{(1)}_{k-1}} = \frac{\mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{(1-p)}{2}k}}\right)}{\Omega_\omega\left(\frac{1}{\mu^{\left(\frac{1}{2}-p\right)k}}\right)} = \mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{p}{2}k}}\right),$$

due to the previous quantitative estimates made on $\Theta^{(1)}_n$ and $\Theta^{(3)}_n$, and the fact that $\Theta^{(2)}_n \longrightarrow 1$ and $\pi^{(1)}_n + \pi^{(3)}_n \longrightarrow 1$. Thus

$$\pi^{(2)}_k = \left(1+\mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{p}{2}k}}\right)\right)\left(\frac{\pi^{(1)}_{k-1}}{\Theta^{(2)}_{k-1}+\Theta^{(3)}_{k-1}}\Theta^{(3)}_{k-1} + \frac{\pi^{(3)}_{k-1}}{\Theta^{(2)}_{k-1}+\Theta^{(1)}_{k-1}}\Theta^{(1)}_{k-1}\right). \tag{4.82}$$

Consider $T^{(1,3)}_{n+1}$. We can verify that the same asymptotics hold, through the same martingale theory used for $T^{(1)}_n$ and $T^{(3)}_n$, by simply repeating the steps in the second part of Proposition 4.48, that is

$$T^{(1,3)}_{n+1} = (1+o_\omega(1))(Y^{(2)}_n + \overline{Y}^{(2)}_n + Y^{(1)}_n + \overline{Y}^{(3)}_n).$$

Noting that

$$Y^{(2)}_n + \overline{Y}^{(2)}_n = \sum_{k=0}^{n}\mu N^{(2)}_k = \sum_{k=0}^{n}\sigma_{k+1}\pi^{(2)}_k,$$

we have that

$$T^{(1,3)}_{n+1} = (1+o_\omega(1))\left[\sum_{k=m}^{n}\sigma_{k+1}\pi^{(2)}_k + \sum_{k=m}^{n}\frac{\sigma_{k+1}}{\tau_k}\left(\frac{\pi^{(3)}_k}{\Theta^{(1)}_k+\Theta^{(2)}_k}T^{(1)}_k + \frac{\pi^{(1)}_k}{\Theta^{(3)}_k+\Theta^{(2)}_k}T^{(3)}_k\right)\right].$$

By applying (4.82), the following asymptotics follow:

$$\begin{aligned}
T_{n+1}^{(1,3)} &= (1+o_\omega(1))\Bigg(\sum_{k=m}^{n}\sigma_{k+1}\left(1+\mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{p}{2}k}}\right)\right)\left(\frac{\pi_{k-1}^{(3)}}{\Theta_{k-1}^{(2)}+\Theta_{k-1}^{(1)}}\Theta_{k-1}^{(1)}\right.\\
&\left.+\frac{\pi_{k-1}^{(1)}}{\Theta_{k-1}^{(2)}+\Theta_{k-1}^{(3)}}\Theta_{k-1}^{(3)}\right)+\sum_{k=m}^{n}\sigma_{k+1}\left(\frac{\pi_k^{(3)}}{\Theta_k^{(2)}+\Theta_k^{(1)}}\Theta_k^{(1)}+\frac{\pi_k^{(1)}}{\Theta_k^{(2)}+\Theta_k^{(3)}}\Theta_k^{(3)}\right)\Bigg)\\
&= (1+o_\omega(1))\Bigg(\sum_{k=m}^{n}\frac{\sigma_{k+1}}{\tau_{k-1}}\left(1+\mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{p}{2}k}}\right)\right)\left(\frac{\pi_{k-1}^{(3)}}{\Theta_{k-1}^{(2)}+\Theta_{k-1}^{(1)}}T_{k-1}^{(1)}\right.\\
&\left.+\frac{\pi_{k-1}^{(1)}}{\Theta_{k-1}^{(2)}+\Theta_{k-1}^{(3)}}T_{k-1}^{(3)}\right)+\sum_{k=m}^{n}\frac{\sigma_{k+1}}{\tau_k}\left(\frac{\pi_k^{(3)}}{\Theta_k^{(2)}+\Theta_k^{(1)}}T_k^{(1)}+\frac{\pi_k^{(1)}}{\Theta_k^{(2)}+\Theta_k^{(3)}}T_k^{(3)}\right)\Bigg)\\
&= (1+o_\omega(1))\Bigg(\sum_{k=m+1}^{n}\frac{\sigma_{k+1}}{\tau_{k-1}}\left(1+\mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{p}{2}k}}\right)\right)\left(\frac{\pi_{k-1}^{(3)}}{\Theta_{k-1}^{(2)}+\Theta_{k-1}^{(1)}}T_{k-1}^{(1)}\right.\\
&\left.+\frac{\pi_{k-1}^{(1)}}{\Theta_{k-1}^{(2)}+\Theta_{k-1}^{(3)}}T_{k-1}^{(3)}\right)+\sum_{k=m}^{n}\frac{\sigma_{k+1}}{\tau_k}\left(\frac{\pi_k^{(3)}}{\Theta_k^{(2)}+\Theta_k^{(1)}}T_k^{(1)}+\frac{\pi_k^{(1)}}{\Theta_k^{(2)}+\Theta_k^{(3)}}T_k^{(3)}\right)\Bigg)\\
&= (1+o_\omega(1))\Bigg(\sum_{k=m}^{n-1}\left[\frac{\sigma_{k+2}}{\tau_k}\left(1+\mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{p}{2}k}}\right)\right)+\frac{\sigma_{k+1}}{\tau_k}\right]\left(\frac{\pi_k^{(3)}}{\Theta_k^{(2)}+\Theta_k^{(1)}}T_k^{(1)}\right.\\
&\left.+\frac{\pi_k^{(1)}}{\Theta_k^{(2)}+\Theta_k^{(3)}}T_k^{(3)}\right)+\frac{\sigma_{n+1}}{\tau_n}\left(\frac{\pi_n^{(3)}}{\Theta_n^{(2)}+\Theta_n^{(1)}}T_n^{(1)}+\frac{\pi_n^{(1)}}{\Theta_n^{(2)}+\Theta_n^{(3)}}T_n^{(3)}\right)\Bigg)\\
&\geq \sum_{k=m}^{n-1}\xi_k T_k^{(1,3)}+\zeta_n T_n^{(1,3)}.
\end{aligned}$$

Recalling that by (4.79) and (4.80),

$$\begin{aligned}
\frac{\sigma_{k+2}}{\tau_k}+\frac{\sigma_{k+1}}{\tau_k} &= \mu(\mu-1)\left(1+\mathcal{O}\left(\frac{1}{\mu^k}\right)\right)+(\mu-1)\left(1+\mathcal{O}\left(\frac{1}{\mu^k}\right)\right)\\
&= (\mu^2-1)\left(1+\mathcal{O}\left(\frac{1}{\mu^k}\right)\right),
\end{aligned}$$

we have that the term in the summations can be rewritten as

$$\begin{aligned}
&(1+o_\omega(1))\left[\frac{\sigma_{k+2}}{\tau_k}\left(1+\mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{p}{2}k}}\right)\right)+\frac{\sigma_{k+1}}{\tau_k}\right]\left(\frac{\pi_k^{(3)}}{\Theta_k^{(2)}+\Theta_k^{(1)}}T_k^{(1)}+\frac{\pi_k^{(1)}}{\Theta_k^{(2)}+\Theta_k^{(3)}}T_k^{(3)}\right)\\
&= (1+o_\omega(1))\left[(\mu^2-1)\left(1+\mathcal{O}\left(\frac{1}{\mu^k}\right)\right)+\mathcal{O}_\omega\left(\frac{1}{\mu^{\frac{p}{2}k}}\right)\right]\left(\left(\frac{1}{2}+o_\omega(1)\right)T_k^{(1)}\right.\\
&\left.+\left(\frac{1}{2}+o_\omega(1)\right)T_k^{(3)}\right)=\left(\frac{\mu^2-1}{2}-\varepsilon_k\right)T_k^{(1,3)}=:\xi_k T_k^{(1,3)},
\end{aligned}$$

while the term outside the summation can similarly be rewritten as

$$(1+o_\omega(1))\frac{\sigma_{n+1}}{\tau_n}\left(\frac{\pi_n^{(3)}}{\Theta_n^{(2)}+\Theta_n^{(1)}}T_n^{(1)}+\frac{\pi_n^{(1)}}{\Theta_n^{(2)}+\Theta_n^{(3)}}T_n^{(3)}\right)\geq\left(\frac{\mu-1}{2}-\varepsilon_n\right)T_n^{(1,3)}=:\zeta_n T_n^{(1,3)}.$$

Let $\xi := (\mu^2-1)/2$ and $\zeta := (\mu-1)/2$ and note that $T_n^{(1,3)}$ is unbounded on $\mathcal{D}_0^{(2)}$, as it tends to infinity. As previously, the term which $T_{n+1}^{(1,3)}$ majorises, $\sum_{k=m}^{n-1}\xi_k T_k^{(1,3)} + \zeta_n T_n^{(1,3)}$, is generated by the following dynamical system, with $T_m = T_m^{(1,3)} > 1$ and

$S_m = 0$, taken as initial values, and the summation resulting from the first component of the orbit of

$$T_{n+1} = S_n + \zeta_n T_n$$
$$S_{n+1} = S_n + \xi_n T_n$$

for all $n \geq m$, which can be rewritten in matrix form as

$$\begin{pmatrix} T_{n+1} \\ S_{n+1} \end{pmatrix} = \begin{pmatrix} \zeta_n & 1 \\ \xi_n & 1 \end{pmatrix} \begin{pmatrix} T_n \\ S_n \end{pmatrix}.$$

Denote the iteration matrix A_n and let $X_n = (T_n, S_n)$. We need to improve on the asymptotic growth of $X_{n+1} = A_n X_n$, specifically of its first component, keeping in mind that in our case X_n is a positive vector (and each of its components diverge). We will achieve this improvement by bootstrapping again through [40, Theorem 1.3]. This is not the most powerful tool available for dynamical systems: better asymptotic growth could be achieved by exploiting results such as *Benzaid-Lutz Theorem*. However our current lack of understanding of the speed of convergence of the residual error term $1+o_\omega(1)$, and the speed with which the random coefficients next to $T_k^{(1)}$ and $T_k^{(3)}$ tend to $1/2$, make it difficult to apply such tools, which we therefore omit to describe (the curious reader is referred to [5] and [17, §8.4, Theorem 8.25] for background).

First of all note that the matrix $A_n \longrightarrow A$, where

$$A = \begin{pmatrix} \zeta & 1 \\ \xi & 1 \end{pmatrix}.$$

Since $\mu^2 > \mu$ for all $\mu > 1$, $\xi > \zeta > 0$ and therefore, starting at m large enough, one can assume $\xi_n > \zeta_n$. The characteristic polynomial of A_n and A are $p_n(\lambda) = \lambda^2 - (1+\zeta_n)\lambda + \zeta_n - \xi_n$ and $p(\lambda) = \lambda^2 - (1+\zeta)\lambda + \zeta - \xi$, yielding two simple eigenvalues $\lambda_n^\pm = [(\zeta_n + 1) \pm \sqrt{(\zeta_n + 1)^2 + 4(\xi_n - \zeta_n)}]/2$ and $\lambda_\pm = [(\zeta + 1) \pm \sqrt{(\zeta + 1)^2 + 4(\xi - \zeta)}]/2$ respectively. Note that

$$\lambda_\pm = \frac{\frac{\mu+1}{2} \pm \sqrt{\left(\frac{\mu+1}{2}\right)^2 + 2(\mu^2 - \mu)}}{2}.$$

Hence $\lambda_+ > (\mu + 1)/2$ (this is the Perron eigenvalue), and also $|\lambda_+| > |\lambda_-|$, with $\lambda_- < 0$. The corresponding eigenvectors of A can be chosen to be $e_\pm = (1, \lambda_\pm - \zeta)$, with $\lambda_+ - \zeta > 1$ (this is the Perron eigenvector) and $\lambda_- - \zeta < 0$. Hence the improvement achieved is that the existing limit $\lim_{n\longrightarrow\infty} \sqrt[n]{\|X_n\|_1} = \lambda_+ > (\mu + 1)/2$ and similarly $\lim_{n\longrightarrow\infty} \sqrt[n]{T_n} = \lambda_+ > (\mu + 1)/2$, which allows for a less restrictive choice of p in the estimated regime of growth of $T_n^{(1,3)}$.

Arguments similar to those exploited in Proposition 4.48, should be possible also on $\mathcal{D}_>^{(i)}$, but they will be technically more difficult, due to the oscillations of π_n. Complementing these findings with a deeper understanding of the dynamical system obeying (2.33) and (2.34), and its behaviour with p_n near the vertices, is likely the next step to succeed in showing that dominance is negligible. Simulations support that the dynamical system does not converge to the vertices, motivating Conjecture 2.3.

Part II

Generalised balls and bins with positive feedback

Chapter 5

Introduction

This introductory chapter is dedicated to generalised BB with positive feedback: we first describe the stochastic process, providing all the details missing from Chapter 1, in terms of a system of iterative equations, and then explain a heuristic argument regarding the onset of dominance. This heuristics is clearly reminiscent of the reduction to a *randomly perturbed dynamical system* exploited in Part I, but the approach followed will not rely directly on dynamical systems techniques. Indeed it will be much more natural to rely on *martingales* instead, due to how they arise from the iterative scheme.

The chapter goes on with a detailed description of our results and concludes with a note for the reader, regarding the relative material in the Appendix, and the introduction of further notation, in view of the more sophisticated use of martingales in this part of the dissertation.

5.1 Iterative equations of the model

Let us start by recalling the probabilistic time-dependent model of generalised BB with positive feedback. Recall that $d \geq 2$ denotes the arbitrary number of bins and that $\{\sigma_n\}$ is the integer-valued positive sequence, representing the number of added balls at times $n \in \mathbb{N}$. Recall that $T_0^{(i)}$ denotes the initial deterministic positive number of balls in the ith bin. For each $n \in \mathbb{N}_0$, let again

$$T_n^{(1)} + \ldots + T_n^{(d)} =: \tau_n = \tau_0 + \sum_{i=1}^{n} \sigma_i$$

be the total number of balls in the bins at time n. Finally recall that

$$\Theta_n^{(i)} := \frac{T_n^{(i)}}{\tau_n}$$

is the proportion of balls in the ith bin at time n. The vector Θ_n is valued in the standard simplex in d dimensions

$$\Delta^{d-1} := \{(x_1, \ldots, x_d) \in [0,1]^d : \ x_1 + \ldots + x_d = 1\},$$

so we let, for every integer $1 \leq i \leq d$, and $x \in \Delta^{d-1}$,

$$\psi^{(i)}(x) := \frac{x_i^{\alpha}}{\sum_{j=1}^{d} x_j^{\alpha}}.$$

Note that for simplicity, when there is no ambiguity, we switch from an upper index notation for the components, to a lower index notation, if the time index is not involved. In a model with no feedback, we set $\alpha = 1$, but, unlike for the ERBRW, in our study of generalised BB we will focus mainly on the model with feedback, that is $\alpha > 1$, which is the more challenging one.

Given the past up to time n, each of the σ_{n+1} balls, which are added to the ith bin independently of each other at random at time $n+1$, will fall in the bin with probability proportional to the number of balls already inside, $T_n^{(i)}$, that is with probability $\psi^{(i)}(\Theta_n)$. The number of balls going in the ith bin at time $n+1$ will be a random variable denoted as $B_{n+1}^{(i)}$. More precisely, according to this model, for every integer $1 \leq i \leq d$,

$$T_{n+1}^{(i)} = T_n^{(i)} + B_{n+1}^{(i)} \tag{5.1}$$

where, conditionally on the past, the random number of balls, which independently of each other are added to the ith bin at time $n+1$, will be modelled as a binomial random variable $B_{n+1}^{(i)} \sim \mathrm{Bin}(\sigma_{n+1}, P_n^{(i)})$, where σ_{n+1} is the size and $P_n^{(i)} := \psi^{(i)}(\Theta_n)$ the probability parameter of the binomial. Therefore the random vector B_{n+1} having d components $B_{n+1}^{(i)}$ is, by construction, a *multinomial* random vector of size σ_{n+1} and vector $P_n = \psi(\Theta_n) = (\psi^{(1)}(\Theta_n), \dots, \psi^{(d)}(\Theta_n))$ of probability parameters (we will use the notation $B_{n+1} \sim \mathrm{Multin}(\sigma_{n+1}, P_n)$). This is true because a multinomial random vector can be seen as a random vector having binomials as marginals, which share the same size and have probability parameters adding up to one, constrained to always add up to the common size. Since in the BB model by construction

$$B_{n+1}^{(i)} = \sigma_{n+1} - \sum_{j \neq i} B_{n+1}^{(j)}$$

and

$$\sum_{i=1}^{d} \psi^{(i)}(x) = 1$$

for every $x \in \Delta^{d-1}$, the claim that B_{n+1} is a multinomial random vector of size σ_{n+1} and vector of probabilities $P_n = \psi(\Theta_n)$ given the past, follows for every $n \geq 0$. We also note that B_{n+1} is independent of $\mathscr{F}_n = \sigma(B_1, ..., B_n)$. Recall that as usual $\{\mathscr{F}_n\}_{n \in \mathbb{N}_0}$ will be the natural filtration on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$, with $\mathscr{F}_0 = \{\emptyset, \Omega\}$ and $\mathscr{F}_\infty := \sigma\left(\bigcup_{n=0}^{\infty} \mathscr{F}_n\right)$; we will denote by $\mathbb{P}_{\mathscr{F}_n}$, $\mathbb{E}_{\mathscr{F}_n}$, $\mathrm{Var}_{\mathscr{F}_n}$ and $\mathrm{Cov}_{\mathscr{F}_n}$ the conditional probability, expectation, variance and covariance.

Lastly we introduce the *normalized fluctuations* of $B_n^{(i)}$. For every $i \in \{1, 2, \dots, d\}$, let

$$\varepsilon_{n+1}^{(i)} := \frac{B_{n+1}^{(i)} - \sigma_{n+1} P_n^{(i)}}{\sqrt{\sigma_{n+1} P_n^{(i)} (1 - P_n^{(i)})}}.$$

Clearly $\mathbb{E}_{\mathscr{F}_n} B_{n+1}^{(i)} = \sigma_{n+1} P_n^{(i)}$ and $\mathrm{Var}_{\mathscr{F}_n} B_{n+1}^{(i)} = \sigma_{n+1} P_n^{(i)} (1 - P_n^{(i)})$, so $\mathbb{E}_{\mathscr{F}_n} \varepsilon_{n+1}^{(i)} = 0$ and $\mathbb{E}_{\mathscr{F}_n} (\varepsilon_{n+1}^{(i)})^2 = 1$.

In this model only two assumptions will be made on the sequence of number of balls added at time n:

$$\sigma_n \text{ is either bounded or diverges to infinity,} \tag{S}$$

$$\rho_n := \frac{\sigma_{n+1}}{\tau_n} \text{ is either bounded or diverges to infinity.} \tag{R}$$

Even though in Chapter 1 we motivated the study of generalisations of BB models with their connection to the ERBRWs, it is also true that in this part of our work, we study these models rather independently, trying to generalise results that hold for any number of bins d to as many regimes of growth of $\{\sigma_n\}$ as possible. Indeed, many technicalities that we will face will arise from regimes quite different from $\sigma_n = \mu^n$, and we will spend no little effort to resolve them (in this sense, Proposition 8.4 will be emblematic). In considering these assumptions, the reader might find helpful to consider a few key examples regarding Assumption R, which gauges how fast the number of balls added grows. If $\{\sigma_n\}$ is bounded, clearly ρ_n vanishes. Similarly, by the power sum formula, if $\{\sigma_n\}$ is polynomial ρ_n vanishes too. If $\{\sigma_n\}$ is exponential, say $\sigma_n = \mu^n$ for some integer $\mu > 1$, then by the geometric formula the limit is the constant $\mu - 1 > 0$. We interpret all such regimes as slow growth, being ρ_n bounded. As an instance of fast growth, one can take $\sigma_n = \mu^{\mu^n}$, for which ρ_n diverges to infinity, since eventually $\tau_n \leq (n+1)\mu^{\mu^n}$ and therefore $\rho_n \geq \mu^{(\mu-1)\mu^n}/(n+1) \longrightarrow \infty$.

We now derive the fundamental iterative equation obeyed by the model. It will be useful, for simplicity, to adopt for a moment a vector notation, instead of a componentwise one, and thus rewrite (5.1) as $T_{n+1} = T_n + B_{n+1}$. Then divide both sides by τ_{n+1}, yielding

$$\frac{T_{n+1}}{\tau_{n+1}} = \frac{T_n + B_{n+1}}{\tau_{n+1}} = \frac{\tau_n \Theta_n + B_{n+1}}{\tau_{n+1}}.$$

Thus we conclude that

$$\Theta_{n+1} = \frac{\tau_n}{\tau_{n+1}}\Theta_n + \frac{1}{\tau_{n+1}} B_{n+1}. \tag{5.2}$$

We can also rewrite (5.2) componentwise, and by exploiting the random fluctuations, it follows that for every $i \in \{1, \dots, d\}$

$$\Theta^{(i)}_{n+1} = \frac{\tau_n}{\tau_{n+1}}\Theta^{(i)}_n + \frac{1}{\tau_{n+1}}\left(\sigma_{n+1}P^{(i)}_n + \varepsilon^{(i)}_{n+1}\sqrt{\sigma_{n+1}P^{(i)}_n(1-P^{(i)}_n)}\right)$$

and therefore

$$\Theta^{(i)}_{n+1} = \frac{\tau_n}{\tau_{n+1}}\Theta^{(i)}_n + \frac{\sigma_{n+1}}{\tau_{n+1}}P^{(i)}_n + \frac{\varepsilon^{(i)}_{n+1}}{\tau_{n+1}}\sqrt{\sigma_{n+1}P^{(i)}_n(1-P^{(i)}_n)}. \tag{5.3}$$

5.2 A useful heuristics for *dominance*

The main objective in this part of our work is finding the probability, for any $d > 2$ and depending on the feedback and regime of growth of $\{\sigma_n\}$, of the events of dominance and monopoly, which are respectively defined as the events $\mathcal{D}$, such that all but one of the proportions of balls in the bins are negligible, and $\mathcal{M}$ such that all but one of the bins eventually stops receiving balls, as per Definitions 1.8 and 1.9. The main focus will be on dominance. An interesting heuristic approach to describing the onset of dominance is the following: in (5.2) we can extract the martingale differences by adding and subtracting $\sigma_{n+1}P_n$ to the multinomials, so as to get

$$\Theta_{n+1} = \frac{\tau_n}{\tau_{n+1}}\Theta_n + \frac{\sigma_{n+1}}{\tau_{n+1}}P_n + \frac{B_{n+1} - \sigma_{n+1}P_n}{\tau_n}.$$

One expects negligibility of the martingale differences, so by neglecting them we are left with the dynamical system $\Theta_{n+1} = \psi_n(\Theta_n)$, where

$$\psi_n(x) := \frac{\tau_n}{\tau_{n+1}}x + \frac{\sigma_{n+1}}{\tau_{n+1}}\psi(x),$$

and $\psi(x) := (\psi^{(1)}(x), \dots, \psi^{(d)}(x))$.

With $\alpha = 1$, by (5.2), $\{\Theta_n\}$ is a martingale, thus almost sure convergence of $\{\Theta_n\}$ with no feedback is established by *Doob's Forward Convergence Theorem*. Can convergence be to the vertices of the simplex, with positive probability? Note that with no feedback, the iteration map $\psi_n(x)$ is the identity map; therefore heuristically, starting within the simplex, $\{\Theta_n\}$ should quickly converge to points within the simplex because the iteration map, which broadly speaking drives the deterministic system, does not have any destabilising effect. Thus one would expect no dominance for $\alpha = 1$.

With $\alpha > 1$, the system has an iteration map, whose equilibrium points are characterised by the fixed point equation $x = \psi_n(x)$, which is equivalent to the equation

$$x_i = \frac{x_i^\alpha}{\sum_{j=1}^d x_j^\alpha},$$

whose solutions on the simplex are the vertices and the centres of the l-faces for $l \in \{1, \dots, d-2\}$

$$E_l := \Big(\underbrace{\frac{1}{l+1}, \dots, \frac{1}{l+1}}_{l+1}, \underbrace{0, \dots, 0}_{d-l-1}\Big),$$

with the centre of the simplex denoted as $E := E_{d-1} = \frac{\mathbf{1}}{\mathbf{d}}$ (we call these centres equilibria due to this dynamical interpretation). If the Jacobian matrix has eigenvalues with modulus greater than 1 at the equilibria of the simplex, one expects the dynamical system not to converge to these equilibria. If the vertices are stable under this dynamics, then they are the likely candidates for the limit of $\{\Theta_n\}$. This interpretation is obviously very loose, since the system is nonautonomous. However, particularly instructive can be the case for $d = 2$, studied in [48]. Here one can effortlessly overcome the difficulty of extending the Jacobian test for stability, as we can perform cobwebbing. In this case we only consider the first component of Θ_n: due to symmetry, the system is essentially univariate. The map $\psi_n^{(1)}(x)$ is then the convex combination of the identity map and the function

$$\psi^{(1)}(x) := \frac{x_1^\alpha}{x_1^\alpha + (1 - x_1)^\alpha}$$

through the coefficients ${}^{\tau_n}/_{\tau_{n+1}}$ and ${}^{\sigma_{n+1}}/_{\tau_{n+1}}$. This observation makes us rely less on the Jacobian criterion, since we can see through *cobwebbing*, where the iteration map leads the system, even as it varies. Although the convex combination is time dependent, the graph of $\psi_n^{(1)}$ will always maintain its qualitative profile (see Figure 5.1), and therefore cobwebbing takes the systems away from $1/2$, towards either 0 or 1. The time-dependency entails that the speed, at which this happens, varies; thus, when considering the original stochastic process, the control over the neglected random fluctuations enters the picture, especially close to $1/2$. One must ensure that they do not overpower the destabilizing effect that the feedback introduces, when close to the equilibrium. The case $d = 2$ is very fortunate: thanks to the symmetry, it is essentially univariate, and it has no such issue as partial equilibria of the bins, which significantly complicate the analysis, as we will see in Chapter 8. Nonetheless, it is this heuristic approach, which inspires the overall strategy we will follow, when trying to prove almost sure dominance for $d > 2$ bins in Theorem 1.11.

5.3 Outline of contents

We start with an overview of the results concerning dominance, which is the main focus of this part. In Chapter 7, Theorem 1.10 shows that dominance is negligible if $\alpha = 1$.

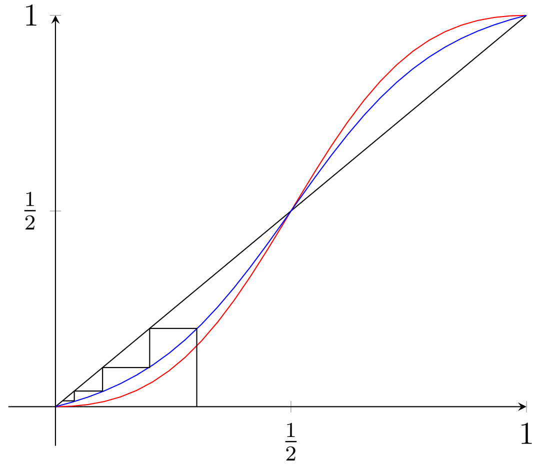


Figure 5.1: $\psi^{(1)}$ (red) and $\psi_n^{(1)}$ (blue) and cobwebbing

The result follows from the convergence of martingales, as previously mentioned, and the fact that with no feedback, the regime is nonmonopolistic (Theorem 1.14). In Chapter 9, Theorem 1.11 shows that when $\alpha > 1$, dominance is almost sure if:

- ρ_n is bounded ($\sigma_n = \mu^n$ belongs to this rate of growth, which is therefore the one we are more interested into, being the regime of the ERBRW);
- $\rho_n \longrightarrow \infty$, $\theta := \lim_{n\to\infty} \frac{\log \tau_n}{\alpha^n} = 0$, $\lambda := \limsup_{n\to\infty} \frac{\sigma_{n+1}\sigma_{n-1}^{\alpha}}{\sigma_n^{\alpha+1}} < 1$.

Even though with feedback $\{\Theta_n\}$ is no longer a martingale, in the argument we exploit martingale techniques on the *Doob's decomposition* of $\{\Theta_n\}$. In addition, we rely on two key facts.

- $\{\Theta_n\}$ undergoes infinitesimal deviations from the equilibria of the simplex under positive feedback (Proposition 8.4, proved in Chapter 8). This means that Θ_n is almost never eventually confined in a δ_n-neighbourhood of the equilibria $\{E_l\}$, where δ_n will be a suitably defined vanishing positive sequence. These deviations are exploited to start the inductive argument of the theorem, which shows that one by one, all components of Θ_n vanish, except the last one.
- In both cases (ρ_n is bounded and $\rho_n \longrightarrow \infty$ with $\theta = 0$, $\lambda < 1$), we have that if $i \in \{1, \ldots, d\}$ is such that $\Theta_n^{(i)} \longrightarrow 0$ as $n \longrightarrow \infty$, then $T_n^{(i)}$ is bounded (respectively Lemmas E.2 and F.1). This allows quantitative estimates on the vanishing rate of $\Theta_n^{(i)}$ that are exploited in the inductive argument of the theorem, so as to ensure that from one step to the next, the components of Θ_n, which have been shown to be vanishing, do so at a fast enough rate. This part of the argument is highly technical, and we rely on applications of the implicit function theorem to carry out the induction step (Lemma 9.1).

These two facts combined, place Θ_n where, informally speaking, the iteration map ψ_n further pushes it away from the equilibria, towards some vertex, as we described in the heuristic section. In Conjecture 1.12 we claim that the conditions $\theta = 0$ and $\lambda < 1$ are not necessary. One ground for this conjecture is [48, Theorem 1.2], that is the fact that for $d = 2$ bins, this has already been proved; a second ground is that the issue arising in the higher-dimensional setting seems to be more of technical nature. Although in the argument of Theorem 1.11 described above we rely on Lemmas E.2 and F.1, this reliance could be relaxed.

Before moving on to describing the results concerning monopoly, it is perhaps useful to provide a more intuitive interpretation of the parameters involved in these arguments. Recall that $\{\rho_n\}$ divides the various regimes of growth of $\{\sigma_n\}$ in two categories: *slow groth* when it is bounded, *fast growth* when it diverges to infinity. We think of the quantity λ as a regularity parameter. For example, if it is a limit, for $\rho_n \longrightarrow \infty$ and $\theta = 0$ it is always zero (see [48, Lemma 6.1]). To get λ large, one can think of irregular examples, such as

$$\sigma_n = \begin{cases} \mu^n, & n \equiv 0 \bmod(3) \\ 1, & n \equiv 1, 2 \bmod(3) \end{cases},$$

where $\mu > 1$, in which case it is not even finite, since for all $k \in \mathbb{N}$, $\lambda_{3k+1} = \sigma_{3k+2}\sigma_{3k}^{\alpha}\sigma_{3k+1}^{-\alpha-1} = \mu^{3\alpha k}$. Note that we implicitly made one additional regularity assumption, namely the existence of $\theta := \lim_{n\to\infty} \alpha^{-n} \log \tau_n \in [0, \infty]$. If it exists, θ is a parameter which reveals finer details of the regime of growth of fast growing $\{\sigma_n\}$, since the logarithm allows to discern between different rates of fast growth. For example, all slow regimes previously mentioned yield $\theta = 0$, but a fast regime like $\sigma_n = \mu^{\mu^n}$ for $\mu > 1$, does not necessarily imply $\theta = \infty$. In fact, note that whenever $\rho_n \longrightarrow \infty$, $\tau_n \sim \sigma_n$ because

$$\frac{\sigma_n}{\tau_n} = \frac{1}{1 - \frac{1}{\rho_{n-1}}} \longrightarrow 1,$$

as $n \longrightarrow \infty$. Therefore, having denoted $\theta_n := \alpha^{-n} \log \tau_n$ so that $\theta = \lim_{n\longrightarrow\infty} \theta_n$, we can see that $\theta_n \sim (\mu/\alpha)^n \log \mu$. Then if $\mu < \alpha$, one has that $\theta = 0$; if $\mu = \alpha$, $\theta = \log \alpha$; if $\mu > \alpha$, $\theta = \infty$.

The assumption on the existence of the limit of $\theta_n := \alpha^{-n} \log \tau_n$ is crucial only for the study of monopoly, to which we now briefly turn our attention to. We consider these results as secondary, so we will not describe the arguments with as much detail as we did for those concerning dominance. Extending the results for $d = 2$, we recall that we show that for $d > 2$ monopoly almost never occurs in the following scenarios:

- if there is no feedback, by Theorem 1.14;
- if there is feedback, in *supercritical regime*, that is $\theta = \infty$ (and thus $\rho_n \longrightarrow \infty$), by Theorem 1.15;
- if there is feedback, in *subcritical regime* (that is $\theta = 0$) with ρ_n diverging to infinity and $\lambda > 1$, by Theorem 1.16.

The proof of these results can be found in Chapter C, and they are a straightforward adaptation from the corresponding results in [48]. In Corollary 1.13 we show that with positive feedback, monopoly occurs almost surely in all regimes of growth covered by Theorem 1.11, that is when:

- ρ_n is bounded;
- $\rho_n \longrightarrow \infty$, $\theta = 0$ and $\lambda < 1$.

We did not conduct the study of monopoly for $d > 2$ under the *critical regime* ($\alpha > 1$, $\theta \in (0, \infty)$), due to our main focus being the application of this approach to the ERBRW. The reader with a keen interest in BB models will find beneficial reading [48, §1], as it contains interesting heuristic remarks regarding monopoly, which we did not discuss here, and [48, §8], which offers an in depth study of monopoly in the critical regime for $d = 2$. Also, note that we did not study the case of negative feedback ($0 < \alpha < 1$) under time-dependence.

5.4 Note for the reader

Part IV contains several technical results that are vital to the main arguments of our proofs, but they are not an original contribution of the present work. They are in the appendix, as they have easily been adapted from the corresponding results shown in [48]. They are included for self-containedness.

5.5 Notation

To conclude we mention some additional notation which we will adopt throughout this part.

- We will denote $[d] \coloneqq \{i \in \mathbb{N} : 1 \leq i \leq d\}$.
- The Euclidean norm will be denoted as $\|\cdot\|$.
- The Hilbert space of square integrable random variables $\mathscr{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ will be denoted simply by $\mathscr{L}^2$. For a random variable X the notation $X \in \mathscr{L}^2(\mathscr{G})$ for some sub-σ-algebra $\mathscr{G}$ of $\mathscr{F}$, means that almost surely $\mathbb{E}_{\mathscr{G}} X^2 < \infty$, where $\mathbb{E}_{\mathscr{G}}$ denotes the expectation conditional on the sub-σ-algebra $\mathscr{G}$, and is therefore a random variable itself. When a sequence of random variables $X_n \in \mathscr{L}^2(\mathscr{G})$ converges to a random variable X in $\mathscr{L}^2(\mathscr{G})$, this means by definition that almost surely $\lim_{n \longrightarrow \infty} \mathbb{E}_{\mathscr{G}}(X_n - X)^2 = 0$.
- To say that a random variable X is measurable with respect to a sub-σ-algebra $\mathscr{G}$ of $\mathscr{F}$, we will use the notation $X \in \mathrm{m}\mathscr{G}$.

Chapter 6

No monopoly

In this section we show a fundamental condition for nonmonopolistic regimes, which we adapt from [48, Lemma 2.1].

Lemma 6.1. *Assume*

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_n^{\alpha}} = \infty.$$

Then $\mathbb{P}(\forall i \in [d],\ T_n^{(i)} \longrightarrow \infty) = 1$, *and therefore* $\mathbb{P}(\mathcal{M}) = 0$.

Proof. For the second part it is sufficient to note that

$$\begin{aligned}\mathcal{M}^c &= \left(\bigcup_{i=1}^{d} \{T_n^{(i)} = \sigma_n,\ \text{ev.}\}\right)^c = \bigcap_{i=1}^{d} \{T_n^{(i)} = \sigma_n,\ \text{ev.}\}^c = \bigcap_{i=1}^{d} \{T_n^{(i)} < \sigma_n,\ \text{i.o.}\} \\ &= \{\forall\, i \in [d],\ T_n^{(i)} < \sigma_n,\ \text{i.o.}\} \supset \{\forall\, i \in [d],\ T_n^{(i)} \to \infty\}\,.\end{aligned}$$

Then $\mathbb{P}(\forall i \in [d],\ T_n^{(i)} \to \infty) = 1$ directly implies $\mathbb{P}(\mathcal{M}) = 0$. We will focus on showing that

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_n^{\alpha}} = \infty$$

implies $\mathbb{P}(\forall i \in [d],\ T_n^{(i)} \to \infty) = 1$ by a componentwise martingale argument. Note that for every $i \in [d]$,

$$T_n^{(i)} = T_0^{(i)} + \sum_{j=1}^{n} B_j^{(i)} = T_0^{(i)} + M_n^{(i)} + Y_n^{(i)},$$

where

$$M_n^{(i)} := \sum_{j=1}^{n} (B_j^{(i)} - \sigma_j P_{j-1}^{(i)})$$

with $M_0^{(i)} = 0$, and

$$Y_n^{(i)} := \sum_{j=1}^{n} \sigma_j P_{j-1}^{(i)}$$

with $Y_0^{(i)} = 0$. Note that for every $i \in [d]$, $Y_n^{(i)}$ almost surely diverges to infinity under the assumption that

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_n^{\alpha}} = \infty.$$

In fact by (8.1) and the nondecreasing monotonicity of $T_n^{(i)} \geq 1$,

$$Y_n^{(i)} = \sum_{j=1}^{n} \sigma_j \psi^{(i)}(\Theta_{j-1}) \geq \sum_{j=1}^{n} \sigma_j (\Theta_{j-1}^{(i)})^\alpha = \sum_{j=1}^{n} \frac{\sigma_j (T_{j-1}^{(i)})^\alpha}{\tau_{j-1}^\alpha} \geq \sum_{j=1}^{n} \frac{\sigma_j}{\tau_{j-1}^\alpha} = \sum_{k=0}^{n} \frac{\sigma_{k+1}}{\tau_k^\alpha}. \tag{6.1}$$

Therefore, proving that

$$\{\forall\, i \in [d],\ Y_n^{(i)} \to \infty\} \subseteq \{\forall\, i \in [d],\ T_n^{(i)} \to \infty\}$$

would yield the claim. In order to do so, first we will show that $M_n^{(i)}$ is a centred $\mathscr{F}_n$-martingale; next, by exploiting the angle bracket process, we show that $Y_n^{(i)} \to \infty$ implies $T_n^{(i)} \to \infty$.

Step 1. Clearly $M_n^{(i)}$ is trivially adapted to $\{\mathscr{F}_n\}$ as $B_n^{(i)}$ is too. Since $\mathbb{E}_{\mathscr{F}_{j-1}} B_j^{(i)} = \sigma_j P_{j-1}^{(i)}$ and $P_{j-1}^{(i)} \in \mathrm{m}\mathscr{F}_{j-1}$, by the tower property it holds that

$$\mathbb{E} M_n^{(i)} = \sum_{j=1}^{n} \mathbb{E}(B_j^{(i)} - \sigma_j P_{j-1}^{(i)}) = \sum_{j=1}^{n} \mathbb{E}\mathbb{E}_{\mathscr{F}_{j-1}}(B_j^{(i)} - \sigma_j P_{j-1}^{(i)}) = 0,$$

thus $M_n^{(i)}$ is centred. Moreover, for every n,

$$\begin{aligned}\mathbb{E}|M_n^{(i)}| &\leq \sum_{j=1}^{n} \mathbb{E}|B_j^{(i)} - \sigma_j P_{j-1}^{(i)}| = \sum_{j=1}^{n} \mathbb{E}\mathbb{E}_{\mathscr{F}_{j-1}}|B_j^{(i)} - \sigma_j P_{j-1}^{(i)}| \leq \\ &\sum_{j=1}^{n} \mathbb{E}\mathbb{E}_{\mathscr{F}_{j-1}}(B_j^{(i)} + \sigma_j P_{j-1}^{(i)}) = 2\sum_{j=1}^{n} \sigma_j \mathbb{E} P_{j-1}^{(i)} \leq 2\sum_{j=1}^{n} \sigma_j = 2(\tau_n - \tau_0) < \infty.\end{aligned}$$

Lastly, since $B_j^{(i)} \in \mathrm{m}\mathscr{F}_{n-1}$ and $P_{j-1}^{(i)} \in \mathrm{m}\mathscr{F}_{n-1}$ for all $1 \leq j \leq n$,

$$\mathbb{E}_{\mathscr{F}_{n-1}} M_n^{(i)} = \sum_{j=1}^{n} \mathbb{E}_{\mathscr{F}_{n-1}}(B_j^{(i)} - \sigma_j P_{j-1}^{(i)}) = 0 + \sum_{j=1}^{n-1}(B_j^{(i)} - \sigma_j P_{j-1}^{(i)}) = M_{n-1}^{(i)},$$

so for every $i \in [d]$, $M_n^{(i)}$ is an $\mathscr{F}_n$-martingale.

Step 2. Consider that $M_n^{(i)} \in \mathscr{L}^2$. Indeed

$$\begin{aligned}\mathbb{E}(M_n^{(i)})^2 &= \mathbb{E}(M_n^{(i)} - M_{n-1}^{(i)} + M_{n-1}^{(i)} - M_0^{(i)})^2 = \mathbb{E}(M_n^{(i)} - M_{n-1}^{(i)})^2 + \mathbb{E}(M_{n-1}^{(i)})^2 \\ &+ 2\mathbb{E}\left[(M_n^{(i)} - M_{n-1}^{(i)})(M_{n-1}^{(i)} - M_0^{(i)})\right] = \mathbb{E}(M_n^{(i)} - M_{n-1}^{(i)})^2 + \mathbb{E}(M_{n-1}^{(i)})^2,\end{aligned}$$

since

$$\mathbb{E}\left[(M_n^{(i)} - M_{n-1}^{(i)})(M_{n-1}^{(i)} - M_0^{(i)})\right] = \mathbb{E}\mathbb{E}_{\mathscr{F}_{n-1}}\left[(M_n^{(i)} - M_{n-1}^{(i)})(M_{n-1}^{(i)} - M_0^{(i)})\right]$$

and

$$\mathbb{E}_{\mathscr{F}_{n-1}}\left[(M_n^{(i)} - M_{n-1}^{(i)})(M_{n-1}^{(i)} - M_0^{(i)})\right] = (M_{n-1}^{(i)} - M_0^{(i)})\mathbb{E}_{\mathscr{F}_{n-1}}(M_n^{(i)} - M_{n-1}^{(i)}) = 0$$

by the martingale property. Hence

$$\begin{aligned}\mathbb{E}(M_n^{(i)})^2 &= \mathbb{E}(M_n^{(i)} - M_{n-1}^{(i)})^2 + \mathbb{E}(M_{n-1}^{(i)})^2 = \mathbb{E}(M_{n-1}^{(i)})^2 + \mathbb{E}\mathbb{E}_{\mathscr{F}_{n-1}}(M_n^{(i)} - M_{n-1}^{(i)})^2 \\ &= \mathbb{E}(M_{n-1}^{(i)})^2 + \mathbb{E}\mathbb{E}_{\mathscr{F}_{n-1}}\left(B_n^{(i)} - \sigma_n P_{n-1}^{(i)}\right)^2 = \mathbb{E}(M_{n-1}^{(i)})^2 + \mathbb{E}\,\mathrm{Var}_{\mathscr{F}_{n-1}} B_n^{(i)} \\ &= \mathbb{E}(M_{n-1}^{(i)})^2 + \sigma_n \mathbb{E}\left[P_{n-1}^{(i)}(1 - P_{n-1}^{(i)})\right] \leq \mathbb{E}(M_{n-1}^{(i)})^2 + \sigma_n.\end{aligned}$$

By iterating n times $\mathbb{E}(M_n^{(i)})^2 - \mathbb{E}(M_{n-1}^{(i)})^2 \leq \sigma_n$, and adding up all inequalities, we obtain

$$\mathbb{E}(M_n^{(i)})^2 \leq \sum_{j=1}^{n} \sigma_j = \tau_n - \tau_0 < \infty.$$

Then the angle bracket process is well defined as the previsible part of the *Doob's decomposition* of the submartingale $(M_n^{(i)})^2$ and denoted as

$$\begin{aligned}\langle M^{(i)} \rangle_n &:= \sum_{j=1}^{n} \mathbb{E}_{\mathscr{F}_{j-1}}(M_j^{(i)} - M_{j-1}^{(i)})^2 = \sum_{j=1}^{n} \mathbb{E}_{\mathscr{F}_{j-1}}(B_j^{(i)} - \sigma_j P_{j-1}^{(i)})^2 = \sum_{j=1}^{n} \operatorname{Var}_{\mathscr{F}_{j-1}} B_j^{(i)} \\ &= \sum_{j=1}^{n} \sigma_j P_{j-1}^{(i)}(1 - P_{j-1}^{(i)}) \leq Y_n^{(i)}.\end{aligned}$$

The standard theory of the angle bracket process now applies. Recall the following facts:

- if $\langle M^{(i)} \rangle_n$ converges, then $M_n^{(i)}$ converges almost surely by [50, §12.13];
- if $\langle M^{(i)} \rangle_n$ diverges, then $M_n^{(i)}/\langle M^{(i)} \rangle_n$ vanishes by [50, §12.14].

By (6.1), the assumption of the claim implies that $Y_n^{(i)} \longrightarrow \infty$. We have two cases to inspect.

Case 1. If $\langle M^{(i)} \rangle_n$ converges, then $M_n^{(i)}$ converges, and $T_n^{(i)} = T_0^{(i)} + M_n^{(i)} + Y_n^{(i)} \longrightarrow \infty$;
Case 2. If $\langle M^{(i)} \rangle_n$ diverges, then $M_n^{(i)}/\langle M^{(i)} \rangle_n$ vanishes, and therefore

$$T_n^{(i)} = Y_n^{(i)} \left(1 + \frac{T_0^{(i)}}{Y_n^{(i)}} + \frac{M_n^{(i)}}{Y_n^{(i)}}\right) = Y_n^{(i)}\left(1 + 2o_\omega(1)\right) \longrightarrow \infty,$$

since

$$\left|\frac{M_n^{(i)}}{Y_n^{(i)}}\right| = \frac{|M_n^{(i)}|}{Y_n^{(i)}} \leq \frac{|M_n^{(i)}|}{\langle M^{(i)} \rangle_n} = \left|\frac{M_n^{(i)}}{\langle M^{(i)} \rangle_n}\right| \longrightarrow 0$$

and therefore $M_n^{(i)}/Y_n^{(i)}$ vanishes.

Hence $T_n^{(i)} \longrightarrow \infty$.

□

This lemma yields the three main results on the negligibility of monopoly. The proofs are found in Chapter C. Recall that $\theta_n := \alpha^{-n} \log \tau_n$ and assume $\lim_{n\to\infty} \theta_n =: \theta \in \mathbb{R} \cup \infty$.

- Since by Lemma C.1, $\sum_{n=1}^{\infty} \sigma_n/\tau_n = \infty$, Theorem 1.14 exploits Lemma 6.1 to show that if there is no feedback, monopoly is negligible.
- When $\alpha > 1$ and $\theta = \infty$, recall that the model is said to be in supercritical regime. In this regime Theorem 1.15 exploits Lemma 6.1 to show the negligibility of monopoly.
- When $\alpha > 1$ and $\theta = 0$, recall that the model is said to be in subcritical regime. In this regime, Theorem 1.16 exploits Lemma 6.1 to show that if

$$1 < \lambda := \limsup_{n\longrightarrow\infty} \frac{\sigma_{n+1}\sigma_{n-1}^{\alpha}}{\sigma_n^{\alpha+1}},$$

monopoly is negligible.

Chapter 7

No dominance

In this chapter we prove that dominance is negligible in absence of feedback.

Proof of Theorem 1.10. Since $\Theta_n \in \Delta^{d-1}$, $0 \leq \|\Theta_n\| \leq 1$, so Θ_n is bounded. By (5.2), it can be seen that Θ_n is a martingale:

$$\mathbb{E}_{\mathscr{F}_{n-1}}(\Theta_n) = \frac{\tau_{n-1}}{\tau_n}\mathbb{E}_{\mathscr{F}_{n-1}}(\Theta_{n-1}) + \frac{1}{\tau_n}\mathbb{E}_{\mathscr{F}_{n-1}}(B_n) = \frac{\tau_{n-1}}{\tau_n}\Theta_{n-1} + \frac{\sigma_n}{\tau_n}P_{n-1}$$
$$= \frac{\tau_{n-1}}{\tau_n}\Theta_{n-1} + \frac{\sigma_n}{\tau_n}\Theta_{n-1} = \Theta_{n-1},$$

since $\Theta_{n-1} \in \mathrm{m}\mathscr{F}_{n-1}$ and $P_{n-1} = \psi(\Theta_{n-1}) = \Theta_{n-1}$, because $\alpha = 1$ and $\tau_{n-1} + \sigma_n = \tau_n$. By *Doob's forward convergence theorem*, Θ_n converges almost surely to a bounded random variable Θ.

To show that dominance is negligible, we actually show a stronger claim: that the event, on which any of the proportions of balls in the bins vanishes, is negligible. Define the event

$$\tilde{\mathcal{D}} := \{\exists i \in [d] : \Theta^{(i)} = 0\} = \bigcup_{i=1}^{d} \tilde{\mathcal{D}}_i$$

where $\tilde{\mathcal{D}}_i := \{\Theta^{(i)} = 0\}$. Note that $\mathcal{D} \subseteq \tilde{\mathcal{D}}$. We show that $\mathbb{P}(\tilde{\mathcal{D}}) = 0$. This holds, since for every $i \in [d]$, on the event $\tilde{\mathcal{D}}_i$, $\sum_{j\neq i} \Theta_n^{(j)}$ can be seen as the proportion of a second bin (resulting from merging all the $d-1$ bins other than the ith, which vanishes) approaching unity. In fact since $\alpha = 1$, trivially

$$\Theta_n = \psi(\Theta_n) := \frac{1}{\sum_{j=1}^{d} \Theta_n^{(j)}} \begin{pmatrix} \Theta_n^{(1)} \\ \vdots \\ \Theta_n^{(d)} \end{pmatrix} = \frac{1}{\Theta_n^{(i)} + \left(\sum_{j\neq i} \Theta_n^{(j)}\right)} \begin{pmatrix} \Theta_n^{(1)} \\ \vdots \\ \Theta_n^{(d)} \end{pmatrix}.$$

In this sense, $\tilde{\mathcal{D}}$ can be seen as the event of dominance in the two bins model, so by Theorem D.1, $\mathbb{P}(\tilde{\mathcal{D}}) = 0$. Hence $0 \leq \mathbb{P}(\mathcal{D}) \leq \mathbb{P}(\tilde{\mathcal{D}}) = 0$, and the claim follows. □

Note that for $\alpha > 1$ merging $d-1$ bins as in Theorem 1.10, would not yield a probabilistic model equivalent to the two bins model, because in general

$$\sum_{j=1}^{d} (\Theta_n^{(j)})^\alpha \neq (\Theta_n^{(i)})^\alpha + \left(\sum_{j\neq i} \Theta_n^{(j)}\right)^\alpha.$$

Chapter 8

Getting away from the equilibrium

This chapter is dedicated to deriving a key result for the BB process with positive feedback: that, informally speaking, the vector of proportions Θ_n does not get stuck in any of the equilibrium points (geometrically, recall that by equilibria we refer to the centres of the faces of the unit simplex). This fact will be the key to reach a proof of dominance in presence of feedback, recalling that dominance consists in Θ_n converging to the vertices of the simplex.

8.1 Preliminaries

Let us begin this section with deriving some facts about the components $\psi^{(i)}$ of the vector field $\psi : \Delta^{d-1} \longrightarrow \Delta^{d-1}$, with $\alpha > 1$. They will be frequently used throughout the rest of the chapter. The first fact is a straightforward bound for $\psi^{(i)}$ on Δ^{d-1}, obtained via the Lagrange multipliers method (since $\Theta_n \in \Delta^{d-1}$ will be the argument of $\psi^{(i)}$ in all proofs, all arguments denoted as x belong to the simplex).

Remark 8.1. *For every $i \in [d]$*

$$x_i^\alpha \leq \psi^{(i)}(x) \leq d^{\alpha-1} x_i^\alpha \tag{8.1}$$

on Δ^{d-1}.

Proof. Consider the denominator of $\psi^{(i)}$,

$$f(x) = \sum_j x_j^\alpha,$$

and find its extrema on $[0,1]^d$, subject to the constraint

$$g(x) = \sum_j x_j - 1 = 0.$$

Then the Lagrangian is

$$\mathscr{L}(x, \lambda) = f(x) - \lambda g(x) = \sum_j (x_j^\alpha - \lambda x_j)$$

and

$$\nabla \mathscr{L} = \begin{pmatrix} \alpha x_1^{\alpha-1} - \lambda \\ \vdots \\ \alpha x_d^{\alpha-1} - \lambda \\ 1 - \sum_j x_j \end{pmatrix}.$$

On $[0,1]^d$, the equation $\nabla \mathscr{L} = 0$ yields the solution

$$\begin{pmatrix} x_1 \\ \vdots \\ x_d \\ \lambda \end{pmatrix} = \begin{pmatrix} \frac{1}{d} \\ \vdots \\ \frac{1}{d} \\ \frac{\alpha}{d^{\alpha-1}} \end{pmatrix}.$$

Note that f is continuous on the compact Δ^{d-1}, and therefore it will have two extrema. Since the Lagrange multipliers provide only one solution, the second extremal value must be on the boundary. But $\alpha > 1$ and $0 \leq x_i \leq 1$, so it is obvious that it will be the maximum value 1 to be on the boundary, at each vertex of Δ^{d-1}. Then it is possible to conclude that what has been found earlier are the coordinates of the minimum, which yield the value

$$f\left(\frac{\mathbf{1}}{\mathbf{d}}\right) = \frac{1}{d^{\alpha-1}}.$$

Thus

$$\frac{1}{d^{\alpha-1}} \leq f(x) \leq 1.$$

In conclusion

$$\psi^{(i)}(x) = \frac{x_i^\alpha}{f(x)} \geq x_i^\alpha$$

and

$$\psi^{(i)}(x) = \frac{x_i^\alpha}{f(x)} \leq d^{\alpha-1} x_i^\alpha.$$

□

The second fact is about the regularity of ψ. Explicit calculation of the partial derivatives shows that the restriction of ψ on $\mathring{\Delta}^{d-1}$ is smooth and, since $\alpha > 1$, it is continuously differentiable at $\partial\Delta^{d-1}$, but not necessarily twice differentiable. The partial derivatives are indeed continuous on the whole Δ^{d-1}, and no issues arise in further differentiating as no component vanishes there:

$$\partial_{x_j} \psi^{(i)}(x) = \begin{cases} \alpha x_i^{\alpha-1} \frac{\sum_{k\neq i} x_k^\alpha}{\left(\sum_k x_k^\alpha\right)^2}, & \text{if } i = j \\ -\alpha x_i^\alpha \frac{x_j^{\alpha-1}}{\left(\sum_k x_k^\alpha\right)^2}, & \text{if } i \neq j. \end{cases} \tag{8.2}$$

However for some values of α the off-diagonal derivatives show irregularities, already as of the second derivative. For example for $\alpha = 5/3$, at

$$E_{d-2} := \left(\frac{1}{d-1}, \ldots, \frac{1}{d-1}, 0\right) \in \partial\Delta^{d-1},$$

the derivative $\partial^2_{x_d}\psi^{(1)}(E_{d-2})$ does not exist. Indeed, since $\partial_{x_d}\psi^{(1)}(E_{d-2}) = 0$, the limit defining $\partial^2_{x_d}\psi^{(1)}(E_{d-2})$ simplifies to

$$-\alpha \lim_{h\to 0^\pm} \frac{\frac{1}{(d-1)^\alpha} \frac{h^{\alpha-1}}{\left(\sum_{k=1}^{d-1} \frac{1}{(d-1)^\alpha} + h^\alpha\right)^2}}{h} = -\frac{5}{3(d-1)^{\frac{5}{3}}} \lim_{h\to 0^\pm} \frac{h^{-\frac{1}{3}}}{\left(\frac{1}{(d-1)^{\frac{2}{3}}} + h^{\frac{5}{3}}\right)^2} = \mp\infty.$$

As a result, at the boundary, we can only rely on continuous differentiability of ψ in general, not twice differentiability.

We now briefly introduce the main result of this chapter, Proposition 8.4, the main tool which, by iteration, leads to proving almost sure dominance with feedback in the following chapter, by Theorem 1.11, whose argument relies on Θ_n almost never eventually being confined to a small δ_n-neighbourhood of the equilibrium E of the simplex (or centre, geometrically speaking) and of the partial equilibrium points in its lower-dimensional faces, and this fact is ensured by Proposition 8.4. To get a mental picture, for example, in Δ^2 we have the equilibrium $\mathbf{\frac{1}{3}}$ to look for, and then in the 1-faces (its edges in this case) there are the partial equilibria $(1/2, 1/2, 0)$, $(1/2, 0, 1/2)$ and $(0, 1/2, 1/2)$. A generic vanishing positive valued δ_n such that

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_{n+1}} \delta_n < \infty,$$

as adopted in [48, Proposition 4.1] will not suffice for any $d > 2$, because the argument around partial equilibria is far more problematic: the case $d = 2$ is the only one that does not have partial equilibria, and in this sense it is fortunate. The general meaning of the lemma that will follow is to ensure that it is possible to choose a more specific

$$\delta_n := \frac{1}{\tau_n^r} \tag{8.3}$$

for some $0 < r < 1/2$ instead, so that the argument of Proposition 8.4 succeeds for all $d > 2$ and all regimes of growth satisfying Assumptions S and R (the second of these conditions will play a significant role also in Theorem 1.11). The introduction of the parameter r is related to both Proposition 8.4 and Theorem 1.11; in Proposition 8.4 it will be further restricted to a range of values so that the argument makes it through all regimes of growth.

Lemma 8.2. *Let* $0 < r < 1/2$*, then*

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_{n+1}\tau_n^r} < \infty.$$

Proof. The two cases given by Assumption R need to be treated separately.

- If ρ_n is bounded by some $\rho > 0$, then noting that

$$\frac{\sigma_{n+1}}{\tau_{n+1}\tau_n^r} = \left(\frac{\sigma_{n+1}}{\tau_{n+1}^{1+r}}\right)\left(\frac{\tau_{n+1}^{1+r}}{\tau_n^{1+r}}\right) = \frac{\sigma_{n+1}}{\tau_{n+1}^{1+r}}\left(\frac{\tau_n + \sigma_{n+1}}{\tau_n}\right)^{1+r} = \frac{\sigma_{n+1}}{\tau_{n+1}^{1+r}}(1+\rho_n)^{1+r},$$

  we obtain

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_{n+1}\tau_n^r} < \sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_{n+1}^{1+r}}(1+\rho_n)^{1+r} \le (1+\rho)^{1+r} \sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_{n+1}^{1+r}} \le (1+\rho)^{1+r} \int_{\tau_0}^{\infty} \frac{dx}{x^{1+r}}$$
$$= \frac{(1+\rho)^{1+r}}{r\tau_0^r} < \infty.$$

- If $\rho_n \longrightarrow \infty$, note that for any constant $b > 1$, there will be an $\bar{m} > 0$, such that for all $n > \bar{m}$, $\rho_n > b$. Then it follows that $\sigma_{n+1} > b\tau_n$ for all $n > \bar{m}$, and therefore $\tau_n > \sigma_n > b\tau_{n-1}$, which yields, by induction,

$$\tau_n > b^{n-\bar{m}}\tau_{\bar{m}}.$$

  Consequently, the sum can be estimated as follows. Since $b^r > 1$,

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_{n+1}\tau_n^r} < \sum_{n=0}^{\infty} \frac{1}{\tau_n^r} \le \sum_{n=0}^{\infty} \frac{1}{(b^{n-\bar{m}}\tau_{\bar{m}})^r} < \frac{b^{\bar{m}}}{\tau_{\bar{m}}^r} \sum_{n=0}^{\infty} \frac{1}{b^{rn}} = \frac{b^{\bar{m}}}{\tau_{\bar{m}}^r} \frac{1}{1-\frac{1}{b^r}} < \infty.$$

□

We conclude this section by briefly showing an elementary fact, which will be used in Proposition 8.4.

Remark 8.3. *For any $\gamma > 1$, uniformly in $n > n_m > m$ and assuming that $m = o(n_m)$ as $m \longrightarrow \infty$,*

$$\sum_{k=m+1}^{n} \frac{1}{k^{\gamma+o(1)}} \sim \frac{1}{(\gamma-1)m^{\gamma-1+o(1)}}$$

Proof. The claim follows from integral estimates. Note that $\gamma > 1$ implies that $\gamma + o(1) > 1$, for all m large enough. Hence on one hand, abusing a little the notation, as $m = o(n)$,

$$\sum_{k=m+1}^{n} \frac{1}{k^{\gamma+o(1)}} \geq \int_{m+1}^{n} \frac{dx}{x^{\gamma+o(1)}} =$$
$$\frac{1}{(\gamma-1)m^{\gamma-1+o(1)}}\left((1+o(1))\left(\frac{m}{m+1}\right)^{\gamma-1+o(1)} - (1+o(1))\left(\frac{m}{n}\right)^{\gamma-1+o(1)}\right) =$$
$$\frac{1}{(\gamma-1)m^{\gamma-1+o(1)}}(1+o(1)),$$

on the other hand

$$\sum_{k=m+1}^{n} \frac{1}{k^{\gamma+o(1)}} \leq \frac{1}{(m+1)^{\gamma+o(1)}} + \int_{m+2}^{n+1} \frac{dx}{(x-1)^{\gamma+o(1)}} = \frac{1}{(m+1)^{\gamma+o(1)}} +$$
$$\frac{1}{(\gamma-1)m^{\gamma-1+o(1)}}(1+o(1)) = \frac{1}{(\gamma-1)(m)^{\gamma-1+o(1)}}\left(\frac{(\gamma-1)m^{\gamma-1+o(1)}}{(m+1)^{\gamma+o(1)}} + 1 + o(1)\right)$$
$$= \frac{1}{(\gamma-1)(m)^{\gamma-1+o(1)}}\left(1+o(1)\right),$$

and therefore

$$\sum_{k=m+1}^{n} \frac{1}{k^{\gamma+o(1)}} \sim \frac{1}{(\gamma-1)(m)^{\gamma-1+o(1)}}.$$

□

8.2 Infinitesimal deviations from the equilibria in presence of feedback

This section is dedicated to showing Proposition 8.4: that is, we show that Θ_n is almost never eventually confined in a δ_n-neighbourhood of the equilibrium point of the l-faces of Δ^{d-1}, where $l \in [d-1]$. By symmetry, without loss of generality, the equilibria we will choose to illustrate the argument in Theorem 1.11, are those belonging to the face obtained by intersecting the hyperplanes $H_{l+1} := \{x_d = \ldots = x_{l+2} = 0\}$ with the simplex: $\Delta^{d-1} \cap H_{l+1}$. Therefore, in Proposition 8.4 a generic partial equilibrium of an l-face will be

$$E_l := \Bigg(\underbrace{\frac{1}{l+1}, \ldots, \frac{1}{l+1}}_{l+1}, \underbrace{0, \ldots, 0}_{d-l-1}\Bigg).$$

The full equilibrium point of Δ^{d-1} is denoted as $E := E_{d-1} = \frac{\mathbf{1}}{\mathbf{d}}$.

Proposition 8.4. *Let $\alpha > 1$,*

$$r_\alpha := \max\left\{\frac{1}{4}, \frac{1}{2\alpha}\right\}$$

and

$$\delta_n := \frac{1}{\tau_n^r},$$

where

$$r_\alpha < r < \frac{1}{2}.$$

Then for every $l \in [d-1]$,

$$\mathbb{P}(\|\Theta_n - E_l\| > \delta_n,\ i.o.) = 1.$$

Proof.

Step 1. Since

$$\{\|\Theta_n - E_l\| > \delta_n,\ \text{i.o.}\}^c = \{\|\Theta_n - E_l\| \leq \delta_n,\ \text{ev.}\},$$

we can equivalently show that

$$\mathbb{P}(\|\Theta_n - E_l\| \leq \delta_n,\ \text{ev.}) = 0,$$

but since

$$\{\|\Theta_n - E_l\| \leq \delta_n,\ \text{ev.}\} = \bigcup_{m=1}^{\infty} \bigcap_{n\geq m} \{\|\Theta_n - E_l\| \leq \delta_n\}$$

and

$$\bigcap_{n\geq m} \{\|\Theta_n - E_l\| \leq \delta_n\} \subseteq \bigcap_{n\geq m+1} \{\|\Theta_n - E_l\| \leq \delta_n\},$$

by monotonicity

$$\bigcup_{m=1}^{\infty} \bigcap_{n\geq m} \{\|\Theta_n - E_l\| \leq \delta_n\} = \lim_{m\to\infty} \bigcap_{n\geq m} \{\|\Theta_n - E_l\| \leq \delta_n\}.$$

Hence by the monotone continuity of the probability measure, we have that

$$\mathbb{P}\left(\lim_{m\to\infty} \bigcap_{n\geq m} \{\|\Theta_n - E_l\| \leq \delta_n\}\right) = \lim_{m\to\infty} \mathbb{P}\left(\|\Theta_n - E_l\| \leq \delta_n\ \forall n \geq m\right).$$

All in all, defining

$$\mathcal{H}_m^l := \{\|\Theta_n - E_l\| \leq \delta_n,\ \forall n \geq m\},$$

we will prove the claim by showing that as $m \longrightarrow \infty$,

$$\mathbb{P}(\mathcal{H}_m^l) \to 0.$$

Step 2. We prove that this probability vanishes by rewriting $\Theta_n^{(1)} - 1/(l+1)$ on the event $\mathcal{H}_m$, by iterating (5.2) and (5.3), with the help of the multivariate *Mean Value Theorem* applied to $P_n^{(1)} := \psi^{(1)}(\Theta_{n-1})$, and Taylor approximations at E_l. We will then achieve an iterative multiplicative representation of $\Theta_n^{(1)} - 1/(l+1)$ in terms of $\Theta_m^{(1)} - 1/(l+1)$, which by a suitable probabilistic argument, will be crucial in the estimation of $\mathbb{P}(\mathcal{H}^l_m)$. Before proceeding, we would like to mention that a vector-valued approach is in theory also possible, and it might even look like a more natural way of generalising the argument in [48, Proposition 4.1]: it would require exploiting a vector-valued generalization of the multivariate mean value theorem, using matrices in the iteration suitably diagonalised, in order to highlight the deterministic scales involved. However it turns out far more problematic. Even though such approach succeeds for $l = d-1$ with exponential growth, uniformity of the results with respect to all parameters involved fails when $l < d-1$ and already with the exponential regime.

Consider now the vector field $\psi : \Delta^{d-1} \longrightarrow \Delta^{d-1}$ as defined at the beginning of this section through its components $\psi^{(i)}(x)$. For $i = 1$, we can apply the standard *multivariate Mean Value Theorem*. For every $\Theta_{n-1} \in \Delta^{d-1}$, there exists a point

$$\xi_{\Theta_{n-1}} = t\Theta_{n-1} + (1-t)E_l,$$

corresponding to some parameter $t \in (0,1)$, such that

$$\psi^{(1)}(\Theta_{n-1}) - \psi^{(1)}(E_l) = \nabla\psi^{(1)}(\xi_{\Theta_{n-1}}) \cdot (\Theta_{n-1} - E_l),$$

where $\cdot$ denotes the scalar product. Since by definition $P^{(1)}_{n-1} = \psi^{(1)}(\Theta_{n-1})$ and $\psi^{(1)}(E_l) = 1/(l+1)$, the equation can be rewritten as

$$P^{(1)}_{n-1} = \frac{1}{l+1} + \nabla\psi^{(1)}(\xi_{\Theta_{n-1}}) \cdot (\Theta_{n-1} - E_l). \tag{8.4}$$

Plugging (8.4) into (5.3) with $i = 1$, after subtracting $1/(l+1)$ both sides, yields that uniformly in $n > m$ and $\omega \in \Omega$,

$$\Theta^{(1)}_n - \frac{1}{l+1} = \frac{\tau_{n-1}}{\tau_n}\Theta^{(1)}_{n-1} + \frac{\sigma_n}{\tau_n}P^{(1)}_{n-1} - \frac{1}{l+1} + \frac{\varepsilon^{(1)}_n}{\tau_n}\sqrt{\sigma_n P^{(1)}_{n-1}(1-P^{(1)}_{n-1})} =$$
$$\frac{\tau_{n-1}}{\tau_n}\Theta^{(1)}_{n-1} + \frac{\sigma_n}{\tau_n}\left(\frac{1}{l+1} + \nabla\psi^{(1)}(\xi_{\Theta_{n-1}}) \cdot (\Theta_{n-1} - E_l)\right) - \frac{1}{l+1} +$$
$$\frac{\varepsilon^{(1)}_n}{\tau_n}\sqrt{\sigma_n P^{(1)}_{n-1}(1-P^{(1)}_{n-1})} = \frac{\tau_{n-1}}{\tau_n}\Theta^{(1)}_{n-1} + \frac{\sigma_n}{(l+1)\tau_n} + \frac{\sigma_n}{\tau_n}\sum_{j=1}^{l+1}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\left(\Theta^{(j)}_{n-1} - \frac{1}{l+1}\right)$$
$$+ \frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\Theta^{(j)}_{n-1} - \frac{1}{l+1} + \frac{\varepsilon^{(1)}_n}{\tau_n}\sqrt{\sigma_n P^{(1)}_{n-1}(1-P^{(1)}_{n-1})} =$$
$$\frac{\tau_{n-1}}{\tau_n}\Theta^{(1)}_{n-1} - \frac{\tau_{n-1}}{(l+1)\tau_n} + \frac{\sigma_n}{\tau_n}\sum_{j=1}^{l+1}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\left(\Theta^{(j)}_{n-1} - \frac{1}{l+1}\right) + \frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\Theta^{(j)}_{n-1}$$
$$+ \frac{\varepsilon^{(1)}_n}{\tau_n}\sqrt{\sigma_n P^{(1)}_{n-1}(1-P^{(1)}_{n-1})},$$

so

$$\Theta^{(1)}_n - \frac{1}{l+1} =$$
$$\frac{\tau_{n-1}}{\tau_n}\left(\Theta^{(1)}_{n-1} - \frac{1}{l+1}\right) + \frac{\sigma_n}{\tau_n}\sum_{j=1}^{l+1}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\left(\Theta^{(j)}_{n-1} - \frac{1}{l+1}\right) + \frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\Theta^{(j)}_{n-1}$$
$$+ \frac{\varepsilon^{(1)}_n}{\tau_n}\sqrt{\sigma_n P^{(1)}_{n-1}(1-P^{(1)}_{n-1})}. \tag{8.5}$$

Since on the event $\mathcal{H}_m^l$ it holds that $\|\Theta_n^{(1)} - 1/(l+1)\| \leq \delta_n$ for all $n \geq m$, and $\delta_n \to 0$, some of the terms can be approximated via first degree Taylor expansion of $\psi^{(1)}(x)$ and $\psi_{x_j}^{(1)}(x)$ at E_l, as $n \longrightarrow \infty$, uniformly on $\mathcal{H}_m^l$. For the moment, the place holder notation for Θ_{n-1} will be x. The dependence on ξ_x is in a certain sense negligible on $\mathcal{H}_m^l$ as $m \to \infty$: since ξ_x is on a segment joining x and E_l, and since $\|\xi_x - E_l\| \to 0$ on $\mathcal{H}_m^l$, as $m \to \infty$, we have that $\xi_x \approx E_l$, with error deterministically bounded by δ_n. More formally, since $\psi^{(1)}$ is continuously differentiable at E_l for all l and on the simplex, the constant Taylor approximation will have a uniform error estimate on $\mathcal{H}_m^l$. Specifically, by the remainder theorem (in Lagrange form), since all partial derivatives (which we will compute explicitly) are continuous over a compact set, they all achieve a maximum, and therefore there will be a constant bounding $\|\nabla\psi^{(1)}(\zeta_x)\|$ over the simplex (ζ_x denotes a point on the segment joining E_l and x, so it is in the simplex and the previous bound applies). Then, by the *Cauchy-Schwartz inequality*,

$$|\nabla\psi^{(1)}(\zeta_x) \cdot (x - E_l)| \leq \|\nabla\psi^{(1)}(\zeta_x)\|\|x - E_l\| = \mathcal{O}\left(\|x - E_l\|\right),$$

so

$$\psi^{(1)}(x) = \psi^{(1)}(E_l) + \mathcal{O}\left(\|x - E_l\|\right) = \frac{1}{l+1} + \mathcal{O}\left(\delta_{n-1}\right)$$

and $P_{n-1}^{(1)} = 1/(l+1) + \mathcal{O}\left(\delta_{n-1}\right)$. Therefore as $n \longrightarrow \infty$, uniformly on $\omega \in \mathcal{H}_m^l$

$$\sqrt{P_{n-1}^{(1)}(1 - P_{n-1}^{(1)})} = \sqrt{\left(\frac{1}{l+1} + \mathcal{O}\left(\delta_{n-1}\right)\right)\left(\frac{l}{l+1} + \mathcal{O}\left(\delta_{n-1}\right)\right)} =$$
$$\sqrt{\frac{l}{(l+1)^2} + \mathcal{O}\left(\delta_{n-1}\right)} = \frac{\sqrt{l + \mathcal{O}\left(\delta_{n-1}\right)}}{l+1} = \frac{\sqrt{l + o(1)}}{l+1}$$

Since ψ may not be twice differentiable at points in the boundary like E_l when $l < d-1$, its linear approximation will be expressed without uniform estimate of the error, that is with Peano form of the remainder:

$$\psi^{(1)}(x) = \psi^{(1)}(E_l) + \nabla\psi^{(1)}(E_l) \cdot (x - E_l) + o\left(\|x - E_l\|\right).$$

While this method does not interfere with approximating $P_{n-1}^{(1)}$, and yields a uniform estimate of the error, it prevents us from approximating some of the partial derivatives $\psi_{x_j}^{(1)}(x)$ for $j > l+1$ at E_l, as $m \to \infty$, with a uniform bound on the error. Thus via a different argument, we will show that continuity of those partial derivatives will suffice in achieving a weaker (yet strong enough) uniform bound on the error, which will play an important role, especially when $1 < \alpha < 2$. Consider that by direct calculation and continuity, it holds that

$$\lim_{m\to\infty} \psi_{x_j}^{(i)}(\xi_x) = \psi_{x_j}^{(i)}(E_l) = \begin{cases} \alpha(E_l^{(i)})^{\alpha-1} \frac{\sum_{k\neq i}(E_l^{(k)})^\alpha}{\left(\sum_k (E_l^{(k)})^\alpha\right)^2} = \frac{l\alpha}{l+1}, & \text{if } i = j \leq l+1 \\ \alpha(E_l^{(i)})^{\alpha-1} \frac{\sum_{k\neq i}(E_l^{(k)})^\alpha}{\left(\sum_k (E_l^{(k)})^\alpha\right)^2} = 0, & \text{if } i = j > l+1 \\ -\alpha(E_l^{(i)})^{\alpha} \frac{(E_l^{(j)})^{\alpha-1}}{\left(\sum_k (E_l^{(k)})^\alpha\right)^2} = -\frac{\alpha}{l+1}, & \text{if } i \neq j,\ i,j \leq l+1 \\ -\alpha(E_l^{(i)})^{\alpha} \frac{(E_l^{(j)})^{\alpha-1}}{\left(\sum_k (E_l^{(k)})^\alpha\right)^2} = 0, & \text{if } i \neq j,\ i \text{ or } j > l+1, \end{cases}$$

so, specifically for $i = 1$,

$$\lim_{m\to\infty}\psi^{(1)}_{x_j}(\xi_x) = \psi^{(1)}_{x_j}(E_l) = \begin{cases} \alpha(E_l^{(1)})^{\alpha-1}\frac{\sum_{k\neq 1}(E_l^{(k)})^\alpha}{\left(\sum_k (E_l^{(k)})^\alpha\right)^2} = \frac{l\alpha}{l+1}, & \text{if } 1 = j \\ -\alpha(E_l^{(1)})^{\alpha}\frac{(E_l^{(j)})^{\alpha-1}}{\left(\sum_k (E_l^{(k)})^\alpha\right)^2} = -\frac{\alpha}{l+1}, & \text{if } 1 \neq j \leq l+1 \\ -\alpha(E_l^{(1)})^{\alpha}\frac{(E_l^{(j)})^{\alpha-1}}{\left(\sum_k (E_l^{(k)})^\alpha\right)^2} = 0, & \text{if } 1 \neq j > l+1. \end{cases}$$

Therefore, as $m \to \infty$,

$$\psi^{(1)}_{x_j}(\xi_x) = \begin{cases} \frac{l\alpha}{l+1} + o(1), & \text{if } 1 = j \\ -\frac{\alpha}{l+1} + o(1), & \text{if } 1 \neq j \leq l+1 \\ o(1), & \text{if } 1 \neq j > l+1. \end{cases}$$

Recall that the placeholder x stands for Θ_{n-1}, and recall that since

$$\|\xi_{\Theta_{n-1}} - E_l\| \leq \|\Theta_{n-1} - E_l\| \leq \delta_{n-1}$$

for all $n \geq m$ on $\mathcal{H}^l_m$, we have that $\xi_x \longrightarrow E_l$. For some of the derivatives, an estimate of the error linear in δ_{n-1} can be derived through Taylor expansion at E_l, uniformly on $\mathcal{H}^l_m$. Namely, when $j \leq l+1$, the continuous differentiability at E_l of $\psi^{(1)}_{x_j}(\xi_x)$ ensures that by first degree Taylor expansion $\psi^{(1)}_{x_j}(\xi_x) = \psi^{(1)}_{x_j}(E_l) + \nabla\psi^{(1)}_{x_j}(\zeta_x)\cdot(\xi_x - E_l)$ (with ζ_x in the segment joining E_l and ξ_x) and therefore

$$\psi^{(1)}_{x_j}(\xi_x) = \psi^{(1)}_{x_j}(E_l) + \mathcal{O}\left(\|\xi_x - E_l\|\right) = \begin{cases} \frac{l\alpha}{l+1} + \mathcal{O}(\delta_{n-1}), & \text{if } 1 = j \\ -\frac{\alpha}{l+1} + \mathcal{O}(\delta_{n-1}), & \text{if } 1 \neq j \leq l+1. \end{cases}$$

We can easily show, by direct calculation of the derivatives, that all components of $\nabla\psi^{(1)}_{x_j}$ are continuous at E_l when $j \leq l+1$, and therefore there is a compact set on which all partial derivatives are continuous, and to which the segment joining E_l and ξ_x belongs (which is the requirement for the uniform estimate of the remainder):

$$\partial_{x_s}\psi^{(i)}_{x_j}(x) = \begin{cases} \frac{\alpha x_i^{\alpha-2}\sum_{k\neq i} x_k^\alpha\left((\alpha-1)\sum_k x_k^\alpha - 2\alpha x_i^\alpha\right)}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } i = j = s \\ \frac{\alpha^2 x_i^{\alpha-1}x_s^{\alpha-1}\left(\sum_k x_k^\alpha - 2\sum_{k\neq i} x_k^\alpha\right)}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } i = j \neq s \\ \frac{\alpha^2 x_i^{\alpha-1}x_j^{\alpha-1}\left(2x_i^\alpha - \sum_k x_k^\alpha\right)}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } s = i \neq j \\ \frac{\alpha x_i^{\alpha}x_j^{\alpha-2}\left(2\alpha x_j^\alpha - (\alpha-1)\sum_k x_k^\alpha\right)}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } i \neq j = s \\ \frac{2\alpha^2 x_i^{\alpha}x_j^{\alpha-1}x_s^{\alpha-1}}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } s \neq i \neq j \neq s, \end{cases}$$

so specifically

$$\partial_{x_s}\psi^{(1)}_{x_j}(x) = \begin{cases} \frac{\alpha x_1^{\alpha-2}\sum_{k\neq 1} x_k^\alpha\left((\alpha-1)\sum_k x_k^\alpha - 2\alpha x_1^\alpha\right)}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } 1 = j = s \\ \frac{\alpha^2 x_1^{\alpha-1}x_s^{\alpha-1}\left(\sum_k x_k^\alpha - 2\sum_{k\neq i} x_k^\alpha\right)}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } 1 = j \neq s \\ \frac{\alpha^2 x_1^{\alpha-1}x_j^{\alpha-1}\left(2x_1^\alpha - \sum_k x_k^\alpha\right)}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } s = 1 \neq j \\ \frac{\alpha x_1^{\alpha}x_j^{\alpha-2}\left(2\alpha x_j^\alpha - (\alpha-1)\sum_k x_k^\alpha\right)}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } 1 \neq j = s \\ \frac{2\alpha^2 x_1^{\alpha}x_j^{\alpha-1}x_s^{\alpha-1}}{\left(\sum_k x_k^\alpha\right)^3}, & \text{if } s \neq 1 \neq j \neq s. \end{cases}$$

However, when $j > l+1$, a weaker nonlinear estimate of the error in terms of δ_{n-1} uniform on $\mathcal{H}_m^l$ can be derived, recalling the main steps of the proof of (8.1) regarding the denominator of $\psi^{(1)}$. Indeed, since

$$\psi^{(1)}_{x_j}(\xi_x) = -\alpha(\xi_x^{(1)})^\alpha \frac{(\xi_x^{(j)})^{\alpha-1}}{\left(\sum_k (\xi_x^{(k)})^\alpha\right)^2},$$

since $E_l^{(k)} = 0$ for all $k > l+1$, and since ξ_x has all coordinates nonnegative and subunitary, as it belongs to the simplex, it follows that

$$|\psi^{(1)}_{x_j}(\xi_x)| = \alpha(\xi_x^{(1)})^\alpha \frac{(\xi_x^{(j)})^{\alpha-1}}{\left(\sum_k (\xi_x^{(k)})^\alpha\right)^2} \leq \alpha d^{2(\alpha-1)}(\xi_x^{(1)})^\alpha(\xi_x^{(j)})^{\alpha-1} \leq \alpha d^{2(\alpha-1)}|\xi_x^{(j)} - E_l^{(j)}|^{\alpha-1}$$
$$\leq \alpha d^{2(\alpha-1)}\,\|\xi_x - E_l\|^{\alpha-1} = \mathcal{O}(\delta_{n-1}^{\alpha-1}).$$

This plays a significant role only when $1 < \alpha < 2$, as in all other cases the uniform bound (at least linear) would be automatically recovered. Since for $\alpha \geq 2$ no such issue arises, and the above calculation shows that all, which follows, can be easily adjusted just by replacing the $\mathcal{O}(\delta_j^{\alpha-1})$ terms with $\mathcal{O}(\delta_j)$, the discussion will be mainly, and implicitly, concerned with $1 < \alpha < 2$, without loss of generality. To sum up, going back to the Θ_{n-1} notation, uniformly in $\omega \in \mathcal{H}_m^l$ we have that

$$\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}}) = \begin{cases} \frac{l\alpha}{l+1} + \mathcal{O}(\delta_{n-1}), & \text{if } 1 = j \\ -\frac{\alpha}{l+1} + \mathcal{O}(\delta_{n-1}), & \text{if } 1 \neq j \leq l+1 \\ \mathcal{O}(\delta_{n-1}^{\alpha-1}), & \text{if } 1 \neq j > l+1. \end{cases} \tag{8.6}$$

This result will be plugged in (8.5), which we first rewrite as:

$$\Theta_n^{(1)} - \frac{1}{l+1} = \frac{\tau_{n-1}}{\tau_n}\left(\Theta_{n-1}^{(1)} - \frac{1}{l+1}\right) + \frac{\sigma_n}{\tau_n}\psi^{(1)}_{x_1}(\xi_{\Theta_{n-1}})\left(\Theta_{n-1}^{(1)} - \frac{1}{l+1}\right)$$
$$+ \frac{\sigma_n}{\tau_n}\sum_{j=2}^{l+1}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\left(\Theta_{n-1}^{(j)} - \frac{1}{l+1}\right) + \frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\Theta_{n-1}^{(j)}$$
$$+ \frac{\varepsilon_n^{(1)}}{\tau_n}\sqrt{\sigma_n P_{n-1}^{(1)}(1-P_{n-1}^{(1)})} = \left(\frac{\tau_{n-1}}{\tau_n} + \frac{\sigma_n}{\tau_n}\psi^{(1)}_{x_1}(\xi_{\Theta_{n-1}})\right)\left(\Theta_{n-1}^{(1)} - \frac{1}{l+1}\right)$$
$$+ \frac{\sigma_n}{\tau_n}\sum_{j=2}^{l+1}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\left(\Theta_{n-1}^{(j)} - \frac{1}{l+1}\right) + \frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\Theta_{n-1}^{(j)}$$
$$+ \frac{\varepsilon_n^{(1)}}{\tau_n}\sqrt{\sigma_n P_{n-1}^{(1)}(1-P_{n-1}^{(1)})}.$$

We now apply the Taylor remainder theorem in Lagrange form, so as to express the partial derivatives, separate the main part from the negligible one, rearrange the main part suitably and proceed with the iteration. Uniformly on the probability space and $n > m$,

$$\Theta_n^{(1)} - \frac{1}{l+1} = \left(\frac{\tau_{n-1}}{\tau_n} + \frac{\sigma_n}{\tau_n}\left(\frac{l\alpha}{l+1} + \nabla\psi^{(1)}_{x_1}(\zeta^1_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}} - E_l)\right)\right)\left(\Theta_{n-1}^{(1)} - \frac{1}{l+1}\right)$$
$$- \frac{\sigma_n}{\tau_n}\sum_{j=2}^{l+1}\left(\frac{\alpha}{l+1} - \nabla\psi^{(1)}_{x_j}(\zeta^j_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}} - E_l)\right)\left(\Theta_{n-1}^{(j)} - \frac{1}{l+1}\right) +$$
$$\frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\Theta_{n-1}^{(j)} + \frac{\sqrt{\sigma_n}}{\tau_n}\varepsilon_n^{(1)}\sqrt{P_{n-1}^{(1)}(1-P_{n-1}^{(1)})}. \tag{8.7}$$

Since

$$\sum_{j=2}^{l+1}\left(\frac{\alpha}{l+1}-\nabla\psi^{(1)}_{x_j}(\zeta^j_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}}-E_l)\right)\left(\Theta^{(j)}_{n-1}-\frac{1}{l+1}\right)=$$
$$\frac{\alpha}{l+1}\sum_{j=2}^{l+1}\left(\Theta^{(j)}_{n-1}-\frac{1}{l+1}\right)-\sum_{j=2}^{l+1}\nabla\psi^{(1)}_{x_j}(\zeta^j_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}}-E_l)\left(\Theta^{(j)}_{n-1}-\frac{1}{l+1}\right)$$

and

$$\frac{\alpha}{l+1}\sum_{j=2}^{l+1}\left(\Theta^{(j)}_{n-1}-\frac{1}{l+1}\right)=\frac{\alpha}{l+1}\left(1-\Theta^{(1)}_{n-1}-\sum_{j=l+2}^{d}\Theta^{(j)}_{n-1}-\frac{l}{l+1}\right)=$$
$$-\frac{\alpha}{l+1}\left(\sum_{j=l+2}^{d}\Theta^{(j)}_{n-1}\right)-\frac{\alpha}{l+1}\left(\Theta^{(1)}_{n-1}-\frac{1}{l+1}\right),$$

from (8.7) it follows that

$$\Theta^{(1)}_{n}-\frac{1}{l+1}=\left(\frac{\tau_{n-1}}{\tau_n}+\frac{\sigma_n}{\tau_n}\frac{l\alpha}{l+1}+\frac{\sigma_n}{\tau_n}\nabla\psi^{(1)}_{x_1}(\zeta^1_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}}-E_l)\right)\left(\Theta^{(1)}_{n-1}-\frac{1}{l+1}\right)$$
$$-\frac{\sigma_n}{\tau_n}\frac{\alpha}{l+1}\sum_{j=2}^{l+1}\left(\Theta^{(j)}_{n-1}-\frac{1}{l+1}\right)+\frac{\sigma_n}{\tau_n}\sum_{j=2}^{l+1}\nabla\psi^{(1)}_{x_j}(\zeta^j_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}}-E_l)\left(\Theta^{(j)}_{n-1}-\frac{1}{l+1}\right)$$
$$+\frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\Theta^{(j)}_{n-1}+\frac{\sqrt{\sigma_n}}{\tau_n}\varepsilon^{(1)}_n\sqrt{P^{(1)}_{n-1}(1-P^{(1)}_{n-1})}=\frac{\sigma_n}{\tau_n}\frac{\alpha}{l+1}\left(\sum_{j=l+2}^{d}\Theta^{(j)}_{n-1}\right)$$
$$+\left(\frac{\tau_{n-1}}{\tau_n}+\frac{\sigma_n}{\tau_n}\frac{l\alpha}{l+1}+\frac{\sigma_n}{\tau_n}\nabla\psi^{(1)}_{x_1}(\zeta^1_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}}-E_l)\right)\left(\Theta^{(1)}_{n-1}-\frac{1}{l+1}\right)$$
$$+\frac{\sigma_n}{\tau_n}\frac{\alpha}{l+1}\left(\Theta^{(1)}_{n-1}-\frac{1}{l+1}\right)+\frac{\sigma_n}{\tau_n}\sum_{j=2}^{l+1}\nabla\psi^{(1)}_{x_j}(\zeta^j_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}}-E_l)\left(\Theta^{(j)}_{n-1}-\frac{1}{l+1}\right)$$
$$+\frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\Theta^{(j)}_{n-1}+\frac{\sqrt{\sigma_n}}{\tau_n}\varepsilon^{(1)}_n\sqrt{P^{(1)}_{n-1}(1-P^{(1)}_{n-1})}$$
$$=\left(\frac{\tau_{n-1}+\alpha\sigma_n}{\tau_n}+\frac{\sigma_n}{\tau_n}\nabla\psi^{(1)}_{x_1}(\zeta^1_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}}-E_l)\right)\left(\Theta^{(1)}_{n-1}-\frac{1}{l+1}\right)$$
$$+\frac{\sigma_n}{\tau_n}\sum_{j=2}^{l+1}\nabla\psi^{(1)}_{x_j}(\zeta^j_{\Theta_{n-1}})\cdot(\xi_{\Theta_{n-1}}-E_l)\left(\Theta^{(j)}_{n-1}-\frac{1}{l+1}\right)+\frac{\alpha}{l+1}\frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\Theta^{(j)}_{n-1}$$
$$+\frac{\sigma_n}{\tau_n}\sum_{j=l+2}^{d}\psi^{(1)}_{x_j}(\xi_{\Theta_{n-1}})\Theta^{(j)}_{n-1}+\frac{\sqrt{\sigma_n}}{\tau_n}\varepsilon^{(1)}_n\sqrt{P^{(1)}_{n-1}(1-P^{(1)}_{n-1})}.$$

Define

$$k_j(\Theta_j):=\frac{\tau_j+\alpha\sigma_{j+1}}{\tau_{j+1}}+\frac{\sigma_{j+1}}{\tau_{j+1}}\nabla\psi^{(1)}_{x_1}(\zeta^1_{\Theta_j})\cdot(\xi_{\Theta_j}-E_l)=$$
$$\frac{\tau_j+\alpha\sigma_{j+1}}{\tau_{j+1}}\left(1+\frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\nabla\psi^{(1)}_{x_1}(\zeta^1_{\Theta_j})\cdot(\xi_{\Theta_j}-E_l)\right)=$$
$$\left(\frac{\tau_j+\alpha\sigma_{j+1}}{\tau_{j+1}}\right)\left(1+\frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\mathcal{O}(\delta_j)\right),$$

for $i \in \{2, \dots, l+1\}$ we have that

$$\lambda_j^i(\Theta_j) := \nabla \psi_{x_i}^{(1)}(\zeta_{\Theta_j}^i) \cdot (\xi_{\Theta_j} - E_l) = \mathcal{O}(\delta_j), \tag{8.8}$$

and for $i \in \{l+2, \dots, d\}$ we have that

$$\nu_j^i(\Theta_j) := \psi_{x_i}^{(1)}(\xi_{\Theta_j}) = \mathcal{O}(\delta_j^{\alpha-1}), \tag{8.9}$$

as $j \longrightarrow \infty$, uniformly in $\omega \in \mathcal{H}_m^l$, by (8.6). Then the basic iteration step is

$$\Theta_n^{(1)} - \frac{1}{l+1} = k_{n-1}(\Theta_{n-1}) \left(\Theta_{n-1}^{(1)} - \frac{1}{l+1} \right) + \frac{\sigma_n}{\tau_n} \sum_{j=2}^{l+1} \lambda_{n-1}^j(\Theta_{n-1}) \left(\Theta_{n-1}^{(j)} - \frac{1}{l+1} \right)$$
$$+ \frac{\alpha}{l+1} \frac{\sigma_n}{\tau_n} \sum_{j=l+2}^{d} \Theta_{n-1}^{(j)} + \frac{\sigma_n}{\tau_n} \sum_{j=l+2}^{d} \nu_{n-1}^j(\Theta_{n-1}) \Theta_{n-1}^{(j)} + \frac{\sqrt{\sigma_n}}{\tau_n} \varepsilon_n^{(1)} \sqrt{P_{n-1}^{(1)}(1 - P_{n-1}^{(1)})}. \tag{8.10}$$

Iterating (8.10) $n-m$ times yields, using empty sum and empty product conventions,

$$\Theta_n^{(1)} - \frac{1}{l+1} = \prod_{j=m}^{n-1} k_j(\Theta_j) \left(\Theta_m^{(1)} - \frac{1}{l+1} \right) + \sum_{k=m+1}^{n} \prod_{j=k}^{n-1} k_j(\Theta_j) \frac{\sigma_k}{\tau_k} \sum_{i=2}^{l+1} \lambda_{k-1}^i(\Theta_{k-1})$$
$$\left(\Theta_{k-1}^{(i)} - \frac{1}{l+1} \right) + \sum_{k=m+1}^{n} \prod_{j=k}^{n-1} k_j(\Theta_j) \frac{\sigma_k}{\tau_k} \sum_{i=l+2}^{d} \nu_{k-1}^i(\Theta_{k-1}) \Theta_{k-1}^{(i)} + \frac{\alpha}{l+1}$$
$$\sum_{k=m+1}^{n} \prod_{j=k}^{n-1} k_j(\Theta_j) \frac{\sigma_k}{\tau_k} \sum_{i=l+2}^{d} \Theta_{k-1}^{(i)} + \sum_{k=m+1}^{n} \prod_{j=k}^{n-1} k_j(\Theta_j) \frac{\sqrt{\sigma_k}}{\tau_k} \varepsilon_k^{(1)} \sqrt{P_{k-1}^{(1)}(1 - P_{k-1}^{(1)})}. \tag{8.11}$$

On a side note, the fact that in the case $l = d-1$, (8.7) yields

$$\Theta_n^{(1)} - \frac{1}{d} = \left(\frac{\tau_{n-1}}{\tau_n} + \frac{\sigma_n}{\tau_n} \left(\frac{(d-1)\alpha}{d} + \nabla \psi_{x_1}^{(1)}(\zeta_{\Theta_{n-1}}^1) \cdot (\xi_{\Theta_{n-1}} - E_{d-1}) \right) \right) \left(\Theta_{n-1}^{(1)} - \frac{1}{d} \right)$$
$$- \frac{\sigma_n}{\tau_n} \sum_{j=2}^{d} \left(\frac{\alpha}{d} - \nabla \psi_{x_j}^{(1)}(\zeta_{\Theta_{n-1}}^j) \cdot (\xi_{\Theta_{n-1}} - E_{d-1}) \right) \left(\Theta_{n-1}^{(j)} - \frac{1}{d} \right) + \frac{\sqrt{\sigma_n P_{n-1}^{(1)}(1 - P_{n-1}^{(1)})}}{\tau_n} \varepsilon_n^{(1)}$$
$$= \left(\frac{\tau_{n-1}}{\tau_n} + \frac{\sigma_n}{\tau_n} \left(\frac{(d-1)\alpha}{d} + \nabla \psi_{x_1}^{(1)}(\zeta_{\Theta_{n-1}}^1) \cdot (\xi_{\Theta_{n-1}} - E_{d-1}) \right) \right) \left(\Theta_{n-1}^{(1)} - \frac{1}{d} \right)$$
$$- \frac{\alpha}{d} \frac{\sigma_n}{\tau_n} \sum_{j=2}^{d} \left(\Theta_{n-1}^{(j)} - \frac{1}{d} \right) + \frac{\sigma_n}{\tau_n} \sum_{j=2}^{d} \nabla \psi_{x_j}^{(1)}(\zeta_{\Theta_{n-1}}^j) \cdot (\xi_{\Theta_{n-1}} - E_{d-1}) \left(\Theta_{n-1}^{(j)} - \frac{1}{d} \right)$$
$$+ \frac{\sqrt{\sigma_n}}{\tau_n} \varepsilon_n^{(1)} \sqrt{P_{n-1}^{(1)}(1 - P_{n-1}^{(1)})},$$

combined with the fact that

$$\sum_{j=2}^{d} \left(\Theta_{n-1}^{(j)} - \frac{1}{d} \right) = \left(1 - \Theta_{n-1}^{(1)} - \frac{d-1}{d} \right) = - \left(\Theta_{n-1}^{(1)} - \frac{1}{d} \right),$$

implies that

$$\Theta_n^{(1)} - \frac{1}{d} = \left(\frac{\tau_{n-1}}{\tau_n} + \frac{\sigma_n}{\tau_n}\left(\frac{(d-1)\alpha}{d} + \nabla\psi_{x_1}^{(1)}(\zeta_{\Theta_{n-1}}^1)\cdot(\xi_{\Theta_{n-1}} - E_{d-1})\right)\right)\left(\Theta_{n-1}^{(1)} - \frac{1}{d}\right)$$
$$+ \frac{\alpha}{d}\frac{\sigma_n}{\tau_n}\left(\Theta_{n-1}^{(1)} - \frac{1}{d}\right) + \frac{\sigma_n}{\tau_n}\sum_{j=2}^{d}\nabla\psi_{x_j}^{(1)}(\zeta_{\Theta_{n-1}}^j)\cdot(\xi_{\Theta_{n-1}} - E_{d-1})\left(\Theta_{n-1}^{(j)} - \frac{1}{d}\right)$$
$$+ \frac{\sqrt{\sigma_n}}{\tau_n}\varepsilon_n^{(1)}\sqrt{P_{n-1}^{(1)}(1-P_{n-1}^{(1)})} = \left(\frac{\tau_{n-1}+\alpha\sigma_n}{\tau_n} + \frac{\sigma_n}{\tau_n}\nabla\psi_{x_1}^{(1)}(\zeta_{\Theta_{n-1}}^1)\cdot(\xi_{\Theta_{n-1}} - E_{d-1})\right)$$
$$\left(\Theta_{n-1}^{(1)} - \frac{1}{d}\right) + \frac{\sigma_n}{\tau_n}\sum_{j=2}^{d}\nabla\psi_{x_j}^{(1)}(\zeta_{\Theta_{n-1}}^j)\cdot(\xi_{\Theta_{n-1}} - E_{d-1})\left(\Theta_{n-1}^{(j)} - \frac{1}{d}\right)$$
$$+ \frac{\sqrt{\sigma_n}}{\tau_n}\varepsilon_n^{(1)}\sqrt{P_{n-1}^{(1)}(1-P_{n-1}^{(1)})},$$

so defining similarly $\lambda_j^i(\Theta_j)$ for $i \in \{2,\ldots,d\}$, the iterative formula is the same, with two of the four series missing, which is compatible with the empty sum convention. The two series left will be dealt with in the same way as the corresponding ones in the more general case, since the same asymptotics hold, thanks to the regularity of the function $\psi^{(1)}$, and its derivatives, at the full equilibrium E. Therefore, in what follows, no distinction will need to be made between the two cases. We now go back to developing the iteration formula.

Since

$$\frac{\prod_{j=k}^{n-1} k_j(\Theta_j)}{\prod_{j=m}^{n-1} k_j(\Theta_j)} = \frac{1}{\prod_{j=m}^{k-1} k_j(\Theta_j)},$$

it follows that factoring it out in (8.11) yields

$$\Theta_n^{(1)} - \frac{1}{l+1} = \prod_{j=m}^{n-1} k_j(\Theta_j)\Bigg[\Theta_m^{(1)} - \frac{1}{l+1} + \sum_{k=m+1}^{n}\frac{1}{\prod_{j=m}^{k-1}k_j(\Theta_j)}\frac{\sigma_k}{\tau_k}\sum_{i=2}^{l+1}\lambda_{k-1}^i(\Theta_{k-1})$$
$$\left(\Theta_{k-1}^{(i)} - \frac{1}{l+1}\right) + \sum_{k=m+1}^{n}\frac{1}{\prod_{j=m}^{k-1}k_j(\Theta_j)}\frac{\sigma_k}{\tau_k}\sum_{i=l+2}^{d}\nu_{k-1}^i(\Theta_{k-1})\Theta_{k-1}^{(i)} +$$
$$\frac{\alpha}{l+1}\sum_{k=m+1}^{n}\frac{1}{\prod_{j=m}^{k-1}k_j(\Theta_j)}\frac{\sigma_k}{\tau_k}\sum_{i=l+2}^{d}\Theta_{k-1}^{(i)} + \sum_{k=m+1}^{n}\frac{1}{\prod_{j=m}^{k-1}k_j(\Theta_j)}\frac{\sqrt{\sigma_k}}{\tau_k}\varepsilon_k^{(1)}\sqrt{P_{k-1}^{(1)}(1-P_{k-1}^{(1)})}\Bigg]. \tag{8.12}$$

Note that having defined

$$\pi_{m,k} := \prod_{j=m}^{k-1}\frac{\tau_j + \alpha\sigma_{j+1}}{\tau_{j+1}},$$

we have that

$$\prod_{j=m}^{k-1} k_j(\Theta_j) = \prod_{j=m}^{k-1}\frac{\tau_j+\alpha\sigma_{j+1}}{\tau_{j+1}}\prod_{j=m}^{k-1}\left(1 + \frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\mathcal{O}(\delta_j)\right) = \pi_{m,k}(1+o(1)) \tag{8.13}$$

as $m \longrightarrow \infty$ uniformly in $\omega \in \mathcal{H}_m^l$ and $k \geq m$. This follows, since by Lemma 8.2,

$$\prod_{j=m}^{k-1}\left(1 + \frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\mathcal{O}(\delta_j)\right) = \exp\sum_{j=m}^{k-1}\log\left(1 + \frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\mathcal{O}(\delta_j)\right),$$

which can be derived by noting that as

$$\frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\mathcal{O}(\delta_j)=o(1),$$

by linear Taylor expansion

$$\sum_{j=m}^{k-1}\log\left(1+\frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\mathcal{O}(\delta_j)\right)=\sum_{j=m}^{k-1}\frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\mathcal{O}(\delta_j)+\sum_{j=m}^{k-1}o\left(\frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\mathcal{O}(\delta_j)\right),$$

and since

$$\sum_{j=0}^{\infty}\frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\delta_j<\sum_{j=0}^{\infty}\frac{\sigma_{j+1}}{\tau_{j+1}}\delta_j<\infty,$$

as $m\longrightarrow\infty$ uniformly in $\omega\in\mathcal{H}_m^l$ and $k\geq m$ we have that

$$\sum_{j=m}^{k-1}\log\left(1+\frac{\sigma_{j+1}}{\tau_j+\alpha\sigma_{j+1}}\mathcal{O}(\delta_j)\right)=o(1).$$

In conclusion plugging (8.13) into (8.12) yields

$$\begin{aligned}
&\Theta_n^{(1)}-\frac{1}{l+1}=\\
&\pi_{m,n}(1+o(1))\Bigg[\Theta_m^{(1)}-\frac{1}{l+1}+\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\sum_{i=2}^{l+1}\lambda_{k-1}^i(\Theta_{k-1})\left(\Theta_{k-1}^{(i)}-\frac{1}{l+1}\right)\\
&+\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\sum_{i=l+2}^{d}\nu_{k-1}^i(\Theta_{k-1})\Theta_{k-1}^{(i)}+\frac{\alpha}{l+1}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\sum_{i=l+2}^{d}\Theta_{k-1}^{(i)}\\
&+\sum_{k=m+1}^{n}\frac{\sqrt{\sigma_k}}{\pi_{m,k}\tau_k}\varepsilon_k^{(1)}\sqrt{P_{k-1}^{(1)}(1-P_{k-1}^{(1)})}\Bigg],
\end{aligned}$$

which we can rearrange into the full iteration formula,holding uniformly in $n>m$ and $\omega\in\mathcal{H}_m^l$, as $m\longrightarrow\infty$:

$$\begin{aligned}
&\Theta_n^{(1)}-\frac{1}{l+1}=\\
&\pi_{m,n}(1+o(1))\Bigg[\Theta_m^{(1)}-\frac{1}{l+1}+\sum_{i=2}^{l+1}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\lambda_{k-1}^i(\Theta_{k-1})\left(\Theta_{k-1}^{(i)}-\frac{1}{l+1}\right)\\
&+\sum_{i=l+2}^{d}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\nu_{k-1}^i(\Theta_{k-1})\Theta_{k-1}^{(i)}+\frac{\alpha}{l+1}\sum_{i=l+2}^{d}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\Theta_{k-1}^{(i)}\\
&+\frac{\sqrt{l+o(1)}}{l+1}\sum_{k=m+1}^{n}\frac{\sqrt{\sigma_k}}{\pi_{m,k}\tau_k}\varepsilon_k^{(1)}\Bigg].
\end{aligned}\tag{8.14}$$

Step 3. In this step we focus on the term

$$\sum_{k=m+1}^{n}\frac{\sqrt{\sigma_k}}{\tau_k\pi_{m,k}}\varepsilon_k^{(1)},$$

we find a uniform estimate for the approximation error of a normalized binomial random variable approaching a standard normal, and show asymptotic normality of this term. The characteristic function, denoted as ϕ, will be our main tool. Given binomial random variables distributed with size n and probability parameter p, $B_n \sim \mathrm{Bin}(n,p)$, where

$$0 < \frac{1}{2(l+1)} < p < \frac{3}{2(l+1)} < 1,$$

define

$$X_n := \frac{B_n - np}{\sqrt{np(1-p)}},$$

and recall that

$$\phi_{B_n}(t) = \left(1 - p + pe^{it}\right)^n.$$

Then it holds that

$$\begin{aligned}
\phi_{X_n}(t) &:= \mathbb{E}e^{itX_n} = \mathbb{E}\exp\left\{i\frac{t(B_n - np)}{\sqrt{np(1-p)}}\right\} = \exp\left\{-i\frac{tnp}{\sqrt{np(1-p)}}\right\} \\
&\cdot \mathbb{E}\exp\left\{i\frac{t}{\sqrt{np(1-p)}}B_n\right\} = \exp\left\{-i\frac{tnp}{\sqrt{np(1-p)}}\right\}\phi_{B_n}\left(\frac{t}{\sqrt{np(1-p)}}\right) \\
&= \exp\left\{-i\frac{tnp}{\sqrt{np(1-p)}}\right\}\left(1 - p + p\exp\frac{it}{\sqrt{np(1-p)}}\right)^n \\
&= \exp\left\{-i\frac{tnp}{\sqrt{np(1-p)}}\right\}\exp\left\{n\log\left(1 - p + p\exp\frac{it}{\sqrt{np(1-p)}}\right)\right\} \\
&= \exp\left[n\log\left(1 - p + p\exp\frac{it}{\sqrt{np(1-p)}}\right) - i\frac{tnp}{\sqrt{np(1-p)}}\right].
\end{aligned}$$

We expand in series the complex exponential as $n \longrightarrow \infty$, assuming that t belongs to a bounded interval of $\mathbb{R}$: t will be kept in the $\mathcal{O}$ estimate, so as to preserve uniformity of the final result with respect to t in the bounded interval $n \longrightarrow \infty$ (because of the obvious explicit bounds that have been imposed on p, all p-terms will be absorbed in the $\mathcal{O}$-constant):

$$\exp\frac{it}{\sqrt{np(1-p)}} = 1 + \frac{it}{\sqrt{np(1-p)}} - \frac{t^2}{2np(1-p)} + \mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right).$$

Then

$$\phi_{X_n}(t) = \exp\left[n\log\left(1 + \frac{\sqrt{p}it}{\sqrt{n(1-p)}} - \frac{t^2}{2n(1-p)} + \mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)\right) - i\frac{tnp}{\sqrt{np(1-p)}}\right].$$

Expanding in series the logarithm yields

$$\log\left(1+\frac{\sqrt{p}it}{\sqrt{n(1-p)}}-\frac{t^2}{2n(1-p)}+\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)\right)=\frac{\sqrt{p}it}{\sqrt{n(1-p)}}-\frac{t^2}{2n(1-p)}$$
$$+\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)-\frac{1}{2}\left(\frac{\sqrt{p}it}{\sqrt{n(1-p)}}-\frac{t^2}{2n(1-p)}+\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)\right)^2$$
$$+\mathcal{O}\left(\frac{\sqrt{p}it}{\sqrt{n(1-p)}}-\frac{t^2}{2n(1-p)}+\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)\right)^3$$
$$=\frac{\sqrt{p}it}{\sqrt{n(1-p)}}-\frac{t^2}{2n(1-p)}+\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)+\frac{pt^2}{2n(1-p)}+\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)$$
$$=\frac{\sqrt{p}it}{\sqrt{n(1-p)}}-\frac{t^2}{2n}+\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right),$$

as the whole term satisfies

$$\mathcal{O}\left(\frac{\sqrt{p}it}{\sqrt{n(1-p)}}-\frac{t^2}{2n(1-p)}+\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)\right)^3=\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right),$$

and so do the ones that have been neglected from the square of the trinomial. We plug also this expansion in, and it yields that, as $n \longrightarrow \infty$, uniformly in t in the bounded interval and $1/[2(l+2)] < p < l/[2(l+1)]$,

$$\phi_{X_n}(t)=\exp\left[\frac{\sqrt{npit}}{\sqrt{(1-p)}}-\frac{t^2}{2}+n\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)-i\frac{tnp}{\sqrt{np(1-p)}}\right]$$
$$=\exp\left[-\frac{t^2}{2}+n\mathcal{O}\left(\frac{t^3}{n\sqrt{n}}\right)\right]. \tag{8.15}$$

The uniformity with respect to the parameters in the intervals chosen make this result flexible enough, to be applied to the series, which now we turn our attention to. Note that $\varepsilon_k^{(1)}$ has the same structure as X_k, with $B_k^{(1)} \sim \text{Bin}(\sigma_k, P_{k-1}^{(1)})$ instead of B_k, and therefore size σ_k instead of size k (σ_k, unlike k, might or might not tend to infinity, hence the necessity of the error term) and probability parameter $P_{k-1}^{(1)}$ instead of p. This justifies the choice of the interval for the parameters: since uniformly on the event $\mathcal{H}_m^l$, $|P_{k-1}^{(1)}-1/(l+1)| = \mathcal{O}(\delta_{k-1})$, a natural range for the probability parameter, on which uniformity is necessary, is any interval containing $1/(l+1)$ and excluding 0 and 1. The next step is to study the characteristic function, conditionally on $\mathscr{F}_{k-1}$ (which gives information about $P_{k-1}^{(1)}$), of $\varepsilon_k^{(1)}$. A technicality is necessary here, since while on $\mathcal{H}_m^l$, $P_{k-1}^{(1)}$ will eventually fall in the range that ensures the convergence result, a modified version of $\varepsilon_k^{(1)}$ needs to be constructed (all the while remaining asymptotically equivalent to the original), in order to allow taking conditional expectation and working out the conditional characteristic function, since in Ω there will be also events on which $P_{k-1}^{(1)}$ drops out of the range fixed for the parameters. As $m \longrightarrow \infty$, due to asymptotic equivalence, the iterative expression of $\Theta_n - 1/(l+1)$, as on $\mathcal{H}_m^l$, will not be affected. We take care of this technicality by defining independent random variables

$$B_k^0 \sim \text{Bin}\left(\sigma_r, \frac{1}{l+1}\right),$$

and consequently defining

$$\varepsilon_k^0 := \frac{B_k^0 - \frac{\sigma_k}{l+1}}{\frac{\sqrt{l\sigma_k}}{l+1}}.$$

Denote

$$\hat{\varepsilon}_k := \varepsilon_k^{(1)} \mathbf{1}_{\left\{|P_{k-1}^{(1)} - \frac{1}{l+1}| < \frac{1}{2(l+1)}\right\}} + \varepsilon_k^0 \mathbf{1}_{\left\{|P_{k-1}^{(1)} - \frac{1}{l+1}| \geq \frac{1}{2(l+1)}\right\}}.$$

Since on $\mathcal{H}_m^l$ we have that

$$\left| P_{k-1}^{(1)} - \frac{1}{l+1} \right| = \mathcal{O}(\delta_m),$$

it follows that as $m \longrightarrow \infty$, on $\mathcal{H}_m^l$ we have that

$$\left| P_{k-1}^{(1)} - \frac{1}{l+1} \right| < \frac{1}{2(l+1)}$$

for all $k \geq m+1$, since $\delta_m \longrightarrow 0$. Therefore as $m \longrightarrow \infty$, on $\mathcal{H}_m^l$ we have that $\hat{\varepsilon}_k \equiv \varepsilon_k^{(1)}$ uniformly in $k > m$, which means that as $m \longrightarrow \infty$, uniformly in $\omega \in \mathcal{H}_m^l$ and $n > m$,

$$\begin{aligned} &\Theta_n^{(1)} - \frac{1}{l+1} \\ &= \pi_{m,n}(1 + o(1)) \Bigg[\Theta_m^{(1)} - \frac{1}{l+1} + \sum_{i=2}^{l+1} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \lambda_{k-1}^i(\Theta_{k-1}) \left(\Theta_{k-1}^{(i)} - \frac{1}{l+1} \right) \\ &+ \sum_{i=l+2}^{d} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \nu_{k-1}^i(\Theta_{k-1}) \Theta_{k-1}^{(i)} + \frac{\alpha}{l+1} \sum_{i=l+2}^{d} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \Theta_{k-1}^{(i)} \\ &+ \frac{\sqrt{l + o(1)}}{l+1} \sum_{k=m+1}^{n} \frac{\sqrt{\sigma_k}}{\pi_{m,k}\tau_k} \hat{\varepsilon}_k \Bigg]. \end{aligned} \tag{8.16}$$

Since all the probability parameters involved in $\hat{\varepsilon}_k$, $P_{k-1}^{(1)}$ are in the range chosen uniformly on the probability space as $k \longrightarrow \infty$; since $1 \leq \sigma_{k-1}$ is either divergent or bounded by Assumption S, and:

- if σ_{k-1} is divergent, all calculations done on the conditional characteristic function are going to be analogous to those for the unconditional one (see [51] for some background on the concept of conditional characteristic function introduced by Loève), as $k \longrightarrow \infty$, uniformly in ω and t in the bounded interval, conditionally on $\mathscr{F}_{k-1}$;
- if σ_{k-1} is bounded, then all $\mathcal{O}$-estimates apply trivially uniformly in k, ω and t in the bounded interval;

we can conclude, by applying (8.15), that as $k \longrightarrow \infty$, uniformly in ω and t in the bounded interval,

$$\phi_{\mathscr{F}_{k-1}}^{\hat{\varepsilon}_k}(t) = \exp\left[-\frac{t^2}{2} + \sigma_k \mathcal{O}\left(\frac{t^3}{\sigma_k \sqrt{\sigma_k}} \right) \right]. \tag{8.17}$$

We apply (8.17) to work out

$$\phi_{\mathscr{F}_m}^{\sum_{k=m+1}^n \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \hat{\varepsilon}_k}(t)$$

as $m \longrightarrow \infty$, pointwise in t. The goal is to asymptotically obtain a standard normal characteristic function, and to do so, it will be clear that it is necessary to divide

$$\sum_{k=m+1}^{n} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \hat{\varepsilon}_k$$

by the following deterministic scale, which ensures unitary variance of the term:

$$\mu_{m,n} := \sqrt{\sum_{k=m+1}^{n} \frac{\sigma_k}{\tau_k^2 \pi_{m,k}^2}}.$$

Denote

$$N_{m,n} := \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \hat{\varepsilon}_k.$$

Then for every fixed t,

$$\phi_{\mathscr{F}_m}^{N_{m,n}}(t) = \mathbb{E}_{\mathscr{F}_m} \exp\left\{ i \frac{t}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \hat{\varepsilon}_k \right\} =$$

$$\mathbb{E}_{\mathscr{F}_m} \mathbb{E}_{\mathscr{F}_{m+1}} \dots \mathbb{E}_{\mathscr{F}_{n-1}} \exp\left\{ i \frac{t}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \hat{\varepsilon}_k \right\} =$$

$$\mathbb{E}_{\mathscr{F}_m} \dots \mathbb{E}_{\mathscr{F}_{n-1}} \prod_{k=m+1}^{n} \exp\left\{ i \frac{t}{\mu_{m,n}} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \hat{\varepsilon}_k \right\} =$$

$$\mathbb{E}_{\mathscr{F}_m} \dots \mathbb{E}_{\mathscr{F}_{n-2}} \prod_{k=m+1}^{n-1} \exp\left\{ i \frac{t}{\mu_{m,n}} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \hat{\varepsilon}_k \right\} \mathbb{E}_{\mathscr{F}_{n-1}} \exp\left\{ i \frac{t}{\mu_{m,n}} \frac{\sqrt{\sigma_n}}{\tau_n \pi_{m,n}} \hat{\varepsilon}_n \right\} =$$

$$\dots = \prod_{k=m+1}^{n} \mathbb{E}_{\mathscr{F}_{k-1}} \exp\left\{ i \frac{t}{\mu_{m,n}} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \hat{\varepsilon}_k \right\},$$

yielding

$$\phi_{\mathscr{F}_m}^{N_{m,n}}(t) = \prod_{k=m+1}^{n} \phi_{\mathscr{F}_{k-1}}^{\hat{\varepsilon}_k}(s_k), \tag{8.18}$$

where

$$s_k := \frac{t}{\mu_{m,n}} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}}.$$

Note that for every $m+1 \leq k \leq n$,

$$|s_k| = \frac{|t| \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}}}{\sqrt{\sum_{j=m+1}^{n} \frac{\sigma_j}{\tau_j^2 \pi_{m,j}^2}}} \leq \frac{|t| \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}}}{\frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}}} = |t|,$$

thus as $m \longrightarrow \infty$, s_k always belongs to the fixed bounded interval $[-t, t]$, and therefore by (8.17), which applies to (8.18) via the dummy variable $t = s_k$ for all k, it is possible to conclude that as $m \longrightarrow \infty$, pointwise in t,

$$\phi_{\mathscr{F}_m}^{N_{m,n}}(t) = \prod_{k=m+1}^{n} \phi_{\mathscr{F}_{k-1}}^{\hat{\varepsilon}_k}(s_k) = \prod_{k=m+1}^{n} \exp\left[-\frac{s_k^2}{2} + \sigma_k \mathcal{O}\left(\frac{s_k^3}{\sigma_k \sqrt{\sigma_k}} \right) \right]$$

$$= \exp \sum_{k=m+1}^{n} \left[-\frac{s_k^2}{2} + \sigma_k \mathcal{O}\left(\frac{s_k^3}{\sigma_k \sqrt{\sigma_k}} \right) \right] = \exp - \frac{\sum_{k=m+1}^{n} \left(\frac{t}{\mu_{m,n}} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \right)^2}{2}$$

$$\exp \sum_{k=m+1}^{n} \sigma_k \mathcal{O}\left(\frac{\left(\frac{t}{\mu_{m,n}} \frac{\sqrt{\sigma_k}}{\tau_k \pi_{m,k}} \right)^3}{\sigma_k \sqrt{\sigma_k}} \right) = \exp - \frac{t^2}{2\mu_{m,n}^2} \sum_{k=m+1}^{n} \frac{\sigma_k}{\tau_k^2 \pi_{m,k}^2}$$

$$\exp \frac{1}{\mu_{m,n}^3} \sum_{k=m+1}^{n} \frac{\sigma_k}{\tau_k^3 \pi_{m,k}^3} \mathcal{O}\left(t^3\right) = \exp - \frac{t^2}{2} \exp \frac{1}{\mu_{m,n}^3} \sum_{k=m+1}^{n} \frac{\sigma_k}{\tau_k^3 \pi_{m,k}^3} \mathcal{O}(1).$$

To achieve asymptotic normality, we only need to show that, regardless of whether $\{\sigma_k\}$ is bounded or divergent to infinity, by Assumption S,

$$\frac{1}{\mu_{m,n}^3}\sum_{k=m+1}^{n}\frac{\sigma_k}{\tau_k^3\pi_{m,k}^3}\longrightarrow 0.$$

We anticipate that this result ceases to be uniform in n in the bounded scenario. Indeed the bounded scenario will require negligibility of m with respect to n. Therefore we will introduce a new notation, $n = n_m$, such that $m = o(n_m)$ (for instance, we can choose $n_m = m^2$ or an even faster regime, depending on an other condition, which will be introduced later) in the final step (we will keep the notation n until then). Once this is done, we will have shown that uniformly on the probability space and conditionally on $\mathscr{F}_m$,

$$N_{m,n_m}\xrightarrow{w}\mathscr{N}(0,1) \tag{8.19}$$

as $m\longrightarrow\infty$.

Case 1. Assume $\sigma_k\longrightarrow\infty$. Since uniformly in m for nonnegative x_i

$$\left(\sum_{i=1}^{m}x_i\right)^{\frac{3}{2}}=\sum_{i=1}^{m}x_i\left(\sum_{i=1}^{m}x_i\right)^{\frac{1}{2}}\geq\sum_{i=1}^{m}x_i\,(x_i)^{\frac{1}{2}}=\sum_{i=1}^{m}x_i^{\frac{3}{2}},$$

it follows that

$$\mu_{m,n}^3=\left(\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}^2\tau_k^2}\right)^{\frac{3}{2}}\geq\sum_{k=m+1}^{n}\frac{\sigma_k^{\frac{3}{2}}}{\pi_{m,k}^3\tau_k^3}\geq\min_{m<k\leq n}\sqrt{\sigma_k}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}^3\tau_k^3},$$

therefore, as $m\longrightarrow\infty$ uniformly in n,

$$\frac{1}{\mu_{m,n}^3}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}^3\tau_k^3}\leq\max_{m<k\leq n}\frac{1}{\sqrt{\sigma_k}}\longrightarrow 0,$$

and the claim follows.

Case 2. Assume that there is a constant σ bounding $\{\sigma_k\}$. In this case it holds that

$$(1+o(1))\left(\frac{k}{m}\right)^{\frac{\alpha-1}{\sigma}+o(1)}\leq\pi_{m,k}\leq(1+o(1))\left(\frac{k}{m}\right)^{\sigma(\alpha-1)+o(1)}. \tag{8.20}$$

We show (8.20) after the conclusion is reached. If one assumes it for now, since by Remark 8.3, for any $\gamma>1$, uniformly in $n>n_m>m$ such that $m=o(n_n)$, it holds that

$$\sum_{k=m+1}^{n}\frac{1}{k^{\gamma+o(1)}}\sim\frac{1}{(\gamma-1)m^{\gamma-1+o(1)}},$$

it follows that, as $k \leq \tau_k \leq \tau_0 + k\sigma$,

$$\mu_{m,n}^3 \geq \left(\sum_{k=m+1}^{n} \frac{\sigma_k}{(1+o(1)) \left(\frac{k}{m}\right)^{2\sigma(\alpha-1)+o(1)} \tau_k^2} \right)^{\frac{3}{2}}$$
$$= m^{3\sigma(\alpha-1)+o(1)} \left(\sum_{k=m+1}^{n} \frac{1}{(1+o(1)) k^{2\sigma(\alpha-1)+o(1)} (\tau_0 + k\sigma)^2} \right)^{\frac{3}{2}}$$
$$= (1+o(1)) m^{3\sigma(\alpha-1)+o(1)} \left(\sum_{k=m+1}^{n} \frac{1}{k^{2\sigma(\alpha-1)+o(1)} k^2 \sigma^2 (1+o(1))} \right)^{\frac{3}{2}}$$
$$\sim \frac{m^{3\sigma(\alpha-1)} + o(1)}{\sigma^3} \left(\sum_{k=m+1}^{n} \frac{1}{k^{2\sigma(\alpha-1)+2+o(1)}} \right)^{\frac{3}{2}} \sim \frac{m^{3\sigma(\alpha-1)+o(1)}}{\sigma^3 (2\sigma(\alpha-1)+1)^{\frac{3}{2}}} \left(\frac{1}{m^{2\sigma(\alpha-1)+1+o(1)}} \right)^{\frac{3}{2}}$$
$$= \frac{1}{\sigma^3 (2\sigma(\alpha-1)+1)^{\frac{3}{2}}} \frac{1}{m^{\frac{3}{2}+o(1)}},$$

and

$$\sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}^3 \tau_k^3} \leq \sum_{k=m+1}^{n} \frac{\sigma}{(1+o(1)) \left(\frac{k}{m}\right)^{3\frac{\alpha-1}{\sigma}+o(1)} k^3}$$
$$= (1+o(1)) \sigma m^{3\frac{\alpha-1}{\sigma}+o(1)} \sum_{k=m+1}^{n} \frac{1}{k^{3\frac{\alpha-1}{\sigma}+3+o(1)}} \sim \frac{\sigma m^{3\frac{\alpha-1}{\sigma}+o(1)}}{3\frac{\alpha-1}{\sigma}+2} \frac{1}{m^{3\frac{\alpha-1}{\sigma}+2+o(1)}}$$
$$= \frac{\sigma}{3\frac{\alpha-1}{\sigma}+2} \frac{1}{m^{2+o(1)}}$$

as $m \longrightarrow \infty$ uniformly in $n \geq m^2$ (choosing $n_m = m^2$ is just an arbitrary choice, so as to avoid carrying on with the condition $m = o(n_m)$). Hence

$$\frac{1}{\mu_{m,n}^3} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}^3 \tau_k^3} \leq \frac{\sum_{k=m+1}^{n} \frac{\sigma}{(1+o(1))\left(\frac{k}{m}\right)^{3\frac{\alpha-1}{\sigma}+o(1)} k^3}}{\left(\sum_{k=m+1}^{n} \frac{\sigma_k}{(1+o(1))\left(\frac{k}{m}\right)^{2\sigma(\alpha-1)+o(1)} \tau_k^2} \right)^{\frac{3}{2}}} \sim \frac{\frac{\sigma}{3\frac{\alpha-1}{\sigma}+2} \frac{1}{m^{2+o(1)}}}{\frac{1}{\sigma^3(2\sigma(\alpha-1)+1)^{\frac{3}{2}}} \frac{1}{m^{\frac{3}{2}+o(1)}}}$$
$$= \frac{\sigma^5 (2\sigma(\alpha-1)+1)^{\frac{3}{2}}}{3(\alpha-1)+2\sigma} \frac{1}{m^{\frac{1}{2}+o(1)}} \longrightarrow 0$$

as $m \longrightarrow \infty$, uniformly in $n \geq m^2$.

The argument used to show (8.20) goes as follows.

$$\pi_{m,k} = \exp \sum_{j=m+1}^{k} \log \left(1 + (\alpha-1) \frac{\sigma_j}{\tau_j} \right) \leq \exp \sum_{j=m+1}^{k} \log \left(1 + (\alpha-1) \frac{\sigma}{j} \right) =$$
$$\exp \sum_{j=m+1}^{k} \left[(\alpha-1) \frac{\sigma}{j} + \mathcal{O}\left(\frac{1}{j^2} \right) \right] = \exp \left\{ [(\alpha-1)\sigma + o(1)] \sum_{j=m+1}^{k} \frac{1}{j} \right\}.$$

Since

$$\sum_{j=m+1}^{k} \frac{1}{j} \geq \int_{m+1}^{k} \frac{dx}{x} = \log k - \log(m+1)$$

and

$$\sum_{j=m+1}^{k} \frac{1}{j} \leq \frac{1}{m+1} + \int_{m+2}^{k+1} \frac{dx}{x-1} = \frac{1}{m+1} + \log k - \log(m+1),$$

it follows that

$$0 \le \sum_{j=m+1}^{k} \frac{1}{j} - (\log k - \log(m+1)) \le \frac{1}{m+1}.$$

Since $\log(m+1) = \log m + \log(1 + {}^{1}\!/\!{}_{m})$,

$$-\log\left(1+\frac{1}{m}\right) \le \sum_{j=m+1}^{k} \frac{1}{j} - (\log k - \log m) \le \frac{1}{m+1} - \log\left(1+\frac{1}{m}\right),$$

then by using the series expansion of the logarithm, we have that

$$\begin{aligned}
&-\frac{1}{m} + \mathcal{O}\left(\frac{1}{m^2}\right) \le \sum_{j=m+1}^{k} \frac{1}{j} - (\log k - \log m) \le \frac{1}{m+1} - \frac{1}{m} + \mathcal{O}\left(\frac{1}{m^2}\right) \\
&= -\frac{1}{m(m+1)} + \mathcal{O}\left(\frac{1}{m^2}\right) = \mathcal{O}\left(\frac{1}{m^2}\right),
\end{aligned}$$

yielding

$$\sum_{j=m+1}^{k} \frac{1}{j} = \log k - \log m + \mathcal{O}\left(\frac{1}{m}\right)$$

uniformly in $k > m+1$, as $m \longrightarrow \infty$. Therefore

$$\begin{aligned}
\pi_{m,k} &\le \exp\{[(\alpha-1)\sigma + o(1)]\,(\log k - \log m + o(1))\} \\
&= \exp\left\{[(\alpha-1)\sigma + o(1)]\left(\log\frac{k}{m} + o(1)\right)\right\} \\
&= \exp\left\{[(\alpha-1)\sigma + o(1)]\left(\log\frac{k}{m}\right)\exp o(1)\right\} = (1+o(1))\frac{k^{(\alpha-1)\sigma+o(1)}}{m^{(\alpha-1)\sigma+o(1)}}.
\end{aligned}$$

Similarly, since $\sigma_j \ge 1$,

$$\begin{aligned}
\pi_{m,k} &\ge \exp\sum_{j=m+1}^{k} \log\left(1+\frac{\alpha-1}{\tau_0 + j\sigma}\right) = \exp\sum_{j=m+1}^{k}\left(\frac{\alpha-1}{\sigma+o(1)}\frac{1}{j}\right) + \mathcal{O}\left(\frac{1}{j^2}\right) \\
&= \exp\left(\frac{\alpha-1}{\sigma} + o(1)\right)\sum_{j=m+1}^{k}\frac{1}{j} = \exp\left(\frac{\alpha-1}{\sigma} + o(1)\right)\left(\log\frac{k}{m} + o(1)\right) \\
&= (1+o(1))\frac{k^{\frac{\alpha-1}{\sigma}+o(1)}}{m^{\frac{\alpha-1}{\sigma}+o(1)}}.
\end{aligned}$$

Step 4. The inductive formula will be, from now on, rewritten uniformly in $n \ge m^2$ and $\omega \in \mathcal{H}_m^l$, as

$$\begin{aligned}
&\Theta_n^{(1)} - \frac{1}{l+1} = \\
&\mu_{m,n}\pi_{m,n}(1+o(1))\Bigg[\frac{\left(\Theta_m^{(1)} - \frac{1}{l+1}\right)}{\mu_{m,n}} + \sum_{i=2}^{l+1}\frac{1}{\mu_{m,n}}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\lambda_{k-1}^{i}(\Theta_{k-1})\left(\Theta_{k-1}^{(i)} - \frac{1}{l+1}\right) \\
&+ \sum_{i=l+2}^{d}\frac{1}{\mu_{m,n}}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\nu_{k-1}^{i}(\Theta_{k-1})\Theta_{k-1}^{(i)} + \frac{\alpha}{l+1}\sum_{i=l+2}^{d}\frac{1}{\mu_{m,n}}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\Theta_{k-1}^{(i)} \\
&+ \frac{\sqrt{l+o(1)}}{l+1}N_{m,n}\Bigg]
\end{aligned} \tag{8.21}$$

as $m \longrightarrow \infty$. The aim of this step is to prove that, apart from the $\mathscr{F}_m$-independent standard normal term and the $\mathscr{F}_m$-measurable ones, every other $\mathscr{F}_m$-non-measurable term in the right-hand side of (8.21) vanishes, almost surely, on $\mathcal{H}_m^l$ as $m \longrightarrow \infty$, uniformly in $n \geq m^2$. The way this argument will be carried out depends on Assumption R, that is depending on whether $\rho_n := \sigma_{n+1}/\tau_n$ is bounded or it diverges to infinity. We recall that, intuitively, we understand this assumption as distinguishing between slow growth of the first regime (since $\rho_n < \rho$ implies $\sigma_n + 1 < \rho\tau_n$ and so $\tau_n < \tau_{n+1} = \tau_n + \sigma_{n+1} < (\rho+1)\tau_n$, which implies $\tau_{n+1} \asymp \tau_n$) and fast growth of the second regime (since $\rho_n \longrightarrow \infty$ implies that $\sigma_{n+1} \sim \tau_{n+1}$, as it will be shown later, $\tau_n = o(\sigma_{n+1})$, and it follows that $\tau_n = o(\tau_{n+1})$). To be precise, the argument for slow growth, which relies on Lemma E.2, also applies to fast growth, when $\lambda < 1$ and $\theta = 0$, through Lemma F.1. Indeed it is based on the fact that the vanishing bins will almost surely get a finite amount of balls. This may not happen, in general, for fast growth. Thus, even if for the particular case of fast growth just mentioned it would be possible to proceed in a similar way, the argument used for fast growth as a whole, will not need any such additional assumptions, which would lead to loss of generality. Instead, we will exploit (5.3) on the vanishing components of Θ_n, to modify (8.21) in such a way, so as to exploit how small these vanishing bins' random fluctuations' almost sure bound, provided by Lemma E.3, becomes, compared to the size of the other terms in the summation. Let us first define some quantities to reduce the length of (8.21): for all $i \in \{2, \ldots, l+1\}$,

$$R_{m,n}^{(i)} := \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \lambda_{k-1}^{i}(\Theta_{k-1}) \left(\Theta_{k-1}^{(i)} - \frac{1}{l+1} \right),$$

while for all $i \in \{l+2, \ldots, d\}$,

$$S_{m,n}^{(i)} := \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \nu_{k-1}^{i}(\Theta_{k-1}) \Theta_{k-1}^{(i)}$$

and

$$T_{m,n}^{(i)} := \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \Theta_{k-1}^{(i)}.$$

Denote $\epsilon_{m,n} := \mu_{m,n}\pi_{m,n}$ and define the $\mathscr{F}_m$-measurable term

$$A_{m,n} := \frac{\left(\Theta_m^{(1)} - \frac{1}{l+1}\right)}{\mu_{m,n}}.$$

Then we can rewrite (8.21) as

$$\Theta_n^{(1)} - \frac{1}{l+1} = (1+o(1))\epsilon_{m,n} \left[A_{m,n} + \sum_{i=2}^{l+1} R_{m,n}^{(i)} + \sum_{i=l+2}^{d} S_{m,n}^{(i)} + \frac{\alpha}{l+1} \sum_{i=l+2}^{d} T_{m,n}^{(i)} + \frac{\sqrt{l+o(1)}}{l+1} N_{m,n} \right]. \tag{8.22}$$

The idea of what we are going to do now is, in both cases arising from Assumption R, whenever necessary, to split the measurable term of the series (or something very similar to it), carrying most of the size, from its nonmeasurable remainder, obtaining an $\mathscr{F}_m$-non-measurable tail of the series, which is then proved to be almost surely vanishing on the event considered.

Case 1. If ρ_n is bounded, denote with ρ the positive constant such that $\rho_n \le \rho$ for all n, and start by showing that $R^{(i)}_{m,n}$ and $S^{(i)}_{m,n}$ vanish uniformly. Since

$$|R^{(i)}_{m,n}| \le \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \mathcal{O}(\delta^2_{k-1})$$

uniformly on $\mathcal{H}^l_m$ by (8.8), for some positive constant C and m large enough we have that

$$|R^{(i)}_{m,n}| \le \frac{C}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k}\delta^2_{k-1} = \frac{C}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k\tau^{2r}_{k-1}}.$$

Since

$$\frac{(1+\rho_{k-1})^{2r}}{\tau^{2r}_k} = \frac{1}{\tau^{2r}_{k-1}},$$

it follows that

$$|R^{(i)}_{m,n}| \le \frac{C}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k\tau^{2r}_{k-1}} = \frac{C}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{(1+\rho_{k-1})^{2r}}{\tau^{2r}_k} \le$$
$$\frac{C(1+\rho)^{2r}}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau^{1+2r}_k} \le \frac{C(1+\rho)^{2r}}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\tau^{1+2r}_k} \le \frac{C(1+\rho)^{2r}}{\mu_{m,n}} \int_{\tau_m}^{\infty} \frac{dx}{x^{1+2r}},$$

because $\pi_{m,k} > 1$, so

$$|R^{(i)}_{m,n}| \le \frac{C(1+\rho)^{2r}}{2r\mu_{m,n}} \frac{1}{\tau^{2r}_m}. \tag{8.23}$$

We need to show that this upper bound vanishes, and this requires a lower bound on $\mu_{m,n}$. Note first that

$$\pi_{m,k} = \exp \sum_{j=m+1}^{k} \log\left(1 + (\alpha-1)\frac{\sigma_j}{\tau_j}\right) \le \exp\left\{(\alpha-1)\sum_{j=m+1}^{k} \frac{\sigma_j}{\tau_j}\right\} \le$$
$$\exp\left\{(\alpha-1)\int_{\tau_m}^{\tau_k} \frac{dx}{x}\right\} = \exp\{(\alpha-1)(\log\tau_k - \log\tau_m)\} = \left(\frac{\tau_k}{\tau_m}\right)^{\alpha-1}.$$

Since

$$\frac{1}{\tau^{2\alpha}_k} = \frac{1}{(1+\rho_{k-1})^{2\alpha}\tau^{2\alpha}_{k-1}} > \frac{1}{(1+\rho)^{2\alpha}\tau^{2\alpha}_{k-1}},$$

it follows that

$$\mu_{m,n} \ge \sqrt{\sum_{k=m+1}^{n} \frac{\sigma_k}{\left(\frac{\tau_k}{\tau_m}\right)^{2(\alpha-1)} \tau^2_k}} = \tau^{\alpha-1}_m \sqrt{\sum_{k=m+1}^{n} \frac{\sigma_k}{\tau^{2\alpha}_k}} > \frac{\tau^{\alpha-1}_m}{(1+\rho)^\alpha} \sqrt{\sum_{k=m+1}^{n} \frac{\sigma_k}{\tau^{2\alpha}_{k-1}}}$$
$$\ge \frac{\tau^{\alpha-1}_m}{(1+\rho)^\alpha} \sqrt{\int_{\tau_m}^{\tau_n} \frac{dx}{x^{2\alpha}}} = \frac{\tau^{\alpha-1}_m}{(1+\rho)^\alpha} \sqrt{\frac{1}{2\alpha-1}\left(\frac{1}{\tau^{2\alpha-1}_m} - \frac{1}{\tau^{2\alpha-1}_n}\right)} =$$
$$\frac{\tau^{\alpha-1}_m}{(1+\rho)^\alpha\sqrt{2\alpha-1}} \frac{1}{\tau^{\alpha-\frac{1}{2}}_m} \sqrt{1 - \left(\frac{\tau_m}{\tau_n}\right)^{2\alpha-1}},$$

which yields

$$\mu_{m,n} \ge \frac{1+o(1)}{(1+\rho)^\alpha\sqrt{2\alpha-1}\sqrt{\tau_m}}, \tag{8.24}$$

since $\tau_k \longrightarrow \infty$ monotonically and $m = o(n)$ (usual abuse of notation), thus $\tau_m = o(\tau_n)$. Since $r_\alpha < r < 1/2$ implies $1/4 < r < 1/2$, which ensures $2r > 1/2$, one gets, by plugging (8.24) into (8.23), that as $m \longrightarrow \infty$, uniformly on $\mathcal{H}^l_m$ and $n \geq m^2$

$$|R^{(i)}_{m,n}| \leq \frac{C(1+\rho)^{2r}}{2r\mu_{m,n}} \frac{1}{\tau_m^{2r}} \leq \frac{(1+o(1))C(1+\rho)^{\alpha+2r}\sqrt{2\alpha-1}\sqrt{\tau_m}}{2r\tau_m^{2r}} \longrightarrow 0.$$

Proceed similarly for $S^{(i)}_{m,n}$ by using (8.9) instead. This yields

$$S^{(i)}_{m,n} = \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \mathcal{O}(\delta^\alpha_{k-1}),$$

so for some other constant, which we will keep denoting informally as C, and by using (8.24) again, it follows that

$$\begin{aligned}
|S^{(i)}_{m,n}| &\leq \frac{C}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k\tau^{\alpha r}_{k-1}} = \frac{C}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{(1+\rho_{k-1})^{\alpha r}}{\tau_k^{\alpha r}} \leq \\
&\frac{C(1+\rho)^{\alpha r}}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k^{1+\alpha r}} \leq \frac{C(1+\rho)^{\alpha r}}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\tau_k^{1+\alpha r}} \leq \frac{C(1+\rho)^{\alpha r}}{\mu_{m,n}} \int_{\tau_m}^{\infty} \frac{dx}{x^{1+\alpha r}} \\
&= \frac{C(1+\rho)^{\alpha r}}{\alpha r} \frac{1}{\mu_{m,n}\tau_m^{\alpha r}} \leq \frac{C(1+\rho)^{\alpha r}}{\alpha r} \frac{(1+o(1))(1+\rho)^{\alpha}\sqrt{2\alpha-1}\sqrt{\tau_m}}{\tau_m^{\alpha r}} \longrightarrow 0,
\end{aligned}$$

since $r_\alpha < r < 1/2$ ensures that $\alpha r > 1/2$.

Lastly, for the last series $T^{(i)}_{m,n}$, consider that for every m, $\mathcal{H}^l_m \subseteq \{\forall\, l+2 \leq i \leq d,\ \Theta^{(i)}_j \longrightarrow 0, \text{ as } j \longrightarrow \infty\}$, and therefore by Lemma E.2, for almost every $\omega \in \mathcal{H}^l_m$, $\{T^{(i)}_j(\omega)\}_i$ are bounded for all $l+2 \leq i \leq d$. Since $\{T^{(i)}_j(\omega)\}_i$ are nondecreasing (in j) and integer valued, they are bounded for all $l+2 \leq i \leq d$ if and only if there is $M = M(\omega) \in \mathbb{N}$ such that for all $l+2 \leq i \leq d$ and for all $j \geq k \geq M$, $T^{(i)}_j(\omega) - T^{(i)}_k(\omega) = 0$. Thus if we rewrite, by adding and subtracting the same term,

$$\begin{aligned}
T^{(i)}_{m,n} &= \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \Theta^{(i)}_{k-1} - \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T^{(i)}_m}{\tau_{k-1}} + \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T^{(i)}_m}{\tau_{k-1}} \\
&= \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T^{(i)}_{k-1} - T^{(i)}_m}{\tau_{k-1}} + \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T^{(i)}_m}{\tau_{k-1}},
\end{aligned}$$

since for all $m \geq M$,

$$T^{(i)}_{m,n} = \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T^{(i)}_m}{\tau_{k-1}} \in \mathrm{m}\mathscr{F}_m,$$

this term will not affect the asymptotic distribution of $N_{m,n}$ conditionally on $\mathscr{F}_m$, as it will be eventually known (in the worst case, except for a negligible event) as $m \longrightarrow \infty$. Thus

$$T^{(i)}_{m,n} = \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T^{(i)}_m}{\tau_{k-1}} + o_\omega(1).$$

In conclusion, following the asymptotic negligibility of $R^{(i)}_{m,n}$ and $S^{(i)}_{m,n}$, and the eventual $\mathscr{F}_m$-measurability of $T^{(i)}_{m,n}$, as $m \longrightarrow \infty$; through the definition of

$$\tau^*_{m,n} := \sum_{i=2}^{l+1} R^{(i)}_{m,n} + \sum_{i=l+2}^{d} S^{(i)}_{m,n} + \frac{\alpha}{l+1} \sum_{i=l+2}^{d} \frac{1}{\mu_{m,n}} \sum_{k=m+1}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T^{(i)}_{k-1} - T^{(i)}_m}{\tau_{k-1}},$$

which vanishes almost surely on $\mathcal{H}_m^l$ as $m \longrightarrow \infty$; (8.22) can be rewritten as

$$\Theta_n^{(1)} - \frac{1}{l+1} = (1+o(1))\epsilon_{m,n}\Bigg[A_{m,n} + \frac{\alpha}{l+1}\sum_{i=l+2}^{d}\frac{1}{\mu_{m,n}}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\frac{T_m^{(i)}}{\tau_{k-1}} + \frac{\sqrt{l+o(1)}}{l+1}N_{m,n} + \tau_{m,n}^*\Bigg],$$

and therefore by defining a $\mathscr{F}_m$-measurable term

$$A_{m,n}^* := A_{m,n} + \frac{\alpha}{l+1}\sum_{i=l+2}^{d}\frac{1}{\mu_{m,n}}\sum_{k=m+1}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\frac{T_m^{(i)}}{\tau_{k-1}}$$

and

$$\eta_{m,n}^* := \mathbb{1}_{\mathcal{H}_m^l}\tau_{m,n}^*,$$

in the case when ρ_n is bounded, almost surely on $\mathcal{H}_m^l$ as $m \longrightarrow \infty$, uniformly in $n > m$, (8.22) takes the form

$$\Theta_n^{(1)} - \frac{1}{l+1} = (1+o(1))\epsilon_{m,n}\left[A_{m,n}^* + \frac{\sqrt{l+o(1)}}{l+1}N_{m,n} + \eta_{m,n}^*\right], \tag{8.25}$$

where $\eta_{m,n}^* \longrightarrow 0$ almost surely on the probability space.

Case 2. If $\rho_n \longrightarrow \infty$, note that then $\sigma_n \sim \tau_n$, since

$$\frac{\tau_n}{\sigma_n} = 1 + \frac{\tau_{n-1}}{\sigma_n} = 1 + \frac{1}{\rho_{n-1}} \longrightarrow 1$$

and $\sigma_{n-1} = o(\sigma_n)$, as

$$\frac{\sigma_{n-1}}{\sigma_n} \le \frac{\tau_{n-1}}{\sigma_n} = \frac{1}{\rho_{n-1}} \longrightarrow 0.$$

This gives us rather detailed information about the asymptotic behaviour of the deterministic scales. As to $\pi_{m,k}$, we can show that there are constant $1 < p < q$ such that, for all m large enough,

$$p^{k-m} \le \pi_{m,k} \le q^{k-m}. \tag{8.26}$$

We can see the upper bound by using $\log(1+x) \le x$ for all $x > 1$ and $\sigma_j < \tau_j$. Then

$$\pi_{m,k} = \exp\sum_{j=m+1}^{k}\log\left(1+(\alpha-1)\frac{\sigma_j}{\tau_j}\right) \le \exp\left\{(\alpha-1)\sum_{j=m+1}^{k}\frac{\sigma_j}{\tau_j}\right\}$$
$$\le \exp\{(\alpha-1)\sum_{j=m+1}^{k}1\} = \exp\{(\alpha-1)(k-m)\} = q^{k-m},$$

where $q := \exp(\alpha-1) > 1$. We can see the lower bound, by using $\log(1+x) \ge x/(1+x)$, along with the fact that since $\sigma_j \sim \tau_j$, for m large enough $\sigma_j/\tau_j > 1/\alpha$, so

$$\pi_{m,k} \ge \exp\left\{(\alpha-1)\sum_{j=m+1}^{k}\frac{\frac{\sigma_j}{\tau_j}}{1+(\alpha-1)\frac{\sigma_j}{\tau_j}}\right\} \ge \exp\left\{(\alpha-1)\sum_{j=m+1}^{k}\frac{\frac{\sigma_j}{\tau_j}}{1+(\alpha-1)}\right\}$$
$$= \exp\left\{\frac{\alpha-1}{\alpha}\sum_{j=m+1}^{k}\frac{\sigma_j}{\tau_j}\right\} > \exp\left\{\frac{\alpha-1}{\alpha^2}(k-m)\right\} = p^{k-m},$$

where $p := \exp[(\alpha-1)/\alpha^2] > 1$. As to $\mu_{m,n}$, in this case it is enough to note that, by keeping only the first term of $\mu_{m,n}$,

$$\mu_{m,n} \geq \frac{\sqrt{\sigma_{m+1}}}{\pi_{m,m+1}\tau_{m+1}} \sim \frac{1}{\alpha\sqrt{\tau_{m+1}}} > \frac{1}{(\alpha+1)\sqrt{\tau_{m+1}}}, \tag{8.27}$$

since in this regime, having $\tau_m = o(\tau_{m+1})$ and $\sigma_{m+1} \sim \tau_{m+1}$,

$$\pi_{m,m+1} = \frac{\tau_m + \alpha\sigma_{m+1}}{\tau_{m+1}} \longrightarrow \alpha. \tag{8.28}$$

We now move on to studying the three series as in the previous scenario. Consider first that

$$R^{(i)}_{m,n} = r^{(i)}_{m,n} + \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \lambda^i_{k-1}(\Theta_{k-1}) \left(\Theta^{(i)}_{k-1} - \frac{1}{l+1} \right),$$

where

$$r^{(i)}_{m,n} = \frac{\sigma_{m+1}}{\mu_{m,n}\pi_{m,m+1}\tau_{m+1}} \lambda^i_m(\Theta_m) \left(\Theta^{(i)}_m - \frac{1}{l+1} \right) \in \mathrm{m}\mathscr{F}_m.$$

The necessity of keeping the measurable term will be clear from the argument showing that the $\mathscr{F}_m$-non-measurable part vanishes on $\mathcal{H}^l_m$. This follows since for some constant, which we denote always by C, by (8.8), (8.26) and (8.27) and $\sqrt{\sigma_{m+1}} \sim \sqrt{\tau_{m+1}}$,

$$\begin{aligned}
&\left| \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \lambda^i_{k-1}(\Theta_{k-1}) \left(\Theta^{(i)}_{k-1} - \frac{1}{l+1} \right) \right| \leq \frac{C}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \delta^2_{k-1} \\
&= \frac{C}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k\tau^{2r}_{k-1}} \leq \frac{C\pi_{m,m+1}\tau_{m+1}}{\sqrt{\sigma_{m+1}}\tau^{2r}_{m+1}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \leq \frac{C(\alpha+1)\tau_{m+1}}{\tau^{2r+\frac{1}{2}}_{m+1}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \\
&\leq \frac{C(\alpha+1)\tau_{m+1}}{\tau^{2r+\frac{1}{2}}_{m+1}} \sum_{k=m+2}^{n} \frac{1}{p^{k-m}} \leq \frac{C(\alpha+1)p}{(p-1)\tau^{2r-\frac{1}{2}}_{m+1}} \longrightarrow 0
\end{aligned}$$

as $m \longrightarrow \infty$, uniformly in $n > m$ (since $r > r_\alpha$ ensures that $2r + 1/2 > 1$) and uniformly on $\mathcal{H}^l_m$. Therefore on $\mathcal{H}^l_m$,

$$R^{(i)}_{m,n} = r^{(i)}_{m,n} + o(1).$$

The reason for extracting the measurable term is to avoid having the ratio $\sqrt{\tau_{m+1}}/\tau^{2r}_m$ at the end of the second line in the estimates above. This ratio would be problematic, at this regime of growth. Similarly,

$$S^{(i)}_{m,n} = s^{(i)}_{m,n} + \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \nu^i_{k-1}(\Theta_{k-1})\Theta^{(i)}_{k-1},$$

where

$$s^{(i)}_{m,n} := \frac{1}{\mu_{m,n}} \frac{\sigma_{m+1}}{\pi_{m,m+1}\tau_k} \nu^i_m(\Theta_m)\Theta^{(i)}_m \in \mathrm{m}\mathscr{F}_m.$$

By a similar argument, since for some constant, which we denote again by C, by (8.9), (8.26) and (8.27) and the fact that $\sqrt{\sigma_{m+1}} \sim \sqrt{\tau_{m+1}}$, we have that

$$\begin{aligned}
&\frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \nu^i_{k-1}(\Theta_{k-1})\Theta^{(i)}_{k-1} \leq \frac{C}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \delta^\alpha_{k-1} \\
&= \frac{C}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k\tau^{2r}_{k-1}} \leq \frac{C\pi_{m,m+1}\tau_{m+1}}{\sqrt{\sigma_{m+1}}\tau^{\alpha r}_{m+1}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \leq C(\alpha+1)\frac{\tau_{m+1}}{\tau^{\alpha r+\frac{1}{2}}_{m+1}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \\
&\leq C(\alpha+1)\frac{\tau_{m+1}}{\tau^{\alpha r+\frac{1}{2}}_{m+1}} \sum_{k=m+2}^{n} \frac{1}{p^{k-m}} \leq \frac{C(\alpha+1)p}{(p-1)\tau^{\alpha r-\frac{1}{2}}_{m+1}} \longrightarrow 0
\end{aligned}$$

as $m \longrightarrow \infty$ uniformly in $n > m$ (since $r > r_\alpha$ ensures that $\alpha r + 1/2 > 1$), and uniformly on $\mathcal{H}^l_m$. Thus

$$S^{(i)}_{m,n} = s^{(i)}_{m,n} + o(1).$$

As to $T^{(i)}_{m,n}$, this term is more problematic than the others, also in the fast growth scenario. First consider the series

$$\sum_{j=1}^{\infty} \sigma_j P^{(i)}_{j-1},$$

and partition the event $\mathcal{H}^l_m$ into

$$\mathscr{E}^{(i)}_0 := \mathcal{H}^l_m \cap \left\{ \sum_{j=1}^{\infty} \sigma_j P^{(i)}_{j-1} < \infty \right\}$$

and

$$\mathscr{E}^{(i)}_\infty := \mathcal{H}^l_m \cap \left\{ \sum_{j=1}^{\infty} \sigma_j P^{(i)}_{j-1} = \infty \right\}.$$

Rewrite the series

$$T^{(i)}_{m,n} = \frac{1}{\mu_{m,n}} \frac{\sigma_{m+1}}{\pi_{m,m+1}\tau_{m+1}} \Theta^{(i)}_m + \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \Theta^{(i)}_{k-1}.$$

- On $\mathscr{E}^{(i)}_0$ define

$$S^{(i)} := \sum_{j=1}^{\infty} \sigma_j P^{(i)}_{j-1} \in \mathbb{R}.$$

Consider that by (5.2) multiplied both side by τ_{n+1} we get

$$T^{(i)}_{n+1} = T^{(i)}_n + \varepsilon^{(i)}_{n+1} \sqrt{\sigma_{n+1} P^{(i)}_n (1 - P^{(i)}_n)} + \sigma_{n+1} P^{(i)}_n,$$

so iterating this equation for every $k > m+1$ yields

$$T^{(i)}_{k-1} = T^{(i)}_m + \sum_{j=m+1}^{k-1} \varepsilon^{(i)}_j \sqrt{\sigma_j P^{(i)}_{j-1} (1 - P^{(i)}_{j-1})} + \sum_{j=m+1}^{k-1} \sigma_j P^{(i)}_{j-1}.$$

By Lemma E.3 and (8.1), almost surely on $\mathscr{E}^{(i)}_0$, for m large enough, we have that

$$\left| \sum_{j=m+1}^{k-1} \varepsilon^{(i)}_j \sqrt{\sigma_j P^{(i)}_{j-1} (1 - P^{(i)}_{j-1})} \right| \le \sqrt{\sigma_{k-1}} \sum_{j=m+1}^{k-1} j \sqrt{P^{(i)}_{j-1}} \le$$

$$d^{\frac{\alpha-1}{2}} \sqrt{\sigma_{k-1}} \sum_{j=m+1}^{k-1} j (\Theta^{(i)}_{j-1})^{\frac{\alpha}{2}} \le d^{\frac{\alpha-1}{2}} \sqrt{\sigma_{k-1}} \sum_{j=m+1}^{k-1} \frac{j}{\tau_{j-1}^{\frac{\alpha r}{2}}} \le d^{\frac{\alpha-1}{2}} \frac{\sqrt{\sigma_{k-1}}}{\tau_m^{\frac{\alpha r}{4}}} \sum_{j=m+1}^{k-1} \frac{j}{\tau_{j-1}^{\frac{\alpha r}{4}}},$$

since, having $\sigma_j = o(\sigma_{j+1})$, eventually $\sigma_{j+1} > \sigma_j$ and therefore, for m large enough, $\sigma_{k-1} > \sigma_{k-2} > \ldots > \sigma_{m+1}$. By the ratio test

$$\sum_{j=m+1}^{\infty} \frac{j}{\tau_{j-1}^{\frac{\alpha r}{4}}} < \infty,$$

as in this regime of fast growth

$$\frac{\frac{j+1}{\tau_j^{\frac{\alpha r}{4}}}}{\frac{j}{\tau_{j-1}^{\frac{\alpha r}{4}}}} = \frac{j+1}{j}\left(\frac{\tau_{j-1}}{\tau_j}\right)^{\frac{\alpha r}{4}} \longrightarrow 0 < 1.$$

Being a series of positive terms, there will be some $\Gamma > 0$ such that

$$\sum_{j=m+1}^{\infty} \frac{j}{\tau_{j-1}^{\frac{\alpha r}{4}}} < \Gamma,$$

and therefore

$$\left|\sum_{j=m+1}^{k-1} \varepsilon_j^{(i)} \sqrt{\sigma_j P_{j-1}^{(i)}(1-P_{j-1}^{(i)})}\right| \leq \Gamma d^{\frac{\alpha-1}{2}} \frac{\sqrt{\sigma_{k-1}}}{\tau_m^{\frac{\alpha r}{4}}}. \tag{8.29}$$

It is now possible to prove that the nonmeasurable tail of $T_{m,n}^{(i)}$ almost surely vanishes on this event. First rewrite this tail as follows:

$$\frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \Theta_{k-1}^{(i)} =$$

$$\frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T_m^{(i)} + \sum_{j=m+1}^{k-1} \varepsilon_j^{(i)} \sqrt{\sigma_j P_{j-1}^{(i)}(1-P_{j-1}^{(i)})} + \sum_{j=m+1}^{k-1} \sigma_j P_{j-1}^{(i)}}{\tau_{k-1}} =$$

$$= \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T_m^{(i)}}{\tau_{k-1}} + \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\sum_{j=m+1}^{k-1} \varepsilon_j^{(i)} \sqrt{\sigma_j P_{j-1}^{(i)}(1-P_{j-1}^{(i)})}}{\tau_{k-1}}$$

$$+ \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\sigma_{m+1} P_m^{(i)}}{\tau_{k-1}} + \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\sum_{j=m+2}^{k-1} \sigma_j P_{j-1}^{(i)}}{\tau_{k-1}},$$

then by the previous estimates, as $m \longrightarrow \infty$, uniformly in $n > m$, by (8.26), (8.27) and (8.29), almost surely on $\mathscr{E}_0^{(i)}$ we have that

$$\left|\frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\sum_{j=m+1}^{k-1} \varepsilon_j^{(i)} \sqrt{\sigma_j P_{j-1}^{(i)}(1-P_{j-1}^{(i)})}}{\tau_{k-1}}\right|$$

$$\leq \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\left|\sum_{j=m+1}^{k-1} \varepsilon_j^{(i)} \sqrt{\sigma_j P_{j-1}^{(i)}(1-P_{j-1}^{(i)})}\right|}{\tau_{k-1}}$$

$$\leq \frac{\Gamma d^{\frac{\alpha-1}{2}}}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\frac{\sqrt{\sigma_{k-1}}}{\tau_m^{\frac{\alpha r}{4}}}}{\tau_{k-1}} \leq \frac{\Gamma d^{\frac{\alpha-1}{2}} \pi_{m,m+1}\tau_{m+1}}{\sqrt{\sigma_{m+1}}\tau_m^{\frac{\alpha r}{4}}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\sqrt{\sigma_{k-1}}}{\tau_{k-1}}$$

$$\leq \Gamma d^{\frac{\alpha-1}{2}}(\alpha+1)\frac{\sqrt{\tau_{m+1}}}{\tau_m^{\frac{\alpha r}{4}}} \sum_{k=m+2}^{n} \frac{1}{\pi_{m,k}} \frac{1}{\sqrt{\tau_{k-1}}} \leq \Gamma d^{\frac{\alpha-1}{2}}(\alpha+1)\frac{\sqrt{\tau_{m+1}}}{\sqrt{\tau_{m+1}}\tau_m^{\frac{\alpha r}{4}}} \sum_{k=m+2}^{n} \frac{1}{p^{k-m}}$$

$$\leq \frac{\Gamma d^{\frac{\alpha-1}{2}}(\alpha+1)p}{(p-1)\tau_m^{\frac{\alpha r}{4}}} \longrightarrow 0.$$

This term's estimate holds only almost surely on the event, so it will contribute with a $\mathcal{O}_\omega(1)$. Similarly, recalling that

$$\sum_{j=m+1}^{k-1} \sigma_j P_{j-1}^{(i)} \le S^{(i)},$$

we have that, by (8.26) and (8.27),

$$\frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\sum_{j=m+2}^{k-1} \sigma_j P_{j-1}^{(i)}}{\tau_{k-1}} \le \frac{S^{(i)} \pi_{m,m+1}\tau_{m+1}}{\sqrt{\sigma_{m+1}}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{1}{\tau_{k-1}}$$
$$\le (\alpha+1) S^{(i)} \sqrt{\tau_{m+1}} \sum_{k=m+2}^{n} \frac{1}{\pi_{m,k}} \frac{1}{\tau_{k-1}} \le \frac{(\alpha+1)S^{(i)}}{\sqrt{\tau_{m+1}}} \sum_{k=m+2}^{n} \frac{1}{p^{k-m}} \le \frac{(\alpha+1)S^{(i)} p}{(p-1)\sqrt{\tau_{m+1}}}$$
$$\longrightarrow 0.$$

This term has, in the estimate, $S^{(i)} = S^{(i)}(\omega)$, so it also contributes with a $\mathcal{O}_\omega(1)$. In conclusion as $m \longrightarrow \infty$, on $\mathscr{E}_0^{(i)}$, uniformly in $n > m$,

$$\frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \Theta_{k-1}^{(i)} = \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T_m^{(i)}}{\tau_{k-1}} + \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\sigma_{m+1} P_m^{(i)}}{\tau_{k-1}}$$
$$+ \mathcal{O}_\omega(1),$$

and therefore, the decomposition into $\mathscr{F}_m$-measurable term and vanishing $\mathscr{F}_m$-non-measurable term is

$$T_{m,n}^{(i)} = t_{m,n}^{(i)} + \mathcal{O}_\omega(1),$$

where

$$t_{m,n}^{(i)} := \frac{1}{\mu_{m,n}} \frac{\sigma_{m+1}}{\pi_{m,m+1}\tau_{m+1}} \Theta_m^{(i)} + \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{T_m^{(i)}}{\tau_{k-1}} + \frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\sigma_{m+1} P_m^{(i)}}{\tau_{k-1}}.$$

- On $\mathscr{E}_\infty^{(i)}$ most of what holds on $\mathscr{E}_0^{(i)}$ still applies, except for the existence of the random variable $S^{(i)} \in \mathbb{R}$. Hence the only difference concerns the proof that the term

$$\frac{1}{\mu_{m,n}} \sum_{k=m+2}^{n} \frac{\sigma_k}{\pi_{m,k}\tau_k} \frac{\sum_{j=m+2}^{k-1} \sigma_j P_{j-1}^{(i)}}{\tau_{k-1}}$$

vanishes. The key fact here is that, on $\mathscr{E}_\infty^{(i)}$, not only $\Theta_k^{(i)}$ vanishes, but it decays fast. Specifically,

$$\Theta_{k+1}^{(i)} = \mathcal{O}_\omega(\Theta_k^{(i)}). \tag{8.30}$$

This follows from a relatively simple fact. Recall that in the conclusion of *Step 2*, in Lemma 6.1, the angle bracket process argument showed that for almost all ω, such that

$$\sum_{j=1}^{k} \sigma_j P_{j-1}^{(i)} \longrightarrow \infty,$$

we have that

$$T_k^{(i)} \sim_\omega \sum_{j=1}^{k} \sigma_j P_{j-1}^{(i)}.$$

But then, by (8.1), which implies $P_k^{(i)} \asymp (\Theta_k^{(i)})^\alpha$, and $\sigma_k \sim \tau_k$, it follows that

$$\frac{\Theta_{k+1}^{(i)}}{\Theta_k^{(i)}} = \frac{T_{k+1}^{(i)}}{T_k^{(i)}}\frac{\tau_k}{\tau_{k+1}} \sim_\omega \frac{\sum_{j=1}^{k+1}\sigma_j P_{j-1}^{(i)}}{\sum_{j=1}^{k}\sigma_j P_{j-1}^{(i)}}\frac{\tau_k}{\tau_{k+1}} = \left(1 + \frac{\sigma_{k+1}P_k^{(i)}}{\sum_{j=1}^{k}\sigma_j P_{j-1}^{(i)}}\right)\frac{\tau_k}{\tau_{k+1}}$$
$$\sim_\omega \left(1 + \frac{\sigma_{k+1}P_k^{(i)}}{T_k^{(i)}}\right)\frac{\tau_k}{\tau_{k+1}} = \frac{\tau_k}{\tau_{k+1}} + \frac{\sigma_{k+1}\tau_k P_k^{(i)}}{\tau_{k+1}T_k^{(i)}} \le \frac{\tau_k}{\tau_{k+1}} + \frac{\tau_k P_k^{(i)}}{T_k^{(i)}} \longrightarrow 0,$$

since $\tau_k = o(\tau_{k+1})$ and

$$\frac{\tau_k P_k^{(i)}}{T_k^{(i)}} \asymp \frac{\tau_k (\Theta_k^{(i)})^\alpha}{T_k^{(i)}} = (\Theta_k^{(i)})^{\alpha-1} \longrightarrow 0.$$

By (8.30), for m large enough, $\Theta_{k+1}^{(i)} < \Theta_k^{(i)}$ for all $k > m$, in particular $\Theta_{k-1}^{(i)} < \Theta_{k-2}^{(i)} < \ldots < \Theta_{m+1}^{(i)}$. Then consider that, having adopted empty sum convention when splitting the sums, for $k = m+2$,

$$\sum_{j=m+2}^{k-1}\sigma_j P_{j-1}^{(i)} = 0.$$

Therefore, by (8.26) and (8.27) and the almost sure eventual monotonicity of Θ_{j-1}, we have that

$$\frac{1}{\mu_{m,n}}\sum_{k=m+2}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\frac{\sum_{j=m+2}^{k-1}\sigma_j P_{j-1}^{(i)}}{\tau_{k-1}} = \frac{1}{\mu_{m,n}}\sum_{k=m+3}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\frac{\sum_{j=m+2}^{k-1}\sigma_j P_{j-1}^{(i)}}{\tau_{k-1}} \le$$
$$\frac{d^{\alpha-1}}{\mu_{m,n}}\sum_{k=m+3}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\frac{\sum_{j=m+2}^{k-1}\sigma_j (\Theta_{j-1}^{(i)})^\alpha}{\tau_{k-1}} \le \frac{d^{\alpha-1}}{\mu_{m,n}}\sum_{k=m+3}^{n}\frac{\sigma_k}{\pi_{m,k}\tau_k}\frac{\sum_{j=m+2}^{k-1}\sigma_{k-1} (\Theta_{m+1}^{(i)})^\alpha}{\tau_{k-1}}$$
$$= \frac{d^{\alpha-1}(\Theta_{m+1}^{(i)})^\alpha}{\mu_{m,n}}\sum_{k=m+3}^{n}\frac{(k-m-2)\sigma_k\sigma_{k-1}}{\pi_{m,k}\tau_k\tau_{k-1}} \le \frac{d^{\alpha-1}\pi_{m,m+1}\tau_{m+1}}{\sqrt{\sigma_{m+1}}\tau_{m+1}^{\alpha r}}\sum_{k=m+3}^{n}\frac{k-m-2}{p^{k-m}}$$
$$\le (\alpha+1)d^{\alpha-1}\frac{\sqrt{\tau_{m+1}}}{\tau_{m+1}^{\alpha r}}\sum_{k=m+3}^{n}\frac{(k-m-2)}{p^{k-m}} < \frac{(\alpha+1)d^{\alpha-1}\sqrt{\tau_{m+1}}}{p(p-1)^2\tau_{m+1}^{\alpha r}} \longrightarrow 0,$$

having used, in the conclusion, also that: we can consider m large enough such that $\sigma_{k-1} > \sigma_{k-2} > \ldots > \sigma_{m+2}$; $r > r_\alpha$, ensuring that $\alpha r > {}^1\!/_2$; the series

$$\sum_{k=m+3}^{n}\frac{(k-m-2)}{p^{k-m}} < \sum_{j=0}^{\infty}\frac{j}{p^{j+2}} = \frac{1}{p^3}\sum_{j=1}^{\infty}\frac{j}{p^{j-1}} = \frac{1}{p^3\left(1-\frac{1}{p}\right)^2} = \frac{1}{p(p-1)^2}.$$

Hence also on $\mathscr{E}_\infty^{(i)}$,

$$T_{m,n}^{(i)} = t_{m,n}^{(i)} + o_\omega(1),$$

which therefore holds on $\mathcal{H}_m^l$ as $m \longrightarrow \infty$ uniformly in $n > m$.

Define

$$\tau_{m,n}^{\circledast} = \sum_{i=2}^{l+1} R_{m,n}^{(i)} - r_{m,n}^{(i)} + \sum_{i=l+2}^{d} S_{m,n}^{(i)} - s_{m,n}^{(i)} + \frac{\alpha}{l+1}\sum_{i=l+2}^{d} T_{m,n}^{(i)} - t_{m,n}^{(i)}.$$

Since $\tau_{m,n}^{\circledast} \longrightarrow 0$ almost surely on $\mathcal{H}_m^l$, since as $m \longrightarrow \infty$, (8.22) can be rewritten as

$$\Theta_n^{(1)} - \frac{1}{l+1} =$$
$$(1+o(1))\epsilon_{m,n}\left[A_{m,n} + \sum_{i=2}^{l+1} r_{m,n}^{(i)} + \sum_{i=l+2}^{d} s_{m,n}^{(i)} + \frac{\alpha}{l+1}\sum_{i=l+2}^{d} t_{m,n}^{(i)} + \frac{\sqrt{l+o(1)}}{l+1}N_{m,n} + \tau_{m,n}\right],$$

defining a new $\mathscr{F}_m$-measurable term

$$A^{\circledast}_{m,n} := A_{m,n} + \sum_{i=2}^{l+1} r^{(i)}_{m,n} + \sum_{i=l+2}^{d} s^{(i)}_{m,n} + \frac{\alpha}{l+1} \sum_{i=l+2}^{d} t^{(i)}_{m,n},$$

and the random term

$$\eta^{\circledast}_{m,n} = \mathbb{1}_{\mathcal{H}^l_m} \tau^{\circledast}_{m,n},$$

we have that as $m \longrightarrow \infty$, also when $\rho_n \longrightarrow \infty$, on $\mathcal{H}^l_m$, uniformly in $n > m$, (8.22) can be rewritten as

$$\Theta^{(1)}_n - \frac{1}{l+1} = (1 + o(1))\epsilon_{m,n} \left[A^{\circledast}_{m,n} + \frac{\sqrt{l + o(1)}}{l+1} N_{m,n} + \eta^{\circledast}_{m,n} \right], \tag{8.31}$$

where $\eta^{\circledast}_{m,n} \longrightarrow 0$ almost surely on the probability space.

Step 5. Fix $n_m \geq m^2$ growing fast enough, such that $\epsilon_{m,n_m} \longrightarrow \infty$ as $m \longrightarrow \infty$. This requirement can be satisfied, and that this is true, is shown after the conclusion of the argument. Note that the iteration formulas (8.25) and (8.31) holding on $\mathcal{H}^l_m$ as $m \longrightarrow \infty$, in the slow growth and fast growth regime respectively, have the same form, so the conclusion in either of the two regimes will follow by the same argument, with a formal exchange of A^*_{m,n_m} and $A^{\circledast}_{m,n_m}$. Since it is notationally lighter, we write the argument explicitly for the slow growth regime, that is, relying on (8.25). Consider that by the triangle inequality

$$\|\Theta_n - E_l\| \geq \left| \Theta^{(1)}_n - \frac{1}{l+1} \right|;$$

that in (8.19) we established that conditionally on $\mathscr{F}_m$, N_{m,n_m} is asymptotically a standard normal, where N_{m,n_m} is independent of $\mathscr{F}_m$; and finally recall that $A^*_{m,n_m} \in \mathrm{m}\mathscr{F}_m$ and η^*_{m,n_m} vanishes almost surely on Ω. Then

$$\begin{aligned}
&\mathbb{P}(\mathcal{H}^l_m) = \mathbb{P}\left(\{\|\Theta_n - E_l\| \leq \delta_n,\ \forall\, n \geq m\}\right) \leq \mathbb{P}\Bigg(\left\{ \left| \Theta^{(1)}_{n_m} - \frac{1}{l+1} \right| \leq \delta_{n_m} \right\} \cap \\
&\left\{ \Theta^{(1)}_{n_m} - \frac{1}{l+1} = (1+o(1))\epsilon_{m,n_m} \left[A^*_{m,n_m} + \frac{\sqrt{l+o(1)}}{l+1} N_{m,n_m} + \eta^*_{m,n_m} \right] \right\} \Bigg) \\
&\leq \mathbb{P}\left((1+o(1))\epsilon_{m,n_m} \left| A^*_{m,n_m} + \frac{\sqrt{l+o(1)}}{l+1} N_{m,n_m} + \eta^*_{m,n_m} \right| \leq \delta_{n_m} \right) \\
&= \mathbb{E}\mathbb{P}_{\mathscr{F}_m}\left((1+o(1)) \left(A^*_{m,n_m} + \frac{\sqrt{l+o(1)}}{l+1} N_{m,n_m} + \eta^*_{m,n_m} \right) \in \left[-\frac{\delta_{n_m}}{\epsilon_{m,n_m}}, \frac{\delta_{n_m}}{\epsilon_{m,n_m}} \right] \right),
\end{aligned}$$

which we now show to be vanishing, as a result of ${}^{\delta_{n_m}}\!/\!_{\epsilon_{m,n_m}} \longrightarrow 0$ and the standard normal density being bounded by ${}^1\!/\!\sqrt{2\pi}$, which bounds, asymptotically, the expression above with the area of a rectangle having vanishing base length. Denote

$$a_m := \frac{l+1}{\sqrt{l+o(1)}} \left(-A^*_{m,n_m} - \frac{\delta_{n_m}}{(1+o(1))\epsilon_{m,n_m}} \right)$$

and

$$b_m := \frac{l+1}{\sqrt{l+o(1)}} \left(-A^*_{m,n_m} + \frac{\delta_{n_m}}{(1+o(1))\epsilon_{m,n_m}} \right).$$

Since ${}^{\delta_{n_m}}/_{\epsilon_{m,n_m}} \longrightarrow 0$, for all m large enough,

$$0 < b_m - a_m < 2(l+2)\frac{\delta_{n_m}}{\epsilon_{m,n_m}} \longrightarrow 0,$$

which follows from $(l+1)/\sqrt{l} < (l+2)$, as the inequality is equivalent to $l^3 + 3l^2 + 2l - 1 > 0$, and $l \geq 1$, thus the cubic, which is positive at $l = 1$ and has derivative $3l^2 + 6l + 2 > 0$ on the positive reals, is positive for all $l \geq 1$. We will show that

$$\mathbb{P}_{\mathscr{F}_m}\left(N_{m,n_m} + \frac{l+1}{\sqrt{l+o(1)}}\eta^*_{m,n_m} \in [a_m, b_m]\right) \longrightarrow 0$$

almost surely, by proving that almost surely

$$\mathbb{E}_{\mathscr{F}_m} e^{it(N_{m,n_m} + \hat{\eta}^*_{m,n_m})} \longrightarrow e^{-\frac{t^2}{2}}, \tag{8.32}$$

having defined

$$\hat{\eta}^*_{m,n_m} := \frac{l+1}{\sqrt{l+o(1)}}\eta^*_{m,n_m},$$

which vanishes almost surely on the probability space as well. First of all consider that, by adding and subtracting $e^{itN_{m,n_m}}$,

$$\mathbb{E}_{\mathscr{F}_m} e^{it(N_{m,n_m} + \hat{\eta}^*_{m,n_m})} = \mathbb{E}_{\mathscr{F}_m} e^{itN_{m,n_m}}\left(e^{it\hat{\eta}^*_{m,n_m}} - 1\right) + \mathbb{E}_{\mathscr{F}_m} e^{itN_{m,n_m}}. \tag{8.33}$$

The second term on the right-hand side of (8.33) converges to $e^{-\frac{t^2}{2}}$ by (8.19). As to the first term, we briefly show that it vanishes. Recall that by a standard estimate, obtained from the complex exponential's Taylor expansion remainder,

$$\left|e^{it\hat{\eta}^*_{m,n_m}} - 1\right| \leq \min\{2, |t\hat{\eta}^*_{m,n_m}|\},$$

which, for every t fixed, almost surely vanishes. Then

$$\left|e^{itN_{m,n_m}}\left(e^{it\hat{\eta}^*_{m,n_m}} - 1\right)\right| \leq \min\{2, |t\hat{\eta}^*_{m,n_m}|\} \longrightarrow 0$$

almost surely. Note that for every t fixed, for all m large enough, $\min\{2, |t\hat{\eta}^*_{m,n_m}|\} \leq 2$ almost surely and therefore, by the *Dominated Convergence Theorem for conditional expectations* (see [16, §4.6 Theorem 4.6.10]), almost surely

$$\mathbb{E}_{\mathscr{F}_m} e^{itN_{m,n_m}}\left(e^{it\hat{\eta}^*_{m,n_m}} - 1\right) \longrightarrow \mathbb{E}_{\mathscr{F}_\infty} 0 = 0.$$

Thus (8.32) follows, from which, by the *conditional Lévy Continuity Theorem* (see for example [3]), one can conclude that

$$\mu^\omega_m \xrightarrow{w} \mathscr{N}(0,1),$$

where μ^ω_m denotes the conditional distribution of $N_{m,n_m} + \hat{\eta}^*_{m,n_m}$ given $\mathscr{F}_m$, which is defined, by the standard theory, as the map

$$\mu^\omega_m(\cdot) := \mu_m(\cdot, \omega) : \mathscr{B}(\mathbb{R}) \times \Omega \longrightarrow [0,1]$$

such that for every $\omega \in \Omega$, $\mu^\omega_m(\cdot) : \mathscr{B}(\mathbb{R}) \longrightarrow [0,1]$ is a probability measure on $\mathscr{B}(\mathbb{R})$; for each $B \in \mathscr{B}(\mathbb{R})$, $\mu^\omega_m(B) = \mathbb{P}_{\mathscr{F}_m}(N_{m,n_m} + \hat{\eta}^*_{m,n_m} \in B)(\omega)$, $\mathbb{P}$-almost surely. This yields that for every fixed $a < b$,

$$\mathbb{P}_{\mathscr{F}_m}\left(N_{m,n_m} + \hat{\eta}^*_{m,n_m} \in [a,b]\right) \longrightarrow \Phi(b) - \Phi(a)$$

almost surely, where Φ denotes the standard normal cumulative distribution function. Since Φ is continuous on $\mathbb{R}$, for almost every fixed ω, the convergence above is uniform on $\mathbb{R}$ by the standard theory (see for example [22, §2.6, Theorem 2.6.1] for the statement, and [41, Satz I] for the first proof given of this elementary result, by Pólya). Therefore, for any given $\varepsilon > 0$, let $M = M_\varepsilon := \max\{M', M''\}$, where $M' = M'_\varepsilon \in \mathbb{N}$ is such that for all $m \geq M'$,

$$\frac{\delta_{n_m}}{\epsilon_{m,n_m}} < \frac{\sqrt{2\pi}}{4(l+2)}\varepsilon$$

and

$$b_m - a_m < 2(l+2)\frac{\delta_{n_m}}{\epsilon_{m,n_m}},$$

while $M'' = M''_\varepsilon$ is such that, for all $m \geq M''$, for all $a, b \in \mathbb{R}$,

$$|\mathbb{P}_{\mathscr{F}_m}\left(N_{m,n_m} + \hat{\eta}^*_{m,n_m} \in [a,b]\right) - (\Phi(b) - \Phi(a))| < \frac{\varepsilon}{2}.$$

Then for all $m \geq M$, and almost every ω fixed,

$$\begin{aligned}
&\mathbb{P}_{\mathscr{F}_m}\left(N_{m,n_m} + \hat{\eta}^*_{m,n_m} \in [a_m, b_m]\right) \leq \\
&|\mathbb{P}_{\mathscr{F}_m}\left(N_{m,n_m} + \hat{\eta}^*_{m,n_m} \in [a_m, b_m]\right) - (\Phi(b_m) - \Phi(a_m))| + \Phi(b_m) - \Phi(a_m) \leq \\
&\frac{\varepsilon}{2} + \frac{1}{\sqrt{2\pi}}(b_m - a_m) < \frac{\varepsilon}{2} + \frac{2(l+2)}{\sqrt{2\pi}}\frac{\delta_{n_m}}{\epsilon_{m,n_m}} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Thus

$$\mathbb{P}_{\mathscr{F}_m}\left(N_{m,n_m} + \hat{\eta}^*_{m,n_m} \in [a_m, b_m]\right) \longrightarrow 0$$

almost surely, yielding the claim, passing to the limit under expectation by the *Bounded Convergence Theorem*:

$$\mathbb{P}(\mathcal{H}^l_m) \leq \mathbb{E}\mathbb{P}_{\mathscr{F}_m}\left(N_{m,n_m} + \hat{\eta}^*_{m,n_m} \in [a_m, b_m]\right) \longrightarrow 0.$$

Lastly, we show that what we assumed so far holds: that n_m can be chosen to grow fast enough, so as to let ϵ_{m,n_m} diverge to infinity, regardless of the regime of growth considered. This follows quite simply from

$$\begin{aligned}
&\pi_{m,n_m}\mu_{m,n_m} \geq \frac{\pi_{m,n_m}\sqrt{\sigma_{m+1}}}{\pi_{m,m+1}\tau_{m+1}} \geq \frac{1}{\alpha\tau_{m+1}} \exp\sum_{j=m+1}^{n_m} \log\left(1 + (\alpha-1)\frac{\sigma_j}{\tau_j}\right) \geq \\
&\frac{1}{\alpha\tau_{m+1}} \exp\sum_{j=m+1}^{n_m} \frac{(\alpha-1)\frac{\sigma_j}{\tau_j}}{1 + (\alpha-1)\frac{\sigma_j}{\tau_j}} \geq \frac{1}{\alpha\tau_{m+1}} \exp\left\{\frac{\alpha-1}{\alpha}\sum_{j=m+1}^{n_m}\frac{\sigma_j}{\tau_j}\right\}.
\end{aligned}$$

Since by Lemma C.1 $\sum_{j=1}^{\infty} {}^{\sigma_j}/_{\tau_j} = \infty$, choosing n_m growing fast enough will enable

$$\exp\left\{\frac{\alpha-1}{\alpha}\sum_{j=m+1}^{n_m}\frac{\sigma_j}{\tau_j}\right\}$$

to grow fast enough, so as to let the ratio of the exponential term and τ_{m+1} diverge, because any speed is achievable, being the series divergent.

□

Chapter 9

Dominance and monopoly

This final chapter is mainly dedicated to showing almost sure dominance with positive feedback. The high-level strategy of the proof is showing, thanks to the results from the previous chapter, that as Θ_n deviates infinitesimally from the equilibria, the martingale parts are powerful enough to push Θ_n to the vertices. As this happens, certain components become small. In order for this strategy to work, it is necessary to gather some analytic information about the vanishing components, that is, it is necessary to show certain upper bounds on their rate of decay. In order to do so, we build a specific tool-box in the following preliminary section. In contrast, the results regarding monopoly follow easily, but the reader is reminded that this is a result of the fact that we omitted the study of the critical regime, since our focus in this work is mainly on dominance.

9.1 Preliminaries

This section is dedicated to a technical result which is a variation on the *Implicit Function Theorem*, and will assist us in the argument for almost sure dominance of Theorem 1.11, in the next section. Lemma 9.1 is more easily read, while following that argument of Theorem 1.11, so that its motivations are clear. Nonetheless we attempt an early description of its role in the argument.

Lemma 9.1 ensures that, on a certain event, which will be suitably defined, for every $l \in \{1, \dots, d-2\}$, $F^{(l+1)}(\Theta_n) \coloneqq \psi^{(l+1)}(\Theta_n) - \Theta_n^{(l+1)}$ stays negative by the time that Θ_n starts drifting away from E_{l+1}, the equilibrium of the $(l+1)$-face, in such a way that it does not approach E_l in the l-face (and then again the argument is repeated iteratively, to show that Θ_n does not get stuck at any of the partial equilibria). There will be no issues, when drifting away from E, so the case $l = d-1$ does not need this lemma, and neither does the model with $d = 2$ as a result, which has no such thing as partial equilibria.

Let $x_{l+2}, \dots, x_d$ be the coordinates that are identically zero on the generic l-face of the simplex considered, according to the order, which will be followed in the argument of Theorem 1.11. It will be clear that there is no loss of generality in following a specific order, as all arguments in Proposition 8.4, Lemma 9.1, and Theorem 1.11 are, *mutatis mutandis*, invariant with respect to the permutations of the order in which the coordinates of Θ_n are considered. Denote $x \coloneqq (x_2, \dots, x_{l+1})$, $y \coloneqq x_{l+2}$ and $x' \coloneqq (x_{l+3}, \dots, x_d)$. Since on the simplex

$$x_1 = x_1(x, y, x') = 1 - \sum_{1<j\neq l+2}^{d} x_j - y,$$

define

$$\psi^{(l+1)}(x_1, x, y, x') - x_{l+1}\Big|_{\Delta^{d-1}} =: F^{(l+1)}(x, y, x') : [0,1]^{d-1} \cap \{0 \le x_2 + \ldots + x_d \le 1\} \longrightarrow \mathbb{R}.$$

In this section we will always implicitly assume that (x, y, x') is in the domain of $F^{(l+1)}$ and, except for *Step 3*, with the x_1-coordinate suppressed. We will work in $\mathbb{R}^{d-1}_{x,y,x'} := \mathbb{R}^{d-1}$, as we have $d-1$ degrees of freedom, by implicitly exploiting the canonical projection mapping $\text{proj}_{x,y,x'} := (\text{proj}_{x_2} \times \ldots \times \text{proj}_{x_d}) \circ \Delta_{d-1}$,

$$\begin{aligned} \text{proj}_{x,y,x'} : \mathbb{R}^d &\longrightarrow \mathbb{R}^{d-1} \\ (x_1, x_2, \ldots x_d) &\longmapsto (x, y, x'), \end{aligned}$$

where Δ_{d-1} denotes the $(d-1)$-fold diagonal embedding, exploited so that the domain is not repeated by the tensor product of the maps. In this new coordinate system we will not change notation, for instance $\text{proj}_{x,y,x'}(E_l)$ will still be denoted as E_l. We will not change the axes labels either, aside from using, instead of the label x_{l+2}, the label y. However, the canonical basis directions' indices will be shifted backwards by 1, with respect to the axes' labels, since the first coordinate is suppressed. For example, the direction of the x_2-axis is e_1 (the vector with 1 in the first coordinate, 0 in all the other $d-2$ coordinates), and so on. The reason for keeping the old labels is related to the context in which we apply the lemma, so it will be clear later on. When considering a point, for example E_l, in this new coordinate system, it will sometimes be necessary to further canonically project it onto the coordinate hyperplane

$$H_{l+1} : \; y = 0,$$

often denoted as H for simplicity, whenever the dependence on l is clear. This is the coordinate hyperplane for the axes corresponding to (x, x') (thus isomorphic to $\mathbb{R}^{d-2}$), onto which we project via the analogously defined projection map $\text{proj}_{x,x'} : \mathbb{R}^{d-1} \longrightarrow H$. When projecting, for example, E_l onto H, we will denote

$$\underset{\sim}{E}_l := \text{proj}_{x,x'}(E_l).$$

Finally, denote the canonical projection (through $\text{proj}_{x,y,x'}$) of the l-face considered, as F_l, that is

$$F_l := \left\{(x, y, x') \in [0,1]^{d-1} \cap \{0 \le x_2 + \ldots + x_d \le 1\} : \; (y, x') = (0, \mathbf{0})\right\},$$

denote

$$F_l^- := \left\{(x, y, x') \in F_l : \; x_{l+1} < \frac{1}{l+1}\right\}$$

the canonically projected portion, of the considered l-face of the simplex, satisfying $x_{l+1} < 1/(l+1)$, and similarly denote

$$F_l^-(\delta_n) := \left\{(x, y, x') \in F_l : \; \frac{1}{l+1} - c_{l+1}\frac{\delta_n}{2} < x_{l+1} < \frac{1}{l+1}\right\}$$

the canonically projected portion, satisfying $1/(l+1) - c_{l+1}\delta_n/2 < x_{l+1} < 1/(l+1)$, where, for all $l \in [d-2]$, $\{c_{l+1}\}$ are arbitrary positive subunitary constants, which will be fixed in Theorem 1.11, and δ_n is the monotonically vanishing sequence introduced in Proposition 8.4. For an intuitive understanding of the statement of the lemma, see Figure 9.1.

Lemma 9.1. *For every integer $l \in [d-2]$ fixed, there exists an open neighbourhood $Y_{l+1} := Y$ of 0 in the y-axis and an open neighbourhood $X_{l+1} := X$ of $\underset{\sim}{E}_l$ in $H_{l+1} := H$, such that in $X \times Y$ the level set $\{(x, y, x') : F^{(l+1)}(x, y, x') = 0\}$ is a hypersurface $G_{l+1} := G$, parametrised by the (unique) continuously differentiable function $y = G(x, x')$, where $G : X \longrightarrow Y$. That is, for every $(x, x') \in X$, $F^{(l+1)}(x, G(x, x'), x') = 0$ and $G(\underset{\sim}{E}_l) = 0$.*

Moreover, there are constants $0 < \gamma^*_{l+1} < \varepsilon^*/2$ and $0 < \varepsilon^* < 1$ small enough, and a time n^*_{l+1} large enough, to ensure that $\delta_{n^*_{l+1}}$ is small enough, such that there exists a connected composite $(d-1)$-dimensional polytope $S^{l+1}_{n^*_{l+1}} \cup P^{l+1}_{n^*_{l+1}} \supset F^-_l$, on which $F^{(l+1)}(x, y, x') < 0$, having defined

$$P^{l+1}_{n^*_{l+1}} := \text{proj}_x \left(F^l_n\right) \times \left[0, \varepsilon^*_{l+1}\right] \times [0, \varepsilon^*_{l+1}]^{d-l-2},$$

and

$$S^{l+1}_{n^*_{l+1}} := \left\{(x, y, x') \in C^{l+1}_{n^*_{l+1}}(\varepsilon^*) : (l+1)\gamma^*_{l+1} x_{l+1} + y + \sum_{j=l+3}^{d} x_j \leq \gamma^*_{l+1}\right\},$$

where

$$C_{n^*_{l+1}}(\varepsilon^*_{l+1}) := \text{proj}_x \left(F^-_l(\delta_{n^*_{l+1}})\right) \times \left[0, \varepsilon^*_{l+1}\right] \times \left[0, \varepsilon^*_{l+1}\right]^{d-l-2}$$

and the hyperplane

$$T^*_{l+1} : (l+1)\gamma^*_{l+1} x_{l+1} + y + \sum_{i=l+3}^{d} x_i = \gamma^*_{l+1}$$

is constructed such that $E_l \in T^*_{l+1}$, T^*_{l+1} intersects the coordinate axes y, x_{l+3}, ..., x_d at distance γ^*_{l+1} from the origin on the positive semi-axes and does not intersect the span of the $(x_2, \dots, x_l)$-coordinate axes.

Proof.

Step 1. $F^{(l+1)}$ is continuously differentiable at E_l and

$$F^{(l+1)}(E_l) = \psi^{(l+1)}(E_l) - E^{(l+1)}_l = \frac{\frac{1}{(l+1)^\alpha}}{(l+1)\frac{1}{(l+1)^\alpha}} - \frac{1}{(l+1)} = 0.$$

Also by direct calculation the partial derivatives of

$$F^{(l+1)}(x, y, x') = \frac{x^\alpha_{l+1}}{(1 - \sum_{i \neq 1, l+2} x_i - y)^\alpha + \sum_{i \neq 1, l+2} x^\alpha_i + y^\alpha} - x_{l+1}$$

are, reminding that $x_1 = x_1(x, y, x') = 1 - \sum_{i \neq 1, l+2} x_i - y$,

$$F^{(l+1)}_{x_j}(x, y, x') = \begin{cases} \frac{\alpha x^\alpha_{l+1}(x^{\alpha-1}_1(x,y,x') - x^{\alpha-1}_j)}{(x^\alpha_1(x,y,x') + \sum_{i \neq 1, l+2} x^\alpha_i + y^\alpha)^2}, & j \neq l+1 \\ \frac{\alpha x^{\alpha-1}_{l+1}\left[x^\alpha_1(x,y,x') + \sum_{i \neq 1, l+2} x^\alpha_i + y^\alpha + x_{l+1}(x^{\alpha-1}_1(x,y,x') - x^{\alpha-1}_{l+1})\right]}{(x^\alpha_1(x,y,x') + \sum_{i \neq 1, l+2} x^\alpha_i + y^\alpha)^2}, & j = l+1, \end{cases}$$

so

$$F^{(l+1)}_{x_j}(E_l) = \begin{cases} \frac{\alpha}{l+1}, & j > l+1 \\ \alpha, & j = l+1 \\ 0, & 1 < j < l+1. \end{cases}$$

Hence $F^{(l+1)}_y(E_l) = \alpha/(l+1) > 0$.

Step 2. By the Implicit Function Theorem applied to $F^{(l+1)}(x, y, x')$, there exist a $(d-2)$-dimensional open neighbourhood X of $\underset{\sim}{E}_l$, and an open interval Y in the y-axis, containing 0, such that there exists a unique continuously differentiable hypersurface $G_{l+1} : X \longrightarrow Y$ (denoted simply as G in this argument) satisfying $F^{(l+1)}(x, G(x, x'), x') = 0$ and $G(\underset{\sim}{E}_l) = 0$. Moreover,

$$\nabla G(x, x') = -\frac{\nabla_{x,x'} F^{(l+1)}(x, y, x')}{F_y^{(l+1)}(x, y, x')}\bigg|_{(x, G(x,x'), x')}$$

and therefore

$$G_{x_j}(\underset{\sim}{E}_l) = \begin{cases} -1, & j > l+2 \\ -(l+1), & j = l+1 \\ 0, & 1 < j < l+1. \end{cases}$$

Step 3. We briefly go back to the standard coordinates for Δ^{d-1}, in order to show that $\psi^{(l+1)}(x_1, \ldots, x_{l+1}, 0, \mathbf{0})$ takes negative values for all x, such that $x_{l+1} < 1/(l+1)$ and $(x_1, \ldots, x_{l+1}, 0, \mathbf{0}) \in \Delta^{d-1}$. Consider first that, given $0 \leq x_i$ for all $i \in [l]$, by *Hölder's inequality*, for every $p, q > 1$ such that $1/p + 1/q = 1$,

$$\sum_{i=1}^{l} x_i = \sum_{i=1}^{l} 1 \cdot x_i \leq \left(\sum_{i=1}^{l} 1^q\right)^{\frac{1}{q}} \left(\sum_{i=1}^{l} x_i^p\right)^{\frac{1}{p}} = l^{\frac{1}{q}} \left(\sum_{i=1}^{l} x_i^p\right)^{\frac{1}{p}}.$$

Take $p = \alpha$ and

$$q = \frac{1}{1 - \frac{1}{p}} = \frac{\alpha}{\alpha - 1}.$$

Then

$$\sum_{i=1}^{l} x_i \leq l^{\frac{\alpha-1}{\alpha}} \left(\sum_{i=1}^{l} x_i^\alpha\right)^{\frac{1}{\alpha}},$$

from which it follows that

$$\left(\sum_{i=1}^{l} x_i\right)^{\alpha} \leq l^{\alpha-1} \sum_{i=1}^{l} x_i^\alpha. \tag{9.1}$$

Consider that

$$\psi^{(l+1)}(x_1, x, 0, \mathbf{0}) < 0 \iff \frac{x_{l+1}^\alpha}{\sum_{i=1}^{l} x_i^\alpha + x_{l+1}^\alpha} < x_{l+1} \iff 1 + \frac{\sum_{i=1}^{l} x_i^\alpha}{x_{l+1}^\alpha} > \frac{1}{x_{l+1}}$$

$$\iff \sum_{i=1}^{l} x_i^\alpha > \left(\frac{1}{x_{l+1}} - 1\right) x_{l+1}^\alpha = \left(\sum_{i=1}^{l} x_i\right) x_{l+1}^{\alpha-1},$$

since $x_{l+2} = \ldots = x_d = 0$. Also since $x_{l+1} < 1/(l+1)$,

$$\sum_{i=1}^{l} x_i = 1 - x_{l+1} > \frac{l}{l+1}.$$

This allows us to verify that

$$\sum_{i=1}^{l} x_i^\alpha > \left(\sum_{i=1}^{l} x_i\right) x_{l+1}^{\alpha-1}, \tag{9.2}$$

since by (9.1)

$$\frac{1}{l^{\alpha-1}}\left(\sum_{i=1}^{l} x_i\right)^{\alpha} \leq \sum_{i=1}^{l} x_i^{\alpha}.$$

We now show that

$$\frac{1}{l^{\alpha-1}}\left(\sum_{i=1}^{l} x_i\right)^{\alpha} \geq \left(\sum_{i=1}^{l} x_i\right) x_{l+1}^{\alpha-1}, \tag{9.3}$$

thus yielding (9.2). (9.3) follows from the assumptions:

$$\frac{1}{l^{\alpha-1}}\left(\sum_{i=1}^{l} x_i\right)^{\alpha} \geq \left(\sum_{i=1}^{l} x_i\right) x_{l+1}^{\alpha-1} \iff \left(\sum_{i=1}^{l} x_i\right)^{\alpha-1} \geq (l x_{l+1})^{\alpha-1},$$

and the second inequality holds since

$$\left(\sum_{i=1}^{l} x_i\right)^{\alpha-1} > \left(\frac{l}{l+1}\right)^{\alpha-1},$$

while

$$(l x_{l+1})^{\alpha-1} < \left(\frac{l}{l+1}\right)^{\alpha-1}.$$

Hence (9.3) follows and (9.2) is satisfied.

Step 4. As a consequence of *Step 3*, for all n large enough, $F^{(l+1)}(x, y, x') < 0$ on F_l^-, and therefore, for any $\varepsilon > 0$ small enough, on the compact $F_n^l := F_l^- \setminus C_n^l(\varepsilon)$, where

$$C_n^l(\varepsilon) := \mathrm{proj}_x\left(F_l^-(\delta_n)\right) \times [0, \varepsilon] \times [0, \varepsilon]^{d-l-2}.$$

We will denote $C_n^l(\varepsilon)$ as $C_n(\varepsilon)$ for simplicity. Note that $C_n(\varepsilon) \subseteq X \times Y$ for all n large enough and ε small enough. Since $F^{(l+1)}(x, y, x')$ is continuous and negative on F_n^l, it will be possible to have ε small enough, such that $F^{(l+1)}(x, y, x') < 0$ on the compact $(d-1)$-dimensional polytope

$$P_n^{l+1} := \mathrm{proj}_x\left(F_n^l\right) \times [0, \varepsilon] \times [0, \varepsilon]^{d-l-2}.$$

Intuitively P_n^{l+1} (denoted simply as P_n) will constitute the main body of the polytope in the claim. Additionally ε will be chosen small enough, to allow $C_n(\varepsilon) \subset X \times Y$ for all n large enough.

We now construct a wedge-like tip for this polytope. Throughout the construction, it will be useful to keep an eye on Figure 9.1. Consider T_{l+1} (denoted as T in this argument), the affine tangent hyperplane at E_l to the hypersurface G, which divides $X \times Y$ into two parts, one on which $F^{(l+1)}(x, y, x') < 0$, and one on which $F^{(l+1)}(x, y, x') \geq 0$, by the Implicit Function Theorem. By the standard theory, we can write the implicit Cartesian equation of T as

$$T: \ \nabla F^{(l+1)}(E_l) \cdot \left(\begin{pmatrix} x \\ y \\ x' \end{pmatrix} - E_l\right) = 0.$$

Since

$$\nabla F^{(l+1)}(E_l) = \begin{pmatrix} \nabla_x F^{(l+1)}(E_l) \\ \nabla_{y,x'} F^{(l+1)}(E_l) \end{pmatrix} = \alpha \begin{pmatrix} \mathbf{0} \\ 1 \\ \frac{\mathbf{1}}{\mathbf{l+1}} \end{pmatrix}$$

from *Step 1*, where we proved that

$$\nabla_x F^{(l+1)}(E_l) = (\mathbf{0}, \alpha)$$

and

$$\nabla_{y,x'} F^{(l+1)}(E_l) = \alpha \frac{\mathbf{1}}{\mathbf{l+1}},$$

the implicit equation reads as

$$T:\ \alpha\left(x_{l+1} - \frac{1}{l+1}\right) + \frac{\alpha}{l+1}\left(y + \sum_{j=l+3}^{d} x_j\right) = 0,$$

which, dividing both sides by $\alpha/(l+1)$ and rearranging, yields

$$T:\ (l+1)x_{l+1} + y + \sum_{j=l+3}^{d} x_j = 1. \tag{9.4}$$

In the following, we adopt as normal vector for T,

$$g := \begin{pmatrix} \mathbf{0} \\ l+1 \\ \mathbf{1} \end{pmatrix}.$$

Note that T intersects each of the coordinate axes $y, x_{l+3}, \ldots, x_d$ at 1. This can be seen easily by putting (9.4) together with the x_j-coordinate axis equation, for all $j \geq l+2$, which reads $x_i = 0$, $\forall i \neq j$; intersecting with the x_{l+1}-axis yields

$$x_i = \begin{cases} \frac{1}{l+1}, & i = l+1 \\ 0, & i \neq l+1, \end{cases}$$

since T goes through E_l; for all $j \geq l+2$, intersecting with the x_j-axis yields

$$x_i = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

Lastly, note that T does not intersect the $\mathrm{span}(e_1, \ldots, e_{l-1})$, its equation being $x_{l+1} = \ldots = x_d = 0$, which turns (9.4) into $0 = 1$.

We will now require n to be large enough to allow $F^{(l+1)}(x, y, x') < 0$ on the whole portion of $C_n(\varepsilon)$ defined as follows. Consider

$$\left\{(x, y, x') \in C_n(\varepsilon) :\ (l+1)x_{l+1} + y + \sum_{j=l+3}^{d} x_j < 1\right\}.$$

There is no guarantee that F^{l+1} does not intersect this set for n large enough and ε small enough. Hence we define a suitable hyperplane, which has the same implicit equation form as that of T:

$$T' :\ \beta x_{l+1} + \sum_{i=l+2}^{d} x_i = \gamma.$$

We first require that $E_l \in T$ and $\beta > 0$, $0 < \gamma < 1$. Then

$$\gamma = \frac{\beta}{l+1},$$

yielding the equation

$$T' : \ (l+1)\gamma x_{l+1} + \sum_{i=l+2}^{d} x_i = \gamma.$$

By a similar argument as for T, it will still not intersect the $\operatorname{span}(e_1, \ldots, e_{l-1})$ as a similar intersection would yield $\gamma = 0$. It clearly intersects the x_{l+1}-coordinate axis at $1/(l+1)$. From the form of its equation, it is also clear that for all $j \geq l+2$, intersecting with the x_j-axis yields

$$x_i = \begin{cases} \gamma, & i = j \\ 0, & i \neq j. \end{cases}$$

For $\gamma = 1$, $T' = T$; as we reduce γ, we lower T' towards the $\operatorname{span}(e_1, \ldots, e_l)$ (γ being the intercept with the $x_{l+2}, \ldots, x_d$ axes, we are reducing these coordinates). Since on $C_n(\varepsilon)$, $x_{l+1} < 1/(l+1)$, by *Step 3* we know that as γ is reduced, T' gets closer to the open neighbourhood of F_l^- on which $F^{(l+1)}(x, y, x') < 0$. For some $\gamma_{l+1}^* < \varepsilon_{l+1}^*/2$, for $\varepsilon_{l+1}^* < 1$ small enough, and n_{l+1}^* large enough, the hyperplane

$$T_{l+1}^* : \ (l+1)\gamma_{l+1}^* x_{l+1} + y + \sum_{i=l+3}^{d} x_i = \gamma_{l+1}^*$$

will be such that $F^{(l+1)}(x, y, x') < 0$ on

$$S_{n_{l+1}^*}^{l+1} := \left\{ (x, y, x') \in C_{n_{l+1}^*}(\varepsilon^*) : \ (l+1)\gamma_{l+1}^* x_{l+1} + y + \sum_{j=l+3}^{d} x_j \leq \gamma_{l+1}^* \right\}.$$

To show this, let us denote S_n^{l+1} as S_n for simplicity, and similarly γ and ε. By *Step 2* we can show that the implicit hypersurface $y = G(x, x')$ decreases strictly as we approach $\utilde{E}_l$ coming from $\operatorname{proj}_{x,x'}(S_{n^*})$, meaning that for any direction $u = (u_x, u_{x'})$ such that $\utilde{E}_l + tu \in \operatorname{proj}_{x,x'}(S_{n^*})$ for some $t > 0$ small enough, the directional derivative

$$\partial_u G(\utilde{E}_l) = \nabla G(\utilde{E}_l) \cdot u = -(l+1)u_{l+1} - \sum_{i=l+3}^{d} u_i > 0. \tag{9.5}$$

This can be shown by noting that, for u to satisfy $\utilde{E}_l + tu \in \operatorname{proj}_{x,x'}(S_{n^*})$ for some $t > 0$, we need to have $u_{l+1} < 0$ (because on this set, the x_{l+1} coordinate is strictly less than $1/(l+1)$) and $u_i \geq 0$. Also, since for all $(x, y, x') \in S_n$ it holds that

$$(l+1)\gamma x_{l+1} + y + \sum_{i=l+3}^{d} x_i \leq \gamma,$$

with $y \geq 0$, it also follows that

$$(l+1)\gamma x_{l+1} + \sum_{i=l+3}^{d} x_i \leq \gamma.$$

We can assume ε, and thus γ, subunitary. Then without loss of generality we can assume that, for all directions u considered, for some $0 < t < 1$,

$$(l+1)\gamma \left(\frac{1}{l+1} + t u_{l+1} \right) + \sum_{i=l+3}^{d} u_i \leq \gamma.$$

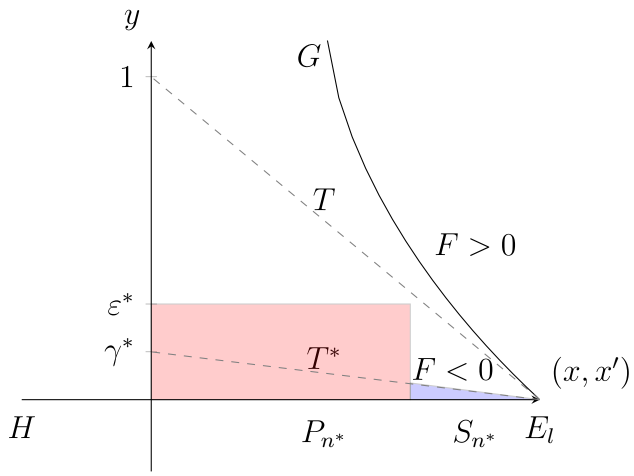


Figure 9.1: P_{n^*} (red) and S_{n^*} (blue)

Therefore

$$\sum_{i=l+3}^{d} u_i \leq \gamma - (l+1)\gamma\left(\frac{1}{l+1} + tu_{l+1}\right) = -(l+1)\gamma t u_{l+1}. \tag{9.6}$$

As a result, by (9.5) and (9.6), $\gamma t < 1$, $u_{l+1} < 0$ and $u_i \geq 0$ for all $l+3 \leq i \leq d$, it follows that

$$\partial_u G(\underset{\sim}{E}_l) \geq -(l+1)u_{l+1} + (l+1)\gamma^* t u_{l+1} = (l+1)(t\gamma^* - 1)u_{l+1} > 0.$$

Since it follows that the surface $y = G(x, x')$ increases in all directions, which go from $\underset{\sim}{E}_l$ to $\mathrm{proj}_{x,x'}(S_n)$. Since T^* is constructed, so as to have lower $y, x_{l+3}, \ldots, x_d$-axes intercept than the tangent hyperplane of G at E_l, it grows at a slower rate than G in all directions going from $\underset{\sim}{E}_l$ to $\mathrm{proj}_{x,x'}(S_n)$. Thus for n large enough and ε small enough, we can fix a suitable γ (respectively denoted as n^*_{l+1}, ε^*_{l+1} and γ^*_{l+1}), such that G will not intersect $S^{l+1}_{n^*_{l+1}}$, meaning that $F^{(l+1)}(x, y, x') < 0$ on the tip of the polytope.

In conclusion, by construction, the composite $(d-1)$-dimensional polytope $P_{n^*} \cup S_{n^*}$ is connected and $F^{(l+1)}(x, y, x') < 0$ on it. Let us explain the construction more informally through Figure 9.1. The horizontal axis is a collapsed representation of the coordinate hyperplane H (the coordinate axis, which has been prioritized in representing it, is x_{l+1}, that is why E_l appears at distance $1/(l+1)$ from the origin, and as to the elevation, we prioritized y, even though the in the higher-dimensional setting the picture is supposed to be homogenous with respect to the $x_{l+3}, \ldots, x_d$ directions, by construction of T^*), the body and the tip of the solid are the coloured regions. The distance between E_l and P_{n^*} is $c_{l+1}\delta_{n^*}/2$, that is the segment, which is the base of the tip, stands actually for $\mathrm{proj}_{x,x'}(S_{n^*}(\varepsilon^*))$. Note that Figure 9.1 is accurate for $d = 3$: $y = x_3$ and $(x, x') = x_2$, with $E_l = E_1 = (1/2, 0)$ and $\underset{\sim}{E}_l = \underset{\sim}{E}_1 = 1/2$. The reason for defining a new set of coordinates, through the projection eliminating the first coordinate, is in fact that the $d = 3$ case becomes univariate in these coordinates, and therefore very easy to handle, offering precious intuition for the general case.

□

The polytope $P^{l+1}_{n^*_{l+1}} \cup P^{l+1}_{n^*_{l+1}}$ extends towards the positive coordinate directions $x_{l+2}, x_{l+3}, \ldots, x_d$, so as to ensure that close enough to the l-face there is space for Θ_n to fit small enough components $\Theta_n^{(l+2)}, \ldots, \Theta_n^{(d)}$ as $\Theta_n^{(l+1)} < 1/(l+1) - \delta_{n^*_{l+1}}/2$, so that $\psi^{(l+1)}(\Theta_n) - \Theta_n^{(l+1)} < 0$, provided n is large enough. This is the gist of the use for

Lemma 9.1 in Theorem 1.11. When making the Cartesian equation of T^*_{l+1} explicit with respect to coordinates $y, x_{l+3}, \ldots, x_d$, we will adopt the notation $y = T^*_{l+1}(x, x')$, $x_{l+3} = T^*_{l+1}(x, y, x_{l+4}, \ldots, x_d)$, and so on. Indeed the explicit equation has always the same form, only the coordinates formally change. For simplicity, we will always make the equation explicit with respect to y. It will be clear that there is no loss of generality in doing this, when exploiting the following corollary in the induction step of Theorem 1.11.

Corollary 9.2. *For all integers $l \in [d-2]$ and $n \geq n^*_{l+1}$, define the sequence*

$$y_{l+1}(\delta_n) := T^*_{l+1}\left(\underset{\sim}{E}_l - c_{l+1}\frac{\delta_n}{2}\underset{\sim}{e}_l\right),$$

where $\underset{\sim}{e}_l$ is the lth element of the canonical basis of $\mathbb{R}^{d-2}_{x,x'}$. By construction, $y_{l+1}(\delta_n) \asymp \delta_n$ and vanishes monotonically.

Proof. Trivially $y_{l+1}(\delta_n) \asymp \delta_n$ by the linearity in δ_n of $y_{l+1}(\delta_n) > 0$, which is obvious from the linearity of $y = T^*_{l+1}(x, x')$ stated in Lemma 9.1. As δ_n is monotone and vanishing by (8.3), trivially ${}^{\delta_n}/_2 < {}^{\delta_{n^*_{l+1}}}/_2$ and

$$\underset{\sim}{E}_l \longleftarrow \underset{\sim}{E}_l - c_{l+1}\frac{\delta_n}{2}\underset{\sim}{e}_l \in \mathrm{proj}_{x,x'}\left(S^{l+1}_{n^*_{l+1}}\right).$$

By Lemma 9.1, $y = T^*_{l+1}(x, x')$ is positive, monotonically decreasing and upper bounded by ${}^{\varepsilon^*_{l+1}}/_2$ along the x_{l+1}-coordinate direction (which is given by $\underset{\sim}{e}_l$, in the canonical projection) and vanishes as $(x, x') \longrightarrow \underset{\sim}{E}_l$, so

$$y_{l+1}(\delta_n) \in \left(0, \frac{\varepsilon^*_{l+1}}{2}\right),$$

with $y_{l+1}(\delta_n) \longrightarrow 0$ as $n \longrightarrow \infty$. □

9.2 Dominance in presence of feedback

In this section we show that with positive feedback, the event of dominance (that is, the event on which one of the components $\Theta^{(i)}_n$ tends to 1) is almost sure in both cases, for which ρ_n is bounded, and for which $\rho_n \longrightarrow \infty$, $\theta = 0$, $\lambda < 1$.

Proof of Theorem 1.11. Define the stopping time $\eta_d := \inf\{n \geq s : \| \Theta_n - E \| > \delta_n\}$, where $s \geq \max\{3, n^*\}$ and

$$n^* = \max_{1 \leq l \leq d-2} n^*_{l+1}$$

can be fixed arbitrarily large, and $\delta_n := {}^1/_{\tau^r_n}$, where $r_\alpha < r < {}^1/_2$, with r_α defined as in Proposition 8.4, n^*_{l+1} is the time defined in Lemma 9.1. For any s, the stopping time η_d is almost surely finite, since by Proposition 8.4, $\mathbb{P}(\| \Theta_n - E \| > \delta_n, \text{ i.o.}) = 1$. By (5.2),

$$\begin{aligned}
\Theta_{n+1} - \Theta_n &= \frac{\tau_n}{\tau_{n+1}}\Theta_n + \frac{1}{\tau_{n+1}}B_{n+1} - \Theta_n \\
&= -\frac{\sigma_{n+1}}{\tau_{n+1}}\Theta_n + \frac{1}{\tau_{n+1}}B_{n+1} - \frac{\sigma_{n+1}}{\tau_{n+1}}P_n + \frac{\sigma_{n+1}}{\tau_{n+1}}\psi(\Theta_n) \\
&= \frac{B_{n+1} - \sigma_{n+1}P_n}{\tau_{n+1}} - \frac{\sigma_{n+1}}{\tau_{n+1}}(\Theta_n - \psi(\Theta_n)).
\end{aligned}$$

Iterating this formula from η_d to $\eta_d + n$ yields

$$\Theta_{\eta_d+n} - \Theta_{\eta_d} = \sum_{k=\eta_d+1}^{\eta_d+n} \frac{B_k - \sigma_k P_{k-1}}{\tau_k} - \sum_{k=\eta_d+1}^{\eta_d+n} \frac{\sigma_k}{\tau_k}(\Theta_{k-1} - \psi(\Theta_{k-1})) = M_n - R_n,$$

where we defined

$$M_n := \sum_{k=\eta_d+1}^{\eta_d+n} \frac{B_k - \sigma_k P_{k-1}}{\tau_k}$$

and

$$R_n := \sum_{k=\eta_d+1}^{\eta_d+n} \frac{\sigma_k}{\tau_k}(\Theta_{k-1} - \psi(\Theta_{k-1})).$$

By the empty sum convention, set $M_0 = 0$ and $R_0 = 0$. We obtained the *Doob-Meyer decomposition* of $\Theta_{\eta_d+n} = \Theta_{\eta_d} + M_n - R_n$. Indeed, trivially $R_n \in \mathrm{m}\mathscr{F}_{\eta_d+n-1}$ (recall that this notation is used for measurability with respect to σ-algebra) and it is therefore previsible with respect to the filtration defined as $\{\mathscr{G}_n\}$, where $\mathscr{G}_n := \mathscr{F}_{\eta_d+n}$, whereas for all i, $M_n^{(i)}$ is a conditional $\{\mathscr{G}_n\}$-martingale. For the definition and basic convergence properties of conditional martingales, see [27]; note that the concepts of martingale and conditional martingale coincide when $\mathscr{G}_0$ is a trivial σ-algebra, and most of the convergence properties are inherited via conditional expectation arguments analogous to the classical ones (for instance the *Conditional Upcrossing Lemma*). Being adapted, $B_k \in \mathrm{m}\mathscr{F}_{\eta_d+n}$ for all $k \leq \eta_d + n$; being previsible, $P_{k-1} \in \mathrm{m}\mathscr{F}_{\eta_d+n}$ for all $k \leq \eta_d + n + 1$; and having $\mathbb{E}_{\mathscr{F}_{\eta_d+n}}(B_{\eta_d+n+1}) = \sigma_{\eta_d+n+1} P_{\eta_d+n}$, it follows that $\mathbb{E}_{\mathscr{G}_n} M_{n+1} = M_n$, since

$$\mathbb{E}_{\mathscr{F}_{\eta_d+n}} M_{n+1} = \sum_{k=\eta_d+1}^{\eta_d+n+1} \frac{1}{\tau_k} \mathbb{E}_{\mathscr{F}_{\eta_d+n}} \left(B_k - \sigma_k P_{k-1}\right) = M_n.$$

Trivially, for all n

$$\begin{aligned}\mathbb{E}_{\mathscr{G}_0}|M_n^{(i)}| = \mathbb{E}_{\mathscr{F}_\eta}|M_n^{(i)}| &\leq \sum_{k=\eta_d+1}^{\eta_d+n+1} \mathbb{E}_{\mathscr{F}_{\eta_d}} \frac{|B_k^{(i)} - \sigma_k P_{k-1}^{(i)}|}{\tau_k} = \sum_{k=\eta_d+1}^{\eta_d+n+1} \mathbb{E}_{\mathscr{F}_{\eta_d}} \frac{\mathbb{E}_{\mathscr{F}_{k-1}}|B_k^{(i)} - \sigma_k P_{k-1}^{(i)}|}{\tau_k} \\ &\leq 2 \sum_{k=\eta_d+1}^{\eta_d+n+1} \mathbb{E}_{\mathscr{F}_{\eta_d}} \frac{\sigma_k P_{k-1}^{(i)}}{\tau_k} \leq \sum_{k=\eta_d+1}^{\eta_d+n+1} 1 = 2n < \infty.\end{aligned}$$

Moreover, $M_n^{(i)}$ is bounded in $\mathscr{L}^2(\mathscr{F}_{\eta_d})$ for all i. In fact, by the tower property,

$$\begin{aligned}\mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)})^2 &= \mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)} - M_{n-1}^{(i)} + M_{n-1}^{(i)} - M_0^{(i)})^2 = \mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)} - M_{n-1}^{(i)})^2 \\ &\quad + \mathbb{E}_{\mathscr{F}_{\eta_d}}(M_{n-1}^{(i)})^2 + 2\mathbb{E}_{\mathscr{F}_{\eta_d}}\left[(M_n^{(i)} - M_{n-1}^{(i)})(M_{n-1}^{(i)} - M_0^{(i)})\right] \\ &= \mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)} - M_{n-1}^{(i)})^2 + \mathbb{E}_{\mathscr{F}_{\eta_d}}(M_{n-1}^{(i)})^2,\end{aligned}$$

since

$$\mathbb{E}_{\mathscr{F}_{\eta_d}}\left[(M_n^{(i)} - M_{n-1}^{(i)})(M_{n-1}^{(i)} - M_0^{(i)})\right] = \mathbb{E}\left[(M_n^{(i)} - M_{n-1}^{(i)})(M_{n-1}^{(i)} - M_0^{(i)})|\mathscr{F}_{\eta_d+n-1}|\mathscr{F}_{\eta_d}\right]$$

and

$$\mathbb{E}_{\mathscr{F}_{\eta_d+n-1}}\left[(M_n^{(i)} - M_{n-1}^{(i)})(M_{n-1}^{(i)} - M_0^{(i)})\right] = (M_{n-1}^{(i)} - M_0^{(i)})\mathbb{E}_{\mathscr{F}_{\eta_d+n-1}}(M_n^{(i)} - M_{n-1}^{(i)}) = 0$$

by the martingale property. Since

$$\mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)})^2 = \mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)} - M_{n-1}^{(i)})^2 + \mathbb{E}_{\mathscr{F}_{\eta_d}}(M_{n-1}^{(i)})^2,$$

and by the tower property

$$\begin{aligned}
\mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)} - M_{n-1}^{(i)})^2 &= \mathbb{E}\left[(M_n^{(i)} - M_{n-1}^{(i)})^2 | \mathscr{F}_{\eta_d+n-1} | \mathscr{F}_{\eta_d}\right] \\
&= \mathbb{E}\left[\left(\frac{B_{\eta_d+n}^{(i)} - \sigma_{\eta_d+n} P_{\eta_d+n-1}^{(i)}}{\tau_{\eta_d+n}}\right)^2 \middle| \mathscr{F}_{\eta_d+n-1} \middle| \mathscr{F}_{\eta_d}\right] \\
&= \mathbb{E}_{\mathscr{F}_{\eta_d}}\left(\frac{\mathrm{Var}_{\mathscr{F}_{\eta_d+n-1}} B_{\eta_d+n}^{(i)}}{\tau_{\eta_d+n}^2}\right) = \mathbb{E}_{\mathscr{F}_{\eta_d}} \frac{\sigma_{\eta_d+n} P_{\eta_d+n-1}^{(i)}(1 - P_{\eta_d+n-1}^{(i)})}{\tau_{\eta_d+n}^2} \\
&\leq \mathbb{E}_{\mathscr{F}_{\eta_d}} \frac{\sigma_{\eta_d+n}}{\tau_{\eta_d+n}^2} = \frac{\sigma_{\eta_d+n}}{\tau_{\eta_d+n}^2},
\end{aligned}$$

we have that

$$\mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)})^2 \leq \mathbb{E}_{\mathscr{F}_{\eta_d}}(M_{n-1}^{(i)})^2 + \frac{\sigma_{\eta_d+n}}{\tau_{\eta_d+n}^2},$$

which can be iterated, yielding

$$\mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)})^2 \leq \sum_{k=\eta_d+1}^{\eta_d+n} \frac{\sigma_k}{\tau_k^2} \leq \int_{\tau_{\eta_d}}^{\infty} \frac{dx}{x^2} = \frac{1}{\tau_{\eta_d}} < 1,$$

and therefore $\sup_n \mathbb{E}_{\mathscr{F}_{\eta_d}}(M_n^{(i)})^2 < \infty$. By the $\mathscr{L}^2$-martingale convergence theorem, M_n converges almost surely in $\mathscr{L}^2(\mathscr{F}_{\eta_d})$.

Step 1. Focus now on the dth component (this justifies the notation with an index d for the stopping time). Consider the event

$$\mathscr{E}_d := \left\{\Theta_{\eta_d}^{(d)} \in \Delta_d\right\},$$

where

$$\Delta_d := \left\{x \in \Delta^{d-1},\ x_d = \min_{1 \leq i \leq d} x_i\right\}.$$

Since $\|\Theta_{\eta_d} - E\| > \delta_{\eta_d}$; since the $(d-1)$-dimensional polytope Δ_d is the convex hull of the vertex E and all other vertices of the simplex having $x_d = 0$; since the angle formed at E by the edges that connect it to any two other vertices in Δ_d is

$$\frac{\pi}{2} < \arccos\left(-\frac{1}{d-1}\right) \leq \frac{2}{3}\pi$$

for every $d \geq 3$; since the angle between the dth basis vector e_d and the normal to the hyperplane to which the simplex belongs, which is $\mathbf{1}$, is

$$\frac{\pi}{4} < \arccos\left(\frac{1}{\sqrt{d}}\right) < \frac{\pi}{2}$$

for every $d \geq 3$; we can see that on $\mathscr{E}_d$ not only $\Theta_{\eta_d} \leq 1/d$ by the *pigeonhole principle*, but

$$\frac{1}{d} - \Theta_{\eta_d}^{(d)} > c_d \delta_{\eta_d},$$

where c_d is a subunitary trigonometric constant bounded away from zero, such that $c_d\delta_{\eta_d}$ is the minimum modulus of the orthogonal projection of

$$\delta_{\eta_d}\frac{x-E}{\|x-E\|}$$

onto the axis spanned by the dth basis vector e_d, for x varying in the compact Δ_d. That is

$$c_d\delta_{\eta_d} = \min_{x\in\Delta_d}\left|\left(\delta_{\eta_d}\frac{x-E}{\|x-E\|}\right)\cdot e_d\right|$$

where $\cdot$ denotes the scalar product in $\mathbb{R}^d$. Hence

$$\mathscr{E}_d \subseteq \left\{\Theta^{(d)}_{\eta_d} < \frac{1}{d} - c_d\delta_{\eta_d}\right\}.$$

To prove that on $\mathscr{E}_d$, $\Theta^{(d)}_{\eta_d+n} \longrightarrow 0$ as $n \longrightarrow \infty$, consider the event

$$\mathscr{S}_d := \left\{\sup_n M^{(d)}_n \leq \frac{c_d\delta_{\eta_d}}{2}\right\}.$$

We will show that $\Theta^{(d)}_n$ vanishes on $\mathscr{S}_d \cap \mathscr{E}_d$. This will be enough to prove that it vanishes almost surely on $\mathscr{E}_d$, thanks to a bound on the probability of the complement $\mathscr{S}^c_d$ that we will derive. We first show this bound. Consider that

$$\mathbb{P}(\mathscr{S}^c_d) = \mathbb{E}(\mathbb{1}_{\mathscr{S}^c_d}) = \mathbb{E}\mathbb{E}_{\mathscr{F}_{\eta_d}}(\mathbb{1}_{\mathscr{S}^c_d}) = \mathbb{E}\mathbb{P}_{\mathscr{F}_{\eta_d}}(\mathscr{S}^c_d),$$

where

$$\mathbb{P}_{\mathscr{F}_{\eta_d}}(\mathscr{S}^c_d) = \mathbb{P}_{\mathscr{F}_{\eta_d}}\left(\sup_n M^{(d)}_n > \frac{c_d\delta_{\eta_d}}{2}\right),$$

which can be estimated by *Doob's submartingale inequality* as follows. Define the event

$$H^d_n := \left\{\max_{k\leq n} M^{(d)}_k > \frac{c_d\delta_{\eta_d}}{2}\right\}.$$

Since $H^d_n \subseteq H^d_{n+1}$, it follows that

$$\mathscr{S}^c_d = \lim_{n\to\infty} H^d_n,$$

and as a consequence of the *monotonicity of probability measures*,

$$\mathbb{P}_{\mathscr{F}_{\eta_d}}(\mathscr{S}^c_d) = \lim_{n\to\infty}\mathbb{P}_{\mathscr{F}_{\eta_d}}(H^d_n).$$

Applying Doob's inequality to the positive submartingale $\{(M^{(d)}_k)^2\}$ yields

$$\begin{aligned}\mathbb{P}_{\mathscr{F}_{\eta_d}}(H^d_n) &= \mathbb{P}_{\mathscr{F}_{\eta_d}}\left(\max_{k\leq n} M^{(d)}_k > \frac{c_d\delta_{\eta_d}}{2}\right) \leq \mathbb{P}_{\mathscr{F}_{\eta_d}}\left(\max_{k\leq n}(M^{(d)}_k)^2 > \frac{c^2_d\delta^2_{\eta_d}}{4}\right)\\ &\leq \frac{4}{c^2_d\delta^2_{\eta_d}}\mathbb{E}_{\mathscr{F}_{\eta_d}}(M^{(d)}_n)^2 \leq \frac{4}{c^2_d\delta^2_{\eta_d}\tau_{\eta_d}} = \frac{4\tau^{2r_\alpha}_{\eta_d}}{c^2_d\tau_{\eta_d}} = \frac{4}{c^2_d\tau^{1-2r}_{\eta_d}} \leq \frac{4}{c^2_d\tau^{1-2r}_s}\end{aligned}$$

almost surely, thus

$$\mathbb{P}_{\mathscr{F}_{\eta_d}}(\mathscr{S}^c_d) \leq \frac{4}{c^2_d\tau^{1-2r}_s}$$

almost surely, and as a result

$$\mathbb{P}(\mathscr{S}_d^c) = \mathbb{E}\mathbb{P}_{\mathscr{F}_{\eta_d}}(\mathscr{S}_d^c) \leq \mathbb{E}\frac{4}{c_d^2 \tau_s^{1-2r_\alpha}} = \frac{4}{c_d^2 \tau_s^{1-2r}}. \tag{9.7}$$

Since s can be chosen arbitrarily large, and since $r_\alpha < r < 1/2$ ensures $1 - 2r > 0$, implying

$$\lim_{s\to\infty} \frac{4}{c_d^2 \tau_s^{1-2}} = 0,$$

we can conclude that if we can show that $\Theta_n^{(d)}$ vanishes on $\mathscr{S}_d \cap \mathscr{E}_d$, by the fact that s can be chosen arbitrarily large and thus $\mathbb{P}(\mathscr{S}_d^c)$ arbitrarily small, we will have shown that it vanishes almost surely on the whole of $\mathscr{E}_d$ (this is a trivial argument by contradiction that we omit).

Consider now the event $\mathscr{S}_d \cap \mathscr{E}_d$ and recall that on $\mathscr{E}_d$,

$$\Theta_{\eta_d}^{(d)} < \frac{1}{d} - c_d \delta_{\eta_d} < \frac{1}{d} - \frac{c_d}{2}\delta_{\eta_d}.$$

By induction, it can be proved that on $\mathscr{S}_d \cap \mathscr{E}_d$, for all n

$$\Theta_{\eta_d+n}^{(d)} < \frac{1}{d} - \frac{c_d}{2}\delta_{\eta_d}.$$

For $n = 0$ it follows trivially from $\mathscr{S}_d \cap \mathscr{E}_d \subseteq \mathscr{E}_d$. Assume that for all $j \in [n-1]$,

$$\Theta_{\eta_d+j}^{(d)} < \frac{1}{d} - \frac{c_d}{2}\delta_{\eta_d}.$$

Then since for all $j \in [n-1]$, $\Theta_{\eta_d+j}^{(d)} < 1/d$, (8.1) yields that for all $j \in [n-1]$,

$$\psi^{(d)}(\Theta_{\eta_d+j}) \leq d^{\alpha-1}(\Theta_{\eta_d+j}^{(d)})^\alpha < \frac{d^{\alpha-1}}{d^{\alpha-1}}\Theta_{\eta_d+j}^{(d)} = \Theta_{\eta_d+j}^{(d)},$$

and therefore, for all $j \in [n-1]$,

$$\Theta_{\eta_d+j}^{(d)} - \psi^{(d)}(\Theta_{\eta_d+j}) > 0.$$

Hence

$$R_n^{(d)} = \sum_{k=\eta_d+1}^{\eta_d+n} \frac{\sigma_k}{\tau_k}(\Theta_{k-1}^{(d)} - \psi^{(d)}(\Theta_{k-1})) > 0$$

on $\mathscr{S}_d \cap \mathscr{E}_d$. The positivity of $R_n^{(d)}$ yields that on $\mathscr{S}_d \cap \mathscr{E}_d$

$$\Theta_{\eta_d+n}^{(d)} = \Theta_{\eta_d}^{(d)} + M_n^{(d)} - R_n^{(d)} < \Theta_{\eta_d}^{(d)} + M_n^{(d)} \leq \frac{1}{d} - c_d\delta_{\eta_d} + \frac{c_d}{2}\delta_{\eta_d} = \frac{1}{d} - \frac{c_d}{2}\delta_{\eta_d}.$$

As a result, since all added terms are strictly positive, $R_n^{(d)}$ is strictly increasing on $\mathscr{S}_d \cap \mathscr{E}_d$. Since $0 \leq \Theta_{\eta_d+n} \leq 1$; since $\{M_n^{(d)}\}$ is almost surely convergent in $\mathscr{L}^2(\mathscr{F}_{\eta_d})$; since $R_n^{(d)}$ is positive and strictly increasing and $R_n^{(d)} = \Theta_{\eta_d}^{(d)} - \Theta_{\eta_d+n}^{(d)} + M_n^{(d)}$, which yields boundedness of $R_n^{(d)}$ as $|R_n^{(d)}| \leq \Theta_{\eta_d}^{(d)} + \Theta_{\eta_d+n}^{(d)} + |M_n^{(d)}| \leq 2 + |M_n^{(d)}|$ and thus of its limit, that is $\sup_n R_n^{(d)} \leq 2 + |M_\infty^{(d)}|$; we can conclude that being bounded positive and increasing, $R_n^{(d)}$ converges almost surely too. This proves that there exists an almost sure limit $\Theta^{(d)} \longleftarrow \Theta_k^{(d)} = \Theta_{\eta_d+n}^{(d)}$. It is left to prove that $\Theta^{(d)} = 0$. By contradiction, assume that there is an $\omega \in \mathscr{S}_d \cap \mathscr{E}_d$ such that $\Theta^{(d)}(\omega) > 0$. But on $\mathscr{S}_d \cap \mathscr{E}_d$, for all $k = \eta_d + j$, where $j \in \mathbb{N}_0$, $\Theta_k^{(d)}(\omega) - \psi^{(d)}(\Theta_k(\omega)) > 0$. Passing to the limit and recalling

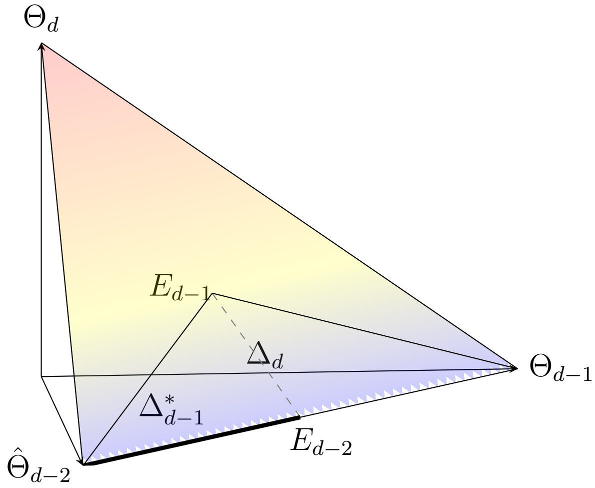


Figure 9.2: Δ_d and Δ^*_{d-1}, where $\hat{\Theta}_{d-2} = (\Theta_1, \ldots, \Theta_{d-2})$.

that by assumption $\Theta^{(d)}(\omega) > 0$, this implies that $\Theta^{(d)}(\omega) - \psi^{(d)}(\Theta(\omega)) > 0$ by the continuity of $\psi^{(d)}$. From which it follows that

$$\frac{\sigma_k}{\tau_k}(\Theta^{(d)}_k(\omega) - \psi^{(d)}(\Theta_k(\omega))) \sim \frac{\sigma_k}{\tau_k}(\Theta^{(d)}(\omega) - \psi^{(d)}(\Theta(\omega)))$$

and that

$$\sum_{k=\eta_d+1}^{\infty} \frac{\sigma_k}{\tau_k}(\Theta^{(d)}(\omega) - \psi^{(d)}(\Theta(\omega))) = (\Theta^{(d)}(\omega) - \psi^{(d)}(\Theta(\omega))) \sum_{k=\eta_d+1}^{\infty} \frac{\sigma_k}{\tau_k} = \infty,$$

as $\Theta^{(d)}(\omega) - \psi^{(d)}(\Theta(\omega)) > 0$ and Lemma C.1 holds. Then by the *limit comparison test* we have that

$$R_\infty(\omega) := \sum_{k=\eta_d+1}^{\infty} \frac{\sigma_k}{\tau_k}(\Theta^{(d)}_{k-1}(\omega) - \psi^{(d)}(\Theta_{k-1}(\omega))) = \infty.$$

Since surely $\Theta^{(d)} \geq 0$, the following contradiction is reached: $\Theta^{(d)}(\omega) = \Theta^{(d)}_\eta(\omega) + M_\infty(\omega) - R_\infty(\omega) = -\infty$. Thus $\Theta^{(d)} = 0$ almost surely on $\mathcal{S}_d \cap \mathscr{E}_d$, and as we argued earlier, by s being allowed to be arbitrarily large and by (9.7), it follows that $\Theta^{(d)} = 0$ almost surely on $\mathscr{E}_d$.

Step 2. By *Step 1* and Lemmas E.2 and F.1 respectively, in the two regimes of growth of the claim, for almost every $\omega \in \mathscr{E}_d$ there exists a finite upper bound $T^{(d)}(\omega)$ for $\{T^{(d)}_n(\omega)\}$. We exploit this bound and define a new stopping time, that similarly to η_d, allows to prove that also $\Theta^{(d-1)}_n$ vanishes on an analogous event, included in the previous one, denoted as $\mathscr{E}_{d-1}$. The overall scheme, besides this step, is to show that, one by one, other $d-2$ components vanish, on a suitable event. The bound $T^{(d)}$ requires information about all future $T^{(d)}_n$. It cannot be used in defining a stopping time. However, since the deterministic function $\log \tau_n$ diverges to infinity, the existence of an almost sure bound allows to conclude that eventually for almost all $\omega \in \mathscr{E}_d$, $T^{(d)}_n(\omega) < \log \tau_n$. As a result for almost every $\omega \in \mathscr{E}_d$, there exists a finite positive integer $N_d = N_d(\omega)$ such that

$$\Theta^{(d)}_n(\omega) < \frac{\log \tau_n}{\tau_n}$$

for all $n \geq N_d$. Recall that $s \geq \max\{3, n^*\}$. It is essential for the argument that this deterministic upper bound on $\Theta^{(d)}_n$ is strictly decreasing and vanishing. In particular, for all $x > 0$, ${}^{\log x}/_x$ is decreasing on $[e, \infty)$ (note that $3 = \lceil e \rceil$), in that

$$\frac{d}{dx}\frac{\log x}{x} = \frac{1 - \log x}{x^2} < 0$$

on (e, ∞) (seen by rearranging the inequality for all $x > 0$ into $\log x > 1$ and taking logarithms). This is the reason for requiring $s \geq 3$. We must also satisfy $s \geq n^*$, so that $s \geq n^*_{d-1}$. This guarantees that, for the constant c_{d-1}, which we will define similarly to c_d,

$$\underset{\sim}{E}_{d-2} - c_{d-1}\frac{\delta_k}{2}\underset{\sim}{e}_{d-2} \in \operatorname{proj}_x\left(S^{d-1}_{n^*_{d-1}}\right),$$

where the notation is with respect to the projected coordinate system

$$(\underbrace{x_2, \ldots, x_{d-1}}_{x}, \underbrace{x_d}_{y})$$

(in this first step there is no coordinate x') used in Lemma 9.1 with $l = d-2$. Finally note that there is a deterministic $M_{d-1} \in \mathbb{N}$, such that

$$\frac{\log \tau_n}{\tau_n} \leq y_{d-1}(\delta_n)$$

for all $n \geq M_{d-1}$, where

$$y_{d-1}(\delta_n) := T^*_{d-1}\left(\underset{\sim}{E}_{d-2} - c_{d-1}\frac{\delta_n}{2}\underset{\sim}{e}_{d-2}\right) \in \left(0, \frac{\varepsilon^*_{d-1}}{2}\right).$$

M_{d-1} is well defined, in that by Corollary 9.2 (with $l = d-2$), $y_{d-1}(\delta_n) \asymp \delta_n$ as $n \longrightarrow \infty$, and

$$\frac{\log \tau_n}{\tau_n} = \frac{\log \tau_n}{\tau_n^{1-r}}\delta_n = o(\delta_n).$$

Thus

$$\frac{\frac{\log \tau_n}{\tau_n}}{y_{d-1}(\delta_n)} \longrightarrow 0,$$

and the ratio is eventually bounded by 1 and the inequality holds eventually, as per the definition of M_{d-1}. Define the stopping time

$$\eta_{d-1} := \inf\left\{n \geq s : \; \Theta^{(d)}_n < \frac{\log \tau_n}{\tau_n} < y_{d-1}(\delta_n), \|\Theta_n - E_{d-2}\| > \delta_n\right\}.$$

It is known, from Proposition 8.4 and the almost sure existence of N_d and M_{d-1} (which are finite), that on $\mathscr{E}_d$, η_{d-1} is almost surely finite. Define

$$\begin{aligned}
\Delta_{d-1} &:= \left\{x \in \Delta^{d-1}, \; x_{d-1} = \min_{1 \leq i \leq d-1} x_i\right\} \\
\Delta^*_{d-1} &= \Delta_{d-1} \cap \Delta_d \\
\mathscr{E}^*_{d-1} &:= \left\{\Theta_{\eta_{d-1}} \in \Delta^*_{d-1}\right\} \cap \mathscr{E}_d \\
G^s_{d-1} &:= \left\{\omega \in \mathscr{E}^*_{d-1} : \; \Theta^{(d)}_n < \frac{\log \tau_n}{\tau_n} < y_{d-1}(\delta_n), \; \forall n \geq s\right\}.
\end{aligned}$$

We have that

$$\mathscr{E}^*_{d-1} \subseteq \left\{\Theta^{(d-1)}_{\eta_{d-1}} < \frac{1}{d-1} - c_{d-1}\delta_{\eta_{d-1}}\right\},$$

because $\left\|\Theta_{\eta_{d-1}} - E_{d-2}\right\| > \delta_{\eta_{d-1}}$; since the $(d-1)$-dimensional polytope Δ^*_{d-1} is the convex hull of the vertex E, E_{d-2} and all other vertices of the simplex having $x_d = x_{d-1} = 0$; since the angle formed at E_{d-2} by the edges that connect it to any two other vertices of Δ^*_{d-1} is either a right angle (if one of the vertices is E, which is always the case if we started with $d = 3$, in which scenario this step is the last step) or is

$$\frac{\pi}{2} < \arccos\left(-\frac{1}{d-2}\right) \leq \frac{2\pi}{3}$$

for every $d > 3$ (if both vertices lie in the $(d-2)$-dimensional face having $x_d = x_{d-1} = 0$, which is Δ^{d-2}); since the angle between the $(d-1)$st basis vector e_{d-1} and the normal to the hyperplane, to which the simplex belongs, is

$$\frac{\pi}{4} < \arccos\left(\frac{1}{\sqrt{d}}\right) < \frac{\pi}{2}$$

for every $d \geq 3$; we can conclude that on $\mathscr{E}_d$,

$$\frac{1}{d-1} - \Theta^{(d-1)}_{\eta_{d-1}} > c_{d-1}\delta_{\eta_{d-1}},$$

where c_{d-1} is a subunitary trigonometric constant bounded away from zero, such that

$$c_{d-1}\delta_{\eta_{d-1}} = \min_{\Theta \in \Delta^*_{d-1}} \left| \left(\delta_{\eta_{d-1}} \frac{\Theta - E_{d-2}}{\|\Theta - E_{d-2}\|} \right) \cdot e_{d-1} \right|.$$

We move on to proving that for all s, on $G^s_{d-1} \cap \mathcal{S}_{d-1}$ almost surely $\Theta^{(d-1)}_k \longrightarrow 0$ with a similar argument as in *Step 1*, but making use of Lemma 9.1 to ensure that $R^{(d-1)}_n$ is positive increasing ((8.1) is no longer sufficient at partial equilibria). Since $G^s_{d-1} \subseteq G^{s+1}_{d-1}$ and since on $\mathscr{E}_d \supseteq \mathscr{E}^*_{d-1}$ there are N_d and M_{d-1} almost surely finite,

$$\bigcup_{s=\max\{3,n^*\}}^{\infty} G^s_{d-1} = \lim_{s\to\infty} G^s_{d-1} = G^{\infty}_{d-1}$$

is well defined and it is such that $\mathbb{P}(\mathscr{E}^*_{d-1} \smallsetminus G^{\infty}_{d-1}) = 0$. Then it follows that

$$\mathbb{P}(\mathscr{E}^*_{d-1} \smallsetminus G^s_{d-1}) \downarrow 0 \tag{9.8}$$

as $s \longrightarrow \infty$. Similarly, by a bound analogous to the one achieved for $\mathcal{S}^c_d$, we show that $\mathbb{P}(\mathcal{S}^c_{d-1})$ is bounded by a function that vanishes as $s \longrightarrow \infty$. Therefore, the conclusion will be that almost surely on $\mathscr{E}^*_{d-1}$, $\Theta^{(d-1)}_k \longrightarrow 0$ (and $\Theta^{(d)}_k \longrightarrow 0$ by the previous step). The argument that follows is technically speaking unnecessary, because the first step of the inductive argument has already been done. However the second step shows the essence of the general inductive step, but with less technicalities. To gain some intuition about this step see Figure 9.2, where all the dimensions $(\Theta_1, \ldots, \Theta_{d-2})$ have been collapsed in one flat subspace of codimension 2, $\hat{\Theta}_{d-2}$. Note that the picture is accurate for the three bins case, where $\hat{\Theta}_{d-2} = \Theta_1$.

Consider the $(d-1)$st components $M^{(d-1)}_n$, $R^{(d-1)}_n$ and the σ-algebra $\mathscr{F}_{\eta_{d-1}}$. Defining $\mathscr{G}_n := \mathscr{F}_{\eta_{d-1}+n}$ by the exact same argument as before, *mutatis mutandis*, it holds that $M^{(d-1)}_n$ is a $\mathscr{G}_n$-martingale and $M^{(d-1)}_n \in \mathscr{L}^2(\mathscr{F}_{\eta_{d-1}})$ and is bounded (the same bound found on the dth component holds), hence it is almost surely convergent in $\mathscr{L}^2(\mathscr{F}_{\eta_{d-1}})$. Define the event

$$\mathcal{S}_{d-1} = \left\{ \sup_n M^{(d-1)}_n \leq \frac{c_{d-1}\delta_{\eta_{d-1}}}{2} \right\}.$$

Analogously to what done in *Step1*, we can find a vanishing bound for $\mathbb{P}(\mathscr{S}^c_{d-1})$ using Doob's inequality. Define

$$H_n^{d-1} = \left\{ \max_{k \le n} M_k^{(d-1)} > \frac{c_{d-1}\delta_{\eta_{d-1}}}{2} \right\}.$$

Then

$$\mathbb{P}_{\mathscr{F}_{\eta_{d-1}}}(H_n^{d-1}) \le \frac{4(M_n^{(d-1)})^2}{c_{d-1}^2\delta_{\eta_{d-1}}^2} \le \frac{4\tau_{\eta_{d-1}}^{2r}}{c_{d-1}^2\tau_{\eta_{d-1}}} = \frac{4}{c_{d-1}^2\tau_{\eta_{d-1}}^{1-2r}} \le \frac{4}{c_{d-1}^2\tau_s^{1-2r}}.$$

As in *Step 1* this implies that

$$\mathbb{P}(\mathscr{S}^c_{d-1}) \le \frac{4}{c_{d-1}^2\tau_s^{1-2r}} \tag{9.9}$$

and therefore this probability can be made arbitrarily small by taking s sufficiently large. The argument proceeds as before, by proving by induction that for all $n \ge 0$, on $G^s_{d-1} \cap \mathscr{S}_{d-1}$,

$$\Theta^{(d-1)}_{\eta_{d-1}+n} < \frac{1}{d-1} - c_{d-1}\frac{\delta_{\eta_{d-1}}}{2}. \tag{9.10}$$

The case $n = 0$ follows by $G^s_{d-1} \cap \mathscr{S}_{d-1} \subseteq G^s_{d-1} \subseteq \mathscr{E}^*_{d-1}$, on which

$$\Theta^{(d-1)}_{\eta_{d-1}} < \frac{1}{d-1} - c_{d-1}\delta_{\eta_{d-1}}.$$

Assume the induction hypothesis true for all indices $0 \le j < n$. The main difference from the induction argument in *Step 1* is that now the inductive hypothesis does not ensure that $\Theta^{(d-1)}_{\eta_{d-1}+j} < 1/d$ (which would allow using (8.1) as in *Step 1*) since $1/(d-1) > 1/d$. To show (9.10), we use Lemma 9.1 instead. For all $0 \le j < n$, set $k = \eta_{d-1} + j$, and recall that

$$F^{(d-1)}(x, y) := F^{(d-1)}(\underbrace{\Theta_k^{(2)}, \dots \Theta_k^{(d-1)}}_{x}, \underbrace{\Theta_k^{(d)}}_{y}) := \psi^{(d-1)}(\Theta_k) - \Theta_k^{(d-1)}.$$

We show that $R_n^{(d-1)}$ is positive, so that the induction step will follow, as

$$\Theta^{(d-1)}_{\eta_{d-1}+n} = \Theta^{(d-1)}_{\eta_{d-1}} + M_n^{(d-1)} - R_n^{(d-1)} < \Theta^{(d-1)}_{\eta_{d-1}} + M_n^{(d-1)} < \frac{1}{d-1} - c_{d-1}\delta_{\eta_{d-1}}$$
$$+ \frac{c_{d-1}\delta_{\eta_{d-1}}}{2} = \frac{1}{d-1} - c_{d-1}\frac{\delta_{\eta_{d-1}}}{2}.$$

The conditions that have been put in the definition of the stopping time η_{d-1} all intervene here, to ensure that the vector $(\Theta_k^{(2)}, \dots, \Theta_k^{(d)}) \in P^{d-1}_{n^*_{d-1}} \cup S^{d-1}_{n^*_{d-1}}$, on which $F^{(d-1)}(x, y) < 0$, so that $-F^{(d-1)}(\Theta_k^{(2)}, \dots, \Theta_k^{(d)}) > 0$ for all $\eta_{d-1} < k < \eta_{d-1} + n$ and therefore $R_n^{(d-1)}$ is positive. Since $k > \eta_{d-1} > s \ge n^*_{d-1}$ and $\omega \in G^s_{d-1} \cap \mathscr{S}_{d-1} \subseteq G^s_{d-1} \subseteq \mathscr{E}_d$,

$$\Theta_k^{(d)} < \frac{\log \tau_k}{\tau_k} < y_{d-1}(\delta_k)$$

for all considered k. The induction hypothesis ensures that

$$\Theta_k^{(d-1)} < \frac{1}{d-1} - c_{d-1}\frac{\delta_{\eta_{d-1}}}{2},$$

but the monotonicity of δ_k, following from its definition by (8.3), tells us that

$$\frac{1}{d-1} - c_{d-1}\frac{\delta_{\eta_{d-1}}}{2} > \frac{1}{d-1} - c_{d-1}\frac{\delta_{n^*_{d-1}}}{2}.$$

Then we cannot conclude simply that $(\Theta_k^{(2)}, \dots, \Theta_k^{(d)}) \in P^{d-1}_{n^*_{d-1}}$, but we can be sure that $(\Theta_k^{(2)}, \dots, \Theta_k^{(d)}) \in P^{d-1}_{n^*_{d-1}} \cup S^{d-1}_{n^*_{d-1}}$. Indeed, the components lower than the $(d-1)$st have no constraints to satisfy on $P^{d-1}_{n^*_{d-1}} \cup S^{d-1}_{n^*_{d-1}}$, besides corresponding to points of the simplex, which is trivial. Whereas, $\Theta_k^{(d)}$ has been shown small enough to ensure that, if it were the case that

$$\Theta_k^{(d-1)} > \frac{1}{d-1} - c_{d-1}\frac{\delta_{n^*_{d-1}}}{2},$$

it would follow that $(\Theta_k^{(2)}, \dots, \Theta_k^{(d)}) \in S^{d-1}_{n^*_{d-1}}$. Hence $F^{(d-1)}(\Theta_k^{(2)}, \dots, \Theta_k^{(d)}) < 0$ and the induction step is completed. Then (9.10) holds on $G^s_{d-1} \cap \mathcal{S}_{d-1}$ for all $n \geq 0$. Thus $\{R_n^{(d-1)}\}$ is positive and increasing on $G^s_{d-1} \cap \mathcal{S}_{d-1}$. At this point, reasoning as in the concluding part of *Step 1*, it is possible to conclude by contradiction that $\Theta_k^{(d-1)} \longrightarrow 0$ almost surely on $G^s_{d-1} \cap \mathcal{S}_{d-1}$. Since s is arbitrary, through (9.8) and (9.9) we have that $\Theta_k^{(d-1)} \longrightarrow 0$ almost surely on $\mathscr{E}^*_{d-1}$. Since $\mathscr{E}^*_{d-1} \subseteq \mathscr{E}_d$, also $\Theta_k^{(d)} \longrightarrow 0$ almost surely on $\mathscr{E}^*_{d-1}$ by *Step 1*. Thus almost surely on $\mathscr{E}^*_{d-1}$ both $\Theta_k^{(d)}$ and $\Theta_k^{(d-1)}$ vanish.

Step 3. To conclude we repeat the argument for a general step $l \in [d-3]$, making use of Lemma 9.1 in its full generality. As induction hypothesis assume that on $\mathscr{E}^*_{l+2}$ (inductively defined as usual), $\Theta_n^{(l+2)}, \dots, \Theta_n^{(d)}$ have been shown to vanish almost surely. Then by Lemmas E.2 and F.1, for almost every $\omega \in \mathscr{E}^*_{l+2}$ there is a finite bound $T^{(i)}(\omega)$ for $\{T_n^{(i)}(\omega)\}$, for all $l+2 \leq i \leq d$, in both regimes of growth respectively. We discussed in *Step 2* that this allows us to conclude that, for almost every $\omega \in \mathscr{E}^*_{l+2}$, there exist finite positive integers $N_i = N_i(\omega)$ such that, if we define $N^*_{l+2} := \max_{l+2 \leq i \leq d} N_i$, it holds that

$$\Theta_n^{(i)}(\omega) < \frac{\log \tau_n}{\tau_n}$$

for all $n \geq N^*_{l+2}$ and $l+2 \leq i \leq d$. Recall that $s \geq \max\{3, n^*\} \geq n^*_{l+1}$. This guarantees that, for the constant c_{l+1}, defined in a moment similarly to the $\{c_i\}$ in the previous steps,

$$\underset{\sim}{E}_l - c_{l+1}\frac{\delta_k}{2}\underset{\sim}{e}_l \in \mathrm{proj}_{x,x'}\left(S^{l+1}_{n^*_{l+1}}\right),$$

where the notation is with respect to the projected coordinate system

$$(\underbrace{x_2, \dots, x_{l+1}}_{x}, \underbrace{x_{l+2}}_{y}, \underbrace{x_{l+3}, \dots, x_d}_{x'})$$

used in Lemma 9.1. Finally note that there is a deterministic $M_{l+1} \in \mathbb{N}$ such that for all $n \geq M_{l+1}$,

$$\frac{\log \tau_n}{\tau_n} \leq y_{l+1}(\delta_n),$$

where

$$y_{l+1}(\delta_n) = T^*_{l+1}\left(\underset{\sim}{E}_l - c_{l+1}\frac{\delta_n}{2}\underset{\sim}{e}_l\right) \in \left(0, \frac{\varepsilon^*_{l+1}}{2}\right).$$

M_{l+1} is well defined, since by Corollary 9.2, $y_{l+1}(\delta_n) \asymp \delta_n$, while ${}^{\log \tau_n}/_{\tau_n} = o(\delta_n)$ as $n \longrightarrow \infty$. Define the stopping time

$$\eta_{l+1} := \inf\left\{n \geq s :\ \Theta_n^{(i)} < \frac{\log \tau_n}{\tau_n} < y_{l+1}(\delta_n),\ \forall\, l+2 \leq i \leq d,\ \|\Theta_n - E_l\| > \delta_n\right\}.$$

It is known, from Proposition 8.4 and the almost sure finiteness of N^*_{l+2} and M_{l+1}, that on $\mathscr{E}^*_{l+2}$, η_{l+1} is almost surely finite. Define

$$\begin{aligned}
\Delta_{l+1} &:= \left\{ x \in \Delta^{d-1}, x_{l+1} = \min_{i\in\{1,\dots,l+1\}} x_i \right\} \\
\Delta^*_{l+1} &:= \Delta_{l+1} \cap \Delta^*_{l+2} \\
\mathscr{E}^*_{l+1} &:= \{\Theta_{\eta_{l+1}} \in \Delta^*_{l+1}\} \cap \mathscr{E}^*_{l+2} \\
G^s_{l+1} &:= \left\{ \omega \in \mathscr{E}^*_{l+1} :\ \Theta^{(i)}_n < \frac{\log \tau_n}{\tau_n} < y_{l+1}(\delta_n)\ \ \forall l+2 \le i \le d,\ \forall n \ge s \right\}.
\end{aligned}$$

We have that

$$\mathscr{E}^*_{l+1} \subseteq \left\{ \Theta^{(l+1)}_{\eta_{l+1}} < \frac{1}{l+1} - c_{l+1}\delta_{\eta_{l+1}} \right\},$$

because $\left\|\Theta_{\eta_{l+1}} - E_l\right\| > \delta_{\eta_{l+1}}$; since the $(d-1)$-dimensional polytope Δ^*_{d-1} is the convex hull of $E_{d-1}, E_{d-2}, \dots, E_l$, and all vertices of the simplex having $x_d = x_{d-1} = \ldots = x_{l+1} = 0$; since the angle formed at E_l by the edges that connect it to any two other vertices of Δ^*_{l+1} is either a right angle (if only one of the vertices is $E_{d-1}, \dots, E_{l+1}$) or is

$$\frac{\pi}{2} < \arccos\left(-\frac{1}{l}\right) \le \frac{2\pi}{3}$$

for $l > 1$ (if both are vertices lying in the l-face having $x_d = \ldots = x_{l+1} = 0$, which is Δ^l: this scenario is not possible for $l = 1$ since, by definition, the only other vertex of the 1-face is excluded already) or is a nondegenerate acute angle (if both vertices are centres of two distinct faces: since the difference between two centres is always orthogonal to the lower-dimensional face, to which one of the centres belongs, the difference between these two centres, and the edges connecting them to E_l, form a nondegenerate right triangle); since the angle between the $(l+1)$st basis vector e_{l+1} and $\mathbf{1}$, the normal to the hyperplane, to which the simplex belongs, is

$$\frac{\pi}{4} < \arccos\left(\frac{1}{\sqrt{d}}\right) < \frac{\pi}{2}$$

for every $d \ge 3$; we can then conclude that on $\mathscr{E}^*_{l+1}$,

$$\frac{1}{l+1} - \Theta^{(l+1)}_{\eta_{l+1}} > c_{l+1}\delta_{\eta_{l+1}},$$

where c_{l+1} is a subunitary trigonometric constant bounded away from zero, such that

$$c_{l+1}\delta_{\eta_{l+1}} = \min_{\Theta\in\Delta^*_{l+1}} \left| \left(\delta_{\eta_{l+1}} \frac{\Theta - E_l}{\|\Theta - E_l\|} \right) \cdot e_{l+1} \right|.$$

We will now prove that for all s, on $G^s_{l+1} \cap \mathcal{S}_{l+1}$ almost surely $\Theta^{(l+1)}_k \longrightarrow 0$ with a similar argument as in *Step 2*, but making use of Lemma 9.1 in its full generality, so as to ensure that $R^{(l+1)}_n$ is positive increasing. By the monotonicity of the events G^s_{l+1} and since on $\mathscr{E}^*_{l+1}$ there are N^*_{l+2} and M_{l+1} almost surely finite, as in *Step 2*, we have that

$$\mathbb{P}(\mathscr{E}^*_{l+1} \smallsetminus G^s_{l+1}) \downarrow 0 \tag{9.11}$$

as $s \longrightarrow \infty$. Similarly, by a bound analogous to that achieved for $\mathcal{S}^c_{d-1}$, we will show that $\mathbb{P}(\mathcal{S}^c_{l+1})$ is bounded by a function that vanishes as $s \longrightarrow \infty$. Therefore the conclusion will be that almost surely on $\mathscr{E}^*_{l+1}$, $\Theta^{(l+1)}_k \longrightarrow 0$ (and $\Theta^{(i)}_k \longrightarrow 0$ for all $l+2 \le i \le d$ by the induction hypothesis). Consider the $(l+1)$st components $M_n^{(l+1)}$, $R_n^{(l+1)}$ and the sigma algebra $\mathscr{F}_{\eta_{l+1}}$. Defining $\mathscr{G}_n = \mathscr{F}_{\eta_{l+1}+n}$ by the exact same argument as in the previous two steps, *mutatis mutandis*, it holds that $M_n^{(l+1)}$ is a $\mathscr{G}_n$-martingale and $M_n^{(l+1)} \in \mathscr{L}^2(\mathscr{F}_{\eta_{l+1}})$ and is bounded (the same bound found on the previous components holds), hence it is almost surely convergent in $\mathscr{L}^2(\mathscr{F}_{\eta_{l+1}})$. Define the event

$$\mathcal{S}_{l+1} := \left\{ \sup_n M_n^{(l+1)} \le \frac{c_{l+1}\delta_{\eta_{l+1}}}{2} \right\}.$$

We first show that $\mathbb{P}(\mathcal{S}_{l+2}^c)$ has a vanishing upper bound, via Doob's inequality. Define

$$H_n^{l+1} := \left\{ \max_{k \le n} M_k^{(l+1)} > \frac{c_{l+1}\delta_{\eta_{l+1}}}{2} \right\}.$$

Then

$$\mathbb{P}_{\mathscr{F}_{\eta_{l+1}}}(H_n^{l+1}) \le \frac{4(M_n^{(l+1)})^2}{c_{l+1}^2\delta_{\eta_{l+1}}^2} \le \frac{4\tau_{\eta_{l+1}}^{2r}}{c_{l+1}^2\tau_{\eta_{l+1}}} = \frac{4}{c_{l+1}^2\tau_{\eta_{l+1}}^{1-2r}} \le \frac{4}{c_{l+1}^2\tau_s^{1-2r}}.$$

As in *Step 2* this implies that

$$\mathbb{P}(\mathcal{S}_{l+1}^c) < \frac{4}{c_{l+1}^2\tau_s^{1-2r}}. \tag{9.12}$$

Therefore, we can proceed as in *Step 2*, by proving by induction that for all $n \ge 0$ on $G_{l+1}^s \cap \mathcal{S}_{l+1}$,

$$\Theta_{\eta_{l+1}+n}^{(l+1)} < \frac{1}{l+1} - c_{l+1}\frac{\delta_{\eta_{l+1}}}{2}. \tag{9.13}$$

The case $n = 0$ follows by $G_{l+1}^s \cap \mathcal{S}_{l+1} \subseteq G_{l+1}^s \subseteq \mathscr{E}_{l+1}^*$, on which

$$\Theta_{\eta_{l+1}}^{(l+1)} < \frac{1}{l+1} - c_{l+1}\delta_{\eta_{l+1}}.$$

Assume the inductive hypothesis for all $j \in [n-1]$. To show (9.13) we will use Lemma 9.1 so as to ensure that $R_n^{(l+1)} \ge 0$. For all $j \in [n-1]$ set $k = \eta_{l+1} + j$, and recall that

$$F^{(l+1)}(x, y, x') := F^{(l+1)}(\underbrace{\Theta_k^{(2)}, \ldots, \Theta_k^{(l+1)}}_{x}, \underbrace{\Theta_k^{(l+2)}}_{y}, \underbrace{\Theta_k^{(l+3)}, \ldots, \Theta_k^{(d)}}_{x'}) = \psi^{(l+1)}(\Theta_k) - \Theta_k^{(l+1)}.$$

If we show that $R_n^{(l+1)}$ is positive, then the induction step will follow, since

$$\begin{aligned}\Theta_{\eta_{l+1}+n}^{(l+1)} &= \Theta_{\eta_{l+1}}^{(l+1)} + M_n^{(l+1)} - R_n^{(l+1)} < \Theta_{\eta_{l+1}}^{(l+1)} + M_n^{(l+1)} < \frac{1}{l+1} - c_{l+1}\delta_{\eta_{l+1}} \\ &+ c_{l+1}\frac{\delta_{\eta_{l+1}}}{2} = \frac{1}{l+1} - c_{l+1}\frac{\delta_{\eta_{l+1}}}{2}.\end{aligned}$$

The conditions that have been put in the definition of the stopping time η_{l+1} all intervene here, to ensure that the vector $(\Theta_k^{(2)}, \ldots, \Theta_k^{(d)}) \in P_{n_{l+1}^*}^{l+1} \cup S_{n_{l+1}^*}^{l+1}$, on which $F^{(l+1)}(x, y, x') < 0$, so that $-F^{(l+1)}(\Theta_k^{(2)}, \ldots, \Theta_k^{(d)}) > 0$ for all $\eta_{l+1} < k < \eta_{l+1} + n$ and therefore $R_n^{(l+1)}$ would be positive. Since $k > \eta_{l+1} > s \ge n_{l+1}^*$ and $\omega \in G_{l+1}^s \cap \mathcal{S}_{d-1} \subseteq G_{l+1}^s \subseteq \mathscr{E}_{l+1}^*$,

$$\Theta_k^{(i)} < \frac{\log \tau_k}{\tau_k} < y_{l+1}(\delta_k)$$

for all $l+2 \le i \le d$, for all considered k. The induction hypothesis ensures that

$$\Theta_k^{(l+1)} < \frac{1}{l+1} - c_{l+1}\frac{\delta_{\eta_{l+1}}}{2},$$

but the monotonicity of δ_k yields

$$\frac{1}{l+1} - c_{l+1}\frac{\delta_{\eta_{l+1}}}{2} > \frac{1}{l+1} - c_{l+1}\frac{\delta_{n^*_{l+1}}}{2}.$$

Thus we cannot conclude simply that $(\Theta_k^{(2)}, \ldots, \Theta_k^{(d)}) \in P^{l+1}_{n^*_{l+1}}$, but we can be sure that the vector $(\Theta_k^{(2)}, \ldots, \Theta_k^{(d)}) \in P^{l+1}_{n^*_{l+1}} \cup S^{l+1}_{n^*_{l+1}}$. Indeed the components lower than the $(l+1)$st have no constraints to satisfy on $P^{l+1}_{n^*_{l+1}} \cup S^{l+1}_{n^*_{l+1}}$, besides corresponding to points of the simplex, which is trivial; whereas the components higher than the $(l+1)$st are all satisfying the same requirement that the $(l+2)$nd component has to satisfy, in order to ensure that the vector is in $S^{l+1}_{n^*_{l+1}}$, were

$$\Theta_k^{(l+1)} > \frac{1}{l+1} - c_{l+1}\frac{\delta_{n^*_{l+1}}}{2}.$$

By the symmetries of the equation of the hyperplane T^*_{l+1} with respect to the $(x_{l+3}, \ldots, x_d)$-coordinate directions, discussed in Lemma 9.1 and in the introduction to Corollary 9.2, this is enough to ensure that also these coordinates of the vector are small enough to ensure $(\Theta_k^{(2)}, \ldots, \Theta_k^{(d)}) \in P^{l+1}_{n^*_{l+1}} \cup S^{l+1}_{n^*_{l+1}}$, because $\Theta_k^{(l+2)}, \ldots, \Theta_k^{(d)}$ would be such that, for every $l+2 \le j \le d$,

$$\Theta_k^{(j)} < T^*_{l+1}(\Theta_k^{(2)}, \ldots, \Theta_k^{(l+2)}, \ldots, \Theta_k^{(j-1)}, \Theta_k^{(j+1)}, \ldots, \Theta_k^{(d)}),$$

if it were true that

$$\Theta_k^{(l+1)} > \frac{1}{l+1} - c_{l+1}\frac{\delta_{n^*_{l+1}}}{2}$$

(that is, the vector would be in the tip of the polytope, in the worst case scenario). Hence $F^{(l+1)}(\Theta_k^{(2)}, \ldots, \Theta_k^{(d)}) < 0$ and the induction step is complete. As (9.13) holds on $G^s_{l+1} \cap \mathcal{S}_{l+1}$ for all $n \ge 0$, $\{R_n^{(l+1)}\}$ is positive and increasing on $G^s_{l+1} \cap \mathcal{S}_{l+1}$. At this point, *mutatis mutandis*, it is possible to conclude, as in *Step 2*, by contradiction, that $\Theta_k^{(l+1)} \longrightarrow 0$ almost surely on $G^s_{l+1} \cap \mathcal{S}_{l+1}$. Since s is arbitrary, through (9.11) and (9.12) we have that $\Theta_k^{(l+1)} \longrightarrow 0$ almost surely on $\mathscr{E}^*_{l+1}$. Also, since $\mathscr{E}^*_{l+1} \subseteq \mathscr{E}^*_{l+2}$, by induction hypothesis $\Theta_k^{(l+2)} \longrightarrow 0, \ldots, \Theta_k^{(d)} \longrightarrow 0$ almost surely on $\mathscr{E}^*_{l+1}$. Thus almost surely on $\mathscr{E}^*_{l+1}$, $\Theta_k^{(d)}, \ldots, \Theta_k^{(l+1)}$ vanish.

Step 4. The induction above terminates with $l = 1$, with the construction of the event $\mathscr{E}^*_2$, on which, in conclusion, $\Theta_k^{(2)}, \ldots, \Theta_k^{(d)}$ all vanish almost surely, meaning that $\Theta_k^{(1)} \longrightarrow 1$ almost surely on $\mathscr{E}^*_2$. Note that $\mathscr{E}^*_2$ is defined in terms of the subset of the simplex $\Delta^*_2 := \{x \in \Delta^{d-1} : x_d \le \ldots \le x_2 \le x_1\}$. Going through all the permutations of the coordinates, defines a covering of the simplex with $d!$ analogous sets (they all have equal Lebesgue measure, since the coordinate permutation elementary matrix has a determinant of 1). To be more rigorous with the notation, we could denote $\Delta^*_{d,\ldots,2,1}$ the set we constructed and $\mathscr{E}^*_{d,\ldots,2,1}$ the corresponding event, and similarly all the other $d!-1$ sets and events, with the corresponding ordering in the indices. Proposition 8.4 and Lemma 9.1 and the arguments in *Steps 1-3* have been presented in the decreasing order from d to 1, but there is nothing special about this order: the arguments can be repeated with respect to any order we were to choose for the coordinates. In each of these events $\mathscr{E}^*_{i_1,\ldots,i_d}$, where the permutation

$$\begin{pmatrix} 1 & 2 & \ldots & d \\ i_1 & i_2 & \ldots & i_d \end{pmatrix} \in S_d,$$

the last index i_d means that $\Theta_n^{(i_d)} \longrightarrow 1$ for almost every $\omega \in \mathscr{E}^*_{i_1,\ldots,i_d}$. The union of the sets $\Delta^*_{i_1,\ldots,i_d}$ covers the simplex and thus the union of all the events $\mathscr{E}^*_{i_1,\ldots,i_d}$ is an almost sure event, on which one component of Θ_n tends to 1. Hence $\mathbb{P}(\mathcal{D}) = 1$.

□

9.3 Monopoly in presence of feedback

Recall that monopoly is the event in which all but one of the bins receive finitely many balls. Then the proof of almost sure monopoly in presence of feedback follows easily for the regimes in which we just showed that dominance is almost sure: ρ_n is bounded; $\rho_n \longrightarrow \infty$, $\theta = 0$ and $\lambda < 1$.

Proof of Corollary 1.13. Since when $\alpha > 1$, if ρ_n is bounded (regime of growth covered by Lemma E.2) or, if it goes to infinity with $\theta = 0$ and $\lambda < 1$ (regimes of growth covered by Lemma F.1), $\mathbb{P}(\mathcal{D}) = 1$ by Theorem 1.11, by Lemmas E.2 and F.1 again, applied to each of the almost surely vanishing components, the corresponding bins get a bounded number of balls almost surely, hence the one bin whose proportions tend to 1 will get eventually all balls almost surely. Thus $\mathbb{P}(\mathcal{M}) = 1$ in the regimes, for which dominance has been shown to be almost sure. □

Part III

Supplements to Part I

Appendix A

The dynamical system with $\rho < 1$

In this section of the appendix we show how to generalise all the results for the dynamical system with $\rho = 1$ obtained in Chapter 3, to the dynamical system described by (2.35) and (2.36). We will not repeat all arguments when the variations are trivial, however, for self-containedness and ease of the reader, we will always rewrite the full statements of all results.

For one-step iterations arguments, a less cumbersome notation will sometimes be used, in order to omit the time index, and (2.35) and (2.36) will often be written as

$$\begin{aligned}\hat{q} &= M_p q\\ \hat{p} &= (1-\rho)p + \rho(\mathbf{1} - q - \hat{q}),\end{aligned}$$

where we recall that

$$M_p := \begin{pmatrix} 0 & \frac{p_3}{p_1+p_3} & \frac{p_2}{p_1+p_2} \\ \frac{p_3}{p_2+p_3} & 0 & \frac{p_1}{p_1+p_2} \\ \frac{p_2}{p_2+p_3} & \frac{p_1}{p_1+p_3} & 0 \end{pmatrix}.$$

A.1 Preliminaries

The same conventions as in Chapter 3 will be adopted, that is $p_0 \notin \partial\Sigma$ and $q_0 \notin \partial\Sigma$. We briefly justify this assumption. In the case of this dynamical system, with $p_0 \in \partial\Sigma \setminus V$ no consistency issue arises for the iteration if $q_0 \in V$. Nonetheless, we will follow the same conventions of Chapter 3, as we can take care of extremal cases easily in the next two remarks.

Remark A.1. *Let $p_0 \in E_i$ and $q_0 = v_i$ for some $i \in \{1, 2, 3\}$. Then $p_n = p_0$ for all $n \in \mathbb{N}$ and q_n is 2-periodic for all $n \in \mathbb{N}_0$.*

Proof. Without loss of generality, by symmetry, assume $i = 1$. Let $p_0 = (0, p_0^{(2)}, 1-p_0^{(2)})$ and denote $a = p_0^{(2)}$, $0 < a < 1$. We have that

$$q_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1-a & 0 & 0 \\ a & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1-a \\ a \end{pmatrix},$$

so

$$p_1 = (1-\rho)p_0 + \rho\left(\mathbf{1} - v_1 - \begin{pmatrix} 0 \\ 1-a \\ a \end{pmatrix}\right) = (1-\rho)p_0 + \rho\begin{pmatrix} 0 \\ a \\ 1-a \end{pmatrix} = p_0$$

and

$$q_2 = \begin{pmatrix} 0 & 1 & 1 \\ 1-a & 0 & 0 \\ a & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1-a \\ a \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = q_0.$$

By induction this shows q_n is 2-periodic, with p_n fixed. □

Remark A.2. *Let $p_0 \in E_i$ and $q_0 = v_j$ for some $i \neq j \in \{1,2,3\}$. Then $p_n = p_1$ for all $n \in \mathbb{N}$ and q_n is 2-periodic for all $n \in \mathbb{N}$.*

Proof. Without loss of generality, by symmetry, assume $i = 1$ and $j = 2$. Let $p_0 = (0, p_0^{(2)}, 1 - p_0^{(2)})$ and denote $a = p_0^{(2)}$, $0 < a < 1$. We have that

$$q_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1-a & 0 & 0 \\ a & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

so

$$p_1 = (1-\rho)p_0 + \rho(\mathbf{1} - v_1 - v_2) = (1-\rho)p_0 + \rho v_3 = \begin{pmatrix} 0 \\ (1-\rho)a \\ 1-(1-\rho)(a) \end{pmatrix}.$$

Note that there is no inconsistency issue, as $p_1 \notin V$. Thus, by letting $b = (1-\rho)a$, having $q_1 = (1,0,0)$ and $p_1 = (0, b, 1-b)$, we fall back in the case studied in Remark A.1. □

The inconsistency issue of Remark 3.3 is avoided through the convex combination with the past: even with q_0 not being the vertex corresponding to the edge on which p_0 lies, the dynamics is essentially the same as that described in Remark A.1. The statement equivalent to Remark 3.5 does not hold for this system, since due to the convex combination with the past, $p_1 \notin \partial\Sigma$ even if $q_0 \in V$. In particular, by a similar argument to that used in Remark 3.7, it is easy to establish the following.

Remark A.3. *Let $p_0 \notin \partial\Sigma$: if $q_0 \in E_i$ for some $i \in \{1,2,3\}$, $p_1 \notin \partial\Sigma$ and $q_1 \notin \partial\Sigma$; if $q_0 \in V$, $p_1 \notin \partial\Sigma$, $p_2 \notin \partial\Sigma$ and $q_2 \notin \partial\Sigma$.*

Proof. By symmetry, without loss of generality we can show the first case via explicit calculation for $q_0 = (0, a, 1-a)$, with $0 < a < 1$. Due to $p_0 \notin \partial\Sigma$, we will have $q_1 \notin \partial\Sigma$, since

$$q_1 = (a\phi(p_0^{(3)}, p_0^{(1)}) + (1-a)\phi(p_0^{(2)}, p_0^{(1)}), (1-a)\phi(p_0^{(1)}, p_0^{(2)}), a\phi(p_0^{(1)}, p_0^{(3)}))$$

and $\phi(p_0^{(i)}, p_0^{(j)}) = 0$ if and only if $p_0^{(i)} = 0$ (which is not allowed). Hence

$$p_1 = (1-\rho)p_0 + \rho(1 - q_1 - q_0) \notin \partial\Sigma$$

since $p_0 \notin \partial\Sigma$. For the second case, without loss of generality assume $q_0 = v_1$. Due to $p_0 \notin \partial\Sigma$, we will have $q_1 \in E_1$ since, by letting this time $a = \phi(p_0^{(2)}, p_0^{(3)})$, $0 < a < 1$, we have that

$$q_1 = \begin{pmatrix} 0 & * & * \\ 1-a & * & * \\ a & * & * \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1-a \\ a \end{pmatrix} \in E_1,$$

and due to the convex combination,

$$p_1 = (1-\rho)p_0 + \rho(1 - q_1 - q_0) \notin \partial\Sigma.$$

This falls back into the first case above, which will lead to $q_2 \notin \partial\Sigma$ and $p_2 \notin \partial\Sigma$. □

This means that also these extremal conditions have been reduced to regular orbits, via the following remark.

Remark A.4. *If $p_0 \notin \partial\Sigma$ and $q_0 \notin \partial\Sigma$, then $p_n \notin \partial\Sigma$ and $q_n \notin \partial\Sigma$ for all $n \in \mathbb{N}$.*

Proof. The proof proceeds as in Remark 3.8, by showing that if $p_n \notin \partial\Sigma$ and $q_n \notin \partial\Sigma$, then $p_{n+1} \notin \partial\Sigma$ and $q_{n+1} \notin \partial\Sigma$. Like in Remark 3.8

$$q_{n+1} = \begin{pmatrix} \phi(p_n^{(3)}, p_n^{(1)})q_n^{(2)} + \phi(p_n^{(2)}, p_n^{(1)})q_n^{(3)} \\ \phi(p_n^{(3)}, p_n^{(2)})q_n^{(1)} + \phi(p_n^{(1)}, p_n^{(2)})q_n^{(3)} \\ \phi(p_n^{(2)}, p_n^{(3)})q_n^{(1)} + \phi(p_n^{(1)}, p_n^{(3)})q_n^{(2)} \end{pmatrix} \notin \partial\Sigma.$$

As to p_{n+1}, note that

$$p_{n+1} = (1-\rho)p_n + \rho(\mathbf{1} - q_{n+1} - q_n) \notin \partial\Sigma,$$

because performing a nontrivial $(0 < \rho < 1)$ convex combination of a nonnegative vector with p_n yields automatically a vector in the interior of the simplex. □

As a result of these introductory remarks, unless otherwise stated, all orbits will be considered with initial conditions $p_0 \notin \partial\Sigma$ and $q_0 \notin \partial\Sigma$, which only yield *regular orbits*, that is orbits for which eventually $p_n \in \mathring{\Sigma}$ for all n. The result given by Remark 3.11 trivially holds true, due to its argument being not dependent on the equations of the system.

Remark A.5. *If $p_{n+1} - p_n \longrightarrow \mathbf{0}$ as $n \longrightarrow \infty$ and p_n does not converge to any of the vertices, then there is a subsequence $\{p_{n_j}\}$ bounded away from V.*

A.2 Fixed points and potential function

From (2.35) and (2.36) it is immediate to derive the fixed point equations

$$\begin{aligned} q &= M_p q \\ p &= (1-\rho)p + \rho(\mathbf{1} - 2q). \end{aligned}$$

Since the second equation is equivalent to $p = \mathbf{1} - 2q$, the same set of equilibrium points as in Chapter 3 is obtained: $\{(p, q_p) : \; p \in \Sigma_0\}$, where

$$q_p := \frac{\mathbf{1} - p}{2}.$$

The same potential $V(p,q) := \|q - q_p\|_1$ works for this system. In fact note that by (2.36) we have

$$\frac{\mathbf{1} - \hat{p}}{2} = \frac{\mathbf{1} - (1-\rho)p - \rho(\mathbf{1} - \hat{q} - q)}{2},$$

which can be rearranged as

$$\frac{\mathbf{1} - \hat{p}}{2} = (1-\rho)\frac{\mathbf{1} - p}{2} + \frac{\rho}{2}(q + \hat{q}),$$

from which it follows that

$$q_{\hat{p}} = (1-\rho)q_p + \frac{\rho}{2}(q + \hat{q}) \tag{A.1}$$

and therefore

$$\hat{q} - q_{\hat{p}} = \hat{q} - (1-\rho)q_p - \frac{\rho}{2}(\hat{q} - q) = \left(1 - \frac{\rho}{2}\right)\hat{q} - (1-\rho)q_p - \frac{\rho}{2}q = \left(1 - \frac{\rho}{2}\right)\hat{q}$$
$$- \left(1 - \frac{\rho}{2}\right) q_p + \frac{\rho}{2}q_p - \frac{\rho}{2}q = \left(1 - \frac{\rho}{2}\right)(\hat{q} - q_p) - \frac{\rho}{2}(q - q_p),$$

which ultimately yields

$$\hat{q} - q_{\hat{p}} = \left[\left(1 - \frac{\rho}{2}\right) M_p - \frac{\rho}{2} I\right](q - q_p). \tag{A.2}$$

Thus we can conclude with a remark analogous to Remark 3.13.

Remark A.6. *Since*

$$\|L_p\|_1 := \left\|\left(1 - \frac{\rho}{2}\right) M_p - \frac{\rho}{2} I\right\|_1 = 1$$

for all $p \in \Sigma_0$*, taking the norm on both sides of* (A.2) *yields* $V(p_{n+1}, q_{n+1}) \leq V(p_n, q_n)$*, and therefore* $V(p, q)$ *defines a* Lyapunov potential function *for this system. Moreover, there is* $0 \leq \ell \in \mathbb{R}$*, dependent on the initial conditions, such that* $V(p_n, q_n) \longrightarrow \ell$ *as* $n \longrightarrow \infty$*.*

Orbits do not get stuck at equilibria for this system either.

Remark A.7. *If* $p_0 \in \Sigma_0$ *and* $q_{p_0} \neq q_0 \in \Sigma_0$*, then for all* $p \in \Sigma_0$*,* $(p_n, q_n) \neq (p, q_p)$ *for all* $n \in \mathbb{N}$*.*

Proof. If $q_{n+1} = q_{p_{n+1}}$, let us write $p_{n+1} = p$ and $q_{n+1} = q_p$ for some $n \in \mathbb{N}_0$, then by (A.2) it follows that

$$\mathbf{0} = L_{p_n}(q_n - q_{p_n}),$$

implying that $q_n - q_{p_n}$ is in the kernel of L_{p_n}. We will now show that L_{p_n} is invertible. Recall that by Lemma 3.19 (a, b) M_{p_n} has an eigenvalue of 1 and two other nonpositive eigenvalues, bigger or equal than -1. Let then λ be one of these eigenvalues. Then L_{p_n} has a corresponding eigenvalue of

$$\left(1 - \frac{\rho}{2}\right)\lambda - \frac{\rho}{2}.$$

Assume any of these eigenvalues of L_{p_n} to be zero, by contradiction. Then since $0 < \rho < 1$,

$$\lambda = \frac{\frac{\rho}{2}}{1 - \frac{\rho}{2}} > 0.$$

Hence $\lambda = 1$. But then

$$1 - \frac{\rho}{2} = \frac{\rho}{2}$$

which is equivalent to $\rho = 1$, which is a contradiction. Since L_{p_n} is invertible, its kernel is trivial and $q_n - q_{p_n} = \mathbf{0}$. Then by (2.36)

$$p = (1-\rho)p_n + \rho(\mathbf{1} - q_p - q_{p_n}) = (1-\rho)p_n + \frac{\rho}{2}(p + p_n),$$

which can be rearranged into

$$\left(1 - \frac{\rho}{2}\right) p = \left(1 - \frac{\rho}{2}\right) p_n,$$

yielding $p_n = p$ and thus $q_n = q_p$. The argument can be concluded by induction as in Remark 3.12. ∎

Lemma A.8. *For every $p \notin \partial\Sigma$ and $q \neq q_p$,*

$$V(\hat{p},\hat{q}) < V(p,q).$$

Therefore, the potential is eventually strictly decreasing along the regular orbits of the dynamical system.

Proof. If $q = q_p$, then $\hat{q} = q = q_p$ and $\hat{p} = p$. Therefore, $V(\hat{p},\hat{q}) = V(p,q) = 0$. Assume $q \neq q_p$. By (A.2), $V(\hat{p},\hat{q}) < V(p,q)$ if and only if

$$\left\| \left[\left(1 - \frac{\rho}{2}\right) M_p - \frac{\rho}{2} I \right] \frac{q - q_p}{\|q - q_p\|_1} \right\|_1 < 1.$$

To show this, one proceeds exactly as in Lemma 3.14, but for the newly defined matrix L_p of this section. Defining

$$v := \frac{q - q_p}{\|q - q_p\|_1}$$

and using the vertex of the hexagon $v = (1/2, 0, -1/2)$ without loss of generality, the computation yields

$$L_p v = \frac{1}{4} \begin{pmatrix} -\rho - (2-\rho)\frac{p_2}{p_1+p_2} \\ (2-\rho)\left(\frac{p_3}{p_2+p_3} - \frac{p_1}{p_1+p_2}\right) \\ \rho + (2-\rho)\frac{p_2}{p_2+p_3} \end{pmatrix}$$

and then, as a result, if $p \notin \partial\Sigma$,

$$\begin{aligned} \|L_p v\|_1 &= \frac{1}{4}\left[2\rho + (2-\rho)\left(\frac{p_2}{p_1+p_2} + \frac{p_2}{p_2+p_3} + \left|\frac{p_3}{p_3+p_2} - \frac{p_1}{p_1+p_2}\right|\right)\right] \\ &= \begin{cases} \frac{1}{2}\left(\rho + (2-\rho)\frac{p_2}{p_1+p_2}\right) < 1 & , \frac{p_3}{p_3+p_2} \geq \frac{p_2}{p_1+p_2} \\ \frac{1}{2}\left(\rho + (2-\rho)\frac{p_2}{p_2+p_3}\right) < 1 & , \frac{p_3}{p_3+p_2} < \frac{p_1}{p_1+p_2} \end{cases} \end{aligned} \tag{A.3}$$

because

$$\frac{1}{2}\left(\rho + (2-\rho)\frac{p_2}{p_1+p_2}\right) = \frac{1}{2}\left[\rho\left(1 - \frac{p_2}{p_1+p_2}\right) + 2\frac{p_2}{p_1+p_2}\right] < \frac{1}{2}\left(1 + \frac{p_2}{p_1+p_2}\right) < 1$$

and

$$\frac{1}{2}\left(\rho + (2-\rho)\frac{p_2}{p_2+p_3}\right) < \frac{1}{2}\left(1 + \frac{p_2}{p_2+p_3}\right) < 1,$$

due to the fact that none of the coordinates of $p \notin \partial\Sigma$ is zero. Note that since by Remark A.4 $p_n \notin \partial\Sigma$ for all $n \in \mathbb{N}_0$ and by Remark A.7 $q_n \neq q_{p_n}$, we can conclude that the potential is strictly decreasing along the orbits. □

We proceed with the same strategy as the one outlined at the end of Section 3.2. In the following sections we will show convergence of $\{p_n\}$ when:

- $\{p_n\}$ is bounded away from the boundary (Section A.3);
- $\ell = 0$ and $\{p_n\}$ is not bounded away from the boundary (Section A.4);
- $\ell > 0$ and $\{p_n\}$ is not bounded away from the boundary (Section A.5).

A.3 Convergence bounded away from the boundary

The main goal of this section is showing the convergence of the dynamical system when $\{p_n\}$ is known to be bounded away from the boundary of the simplex.

Proposition A.9. *If $\{p_n\}$ is bounded away from $\partial\Sigma$, there is a constant $0 < \mathcal{c} < 1$, dependent on the initial condition, such that*

$$V(p_{n+1}, q_{n+1}) < \mathcal{c}V(p_n, q_n).$$

Hence

$$\ell := \lim_{n\longrightarrow\infty} V(p_n, q_n) = 0,$$

and the dynamical system converges to an internal equilibrium.

Proof. The proof of the geometric decaying upper bond and that $\ell = 0$ is the same as that in Proposition 3.15, with the same definition of $\mathcal{c}$. As to the convergence of the dynamical system to an internal equilibrium, noting that

$$\hat{p} - p = \rho(2q_p - q - \hat{q}) = \rho[M_p(q_p - q) + q_p - q],$$

it similarly follows that

$$\|\hat{p} - p\|_1 \leq 2\rho V(p, q). \tag{A.4}$$

The geometric decaying upper bound ensured by $\{p_n\}$ being bounded away from $\partial\Sigma$, ensures that $\sum_{n=0}^{\infty} V(p_n, q_n) < \infty$ as in Proposition 3.15. □

A result analogous to Proposition 3.17 still holds. It is enough to invoke Proposition A.9 instead of Proposition 3.15 to get the constant $\mathcal{c}$ of the geometric decay, as it is similarly defined, and to note that the two other crucial estimates of the argument in Proposition 3.17 were: (3.2), which we also know to hold in this case by (A.2); $\|p_{n+1} - p_n\|_1 \leq 2\|q_n - q_{p_n}\|_1$, which is still true for $\rho < 1$, as we have just shown in the proof of Proposition A.9 that $\|p_{n+1} - p_n\|_1 \leq 2\rho\|q_n - q_{p_n}\|_1 < 2\|q_n - q_{p_n}\|_1$. Hence, with no substantial changes to the proof, just by replacing the L_{p_n} of Chapter 3 with the one defined in this section, the corresponding result holds. Recall that $U((p, q), r, r') := B(p, r) \times B(q, r')$ and dist are with respect to the 1-norm, where $U((p, q), r) := U((p, q), r, r)$ and $B(p, r)$ is the ball centred at p of radius r in 1-norm.

Proposition A.10. *For every $p \notin \partial\Sigma$ and a small enough $0 < \varepsilon' < \operatorname{dist}(p, \partial\Sigma)$, there is a $\delta' > 0$ suitably smaller than ε' such that, if $(p_0, q_0) \in U((p, q_p), \delta')$ then $(p_n, q_n) \in B((p, q_p), \varepsilon', \varepsilon'/2)$ for all $n \in \mathbb{N}$.*

By Proposition A.10 and the final convergence claim in Proposition A.9, the following holds.

Corollary A.11. *If (p_0, q_0) is close enough to an internal equilibrium (p, q_p), the system converges to a (possibly different) internal equilibrium.*

A.4 Convergence to the boundary with $\ell = 0$

The main goal of this section is to show that if $\{p_n\}$ approaches the boundary and the limit of the potential $\ell := \lim_{n\longrightarrow\infty} V(p_n, q_n) = 0$, the dynamical system converges. The setup will be the same as the one at the beginning of Section 3.4, with the same notation, since we have recovered both Remark A.5 and (A.4). Then we can immediately proceed to establish similar eigencoordinate iteration formulas.

Lemma A.12.

$$\hat{\alpha} = \alpha\left[\left(1-\frac{\rho}{2}\right)\lambda_0(p) - \frac{\rho}{2}\right]\frac{\begin{vmatrix} e_0^{(i)}(p) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(p) & e_{-1}^{(j)}(\hat{p})\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p})\end{vmatrix}} + \beta\left[\left(1-\frac{\rho}{2}\right)\lambda_{-1}(p) - \frac{\rho}{2}\right]\frac{\begin{vmatrix} e_{-1}^{(i)}(p) & e_{-1}^{(i)}(\hat{p}) \\ e_{-1}^{(j)}(p) & e_{-1}^{(j)}(\hat{p})\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p})\end{vmatrix}} \tag{A.5}$$

$$\hat{\beta} = \alpha\left[\left(1-\frac{\rho}{2}\right)\lambda_0(p) - \frac{\rho}{2}\right]\frac{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_0^{(i)}(p) \\ e_0^{(j)}(\hat{p}) & e_0^{(j)}(p)\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p})\end{vmatrix}} + \beta\left[\left(1-\frac{\rho}{2}\right)\lambda_{-1}(p) - \frac{\rho}{2}\right]\frac{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(p) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(p)\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p})\end{vmatrix}} \tag{A.6}$$

Proof. As in Corollary 3.39, by (A.2) we obtain a system of three linear equations in two variables $(\hat{\alpha}, \hat{\beta}) \in \mathbb{R}^2$,

$$\hat{\alpha}e_0(\hat{p}) + \hat{\beta}e_{-1}(\hat{p}) = \alpha\left[\left(1-\frac{\rho}{2}\right)\lambda_0(p) - \frac{\rho}{2}\right]e_0(p) + \beta\left[\left(1-\frac{\rho}{2}\right)\lambda_{-1}(p) - \frac{\rho}{2}\right]e_{-1}(p),$$

which can therefore be solved by picking any two of the three equations as per the discussion in Corollary 3.39. Then we have that, for any $i \neq j \in \{1, 2, 3\}$ chosen, the linear system

$$\begin{pmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p})\end{pmatrix}\begin{pmatrix}\hat{\alpha} \\ \hat{\beta}\end{pmatrix} = \begin{pmatrix} \alpha\left[\left(1-\frac{\rho}{2}\right)\lambda_0(p) - \frac{\rho}{2}\right]e_0^{(i)}(p) + \beta\left[\left(1-\frac{\rho}{2}\right)\lambda_{-1}(p) - \frac{\rho}{2}\right]e_{-1}^{(i)}(p) \\ \alpha\left[\left(1-\frac{\rho}{2}\right)\lambda_0(p) - \frac{\rho}{2}\right]e_0^{(j)}(p) + \beta\left[\left(1-\frac{\rho}{2}\right)\lambda_{-1}(p) - \frac{\rho}{2}\right]e_{-1}^{(j)}(p)\end{pmatrix}$$

has a unique solution, which can be calculated by Cramer's rule,

$$\hat{\alpha} = \frac{\begin{vmatrix} \alpha\left[\left(1-\frac{\rho}{2}\right)\lambda_0(p) - \frac{\rho}{2}\right]e_0^{(i)}(p) + \beta\left[\left(1-\frac{\rho}{2}\right)\lambda_{-1}(p) - \frac{\rho}{2}\right]e_{-1}^{(i)}(p) & e_{-1}^{(i)}(\hat{p}) \\ \alpha\left[\left(1-\frac{\rho}{2}\right)\lambda_0(p) - \frac{\rho}{2}\right]e_0^{(j)}(p) + \beta\left[\left(1-\frac{\rho}{2}\right)\lambda_{-1}(p) - \frac{\rho}{2}\right]e_{-1}^{(j)}(p) & e_{-1}^{(j)}(\hat{p})\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p})\end{vmatrix}}$$

$$\hat{\beta} = \frac{\begin{vmatrix} e_0^{(i)}(\hat{p}) & \alpha\left[\left(1-\frac{\rho}{2}\right)\lambda_0(p) - \frac{\rho}{2}\right]e_0^{(i)}(p) + \beta\left[\left(1-\frac{\rho}{2}\right)\lambda_{-1}(p) - \frac{\rho}{2}\right]e_{-1}^{(i)}(p) \\ e_0^{(j)}(\hat{p}) & \alpha\left[\left(1-\frac{\rho}{2}\right)\lambda_0(p) - \frac{\rho}{2}\right]e_0^{(j)}(p) + \beta\left[\left(1-\frac{\rho}{2}\right)\lambda_{-1}(p) - \frac{\rho}{2}\right]e_{-1}^{(j)}(p)\end{vmatrix}}{\begin{vmatrix} e_0^{(i)}(\hat{p}) & e_{-1}^{(i)}(\hat{p}) \\ e_0^{(j)}(\hat{p}) & e_{-1}^{(j)}(\hat{p})\end{vmatrix}},$$

which yields the iteration given in (A.5) and (A.6), where the ratios of the determinants do not depend on the choice of $i \neq j$, by the same reasoning as in Corollary 3.39. □

The Taylor expansion about $(p_1, \alpha, \beta) = (0, 0, 0)$ will now be possible as in Section 3.4, having analogous iteration formulas.

Lemma A.13.

$$\hat{p}_1 = p_1 + \rho_1(r) \tag{A.7}$$
$$\hat{p}_2 = p_2 - 2\rho(1-p_2)p_1\beta + \rho_2(r), \tag{A.8}$$

where $\rho_1(r) = \mathcal{O}(\beta p_1, \alpha p_1)$ *and* $\rho_2(r) = \mathcal{O}(\alpha, \beta p_1^2)$

Proof. Since

$$\hat{p} - p = \rho(\mathbf{1} - p - \hat{q} - q) = \rho(2q_p - M_p q - q) = -\rho(M_p + I)(q - q_p),$$

we have that

$$\hat{p} = p - \rho\alpha(1 + \lambda_0(p))e_0(p) - \rho\beta(1 + \lambda_{-1}(p))e_{-1}(p), \tag{A.9}$$

from which, reading off the first two components and applying Lemma 3.19 (f, g, h), it follows that

$$\begin{aligned}
\hat{p}_1 &= p_1 - \rho\alpha(1 - 2p_1 + \mathcal{O}(p_1^2))\mathcal{O}(p_1) - \rho\beta(2p_1 + \mathcal{O}(p_1^2))(-1 + \mathcal{O}(p_1)) \\
&= p_1 + \mathcal{O}(\alpha p_1, \beta p_1) \\
\hat{p}_2 &= p_2 - \rho\alpha(1 - 2p_1 + \mathcal{O}(p_1^2))(1 + \mathcal{O}(p_1)) - \rho\beta(2p_1 + \mathcal{O}(p_1^2))(1 - p_2 + \mathcal{O}(p_1)) \\
&= p_2 - 2\rho(1 - p_2)p_1\beta + \mathcal{O}(\alpha, \beta p_1^2),
\end{aligned}$$

having used the smoothness of the eigenvectors to linearise, as p approaches the edge E_1, and the relative compactness of $\mathscr{K}^*_{\varepsilon,\frac{\delta}{8}}$ to estimate uniformly the Jacobian term as in the conclusion of Lemma 3.21. □

Lemma A.14.

$$\hat{\alpha} = -\frac{\rho}{2}\alpha(1 + \rho_3(r)) + \rho_4(r) \tag{A.10}$$

$$\hat{\beta} = -\beta\left[1 - (2 - \rho)\, p_1\right] + \rho_5(r), \tag{A.11}$$

where $\rho_3(r) = \mathcal{O}(\alpha, p_1)$*,* $\rho_4(r) = \mathcal{O}(\beta\alpha, \beta^2 p_1)$*,* $\rho_5(r) = \mathcal{O}(\alpha^2, \alpha\beta, \beta^2 p_1, \beta p_1^2)$*.*

Proof. As in Lemma 3.22, by Lemma A.13 it follows that $\hat{p} = p + \mathcal{O}(\boldsymbol{\alpha}, \boldsymbol{\beta p_1})$. We plug this estimate, along with that of Lemma 3.19 (g), in (A.5) and (A.6), precisely in the terms next to α and β, exactly as done in Lemma 3.22. The only difference are the coefficients involving ρ, so there is no need to show the explicit calculation again. It yields, due to smoothness of the eigenvectors' components and relative compactness of $K^*_{\varepsilon,\frac{\delta}{8}}$, the following estimates for those terms involved in (A.5),

$$\begin{aligned}
\hat{\alpha} &= \alpha\left(-\frac{\rho}{2} + \mathcal{O}(p_1)\right)(1 + \mathcal{O}(\alpha, \beta p_1)) + \beta\left(-\left(1 - \frac{\rho}{2}\right) - \frac{\rho}{2} + \mathcal{O}(p_1)\right)\mathcal{O}(\alpha, \beta p_1) \\
&= \alpha\left(-\frac{\rho}{2} + \mathcal{O}(p_1)\right)(1 + \mathcal{O}(\alpha, \beta p_1)) + \beta\left(-1 + \mathcal{O}(p_1)\right)\mathcal{O}(\alpha, \beta p_1) \\
&= \alpha\left(-\frac{\rho}{2} + \mathcal{O}(\alpha, p_1)\right) + \beta\mathcal{O}(\alpha, \beta p_1) = -\frac{\rho}{2}\alpha(1 + \mathcal{O}(\alpha, p_1)) + \mathcal{O}(\beta\alpha, \beta^2 p_1),
\end{aligned}$$

and the following estimates for those terms in (A.6),

$$\begin{aligned}
\hat{\beta} &= \alpha\left(-\frac{\rho}{2} + \mathcal{O}(p_1)\right)\mathcal{O}(\alpha, \beta p_1) + \beta\left[\left(1 - \frac{\rho}{2}\right)\left(-1 + 2p_1 + \mathcal{O}(p_1^2)\right) - \frac{\rho}{2}\right](1 + \mathcal{O}(\alpha, \beta p_1)) \\
&= \mathcal{O}(\alpha^2, \alpha\beta p_1) + \beta\left[-1 + (2 - \rho)\, p_1 + \mathcal{O}(p_1^2) + \mathcal{O}(\alpha, \beta p_1)\right] \\
&= \mathcal{O}(\alpha^2, \alpha\beta p_1) - \beta\left[1 - (2 - \rho)\, p_1\right] + \mathcal{O}(\alpha\beta, \beta^2 p_1, \beta p_1^2) \\
&= -\beta\left[1 - (2 - \rho)\, p_1\right] + \mathcal{O}(\alpha^2, \alpha\beta, \beta^2 p_1, \beta p_1^2),
\end{aligned}$$

since the same considerations as in Lemma 3.22 hold, about the determinantal terms. □

Due to Lemmas A.13 and A.14 holding as in Section 3.4, no substantial changes from Lemma 3.23 are involved in the proof of the errors' estimates.

Lemma A.15. *Let the constant* $\theta \coloneqq 1/16$. *There is a constant* $c > 0$ *such that for all sufficiently small* ε, *on the closure* $\overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$ *we have that*

$$\begin{aligned}
&|\rho_1(r)| < \theta p_1\\
&|\rho_2(r)| < c|\alpha| + p_1|\beta|\\
&|\rho_3(r)| < \theta\\
&|\rho_4(r)| < \theta|\alpha| + \theta p_1|\beta|\\
&|\rho_5(r)| < \theta|\alpha| + \theta p_1|\beta|.
\end{aligned}$$

It will be necessary, for further arguments, to add a requirement on ε, that given δ, c, θ,

$$\varepsilon < \min\left\{\theta, \frac{\delta(1-2\theta)}{16(3+c)}\right\}.$$

We continue by adopting the same exact setup as in the corresponding part of Section 3.4 immediately preceding Lemma 3.24. Indeed the counterpart of Lemma 3.24 holds with no changes to the proof, due to Proposition 3.17 having been extended to this case by Proposition A.10 with no changes either, the constant $\mathscr{c}$ being similarly defined in Lemma A.8 after redefining L_{p_n} with the new corresponding matrix.

Corollary A.16. *There is a* $k \geq K$ *large enough such that, having defined* $m \coloneqq n_k$, *for all* $m \leq n < \tau_k$, $p_n^{(1)} \leq \varepsilon$.

In the following, all the proofs are made with respect to this large enough $m = n_k$, and therefore the corresponding τ_k will be simply denoted as τ.

Lemma A.17. *For all* $m \leq n \leq \tau$,

$$|\alpha_n| \leq \max\left\{\left(\frac{3}{4}\right)^{n-m}|\alpha_m|, p_n^{(1)}|\beta_n|\right\}.$$

Proof. The same proof by induction of Lemma 3.25 works with the new iterative formulas established in Lemma A.14. If $n = m$, the statement $|\alpha_m| \leq \max\{|\alpha_m|, p_m^{(1)}|\beta_m|\}$ is trivially true. If $n = m+1 < \tau$, it hods that

$$|\alpha_{m+1}| \leq \frac{\rho}{2}|\alpha_m|(1+|\rho_3(r_m)|) + |\rho_4(r_m)| \leq \frac{|\alpha_m|}{2}(1+\theta) + \theta|\alpha_m| + \theta p_m^{(1)}|\beta_m|$$

by definition of m. Then

$$|\alpha_{m+1}| \leq |\alpha_m|\frac{1+3\theta}{2} + \theta p_m^{(1)}|\beta_m|.$$

If $|\alpha_m| \geq p_m^{(1)}|\beta_m|$, then

$$|\alpha_{m+1}| \leq \frac{1+5\theta}{2}|\alpha_m| \leq \frac{3}{4}|\alpha_m|,$$

as $\theta \leq 1/16 < 1/10$ implies $(1+5\theta)/2 < 3/4$. If instead $|\alpha_m| < p_m^{(1)}|\beta_m|$, then it holds that

$$|\alpha_{m+1}| \leq \frac{1+5\theta}{2} p_m^{(1)}|\beta_m|. \tag{A.12}$$

By definition of m and by Lemmas A.13 and A.15 the exact same as the corresponding argument in Lemma 3.25 holds, hence

$$p_{m+1}^{(1)} \geq (1-\theta)p_m^{(1)}, \tag{A.13}$$

since at time m the orbit is in $\overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$. Note also that since by hypothesis $|\alpha_m| < |\beta_m|$ and by assumption $\varepsilon \le \theta$, it follows, by definition of m, that $p^{(1)}_m < \varepsilon \le \theta$. This yields, by applying Lemmas A.14 and A.15, that

$$\begin{aligned}|\beta_{m+1}| &\ge |\beta_m|\,[1-(2-\rho)\,p^{(1)}_m] - |\rho_5(r_m)| > |\beta_m|(1-2p^{(1)}_m) - |\rho_5(r_m)| \\ &> |\beta_m|(1-2p^{(1)}_m) - \theta|\alpha_m| - \theta p^{(1)}_m|\beta_m| > |\beta_m|(1-2p^{(1)}_m) - 2\theta p^{(1)}_m|\beta_m| \\ &= \beta_m|(1-2(1+\theta)p^{(1)}_m) > |\beta_m|(1-2\theta(1+\theta)),\end{aligned}$$

and thus

$$|\beta_{m+1}| > |\beta_m|(1-3\theta), \tag{A.14}$$

because $0 < \theta < 1/2$ ensures that $2\theta(1+\theta) < 3\theta$. Plugging the bounds of (A.13) and (A.14) into (A.12) yields that

$$\alpha_{m+1} \le \frac{1+5\theta}{2}\frac{p^{(1)}_{m+1}|\beta_{m+1}|}{(1-\theta)(1-3\theta)} < p^{(1)}_{m+1}|\beta_{m+1}|.$$

Assume now the thesis for any $m+1 \le n < \tau$. This time, when carrying out the inductive step, it will not be possible to appeal to the definition of m, but it will be necessary to rely on Corollary A.16, which ensures that $(p_n, q_n) \in \overline{\mathscr{K}}^*_{\varepsilon,\frac{\delta}{8}}$, and therefore makes it possible for the same aforementioned lemmas to apply in the corresponding steps, which are repeated trivially as in Lemma 3.25, with very elementary changes, similar to those shown in the base case for the induction. It will not be necessary to show them explicitly. □

The time σ is also defined as in Section 3.4.

Lemma A.18.

a) If $\tau < \infty$, then $\sigma < \tau$ and $p^{(2)}_\sigma \in [\delta/4, 1-\delta/4]$.

b) If $\sigma < \infty$, then for all $\sigma \le n \le \tau$,

$$\left(\frac{3}{4}\right)^{n-m}|\alpha_m| \le p^{(1)}_n|\beta_n|.$$

Proof.

a) For all $m \le n \le \sigma\wedge\tau$, by definition of σ, $p^{(1)}_n|\beta_n| < (3/4)^{n-m}|\alpha_m|$, and as a result by Lemma A.17, $\alpha_n \le (3/4)^{n-m}|\alpha_m|$. By Lemmas A.13 and A.15 and Corollary A.16 it follows that

$$\begin{aligned}|p^{(2)}_{\sigma\wedge\tau} - p^{(2)}_m| &\le \sum_{n=m}^{\sigma\wedge\tau-1}|p^{(2)}_{n+1} - p^{(2)}_n| = \sum_{n=m}^{\sigma\wedge\tau-1}|2\rho(1-p^{(2)}_n)p^{(1)}_n\beta_n - \rho_2(r_n)| \\ &< 2\sum_{n=m}^{\sigma\wedge\tau-1}p^{(1)}_n|\beta_n| + \sum_{n=m}^{\sigma\wedge\tau-1}c|\alpha_n| + p^{(1)}_n|\beta_n| = 3\sum_{n=m}^{\sigma\wedge\tau-1}p^{(1)}_n|\beta_n| + \\ &\quad c\sum_{n=m}^{\sigma\wedge\tau-1}|\alpha_n| \le (3+c)|\alpha_m|\sum_{n=m}^{\sigma\wedge\tau-1}\left(\frac{3}{4}\right)^{n-m} < 4(3+c)\varepsilon \\ &< 4(3+c)\frac{\delta(1-2\theta}{16(3+c)} < \frac{\delta}{4}.\end{aligned}$$

The same argument as in Lemma 3.26 implies that $\sigma < \tau$, and in particular that $p^{(2)}_\sigma \in [\delta/4, 1-\delta/4]$.

b) Same proof as in Lemma 3.26 (b).

□

Theorem A.19. *$\tau = \infty$ and the dynamical system converges.*

Proof. Suppose first $\sigma = \infty$. By Lemma A.18 (a), if $\tau < \infty$, $\sigma < \tau < \infty$ against the hypothesis, hence $\tau = \infty$. By Lemmas A.13, A.15 and A.17 and Corollary A.16 and the definition of σ, it follows that

$$\sum_{n=m}^{\infty} |p_{n+1}^{(2)} - p_n^{(2)}| = \sum_{n=m}^{\infty} |2\rho(1-p_n^{(2)})p_n^{(1)}\beta_n + \rho_2(r_n)| \leq 3\sum_{n=m}^{\infty} p_n^{(1)}|\beta_n| + c\sum_{n=m}^{\infty}|\alpha_n|$$
$$\leq (3+c)|\alpha_m| \sum_{n=m}^{\infty} \left(\frac{3}{4}\right)^{n-m} < \frac{\delta}{4},$$

so $p_n^{(2)}$ converges within $[\delta/4, 1-\delta/4] \subset [\delta/8, 1-\delta/8]$. Then the same conclusion, as in the corresponding case in Theorem 3.27, follows yielding the claim.

Suppose now $\sigma < \infty$. By Lemma A.18 (b), $\sigma < \tau$. By Lemma A.14, for any $\sigma \leq k < \tau$,

$$\beta_{k+1} + (-1)^{k-\sigma}\beta_\sigma = \sum_{n=\sigma}^{k}(-1)^{k-n}(\beta_{n+1}+\beta_n) = \sum_{n=\sigma}^{k}(-1)^{k-n}\left[(2-\rho)\,\beta_n p_n^{(1)} + \rho_5(r_n)\right].$$

By the definition of σ, Lemmas A.15 and A.17 and Corollary A.16, for all $\sigma \leq n \leq k$, $|\rho_5(r_n)| < 2\theta p_n^{(1)}|\beta_n|$, so it follows that $\{\beta_n\}_{n=\sigma}^{k}$ has alternating signs, since if $\beta_n \neq 0$,

$$\beta_{n+1}\beta_n = \left[-1 + (2-\rho)\,p_n^{(1)}\right]\beta_n^2 + \beta_n\rho_5(r_n) < \left[-1 + 2p_n^{(1)}\right]\beta_n^2 + |\beta_n||\rho_5(r_n)|$$
$$\leq (-1 + 2p_n^{(1)} + 2\theta p_n^{(1)})\beta_n^2 < [-1 + 2\theta(1+\theta)]\beta_n^2 < 0$$

by Corollary A.16 and $\theta < 1/4$, which ensures that $2\theta(1+\theta) < 1$, because $2x(1+x)$ is increasing on the positive reals and valued 0 at 0 and 1 at $(\sqrt{3}-1)/2 > 1/4$, which follows from $\sqrt{3} > 3/2$ (which is equivalent to $12 > 9$). Clearly if $\beta_n = 0$ there is no contribution made towards the sum we are interested in estimating, which is $\sum_{n=\sigma}^{k} p_n^{(1)}|\beta_n|$, and the zero term can just be neglected. Since for all $k \geq m$, $|\beta_k| < \varepsilon$, and the sign alternates as aforementioned,

$$2\varepsilon > |\beta_{k+1} + (-1)^{k-\sigma}\beta_\sigma| = \left|\sum_{n=\sigma}^{k}(-1)^{k-n}\left((2-\rho)\,\beta_n p_n^{(1)} + \rho_5(r_n)\right)\right|$$
$$= \left|\sum_{n=\sigma}^{k}(-1)^{k-n}\mathrm{sign}(\beta_n)\,(2-\rho)\,|\beta_n|p_n^{(1)} + \sum_{n=\sigma}^{k}(-1)^{k-n}\rho_5(r_n)\right| \geq$$
$$(2-\rho)\sum_{n=\sigma}^{k}|\beta_n|p_n^{(1)} - \left|\sum_{n=\sigma}^{k}(-1)^{k-n}\rho_5(r_n)\right| > \sum_{n=\sigma}^{k}|\beta_n|p_n^{(1)} - \sum_{n=\sigma}^{k}|\rho_5(r_n)|$$
$$\geq (1-2\theta)\sum_{n=\sigma}^{k}|\beta_n|p_n^{(1)} \geq 0.$$

In conclusion it has been shown that

$$\sum_{n=\sigma}^{k} p_n^{(1)}|\beta_n| < \frac{2\varepsilon}{1-2\theta}. \tag{A.15}$$

The main argument is the same as that of the second part of Theorem 3.27 then: we show $\tau = \infty$ by contradiction. If $\tau < \infty$ one can use (A.15) with $k = \tau - 1$, and therefore by Lemmas A.13, A.15, A.17 and A.18 and Corollary A.16, we have, by the same estimate as in the previous case ($\sigma = \infty$, more precisely Lemma A.18 (a) ensures that $\sigma < \tau$, while Lemma A.18 (b) ensures that $(3/4)^{n-m}|\alpha_m| \le p_n^{(1)}|\beta_n|$), that

$$|p_\sigma^{(2)} - p_\tau^{(2)}| \le \sum_{n=\sigma}^{\tau-1} |p_{n+1}^{(2)} - p_n^{(2)}| < 3\sum_{n=\sigma}^{\tau-1} p_n^{(1)}|\beta_n| + c\sum_{n=\sigma}^{\tau-1} |\alpha_n| \le 3\sum_{n=\sigma}^{\tau-1} p_n^{(1)}|\beta_n|$$
$$+ c\sum_{n=\sigma}^{\tau-1} \max\left\{\left(\frac{3}{4}\right)^{n-m}|\alpha_m|, p_n^{(1)}|\beta_n|\right\} = 3\sum_{n=\sigma}^{\tau-1} p_n^{(1)}|\beta_n| + c\sum_{n=\sigma}^{\tau-1} p_n^{(1)}|\beta_n|$$
$$= (3+c)\sum_{n=\sigma}^{\tau-1} p_n^{(1)}|\beta_n| \le \frac{(3+c)2\varepsilon}{1-2\theta} < \frac{\delta}{8}.$$

But if $\tau < \infty$ by Lemma A.18 (a), $p_\sigma^{(2)} \in [\delta/4, 1-\delta/4]$. This yields that $p_\tau^{(2)} \in [\delta/8, 1-\delta/8]$, in contradiction with the definition of τ. Hence $\tau = \infty$ and the conclusion is reached as in the corresponding part of Theorem 3.27. □

Remark A.20. *Repeating this argument for $p_* \in E_i$ with $i \in \{2,3\}$, by exploiting the symmetry of the model, defining σ and τ accordingly in terms of the corresponding coordinates and showing an analogous version of Theorem A.19 for $i \in \{2,3\}$ as well, yields convergence of any orbit approaching the boundary with $\ell = 0$.*

A.5 Convergence to the boundary with $\ell > 0$

The main goal of this section is to show that if $\{p_n\}$ approaches the boundary and the limit of the potential $\ell := \lim_{n\longrightarrow\infty} V(p_n, q_n) > 0$, the dynamical system converges. With the same introductory remarks as those in Section 3.5, we show how the arguments of this section generalise.

A.5.1 Convergence of boundary orbits

We will assume $q_0 \in \Sigma_0$ since by Remarks A.1 and A.2 it is already known that if $q_0 \in V$ and $p_0 \in \partial\Sigma \setminus V$, $\{p_n\}$ eventually becomes a constant point of the edge to which p_0 belongs, and q_n is 2-periodic.

Lemma A.21. *Let $p_0 \in E_i$ for some $i \in \{1,2,3\}$. Then $p_n \longrightarrow p_* \in E_i$ as $n \longrightarrow \infty$.*

Proof. By symmetry, without loss of generality, assume $i = 1$, that is $p_0 \in E_1$, or equivalently $p_0^{(1)} = 0$ and $0 < p_0^{(2)} < 1$. It follows that

$$q_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1-p_0^{(2)} & 0 & 0 \\ p_0^{(2)} & 0 & 0 \end{pmatrix} q_0 = \begin{pmatrix} q_0^{(2)} + q_0^{(3)} \\ (1-p_0^{(2)})q_0^{(1)} \\ p_0^{(2)}q_0^{(1)} \end{pmatrix},$$

$$p_1 = (1-\rho)\begin{pmatrix} 0 \\ p_0^{(2)} \\ 1-p_0^{(2)} \end{pmatrix} + \rho\left(\mathbf{1} - q_0 - \begin{pmatrix} q_0^{(2)} + q_0^{(3)} \\ (1-p_0^{(2)})q_0^{(1)} \\ p_0^{(2)}q_0^{(1)} \end{pmatrix}\right).$$

We obtain immediately that $p_1^{(1)} = 0$ and therefore, by induction, that for all $n \in \mathbb{N}_0$, $p_n^{(1)} = 0$, that is we are in presence of a boundary orbit. Also we can see that

$$p_1^{(2)} = (1-\rho)p_0^{(2)} + \rho(1 - q_0^{(2)} - (1-p_0^{(2)})q_0^{(1)}) > 0,$$

since $0 < p_0^{(2)} < 1$. Hence by induction $0 < p_n^{(2)} < 1$ for all $n \in \mathbb{N}_0$. In conclusion for all $n \in \mathbb{N}_0$,

$$M_{p_n} = \begin{pmatrix} 0 & 1 & 1 \\ 1 - p_n^{(2)} & 0 & 0 \\ p_n^{(2)} & 0 & 0 \end{pmatrix}.$$

The convex combination will make the proof of convergence easier than with $\rho = 1$. As in Lemma 3.29, we proceed by estimating $p_{n+1}^{(2)} - p_n^{(2)}$. Since we have just shown that for all n, M_{p_n} takes a form such that $q_{n+1}^{(2)} = (1 - p_n^{(2)})q_n^{(1)}$, we have that

$$\begin{aligned} p_{n+1}^{(2)} - p_n^{(2)} &= (1-\rho)p_n^{(2)} + \rho[1 - (1 - p_n^{(2)})q_n^{(1)} - q_n^{(2)}] - p_n^{(2)} \\ &= \rho[1 - (1 - p_n^{(2)})q_n^{(1)} - q_n^{(2)} - p_n^{(2)}] = \rho\left[(1 - p_n^{(2)})(1 - q_n^{(1)}) - q_n^{(2)}\right]. \end{aligned}$$

Similarly, due to the form of M_{p_n} for all $n \in \mathbb{N}_0$,

$$q_n^{(1)} = q_{n-1}^{(2)} + q_{n-1}^{(3)} = 1 - q_{n-1}^{(1)}$$

and

$$q_n^{(2)} = (1 - p_{n-1}^{(2)})q_{n-1}^{(1)},$$

which yields

$$q_n^{(2)} = (1 - p_{n-1}^{(2)})(1 - q_n^{(1)}).$$

Thus

$$p_{n+1}^{(2)} - p_n^{(2)} = \rho\left[(1 - p_n^{(2)})(1 - q_n^{(1)}) - (1 - p_{n-1}^{(2)})(1 - q_n^{(1)})\right] = -\rho(1 - q_n^{(1)})(p_n^{(2)} - p_{n-1}^{(2)}).$$

Due to the factor ρ, this already yields convergence, since $|p_{n+1}^{(2)} - p_n^{(2)}| \leq \rho|p_n^{(2)} - p_{n-1}^{(2)}|$ already produces a geometric series bound on the series of increments of the $p^{(2)}$-component, while $p_n^{(1)}$ is identically zero. We can use the monotonicity and the alternating sign to conclude, like in Lemma 3.29, that convergence is not at the vertices. □

Corollary A.22. *If $p_0 \in E_i$, then for all $q_0 \in \Sigma_0$ there exists $p_* \in E_i$ and some $\beta \geq 0$ (dependent on p_0 and q_0) such that the set of accumulation points of the boundary orbit is $\{(p_*, q_{p_*} \pm \beta e_{-1}(p_*))\}$. Moreover, if $\lim_{n\to\infty} V(p_n, q_n) = \ell$ then $\beta = \ell/2$.*

Proof. By symmetry, without loss of generality, set $i = 1$. Note that

$$\Phi_\rho(p, q) := \begin{pmatrix} (1-\rho)p + \rho(\mathbf{1} - q - M_p q) \\ M_p q \end{pmatrix}$$

is continuous on $E_1 \times \Sigma$. Rearranging (2.36) yields

$$(I + M_{p_n})q_n = \frac{1-\rho}{\rho}p_n - \frac{1}{\rho}p_n + \mathbf{1} = \mathbf{1} - p_n + \frac{1}{\rho}(p_n - p_{n+1}) = 2q_{p_n} + \frac{1}{\rho}(p_n - p_{n+1}).$$

By Lemma A.21 $p_n \longrightarrow p_*$, so $p_n - p_{n+1} \longrightarrow 0$ and $q_{p_n} \longrightarrow q_{p_*}$ as $n \longrightarrow \infty$, yielding

$$\frac{I + M_{p_n}}{2}q_n \longrightarrow q_{p_*}$$

as in Corollary 3.30. From this we obtain

$$\frac{I + M_{p_*}}{2}(q_n - q_{p_*}) \longrightarrow \mathbf{0}.$$

From here on, the conclusion is the same as in Corollary 3.30. □

Proposition A.23. *If $p_0 \in E_i$, then for all $q_0 \in \Sigma_0$ there exists $p_* \in E_i$ and $\beta \geq 0$ (dependent on initial conditions) such that the boundary orbit approaches the 2-cycle $\{(p_*, q_{p_*} \pm \beta e_{-1}(p_*))\}$, with $\beta = |\beta_0|$, where $q_0 = q_{p_0} + \alpha_0 e_0(p_0) + \beta_0 e_{-1}(p_0)$.*

Proof. The proof is as in Proposition 3.31, with a few changes, which we now point out. Given the same set-up as in Proposition 3.31 with $i = 1$ and the same eigensystem, noting that for a boundary orbit as such, $e_0(p_n) \coloneqq e_0 = (0, 1, -1)$ and

$$e_{-1}(p_{n+1}) - e_{-1}(p_n) = (p_{n+1}^{(2)} - p_n^{(2)})e_0,$$

it is straightforward to rewrite (A.2) by exploiting the eigencoordinates:

$$\alpha_{n+1}e_0 + \beta_{n+1}e_{-1}(p_{n+1}) = -\frac{\rho}{2}\alpha_n e_0 - \beta_n e_{-1}(p_n). \tag{A.16}$$

yielding the system

$$\begin{aligned} -\beta_{n+1} &= \beta_n \\ \alpha_{n+1} + (1 - p_{n+1}^{(2)})\beta_{n+1} &= -\frac{\rho}{2}\alpha_n - (1 - p_n^{(2)})\beta_n \\ -\alpha_{n+1} + p_{n+1}^{(2)}\beta_{n+1} &= \frac{\rho}{2}\alpha_n - p_n^{(2)}\beta_n. \end{aligned}$$

Since the first equation plugged into the second makes the latter a scalar multiple of the third equation, the system is consistent and overdetermined, so we solve it by keeping the first and third equation only, and finally use the first equation to simplify the third, obtaining

$$\alpha_{n+1} = -\frac{\rho}{2}\alpha_n + \beta_n(p_n^{(2)} - p_{n+1}^{(2)}) \tag{A.17}$$

$$\beta_{n+1} = -\beta_n. \tag{A.18}$$

By rearranging (2.36), we can also express $p_{n+1} - p_n$ in eigencoordinates. We have that

$$p_{n+1} - p_n = \rho(2q_{p_n} - q_{n+1} - q_n) = -\rho(M_{p_n} + I)(q_n - q_{p_n}) = -\rho\alpha_n e_0.$$

Thus $p_{n+1}^{(2)} - p_n^{(2)} = -\rho\alpha_n$, which allows us to rewrite (A.17) and (A.18) as

$$\alpha_{n+1} = \rho\left(\beta_n - \frac{1}{2}\right)\alpha_n \tag{A.19}$$

$$\beta_{n+1} = -\beta_n. \tag{A.20}$$

Thus we argue exactly as in Corollary 3.30 to prove the claim, thanks to the geometric decay of α_n. The only difference in the bounds is that they will have a ρ^2 factor in the two-steps iterates, that is

$$\alpha_{n+1} = -\rho^2\left(\beta_{n-1}^2 - \frac{1}{4}\right)\alpha_{n-1},$$

yielding

$$|\alpha_{n+1}| < \frac{\rho^2}{12}|\alpha_{n-1}|$$

and in conclusion resulting into

$$|\alpha_n| < M\left(\frac{\rho}{2\sqrt{3}}\right)^n,$$

having defined $M \coloneqq \max\{|\alpha_0|, |\alpha_1|\}$. □

Remark A.24. *The geometric upper bound on the decay of the* $|\alpha_n|$ *for a boundary orbit is uniform, since* $M := \max\{|\alpha_0|, |\alpha_1|\}$ *holds uniformly, by the boundedness of the simplex. Hence there exists a uniform constant* $\tilde{M}$ *such that for any boundary orbit*

$$|\alpha_n| < \tilde{M}\left(\frac{\rho}{2\sqrt{3}}\right)^n. \tag{A.21}$$

The following is immediate from Proposition A.23. Recall that Σ^* denotes the medial triangle (boundary excluded).

Corollary A.25. *If* $p_0 \in E_i$ *and* $q_0 = q_{p_0} + \alpha_0 e_0(p_0) + \beta_0 e_{-1}(p_0) \in \Sigma_0$*,* $p_n \longrightarrow p_* \in E_i$*, while* $\{q_n\}$ *either converges to* $q_{p_*} \in \partial\Sigma^*$ *if* $\beta_0 = 0$ *or approaches the* 2*-cycle* $\{(p_*, q_{p_*} \pm \beta e_{-1}(p_*))\}$*, with* $\beta = |\beta_0|$*.*

A.5.2 Structure of the set of accumulation points

In this section we study the set of accumulation points of regular orbits approaching the boundary, and therefore assume $p_0 \notin \partial\Sigma$. Since we will approach the problem from the point of view of boundary orbits, it is useful to describe the dynamical system in terms of its iteration map $\Phi(p,q) := \Phi_\rho(p,q)$ (the ρ will always be omitted in this section, as there is no confusion with Section 3.5), which is continuous on $\Sigma_0 \times \Sigma$, defined as

$$\begin{pmatrix} p_{n+1} \\ q_{n+1} \end{pmatrix} = \Phi(p_n, q_n) = \begin{pmatrix} \Phi_p(p_n, q_n) \\ \Phi_q(p_n, q_n) \end{pmatrix} := \begin{pmatrix} (1-\rho)p_n + \rho(\mathbf{1} - q_n - M_{p_n} q_n) \\ M_{p_n} q_n \end{pmatrix}.$$

Lemma A.26. *Let* $\{(p_n, q_n)\}$ *be an orbit such that* $\{p_n\}$ *is not bounded away from the vertices, that is such that there is* $\{n_k\}$*, with* $p_{n_k} \longrightarrow v_i$ *for some* $i \in \{1,2,3\}$*. Then the set of accumulation points of* $\{(p_{n_k-1}, q_{n_k-1})\}$*,* $\{(p_{n_k}, q_{n_k})\}$ *and* $\{(p_{n_k+1}, q_{n_k+1})\}$ *is a subset of* $\left\{(v_i, q_{v_i} \pm \frac{\ell}{2} e_{-1}(v_i))\right\}$*. Moreover if* $\{n_k\}$ *is such that* $\{q_{n_k}\}$ *converges,* $\{q_{n_k-1}\}$*,* $\{q_{n_k}\}$*, and* $\{q_{n_k+1}\}$ *asymptotically oscillate about* $q_* = q_{v_i} \pm \frac{\ell}{2} e_{-1}(v_i)$ *and* $\hat{q}_* = q_{v_i} \mp \frac{\ell}{2} e_{-1}(v_i)$*, while* $\{p_{n_k-1}\}$*,* $\{p_{n_k}\}$*, and* $\{p_{n_k+1}\}$ *all tend to* v_i*, that is* $(p_{n_k-1}, q_{n_k-1}) \longrightarrow (v_i, \hat{q}_*)$*,* $(p_{n_k}, q_{n_k}) \longrightarrow (v_i, q_*)$ *and* $(p_{n_k+1}, q_{n_k+1}) \longrightarrow (v_i, \hat{q}_*)$*.*

Proof. By symmetry, without loss of generality, assume $i = 2$ and let $(p_{n_{k_r}}, q_{n_{k_r}}) \longrightarrow (v_2, q_*)$ as $r \longrightarrow \infty$ as in Lemma 3.34. Note that as $r \longrightarrow \infty$

$$(1-\rho)p^{(1)}_{n_r-1} + \rho(1 - q^{(1)}_{n_r-1} - q^{(1)}_{n_r}) \longrightarrow 0 \tag{A.22}$$
$$(1-\rho)p^{(2)}_{n_r-1} + \rho(1 - q^{(2)}_{n_r-1} - q^{(2)}_{n_r}) \longrightarrow 1 \tag{A.23}$$
$$(1-\rho)p^{(3)}_{n_r-1} + \rho(1 - q^{(3)}_{n_r-1} - q^{(3)}_{n_r}) \longrightarrow 0, \tag{A.24}$$

and therefore (A.23) implies that $q^{(2)}_{n_r-1} + q^{(2)}_{n_r} \longrightarrow 0$, since $0 \le p^{(2)}_{n_r-1} \le 1$, and therefore if $1 - q^{(2)}_{n_r-1} - q^{(2)}_{n_r}$ did not tend to 1, the convex combination of the two would not either. Hence $q^{(2)}_{n_r} \longrightarrow 0 = q^{(2)}_*$ and $p^{(2)}_{n_r-1} \longrightarrow 1$, which also implies that for $i \in \{1,3\}$, $p^{(i)}_{n_r-1} \longrightarrow 0$, that is $p_{n_r-1} \longrightarrow v_2$ too. (A.22) and (A.24) directly imply also that for $i \in \{1,3\}$, $q^{(i)}_{n_r} + q^{(i)}_{n_r-1} \longrightarrow 1$ as $r \longrightarrow \infty$, because if the convex combinations tend to zero, then for all $i \in \{1,3\}$, $1 - q^{(i)}_{n_r} - q^{(i)}_{n_r-1} \longrightarrow 0$ as $r \longrightarrow \infty$. From here on, one can repeat the same exact argument in the corresponding part of Lemma 3.34 as the same properties have been shown, with, in addition, the fact that $p_{n_r-1} \longrightarrow v_2$ as $r \longrightarrow \infty$ (this was not possible to be shown in Lemma 3.34). The part of the argument that finds the two possible forms of the limit of $\{q_{n_r}\}$, q_* and $\hat{q}_*$, given the potential limit ℓ, stays the same since the potential is the same. Also the part in which, without loss of generality, assuming $q_* = ((1-\ell)/2, 0, (1+\ell)/2)$, we showed that

$q_{n_r-1} \longrightarrow \hat{q}_* := ((1+\ell)/2, 0, (1-\ell)/2)$ first and $q_{n_r+1} \longrightarrow \hat{q}_*$ secondly, stays the same, as it relies only on (2.35), which is the same as (2.37). The only difference is in how we show the conclusion for p_{n_r+1}. As $r \longrightarrow \infty$, $q_{n_r+1} \longrightarrow \hat{q}_*$ and $q_{n_r} + q_{n_r+1} \longrightarrow (1,0,1)$ so

$$p_{n_r+1} = (1-\rho)p_{n_r} + \rho[\mathbf{1} - (q_{n_r} + q_{n_r+1})] \longrightarrow (1-\rho)v_2 + \rho(\mathbf{1} - (1,0,1)) = v_2,$$

and $q_{p_{n_r+1}} \longrightarrow q_{v_2} = (1/2, 0, 1/2)$. Like in Lemma 3.34, we now need to start from an arbitrary convergent subsubsequence of $\{(p_{n_k+1}, q_{n_k+1})\}$ and in addition, also from one of $\{p_{n_k-1}, q_{n_k-1}\}$. We will denote them $\{(p_{n_r+1}, q_{n_r+1})\}$ and $\{(p_{n_r-1}, q_{n_r-1})\}$, and their limit will be, in each case separately, denoted as $(p_*, \hat{q}_*)$, to be determined. The underlying hypothesis is that $p_{n_r} \longrightarrow v_2$. In the case of the convergent forward shift subsubsequence, by (2.36) we can see that

$$q_{n_r} = \frac{(1-\rho)p_{n_r} - p_{n_r+1}}{\rho} + \mathbf{1} - q_{n_r+1} \longrightarrow \frac{(1-\rho)v_2 - p_*}{\rho} + \mathbf{1} - \hat{q}_* =: q_*,$$

so we have again $(p_{n_r}, q_{n_r}) \longrightarrow (v_2, q_*)$ and we can proceed, through the same argument shown above, with showing that the forms of q_* and $\hat{q}_*$ are the claimed ones, and that $p_* = v_2$. Similarly, in the case of the convergent backward shift subsubsequence, by (2.36) we can see that

$$q_{n_r} = \frac{(1-\rho)p_{n_r-1} - p_{n_r}}{\rho} + \mathbf{1} - q_{n_r} \longrightarrow \frac{(1-\rho)p_* - v_2}{\rho} + \mathbf{1} - \hat{q}_* =: q_*,$$

thus $(p_{n_r}, q_{n_r}) \longrightarrow (v_2, q_*)$ and we can repeat the same strategy just discussed in the previous case. The second part of the claim trivially follows by taking $n_r = n_k$ in the argument above. □

The next proposition is proved the same as Proposition 3.35, by swapping for the iteration map $\Phi(p,q) = \Phi_\rho(p,q)$ in the argument, and all the references to Lemma 3.34 and Remark 3.32 being replaced by Lemma A.26 and Remark A.24.

Proposition A.27. *Let $\{(p_n, q_n)\}$ be an orbit such that $\ell > 0$. The set of its accumulation points is a subset of*

$$\{(p, q_p \pm \beta e_{-1}(p)) : \ p \in \partial\Sigma, \ \beta > 0, \ V(p, q_p \pm \beta e_{-1}(p)) = \ell\}.$$

Remark A.28. *Consider an orbit such that $\ell > 0$ and, for some $\{n_k\}_{k\in\mathbb{N}}$, $p_{n_k} \longrightarrow p_* \in E_i$ for some $i \in \{1,2,3\}$. The set of accumulation points of $\{(p_{n_k}, q_{n_k})\}$ and $\{(p_{n_k+1}, q_{n_k+1})\}$ is a subset of $\{(p_*, q_{p_*} \pm \beta e_{-1}(p_*)) : \ \beta > 0, \ V(p_*, q_{p_*} \pm \beta e_{-1}(p_*)) = \ell\}$. Moreover, if $\{n_k\}$ is such that also $\{q_{n_k}\}$ converges, $\{q_{n_k}\}$ and its shift $\{q_{n_k+1}\}$ asymptotically oscillate between $q_* = q_{p_*} \pm \beta e_{-1}(p_*)$ and $\hat{q}_* = q_{p_*} \mp \beta e_{-1}(p_*)$, that is if $(p_{n_k}, q_{n_k}) \longrightarrow (p_*, q_*)$, then $(p_{n_k+1}, q_{n_k+1}) \longrightarrow (p_*, \hat{q}_*)$ as $k \longrightarrow \infty$.*

Proof. The claim for the q-component's shift follows as in Remark 3.36 by using Proposition A.27 instead of Proposition 3.35, since (2.35) is the same as (2.33). The only explicit difference is in (2.36), which now yields - when starting with a subsubsequence $\{(p_{n_r+1}, q_{n_r+1})\}$ convergent to some $(p, \hat{q}_*)$ to be determined - that

$$q_{n_r} = \frac{(1-\rho)p_{n_r} - p_{n_r+1}}{\rho} + \mathbf{1} - q_{n_r+1} \longrightarrow \frac{(1-\rho)p_* - p}{\rho} + \mathbf{1} - \hat{q}_* = q_*$$

by Proposition A.27. The rest of the argument is exactly the same. The claim about the p-component's shift follows also as in Remark 3.36, since using again $\hat{q}_*+q_* = 2q_{p_*}$, if $p_{n_r} \longrightarrow p_*$ as $r \longrightarrow \infty$,

$$\begin{aligned} p_{n_r+1} &= (1-\rho)p_{n_r} + \rho(\mathbf{1} - q_{n_r+1} - q_{n_r}) \longrightarrow (1-\rho)p_* + \rho[\mathbf{1} - (\hat{q}_* + q_*)] \\ &= (1-\rho)p_* + \rho(\mathbf{1} - 2q_{p_*}) = (1-\rho)p_* + \rho p_* = p_*. \end{aligned}$$

The second part of the claim is shown by taking $n_r = n_k$. □

The next corollary is proved as its counterpart Corollary 3.37, as the argument does not depend on ρ. It is enough to replace the references to (3.4), Remark 3.36, and Lemma 3.34 with (A.4), Remark A.28, and Lemma A.26 .

Corollary A.29. *Let $\{(p_n, q_n)\}$ be an orbit. Then $p_{n+1} - p_n \longrightarrow \mathbf{0}$ as $n \longrightarrow \infty$.*

Remark A.30. *Let $\{(p_n, q_n)\}$ be an orbit with $\{p_n\}$ not convergent to any of the vertices. By Corollary A.29 and Remark A.5, there is a subsequence $\{p_{n_j}\}$ bounded away from V.*

The following claim is trivially true if $\ell = 0$.

Corollary A.31. *Let $\{(p_n, q_n)\}$ be an orbit not convergent to the vertices and such that $\ell > 0$, that is $V(p_n, q_n) = \|\alpha_n e_0(p_n) + \beta_n e_{-1}(p_n)\|_1 \longrightarrow \ell > 0$. Then $\alpha_n \longrightarrow 0$ and $|\beta_n| \longrightarrow \ell/2$ as $n \longrightarrow \infty$.*

Proof. For the first part of the statement, consider that, following the notation of Lemma 3.19, in eigencoordinates (A.9) holds, which we recall below,

$$p_{n+1} - p_n = -\rho[\alpha_n(1 + \lambda_0(p_n))e_0(p_n) + \beta_n(1 + \lambda_{-1}(p_n))e_{-1}(p_n)],$$

and as $p_n \longrightarrow \partial\Sigma$, $1+\lambda_{-1}(p_n) \longrightarrow 0$ and $1+\lambda_0(p_n) \longrightarrow 1$; as a result of Corollary A.29 and Lemma 3.19 (h), we obtain that $\alpha_n \longrightarrow 0$ as $n \longrightarrow \infty$ exactly as in Corollary 3.39, since the factor ρ in (A.9) does not change substantially the argument. Therefore the second part of the proof regarding $\{|\beta_n|\}$ remains unaltered. □

We conclude this section with the remark corresponding to Remark 3.40. Showing that the potential is eventually subunitary requires a different argument in this case, due to (A.1) taking the place of (3.1). In particular, it will not be possible to show that the potential is already subunitary at the second iterate. We will have to show that eventually, at a time large enough, it will become subunitary, by exploiting the accumulation points of the orbits.

Remark A.32. *For any orbit with $\ell > 0$, which does not converge to a vertex, $\ell < 1$.*

Proof. By Remark A.5 and Corollary 3.37, exploiting the boundedness of $\{p_n\}$, we can assume the existence of a subsequence $\{p_{n_j}\}$ converging to a point $p_* \in E_i$ for some $i \in \{1, 2, 3\}$. Without loss of generality, by symmetry, assume $i = 1$. By Remark A.28 assume that the limit points for $\{q_{n_j}\}$ and $\{q_{n_j+1}\}$ are, without loss of generality, $q_* = q_{p_*} - \beta e_{-1}(p_*)$ and $\hat{q}_* = q_{p_*} + \beta e_{-1}(p_*)$, for some $0 < |\beta| = \ell/2$, by Corollary A.31. Note that $q^{(1)}_{p_*} = 1/2$ and that $q^{(1)}_{p_{n_j+1}} = (1 - p^{(1)}_{n_j+1})/2 < 1/2$ for all $j \in \mathbb{N}$ by Remark A.4. Note also that for all j large enough, due to the nonzero angle that the eigendirection $e_{-1}(p_*)$ forms with E_2 and E_3, due to $\beta > 0$ and due to the elementary geometry of the simplex, we have that $q^{(1)}_{n_j} - q^{(1)}_{p_{n_j}} > 0$, $q^{(2)}_{n_j} - q^{(2)}_{p_{n_j}} < 0$,

$-(q^{(1)}_{n_j} - q^{(1)}_{p_{n_j}} + q^{(2)}_{n_j} - q^{(2)}_{p_{n_j}}) = q^{(3)}_{n_j} - q^{(3)}_{p_{n_j}} < 0$ and $q^{(1)}_{n_j+1} - q^{(1)}_{p_{n_j+1}} < 0$, $q^{(2)}_{n_j+1} - q^{(2)}_{p_{n_j+1}} > 0$, $-(q^{(1)}_{n_j+1} - q^{(1)}_{p_{n_j+1}} + q^{(2)}_{n_j+1} - q^{(2)}_{p_{n_j+1}}) = q^{(3)}_{n_j+1} - q^{(3)}_{p_{n_j+1}} > 0$. Therefore

$$\begin{aligned}\|q_{n_j+1} - q_{p_{n_j+1}}\|_1 = \sum_{i\in\{1,2,3\}} |q^{(i)}_{n_j+1} - q^{(i)}_{p_{n_j+1}}| &= -(q^{(1)}_{n_j+1} - q^{(1)}_{p_{n_j+1}}) + q^{(2)}_{n_j+1} - q^{(2)}_{p_{n_j+1}} \\ - (q^{(1)}_{n_j+1} - q^{(1)}_{p_{n_j+1}} + q^{(2)}_{n_j+1} - q^{(2)}_{p_{n_j+1}}) &= 2(q^{(1)}_{p_{n_j+1}} - q^{(1)}_{n_j+1}) < 2\cdot\frac{1}{2} = 1,\end{aligned}$$

and we can conclude by Remark A.6, that eventually the potential is subunitary and $\ell < 1$. ☐

A.5.3 Convergence of regular orbits

We will adopt the same exact set-up as in Section 3.5.3, since all the corresponding theorems involved in the set-up have been shown to apply: Lemma A.26, Remarks A.28 and A.30, Proposition A.27, Corollary A.31, and (A.9). Even the estimate $\|\hat{p} - p\|_1 \leq 3B(|\alpha| + p_1)$ for all $p_1 < \varepsilon$ does not change, due to the factor ρ in (A.9) being subunitary. We now list the restrictions that change. Define

$$R \coloneqq 1 + \rho\left(\frac{1}{\delta} - 1\right),$$

then we require

$$\varepsilon' < \min\left\{\frac{\varepsilon}{R}, \frac{\varepsilon}{12B}\right\},$$

(ε' is to be further restricted too). Having defined

$$\mathcal{K}^{\ell}_{\varepsilon',\delta} \coloneqq \left\{(p,q) \in \Sigma^2 :\ 0 < p^{(1)},\ |\alpha|,\ \left||\beta| - \frac{\ell}{2}\right| \leq \varepsilon',\ \delta \leq p^{(2)} \leq 1-\delta\right\},$$

and similarly $\mathcal{K}^{\ell}_{\varepsilon,\delta'}$, by construction of m, $(p_m, q_m) \in \mathcal{K}^{\ell}_{\varepsilon',\delta}$, and

$$\|p_{m+1} - p_m\|_1 < 6B\varepsilon' < \frac{\varepsilon}{2}, \tag{A.25}$$

thus ensuring, exactly as in Section 3.5.3, that

$$p^{(1)}_{m+1} < \varepsilon. \tag{A.26}$$
$$p^{(2)}_{m+1} < 1 - \delta', \tag{A.27}$$
$$p^{(2)}_{m+1} > \delta'. \tag{A.28}$$

Define a hitting time

$$\sigma \coloneqq \inf\left\{n \geq m :\ p^{(2)}_n \notin \left[\frac{\delta}{2}, 1 - \frac{\delta}{2}\right]\right\} \in \mathbb{N} \cup \infty.$$

Note that $\sigma > m + 1$ by construction of m. We now derive some iterative formulas and bounds.

Remark A.33. *For all $m \leq n < \sigma$, $p^{(1)}_{n+1} < Rp^{(1)}_n$.*

Proof. Since for all $m \leq n < \sigma$,

$$1 - q^{(1)}_{n+1} - q^{(1)}_n = p^{(1)}_n\left(\frac{q^{(2)}_n}{1-p^{(2)}_n} + \frac{q^{(3)}_n}{1-p^{(3)}_n}\right) < \frac{2}{\delta}(1 - q^{(1)}_m)p^{(1)}_n \leq \frac{2}{\delta}p^{(1)}_n,$$

$$p^{(1)}_{n+1} < (1-\rho)p^{(1)}_n + \frac{2\rho}{\delta}p^{(1)}_n,$$

and the claim follows by the definition of R. ☐

Remark A.34. *For any $n \geq 0$,*

$$p^{(1)}_{n+2} = p^{(1)}_{n+1}\left\{(1-\rho)+\rho\left[(1-q^{(1)}_{n+1})\vartheta_{n+1} - p^{(1)}_{n+1}\frac{\vartheta_{n+1}}{\rho} + p^{(1)}_n\vartheta''_{n+1}\right]\right\},$$

where ϑ_{n+1} and ϑ'_{n+1} are defined as in (3.33) and (3.34) in Remark 3.42, and

$$\vartheta''_{n+1} := \frac{1-\rho}{\rho}\vartheta_{n+1} + \vartheta'_{n+1}. \tag{A.29}$$

Proof. For any $n \geq 0$,

$$\begin{aligned} p^{(1)}_{n+2} &= (1-\rho)p^{(1)}_{n+1} + \rho p^{(1)}_{n+1}\left(\frac{q^{(2)}_{n+1}}{1-p^{(2)}_{n+1}} + \frac{q^{(3)}_{n+1}}{p^{(1)}_{n+1}+p^{(2)}_{n+1}}\right) \\ &= p^{(1)}_{n+1}\{(1-\rho)+\rho(q^{(1)}_n\vartheta_{n+1} + p^{(1)}_n\vartheta'_{n+1})\} \end{aligned}$$

by the second step (in which we rearranged the factor in the brackets) in the proof of Remark 3.42, and since by (2.36)

$$q^{(1)}_n = (1-q^{(1)}_{n+1}) + \frac{1-\rho}{\rho}p^{(1)}_n - \frac{1}{\rho}p^{(1)}_{n+1},$$

the claim follows. □

Remark A.35. *For any $n \geq 0$,*

$$p^{(2)}_{n+2} - p^{(2)}_{n+1} = -\rho q^{(1)}_n\left(p^{(2)}_{n+1} - p^{(2)}_n + \xi_{n+1} - \xi'_{n+1} - \eta'_{n+1} - \eta_n + \eta''_{n+1} + \eta'''_{n+1}\right),$$

where η'_{n+1}, η_n, η''_{n+1} and η'''_{n+1} are defined as in (3.35) to (3.38) in Remark 3.43, and

$$\xi_{n+1} := p^{(1)}_{n+1}\frac{p^{(2)}_{n+1}}{1-p^{(1)}_{n+1}}\left(1+\frac{\frac{1}{\rho}-1}{q^{(1)}_n}\right) \tag{A.30}$$

$$\xi'_{n+1} := p^{(1)}_n p^{(2)}_{n+1}\frac{\frac{1}{\rho}-1}{q^{(1)}_n(1-p^{(1)}_{n+1})}. \tag{A.31}$$

Proof. For any $n \geq 0$,

$$\begin{aligned} p^{(2)}_{n+2} - p^{(2)}_{n+1} &= \rho\left(\frac{p^{(2)}_{n+1}}{1-p^{(1)}_{n+1}}q^{(1)}_{n+1} + \frac{p^{(2)}_{n+1}}{p^{(1)}_{n+1}+p^{(2)}_{n+1}}q^{(3)}_{n+1} - p^{(2)}_{n+1}\right) \\ &= \rho\left[p^{(2)}_{n+1}\left(\frac{q^{(1)}_{n+1}}{1-p^{(1)}_{n+1}} - 1\right) + \frac{p^{(2)}_{n+1}}{p^{(1)}_{n+1}+p^{(2)}_{n+1}}\left(q^{(1)}_n\frac{p^{(2)}_n}{1-p^{(1)}_n} + q^{(2)}_n\frac{p^{(1)}_n}{1-p^{(2)}_n}\right)\right] \\ &= \rho\left\{p^{(2)}_{n+1}\left[\frac{q^{(1)}_{n+1}}{1-p^{(1)}_{n+1}} - 1 + \frac{p^{(1)}_n}{(p^{(1)}_{n+1}+p^{(2)}_{n+1})(1-p^{(2)}_n)}q^{(2)}_n\right]\right. \\ &\quad \left. + p^{(2)}_n\frac{q^{(1)}_n}{1-p^{(1)}_n}\frac{p^{(2)}_{n+1}}{p^{(1)}_{n+1}+p^{(2)}_{n+1}}\right\} = \rho\left\{p^{(2)}_{n+1}\left[-q^{(1)}_n - p^{(1)}_{n+1}\frac{q^{(1)}_n+\frac{1}{\rho}-1}{1-p^{(1)}_{n+1}}\right.\right. \\ &\quad \left. - p^{(1)}_n\frac{1-\frac{1}{\rho}}{1-p^{(1)}_{n+1}} + \frac{p^{(1)}_n}{(p^{(1)}_{n+1}+p^{(2)}_{n+1})(1-p^{(2)}_n)}q^{(2)}_n\right] + p^{(2)}_n q^{(1)}_n\left(1+\frac{p^{(1)}_n}{1-p^{(1)}_n}\right) \\ &\quad \left.\left(1-\frac{p^{(1)}_{n+1}}{p^{(1)}_{n+1}+p^{(2)}_{n+1}}\right)\right\} = -\rho q^{(1)}_n\left[p^{(2)}_{n+1} - p^{(2)}_n + p^{(1)}_{n+1}\frac{p^{(2)}_{n+1}}{1-p^{(1)}_{n+1}}\left(1+\frac{\frac{1}{\rho}-1}{q^{(1)}_n}\right)\right. \\ &\quad + p^{(1)}_n p^{(2)}_{n+1}\frac{1-\frac{1}{\rho}}{q^{(1)}_n(1-p^{(1)}_{n+1})} - p^{(1)}_n\frac{p^{(2)}_{n+1}}{(p^{(1)}_{n+1}+p^{(2)}_{n+1})(1-p^{(2)}_n)}\frac{q^{(2)}_n}{q^{(1)}_n} \\ &\quad \left. - p^{(2)}_n\frac{p^{(1)}_n}{1-p^{(1)}_{n+1}} + p^{(1)}_{n+1}\frac{p^{(2)}_n}{p^{(1)}_{n+1}+p^{(2)}_{n+1}} + p^{(2)}_n\frac{p^{(1)}_n}{1-p^{(1)}_n}\frac{p^{(1)}_{n+1}}{p^{(1)}_{n+1}+p^{(2)}_{n+1}}\right], \end{aligned}$$

since by (2.36),

$$1-q_{n+1}^{(1)}=q_n^{(1)}+\left(1-\frac{1}{\rho}\right)p_n^{(1)}+\frac{p_{n+1}^{(1)}}{\rho},$$

and therefore

$$\frac{q_{n+1}^{(1)}}{1-p_{n+1}^{(1)}}-1=-\frac{1-q_{n+1}^{(1)}}{1-p_{n+1}^{(1)}}+\frac{p_{n+1}^{(1)}}{1-p_{n+1^{(1)}}}=-\frac{q_n^{(1)}}{1-p_{n+1}^{(1)}}-\frac{1-\frac{1}{\rho}}{1-p_{n+1}^{(1)}}p_n^{(1)}-\frac{p_{n+1}^{(1)}}{\rho(1-p_{n+1}^{(1)})}$$
$$+\frac{p_{n+1}^{(1)}}{1-p_{n+1}^{(1)}}=-q_n^{(1)}-p_{n+1}^{(1)}\frac{q_n^{(1)}+\frac{1}{\rho}-1}{1-p_{n+1}^{(1)}}-p_n^{(1)}\frac{1-\frac{1}{\rho}}{1-p_{n+1}^{(1)}}.$$

Thus the claim follows. □

Finally we require that $\delta<\rho^4/45$, and define

$$\gamma=\gamma(\delta'):=\sqrt{:=1-4\frac{(\delta')^2}{\rho^2}+240\frac{(\delta')^3}{\rho^2}}.$$

The constant $\gamma':=\gamma^2=1-4\rho^2(\delta')^2+240(\delta')^3/\rho^2$ is positive subunitary, since $-4\rho^2x^2+240x^3/\rho^2$ is negative monotone decreasing on $(0,\rho^4/90)$, and it attains minimum at $\rho^4/90$ of value $-\rho^{10}/6075>-1/6075$. Recall that

$$R:=1+\rho\left(\frac{2}{\delta}-1\right).$$

Further require

$$\varepsilon<\min\left\{(\delta')^5,\frac{\rho\delta'}{2(2-\rho)R}(1-\gamma^2),\frac{\rho\delta'}{2-\rho}\frac{1}{1+\frac{R}{1-\gamma}}\right\}.$$

Define also

$$D:=2+\frac{1}{\delta'}\left(\frac{1}{\delta'}+\frac{1}{\rho}-1\right)+R\left(2+\frac{1}{\delta'\rho}\right)$$

and let Γ be a constant such that

$$0<\Gamma<\frac{\delta'}{D(1-\delta')}.$$

Finally let $\lambda:=\max\{\gamma,\rho\}$ and further restrict

$$\varepsilon'<\min\left\{\frac{\varepsilon}{12B},\frac{\varepsilon}{4RD}(1-\gamma^2),\frac{\delta\Gamma\lambda(1-\lambda)}{2(R+\Gamma)}\right\}.$$

Lemma A.36. *Let $\gamma':=\gamma^2$, assume that $p_{m+l}\le R(\gamma')^{\lfloor\frac{l}{2}\rfloor}p_m^{(1)}$ for all $0\le l\le 2k-b$, where $b\in\{0,1\}$. Then for all $b\le j\le 2k$,*

$$\delta'<q_{m+2k-j}^{(1)}<\frac{1}{2}-\delta'$$

if j is even and

$$\frac{1}{2}+\delta'<q_{m+2k-j}^{(1)}<1-\delta'$$

if j is odd.

Proof. We iterate the first component of (2.36) after rearranging it as

$$q^{(1)}_{m+l} = 1 - q^{(1)}_{m+l-1} + \frac{1-\rho}{\rho} p^{(1)}_{m+l-1} - \frac{1}{\rho} p^{(1)}_{m+l},$$

for every $l \geq 0$. It yields

$$q^{(1)}_{m+l} = \begin{cases} q^{(1)}_m - \frac{1-\rho}{\rho} p^{(1)}_m + \frac{2-\rho}{\rho} \sum_{j=1}^{l-1} (-1)^{j+1} p^{(1)}_{m+j} - \frac{1}{\rho} p^{(1)}_{m+l} & l \text{ even} \\ 1 - q^{(1)}_m + \frac{1-\rho}{\rho} p^{(1)}_m + \frac{2-\rho}{\rho} \sum_{j=1}^{l-1} (-1)^{j} p^{(1)}_{m+j} - \frac{1}{\rho} p^{(1)}_{m+l} & l \text{ odd.} \end{cases} \tag{A.32}$$

Recall that by construction

$$\varepsilon < \frac{\rho\delta'}{2(2-\rho)R}(1-\gamma').$$

Since $q^{(1)}_m < 1/2 - \delta$,

$$\begin{aligned} q^{(1)}_{m+2k-j} &= q^{(1)}_m - \frac{1-\rho}{\rho} p^{(1)}_m + \frac{2-\rho}{\rho} \sum_{l=1}^{2k-j-1} (-1)^{l+1} p^{(1)}_{m+l} - \frac{1}{\rho} p^{(1)}_{m+2k-j} \\ &< \frac{1}{2} - \delta + \frac{2-\rho}{\rho} \sum_{l=0}^{2k-j} p^{(1)}_{m+l} < \frac{1}{2} - \delta + 2\frac{(2-\rho)R}{\rho}\varepsilon \sum_{l=0}^{\lfloor \frac{2k-j}{2} \rfloor} (\gamma')^l \\ &< \frac{1}{2} - \delta + \frac{2(2-\rho)R\varepsilon}{\rho(1-\gamma')} < \frac{1}{2} - \delta' \end{aligned} \tag{A.33}$$

for all even $b \leq j \leq 2k$ (with the bound for $j = 2k$ holding also with δ, by adopting empty sum convention) and

$$\begin{aligned} q^{(1)}_{m+2k-j} &= 1 - q^{(1)}_m + \frac{1-\rho}{\rho} p^{(1)}_m + \frac{2-\rho}{\rho} \sum_{l=1}^{2k-j-1} (-1)^{l} p^{(1)}_{m+l} - \frac{1}{\rho} p^{(1)}_{m+2k-j} \\ &> \frac{1}{2} + \delta - \frac{2-\rho}{\rho} \sum_{l=0}^{2k-j} p^{(1)}_{m+l} > \frac{1}{2} + \delta - 2\frac{(2-\rho)R\varepsilon}{\rho} \sum_{l=0}^{\lfloor \frac{2k-j}{2} \rfloor} (\gamma')^l \\ &> \frac{1}{2} + \delta - 2\frac{(2-\rho)R\varepsilon}{\rho(1-\gamma')} > \frac{1}{2} + \delta' \end{aligned} \tag{A.34}$$

for all odd $b \leq j \leq 2k$. Similarly, since $\delta < q^{(1)}_m < 1 - \delta$,

$$\begin{aligned} q^{(1)}_{m+2k-j} &= q^{(1)}_m - \frac{1-\rho}{\rho} p^{(1)}_m + \frac{2-\rho}{\rho} \sum_{l=1}^{2k-j-1} (-1)^{l+1} p^{(1)}_{m+l} - \frac{1}{\rho} p^{(1)}_{m+2k-j} \\ &> \delta - \frac{2-\rho}{\rho} \sum_{l=0}^{2k-j} p^{(1)}_{m+l} > \delta - 2\frac{(2-\rho)R}{\rho}\varepsilon \sum_{l=0}^{\lfloor \frac{2k-j}{2} \rfloor} (\gamma')^l \\ &> \delta - \frac{2(2-\rho)R\varepsilon}{\rho(1-\gamma')} > \delta' \end{aligned} \tag{A.35}$$

for all even $b \leq j \leq 2k$ and

$$\begin{aligned} q^{(1)}_{m+2k-j} &= 1 - q^{(1)}_m + \frac{1-\rho}{\rho} p^{(1)}_m + \frac{2-\rho}{\rho} \sum_{l=1}^{2k-j-1} (-1)^{l} p^{(1)}_{m+l} - \frac{1}{\rho} p^{(1)}_{m+2k-j} \\ &< 1 - \delta + \frac{2-\rho}{\rho} \sum_{l=0}^{2k-j} p^{(1)}_{m+l} < 1 - \delta + 2\frac{(2-\rho)R\varepsilon}{\rho} \sum_{l=0}^{\lfloor \frac{2k-j}{2} \rfloor} (\gamma')^l \\ &< 1 - \delta + 2\frac{(2-\rho)R\varepsilon}{\rho(1-\gamma')} < 1 - \delta' \end{aligned} \tag{A.36}$$

for all odd $b \le j \le 2k$. □

Lemma A.37. *Let $\gamma' := \gamma^2$, assume that $p_{m+l} \le R(\gamma')^{\lfloor \frac{l}{2} \rfloor} p^{(1)}_m$ for all $0 \le l \le 2k-b$, where $b \in \{0,1\}$, and $\delta' < p^{(2)}_{m+l} < 1-\delta'$. Then for all $b \le j \le 2k-1$,*

$$|p^{(2)}_{m+2k-j} - p_{m+2k-j-1}| < \varepsilon.$$

Proof. Recall that

$$\varepsilon' < \frac{\varepsilon}{4RD}(1-\gamma').$$

Iterate Remark A.35 applied to $n = m + 2(k-1) - j$, down to time m. It yields

$$|p^{(2)}_{m+2k-j} - p^{(2)}_{m+2k-j-1}| \le |p^{(2)}_{m+1} - p^{(2)}_m| + \sum_{l=1}^{2k-j-1} E_{m+l} < \frac{\varepsilon}{2} + \sum_{l=1}^{2k-2} E_{m+l},$$

where $E_{m+l} := \xi_{m+l} + \xi'_{m+l} + \eta_{m+l} + \eta'_{m+l} + \eta_{m+l-1} + \eta''_{m+l} + \eta'''_{m+l}$. Note that the hypotheses allow to apply Lemma A.36, thus $\delta' < q^{(1)}_{m+l-1} < 1-\delta'$ for all $1 \le l \le 2k-j-1$. This implies that, by using Remark A.33 and the assumptions, for all $1 \le l \le 2k-j-1$, we have that

$$\eta_{m+l-1} < p^{(1)}_{m+l-1} < R(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m \tag{A.37}$$

$$\eta_{m+l} < R p^{(1)}_{m+l-1} < R^2(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m \tag{A.38}$$

$$\eta'_{m+l} < \frac{1}{(\delta')^2} p^{(1)}_{m+l-1} < \frac{R}{(\delta')^2}(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m \tag{A.39}$$

$$\eta''_{m+l} < \frac{R}{\delta'} p^{(1)}_{m+l-1} < \frac{R^2}{\delta'}(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m \tag{A.40}$$

$$\eta'''_{m+l} < p^{(1)}_{m+l-1} < R(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m, \tag{A.41}$$

and we can estimate similarly the two error terms

$$\xi_{m+l} < \left[1 + \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right)\right] R^2(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m \tag{A.42}$$

$$\xi'_{m+l} < \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right) R(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m. \tag{A.43}$$

Hence

$$E_{m+l} < RD(\gamma')^{\lfloor \frac{l-1}{2} \rfloor} p^{(1)}_m, \tag{A.44}$$

which yields a bound, uniform in k, on the increments of the $p^{(2)}$-component:

$$|p^{(2)}_{m+2k-1} - p^{(2)}_{m+2k-2}| < \frac{\varepsilon}{2} + RD\varepsilon' \sum_{l=1}^{\infty} (\gamma')^{\lfloor \frac{l-1}{2} \rfloor} = \frac{\varepsilon}{2} + 2RD\frac{\varepsilon'}{1-\gamma'} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon. \tag{A.45}$$

□

Lemma A.38. *For all $m \le n \le \sigma$, $p^{(1)}_n \le R\gamma^{n-m-1} p^{(1)}_m$.*

Proof. Let

$$\gamma' := \gamma^2 = 1 - 4\frac{(\delta')^2}{\rho^2} + 240\frac{(\delta')^3}{\rho^2},$$

we will first show that for every $k \ge 0$ such that $m + 2k \le \sigma$,

$$p^{(1)}_{m+2k} \le R(\gamma')^k p^{(1)}_m \tag{A.46}$$

$$p^{(1)}_{m+2k+1} \le R(\gamma')^k p^{(1)}_m \tag{A.47}$$

and if $m+2k=\sigma$, we only show the claim up to (A.46). Recall that by construction $\delta' < \rho^4/90$,

$$\varepsilon < \frac{\rho\delta'}{2(2-\rho)R}(1-\gamma'),$$

and

$$\varepsilon' < \frac{\varepsilon}{4RD}(1-\gamma^2).$$

Choosing $n=m$ in Remark A.33, yields the first odd case (the even case is trivial as $R>1$) for $k=0$:

$$p^{(1)}_{m+1} \leq R p^{(1)}_m.$$

Let $n=m$ in (3.33) and (3.34), and apply the hypothesis made in (A.26) to (A.28), and the fact that $|p^{(2)}_{m+1}-p^{(2)}_m| < \varepsilon/2$, which follows from (A.25). Then the following estimates follow, as in the corresponding step of Lemma 3.46:

$$|\vartheta_{m+1}| \leq \frac{2}{\delta'(1-\varepsilon)}, \tag{A.48}$$

$$|\vartheta_{m+1}| \leq 2 + \frac{\varepsilon}{1-\varepsilon}\left(2+\frac{3}{\delta'}\right), \tag{A.49}$$

$$|\vartheta'_{m+1}| \leq \frac{1}{\delta'}\left(\frac{1}{1-\varepsilon}+\frac{2}{\delta'+\varepsilon}\right), \tag{A.50}$$

yielding

$$|\vartheta''_{m+1}| \leq \left(\frac{1}{\rho}-1\right)|\vartheta_{m+1}| + |\vartheta'_{m+1}| \leq \left(\frac{1}{\rho}-1\right)\frac{2}{\delta'(1-\varepsilon)} + \frac{1}{\delta'}\left(\frac{1}{1-\varepsilon}+\frac{2}{\delta'+\varepsilon}\right),$$

and therefore

$$|\vartheta''_{m+1}| < \frac{1}{\delta'}\left[\left(\frac{2}{\rho}-1\right)\frac{1}{1-\varepsilon}+\frac{2}{\delta'}\right]. \tag{A.51}$$

Plug the estimates (A.48) to (A.51) into Remark A.34 applied to $n=m$, it yields

$$\begin{aligned}
p^{(1)}_{m+2} &\leq p^{(1)}_{m+1}\Bigg\{(1-\rho)+\rho\Bigg[2(1-q^{(1)}_{m+1})+\frac{\varepsilon}{1-\varepsilon}\left(2+\frac{3}{\delta'}\right)+\frac{p^{(1)}_{m+1}}{\rho}\frac{2}{\delta'(1-\varepsilon)} \\
&\quad + p^{(1)}_m\frac{1}{\delta'}\left[\left(\frac{2}{\rho}-1\right)\frac{1}{1-\varepsilon}+\frac{2}{\delta'}\right]\Bigg]\Bigg\} \leq p^{(1)}_{m+1}\Bigg\{(1-\rho)+\rho\Bigg[2(1-q^{(1)}_{m+1}) \\
&\quad + \varepsilon\Bigg\{\frac{1}{1-\varepsilon}\left(2+\frac{3}{\delta'}\right)+\frac{1}{\rho}\frac{2}{\delta'(1-\varepsilon)}+\frac{1}{\delta'}\left[\left(\frac{2}{\rho}-1\right)\frac{1}{1-\varepsilon}+\frac{2}{\delta'}\right]\Bigg\}\Bigg]\Bigg\} \\
&= p^{(1)}_{m+1}\Bigg\{(1-\rho)+\rho\left[2(1-q^{(1)}_{m+1})+2\varepsilon\left(\frac{1}{1-\varepsilon}+\left(1+\frac{2}{\rho}\right)\frac{1}{\delta'(1-\varepsilon)}+\frac{1}{(\delta')^2}\right)\right]\Bigg\}.
\end{aligned}$$

Recall that $\varepsilon < (\delta')^5 < 1/90^5$. Then $q^{(1)}_m < 1/2-\delta$ implies that

$$\begin{aligned}
q^{(1)}_{m+1} &= 1-q^{(1)}_m+\frac{1-\rho}{\rho}p^{(1)}_m-\frac{1}{\rho}p^{(1)}_{m+1} > \frac{1}{2}+\delta-\frac{1-\rho}{\rho}p^{(1)}_m-\frac{R}{\rho}p^{(1)}_m \\
&> \frac{1}{2}+\delta-2\frac{R}{\rho}\varepsilon > \frac{1}{2}+\delta'
\end{aligned}$$

by construction of ε, since $2-\rho>1$. From

$$\begin{aligned}
&1-\rho+2\rho(1-q^{(1)}_{m+1})+2\varepsilon\left(\frac{1}{\delta'}+\left(1+\frac{2}{\rho}\right)\frac{1}{(\delta')^2}+\frac{1}{(\delta')^2}\right) \leq 1-\rho+2\rho(1-q^{(1)}_{m+1}) \\
&+2\frac{\varepsilon}{(\delta')^2}\left(3+\frac{2}{\rho}\right) < 1-2\rho\delta'+\left(6+\frac{4}{\rho}\right)(\delta')^3 < 1-2\rho\delta'+\frac{10}{\rho}(\delta')^3,
\end{aligned}$$

it follows that

$$p^{(1)}_{m+2} < Rp^{(1)}_m \left(1 - 2\rho\delta' + \frac{10}{\rho}(\delta')^3\right) < R\gamma' p^{(1)}_m.$$

Note that for any $\delta' > 0$, $\gamma' = 1 - 4\rho^2(\delta')^2 + 240(\delta')^3/\rho^2 > 1 - 2\rho\delta' + 10(\delta')^3/\rho$ since it is equivalent to $\rho - 2\rho^2\delta' + (\delta')^2({}^{120}/_{\rho^2} - {}^{5}/_{\rho}) > 0$, and $\rho - 2\rho^2 x + x^2({}^{120}/_{\rho^2} - {}^{5}/_{\rho})$ is a convex parabola (since ${}^{120}/_{\rho^2} - {}^{5}/_{\rho} > 0$) with symmetry axis parallel to the y-axis taking value $\rho > 0$ at 0 and having negative discriminant $\rho^4 + 5 - {}^{120}/_{\rho}$ (since $\rho < 1$, the discriminant is less than -114). We have thus shown that $p^{(1)}_{m+2} \leq R\gamma' p^{(1)}_m$. If $\sigma > m + 2$, the case $k = 1$ is not yet concluded. Since by the geometric decay proved so far

$$p^{(1)}_{m+2} < R\gamma' p^{(1)}_m < \varepsilon, \tag{A.52}$$

and by the definition of σ, $\delta' < p^{(2)}_{m+2} < 1 - \delta'$, the same estimates in (A.48), (A.50) and (A.51) apply to ϑ_{m+2}, ϑ'_{m+2} and ϑ''_{m+2} with the due shift of time indices. However, (A.49) does not apply automatically, since nothing guarantees that the same bound applies on the shifted increments of the $p^{(2)}$-component. Let us first assume that indeed also $|p^{(2)}_{m+2} - p^{(2)}_{m+1}| < \varepsilon$ and therefore that also (A.49) applies, with the due shift of indices. Then plugging them into Remark A.34 applied to $n = m + 1$ yields

$$\begin{aligned}
p^{(1)}_{m+3} &\leq p^{(1)}_{m+2}\left[1 - \rho + 2\rho(1 - q^{(1)}_{m+2}) + 2\frac{\varepsilon}{(\delta')^2}\left(3 + \frac{2}{\rho}\right)\right] \\
&= p^{(1)}_{m+2}\left[1 - \rho + 2\rho\left(q^{(1)}_{m+1} - \frac{1-\rho}{\rho}p^{(1)}_{m+1} + \frac{1}{\rho}p^{(1)}_{m+2}\right) + 2\frac{\varepsilon}{(\delta')^2}\left(3 + \frac{2}{\rho}\right)\right] \\
&\leq p^{(1)}_{m+1}\left[1 - \rho + 2\rho(1 - q^{(1)}_{m+1}) + 2\frac{\varepsilon}{(\delta')^2}\left(3 + \frac{2}{\rho}\right)\right]\Bigg[1 - \rho + 2\rho q^{(1)}_{m+1} \\
&+ 2(1-\rho)p^{(1)}_{m+1} + 2p^{(1)}_{m+2} + 2\frac{\varepsilon}{(\delta')^2}\left(3 + \frac{2}{\rho}\right)\Bigg] < p^{(1)}_{m+1}\Bigg[1 - \rho + 2\rho(1 - q^{(1)}_{m+1}) \\
&+ 2\frac{\varepsilon}{(\delta')^2}\left(3 + \frac{2}{\rho}\right)\Bigg]\left[1 - \rho + 2\rho q^{(1)}_{m+1} + 2\varepsilon\left(2 + \frac{1}{(\delta')^2}\left(3 + \frac{2}{\rho}\right)\right)\right].
\end{aligned}$$

Noting that

$$[1 - \rho + 2\rho(1 - q^{(1)}_{m+1})](1 - \rho + 2\rho q^{(1)}_{m+1}) = 1 - \rho^2 + 4\rho^2 q^{(1)}_{m+1}(1 - q^{(1)}_{m+1})$$

and that $1 - \rho + 2\rho(1 - q^{(1)}_{m+1}), 1 - \rho + 2\rho q^{(1)}_{m+1} < 1 + \rho$, we can multiply out the two factors and get the following estimate:

$$\begin{aligned}
p^{(1)}_{m+3} &< p^{(1)}_{m+1}\Bigg[1-\rho^2+4\rho^2 q^{(1)}_{m+1}(1-q^{(1)}_{m+1})+2(1+\rho)\varepsilon\left(1+\frac{1}{(\delta')^2}\left(3+\frac{2}{\rho}\right)\right)\\
&+4\frac{\varepsilon^2}{(\delta')^2}\left(3+\frac{2}{\rho}\right)\left(2+\frac{1}{(\delta')^2}\left(3+\frac{2}{\rho}\right)\right)\Bigg] < p^{(1)}_{m+1}\Bigg[1-4\rho^2(\delta')^2\\
&+4(\delta')^5+4(\delta')^3\left(3+\frac{2}{\rho}\right)+4(\delta')^8\left(3+\frac{2}{\rho}\right)\left(2+\frac{1}{(\delta')^2}\left(3+\frac{2}{\rho}\right)\right)\Bigg]\\
&< p^{(1)}_{m+1}\Bigg\{1-4\rho^2(\delta')^2+4(\delta')^3\Bigg[\left(3+\frac{2}{\rho}\right)+(\delta')^2+(\delta')^5\left(3+\frac{2}{\rho}\right)\\
&\left(2+\frac{1}{(\delta')^2}\left(3+\frac{2}{\rho}\right)\right)\Bigg]\Bigg\} \le p^{(1)}_{m+1}\Bigg\{1-4\rho^2(\delta')^2+4(\delta')^3\\
&\Bigg[\left(3+\frac{2}{\rho}\right)+(\delta')^2+\left(3+\frac{2}{\rho}\right)^2(\delta')^3+2\left(3+\frac{2}{\rho}\right)(\delta')^5\Bigg]\Bigg\}\\
&< p^{(1)}_{m+1}\left\{1-4\rho^2(\delta')^2+16(\delta')^3\left(5+\frac{5}{\rho}+\frac{1}{\rho^2}\right)\right\}\\
&< p^{(1)}_{m+1}\left\{1-4\rho^2(\delta')^2+\frac{240}{\rho^2}(\delta')^3\right\} \le R\gamma' p^{(1)}_m,
\end{aligned}$$

where in the second last inequality, we used trivially that $\delta' < 1$, as the factor next to the cubic term, for δ' small, becomes negligible, considering that regimes with ρ small are also allowed; in the last inequality we used trivially that $5 < 5/\rho < 5/\rho^2$ due to $0 < \rho < 1$. We now show that the upper bound on the $p^{(2)}$-component keeps applying uniformly, by using Remark A.35 applied to $n = m$, yielding the same upper bound as Lemma 3.45, by defining $E_{m+1} := \xi_{m+1} + \xi'_{m+1} + \eta_{m+1} + \eta'_{m+1} + \eta_m + \eta''_{m+1} + \eta'''_{m+1}$, that is obtaining

$$|p^{(2)}_{m+2} - p^{(2)}_{m+1}| \le |p^{(2)}_{m+1} - p^{(2)}_m| + E_{m+1}. \tag{A.53}$$

Since $p^{(1)}_{m+2} < R\gamma' p^{(1)}_m < \varepsilon$, by the definition of σ, which ensures that also $\delta' < p^{(2)}_{m+1} < 1-\delta'$, and by exploiting Remark A.33 with $n = m$, we can estimate

$$\eta_m = \frac{p^{(2)}_m p^{(1)}_m}{p^{(2)}_m + p^{(3)}_m} < p^{(1)}_m \tag{A.54}$$

$$\eta_{m+1} = \frac{p^{(2)}_{m+1} p^{(1)}_{m+1}}{p^{(2)}_{m+1} + p^{(3)}_{m+1}} < R p^{(1)}_m \tag{A.55}$$

$$\eta'_{m+1} := \frac{q^{(2)}_m}{q^{(1)}_m}\frac{p^{(2)}_{m+1}}{p^{(1)}_{m+1} + p^{(2)}_{m+1}}\frac{p^{(1)}_m}{1-p^{(2)}_m} < \frac{1}{(\delta')^2} p^{(1)}_m \tag{A.56}$$

$$\eta''_{m+1} := p^{(2)}_m \frac{p^{(1)}_{m+1}}{p^{(1)}_{m+1} + p^{(2)}_{m+1}} < \frac{1-\delta'}{\delta'} R p^{(1)}_m < \frac{R}{\delta'} p^{(1)}_m \tag{A.57}$$

$$\eta'''_{m+1} = p^{(2)}_m \frac{p^{(1)}_{m+1}}{p^{(1)}_{m+1} + p^{(2)}_{m+1}}\frac{p^{(1)}_m}{p^{(2)}_m + p^{(3)}_m} < p^{(1)}_m \tag{A.58}$$

$$\xi_{m+1} = \frac{p^{(2)}_{m+1}}{p^{(2)}_{m+1} + p^{(3)}_{m+1}}\left(1+\frac{\frac{1}{\rho}-1}{q^{(1)}_m}\right) p^{(1)}_{m+1} < \left[1+\frac{1}{\delta'}\left(\frac{1}{\rho}-1\right)\right] R p^{(1)}_m \tag{A.59}$$

$$\xi'_{m+1} = \frac{p^{(2)}_{m+1}}{p^{(2)}_{m+1} + p^{(3)}_{m+1}}\frac{\frac{1}{\rho}-1}{q^{(1)}_m} p^{(1)}_m < \frac{1}{\delta'}\left(\frac{1}{\rho}-1\right) p^{(1)}_m, \tag{A.60}$$

yielding

$$E_{m+1} \le D p^{(1)}_m.$$

Note that $R > 1$, and therefore from

$$\varepsilon' < \frac{\varepsilon}{4RD}(1-\gamma'),$$

we get

$$E_{m+1} \leq \frac{\varepsilon}{4R}.$$

Since $|p^{(2)}_{m+1} - p^{(2)}_m| < \varepsilon/2$, this yields

$$|p^{(2)}_{m+2} - p^{(2)}_{m+1}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{4R} < \varepsilon.$$

Apart from the base cases, this estimate will be less immediate in further steps, and we will be relying on Lemma A.37.

To summarise what has been shown in this two steps argument: there is a constant $\gamma' = \gamma(\delta')$ holding uniformly on $\mathcal{K}^*$ for both cases, $\sigma = m+2$ and $\sigma > m+2$. In the first case $p^{(1)}_{m+1} < R(\gamma')^0 p^{(1)}_m$ (case $k=0$), and $p^{(1)}_{m+2} < R\gamma' p^{(1)}_m$ (half case $k=1$); in the second case both $p^{(1)}_{m+1} < R(\gamma')^0 p^{(1)}_m$ (case $k=0$), and $p^{(1)}_{m+2} < R\gamma' p^{(1)}_m$ and $p^{(1)}_{m+3} < R\gamma' p^{(1)}_m$ (full case $k=1$). Note, before proceeding, that the estimate on $q^{(1)}_n$'s oscillations above and below $1/2$ has to iterate at each step. For example, for the next step it will hold, because

$$\begin{aligned} q^{(1)}_{m+2} &= 1 - q^{(1)}_{m+1} + \frac{1-\rho}{\rho} p^{(1)}_{m+1} - \frac{1}{\rho} p^{(1)}_{m+2} = q^{(1)}_m - \frac{1-\rho}{\rho} p^{(1)}_m + \frac{2-\rho}{\rho} p^{(1)}_{m+1} - \frac{1}{\rho} p^{(1)}_{m+2} \\ &< \frac{1}{2} - \delta + \frac{(2-\rho)R}{\rho} p^{(1)}_m < \frac{1}{2} - \delta' \end{aligned}$$

by construction of ε. Apart from the first few steps, this condition will not be so immediate to verify, because geometric terms will start adding up, and we will rely on Lemma A.36. Assume that $m+3 < n < \sigma$, for some n, and let us prove the claim for $n+1$. There are two cases to consider: the even step $n = m+2k-1$ to $n+1 = m+2k$ first, and the odd step $n = m+2k$ to $n+1 = m+2k+1$ afterwards, for all $k \in \mathbb{N}$ such that n is in the mentioned range.

- In the even step one has the induction hypothesis that for all $1 \leq j \leq 2k$,

$$p^{(1)}_{m+2k-j} < R(\gamma')^{\lfloor \frac{2k-j}{2} \rfloor} p^{(1)}_m \tag{A.61}$$

  and (A.46) needs to be shown. As to the oscillations of $q^{(1)}$, they are δ'-bounded away from $1/2$ in the correct order, thanks to (A.61) and Lemma A.36 applied with $b = 1$:

$$q^{(1)}_{m+2k-j} < \frac{1}{2} - \delta'$$

  for all even $1 \leq j \leq 2k$ (with bound for $j = 2k$ holding also with δ, by adopting empty sum convention) and

$$q^{(1)}_{m+2k-j} > \frac{1}{2} + \delta'$$

  for all odd $1 \leq j \leq 2k$. All that remains to be shown is that

$$p^{(1)}_{m+2k} < (\gamma') p^{(1)}_{m+2k-1},$$

  by using

$$q^{(1)}_{m+2k-1} > \frac{1}{2} + \delta'.$$

Since by (A.61),

$$p^{(1)}_{m+2k-1} < R(\gamma')^{k-1} p^{(1)}_m,$$

(A.52) holds for $p^{(1)}_{m+2k-1}$ too, and $\sigma > m+2k-1$ implies that $\delta' < p^{(2)}_{m+2k-1} < 1-\delta'$; the estimates in (A.48) to (A.51) apply also to ϑ_{m+2k-1}, ϑ'_{m+2k-1} and ϑ''_{m+2k-1} (with the due shift of time indices) since by Lemma A.37 with $b=1$, $|p^{(2)}_{m+2k-1} - p^{(2)}_{m+2k-1}| < \varepsilon$. Plugging the aforementioned estimates into Remark A.34 applied to $n = m+2k-2$, yields the same estimate as that obtained for $p^{(1)}_{m+2}$,

$$p^{(1)}_{m+2k} \le p^{(1)}_{m+2k-1}\left\{(1-\rho) + \rho\left[2(1-q^{(1)}_{m+2k-1}) + 2\varepsilon\left(\frac{1}{1-\varepsilon} + \left(1+\frac{2}{\rho}\right)\frac{1}{\delta'(1-\varepsilon)} + \frac{1}{(\delta')^2}\right)\right]\right\} < p^{(1)}_{m+2k-1}\left[1 - 2\rho\delta' + \frac{10}{\rho}(\delta')^3\right] < \gamma' p^{(1)}_{m+2k-1},$$

resulting into (A.46) by (A.61).

- In the odd step one has (A.61) holding for all $0 \le j \le 2k$, and (A.47) needs to be shown.For the oscillations of $q^{(1)}$ we proceed similarly but, with a different range for j, by exploiting Lemma A.36 applied with $b=0$:

$$q^{(1)}_{m+2k-j} < \frac{1}{2} - \delta'$$

for all even $0 \le j \le 2k$ (with bound for $j = 2k$ holding also with δ, by adopting empty sum convention) and

$$q^{(1)}_{m+2k-j} > \frac{1}{2} + \delta'$$

for all odd $0 \le j \le 2k$. In particular, being known that $q^{(1)}_{m+2k-1} > 1/2 + \delta'$, the part concerning p is obvious from the same calculation performed in the third step, since knowing $p^{(1)}_{m+2k} < R(\gamma')^k p^{(1)}_m$, using the same bounds for ϑ_{m+2k}, ϑ'_{m+2k} and ϑ''_{m+2k}, yields $p^{(1)}_{m+2k+1} < R(\gamma')^k p^{(1)}_m$. All that has to be shown explicitly is that

$$p^{(1)}_{m+2k+1} < \gamma' p^{(1)}_{m+2k-1},$$

by using

$$q^{(1)}_{m+2k-1} > \frac{1}{2} + \delta'.$$

Since by the geometric decay in (A.61), for the new range of indices,

$$p^{(1)}_{m+2k-1} < R(\gamma')^{k-1} p^{(1)}_m,$$

extending (A.52), and $\sigma > m+2k$ implies $\delta' < p^{(2)}_{m+2k} < 1-\delta'$; then the estimates in (3.53) to (3.55) apply also to ϑ_{m+2k}, ϑ'_{m+2k} and ϑ''_{m+2k} (with the due shift of time indices) since $|p^{(2)}_{m+2k} - p^{(2)}_{m+2k-1}| < \varepsilon$, because we can apply Lemma A.37 with $b=0$; also the previous step's estimates for ϑ_{m+2k-1} and ϑ'_{m+2k-1} keep applying, and they are vital, since in this step the bound needed, is yielded by iterating the previous even step into the current odd one, producing a two-step estimate, because a one step estimate would not yield a subunitary constant, due to $q^{(1)}_{m+2k} < 1/2 - \delta'$, which would imply $2(1-q^{(1)}_{m+2k}) > 1$. Therefore, by plugging these estimates into Remark A.34 applied to $n = m+2k-1$, and also using the estimate from the previous even step, yields the same estimate as that obtained for $p^{(1)}_{m+3}$:

$$p^{(1)}_{m+2k+1} \leq p^{(1)}_{m+2k-1}\left[1-\rho+2\rho(1-q^{(1)}_{m+2k-1})+2\frac{\varepsilon}{(\delta')^2}\left(3+\frac{2}{\rho}\right)\right]$$
$$\left[1-\rho+2\rho q^{(1)}_{m+2k-1}+2\varepsilon\left(2+\frac{1}{(\delta')^2}\left(3+\frac{2}{\rho}\right)\right)\right]$$
$$< p^{(1)}_{m+2k-1}\left\{1-4\rho^2(\delta')^2+\frac{240}{\rho^2}(\delta')^3\right\} = \gamma' p^{(1)}_{m+2k-1}$$

resulting into (A.47) by (A.61).

Having shown (A.46) and (A.47), we can easily derive the main claim by simply setting $\gamma := \sqrt{\gamma'}$, so as to express the two-steps geometric decaying upper bound as a one-step geometric decaying one. Equivalently, it has been shown that for all $1 \leq l \leq \sigma - m$,

$$p^{(1)}_{m+l} < R(\gamma')^{\lfloor \frac{l}{2} \rfloor} p^{(1)}_m.$$

Since

$$\left\lfloor \frac{l}{2} \right\rfloor \geq \frac{l-1}{2},$$

it follows that

$$p^{(1)}_{m+l} < R\sqrt{\gamma'}^{\,l-1} p^{(1)}_m,$$

hence for the uniform constant γ, we have that for all $m < n \leq \sigma$,

$$p^{(1)}_n < R\gamma^{n-m-1} p^{(1)}_m.$$

□

For any $\tau \geq m$ we define

$$\zeta := \inf\left\{n > \tau : \frac{|p^{(2)}_{n+1} - p^{(2)}_n|}{p^{(1)}_n} < \frac{1}{\Gamma}\right\}.$$

Lemma A.39. *Suppose that there exists $m \leq \tau < \sigma$, such that*

$$\frac{p^{(1)}_\tau}{|p^{(2)}_{\tau+1} - p^{(2)}_\tau|} \leq \Gamma.$$

Then for all $\tau \leq n \leq \zeta \wedge \sigma$,

$$|p^{(2)}_{n+1} - p^{(2)}_n| < \rho^{n-m}|p^{(2)}_{m+1} - p^{(2)}_m|.$$

Proof. We proceed like in Lemma 3.47, by showing the claim explicitly only for $\tau = m$. If $\zeta = m+1$ we need to show the claim only for $n = m+1$, since for $n = m$ it is trivial. It is known that only the condition $p^{(1)}_m \leq \Gamma|p^{(2)}_{m+1} - p^{(2)}_m|$ holds, and $p^{(1)}_{m+1} < Rp^{(1)}_m$ by Remark A.33 applied to $n = m$, along with the hypotheses made in (A.26) to (A.28).

Then by (A.54) to (A.58), it follows that

$$\eta_m < p_m^{(1)} \leq \Gamma|p_{m+1}^{(2)} - p_m^{(2)}| \tag{A.62}$$
$$\eta_{m+1} < Rp_m^{(1)} \leq R\Gamma|p_{m+1}^{(2)} - p_m^{(2)}| \tag{A.63}$$
$$\eta'_{m+1} < \frac{1}{(\delta')^2}p_m^{(1)} \leq \frac{\Gamma}{(\delta')^2}|p_{m+1}^{(2)} - p_m^{(2)}| \tag{A.64}$$
$$\eta''_{m+1} < \frac{R}{\delta'}p_m^{(1)} \leq \frac{R}{\delta'}\Gamma|p_{m+1}^{(2)} - p_m^{(2)}| \tag{A.65}$$
$$\eta'''_{m+1} < p_m^{(1)} \leq \Gamma|p_{m+1}^{(2)} - p_m^{(2)}| \tag{A.66}$$
$$\xi_{m+1} < \left[1 + \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right)\right] Rp_m^{(1)} \leq \left[1 + \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right)\right] R\Gamma|p_{m+1}^{(2)} - p_m^{(2)}| \tag{A.67}$$
$$\xi'_{m+1} < \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right) p_m^{(1)} \leq \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right) \Gamma|p_{m+1}^{(2)} - p_m^{(2)}|. \tag{A.68}$$

Plugging these estimates into Remark A.35 applied to $n = m$ yields

$$|p_{m+2}^{(2)} - p_{m+1}^{(2)}| < \rho q_m^{(1)} (1 + \Gamma D) |p_{m+1}^{(2)} - p_m^{(2)}|.$$

Since by construction $\Gamma < \delta'/[D(1-\delta')]$,

$$q_m^{(1)}(1 + \Gamma D) < q_m^{(1)} \left(1 + \frac{\delta'}{1-\delta'}\right) = \frac{q_m^{(1)}}{1-\delta'} \leq 1,$$

and therefore it follows that

$$|p_{m+2}^{(2)} - p_{m+1}^{(2)}| \leq \rho |p_{m+1}^{(2)} - p_m^{(2)}|,$$

and the claim for $\zeta = m+1$ and the first step of the induction is complete.

Assuming now $\zeta > m+1$, we show the rest by induction as in Lemma 3.47. Recall that $\delta' < \delta < q_m^{(1)} < 1-\delta < 1-\delta'$ and that

$$\varepsilon < \frac{\rho\delta'}{(2-\rho)\left(1 + \frac{R}{1-\gamma}\right)}.$$

Since for all $m \leq n \leq \zeta \wedge \sigma$, by Lemma A.38, it holds that $p_n^{(1)} < R\gamma^{n-m-1}p_m^{(1)}$, or equivalently that $p_{m+k}^{(1)} < R\gamma^{k-1}\varepsilon$ for all $k \in \mathbb{N}$ such that $n = m+k$ is within the bounds above; by (A.32), for all such k,

$$q_{m+k}^{(1)} \geq \begin{cases} q_m^{(1)} - \frac{2-\rho}{\rho}\sum_{j=0}^k p_{m+j}^{(1)} \geq \delta - \frac{(2-\rho)}{\rho}\varepsilon\left(1 + R\sum_{j=0}^k \gamma^j\right) \\ > \delta - \frac{2-\rho}{\rho}\varepsilon\left(1 + \frac{R}{1-\gamma}\right) > \delta' & k \text{ even} \\ 1 - q_m^{(1)} - \frac{2-\rho}{\rho}\sum_{j=0}^k p_{m+j}^{(1)} \geq \delta - \frac{(2-\rho)}{\rho}\varepsilon\left(1 + R\sum_{j=0}^k \gamma^j\right) \\ \geq \delta - \frac{2-\rho}{\rho}\varepsilon\left(1 + \frac{R}{1-\gamma}\right) > \delta' & k \text{ odd.} \end{cases}$$

This ensures that estimating η'_n, ξ_n and ξ'_n with the constants, which upper bound reciprocals of $q^{(1)}$, can carry out during the induction step. As to the constants, which lower bound reciprocals involving $q^{(1)}$, one can proceed analogously:

$$q_{m+k}^{(1)} \leq \begin{cases} q_m^{(1)} + \frac{2-\rho}{\rho}\sum_{j=0}^k p_{m+j}^{(1)} \leq 1 - \delta + \frac{(2-\rho)}{\rho}\varepsilon\left(1 + R\sum_{j=0}^k \gamma^j\right) \\ < 1 - \delta + \frac{2-\rho}{\rho}\varepsilon\left(1 + \frac{R}{1-\gamma}\right) < 1 - \delta' & k \text{ even} \\ 1 - q_m^{(1)} + \frac{2-\rho}{\rho}\sum_{j=0}^k p_{m+j}^{(1)} \leq 1 - \delta + \frac{(2-\rho)}{\rho}\varepsilon\left(1 + R\sum_{j=0}^k \gamma^j\right) \\ < 1 - \delta + \frac{2-\rho}{\rho}\varepsilon\left(1 + \frac{R}{1-\gamma}\right) < 1 - \delta' & k \text{ odd.} \end{cases}$$

The inductive hypothesis is then like in Lemma 3.47: assuming that for some $k \geq 0$ such that $m+k+1 < \zeta \wedge \sigma$, $|p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| < \rho^k |p^{(2)}_{m+1} - p^{(2)}_{m}|$, we need to show that $|p^{(2)}_{m+k+2} - p^{(2)}_{m+k+1}| < \rho^{k+1} |p^{(2)}_{m+1} - p^{(2)}_{m}|$, and it will be done by showing that $|p^{(2)}_{m+k+2} - p^{(2)}_{m+k+1}| < \rho |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|$. Since by the definition of σ it still holds that $\delta' \leq p^{(2)}_{m+k+1} \leq 1-\delta'$, by the definition of ζ it still holds that $p^{(1)}_{m+k} \leq \Gamma |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|$ and the geometric decay of the first component ensures the bounds on $q^{(1)}_{m+k}$ and $q^{(1)}_{m+k+1}$ as shown above, it follows that

$$
\begin{aligned}
\eta_{m+k} &< p^{(1)}_{m+k} \leq \Gamma |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\eta_{m+k+1} &< R p^{(1)}_{m+k} \leq R\Gamma |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\eta'_{m+k+1} &< \frac{1}{(\delta')^2} p^{(1)}_{m+k} \leq \frac{\Gamma}{(\delta')^2} |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\eta''_{m+k+1} &< \frac{R}{\delta'} p^{(1)}_{m+k} \leq \frac{R}{\delta'} \Gamma |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\eta'''_{m+k+1} &< p^{(1)}_{m+k} \leq \Gamma |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\xi_{m+k+1} &< \left[1 + \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right)\right] R p^{(1)}_{m+k} \leq \left[1 + \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right)\right] R\Gamma |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}| \\
\xi'_{m+k+1} &< \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right) p^{(1)}_{m+k} \leq \frac{1}{\delta'}\left(\frac{1}{\rho} - 1\right) \Gamma |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|.
\end{aligned}
$$

Plugging these estimates into Remark A.35 applied to $n = m+k$ yields

$$
|p^{(2)}_{m+k+2} - p^{(2)}_{m+k+1}| < \rho q^{(1)}_{m+k} (1 + \Gamma D) |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|.
$$

Since by construction $\Gamma < \delta'/[D(1-\delta')]$,

$$
q^{(1)}_{m+k}(1+\Gamma D) < q^{(1)}_{m+k}\left(1 + \frac{\delta'}{1-\delta'}\right) = \frac{q^{(1)}_{m+k}}{1-\delta'} \leq 1,
$$

and therefore it follows that

$$
|p^{(2)}_{m+k+2} - p^{(2)}_{m+k+1}| \leq \rho |p^{(2)}_{m+k+1} - p^{(2)}_{m+k}|,
$$

and the induction is complete. If $m < \tau < \sigma$, the same concluding remarks apply, as in Lemma 3.47. □

In the following theorem, the proof of convergence, given that the previous results (Lemmas A.38 and A.39) have been rederived in correspondingly to the old ones (Lemmas 3.46 and 3.47), plays out exactly as Theorem 3.48, with only a small change: whenever the factor $1+2/\Gamma$ appears, it must be replaced with the factor $1+R/\Gamma$, because the factor of 2 in the statement of Lemma 3.46 has been replaced by the factor R in Lemma A.38. Also $c = \rho$, in the notation of Section 3.5.3. The rest of the proof then does not change, since in the condition defining ε' the term

$$
\frac{\delta\Gamma\lambda(1-\lambda)}{2(R+\Gamma)}
$$

has now replaced

$$
\frac{\delta\Gamma\lambda(1-\lambda)}{2(2+\Gamma)}.
$$

Hence it is not necessary to repeat the modified argument.

Theorem A.40. *$\{p_n\}$ converges to some $p_* \in E_1$.*

The proof of the next corollary also does not change. It is enough to appeal to the corresponding lemmas used from Section 3.5.3 (Theorem 3.48, Corollary 3.39, Remark 3.36, and Lemmas 3.19 and 3.46) which have been rederived in this section (Theorem A.40, Corollary A.31, Remark A.28, and Lemmas A.38 and 3.19).

Corollary A.41. *As p_n converges to some $p_* \in E_1$, q_n is asymptotically 2-periodic to $\{q_{p_*} \pm \frac{\ell}{2} e_{-1}(p_*)\}$.*

Then the same conclusive remarks of Section 3.5.3 apply, *mutatis mutandi*, leading to the following.

Remark A.42. *If $(p_0, q_0) \in \mathcal{K}_{\varepsilon',\delta}$, then by Theorem A.40, for some $p_* \in E_1$, $p_n \longrightarrow p_*$ and $\{q_n\}$ diverges.*

The same is not true for the argument in Section A.4, as explained in Section 3.5.3.

Remark A.43. *If $p_* \in E_i$ for $i \in \{2, 3\}$, one can proceed by exploiting the symmetry of the model, define σ, ζ_i and τ_i accordingly in terms of the corresponding coordinates, and show an analogous version of Theorem A.40 and Remark A.42 for $i \in \{2, 3\}$ as well, thus yielding convergence of $\{p_n\}$ to some $p_* \in \partial\Sigma \setminus V$ and asymptotic 2-periodicity of $\{q_n(\omega)\}$ to $\{q_{p_*} \pm \frac{\ell}{2} e_{-1}(p_*)\}$ for any orbit having $\ell > 0$ and a subsequence bounded away from the vertices.*

A.6 Convergence of the dynamical system

In this section we finally put together all the convergence results gathered so far to show firstly the convergence of $\{p_n\}$, secondly that $\{q_n\}$ may or may not converge.

Proof of Theorem 2.1. Let $p_0 \notin \partial\Sigma$. By Lemma A.8 the limit ℓ of the potential function exists. If $\{p_n\}$ is bounded away from the boundary, it converges by Proposition A.9. If $\ell = 0$ and $\{p_n\}$ is not bounded away from the boundary, it converges by Remark A.20. If $\ell > 0$ and $\{p_n\}$ is not bounded away from the boundary, it converges by Remark A.43. By mutual exclusion the only case left is convergence to a vertex. Let $p_0 \in \partial\Sigma \setminus V$ and $q_0 \in \Sigma_0$. Then $\{p_n\}$ converges by Lemma A.21. Let $p_0 \in E_i$ and $q_0 \in V$. Then $\{p_n\}$ converges by Remarks A.1 and A.2. □

Proof of Corollary 2.2. Let $p_0 \notin \partial\Sigma$. By Lemma A.8 the limit ℓ of the potential function exists. By Theorem 2.1 if $\ell = 0$ the convergence to $\overline{\Sigma}^*$ is trivial. If $\ell > 0$ the convergence to the limit 2-cycle follows either by Remark A.43 if $p_* \in V$, or by the introductory remarks to Section A.5.3. Let $p_0 \in \partial\Sigma \setminus V$ and $q_0 \in \Sigma_0$. Then $\{q_n\}$ either converges in $\partial\Sigma^* \subset \overline{\Sigma}^*$ or is asymptotic to a 2-cycle by Corollary A.25. Let $p_0 \in E_i$ and $q_0 \in V$. Then $\{q_n\}$ is 2-periodic by Remarks A.1 and A.2, and thus trivially asymptotic to a 2-cycle. □

Appendix B

Construction of $\vartheta(\mu)$ and $\theta(\mu)$

In Section 4.4, prior to Lemma 4.13, we discussed extensively the reasons behind choosing a $1 < \nu < \sqrt{\mu}$, as a function of $\mu > 1$, defined as

$$\nu = \nu(\mu) := \begin{cases} \frac{7}{5}, & \mu \geq 2 \\ \mu^{\frac{1}{2}-\vartheta}, & 1 < \mu < 2, \end{cases}$$

where $\vartheta = \vartheta(\mu)$ can be determined such that $0 < \vartheta < 1/2$. While the case $\mu \geq 2$ is an easy guess, the case $1 < \mu < 2$ is not so easy to guess, so a constructive approach, which requires determining the functions $\vartheta(\mu)$ and $\theta(\mu)$, is what leads to the guess made in Lemma 4.13. In this appendix we show how to construct them.

Let $\eta = \eta(\mu) := \min\{\rho, 1-\rho\}$. Clearly m can be so large to ensure $\rho - \eta < \rho_{n+1} < \rho + \eta$ for all $n \geq m$, where

$$\eta = \eta(\mu) := \min\{\rho, 1-\rho\} = \begin{cases} 1-\rho, & \mu \geq 2 \\ \rho, & 1 < \mu < 2, \end{cases}$$

since for all $\mu > 1$, $\mu \geq 2$ is equivalent to $1 - \rho \leq \rho$ as $\rho = (\mu-1)/\mu$, thus implying also that $\mu < 2$ is equivalent to $\rho < 1 - \rho$. Define

$$a = a(\theta, \mu) := \frac{\rho + \eta + (4 + \rho + \eta)\theta}{2}$$

and

$$b = b(\theta, \mu) := \frac{a(\theta, \mu)}{d(\theta, \mu)},$$

where

$$d(\theta, \mu) := \left[1 - \left(1 - \frac{1}{\sqrt[3]{\mu}}\right)\theta\right]\left[1 - 2\frac{2 - \rho + \eta + 2\theta}{3(1 + \sqrt[3]{\mu})}(\sqrt[3]{\mu} - 1)\theta\right].$$

We have seen, in Lemma 4.13, how the requirement $0 < \theta < 1/2$ implies the positivity of the factors appearing in d. The constant $\theta = \theta(\mu)$ is to be determined and thus fixed along with $\vartheta(\mu)$, such that, for all $\mu \geq 2$, $0 < a < 3/4$ and $0 < b < 1$ and $3/4 > 1/\nu$; while for all $1 < \mu < 2$ we require the conditions $0 < a < b < 1$ and $a > 1/\nu$. That this is possible for all $\mu \geq 2$ has been shown in Lemma 4.13. The construction that leads to all these conditions to be met also for $1 < \mu < 2$, requires more work.

First of all it needs to be shown that it is possible to find $0 < \vartheta, \theta < 1/2$ such that

$$a(\theta, \mu) > \frac{1}{\nu(\mu)} = \mu^{\vartheta(\mu) - \frac{1}{2}}. \tag{B.1}$$

Next one needs to show that the θ thus found ensures that $0 < a, b < 1$. (B.1) is equivalent to

$$\frac{1}{2}\log\mu + \log a > \vartheta\log\mu,$$

which is equivalent to

$$\vartheta < \frac{1}{2} + \frac{\log a}{\log\mu}.$$

Thus the existence of ϑ can be ensured by requiring that $\theta = \theta(\mu)$ be such that

$$-\frac{1}{2} < \frac{\log a(\theta,\mu)}{\log\mu} < 0,$$

or equivalently, since $\eta = \rho$, that

$$\frac{1}{\sqrt{\mu}} < \rho + (2+\rho)\theta < 1,$$

and then by finding explicitly a parametrisation $\theta = \theta(\mu)$ that satisfies it for all $1 < \mu < 2$ (on top of ensuring $0 < a, b < 1$). Once this is done, we can use any

$$0 < \vartheta(\mu) < \frac{1}{2} + \frac{1}{\log\mu}\log\left[\rho + (2+\rho)\theta\right],$$

since the specific form will not be playing an explicit role in the argument of Theorem 4.18, so we will just take, for example, half the upper bound as our ϑ. On the other hand, we will determine θ explicitly, as it will play an important role in all estimates from Lemma 4.14 onward. The condition that ensures the existence of ϑ can be rephrased as a system of two inequalities:

$$\rho + (2+\rho)\,\theta < 1 \tag{B.2}$$

$$\rho + (2+\rho)\,\theta > \frac{1}{\sqrt{\mu}}. \tag{B.3}$$

We will first study this as a subset of the (θ,ρ)-plane's first quadrant, due to $0 < \rho, \theta < 1/2$. First we rearrange (B.2) and (B.3), so that they are shown to describe regions of the plane delimited by rectangular hyperbolic boundaries. Indeed, (B.2) is equivalent, by adding and subtracting 1 to θ and rearranging, to

$$(\rho+2)(\theta+1) < 3.$$

This is the part of the plane strictly between the two branches of the rectangular hyperbola of equation $(\rho+2)(\theta+1) = 3$, which has a centre of coordinates $(-1,-2)$. One branch is on the third quadrant, the only relevant branch goes through the first quadrant and meets the θ-axis at $(1/2, 0)$ and the ρ-axis at $(0,1)$. In Figure B.1 we can see the the rectangular hyperbola defining the boundary of the region corresponding to the inequality (B.2) in solid gray (concealed partially by the dashed blue line, as they coincide). The region corresponding to (B.2), which is independent of μ, is the one closer to the origin, between the θ-axis, the ρ-axis and the gray hyperbola. Through a similar manipulation as that used for (B.2), (B.3) is shown to be equivalent to

$$(2+\rho)(1+\theta) > 2 + \frac{1}{\sqrt{\mu}}.$$

This is the part of the plane strictly outside the two branches of the rectangular hyperbola of equation $(2+\rho)(1+\theta) = 2 + 1/\sqrt{\mu}$, which has a centre of coordinates

$(-1,-2)$. The simultaneous inequalities (B.2) and (B.3) admit therefore a solution, with positive values for θ and ρ, due to the fact that $2+{}^{1}/\sqrt{2} < 2+{}^{1}/\sqrt{\mu} < 3$. Indeed, one branch is on the negative quadrant, the only relevant branch goes through the positive quadrant and meets the θ-axis at $({}^{1}/2\sqrt{\mu}, 0)$ and the ρ-axis at $(0, {}^{1}/\sqrt{\mu})$. In Figure B.1 we can see the hyperbolas defining the boundary of the region corresponding to the inequality (B.3) and positive values of θ and ρ, for the limit parameters $\mu = 1$, in dashed blue (coinciding with the hyperbola for (B.2)), and $\mu = 2$, in dashed red. The region corresponding to Figure B.1 is obviously always the one further from the origin, between the θ-axis, the ρ-axis and above the dashed hyperbolic branch corresponding to the value of μ at hand (clearly all the branches for $1 < \mu < 2$ are in between the limit ones). The solution to (B.2) and (B.3) is thus represented by the area between the two branches (the dashed one corresponding to the value of μ at hand and the gray one, which is independent of μ, see Figure B.1 for the solution corresponding to $\mu = {}^{3}/2$) and the coordinate axes. For $\mu = 2$ this area is the largest and vanishes as $\mu \longrightarrow 1$. Due to $0 < {}^{1}/\sqrt{\mu} < 1$ for all $\mu > 1$, the area in between the solid gray and dashed branch is always nonnegligible and represents the solution for positive values. What we need to find is a parametrisation of θ in terms of μ that always falls within this feasible area. Clearly, since $\sqrt[3]{\mu} < \sqrt{\mu}$ for all $\mu > 1$, ${}^{1}/\sqrt{\mu} < {}^{1}/\sqrt[3]{\mu} < 1$, so an option is to choose all $\theta > 0$ such that $(2+\rho)(1+\theta) = 2 + {}^{1}/\sqrt[3]{\mu}$. This is a consistent choice, since we have fixed $\eta = \rho = 1 - {}^{1}/\mu < {}^{1}/2$, and the hyperbola chosen allows for $\rho \in (0, {}^{1}/\sqrt[3]{\mu})$. Indeed, having $1 < \mu < 2$,

$$\rho = 1 - \frac{1}{\mu} < \frac{1}{\sqrt[3]{\mu}}$$

since while $1 - {}^{1}/\mu < {}^{1}/2$, ${}^{1}/\sqrt[3]{\mu} > {}^{1}/\sqrt[3]{2} > {}^{1}/2$. Hence the intersection between the horizontal line, drawn at the height of the fixed value given by ρ and the hyperbola in question, is always well defined. Hence for $1 < \mu < 2$ we can set

$$\theta(\mu) := \frac{2 + \frac{1}{\sqrt[3]{\mu}}}{2+\rho} - 1 = \frac{2 + \frac{1}{\sqrt[3]{\mu}}}{3 - \frac{1}{\mu}} - 1$$

(see Figure B.2 for an intuitive representation depicting the case $\mu = {}^{3}/2$, where the solid black hyperbola represents the hyperbolic subset of solution chosen to derive $\theta(\mu)$ explicitly from the $\rho(\mu)$). Note that this definition is consistent with $0 < \theta(\mu) < {}^{1}/2$, as $1 < \mu < 2$. As aforementioned, one can let, for example,

$$\vartheta(\mu) := \frac{1}{4} + \frac{1}{2\log\mu}\log a.$$

Note that by adding and subtracting 1 to θ and recalling that we chose $\theta(\mu)$ such that

$$(2+\rho)(1+\theta) = 2 + \frac{1}{\sqrt[3]{\mu}},$$

we obtain

$$a = \rho + (2+\rho)\theta = \rho + (2+\rho)(1+\theta) - (2+\rho) = \rho + 2 + \frac{1}{\sqrt[3]{\mu}} - 2 - \rho = \frac{1}{\sqrt[3]{\mu}}.$$

As a result of $a = {}^{1}/\sqrt[3]{\mu}$, and it follows that

$$\vartheta(\mu) := \frac{1}{4} + \frac{1}{2\log\mu}\log\left(\mu^{-\frac{1}{3}}\right) = \frac{1}{12},$$

thus yielding

$$\nu = \mu^{\frac{1}{2}-\vartheta} = \mu^{\frac{5}{12}},$$

and the set up of the constants is thus complete. The final check that $0 < a < b < 1$ has been done in Lemma 4.13.

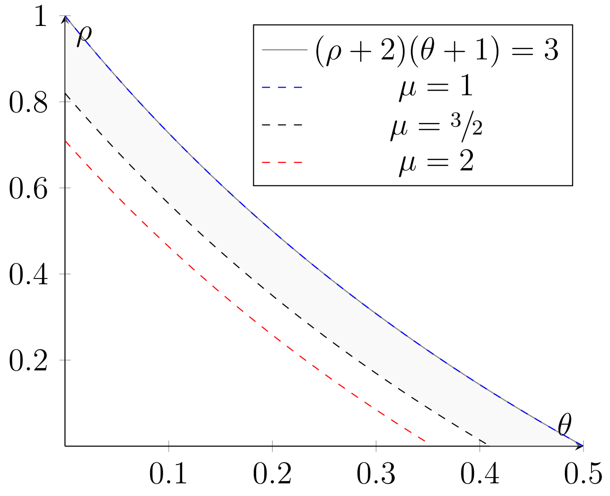


Figure B.1: Nonnegligibility of the positive solution to (B.2) and (B.3)

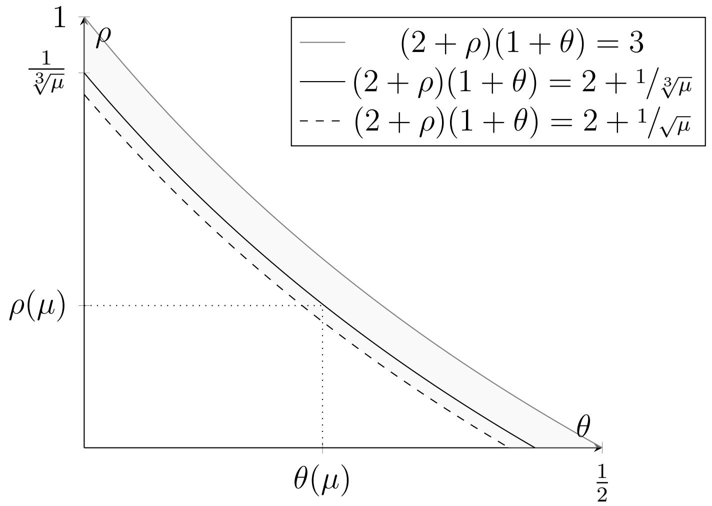


Figure B.2: Finding $\theta(\mu)$ given $\rho(\mu)={}^1\!/_3$ for $\mu={}^3\!/_2$

Part IV

Supplements to Part II

Appendix C

Supplements to Chapter 6

In this chapter, for self-containedness, we include a few brief results regarding monopoly, which can be found, respectively, in [48, Lemma 2.3, Lemma 2.2, Theorem 1.3, Theorem 1.4]. Recall the usual notation for the BB model: α is the feedback parameter, σ_n is the integer-valued time-dependent number of balls thrown at the bins, $\tau_n := \tau_0 + \sum_{i=1}^{n} \sigma_i$, $\rho_n := {}^{\sigma_{n+1}}/_{\tau_n}$, $\theta \longleftarrow \theta_n := \alpha^{-n} \log \tau_n$, $\lambda = \limsup_{n\to\infty} \sigma_{n+1}\sigma_{n-1}^{\alpha}\sigma_n^{-\alpha-1}$, and monopoly is the event $\mathcal{M}$ for which all but one bin receive finitely many balls. Recall also the claim proved in Lemma 6.1: if

$$\sum_{n=1}^{\infty} \frac{\sigma_{n+1}}{\tau_n^{\alpha}} = \infty,$$

then $\mathbb{P}(\mathcal{M}) = 0$.

Lemma C.1.

$$\sum_{n=1}^{\infty} \frac{\sigma_n}{\tau_n} = \infty.$$

Proof. Note first that

$$\tau_n = \tau_{n-1}\left(\frac{\tau_{n-1}}{\tau_n}\right)^{-1} = \tau_{n-1}\left(1 - \frac{\sigma_n}{\tau_n}\right)^{-1} = \tau_0 \prod_{k=1}^{n}\left(1 - \frac{\sigma_k}{\tau_k}\right)^{-1} =$$

$$\tau_0 \exp - \sum_{k=1}^{n} \log\left(1 - \frac{\sigma_k}{\tau_k}\right).$$

Thus as $n \longrightarrow \infty$,

$$-\sum_{k=1}^{n} \log\left(1 - \frac{\sigma_k}{\tau_k}\right) = \log\frac{\tau_n}{\tau_0} \longrightarrow \infty.$$

Without loss of generality, assume ${}^{\sigma_n}/_{\tau_n} \longrightarrow 0$ (otherwise, there is nothing to prove, the series would be divergent by assumption). It is then possible to expand in Taylor series the logarithm and get the following:

$$-\log\left(1 - \frac{\sigma_k}{\tau_k}\right) = \frac{\sigma_k}{\tau_k} + \frac{\sigma_k^2}{2\tau_k^2} + o\left(\frac{\sigma_k^2}{\tau_k^2}\right).$$

It is now possible to apply the *limit comparison* test to the two series $\sum_{n=1}^{\infty} {}^{\sigma_n}/_{\tau_n}$ and $-\sum_{n=1}^{\infty} \log\left(1 - {}^{\sigma_n}/_{\tau_n}\right)$:

$$\frac{\frac{\sigma_n}{\tau_n}}{-\log\left(1 - \frac{\sigma_n}{\tau_n}\right)} = \frac{\frac{\sigma_n}{\tau_n}}{\frac{\sigma_n}{\tau_n} + \frac{\sigma_n^2}{2\tau_n^2} + o\left(\frac{\sigma_n^2}{\tau_n^2}\right)} = \frac{1}{1 + \frac{\sigma_n}{2\tau_n} + o\left(\frac{\sigma_n}{\tau_n}\right)} \longrightarrow 1.$$

Thus either both series diverge or both converge. From what has been noted first, the claim follows. □

In the next theorem we show that there is no monopoly without feedback.

Proof of Theorem 1.14. Lemma C.1 and $\tau_{n-1} < \tau_n$ imply that

$$\infty = \sum_{n=1}^{\infty} \frac{\sigma_n}{\tau_n} < \sum_{n=1}^{\infty} \frac{\sigma_n}{\tau_{n-1}} = \sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_n},$$

so by Lemma 6.1 the claim follows. □

In the next theorem we show that if there is feedback and $\theta = \infty$ (supercritical regime), then there is no monopoly.

Proof of Theorem 1.15. Assume by contradiction that

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_n^\alpha} < \infty.$$

Then $\sigma_n/\tau_{n-1}^\alpha$ vanishes, hence there is a k such that for all $n \geq k$, $\sigma_n < \tau_{n-1}^\alpha$. Hence for all such n,

$$\tau_n = \tau_{n-1} + \sigma_n \leq \tau_{n-1} + \tau_{n-1}^\alpha \leq 2\tau_{n-1}^\alpha.$$

Iterating this $n-k$ times yields

$$\tau_n \leq 2^{\sum_{i=0}^{n-k-1} \alpha^i} \tau_k^{\alpha^{n-k}} = 2^{\frac{\alpha^{n-k}-1}{\alpha-1}} \tau_k^{\alpha^{n-k}} \leq 2^{\frac{\alpha^{n-k}}{\alpha-1}} \tau_k^{\alpha^{n-k}} = \left(2^{\frac{1}{\alpha-1}} \tau_k\right)^{\alpha^{n-k}},$$

but then, for all $n \geq k$,

$$\theta_n \leq \frac{\alpha^{n-k} \log\left(2^{\frac{1}{\alpha-1}} \tau_k\right)}{\alpha^n} = \frac{\log\left(2^{\frac{1}{\alpha-1}} \tau_k\right)}{\alpha^k},$$

and therefore the contradiction

$$\theta \leq \frac{\log\left(2^{\frac{1}{\alpha-1}} \tau_k\right)}{\alpha^k} < \infty$$

is reached. The claim now follows by Lemma 6.1. □

In the next theorem we show that with feedback and $\theta = 0$ (subcritical regime), if $\rho_n \longrightarrow \infty$ and $\lambda > 1$, there is no monopoly.

Proof of Theorem 1.16. Rewrite

$$\frac{\frac{\sigma_{n+1}}{\tau_n^\alpha}}{\frac{\sigma_n}{\tau_{n-1}^\alpha}} = \frac{\rho_n}{\rho_{n-1}} \left(\frac{\tau_{n-1}}{\tau_n}\right)^{\alpha-1} = \frac{\rho_n}{\rho_{n-1}^\alpha} \left(\frac{1}{1+\rho_{n-1}}\right)^{\alpha-1} = \frac{\rho_n}{\rho_{n-1}^\alpha} \left(\frac{\rho_{n-1}}{1+\rho_{n-1}}\right)^{\alpha-1}.$$

Since $\rho_n \longrightarrow \infty$,

$$\left(\frac{\rho_{n-1}}{1+\rho_{n-1}}\right)^{\alpha-1} \longrightarrow 1,$$

and, noting that

$$\frac{\sigma_n}{\tau_n} = \frac{1}{1 - \frac{1}{\rho_{n-1}}},$$

$\sigma_n \sim \tau_n$, so that

$$\frac{\rho_n}{\rho_{n-1}^\alpha} = \frac{\sigma_{n+1}\tau_{n-1}^\alpha}{\tau_n \sigma_n^\alpha} \sim \frac{\sigma_{n+1}\sigma_{n-1}^\alpha}{\sigma_n^{\alpha+1}}.$$

Then since

$$1 < \limsup_{n\to\infty} \frac{\sigma_{n+1}\sigma_{n-1}^\alpha}{\sigma_n^{\alpha+1}} = \limsup_{n\to\infty} \frac{\rho_n}{\rho_{n-1}^\alpha}\left(\frac{\rho_{n-1}}{1+\rho_{n-1}}\right)^{\alpha-1} = \limsup_{n\to\infty} \frac{\frac{\sigma_{n+1}}{\tau_n^\alpha}}{\frac{\sigma_n}{\tau_{n-1}^\alpha}},$$

it follows that

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_n^\alpha} = \infty,$$

and therefore, $\mathbb{P}(\mathcal{M}) = 0$ by the ratio test and Lemma 6.1. $\square$

Appendix D

Supplements to Chapter 7

In this section, for self-containedness, we outline the idea of the proof of no dominance in absence of feedback for two bins, which can be read in [48, Theorem 1.1] in full detail. The two bins scenario is essentially univariate, since the vector $\Theta_n = (\Theta_n^{(1)}, 1 - \Theta_n^{(1)})$. Thus by symmetry it is possible to focus only on $\Theta_n^{(1)} := T_n^{(1)}/\tau_n$ for all arguments, and therefore we denote $\Theta_n^{(1)}$ as $\Theta_n := T_n/\tau_n$, just for this chapter. For two bins the event of dominance, that there is one of the two proportions converging to 1, can equivalently be stated as the event that there is one of the two proportions converging to 0.

Theorem D.1. *Let $\alpha = 1$. Then Θ_n converges almost surely to a random variable Θ and $\mathbb{P}(\mathcal{D}) = 0$.*

Idea of the proof. By the first part of the argument of Theorem 1.10 applied to $d = 2$, we have already seen that the almost sure convergence holds by the martingale convergence theorem, so we only need to show the idea of the proof behind the second part of the argument. By symmetry, it suffices to show that $\mathbb{P}(\Theta = 0) = 0$.

Denote by $f_n(\lambda) := \mathbb{E}e^{-\lambda\Theta_n}$ and $f(\lambda) := \mathbb{E}e^{-\lambda\Theta}$, for $\lambda \in \mathbb{R}$. Since $f(\lambda) \geq \mathbb{E}\left[e^{-\lambda\Theta}\mathbb{1}_{\{\Theta=0\}}\right] = \mathbb{P}(\Theta = 0)$ for all λ, the claim will follow by showing that there is a sequence $\{\lambda_m\}_{m\in\mathbb{N}}$ such that

$$\lim_{m\longrightarrow\infty} f(\lambda_m) = 0.$$

Let $c \in (0, 1)$ be such that

$$e^{-x} \leq 1 - x + \frac{x^2}{2}$$

for all $x \in [0, c]$, and define $\lambda_m = c\tau_m$. The bulk of the technical work is showing by induction on k, by relying on the monotonicity of all functions $f_{n-k}(\lambda)$, that

$$f_n(\lambda_m) \leq f_{n-k}\left(\lambda_m - \lambda_m^2 \sum_{i=n-k+1}^{n} \frac{\sigma_i}{\tau_i^2}\right)$$

for all m, $n > m$ and $1 \leq k \leq n - m$. Substituting $k = n - m$ in the inductive upper bound obtained, and exploiting the monotonicity of $f_m(\lambda)$, yields

$$f_n(\lambda_m) \leq f_m\left(\lambda_m - \lambda_m^2 \sum_{i=m+1}^{n} \frac{\sigma_i}{\tau_i^2}\right) \leq f_m(\lambda_m(1-c)) = \mathbb{E}e^{-c(1-c)\tau_m\Theta_m} = \mathbb{E}e^{-c(1-c)T_m}$$

for all m and $n > m$. The second inequality follows since, noting that

$$\sum_{i=n-k+1}^{n} \frac{\sigma_i}{\tau_i^2} = \sum_{i=n-k+1}^{n} \frac{\tau_i - \tau_{i-1}}{\tau_i^2} \leq \int_{\tau_{n-k}}^{\tau_n} \frac{dx}{x^2}$$

and that $\tau_m \le \tau_{n-k}$ and $c < 1$, it holds that

$$\lambda_m - \lambda_m^2 \sum_{i=m+1}^{n} \frac{\sigma_i}{\tau_i^2} \ge \lambda_m \left(1 - c\tau_m \int_{\tau_m}^{\infty} \frac{dx}{x^2}\right) = \lambda_m(1-c).$$

By the *Dominated Convergence Theorem*, knowing that $f_n(\lambda) \longrightarrow f(\lambda)$ as $n \longrightarrow \infty$ for all $\lambda > 0$, we can take limits in the inequality just derived, yielding $f(\lambda_m) \le \mathbb{E}e^{-c(1-c)T_m}$ for all m. As by Theorem 1.14 it is known that $\mathbb{P}(\mathcal{M}) = 0$, almost surely none of the bins eventually stops receiving balls, that is, almost surely $T_m \longrightarrow \infty$. By the *Dominated Convergence Theorem* again, we have that $\mathbb{E}e^{-c(1-c)T_m} \longrightarrow 0$ as $m \longrightarrow \infty$, which shows the fact that $f(\lambda_m) \longrightarrow 0$ as $m \longrightarrow \infty$. Since λ is arbitrary in the inequalities $0 \le \mathbb{P}(\Theta = 0) \le f(\lambda)$ previously established, we have that this last fact implies that $\mathbb{P}(\Theta = 0)$. □

Appendix E

Supplements to Chapter 8

In this section, for self-containedness, we include the proof of three useful lemmas that are adaptations of [48, Lemma 6.4, Proposition 7.1, Proposition 6.3] respectively. Recall that for the BB model: α is the feedback parameter, σ_n is the integer-valued time-dependent number of balls thrown at the bins, $\tau_n = \tau_0 + \sum_{i=1}^n \sigma_i$, $\rho_n := \sigma_{n+1}/\tau_n$ and the normalised fluctuations are defined as

$$\varepsilon_{n+1}^{(i)} := \frac{B_{n+1}^{(i)} - \sigma_{n+1} P_n^{(i)}}{\sqrt{\sigma_{n+1} P_n^{(i)} (1 - P_n^{(i)})}}.$$

Lemma E.1. *Let $\{\xi_n\}$ be a sequence of random variables adapted to the filtration $\{\mathscr{F}_n\}$. Suppose $|\xi_i| \le \zeta_n a_i$ for all $i \ge n$, for all $n \in \mathbb{N}$ almost surely, where $\{\zeta_n\}$ is an almost surely positive and square-integrable $\{\mathscr{F}_n\}$-adapted sequence, and $\{a_n\}$ is a deterministic square-summable sequence. Then for all $j \in \{1, \dots, d\}$, almost surely,*

$$\sum_{i=0}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} < \infty$$

and

$$\liminf_{n \to \infty} \frac{1}{\zeta_n \sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} < \infty,$$

where $A_n = \sum_{i=n}^{\infty} a_i^2$.

Proof. Let $n \in \mathbb{N}_0$. For all $m \ge n$ denote

$$S_m^n := \sum_{i=n}^{m-1} \xi_i \varepsilon_{i+1}^{(j)},$$

where the empty sum convention is adopted (that is $S_n^n = 0$). Then $S^n = \{S_m^n\}_m$ is a martingale with respect to $\{\mathscr{F}_m\}$. Indeed, since $\{\xi_i\}$ are adapted and $\{\varepsilon_{i+1}^{(j)}\}$ are centred conditionally on this filtration, as well as adapted to it, S_m^n is adapted and

$$\mathbb{E}_{\mathscr{F}_{m-1}} S_m^n = \sum_{i=n}^{m-1} \mathbb{E}_{\mathscr{F}_{m-1}} (\xi_i \varepsilon_{i+1}^{(j)}) = \sum_{i=n}^{m-2} \xi_i \varepsilon_{i+1}^{(j)} + \xi_{m-1} \mathbb{E}_{\mathscr{F}_{m-1}} \varepsilon_m^{(j)} = \sum_{i=n}^{m-2} \xi_i \varepsilon_{i+1}^{(j)} = S_{m-1}^n.$$

Since boundedness in L^2 holds, thanks to the martingale property of orthogonality of the increments

$$\sum_{m=n}^{\infty} \mathbb{E}(S_{m+1}^n - S_m^n)^2 = \sum_{m=n}^{\infty} \mathbb{E}\xi_m^2 (\varepsilon_{m+1}^{(j)})^2$$

is finite, which follows from the two square-summability and square-integrability hypotheses made on $\{\xi_i\}$ and $\{\zeta_n\}$ respectively:

$$\sum_{m=n}^{\infty} \mathbb{E}\xi_m^2(\varepsilon_{m+1}^{(j)})^2 \leq \sum_{m=n}^{\infty} a_m^2 \mathbb{E}\zeta_n^2(\varepsilon_{m+1}^{(j)})^2 = \sum_{m=n}^{\infty} a_m^2 \mathbb{E}\zeta_n^2 \mathbb{E}_{\mathscr{F}_m}(\varepsilon_{m+1}^{(j)})^2 = \mathbb{E}\zeta_n^2 \sum_{m=n}^{\infty} a_m^2 < \infty,$$

where it has been made use of the null conditional mean and unitary conditional variance of $\varepsilon_{m+1}^{(j)}$, on the top of the random variables ζ_n being adapted. This implies the almost sure convergence to a finite limit of the martingale S_m^n, as $m \longrightarrow \infty$, by the standard theory. Hence

$$\sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} < \infty,$$

which implies the first claim that

$$\sum_{i=0}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} < \infty$$

almost surely.

In order to show that almost surely

$$\liminf_{n\to\infty} \frac{1}{\zeta_n\sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} < \infty,$$

we equivalently prove that

$$\mathbb{P}\left(\liminf_{n\to\infty} \frac{1}{\zeta_n\sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} = \infty\right) = 0.$$

Having defined $E_n^k := \left\{\frac{1}{\zeta_n\sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i\varepsilon_{i+1} > k\right\}$, rewrite the event in the following way:

$$\left\{\liminf_{n\to\infty} \frac{1}{\zeta_n\sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} = \infty\right\} = \bigcap_{k=1}^{\infty} \left\{\liminf_{n\to\infty} \frac{1}{\zeta_n\sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} > k\right\} =$$
$$\bigcap_{k=1}^{\infty} \left\{\exists N \in \mathbb{N} : \forall n \geq N,\ \frac{1}{\zeta_n\sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} > k\right\} =$$
$$\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \left\{\forall n \geq N\ \frac{1}{\zeta_n\sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} > k\right\} = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} \left\{\frac{1}{\zeta_n\sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} > k\right\}$$
$$= \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} E_n^k = \bigcap_{k=1}^{\infty} H_k,$$

where $H_k = \bigcup_{N=1}^{\infty} G_N$, with $G_N = \bigcap_{n\geq N} E_n^k$. In order to prove that $\mathbb{P}\left(\bigcap_{k=1}^{\infty} H_k\right) = 0$, it is enough to note first that $H_k \supseteq H_{k+1}$, because trivially $E_n^k \supseteq E_n^{k+1}$, and therefore $\lim_{k\to\infty} H_k = \bigcap_{k=1}^{\infty} H_k$ is well defined, and by the monotonicity of the probability measure

$$\mathbb{P}\left(\bigcap_{k=1}^{\infty} H_k\right) = \lim_{k\to\infty} \mathbb{P}(H_k).$$

By a similar reasoning, note also that $\mathbb{P}(H_k) = \mathbb{P}\left(\bigcup_{N=1}^{\infty} G_N\right)$, and by the definition of G_N, $G_N \subseteq G_{N+1}$, implying that $\lim_{N\to\infty} G_N = \bigcup_{N=1}^{\infty} G_N$ and therefore, by monotonicity of the probability measure, $\mathbb{P}\left(\bigcup_{N=1}^{\infty} G_N\right) = \lim_{N\to\infty} \mathbb{P}(G_N)$. All in all, we have that

$$\mathbb{P}\left(\liminf_{n\to\infty} \frac{1}{\zeta_n\sqrt{A_n}} \sum_{i=n}^{\infty} \xi_i \varepsilon_{i+1}^{(j)} = \infty\right) = \lim_{k\to\infty} \lim_{N\to\infty} \mathbb{P}(G_N),$$

thus all there is left to prove is that $\lim_{k\to\infty} \lim_{N\to\infty} \mathbb{P}(G_N) = 0$. Proceed with estimating $\mathbb{P}(G_N)$ from above. Since trivially $G_N \subseteq E_N^k$, raising to the square yields, conditionally on $\mathscr{F}_N$, that

$$\begin{aligned}\mathbb{P}(G_N) \le \mathbb{P}(E_N^k) &\le \mathbb{P}\left(\frac{1}{\zeta_N^2 A_N}\left(\sum_{i=N}^{\infty} \xi_i \varepsilon_{i+1}^{(j)}\right)^2 > k^2\right) \\ &= \mathbb{E}\mathbb{P}_{\mathscr{F}_N}\left(\frac{1}{\zeta_N^2 A_N}\left(\sum_{i=N}^{\infty} \xi_i \varepsilon_{i+1}^{(j)}\right)^2 > k^2\right),\end{aligned}$$

with $\zeta_N^2 \in \mathrm{m}\mathscr{F}_N$ and A_N deterministic, allowing to apply *Markov's inequality* conditionally on $\mathscr{F}_N$ as follows:

$$\mathbb{P}_{\mathscr{F}_N}\left(\frac{1}{\zeta_N^2 A_N}\left(\sum_{i=N}^{\infty} \xi_i \varepsilon_{i+1}^{(j)}\right)^2 > k^2\right) = \mathbb{P}_{\mathscr{F}_N}\left(\left(\sum_{i=N}^{\infty} \xi_i \varepsilon_{i+1}^{(j)}\right)^2 > k^2 A_N \zeta_N^2\right) \le$$
$$\frac{\mathbb{E}_{\mathscr{F}_N}\left(\sum_{i=N}^{\infty} \xi_i \varepsilon_{i+1}^{(j)}\right)^2}{k^2 A_N \zeta_N^2} \le \frac{1}{k^2}.$$

The last step follows from the very same calculation done earlier, where we showed $\mathscr{L}^2$-boundedness of S^n, but conditionally on $\mathscr{F}_N$ this time. This shows that S_m^N is a martingale bounded in $\mathscr{L}^2(\mathscr{F}_N)$. The orthogonality of the increments implies the same upper bound for all $m > N$:

$$\mathbb{E}_{\mathscr{F}_N}\left(\sum_{i=N}^{m-1} \xi_i \varepsilon_{i+1}^{(j)}\right)^2 = \mathbb{E}_{\mathscr{F}_N}(S_m^N)^2 = \mathbb{E}_{\mathscr{F}_N}(S_N^N)^2 + \sum_{i=N}^{m-1} \mathbb{E}_{\mathscr{F}_N}(S_{i+1}^N - S_i^N)^2 =$$
$$\sum_{i=N}^{m-1} \mathbb{E}_{\mathscr{F}_N} \xi_i^2 (\varepsilon_{i+1}^{(j)})^2 \le \zeta_N^2 \sum_{i=N}^{m-1} a_i^2 \le \zeta_N^2 A_N.$$

Hence

$$\mathbb{E}_{\mathscr{F}_N}\left(\sum_{i=N}^{\infty} \xi_i \varepsilon_{i+1}^{(j)}\right)^2 \le \sup_{m\ge N} \mathbb{E}_{\mathscr{F}_N}(S_m^N)^2 \le \zeta_N^2 A_N,$$

and the result follows from taking expectation. Since $\mathbb{P}(G_N) \le {}^1\!/_{k^2}$, and $\lim_{k\to\infty} \lim_{N\to\infty} \mathbb{P}(G_N) \le \lim_{k\to\infty} {}^1\!/_{k^2} = 0$, the conclusion of the argument is reached. □

Lemma E.2. *Let $i \in [d]$. If ρ_n is bounded, on the event $\{\Theta_n^{(i)} \longrightarrow 0\}$, $T_n^{(i)}$ is almost surely bounded.*

Proof. The aim is to prove, by contradiction, that $T_k^{(i)}(\omega)$ is bounded for almost all

$\omega \in \left\{\Theta_k^{(i)} \longrightarrow 0\right\}$. Assume by contradiction that on $\Theta_k^{(i)} \longrightarrow 0$, $T_k^{(i)} \longrightarrow \infty$. Since

$$T_i^{(i)} = T_{j-1}^{(i)} + B_j^{(i)} = T_{j-1}^{(i)} + \sigma_j P_{j-1}^{(i)} + \left(B_j^{(i)} - \sigma_j P_{j-1}^{(i)}\right) = T_{j-1}^{(i)} + \sigma_j P_{j-1}^{(i)} + \varepsilon_j^{(i)} \sqrt{\sigma_j P_{j-1}^{(i)} \left(1 - P_{j-1}^{(i)}\right)} \leq T_{j-1}^{(i)} + d^{\alpha-1} \sigma_j \left(\Theta_{j-1}^{(i)}\right)^\alpha + \varepsilon_j^{(i)} \sqrt{\sigma_j P_{j-1}^{(i)} \left(1 - P_{j-1}^{(i)}\right)} = T_{j-1}^{(i)} + d^{\alpha-1} \sigma_j \left(\frac{T_{j-1}^{(i)}}{\tau_{j-1}}\right)^\alpha + \varepsilon_j^{(i)} \sqrt{\sigma_j P_{j-1}^{(i)} \left(1 - P_{j-1}^{(i)}\right)}$$

by the bound in (8.1),

$$\frac{T_j^{(i)} - T_{j-1}^{(i)}}{\left(T_{j-1}^{(i)}\right)^\alpha} \leq d^{\alpha-1} \frac{\sigma_j}{\tau_{j-1}^\alpha} + \varepsilon_j^{(i)} \frac{\sqrt{\sigma_j P_{j-1}^{(i)} \left(1 - P_{j-1}^{(i)}\right)}}{\left(T_{j-1}^{(i)}\right)^\alpha} = d^{\alpha-1} \frac{\sigma_j}{\tau_{j-1}^\alpha} + \varepsilon_j^{(i)} \xi_{j-1}^{(i)}, \tag{E.1}$$

having defined

$$\xi_{j-1}^{(i)} = \frac{\sqrt{\sigma_j P_{j-1}^{(i)} \left(1 - P_{j-1}^{(i)}\right)}}{\left(T_{j-1}^{(i)}\right)^\alpha}.$$

Summing (E.1) from $n+1$ to infinity yields the following upper bound:

$$\sum_{j=n+1}^{\infty} \frac{T_j^{(i)} - T_{j-1}^{(i)}}{\left(T_{j-1}^{(i)}\right)^\alpha} \leq d^{\alpha-1} \sum_{j=n+1}^{\infty} \frac{\sigma_j}{\tau_{j-1}^\alpha} + \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)}. \tag{E.2}$$

We also find a lower bound for the same summation, since as $T_k^{(i)} \longrightarrow \infty$, $\left(T_{j-1}^{(i)}\right)^{-\alpha}$ is decreasing, and in particular vanishing. Therefore, arguing as in the *integral test* for series, we have that

$$\sum_{j=n+1}^{\infty} \frac{T_j^{(i)} - T_{j-1}^{(i)}}{\left(T_{j-1}^{(i)}\right)^\alpha} \geq \int_{T_n^{(i)}}^{\infty} \frac{dx}{x^\alpha} = \frac{1}{(\alpha - 1)\left(T_n^{(i)}\right)^{\alpha-1}}. \tag{E.3}$$

In conclusion putting (E.2) and (E.3) together yields

$$\frac{1}{(\alpha - 1)\left(T_n^{(i)}\right)^{\alpha-1}} \leq d^{\alpha-1} \sum_{j=n+1}^{\infty} \frac{\sigma_j}{\tau_{j-1}^\alpha} + \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)}. \tag{E.4}$$

A bound can similarly be found for the first summation involved, as ρ_{j-1} is bounded by a constant ρ:

$$\sum_{j=n+1}^{\infty} \frac{\sigma_j}{\tau_{j-1}^\alpha} = \sum_{j=n+1}^{\infty} \frac{(1+\rho_{j-1})^\alpha \sigma_j}{\tau_j^\alpha} \leq (1+\rho)^\alpha \sum_{j=n+1}^{\infty} \frac{\sigma_j}{\tau_j^\alpha} = (1+\rho)^\alpha \sum_{j=n+1}^{\infty} \frac{\tau_j - \tau_{j-1}}{\tau_j^\alpha} \leq (1+\rho)^\alpha \int_{\tau_n}^{\infty} \frac{dx}{x^\alpha} = \frac{(1+\rho)^\alpha}{(\alpha-1)\tau_n^{\alpha-1}}.$$

Plugging

$$\sum_{j=n+1}^{\infty} \frac{\sigma_j}{\tau_{j-1}^\alpha} \leq \frac{(1+\rho)^\alpha}{(\alpha-1)\tau_n^{\alpha-1}} \tag{E.5}$$

in (E.4) yields

$$\frac{1}{(\alpha - 1)\left(T_n^{(i)}\right)^{\alpha-1}} \leq \frac{d^{\alpha-1}(1+\rho)^\alpha}{(\alpha-1)\tau_n^{\alpha-1}} + \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)} = \frac{\gamma}{(\alpha-1)\tau_n^{\alpha-1}} + \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)},$$

where $\gamma = d^{\alpha-1}(1+\rho)^{\alpha}$. Rearranging it, yields

$$\frac{1}{\left(\Theta_n^{(i)}\right)^{\alpha-1}} = \left(\frac{\tau_n}{T_n^{(i)}}\right)^{\alpha-1} \leq \gamma + (\alpha-1)\tau_n^{\alpha-1} \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)},$$

thus

$$\frac{1}{\left(\Theta_n^{(i)}\right)^{\alpha-1}} - \gamma \leq (\alpha-1)\tau_n^{\alpha-1} \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)}. \tag{E.6}$$

It is now possible to reach a contradiction by invoking Lemma E.1. Consider that by definition of $\xi_j^{(i)}$, (8.1) and $T_j^{(i)}$ being nondecreasing, we have that

$$\xi_{j-1}^{(i)} = \frac{\sqrt{\sigma_j P_{j-1}^{(i)}\left(1-P_{j-1}^{(i)}\right)}}{\left(T_{j-1}^{(i)}\right)^{\alpha}} \leq \frac{1}{\left(T_{j-1}^{(i)}\right)^{\frac{\alpha}{2}}} \sqrt{\frac{\sigma_j d^{\alpha-1}\left(\Theta_{j-1}^{(i)}\right)^{\alpha}}{\left(T_{j-1}^{(i)}\right)^{\alpha}}} \leq \frac{1}{\left(T_n^{(i)}\right)^{\frac{\alpha}{2}}} \sqrt{\frac{d^{\alpha-1}\sigma_j}{\tau_{j-1}^{\alpha}}}.$$

Define

$$\zeta_n := \frac{1}{\left(T_n^{(i)}\right)^{\frac{\alpha}{2}}} \in \mathscr{L}_2$$
$$a_j := \sqrt{\frac{d^{\alpha-1}\sigma_j}{\tau_{j-1}^{\alpha}}} \in l_2,$$

where $\{\zeta_n\}$ is trivially adapted, since $\zeta_n \in \mathrm{m}\mathscr{F}_n$, and square integrable on Ω, as $0 < \zeta_n < 1$. The square-summability of $\{a_j\}$ follows trivially from (E.5). Then

$$|\xi_{j-1}| = \xi_{j-1} \leq \zeta_n a_j.$$

Define also

$$A_n := \sum_{j=n+1}^{\infty} a_j^2 = d^{\alpha-1} \sum_{j=n+1}^{\infty} \frac{\sigma_j}{\tau_{j-1}^{\alpha}} \leq \frac{\gamma}{(\alpha-1)\tau_n^{\alpha-1}}$$

and further estimate

$$\frac{1}{\left(\Theta_n^{(i)}\right)^{\alpha-1}} - \gamma \leq (\alpha-1)\tau_n^{\alpha-1} \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)} \leq (\alpha-1)\tau_n^{\alpha-1} \frac{\sqrt{\frac{\gamma}{(\alpha-1)\tau_n^{\alpha-1}}}}{\left(T_n^{(i)}\right)^{\frac{\alpha}{2}} \zeta_n \sqrt{A_n}} \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)}$$
$$= \sqrt{\frac{(\alpha-1)\gamma}{\tau_n}} \left(\frac{\tau_n}{T_n^{(i)}}\right)^{\frac{\alpha}{2}} \frac{1}{\zeta_n\sqrt{A_n}} \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)} = \sqrt{\frac{(\alpha-1)\gamma}{\tau_n\left(\Theta_n^{(i)}\right)^{\alpha}}} \frac{1}{\zeta_n\sqrt{A_n}} \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)}.$$

Thus

$$\left(\frac{1}{\left(\Theta_n^{(i)}\right)^{\alpha-1}} - \gamma\right) \sqrt{\tau_n\left(\Theta_n^{(i)}\right)^{\alpha}} \leq \frac{\sqrt{(\alpha-1)\gamma}}{\zeta_n\sqrt{A_n}} \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)},$$

Lemma E.1 ensures that

$$\liminf_{n\longrightarrow\infty} \frac{1}{\zeta_n\sqrt{A_n}} \sum_{j=n+1}^{\infty} \varepsilon_j^{(i)} \xi_{j-1}^{(i)} < \infty,$$

and therefore

$$\liminf_{n\longrightarrow\infty} \left(\frac{1}{\left(\Theta_n^{(i)}\right)^{\alpha-1}} - \gamma\right) \sqrt{\tau_n\left(\Theta_n^{(i)}\right)^{\alpha}} < \infty.$$

Since
$$\left(\frac{1}{\left(\Theta_n^{(i)}\right)^{\alpha-1}}-\gamma\right)\sqrt{\tau_n\left(\Theta_n^{(i)}\right)^{\alpha}}=\left(\frac{1}{\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}-1}}-\gamma\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}}\right)\sqrt{\tau_n}$$
and
$$\frac{1}{\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}-1}}-\gamma\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}}\sim\frac{1}{\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}-1}}$$
as $\Theta_n^{(i)}\longrightarrow 0$ by hypothesis, and γ is a constant, it follows that almost surely
$$\infty>\liminf_{n\longrightarrow\infty}\left(\frac{1}{\left(\Theta_n^{(i)}\right)^{\alpha-1}}-\gamma\right)\sqrt{\tau_n\left(\Theta_n^{(i)}\right)^{\alpha}}=\liminf_{n\longrightarrow\infty}\left(\frac{1}{\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}-1}}-\gamma\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}}\right)\sqrt{\tau_n}$$
$$=\liminf_{n\longrightarrow\infty}\frac{\sqrt{\tau_n}}{\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}-1}}.$$
The contradiction follows from
$$\liminf_{n\longrightarrow\infty}\frac{\sqrt{\tau_n}}{\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}-1}}<\infty.$$
We have indeed two cases.

Case 1. If $\alpha\geq 2$, having $(\Theta_n^{(i)})^{\frac{\alpha}{2}-1}\leq 1$, it trivially follows that
$$\infty>\liminf_{n\longrightarrow\infty}\frac{\sqrt{\tau_n}}{\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}-1}}\geq\liminf_{n\longrightarrow\infty}\sqrt{\tau_n}=\infty,$$
which is a contradiction.

Case 2. If $1<\alpha<2$, then $0<(2-\alpha)/2<1/2$ and $(\alpha-1)/(2-\alpha)>0$. Note that
$$\frac{\sqrt{\tau_n}}{\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}-1}}=\left(T_n^{(i)}\right)^{\frac{2-\alpha}{2}}\tau_n^{\frac{\alpha-1}{2}}=\left(T_n^{(i)}\tau_n^{\frac{\alpha-1}{2-\alpha}}\right)^{\frac{2-\alpha}{2}}.$$
But since $T_n^{(i)}\longrightarrow\infty$,
$$\left(T_n^{(i)}\tau_n^{\frac{\alpha-1}{2-\alpha}}\right)^{\frac{2-\alpha}{2}}\longrightarrow\infty,$$
and it follows that
$$\infty>\liminf_{n\longrightarrow\infty}\frac{\sqrt{\tau_n}}{\left(\Theta_n^{(i)}\right)^{\frac{\alpha}{2}-1}}=\liminf_{n\longrightarrow\infty}\left(T_n^{(i)}\tau_n^{\frac{\alpha-1}{2-\alpha}}\right)^{\frac{2-\alpha}{2}}=\infty,$$
which is a contradiction.

Hence $T_k^{(i)}$ must be almost surely bounded on the event $\Theta_k^{(i)}\longrightarrow 0$. ☐

Lemma E.3. $\mathbb{P}(\varepsilon_{n+1}^{(i)}\leq n,\ \text{ev.})=1$.

Proof. Equivalently, we can prove that $\mathbb{P}(\varepsilon_{n+1}^{(i)}>n,\ \text{i.o.})=0$. Let $E_{n+1}:=\left\{\varepsilon_{n+1}^{(i)}>n\right\}\in\mathscr{F}_{n+1}$. By *Lévy's extension of Borel-Cantelli Lemma*, if
$$\sum_{n=0}^{\infty}\mathbb{P}_{\mathscr{F}_n}(E_{n+1})<\infty,$$

then

$$\sum_{n=0}^{\infty} \mathbb{1}_{E_{n+1}} < \infty,$$

that is E_{n+1} does not occur infinitely often. *Vice versa*, if

$$\sum_{n=0}^{\infty} \mathbb{P}_{\mathscr{F}_n}(E_{n+1}) = \infty,$$

then

$$\sum_{n=0}^{\infty} \mathbb{1}_{E_{n+1}} \sim \sum_{n=0}^{\infty} \mathbb{P}_{\mathscr{F}_n}(E_{n+1}),$$

that is E_{n+1} does not occur infinitely often, as

$$\sum_{n=0}^{\infty} \mathbb{1}_{E_{n+1}}$$

must diverge. Summing up,

$$\{E_{n+1},\ \text{i.o.}\} = \left\{ \sum_{n=0}^{\infty} \mathbb{P}_{\mathscr{F}_n}(E_{n+1}) = \infty \right\}.$$

Since by the *conditional Markov's inequality*

$$\mathbb{P}_{\mathscr{F}_n}(E_{n+1}) \leq \mathbb{P}_{\mathscr{F}_n}(|\varepsilon^{(i)}_{n+1}| > n) = \mathbb{P}_{\mathscr{F}_n}((\varepsilon^{(i)}_{n+1})^2 > n^2) \leq \frac{\mathbb{E}_{\mathscr{F}_n}(\varepsilon^{(i)}_{n+1})^2}{n^2} = \frac{1}{n^2},$$

the series converges almost surely, and the claim follows. $\square$

Appendix F

Supplements to Chapter 9

In this chapter, for self-containedness, we include the proof of an adaptation of [48, 7.2]. Recall that for the BB model α is the feedback parameter, σ_n is the integer-valued time-dependent number of balls thrown at the bins, $\tau_n = \tau_0 + \sum_{i=1}^n \sigma_i$, $\rho_n := {}^{\sigma_{n+1}}/_{\tau_n}$, $\theta \longleftarrow \theta_n := \alpha^{-n}\log\tau_n$, $\lambda := \limsup_{n\to\infty}\sigma_{n+1}\sigma_{n-1}^{\alpha}\sigma_n^{-\alpha-1}$ and the normalised fluctuations are defined as

$$\varepsilon_{n+1}^{(i)} := \frac{B_{n+1}^{(i)} - \sigma_{n+1}P_n^{(i)}}{\sqrt{\sigma_{n+1}P_n^{(i)}(1-P_n^{(i)})}}.$$

Lemma F.1. *Assume $\rho_n \longrightarrow \infty$, $\lambda < 1$ and that there exists $\lim_{n\to\infty}\theta_n =: \theta = 0$. Let $i \in [d]$. Then on the event $\{\Theta_n^{(i)} \longrightarrow 0\}$, $T_n^{(i)}$ is almost surely bounded.*

Proof. The argument will show that the event

$$\mathscr{E} = \{\Theta_n^{(i)} \longrightarrow 0\} \cap \{T_n^{(i)} \longrightarrow \infty\}$$

almost never occurs. In order to do this a number of fixed parameters will be needed: ε, q and δ. The conditions they will have to satisfy will be specified in due time. Recall that by (8.1) the following iterative upper bound holds:

$$T_{n+1}^{(i)} \le T_n^{(i)} + d^{\alpha-1}\sigma_{n+1}\frac{(T_n^{(i)})^\alpha}{\tau_n^\alpha} + \varepsilon_{n+1}^{(i)}\sqrt{\sigma_{n+1}P_n^{(i)}(1-P_n^{(i)})}. \tag{F.1}$$

Define $k_0 = 0$ and $\delta > 0$ small enough, to satisfy all conditions, which will be imposed in due time. For every $n \in \mathbb{N}$ define a sequence of stopping times

$$k_n = \inf\left\{ j > k_{n-1} : \frac{\sigma_{j+1}}{\tau_j^\alpha}(T_j^{(i)})^\alpha \le \delta T_j^{(i)} \right\},$$

that is the successive times, at which the second term of the iteration is $d^{\alpha-1}\delta$-smaller than the first. Let $\varepsilon > 0$ be a small parameter, the conditions upon which will be clear later in *Step 3*. Similarly, δ is assumed small enough, so that $d^{\alpha-1}\delta < \varepsilon$ (first condition on δ).

Step 1. In this first step we prove that on $\mathscr{E}$ the $d^{\alpha-1}\delta$-negligibility of the second term of the iteration occurs infinitely often, that is, for all n, k_n is almost surely finite. Proceeding by contradiction, suppose that it is not true. Then a random variable $\bar{n}$ expressing the last time at which the $d^{\alpha-1}\delta$-negligibility happened can be defined, which is finite on $\mathscr{E}$ with positive probability:

$$\bar{n} = \sup\left\{ j \in \mathbb{N} : \frac{\sigma_{j+1}}{\tau_j^\alpha}(T_j^{(i)})^\alpha \le \delta T_j^{(i)} \right\}.$$

This means that with positive probability eventually on $\mathscr{E}$ the second term of the iteration plays the main role (ignoring for the moment the contributions of the third term). Using the hypothesis that $\theta = 0$, this will yield a contradiction with $T_n^{(i)} \geq 1$. The argument is as follows. Since the intersection $\{\bar{n} < \infty\} \cap \{\mathscr{E}\}$ has already been shown to be an event having positive probability, consider that for all ω in this event, for all $n \geq \bar{n}$,

$$\frac{\sigma_{n+1}}{\tau_n^\alpha}(T_n^{(i)})^\alpha > \delta T_n^{(i)}.$$

Plugging this into the first term of (F.1) yields

$$T_{n+1}^{(i)} \leq \frac{\sigma_{n+1}}{\delta \tau_n^\alpha}(T_n^{(i)})^\alpha + d^{\alpha-1}\sigma_{n+1}\frac{(T_n^{(i)})^\alpha}{\tau_n^\alpha} + \varepsilon_{n+1}^{(i)}\sqrt{\sigma_{n+1}P_n^{(i)}(1-P_n^{(i)})} =$$
$$c\frac{\sigma_{n+1}}{\tau_n^\alpha}(T_n^{(i)})^\alpha + \varepsilon_{n+1}^{(i)}\sqrt{\sigma_{n+1}P_n^{(i)}(1-P_n^{(i)})} = c\frac{\sigma_{n+1}}{\tau_n^\alpha}(T_n^{(i)})^\alpha\left(1+\hat{\xi}_n\varepsilon_{n+1}^{(i)}\right),$$

having defined $c := d^{\alpha-1} + 1/\delta$, and

$$\hat{\xi}_n := \frac{\tau_n^\alpha}{c(T_n^{(i)})^\alpha\sqrt{\sigma_{n+1}}}\sqrt{P_n^{(i)}(1-P_n^{(i)})}.$$

Thus

$$T_n^{(i)} \leq c\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^\alpha\left(1+\hat{\xi}_{n-1}\varepsilon_n^{(i)}\right). \tag{F.2}$$

Iterating (F.2) $n-k-1$ times yields, for all $k > \bar{n}$ large enough and $n > k$, that

$$T_n^{(i)} \leq c^{1+\alpha+\ldots+\alpha^{n-k-1}}(T_k^{(i)})^{\alpha^{n-k}}\frac{\sigma_n}{\tau_{n-1}^\alpha}\left(\frac{\sigma_{n-1}}{\tau_{n-2}^\alpha}\right)^\alpha\cdots\left(\frac{\sigma_{k+1}}{\tau_k^\alpha}\right)^{\alpha^{n-k-1}}\prod_{j=k}^{n-1}\left(1+\hat{\xi}_j\varepsilon_{j+1}^{(i)}\right)^{\alpha^{n-j-1}}$$
$$\leq \left(c^{1+\alpha^{-1}+\alpha^{-2}+\ldots+\alpha^{-(n-k-1)}}\right)^{\alpha^{n-k-1}}\tau_n\left(\frac{T_k^{(i)}}{\tau_k}\right)^{\alpha^{n-k}}\prod_{j=k}^{n-1}\left(1+\hat{\xi}_j\varepsilon_{j+1}^{(i)}\right)^{\alpha^{n-j-1}} \leq$$
$$\left(c^{\sum_{j=0}^\infty \alpha^{-j}}\right)^{\alpha^{n-k-1}}\tau_n\left(\Theta_k^{(i)}\right)^{\alpha^{n-k}}\prod_{j=k}^{n-1}\left(1+\hat{\xi}_j\varepsilon_{j+1}^{(i)}\right)^{\alpha^{n-j-1}} = \left(c^{\frac{\alpha}{\alpha-1}}\right)^{\alpha^{n-k-1}}\tau_n\left(\Theta_k^{(i)}\right)^{\alpha^{n-k}}$$
$$\left(\prod_{j=k}^{n-1}\left(1+\hat{\xi}_j\varepsilon_{j+1}^{(i)}\right)^{\alpha^{k-j-1}}\right)^{\alpha^{n-k}} = \tau_n\left[c^{\frac{1}{\alpha-1}}\Theta_k^{(i)}\exp\left(\sum_{j=k}^{n-1}\alpha^{k-j-1}\log\left(1+\hat{\xi}_j\varepsilon_{j+1}^{(i)}\right)\right)\right]^{\alpha^{n-k}}$$
$$\leq \tau_n\left[c^{\frac{1}{\alpha-1}}\Theta_k^{(i)}\exp\left(\alpha^{k-1}\sum_{j=k}^{n-1}\frac{\hat{\xi}_j}{\alpha^j}\varepsilon_{j+1}^{(i)}\right)\right]^{\alpha^{n-k}} = \tau_n\left[c^{\frac{1}{\alpha-1}}\Theta_k^{(i)}\exp\left(\alpha^{k-1}\sum_{j=k}^{n-1}\xi_j\varepsilon_{j+1}^{(i)}\right)\right]^{\alpha^{n-k}},$$

where $\xi_j = \hat{\xi}_j/\alpha^j$, and the following facts have been used, that $\sigma_j \leq \tau_j$ and $\log(1+x) \leq x$. For the second inequality to hold, $x := \hat{\xi}_j\varepsilon_{j+1}^{(i)}$ needs to satisfy the hypothesis $x > -1$, as the fluctuations $\varepsilon_{j+1}^{(i)}$ can take negative values. However, this is immediately shown by recalling (8.1) and that $\delta > 0$:

$$\hat{\xi}_j\varepsilon_{j+1}^{(i)} = \frac{\tau_j^\alpha}{c(T_j^{(i)})^\alpha\sqrt{\sigma_{j+1}}}\sqrt{P_j^{(i)}(1-P_j^{(i)})}\frac{B_{j+1}^{(i)}-\sigma_{j+1}P_j^{(i)}}{\sqrt{\sigma_{j+1}P_j^{(i)}(1-P_j^{(i)})}} \geq$$
$$-\frac{\tau_j^\alpha}{c(T_j^{(i)})^\alpha}P_j^{(i)} \geq -\frac{d^{\alpha-1}\tau_j^\alpha}{c(T_j^{(i)})^\alpha}(\Theta_j^{(i)})^\alpha = -\frac{d^{\alpha-1}}{c} = -\frac{d^{\alpha-1}}{d^{\alpha-1}+1/\delta} > -1.$$

Thus, for all $\bar{n} < k < n$,

$$T_n^{(i)} \leq \tau_n \left[c^{\frac{1}{\alpha-1}} \Theta_k^{(i)} \exp\left(\alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} \right) \right]^{\alpha^{n-k}},$$

then

$$\log T_n^{(i)} \leq \log \tau_n + \alpha^{n-k} \log \left[c^{\frac{1}{\alpha-1}} \Theta_k^{(i)} \exp\left(\alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} \right) \right],$$

and therefore

$$\alpha^{-n} \log T_n^{(i)} \leq \alpha^{-n} \log \tau_n + \alpha^{-k} \log \left[c^{\frac{1}{\alpha-1}} \Theta_k^{(i)} \exp\left(\alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} \right) \right].$$

Recall that $0 = \theta := \lim_{n\to\infty} \alpha^{-n} \log \tau_n$, hence

$$\limsup_{n\to\infty} \alpha^{-n} \log T_n^{(i)} \leq \limsup_{n\to\infty} \alpha^{-n} \log \tau_n + \alpha^{-k} \log \left[c^{\frac{1}{\alpha-1}} \Theta_k^{(i)} \exp\left(\alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} \right) \right]$$
$$= \alpha^{-k} \log \left[c^{\frac{1}{\alpha-1}} \Theta_k^{(i)} \exp\left(\alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} \right) \right]$$

for all $k > \bar{n}$. However by Lemma E.1 it can be shown that the right-hand side is negative on $\mathscr{E}$ for some k (depending on ω). This would be possible only if $T_n^{(i)} < 1$, a contradiction with $T_n^{(i)} \geq T_0^{(i)} > 1$. More precisely, the argument is as follows. Since on $\mathscr{E}$, $\Theta_k^{(i)} \longrightarrow 0$, it is enough to show that

$$\exp\left(\alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} \right)$$

is infinitely often bounded above (that is, for infinitely many k, there is a finite uniform upper bound). Clearly this amounts to show that

$$\alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)}$$

is infinitely often bounded above. Since, if it diverges to infinity, then its limit inferior will do so too, it is enough to prove that almost surely

$$\liminf_{k\to\infty} \alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} < \infty,$$

in order to achieve the desired boundedness infinitely often. Here is where Lemma E.1 is used. The hypotheses of the lemma are satisfied on the whole of Ω, since for all $j \geq k > \bar{n}$,

$$|\xi_j| = \frac{|\hat{\xi}_j|}{\alpha^j} = \frac{\tau_j^\alpha}{c\alpha^j (T_j^{(i)})^\alpha \sqrt{\sigma_{j+1}}} \sqrt{P_j^{(i)}(1 - P_j^{(i)})} \leq \frac{d^{\frac{\alpha-1}{2}} \tau_j^\alpha}{c\alpha^j (T_j^{(i)})^\alpha \sqrt{\sigma_{j+1}}} \sqrt{\frac{(T_j^{(i)})^\alpha}{\tau_j^\alpha}}$$
$$= \frac{d^{\frac{\alpha-1}{2}}}{c\alpha^j} \sqrt{\frac{(T_j^{(i)})^\alpha}{\sigma_{j+1} \tau_j^\alpha}} < \frac{d^{\frac{\alpha-1}{2}}}{c\alpha^j \sqrt{\delta T_j^{(i)}}} \leq \frac{d^{\frac{\alpha-1}{2}}}{c\alpha^j \sqrt{\delta}}.$$

Define $a_j := \alpha^{-j}$ and

$$\zeta_k := \frac{d^{\frac{\alpha-1}{2}}}{c\sqrt{\delta}} > 0.$$

Note that $\{a_j\}$ are square-summable, since $\alpha^2 > 1$ yields a convergent geometric series. Denote

$$A_k := \sum_{j=k}^{\infty} a_j^2 = \left(\frac{1}{1-\alpha^{-2}} - \frac{1-\alpha^{-2k}}{1-\alpha^{-2}}\right) = \frac{\alpha^{-2k}}{1-\alpha^{-2}} \le \frac{1}{(1-\alpha^{-2})}.$$

Note that $\{\zeta_k\}$ is trivially square-integrable and adapted, since it is constant. Finally, since $|\xi_j| \le \zeta_k a_j$, by Lemma E.1 it holds that

$$\liminf_{k\to\infty} \frac{1}{\zeta_k\sqrt{A_k}} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} < \infty.$$

Since

$$\frac{1}{\zeta_k\sqrt{A_k}} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} = \frac{c\sqrt{\delta}}{d^{\frac{\alpha-1}{2}}} \frac{\sqrt{1-\alpha^{-2}}}{\alpha^{-k}} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} = c\sqrt{\frac{\delta(\alpha^2-1)}{d^{\alpha-1}}} \left(\alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)}\right),$$

factor the positive constant $c\sqrt{\delta(\alpha^2-1)/d^{\alpha-1}}$ out of the $\liminf$, and the claim that

$$\liminf_{k\to\infty} \alpha^{k-1} \sum_{j=k}^{\infty} \xi_j \varepsilon_{j+1}^{(i)} < \infty$$

follows.

Step 2. In this step we provide an upper bound for the random noise terms $\varepsilon_j^{(i)}$, for all $j \ge k_n$, as a function of j and n, which applies almost surely, infinitely often. For all $n \in \mathbb{N}$ let $c_n \in \mathbb{R}$ be a sequence diverging to infinity, and let

$$\mathscr{E}_n := \left(\{k_n < \infty\} \cap \left\{\varepsilon_j^{(i)} \le c_j(j-k_n)\ \forall j > k_n\right\}\right) \cup \{k_n = \infty\},$$

where $c_j(j-k_n)$ is the bound aforementioned. Then we show that

$$\mathbb{P}(\mathscr{E}_n,\ \text{i.o.}) = 1. \tag{F.3}$$

Let $G_n := \bigcup_{k\ge n} \mathscr{E}_k$ and note that $G_{n+1} \subseteq G_n$, then

$$\{\mathscr{E}_n,\ \text{i.o.}\} := \bigcap_{n=1}^{\infty} \bigcup_{k\ge n} \mathscr{E}_k = \bigcap_{n=1}^{\infty} G_n = \lim_{n\to\infty} G_n,$$

and therefore $\mathbb{P}(\mathscr{E}_n,\ \text{i.o.}) = 1$ if and only if

$$\lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\ge n} \mathscr{E}_k\right) = 1.$$

Since $\mathbb{P}\left(\bigcup_{k\ge n} \mathscr{E}_k\right) \ge \mathbb{P}(\mathscr{E}_n)$, it will suffice to prove that

$$\lim_{n\to\infty} \mathbb{P}(\mathscr{E}_n) = 1.$$

Recall that $\mathbb{E}_{\mathscr{F}_{j-1}}\varepsilon_j^{(i)} = 0$, $\mathbb{E}_{\mathscr{F}_{j-1}}(\varepsilon_j^{(i)})^2 = 1$; then for all n large enough, so that $c_j(j - k_n) \geq 0$ and $(c_j(j-k_n))^{-2} < 1$ (recall that $c_j \longrightarrow \infty$), by *Markov's inequality*, the *tower property* of the conditional expectation applied iteratively, the measurability of the stopping time k_n with respect to its own stopped σ-algebra $\mathscr{F}_{k_n}$ and finally by the *monotone convergence theorem*, it follows that

$$\mathbb{P}(\mathscr{E}_n) = \mathbb{P}\left(\{k_n < \infty\} \cap \left\{\varepsilon_j^{(i)} \leq c_j(j-k_n)\ \forall j > k_n\right\}\right) + \mathbb{P}(k_n = \infty) =$$

$$\mathbb{P}\left(\bigcap_{l=1}^{\infty} \left\{\varepsilon_j^{(i)} \leq c_j(j-k_n)\ \forall k_n < j \leq k_n + l\right\} \cap \{k_n < \infty\}\right) + \mathbb{P}(k_n = \infty) =$$

$$\lim_{l\to\infty} \mathbb{P}\left(\left\{\varepsilon_j^{(i)} \leq c_j(j-k_n)\ \forall k_n < j \leq k_n + l\right\} \cap \{k_n < \infty\}\right) + \mathbb{P}(k_n = \infty) =$$

$$\lim_{l\to\infty} \mathbb{E}\left(\mathbb{1}_{\left\{\varepsilon_j^{(i)} \leq c_j(j-k_n)\ \forall k_n < j \leq k_n + l\right\}} \mathbb{1}_{\{k_n<\infty\}}\right) + \mathbb{P}(k_n = \infty) =$$

$$\lim_{l\to\infty} \mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}} \prod_{j=k_n+1}^{k_n+l} \mathbb{1}_{\left\{\varepsilon_j^{(i)} \leq c_j(j-k_n)\right\}}\right) + \mathbb{P}(k_n = \infty) =$$

$$\lim_{l\to\infty} \mathbb{E}\mathbb{E}_{\mathscr{F}_{k_n+l-1}}\left(\mathbb{1}_{\{k_n<\infty\}} \prod_{j=k_n+1}^{k_n+l} \mathbb{1}_{\left\{\varepsilon_j^{(i)} \leq c_j(j-k_n)\right\}}\right) + \mathbb{P}(k_n = \infty) =$$

$$\lim_{l\to\infty} \mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}} \prod_{j=k_n+1}^{k_n+l-1} \mathbb{1}_{\left\{\varepsilon_j^{(i)} \leq c_j(j-k_n)\right\}} \mathbb{E}_{\mathscr{F}_{k_n+l-1}} \mathbb{1}_{\left\{\varepsilon_{k_n+l}^{(i)} \leq c_{k_n+l} l\right\}}\right) + \mathbb{P}(k_n = \infty)$$

$$= \lim_{l\to\infty} \mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}} \prod_{j=k_n+1}^{k_n+l-1} \mathbb{1}_{\left\{\varepsilon_j^{(i)} \leq c_j(j-k_n)\right\}} \mathbb{P}_{\mathscr{F}_{k_n+l-1}}\left(\varepsilon_{k_n+l}^{(i)} \leq c_{k_n+l} l\right)\right) + \mathbb{P}(k_n = \infty)$$

$$\geq \lim_{l\to\infty} \mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}} \prod_{j=k_n+1}^{k_n+l-1} \mathbb{1}_{\left\{\varepsilon_j^{(i)} = c_j(j-k_n)\right\}} \left[1 - \mathbb{P}_{\mathscr{F}_{k_n+l-1}}\left(|\varepsilon_{k_n+l}^{(i)}| > c_{k_n+l} l\right)\right]\right)$$

$$+ \mathbb{P}(k_n = \infty) = \lim_{l\to\infty} \mathbb{E}\Bigg(\mathbb{1}_{\{k_n<\infty\}} \prod_{j=k_n+1}^{k_n+l-1} \mathbb{1}_{\left\{\varepsilon_j^{(i)} = c_j(j-k_n)\right\}} [1 - \mathbb{P}_{\mathscr{F}_{k_n+l-1}}((\varepsilon_{k_n+l}^{(i)})^2$$

$$> c_{k_n+l}^2 l^2)]\Bigg) + \mathbb{P}(k_n = \infty) \geq \lim_{l\to\infty} \mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}} \prod_{j=k_n+1}^{k_n+l-1} \mathbb{1}_{\left\{\varepsilon_j^{(i)} = c_j(j-k_n)\right\}} \left(1 - \frac{1}{c_{k_n+l}^2 l^2}\right)\right)$$

$$+ \mathbb{P}(k_n = \infty) \geq \ldots \geq \lim_{l\to\infty} \mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}} \prod_{j=k_n+1}^{k_n+l} \left(1 - \frac{1}{c_j^2(j-k_n)^2}\right)\right) + \mathbb{P}(k_n = \infty) =$$

$$\lim_{l\to\infty} \mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}} \prod_{j=1}^{l} \left(1 - \frac{1}{c_{k_n+j}^2 j^2}\right)\right) + \mathbb{P}(k_n = \infty) =$$

$$\lim_{l\to\infty} \mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}} \exp \sum_{j=1}^{l} \log\left(1 - \frac{1}{c_{k_n+j}^2 j^2}\right)\right) + \mathbb{P}(k_n = \infty) =$$

$$\mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}} \exp \sum_{j=1}^{\infty} \log\left(1 - \frac{1}{c_{k_n+j}^2 j^2}\right)\right) + \mathbb{P}(k_n = \infty).$$

Consider now that for all $x \in [0, 1/2)$,

$$\log(1-x) \geq -x - x^2,$$

as the two concave functions meet at 0, where their derivatives, respectively $(x-1)^{-1}$ and $-(1+2x)$ take the common value -1. Consider also that for all $0 < x < 1$, the difference of such derivatives $(x-1)^{-1}+(1+2x)>0$ if and only if $0<x<1/2$, meaning that for all x in this interval, the parabola decays faster than the logarithm. Having started at zero, where both of the two functions and the corresponding derivatives coincide, the inequality follows. Since

$$\hat{c}_n := \min_{j\geq n}\{c_j\}$$

diverges to infinity, for n sufficiently large, to satisfy also $\hat{c}_n^{-2} < 1/2$, it holds that for all $x\in[0,\hat{c}_n^{-2}]$, the logarithmic inequality is also satisfied. Note that since $k_n+j\geq n$ for all n and j, $c_{k_n+j}\geq \hat{c}_n$. Hence for all such n large enough, having $1/(\hat{c}_n^2 j^2)\in[0,\hat{c}_n^{-2}]$, it follows that, as $n\longrightarrow\infty$,

$$\begin{aligned}
&\mathbb{P}(\mathscr{E}_n) \geq \mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}}\exp\sum_{j=1}^{\infty}\log\left(1-\frac{1}{\hat{c}_n^2 j^2}\right)\right)+\mathbb{P}(k_n=\infty)\geq \\
&\mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}}\exp-\left\{\sum_{j=1}^{\infty}\frac{1}{\hat{c}_n^2 j^2}+\sum_{j=1}^{\infty}\frac{1}{\hat{c}_n^4 j^4}\right\}\right)+\mathbb{P}(k_n=\infty)= \\
&\exp-\hat{c}_n^{-2}\left\{\sum_{j=1}^{\infty}\frac{1}{j^2}+\hat{c}_n^{-2}\sum_{j=1}^{\infty}\frac{1}{j^4}\right\}\mathbb{E}\left(\mathbb{1}_{\{k_n<\infty\}}\right)+\mathbb{P}(k_n=\infty)= \\
&(1+o(1))\mathbb{P}(k_n<\infty)+\mathbb{P}(k_n=\infty)\longrightarrow 1.
\end{aligned}$$

Step 3. In this step, using the stopping times k_n, at which the second term of (F.1) is $d^{\alpha-1}\delta$-smaller than the first, and combining them with the new bound obtained for the random fluctuations $\varepsilon_n^{(i)}$ in (F.3), a stopping time k_ν will be determined such that, for all $n\geq k_\nu$, the second term of (F.1) will be $d^{\alpha-1}\delta q^{n-k_\nu}$-smaller than the first one. This is a much stronger domination, as q is subunitary (further hypotheses will be needed on q). More precisely, the argument is that on $\mathscr{E}$ there is a random variable ν such that for all $n\geq k_\nu$, it holds that $\varepsilon_n^{(i)}\leq n$ and

$$\frac{\sigma_{n+1}}{\tau_n^{\alpha}}(T_n^{(i)})^{\alpha}\leq \delta q^{n-k_\nu}T_n^{(i)},$$

that is an eventually exponentially decaying upper bound. One more parameter γ will be needed to prove this, and further hypotheses on δ, ε and q.

- $q\in(\lambda,1)$, which is consistent since $0\leq\lambda<1$.
- $\varepsilon>0$ small enough, to let $(\lambda+\varepsilon)(1+2\varepsilon)^{\alpha-1}<q$, which is consistent since the function $f(\varepsilon)=(\lambda+\varepsilon)(1+2\varepsilon)^{\alpha-1}$ has $\lim_{\varepsilon\to0^+}f(\varepsilon)=\lambda$ and is increasing for $\varepsilon>0$, since $f'(\varepsilon)=(1+2\varepsilon)^{\alpha-1}+2(\alpha-1)(\lambda+\varepsilon)(1+2\varepsilon)^{\alpha-2}$ and therefore, from the right, it approaches λ from above, and since $0\leq\lambda<q<1$, an $\varepsilon>0$ small enough, to satisfy the condition, exists by the continuity of $f(\varepsilon)$ on $\varepsilon\geq0$.
- $\gamma\in(\max\{0,\alpha-2\},\alpha-1)$.
- δ should be small enough, such that

$$\sqrt{d^{\alpha-1}\delta^{\frac{\gamma}{\alpha-1}}}\max_{m\in\mathbb{N}_0}\left[(m+1)\sqrt{q^{\frac{\gamma m}{\alpha-1}}}\right]<\varepsilon.$$

  This last condition is well posed, since the continuous function $h(x)=(x+1)q^{\gamma x/[2(\alpha-1)]}$, for $x\in\mathbb{R}$, is such that $h(0)=1$ and $h(x)$ vanishes as $x\longrightarrow\infty$, due to the exponential q^x beating the linear term (recall that $0<q<1$).

> By a standard compactness argument (split the nonnegative half line in two at some large enough value, after which $h(x)$ is smaller than some subunitary value, and consider the compact subinterval between zero and this large value, on which the maximum is attained by continuity), $h(x)$ achieves its maximum on the nonnegative half-line, and evaluating for $x = m \in \mathbb{N}_0$, the maximum of the sequence $\{h(m)\}$ will be also achieved at one of such nonnegative integers m and, making δ small enough, the factor $\delta^{\gamma/[2(\alpha-1)]}$ will be able to make the maximum arbitrarily small, making possible the condition that it be smaller than ε.

In order to suitably apply (F.3), define the sequence

$$c_n = \sqrt{\frac{\sigma_n}{\tau_{n-1}^{\alpha}}}^{\frac{\gamma}{\alpha-1}-1}.$$

Since $\gamma < \alpha - 1$, $c_n \longrightarrow \infty$ because ${\sigma_n}/{\tau_{n-1}^{\alpha}} \longrightarrow 0$. This last fact is a consequence of $\rho_n \longrightarrow \infty$ and $\lambda < 1$. Indeed, in this regime $\sigma_n \sim \tau_n$, since

$$\frac{\sigma_n}{\tau_n} = \frac{1}{\rho_{n-1}} + 1 \longrightarrow 1.$$

Next rewrite

$$\frac{\frac{\sigma_{n+1}}{\tau_n^{\alpha}}}{\frac{\sigma_n}{\tau_{n-1}^{\alpha}}} = \frac{\rho_n}{\rho_{n-1}}\left(\frac{\tau_{n-1}}{\tau_n}\right)^{\alpha-1} = \frac{\rho_n}{\rho_{n-1}^{\alpha}}\left(\frac{1}{1+\rho_{n-1}}\right)^{\alpha-1} = \frac{\rho_n}{\rho_{n-1}^{\alpha}}\left(\frac{\rho_{n-1}}{1+\rho_{n-1}}\right)^{\alpha-1}.$$

Having $\rho_n \longrightarrow \infty$,

$$\left(\frac{\rho_{n-1}}{1+\rho_{n-1}}\right)^{\alpha-1} \longrightarrow 1,$$

also,

$$\frac{\rho_n}{\rho_{n-1}^{\alpha}} = \frac{\sigma_{n+1}\tau_{n-1}^{\alpha}}{\tau_n\sigma_n^{\alpha}} \sim \frac{\sigma_{n+1}\sigma_{n-1}^{\alpha}}{\sigma_n^{\alpha+1}},$$

therefore

$$\limsup_{n\to\infty} \frac{\frac{\sigma_{n+1}}{\tau_n^{\alpha}}}{\frac{\sigma_n}{\tau_{n-1}^{\alpha}}} = \limsup_{n\to\infty} \frac{\rho_n}{\rho_{n-1}^{\alpha}} = \limsup_{n\to\infty} \frac{\sigma_{n+1}\sigma_{n-1}^{\alpha}}{\sigma_n^{\alpha+1}} = \lambda < 1,$$

thus

$$\sum_{n=0}^{\infty} \frac{\sigma_{n+1}}{\tau_n^{\alpha}} < \infty,$$

which yields that ${\sigma_{n+1}}/{\tau_n^{\alpha}}$ vanishes. From this discussion, since by *Step 1* almost surely on $\mathscr{E}$ all k_n are finite; since by Lemma E.3 almost surely eventually $\varepsilon_n^{(i)} \leq n-1 < n$; since by *Step 2*, $\mathscr{E}_n$ infinitely often almost surely occurs on Ω; we can conclude that for every $\omega \in \Omega$, there is an index $\nu(\omega)$ such that $\mathscr{E}_\nu$ occurs and (think of it large enough) such that $k_\nu < \infty$ almost surely on $\mathscr{E}$ (by *Step 1* again), $\varepsilon_n^{(i)} \leq n$ for all $n \geq k_\nu$ (by Lemma E.3) and such that for all $n > k_\nu$,

$$\frac{\sigma_{n+1}\tau_{n-1}^{\alpha}}{\tau_n\sigma_n^{\alpha}} < \lambda + \varepsilon,$$

which is possible because

$$\limsup_{n\to\infty} \frac{\sigma_{n+1}\tau_{n-1}^{\alpha}}{\tau_n\sigma_n^{\alpha}} = \lambda.$$

We now show that, by induction, this last condition yields that for all $n \geq k_\nu$,

$$\frac{\sigma_{n+1}}{\tau_n^\alpha}(T_n^{(i)})^\alpha \leq \delta q^{n-k_\nu} T_n^{(i)}. \tag{F.4}$$

For $n = k_\nu$, this follows by the definition of k_ν given in *Step 1*, as trivially $q^{n-k_\nu} = 0$. For $n > k_\nu$, assuming the induction hypothesis satisfied for $n-1$, yields that

$$(T_{n-1}^{(i)})^{\alpha-1} \leq \delta q^{n-1-k_\nu} \frac{\tau_{n-1}^\alpha}{\sigma_n},$$

and knowing that $\omega \in \mathscr{E}_\nu$, which means $\varepsilon_n^{(i)} \leq c_n(n-k_\nu)$ for all $n > k_n u$, all these results can be plugged into (F.1). Rewrite the claim to be proved as

$$\frac{\sigma_{n+1}}{\tau_n^\alpha}(T_n^{(i)})^{\alpha-1} \leq \delta q^{n-k_\nu},$$

then use (F.1) and, as aforementioned,

$$\begin{aligned}
\frac{\sigma_{n+1}}{\tau_n^\alpha}(T_n^{(i)})^{\alpha-1} &\leq \frac{\sigma_{n+1}}{\tau_n^\alpha}\left(T_{n-1}^{(i)} + d^{\alpha-1}(T_{n-1}^{(i)})^\alpha \frac{\sigma_n}{\tau_{n-1}^\alpha} + \varepsilon_n^{(i)}\sqrt{\sigma_n P_{n-1}^{(i)}(1-P_{n-1}^{(i)})}\right)^{\alpha-1} \\
&\leq \frac{\sigma_{n+1}}{\tau_n^\alpha}\left(T_{n-1}^{(i)} + d^{\alpha-1}(T_{n-1}^{(i)})^\alpha \frac{\sigma_n}{\tau_{n-1}^\alpha} + c_n(n-k_\nu)\sqrt{d^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^\alpha}\right)^{\alpha-1} \\
&= \frac{\sigma_{n+1}}{\tau_n^\alpha}(T_{n-1}^{(i)})^{\alpha-1}\left(1 + d^{\alpha-1}(T_{n-1}^{(i)})^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha} + c_n(n-k_\nu)\sqrt{d^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^{\alpha-2}}\right)^{\alpha-1} \\
&\leq \frac{\sigma_{n+1}}{\tau_n^\alpha}\delta q^{n-1-k_\nu}\frac{\tau_{n-1}^\alpha}{\sigma_n}\left(1 + d^{\alpha-1}\delta q^{n-1-k_\nu} + c_n(n-k_\nu)\sqrt{d^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^{\alpha-2}}\right)^{\alpha-1}.
\end{aligned}$$

Since it is possible to find $\max\{0, \alpha-2\} < \gamma < \alpha - 1$, which is always positive, the direction of the bound can be preserved (the definition of gamma given earlier comes precisely from this argument, to avoid the possible negativity of $\alpha - 2$, which would otherwise require to proceed by cases):

$$(T_{n-1}^{(i)})^{\alpha-2} \leq (T_{n-1}^{(i)})^\gamma = \left((T_{n-1}^{(i)})^{\alpha-1}\right)^{\frac{\gamma}{\alpha-1}} \leq \left(\delta q^{n-1-k_\nu}\frac{\tau_{n-1}^\alpha}{\sigma_n}\right)^{\frac{\gamma}{\alpha-1}},$$

hence

$$\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^{\alpha-2} \leq \left(\delta q^{n-1-k_\nu}\frac{\tau_{n-1}^\alpha}{\sigma_n}\right)^{\frac{\gamma}{\alpha-1}} \frac{\sigma_n}{\tau_{n-1}^\alpha} = \left(\delta q^{n-1-k_\nu}\right)^{\frac{\gamma}{\alpha-1}}\left(\frac{\sigma_n}{\tau_{n-1}^\alpha}\right)^{1-\frac{\gamma}{\alpha-1}},$$

then

$$\begin{aligned}
&\sqrt{d^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^{\alpha-2}} \leq d^{\frac{\alpha-1}{2}}\sqrt{\left(\delta q^{n-1-k_\nu}\right)^{\frac{\gamma}{\alpha-1}}\left(\frac{\sigma_n}{\tau_{n-1}^\alpha}\right)^{1-\frac{\gamma}{\alpha-1}}} = \\
&d^{\frac{\alpha-1}{2}}\left(\delta q^{n-1-k_\nu}\right)^{\frac{\gamma}{2(\alpha-1)}}\sqrt{\left(\frac{\sigma_n}{\tau_{n-1}^\alpha}\right)^{1-\frac{\gamma}{\alpha-1}}}
\end{aligned}$$

and therefore, to eliminate this nonconstant term, we define

$$c_n = \sqrt{\left(\frac{\sigma_n}{\tau_{n-1}^\alpha}\right)^{\frac{\gamma}{\alpha-1}-1}},$$

so that

$$
\begin{aligned}
& c_n\sqrt{d^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^{\alpha-2}} \le d^{\frac{\alpha-1}{2}}\left(\delta q^{n-1-k_\nu}\right)^{\frac{\gamma}{2(\alpha-1)}} c_n\sqrt{\left(\frac{\sigma_n}{\tau_{n-1}^\alpha}\right)^{1-\frac{\gamma}{\alpha-1}}} \\
& = d^{\frac{\alpha-1}{2}}\left(\delta q^{n-1-k_\nu}\right)^{\frac{\gamma}{2(\alpha-1)}} .
\end{aligned}
$$

In conclusion,

$$
\begin{aligned}
& \frac{\sigma_{n+1}}{\tau_n^\alpha}(T_n^{(i)})^{\alpha-1} \\
& \le \delta q^{n-1-k_\nu}\frac{\sigma_{n+1}\tau_{n-1}^\alpha}{\sigma_n\tau_n^\alpha}\left(1+d^{\alpha-1}\delta q^{n-1-k_\nu}+(n-k_\nu)d^{\frac{\alpha-1}{2}}\left(\delta q^{n-1-k_\nu}\right)^{\frac{\gamma}{2(\alpha-1)}}\right)^{\alpha-1} \\
& \le \delta q^{n-1-k_\nu}\frac{\sigma_{n+1}\tau_{n-1}^\alpha}{\sigma_n\tau_n^\alpha}\left(1+d^{\alpha-1}\delta+d^{\frac{\alpha-1}{2}}\delta^{\frac{\gamma}{2(\alpha-1)}}(n-k_\nu)q^{\frac{\gamma(n-k_\nu-1)}{2(\alpha-1)}}\right)^{\alpha-1} \\
& \le \delta q^{n-1-k_\nu}\frac{\sigma_{n+1}\tau_{n-1}^\alpha}{\sigma_n\tau_n^\alpha}\left(1+\varepsilon+d^{\frac{\alpha-1}{2}}\delta^{\frac{\gamma}{2(\alpha-1)}}\max_{m\in\mathbb{N}_0}\left[(m+1)q^{\frac{\gamma m}{2(\alpha-1)}}\right]\right)^{\alpha-1} \\
& \le \delta q^{n-1-k_\nu}\frac{\sigma_{n+1}\tau_{n-1}^\alpha}{\sigma_n\tau_n^\alpha}(1+2\varepsilon)^{\alpha-1} \le \delta q^{n-1-k_\nu}(\lambda+\varepsilon)(1+2\varepsilon)^{\alpha-1} \le \delta q^{n-k_\nu}
\end{aligned}
$$

which is equivalent to (F.4).

Step 4. The exponential decaying bound is strong enough to allow a proof of the boundedness of $T_n^{(i)}$ on $\mathscr{E}$, a contradiction that implies $\mathbb{P}(\mathscr{E})=0$. Since on $\mathscr{E}$, for all $n\ge k_\nu$, it holds that the usual iterative bound becomes, iterated $n-k_\nu$ times after applying (F.4),

$$
\begin{aligned}
T_n^{(i)} & \le T_{n-1}^{(i)} + d^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^\alpha + \varepsilon_n^{(i)}\sqrt{\sigma_n P_{n-1}^{(i)}(1-P_{n-1}^{(i)})} \\
& \le T_{n-1}^{(i)}\left(1+d^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^{\alpha-1}+\frac{n}{T_{n-1}^{(i)}}\sqrt{\sigma_n P_{n-1}^{(i)}(1-P_{n-1}^{(i)})}\right) \\
& \le T_{n-1}^{(i)}\left(1+d^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^{\alpha-1}+\frac{n}{T_{n-1}^{(i)}}\sqrt{d^{\alpha-1}\sigma_n\frac{(T_{n-1}^{(i)})^\alpha}{\tau_{n-1}^\alpha}}\right) \\
& = T_{n-1}^{(i)}\left(1+d^{\alpha-1}\frac{\sigma_n}{\tau_{n-1}^\alpha}(T_{n-1}^{(i)})^{\alpha-1}+\frac{n}{\sqrt{T_{n-1}^{(i)}}}\sqrt{d^{\alpha-1}\sigma_n\frac{(T_{n-1}^{(i)})^{\alpha-1}}{\tau_{n-1}^\alpha}}\right) \\
& \le T_{n-1}^{(i)}\left(1+d^{\alpha-1}\delta q^{n-k_\nu-1}+\frac{n}{\sqrt{T_{n-1}^{(i)}}}\sqrt{d^{\alpha-1}\delta q^{n-k_\nu-1}}\right) \\
& \le T_{n-1}^{(i)}\left(1+d^{\alpha-1}\delta q^{n-k_\nu-1}+n\sqrt{d^{\alpha-1}\delta q^{n-k_\nu-1}}\right) \le \dots \\
& \le T_{k_\nu}^{(i)}\prod_{j=1}^{n-k_\nu}\left(1+d^{\alpha-1}\delta q^{j-1}+(k_\nu+j)\sqrt{d^{\alpha-1}\delta q^{j-1}}\right) \le \\
& \le T_{k_\nu}^{(i)}\prod_{j=1}^{\infty}\left(1+d^{\alpha-1}\delta q^{j-1}+(k_\nu+j)\sqrt{d^{\alpha-1}\delta q^{j-1}}\right) .
\end{aligned}
$$

The infinite product converges, since

$$\begin{aligned}
&\prod_{j=1}^{\infty}\left(1+d^{\alpha-1}\delta q^{j-1}+(k_\nu+j)\sqrt{d^{\alpha-1}\delta q^{j-1}}\right)\\
&=\exp\sum_{j=1}^{\infty}\log\left(1+d^{\alpha-1}\delta q^{j-1}+(k_\nu+j)\sqrt{d^{\alpha-1}\delta q^{j-1}}\right)\\
&\leq \exp\left(\sum_{j=1}^{\infty}d^{\alpha-1}\delta q^{j-1}+\sum_{j=1}^{\infty}(k_\nu+j)\sqrt{d^{\alpha-1}\delta q^{j-1}}\right)\\
&\leq \exp\left(d^{\alpha-1}\delta\sum_{j=1}^{\infty}q^{j-1}\right)\exp\left(\sqrt{d^{\alpha-1}\delta}\sum_{j=1}^{\infty}(k_\nu+j)\sqrt{q}^{j-1}\right)\\
&=\exp\left(d^{\alpha-1}\delta\sum_{j=1}^{\infty}q^{j-1}\right)\exp\left(\sqrt{d^{\alpha-1}\delta}\sum_{j=1}^{\infty}j\sqrt{q}^{j-1}\right)\exp\left(\sqrt{d^{\alpha-1}\delta}k_\nu\sum_{j=1}^{\infty}\sqrt{q}^{j-1}\right),
\end{aligned}$$

and all series converge, since $\sum_{j=1}^{\infty} q^{j-1}$ and $\sum_{j=1}^{\infty}\sqrt{q}^{j-1}$ are geometric series with q and $\sqrt{q}$ subunitary, and $\sum_{j=1}^{\infty} j\sqrt{q}^{j-1}$ converges by the *ratio test*, since as $j \longrightarrow \infty$,

$$\frac{(j+1)\sqrt{q}^{j}}{j\sqrt{q}^{j-1}} \longrightarrow \sqrt{q} < 1.$$

Since for all $n \geq k_\nu$ there is a random variable that uniformly bounds

$$T_n^{(i)} \leq T_{k_\nu}^{(i)}\prod_{j=1}^{\infty}\left(1+d^{\alpha-1}\delta q^{j-1}+(k_\nu+j)\sqrt{d^{\alpha-1}\delta q^{j-1}}\right) < \infty,$$

the contradiction with $T_n^{(i)} \longrightarrow \infty$ follows. As aforementioned, this implies, through $\mathbb{P}(\mathscr{E}) = 0$, that on $\{\Theta_n^{(i)} \longrightarrow 0\}$, $T_n^{(i)}$ is almost surely bounded.

□

Bibliography

[1] O. Angel, N. Crawford, and G. Kozma. Localization for linearly reinforced random walks. *Duke Math. J.*, 163(5):889-921, 2014.

[2] B. Arthur. Increasing Returns and Path Dependence on the Economy. The University of Michigan Press, 1994.

[3] A. M. Awad. Conditional Levy Continuity Theorem. *Univ. Nac. Tucuman Rev. Ser. A*, 29(2):1-6, 1992.

[4] M. Benaïm, Dynamics of stochastic approximation algorithms. In *Séminaire de probabilités XXXIII*, Vol. 33, 1-68, 1999.

[5] Z. Benzaid and D. A. Lutz. Asymptotic representation of solutions of perturbed systems of linear difference equations. *Stud. Appl. Math.*, 77:195-221, 1987.

[6] D. Bertacchi and F. Zucca. Recent results on branching random walks. In *Statistical Mechanics and Random Walks: Principles, Processes and Applications*, 289–340, Nova Science Publishers, 2012.

[7] I. Bordeu, S. Amarteifio, R. Garcia-Millan, B. Walter, N. Wei, and G. Pruessner. Volume explored by a branching random walk on general graphs. *Sci. Rep.*, 9:15590, 2019.

[8] A. Collevecchio, C. Cotar, and M. LiCalzi. On a preferential attachment and generalized Pólya urn model. *Ann. Appl. Probab.*, 23(3):1219-1253, 2013.

[9] D. Coppersmith and P. Diaconis. Random walks with reinforcement. *Unpublished manuscript*, 1987.

[10] C. Cotar and D. Thacker. Edge- and vertex-reinforced random walks with superlinear reinforcement on infinite graphs. *Ann. Probab.*, 45(4):2655-2706, 2017.

[11] B. Davis. Reinforced random walk. *Probab. Th. Rel. Fields*, 84:203-229, 1990.

[12] P. Diaconis. Group representations in probability and statistics. Institute of Mathematical Statistics, 1988.

[13] P. Diaconis. Recent progress on de Finetti's notion of exchangeability. In *Bayesian Statistics*, 111-125, Oxford University Press, 1988.

[14] E. Drinea, A. Frieze, and M. Mitzenmacher. Balls in bins process with feedback. In *SODA*, 308-315, SIAM, 2002.

[15] E. Dumonteil, S. N. Majumdar, A. Rosso, and A. Zoia. Spatial extent of an outbreak in animal epidemics. *Proc. Natl. Acad. Sci. USA*, 110:4239–4244, 2013.

[16] R. Durret. Probability: Theory and examples. 5th ed. Cambridge University Press, 2019.

[17] S. Elaydi. Introduction to difference equations. Springer, 2005.

[18] W. Feller. An Introduction to Probability Theory and its Applications, Vol. II. 2nd ed. Wiley, 1971.

[19] K. Khanin and R. Khanin. A probabilistic model for the establishment of neuron polarity. *J. Math. Biol.*, 42(1):26-40, 2001.

[20] M. Keane and S. Rolles. Edge-reinforced random walk on finite graphs. In *Infinite Dimensional Stochastic Analysis*, Vol. 52 of *Verhandelingen, Afdeling Natuurkunde. Eerste Reeks. Koninklijke Nederlandse Akademie van Wetenschappen*, 217-234, Royal Netherland Academy of Arts and Sciences, 2000.

[21] G. Kozma. Reinforced random walk. In *ECM*, 429-443, EMS, 2014.

[22] E. L. Lehmann. Elements of large-sample theory. Springer, 1999.

[23] V. Limic. Attracting edge property for a class of reinforced random walks. *Ann. Probab.*, 31:1615-1654, 2003.

[24] V. Limic and P. Tarrès. Attracting edge and strongly edge reinforced random walks. *Ann. Probab.*, 35:1783-1806, 2007.

[25] V. Limic and P. Tarrès. What is the difference between a square and a triangle? In *In and Out of Equilibrium 2. Progress in Probability*, Vol. 60, 481-495, Birkhäuser, 2008.

[26] L. Lovász. Random walks on graphs: a survey. In *Combinatorics, Paul Erdös is eighty*, Vol. 2, 353-397, Janos Bolyai Mathematical Society, 1996.

[27] D. Majerek and W. Zięba. Conditional martingales. *Acta Math.*, 32(1):41-50, 2007.

[28] F. Merkl and S. Rolles. Linearly edge-reinforced random walks. In *Dynamics & Stochastics: Festschrift in honor of M. S. Keane*, Vol. 48 of *IMS Lecture Notes Monogr. Ser.*, 66-77, Institute of Mathematical Statistics Press, 2006.

[29] F. Merkl and S. Rolles. A random environment for linearly edge reinforced random walks on inifinite graphs. *Probab. Th. Rel. Fields*, 138:157-176, 2007.

[30] M. Menshikov and V. Shcherbakov. Balls-in-bins models with asymmetric feedback and reflection. *ALEA Lat. Am. J. Probab. Math. Stat.*, 20:1–19, 2023.

[31] M. Mitzenmacher, R. Oliveira, and J. Spencer. A scaling result for explosive processes. *Electron. J. Combin.*, 11(1):R31, 2004.

[32] R. Oliveira. The onset of dominance in balls-in-bins processes with feedback. *Random Struct. Algorithms*, 34(4):419-526, 2009.

[33] H. G. Othmer and A. Stevens. Aggregation, blowup, and collapse: the ABCs of taxis in reinforced random walks. *SIAM J. Appl. Math.*, 57:1044-1081, 1997.

[34] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979, 2015.

[35] R. Pemantle. Random Processes with reinforcement. *Doctoral dissertation*, M.I.T., 1988.

[36] R. Pemantle. Phase transition of reinforced random walk and RWRE on trees. *Ann. Prob.*, 16:1229-1241, 1988.

[37] R. Pemantle. A time-dependent version of Pólya's urn. *J. Theoret. Probab.*, 3:627-637, 1990.

[38] R. Pemantle. Vertex-reinforced random walk. *Probab. Th. Rel. Fields*, 92:117-136, 1992.

[39] R. Pemantle. A survey of random processes with reinforcement, *Probab. Surveys*, 4:1-79, 2007.

[40] M. Pituk. A link between Perron-Frobenius theorem and Perron theorem for difference equations. *Linear Algebra Appl.*, 434(2):490-500, 2011.

[41] G. Pólya. Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und Momenten problem. *Math. Zeitschrift*, 8, 1920.

[42] S. Rolles. How edge-reinforced random walk arises naturally. *Probab. Th. Rel. Fields*, 126(2):243-260, 2003.

[43] S. Rolles. On the recurrence of edge-reinforced random walk on $\mathbb{Z} \times G$. *Probab. Th. Rel. Fields*, 135(2):216–264, 2006.

[44] C. Sabot and P. Tarrès. Edge-reinforced random walk, Vertex-Reinforced Jump Process and the supersymmetric hyperbolic sigma model. *J. Eur. Math. Soc.*, 17:2353-2378, 2015.

[45] Q. Sattentau. Avoiding the void: cell-to-cell spread of human viruses. *Nat. Rev. Microbiol.*, 6:815–826, 2008.

[46] T. Sellke. Reinforced random walk on the d-dimensional integer lattice. *Technical Report*, 94-26, Purdue University, 1994.

[47] C. Shapiro and H. R. Varian. Information Rules. Harvard Business School Press, 1999.

[48] N. Sidorova. Time-dependent balls and bins model with positive feedback. arXiv:1809.02221v1 [math.PR], *Preprint*, 2018.

[49] N. Sidorova. Time-dependent Pólya urn. arXiv:1807.04844 [math.PR], *Preprint*, 2018.

[50] D. Williams. Probability with martingales. Cambridge University Press, 1991.

[51] D. Yuan and L. Lei. Some results following from conditional characteristic functions. *Commun. Stat. - Theory Methods*, 45(12):3706-3720, 2016.