universität wien

DISSERTATION

Titel der Dissertation

Co-evolutionary dynamics of networks and play

Verfasser

Mag. Mathias Staudigl

Angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, im Dezember 2009

Studienkennzahl lt. Studienblatt:	084 140
Dissertationsgebiet lt. Studienblatt:	Economics
Betreuer / Betreuerin:	Univ.-Prof. Dr. Immanuel Bomze
	Univ.-Prof. Dr. Manfred Nermuth

Co-evolutionary dynamics of networks and play

Mathias Staudigl

Preface

This thesis is the outgrowth of a long struggle trying to combine evolutionary game theory with the fascinating science of networks. Reading all the papers and books from the various disciplines, encompassing such completely disperse fields such as mathematical sociology, evolutionary biology, statistical physics, mathematics, and economics, often made me more confused than clarified how I should think about network theory. My pragmatic conclusion, at least from my point of view today, is that there is no other way to think about networks than as abstract modeling devices, with which large and complex systems can be translated into a mathematical language. And there we have the connection! Just as game theory is simply a language to model strategic *interactions* among agents with *interdependent* utility functions, so is network theory a language to model dependency structures among *interacting* entities. Note the words: Interaction, Interdependency. Now it is clear that these two disciplines must have something in common. My objective in this dissertation is to elaborate more on this close relationship between ideas coming from evolutionary game theory and models from random graph theory. Indeed, one of the central findings of this dissertation is that evolutionary game theory can be used to give a behavioral micro-foundation to a fairly large class of random graph models, known as inhomogeneous random graphs. This interesting connection makes it possible to endogenize interaction structure into a local interaction system in a transparent way. While random graph theory is more in the domain of statistical physics and mathematics, local interaction systems have received much interest in economic theory. Combining these two fields is an exciting challenge, and this thesis is just a possible first step to fully explore the scope of the interrelations. We are still very far from a completely satisfying understanding of the co-evolution of networks and play and I hope that future research will find one or the other result presented in this thesis to be valuable.

Writing these essays was not an easy task. At this stage I should like to spell out my gratefulness to all those persons who helped me on this partly painful, but most of the time enormously exciting, journey.
First, I would like to thank my two supervisors, Immanuel Bomze and Manfred Nermuth, for reading and discussing all the material. To my principal supervisor, Immanuel Bomze, thanks for giving me the freedom to pursue my work and listening to my ideas. Christina Pawlowitsch deserves many credits for reading some of the papers of this dissertation and making many astute comments. Thanks also to Stefano DeMichelis for bringing me in contact with the mathematical literature on random graphs, and for having organized this beautiful accommodation in Pavia. Michael König has to be mentioned for bringing the physics literature closer to my attention and for many helpful discussions on this topic. Carlos Alós-Ferrer must be thanked for giving me the right critique at the right time, and William H. Sandholm for a detailed list of comments and suggestions. Last, but certainly not least, I would like to thank the person without whom I would have never been able to write even a word (beside doing a PhD!). This goes out to you Petra, and I am very happy to dedicate this work to you.

Contents

Chapter 1

Introduction

1.1 Co-evolution of networks and play

Traditionally, evolutionary game theory focused on settings involving large populations of anonymous and (strategically) interacting agents (Sandholm, 2009b). The simplest setting considers a single population of individuals who are recurrently and randomly matched with other agents to play some given normal form game (for a critical account of this random matching hypothesis see Boylan, 1992, Alós-Ferrer, 1999). After the matching took place, the agents observe an outcome, which is, in most cases, the utility generated from the interaction. An evolutionary process describes how agents respond to this outcome by adapting their behavior. The typical story is that agents have some cognitive ability to evaluate the outcome of the matching, and they employ *behavioral rules* which prescribe their behavior in the next round of matching. Such behavioral rules can be very sophisticated, such as always best responding to the previous outcome of the matching, or can be rather simple such as variants of reinforcement learning (see Weibull, 1995, Hofbauer and Sigmund, 1998, Sandholm, 2009b, and the references therein). Whatever the specific assumption is, one can model this strategy adjustment process as a random process, which may be approximated by a deterministic dynamical system when the number of interacting and updating agents is sufficiently large. The technically most convenient situation arises if one assumes a continuum population at the outset. In such a case, the description of the long-run evolution of the society boils effectively down to an aggregate population dynamics, capturing the most essential statistical regularities of the underlying process, e.g. the frequency distribution over pure actions used by the agents. These "mean-field dynamics" capture the expected law of motion of the aggregate frequency of actions, and its rest points reflect situations where almost all agents in the population have no interest to change their behavior (see Benaïm and Weibull, 2003, Sandholm, 2009b, for a formal justification of the use of mean-field methods). Of course, real populations are finite, and the main theme of this dissertation will be that the pattern of interaction, i.e. the *design* of the matching technology, is of high importance in these models. To give an illustration on this point, let us consider the following simple 2×2 game:

	a_1	a_2
a_1	(e, e)	(f, g)
a_2	(g, f)	(h, h)

(1.1.1)

The exact values of the entries in this table are not of importance. We require however that $e > g$ and $h > f$, so that we have a coordination game with strict Nash equilibria $(a_1, a_1), (a_2, a_2)$. Applying any reasonable evolutionary dynamic on this game produces the vector field depicted in Figure 1.1. Calling x the global frequency of a_1-players in the population,

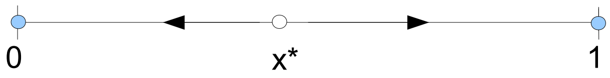

Figure 1.1: Vector Field of any reasonable deterministic evolutionary dynamic generated by the game (1.1.1).

we see that the heterogeneity equilibrium, where a fraction $x^* = \frac{h-f}{e-g+h-f}$ of individuals play a_1 and the remaining fraction plays a_2, is an unstable fixed point of the dynamic. The two strict Nash equilibria are asymptotically stable fixed points, with basins of attraction separated exactly at x^*. Hence, from a dynamic point of view, the society will find itself in a situation where every player chooses the same action. This global conformity prediction is one the one hand confirming and on the other hand disappointing. It is confirming, since we want that the agents can agree in a decentralized way to play a reasonable equilibrium, which is in (1.1.1) certainly only a pure strategy equilibrium. On the other hand, it is disappointing since a prediction in which all agents behave in the very same way misses some fundamental observations made in the real world: *persistent* heterogeneity in the modes of behavior of the agents at the aggregate level of society. Once we recognize that the way how people interact shapes the way how they behave (and vice versa), we can easily model a situation where players who interact more frequently (i.e. they are matched on a regular basis) with each other will coordinate on the same action, but agents who never meet my play different strategies. To illustrate the point, let me present a very simple "model" which is already capable to produce a picture of *local conformity*, but *global heterogeneity*. The population we consider is a finite one, but is best thought to be fairly large, say N. These N individuals play the game (1.1.1) on a regular basis with a subset of players, which we call, in a rather informal way, their *neighbors*. The agents do not need to know much about the specific attributes of their neighbors; They just know that the common game in question is (1.1.1). Each player starts with some action in the game, and only at some random points in time a randomly chosen individual gets the chance to revise his decision what to play in subsequent periods. Suppose that player i received such an opportunity. Conditional on this event, she employs a very simple (imitative) behavioral rule, namely the following:

> "Pick the most prominent action in the neighborhood. In case of a tie, stick to your action."

Over time, the neighborhoods change, as do social relationships in the real world. To capture this phenomenon, we introduce two further adjustment rules, which directly act on the *social network* of neighborhood relations. At some random moments of time, let one randomly chosen individual get the opportunity to search for a new agent, who is not currently a neighbor. Conditional on this event, she does this in the following way:

> "Take the first best individual you meet on your search. If he plays the same action as you do, become a neighbor of him with probability $1 - \varepsilon$. Otherwise, become a neighbor of him only with probability ε."

Such simple *global search protocols* are frequently used in recent models on the evolution of social networks (see Vega-Redondo, 2007, and the reference therein). The probability ε can be interpreted as the rate of experimentation of an agent.

To complete the description of the network evolution process, we need some force counteracting the repeated creation of new connections. The simplest way to introduce such a "death-element" is by assuming that every existing connection can be destroyed with equal probability. Such assumptions have been made in "phenomenological" models (Ehrhardt et al., 2006b) of network evolution, and there this effect is often traced back to some exogenous force of *network volatility*. Volatility is a crucial element in the models of the subsequent chapters. While many papers in the network formation literature model volatility as an unguided drift term destroying any currently existing edge in the network at a constant rate, it may also be used as a modeling device of the network formation mechanism. A very general model, allowing for all these interpretations of volatility is presented in chapter 2.[1]

These simple rules of adjustment give rise to a well defined Markov chain acting on the tuple of action profiles and neighborhood relationships, which models a *stochastic co-evolutionary process of networks and play*. To get a rough idea on the possible scenarios generated by such a process we may employ computer simulations. The results of a simulation of such a process are

[1] It is interesting to see that a sort of volatility has also been used by Blume (1995). In his continuous flow model players are randomly matched and the relationship holds for a random amount of time until it is dissolved again. This is a formulation of volatility in form of unguided drift.

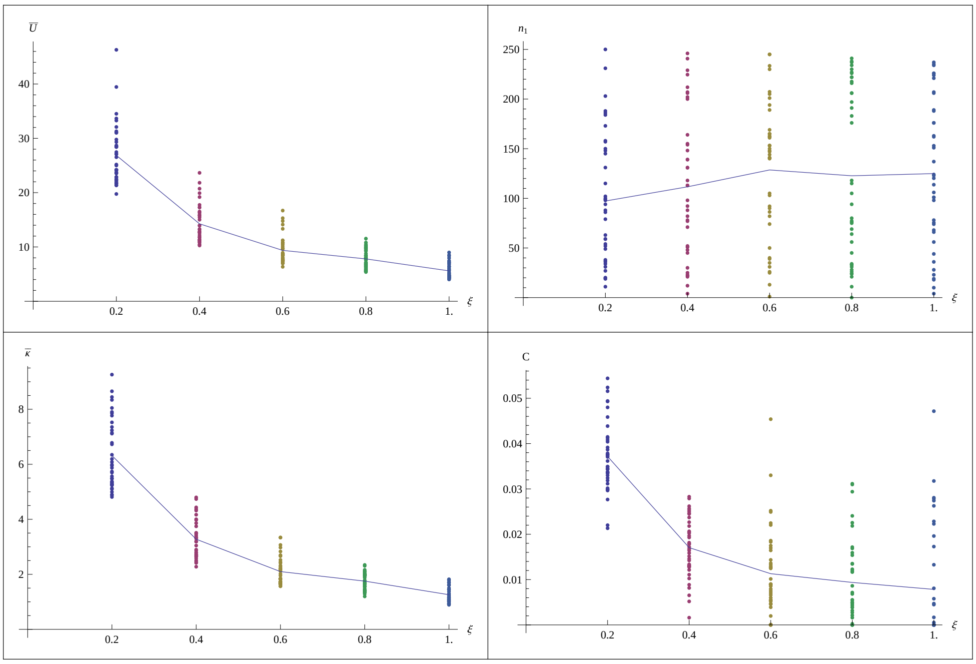

Figure 1.2: Results from a computer simulation of a simple co-evolutionary process with payoffs $[e, f, g, h] = [5, 2, 1, 4]$ in game (1.1.1). The plots are summary statistics from 40 independent computer simulations performed with the program NetLogo 4.0.4 (Wilensky, 1999). With rate $\nu = 4$ a randomly chosen player receives an action adjustment opportunity, with rate $\lambda = 1$ an agent, who is not completely connected yet, may form a new link, and with the variable rate $\xi \in \{0.2, 0.4, 0.6, 0.8, 1\}$ a randomly chosen link becomes destroyed. The rate of experimentation is $\varepsilon = 0.05$. Each simulation run was given 40.000 rounds for relaxation, and the periods 40.001-40.201 were used for data taking. Dots are the averages over this time span for each run. The closed curve connects averages over all simulations time average. Population size is 250. The initial network is a degree-regular graph with mean degree of 5. The initial number of a_1 players is chosen at random.

depicted in Figure 1.2. We see there the outcome of a series of independent computer experiments where a society of 250 individuals is involved in the co-evolutionary process sketched above. We evaluate these numerical experiments in terms of 4 aggregate statistics: the average utility of the agents ($\bar{U}$), the number of a_1-players (n_1), the average number of neighbors an agent has ($\bar{\kappa}$, called the average *degree* of an individual), and the clustering coefficient of the network (C).[2] The following general picture emerges. The larger in magnitude the effect of volatility, the more difficult it is for the players to achieve coordination. We may see this from the inverse relationship observed in all four plots. In terms of network characteristics this is rather obvious, since a higher rate of volatility implies that more links will be destroyed on average, and in equilibrium this must result in a sparser graph. However, the interesting thing is that higher volatility leads also to a drastic drop in the mean utility the agents obtain in the long run. From the frequency statistic on the number of a_1-players we also see, that the society consists, notably in the long-run, of a_1 as well as a_2-players. This is the co-existence result we wanted. The fall in utility might have two sources: Persistent mis-coordination among the agents, or "under-connectedness" for too high levels of volatility. The "movie" displayed in Figures 1.3 and 1.4 illustrates that it is the second effect which causes the loss in welfare.

Taking different snapshots of the social network for different environmental scenarios reveals the full working of volatility in the model. The upper row of plots in Figure 1.3 depict a scenario of low environmental volatility. We see an emerging pattern of two large blocks in the society, corresponding to the two different types of behavior in the game. The degree of separation becomes more accentuated as time proceeds, and in the case of no experimentation, i.e. $\varepsilon = 0$, we would observe two disjoints components in the network. For higher levels of volatility a different picture emerges. The networks will be fairly sparse, meaning that every individual will interact only with few neighbors, and we observe a tree-like architecture at the aggregate level. This pattern holds for a large range of intermediate values of volatility, as can be seen form the lower row of Figure 1.3 and the upper row of Figure 1.4. For very high levels of volatility, most agents will be loners, making the model not so interesting. We also see from Figures 1.3, 1.4 that it is not the agents' inability to coordinate successfully. Much more the contrary is the case. Almost all connected pairs of agents displays the same type of behavior. We will see that this remains to be true in coordination games, even if we assume more general (and sophisticated) behavioral rules. We learn from this "toy model", that already simple rules of behavior may lead to interesting phenomena, such as local homogeneity but global diversity, and complicated patterns of interactions. In view of this, we might hope to build more sophisticated models, driven by behaviorally sensible assumptions on the way the players react to the behavior of their neighbors, which yield interesting

[2] Clustering is a standard measure in network analysis, and heuristically measures the number of closed triangles in a graph. Equivalently, one can think of it as the fraction of neighbors of neighbors of an agent who are again his neighbors. We use for C the average over all individual node clustering coefficients, as proposed by Watts and Strogatz (1998).

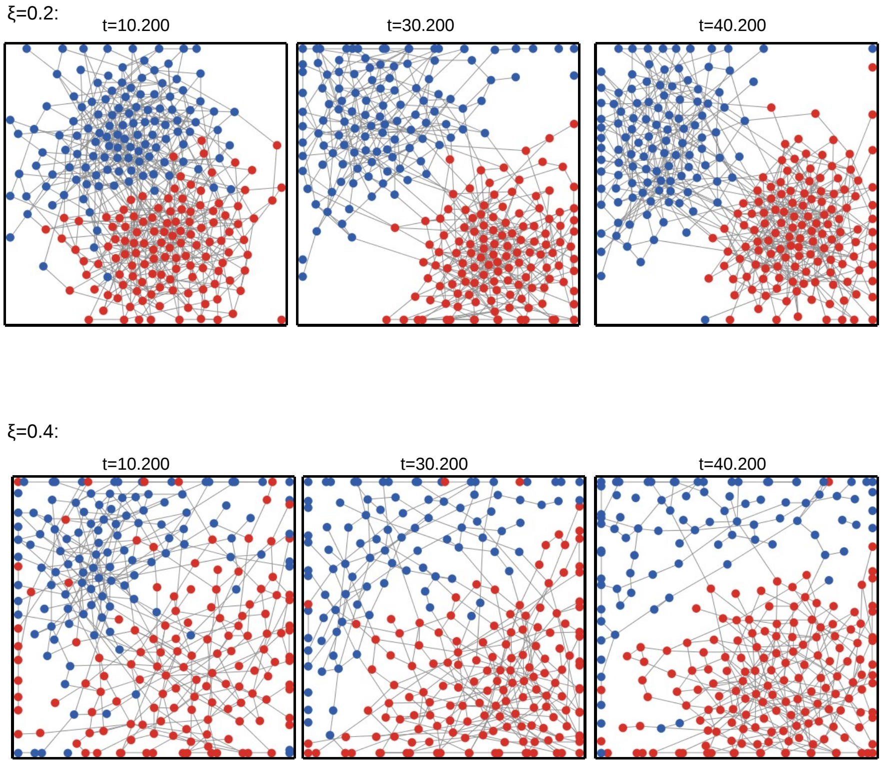

Figure 1.3: Time evolution of the network for different values of volatility ξ. Initially there are 75 a_1-players. Other parameter values are $\lambda = 1, \nu = 4, \varepsilon = 0.05$. Red dots symbolize a_1-players.

outcomes at the macroscopic level of society. The results generated by these models will, in general, be distributions over action profiles and networks, inducing a *statistical ensemble of random graphs*. The characterization of the probability measures of such random graph ensembles is the new element of such models, giving us a more refined picture of aggregate population dynamics.

1.2 Inhomogeneous random graphs

This section should give a rough idea what we actually mean when we speak of a random graph. A rigorous account of this fascinating topic would require the writing of at least one book, and so we will only convey basic ideas. The classical reference on random graphs is Bollobás (2001), which exclusively

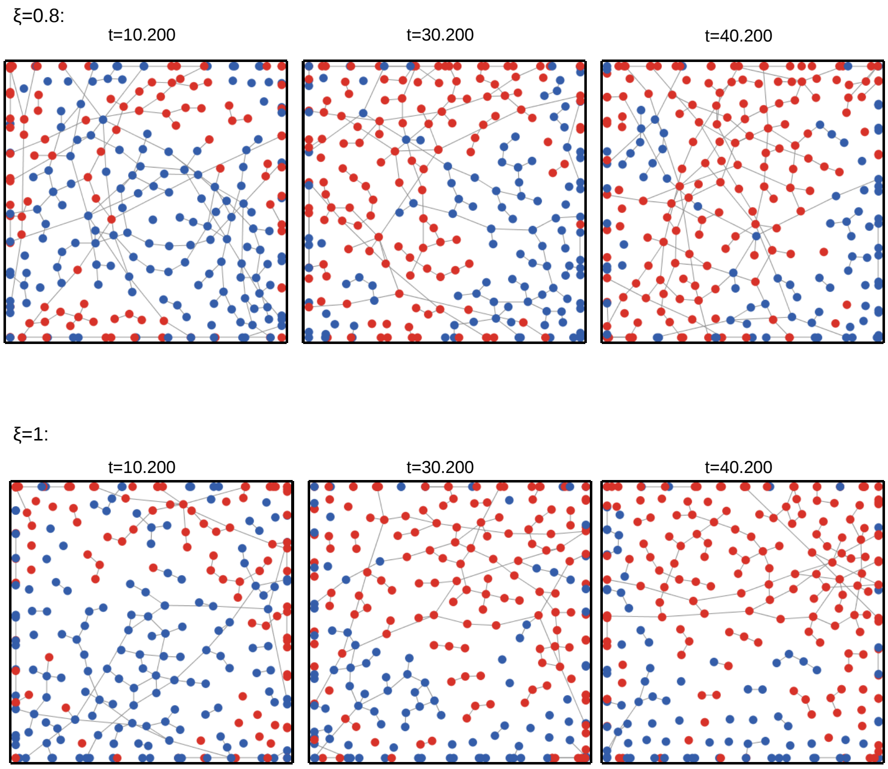

Figure 1.4: Time evolution of the network for different values of volatility ξ. Initially there are 75 a_1-players. Other parameter values are $\lambda = 1, \nu = 4, \varepsilon = 0.05$. Red dots symbolize a_1-players.

studies the Erdös-Rényi model (Erdös and Rényi, 1960). Recent, mathematically rigorous, accounts are the books by Durrett (2007) and Chung and Lu (2006), which go beyond the Erdös-Rényi model. To introduce the necessary ideas needed for the subsequent chapters of this thesis, we follow the ingenious approach of Park and Newman (2004), giving an integrated derivation of generalized, or inhomogeneous, random graphs, in a statistical mechanics fashion.

In this thesis we will speak of networks and undirected graphs synonymously. An undirected graph is a pair of sets $G = (\mathcal{I}, \mathcal{E})$, where $\mathcal{I}$ is a countable set of vertices and $\mathcal{E}$ the set of unordered pairs of connected vertices, called the edge set of the graph G. The set of unordered pairs of members of $\mathcal{I}$ is denoted as $\mathcal{I}^{(2)}$. Elements of this set are pairs $(i,j) \equiv (j,i)$ with $i,j \in \mathcal{I}$. Most of the time we will work with finite graphs, meaning the there is only a finite number of vertices. The vertex set is then $\{1, 2, \ldots, N\}$. Moreover, the set $\mathcal{I}$ is in this thesis assumed to be time-invariant, so that the only way how to discriminate between different graphs on the vertex set $\mathcal{I}$ is via its edge set. To indicate this, we write the edge set for a graph G as $\mathcal{E}(G)$. The complete graph on $\mathcal{I}$ is the network where all nodes are interconnected, i.e. if $|\mathcal{I}| = N$ and G^c is the complete graph[3], then $|\mathcal{E}(G)| = N(N-1)/2$. Given a graph $G = (\mathcal{I}, \mathcal{E})$, we call $G' = (\mathcal{V}, \mathcal{E}')$ a *subgraph* of G, if $\mathcal{V} \subseteq \mathcal{I}$ and $\mathcal{E}' \subset \mathcal{E}$. All networks in our models are in this sense subgraphs (with $\mathcal{V} = \mathcal{I}$) of the complete graph G^c, and we call $\mathcal{G}[\mathcal{I}]$ the set of all such graphs. A more algebraic approach to graph theory is provided by working with the edge indicator function $g : \mathcal{I}^{(2)} \times \mathcal{G}[\mathcal{I}] \to \{0, 1\}$, defined as

$$g((i,j), G) = \begin{cases} 1 & \text{if } (i,j) \in \mathcal{E}(G), \\ 0 & \text{otherwise.} \end{cases} \tag{1.2.1}$$

For a graph $G \in \mathcal{G}[\mathcal{I}]$ we may abbreviate $g((i,j), G)$ as $g_{ij}(G)$. The edge indicators establish a one-to-one correspondence[4] between the ensemble of graphs $\mathcal{G}[\mathcal{I}]$ and the set $\{0,1\}^{\mathcal{I}^{(2)}}$, members of which are vectors $g = (g_{ij})_{1 \le i < j \le N}$. Therefore, we will be safe by identifying $\mathcal{G}[\mathcal{I}]$ with the set of edge-indicators. In this sense, we call the vector g a network, or a graph. In the following definition we fix the meaning of an *ensemble of random graphs* (see also the definitions in Vega-Redondo, 2007, Dorogovtsev and Mendes, 2003).

Definition 1.2.1. *A random graph ensemble is a discrete probability space $(\mathcal{G}[\mathcal{I}], \mu)$, where every graph $g \in \mathcal{G}[\mathcal{I}]$ is assigned a statistical weight $\mu(\{g\}) = \mu(g)$. An ensemble of random graphs is called uncorrelated, if all edges appear with independent probability. The edge-success probability of vertices i and j in the uncorrelated random graph ensemble $(\mathcal{G}[\mathcal{I}], \mu)$ is the probability of the event $\{g_{ij} = 1\}$. It is determined by*

$$p_{ij} = \sum_{g \in \mathcal{G}[\mathcal{I}] : g_{ij} = 1} \mu(g) = \sum_{g \in \mathcal{G}[\mathcal{I}]} g_{ij} \mu(g) = \mathbb{E}_{\mu}[g_{ij}]. \tag{1.2.2}$$

If $p_{ij} = p$ for all $i, j \in \mathcal{I}$, we call the uncorrelated ensemble $(\mathcal{G}[\mathcal{I}], \mu)$ the Erdös-Rényi model. Otherwise, we call the ensemble an inhomogeneous random graph.

[3] In graph theory a complete graph on N vertices is normally denoted by K_N. By choosing the above notation, we follow the literature in economics.

[4] This equivalence holds only up to a bijective relabeling of the vertices, since graphs that preserve adjacency under a permutation of the labels of the vertices are regarded as equivalent in graph theory. See Bollobás (1998).

The definition makes it clear that in the Erdös-Rényi model all vertices are statistically equivalent objects, and the graph measure μ is the Bernoulli distribution $\mu(g) = p^{e(g)}(1-p)^{N(N-1)/2-e(g)}$, where $e(g) := \sum_{i,j>i} g_{ij}$ is the number of active edges in the graph g. Hence, the Erdös-Rényi-model uses only information on the number of edges in a graph, and all graphs with the same number of edges appear with the same probability. It is therefore the most basic random graph model and serves as a useful benchmark.
For inhomogeneous random graphs, it follows from the independence assumption that the graph measure must be of the form

$$\mu(g) = \prod_{i,j>i} p_{ij}^{g_{ij}}(1-p_{ij})^{1-g_{ij}}. \tag{1.2.3}$$

We see that in such uncorrelated models the edge-success probabilities provide all information on the random graph ensemble.
The following procedure, outlined in Park and Newman (2004), is a simple way to estimate a general random graph model $(\mathcal{G}[\mathcal{I}], \mu)$ from empirical observations, and to determine the edge-success probabilities. Let $\mathcal{G}[\mathcal{I}]$ be the set of feasible networks in the sequel of our experiments. For every graph $g \in \mathcal{G}[\mathcal{I}]$ we may have different kinds of observations which depend on g. Let $\vec{x} := (x_k)_{k=1,\dots,r}$ denote a set of observable attributes (i.e. the level of investments and profits of a set of firms in a trade network), and $x_k(g)$ the value of the observable k given the graph g. Let $\bar{x}_k$ denote the empirical average of the observations of characteristic $k = 1, 2, \dots, r$. We want to determine a probability measure μ on $\mathcal{G}[\mathcal{I}]$, that fits the empirical averages of the graph observables as best as possible. One approach to solve this problem is by writing down the negated entropy function

$$h := \sum_{g \in \mathcal{G}[\mathcal{I}]} \mu(g) \log \mu(g),$$

and formulating the convex minimization problem

$$\begin{aligned} &\min h, \\ \text{s.t.} &\sum_{g \in \mathcal{G}[\mathcal{I}]} \mu(g) x_k(g) = \bar{x}_k \quad 1 \le k \le r, \\ &\sum_{g \in \mathcal{G}[\mathcal{I}]} \mu(g) = 1. \end{aligned}$$

Remember that the argument with respect to which we want to minimize h is the probability distribution μ. We can solve this optimization problem by standard methods, by assigning to each graph observable the Lagrangian multiplier θ_k. Let $\vec{\theta} := (\theta_1, \dots, \theta_r)$, and denote by $\langle \vec{\theta}, \vec{x} \rangle := \sum_{k=1}^{r} \theta_k x_k$ the standard inner product on $\mathbb{R}^r$. The first-order conditions read as

$$1 + \log \mu(g) + \langle \vec{\theta}, \vec{x}(g) \rangle = 0$$

for all $g \in \mathcal{G}[\mathcal{I}]$. Let us define the *graph Hamiltonian* $H : \mathcal{G}[\mathcal{I}] \to \mathbb{R}$, as

$$H(g) := \langle \vec{\theta}, \vec{x}(g) \rangle.$$

Then we can reformulate the first-order conditions as

$$(\forall g \in \mathcal{G}[\mathcal{I}]) : \quad 1 + \log \mu(g) = H(g),$$

and consequently, for any pair of graphs $g, g' \in \mathcal{G}[\mathcal{I}]$,

$$\frac{\mu(g')}{\mu(g)} = \exp[H(g') - H(g)].$$

Summing over all $g' \in \mathcal{G}[\mathcal{I}]$, and using the normalization condition $\sum_{g' \in \mathcal{G}[\mathcal{I}]} \mu(g') = 1$, we get

$$(\forall g \in \mathcal{G}[\mathcal{I}]) : \quad \mu(g) = \frac{\exp(H(g))}{\sum_{g' \in \mathcal{G}[\mathcal{I}]} \exp(H(g'))}. \tag{1.2.4}$$

Such a probability measure is known in statistical mechanics as a *Gibbs measure*, and will appear frequently in the papers presented in chapters 4 and 5. In the statistical analysis of social networks, such a probability measure is also a prominent choice. In this literature it is known as the exponential random graph model (ERGM), or p^*-model (Wasserman and Pattison, 1996, Anderson et al., 1999, Snijders et al., 2006). These models have their roots in the Markov graphs of Strauss and Frank (1986) and Strauss (1986).
To demonstrate the generality of this approach, we will show that the Erdös-Rényi model and inhomogeneous random graphs can be obtained by suitably specifying the graph Hamiltonian. For the Erdös-Rényi model, we put $H(g) := \theta e(g)$ (emphasizing again that the only variable of this ensemble is the edge-count). Plugging this into the probability measure (1.2.4) gives

$$\mu(g) = p^{e(g)}(1-p)^{N(N-1)/2 - e(g)}.$$

To see this, we need to compute the normalizing factor (called the partition function in statistical physics)

$$\begin{aligned}
\sum_{g\in\mathcal{G}[\mathcal{I}]} \exp(H(g)) &= \sum_{g\in\mathcal{G}[\mathcal{I}]} \exp(\theta \sum_{i=1}^{N}\sum_{j>i} g_{ij}) \\
&= \prod_{i=1}^{N}\prod_{j>i} \sum_{g_{ij}\in\{0,1\}} \exp(\theta g_{ij}) \\
&= \prod_{i=1}^{n}\prod_{j>i} (1+\exp(\theta)) \\
&= (1+\exp(\theta))^{N(N-1)/2}
\end{aligned}$$

Let $p := (1+\exp(-\theta))^{-1} = \frac{\exp(\theta)}{1+\exp(\theta)}$, so that $1-p = \frac{1}{1+\exp(\theta)}$ and $\frac{p}{1-p} = \exp(\theta)$, to get

$$\begin{aligned}
\mu(g) &= \exp(\theta)^{e(g)}(1+\exp(\theta))^{-N(N-1)/2} \\
&= p^{e(G)}(1-p)^{N(N-1)/2-e(g)}.
\end{aligned}$$

For the inhomogeneous random graph, we assign to each element in $\mathcal{I}^{(2)}$ an edge weight θ_{ij}, and set $H(g) := \sum_{i=1}^{N}\sum_{j>i}\theta_{ij}g_{ij}$. The partition function for this model is

$$\sum_{g'\in\mathcal{G}[\mathcal{I}]} \exp(H(g')) = \prod_{i=1}^{N}\prod_{j>i}(1+\exp(\theta_{ij})).$$

Defining $p_{ij} := \frac{\exp(\theta_{ij})}{1+\exp(\theta_{ij})}$, we obtain for all $g \in \mathcal{G}[\mathcal{I}]$

$$\mu(g) = \prod_{i=1}^{N}\prod_{j>i} p_{ij}^{g_{ij}}(1-p_{ij})^{1-g_{ij}}.$$

In a co-evolutionary model, the state variable is the complete profile of actions *together* with the network. Hence, as one of the long-run characteristics of such a process we will observe an induced ensemble of networks. The purpose of this thesis is to present models were these ensembles can be characterized to a large extent. It will turn out that the statistical mechanics approach of Park and Newman (2004) is quite fruitful for these purposes, and we will further show that co-evolutionary models, as defined in this thesis, have a strong relation with inhomogeneous random graphs. To give a first impression of this, we will compare the statistical outcome of the experiment performed in Section 1.1 with the Erdös-Rényi graph. The only two graph statistics,

ξ	0.2	0.4	0.6	0.8	1
$\bar{\kappa}$	6.32	3.28	2.10	1.75	1.26
C	0.0372	0.0170	0.0113	0.0094	0.0078
C_{ER}	0.0253	0.013	0.009	0.007	0.005

Table 1.1: Numerical values of plot 1.2. For a given graph $g \in \mathcal{G}[\mathcal{I}]$, the average degree is defined by $\bar{\kappa}(g) := \frac{1}{N}\sum_{i\in\mathcal{I}}\kappa^i(g) = \frac{1}{N}\sum_{i,j\in\mathcal{I}} g_{ij} = \frac{2e(g)}{N}$. The clustering-coefficient for a single vertex is calculated as $C^i(g) := \frac{\sum_{j\in\mathcal{I}}\sum_{k>j} g_{ij}g_{jk}g_{ki}}{\sum_{j\in\mathcal{I}}\sum_{k>j} g_{ij}g_{ik}}$, and $C(g) := \frac{1}{N}\sum_{i\in\mathcal{I}} C^i(g)$.

visualized in Figure 1.2, were the clustering coefficient and the average degree. Its precise values are given in Table 1.1.

In an Erdös-Rényi graph, an asymptotically sharp upper bound for the clustering coefficient is $\bar{\kappa}/N$ (see Newman, 2002). This bound is used for C_{ER} in Table 1.1. We see that clustering produced by the model is only slightly above clustering predicted by the simple Erdös-Rényi model. This suggests that the generated ensemble should be close to an uncorrelated random graph. However, Figures 1.3 and 1.4 showed that the network has a clearly observable structure (e.g. block building between a_1 and a_2 players). This distinguishes it from an Erdös-Rényi graph, as can be seen in Figures 1.5, 1.6 where Erdös-Rényi graphs with similar average degrees are visualized.[5] We see that these graphs miss the structure found in Figures 1.3, 1.4. This suggests that there are at least two probabilistic laws governing the interaction structure in the society (corresponding to the two actions in the game (1.1.1)), making it unlikely that the network as a whole is an Erdös-Rényi ensemble.

1.3 Structure of the Thesis

This dissertation presents three interrelated models, written in the spirit of the co-evolutionary "model" sketched in Section 1.1. The purpose of this thesis is to convince the reader that studying the interaction structure, jointly with the behavioral profile of a population, gives us a much richer picture on the long-run outcomes of evolutionary processes. I hope that this goal has been at least partly accomplished.

[5]These plots have been generated with Mathematica, using the Spring Electrical Embedding Algorithm for visualization.

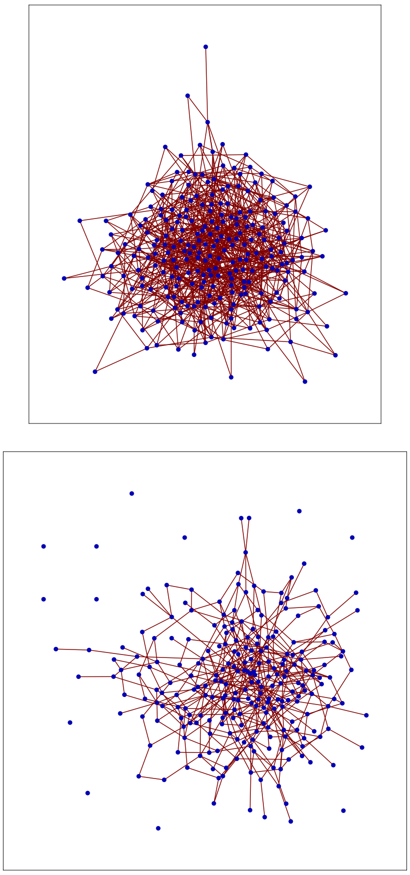

Figure 1.5: Erdös-Rényi graph with mean degree $\bar{\kappa} = 6.25$ and $\bar{\kappa} = 3.25$.

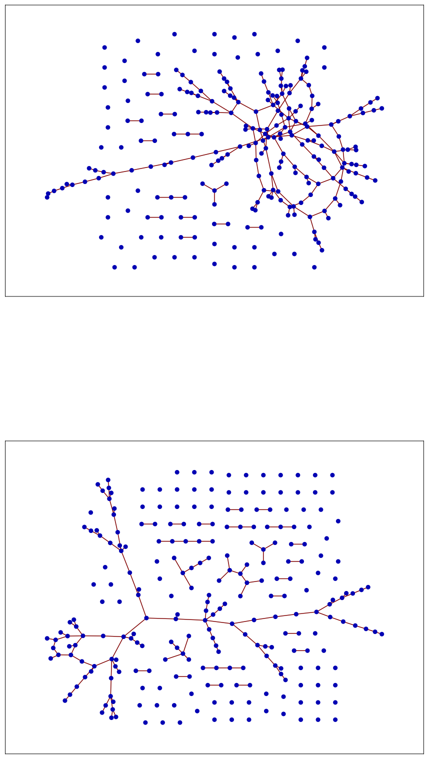

Figure 1.6: Erdös-Rényi graph with mean degree $\bar{\kappa} = 1.75$ and $\bar{\kappa} = 1.25$.

Chapter 2 defines the general mathematical framework of the models presented in this thesis. The innovation of the article presented in this chapter is the definition of *stochastic co-evolutionary models with noise*, which extends the models of evolution with noise of Kandori et al. (1993), Young (1993) an Ellison (2000). We do so by imposing a list of "axioms" on the structure of the behavioral rules the agents may employ. These "axioms" are of a technical nature; any behavioral assumption should be made on top of these. We show that, equipped with this set of "axioms" alone, the process possess already sufficient structure to crystal out some important features. Building on a rich literature on stochastic evolutionary game dynamics (see the textbooks of Samuelson, 1997, Young, 1998, Sandholm, 2009b), and simulated annealing (in particular the work of Catoni, 1999; 2001), we provide a tree-characterization of the long-run invariant distribution of such co-evolutionary models. This allows us to propose a procedure which identifies stochastically stable states in these large models. Influential literature therefore have been Beggs (2005) and Alós-Ferrer and Netzer (2007). Beyond that, the general framework is sufficiently rich to encompass various recent models on the co-evolution of networks and play, as will be shown by means of two examples. We also establish a rather deep and interesting connection between the just mentioned co-evolutionary literature, and the mathematical literature on *inhomogeneous random graphs* (see Section 1.2). Particular important references are here, beside the ones already mentioned, Söderberg (2002) and the impressive account by Bollobás et al. (2007). As defined in Section 1.2, in an inhomogeneous random graph the probabilities with which two randomly chosen individuals become successfully connected may depend on certain characteristics of the agents (such as their action chosen in a game, or their idiosyncratic preferences). Söderberg (2002) and Bollobás et al. (2007) assume some form of exogenously given heterogeneity in the population. We provide a closed-form expression of the edge-success probabilities as a corollary of the long-run dynamic equilibrium of the co-evolutionary process. Hence, the paper provides a *micro-foundation*, and therefore a new and independent definition, of the class of inhomogeneous random graphs. This completely new result opens the way to explore many more connections between the theory of stochastic game dynamics and random graph theory.

Chapters 4 and 5 present more specific models, but are nevertheless in the spirit of the general framework of chapter 2. The paper presented in chapter 4, entitled "Potential games played in volatile environments", is heavily influenced by the path-breaking work of Ehrhardt et al. (2008a). We extend their setting to more general behavioral rules (both in the action and the linking dynamics) and study general potential games (Monderer and Shapley, 1996). In this, admittedly simple, framework many properties of the long-run behavior of the system can be expressed in beautiful closed-form expressions. Among other things, this paper presents the degree-distribution of the induced ensemble of random graphs, which is - obviously in view of chapter 2 - an inhomogeneous random graph. A complete description of the generated random graph ensemble is provided in Theorem 4.5.1, called the *Erdös-Rényi-decomposition*. Further, we present a rather general theorem on the concentration of the invariant measure on the set of potential maximizers, thereby generalizing classical results of Blume (1993; 1997) and Young (1998). The proof of this theorem differs from the proofs presented in these references, by relying on Markov's inequality and the structure of the invariant measure alone. Inspiration for the argument came from the study of simulated annealing (Brémaud, 1998, Catoni, 1999), and gives a clearer picture on the mechanics of the process. An application to pure-coordination games illustrates the introduced concepts.

Chapter 5 is an extension of the model of chapter 4, but still considers potential games. In this paper, entitled "Co-evolutionary dynamics and Bayesian interaction games", we introduce a new class of games, called *structured Bayesian interaction games*. This class of games merges the interaction games of Morris (1997; 2000) and the Bayesian population games of Ely and Sandholm (2005) and Sandholm (2007a). An interaction game is an N-player game, where interaction is modeled by an arbitrary graph, or network. In our model, we always assume interaction to be binary, i.e. two agents play per interaction. The term "structured" highlights that the matching will in general be distinct from the global random matching framework briefly mentioned in Section 1.1. A Bayesian population game is a population game where the agents' preferences are diverse. The seminal paper of Ely and Sandholm (2005) introduced this class of games in an unstructured continuum population. Sandholm (2007a) uses this framework to give a new (evolutionary) proof of the classical Harsanyi purification theorem (Harsanyi, 1973). In this beautiful paper, the type of a player is his preference bias. We adapt this setup to structured Bayesian interaction games. The model is therefore capable to describe the long-run evolution of a society with significant heterogeneity, and we are still able to analytically pin down the invariant distribution of this process, and to demonstrate the connection with the class of inhomogeneous random graphs. The latter is derived in form of another *Erdös-Rényi decomposition*, as in the paper presented in chapter 4.

The models of chapter 4 and 5 are continuous-time Markov processes, while the general framework is formulated for discrete-time Markov chains. This poses no technical problems, once one realizes that every finite-state continuous-time Markov process can be fully studied by its so-called *embedded jump chain*. Chapter 3 establishes the connection with these models in a self-contained way, and provides a full description on the construction of the generator and the stochastic semi-group modeling evolution in the continuous-time models.

Chapter 6 summarizes our results and discusses some ideas for future refinements and extensions of the models. Although the chapters are related, starting from being very general, and then becoming more specific, each paper stands for its own. Thus, every chapter of this thesis can be read separately without having looked at previous ones.

Chapter 2

On a general class of stochastic co-evolutionary dynamics

2.1 Introduction

Recently there has been an attempt to apply stochastic evolutionary game dynamics to models on the co-evolution of networks and play. Broadly speaking, one may divide these models in two classes. There is one branch of literature which extends the mistakes model of Kandori et al. (1993) and Young (1993) to a random process of action adjustment and link creation/destruction. Jackson and Watts (2002), Goyal and Vega-Redondo (2005), Hojman and Szeidl (2006) are models in this direction, and we call them, due to their ancestry, "classical" models. Another type of models assume that the network is under a recurrent attack of unguided drift, which is interpreted as environmental volatility. Marsili et al. (2004), Ehrhardt et al. (2006a; 2008a) are models in this direction, which we will call "volatility" models. The aim of this paper is to present an unified framework, that is rich enough to incorporate classical, as well as volatility models. We do so by presenting a rather general class of co-evolutionary models, called $\mathcal{M}^{\beta}$. In essence $\mathcal{M}^{\beta}$ is a family of perturbed Markov chains taking values on some finite state space Ω, which consists of all pairs of action profiles (α) and networks (g). We give an "axiomatic" definition of processes $\mathcal{M}^{\beta}$ which models the co-evolution of networks and play in an integrated way. At a heuristic level, the algorithm works as follows:
Suppose the system starts from some point $\omega = (\alpha, g)$. Departing from this state, the system may evolve via three possible routes. With some probability a randomly chosen individual gets the opportunity to change his action. This causes a change in the action profile α. With complementary probability the network changes, resulting in the creation of a new edge, or the destruction of an existing edge. The characterizing feature of the process is that the *behavioral rules*, describing how agents change their action, or how they create or delete links depend, in general, on the benefits of the bilateral interaction, which, in turn, is modeled by a game in normal form. This produces an interesting coupling between the evolution of the action profile α and the evolution of the network g. After one of these events, the process arrives at a new state, and the algorithm repeats these steps infinitely often.

The objective of this paper is to investigate the asymptotic properties of this stochastic algorithm. We assume that the rules defining the transition probabilities of $\mathcal{M}^{\beta}$ are governed by a *noise parameter* $\beta \in \mathbb{R}_+$, as is by now standard in stochastic evolutionary models.[1] For $\beta > 0$ the process will

[1] On the importance of noise in game theory see Binmore and Samuelson (1999) and Blume (2003).

be ergodic, and the long-run predictions are given by its unique invariant distribution $\mu^\beta \in \Delta(\Omega)$. In principle, the invariant distribution contains all information one needs to deduce more specific information about the long-run behavior of the system, such as the *marginal* probability distribution over action configurations (the object studied in "classical" evolutionary game theory with fixed interaction structure) and the *conditional* probability distribution over networks.[2] Particularly interesting is the behavior of the invariant distribution as noise vanishes. This leads to the study of stochastically stable states, which is one of the most prominent selection criteria of evolutionary game theory. The traditional way to perform stochastic stability analysis is by viewing the Markov process as a weighted and directed graph and looking for paths with least resistance. Kandori et al. (1993) and Young (1993) pioneered this approach, by adapting tools developed by Freidlin and Wentzell (1998). The first contribution of this paper is the presentation of a tree-characterization algorithm to compute stochastically stable states in general co-evolutionary models. Thereby we obtain a selection criterion of recurrent classes of states consisting of profiles of actions and architectures of interaction, extending traditional models of evolutionary game theory where only the action profiles are considered as state variable. The "classical" models of Jackson and Watts (2002) and Goyal and Vega-Redondo (2005) are also concerned with this task. Our general model provides a systematic tool kit to find stochastically stable states in a transparent way. We show by means of two examples, a "volatility" model and a "classical" model based on Jackson and Watts (2002), that such a stochastic stability analysis is still tractable in co-evolutionary models.

The second, and truly original, contribution of this paper is the characterization of the generated *random graph ensemble*, conditional on a fixed profile of actions. For this characterization we impose 3 additional "axioms". We show that any stochastic process, satisfying the stated assumptions, will converge in the long run to the probability ensemble of so-called *inhomogeneous random graphs* (Söderberg, 2002, Bollobás et al., 2007). Inhomogeneous random graphs are a straightforward extension of the classical Erdös-Rényi model (Erdös and Rényi, 1960), by allowing edge success probabilities to be vertex specific. These models are very popular in the literature on random graphs, and to the best of our knowledge, this interesting connection between evolutionary game dynamics and random graph theory is novel. A co-evolutionary model with noise provides therefore a new and independent derivation of in-

[2]Due to the coupling of the behavior dimension with the network dimension it would make no sense to study a marginal distribution over networks. Only a conditional distribution, i.e. the probability distribution over networks for a *fixed* action profile, makes sense in these models.

homogeneous random graphs.

The class of Markov chains we study is known in the literature on stochastic optimization as a "generalized Metropolis algorithm", and is rigorously surveyed by Catoni (1999; 2001). Beggs (2005) was among the first to recognize the close relationship between this class of random processes, and the stochastic dynamics used in evolutionary game theory. We also exploit this analogy and show that it provides a flexible language to study many models on the co-evolution of networks and play. To underline this, we devote a whole section to show that the models presented in Ehrhardt et al. (2008a;b) and chapters 4, 5 fit perfectly into our framework. A minor modification of the process also allows us to study the model of Jackson and Watts (2002). Related to our work is also the recent paper by Alós-Ferrer and Netzer (2007). However, these authors fix the behavioral rules of the agents at the outset, by assuming that strategy revisions are governed by the logit dynamics, introduced by Blume (1993) into game theory. Moreover, their paper assumes an exogenously fixed interaction structure.
The rest of the paper is organized as follows. Section 2.2 introduces our theoretical framework. In Section 2.2.2 we derive a general form of the invariant distribution, and an algorithm to detect stochastically stable states. Section 2.3 presents a "classical model" and a "volatility" model. The characterization of the generated random graph ensemble is presented in Section 2.4. Section 2.5 concludes. Sections 2.6 and 2.7 collect some well-known facts on stochastic stability analysis in a self-contained way.

2.2 A class of Markov processes

We consider a finite population of individuals $\mathcal{I} = \{1, 2, \dots, N\}$. The set of all unordered pairs of individuals will be denoted by $\mathcal{I}^{(2)}$. The set of ordered pairs of a finite set Ω is denoted as $\Omega \times \Omega = \Omega^2$. In this paper we identify networks with simple and undirected graphs on the vertex set $\mathcal{I}$. Call $\mathcal{G}[\mathcal{I}]$ the set of all such graphs, members of which are pairs $G = (\mathcal{I}, \mathcal{E})$, where $\mathcal{E} = \mathcal{E}[G] \subseteq \mathcal{I}^{(2)}$ is the set of edges (links). Another convenient representation of a network is via a tuple $g = (g_{ij})_{1 \leq i < j \leq N} \in \{0, 1\}^{\mathcal{I}^{(2)}} \equiv \mathcal{G}[\mathcal{I}]$. If $g_{ij} = 1$ we say that individual i is connected to individual j, or j is a neighbor of i (and vice versa). Another terminology for connectedness will be that the edge (i, j) is active. If $g_{ij} = 0$ then i and j are not connected, or edge (i, j) is neutral. The neighbors of player i in the network g are contained in the set $\mathcal{N}^i(g) := \{j \in \mathcal{I} | g_{ij} = 1\}$. Call $\bar{\mathcal{N}}^i(g) := \mathcal{N}^i(g) \cup \{i\}$. The number of neighbors of player i defines his degree $\kappa^i(g) := |\mathcal{N}^i(g)|$. Given a network g

and a subset of players $\mathcal{V} \subseteq \mathcal{I}$ denote the restriction of g on $\mathcal{V}$ as $g[\mathcal{V}]$, which is an element of $\mathcal{G}[\mathcal{V}]$. The complete network on the subset $\mathcal{V}$ is denoted by $g^c[\mathcal{V}]$. Hence, for every $g \in \mathcal{G}[\mathcal{I}]$ and a partition of $\mathcal{I}$ into sets $\mathcal{V}_1, \mathcal{V}_2$, we can write $g = g[\mathcal{V}_1] \oplus g[\mathcal{V}_2]$, where $\oplus$ is interpreted as the concatenation of two lists of binary valued functions (after possibly relabeling the players). In this notation $g' = g \oplus g^c[\{(i,j)\}] \equiv g \oplus (i,j)$ is the network obtained by adding the edge (i,j) to g. Analogously, $g' = g \ominus (i,j)$ is the network obtained from g by deleting edge (i,j). Denote by $e(g) = \sum_{i,j>i} g_{ij}$ the number of edges in the network g.
Each individual possesses a utility function u^i, describing her preferences over some finite set of common actions $\mathcal{A} = \{a_1, \ldots, a_q\}$.[3] This defines a *base game* $\Gamma = (\mathcal{I}, \mathcal{A}, (u^i)_{i\in\mathcal{I}})$.
The utility player i gets from choosing one of these actions depends on the behavior of his neighboring players. Let $\alpha = (\alpha^i)_{i\in\mathcal{I}} \in \mathcal{A}^{\mathcal{I}}$ denote an action profile of the population. A *population state* is a pair $\omega = (\alpha, g) \in \mathcal{A}^{\mathcal{I}} \times \mathcal{G}[\mathcal{I}] \equiv \Omega$. Given an action profile α let $\alpha_i^a = (a, \alpha_{-i}) = (\alpha^1, \ldots, \alpha^{i-1}, a, \alpha^{i+1}, \ldots, \alpha^N)$. Utility of player i at state ω is defined as

$$\pi^i(\alpha, g) \equiv \pi^i(\omega) := \sum_{j\in\mathcal{N}^i(g)} u^i(\alpha^i, \alpha^j). \tag{2.2.1}$$

2.2.1 Co-evolution with noise

In the spirit of Young (1993) and Ellison (2000), we call a *co-evolutionary model with noise* a family of perturbed time-homogeneous Markov chains

$$\mathcal{M}^\beta = \left(\Omega, \mathcal{F}, \mathbb{P}, (X_n^\beta)_{n\in\mathbb{N}_0}\right)_{\beta\in\mathbb{R}_+},$$

where $X^\beta = (X_n^\beta)_{n\in\mathbb{N}_0}$ is a family of Ω-valued random variables, indexed by a discrete time parameter n and a noise parameter β; $\mathcal{F}$ is a σ-algebra, and $\mathbb{P} : \mathcal{F} \to [0,1]$ a probability measure. A realization $\{X_n^\beta = \omega\}$ defines an action profile α and a network g. The Markov property states that for any history $A_{n-1} = \{X_0^\beta, \ldots, X_{n-1}^\beta\}$ on which $\{X_{n-1}^\beta = \omega\}$ holds, the probability that the process visits state ω' in the next period depends only on ω, i.e.

$$\mathbb{P}(X_n^\beta = \omega' | A_{n-1}) = \mathbb{P}(X_n^\beta = \omega' | X_{n-1}^\beta = \omega) \equiv K^\beta(\omega, \omega'), \tag{2.2.2}$$

where $K^\beta : \Omega^2 \to [0,1]$ is the transition probability function of the stochastic process X^β. Denote by $\mathbf{K}^\beta := [K^\beta(\omega, \omega')]_{(\omega,\omega')\in\Omega^2}$ the transition matrix of the process X^β. Assume that these probabilities vary continuously with the

[3] In principle every individual could have his own action set. This would require more notation, and does not contribute anything to this paper.

noise parameter β. For $\beta \to 0$ we obtain the *unperturbed Markov chain* $\mathcal{M} = (\Omega, \mathcal{F}, \mathbb{P}, (X_n)_{n \in \mathbb{N}_0})$, with corresponding transition matrix $\mathbf{K}$. Denote by $\Re$ the collection of recurrent classes of the unperturbed chain, and $\mathcal{L}_1, \ldots, \mathcal{L}_k$ the k-recurrent classes of $\mathcal{M}$. By the decomposition theorem $\Omega = \mathcal{Q} \cup \Re$, where $\mathcal{Q}$ is the class of transient states in the unperturbed process.
Given the current state $\{X_n^\beta = \omega\}$, the following 3 events may take place:

Action adjustment: With probability $q_1(\omega) \in [0, 1]$ the action configuration α changes. Let $\nu \geq 0$ denote the rate with which player i receives an action revision opportunity.[4] Define the *volume* of the action adjustment process as $N\nu$. The probability that player i gets a revision opportunity is defined as $1/N$. Denote by $b^{i,\beta}(\cdot|\omega)$ a *probabilistic behavioral rule* describing how player i selects an action, given the population state ω. Specifically, assume that this behavioral rule satisfies the two "axioms":

(A1) For all $i \in \mathcal{I}$ and $\beta > 0$, $b^{i,\beta}(\cdot|\omega)$ is a full support distribution on $\mathcal{A}$.

(A2) For all $i \in \mathcal{I}$ there exists a *cost function* $c_1^i : \Omega^2 \to \mathbb{R}_+$ satisfying

$$-\lim_{\beta \to 0} \beta \log b^{i,\beta}(a|\omega) = c_1^i(\omega, (\alpha_i^a, g)). \tag{2.2.3}$$

This can be alternatively written as

$$b^{i,\beta}(a|\omega) = \exp\left[-\frac{1}{\beta}(c_1^i(\omega, (\alpha_i^a, g) + o(1))\right]$$

where $o(1)$ represents terms that go to 0 as $\beta \to 0$.

As $\beta \to 0$ the probability that player i makes a costly decision converges to 0 at exponential rate. A costless transition will be made even in the zero noise limit. Observe that the revision processes of Kandori et al. (1993) and Blume (1993), or adaptive learning of Young (1993) satisfy all these assumptions.[5]

[4] Assuming that this rate is heterogeneous is possible, but this is the basic assumption made in the literature.

[5] (A1) and (A2) are the most basic assumptions. An appealing additional requirement would be

(A3) $(\forall i \in \mathcal{I}) : c_1^i(\omega, (\alpha_i^a, g)) > 0$ iff $a \notin \arg\max_{a' \in \mathcal{A}} \pi^i(\alpha_i^{a'}, g)$.

which says that only suboptimal choices have positive transition costs. In this sense, players use noisy best response rules (see Sandholm, 2009b). However, for the general discussion such an assumption is not necessary.

Link creation: With unconditional probability $q_2(\omega)$ the process allows the network to expand. For all $i \in \mathcal{I}$ define a *rate function* $\lambda^i : \Omega \to \mathbb{R}_+$, satisfying $\kappa^i(\omega) = N-1 \Rightarrow \lambda^i(\omega) = 0$. The *volume* of the link creation process is defined as the sum of all rate functions $\bar{\lambda}(\omega) := \sum_{i\in\mathcal{I}} \lambda^i(\omega)$. The conditional probability that player i receives the chance to form a link is $\lambda^i(\omega)/\bar{\lambda}(\omega)$. Conditional on this event, player i computes a tuple $w^{i,\beta}(\omega) := (w_j^{i,\beta}(\omega))_{j\in\mathcal{I}}$, satisfying:

(L1) If $g_{ij} = 0$ and $\beta > 0$, then $\min\{w_j^{i,\beta}(\omega), w_i^{j,\beta}(\omega)\} > 0$. If $g_{ij} = 1$ or $i = j$, then $w_j^{i,\beta}(\omega) = w_i^{j,\beta}(\omega) = 0$ for all β,

(L2) $(\forall i \in \mathcal{I})(\forall \omega \in \Omega) : \sum_{j\in\mathcal{I}} w_j^{i,\beta}(\omega) = 1$

(L3) $(\forall i, j \in \mathcal{I})(\forall \omega \in \Omega) : -\lim_{\beta\to 0} \beta \log w_j^{i,\beta}(\omega) = c_2^i(\omega, (\alpha, g \oplus (i,j)))$.

$c_2^i : \Omega^2 \to \mathbb{R}_+$ is again a cost function for player i. Condition (L1) says that all neutral edges have a positive probability of becoming created for $\beta > 0$. This is an irreducibility assumption. (L3) is a large deviation assumption on the link creation probability.

Let $\mathbf{W}^\beta(\omega) = \bar{\lambda}(\omega)^{-1}\operatorname{diag}[\lambda^1(\omega), \ldots, \lambda^N(\omega)][w_j^{i,\beta}]_{i,j\in\mathcal{I}}$ denote the matrix of link creation probabilities at state ω.[6] The i-th row of this matrix is $\left(\lambda^i(\omega)/\bar{\lambda}(\omega)\right) w^{i,\beta}(\omega)$.[7] Next, define the symmetric matrix $\bar{\mathbf{W}}^\beta(\omega) := [\bar{w}_{ij}^\beta(\omega)]_{i,j\in\mathcal{I}} = \mathbf{W}^\beta(\omega) + \mathbf{W}^\beta(\omega)^\top$.[8] The scalar $\bar{w}_{ij}^\beta(\omega)$ is the conditional probability that the passive edge (i,j) is formed, starting from ω.

Link destruction: With unconditional probability $q_3(\omega)$ a link becomes destroyed. Let $\xi \geq 0$ denote the constant rate of link destruction.[9] A positive level of volatility will imply that, independent of β, there is always a chance that a link becomes destroyed. Additionally to this drift term, let us assign to each edge (i,j) a weight $v_{ij}^\beta(\omega)$. The higher the weight of an active edge, the larger will be the conditional probability that it becomes destroyed. Let $\mathbf{V}^\beta(\omega) = [v_{ij}^\beta(\omega)]_{1\leq i,j\leq N}$ the $N \times N$ matrix of edge weights, satisfying:

[6] $\operatorname{diag}[x_1, \ldots, x_n]$ is the $n \times n$ diagonal matrix having x_i as entry in its i-th principal diagonal and 0 off the principal diagonal.

[7] Note that the above conditions on the distribution $w^{i,\beta}$ requires that a completely connected individual puts weight 1 one himself. This causes no trouble because such players do not get a link creation opportunity by default. Hence the algorithm produces simple graphs, i.e. graphs that have no multiple connections and self-loops, as desired.

[8] $\mathbf{W}^\top$ is the transposition of $\mathbf{W}$.

[9] This is exactly the volatility parameter of Marsili et al. (2004), Ehrhardt et al. (2006a; 2008a;b).

(D1) $\mathbf{V}^\beta(\omega)$ is a symmetric matrix, and, for $\beta > 0$, $v^\beta_{ij}(\omega) > 0$ if $g_{ij} = 1$, and $v^\beta_{ij}(\omega) = 0$ for $g_{ij} = 0$,

(D2) $\sum_{i,j>i} v^\beta_{ij}(\omega) = 1$,

(D3) $(\forall i \in \mathcal{I})(\forall \omega \in \Omega) : -\lim_{\beta\to 0} \beta \log v^\beta_{ij}(\omega) = c_3^{(i,j)}(\omega, (\alpha, g \ominus (i,j)))$.

(D1) says that edges (i,j) and (j,i) are treated symmetrically. This is a natural assumption for undirected graphs. Moreover, it requires that all currently active edges are destroyed with positive probability if $\beta > 0$. (D2) requires that, conditional on the event of link destruction, the expected number of destroyed edges is 1. (D3) is our large deviation assumption. The *volume* of the link destruction process is defined as $\bar{\xi}(\omega) := \xi f(\omega, \mathbf{V}^\beta)$, where $f(\cdot,\cdot)$ is a bounded non-negative function, normalized by the condition $f(\omega, \mathbf{V}^\beta) = 0$ if the network is the empty graph at ω.[10]

Let $\omega = (\alpha, g)$ be the current population state. Define

$$\Lambda(\omega) = N\nu + \bar{\lambda}(\omega) + \bar{\xi}(\omega). \tag{2.2.4}$$

By the frequency interpretation of probabilities, one can interpret the number $N\nu =: \tau_a$ as the time scale of action adjustment events, and $\bar{\lambda}(\omega) + \bar{\xi}(\omega) =: \tau_g$ as the time scale of network evolution. The ratio $\tau = \tau_g/\tau_a$ measures how fast network evolution is, relative to action adjustment. If τ is much larger than 1, network evolution will proceed at a faster time scale than action adjustment. If τ is much smaller than 1, then action adjustment opportunities arrive much more frequently to the population. The probabilities $q_\sigma(\omega), \sigma = 1,2,3$, specifying the timing of evolution, are defined as

$$q_1(\omega) = \frac{N\nu}{\Lambda(\omega)},\; q_2(\omega) = \frac{\bar{\lambda}(\omega)}{\Lambda(\omega)},\; q_3(\omega) = 1 - q_1(\omega) - q_2(\omega). \tag{2.2.5}$$

The elements of the transition matrix $\mathbf{K}^\beta$ are then given by

$$K^\beta(\omega, \omega') = \begin{cases} q_1(\omega)\frac{1}{N} b^{i,\beta}(a|\omega) & \text{if } \omega' = (\alpha_i^a, g), \\ q_2(\omega)\bar{w}^\beta_{i,j}(\omega) & \text{if } \omega' = (\alpha, g \oplus (i,j)), \\ q_3(\omega)v^\beta_{ij}(\omega) & \text{if } \omega' = (\alpha, g \ominus (i,j)), \\ 0 & \text{otherwise.} \end{cases} \tag{2.2.6}$$

[10]The reason why a positive rate of link destruction is needed is to exclude trivial stationary states where all players are completely connected, simply because all edges are formed with positive probability. Of course, assuming $\xi > 0$ does not exclude the complete graph of being a stationary state. Henceforth assume that $\xi > 0$ and fixed, so that β is the only varying parameter.

It is easy to verify that $\sum_{\omega' \in \Omega} K^\beta(\omega, \omega') = q_1(\omega) + q_2(\omega) + q_3(\omega) = 1$ for all $\omega \in \Omega$. By the irreducibility assumptions (A1), (L1) and (D1), the matrix $\mathbf{K}^\beta$ is irreducible for $(\beta, \xi) \gg (0, 0)$. Further, it is easy to see that the chain is aperiodic. Since Ω is a finite set, ergodicity of the process X^β is guaranteed. Hence, provided $\beta > 0$, there exists a unique invariant distribution $\mu^\beta \in \Delta(\Omega)$. It is well known that for $\beta \to 0$ the process concentrates on a subset of $\Re$. To classify such states, we use the following definition of stochastic stability.[11]

Definition 2.2.1 (Sandholm (2009b)). *Given a co-evolutionary model with noise $\mathcal{M}^\beta$, we call a state $\omega \in \Omega$ stochastically stable if*

$$\lim_{\beta \to 0} \beta \log \mu^\beta(\omega) = 0. \tag{2.2.7}$$

Let Ω^* denote the set of stochastically stable states.

2.2.2 On trees, graphs and stochastic stability

At every point of time the process may undertake one of three different transitions. The most appealing way to think about the stochastic dynamic is in terms of directed graphs, as done by Kandori et al. (1993), Young (1993), building on the work of Freidlin and Wentzell (1998). Every co-evolutionary model with noise $\mathcal{M}^\beta$ can be analyzed via directed graphs of the form $T = (\Omega, \vec{E})$. The vertex set of such graphs is the state space and the edge set is a subset of Ω^2. A graph T will be called a *revision graph*, and we will henceforth identify every revision graph with its edge set $\vec{E}(T)$.

Definition 2.2.2. *Given a co-evolutionary model with noise $\mathcal{M}^\beta$ and a revision graph T, define the* ***reach*** *of state $\omega \in \Omega$ under T as the set*

$$\mathcal{R}_T(\omega) := \{\omega' \in \Omega | (\exists \vec{e} \in \vec{E}(T)) : \vec{e} = (\omega, \omega')\}.$$

The reach of a state is the collection of states that the process may visit after one step under the revision graph T, starting from ω. The reach of a state ω can be subdivided as follows; Call $\mathcal{R}_{T,1}(\omega)$ the set of states in the reach of ω that differ in the action configuration, $\mathcal{R}_{T,2}(\omega)$ the set of states that are reachable from ω by creation of a single link, and finally $\mathcal{R}_{T,3}(\omega)$ the

[11] Most models using stochastic evolutionary dynamics call a state stochastically stable if it receives *positive* weight in the limit distribution. Definition 2.2.1 says that ω is stochastically stable if $\log \mu^\beta(\omega) \to a \leq 0$ as $\beta \downarrow 0$. This is a weaker requirement than the conventional stochastic stability criterion, since it may well be that the mass converges to 0 at a sub-exponential rate. See Sandholm (2009b, ch. 12), for a detailed discussion.

set of states reachable from ω by deleting a single link.[12] We will work with the following special class of graphs. Their role has also been emphasized by Samuelson (1997), Catoni (1999), Beggs (2005) and Alós-Ferrer and Netzer (2007).

Definition 2.2.3. *Consider a non-empty set $\mathcal{X} \subset \Omega$. A revision graph T is called a $\mathcal{X}$-revision graph if it is an element of the class of graphs $\mathcal{T}(\mathcal{X})$, satisfying*

(i) $(\forall \omega \in \Omega) : |\mathcal{R}_T(\omega)| = \mathbb{1}_{\{\omega \notin \mathcal{X}\}}$,

(ii) T does not contain a cycle.

A labeled ω-revision tree (T_ω, ℓ) is a $\{\omega\}$-revision graph $T_\omega \in \mathcal{T}(\{\omega\}) \equiv \mathcal{T}_\omega$ and a function $\ell : \vec{E}(T_\omega) \to \mathcal{I}^{(2)}$ with the property that

(iii) for all edges $\vec{e}$ the labeling function $\ell(\vec{e})$ returns the pair of players (i, j) involved in the transition modeled by the edge $\vec{e} \in \vec{E}(T_\omega)$. If $j = i$ then we interpret the pair (i, i) as i.

A $\mathcal{X}$-revision graph $T \in \mathcal{T}(\mathcal{X})$ joins every point in $\Omega \backslash \mathcal{X}$ to $\mathcal{X}$, without loops. In the main text of the paper we will only need the concept of labeled revision trees. More general $\mathcal{X}$-revision graphs will be of importance in Section 2.7. For this class of revision graphs conditions (i) and (ii) are a version of the standard graph-constructs of Freidlin and Wentzell (1998). They merely assert that T_ω is a tree with root ω.[13] All paths from the branches of the tree lead in a unique way to ω. The distinguishing point in the definition of a labeled revision tree is exactly the labeling function, whose purpose will become clear later on.[14] For a given revision tree $T_\omega \in \mathcal{T}_\omega$, define the set

$$\mathcal{S}_{T_\omega,\sigma} := \{\vec{e} = (\omega', \omega'') \in \vec{E}(T_\omega) | \omega'' \in \mathcal{R}_{T_\omega,\sigma}(\omega')\}, \ \sigma \in \{1, 2, 3\},$$

which is the collection of all edges used on a transition of type $\sigma \in \{1, 2, 3\}$. By definition we have $\vec{E}(T_\omega) = \bigcup_{\sigma=1}^{3} \mathcal{S}_{T_\omega,\sigma}$.
Following Freidlin and Wentzell (1998) we can now completely characterize

[12] Obviously $\mathcal{R}_T(\omega) = \mathcal{R}_{T,1}(\omega) \cup \mathcal{R}_{T,2}(\omega) \cup \mathcal{R}_{T,3}(\omega)$.

[13] Contrary to the visual appearance of trees in nature, a root is here a sink instead of a source.

[14] Note that for the current type of stochastic process, the labeling function is uniquely defined for a given revision tree T_ω. See Alós-Ferrer and Netzer (2007) for a process where this need not be the case.

the invariant distribution of the co-evolutionary process. With a slight abuse of notation define the numbers

$$\begin{aligned} K^{\beta}(T_{\omega}) &:= \prod_{\vec{e} \in \vec{E}(T_{\omega})} K^{\beta}(\vec{e}) = \prod_{\sigma=1}^{3} \prod_{\vec{e} \in \mathcal{S}_{T_{\omega},\sigma}} K^{\beta}(\vec{e}), \\ \rho^{\beta}(\omega) &:= \sum_{(T_{\omega},\ell) \in \mathcal{T}_{\omega}} K^{\beta}(T_{\omega}). \end{aligned}$$

Theorem 2.2.1 (The Markov chain tree theorem). *For $\beta > 0$, the unique invariant distribution of the co-evolutionary model with noise $\mathcal{M}^{\beta}$ is given by*

$$(\forall \omega \in \Omega) : \mu^{\beta}(\omega) = \frac{\rho^{\beta}(\omega)}{\sum_{\omega' \in \Omega} \rho^{\beta}(\omega')}. \tag{2.2.8}$$

Proof. See section 2.7. This is Lemma 3.1 of Freidlin and Wentzell (1998). This representation holds for *every* irreducible Markov chain, and is not restricted to the current model. See Young (1998) or Sandholm (2009b) for alternative elegant proofs of this fact. □

Consider a state $\omega \in \Omega$ with revision tree (T_{ω}, ℓ). By construction of the transition probabilities, for every edge $\vec{e} \in \mathcal{S}_{T_{\omega},\sigma}, \sigma = 1, 2, 3$ there exists a *derived* cost function $\hat{c}_{\sigma} : \Omega^2 \to \mathbb{R}_+ \cup \{+\infty\}$, such that

$$K^{\beta}(\vec{e}) = \exp\left[-\frac{1}{\beta}(\hat{c}_{\sigma}(\vec{e}) + o(1))\right], \ \sigma \in \{1, 2, 3\},$$

depending on the type of transition under the edge $\vec{e}$.[15] If the transition $\vec{e} \in \mathcal{S}_{T_{\omega},\sigma}$ is not possible for $\beta > 0$, then set $\hat{c}_{\sigma}(\vec{e}) = \infty$. Define the derived costs of a revision tree (T_{ω}, ℓ) as

$$\hat{C}(T_{\omega}) = \sum_{\sigma=1}^{3} \sum_{\vec{e} \in \mathcal{S}_{T_{\omega},\sigma}} \hat{c}_{\sigma}(\vec{e}), \tag{2.2.9}$$

so that $K^{\beta}(T_{\omega}) = \exp\left[\frac{1}{\beta}(\hat{C}(T_{\omega}) + o(1))\right]$. The *stochastic potential* of state ω is the lowest cost of reaching it, i.e.

$$\gamma(\omega) := \min_{(T_{\omega},\ell) \in \mathcal{T}_{\omega}} \hat{C}(T_{\omega}). \tag{2.2.10}$$

[15] Derived cost functions will be used in this paper only for the link creation process. In the action revision process one would also need a derived cost function to account for the unlikelihood of a transition when one would apply the learning model of Alós-Ferrer and Netzer (2007) (see their concept of the waste of a labeled revision tree).

We are now ready to present a fairly general result characterizing the low-noise behavior of the invariant distribution (see also Catoni, 1999, Beggs, 2005).

Proposition 2.2.1. *Consider a co-evolutionary model with noise $\mathcal{M}^\beta$ with derived cost functions $\hat{c} = (\hat{c}_1, \hat{c}_2, \hat{c}_3)$ and invariant distribution μ^β. Let $\gamma : \Omega \to \mathbb{R}_+$ be the potential function defined in eq. (2.2.10). For all $\omega \in \Omega$ we have*

$$-\lim_{\beta \to 0} \beta \log \mu^\beta(\omega) = \gamma(\omega) - \min_{\omega' \in \Omega} \gamma(\omega'). \tag{2.2.11}$$

Before proving this proposition we need some additional facts. Order the factors in the invariant measure ρ^β according to their leading terms as $\beta \to 0$. This leads to the low-noise expression

$$\begin{aligned}
\rho^\beta(\omega) &= \sum_{(T_\omega, \ell) \in \mathcal{T}_\omega} \exp\left[\frac{1}{\beta}\left(\hat{C}(T_\omega) + o(1)\right)\right] \\
&= B_\omega \exp(-\gamma(\omega)/\beta)(1 + o(1))
\end{aligned}$$

For sufficiently small β, the invariant distribution can therefore be written as

$$\mu^\beta(\omega) = \frac{B_\omega \exp(-\gamma(\omega)/\beta)(1 + o(1))}{\displaystyle\sum_{\omega' \in \Omega} B_{\omega'} \exp(-\gamma(\omega')/\beta)(1 + o(1))}. \tag{2.2.12}$$

The following simple fact is a useful intermediate result.

Lemma 2.2.1. *Given two finite sequences $(f(1), \ldots, f(n))$, $(B_1, \ldots, B_n)$ of non-negative real numbers, then*

$$\lim_{\beta \to 0} \frac{\log\left(\displaystyle\sum_{i=1}^n B_i \exp(-f(i)/\beta)\right)}{\displaystyle\max_{i=1}^n \log(B_i \exp(-f(i)/\beta))} = 1.$$

Proof. Without loss of generality, let $f(n) = \min_{i=1}^n f(i)$. By absorbing states with equal values of $f(i)$ in the constant B_i we can, without loss of generality, assume that all values are different. The denominator is thus $\log(B_n \exp(-f(n)/\beta))$. Write the polynomial inside $\log(\cdot)$ in the numerator by collecting the terms of highest order, i.e.

$$\begin{aligned}
\sum_{i=1}^n B_i \exp(-f(i)/\beta) &= B_n \exp(-f(n)/\beta)\left(1 + \sum_{i=1}^{n-1} \frac{B_i}{B_n} \exp(-(f(i) - f(n))/\beta)\right) \\
&= B_n \exp(-f(n)/\beta) r(\beta)
\end{aligned}$$

and $\beta \log r(\beta) \to 0$ as $\beta \to 0$. Hence, the ratio can be written as

$$\frac{\beta \log B_n + \beta \log r(\beta) - f(n)}{\beta \log B_n - f(n)} \to 1, \text{as } \beta \to 0.$$

$\square$

Proof of Proposition 2.2.1. Start from eq. (2.2.12). Take logarithms and multiply both sides by $-\beta$ to arrive at

$$\begin{aligned} -\beta \log \mu^\beta(\omega) &= -\beta \log(B_\omega \exp(-\gamma(\omega)/\beta)) \\ &\quad + \beta \log \left(\sum_{\omega' \in \Omega} B_{\omega'} \exp(-\gamma(\omega')/\beta) \right) + O(\beta). \end{aligned}$$

The claim now follows from Lemma 2.2.1. $\square$

This shows that a state is stochastically stable according to Definition 2.2.1 iff it is a state with minimal stochastic potential.

Corollary 2.2.1. $\Omega^* = \{\omega \in \Omega | \gamma(\omega) = \min_{\omega' \in \Omega} \gamma(\omega')\}$.

We see that the main difference between a co-evolutionary model with noise and a classical evolutionary model is the addition of two further cost functions, corresponding to the two added processes modeling the evolution of the network. Departing from here it is easy to see that all well-known results on stochastic stability are applicable. Referring to section 2.7 for proofs of these facts, we just introduce some concepts in order to fix the notation.[16] Let $\mathcal{X}, \mathcal{X}'$ be non-empty subsets of Ω. A *path* from $\mathcal{X}$ to $\mathcal{X}'$ is a directed graph whose vertex set is a non-repeating sequence of states $\{\omega_1, \dots, \omega_l\}$ such that $\omega_1 \in \mathcal{X}, \omega_l \in \mathcal{X}'$ and $\omega_t \notin \mathcal{X}'$ for all $t = 2, \dots, l-1$. The set of all $\mathcal{X}, \mathcal{X}'$ paths is denoted as $\mathcal{P}(\mathcal{X}, \mathcal{X}')$. For two states $\omega \in \mathcal{X}, \omega' \in \mathcal{X}'$ call a path $P \in \mathcal{P}(\mathcal{X}, \mathcal{X}')$ a (ω, ω')-*revision path* if $\omega_1 = \omega$ and $\omega_l = \omega'$, and whose edges are the transitions $(\omega_i, \omega_{i+1}), 1 \leq i \leq l-1$. Denote by $\mathcal{P}_{\omega,\omega'}(\mathcal{X}, \mathcal{X}')$ the set of (ω, ω')-revision paths and P one particular path. To each ω, ω'-path there corresponds a labeling function ℓ, as in Definition 2.2.3. For $P \in \mathcal{P}_{\omega,\omega'}(\mathcal{X}, \mathcal{X}')$ let (P, ℓ) denote a (ω, ω')-revision path. Let $\mathcal{L}, \mathcal{L}' \in \Re$, and denote the cost of transition from recurrent class $\mathcal{L}$ to $\mathcal{L}'$ by $C(\mathcal{L}, \mathcal{L}')$. The cost of a transition from recurrent class $\mathcal{L}$ to $\mathcal{L}'$ is

$$C(\mathcal{L}, \mathcal{L}') = \min_{\omega \in \mathcal{L}} \min_{\omega' \in \mathcal{L}'} \min_{(P,\ell): P \in \mathcal{P}_{\omega,\omega'}(\mathcal{L}, \mathcal{L}')} \hat{C}(P), \tag{2.2.13}$$

[16] See also Samuelson (1997), Young (1998) or Sandholm (2009b) for textbook treatments of this, or the papers by Young (1993) and Ellison (2000).

where $\hat{C}(P)$ is defined as in (2.2.9), applied to a (ω, ω')-revision path. In section 2.7 we argue that all states within one recurrent class are connected by a null cost path. This allows one to study revision graphs between recurrent classes. Therefore, we introduce the class of revision graphs $\hat{T} = (\{\mathcal{L}_1, \dots, \mathcal{L}_k\}, \vec{E})$, where $\vec{E}(\hat{T}) \subseteq \{\mathcal{L}_1, \dots, \mathcal{L}_k\}^2$. A $\mathcal{L}$-revision tree $\hat{T} \in \hat{\mathcal{T}}(\mathcal{L})$ is a revision graph in the sense of Definition 2.2.3. The costs of such a revision tree are defined as $C(\hat{T}) = \sum_{\vec{e} \in \vec{E}(\hat{T})} C(\vec{e})$, with $\vec{e} = (\mathcal{L}', \mathcal{L}'')$. Letting $\hat{\gamma} : \Re \to \mathbb{R}_+$ be a potential function on the set of recurrent classes, one can show (see section 2.7) that for all $\omega \in \mathcal{L}$

$$\gamma(\omega) = \hat{\gamma}(\mathcal{L}) = \min_{\hat{T} \in \hat{\mathcal{T}}(\mathcal{L})} C(\hat{T}). \tag{2.2.14}$$

2.3 Applications

In this section we apply the above general framework to some recent models. In both models we consider the base game $\Gamma = (\mathcal{I}, \{a_1, a_2\}, u)$, with normal form

	a_1	a_2
a_1	$(e-\phi, e-\phi)$	$(f-\phi, g-\phi)$
a_2	$(g-\phi, f-\phi)$	$(h-\phi, h-\phi)$

(2.3.1)

Assume that $h > e > f > g$ but $e + f > h + g$. This means that (a_1, a_1) is a (strictly) risk-dominant Nash equilibrium, while (a_2, a_2) is a Pareto efficient strict Nash equilibrium. The number $\phi \geq 0$ is a fee two incident players have to pay in order to play the game. It does not alter the nature of the game, but possibly affects the way how players form their social network. There is also a mixed strategy equilibrium where a_1 is played with probability $x = \frac{h-f}{e-g+h-f} < 1/2$.

2.3.1 A volatility model

In chapter 4 a volatility model for general potential games is presented. Here we study a version of this model in the context of the symmetric coordination game (2.3.1) with $\phi = 0$. The co-evolutionary model with noise $\mathcal{M}^\beta$, which is essentially a version of the model presented in chapter 4, is the following;

Action adjustment: Assume that

$$b^{i,\beta}(a_\sigma|\omega) = \frac{\exp(\pi^i(\alpha_i^{a_\sigma}, g)/\beta)}{\sum_{r=1}^2 \exp(\pi^i(\alpha_i^{a_r}, g)/\beta)}, \quad \sigma = 1, 2. \tag{2.3.2}$$

This behavioral rule satisfies (A2) with cost function

$$\hat{c}_1(\omega, (\alpha_i^a, g)) = c_1^i(\omega, (\alpha_1^a, g)) = \max_{a' \in \mathcal{A}} \pi^i(\alpha_i^{a'}, g) - \pi^i(\alpha_i^a, g). \tag{2.3.3}$$

Link creation: Assume that $\lambda^i(\omega) = \lambda \mathbb{1}_{\{\kappa^i(\omega) < N-1\}}$, so that every incompletely connected player receives a link creation opportunity with rate $\lambda \geq 0$. Conditional on this event player i samples player j with probability

$$w_j^{i,\beta}(\omega) = \frac{\exp(u(\alpha^i, \alpha^j)/\beta)}{\sum_{k \notin \bar{\mathcal{N}}^i(\omega)} \exp(u(\alpha^i, \alpha^k)/\beta)}.$$

(L4) is satisfied with cost function

$$c_2^i(\omega, (\alpha, g \oplus (i,j)) = \max_{k \notin \bar{\mathcal{N}}^i(\omega)} u(\alpha^i, \alpha^k) - u(\alpha^i, \alpha^j).$$

Link destruction: Once a link is selected by the process (an event with rate ξ) it becomes destroyed at rate 1. Hence $v_{ij}^{\beta}(\omega) = \frac{g_{ij}}{e(\omega)}$. (D3) is satisfied with

$$\hat{c}_3(\omega, (\alpha, g \ominus (i,j))) = c_3^{(i,j)}(\omega, (\alpha, g \ominus (i,j))) \equiv 0.$$

The volume of this subprocess is given by $\bar{\xi} = \xi e(\omega)$, so that $f(\omega, \mathbf{V}^{\beta}) = e(\omega)$.[17]

It remains to determine the derived cost function $\hat{c}_2$. When a link becomes created, a pair of players (i,j) is involved with $j > i$. Suppose this event is on the ω-revision tree (T_ω, ℓ), and call the edge of transition corresponding to this event $\vec{e}$. The labeling function returns the pair of players $\ell(\vec{e}) = (i,j)$. Let $\ell(\vec{e})^-$ be the player with the lower index involved in the transition $\vec{e}$, i.e. i, and $\ell(\vec{e})^+$ the player with the higher index, i.e. j.[18]

Lemma 2.3.1. *For every $\omega \in \Omega$ and $(T_\omega, \ell) \in \mathcal{T}_\omega$, the derived cost of a transition $\vec{e} \in \mathcal{S}_{T_\omega,2}$ is*

$$\hat{c}_2(\vec{e}) = \min\{c_2^{\ell(\vec{e})^-}(\vec{e}), c_2^{\ell(\vec{e})^+}(\vec{e})\}. \tag{2.3.4}$$

[17]The paper to be presented in chapter 5 models a situation where the agents may have idiosyncratic preferences over the actions, which is interpreted as the "type" of the agent. Link decay probabilities are then functions of the types of the involved players. Particularly, it is assumed that $v_{ij}^{\beta}(\omega) = \hat{\xi}_{ij}^{\beta} / \sum_{k>l} \hat{\xi}_{kl}^{\beta} g_{kl}$, for given functions $\{\hat{\xi}_{ij}^{\beta}\}$, which depend on the realized types of the agents and on $\beta > 0$. The corresponding volume is now $\bar{\xi}(\omega) = \sum_{j>i} \hat{\xi}_{ij}^{\beta} g_{ij}$.

[18]I thank Stefano DeMichelis for giving me the right hint for the proof of the following Lemma.

Proof. The probability that edge (i,j) becomes created is

$$\bar{w}_{ij}^{\beta} = \frac{\lambda}{\bar{\lambda}(\omega)}(w_j^{i,\beta}(\omega) + w_i^{j,\beta}(\omega)).$$

By the large deviation principle (L4), for small β we have

$$\begin{aligned} w_j^{i,\beta}(\omega) + w_i^{j,\beta}(\omega) = \exp&\left[-\frac{1}{\beta}(c_2^i(\omega,(\alpha, g \oplus (i,j))) + o(1))\right] \\ + \exp&\left[-\frac{1}{\beta}(c_2^j(\omega,(\alpha, g \oplus (i,j))) + o(1))\right], \end{aligned}$$

and so we can apply Lemma 2.2.1, which gives us the desired result. □

Thus, for every $\omega \in \Omega$ the cost of a revision tree $(T_\omega, \ell) \in \mathcal{T}_\omega$ is $\hat{C}(T_\omega) = \sum_{\sigma=1}^{2} \sum_{\vec{e} \in \mathcal{S}_{T_\omega,\sigma}} \hat{c}_\sigma(\vec{e})$.

Recurrent classes and stochastic stability

Define the set

$$\tilde{\Omega} = \{\omega \in \Omega | g_{ij} = 1 \Rightarrow \alpha^i = \alpha^j\}.$$

A network in this set has only edges between two coordinated players. It may have several connected components and, in particular, it may not be completely connected. Distinguished classes of states in $\tilde{\Omega}$ are the *global conformity sets*

$$\mathcal{L}_\sigma = \{\omega \in \tilde{\Omega} | (\forall i \in \mathcal{I}) : \alpha^i = a_\sigma\},\ \sigma = 1, 2.$$

Let $\mathcal{L}_{1,2} = \tilde{\Omega} \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)$, the *co-existence set*. The following Lemma characterizes the recurrent classes $\Re$ of the unperturbed process.

Lemma 2.3.2. *Consider the unperturbed co-evolutionary model $\mathcal{M}$ of chapter 4. We have $\Re = \tilde{\Omega}$.*

Proof. The proof proceeds by a fairly general constructive argument, which is presented in section 2.6. □

We see that the process allows for *global heterogeneity*, since there may be multiple connected components displaying different types of conventions. However, within every connected component we must have *local conformity*. Due to this large number of equilibria, we hope that the concept of stochastic stability gives us some hint which states are more likely to be observed in the long-run. The following proposition shows that this is not the case.

Proposition 2.3.1. *Consider the coordination game (2.3.1) with $\phi = 0$, and the co-evolutionary model with noise $\mathcal{M}^\beta$ of chapter 4. We have $\Omega^* = \Re$.*

Proof. Fix $\omega \in \mathcal{L}_1, \omega' \in \mathcal{L}_2$. We will construct a zero cost path $P \in \mathcal{P}_{\omega,\omega'}(\mathcal{L}_1, \mathcal{L}_2)$, which implies $C(\mathcal{L}_1, \mathcal{L}_2) = 0$. A symmetric argument shows that $C(\mathcal{L}_2, \mathcal{L}_1) = 0$, so that $\gamma(\omega) = \gamma(\omega') = 0$. From this it follows that $\gamma(\omega'') = 0$ for all $\omega'' \in \mathcal{L}_{1,2}$, since all paths from ω to ω' must pass through some state $\omega'' \in \mathcal{L}_{1,2}$.

Step 1: From ω apply a sequence of link destruction events. All this has zero costs and in finitely many steps we arrive at state $\hat{\omega} \in \mathcal{L}_1$ with the empty network.

Step 2: Give two randomly chosen players sequentially an action adjustment opportunity where they switch to a_2. This has zero costs, since a loner selects both actions with equal probability.

Step 3: Give one of the two players a link creation opportunity. Under **K** a link between them will be established. We are now at a state in $\mathcal{L}_{1,2}$.

Step 4: Give the remaining players action adjustment opportunities where they switch to a_2, and then a link creation opportunity. Iterate this until the process arrives at the desired state $\omega' \in \mathcal{L}_2$.

Steps 1-4 defines a path from ω to ω' having zero costs. Clearly all steps are reversible, i.e. steps 4-1 define a path from ω' to ω having zero costs. This demonstrates $\gamma(\omega) = \gamma(\omega') = 0$. □

2.3.2 A classical model

We discuss a slight variation of Jackson and Watts (2002). To get the most interesting scenario, we reduce the set of admissible parameters in requiring that $x > \frac{1}{N-1}$ and $\phi \in (g, e)$. Jackson and Watts (2002) take the mistakes model of Kandori et al. (1993) and Young (1993) as universal behavioral rule. Parameterizing noise as $\varepsilon = \exp(-1/\beta)$, and henceforth calling ε the noise parameter, allows us to study this behavioral rule. The co-evolutionary model with noise $\mathcal{M}^\varepsilon$ is as follows;

Action adjustment: Assume that each player receives with uniform probability $1/N$ the opportunity to change his action. Conditional on this event he selects action $a \in \mathcal{A}$ with probability

$$b^{i,\varepsilon}(a|\omega) = \begin{cases} 1 - \frac{\varepsilon}{2} & \text{if } \alpha^i \neq a \text{ and } \{a\} = \arg\max_{a' \in \mathcal{A}} \pi^i(\alpha_i^{a'}, g), \\ 1 - \frac{\varepsilon}{2} & \text{if } \alpha^i = a \text{ and } \{\alpha^i\} = \arg\max_{a' \in \mathcal{A}} \pi^i(\alpha_i^{a'}, g), \\ \frac{\varepsilon}{2} & \text{otherwise.} \end{cases}$$

This behavioral rule says that a player abandons his currently used action with relatively high probability, if there exists a strictly better action. Otherwise he sticks to his action and switches only with the relatively small probability ε. This behavioral rules satisfies condition (A2) with cost function

$$\hat{c}_1(\omega,(\alpha_i^a,g)) = c_1^i(\omega,(\alpha_i^a,g)) = \begin{cases} 0 & \text{if } \alpha^i \neq a \text{ and } \{a\} = \arg\max_{a'\in\mathcal{A}} \pi^i(\alpha_i^{a'},g), \\ 0 & \text{if } \alpha^i = a \text{ and } \{\alpha^i\} = \arg\max_{a'\in\mathcal{A}} \pi^i(\alpha_i^{a'},g), \\ 1 & \text{otherwise.} \end{cases}$$

Link creation: Jackson and Watts (2002) introduce a cooperative element into the link creation process. To capture this, we have to make a slight modification in the construction of our co-evolutionary model with noise. Let $\mathcal{D}(\omega)$ denote the set of neutral links at ω and $d(\omega)$ its cardinality. Instead of the individual players' rate functions, assume that the event of link creation arrives to the society at the constant rate $\bar{\lambda}(\omega) := \lambda d(\omega)$, where λ is a positive constant. Define the events

$$(1 \leq i,j \leq N) : A_j^i(\phi) := \{\omega \in \Omega | u(\alpha^i,\alpha^j) \geq \phi\}.$$

If $\omega \in A_j^i(\phi)$ then the edge (i,j) is profitable from the point of view of player i at ω. The number of mutually profitable neutral links is

$$m(\omega) = \sum_{(i,j)\in\mathcal{D}(\omega)} \mathbb{1}_{A_j^i(\phi)\cap A_i^j(\phi)}(\omega).$$

Following the spirit of pairwise stability (Jackson and Wolinsky, 1996), assume that a neutral link is set to be active with probability $1-\varepsilon$ if both players mutually agree. With the small probability ε assume that all links have a chance to be formed. The (conditional) probability that a neutral edge (i,j) will be added is

$$(\forall (i,j) \in \mathcal{D}(\omega)) : \bar{w}_{i,j}^{\varepsilon}(\omega) := \begin{cases} \frac{1-\varepsilon}{m(\omega)} + \frac{\varepsilon}{d(\omega)} & \text{if } \mathbb{1}_{A_j^i(\phi)\cap A_i^j(\phi)}(\omega) = 1, \\ \frac{\varepsilon}{d(\omega)} & \text{otherwise.} \end{cases} \tag{2.3.5}$$

The term $\varepsilon/d(\omega)$ is the "background noise" of the system, and gives the uniform probability that a link will be formed. If edge (i,j) is neutral at ω, but both players are not hurt by the creation of the link, then they will independently agree to form it with the high probability $1-\varepsilon$, which increases their chance of being formed.[19] Let $\bar{A}$ denote the complementary set of A. The cost function of this sub-process is

$$\hat{c}_2(\omega,(\alpha, g \oplus (i,j))) = \mathbb{1}_{\bar{A}_j^i(\phi)\cup\bar{A}_i^j(\phi)}(\omega).$$

[19] Jackson and Watts (2002) assume that a link is created with probability $1-\varepsilon$ iff it is a

Link destruction: With rate $\xi > 0$ links become destroyed. Conditional on this event, pick one edge $(i,j) \in \mathcal{E}(\omega)$ with uniform probability, and allow the incident players to re-evaluate the benefits arising from this connection. This leads to $\bar{\xi}(\omega) := \xi e(\omega)$. Denote by

$$\bar{m}(\omega) = \sum_{(i,j)\in\mathcal{E}(\omega)} \mathbb{1}_{\bar{A}^i_j(\phi)\cup\bar{A}^j_i(\phi)}(\omega)$$

the number of active links where at least one player benefits from the deletion of the link. If (i,j) is a link where at least one player is better off after its destruction, suppose that with large probability $1-\varepsilon$ it will be destroyed. With the small probability ε every link can be destroyed once it has been selected. This leads to the following version of link destruction probabilities

$$(\forall (i,j) \in \mathcal{E}(\omega)) : v^{\varepsilon}_{i,j}(\omega) = \begin{cases} \frac{1-\varepsilon}{\bar{m}(\omega)} + \frac{\varepsilon}{e(\omega)} & \text{if } \mathbb{1}_{\bar{A}^i_j(\phi)\cup\bar{A}^j_i(\phi)}(\omega) = 1, \\ \frac{\varepsilon}{e(\omega)} & \text{otherwise.} \end{cases}$$

The cost function of this process is given by

$$\hat{c}_3(\omega, (\alpha, g \ominus (i,j))) = \mathbb{1}_{A^i_j(\phi)\cap A^j_i(\phi)}(\omega).$$

Recurrent classes and stochastically stable states

Define $\mathcal{I}_1(\omega) = \{i \in \mathcal{I} | \alpha^i = a_1 \text{ on } \omega\}$, and for every $2 \le n \le N-2$,

$$\mathcal{L}^n_{1,2} = \{\omega \in \Omega | g = g^c[\mathcal{I}_1(\omega)] \oplus g^c[\mathcal{I}_2(\omega)] \;\&\; |\mathcal{I}_1(\omega)| = n\},$$
$$\mathcal{L}_{1,2} = \bigcup_{n=2}^{N-2} \mathcal{L}^n_{1,2}.$$

Let

$$\mathcal{L}_\sigma = \{\omega \in \Omega | (\forall i \in \mathcal{I}) : \alpha^i = a_\sigma \;\&\; g = g^c[\mathcal{I}]\},\ \sigma = 1, 2.$$

Lemma 2.3.3. *Let $\mathcal{M}$ be the unperturbed co-evolutionary process of Jackson and Watts (2002) with $\phi \in (g, e)$. Then*

$$\Re = \mathcal{L}_1 \cup \mathcal{L}_2 \cup \mathcal{L}_{1,2}.$$

strict Pareto improvement, i.e. at least one player is strictly better off after the connection has been established and no party is hurt by the creation of the link. We assume that a link is already formed if it is a weak Pareto improvement, i.e. no party is hurt by the formation of the link. Additionally, Jackson and Watts (2002) assume that the error probability ε is not the same in the action evolution process as it is in the graph evolution process. However, it is required that the error probabilities go to zero at the same rate so not to get "twisted" equilibrium selection results, as argued by Bergin and Lipman (1996).

Proof. The algorithm in section 2.6 shows that in finite time there are no links between agents playing different actions. Call ω_m the state at which the algorithm stops. In the unperturbed model, with probability 1, only links which are mutually profitable are formed and links which harm at least one player are destroyed. At ω_m no player has an incentive to change his action. If a link creation event takes place, with conditional probability 1 only an edge is formed if the selected pair is in the same action class $\mathcal{I}_\sigma(\omega_m), \sigma = 1, 2$. Moreover, these links never become destroyed. A link destruction event at ω_m leaves the state invariant with probability 1, since at this state no edges between players from different action classes exist. For the same reason, an action adjustment event leaves ω_m invariant, with probability 1. If $\mathcal{I}_\sigma(\omega_m) = \emptyset$ for a $\sigma = 1, 2$, then the process arrives at a state where global conformity prevails. Otherwise the process leads to a state in the co-existence set $\mathcal{L}_{1,2}$. $\square$

To select among the recurrent sets, we now perform an analysis via stochastic stability.

Proposition 2.3.2. *Let $\mathcal{M}^\varepsilon$ be the co-evolutionary model with noise of Jackson and Watts (2002). Then for $\phi \in (g, e)$ and $x \geq \frac{1}{N-1}$*

$$\Omega^* = \mathcal{L}_1 \cup \mathcal{L}_2.$$

Proof. We will explicitly calculate the potentials of the three recurrent classes and show that, under the stated parameter assumptions, $\hat{\gamma}(\mathcal{L}_1) = \hat{\gamma}(\mathcal{L}_2) < \hat{\gamma}(\mathcal{L}^n_{1,2})$ for all $2 \leq n \leq N-2$.
Let $\omega \in \mathcal{L}_1, \omega' \in \mathcal{L}_2$. Observe that, under the assumption $x \geq \frac{1}{N-1}$, at least two agents must change their action to enter $\mathcal{L}_{1,2}$ via an unperturbed move, i.e. $C(\mathcal{L}_\sigma, \mathcal{L}_{1,2}) = 2, \sigma = 1, 2$.[20] Further observe that for $3 \leq j \leq N-3, C(\mathcal{L}^j_{1,2}, \mathcal{L}^{j\pm1}_{1,2}) = 1$, since a single deviator reduces/increases the set of a_1 players, and applying then link creation/destruction leads to some state in the desired recurrent class. Moreover, the sets $\mathcal{L}^1_{1,2}$ and $\mathcal{L}^{N-1}_{1,2}$ where one loner plays a_1, respectively a_2, are transient since the loner can switch to the different action at zero cost and then we apply the link formation process

[20] To see this, note that $C(\mathcal{L}_1, \mathcal{L}^2_{1,2}) = 2$ by definition of risk dominance. To get $C(\mathcal{L}_2, \mathcal{L}^{N-2}_{1,2}) = 2$ suppose that one player deviates from $\omega' \in \mathcal{L}_2$ and plays a_1. The network remains unchanged. Apply the action adjustment process in the next period to a current a_2 player. This player will switch to a_1 iff $e+(N-2)f-\phi(N-1) > g+(N-2)h+\phi(N-1)$, or iff $x < 1/(N-1)$. Since we assume that $x \geq 1/(N-1)$ another tremble is needed to make a_1 a best response.

which leads directly to a state in $\mathcal{L}_1$, respectively $\mathcal{L}_2$. It follows that

$$\begin{aligned}
\hat{\gamma}(\mathcal{L}_1) &= C(\mathcal{L}_2, \mathcal{L}_{1,2}^2) + \sum_{j=2}^{N-3} C(\mathcal{L}_{1,2}^j, \mathcal{L}_{1,2}^{j+1}) + C(\mathcal{L}_{1,2}^{N-2}, \mathcal{L}_1) \\
&= 2 + N - 4 + 1 = N - 1 \\
\hat{\gamma}(\mathcal{L}_2) &= C(\mathcal{L}_1, \mathcal{L}_{1,2}^{N-2}) + \sum_{j=2}^{N-3} C(\mathcal{L}_{1,2}^{j+1}, \mathcal{L}_{1,2}^j) + C(\mathcal{L}_{1,2}^2, \mathcal{L}_2) \\
&= N - 1 \\
\hat{\gamma}(\mathcal{L}_{1,2}^n) &= C(\mathcal{L}_1, \mathcal{L}_{1,2}^2) + C(\mathcal{L}_2, \mathcal{L}_{1,2}^{N-2}) \\
&\quad + \sum_{j=2}^{n-1} C(\mathcal{L}_{1,2}^j, \mathcal{L}_{1,2}^{j+1}) + \sum_{j=2}^{N-n-1} C(\mathcal{L}_{1,2}^{N-j}, \mathcal{L}_{1,2}^{N-j-1}) \\
&= 2 + 2 + (n - 1 - 2 + 1) + (N - n - 1 - 2 + 1) = N
\end{aligned}$$

□

This is Case (iii) of Proposition 1 of Jackson and Watts (2002), which they point out to be the most interesting scenario. It emphasizes the role of the network in the equilibrium selection procedure. In a model with fixed interaction structure one of the robust results of stochastic evolutionary game dynamics is the selection of the risk dominant equilibrium at every connected component. In calculating the transition costs above we have implicitly made use of the flexibility of the interaction structure. Since a change in the action of a player may also cause a change in the network structure, transitions between the strict Nash equilibria of the coordination game become possible. However, we also observe that the "classical" model of Jackson and Watts (2002) is sensitive to a refinement made by stochastic stability, whereas in the "volatility" model we were not able to discriminate among the recurrent classes by applying the stochastic stability criterion. Whether this is a drawback of the volatility model or not must be judged by the application at hand.

2.4 A micro-founded model for inhomogeneous random graphs.

The theory of random graphs provides in essence 2 classes of models; the "randomly grown graphs", mostly using a version of preferential attachment (Barabási and Albert, 1999), and generalized random graphs (Newman, 2003).

Under some additional assumptions on the structure of the Markov chain (2.2.6), we are able to characterize the induced ensemble of random graphs for general behavioral rules.
Let us add the following two assumptions on the structure of the transition probabilities:

(L4) $(\forall i \in \mathcal{I}) : \lambda^i(\omega) = \lambda \mathbb{1}_{\{\kappa^i(\omega)<N-1\}}$;

(L5) $(\forall i.j \in \mathcal{I}) : w_j^{i,\beta}(\omega) = \hat{w}_j^{i,\beta}(\alpha)(1-g_{ij})$, where $\hat{w}_j^{i,\beta}(\cdot)$ satisfies (L1)-(L3).

(L4) defines the volume of the link creation subprocess as $\bar{\lambda}(\omega) = \lambda \sum_{i\in\mathcal{I}} \mathbb{1}_{\{\kappa^i(\omega)<N-1\}}$.
In the link destruction process we demand additionally

(D4) $(\forall i, j \in \mathcal{I}) : v_{ij}^{\beta}(\omega) = \frac{\hat{v}_{ij}^{\beta}(\alpha)g_{ij}}{f(\omega,\mathbf{V}^{\beta})}$.

(D2) tells us that $f(\omega, \mathbf{V}^{\beta}) = \sum_{j>i} \hat{v}_{ij}^{\beta}(\alpha)g_{ij}$.
Using these additional assumptions we will derive a *random graph process*, modeling the evolution of the network for a fixed action profile α.[21] Let $(\tilde{G}_n^{\beta})_{n=0}^{\infty}$ denote the random graph process with transition probabilities $K_{2,3}^{\beta} : \mathcal{G}[\mathcal{I}] \times \mathcal{G}[\mathcal{I}] \to [0,1]$, defined as

$$\begin{aligned}
K_{2,3}^{\beta}(g,g') &= \mathbb{P}(\tilde{G}_{n+1}^{\beta} = g'|\tilde{G}_n^{\beta} = g) \\
&= \mathbb{P}(X_{n+1}^{\beta} = (\alpha, g')|X_n^{\beta} = (\alpha, g), \text{Network evolution}) \\
&= \frac{1}{\mathbb{P}(\text{Network evolution}|X_n^{\beta} = (\alpha, g))}\mathbb{P}(X_{n+1}^{\beta} = (\alpha, g')|X_n^{\beta} = (\alpha, g)) \\
&= \frac{1}{q_2(\alpha, g) + q_3(\alpha, g)}K^{\beta}((\alpha, g), (\alpha, g'))
\end{aligned}$$

Using (2.2.5), (2.2.6) and (L4), (L5), (D4), the transition probabilities are given by

$$K_{2,3}^{\beta}(g,g') = \frac{1}{q_2(\alpha,g)+q_3(\alpha,g)}\begin{cases} \lambda(\hat{w}_j^{i,\beta}(\alpha) + \hat{w}_i^{j,\beta}(\alpha)) & \text{if } g' = g \oplus (i,j), \\ \xi \hat{v}_{ij}^{\beta}(\alpha) & \text{if } g' = g \ominus (i,j), \\ 0 & \text{otherwise.} \end{cases}$$

This chain is irreducible but no longer aperiodic. It serves as a jump chain of the *continuous-time* random graph process[22] $(\tilde{G}^{\beta}(t))_{t\geq 0}$ with generator

$$\eta_{2,3}^{\beta}(g \to g') = (q_2(\alpha, g) + q_3(\alpha, g))(K_{2,3}^{\beta}(g, g') - \delta_{g,g'}) \tag{2.4.1}$$

[21] An interpretation of such a process can be given by assuming that action adjustment is a relatively fast process compared to network evolution. In this case, it makes sense to assume that the profile α reaches a temporary stationary state for a given network g, and when evolution shapes the network the profile α is fixed.

[22] See chapter 3 for a thorough explanation of these terms.

where $\delta_{g,g'} = 1$ iff $g = g'$, and 0 otherwise. This continuous time process allows us to identify the invariant distribution of the original Markov chain in a simple way. Let $\mathbf{Id}$ denote the identity matrix on $\mathcal{G}[\mathcal{I}]$, and define the matrix $\boldsymbol{\eta}^{\beta}_{2,3} := \left[\eta^{\beta}_{2,3}(g \to g')\right]_{g,g' \in \mathcal{G}[\mathcal{I}]}$. Additionally, call $\hat{q}(g) := q_2(\alpha, g) + q_3(\alpha, g)$, and $\hat{\boldsymbol{q}} := [\hat{q}(g)]_{g \in \mathcal{G}[\mathcal{I}]}$. The generator of the continuous-time process $(\tilde{G}^{\beta}(t))_{t \geq 0}$ is defined by $\boldsymbol{\eta}^{\beta}_{2,3} = \hat{\boldsymbol{q}}(\mathbf{K}^{\beta}_{2,3} - \mathbf{Id})$. A measure ν on $\mathcal{G}[\mathcal{I}]$ is said to be invariant with respect to the generator $\boldsymbol{\eta}^{\beta}_{2,3}$ if $\nu \boldsymbol{\eta}^{\beta}_{2,3} = \mathbf{0}$.

Lemma 2.4.1. *The following are equivalent:*

(a) ν *is invariant under* $\boldsymbol{\eta}^{\beta}_{2,3}$,

(b) $\mu \mathbf{K}^{\beta}_{2,3} = \mu$ *where* $\mu(g) = \nu(g)\hat{q}(g)$.

Proof. Define the measure $\mu(g) := \nu(g)\hat{q}(g)$ for all $g \in \mathcal{G}[\mathcal{I}]$. For all g, g' we have $\eta^{\beta}_{2,3}(g \to g') = \hat{q}(g)(K^{\beta}_{2,3}(g, g') - \delta_{g,g'})$. Thus,

$$\begin{aligned}\sum_{g \in \mathcal{G}[\mathcal{I}]} \mu(g)(K^{\beta}_{2,3}(g, g') - \delta_{g,g'}) &= \sum_{g \in \mathcal{G}[\mathcal{I}]} \nu(g)\hat{q}(g)(K^{\beta}_{2,3}(g, g') - \delta_{g,g'}) \\ &= \sum_{g \in \mathcal{G}[\mathcal{I}]} \nu(g)\eta^{\beta}_{2,3}(g \to g').\end{aligned}$$

□

The next proposition characterizes the invariant distribution of the continuous-time random graph process. Its proof is a surprisingly simple calculation, which uses many ideas spelled out in Section 1.2.

Proposition 2.4.1. *Consider the random graph process* $(\tilde{G}^{\beta}(t))_{t \geq 0}$ *with generator (2.4.1). Its unique invariant distribution is the product measure*

$$\hat{\mu}^{\beta}_{2,3}(g|\alpha) = \prod_{i=1}^{N} \prod_{j>i} p^{\beta}_{ij}(\alpha)^{g_{ij}}(1 - p^{\beta}_{ij}(\alpha))^{1-g_{ij}}, \tag{2.4.2}$$

with the edge-success probability

$$(\forall i, j \in \mathcal{I}) : p^{\beta}_{ij}(\alpha) = \frac{\lambda(\hat{w}^{i,\beta}_{j}(\alpha) + \hat{w}^{j,\beta}_{i}(\alpha))}{\lambda(\hat{w}^{i,\beta}_{j}(\alpha) + \hat{w}^{j,\beta}_{i}(\alpha)) + \xi \hat{v}^{\beta}_{ij}(\alpha)}. \tag{2.4.3}$$

Proof. The Markov process $(\tilde{G}^{\beta}(t))_{t \geq 0}$ is irreducible for $\beta > 0$ and reversible by the symmetry assumption (D1). Solving the detailed balance conditions

$$\hat{\mu}^{\beta}_{2}(g|\alpha)\eta^{\beta}_{2,3}(g \to g \oplus (i, j)) = \hat{\mu}^{\beta}_{2,3}(g \oplus (i, j)|\alpha)\eta^{\beta}_{2,3}(g \oplus (i, j) \to g)$$

for all $g \in \mathcal{G}[\mathcal{I}]$ gives us

$$\hat{\mu}^{\beta}_{2,3}(g|\alpha) = \frac{1}{Z^{\beta}_{2,3}(\alpha)} \prod_{i=1}^{N} \prod_{j>i} \left(\frac{\lambda}{\xi} \frac{\hat{w}_j^{i,\beta}(\alpha) + \hat{w}_i^{j,\beta}(\alpha)}{\hat{v}^{\beta}_{ij}(\alpha)} \right)^{g_{ij}}.$$

Let $\bar{\hat{w}}^{\beta}_{ij}(\alpha) := \hat{w}_j^{i,\beta}(\alpha) + \hat{w}_i^{j,\beta}(\alpha)$ and define $\theta^{\beta}_{ij}(\alpha) := \log\left(\frac{\lambda}{\xi}\frac{\bar{\hat{w}}^{\beta}_{ij}(\alpha)}{\hat{v}^{\beta}_{ij}(\alpha)}\right)$. Further, define the Hamiltonian $H^{\beta}(g|\alpha) := \sum_{i,j>i} \theta^{\beta}_{ij}(\alpha) g_{ij}$, so that $\hat{\mu}^{\beta}_{2,3}(g|\alpha) = \frac{\exp(H^{\beta}(g|\alpha))}{\sum_{g' \in \mathcal{G}[\mathcal{I}]} \exp(H^{\beta}(g'|\alpha))}$. The constant of normalization can then be written as

$$Z^{\beta}_{2,3}(\alpha) = \sum_{g' \in \mathcal{G}[\mathcal{I}]} \exp(H^{\beta}(g'|\alpha)) = \sum_{i,j>i} \sum_{g_{ij}=0}^{1} \left(\prod_{i,j>i} \exp(\theta^{\beta}_{ij}(\alpha) g_{ij}) \right) = \prod_{i,j>i} (1 + \exp(\theta^{\beta}_{ij})).$$

The probability that edge (i,j) is active in the long run is

$$\begin{aligned} p^{\beta}_{ij}(\alpha) &= \sum_{g' \in \mathcal{G}[\mathcal{I}]: g'_{ij}=1} \hat{\mu}^{\beta}_{2,3}(g'|\alpha) = \sum_{g' \in \mathcal{G}[\mathcal{I}]} g'_{ij} \hat{\mu}^{\beta}_{2}(g'|\alpha) = \frac{\partial \log Z^{\beta}_{2,3}(\alpha)}{\partial \theta^{\beta}_{ij}(\alpha)} = \frac{\exp(\theta^{\beta}_{ij}(\alpha))}{1 + \exp(\theta^{\beta}_{ij}(\alpha))} \\ &= \frac{\lambda \bar{\hat{w}}^{\beta}_{ij}(\alpha)}{\lambda \bar{\hat{w}}^{\beta}_{ij}(\alpha) + \xi \hat{v}^{\beta}_{ij}(\alpha)}. \end{aligned}$$

Collecting terms and doing some simple manipulations gives the desired result. □

This strong result gives a full characterization of the induced ensemble of random graphs for volatility models such as Marsili et al. (2004), Ehrhardt et al. (2008a;b), and establishes an interesting connection with the random graph processes proposed by Söderberg (2002), Park and Newman (2004) or Bollobás et al. (2007). Any co-evolutionary model with noise, satisfying the set of assumptions (A1)-(A2), (L1)-(L5) and (D1-D4) will generate an inhomogeneous random graph, with edge-success probabilities (2.4.3)

Corollary 2.4.1. *The unique invariant distribution of the discrete-time random graph process* $(G^{\beta}_n)_{n \geq 0}$ *is*

$$\frac{\hat{q}(g)\nu^{\beta}(g|\alpha)}{\sum_{g' \in \mathcal{G}[\mathcal{I}]} \hat{q}(g')\nu^{\beta}(g'|\alpha)}$$

with

$$\nu^{\beta}(g|\alpha) := \prod_{i=1}^{N} \prod_{j>i} \left(\frac{\lambda}{\xi} \frac{\hat{w}_j^{i,\beta}(\alpha) + \hat{w}_i^{j,\beta}(\alpha)}{\hat{v}^{\beta}_{ij}(\alpha)} \right)^{g_{ij}}.$$

Proof. This follows form Lemma 2.4.1. □

2.5 Conclusion

This paper presented a general framework for studying co-evolutionary models with noise. We gave a complete characterization of the invariant distribution of such a model, which is a joint probability distribution on the set of action profiles and the set of networks. By means of two examples, a volatility model akin to Ehrhardt et al. (2008b) and a classical model based on Jackson and Watts (2002), we have shown how the unified approach is useful to make a systematic investigation of co-evolutionary models. Beside presenting a unified formalism to perform the by now important equilibrium selection technique of stochastic stability, we have demonstrated that a co-evolutionary model with noise generates an inhomogeneous random graph ensemble for the long run interaction structure of the population. The main result in this direction provides a closed form solution for the probability measure of this graph ensemble, and presents the general form for edge-success probabilities. Based on this novel insight, many new questions arise.

First, the edge success probabilities depend only on the behavioral rules the agents are assumed to use. It would be interesting to see what differences between networks arise by assuming different behavioral rules. For instance, do best-responding agents tempt to self-organize in more structured and/or efficient network topologies as imitative agents? What role plays the underlying noise structure of the model (meant here as the interplay between behavioral noise β and overall network volatility ξ)? Second, the literature on social and epidemic diffusion (see e.g. Morris, 2000, Alós-Ferrer and Weidenholzer, 2008, Pastor-Satorras and Vespignani, 2001) have emphasized the importance of the network architecture in order to understand the phenomenon of contagion. In particular, notions of network clustering and cohesiveness have turned out to be important. We do not yet know the statistics produced by a co-evolutionary model. Third, in the context of volatility models Ehrhardt et al. (2008a) find three interesting dynamic effects; Resilience, Equilibrium co-existence and phase transitions (i.e. a discontinuous switch in the connectivity of the network by a slight change of the parameters affecting the edge success probability). Under what parameter configurations are these phenomena reproduced in the framework of a co-evolutionary model? The recent work by Bollobás et al. (2007) studies inhomogeneous random graphs and detects also a phase transition in network connectivity by exploring the size of components with a branching process approximation. Future work should continue in this direction in order to explore the fine details of the random graph ensemble.

2.6 Proof of Lemma 2.3.2

We first show that if $\omega \in \Re$, then there is no positive probability path under $\mathbf{K}$ that leads out of this set. Under ω every connected pair of players is coordinated. Let i be a current a_1 player. Every player in the component to which i belongs must then also play a_1.[23] Hence, every graph corresponding to $\omega \in \Omega^*$ must consist of finitely many components, each characterized by behavioral conformity. By definition, applying $\mathbf{K}$ to such a state will not lead to a state outside Ω^*.
Now consider a state $\omega \notin \Re$. To show that such a state is transient under $\mathbf{K}$, we have to find a positive probability path under $\mathbf{K}$ that leads to some state $\omega' \in \Re$. The following algorithm constructs such a path;
Let $\omega_0 = \omega$ be our initial state. The set of uncoordinated edges $\mathcal{E}(\mathcal{I}_1(\omega_0), \mathcal{I}_2(\omega_0)) \neq \emptyset$, by hypothesis.[24] Let $t = 0, 1, 2, \ldots, m$ measure the number of iterations of the algorithm. Start from $t = 0$. The algorithm generates a sequence $\{\omega_t\}_{t=0}^{m}$, where the transition from ω_t to ω_{t+1} proceeds as follows:

Step 1: Pick the first edge from this set. Let one of the two involved players receive an action adjustment opportunity where he switches only to an action that gives him a strictly larger payoff compared to ω_t.[25] If this player changes his action, delete the edge from the list of uncoordinated edges, and call the resulting state ω_{t+1}. Then repeat Step 1. If the player does not change his action, go to Step 2.

Step 2: Give the other player an action adjustment opportunity as in Step 1. If he changes his action, delete the edge from the list of uncoordinated edges and call the resulting state ω_{t+1}. Then repeat Step 1. If the player does not change his action, go to Step 3.

Step 3: Delete the edge by a targeted link destruction event.[26] Call the resulting state ω_{t+1} and note that the set of uncoordinated edges decreased by 1. Go to Step 4.[27]

[23]If j would be a player in the component who plays a_2 he cannot be linked with a player who is path connected with i.
[24] This is the set of links that connect players from $\mathcal{I}_1(\omega)$ to players in $\mathcal{I}_2(\omega)$. $\mathcal{I}_\sigma(\omega)$ is the set of a_σ-players at ω.
[25]In 2×2 games with finite populations this choice rule is generically equivalent to demanding that a play switches to a best-response.
[26]Note that this is always a zero-cost step.
[27]An intermediate stage could be added to the algorithm, where we apply $\mathbf{K}$ to ω_{t+1} by letting the involved players create a link. This will lead to the creation of maximally 2 coordinated links.

Step 4: Order the edges in $\mathcal{E}(\mathcal{I}_1(\omega_{t+1}), \mathcal{I}_2(\omega_{t+1}))$ in some way. If this set is empty, exit the algorithm. Otherwise, go to Step 1.

2.7 The Markov chain tree theorem and set-valued cost functions

To proof (2.2.8) we will make use of some general results from the theory of Markov chains and simulated annealing. Norris (1997) and Grimmett and Stirzaker (2001) are good references for the theory of finite Markov chains, and Catoni (1999; 2001) collects the relevant results from simulated annealing. Let $\omega \in \Omega, x, y, z \in \Omega \setminus \{\omega\}$ and $\mathcal{X} \subset \Omega$ a nonempty set. Denote by $\mathbf{K}^{\beta,n}$ the n-fold Matrix product of $\mathbf{K}^{\beta}$. The interpretation of this operation is that $K^{\beta,n}(x,y) = \mathbb{P}(X_n^\beta = y | X_0^\beta = x)$. Let $\omega \in \Omega$ be an arbitrary fixed state and define its *first passage time* as the random variable

$$\tau(\omega) := \inf\{n \geq 1 | X_n^\beta = \omega\}.$$

Since ω is recurrent we have $\mathbb{P}(\tau(\omega) < \infty | X_0^\beta = z) = 1$ for all z. Hence, the process returns to state ω almost surely, independent from where it takes off. Suppose we start the process from y and want to keep track of the number of times the chain visits x before it returns to ω. Phrased in probabilistic terms this amounts to calculate

$$\mathbb{E}\left(\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n^\beta = x\} \cap \{\tau(\omega) > n\}} | X_0^\beta = y\right) \tag{2.7.1}$$

The graph description of finite Markov chains is useful to calculate this seemingly complicated expression. Recall that a $\mathcal{X}$-graph is an element of the set of graphs $\mathcal{T}(\mathcal{X})$, connecting every point in $\Omega \setminus \mathcal{X}$ to a point in $\mathcal{X}$, without loops. If $\mathcal{X} = \{\omega, x\}$ then $\mathcal{T}(\{\omega, x\})$ contains all graphs which connect points from $\Omega \setminus \{\omega, x\}$ in a unique way either to ω or x. Denote by $\mathcal{T}_{y,x}(\mathcal{X})$ the set of $\mathcal{X}$-graphs which contain a path $\{\omega_1, \ldots, \omega_l\}$ such that $\omega_1 = y, \omega_l = x$ and $\omega_t \notin \mathcal{X}$, for all $2 \leq t \leq l-1$. If $y = x$ set $\mathcal{T}_{x,x}(\mathcal{X}) = \mathcal{T}(\mathcal{X})$. If $y \in \mathcal{X}$ set $\mathcal{T}_{y,x}(\mathcal{X}) = \emptyset$. It is intuitive that (2.7.1) should be proportional to the probability of graphs $T \in \mathcal{T}_{y,x}(\{\omega, x\})$. However, we also require to return to ω, so not all possible paths are allowed. We have to condition on the ω-trees, since these are the paths that lead in a unique way to ω. This heuristic argument suggests that (2.7.1) can be calculated as

$$\frac{\sum_{T \in \mathcal{T}_{y,x}(\{\omega,x\})} K^\beta(T)}{\sum_{T_\omega \in \mathcal{T}_\omega} K^\beta(T_\omega)} = \frac{1}{\rho^\beta(\omega)} \sum_{T \in \mathcal{T}_{y,x}(\{\omega,x\})} K^\beta(T). \tag{2.7.2}$$

Lemma 3.1 of Catoni (1999) gives a rigorous proof of this heuristic.[28]

Lemma 2.7.1 (Lemma 3.1, Catoni (1999)). *Let $\bar{\mathbf{K}}^\beta$ denote the matrix $\mathbf{K}^\beta$ restricted to the set $\Omega \setminus \{\omega\}$. Then*

$$\begin{aligned}\sum_{n=0}^{\infty} \bar{K}^{\beta,n}(y,x) &= \mathbb{E}\left[\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n^\beta = x\} \wedge \{\tau(\omega) > n\}} | X_0^\beta = y\right] \\ &= \frac{\sum_{T \in \mathcal{T}_{y,x}(\{\omega,x\})} K^\beta(T)}{\rho^\beta(\omega)}.\end{aligned}$$

Before proving this result, we need the following simple observation.

Lemma 2.7.2. *For all $y, x \neq \omega$, we have $\lim_{n\to\infty} \bar{K}^{\beta,n}(y,x) = 0$.*

Proof.

$$\lim_{n\to\infty} \bar{K}^{\beta,n}(y,x) \leq \mathbb{P}(\tau(\omega) = \infty | X_0^\beta = y) = 0$$

since ω is a recurrent state. □

This simple fact has the consequence that $(\mathbf{Id} - \bar{\mathbf{K}}^\beta)$ is invertible. This is interesting, because for all $y, x \neq \omega$

$$\begin{aligned}(\mathbf{Id} - \bar{\mathbf{K}}^\beta)^{-1}(y,x) &= \sum_{n=0}^{\infty} \bar{K}^{\beta,n}(y,x) \\ &= \sum_{n=0}^{\infty} \mathbb{E}\left(\mathbb{1}_{\{X_n^\beta = x\} \wedge \{\tau(\omega) > n\}} | X_0^\beta = y\right).\end{aligned}$$

Hence, this gives us the expected number of times the process visits x before hitting ω, which is the quantity we want to compute in Lemma 2.7.1.

Proof of Lemma 2.7.1. For all $y, x \neq \omega$, let us define

$$M(y,x) := \frac{1}{\rho^\beta(\omega)} \sum_{T \in \mathcal{T}_{y,x}(\{\omega,x\})} K^\beta(T).$$

Define the Kronecker-delta function as $\delta_{y,x} = 1$ if $y = x$ and 0 otherwise. We have to show that for all $y, x \neq \omega$

$$\sum_{z \neq \omega} (\delta_{y,z} - \bar{K}^\beta(y,z)) M(z,x) = \delta_{y,x}.$$

[28] The proof, which is taken from Catoni (1999), extends literally to the case where the singleton is replaced by a non-empty subset $\mathcal{X}$.

This can be written as

$$\sum_{z\neq y} K^\beta(y,z)M(y,x) = \delta_{y,x} + \sum_{z\in\Omega\setminus\{\omega,y\}} K^\beta(y,z)M(z,x)$$

$$\Leftrightarrow \sum_{z\neq y} K^\beta(y,z) \sum_{T\in\mathcal{T}_{y,x}(\{\omega,x\})} K^\beta(T) = \delta_{y,x}\rho^\beta(\omega) + \sum_{z\in\Omega\setminus\{\omega,y\}} K^\beta(y,z) \sum_{T\in\mathcal{T}_{z,x}(\{\omega,x\})} K^\beta(T)$$

Define the sets $\mathcal{C}_1 := \{(z,T)|z\neq y,\ T\in\mathcal{T}_{y,x}(\{\omega,x\})\}$ and $\mathcal{C}_2 := \{(z,T)|z\in \Omega\setminus\{\omega,y\},\ T\in\mathcal{T}_{z,x}(\{\omega,x\})\}$, so that we can equivalently write

$$\sum_{(z,T)\in\mathcal{C}_1} K^\beta(x,z)K^\beta(T) = \delta_{y,x}\rho^\beta(\omega) + \sum_{(z,T)\in\mathcal{C}_2} K^\beta(y,z)K^\beta(T) \tag{2.7.3}$$

Let us consider the case $y = x$ first, so that $\mathcal{C}_1$ is defined by the revision graphs in $\mathcal{T}(\{\omega,x\})$. Then $\mathcal{C}_2 \subset \mathcal{C}_1$, since every $\{\omega,x\}$-revision tree that contains an (z,x)-path is a $\{\omega,x\}$-revision graph. The converse, of course, need not apply. Define the map

$$\varphi : \mathcal{C}_1\setminus\mathcal{C}_2 \to \mathcal{T}_\omega, (z,T)\mapsto \varphi(z,T) = (\Omega, \vec{E}(T)\cup\{(x,z)\}).$$

This operation takes a $\{\omega,x\}$-revision tree, not containing a (z,x)-path, and adds the edge (x,z). Thus, from the point z we have to arrive at ω in a unique way. By adding the edge (x,z) we create a ω-revision tree. This is illustrated in Figure 2.1. If we can show that φ is bijective, then we can move between $\mathcal{C}_1\setminus\mathcal{C}_2$ and $\mathcal{T}_\omega$ without losing any information. For $T' = \varphi(z,T)\in\mathcal{T}_\omega$, the inverse mapping is

$$\varphi^{-1}(T') = (\varphi_1^{-1}(T'),\varphi_2^{-1}(T')) = (\mathcal{R}_{T'}(x), \vec{E}(T')\setminus\{(x,\mathcal{R}_{T'}(x))\}) = (z,\vec{E}(T')\setminus\{(x,z)\}).$$

The left-hand side of eq. (2.7.3) turns then to[29]

$$\begin{aligned}
\sum_{(z,T)\in\mathcal{C}_1} K^\beta(x,z)K^\beta(T) &= \sum_{(z,T)\in\mathcal{C}_1\setminus\mathcal{C}_2} K^\beta(x,z)K^\beta(T) + \sum_{(z,T)\in\mathcal{C}_2} K^\beta(x,z)K^\beta(T)\\
&= \sum_{T'\in\mathcal{T}_\omega} K^\beta(x,\varphi_1^{-1}(T'))K^\beta(\varphi_2^{-1}(T')) + \sum_{(z,T)\in\mathcal{C}_2} K^\beta(x,z)K^\beta(T)\\
&= \sum_{T'\in\mathcal{T}_\omega} K^\beta(x,\mathcal{R}_{T'}(x))\frac{K^\beta(T')}{K^\beta(x,\mathcal{R}_{T'}(x))} + \sum_{(z,T)\in\mathcal{C}_2} K^\beta(x,z)K^\beta(T)\\
&= \rho^\beta(\omega) + \sum_{(z,T)\in\mathcal{C}_2} K^\beta(x,z)K^\beta(T)
\end{aligned}$$

what coincides with the right-hand side of this equation. Now, consider the

[29] Define $0\cdot\infty = 0$.

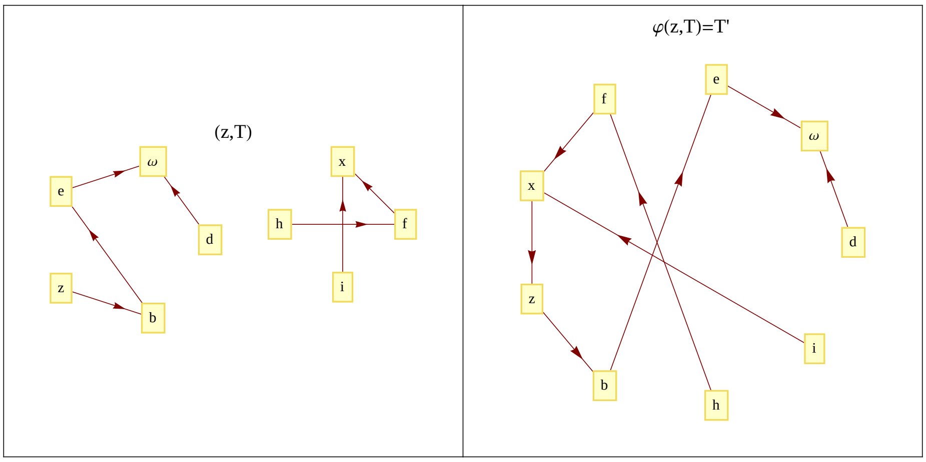

Figure 2.1: Starting from a pair $(z, T) \in \mathcal{C}_1 \setminus \mathcal{C}_2$ in the left panel, applying φ creates a graph T_ω shown in the right panel.

case $y \neq x$. Define the map $\varphi : \mathcal{C}_1 \to \mathcal{C}_2$ by

$$\varphi(z, T) = \begin{cases} (z, T) & \text{if } T \in \mathcal{T}_{z,x}(\{\omega, x\}), \\ (\mathcal{R}_T(y), (\vec{E}(T) \cup \{(y, z)\}) \setminus \{(y, \mathcal{R}_T(y))\}) & \text{if } T \notin \mathcal{T}_{z,x}(\{\omega, x\}). \end{cases}$$

φ maps the pair (z, T) onto itself if T contains a (z, x)-path. If such a path does not exist, then it connects y with z, deletes the (unique) outgoing edge from y, and shifts the initial vertex of the path from y to its unique neighbor under T, $\mathcal{R}_T(y)$. Since there exists a path connecting y with x (because $(z, T) \in \mathcal{C}_1$), the (unique) neighbor of y on T is also connected with x. Hence, we have constructed a revision tree $T' \in \mathcal{T}_{\mathcal{R}_T(y),x}(\{\omega, x\})$, with $\mathcal{R}_T(y) \in \Omega \setminus \{\omega, y\}$.[30] If we can show that φ is bijective, then $\mathcal{C}_1 = \mathcal{C}_2$ follows and we are done. We claim

$$\varphi^{-1}(z, T) = \begin{cases} (z, T) & \text{if } T \in \mathcal{T}_{y,x}(\{\omega, x\}), \\ (\mathcal{R}_T(y), (\vec{E}(T) \cup \{(y, z)\}) \setminus \{(y, \mathcal{R}_T(y))\}) & \text{if } T \notin \mathcal{T}_{y,x}(\{\omega, x\}). \end{cases}$$

Then $\varphi^{-1}(\varphi(z, T)) = (z, T)$ for all $(z, T) \in \mathcal{C}_1$. To see this, start with $(z, T) \in \mathcal{T}_{z,x}(\{\omega, x\})$. Then $\varphi(z, T) = (z, T) \in \mathcal{C}_2$ and $T \in \mathcal{T}_{y,x}(\{\omega, x\})$, hence $\varphi^{-1}(\varphi(z, T)) = (z, T)$. In the case where $T \notin \mathcal{T}_{z,x}(\{\omega, x\})$, let us call $\varphi(z, T) = (z', T') \in \mathcal{C}_2$. Then $T' \notin \mathcal{T}_{y,x}(\{\omega, x\})$, and consequently

$$\begin{aligned} \varphi(z', T') &= (\mathcal{R}_{T'}(y), (\vec{E}(T') \cup \{(y, z')\}) \setminus \{(y, \mathcal{R}_{T'}(y))\}) \\ &= (z, (\vec{E}(T') \cup \{(y, \mathcal{R}_T(y))\}) \setminus \{(y, z)\}) \\ &= (z, T) \end{aligned}$$

[30] If $\mathcal{R}_T(y) = x$ then we get the pair (x, T) with $T \in \mathcal{T}(\{\omega, x\})$ which lies in $\mathcal{C}_2$ for $z = x$.

□

The expected time spent in some state x before the system returns to ω is given by

$$v_x(\omega) = \mathbb{E}\left(\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n^\beta = x\}\wedge\{\tau(\omega)>n\}} | X_0^\beta = \omega\right). \tag{2.7.4}$$

Intuitively, this is the average length of ω-cycles in the revision graph T, on which x is visited.

Lemma 2.7.3. *Let $v(\omega)$ denote the vector whose elements are defined by (2.7.4). Then*

(i) $v_\omega(\omega) = 1$*;*

(ii) $v(\omega)\mathbf{K}^\beta = v(\omega)$*;*

(iii) $v(\omega)$ *is bounded and positive.*

Proof. (i) By definition.

(ii) By the Markov property and time-homogeneity we have

$$\begin{aligned}
v_x(\omega) &= \sum_{n=1}^{\infty} \mathbb{P}(X_n^\beta = x, \tau(\omega) \geq n | X_0^\beta = \omega) \\
&= \sum_{n=1}^{\infty} \sum_{\omega' \in \Omega} \mathbb{P}(X_{n-1}^\beta = \omega', X_n^\beta = x, \tau(\omega) \geq n | X_0^\beta = \omega) \\
&= \sum_{n=1}^{\infty} \sum_{\omega' \in \Omega} \mathbb{P}(X_{n-1}^\beta = \omega', \tau(\omega) \geq n | X_0^\beta = \omega) K^\beta(\omega', x) \\
&= \sum_{n=0}^{\infty} \sum_{\omega' \in \Omega} \mathbb{P}(X_n^\beta = \omega', \tau(\omega) - 1 \geq n | X_0^\beta = \omega) K^\beta(\omega', x) \\
&= \sum_{\omega' \in \Omega} v_{\omega'}(\omega) K^\beta(\omega', x)
\end{aligned}$$

(iii) Suppose there exists a state x such that $v_x(\omega) = 0$. Then, for all $n \geq 1$,

$$0 = \sum_{\omega' \in \Omega} v_{\omega'}(\omega) K^{\beta,n}(\omega', x) = K^{\beta,n}(\omega, x) + \sum_{y \neq \omega} v_y(\omega) K^{\beta,n}(y, x),$$

and so $K^{\beta,n}(\omega, x) = 0$, contradicting irreducibility. Essentially the same argument can be used to see that $v_x(\omega) < \infty$ for all x.

□

The *expected return time* to ω is $\bar{v}(\omega) = \sum_{\omega' \in \Omega} v_{\omega'}(\omega)$. This is a measure of the average length of ω-cycles. A state ω is called *positive recurrent* if $\bar{v}(\omega) < \infty$.

Lemma 2.7.4. *Let $\mathbf{K}^\beta$ be irreducible and recurrent. Then $\mathbf{K}^\beta$ has an invariant distribution μ^β such that $\mu^\beta(\{\omega\}) = \mu^\beta(\omega) = \frac{1}{\bar{v}(\omega)}$.*

Proof. Since Ω is finite, there exists a positive recurrent state $\omega \in \Omega$. From irreducibility, it follows that all states are positive recurrent. Then $\bar{v}(\omega) = \sum_{\omega' \in \Omega} v_{\omega'}(\omega) < \infty$. Since $v(\omega)$ defines an invariant measure for $\mathbf{K}^\beta$, $\mu^\beta = \frac{1}{\bar{v}(\omega)} v(\omega)$ is an invariant distribution for $\mathbf{K}^\beta$, satisfying $\mu^\beta(\omega) = 1/\bar{v}(\omega)$. $\square$

Using this Lemma, observe that

$$\begin{aligned}
\mu^\beta(\omega) &= \left(\sum_{\omega' \in \Omega} v_{\omega'}(\omega) \right)^{-1} = \left(1 + \sum_{x \neq \omega} v_x(\omega) \right)^{-1} \\
&= \left[1 + \sum_{n=1}^{\infty} \mathbb{P}(X_n^\beta \neq \omega, \tau(\omega) \geq n+1 | X_0^\beta = \omega) \right]^{-1} \\
&= \left[1 + \sum_{n=1}^{\infty} \sum_{y \neq \omega} \mathbb{P}(X_1^\beta = y, \tau(\omega) \geq n+1 | X_0^\beta = \omega) \right]^{-1} \\
&= \left[1 + \sum_{y \neq \omega} K^\beta(\omega, y) \sum_{n=1}^{\infty} \mathbb{P}(\tau(\omega) \geq n | X_0^\beta = y) \right]^{-1} \\
&= \left[1 + \sum_{y \neq \omega} K^\beta(\omega, y) \mathbb{E}(\tau(\omega) | X_0^\beta = y) \right]^{-1}.
\end{aligned}$$

We have for all $y \neq \omega$ the identity

$$\begin{aligned}
\mathbb{E}(\tau(\omega) | X_0^\beta = y) &= \mathbb{E}\left(\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n^\beta \neq \omega\} \cap \{\tau(\omega) > n\}} | X_0^\beta = y \right) \\
&= \mathbb{E}\left(\sum_{x \neq \omega} \sum_{n=0}^{\infty} \mathbb{1}_{\{X_n^\beta = x\} \cap \{\tau(\omega) > n\}} | X_0^\beta = y \right) \\
&= \sum_{x \neq \omega} \bar{K}^{\beta, x}(y, x).
\end{aligned}$$

The last equality follows from Lemma 2.7.1. Using this identity gives

$$\begin{aligned}
\mu^\beta(\omega) &= \left[1 + \frac{1}{\rho^\beta(\omega)} \sum_{y,x\neq\omega} K^\beta(\omega,y) \sum_{T\in\mathcal{T}_{y,x}(\{\omega,x\})} K^\beta(T)\right]^{-1}. \\
&= \frac{\rho^\beta(\omega)}{\rho^\beta(\omega) + \sum_{x\neq\omega}\sum_{y\neq\omega} K^\beta(\omega,y)\sum_{T\in\mathcal{T}_{y,x}(\{\omega,x\})} K^\beta(T)} \\
&= \frac{\rho^\beta(\omega)}{\rho^\beta(\omega) + \sum_{x\neq\omega}\sum_{T_x\in\mathcal{T}_x} K^\beta(T_x)}
\end{aligned}$$

which is eq. (2.2.8).
We now provide some justifications for the cost functions (2.2.13). The results presented here are due to Beggs (2005), who in turn builds on the work of Catoni (1999). The clue is to consider a modified Markov chain, which monitors only transitions in a suitably chosen subset $\mathcal{X} \subset \Omega$. Therefore, for $m \in \mathbb{N}_0$, define the stopping times of successive visitations of the set $\mathcal{X}$ as $\tau_{-1}(\mathcal{X}) \equiv 0$, $\tau_m(\mathcal{X}) := \inf\{n \geq \tau_{m-1}(\mathcal{X}) + 1 | X^\beta_n \in \mathcal{X}\}$. The Markov chain $Z^\beta_m := X^\beta_{\tau_m(\mathcal{X})}$ records all visitations of X^β to the set $\mathcal{X}$.

Lemma 2.7.5. *Let $\mathcal{X} \subset \Omega$ be a non-empty set. $(Z^\beta_m)_{m\geq 0}$ is an irreducible, recurrent and time-homogeneous Markov chain on $\mathcal{X}$. Its unique invariant distribution is given by $\mu^\beta(\cdot|\mathcal{X})$ and its transition probabilities are for all $y, x \in \mathcal{X}$*

$$\mathbb{P}(Z^\beta_{m+1} = x | Z^\beta_m = y) = K^\beta(y,x) + \sum_{z\in\Omega\setminus\mathcal{X}} K^\beta(y,z) Q_{\Omega\setminus\mathcal{X}}(z \to x) \qquad (2.7.5)$$

where

$$Q^\beta_{\Omega\setminus\mathcal{X}}(z \to x) := \frac{\sum_{T\in\mathcal{T}_{z,x}(\mathcal{X})} K^\beta(T)}{\sum_{T\in\mathcal{T}(\mathcal{X})} K^\beta(T)}.$$

Proof. That the restricted process is a Markov chain with these properties can be proved quite easily. See Proposition 7.2.1 in Catoni (2001). For the second claim, note that the strong Markov property (see Norris, 1997), applied to the stopping times $\tau_m(\mathcal{X})$, implies that

$$\begin{aligned}
\mathbb{P}(Z^\beta_{m+1} = x | Z^\beta_m = y) &= \mathbb{P}(X^\beta_{\tau_{m+1}(\mathcal{X})} = x | X^\beta_{\tau_m(\mathcal{X})} = y) = \mathbb{P}(X^\beta_{\tau_1(\mathcal{X})} = x | X^\beta_0 = y) \\
&= K^\beta(y,x) + \sum_{n=2}^{\infty} \sum_{z \notin \mathcal{X}} \mathbb{P}(X^\beta_s \notin \mathcal{X} \; \forall s \in [1, n-1], X^\beta_n = x | X^\beta_0 = y) \\
&= K^\beta(y,x) + \sum_{n=1}^{\infty} \sum_{z,\omega \notin \mathcal{X}} K^\beta(y,z) \mathbb{P}(X^\beta_n = \omega, \tau(\mathcal{X}) \geq n | X^\beta_0 = z) K^\beta(\omega, x) \\
&= K^\beta(y,x) + \sum_{n=1}^{\infty} \sum_{z,\omega \notin \mathcal{X}} K^\beta(y,z) \frac{\sum_{T \in \mathcal{T}_{z,\omega}(\mathcal{X} \cup \{\omega\})} K^\beta(T)}{\sum_{T \in \mathcal{T}(\mathcal{X})} K^\beta(T)} K^\beta(\omega, x) \\
&= K^\beta(y,x) + \sum_{n=1}^{\infty} \sum_{z \notin \mathcal{X}} K^\beta(y,z) \frac{\sum_{T \in \mathcal{T}_{z,x}(\mathcal{X})} K^\beta(T)}{\sum_{T \in \mathcal{T}(\mathcal{X})} K^\beta(T)} \\
&= K^\beta(y,x) + \sum_{z \in \Omega \setminus \mathcal{X}} K^\beta(y,z) Q^\beta_{\Omega \setminus \mathcal{X}}(z \to x)
\end{aligned}$$

where we have used in the fourth line Lemma 2.7.1. □

We will apply this result to derive the set-wise cost functions (2.2.13). Let $\mathcal{L}_1, \ldots, \mathcal{L}_k$ denote the recurrent classes of the unperturbed model $\mathcal{M}$ and $\Re = \bigcup_{i=1}^k \mathcal{L}_i$ the union of recurrent classes. The literature often refers to the sets $\mathcal{L}_i$ as limit sets. From each limit set we make an arbitrary selection $x_i \in \mathcal{L}_i, 1 \leq i \leq k$, and define $\mathcal{X} := \{x_1, \ldots, x_k\}$. Note that $\mathcal{X}$ contains the absorbing states (i.e. the singleton recurrent sets). For $y, x \in \mathcal{X}$, let

$$c^{\mathcal{X}}(y,x) := -\lim_{\beta \to 0} \beta \log \mathbb{P}(Z^\beta_{m+1} = x | Z^\beta_m = y)$$

be the cost function of the restricted process $(Z^\beta_m)_{m \geq 0}$. Further, define $\hat{c}^*(\omega) := \min_{y \in \Omega \setminus \{\omega\}} \hat{c}(\omega, y)$ the least cost transition from some state $\omega \in \Omega$ (omitting the type of transition).

Lemma 2.7.6. *Let $\mathcal{X} = \{x_1, \ldots, x_k\}, x_i \in \mathcal{L}_i, 1 \leq i \leq k$. Then, for all $y, x \in \mathcal{X}$, the costs of transiting from y to x are given by*

$$c^{\mathcal{X}}(y,x) = \min_{P \in \mathcal{P}_{y,x}(\bar{\mathcal{X}} \cup \{y\}, \mathcal{X})} \hat{C}(P), \tag{2.7.6}$$

where for any path P, $\hat{C}(P) := \sum_{\vec{e} \in \vec{E}(P)} \hat{c}(\vec{e})$, and $\bar{\mathcal{X}} = \Omega \setminus \mathcal{X}$.

Proof. The proof is based on Lemma 2.2.1 and the transition probability of the restricted process (Z^β_m) found in Lemma 2.7.5. We know that $K^\beta(y,x) =$

$\exp\left[-\frac{1}{\beta}(\hat{c}(y,x)+o(1))\right]$. For a given point $z \in \Omega \setminus \mathcal{X}$, we have to find an asymptotic bound for $K^\beta(y,z)Q^\beta_{\Omega\setminus\mathcal{X}}(z \to x)$. For sufficiently small β this probability can be written as

$$\exp\left[-\frac{1}{\beta}(\hat{c}(y,z)+o(1))\right]\frac{\sum_{T\in\mathcal{T}_{z,x}(\mathcal{X})}\exp\left[-\frac{1}{\beta}(\hat{C}(T)+o(1))\right]}{\sum_{T\in\mathcal{T}(\mathcal{X})}\exp\left[-\frac{1}{\beta}(\hat{C}(T)+o(1))\right]}.$$

Taking logarithms, and multiplying by $-\beta$ gives us the leading order term

$$\hat{c}(y,z)-\beta\log[Q^\beta_{\Omega\setminus\mathcal{X}}(z\to x)]. \tag{2.7.7}$$

The second terms is governed by

$$\log\left[\sum_{T\in\mathcal{T}_{z,x}(\mathcal{X})}\exp\left(-\frac{1}{\beta}(\hat{C}(T)+o(1))\right)\right]-\log\left[\sum_{T\in\mathcal{T}(\mathcal{X})}\exp\left(-\frac{1}{\beta}(\hat{C}(T)+o(1))\right)\right].$$

Lemma 2.2.1 tells us that for $\beta\downarrow 0$ this number is asymptotically equivalent to

$$\max_{T\in\mathcal{T}_{z,x}(\mathcal{X})}\exp(-\hat{C}(T)/\beta)-\max_{T\in\mathcal{T}(\mathcal{X})}\exp(-\hat{C}(T)/\beta).$$

(2.7.7) boils then down to

$$\hat{c}(y,z)+\min_{T\in\mathcal{T}_{z,x}(\mathcal{X})}\hat{C}(T)-\min_{T\in\mathcal{T}(\mathcal{X})}\hat{C}(T). \tag{2.7.8}$$

Call $T^*_{z,x}\in\mathcal{T}_{z,x}(\mathcal{X})$ a least cost $\mathcal{X}$-graph containing a (z,x)-path, and $T^*_{\mathcal{X}}\in\mathcal{T}(\mathcal{X})$ a least cost $\mathcal{X}$-graph. Call P^* the (z,x)-path used on $T^*_{z,x}$. We claim that all edges in $T^*_{z,x}$, which are not on the path P^*, are also used under $T^*_{\mathcal{X}}$. This follows from the fact that $\mathcal{T}_{z,x}(\mathcal{X})\subset\mathcal{T}(\mathcal{X})$. The only difference between the graphs $T^*_{z,x}$ and $T^*_{\mathcal{X}}$ are the edges on the path $P^*=\{\omega_1,\ldots,\omega_l\},\omega_1=z,\omega_l=x,\omega_t\notin\mathcal{X},\forall t=2,\ldots,l-1$. The edge (ω_{t-1},ω_t) in P^* need not be globally optimal, so that this edge causes supplementary costs $\hat{c}(\omega_{t-1},\omega_t)-\hat{c}^*(\omega_{t-1})$. The term $\hat{c}^*(\omega_{t-1})$ is the cost of the edge leaving ω_{t-1} under $T^*_{\mathcal{X}}$. Hence, for any $z\notin\mathcal{X}$ we can pin down the costs of a transition from y to x, via z, as

$$\hat{c}(y,z)+\min_P\left\{\sum_{t=2}^{l}[\hat{c}(\omega_{t-1},\omega_t)-\hat{c}^*(\omega_{t-1})]:P=\{\omega_1,\ldots,\omega_l\}\in\mathcal{P}_{z,x}(\bar{\mathcal{X}},\mathcal{X})\right\}.$$

Call this $\hat{C}^{\mathcal{X}}(y\to x|z)$. It follows that

$$c^{\mathcal{X}}(y,x)=\min\left\{\hat{c}(y,x),\min_{z\in\Omega\setminus\mathcal{X}}\hat{C}^{\mathcal{X}}(y\to x|z)\right\}.$$

Next, we claim that if ω is used on the optimal path P^*, then $\hat{c}^*(\omega) = 0$. To see this, observe that by definition of such paths, ω is either a transient state, or it is a recurrent state, not contained in the selection $\mathcal{X}$. In the first case, $\hat{c}^*(\omega) = 0$, since any transient state can be appended to a zero-cost path leading to some recurrent state. In the second case we also have $\hat{c}^*(\omega) = 0$, since ω cannot be absorbing, hence communicates with another state in the same recurrent class. Hence, if ω_1 is the first state on the optimal path P^* then $\hat{c}^*(\omega_1) = 0$, and iteration gives $\hat{c}^*(\omega_{t-1}) = 0$ for all $2 \leq t \leq l-1$. Consequently, calling $\hat{C}(P) = \sum_{t=1}^{l-1} \hat{c}(\omega_t, \omega_{t+1})$ for a path $P \in \mathcal{P}_{z,x}(\bar{\mathcal{X}}, \mathcal{X})$, we have

$$\begin{aligned} c^{\mathcal{X}}(y,x) &= \min\left\{\hat{c}(y,x), \min_{z\in\Omega\setminus\mathcal{X}} \min_{P\in\mathcal{P}_{z,x}(\Omega\setminus\mathcal{X},\mathcal{X})} \Big(\hat{c}(y,z) + \hat{C}(P)\Big)\right\} \\ &= \min_{P\in\mathcal{P}_{y,x}(\bar{\mathcal{X}}\cup\{y\},\mathcal{X})} \hat{C}(P). \end{aligned}$$

□

This Lemma gives us the costs of a transition between two recurrent states y, x. If $y \in \mathcal{L}$ and $x \in \mathcal{L}'$, then we can extend the above argument to a set-wise cost functions, measuring the difficulty of a transition from recurrent class $\mathcal{L}$ to recurrent class $\mathcal{L}'$. Let $\omega \in \mathcal{L}, \omega' \in \mathcal{L}'$. There is a zero-cost path connecting y with ω, and a zero-cost path connecting x with ω'. Hence, the least cost of moving from $\mathcal{L}$ to $\mathcal{L}'$ is exactly (2.2.13). This in turn shows that the least cost of reaching a state $\omega \in \mathcal{L}$ coincides with the minimal cost needed to reach the limit set $\mathcal{L}$ from all other limit sets, justifying eq. (2.2.14). Hence, if ω is stochastically stable, so must all states in the same recurrent class. This gives us the following result.

Corollary 2.7.1.

$$\Omega^* = \bigcup\{\mathcal{L} | (\exists \omega \in \mathcal{L}) : \gamma(\omega) = \min_{\omega'\in\Omega} \gamma(\omega')\}. \tag{2.7.9}$$

One can also use the argument in Lemma 2.7.6 to establish a connection with the radius/co-radius formulas of Ellison (2000). I refer to Beggs (2005) for further discussions.

Chapter 3

Continuous-time co-evolutionary models

3.1 Introduction

Almost all papers on the co-evolution of networks and play study finite-state time-homogeneous Markov processes which have a "birth-death" structure. A good summary of this work can be found in Vega-Redondo (2007). Such processes are very well studied, and the distinction between continuous time and discrete time theory is, more or less, a matter of taste. For some purposes the continuous-time theory offers a more intuitive way to describe the process and some results are more straightforward to derive. This note will discuss a class of continuous-time Markov processes which are going to be used in chapters 4 and 5. In this chapter we will not present any new results. It is much more intended to serve as a sort of reference for constructing new co-evolutionary models. I have tried to develop the theory in a self-contained way, and attempted to be as precise and concise as possible. Section 3.2 presents a class of co-evolutionary models with noise, $\mathcal{M}^\beta$, in continuous time. The connection to the family $\mathcal{M}^\beta$ evolving in discrete time, as set forth in the previous chapter, is presented in 3.3.

3.2 Construction of the family $\mathcal{M}^\beta$ in continuous time

We will use the notation already introduced in the previous two chapters. Chapters 4 and 5 deal with continuous time co-evolutionary models with noise

$$\mathcal{M}^\beta = (\Omega, \mathcal{F}, \mathbb{P}, (Y^\beta(t))_{t\geq 0})_{\beta\in\mathbb{R}_+},$$

where $\Omega = \mathcal{A}^\mathcal{I} \times \mathcal{G}[\mathcal{I}]$ is the finite state space containing pairs of action profiles α, and networks g. An element of this set is denoted as $\omega := (\alpha, g)$. $\mathcal{F}$ is a σ-algebra, which can be chosen to be the set of all subsets 2^Ω. $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a probability measure, and $(Y^\beta(t))_{t\geq 0}$ is a family of Ω-valued random variables, indexed by a continuous-time parameter $t \geq 0$ and a noise parameter $\beta \geq 0$. If desired one could add the population size as additional parameter (as will be done in chapter 5). Hence, the only technical difference between $\mathcal{M}^\beta$, as defined in chapter 2 and the family of random processes discussed here, is that time is continuous. The process satisfies the Markov property, meaning that for any sequence $0 \leq t_0 \leq t_1 \leq \ldots \leq t_k \leq t$, we have

$$\mathbb{P}(Y^\beta(t) = \omega | Y^\beta(t_j), 1 \leq j \leq k) = \mathbb{P}(Y^\beta(t) = \omega | Y^\beta(t_k)).$$

A sample path is characterized by its jump times $(J_n)_{n\in\mathbb{N}_0}$, and its holding times $S_{n+1} := J_{n+1} - J_n$ (Norris, 1997). We define these random variables

inductively as $J_0 = 0$, and for every $n \geq 1$, $J_n := \inf\{t \geq J_{n-1} : Y^\beta(t) \neq Y^\beta(J_{n-1})\}$. A technical difficulty would arise if $J_n = \infty$ for some n. This will not play any role for the process described in this chapter, and so we do not need to discuss this possibility.
The Markov process $(Y^\beta(t))_{t\geq 0}$ builds on the following ideas:

- Players adapt their action to their interaction neighborhood $\mathcal{N}^i$, which in turn is partially under control of the agents. Partial refers to the process of link creation where a link between two players can be formed following the initiative of a single player.

- The decisions once made should last for some time, so that their consequences can be realized by the player. This phenomenon of *inertia* is admittedly formulated in a rather extreme way, since most of the time a single player will not do anything. Further, when a player is allowed to change some of his/her characteristics, it is assumed that only a single attribute may be changed at one update event. Moreover, the arrival of updating events is random, hence there is no strategic delay. It is like the players walk through a labyrinth with closed doors. At some moments of time a door suddenly opens and the player closest to this door goes through it.

- The process can be studied on two "layers". There is a macro level of the process and a micro level. On the macro level there are only 3 elementary events which may be observed, changing the state variable $\omega = (\alpha, g)$: A change in the action profile α might appear, but we do not know at which position of this list the change took place. A change in the network g may take place, so that one existing edge disappears from the network, or a previously non-existing edge is added to the network. Again, from the macro-perspective we don't know the pair of vertices involved in this change. The Micro-level gives us the missing information. Here the probabilities are determined which govern the pattern of the process $(Y^\beta(t))_{t\geq 0}$.

On the macro level 3 elementary events may be observed changing the state variable ω. These events define how evolution forms the system over time. One possible event is that the action profile α changes at some point of time. This defines a sub-process called "action update", or "action adjustment". Other possible routes for evolution to come in is via changes in the current network g. There we have two possibilities; Either the network expands, in the sense that a link becomes added, or the network shrinks, in the sense that a link becomes destroyed. In the first case we will speak of the event of “link creation”, while in the latter case we will observe an event of “link destruction”. Each of these events appear with an independent probability which is calculated from an aggregation procedure of the micro-level; See Figure 3.1 for an illustration on the workings of the dynamic. Hence, the

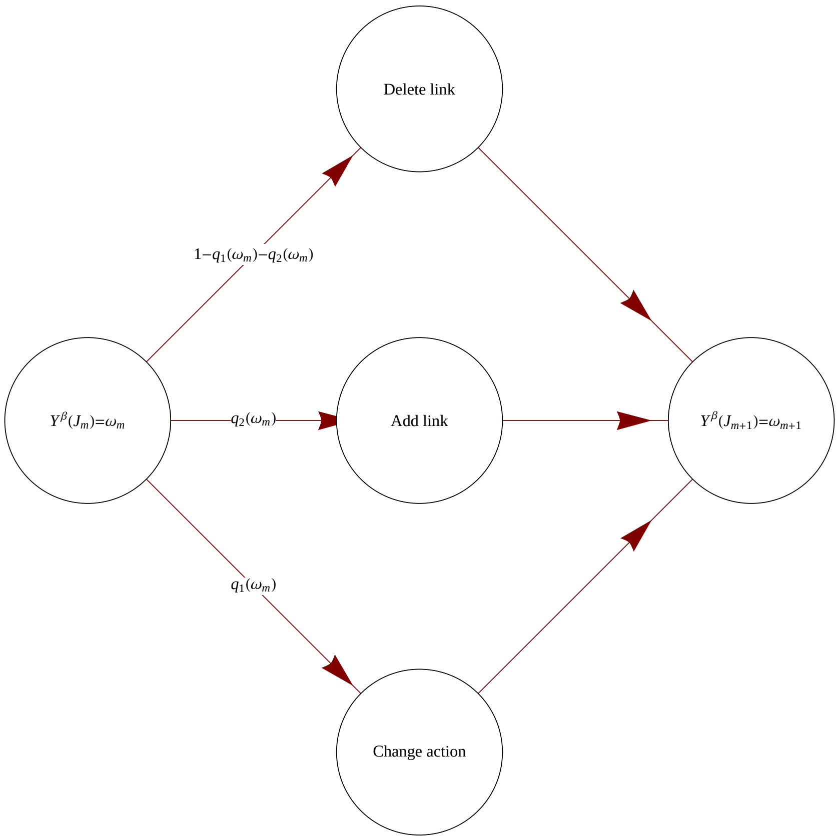

Figure 3.1: A diagrammatic sketch of the functioning of the Markov process $(Y^\beta(t))_{t\geq 0}$.

process $(Y^\beta(t))_{t\geq 0}$ combines an evolutionary dynamic acting on the action profile α with a graph dynamic which shapes the interaction structure of the population. The two processes proceed at characteristic time scales τ_a, τ_g, respectively. The ratio $\tau := \tau_g/\tau_a$ determines the relative speed of the two processes. If τ is much larger than 1, network evolution will proceed at a faster time scale than action adjustment. If τ is much smaller than 1, then action adjustment opportunities arrive with a higher frequency to the population. Both cases correspond to a situation were the population dynamics is a two-tiered process. If we want to have a truly co-evolutionary model, we must time the two processes so that $\tau \approx 1$. One can interpret the macro-level of the process as the regulation of the two time-scales. It tells us when which type of event will cause the system to move. At the micro-level (see Section 3.2.2) we analyze the relevant processes within their characteristic time scale. In essence, all these processes are modeled as simple Poisson processes, which we will call "counting processes".

3.2.1 On the use of the term "counting process"

The words "arrival process" and "counting process" are used for any time homogeneous continuous-time random process $X = (X(t))_{t\geq 0}$, with the following properties[1]:

(i) $t \mapsto X(t)$ is right-continuous, non-negative, non-decreasing and integer-valued;

(ii) at the points of discontinuity a jump of magnitude 1 is observed;

(iii) the increments at points of time $0 < t_1 < t_2 < \ldots < t_k : X(t_1), X(t_2) - X(t_1), \ldots, X(t_{k-1}) - X(t_k)$ are independent.

(iv) The distribution of $X(t) - X(s)$ depends only on $t - s$.

Starting with this definition, let us define random variables

$$(\forall n \in \mathbb{N}) : \; J_n := \inf\{t : X(t) \geq n\}, \; T_n := J_n - J_{n-1}. \tag{3.2.1}$$

$(J_n)_{n\in\mathbb{N}}$ are called the *jump times* of the random process X. These are distinguished points of time - the only points at which a change in the system X is observed. If $J_n = t$ then the process X makes its n-th jump at time t. For any $\varepsilon > 0$ we have $X(J_n - \varepsilon) < n$. The numbers $(T_n)_{n\in\mathbb{N}}$ are the *holding times* of the process X, measuring the length of time between two consecutive jumps.

By right-continuity it follows that for every $t > 0$, there exists a $\varepsilon > 0$ such that $X(t) = X(t + \varepsilon)$. So for every n we get $T_n > 0$. By definition $J_n = T_0 + \sum_{j=1}^{n} T_j$, and T_0 is interpreted as the starting time of the process which will be thought of as "time 0". Now consider the comparison process

$$N(t) := \sup\{n \in \mathbb{N} : J_n \leq t\}. \tag{3.2.2}$$

[1]See also Grimmett and Stirzaker (2001), p.247 or Norris (1997), chapter 2.

Its interpretation is that it counts the largest number of jumps the process can make in the given time interval $[0, t]$. We will show that $N(t) = X(t)$ for all $t \geq 0$, and so the above definition of an arrival process is equivalent to the construction of a random process having positive inter-arrival times.

Proposition 3.2.1. *For all $t \geq 0$ we have $N(t) = X(t)$.*

Proof. The sequences $(J_n)_{n\in\mathbb{N}}, (N(t))_{t\geq 0}$ are non-decreasing by definition. If $N(t) = n$ then $J_n \leq t < J_{n+1}$. Since $(X(t))_{t\geq 0}$ is non-decreasing, it follows $n \leq X(t) < n + 1$. Hence, $X(t) = n$, since X is integer-valued.
Suppose now that $X(t) = n$. Then $N(t) = n$, because $N(t) > n$ leads to the contradiction that $J_n < t$ and $N(t) < n$ generates the counter-factual $J_n > t$. □

This line of reasoning shows that our definition of an arrival process means that we think of a random process having positive inter-arrival times, and discrete points of time at which its value increases by one unit. Drawing a realization of the sequences $(J_n)_{n\in\mathbb{N}_0}$ and $(N(t))_{t\geq 0}$ in the $(\mathbb{R}, \mathbb{N}_0)$- plane shows that $(J_n)_{n=0}^{\infty}$ is the inverse of $N(t)$ in the sense that if $J_n \leq t$ then $N(t) \leq n$.
Points (iii) and (iv) in the definition of X are needed to determine the distribution of the holding times. The event $\{T_n > t + h\}$ is probabilistically equivalent to the event $\{J_n > t+h+J_{n-1} \& \ J_n > t+J_{n-1}\}$ for every $t, h > 0$. Hence, on $\{J_{n-1} \leq s < J_n\}$ and point (iv) in the definition, this event depends only on the behavior of the process in the interval $[s, s + t + h]$. The probability that the process performs its n-th jump after $t + h$ is equivalent to saying that it does not make a jump in the interval $[s, s + t]$ and $(s + t, s + t + h]$. It follows that

$$
\begin{aligned}
&\mathbb{P}(T_n > t + h|X(\tau); \tau \leq s) = \mathbb{P}(X(s + t + h) - X(s) = 0|X(\tau); \tau \leq s) \\
&= \mathbb{P}(J_n > s + t + h \ \& \ J_n > t + s|X(\tau); \tau \leq s) \\
&= \mathbb{P}(X(s + t) - X(s) = 0 \ \& \ X(s + t + h) - X(s + t) = 0|X(\tau); \tau \leq s) \\
&= \mathbb{P}(X(s + t + h) - X(s + t) = 0|X(\tau); \tau \leq s \ \& \ X(s + t) - X(s) = 0) \\
&\times \mathbb{P}(X(s + t) - X(s) = 0|X(\tau); \tau \leq s) \\
&= \mathbb{P}(T_n > h|X(\tau); \tau \leq s + t)\mathbb{P}(T_n > t|X(\tau); \tau \leq s \ \& \ T_n > t) \ |\text{(iii) and (iv)} \\
&= \psi(t)\psi(h)
\end{aligned}
$$

Hence, we obtain $\psi(t+h) = \psi(t)\psi(h)$, $\psi(0) = 1$. We can solve this functional equation with the help of the following well-known Lemma, whose proof is presented just for sake of completeness.

Lemma 3.2.1. *For all $a \in \mathbb{R}$ there is a unique solution $\psi : \mathbb{R} \to \mathbb{R}$ to the functional equation $\psi(t+h) = \psi(t)\psi(h)$ for $t, h \in \mathbb{R}$, satisfying $\psi(1) = \exp(-a)$. This solution is $\psi(t) = \exp(-at)$.*

Proof. The proof proceeds by finding a function restricted to the rationales $\mathbb{Q}$ satisfying the functional equation, and then, by a suitable limiting argument, showing that we can extend the domain to $\mathbb{R}$. Let $\phi : \mathbb{Q} \to \mathbb{R}$ be a function satisfying $\phi(p+q) = \phi(p)\phi(q)$ for all $p, q \in \mathbb{Q}$ and $\phi(1) = \exp(-a)$. For $p = 1, q = 0$ we get $\phi(1) = \phi(1)\phi(0)$, hence $\phi(0) = 1$. For $p = -q$ we have $\phi(0) = 1 = \phi(p)\phi(-p)$, and so $\phi(p) \neq 0$ and $\phi(-p) = 1/\phi(p)$. For $p = q$ $\phi(2p) = \phi(p)^2 > 0$, and so $\phi(p) > 0$ for all rational p. Starting from this identity, we have, by induction on n, that $\phi(np) = \phi(p)^n$ for all $n \in \mathbb{N}$. In particular, we observe that

$$\exp(-ap) = \phi(1)^p = \phi(p).$$

To check uniqueness, assume that ϕ_1, ϕ_2 are two solutions of the functional equation satisfying $\phi_2(1) = \phi_1(1) = \exp(-a)$. Then for all $p \in \mathbb{Q}$

$$\begin{aligned}\phi_1(p) &= \phi_1(1)^p = \exp(-ap) \\ &= \phi_2(1)^p = \phi_2(p) \Rightarrow \phi_1(p) = \phi_2(p).\end{aligned}$$

Now we extend the domain of ϕ to $\mathbb{R}$. If $t \in \mathbb{R}$ is rational define $\psi(t) := \phi(t)$. If t is irrational, let $(t_n)_{n=0}^{\infty}$ denote a sequence of rational numbers converging to t. If $a = 0$ then $\phi(1) = 1$ and so $\phi(t_n) = 1$ for all n, independently of how we choose the sequence. Thus, if $t_n \to t$, it follows $\phi(t_n) \to \phi(t) = 1 =: \psi(t)$ for $n \to \infty$. Now suppose $a > 0$. Then $(\phi(t_n))_{n=0}^{\infty}$ is a monotonically falling sequence in $[0,1] \cap \mathbb{Q}$ if we choose $t_n \nearrow t$. Similarly, if $a < 0$, we can choose $t_n \nearrow t$ to make $(\phi(t_n))_{n=0}^{\infty}$ a monotonically increasing sequence in $[1, \infty) \cap \mathbb{Q}$, bounded from above. Hence, in both cases the limit $\psi(t) := \lim_{n\to\infty} \phi(t_n)$ exists. Now we show that $\psi(t)$ is independent of the choice of the sequence (t_n). Let $(t_n), (s_n)$ denote two sequences such that $t_n, s_n \nearrow t$. If $a > 0$, then $\sup_n \phi(t_n) = \phi(t_0) =: L$, and if $a < 0$ then $\sup_n \phi(t_n) = \psi(t) =: L$. Further $|t_n - s_n| \to 0$ as $n \to \infty$. Hence

$$\begin{aligned}|\phi(t_n) - \phi(s_n)| &= \exp(-at_n)|1 - \exp(-a(t_n - s_n))| \\ &\leq L|a||t_n - s_n| + o(1)\end{aligned}$$

where $o(1)$ represents non-negative terms which go to zero as $n \to \infty$ at a higher order than $|t_n - s_n|$. This shows that ϕ satisfies a Lipschitz condition, and as $n \to \infty$ we observe $|\phi(t_n) - \phi(s_n)| \to 0$. $\square$

This shows that $\psi(t) = \exp(-ct)$ for $c := -\log \psi(1) \geq 0$.

3.2.2 The Micro Level

Figure 3.1 describes the phenomena $\mathcal{M}^\beta$ captures at the macro-level of the society: Action adjustment, link creation and link destruction. In this section we describe the working of the process on these sub-processes. The presentation is tailored to the models of chapters 4 and 5, so that some specific assumptions on functional forms are discussed. This is done only to establish a clear connection between the general framework of chapter 2, and the more specific forthcoming models.

Action adjustment

This sub-process is the easiest to understand, so we should start with it. The general idea carries over to the other 2 processes. Every player possesses a Poisson alarm clock running with constant intensity $\nu > 0$, which is common to all players. I denote this process as $(R^i(t))_{t\geq 0}$ for every player $i \in \mathcal{I}$. This is a counting process as defined in Section 3.2.1. For an integer m, the event $\{R^i(t) = m\}$ means that player i had at time t already m action update opportunities. The infinitesimal description of Poisson processes is very convenient to work with. It states that for a very small time interval $[t, t+h), h \to 0$, the clock of player i rings with probability[2]

$$\mathbb{P}(R^i(t+h) - R^i(t) = m | R^i(s); 0 \leq s \leq t) = \begin{cases} \nu h + o(h) & \text{if } m = 1, \\ 1 - \nu h + o(h) & \text{if } m = 0. \end{cases} \tag{3.2.3}$$

uniformly in t. This assumption, combined with our definition of a counting process, gives us the following Proposition.

Proposition 3.2.2. *$R^i(t) \sim \mathrm{POI}(\nu t)$ for all $t \geq 0$ and $i \in \mathcal{I}$.*

Proof. Write $p_m^{(i)}(t) := \mathbb{P}(R^i(t) = m)$. Then, by rules of conditional probability, for $t > 0$ and $m \in \mathbb{N}$ we get

$$\begin{aligned} p_m^{(i)}(t+h) &= p_m^{(i)}(t)\mathbb{P}(R^i(t+h) - R^i(t) = 0 | R^i(t) = m) \\ &+ p_{m-1}^{(i)}(t)\mathbb{P}(R^i(t+h) - R^i(t) = 1 | R^i(t) = m-1) \\ &+ \sum_{k=2}^{m} p_{m-k}^{(i)}(t)\mathbb{P}(R^i(t+h) - R^i(t) = k | R^i(t) = m-k). \end{aligned}$$

The last sum is $o(h)$ for $m > 2$ and absent for $m \leq 1$. For $m = 0$ only the first sum appears. Inserting the relevant expressions, and doing some

[2]With the symbol $o(h)$ we collect remainder terms that go to zero as $h \to 0$.

elementary manipulations, gives us

$$\frac{p_m^{(i)}(t+h) - p_m^{(i)}(t)}{h} = -\nu[p_m^{(i)}(t) - p_{m-1}^{(i)}(t)] + o(h). \tag{3.2.4}$$

If (3.2.3) holds uniformly in t, we can replace $t = s - h$ and get

$$\frac{p_m^{(i)}(s) - p_m^{(i)}(s-h)}{h} = -\nu[p_m^{(i)}(s-h) - p_{m-1}^{(i)}(s-h)] + o(h). \tag{3.2.5}$$

Now we see that $|p_m^{(i)}(t+h) - p_m^{(i)}(t)| \leq o(h) \to 0$ as $h \downarrow 0$. So $p_m^i(\cdot)$ is continuous. Then we see that

$$\lim_{h \nearrow 0} \left| \frac{p_m^{(i)}(s) - p_m^{(i)}(s-h)}{h} \right| = \lim_{h \searrow 0} \left| \frac{p_m^{(i)}(t+h) - p_m^{(i)}(t)}{h} \right|,$$

showing differentiability for $t > 0$. At $t = 0$ only the right-derivative exists. Hence, for all $m \in \mathbb{N}, t \geq 0, i \in \mathcal{I}$ we have to solve the dynamical system

$$\dot{p}_m^{(i)}(t) = -\nu[p_m^{(i)}(t) - p_{m-1}^{(i)}(t)], \quad p_m^{(i)}(0) = \delta_{m,0} \tag{3.2.6}$$

$$\dot{p}_0^{(i)}(t) = -\nu p_0^{(i)}(t) \tag{3.2.7}$$

The most elegant approach to solve this countable infinite system of *Kolmogorov forward equations* is by using generating function techniques (see e.g. Grimmett and Stirzaker, 2001, ch. 5). Let $G^{(i)}(t,x) := \sum_{k=0}^{\infty} x^k p_k^{(i)}(t)$ be the generating function of the probability mass function $\{p_m^{(i)}(t)\}_{m=0}^{\infty}$, $x \in (0,1]$. Then

$$\begin{aligned}
\frac{\partial G^{(i)}(t,x)}{\partial t} &= \sum_{k=0}^{\infty} x^k \dot{p}_k^{(i)}(t) \\
&= \sum_{k=0}^{\infty} x^k \nu[-p_k^{(i)}(t) + p_{k-1}^{(i)}(t)\mathbb{1}_{k\geq 1}] \\
&= -\nu \sum_{k=0}^{\infty} x^k p_k^{(i)}(t) + \nu \sum_{k=1}^{\infty} x^k p_{k-1}^{(i)}(t) \\
&= -\nu G^{(i)}(t,x) + \nu x G^{(i)}(t,x) = \nu(x-1)G^{(i)}(t,x)
\end{aligned}$$

with boundary condition (for proper normalization) $G^{(i)}(t,1) = 1$ for all $t \geq 0$. Solving for $G^{(i)}(t,x)$, we get

$$\begin{aligned}
&\frac{\partial}{\partial t}\left[G^{(i)}(t,x)\exp(-\nu(x-1)t)\right] = 0 \quad |\text{integrate over } [0,T] \\
&\Rightarrow G^{(i)}(T,x)\exp(-\nu(x-1)T) - G^{(i)}(0,x) = A \\
&\Leftrightarrow G^{(i)}(t,x) = \exp(\nu(x-1)t)
\end{aligned}$$

Suppose $R^i(t)$ is Poisson distributed with parameter νt. Then the generating function is

$$\text{Poiss}(t,x) = \sum_{k=0}^{\infty} x^k \exp(-\nu t)\frac{(\nu t)^k}{k!} = \exp(-\nu t)\sum_{k=0}^{\infty}\frac{(\nu x t)^k}{k!} = \exp(\nu t(x-1)).$$

showing that $\text{Poiss}(t,x) = G^{(i)}(t,x)$ for all $t \geq 0$ and so $(R^i(t))_{t\geq 0}$ is a Poisson process. □

To establish a connection with the macro level, define the compound process $R(t) := \sum_{i=1}^{N} R^i(t)$. The distribution of this process can be easily derived from the infinitesimal descriptions of all individual processes, and is in fact a simple consequence of the superposition principle of independent Poisson processes (again see Grimmett and Stirzaker (2001) for background information).

$$\mathbb{P}(R(t+h)-R(t)=0|R(s);0\leq s\leq t) = \mathbb{P}\left(\bigcap_{i=1}^{N}\{R^i(t+h)-R^i(t)=0\}|R(s);0\leq s\leq t\right)$$

$$\begin{aligned}
&= \prod_{i=1}^{N}\mathbb{P}(R^i(t+h)-R^i(t)=0|R(s);0\leq s\leq t) \quad |\text{by independence}\\
&= (1-\nu h+o(h))^N \quad |\text{by (3.2.3)}\\
&= (1-\nu h)^N + o(h)\\
&= 1 - N\nu h + o(h).
\end{aligned}$$

$$\begin{aligned}
&\mathbb{P}(R(t+h)-R(t)=1|R(s);0\leq s\leq t)\\
&= \mathbb{P}\left(\bigcup_{i=1}^{N}\bigcap_{j\neq i}\{R^i(t+h)-R^i(t)=1\}\cap\{R^j(t+h)-R^j(t)=0\}|R(s);0\leq s\leq t\right)\\
&= \sum_{i=1}^{N}\mathbb{P}(R^i(t+h)-R^i(t)=1|R(s);0\leq s\leq t)\\
&\times \mathbb{P}\left(\bigcap_{j\neq i}\{R^j(t+h)-R^j(t)=0\}|R(s);0\leq s\leq t\right)\\
&= \sum_{i=1}^{N}\mathbb{P}(R^i(t+h)-R^i(t)=1|R(s);0\leq s\leq t)\\
&\times \prod_{j\neq i}\mathbb{P}(R^j(t+h)-R^j(t)=0|R(s);0\leq s\leq t)\\
&= N(\nu h+o(h))(1-\nu h+o(h))^{N-1}\\
&= N\nu h(1-(N-1)\nu h)+o(h) = N\nu h + o(h)
\end{aligned}$$

From the construction of the aggregate counting process $R(t)$, it is clear that, with probability 1, the following statement holds:

"$R(t)$ jumps by 1$\Leftrightarrow$ $(\exists! i \in \mathcal{I})$: $R^i(t)$ jumps by 1"

Further, we have just shown that the aggregate process $R(t)$ is Poisson with parameter $N\nu$. The same computations we have employed to determine the distribution of the individual counting process $(R^i(t))$ go through for the process $(R(t+s)-R(t))_{s\geq 0}$. To see this, let $\mathcal{F}_t = \sigma(\{R(\tau) : \tau \in [0,t]\})$ be the σ-algebra generated by sets $\{R(\tau) = m_\tau\}$, $\tau \in \mathbb{Q}$ and $m_\tau \in \mathbb{N}_0$, set $\mu := N\nu$ and define $\Delta_h := R(t+h)-R(t)$ the increment of the process at t. Note that, by right-continuity of the process $R(t)$ the consideration of such a σ-algebra makes sense, because any sample trajectory $\{R(\tau), \tau \leq t\}$ can be represented as a countable union of sets of the form $\{R(\tau_1) = m_{\tau_1}, R(\tau_2) = m_{\tau_2}, \ldots\}$ for some sequence of rational numbers $(\tau_q)_{q=1}^{\infty}$ lying in $[0,t]$, and a non-decreasing (and bounded) sequence of integers $(m_{\tau_q})_{q=1}^{\infty}$. Further, call now $p_m(s) = \mathbb{P}(R(t+s) - R(t) = m|\mathcal{F}_t)$. We deduce that

$$p_n(s+h) = \sum_{k=0}^{n} p_{n-k}(s)\mathbb{P}(\Delta_{s+h} - \Delta_s = k|\mathcal{F}_t, \Delta_s = n-k).$$

Given the infinitesimal description of our process as above, we can derive a system of differential equations

$$\dot{p}_n(s) = -\mu[p_{n-1}(s) - p_n(s)]$$

whose unique solution we already know. Hence, we have shown that the process $(\tilde{R}(s))_{s\geq 0} := (R(t+s)-R(t))_{s\geq 0}$ is a Poisson process with parameter μ.[3]

Conditional on the event that the counter $R(t)$ moved by 1 unit at time t, we can determine the probability with which player i has caused this increase in the counter. To do so, we introduce two other families of random variables which highlight the structure of a counting process. Let $J_m^A := \inf\{t \geq 0 : R(t) = m\}$ be the time at which m action update opportunities arrived to the society. From this define $T_m^A := J_m^A - J_{m-1}^A$ as the time the population has to wait between the $(m-1)$-th and the m-th action update opportunity. We derive now the distribution of these holding times, using some properties established in Section 3.2.1. Note that the event $\{R(t+s) - R(t) = 1\}$ is equivalent to the event $\{T_{m+1}^A < s\}$ on $\{R(t) = m\}$, for some $m \in \mathbb{N}_0$. It follows that

$$\mathbb{P}(T_1^A > t) = \mathbb{P}(R(t) = 0) = \exp(-\mu t).$$

[3]In fact, what we have just shown is the Markov property of the Poisson process.

Hence, in the notation of Section 3.2.1, $\psi^A(t+h) = \psi^A(t)\psi^A(h)$, and so $(T_n^A)_{n=0}^{\infty}$ is a sequence of i.i.d exponentially distributed random variables with mean $1/\mu$.
Since the aggregate process $(R(t))_{t\geq 0}$ and the individual processes $(R^i(t))_{t\geq 0}, i \in \mathcal{I}$ are fundamentally linked (the first increases by one iff exactly one individual counter increased by 1), we can also determine the probability with which the counter of player i increased by 1, conditional that R jumped by 1. Let $(T_m^{A,j})_{m\geq 1}$ be the family of waiting times of player j. Then the event $\{T_m^A > t\} = \bigcap_{i=1}^N \{T_{m^i}^{A,i} > t\}$ where the integers $m^1, \ldots, m^N$ satisfy $\sum_{i=1}^N m^i = m$. By definition of the aggregate process $\sum_{i=1}^N \sup_{s<J_m^A} R^i(s) = m-1$. Thus, we can alternatively define $T_m^A = \min\{T_{m^1}^{A,1}, \ldots, T_{m^N}^{A,N}\}$. Still differently, one can define the random variable so that $T_{(t)}^A = \min_{i\in\mathcal{I}}\{T_{R^i(t)}^{A,i}\}$. Take a sequence $0 < t_1 < t_2 < \ldots < t_n$ and let $\mathcal{F}_{t_n} = \{R(t_1) = m_1, \ldots, R(t_n) = m_n\}$. The event $B = \{R^i(t_n+s) - R^i(t_n) = 1\} \cap \{R(t_n+s) - R(t_n) = 1\}$ gives the event that player i is the one who caused the process R to move. By construction we have $B = \{T_{R^i(t_n)}^{A,i} = T_{(t_n)}^A\} \cap \{T_{(t_n)}^A \leq s\}$ on $\mathcal{F}_{t_n}$. Hence

$$
\begin{aligned}
\mathbb{P}(B|\mathcal{F}_{t_n}) &= \mathbb{P}\left(T_{R^i(t_n)}^{A,i} = T_{(t_n)}^A \;\&\; T_{(t_n)}^A \leq s|\mathcal{F}_{t_n}\right) \\
&= \mathbb{P}\left(\bigcap_{j\neq i}\{T_{R^j(t_n)}^{A,j} > T_{R^i(t_n)}^{A,i}\} \cap \{T_{R^i(t_n)}^{A,i} \leq s\}|\mathcal{F}_{t_n}\right) \\
&= \int_0^s \left[\mathbb{P}\left(\bigcap_{j\neq i}\{T_{R^j(t_n)}^{A,j} > \tau\}|\mathcal{F}_{t_n}\right)\right] \nu\exp(-\nu\tau)\,\mathrm{d}\,\tau \\
&= \int_0^s \left[\prod_{j\neq i}\mathbb{P}(T_{R^j(t_n)}^{A,j} \geq \tau|\mathcal{F}_{t_n})\right] \nu\exp(-\nu\tau)\,\mathrm{d}\,\tau \\
&= \int_0^s \left[\exp(-(N-1)\nu\tau)\right] \nu\exp(-\nu\tau)\,\mathrm{d}\,\tau \\
&= \frac{1}{N}\left(1-\exp(-\mu s)\right)
\end{aligned}
$$

for all $s \geq 0$. For $s \to \infty$ the probability goes to $1/N$. We see that the event indicating that player i is the one who made the action update, and the actual waiting time before he does so, are independent on $\mathcal{F}_{t_n}$. Every player is equally likely to be the one to update, and the aggregate waiting time $T_m^A \sim$Exp$(N\nu)$. The rate of this distribution defines the time scale of the process of action adjustment, $\tau_a := N\nu$. Observe that in chapter 2 we have called this the *volume* of the action adjustment subprocess.
Conditional on the event of an action revision opportunity, player i employs

a behavioral rule $b^i(\cdot|\omega)$, which is a probability distribution on the common action set $\mathcal{A}$. The models of chapters 4 and 5 will assume that this choice function is a smoothed best response, where the smoothing is done by a logistic function. The derivation of this behavioral rule, known as the logit choice-function, is well-known from the literature on discrete choice (see Anderson et al. (1992), or Hofbauer and Sandholm (2002) for a recent discussion).

Link creation

The process of link creation proposed in chapters 4 and 5 follows essentially the "stochastic-actor model", developed by Snijders (2001), which provides a micro-foundation for the p^*-models of social network analysis.[4] We will see that the rate function of the individuals of chapter 2 corresponds now to the intensity of the players' Poisson processes. The random vectors $(w_j^{i,\beta})_{\in\mathcal{I}}$ are derived from a random utility model.
When asked to create a fresh link, player i will maximize the random utility

$$u(\alpha^i, \alpha^j) + \varepsilon^i_j$$

for all other players j, which are currently not in his neighborhood. Conditional on the event that player i has the chance to form a link, she will do so with probability 1 (no chance is wasted). The probability that the edge (i, j) is added to the network is given by

$$w^i_j(\omega) := \mathbb{P}\left(\forall k \notin \bar{\mathcal{N}}^i(\omega) : u(\alpha^i, \alpha^j) + \varepsilon^i_j \geq u(\alpha^i, \alpha^k) + \varepsilon^i_k|\omega\right), \qquad (3.2.8)$$

where $\omega = (\alpha, g)$ is the population state at the time where player i gets the chance to create a link.[5] The models presented in chapters 4 and 5 propose a very specific form for the choice probabilities (3.2.8). There we assume that

$$w^{i,\beta}_j(\omega) = \frac{\exp(u(\alpha^i, \alpha^j)/\beta)}{\sum_{k\notin\bar{\mathcal{N}}^i(\omega)} \exp(u(\alpha^i, \alpha^k)/\beta)}.$$

This functional form assumes that a player samples another agent from the society with a probability that is monotonically increasing in the per-interaction payoff. The scale parameter β controls the influence of interaction

[4]See the discussion in section 1.2 of chapter 1.

[5]Note that for a player who is completely connected to the rest of the society (i.e. $|\bar{\mathcal{N}}^i(\omega)| = N-1$) the set determining the probability above is the empty set. Anyway, we will time the stochastic process $(Y^\beta(t))_{t\geq 0}$ in such a way that no completely connected player will receive the opportunity to create a link, and so this situation must not be treated separately.

payoffs on the probability. Very high levels of β mean that all players have more or less the same probability of being selected. Very low levels of β imply that the probability puts more mass to players which produce relatively high per-interaction utility. To arrive at this logit formula one assumes that the noise terms ε_j^i are i.i.d. double exponential distributed, hence have a density

$$f(x) = \frac{1}{\beta}\exp(-x/\beta - \gamma)\exp\left[-\exp(-x/\beta - \gamma)\right]$$

where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant, and cumulative distribution function $F(x) = \exp\left[-\exp\left(-x/\beta - \gamma\right)\right]$.[6] One can show that this distribution has mean 0 and variance $\frac{\beta^2\pi^2}{6}$. It is helpful to remember the formula of the variance, especially in view of stochastic stability, i.e. when $\beta \downarrow 0$. In case of this special class of perturbation, we can compute the probability w_j^i explicitly as follows;

$$\begin{aligned}
&\mathbb{P}\left(\varepsilon_k^i \leq u(\alpha^i, \alpha^j) - u(\alpha^i, \alpha^k) + \varepsilon_j^i \ \forall k \notin \bar{\mathcal{N}}^i(\omega)|\omega\right) \\
&= \int_{-\infty}^{+\infty} \mathbb{P}(\bigcap_{k \neq j; k \notin \bar{\mathcal{N}}^i(\omega)} \{\varepsilon_k^i \leq u(\alpha^i, \alpha^j) - u(\alpha^i, \alpha^k) + x\}|\omega) f(x)\,\mathrm{d}\,x \\
&= \int_{-\infty}^{+\infty} \left[\prod_{k \neq j; k \notin \bar{\mathcal{N}}^i(\omega)} \mathbb{P}(\varepsilon_k^i \leq u(\alpha^i, \alpha^j) - u(\alpha^i, \alpha^k) + x|\omega)\right] f(x)\,\mathrm{d}\,x
\end{aligned}$$

Now make a substitution of variables:

$$t(x) := \exp(-x/\beta - \gamma), \ (\forall j \notin \bar{\mathcal{N}}^i(\omega)) : y_{ij} := \exp(u(\alpha^i, \alpha^j)/\beta)$$

Then we see that $\mathrm{d}\,t = (-1/\beta)\exp(-x/\beta - \gamma)\,\mathrm{d}\,x$ and $t(+\infty) = 0, t(-\infty) = +\infty$. Hence, plugging this into the integral gives

$$\begin{aligned}
&\int_0^{+\infty} \left[\prod_{k \neq j; k \notin \bar{\mathcal{N}}^i(\omega)} \exp\left(-t\frac{y_{ik}}{y_{ij}}\right)\right] \exp(-t)\,\mathrm{d}\,t \\
&= \int_0^{\infty} \prod_{k \notin \bar{\mathcal{N}}^i(\omega)} \exp\left(-t\frac{y_{ik}}{y_{ij}}\right) \mathrm{d}\,t \\
&= \frac{y_{ij}}{\sum_{k \notin \bar{\mathcal{N}}^i(\omega)} y_{ik}} = \frac{\exp(u(\alpha^i, \alpha^j)/\beta)}{\sum_{k \notin \bar{\mathcal{N}}^i(\omega)} \exp(u(\alpha^i, \alpha^k)/\beta)}
\end{aligned}$$

By pre-multiplying this expression with the indicator $(1 - g_j^i)$, we can extend the choice probability $w_j^i(\omega)$ to hold for all $N - 1$ potential linking partners.

[6] It is also known as the Gumbel distribution or type 1 extreme-value distribution.

The second ingredient of the stochastic-actor model is a so-called rate function, as has been introduced in chapter 2. The rate function is an agent-specific mapping $\lambda^i : \Omega \to \mathbb{R}_+$. The value of this rate function is coupled with the sample path $t \mapsto Y^\beta(t)$. For a random vector $(J_1, \dots, J_m) \in \mathbb{R}_+^m$ of jump times, let $(t_1, t_2, \dots, t_m)$ be one particular realization with corresponding states $(\omega_1, \dots, \omega_m) \in \Omega^m$ visited by the process at these points of time. Then the rate function of any player $i \in \mathcal{I}$ has the following "sample path":

$$
\begin{aligned}
&[0, t_1) : \ \lambda^i(\omega_0) \quad \omega_0 = Y^\beta(0) \text{ the given initial state} \\
&[t_1, t_2) : \ \lambda^i(\omega_1) \\
&\vdots \\
&[t_{m-1}, t_m) : \ \lambda^i(\omega_{m-1})
\end{aligned}
$$

Hence, there is a coupling between the process $(Y^\beta(t))_{t\geq 0}$ and a processes $(L^i(t))_{t\geq 0}$, where the integer $L^i(t)$ measures the number of link creation opportunities player i had up to time t. Let $(T^{L,i}_{Y^\beta(t)})_{t\geq 0}$ denote the waiting times of player i, in dependence of the state $Y^\beta(t)$. The functioning of the link creation process of the individual players is analogous to the action update process, though calculations cannot be done in such an explicit way. On the event $\{J_m \leq t < J_{m+1}\}$ and given the information $\mathcal{F}_t = \sigma(\{Y^\beta(s); s \in [0, t]\})$, a randomly selected individual gives "birth" to a new link at an independent rate $\lambda^i(Y^\beta(t))$, meaning that the conditional probability with which player i creates a link in a small time interval $[t, t+h], h \downarrow 0$ is $\lambda^i(Y^\beta(t))h + o(h)$. Formally, on the set $\{Y^\beta(t) = \omega_m\}$, we have

$$
\begin{aligned}
&\mathbb{P}(L^i(t+h) - L^i(t) = 1 | Y^\beta(s); 0 \leq s \leq t \ \& \ J_m \leq t < J_{m+1}) = \lambda^i(\omega_m)h + o(h) \\
&\mathbb{P}(L^i(t+h) - L^i(t) > 1 | Y^\beta(s); 0 \leq s \leq t \ \& \ J_m \leq t < J_{m+1}) = o(h)
\end{aligned}
$$

for all $i \in \mathcal{I}$. For general $h \geq 0$, this implies that the distribution of waiting times $T^{L,i}$ has to be defined for the several intervals $[J_m, J_{m+1})$. For a given $t \in [J_m, J_{m+1})$, or equivalently $S_{m+1} > t - J_m$, the probability that any player $i \in \mathcal{I}$ has to wait for additional $h \geq 0$ time units, is consequently

$$
\begin{aligned}
&\mathbb{P}(L^i(t+h) - L^i(t) = 0 | Y^\beta(s), 0 \leq s \leq t \ \& \ J_m \leq t < J_{m+1}) = \\
&= \exp(-\lambda^i(\omega_m)h)\mathbb{1}_{\{S_{m+1} > t+h-J_m\}}.
\end{aligned}
$$

Would h be so large that $S_{m+1} \leq t + h - J_m$, or equivalently, $J_{m+1} \leq t + h$ then the state ω must have changed in the meanwhile and consequently the distribution of the waiting times must be adjusted accordingly.
On $\mathcal{F}_t$ the waiting times of the players are independent of each other. Hence,

we can again consider a compound arrival process $L(t) := \sum_{i=1}^{N} L^i(t)$, indicating the total number of link creation opportunities the society had up to time t. To determine the distribution of this process, one can proceed as in the action update process, but more information is needed. One has to condition on $\mathcal{F}_t$ and $\{t \geq J_m\} \cap \{S_{m+1} > t + h - J_m\}$, so to be sure that the process did not perform its m-th jump yet. On this event, the processes $(L^i(t))_{t\geq 0}$ run independently of each other, and the superposition principle of independent Poisson processes applies to tell us that $(L(t))_{t\in[J_m,J_{m+1})}$ is a Poisson process with rate $\bar{\lambda}(\omega_m) := \sum_{i\in\mathcal{I}} \lambda^i(\omega_m)$. This relation holds for all $m \in \mathbb{N}$. This is exactly the *volume* of the link creation process, as defined in chapter 2. By the infinitesimal characterization of Poisson processes we see, for $h \downarrow 0$, that

$$\mathbb{P}(L(t+h) - L(t) = 1 | Y^\beta(s); 0 \leq s \leq t \,\&\, J_m \leq t < J_{m+1}) = \bar{\lambda}(\omega_m)h + o(h),$$
$$\mathbb{P}(L(t+h) - L(t) > 1 | Y^\beta(s); 0 \leq s \leq t \,\&\, J_m \leq t < J_{m+1}) = o(h).$$

By construction of the compound process, if a link creation opportunity comes to the society, almost surely only a single player will get the chance to exhibit it. Thus, analogous to action updating, the micro and the macro level are fundamentally linked by the statement:

> "If $L(t) - \sup_{s<t} L(s) = 1 \Leftrightarrow (\exists!\, i \in \mathcal{I}) : L^i(t) - \sup_{s<t} L^i(s) = 1$ with probability 1."

Given this relationship, the waiting times of the process L, denoted by T^L, can be defined by $T^L_{(t)} = \min\{T^{L,1}_{(t)}, \ldots, T^{L,N}_{(t)}\}$, where the notation $T^L_{(t)} \equiv T^L_{Y^\beta(t)}$ is chosen for convenience. Let $G_t = \{t \geq J_m\} \cap \{S_{m+1} > t + h - J_m\}$ and $B = \{L^i(t+h) - L^i(t) = 1\} \cap \{L(t+h) - L(t) = 1\}$, i.e. the event that a link creation opportunity arrived at the population at time $t+h$, and player i is the one who went for it. Since $\{L(t+h) - L(t) \geq 1\} = \{T^L_{(t)} \leq h\}$, we get the following result

$$\begin{aligned}
\mathbb{P}(B|\mathcal{F}_t, G_t) &= \mathbb{P}(\{T^L_{(t)} = T^{L,i}_{(t)}\} \cap \{T^L_{(t)} \leq h\} | \mathcal{F}_t, G_t) \\
&= \mathbb{P}(\{T^{L,i}_{(t)} \leq h\} \cap \bigcap_{j\neq i} \{T^{L,j}_{(t)} > T^{L,i}_{(t)}\}) | \mathcal{F}_t, G_t) \\
&= \int_0^h \prod_{j\neq i} \mathbb{P}(T^{L,j}_{(t)} > s | \mathcal{F}_t, G_t) \lambda^i(\omega_m) \exp(-\lambda^i(\omega_m)s)\, \mathrm{d}\, s \\
&= \frac{\lambda^i(\omega_m)}{\bar{\lambda}(\omega_m)} (1 - \exp(-\bar{\lambda}(\omega_m)h))
\end{aligned}$$

where one uses (i) rules of conditional probability, (ii) conditional independence of the waiting times and (iii) conditional on $\mathcal{F}_t$ and G_t each waiting time is $T^{A,j}_{(t)} \sim \text{Exp}(\lambda^j(\omega_m))$, in this ordering. This tells us that the probability with which player i receives the link creation opportunity is proportional to her rate function, and the distribution of link creation opportunities is conditionally independent of this event.

Link destruction

A network is a list of edges $g = (g^i_j)_{1 \leq i \leq j \leq N}$. The degree of individual i is defined as $\kappa^i(g) := \sum_{j=1}^N g^i_j$. The total degree of a network is defined as the sum of all degrees, $\sum_{i=1}^N \kappa^i(g) = 2\sum_{i=1}^N \sum_{j>i} g^i_j$. Hence, the total degree of a network is a statistic where edges are counted twice in an undirected graph. To avoid this double counting, define the "effective degree" of the network prevailing at population state ω as $e(\omega) := \sum_{i=1}^N \sum_{j>i} g^i_j$. The set $\mathcal{E}(\omega)$ collects all the unordered pairs of links in the network g at ω. For every such edge, define the *survival time* $T^{D,(i,j)}_{(t)}$ on $\mathcal{F}_t$ and G_t. In chapter 4 we will assume that the survival times are i.i.d Exp(ξ) distributed. The model in chapter 5 extends this, by allowing the survival times to be edge-specific, i.e. the expected life-time of a currently active edge is ξ_{ij} (and this function may also depend on the noise level β). We derive the distribution of the survival terms for the more general case, since it comes without any additional costs. Let $(D(t))_{t \geq 0}$ denote the random "death" process acting on the set of edges $\mathcal{E}$, with the interpretation that $\{D(t) = (i,j)\}$ is the event that the edge $(i,j) \in \mathcal{E}$ is destroyed at time t. Destructive events depend on survival times. Therefore, the smallest survival time, measured at some time $t > 0$, determines the distribution of $(D(t))_{t \geq 0}$. This random variable is defined as $T^D_{(t)} := \inf_{i \in \mathcal{I}} \left\{ \inf_{j \in \{i,\dots,N\}} T^{D,(i,j)}_{(t)} \right\}$ where we choose $\inf \emptyset = \infty$.[7] On $\{Y^\beta(t) = \omega\}$ define $\mathcal{E}^*(\omega) := \mathcal{E}(\omega) \setminus \{(i,j)\}$, so that $|\mathcal{E}^*(\omega)| = e(\omega) - 1$.

[7] This simply states that a loner does not lose any (non-existing) links.

Then, we can calculate

$$
\begin{aligned}
&\mathbb{P}(\{T^D_{(t)} \leq h\} \cap \{T^D_{(t)} = T^{D,(i,j)}_{(t)}\}|\mathcal{F}_t, G_t) = \\
&= \mathbb{P}[\{T^{D,(i,j)}_{(t)} \leq h\} \cap \bigcap_{(k,l)\in\mathcal{E}^*(Y^\beta(t))} \{T^{D,(k,l)}_{(t)} > T^{D,(i,j)}_{(t)}\}|\mathcal{F}_t, G_t] \\
&= \int_0^h \mathbb{P}[\bigcap_{(k,l)\in\mathcal{E}^*(Y^\beta(t))} \{T^{D,(k,l)}_{(t)} > s\}|\mathcal{F}_t, G_t]\xi_{ij}\exp(-\xi_{ij}s)\,\mathrm{d}\,s \\
&= \int_0^h \prod_{(k,l)\in\mathcal{E}^*(Y^\beta(t))} \mathbb{P}[T^{D,(k,l)}_{(t)} > s|\mathcal{F}_t, G_t]\xi_{ij}\exp(-\xi_{ij}s)\,\mathrm{d}\,s \\
&= \int_0^h \exp\left(-\sum_{(k,l)\in\mathcal{E}^*(Y^\beta(t))} \xi_{kl}s\right)\xi_{ij}\exp(-\xi_{ij}s)\,\mathrm{d}\,s \\
&= \frac{\xi_{ij}g_{ij}}{\sum_{(k,l)\in\mathcal{E}(Y^\beta(t))}\xi_{kl}}\left[1-\exp\left(-\sum_{(k,l)\in\mathcal{E}(Y^\beta(t))}\xi_{kl}h\right)\right]
\end{aligned}
$$

It follows that the conditional probability that edge (i, j) becomes destroyed is exactly

$$
\frac{\xi_{ij}g_{ij}}{\sum_{(k,l)\in\mathcal{E}(Y^\beta(t))}\xi_{kl}}.
$$

3.3 Construction of the family $\mathcal{M}^\beta$ in discrete time

In this section we introduce a discrete-time version of a co-evolutionary model with noise, based on the general discussion of chapter 2. There are two good reasons why a discrete-time dynamics is interesting. First, one might be skeptical against the use of continuous-time models per se, if they are not derived from suitable limiting operations (see Binmore et al. (1995), Binmore and Samuelson (1997) for a discussion on this point). Second, to implement numerical simulations one needs a more algorithmic discrete-time dynamic than the continuous-time dynamic. A general way to couple discrete-time Markov chains to continuous-time Markov processes is via its so-called *embedded jump chain*. Section 3.4 is a self-contained discussion on this topic, and I refer to the books by Stroock (2005) and Norris (1997) for more details. In chapter 2 a Markov chain $X^\beta = (X^\beta_n)_{n\in\mathbb{N}}$ has been defined whose transition matrix was denoted by $\mathbf{K}^\beta = [K^\beta(\omega, \omega')]_{\omega,\omega'\in\Omega}$. Elements of this matrix are the one-step transition probabilities of X^β, i.e. for all $\omega, \omega' \in \Omega$

$$\mathbb{P}(X^\beta_{n+1} = \omega' | X^\beta_n = \omega) = K^\beta(\omega, \omega'). \tag{3.3.1}$$

Like the continuous-time co-evolutionary model of Section 3.2, this discrete-time process can be viewed from a micro-perspective and a macro-perspective. At the micro level we specify what happens when a player receives an action adjustment/link creation opportunity, or when a link destruction event takes place. At the macro-level the probabilities of these events is specified. Hence, at the macro-level we specify a distribution $q(\omega) = (q_1(\omega), q_2(\omega), q_3(\omega))$, where $q_1(\omega)$ is the probability that *some* player receives an action adjustment opportunity when the current population state is ω, and $q_2(\omega)+q_3(\omega) = 1 - q_1(\omega)$ is the probability that the network changes. The micro-level is, as in the continuous-time theory, separated into the following three categories.

Action adjustment: Conditional on an action adjustment event, every player receives a switching opportunity with equal probability $1/N$. The volume of the action revision process is defined as $N\nu$. Conditional on receiving a revision opportunity, player i selects action a_v with probability $b^i(a_v|\omega)$.

Link creation: As in the continuous-time theory of Section 3.2 assume that each player possesses a rate function λ^i. Define the volume of the link creation process $\bar{\lambda}$. The conditional probability that player i receives a link creation opportunity, starting from ω, is $\lambda^i(\omega)/\bar{\lambda}(\omega)$. The conditional probability that the edge (i, j) is formed is then

$$\bar{w}^\beta_{ij}(\omega) := \frac{1}{\bar{\lambda}(\omega)} \left(\lambda^i(\omega) w^{i,\beta}_j(\omega) + \lambda^j(\omega) w^{j,\beta}_i(\omega) \right). \tag{3.3.2}$$

Link destruction: The probability that link (i, j) becomes destroyed is given by a conditional probability $v^\beta_{ij}(\omega)$, satisfying certain conditions stated in chapter 2. These functions are a weighting system on the set of existing edges. The higher the weight of an edge, the more likely it is that the edge will be destroyed. All edge weights are collected in a matrix $\mathbf{V}^\beta$, and I assume that the total volume of the link destruction process is of the form $\bar{\xi}(\omega) := \xi f(\omega, \mathbf{V}^\beta)$, for some bounded function f which is only required to take the value 0 at the empty graph. Certainly, more structural properties would make the model more meaningful, but for the general discussion this is all we need. In chapter 4 we will assume that a selected link is destroyed at rate 1, i.e. $v^\beta_{ij}(\omega) \equiv \frac{g_{ij}}{e(\omega)}$ and $f(\omega, \mathbf{V}^\beta) = e(\omega)$.

To summarize, the transition Matrix of X^β is given by

$$K^\beta(\omega,\hat{\omega}) = \begin{cases} q_1(\omega)\frac{1}{N}b^\beta(a|\omega) & \text{if } \hat{\omega} = (\alpha_i^a, g), \\ q_2(\omega)\bar{w}_{i,j}^\beta(\omega) & \text{if } \hat{\omega} = (\alpha, g \oplus (i,j)), \\ q_3(\omega)v_{ij}^\beta(\omega) & \text{if } \hat{\omega} = (\alpha, g \ominus (i,j)), \\ 0 & \text{otherwise.} \end{cases} \tag{3.3.3}$$

See chapter 2 for more details.

3.4 The law of $\mathcal{M}^\beta$ in continuous time

Coming back to the continuous-time Markov process $(Y^\beta(t))_{t\geq 0}$ we have determined the several conditional probabilities that a player receives an action adjustment opportunity, a link gets activated, or a link becomes destroyed. Now we turn to the macro level of the process. Section 3.4.1 uses the distribution of the waiting times T^R, T^L, T^D to derive the distribution of the holding times $(S_n)_{n\in\mathbb{N}}$ of the continuous-time process $(Y^\beta(t))_{t\geq 0}$. This will then be combined with the discrete-time process X^β of Section 3.3 in order to derive the *stochastic-semi group* of the continuous-time model $\mathcal{M}^\beta$.

3.4.1 The distribution of holding times

We have just derived the distribution of the waiting times T^R, T^L, T^D. Now we put these 3 sub-processes together in order to deduce some information on the shape of sample paths of the process $(Y^\beta(t))_{t\geq 0}$. Proposition 3.4.1 shows that the holding times of the process are exponentially distributed random variables, with state-dependent means. For simplicity, we do this for the case of constant volatility rates ξ, such as is also done in many other models on the evolution of networks and play (see e.g. Marsili et al., 2004, Ehrhardt et al., 2006a; 2008b). In chapter 2 we have called such models "volatility" models and referred to ξ as the rate of volatility.

Proposition 3.4.1. *Consider a volatility model with rate ξ. Conditional on the events $\mathcal{F}_t = \sigma\left(\{Y^\beta(s); 0 \leq s \leq t\}\right)$, $\{J_m \leq t < J_{m+1}\}$, and $\{Y^\beta(J_m) = \omega\}$, we have $S_m \sim \mathrm{Exp}(\Lambda(\omega))$, where $\Lambda(\omega) = \bar{\lambda}(\omega) + N\nu + e(\omega)\xi$.*

Proof. Conditional on $\mathcal{F}_t = \sigma\left(\{Y^\beta(s); 0 \leq s \leq t\}\right)$ and $\{J_m \leq t < J_{m+1}\}$ the event $\{S_{m+1} > h\} = \{Y^\beta(t+h) = Y^\beta(J_m)\}$ has the following probability distribution:

$$\begin{aligned}
&\mathbb{P}(S_{m+1} > h|\mathcal{F}_t \,\&\, J_m \leq t < J_{m+1}) = \mathbb{P}(\min\{T^A_{(t)}, T^L_{(t)}, T^D_{(t)}\} > h|\mathcal{F}_t \,\&\, J_m \leq t < J_{m+1}) \\
&= \mathbb{P}(\{S_{m+1} = T^A_{(t)}\} \cap \{T^A_{(t)} > h\}|\mathcal{F}_t \,\&\, J_m \leq t < J_{m+1}) \quad |(a) \\
&+ \mathbb{P}(\{S_{m+1} = T^L_{(t)}\} \cap \{T^L_{(t)} > h\}|\mathcal{F}_t \,\&\, J_m \leq t < J_{m+1}) \quad |(b) \\
&+ \mathbb{P}(\{S_{m+1} = T^D_{(t)}\} \cap \{T^D_{(t)} > h\}|\mathcal{F}_t \,\&\, J_m \leq t < J_{m+1}) \quad |(c)
\end{aligned}$$

Start with calculating (a):

$$\begin{aligned}
&\mathbb{P}(\{S_{m+1} = T^A_{(t)}\} \cap \{T^A_{(t)} > h\}|\mathcal{F}_t \,\&\, J_m \leq t < J_{m+1}) \\
&= \mathbb{P}(\{T^D_{(t)} > T^A_{(t)} \,\&\, T^L_{(t)} > T^A_{(t)}\} \cap \{T^A_{(t)} > h\}|\mathcal{F}_t \,\&\, J_m \leq t < J_{m+1}) \\
&= \int_h^\infty N\nu \exp(-\Lambda(Y^\beta(t))s)\,\mathrm{d}\,s \\
&= \frac{N\nu}{\Lambda(Y^\beta(t))} \exp(-\Lambda(Y^\beta(t))h)
\end{aligned}$$

By essentially the same technique, we can calculate (b) and (c) to get

$$\begin{aligned}
(b) &\Rightarrow \frac{\bar{\lambda}(Y^\beta(t))}{\Lambda(Y^\beta(t))} \exp(-\Lambda(Y^\beta(t))h) \\
(c) &\Rightarrow \frac{e(Y^\beta(t))\xi}{\Lambda(Y^\beta(t))} \exp(-\Lambda(Y^\beta(t))h)
\end{aligned}$$

and $\Lambda(Y^\beta(t)) := N\nu + e(Y^\beta(t))\xi + \bar{\lambda}(Y^\beta(t))$. Hence

$$\mathbb{P}(S_{m+1} > h|\mathcal{F}_t \,\&\, J_m \leq t < J_{m+1}) = \exp(-\Lambda(Y^\beta(t))h) \tag{3.4.1}$$

$\square$

3.4.2 The stochastic semi-group and its generator

Consider the co-evolutionary model with noise $\mathcal{M}^\beta$, whose sample path is recorded by functions $t \mapsto Y^\beta(t)$. From the construction presented in Section 3.2, we can deduce that $Y^\beta(\cdot)$ is right-continuous in t, and a sample path makes discrete jumps at random points of time, measured by the countable sequence of jump times $(J_n)_{n=0}^\infty$. The discrete-time process $X^\beta = (X^\beta_n)_{n=0}^\infty$, defined as $X^\beta_n = Y^\beta(J_n)$ for all $n \geq 0$, is called the *jump chain* of the process $(Y^\beta(t))_{t\geq 0}$. The distributional law of $(Y^\beta(t))_{t\geq 0}$ is closely related to the law of X^β and the distribution of the holding times $(S_n)_{n\in\mathbb{N}}$, which has been determined in Proposition 3.4.1. To be specific, the distribution of the class of Markov processes of Section 3.2 can be derived from the joint distribution of a Markov chain X^β and the holding times (S_n), by the identity

$$\begin{aligned}\mathbb{P}(Y^\beta(t) = \omega) &= \sum_{n=0}^{\infty} \mathbb{P}(X_n^\beta = \omega \ \& \ J_n \leq t < J_{n+1}) \\ &= \mathbb{P}(X_n^\beta = \omega \ \& \ J_n \leq t < J_{n+1} \text{ for some } n \geq 0).\end{aligned}$$

From Proposition 3.4.1 one sees that the distribution of the n-th holding time depends only on the state of the process after its $(n-1)$-st jump. Let $\mathcal{F}_{J_m} := \{A \in \mathcal{F}_t : A \cap \{J_m \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\}$ the σ-algebra generated by all sets $\{Y^\beta(s) = \omega\}$ for $\omega \in \Omega$ and $s \leq t$, and where the process makes exactly m-jumps.[8] The *strong Markov property* states that, conditional on $\mathcal{F}_{J_m}$, the Markov chain $(X_{m+n}^\beta)_{n\geq 1}$ and the holding times $(S_{m+n})_{n\geq 1}$ are independent of each other. From this conditional independence, it follows that

$$\begin{aligned}\mathbb{P}(Y^\beta(t) = \omega | \mathcal{F}_{J_m}) &= \mathbb{P}(X_{m+1}^\beta = \omega \ \& \ J_{m+1} - J_m \leq t | \mathcal{F}_{J_m}) \\ &= \mathbb{P}(S_{m+1} \leq t | \mathcal{F}_{J_m}) \mathbb{P}(X_{m+1}^\beta = \omega | \mathcal{F}_{J_m}).\end{aligned}$$

Let $\mathbf{K}^\beta = \left[K^\beta(\omega, \hat{\omega})\right]_{\omega,\hat{\omega}\in\Omega}$ be the transition matrix of the Markov chain X^β. Then, using Proposition 3.4.1, we can rewrite the above equation in the more concise version

$$\mathbb{P}(Y^\beta(t) = \omega | \mathcal{F}_{J_m}) = \left[1 - \exp(-\Lambda(Y^\beta(J_m))(t - J_m))\right] K^\beta(Y^\beta(J_m), \omega). \tag{3.4.2}$$

This shows that the distribution of the continuous-time Markov process $(Y^\beta(t))_{t\geq 0}$ is time-homogeneous, in the sense that only the time difference $(t - J_m)$ plays a role. Thus, let us define the transition probability of $(Y^\beta(t))_{t\geq 0}$ as

$$p_{\omega,\omega'}(t) := \mathbb{P}(Y^\beta(t) = \omega' | Y^\beta(0) = \omega). \tag{3.4.3}$$

On $\{Y^\beta(J_m) = \omega\}$, eq. (3.4.2) defines completely the *stochastic semi-group* $(\mathbf{P}(t))_{t\geq 0}$, with $\mathbf{P}(t) := [p_{\omega,\omega'}(t)]_{\omega,\omega'\in\Omega}$, as

$$p_{\omega,\omega'}(h) := [1 - \exp(-\Lambda(\omega)h)]\, K^\beta(\omega, \omega') \tag{3.4.4}$$

for all $\omega \neq \omega' \in \Omega$ and $h := t - J_m \geq 0$. Eq. (3.4.4) gives the probability that the process makes a jump in the time interval $[J_m, t]$, which, by time-homogeneity, can be translated to the interval $[0, h]$. Since the Markov matrix $\mathbf{K}^\beta$ has been explained in Section 3.3, the distribution of the continuous-time

[8] That this set is a meaningful objects relies on the deep result that jump times (and thus holding times) are *stopping times* of $(Y^\beta(t))_{t\geq 0}$. See Norris (1997) for a very accessible measure-theoretic introduction to stopping times.

co-evolutionary model with noise is completely specified.
A particular elegant formulation of the stochastic semi-group is obtained when one considers transitions in a very small time interval $[t, t+h]$. In this case, eq. (3.4.4) allows us to derive an illuminating connection between the transition probabilities and the *infinitesimal generator* of the process $(Y^{\beta}(t))_{t\geq 0}$. Consider a mapping $\eta^{\beta} : \Omega^2 \to \mathbb{R}$, with the properties

(G1) $-\infty < \eta^{\beta}(\omega \to \omega) \leq 0$, and

(G2) for all $\omega' \neq \omega$, $\eta^{\beta}(\omega \to \omega') \geq 0$, and

(G3) $\sum_{\omega' \in \Omega} \eta^{\beta}(\omega \to \omega') = 0$.

Proposition 3.4.2 (Infinitesimal definition of Markov processes)**.** *Consider the continuous-time co-evolutionary model with noise $\mathcal{M}^{\beta}$, with stochastic semi-group $(\mathbf{P}(t))_{t\geq 0}$ defined by (3.4.4). For all $t, h \geq 0$, conditional on $\{Y^{\beta}(t) = \omega\}$, $Y^{\beta}(t+h)$ is independent of $\mathcal{F}_t$, and as $h \downarrow 0$*

$$p_{\omega,\omega'}(h) = \delta_{\omega,\omega'} + \eta^{\beta}(\omega \to \omega')h + o(h) \tag{3.4.5}$$

for all ω, ω'. The operator $\boldsymbol{\eta}^{\beta} := [\eta^{\beta}(\omega \to \omega')]_{(\omega,\omega')\in\Omega^2}$ is called the infinitesimal generator of the Markov process $\mathcal{M}^{\beta}$.

Proof. Conditional independence follows from the definition of the stochastic semi-group. To see eq. (3.4.5), consider eq. (3.4.4) for $\omega \neq \omega'$, and let $h \downarrow 0$. This gives

$$\begin{aligned} p_{\omega,\omega'}(h) &= \Lambda(\omega)K^{\beta}(\omega,\omega')h + o(h) \\ &= \eta^{\beta}(\omega \to \omega')h + o(h) \end{aligned}$$

with $\eta^{\beta}(\omega \to \omega') := \Lambda(\omega)K^{\beta}(\omega,\omega')$. Now consider $\omega = \omega'$. The probability that the process stays constant in the interval $[t, t+h]$ is determined by two disjoint events. First, the process will stay constant if the subsequent jump occurs at some time after $t+h$. On $\{J_m \leq t < J_{m+1}\}$, for some $m \in \mathbb{N}$, this event is equivalent to the event $\{J_{m+1} - J_m > t+h\} = \{S_{m+1} > t+h\}$. Conditional on $\{Y^{\beta}(t) = \omega\}$, we know that $S_{m+1} \sim \mathrm{Exp}(\Lambda(\omega))$, and so the event $\{S_{m+1} > t+h\}$ appears with probability $\exp(-\Lambda(\omega)h)$. Second, the process will stay constant if the event $\{J_{m+1} - J_m \leq t+h \;\&\; X^{\beta}_{m+1} = \omega\}$ takes place. On $\{Y^{\beta}(t) = \omega\}$ this event has probability $[1 - \exp(-\Lambda(\omega)h)]\, K^{\beta}(\omega,\omega)$. Summarizing, we have

$$\begin{aligned} p_{\omega,\omega}(h) &= \exp(-\Lambda(\omega)h) + [1 - \exp(-\Lambda(\omega)h)]\, K^{\beta}(\omega,\omega) \\ &= 1 - \Lambda(\omega)\left(1 - K^{\beta}(\omega,\omega)\right)h + o(h) \quad |\text{for } h \downarrow 0 \\ &= 1 + \eta^{\beta}(\omega \to \omega)h + o(h) \end{aligned}$$

with $\eta^\beta(\omega \to \omega) := -\Lambda(\omega)\left(1 - K^\beta(\omega,\omega)\right)$. It is easy to verify that the generator satisfies the conditions (G1)-(G3). $\square$

Normally, the connection between the stochastic semi-group of a continuous-time process and its embedded jump chain is established by defining the latter with zero in its main diagonal, i.e. $K^\beta(\omega,\omega) = 0$. The resulting Markov chain would then not be aperiodic, making it a bad model for separate studies.[9] To avoid this, we have assumed that $K^\beta(\omega,\omega) > 0$. This might appear as a redundant condition, since a chain with $K^\beta(\omega,\omega) > 0$ can always be transformed into a chain with $K^\beta(\omega,\omega) = 0$.[10] However, we do this step in order to provide an alternative discrete-time model, having in mind this redundancy argument. In the subsequent chapters we will work with continuous-time Markov processes whose generators will have the structure as depicted above. A co-evolutionary model with noise as defined in chapter 2 was a discrete-time Markov chain with transition matrix $\mathbf{K}^\beta$. The generator of a continuous-time co-evolutionary model with noise is connected with this transition matrix by the general formula $\mathrm{diag}[\Lambda(\omega) : \omega \in \Omega](\mathbf{K}^\beta - \mathbf{Id})$.

[9] Aperiodicity ensures convergence of the Markov chain. The uniqueness of the invariant distribution is guaranteed by irreducibility alone.

[10] Simply define the transition probabilities of the new chain as $\tilde{K}^\beta(\omega,\omega) = 0$, and $\tilde{K}^\beta(\omega,\omega') = K^\beta(\omega,\omega')/(1 - K^\beta(\omega,\omega))$ for all $\omega \neq \omega'$.

Chapter 4

Potential games played in volatile environments

4.1 Introduction

The analysis of social networks has recently gained interest in various fields in the sciences and social sciences. By now there is a rich literature on social networks in economics; the textbooks by Jackson (2008) and Vega-Redondo (2007) give a concise overview on this emerging field. Recently, tools from evolutionary game theory have been used to study the co-evolution of networks and play. Models in this vein are Jackson and Watts (2002), Goyal and Vega-Redondo (2005), and Hojman and Szeidl (2006). Another type of model, which is more in the tradition of statistical physics, puts more weight on modeling the evolution of the network, without paying too much attention to the role of strategic interactions. Prominent examples are Ehrhardt et al. (2006b; 2008a;b). This paper aims to combine these two streams of literature in a very simplistic model. I present a stochastic co-evolutionary model which includes three sub-processes: action adjustment, link creation, or link destruction. These three sub-processes are combined into one continuous-time Markov process called a co-evolutionary model with noise. For positive noise levels the process is ergodic. For the class of potential games (Monderer and Shapley, 1996) many fundamental characteristics of the system are explicitly computable. Key to all the results in this paper is the closed-form expression of the invariant distribution. This probability distribution describes the long-run behavior of the system in two complementary ways. First, it gives us complete information on the joint probability distribution over action profiles and networks which governs the "equilibrium" of the stochastic dynamics. Second, by virtue of ergodicity, it gives us complete information on which states are more frequently observed over time compared to others. From the invariant distribution one can deduce the conditional probability distribution over networks for a fixed profile of actions. In the parlance of random graph theory this gives us the ensemble of random graphs. The interesting result is that the model generates so-called *inhomogeneous random graphs*. Inhomogeneous random graphs are a straightforward extension of the classical random graph model proposed by Erdös and Rényi (1960), where the probability with which two vertices are linked depend in some way on the characteristics of the vertices. Söderberg (2002) and Bollobás et al. (2007) are models in this direction. These papers fix the edge success probability at the outset. In this paper the edge success probability are a product of the long-run equilibrium of the system, hence we obtain them *endogenously*. Next, I provide an expression for the marginal distribution over action profiles in the population. This measure is interesting if one is not interested in the effects of the interaction structure. Finally, we explore the well-known relationship between potential maximizers and stochastic stability (for early work in this direction see for instance Blume, 1993, Young, 1998, ch.6). A fairly general argument is provided, showing that as noise vanishes the invariant distribution concentrates on the set of potential maximizers. At first sight, this might not be a too surprising result. However, former models were only concerned with fixed interaction structures, so the conclusion of our theorem extends the previous ones. Moreover, the argument presented in this paper is much more general than the proofs in the just mentioned literature. This technique allows to study the low-noise behavior of the invariant distribution also in more complicated models, as, for instance, the one which is going to be presented in chapter 5 of this thesis. Since the class of potential games is rather narrow, I also sketch briefly how the results obtained extend if the potential game assumption is dropped. In chapter 2 a rather general class of co-evolutionary dynamics has been presented, and I refer to this article for further details. However, many games arising in economic applications have this special structure. The most prominent class of potential games are congestion games (Rosenthal, 1973). They also arise in Cournot oligopoly models with linear inverse demand functions (Monderer and Shapley, 1996). Recently, Sandholm (2007b) studied a mechanism design problem where the planner can construct a pricing scheme, so that the transformed game is a potential game, which leads, in his model, to the long-run selection of socially efficient outcomes. Ui (2000) has shown interesting connections between the Shapley value and potential functions, and Morris and Ui (2005) use potential methods to study equilibria which are robust to incomplete information.

Closest to the present work is a recent paper by Ehrhardt et al. (2008b), who study a similar dynamic process. Their link formation mechanism is designed in such a way that only players who play the same action form a link. This is interpreted as a pure homophily based linking process. They also characterize the induced ensemble of random graphs, and find that the network consists of disjoint components, each following the distribution of an Erdös-Rényi ensemble. This paper extends their result by allowing for much more general behavioral rules, both in the action adjustment and the link creation process, which results in a richer interaction structure.

The rest of the paper is organized as follows. In Section 4.2, the model framework is explained in detail. In sections 4.3 and 4.4, I derive the asymptotic characteristics of the model. Sections 4.5 and 4.6 present an analysis of the joint distribution of action profiles and social networks as well as the induced marginal distributions. Section 4.7 characterizes stochastically stable states. Section 4.8 sketches a general class of stochastic processes on the co-evolution of networks and play. 4.10 collects lengthy and technical proofs.

4.2 The model

Consider a finite population of individuals $i, j, k \in \mathcal{I} = \{1, 2, \ldots, N\}$, members of which are called players or agents. Each player can choose one out of q different pure actions from the set $\mathcal{A} = \{a_1, a_2, \ldots, a_q\}$. I will also say "playing action r" with the understanding that this is action a_r. An *action profile* (configuration) is a tuple $\alpha = (\alpha^i)_{i \in \mathcal{I}} \in \mathcal{A}^{\mathcal{I}}$. When individual i meets individual j, they engage in a 2-player game defined by the payoff function $u : \mathcal{A}^2 \to \mathbb{R}$. We assume that this function is symmetric in the following sense:

Assumption 4.2.1.

$$(\forall a, a' \in \mathcal{A}) : u(a, a') = u(a', a) \tag{4.2.1}$$

Games with this special property are known as (exact) *potential games* (Monderer and Shapley, 1996). This defines the *base game* $\Gamma^b := (\mathcal{A}, u)$.

The interaction structure is modeled as an undirected graph (network). Let $\mathcal{I}^{(2)}$ denote the set of unordered pairs of players. There are $N(N-1)/2$ such pairs. A graph is a pair $G = (\mathcal{I}, \mathcal{E})$, where we interpret $\mathcal{I}$ as the set of vertices (nodes) and $\mathcal{E} = \mathcal{E}(G) \subseteq \mathcal{I}^{(2)}$ the set of *edges* (links). An edge is an unordered pair of players $(i, j) \equiv (j, i)$ with the interpretation that if $(i, j) \in \mathcal{E}$, then players i and j play against each other. If $\mathcal{E} = \mathcal{I}^{(2)}$ we obtain the *complete graph* on $\mathcal{I}$, denoted by G^c. In this graph each individual is connected to everybody else and we obtain the standard global matching model. If $\mathcal{E} = \emptyset$ then we speak of the *empty graph* G^e. A graph $G' = (\mathcal{I}', \mathcal{E}')$ is a *subgraph* of $G = (\mathcal{I}, \mathcal{E})$ if $\mathcal{I}' \subseteq \mathcal{I}$ and $\mathcal{E}' \subseteq \mathcal{E}$. For two disjoint subsets of players $\mathcal{V}, \mathcal{V}' \subset \mathcal{I}$ denote the set of edges that join players from $\mathcal{V}$ with players belonging to $\mathcal{V}'$ (and vice-versa) as $\mathcal{E}(\mathcal{V}, \mathcal{V}')$.

All graphs on $\mathcal{I}$ differ only in terms of their edge set $\mathcal{E}$. Let $\mathcal{G}[\mathcal{I}]$ denote the set of graphs that can be formed on the vertex set $\mathcal{I}$. It is often more convenient to work with networks via the function $g : \mathcal{I}^{(2)} \times \mathcal{G}[\mathcal{I}] \to \{0, 1\}$, assigning to each pair $(i, j) \in \mathcal{I}^{(2)}$ the value $g((i, j), G) \equiv g((j, i), G) \equiv g^i_j(G) \in \{0, 1\}$. If $g^i_j(G) = 1$ then players i and j are linked under the graph G and play against each other. Thus, we have the identity $\mathcal{E}(G) = \{(i, j) \in \mathcal{I}^{(2)} | g^i_j(G) = 1\}$ for all graphs $G \in \mathcal{G}[\mathcal{I}]$. It follows that every graph $G \in \mathcal{G}[\mathcal{I}]$ can be identified through the *realization* of links $g(G) = (g^i_j(G))_{1 \leq i < j \leq N} \in \{0, 1\}^{\mathcal{I}^{(2)}}$. In view of this equivalence, we will identify the space $\mathcal{G}[\mathcal{I}] \equiv \mathcal{G}$ as the set of all possible edge realizations $\{0, 1\}^{\mathcal{I}^{(2)}}$, members of which are vectors $g = (g^i_j)_{1 \leq i \leq j \leq N}$. The number of edges of the graph g is $e(g) := \sum_{i=1}^N \sum_{j > i} g^i_j$.

A *population state* is the pair $\omega = (\alpha, g) \in \Omega \equiv \mathcal{A}^{\mathcal{I}} \times \mathcal{G}$. It contains an action

profile and a network. Let $\alpha_i^{a_v} := (\alpha^1, \ldots, \alpha^{i-1}, a_v, \alpha^i, \ldots, \alpha^N)$. Let $g \oplus (i,j)$ denote the network that we obtain if the (previously non-existing) edge connecting players i and j is created, and $g \ominus (i,j)$ be the network resulting from the deletion of the edge connecting players i and j.
Given a population state ω, define for every player $i \in \mathcal{I}$ the (open) *interaction neighborhood*

$$\mathcal{N}^i(\omega) = \bigcup_{r=1}^{q} \{j \in \mathcal{I} | g_j^i = 1 \ \&\ \alpha^j = a_r\}.$$

The set $\mathcal{N}^i \cup \{i\} \equiv \bar{\mathcal{N}}^i$ defines the closed interaction neighborhood of a player. There are $\kappa_r^i(\omega) := |\{j \in \mathcal{I} | g_j^i = 1 \ \&\ \alpha^j = a_r\}|$ r-players against which player i has to play. The total number of games in which player i is involved is given by his *degree* $\kappa^i(\omega) = \sum_{r=1}^q \kappa_r^i(\omega)$. From all these interactions, player i receives the total payoff

$$\pi^i(\alpha, g) \equiv \pi^i(\omega) := \sum_{j \in \mathcal{N}^i(\omega)} u(\alpha^i, \alpha^j) = \sum_{r=1}^{q} u(\alpha^i, a_r)\kappa_r^i(\omega). \tag{4.2.2}$$

In analogy with standard population games, I will call the collection of payoff functions $\pi = (\pi^i)_{i \in \mathcal{I}}$ the *structured population game*.

4.3 Co-evolution with noise

Consider the family of perturbed Markov processes

$$\mathcal{M}^\beta = (\Omega, \mathcal{F}, \mathbb{P}, (Y^\beta(t))_{t \geq 0})_{\beta \in \mathbb{R}_+},$$

where Ω is the finite states space of pairs $\omega = (\alpha, g)$, $\mathcal{F}$ a suitably chosen σ-algebra (e.g. 2^Ω), $\mathbb{P} : \mathcal{F} \to [0,1]$ a probability measure, and $(Y^\beta(t))_{t \geq 0}$ a family of Ω-valued random variables indexed by a noise parameter $\beta \geq 0$ and a continuous time parameter t. $\mathcal{M}^\beta$ will define a co-evolutionary model with noise. The time evolution of this process can be studied by its infinitesimal generator. Define the operator $\boldsymbol{\eta}^\beta := \left[\eta^\beta(\omega \to \omega')\right]_{\omega, \omega' \in \Omega}$ whose components are mappings $\eta^\beta : \Omega \times \Omega \to \mathbb{R}$ satisfying $0 \leq \eta^\beta(\omega \to \hat{\omega}) < \infty$ for all $\hat{\omega} \neq \omega$, and $\sum_{\hat{\omega}} \eta^\beta(\omega \to \hat{\omega}) = 0$ for all $\omega \in \Omega$. The value $\eta^\beta(\omega \to \hat{\omega})$ is interpreted as the *rate* with which the process moves from state ω to some other state $\hat{\omega}$.[1] The generator is defined by the following sub-processes.

[1]In a very small time interval $[t, t+h)$, the probability that the process moves from ω to $\hat{\omega}$ is then approximately $\eta^\beta(\omega \to \hat{\omega})h$.

Action update: The way how players update their actions is modeled as in Blume (2003) or Hofbauer and Sandholm (2007). Players are endowed with independent Poisson alarm clocks, ringing at the common rate $\nu > 0$. The total rate of this subprocess is thus $N\nu$. Conditional on the event of a revision opportunity, player i receives the chance to adjust his action with probability $1/N$. When player i gets a revision opportunity he calculates the current expected payoff of all of his pure actions, given the set of neighbors, but his computations are perturbed by some random shock $\varepsilon^i = (\varepsilon^i_a)_{a\in\mathcal{A}}$. Assume that these perturbations are i.i.d. type 1 extreme value distributed,[2] and that i selects action $a_r \in \mathcal{A}$ with probability

$$b^i(a_r|\omega) := \mathbb{P}\left(a_r \in \arg\max_{a_v\in\mathcal{A}}(\pi^i(\alpha_i^{a_v}, g) + \varepsilon^i_{a_v})|\omega\right). \tag{4.3.1}$$

Computing this probability explicitly leads to

$$b^{i,\beta}(a_r|\omega) = \frac{\exp(\pi^i(\alpha_i^{a_r}, g)/\beta)}{\sum_{v=1}^q \exp(\pi^i(\alpha_i^{a_v}, g)/\beta)}. \tag{4.3.2}$$

The transition $\omega = (\alpha, g) \to \hat{\omega} = (\alpha_i^{a_r}, g) \neq \omega$ proceeds therefore at a rate

$$\eta^\beta(\omega \to \hat{\omega}) = \nu b^i(a_r|\omega). \tag{4.3.3}$$

Link creation: Here ideas of the *stochastic-actor model*, developed in Snijders (2001), are used. The key-ingredients of this model are a *rate function*, governing the pace at which individuals update their connections, and an *objective function*, capturing the preferences of the individuals concerning link creation. For the rate function, I make the following assumption:

Assumption 4.3.1. *The rate functions of individuals take the form*

$$(\forall i \in \mathcal{I})(\forall \omega \in \Omega) : \lambda^{i,\beta}(\omega) = \sum_{k\notin\bar{\mathcal{N}}^i(\omega)} \exp(u(\alpha^i, \alpha^k)/\beta). \tag{4.3.4}$$

[2] This formulation of stochastically perturbed payoffs has a very long tradition in the theory of discrete choice, see e.g. Anderson et al. (1992). For a more recent treatise and alternative interpretation see van Damme and Weibull (2002). The cumulative distribution function of a doubly exponential distributed random variable with mean 0 and variance $\frac{\beta^2\pi^2}{6}$ is $F(x) = \exp\left[-\exp(-x/\beta - \gamma)\right]$. Beside its importance in theoretical economics, it has also been used in experimental studies, see for instance McKelvey and Palfrey (1995), where it is known as the "quantal response function".

This formulation reflects the intuitive idea that players, who expect a large profit from interactions with currently unknown players, should be relatively fast in creating their network. Let

$$\bar{\lambda}^{\beta}(\omega) := \sum_{i\in\mathcal{I}} \lambda^{i,\beta}(\omega) = 2\sum_{i,j>i} \exp(u(\alpha^i,\alpha^j)/\beta)(1-g^i_j),$$

so that the conditional probability that player i receives a link creation opportunity is simply $\lambda^{i,\beta}(\omega)/\bar{\lambda}^{\beta}(\omega)$. Conditional on this event, player i screens the set of unknown players (i.e. those player who are not neighbors yet) and picks one player from this set who yields the highest per-interaction payoff, perturbed by a noisy signal $\zeta^i = (\zeta^i_k)_{k\notin\bar{\mathcal{N}}^i(\omega)}$. Hence, the conditional probability that i selects j for a linking partner is

$$w^i_j(\omega) := \mathbb{P}\left(u(\alpha^i,\alpha^j) + \zeta^i_j \geq u(\alpha^i,\alpha^k) + \zeta^i_k \ \forall k \notin \bar{\mathcal{N}}^i|\omega\right). \qquad (4.3.5)$$

If we assume that the random perturbation follows the same distributional law as in the action adjustment process one obtains the logit formula

$$(\forall i \in \mathcal{I})(\forall j \notin \bar{\mathcal{N}}^i(\omega)) : w^{i,\beta}_j(\omega) = \frac{\exp(u(\alpha^i,\alpha^j)/\beta)}{\sum_{k\notin\bar{\mathcal{N}}^i(\omega)} \exp(u(\alpha^i,\alpha^k)/\beta)}. \qquad (4.3.6)$$

For general link creation probabilities (4.3.5), the rate of transiting from state $\omega = (\alpha, g)$ to state $\hat{\omega} = (\alpha, g \oplus (i,j))$ is

$$\eta^{\beta}(\omega \to \hat{\omega}) = \lambda^i(\omega)w^i_j(\omega) + \lambda^j(\omega)w^j_i(\omega). \qquad (4.3.7)$$

Using Assumption 4.3.1 and (4.3.6) gives us

$$\eta^{\beta}(\omega \to \hat{\omega}) = 2\exp(u(\alpha^i,\alpha^j)/\beta). \qquad (4.3.8)$$

Link destruction: To make the dynamic interesting, we need a process that counteracts the creation of links. Following recent papers by Ehrhardt et al. (2006b; 2008b), I assume that there exists an *exogenous* random shock removing any of these links. This unguided drift term models the phenomenon of *environmental volatility*, and is a key ingredient of the model. It captures the idea that connections are not everlasting, but as time goes by and players change their behavior, the profitability of links will also change, making some connections obsolete. The rate at which the link (i,j) disappears is given by $\xi > 0$.[3] Hence, in a very small time interval $[t, t+h)$ the *probability of survival* of a currently existing edge (i,j) is $\xi h + o(h)$. The expected life time of an edge is $1/\xi$. Hence, starting from $\omega = (\alpha, g)$, the transition rate to $\hat{\omega} = (\alpha, g \ominus (i,j))$ is

$$\eta^{\beta}(\omega \to \hat{\omega}) = \xi. \tag{4.3.9}$$

The last case we have to consider is a "phantom switch", i.e. the transition rate $\eta^{\beta}(\omega \to \omega)$. Define the rate of such an event as

$$\eta^{\beta}(\omega \to \omega) = -\Lambda^{(\beta,\xi)}(\omega), \tag{4.3.10}$$

where

$$\Lambda^{(\beta,\xi)}(\omega) := \nu \sum_{i=1}^{N} \sum_{a \in \mathcal{A} \setminus \{\alpha^i\}} b^{i,\beta}(a|\omega) + \xi e(g) + \bar{\lambda}^{\beta}(\omega). \tag{4.3.11}$$

To summarize, the infinitesimal generator of the co-evolutionary model with noise $\mathcal{M}^{\beta}$ is defined as

$$\eta^{\beta}(\omega \to \hat{\omega}) = \begin{cases} \nu b^{i,\beta}(a|\omega) & \text{if } \hat{\omega} = (\alpha_i^a, g) \neq \omega, \\ 2\exp(u(\alpha^i, \alpha^j)/\beta) & \text{if } \hat{\omega} = (\alpha, g \oplus (i,j)), \\ \xi & \text{if } \hat{\omega} = (\alpha, g \oplus (i,j)), \\ -\Lambda^{(\beta,\xi)}(\omega) & \text{if } \hat{\omega} = \omega, \\ 0 & \text{otherwise.} \end{cases} \tag{4.3.12}$$

It is easily verified that $\sum_{\hat{\omega} \in \Omega} \eta^{\beta}(\omega \to \hat{\omega}) = 0$ for all $\omega \in \Omega$. For $\beta > 0$ we observe that $\eta^{\beta}(\omega \to \omega') > 0$ for $\omega \neq \omega'$, meaning that there can be no single state that is absorbing. Irreducibility of the generator follows from this easily. Furthermore, in view of the finiteness of the state space, positive recurrence of the process follows. Hence, the Markov process is ergodic.

4.4 The invariant distribution

By ergodicity, the co-evolutionary model with noise admits a unique invariant distribution $\mu^{(\beta,\xi)} = (\mu^{(\beta,\xi)}(\omega))_{\omega \in \Omega}$. In terms of the generator $\boldsymbol{\eta}^{\beta}$, this

[3]The assumption of constant link decay rates is less restrictive as it may seem. Since link creation probabilities are payoff driven, players will be more likely to establish links which are associated with higher per-interaction payoff. Hence, if a highly valuable link disappears, ceteris paribus, there is a relatively high probability that it will be re-established in future periods. Extending to heterogeneous link destruction rates is straightforward and will be done in chapter 5.

probability distribution satisfies the global balance equation $\mu^{(\beta,\xi)}\boldsymbol{\eta}^{\beta} = \mathbf{0}$. Determining this probability vector is facilitated in the special class of reversible Markov processes. Given the model $\mathcal{M}^{\beta}$ with generator $\boldsymbol{\eta}^{\beta}$, we can define for a given $T > 0$ its *time reversal* as the process $(\hat{Y}^{\beta}(t))_{0\leq t\leq T}$, with $\hat{Y}^{\beta}(t) = Y^{\beta}(T-t)$. A Markov process $(Y^{\beta}(t))_{t\geq 0}$ is said to be *reversible*, if its time reversal has the same distribution as the original process (see Stroock, 2005, ch. 5). In our case, reversibility will appear as an equilibrium phenomenon (i.e. $(Y^{\beta}(t))_{t\geq 0}$ is *reversible in equilibrium*). The *detailed balance condition*, relative to the infinitesimal generator $\boldsymbol{\eta}^{\beta}$, gives a sufficient condition for $\mu^{(\beta,\xi)}$ being an invariant distribution. The measure $\mu^{(\beta,\xi)}$ is in detailed balance with the generator $\boldsymbol{\eta}^{\beta}$ if

$$(\forall \omega, \hat{\omega} \in \Omega) : \mu^{(\beta,\xi)}(\omega)\eta^{\beta}(\omega \to \hat{\omega}) = \mu^{(\beta,\xi)}(\hat{\omega})\eta^{\beta}(\hat{\omega} \to \omega). \tag{4.4.1}$$

A probability distribution satisfying the detailed balance condition (4.4.1) must be an invariant distribution. Conversely, a probability distribution satisfying (4.4.1) implies reversibility of the corresponding Markov process.

Theorem 4.4.1. *Given $(\beta,\xi) \gg (0,0)$, the unique invariant distribution of the co-evolutionary model with noise $\mathcal{M}^{\beta}$ equals*

$$(\forall \omega \in \Omega) : \mu^{(\beta,\xi)}(\omega) = \frac{1}{Z^{(\beta,\xi)}} \prod_{i=1}^{N} \prod_{j>i} \left[\frac{2}{\xi} \exp\left(\frac{u(\alpha^i, \alpha^j)}{\beta} \right) \right]^{g_j^i}, \tag{4.4.2}$$

where $Z^{(\beta,\xi)} \equiv \sum_{\omega\in\Omega} \prod_{i=1}^{N} \prod_{j>i} \left[\frac{2}{\xi} \exp\left(\frac{u(\alpha^i,\alpha^j)}{\beta} \right) \right]^{g_j^i}$ is the the partition function.

Proof. See Section 4.10. □

A consequence of ergodicity is the convergence of long-run averages of sample paths to the invariant distribution. Formally, this means

$$\mathbb{P}\left(\lim_{t\to\infty} \frac{1}{t} \int_0^t \mathbb{1}_{\{Y^{\beta}(s)=\omega\}} \,\mathrm{d}\, s = \mu^{(\beta,\xi)}(\omega) \right) = 1,$$

where $\mathbb{1}_A$ is the indicator function of a measurable set $A \subseteq \Omega$, and for any integrable function $f : \Omega \to \mathbb{R}$

$$\mathbb{P}\left(\lim_{t\to\infty} \frac{1}{t} \int_0^t f(Y^{\beta}(s)) \,\mathrm{d}\, s = \mathbb{E}_{\mu^{(\beta,\xi)}}[f] \right) = 1,$$

where $\mathbb{E}_{\mu^{(\beta,\xi)}}[f] = \sum_{\omega\in\Omega} f(\omega)\mu^{(\beta,\xi)}(\omega)$ is the expected value of the function f under the invariant distribution $\mu^{(\beta,\xi)}$.

Observe that for $\beta > 0$ $\mu^{(\beta,\xi)}$ is a full support distribution on Ω. Thus, the only thing one may be able to deduce from it is to classify a subset of states which receive more mass than others. The subsequent chapters are devoted to this exercise.
Define an aggregate utility index as the sum of individual utilities,

$$(\forall \omega \in \Omega) : U(\omega) = \sum_{i=1}^{N} \pi^i(\omega) = 2 \sum_{i=1}^{N} \sum_{j>i} u(\alpha^i, \alpha^j) g_j^i \tag{4.4.3}$$

Efficiency, in terms of this index, is a state in the argmax set of (4.4.3). Lemma 4.4.1 shows that one can construct from eq. (4.4.3) a real-valued function, which captures the effects of individual utilities due to a single change in the state variable ω. In game theory such a function is known as an *exact potential* (Monderer and Shapley, 1996). Since the state variable encompasses the connections among the players, but these are not part of the strategy of a single player, a potential function for $\pi = (\pi^i)_{i\in\mathcal{I}}$ is not a potential function in its game-theoretic sense. However, it fulfills the same role in the dynamic analysis to come as a conventional potential function in the sense of Monderer and Shapley (1996), and so we will still call such a function a potential function for the structured population game, having in mind that this does not conform with its established use in game theory.

Lemma 4.4.1. *The structured population game $(\pi^i)_{i\in\mathcal{I}}$ is a potential game with exact potential function*

$$(\forall \omega \in \Omega) : P(\omega) = \frac{1}{2} \sum_{i=1}^{N} \pi^i(\omega) = \sum_{i=1}^{N} \sum_{j>i} u(\alpha^i, \alpha^j) g_j^i. \tag{4.4.4}$$

Proof. We have to show that

$$\begin{aligned} &P(\alpha_i^{a_v}, g) - P(\alpha, g) = \pi^i(\alpha_i^{a_v}, g) - \pi^i(\alpha, g), \text{ and} \\ &P(\alpha, g \oplus (i,j)) - P(\alpha, g) = u(\alpha^i, \alpha^j) \end{aligned}$$

Let us start with the event of a link creation between players i and j. The destruction of such a link has the same consequences. A direct computation shows that

$$P(\alpha, g \oplus (i,j)) - P(\alpha, g) = \frac{1}{2}(u(\alpha^i, \alpha^j) + u(\alpha^j, \alpha^i)) = u(\alpha^i, \alpha^j),$$

by symmetry of the payoff function u. Now, concerning a change in action of player i, we now have to take care of the environment of this player. All players in the set $\mathcal{I}\backslash\bar{\mathcal{N}}^i(\omega)$ are not affected by the change in player i's action. Fix the state ω and suppose player i switches to action a_v. The new state is therefore $(\alpha_i^{a_v}, g)$, and we can write

$$\begin{aligned}
U(\alpha_i^{a_v}, g) &= \sum_{k=1}^{N} \sum_{j:g_j^k=1} u(\alpha^k, \alpha^j) \\
&= U(\alpha, g) + \sum_{j:g_j^i=1} [u(a_v, \alpha^j) - u(\alpha^i, \alpha^j)] + \sum_{\ell:g_i^\ell=1} [u(\alpha^\ell, a_v) - u(\alpha^\ell, \alpha^i)] \\
&= U(\alpha, g) + 2 \sum_{j:g_j^i=1} [u(a_v, \alpha^j) - u(\alpha^i, \alpha^j)]
\end{aligned}$$

□

Now consider the function $H : \Omega \times \mathbb{R}_+ \times \mathbb{R}_{++} \to \mathbb{R}$, defined as

$$H(\omega, \beta, \xi) := P(\omega) + \beta e(\omega) \log\left(\frac{2}{\xi}\right) \tag{4.4.5}$$

This function acts as a a *graph Hamiltonian* for the invariant distribution.[4] One sees that there are two components combined in the graph Hamiltonian. The first component is the potential function of the game, which measures (up to a linear scaling) the aggregate utility of the population. The second part is a size measure of the interaction graph, weighted by the volatility parameter ξ. If $\xi > 2$ then too large graphs (measured by the number of edges) lead to a reduction in the value of the Hamiltonian. This effect is in turn weighted by the noise level β. Proposition 4.4.1 shows that it contains all the information one needs to determine the invariant distribution of the process $\mathcal{M}^\beta$. Its proof is straightforward and therefore omitted.

Proposition 4.4.1. *The stationary distribution of the co-evolutionary model with noise $\mathcal{M}^\beta$ is the Gibbs measure*

$$\mu^{(\beta,\xi)}(\omega) = \frac{e^{\frac{1}{\beta}H(\omega,\beta,\xi)}}{\sum_{\hat{\omega}\in\Omega} e^{\frac{1}{\beta}H(\hat{\omega},\beta,\xi)}}. \tag{4.4.6}$$

From the definition of the Hamiltonian (4.4.5), one can see that a large value of β, combined with $\xi > 2$, implies that too large graphs will not receive

[4]For a general discussion of this concept see Park and Newman (2004). In statistical mechanics a Hamiltonian is, roughly, a measure of the energy of a system. In the simplest case it is the sum of the potential energy and kinetic energy. This description fits also perfectly to the form of the Hamiltonian (4.4.5).

too much weight in the long run. A small value of β means in turn that, for any given volatility level ξ, the penalty of densely connected societies has a small influence on the invariant distribution. It is exactly this trade-off between β and volatility ξ which makes the form of the invariant distribution interesting. High environmental volatility, accompanied with moderate noise will lead to a sparsely connected society. On the other side, a small value of β will dominate any value of volatility ξ, and the value of the potential function will dominate the shape of the invariant distribution.

4.5 The ensemble of random graphs

Given $\omega = (\alpha, g) \in \Omega$, define the set of r-players as $\mathcal{I}_r(\omega) := \{i \in \mathcal{I} | \alpha^i = a_r\}$. Sets of this form will be called action classes. Every state assigns each player to a single action class. Hence, the family $\{\mathcal{I}_r\}_{1\leq r\leq q}$ defines a partition on the set $\mathcal{I}$. Fix a partition $\boldsymbol{\mathcal{I}} \equiv \{\mathcal{I}_r\}_{1\leq r\leq q}$ and define the subspace

$$\Omega(\boldsymbol{\mathcal{I}}) := \{\omega \in \Omega | \mathcal{I}_r(\omega) = \mathcal{I}_r,\ 1 \leq r \leq q\}.$$

We say that state ω agrees with the action partition $\boldsymbol{\mathcal{I}}$, if it is contained in $\Omega(\boldsymbol{\mathcal{I}})$. Note that the definition of the set $\Omega(\boldsymbol{\mathcal{I}})$ does not say anything about network structures. Once we condition on an action partition, we fix a strategy configuration $\alpha \in \mathcal{A}^{\mathcal{I}}$, but allow for all potential networks. In other words, $\mu^{(\beta,\xi)}(\omega|\boldsymbol{\mathcal{I}}) \equiv \mu^{(\beta,\xi)}(g|\alpha)$.
Given a partition $\boldsymbol{\mathcal{I}}$, the product operator $\prod_{i=1}^N \prod_{j>i}$ has the same meaning as the product operator $\prod_{r=1}^q \prod_{i\in\mathcal{I}_r(\omega)} \prod_{v\geq r} \prod_{j\in\mathcal{I}_v(\omega);j>i}$. This implies that we are able to re-formulate the stationary distribution in terms of action classes, so that for all $\omega \in \Omega$

$$\mu^{(\beta,\xi)}(\omega|\boldsymbol{\mathcal{I}}) \propto \prod_{r=1}^{q} \prod_{i\in\mathcal{I}_r} \left\{ \prod_{v\geq r} \prod_{j\in\mathcal{I}_v;j>i} \left[\frac{2}{\xi} \exp\left(\frac{u(a_r, a_v)}{\beta}\right)\right]^{g_j^i} \right\} \mathbb{1}_{\{\omega\in\Omega(\boldsymbol{\mathcal{I}})\}}. \tag{4.5.1}$$

For proper normalization of this measure, one has to compute the total mass received by the set $\Omega(\boldsymbol{\mathcal{I}})$, which is

$$\mu^{(\beta,\xi)}(\Omega(\boldsymbol{\mathcal{I}})) = \sum_{g\in\mathcal{G}} \mu^{(\beta,\xi)}(\alpha, g).$$

Let $e_{r|v}(\omega) := \sum_{i\in\mathcal{I}_r(\omega)} \sum_{j\in\mathcal{I}_v(\omega),j>i} g_j^i$, denote the number of edges connecting r-players with v-players at state ω, and define for all $(i, j) \in \mathcal{I}^{(2)}$

$$p_{ij}^{(\beta,\xi)}(\omega) := \frac{2\exp(u(\alpha^i, \alpha^j)/\beta)}{2\exp(u(\alpha^i, \alpha^j)/\beta) + \xi}, \quad \theta_{ij}^{(\beta,\xi)}(\omega) := \log\left(\frac{p_{i,j}^{(\beta,\xi)}(\omega)}{1 - p_{i,j}^{(\beta,\xi)}(\omega)}\right).$$

Setting $\theta^{(\beta,\xi)}(\omega) := (\theta_{ij}^{(\beta,\xi)}(\omega))_{(i,j)\in\mathcal{I}^{(2)}}$, we get

$$\begin{aligned}
\frac{1}{\beta}H(\omega;\beta,\xi) &= \frac{1}{\beta}\sum_{i=1,j>i}^{N} u(\alpha^i,\alpha^j)g_j^i + \log(2/\xi)\sum_{i=1,j>i}^{N} g_j^i \\
&= \sum_{i=1,j>i}^{N}\left[\log(\exp(u(\alpha^i,\alpha^j)/\beta) + \log(2/\xi)\right] g_j^i \\
&= \sum_{i=1,j>i}^{N}\theta_{ij}^{(\beta,\xi)} g_j^i =: h\left(\omega,\theta^{(\beta,\xi)}\right)
\end{aligned}$$

Given an action partition $\boldsymbol{\mathcal{I}}$, consider the subgraph $G_{r|v} := (\mathcal{I}_r \cup \mathcal{I}_v, \mathcal{E}_{r|v})$, where $\mathcal{E}_{r|v} = \mathcal{E}(\mathcal{I}_r,\mathcal{I}_v)$. For all $(i,j) \in [\mathcal{I}_r \cup \mathcal{I}_v]^{(2)}$, the numbers $p_{ij}^{(\beta,\xi)}, \theta_{ij}^{(\beta,\xi)}$ are constant, so that we may write $p_{r|v}^{(\beta,\xi)}$ and $\theta_{r|v}^{(\beta,\xi)}$. Lemma 4.10.1 of Section 4.10 shows that

$$\mu^{(\beta,\xi)}(\Omega(\boldsymbol{\mathcal{I}})) \propto \prod_{r=1}^{q}\prod_{v\geq r}\left(1-p_{r|v}^{(\beta,\xi)}\right)^{-\frac{|\mathcal{I}_r|(|\mathcal{I}_v|-\delta_{r,v})}{1+\delta_{r,v}}} \tag{4.5.2}$$

where $\delta_{x,y} = 1$ if, and only if, $x = y$, and 0 otherwise. The main result of this section is then the following.

Theorem 4.5.1 (The Erdös-Rényi Decomposition)**.** *Fix an action partition* $\boldsymbol{\mathcal{I}}$ *and* $(\beta,\xi) \gg (0,0)$.

(a) The measure (4.5.1) is the conditional distribution over graphs $g \in \mathcal{G}$ *and factorizes to*

$$\mu^{(\beta,\xi)}(\omega|\boldsymbol{\mathcal{I}}) = \prod_{r=1,v\geq r}\left[p_{r|v}^{(\beta,\xi)}\right]^{e_{r|v}(\omega)}\left[1-p_{r|v}^{(\beta,\xi)}\right]^{\frac{|\mathcal{I}_r|(|\mathcal{I}_v|-\delta_{r,v})}{1+\delta_{r,v}}-e_{r|v}(\omega)}. \tag{4.5.3}$$

(b) There exists a continuously differentiable function $F_{\boldsymbol{\mathcal{I}}}(\theta^{(\beta,\xi)})$ *such that for all* $1 \leq r \leq l \leq q$

$$\frac{\partial F_{\boldsymbol{\mathcal{I}}}(\theta^{(\beta,\xi)})}{\partial\theta_{r|v}^{(\beta,\xi)}} = \mathbb{E}_{\mu^{(\beta,\xi)}}[e_{r|v}|\boldsymbol{\mathcal{I}}],\quad \frac{\partial^2 F_{\boldsymbol{\mathcal{I}}}(\theta^{(\beta,\xi)})}{\partial\theta_{r|v}^{(\beta,\xi)}\partial\theta_{r|l}^{(\beta,\xi)}} = \mathrm{Cov}_{\mu^{(\beta,\xi)}}[e_{r|v}, e_{r|l}|\boldsymbol{\mathcal{I}}]$$

$$2\frac{\partial F_{\boldsymbol{\mathcal{I}}}(\theta^{(\beta,\xi)})}{\partial(1/\beta)} = \mathbb{E}_{\mu^{(\beta,\xi)}}[U|\boldsymbol{\mathcal{I}}]$$

(c) The statistical ensemble of subgraphs $\mathcal{G}[\mathcal{I}_r \cup \mathcal{I}_v]$ is an Erdös-Rényi graph with edge success probability

$$p_{r|v}^{(\beta,\xi)} = \frac{2\exp(u(a_r,a_v)/\beta)}{2\exp(u(a_r,a_v)/\beta)+\xi}.$$

Proof. See Section 4.10. □

Part (a) of the Theorem shows that the equilibrium ensemble of graphs boils down to an *inhomogeneous random graph* (Söderberg, 2002, Bollobás et al., 2007). For an arbitrary action profile α, eq. (4.5.3) gives us complete information about the probability with which an r-strategist interacts with players from other action classes. Thus, if one wants to make a probabilistic prediction about the interaction pattern between r-players and v-players, all one has to do is to look at the factor

$$[p_{r|v}^{(\beta,\xi)}]^{e_{r|v}(\omega)}\left[1-p_{r|v}^{(\beta,\xi)}\right]^{\frac{|\mathcal{I}_r|(|\mathcal{I}_v|-\delta_{r,v})}{1+\delta_{r,v}}-e_{r|v}(\omega)}$$

what is exactly the probability measure of the random graph model of Erdös and Rényi (1960) (see also Section 1.2 of this thesis). Part (b) is a standard result for Gibbs measures (see e.g. Stroock, 2005). Using the explicit from for the marginal distribution $\mu^{(\beta,\xi)}(\Omega(\boldsymbol{\mathcal{I}}))$, a simple computation shows that for all $1 \leq r \leq v \leq q$

$$\mathbb{E}_{\mu^{(\beta,\xi)}}[e_{r|v}|\boldsymbol{\mathcal{I}}] = \frac{|\mathcal{I}_r|(|\mathcal{I}_v|-\delta_{r,v})}{1+\delta_{r,v}}p_{r|v}^{(\beta,\xi)}.$$

For the covariances we see that

$$\mathrm{Cov}_{\mu^{(\beta,\xi)}}[e_{r|v},e_{r|l}|\boldsymbol{\mathcal{I}}] = \begin{cases} 0 & \text{if } v \neq l, \\ \frac{|\mathcal{I}_r|(|\mathcal{I}_v|-\delta_{r,v})}{1+\delta_{r,v}}p_{r|v}^{(\beta,\xi)}(1-p_{r|v}^{(\beta,\xi)}) & \text{if } v = l. \end{cases}$$

The fact that the total graph can be regarded as a collection of independent Erdös-Rényi graphs (with different edge success probabilities) makes it possible to derive a probability distribution for the degree of a randomly selected individual $i \in \mathcal{I}_r$. Since $\kappa^i = \sum_{v=1}^{q}\kappa_v^i$, we first have to determine the distribution of the random variables $\kappa_v^i, 1 \leq v \leq q$. Theorem 4.5.1 tells us that κ_v^i has a Binomial distribution with parameters $(|\mathcal{I}_v|-\delta_{r,v}, p_{r|v}^{(\beta,\xi)})$ (see e.g. Bollobás, 1998).

Proposition 4.5.1. *Given an action partition $\boldsymbol{\mathcal{I}}$ pick a player $i \in \mathcal{I}_r$ and let $n_v := |\mathcal{I}_v|, 1 \leq v \leq q$. The degree of player i is distributed according to the mass function*

$$f_{r,\kappa}(k|\boldsymbol{\mathcal{I}}) := \frac{1}{R^{(\beta,\xi)}(\boldsymbol{\mathcal{I}})} \sum_{k_1+\ldots+k_q=k} \frac{k!}{k_1! \cdots k_q!} \prod_{v=1}^{q} \left[f_{r|v}^{(\beta,\xi)}(k_v) \right]^{k_v}, \tag{4.5.4}$$

$$f_{r|v}^{(\beta,\xi)}(k_v) := \binom{n_v - \delta_{r,v}}{k_v}^{1/k_v} \left(\frac{p_{r|v}^{(\beta,\xi)}}{1 - p_{r|v}^{(\beta,\xi)}} \right). \tag{4.5.5}$$

where $R^{(\beta,\xi)}(\boldsymbol{\mathcal{I}})$ is the normalizing factor.

Proof. See Section 4.10. □

Observe that for the degree distribution it suffices to know the number of players in the various action classes, not their identity. Hence, all action partitions $\boldsymbol{\mathcal{I}}$ that put the same number of players into the various classes are equivalent in terms of the connectivity structure of the network. Thus, instead of looking at a specific action partition $\boldsymbol{\mathcal{I}}$, it is sufficient to work with less information contained in a tuple $n = (n_1, \ldots, n_q)$ such that $n_v = |\mathcal{I}_v|$ for all v and $\sum_{v=1}^{q} n_v = N$.

Example 4.5.1. *Consider the coordination game*

	a_1	a_2
a_1	$(3,3)$	$(0,0)$
a_2	$(0,0)$	$(1,1)$

We will examine the degree distribution for 1*-players under various parameter constellations (β, ξ) for the frequency vector $n = (80, 20)$. Figure 4.1 shows the degree distribution for a typical 1-player when $(\beta, \xi) = (0.5, 70)$. The mean degree of 1-players is seen to be 78. However, we cannot say to which action class most of this links lead to since we only look at the distribution of the total degree κ. Applying Theorem 4.5.1, we get complete information about the inter-group connectivity pattern by inspecting the two numbers*

$$p_{1|1} = \frac{2\exp(3/\beta)}{2\exp(3/\beta) + \xi}, \quad p_{1|2} = \frac{2}{2 + \xi}$$

Note that $p_{1|1} \to 1$ as $\beta \to 0$, implying that in this limit only links within the same action class exist with probability 1. Consequently, for small noise levels the majority of the 78 neighbors will be 1-players as well. For larger levels of noise (the right figure with $\beta = 1.5$) we observe a drastically smaller average degree. This implies that the effect of the parameter values β and ξ goes into the same direction. Increasing β with constant ξ will have qualitatively the same effect as increasing ξ with constant β.

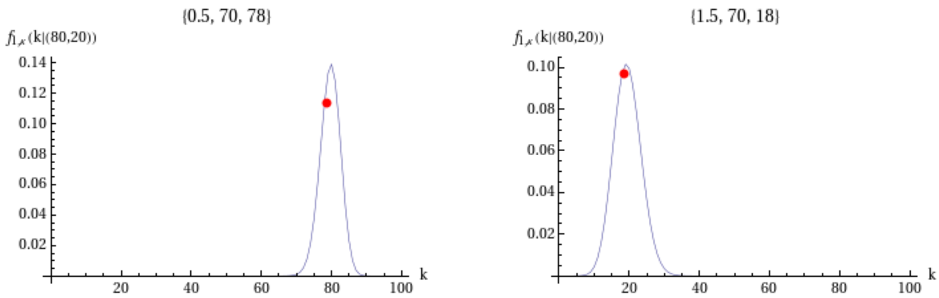

Figure 4.1: Degree distributions for 1-players under various parameter constellations. The triple at the top of each plot is $(\beta, \xi, \bar{k})$, i.e. the noise and volatility rate and the resulting average degree for this action class. The point marks the position of the mean of this distribution.

4.6 An invariant distribution over action profiles

Having derived a probability distribution on the set of networks, we will now derive a probability distribution on the set of action frequency vectors $n = (n_1, \ldots, n_q)$. Let $\mathcal{D} := \{n \in \mathbb{N}^q | \sum_{r=1}^q n_r = N\}$ denote the set of admissible action frequency vectors and define the correspondence $\Psi : \mathcal{D} \to 2^\Omega$ as $\Psi(n) = \{\omega \in \Omega | (\forall r = 1, 2, \ldots, q) : |\mathcal{I}_r(\omega)| = n_r\}$.

Proposition 4.6.1. *The invariant distribution over action frequency vectors $n \in \mathcal{D}$ is given by the mapping $\rho^{(\beta,\xi)} = \mu^{(\beta,\xi)} \circ \Psi : \mathcal{D} \to [0, 1]$, defined as*

$$\rho^{(\beta,\xi)}(n) := \Upsilon^{-1} \frac{N!}{\prod_{r=1}^q n_r!} \prod_{r=1}^q \left[z_r^{(\beta,\xi)}(n)\right]^{n_r}, \tag{4.6.1}$$

where

$$z_r^{(\beta,\xi)}(n) := \prod_{v \geq r} \left[1 + \frac{2}{\xi} \exp\left(\frac{u(a_r, a_v)}{\beta}\right)\right]^{\frac{n_v - \delta_{r,v}}{1+\delta_{r,v}}}, \quad 1 \leq r \leq q \tag{4.6.2}$$

$$\Upsilon = \sum_{n \in \mathcal{D}} \frac{N!}{\prod_{r=1}^q n_r!} \prod_{r=1}^q \prod_{v \geq r} \left[1 + \frac{2}{\xi} \exp\left(\frac{u(a_r, a_v)}{\beta}\right)\right]^{\frac{n_r(n_v - \delta_{r,v})}{1+\delta_{r,v}}}. \tag{4.6.3}$$

Proof. The proof starts from the distribution over action classes $\boldsymbol{\mathcal{I}}$ (4.5.2). The rest is a simple combinatorial exercise. The population consists of N distinct elements. There are q different boxes over which we want to distribute the N elements, and in each box $r = 1, \ldots, q$ there should be n_r

elements at the end of the day, and all N elements must be in one box, so that $\sum_{r=1}^{q} n_r = N$ holds. There are $\frac{N!}{n_1! \dots n_q!}$ different ways of solving this allocation problem. Counting all states ω that agree with a given action class size profile n leads to a probability distribution having the form

$$\mu^{(\beta,\xi)}(\Psi(n)) \propto \frac{N!}{\prod_{r=1}^{q} n_r!} \prod_{r=1}^{q} \prod_{v \geq r} \left[1 + \frac{2}{\xi} \exp\left(\frac{u(a_r, a_v)}{\beta}\right)\right]^{\frac{n_r(n_v - \delta_{r,v})}{1+\delta_{r,v}}} . \quad (4.6.4)$$

Using the respective definitions of the maps $\rho^{(\beta,\xi)}$ and $z_r^{(\beta,\xi)}(n)$ yields the desired result. □

4.7 Stochastic stability

Stochastic game dynamics have become important due to their power concerning equilibrium selection. The concept of stochastic stability, introduced by Foster and Young (1990), Young (1993) and Kandori et al. (1993) into game theory, gives a selection criterion based on the underlying dynamic process.

Definition 4.7.1. *A state $\omega \in \Omega$ is a stochastically stable state if $\lim_{\beta \to 0} \mu^{(\beta,\xi)}(\omega) > 0$. The set of stochastically stable states is*

$$\Omega^* := \{\omega \in \Omega | \lim_{\beta \to 0} \mu^{(\beta,\xi)}(\omega) > 0\}.$$

It has been shown by Blume (1993; 1997) and Young (1998) that the logit dynamics concentrates on the set of potential maximizers as the noise level goes to zero. However, their results are not directly applicable in the current context, since the graph is itself part of the state variable.

4.7.1 Selection of Potential maximizers

The following Theorem, the proof of which is based on the general discussion in Catoni (1999), is the main result of this section.

Theorem 4.7.1. *The Gibbs distribution (4.4.6) concentrates on the set $\mathcal{P} := \arg\max_{\omega \in \Omega} P(\omega)$ as $\beta \to 0$.*

Proof. See Section 4.10. □

This shows that in the limit of vanishingly small noise the process will spend almost all of its time in the vicinity of potential maximizers. In view of the relation between the potential function and aggregate utility, this gives an efficiency result for long run behavior. Furthermore, in view of the ergodic theorem, which has been mentioned briefly in Section 4.4, we know that long run averages of the potential function converge to the expected value under the invariant distribution $\mu^{(\beta,\xi)}$. Since this expected value converges to $\max_{\omega\in\Omega} P(\omega)$ as $\beta \to 0$, we get the following corollary.

Corollary 4.7.1. *Let* $\mathcal{U} := \arg\max_{\omega\in\Omega} U(\omega) = \mathcal{P}$. *Then*

$$\lim_{\beta\to 0} \mu^{(\beta,\xi)}(\mathcal{U}) = 1. \tag{4.7.1}$$

Almost surely therefore the process arrives at states where social welfare is maximized.

4.7.2 Efficiency in pure coordination games

Consider the class of games with payoff function $u(a, a') := \phi(a, a') - c$, that satisfies condition (4.2.1), as well as

$$\begin{gathered}(\forall r = 1, 2, \ldots, q) : \max_{1\leq v\leq q} \phi(a_v, a_r) = \phi(a_r, a_r),\\ \phi(a_1, a_1) \leq \phi(a_2, a_2) \leq \ldots \leq \phi(a_q, a_q).\end{gathered} \tag{4.7.2}$$

The first condition states that matching the action chosen by the opponent is always a best reply. The second condition imposes an ordering on the payoffs of actions, where a_q denotes the payoff dominant action. From the symmetry of the payoff function, eq. (4.2.1), it follows that there are q strict Nash equilibria in the base game where the two players choose the same action. The constant $c \geq 0$ has no strategic effect, and can be interpreted as the costs of a link.[5] To keep notation simple, suppose that all strict Nash equilibria have different payoffs. Let $g^e = g(G^e), g^c = g(G^c)$ denote the empty and the complete graph, respectively.

Proposition 4.7.1. *Let* $P : \Omega \to \mathbb{R}$ *be the potential function (4.4.4), and suppose that the payoff function of the base game* G^b *satisfies (4.7.2). Then*

$$\mathcal{P} = \begin{cases} \{(a_q, \ldots, a_q)\} \times \{g^c\} & ,\text{if } u(a_q, a_q) > 0 \\ (\mathcal{A}^{\mathcal{I}} \times \{g^e\}) \cup \{\omega \in \Omega | \alpha = (a_q, \ldots, a_q)\} & ,\text{if } u(a_q, a_q) = 0 \\ \mathcal{A}^{\mathcal{I}} \times \{g^e\} & ,\text{if } u(a_q, a_q) < 0 \end{cases}$$

[5] Jackson and Watts (2002), Goyal and Vega-Redondo (2005) consider symmetric 2×2 coordination games, which are potential games, having this payoff structure.

Proof. It is straightforward to see that the potential function (4.4.4) can be written as

$$P(\omega) = \sum_{r=1}^{q} \sum_{v \geq r} u(a_r, a_v) \left(\sum_{i \in \mathcal{I}_r(\omega)} \sum_{j \in \mathcal{I}_v(\omega); j > i} g_j^i \right).$$

From this one can immediately see the validity of the claim for the high-cost scenario $u(a_q, a_q) < 0$.
Now consider the case where $u(a_q, a_q) = 0$. Clearly $P(\omega) \leq 0$ for all $\omega \in \Omega$, with equality only at the states that are in the set described in the text of the Proposition.
Finally, consider the case $u(a_q, a_q) > 0$. Since this is the largest payoff obtainable from the base game, and the potential function is linear in the links, the claim follows. This is also the unique maximizer of the potential function. $\square$

Corollary 4.7.2. *Consider the co-evolutionary model $\mathcal{M}^\beta$, with base game from the class of pure-coordination games (4.7.2). Then $\Omega^* = \mathcal{P}$.*

4.8 A general class of stochastic co-evolutionary dynamics

The model presented so far relied on the assumptions that the base game has an exact potential, and the rate functions of the individual players have the particular form (4.3.4). These assumptions make the model very tractable, and we were able to deduce many fundamental characteristics of the long-run behavior of the system. On the other hand, one may say that these assumptions are too strict. Let me shortly discuss how the model can be extended to a rather general class of co-evolutionary models with noise. For a detailed discussion let me refer to chapter 2. For sake of completeness, let me just sketch what the long run behavior of this model would be, if one drops Assumptions 1 and 2. Instead of (4.3.4), assume that the players' rate function equals $\lambda^i(\omega) = \lambda \mathbb{1}_{\{\kappa^i(\omega) < N-1\}}$, and λ is a positive constant. For sake of illustration, suppose the base game is a symmetric 2×2 coordination game with one Pareto efficient equilibrium (a_1, a_1), and one risk-dominant equilibrium (a_2, a_2). The specific payoffs are not important.[6] I claim that these small alterations of the model lead to a *non-selection* result. Any pair of players, which use the same action, may be connected in the long-run

[6] Of course, this is still a potential game.

equilibrium; putting it differently, as β goes to zero we do not obtain a point prediction as in Section 4.7, but the limit distribution will (in general) put positive weight on a proper subset of Ω. The heuristic explanation of this "negative" result is the following.

- Since the rate function of players is uncoupled with the noise parameter, the speed of the link creation process is unaffected by the level of noise. Looking back at (4.3.4), we see that as β goes to 0 the link creation process becomes arbitrary fast.

- The link destruction process deletes any edge with the constant rate ξ. This process is pure drift, i.e. it is independent of the base game, and in particular of the noise level β. In the terminology of stochastic stability calculus, this implies that link destructions are zero cost events. However, it turns out that the rate-ratio λ/ξ determines the number of links the system can carry in the long run.

- The logit choice function of the action adjustment process (4.3.2) puts equal probability on all actions a loner may choose. However, if a player has at least one neighbor and if β goes to zero, this player will play a best response against the neighbors' behavior with probability arbitrary close to 1.

- Suppose the system is currently in a full coordination state, say the population coordinates on the efficient equilibrium $(a_1, \ldots, a_1)$. The network will not be complete in general, but one can derive a distribution over networks, given this action configuration. If there are some loners in the current state, let them switch to a_2, and give them a link creation opportunity. These steps can be made with zero costs. Now, by definition of the coordination game, an optimal decision in the link creation process is to connect the a_2 players. We are then already in a state where a_1 and a_2 co-exist. At this state no player has an incentive to change his action, so we will not return to the state we were coming from. If there are no loners, we can construct a sequence of link destruction, action adjustment and link creation events, all causing no costs, which leads to a state where two coordination equilibria co-exist, as follows: Destroy the links of player i. Give him an action adjustment opportunity where he chooses a_2. Since a loner may choose any action with equal probability without making an error, this causes no costs. Do the same thing with player $j \neq i$. Then give them a link creation opportunity. Since i and j are the only agents playing a_2, an optimal decision in the link creation process is to create the link (i, j). Now we are in a co-existence state and no player has an incentive to change his action.

- In the same vein we can walk through the set

$$\Omega^* = \{\omega \in \Omega | g^i_j = 1 \Rightarrow \alpha^i = \alpha^j\}$$

without any costs, in the sense of stochastic stability analysis. As a result, all states contained in this set are stochastically stable.

A similar result, but with admittedly sharper limit predictions, is obtained by Jackson and Watts (2002). These authors add to the drift term ξ a direction, by assuming that only links where at least one player is better off after the destruction of the link, are very likely to become destroyed. For a fairly large set of parameters (such as linking costs as in Section 4.7) they also get a co-existence result. However, due to this directionality in the link destruction process, they get sharper limit results in the network dimension under the assumptions that the costs per link are constant.

4.9 Conclusion

This paper presented a stylized model on the co-evolution of networks and play in the class of potential games. Assumption 4.3.1 was crucial to derive a closed-form solution of the unique invariant distribution and to obtain sharp predictions as the noise in the players' decision rules goes to zero. A general selection theorem of potential maximizers applies in this case. Without Assumption 4.3.1 the invariant distribution can still be completely characterized, but the model loses its predictive power in the low-noise limit. It seems therefore that some assumptions in this direction are needed if one wants to obtain sharp limiting predictions.
There are many possible routes for extensions. In the next chapter I analyze the current model with Assumption 4.3.1, but assuming an inverse relationship in the rate function with the size of the population. The intuition is that a larger population should make it less likely that a single agent receives the chance to create a link. In the infinitely large population limit and small positive noise the generated networks do not converge to complete graphs anymore. Hence, nicer asymptotic results are obtained without losing much in analytical power.
A more fundamental question is, however, which class of networks (in the sense of random graph theory) such co-evolutionary models are capable to create. The general framework of chapter 2 is a first step in this direction.

There we have seen that such models seem to generate, under fairly mild assumptions on the structure of the random process, so-called inhomogeneous random graphs (see e.g. the nice survey by Newman, 2003). These models are straightforward extensions of the classical Erdös-Rényi model, where the edge-success probabilities depend on the attributes of the individual vertices. It would be interesting to see how deep this connection indeed is.

4.10 Proofs of selected Theorems and Propositions

Proof of Theorem 4.4.1. Uniqueness follows from irreducibility and recurrence of $\boldsymbol{\eta}^{\beta}$.
By construction of the dynamics, we know that changes occur in the process only in one "coordinate": either a single change in the links of the network takes place, or one, and almost surely only one, player switches to another action. By statistical independence of these two processes we can treat them separately. Start with a change in the network structure. It suffices to consider the creation of a fresh link. Let $\omega = (\alpha, g), \hat{\omega} = (\alpha, g \oplus (i,j)) \in \Omega$. The rate of link creation between players i and j is given by eq. (4.3.8). The rate with which one returns to the state ω is eq. (4.3.9). Detailed balance (4.4.1) demands that

$$\frac{\mu^{(\beta,\xi)}(\hat{\omega})}{\mu^{(\beta,\xi)}(\omega)} = \frac{2}{\xi}\exp(u(\alpha^i,\alpha^j)/\beta). \tag{4.10.1}$$

It is easy to see that the measure (4.4.2) satisfies this condition.
Now consider the event of action adjustment. Let player k be the one who receives such an opportunity and suppose she switches to action $a_v \in \mathcal{A}$. Let $\omega, \hat{\omega} = (\alpha_k^{a_v}, g) \in \Omega$ be the states involved in this transition. The associated rate ratio is

$$\begin{aligned}
\frac{\eta^{\beta}(\omega \to \hat{\omega})}{\eta^{\beta}(\hat{\omega} \to \omega)} &= \frac{\nu b^{k,\beta}(a_v|\omega)}{\nu b^{k,\beta}(\alpha^k|\omega)} \\
&= \exp\left[\frac{1}{\beta}\left(\sum_{j:g_j^k=1}[u(a_v,\alpha^j) - u(\alpha^k,\alpha^j)]\right)\right].
\end{aligned} \tag{4.10.2}$$

Rewrite the invariant distribution as

$$\mu^{(\beta,\xi)}(\omega) \propto \prod_{i=1}^{k}\prod_{j>i}\left[\frac{2}{\xi}\exp(u(\alpha^i,\alpha^j)/\beta)\right]^{g_j^i} \times \prod_{i=k+1}^{N}\prod_{j>i}\left[\frac{2}{\xi}\exp(u(\alpha^i,\alpha^j)/\beta)\right]^{g_j^i}.$$

Note that the second term on the right-hand side does not depend on player k, and thus the change in the action of this player does have no effect on this term. Hence, we see that the probability ratio boils down to

$$\frac{\mu^{(\beta,\xi)}(\hat{\omega})}{\mu^{(\beta,\xi)}(\omega)} = \prod_{i=1}^{k}\prod_{j>i}\left[\exp\left(\frac{u(\hat{\alpha}^i,\hat{\alpha}^j)-u(\alpha^i,\alpha^j)}{\beta}\right)\right]^{g_j^i}.$$

Since $\hat{\alpha}^i = \alpha^i$ for all $i \neq k$, $\hat{\alpha}^k = a_v$, and payoffs as well as the indicators g_j^i are symmetric, we see that

$$\begin{aligned}\frac{\mu^{(\beta,\xi)}(\hat{\omega})}{\mu^{(\beta,\xi)}(\omega)} &= \prod_{j=1}^{N}\left[\exp\left(\frac{u(a_v,\alpha^j)-u(\alpha^k,\alpha^j)}{\beta}\right)\right]^{g_j^k} \\ &= \exp\left[\frac{1}{\beta}\left(\sum_{j:g_j^k=1}[u(a_v,\alpha^j)-u(\alpha^k,\alpha^j)]\right)\right].\end{aligned}$$

This is the rate ratio (4.10.2). □

Lemma 4.10.1. *Fix an action partition $\boldsymbol{\mathcal{I}}$ and let $\omega \in \Omega(\boldsymbol{\mathcal{I}})$. Define*

$$m^{(\beta,\xi)}(\omega|\boldsymbol{\mathcal{I}}) = \prod_{r=1}^{q}\prod_{i\in\mathcal{I}_r}\left\{\prod_{v\geq r}\prod_{j\in\mathcal{I}_v;j>i}\left[\frac{2}{\xi}\exp\left(\frac{u(a_r,a_v)}{\beta}\right)\right]^{g_j^i}\right\}\mathbb{1}_{\{\omega\in\Omega(\boldsymbol{\mathcal{I}})\}} \tag{4.10.3}$$

the mass of state ω, conditional on the event that the action partition $\boldsymbol{\mathcal{I}}$ is realized. Then the mass received by the set $\Omega(\boldsymbol{\mathcal{I}})$ in the long run is given by

$$m^{(\beta,\xi)}(\Omega(\boldsymbol{\mathcal{I}})) = \prod_{r=1,v\geq r}^{q}\left(1-p_{r|v}^{(\beta,\xi)}\right)^{-\frac{|\mathcal{I}_r|(|\mathcal{I}_v|-\delta_{r,v})}{1+\delta_{r,v}}}. \tag{4.10.4}$$

Proof. We have to compute $\sum_{\omega\in\Omega} m^{(\beta,\xi)}(\omega|\boldsymbol{\mathcal{I}})$. On $\Omega(\boldsymbol{\mathcal{I}})$ the action profile is fixed, and all states differ only in the number of edges. We can write

$$m^{(\beta,\xi)}(\omega|\boldsymbol{\mathcal{I}}) = \prod_{r=1,v\geq r}^{q}\left(\frac{p_{r|v}^{(\beta,\xi)}}{1-p_{r|v}^{(\beta,\xi)}}\right)^{e_{r|v}(\omega)}.$$

Hence

$$\begin{aligned} m^{(\beta,\xi)}(\Omega(\boldsymbol{\mathcal{I}})) &= \sum_{\omega\in\Omega(\boldsymbol{\mathcal{I}})} m^{(\beta,\xi)}(\omega|\boldsymbol{\mathcal{I}}) \\ &= \prod_{r=1,v\geq r}^{q} \sum_{k=0}^{\frac{|\mathcal{I}_r|(|\mathcal{I}_v|-\delta_{r,v})}{1+\delta_{r,v}}} \binom{\frac{|\mathcal{I}_r|(|\mathcal{I}_v|-\delta_{r,v})}{1+\delta_{r,v}}}{k} \left(\frac{p_{r|v}^{(\beta,\xi)}}{1-p_{r|v}^{(\beta,\xi)}}\right)^k \\ &= \prod_{r=1,v\geq r}^{q} \left(1-p_{r|v}^{(\beta,\xi)}\right)^{-\frac{|\mathcal{I}_r|(|\mathcal{I}_v|-\delta_{r,v})}{1+\delta_{r,v}}}. \end{aligned}$$

□

Proof of Theorem 4.5.1. (a) Lemma 4.10.1 shows that (4.5.1) is given by

$$\mu^{(\beta,\xi)}(\omega|\boldsymbol{\mathcal{I}}) = \frac{m^{(\beta,\xi)}(\omega|\boldsymbol{\mathcal{I}})}{m^{(\beta,\xi)}(\Omega(\boldsymbol{\mathcal{I}}))}.$$

A direct calculation of this ratio gives Eq. (4.5.3).

(b) For ease of notation, I omit the superscripts (β,ξ). Define

$$F_{\boldsymbol{\mathcal{I}}}(\theta) := \log \mu(\Omega(\boldsymbol{\mathcal{I}})). \tag{4.10.5}$$

Using the representation via the Gibbs measure (4.4.6), one gets

$$\mu(\Omega(\boldsymbol{\mathcal{I}})) = \sum_{\omega\in\Omega(\boldsymbol{\mathcal{I}})} \mu(\omega|\boldsymbol{\mathcal{I}}) = \sum_{\omega\in\Omega(\boldsymbol{\mathcal{I}})} e^{h(\omega,\theta)}.$$

Taking the partial derivative with respect to $\theta_{r|v}$ gives

$$\begin{aligned} \frac{\partial F_{\boldsymbol{\mathcal{I}}}(\theta)}{\partial\theta_{r|v}} &= \frac{1}{\mu(\Omega(\boldsymbol{\mathcal{I}}))} \sum_{\omega\in\Omega(\boldsymbol{\mathcal{I}})} \frac{\partial h(\omega,\theta)}{\partial\theta_{r|v}} e^{h(\omega,\theta)} \\ &= \sum_{\omega\in\Omega(\boldsymbol{\mathcal{I}})} \frac{\partial h(\omega,\theta)}{\partial\theta_{r|v}} \mu(\omega|\boldsymbol{\mathcal{I}}) \\ &= \mathbb{E}_\mu\left[\frac{\partial h(\omega,\theta)}{\partial\theta_{r|v}} \mid \boldsymbol{\mathcal{I}}\right]. \end{aligned}$$

Since the Hamiltonian h is a linear function in the edge parameter θ, the first equation follows. Moreover, linearity of h in θ implies that all higher-order derivatives vanish. To obtain the covariance relationship observe that

$$\frac{\partial^2 F_{\boldsymbol{\mathcal{I}}}(\theta)}{\partial\theta_{r|v}\partial\theta_{r|l}} = \sum_{\omega\in\Omega(\boldsymbol{\mathcal{I}})} \frac{\partial h(\omega,\theta)}{\partial\theta_{r|v}} \frac{\partial\mu(\omega|\boldsymbol{\mathcal{I}})}{\partial\theta_{r|l}}. \tag{4.10.6}$$

Compute

$$\frac{\partial \mu(\omega|\boldsymbol{\mathcal{I}})}{\partial \theta_{r|l}} = \frac{\partial h(\omega,\theta)}{\partial \theta_{r|l}}\mu(\omega|\boldsymbol{\mathcal{I}}) - \mu(\omega|\boldsymbol{\mathcal{I}})\frac{\partial F_{\boldsymbol{\mathcal{I}}}(\theta)}{\partial \theta_{r|l}}.$$

Plugging this into Eq. (4.10.6) one gets

$$\begin{aligned}\frac{\partial^2 F_{\boldsymbol{\mathcal{I}}}(\theta)}{\partial \theta_{r|v}\partial \theta_{r|l}} &= \sum_{\omega\in\Omega(\boldsymbol{\mathcal{I}})} e_{r|v}(\omega)e_{r|l}(\omega)\mu(\omega|\boldsymbol{\mathcal{I}}) - \mathbb{E}[e_{r|v}|\boldsymbol{\mathcal{I}}]\mathbb{E}[e_{r|l}|\boldsymbol{\mathcal{I}}] \\ &= \mathrm{Cov}_{\mu^{(\beta,\xi)}}[e_{r|v}, e_{r|l}|\boldsymbol{\mathcal{I}}].\end{aligned}$$

To obtain the third equation, note that by definition of the function h we have $\frac{\partial h(\omega,\theta)}{\partial(1/\beta)} = P(\omega) = U(\omega)/2$.

(c) This follows directly from the product measure (4.5.3) and the definition of the Erdös-Rényi-model.

$\square$

Proof of Proposition 4.5.1. For ease of notation I skip again the parameters (β,ξ). $\kappa_1^i,\ldots\kappa_q^i$ are independent Binomially distributed random variables with respective parameters $(n_v-\delta_{r,v}, p_{r|v}), 1\le v\le q$. Thus

$$\mathbb{P}(\kappa^i = k_1,\ldots,\kappa_q^i = k_q|\boldsymbol{\mathcal{I}}, i\in\mathcal{I}_r) = \prod_{v=1}^{q}\mathbb{P}(\kappa_v^i = k_v|\boldsymbol{\mathcal{I}}, i\in\mathcal{I}_r),$$

where for $1\le v\le q$

$$\begin{aligned}\mathbb{P}(\kappa_v^i = k_v|\boldsymbol{\mathcal{I}}, i\in\mathcal{I}_r) &= \binom{n_v-\delta_{r,v}}{k_v} p_{r|v}^{k_v}(1-p_{r|v})^{n_v-\delta_{r,v}-k_v} \\ &= \left[f_{r|v}(k_v)\right]^{k_v}(1-p_{r|v})^{n_v-\delta_{r,v}}.\end{aligned}$$

There are $\frac{k!}{k_1!\cdots k_q!}$ ways to construct a list $(k_1,\ldots,k_q)$ whose sum equals k. Hence

$$\mathbb{P}(\kappa^i = k|\boldsymbol{\mathcal{I}}, i\in\mathcal{I}_r) \propto \sum_{k_1+\ldots+k_q=k}\frac{k!}{k_1!\cdots k_q!}\prod_{v=1}^{q}\mathbb{P}(\kappa^i = k_v|\boldsymbol{\mathcal{I}}, i\in\mathcal{I}_r).$$

In each of the products on the right hand side, the factor $(1-p_{r|v})^{n_v-\delta_{r,v}}$ is a constant and so cancels out after normalization. Hence, define the normalization factor

$$R(\boldsymbol{\mathcal{I}}) = \sum_{k=0}^{N-1}\mathbb{P}(\kappa^i = k|\boldsymbol{\mathcal{I}}, i\in\mathcal{I}_r),$$

and call $f_{r,\kappa}(k|\boldsymbol{\mathcal{I}}) := \mathbb{P}(\kappa^i = k|\boldsymbol{\mathcal{I}}, i\in\mathcal{I}_r)$ to get the desired result. $\square$

Proof of Theorem 4.7.1. For any $\varepsilon > 0$ consider the set

$$A_\varepsilon := \{\omega \in \Omega | P(\omega) < \max_{\omega' \in \Omega} P(\omega') - \varepsilon\}.$$

I will show that $\lim_{\beta \to 0} \mu^{(\beta,\xi)}(A_\varepsilon) = 0$. Let $P^* := \max_{\omega' \in \Omega} P(\omega')$ the global maximum value of the potential function, and $\mathcal{P} := \arg\max_{\omega \in \Omega} P(\omega)$. Define the measure $\mu_0^\xi : \mathcal{G} \to [0, \infty], g \mapsto \mu_0^\xi(g) := (2/\xi)^{e(g)}$. Since the Hamiltonian of the Gibbs measure is additive separable in the measure μ_0^ξ and the potential function P, we get for all $\omega = (\alpha, g)$

$$\mu^{(\beta,\xi)}(\alpha, g) \propto e^{\frac{1}{\beta} H(\omega,\beta,\xi)} = e^{\frac{1}{\beta} P(\omega)} \mu_0^\xi(g).$$

The set A_ε can be written as

$$A_\varepsilon = \{\omega \in \Omega | e^{-\frac{1}{\beta} P(\omega)} > e^{-\frac{1}{\beta}(P^* - \varepsilon)}\}.$$

Markov's inequality[7] gives us

$$\mu^{(\beta,\xi)}(A_\varepsilon) \leq e^{\frac{1}{\beta}(P^* - \varepsilon)} \mathbb{E}_{\mu^{(\beta,\xi)}} \left[e^{-\frac{1}{\beta} P} \right],$$

where

$$\mathbb{E}_{\mu^{(\beta,\xi)}}[e^{-\frac{1}{\beta} P}] = \sum_{\omega = (\alpha, g) \in \Omega} \mu^{(\beta,\xi)}(\omega) e^{-\frac{1}{\beta} P(\omega)} = \frac{1}{Z} \sum_{\omega = (\alpha, g) \in \Omega} \mu_0^\xi(g),$$

with $Z = \sum_{\omega \in \Omega} e^{\frac{1}{\beta} H(\omega,\beta,\xi)} \geq |\mathcal{P}| e^{\frac{1}{\beta} H^*}$, where H^* is the minimum value of the Hamiltonian at states $\omega \in \mathcal{P}$, i.e. if $\omega = (\alpha, g) \in \mathcal{P}$ then $H^* = P^* + \min_{\omega \in \mathcal{P}} \beta \log \mu_0^\xi(g)$. Let $K := \min\limits_{\omega = (\alpha, g) \in \mathcal{P}} \mu_0^\xi(g) > 0$ be the minimum value of the graph measure on the set of potential maximizers. Thus, $H^* = P^* + \beta \log K$, and so $Z \geq |\mathcal{P}| K e^{\frac{1}{\beta} P^*} > 0$. Next, we compute

$$\begin{aligned}
\sum_{\omega = (\alpha, g) \in \Omega} \mu_0^\xi(g) &= \sum_{n \in \mathcal{D}} \frac{N!}{n_1! \cdots n_q!} \prod_{r=1}^{q} \prod_{v \geq r} \sum_{i \in \mathcal{I}_r, j \in \mathcal{I}_v : g_j^i \in \{0,1\}} (2/\xi)^{g_j^i} \\
&= \sum_{n \in \mathcal{D}} \frac{N!}{n_1! \cdots n_q!} \prod_{r=1}^{q} \prod_{v \geq r} (1 + 2/\xi)^{\frac{n_r(n_v - \delta_{r,v})}{1 + \delta_{r,v}}} \\
&= q^N (1 + 2/\xi)^{\frac{N(N-1)}{2}}.
\end{aligned}$$

[7] See e.g. Grimmett and Stirzaker (2001), p.319.

where in the last step we have made use of the Multinomial Theorem. It follows that

$$\mu^{(\beta,\xi)}(A_\varepsilon) \leq \frac{1}{Z} q^N (1+2/\xi)^{\frac{N(N-1)}{2}} e^{\frac{1}{\beta}(P^*-\varepsilon)} \leq \frac{q^N(1+2/\xi)^{\frac{N(N-1)}{2}}}{K|\mathcal{P}|e^{\frac{1}{\beta}P^*}} e^{\frac{1}{\beta}(P^*-\varepsilon)}$$
$$= K_1 e^{-\frac{\varepsilon}{\beta}}.$$

where $K_1 := \frac{q^N(1+2/\xi)^{\frac{N(N-1)}{2}}}{|\mathcal{P}|K} > 0$ a factor independent of β and ε. For $\beta \to 0$ the upper bound goes to zero, establishing the result. $\square$

Chapter 5

Co-evolutionary dynamics and Bayesian interaction games

5.1 Introduction

In large populations of interacting agents, the decisions made by the individuals are often influenced by the decisions made by other individuals, which are met on a frequent basis. Such situations are most naturally modeled by *local interaction games* (Morris, 2000). Much of the early literature on interaction games (e.g. Ellison, 1993, Blume, 1993) assumed very specific forms on the architecture of the interaction structure of the population, but devoted much effort to study the long-run evolution of the actions chosen by the interacting individuals. This was mainly done in the class of symmetric 2×2 coordination games where a conflict between Pareto efficiency and risk considerations exists. A seemingly robust result of this literature is the selection of states where all individuals play the risk-dominant strategy. However, a recent literature has emphasized that also the *architecture* of the interaction structure may have an influence on the long-run behavior of the population. On the one hand Hofbauer (1999) and Morris (2000) have shown that in symmetric 2×2 coordination games a strict Nash equilibrium is *spatially dominant*, or *contagious*, if it is risk-dominant. On the other hand, Young (1998; 2003) and Morris (2000) show that such contagious dynamics depend on the geometry of the interaction structure, i.e. the social network on which interaction takes place. These authors give conditions under which a society is very likely to become victim of contagious dynamics. In a nutshell, these are low *clustering* and a certain form of *cohesiveness* of the network.[1] Morris (2000) asks then to go one step further and understand how likely these critical properties are to emerge. This paper is one small step in this direction. Building on earlier work by Ehrhardt et al. (2006a; 2008a) - and the model presented in chapter 4 - we present a co-evolutionary model of networks and play where agents' preferences are diverse. The static framework for these models will be a new class of games called *structured Bayesian interaction games* with N players, for short BG^N. Formally, a structured Bayesian interaction game is a pair $(\Gamma, \boldsymbol{m})$, where Γ is an interaction game and $\boldsymbol{m}$ is the distribution of types in the population. As in standard games of incomplete information we assume that the types of the players are determined before play actually starts. Once a type distribution has been realized we assume that the types the players are fixed and the co-evolutionary model with noise starts to shape the system. The co-evolutionary process will be ergodic, and so almost surely the process will reach its unique stationary distribution. Such a stationary distribution describes the probability distribution over states for an *interim* version of the Bayesian interaction game, which we would like to call an structured interim Bayesian interaction game, for short IBG^N. The game IBG^N is described by the interaction game Γ, and a *preference decomposition* $\mathcal{T}$, consisting of all sets of players who are of the same type. In the population game, every structure interim Bayesian interaction game can be seen as a random realization of a structured Bayesian interaction game BG^N, given the measure $\boldsymbol{m}$. A preference decomposition records the names of the players who are of a certain type. Dropping the names leads to a less informative version of an interim Bayesian interaction game, which is still useful for our purposes. We will consider anonymous structured Bayesian interaction games, $\mathrm{I\bar{B}G}^N$, where the preference decomposition $\mathcal{T}$ is replaced by a tuple $\boldsymbol{N}$ counting the number of players of a certain type.

The co-evolutionary model with noise is essentially the same as the process introduced in chapter 4. However, we make the transition probabilities dependent on the population size. This seemingly minor modification of the

[1] Clustering is a standard measure in the literature on networks. Intuitively it measures the number of triangles in a graph, or differently speaking, the probability that a neighbor of a neighbor of a vertex is again a neighbor of that vertex. Morris (2000) emphasizes p-cohesiveness. A subset of vertices is called p-cohesive if at least a fraction p of the neighbors of a vertex in this set are again members of the set. Young (1998), ch. 6, introduces close-knit graphs. A subset of vertices $\mathcal{V}$ is called p-close-knit if each subset $\mathcal{V}' \subset \mathcal{V}$ has at least fraction p of neighbors in $\mathcal{V}$. See also Jackson (2008), Vega-Redondo (2007). There is also a very recent literature on diffusion on networks with given degree distributions; See López-Pintado (2008), Jackson and Yariv (2008).

stochastic process will have huge consequences on the long-run behavior of the dynamics. On the one hand, it allows us to investigate the model in the limit of infinitely large populations, while holding the noise level constant. On the other hand, it improves the results found in chapter 4 in the sense that we obtain much nicer results on the network topology, provided we allow for some positive noise. We show in this paper that the co-evolutionary process is analytically solvable, and provide a closed-form solution for its unique invariant measure. This measure can be studied by introducing interaction Hamiltonians, a concept taken from Markov random fields (see Brémaud, 1998). One can think of these functions as the analog to the interaction potentials of Ui (2000). They allow us to write the distribution as a Gibbs measure. Starting from this, we are able to characterize the induced ensemble of random graphs, and we state a closed-form solution for the long-run measure governing the distribution of actions in the population.

5.1.1 Related literature

The definition of Bayesian interaction games is closely related to the definition of *Bayesian population games*, introduced by Ely and Sandholm (2005) and Sandholm (2007a), and the illuminating presentation of *interaction games* in Morris (1997; 2000). Ely and Sandholm (2005) define a Bayesian population game as a population version of a game with incomplete information. They assume a continuum population of mass 1, where nature distributes the types according to some given measure before play starts. Agents of a certain type form again a continuum, and the set of agents having a certain type forms a sub-populations with mass depending on the probability measure used by nature. Ely and Sandholm (2005) define a Bayesian strategy as a mapping assigning to each sub-population a mixed strategy, where the term mixed strategy has again a population interpretation in recording the relative frequency of agents playing some action. Given a distribution of types, they introduce a new class of evolutionary dynamics, the aggregate Bayesian best response dynamic, defined on the function space of Bayesian strategies. In a sense it can be seen as an extension of the well-known best response dynamic studied by Gilboa and Matsui (1991). Ely and Sandholm (2005) go on in showing that their Bayesian best response dynamic can be studied by means of a finite-dimensional best-response dynamic, which in turn has some similarities with the perturbed best response dynamics of Fudenberg and Levine (1998) and Hofbauer and Sandholm (2002). Sandholm (2007a) uses this dynamic to give a new (evolutionary) version of Harsanyi's purification theorem (Harsanyi, 1973). As in Harsanyi (1973), he assumes that all players share a common payoff function, but each individual may have an id-

iosyncratic preference over the common set of actions, which is interpreted as the type of a player. We follow this set-up, but propose a different definition of Bayesian population games, where *population structure*, embodied by an undirected graph, is a characterizing feature of the model. Beside this crucial fact, we differ from Sandholm (2007a) in two other aspects. First, while Sandholm (2007a) studies a deterministic mean-field dynamic, we characterize completely the long-run behavior of an ergodic stochastic process.
Second, although we share the interest in perturbed best response dynamics, the perturbation structure is completely different. Sandholm (2007a) uses the type distribution as perturbation device, while we *fix* the type distribution and assume that the interim payoffs of the players, i.e. after they have learned their own type, but not the type of any other agent, are perturbed in a specific stochastic way. Hence, we emphasize the appropriate interim definitions of Bayesian games, IBG^N and $\mathrm{I\bar{B}G}^N$, respectively. In this model we are able to investigate the behavior of the invariant measure as the population gets large. To make this limiting operation meaningful we consider a sequence of *Bayesian strategy* profiles, generated by a sequence of anonymous preference decompositions. As in Ely and Sandholm (2005) and Sandholm (2007a) we call a Bayesian strategy profile a strategy distribution describing the aggregate behavior of the players of each type.
The rest of the paper is organized as follows. Section 5.2 defines structured Bayesian interaction games, and thereby fixes the static framework of the model. Section 5.3 introduces co-evolutionary dynamics with noise as a collection of perturbed continuous-time Markov processes, inspired by Catoni (1999). Section 5.4 discusses the asymptotic properties of the process and section 5.5 studies the small-noise behavior of the invariant distribution for finite populations. Section 5.6 shows that the generated random graphs are so-called inhomogeneous random graphs, as studied by Söderberg (2002) and Bollobás et al. (2007). Finally, section 5.7 provides an expression for the long-run measure over action profiles for all levels of noise and population sizes, as well the behavior of the measure in the limit of infinite populations. Section 5.8 concludes. A technical Appendix at the end of the paper collects the relevant proofs.

5.2 Structured Bayesian interaction games

Following Morris (2000), we define an *interaction game* as a tuple $\Gamma = (\mathcal{I}, \mathcal{G}[\mathcal{I}], \mathcal{A}, \pi)$. The set $\mathcal{I} = \{1, \ldots, N\}$ denotes the finite set of players, which is interpreted as a single population of interacting agents. Interaction is modeled by an evolving undirected graph $G \in \mathcal{G}[\mathcal{I}]$. Denote by $\mathcal{I}^{(2)}$ the set of unordered pairs of players. In this paper we identify a network as a graph in the set $\mathcal{G}[\mathcal{I}]$. An element of this set is a pair $G = (\mathcal{I}, \mathcal{E})$, where $\mathcal{E} = \mathcal{E}(G) \subset \mathcal{I}^{(2)}$ is the set of undirected edges (links). The *size* of a graph G is its number of edges $|\mathcal{E}(G)| =: e(G)$. For $G \in \mathcal{G}[\mathcal{I}]$, we call $G' \in \mathcal{G}[\mathcal{I}]$ a *subgraph* of G if $\mathcal{E}(G') \subset \mathcal{E}(G)$. For two disjoint subsets $\mathcal{V}, \mathcal{V}' \subset \mathcal{I}$, denote the set of edges that connect vertices from $\mathcal{V}$ to vertices from $\mathcal{V}'$ (and vice versa) as $\mathcal{E}(\mathcal{V}, \mathcal{V}')$. Fixing the vertex set, we can identify every graph on $\mathcal{G}[\mathcal{I}]$ through its edge set. Therefore, we can give an alternative representation of an undirected graph via a tuple $g = (g_{ij})_{1 \le i < j \le N} \in \{0, 1\}^{\mathcal{I}^{(2)}}$, where the functions g_{ij} are indicator functions on the set of edges $\mathcal{E}$. If $g_{ij} = 1$ then individual i is connected to individual j, or j is a neighbor of i (and vice versa). If $g_{ij} = 0$ then i and j are not connected. By defining g as

$$(i, j) \in \mathcal{E}(G) \Leftrightarrow g_{ij} = 1$$

we can establish a one-to-one correspondence (up to a permutation of the players' labels) between $\mathcal{G}[\mathcal{I}]$ and the set of edge realizations $\{0, 1\}^{\mathcal{I}^{(2)}}$, members of which are vectors g. In view of this equivalence let us identify a graph by the corresponding tuple g, and call henceforth $\mathcal{G}[\mathcal{I}]$ the set of all edge realizations $g = (g_{ij})_{1 \le i \le j \le N}$.
The neighbors of player i are contained in the set $\mathcal{N}^i(g) = \{j \in \mathcal{I} | g_{ij} = 1\}$. Call $\bar{\mathcal{N}}^i(g) = \mathcal{N}^i(g) \cup \{i\}$. The number of neighbors of player i defines his degree $\kappa^i(g) := |\mathcal{N}^i(g)|$. Given a graph g and a subset of players $\mathcal{V} \subseteq \mathcal{I}$ denote the restriction of g on $\mathcal{V}$ as $g[\mathcal{V}]$, which is an element of $\mathcal{G}[\mathcal{V}]$. The complete graph on the subset $\mathcal{V}$ is denoted by $g^c[\mathcal{V}]$. For every $g \in \mathcal{G}[\mathcal{I}]$ and a partition of $\mathcal{I}$ into sets $\mathcal{V}_1, \mathcal{V}_2$ one can write $g = g[\mathcal{V}_1] \oplus g[\mathcal{V}_2]$, where $\oplus$ is interpreted as the concatenation of two lists of binary valued functions (after possibly relabeling the players). In this notation, the graph $g' = g \oplus g^c[\{(i, j)\}] =: g \oplus (i, j)$ is the graph we obtain by adding the edge (i, j) to g. Analogously, $g' = g \ominus (i, j)$ is the graph obtained from g by deleting edge (i, j).
$\mathcal{A} = \{a_1, \ldots, a_n\}$ is the common set of actions among which players can choose. The utility player i gets from action a_v depends on the prevailing interaction structure g, and the action profile $\alpha = (\alpha^i)_{i \in \mathcal{I}}$, which specifies one action for every player. The pair $\omega := (\alpha, g) \in \mathcal{A}^{\mathcal{I}} \times \mathcal{G}[\mathcal{I}] \equiv \Omega^N$ defines a *population state*. Call $\alpha_i^{a_v}$ the action profile where all players choose the same action as prescribed in α, but player i plays action a_v. Utility of player i at population state ω in the interaction game Γ is defined as

$$u^i(\alpha, g) \equiv u^i(\omega) := \sum_{j \in \mathcal{N}^i(g)} \pi(\alpha^i, \alpha^j). \tag{5.2.1}$$

In this paper we assume that the common payoff function π defines an exact potential game (Monderer and Shapley, 1996). A sufficient condition for this is that it has the *partnership structure* (Hofbauer and Sigmund, 1998)

$$(\forall a, a' \in \mathcal{A}): \ \pi(a, a') = \pi(a', a). \tag{5.2.2}$$

Example 5.2.1. *The model allows for a much richer payoff structure than the one postulated by eq. (5.2.1). To the function π, one can add terms that do not depend on the action of a single player without destroying the structure of a potential game.*[2] *An interesting setting would be the following. Assume that for every link player i is incident to, a fee $\frac{c(\kappa^i)}{\kappa^i}$ has to be paid.*[3] *The function $c : \{0, 1, 2 \ldots, N-1\} \to \mathbb{R}_+$ measures the total fee a player with κ neighbors has to pay. Define the utility function of player i in the interaction game as*

$$u^i(\alpha, g) := \sum_{j \in \mathcal{N}^i(g)} \left(\pi(\alpha^i, \alpha^j) - \frac{c(\kappa^i(g))}{\kappa^i(g)} \right),$$

where π satisfies (5.2.2). All the following results would go through if we work with such a per-interaction payoff function. One could go further by assuming that every player has its own cost function c^i which depends exclusively on his degree κ^i. Similar functional forms for the per-interaction payoff function have been used by Bala and Goyal (2000), Jackson and Watts (2002) and Goyal and Vega-Redondo (2005).

Given an interaction game Γ, each player may have idiosyncratic preferences over the elements in the common action set $\mathcal{A}$. Beside the interaction structure g, this introduces a second source of heterogeneity into the population. Formally, let us assume that if player i chooses action α^i in the interaction game Γ, his *subjective utility* is given by

$$U^i(\alpha, g) \equiv U^i(\omega) := u^i(\omega) + \theta^i(\alpha^i). \tag{5.2.3}$$

The map $\theta^i : \mathcal{A} \to \mathbb{R}$ is the idiosyncratic preference of player i, which we will interpret as the *type* of player i. The *type space* is the set of functions $\Theta := \{\theta : \mathcal{A} \to [a, b]\}$. We assume a finite variety of types, so that Θ consists of $0 \leq K \leq N$ different functions, and K is exogenously given.[4]

Following Ely and Sandholm (2005) and Sandholm (2007a), this notion of

[2] See Sandholm (2009b) and Sandholm (2009a) for characterizations of potential games in this direction.

[3] Define $\frac{0}{0} = 0$.

[4] In case of finite populations this assumption is vacuous.

types allows us to define a *structured Bayesian interaction game* with finite type space $\Theta = \{\theta_1, \ldots, \theta_K\}$ and population size N as the pair $\mathrm{BG}^N := (\Gamma, \boldsymbol{m})$. Γ is an interaction game, and $\boldsymbol{m} := (m_r)_{r=1}^K$ is a probability distribution, with m_r the prior probability that a player is of type θ_r.
We implicitly assume that agents learn their types before the evolutionary process starts to shape the system. Thus, we can alternatively represent a *realized* type distribution via a *preference decomposition* $\mathcal{T} = \{\mathcal{T}_r\}_{r=1}^K$. Each set $\mathcal{T}_r$ is the set of players of type r. Once a type distribution is realized a preference decomposition is fixed. This would indeed correspond to an interim definition of a game with incomplete information, where exactly the players in the set $\mathcal{T}_r$ have learned that they are of type r, but do not know the types of the other players. The resulting *interim structured Bayesian interaction game* is the tuple $\mathrm{IBG}^N = (\Gamma, \mathcal{T})$. It can be viewed as the empirical counterpart to BG^N. Conditional on IBG^N, we will write the interim payoffs of a player i as $U^i(\omega|\mathcal{T})$.[5]
Example 5.2.2 will be our "role model", which is used recurrently in order to illustrate the used concepts more clearly.

Example 5.2.2. *The specification of utility functions (5.2.3) is particularly interesting in situations where the interaction game models a game with strategic complements. Suppose the players may choose between two alternatives $\mathcal{A} = \{a_1, a_2\}$. The function π models the externalities individuals exert on their direct neighbors. For instance, assume that a_2 represents Microsoft Windows and a_1 Ubuntu Linux, and we want to model a situation where an individuals has to decide which operating system to use. Two connected players communicate most efficiently if both use the same operating system.[6] Windows is said to be more user-friendly than Linux, thus networks much more efficiently than the latter. To capture this formally, assume that the individuals play the following coordination game:*

	a_1	a_2
a_1	$(1,1)$	$(0,0)$
a_2	$(0,0)$	$(4,4)$

If two users of the same operating system meet, they are able to communicate easily. Windows users have however the advantage that their operating sys-

[5] The conditioning on the complete preference realization $\mathcal{T}$ is redundant, since a player learns only his own type, not the types of the other agents in the population. This notation is chosen for convenience, and will hopefully not cause any confusion.

[6] See Katz and Shapiro (1994) for more on competing technologies and networked markets. In particular it is mentioned in this work that product choices for a "network good" (e.g. the choice of an operating system), depend crucially on the expected size of the network, a feature that is nicely captured in this example.

tem networks more efficiently, since it is much easier to use. This advantage is translated in the game by assuming that Windows is the risk dominant action.
Agents can also have their own tastes about the operating system they want to use. For instance, people with profound computer skills may have a strong preference for using Linux. Hence, let us assume the population consists of two separate groups $\mathcal{T} = \{\mathcal{T}_1, \mathcal{T}_2\}$, which are characterized through the preference biases

	θ_1	θ_2
a_1	8	1
a_2	0	1

Agents from group 1 are the Linux aficionados; In the absence of any network effects they would prefer to use Linux. Agents of group 2 have no real opinion about the two operating systems. Let i be a representative agent from group $\mathcal{T}_1$, and let ω be the current population state. Denote by $\kappa_2^i(\omega)$ the number of i's neighbors who use Windows. His payoff from using Linux is $U^i(\alpha_i^{a_1}, g|\mathcal{T}) = \kappa^i(\omega) - \kappa_2^i(\omega) + 8$, while utility from using Windows is given by $U^i(\alpha_i^{a_2}, g|\mathcal{T}) = 4\kappa_2^i(\omega)$. Adopting Windows becomes a best response for i iff $\frac{\kappa_2^i(\omega)}{\kappa^i(\omega)} > \frac{1}{5} + \frac{8}{5\kappa^i(\omega)}$. Hence, if player i communicates with many people, the critical ratio of Windows users in his neighborhood, which induces him to become a Windows user too, is close to 0.2. Suppose that $\kappa^i(\omega) = 5$. Then it suffices that 3 neighbors use Windows to make Windows the optimal decision for i. If the preference for Linux would be absent, a single neighbor is enough to make Windows optimal for player i.

This example demonstrates a general feature of the effect of idiosyncratic preferences. A player who has only few neighbors will show a choice behavior, which may drift away significantly from the interaction game Γ. This captures the intuition that a player with few interactions should be less dependent on the behavior of his "peers". However, a player with large degree must coordinate his behavior with many people. This interaction effect might dominate the idiosyncratic preferences.
Instead of taking a preference decomposition as the primitive in the definition of a structured interim Bayesian interaction game, we might want to have a more quantitative description of the model. Therefore, we will alternatively consider the "anonymous" version of a structured interim Bayesian interaction game, where the realized type distribution $\mathcal{T}$ is replaced by the vector $\boldsymbol{N} = (N_1, \ldots, N_K)$. Each entry of this vector measures the absolute size of the subpopulation of players with type r, and $\sum_r N_r = N$. An anonymous structured interim Bayesian interaction game with population size N is the pair $\bar{\text{IBG}}^N = (\Gamma, \boldsymbol{N})$.

Remark 5.2.1. *In Fudenberg and Tirole (1991), pp. 227, a discussion on the interpretation of the notion of types in games with incomplete information is provided. The interim approach to Bayesian games is very much in line with our definition of (anonymous) interim Bayesian interaction games. The anonymous version also allows us to study games with large player sets. Suppose there is a countable infinite sampling population, say* $\mathbb{N}$*, from which nature independently samples players. A sampled player* $i \in \mathbb{N}$ *is independently assigned to be of type* θ_r *with probability* m_r*. Nature repeats this sampling/assignment-procedure* N *times, without replacement.*[7] *The probability that there are* N_r *players of type* $1 \leq r \leq K$ *after* N *independent trials, such that* $\sum_{r=1}^{K} N_r = N$*, is therefore multinomial distributed with mass function*

$$P_{\boldsymbol{m}}(\boldsymbol{N}) = \frac{N!}{N_1! \cdots N_k!} m_1^{N_1} \cdots m_K^{N_K}. \tag{5.2.4}$$

A random vector $\boldsymbol{N} = (N_1, \ldots, N_k)$ *defines an anonymous preference decomposition. The empirical frequency* $\hat{m}_r^N := N_r/N$ *is a maximum likelihood estimator of the true parameter* m_r*, and the strong law of large numbers implies that* $\hat{m}_r^N \to m_r$ *almost surely, as* $N \to \infty$*.*

Consider the model $\bar{\text{IBG}}^N = (\Gamma, \boldsymbol{N})$. Given the preference decomposition $\boldsymbol{N}$, let Δ denote the $n-1$ dimensional unit simplex, and define

$$\mathcal{X}_r^N(\boldsymbol{N}) := \{\sigma = (\sigma(s))_{1 \leq s \leq n} \in \mathbb{R}^n | \sigma(s) \in \{0, \frac{1}{N_r}, \ldots, \frac{N_r - 1}{N_r}, 1\} \ \& \ \sum_{s=1}^{n} \sigma(s) = 1\}$$

for all $1 \leq r \leq K$ with $N_r > 0$. If $N_r = 0$, so that type r does not appear in the population, set $\mathcal{X}_r^N(\boldsymbol{N}) = \emptyset$. The sets $\mathcal{X}_r^N(\boldsymbol{N})$ are discrete (inner) approximations of the simplex Δ, with coarseness factor N_r. If $N_r \to \infty$ as $N \to \infty$, we have $\mathcal{X}_r^N(\boldsymbol{N}) \to \Delta$. We call a $n \times 1$ column vector $\sigma_r^N \in \mathcal{X}_r^N(\boldsymbol{N})$ an *interim Bayesian strategy*. It completely describes the aggregate behavior of the players of type r, given the preference decomposition $\boldsymbol{N}$, by recording the frequencies with which the actions are played in this subpopulation. An *interim Bayesian strategy profile* is the $n \times K$ matrix

$$\boldsymbol{\sigma}^N = \begin{bmatrix} \sigma_1^N(1) & \sigma_2^N(1) & \ldots & \sigma_K^N(1) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1^N(n) & \sigma_2^N(n) & \ldots & \sigma_K^N(n) \end{bmatrix} \in \prod_{r=1}^{K} \mathcal{X}_r^N(\boldsymbol{N}). \tag{5.2.5}$$

Note that, since K is fixed, we can define Bayesian strategy profiles for any population size N.

[7]This is not important, since the sample is infinitely large so that sampling with replacement is equivalent to sampling without replacement.

5.3 Co-evolution with noise

Based on the general discussion in chapter 2, let us define a co-evolutionary model with noise as a tuple

$$\mathcal{M}^{\beta,N} = (\Omega^N, \mathcal{F}, \mathbb{P}, (Y^{\beta,N}(t))_{t\geq 0})_{\beta\in\mathbb{R}_+},$$

which consists of a probability measure $\mathbb{P} : \mathcal{F} \to [0,1]$, a family of Ω^N-valued random variables $Y^{\beta,N} = (Y^{\beta,N}(t))_{t\geq 0}$ indexed by a noise parameter $\beta \geq 0$, the population size N, and a continuous-time parameter $t \geq 0$. The dynamic will be active on three levels: action adjustment, link creation and link destruction. To regulate the frequency with which the process undertakes one of these steps, we introduce random variables $T^A_{(t)}, T^L_{(t)}, T^D_{(t)}$, which describe the random amount of time that elapses before any of the mentioned events take place, conditional on the population state at time t. The workings of the process is as follows:[8]

Action adjustment: We adapt the revision dynamics of Blume (1993; 2003) and Hofbauer and Sandholm (2007). For each player define a constant intensity Poisson process with parameter $\nu \geq 0$. By the superposition principle of the Poisson process (see e.g. Grimmett and Stirzaker, 2001) the random variable $T^A_{(t)}$ has an exponential distribution with rate $N\nu$. Conditional on the event of an action adjustment, each player faces the same probability $1/N$ to be granted a switching opportunity. If player i is selected for a revision opportunity, assume he switches to action $a \in \mathcal{A}$ with probability determined by the logit choice function

$$b^{i,\beta}(a|\omega) = \frac{\exp(\frac{1}{\beta}U^i(\alpha^a_i, g))}{\sum_{a'\in\mathcal{A}} \exp(\frac{1}{\beta}U^i(\alpha^{a'}_i, g))}. \tag{5.3.1}$$

The *rate* of the transition $(\alpha, g) = \omega \to (\alpha^a_i, g) = \omega'$ is therefore

$$\eta^{\beta,N}_{\omega,\omega'} = \nu b^{i,\beta}(a|\omega). \tag{5.3.2}$$

Link creation: Consider the following version of the *stochastic actor-model*, introduced by Snijders (2001). For each player i, let us define a *rate function* $\lambda^i : \Omega^N \to \mathbb{R}_+$. This function regulates the frequency with which player i will become active in the event of link creation. To be specific, define for every player i a conditionally independent Poisson process with intensity $\lambda^i(\cdot)$, which will depend on the current population state $Y^{\beta,N}(t)$. Hence, conditional on $\{Y^{\beta,N}(t) = \omega\}$, the expected

[8] A full description of the Markov process is given in chapter 3.

time player i has to wait in the link creation process is $1/\lambda^i(\omega)$. The conditional distribution of the random variable $T^L_{(t)}$ is the exponential with rate

$$\bar{\lambda}^\beta(\omega) := \sum_{j\in\mathcal{I}} \lambda^j(\omega).$$

One can calculate that the conditional probability with which player i receives a link creation opportunity is exactly $\lambda^{i,\beta}(\omega)/\bar{\lambda}^\beta(\omega)$. We make the following assumption on the functional form of the rate functions of the players:

Assumption 5.3.1. *The players' rate function have the form*

$$(\forall i \in \mathcal{I}) : \lambda^{i,\beta}(\omega) := \frac{1}{N} \sum_{j \notin \bar{\mathcal{N}}^i(\omega)} \exp\left(\frac{1}{\beta}\pi(\alpha^i, \alpha^j)\right). \tag{5.3.3}$$

The intuition behind this formulation is the following; the larger the population, the less frequent a single agent should receive a link creation opportunity. Therefore the inverse relationship with the population size N. On the other hand, we would like to capture a situation where those agents who expect a high utility from interactions are relatively fast in their network formation activities. This is modeled via the sum of exponentials.[9]

If player i receives the opportunity to form a link, we assume that he creates a new link with probability depending on the marginal payoff he would get from this connection. Given the functional assumption (5.2.3), this marginal utility is $U^i(\alpha, g \oplus (i,j)) - U^i(\alpha, g) = \pi(\alpha^i, \alpha^j)$. We assume that player i selects player j with probability

$$w_j^{i,\beta}(\omega) := \mathbb{P}\left(\pi(\alpha^i, \alpha^j) + \varepsilon^i_j \geq \pi(\alpha^i, \alpha^k) + \varepsilon^i_k \ \forall k \notin \bar{\mathcal{N}}^i(\omega) | \omega\right).$$

Assuming that the random variables $(\varepsilon^i_k)_{k\notin\bar{\mathcal{N}}^i(\omega)}$ are independent type 1 extreme value distributed, leads to the choice probability function[10]

$$w_j^{i,\beta}(\omega) = \frac{\exp\left(\pi(\alpha^i, \alpha^j)/\beta\right)}{\sum_{k\notin\bar{\mathcal{N}}^i(\omega)} \exp\left(\pi(\alpha^i, \alpha^k)/\beta\right)}. \tag{5.3.4}$$

[9] Observe that the *conditional* probability that player i gets a link creation opportunity is not affected by the population size. It is only the overall frequency of link creation opportunities which is scaled down by the population size.

[10] Note that this particular function is translation invariant, meaning that adding a constant term to the functions under the exponential cause no change in the probabilities. As a consequence, we could add to the interaction game payoff function (5.2.1) a player-specific cost function $c^i(\kappa)$, modeling a fee that player i has to pay in order to support κ links in the graph. See Example 5.2.1.

Since the creation of the link (i,j) can be caused by player i or player j (and almost surely by only one of these players), the *rate* of the transition $(\alpha, g) = \omega \to (\alpha, g \oplus (i,j)) = \omega'$ is given by

$$\eta^{\beta,N}_{\omega,\omega'} = \lambda^{i,\beta}(\omega) w^{i,\beta}_j(\omega) + \lambda^{j,\beta}(\omega) w^{j,\beta}_i(\omega) = \frac{2}{N} \exp\left(\pi(\alpha^i, \alpha^j)/\beta\right). \tag{5.3.5}$$

As in action adjustments, the likelihood ratio that player i chooses to link with j and not with k is $\exp(\frac{\pi(\alpha^i,\alpha^j)-\pi(\alpha^i,\alpha^k)}{\beta})$, i.e. an increasing function in the direct (marginal) benefit of the link (i,j) compared to (i,k). In the limit $\beta \to 0$, again, only links with the highest per-interaction payoff are going to be formed with probability 1. Hence, if one interprets the value of the net-payoff function as a distance measure between two individuals of the society, this choice rule reflects in a natural and intuitive way that people who are "closer" (in the sense of closeness defined by the game) are more likely to interact. There are two potential criticisms of the link formation process. First, we formulate the process as purely symmetric in incentives, since only the payoff function π determines the probability with which a free link becomes created. The idiosyncratic preferences of the players do not appear in this formulation. However, in terms of individual rationality, it is certainly only π which decides how valuable a new link for a player is. Therefore, it is rational for a player not to look at the idiosyncratic preference term in this subprocess. A second, and more substantial, criticism is the huge amount of information agents are assumed to possess. In large populations it is not likely that individuals have global information (even subject to noisy data) about the population state. We implicitly assume this however, since otherwise the agents would not be able to compute the probability distribution (5.3.4).[11]

Link destruction: If we would only consider a link creation process, then the dynamics will almost surely lead to the trivial stationary state where all agents interact with each other, so that the interaction structure will be the complete graph. Therefore, to describe an interesting dynamic, we need to add a process of link removal. Here we build on recent work by Marsili et al. (2004), Ehrhardt et al. (2006a; 2008a;b) by introducing a *volatility scheme*, meeting the following requirements;

[11] A more natural formulation would be a mixture between global search for partners (a rare event) and a local, network dependent, process of search (a frequent event). In Vega-Redondo (2007), ch. 6, one can find models in this direction, which are analytically much more complicated to handle (if at all). See also Vega-Redondo (2006) for an interesting simulation study of such a process.

Definition 5.3.1. *A $N \times N$ matrix $\Xi^\beta := [\xi^\beta_{ij}]_{1\leq i,j\leq N}$ is called a volatility scheme if*

- *Ξ^β is a symmetric matrix,*
- *$(\forall i,j \in \mathcal{I}) : \xi^\beta_{ij} \geq \xi > 0$, for some exogenous background volatility level ξ.*

The function ξ^β_{ij} is the rate with which edge (i,j) becomes destroyed.

If Ξ^β defines a volatility scheme then every currently existing edge has the expected survival time $1/\xi^\beta_{ij}$.[12] For analytical tractability we additionally require the following:

Definition 5.3.2. *Consider the structured interim Bayesian interaction game* $\mathrm{IBG}^N = (\Gamma, \mathcal{T})$. *Given a volatility scheme Ξ^β we call the $K \times K$ matrix $\hat{\Xi}^\beta$ a reduced form volatility scheme if*

$$\xi^\beta_{ij} = \xi^\beta_{rl} \text{ if } i \in \mathcal{T}_r \text{ and } j \in \mathcal{T}_l. \tag{5.3.6}$$

Let the *volume* of the link destruction process be $\bar{\xi}(\omega) = \sum_{j>i} \xi^\beta_{ij} g_{ij}$. By the Poisson structure of the process, the conditional probability that the link (i,j) becomes selected for destruction is $\frac{\xi^\beta_{ij} g_{ij}}{\bar{\xi}(\omega)}$. The random variable $T^D_{(t)}$ follows the exponential distribution with rate $\bar{\xi}(\omega)$. The *rate* of the transition $(\alpha, g) = \omega \to (\alpha, g \ominus (i,j)) = \omega'$ is therefore

$$\eta^{\beta,N}_{\omega,\omega'} = \xi^\beta_{ij} g_{ij}. \tag{5.3.7}$$

This, rather informal, description of the stochastic dynamics, will now be used to derive the law of the random variables $(Y^{\beta,N}(t))_{t\geq 0}$. The total *volume* of the process is measured by the rate function $\Lambda^{\beta,N} : \Omega \to \mathbb{R}_+$ defined as

$$\Lambda^{\beta,N}(\omega) = \nu \sum_{i=1}^{N} \sum_{a\in\mathcal{A}\setminus\{\alpha^i\}} b^{i,\beta}(a|\omega) + \bar{\lambda}^\beta(\omega) + \bar{\xi}(\omega).$$

Intuitively, this function aggregates all the rates of possible transitions away from ω. For positive noise β and finite N, the volume of the process cannot explode, and is bounded away from zero.

Observation 5.3.1. *For $\beta > 0$ and finite N the volume of the process $(Y^{\beta,N}(t))_{t\geq 0}$ has the following properties:*

[12] Again see chapter 3 for derivations of this fact.

(i) $\sup_{\omega\in\Omega}\Lambda^{\beta,N}(\omega)<\infty$*;*

(ii) $(\forall\omega\in\Omega):\Lambda^{\beta,N}(\omega)>0$.

No complications will arise by defining the distribution of the process in terms of its infinitesimal generator (see Stroock, 2005, for further details). This is a linear operator $\boldsymbol{\eta}^{\beta,N}:\Omega\times\Omega\to\mathbb{R}$, with the properties

(G1) $-\infty<\eta^{\beta,N}_{\omega,\omega}\leq 0$,

(G2) for all $\omega'\neq\omega$, $\eta^{\beta,N}_{\omega,\omega'}\geq 0$, and

(G3) $\sum_{\omega'\in\Omega}\eta^{\beta,N}_{\omega,\omega'}=0$.

Equations (5.3.2), (5.3.5), (5.3.7) already define the mappings of this operator. For future reference, let us summarize them as

$$\eta^{\beta,N}_{\omega,\omega'}=\begin{cases}\nu b^{i,\beta}(a|\omega) & \text{if}\omega'=(\alpha_i^a,g)\text{ and }a\neq\alpha^i,\\ \frac{2}{N}\exp\left(\pi(\alpha^i,\alpha^j)/\beta\right) & \text{if }\omega'=(\alpha,g\oplus(i,j)),\\ \xi^{\beta}_{ij}g_{ij} & \text{if }\omega'=(\alpha,g\ominus(i,j)),\\ -\Lambda^{\beta,N}(\omega) & \text{if }\omega'=\omega,\\ 0 & \text{otherwise.}\end{cases}\tag{5.3.8}$$

It is easy to verify that (5.3.8) has the properties (G1)-(G3). Moreover, for $\beta>0$ and finite N, the generator is irreducible and recurrent. Hence, the process is an ergodic continuous-time Markov process. The transition matrix of the process $(Y^{\beta,N}(t))_{t\geq 0}$, denoted as $\mathbf{K}^{\beta,N}$, is implicitly defined as $\boldsymbol{\eta}^{\beta,N}=\boldsymbol{\Lambda}^{\beta,N}(\mathbf{K}^{\beta,N}-\mathbf{Id})$, where $\boldsymbol{\Lambda}^{\beta,N}:=\text{diag}[\Lambda^{\beta,N}(\omega)]_{\omega\in\Omega}$ and $\mathbf{Id}$ is the identity on $\Omega\times\Omega$.

5.4 The invariant distribution

We are able to completely characterize the unique invariant distribution of the Markov process $\{Y^{\beta,N}(t)\}_{t\geq 0}$. This is a consequence of the fact that the constructed dynamic is reversible in equilibrium, and so we can try as an *Ansatz* for the characterization of the equilibrium distribution the detailed balanced equations (Stroock, 2005, ch.5). Let $\mu^{\beta,N}$ be a probability distribution with support Ω^N. This measure is said to be in detailed balance with the generator $\boldsymbol{\eta}^{\beta,N}$ if

$$(\forall\omega,\hat{\omega}\in\Omega^N):\mu^{\beta,N}(\omega)\eta^{\beta,N}_{\omega,\hat{\omega}}=\mu^{\beta,N}(\hat{\omega})\eta^{\beta,N}_{\hat{\omega},\omega}.\tag{5.4.1}$$

One can show that (5.4.1) is a sufficient condition for determining an invariant measure (and hence an invariant distribution) for the process $(Y^{\beta,N}(t))_{t\geq 0}$ (again see Stroock, 2005, for more details).

Theorem 5.4.1. *For any noise level $\beta > 0$ and irrespective of the realization of types, the Markov process $(Y^{\beta,N}(t))_{t\geq 0}$ admits the unique invariant distribution*

$$\mu^{\beta,N}(\omega) = Z^{-1}\prod_{j=1}^{N}\prod_{k>j}\left[\frac{2\exp(\pi(\alpha^j,\alpha^k)/\beta)}{N\xi^{\beta}_{jk}}\right]^{g_{jk}}\exp(\theta^j(\alpha^j)/\beta) \qquad (5.4.2)$$

for every $\omega \in \Omega^N$. The partition function is

$$Z := \sum_{\hat{\omega}\in\Omega^N}\prod_{j=1}^{N}\prod_{k>j}\left[\frac{2\exp(\pi(\hat{\alpha}^i,\hat{\alpha}^j)/\beta)}{N\xi^{\beta}_{jk}}\right]^{g_{jk}}\exp(\theta^j(\hat{\alpha}^j)/\beta).$$

Proof. The proof is virtually identical to the proof of Theorem 4.4.1 of chapter 4, and therefore omitted. □

In the following derivations, we will fix one *arbitrary* preference decomposition $\mathcal{T}$ and extract much information from the invariant distribution. By conditioning on one particular preference decomposition, the volatility scheme is fixed by Definition 5.3.2. Hence, one part of the Markov process is deterministically given, leaving the action profile as the only remaining free variable which will determine the asymptotic connectivity pattern of the population. Since the normalization of the probability distribution is rather uninteresting, we focus in the following on the measure

$$\rho^{\beta,N}(\omega|\mathcal{T}) := \prod_{j=1}^{N}\prod_{k>j}\left[\frac{2\exp(\pi(\alpha^j,\alpha^k)/\beta)}{N\xi^{\beta}_{jk}}\right]^{g_{jk}}\exp(\theta^j(\alpha^j)/\beta).$$

Using the information contained in a preference decomposition, and the model reduction hypothesis of the link decay rates (5.3.6), the measure factorizes to

$$\begin{aligned}\rho^{\beta,N}(\omega|\mathcal{T}) &= \prod_{r=1,l\geq r}^{K}\prod_{i\in\mathcal{T}_r}\prod_{j\in\mathcal{T}_l;j>i}\left[\frac{2\exp(\pi(\alpha^j,\alpha^k)/\beta)}{N\xi^{\beta}_{rl}}\right]^{g_{jk}}\exp(\theta_r(\alpha^i)/\beta)\\ &= \prod_{r=1,l\geq r}^{K}\rho^{\beta,N}_{rl}(\omega|\mathcal{T}).\end{aligned}$$

The individual functions $\rho_{rl}^{\beta,N}(\omega)$ can be interpreted as the mass the process puts on interactions between players of type r who interact with players of type l. We will call these measures the (r,l) *interaction mass*. Formulated in terms of interaction masses, the invariant distribution admits the representation

$$\mu^{\beta,N}(\omega|\mathcal{T}) = \frac{\rho^{\beta,N}(\omega|\mathcal{T})}{\sum_{\omega'\in\Omega^N}\rho^{\beta,N}(\omega'|\mathcal{T})}.$$

5.4.1 Interaction potentials and interaction Hamiltonians

Corresponding to the interaction masses, we can define group specific *interaction Hamiltonians* (Brémaud, 1998, ch.7). These are real-valued functions capturing the contribution of the factor $\rho_{rl}^{\beta,N}(\cdot)$ to the overall mass $\rho^{\beta,N}(\cdot)$. Interaction Hamiltonians consist of two elements; *interaction potentials* and a "distortion term" coming from the volatility affecting the network.[13] Given a preference decomposition $\mathcal{T}$, consider the functions

$$\begin{aligned}
P_{rr}(\omega|\mathcal{T}) &:= \sum_{i\in\mathcal{T}_r}\theta_r(\alpha^i) + \sum_{i\in\mathcal{T}_r;j>i}\pi(\alpha^i,\alpha^j)g_{ij} \quad (1\leq r\leq K);\\
P_{rl}(\omega|\mathcal{T}) &:= \sum_{i\in\mathcal{T}_r}\sum_{j\in\mathcal{T}_l}\pi(\alpha^i,\alpha^j)g_{ij} \quad (l>r);\\
P(\omega|\mathcal{T}) &:= \sum_{r=1}^{K}\left[P_{rr}(\omega|\mathcal{T}) + \sum_{l>r}P_{rl}(\omega|\mathcal{T})\right].
\end{aligned}$$

Lemma 5.4.1. *For any preference decomposition $\mathcal{T}$, the function $P:\Omega^N\to\mathbb{R}$ is a potential function in the sense that*

(i) If $\omega=(\alpha,g)$ and $\hat{\omega}=(\alpha_i^{a_v},g)$, then

$$P(\omega|\mathcal{T}) - P(\hat{\omega}|\mathcal{T}) = U^i(\omega|\mathcal{T}) - U^i(\hat{\omega}|\mathcal{T}).$$

(ii) If $\omega=(\alpha,g)$ and $\hat{\omega}=(\alpha,g\oplus(i,j))$ and say player i was the one who initiated the link, then

$$P(\omega|\mathcal{T}) - P(\hat{\omega}|\mathcal{T}) = U^i(\omega) - U^i(\hat{\omega}|\mathcal{T}).$$

[13] This is closely related to Hamiltonian statistical mechanics. Just interpret "volatility" as kinetic energy, and the potential function as a measure of the potential energy of the system. Ui (2000) establishes a beautiful connection between interaction potentials and the Shapley value.

Proof. For (i), let us consider two states $\omega, \hat{\omega}$ as required in the text. Without loss of generality assume that $i \in \mathcal{T}_r$. Then it follows that the symmetric interaction potentials $P_{\tilde{r}\tilde{r}}, \tilde{r} \neq r$ are unaffected by this change. We count

$$\begin{aligned}
P(\omega|\mathcal{T}) - P(\hat{\omega}|\mathcal{T}) &= \sum_{j\in\mathcal{T}_r} \left[\theta_r(\alpha^j) - \theta_r(\hat{\alpha}^j)\right] + \sum_{j,k\in\mathcal{T}_r;k>j} \left[\pi(\alpha^j,\alpha^k) - \pi(\hat{\alpha}^j,\hat{\alpha}^k)\right] g_{jk} \\
&+ \sum_{l=1}^{r-1}\sum_{j\in\mathcal{T}_l}\sum_{k\in\mathcal{T}_r} \left[\pi(\alpha^j,\alpha^k) - \pi(\hat{\alpha}^j,\hat{\alpha}^k)\right] g_{jk} \\
&+ \sum_{l=r+1}^{K}\sum_{j\in\mathcal{T}_r}\sum_{k\in\mathcal{T}_l} \left[\pi(\alpha^j,\alpha^k) - \pi(\hat{\alpha}^j,\hat{\alpha}^k)\right] g_{jk} \\
&= \theta_r(\alpha^i) - \theta_r(\hat{\alpha}^i) + \sum_{j\in\mathcal{T}_r} \left[\pi(\alpha^i,\alpha^j) - \pi(\hat{\alpha}^i,\alpha^j)\right] g_{ij} \\
&+ \sum_{l=1}^{r-1}\sum_{j\in\mathcal{T}_l} \left[\pi(\alpha^i,\alpha^j) - \pi(a_v,\alpha^j)\right] g_{ij} + \sum_{l=r+1}^{K}\sum_{j\in\mathcal{T}_l} \left[\pi(\alpha^i,\alpha^j) - \pi(a_v,\alpha^j)\right] g_{ij} \\
&= U^i(\alpha, g|\mathcal{T}) - U^i(\alpha_i^{a_v}, g|\mathcal{T}).
\end{aligned}$$

Now for (ii), take $\omega, \hat{\omega}$ as required and suppose, without loss of generality, that $j \in \mathcal{T}_l$. It is clear that from the link creation only the (r,l) interaction potential can be affected. Hence

$$\begin{aligned}
P(\omega|\mathcal{T}) - P(\hat{\omega}|\mathcal{T}) &= P_{rl}(\omega|\mathcal{T}) - P_{rl}(\hat{\omega}|\mathcal{T}) \\
&= \pi(\alpha^i,\alpha^j) = U^i(\omega|\mathcal{T}) - U^i(\hat{\omega}|\mathcal{T}).
\end{aligned}$$

$\square$

To get from the interaction potential to the interaction Hamiltonian, we have to take care of the effect of network volatility. Therefore we need a measure for the size of the graph. The total number of edges between players of preference group r and players of preference group l is given by

$$e_{rl}(\omega) := \sum_{i\in\mathcal{T}_r}\sum_{j\in\mathcal{T}_l;j>i} g_{ij} = |\mathcal{E}(\mathcal{T}_r,\mathcal{T}_l)|.$$

It turns out that counting the number of edges connecting players of two preference groups is a too crude measure for our purposes. We need to work with the graph measure $\mu_0^{\beta,N}(\cdot|\mathcal{T}) : \mathcal{G}[\mathcal{I}] \to [0,\infty]$, defined as

$$\mu_0^{\beta,N}(g|\mathcal{T}) := \prod_{r=1}^{K}\prod_{l\geq r}\left(\frac{2}{N\xi_{rl}^{\beta}}\right)^{e_{rl}(g)}$$

for all $g \in \mathcal{G}[\mathcal{I}]$. Consider now the set of functions

$$H_{rl}(\omega|\mathcal{T}) := P_{rl}(\omega|\mathcal{T}) + \beta \log\left(\frac{2}{N\xi_{rl}^{\beta}}\right) e_{rl}(\omega) \quad (1 \leq r \leq l \leq K) \tag{5.4.3}$$

$$H(\omega|\mathcal{T}) := \sum_{r=1}^{K}\sum_{l\geq r} H_{rl}(\omega|\mathcal{T}) = P(\omega|\mathcal{T}) + \beta \log \mu_0^{\beta,N}(g|\mathcal{T}) \tag{5.4.4}$$

We then get the following "statistical mechanics" version of Theorem 5.4.1.

Theorem 5.4.2. *Given the game* $\mathrm{IBG}^N = (\Gamma, \mathcal{T})$*, the unique invariant distribution* $\mu^{\beta,N}$ *is the Gibbs measure*

$$\mu^{\beta,N}(\omega|\mathcal{T}) = \frac{\exp(\frac{1}{\beta}H^{\beta}(\omega|\mathcal{T}))}{\sum_{\omega'\in\Omega^N}\exp(\frac{1}{\beta}H^{\beta}(\omega'|\mathcal{T}))} \tag{5.4.5}$$

for all $\omega \in \Omega^N$.

Proof. This follows directly from the definition of the Hamiltonian and Lemma 5.4.1. □

5.5 When volatility matters and weak stochastic stability

The Hamiltonian of eq. (5.4.4) provides a complete description of the system, and eq. (5.4.5) connects it with the invariant distribution of the process. Volatility enters the Hamiltonian only via the graph measure $\mu_0^{\beta,N}$. A fundamental question is now when this graph measure has an influence on the limiting behavior of the invariant distribution as the noise level β vanishes. If the effect coming from volatility is negligible as $\beta \to 0$, then the asymptotic prediction of the model boils down to finding the maximizers of the potential function P. The following Proposition provides a sufficient condition when this is the case.

Proposition 5.5.1. *Consider the structured Bayesian interaction game* $\mathrm{IBG}^N = (\Gamma, \mathcal{T})$*. If the volatility scheme* $\hat{\Xi}^{\beta}$ *is uniformly bounded as* $\beta \to 0$*, then*

$$\lim_{\beta\to 0}\sup_{\omega\in\Omega^N} |H^{\beta,N}(\omega|\mathcal{T}) - P(\omega|\mathcal{T})| = 0.$$

Proof. The proof is a simple consequence of the fact that the Hamiltonian is the sum of a potential function and the log-transformed β-weighted graph measure. Let $A^\beta := \sup_{g\in\mathcal{G}[\mathcal{I}]} |\log \mu_0^{\beta,N}(g|\mathcal{T})|$. Eq. (5.4.4) gives us the uniform bound

$$|H^{\beta,N}(\omega|\mathcal{T}) - P(\omega|\mathcal{T})| = \beta|\log \mu_0^{\beta,N}(g|\mathcal{T})| \leq \beta A^\beta,$$

for all $\omega = (\alpha, g) \in \Omega^N$. By assumption, the volatility rates are uniformly bounded from below. If they are additionally uniformly bounded from above, then $A^\beta \to A \in (0,\infty)$ as $\beta \to 0$, and so the claim follows. □

This proposition shows that the underlying volatility scheme must show explosive behavior for $\beta \to 0$ in order to have some influence on the shape of the limiting invariant distribution, and henceforth on the set of stochastically stable states. We use the following (weaker) notion of stochastic stability in our model.

Definition 5.5.1 (Sandholm (2009b))**.** *A state $\omega \in \Omega^N$ is weakly stochastically stable in the small noise limit if*

$$\lim_{\beta\to 0} \beta \log \mu^{\beta,N}(\omega|\mathcal{T}) = 0. \tag{5.5.1}$$

The following Theorem shows that, under the uniform boundedness assumption, stochastically stable states are determined by global maximizers of the potential function, as it is also the case in the fixed interaction models of Young (1998) and Blume (1993). Hence, in an extreme world where players almost never make mistakes, an extreme version of network volatility is needed to break the strong connection between stochastic stability and potential maximization. This finding is summarized as Theorem 5.5.1.

Theorem 5.5.1. *Under the assumptions of Proposition 5.5.1, let $\omega^* \in \arg\max_{\omega\in\Omega^N} P(\omega|\mathcal{T})$. Then for all $\omega \in \Omega^N$ we have*

$$\lim_{\beta\to 0} \beta \log \mu^{\beta,N}(\omega|\mathcal{T}) = P(\omega|\mathcal{T}) - P(\omega^*|\mathcal{T}).$$

Proof. This can be shown by adapting the arguments of the proof of Theorem 12.2.2 in Sandholm (2009b), p. 463. Using the uniform boundedness assumption of Proposition 5.5.1, we see that for any two states $\omega, \omega' \in \Omega^N$, and for β sufficiently small

$$\begin{aligned}
\beta \log \frac{\mu^{\beta,N}(\omega|\mathcal{T})}{\mu^{\beta,N}(\omega'|\mathcal{T})} &= H^{\beta,N}(\omega|\mathcal{T}) - H^{\beta,N}(\omega'|\mathcal{T}) \\
&= P(\omega|\mathcal{T}) - P(\omega'|\mathcal{T}) + \beta\left[\sum_{l\geq r} \log\left(\frac{2}{N\xi^\beta_{rl}}\right)(e_{rl}(g) - e_{rl}(g'))\right] \\
&= P(\omega|\mathcal{T}) - P(\omega'|\mathcal{T}) + o(1).
\end{aligned}$$

$o(1)$ represents terms that go to 0 as $\beta \to 0$. Hence, for all $\omega \in \Omega^N$

$$\lim_{\beta\to 0} \beta \log \mu^{\beta,N}(\omega|\mathcal{T}) = \lim_{\beta\to 0} \left[\beta \log \frac{\mu^{\beta,N}(\omega|\mathcal{T})}{\mu^{\beta,N}(\omega'|\mathcal{T})} - \log \frac{\mu^{\beta,N}(\omega^*|\mathcal{T})}{\mu^{\beta,N}(\omega'|\mathcal{T})} + \beta \log \mu^{\beta,N}(\omega^*|\mathcal{T})\right]$$
$$= P(\omega|\mathcal{T}) - P(\omega^*|\mathcal{T}) + \lim_{\beta\to 0} \beta \log \mu^{\beta,N}(\omega^*|\mathcal{T}).$$

We claim $\lim_{\beta\to 0} \beta \log \mu^{\beta,N}(\omega^*|\mathcal{T}) = 0$. Assume that the limit would be a number $c > 0$. Then we would be able to find null sequences $(\beta^k)_{k=0}^\infty$, $(\varepsilon^k)_{k=0}^\infty$ such that

$$\mu^{\beta^k,N}(\omega^*|\mathcal{T}) = \exp\left(\frac{1}{\beta^k}(c+\varepsilon^k)\right) \to +\infty.$$

This is a contradiction. Now assume that the limit is a constant $c < 0$. Then, for all $\omega \in \Omega^N$, we have

$$\lim_{\beta\to 0} \beta \log \mu^{\beta,N}(\omega|\mathcal{T}) = P(\omega|\mathcal{T}) - P(\omega^*|\mathcal{T}) + c \leq c < 0.$$

It follows that

$$\sum_{\omega\in\Omega^N} \mu^{\beta,N}(\omega|\mathcal{T}) = \sum_{\omega\in\Omega^N} \exp\left(\frac{1}{\beta^k}\beta^k \log \mu^{\beta^k,N}(\omega|\mathcal{T})\right)$$
$$\leq |\Omega^N| \exp\left(\frac{c}{2\beta^k}\right) \to 0$$

for $\beta \to 0$, where for the upper bound we have used the fact that $c/2 < 0$ is a uniform upper bound for the limit $\lim\limits_{\beta\to 0} \beta \log \mu^{\beta,N}(\omega|\mathcal{T})$. This is again a contradiction, leaving $c = 0$ as the only remaining possibility. We end up with the conclusion that

$$\lim_{\beta\to 0} \beta \log \mu^{\beta,N}(\omega|\mathcal{T}) = P(\omega|\mathcal{T}) - P(\omega^*|\mathcal{T})$$

for all $\omega \in \Omega^N$. □

Corollary 5.5.1. *Let*

$$\Omega^{*,N}(\mathcal{T}) := \{\omega \in \Omega^N | \lim_{\beta\to 0} \beta \log \mu^{\beta,N}(\omega|\mathcal{T}) = 0\}$$

denote the set of weakly stochastically stable states in the small noise limit for the realized preference decomposition $\mathcal{T}$. Under the assumptions of Proposition 5.5.1 we get the equivalence

$$\Omega^{*,N}(\mathcal{T}) = \arg\max_{\omega\in\Omega} P(\omega|\mathcal{T}).$$

This observation recovers well-known results from the theory of stochastic evolutionary game dynamics. See Blume (1993; 1997) and Young (1998), ch. 6.

5.6 The ensemble of random graphs

As in the previous models, we attempt to characterize the induced random graph ensemble as much as we can. In this section we demonstrate another version of the canonical result

> "A co-evolutionary model with noise generates an inhomogeneous random graph ensemble."

The derivations are however much more intricate as they were in chapter 4. The main result presented in this section is a version of an Erdös-Rényi decomposition, such as Theorem 4.5.1. Lemma 5.6.2 (The Factorization Lemma) and Proposition 5.6.1 are necessary preparatory results. The basic idea is the following; In order to compute a probability measure over networks, we need to fix the actions employed by the agents. We will do this, as in chapter 4, by action class partitions. For each action class partition we can calculate a probability measure over graphs, which will be of a similar form as the one found by Erdös and Rényi (1960), but the edge-success probabilities will have to be written for pairs of players belonging to different preference and action groups. As a result we obtain a random graph model on the deterministic vertex set $\mathcal{I}$ with an array of edge success probabilities given by $(p_{rl}^{sv}(\beta, N))_{1\leq r,s\leq K;1\leq l,v\leq n}$. This is exactly the inhomogeneous random graph model of Söderberg (2002) and a particular version of the very general model of Bollobás et al. (2007). We now present the formal derivations.

Consider the Bayesian interaction game $\text{IBG}^N = (\Gamma, \mathcal{T})$. For $1 \leq r \leq K$ and $1 \leq v \leq n$, define the set of type r-individuals who play action a_v,

$$I_r(v) := \{i \in \mathcal{I} | \alpha^i = a_v \ \& \ \theta^i = \theta_r\}.$$

The collection $\mathbf{I} := \{\{I_r(v)\}_{v=1}^n\}_{r=1}^K$ induces another partition on the society, which subdivides the several preference classes additionally into action classes. It gives us all the information about the behavior of the players, but no information about the interaction structure. A player from the set $I_r(v)$ will be called an (r, v)-player. Every population state ω induces a subdivision $\mathbf{I}(\omega)$. Since we have no information about the network when we look at $\mathbf{I}$, there will be, in general, several states ω which correspond to a subdivision $\mathbf{I}$. Therefore, for a fixed population subdivision $\mathbf{I}$ we can collect all states that agree with that subdivision, and write them as a set

$$\begin{aligned}\Omega^N(\mathbf{I}) &= \{\omega \in \Omega^N | I_r(v, \omega) = I_r(v), 1 \leq r \leq K \ , 1 \leq v \leq n\} \\ &= \{\omega \in \Omega^N | \mathbf{I}(\omega) = \mathbf{I}\}.\end{aligned}$$

All states in this set share the same action profile α, but differ in the network g. Let $e_{rl}^{sv}(g) := \sum_{i\in I_r(s)}\sum_{j\in I_l(v);j\geq i} g_{ij}$ count the number of edges between (r,v) and (l,s)-players. If $\omega \in \Omega^N(\mathbf{I})$ then the range of $e_{rl}^{vs}(\omega)$ is the fixed set $\{0,1,\dots,\frac{|I_r(s)|(|I_l(v)|-\delta_{r,l}\delta_{s,v})}{1+\delta_{r,l}\delta_{s,v}}\}$.

Lemma 5.6.1. *Let $\mathbf{I},\mathbf{I}'$ denote two subdivisions. If $\mathbf{I}\neq\mathbf{I}'$ then $\Omega^N(\mathbf{I})\cap\Omega^N(\mathbf{I}')=\emptyset$.*

Proof. If $\omega\in\Omega^N(\mathbf{I})\cap\Omega^N(\mathbf{I}')$, then the action profile described by the partitions $\mathbf{I},\mathbf{I}'$ must be the same, and so they contain the same population states. □

This simple observation allows us to give still another representation of the invariant distribution in terms of the sets $\Omega^N(\mathbf{I})$ generated by subdivisions $\mathbf{I}$. For all $\omega\in\Omega^N$ we can write

$$\mu^{\beta,N}(\omega|\mathcal{T}) = \frac{\rho^{\beta,N}(\omega|\mathcal{T})}{\sum_{\mathbf{I}}\rho^{\beta}(\Omega^N(\mathbf{I})|\mathcal{T})},$$

where the sum in the denominator is over all possible subdivisions $\mathbf{I}$.

Lemma 5.6.2 (The Factorization Lemma)**.** *Consider the preference decomposition $\mathcal{T}$, and a state $\omega\in\Omega^N$ with corresponding population subdivision $\mathbf{I}(\omega)$. For all $\omega\in\Omega^N$, the invariant measure $\rho^{\beta,N}$ can be represented as*

$$\rho^{\beta,N}(\omega|\mathcal{T}) := \prod_{s=1}^{n}\phi_1^s(\omega,\beta,N|\mathcal{T})\cdots\phi_K^s(\omega,\beta,N|\mathcal{T}),$$

where for all $1\leq r\leq K$, $l>r$ and $1\leq s\leq n$,

$$\phi_r^s(\omega,\beta,N|\mathcal{T}) := \exp\left(\frac{\theta_r(a_s)}{\beta}|I_r(s,\omega)|\right)\prod_{l\geq r}\phi_{rl}^s(\omega,\beta,N|\mathcal{T}),$$

$$\phi_{rr}^s(\omega,\beta,N|\mathcal{T}) := \prod_{v\geq s}\left(\frac{2\exp(\pi(a_s,a_v)/\beta)}{N\xi_{rr}^{\beta}}\right)^{e_{rr}^{sv}(\omega)},$$

$$\phi_{rl}^s(\omega,\beta,N|\mathcal{T}) := \prod_{v=1}^{n}\left(\frac{2\exp(\pi(a_s,a_v)/\beta)}{N\xi_{rl}^{\beta}}\right)^{e_{rl}^{sv}(\omega)}.$$

Proof. See Section 5.9. □

The trick behind computing the product is illustrated in the multiplication scheme depicted in Figure 5.1. The population subdivision $\mathbf{I}$ induces a block-

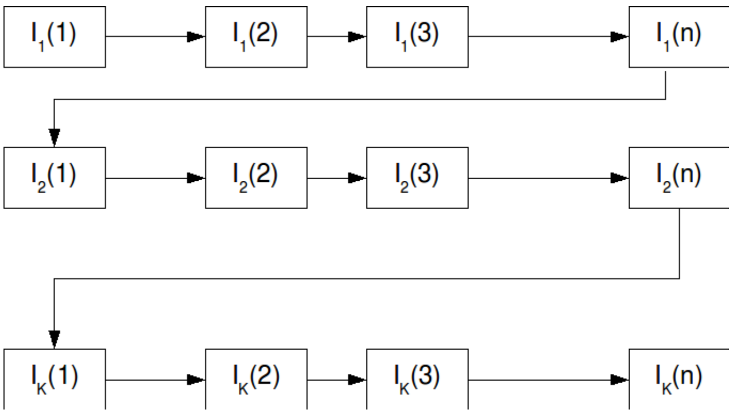

Figure 5.1: Multiplication scheme of the invariant measure $\rho^{\beta,N}$.

like structure on the population. Figure 5.1 shows how to calculate all the combinations between the blocks of the subdivision. To read this graphic, start with the outer-left block $I_1(1)$. Agents belonging to this block can interact with all other agents from blocks $I_1(v), v \geq 1$, and also with all other agents belonging to blocks $I_l(v), l > 1, 1 \leq v \leq n$. To get all these products, simply follow the lines with a ruler, starting from the block $I_1(1)$. Now, turn to the next block $I_1(2)$. We have already captured the interactions with $I_1(1)$, so we don't look backward, but forward. Again, follow the drawn lines with a ruler, starting at block $I_2(1)$. After n steps we have finished the multiplications beginning from blocks inside preference class 1. Then turn to the next line, i.e. $I_2(1)$. All the incoming lines of this block are terms which are already included in our product. So, again follow all lines going out of this block with a ruler, as indicated in the figure. The multiplication procedure stops in finite time. What we can learn from the Factorization Lemma is that the factor ϕ^s_{rl} is increasing in the number of edges with players from block $I_l(v)$ iff the *interaction weight*

$$d^{sv}_{rl}(\beta, N) := \frac{2\exp(\pi(a_s, a_v)/\beta)}{N\xi^{\beta}_{rl}} \tag{5.6.1}$$

is larger than 1. Phrased differently, the process favors interactions between player groups for which the "birth rate" (this is $\frac{2}{N}\exp(\pi(\alpha^i, \alpha^j)/\beta)$) of links exceeds their "death rate" (this is ξ^{β}_{ij}). This observation is very important to get more information about the role of the noise parameter β on the probabilistic interaction structure of the society and further shows how the interaction topology depends on the game the agents play. In particular, fixing a subdivision $\mathbf{I}$ restricts the state space to the subset $\Omega^N(\mathbf{I})$. On this set the action profile α is constant and only the network g is a free variable. Furthermore, the interaction weights between types are fixed on $\Omega^N(\mathbf{I})$. Only in this case it makes therefore sense to calculate the likelihood of certain networks g. In order to do this, we need to evaluate the mass a subset $\Omega^N(\mathbf{I})$ receives in the long-run by the process. Proposition 5.6.1 shows that this measure is a product over functions depending on the interaction weight $d_{rl}^{sv}(\beta, N)$ and the size of the sets $I_r(s)$, fixed by the subdivision $\mathbf{I}$.

Proposition 5.6.1. *Given a preference decomposition $\mathcal{T}$ and a population subdivision $\mathbf{I}$, the mass received by $\Omega^N(\mathbf{I})$ in the long-run equals*

$$\rho^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T}) = \sum_{\omega:\mathbf{I}(\omega)=\mathbf{I}} \rho^\beta(\omega|\mathcal{T}) = \prod_{r=1}^{K}\prod_{s=1}^{n} \Phi_r^s(\mathbf{I},\beta,N)^{|I_r(s)|} \tag{5.6.2}$$

where, for all types $1 \leq r < l \leq K$, and actions $1 \leq s \leq n$, the functions $\Phi_r^s(\cdot,\cdot,\cdot)$ are defined as

$$\begin{aligned}
\Phi_r^s(\mathbf{I},\beta,N) &:= \prod_{l\geq r} \Phi_{rl}^s(\mathbf{I},\beta,N),\\
\Phi_{rr}^s(\mathbf{I},\beta,N) &:= \exp\left(\frac{\theta_r(a_s)}{\beta}\right)\prod_{v\geq s}\left(1+\frac{2\exp(\pi(a_s,a_v)/\beta)}{N\xi_{rr}^\beta}\right)^{\frac{|I_r(v)|-\delta_{s,v}}{1+\delta_{s,v}}},\\
\Phi_{rl}^s(\mathbf{I},\beta,N) &:= \prod_{v=1}^{n}\left(1+\frac{2\exp(\pi(a_s,a_v)/\beta)}{N\xi_{rl}^\beta}\right)^{|I_l(v)|}.
\end{aligned}$$

Proof. See Section 5.9. □

The strategy of the proof is the following; Fixing a population subdivision automatically fixes the number of edges that can be formed between (r,s) and (l,v) players. Further, we do not put *any* restriction on the interaction structure when we calculate the mass that the set $\Omega^N(\mathbf{I})$ receives in equilibrium. This implies that we have to sum over all possible edges that can be formed between block (r,s) and (l,v), by accounting for the number of possibilities to form these edges.
Proposition 5.6.1 completely specifies an invariant distribution on the family of sets $\{\Omega^N(\mathbf{I})\}$. Furthermore, it shows that the invariant measure, determining such an invariant distribution, is of a particularly nice product form. Theorem 5.6.1 follows easily from Proposition 5.6.1, showing that the subgraphs on $\Omega^N(\mathbf{I})$[14] are Erdös-Rényi random graphs (Erdös and Rényi, 1960, Bollobás, 1998).

[14]See the proof of Proposition 5.6.1 in Section 5.9.

Theorem 5.6.1 (The Erdös-Rényi decomposition)**.** *Given a preference decomposition $\mathcal{T}$, a population subdivision $\mathbf{I}$ and a population size N. On $\Omega^N(\mathbf{I})$ the set of subgraphs $\mathcal{G}[I_r(s) \cup I_l(v)] \subset \mathcal{G}[\mathcal{I}]$ is an Erdös-Rényi random graph ensemble, with edge success probability $p_{rl}^{sv}(\beta, N) = \frac{2\exp(\pi(a_s,a_v)/\beta)}{2\exp(\pi(a_s,a_v)/\beta)+N\xi_{rl}^{\beta}}$.*

Proof. See Section 5.9. □

Example 5.6.1. *We continue with Example 5.2.2, where the Windows users and the Linux users interact. Consider an arbitrary population subdivision $\mathbf{I} = \{I_1(1), I_1(2), I_2(1), I_2(2)\}$. Suppose we have the following estimates on the underlying volatility scheme*

$$\hat{\Xi}^{\beta} = \xi \begin{bmatrix} e^{1/\beta} & e^{7/\beta} \\ e^{7/\beta} & e^{1/\beta} \end{bmatrix}, \ \xi > 0.$$

From the given data, we can compute a 4×4 matrix of edge success probabilities, displayed in Table 5.1. For positive β and finite N, we see that

	$I_1(1)$	$I_1(2)$	$I_2(1)$	$I_2(2)$
$I_1(1)$	$\frac{2\exp(1/\beta)}{2\exp(1/\beta)+N\xi\exp(1/\beta)}$	$\frac{2}{2+N\xi\exp(1/\beta)}$	$\frac{2\exp(1/\beta)}{2\exp(1/\beta)+N\xi\exp(7/\beta)}$	$\frac{2}{2+N\xi\exp(7/\beta)}$
$I_1(2)$	$\frac{2}{2+N\xi\exp(1/\beta)}$	$\frac{2\exp(4/\beta)}{2\exp(4/\beta)+N\xi\exp(1/\beta)}$	$\frac{2}{2+N\xi\exp(7/\beta)}$	$\frac{2\exp(4/\beta)}{2\exp(4/\beta)+N\xi\exp(7/\beta)}$
$I_2(1)$	$\frac{2\exp(1/\beta)}{2\exp(1/\beta)+N\xi\exp(7/\beta)}$	$\frac{2}{2+N\xi\exp(7/\beta)}$	$\frac{2\exp(1/\beta)}{2\exp(1/\beta)+N\xi\exp(1/\beta)}$	$\frac{2}{2+N\xi\exp(1/\beta)}$
$I_2(2)$	$\frac{2}{2+N\xi\exp(7/\beta)}$	$\frac{\exp(4/\beta)}{2\exp(4/\beta)+N\xi\exp(7/\beta)}$	$\frac{2}{2+N\xi\exp(1/\beta)}$	$\frac{2\exp(4/\beta)}{2\exp(4/\beta)+N\xi\exp(1/\beta)}$

Table 5.1: Edge-success probabilities of the subgraphs for $\beta > 0$.

all the edge success probabilities are positive, although interactions between two players of the same type are expected to occur much more frequent, than interactions across two preference groups. This well-studied phenomenon of homophily (Powell et al., 2005, Currarini et al., 2008) is frequently observed in real-world social networks. The noise parameter β controls the strength of homophily. For very large β, the edge success probabilities approach the constant $2/(2 + N\xi)$, independent of the preference group. Hence, in this extreme scenario, where a lot of noise exists in the economy, the taste difference of the two population groups play no role.[15] *Such a scenario is very likely to be observed in the initial phase of network evolution, where players don't know anything about each other. For very low β, formally $\beta \searrow 0$, a*

[15] Observe that in this extreme case all subgraphs follow the same distribution. Hence, as $\beta \to \infty$ we obtain the homogeneous random graph model of Erdös-Rényi as special case.

completely different prediction can be made. In this case the above matrix of edge success probabilities converges to a diagonal matrix as in Table 5.2. Only "same-type" links can survive in this extreme world.

	$I_1(1)$	$I_1(2)$	$I_2(1)$	$I_2(2)$
$I_1(1)$	$\frac{2}{2+N\xi}$	0	0	0
$I_1(2)$	0	1	0	0
$I_2(1)$	0	0	$\frac{2}{2+N\xi}$	0
$I_2(2)$	0	0	0	1

Table 5.2: Edge-success probabilities of the subgraphs for $\beta \to 0$.

5.7 Long-run action profiles

In order to characterize the generated ensemble of random graphs it was necessary to fix the types attached to all players' labels. This was done by taking a realized preference decomposition as the primitive of the model. However, as mentioned in Section 5.2, such a partition might contain more information than is necessary. Particularly, if one is only interested in an aggregate description of the population in terms of frequency statistics, the specific realized set of player types does not provide interesting information. Additionally, if one is only interested in a description of aggregate population behavior, again the realized set of types is not interesting. Here, the less demanding notion of anonymous interim Bayesian interaction games $\bar{\text{IBG}}^N$, introduced at the end of Section 5.2, is sufficient to work with. Henceforth, we take a vector $\boldsymbol{N} = (N_1, \dots, N_K)$ as information on the distribution of types in the population. Next, we consider the *behavior matrix* at state $\omega \in \Omega^N$

$$\mathbf{Z}(\omega) := \begin{bmatrix} z_1(1,\omega) & z_2(1,\omega) & \dots & z_K(1,\omega) \\ z_1(2,\omega) & z_2(2,\omega) & \dots & z_K(2,\omega) \\ \vdots & \vdots & \ddots & \vdots \\ z_1(n,\omega) & z_2(n,\omega) & \dots & z_K(n,\omega) \end{bmatrix},$$

where each entry of this matrix is defined as $z_r(s,\omega) := |I_r(s,\omega)|$. Hence, the r-th column of this matrix gives a frequency description of aggregate play in the subpopulation of players of type r.[16] Conditional on the type distribution $\boldsymbol{N}$, it must be true that $\mathbf{1}^T\mathbf{Z} = \boldsymbol{N}$, where $\mathbf{1}$ is a column vector of dimension

[16] The mapping $\omega \mapsto \mathbf{Z}(\omega)$ is in general many-to-one.

n consisting only of 1. Therefore, conditional on the Bayesian interaction game $\mathrm{I\bar{B}G}^N = (\Gamma, \boldsymbol{N})$, define the set of feasible behavior matrices as

$$\mathscr{Z}(\boldsymbol{N}) := \{\mathbf{Z} \in \mathbb{N}^{n\times K} | \mathbf{1}^T\mathbf{Z} = \boldsymbol{N}\}.$$

To get a probability distribution over matrices in this set, we first establish a connection between the set $\Omega^N(\mathbf{I})$, and the subset

$$\begin{aligned}\Omega^N(\mathbf{Z}) :&= \{\omega \in \Omega^N | \mathbf{Z}(\omega) = \mathbf{Z}\} \\ &= \{\omega \in \Omega^N | \, |I_r(s,\omega)| = z_r(s), 1 \leq r \leq K; 1 \leq s \leq n\}.\end{aligned}$$

When we were working on the set $\Omega^N(\mathbf{I})$, we have made use of the fact that the action profile α is constant. Consider the behavior matrix $\mathbf{Z}$ and let $\mathbf{I}$ be a subdivision which agrees with this matrix. Then every permutation of the players' labels still generates the matrix $\mathbf{Z}$. Hence, the mass received by the set $\Omega^N(\mathbf{Z})$ equals the mass of $\Omega^N(\mathbf{I})$, up to a permutation of the players' labels. This leads to the following result.

Proposition 5.7.1. *For a given $\boldsymbol{N}$, let $\mathbf{Z} \in \mathscr{Z}(\boldsymbol{N})$ be a feasible behavior matrix. The long-run mass on $\Omega^N(\mathbf{Z})$ is given by*

$$\gamma^{\beta,N}(\mathbf{Z}|\boldsymbol{N}) = \prod_{r=1}^{K} \frac{N_r!}{z_r(1)!\cdots z_r(n)!}\Phi_r^1(\mathbf{Z},\beta,N)^{z_r(1)}\cdots\Phi_r^n(\mathbf{Z},\beta,N)^{z_r(n)} \quad (5.7.1)$$

where the factors Φ are defined in Proposition 5.6.1.

Proof. Let $\mathbf{I}$ be a subdivision that agrees with the matrix $\mathbf{Z}$ and $\omega \in \Omega^N(\mathbf{I})$. From Proposition 5.6.1, we know that the mass concentrated on $\Omega^N(\mathbf{I})$ is given by eq. (5.6.2). All states in this set generate the same action profile α. Since $|I_r(s,\omega)| = z_r(s)$ for all $1 \leq r \leq K, 1 \leq s \leq n$ and $\omega \in \Omega^N(\mathbf{I})$ (recall that $\mathbf{I}$ agrees with $\mathbf{Z}$), we see that $\Omega^N(\mathbf{I}) \subseteq \Omega^N(\mathbf{Z})$. A permutation of the indices $\{1,\ldots,N\}$ changes the partition $\mathbf{I}$ to some other partition $\mathbf{I}'$ which still must agree with $\mathbf{Z}$. Thus, we see that $\Omega^N(\mathbf{I}') \subseteq \Omega^N(\mathbf{Z})$. In general, if we let $\mathcal{P}(\mathbf{I})$ denote the partitions obtained from $\mathbf{I}$ by permuting the players' labels, we see that $\mathcal{P}(\mathbf{I}) = \Omega^N(\mathbf{Z})$. Thus,

$$\gamma^{\beta,N}(\mathbf{Z}|\boldsymbol{N}) = \prod_{r=1}^{K} C_{r,\boldsymbol{N}}(\mathbf{Z}) \prod_{s=1}^{n} \Phi_r^s(\mathbf{Z},\beta,N)^{z_r(s)}$$

where $\mathcal{T}$ is a preference decomposition that agrees with $\boldsymbol{N}$, and $C_{r,\boldsymbol{N}}(\mathbf{Z})$ is a combinatorial term counting the number of permutations of the players' labels in the various sub-populations. There are $C_{r,\boldsymbol{N}}(\mathbf{Z}) = \frac{N_r!}{z_r(1)!\cdots z_r(n)!}$ permutations of the labels in the subpopulation of type r players which all agree with the r-th column of the matrix $\mathbf{Z}$, and so the claim follows. $\square$

Example 5.7.1. *Let us investigate the shape of the measure (5.7.1) in the case of Examples 5.2.2 and 5.6.1. We have seen in Example 5.2.2 that large neighborhoods make it more likely that Windows will be adopted, even by individuals who have a strong bias for Linux. In Example 2 we have shown that any finite population size N and positive noise level β, symmetric interactions will be observed with a higher probability. This pattern of homophily becomes stronger the smaller β. We will now investigate the dependency of the long run distribution over strategies on β. Fix the population size $N = 100$ and assume that both types appear in equal proportions in the population, i.e. $N_1 = N_2 = 50$. Assume the same data as we did in Example 5.2.2, and put $\xi = 1$. Since $z_r(1) + z_r(2) = N_r, 1 \leq r \leq 2$, there are only 2 independent variables, allowing us to visualize the invariant distribution in 3-dimensional space. In Figure 5.2 we present the invariant distribution for rather high β.*

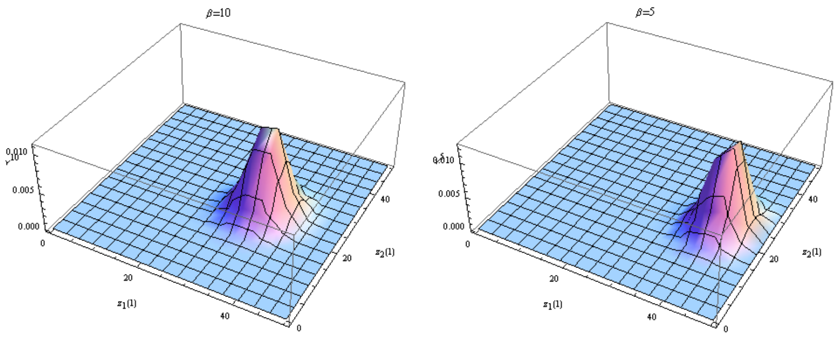

Figure 5.2: Invariant distribution over state matrices $\mathcal{Z}((50, 50))$ induced by the measure (5.7.1).

We see a nice bell shaped distribution, slightly biased in the population group of the Linux fans towards Linux. The bell shaped distribution is simply a reflection of the fact that high β means that individuals decisions are driven to a large extent by randomness. In Figure 5.3 we reduce β, leading to a slightly different image. We observe a much higher degree of specialization in the two population groups. While almost everybody of the 50 Linux fans indeed use Linux, the individuals without a real preference bias decided to use Windows as their operating system. For $\beta = 1$ the probability distribution over the possible strategy configurations is already heavily concentrated on a small subset of $\mathcal{Z}((50, 50))$. Finally, we decrease β to the very small value 0.5 (Figure 5.4). Suddenly a completely different behavior emerges in the population. Windows has emerged as the unique standard in the population.

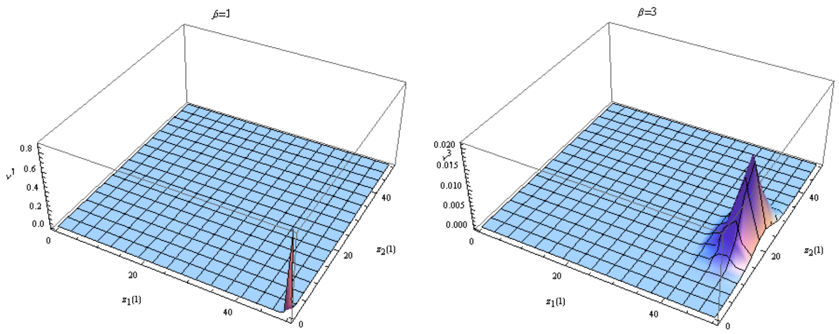

Figure 5.3: Invariant distribution over state matrices $\mathcal{Z}((50,50))$ induced by the measure (5.7.1).

The reason for this drastic change in the behavioral pattern can be found in Example 2. There we have seen that as $\beta \to 0$ with probability 1 only interactions among Windows users from the same preference group will take place. Again we see how the formulation of the dynamic process generating the graph of interaction influences the way how players behave in the model.

5.7.1 Infinite Population behavior

By now we have always considered the case N fixed, β variable (and maybe small). In this section we revert this, by taking $N \to \infty$, and β fixed and positive. In this large population environment we need to study a sequence of Bayesian interaction games $\{\mathrm{IBG}^N\}_{N=N^0}^{\infty}$, indexed by the size of the population. Since in a growing population a behavioral matrix will not be a well-defined object, we will shift our attention to Bayesian strategies, as defined in (5.2.5). Further, instead of an anonymous preference decomposition $\boldsymbol{N}$, we will look at the empirical preference distribution $\hat{\boldsymbol{m}} := \frac{\boldsymbol{N}}{N}$. In a sense, we consider a sequence of plays in the Bayesian interaction game $\mathrm{BG}^N = (\Gamma, \hat{\boldsymbol{m}})$ when the population grows gradually over time to infinity. To make such a process meaningful, we choose a constructive approach, building on the ideas mentioned in remark 5.2.1. We will define a *population growth process*, where in every "period" one new agent joins the population. An exact definition of a "period" is not important. What is however important is that the time scale of population growth does not interfere with the time scale at which the co-evolutionary process $\mathcal{M}^{\beta,N}$ reaches its stationary distribution. To ensure this, we will, in a rather informal way, assume that $\mathcal{M}^{\beta,N}$ is a relatively fast

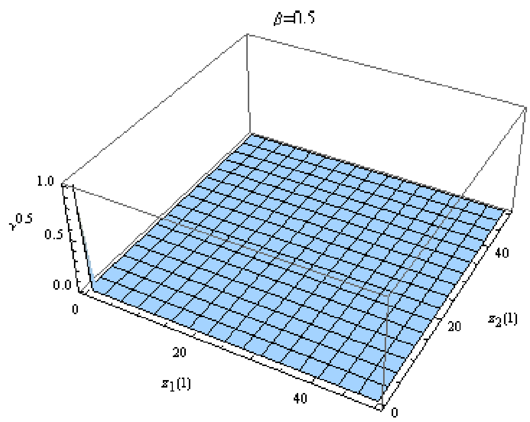

Figure 5.4: Invariant distribution over state matrices $\mathcal{Z}((50, 50))$ induced by the measure (5.7.1).

process, so that the population process is close to its long run prediction $\mu^{\beta,N}$ when population growth sets in.

Good population growth processes

Assume there exists a full-support probability measure $\boldsymbol{m} = (m_1, \ldots, m_K)$ on the type space Θ, according to which nature assigns each "period" τ the types to the players independently. The realized anonymous preference decomposition is $\boldsymbol{N}^\tau = (N_1^\tau, \ldots, N_k^\tau)$ and is a random vector following the multinomial distribution (5.2.4). The realized relative frequency of players of type $r = 1, 2, \ldots, K$ is $\hat{m}_r^\tau := N_r^\tau / N^\tau$. In between the time loop $\tau \to \tau + 1$, agents play the game $\bar{\text{IBG}}^{N^\tau} = (\Gamma, \boldsymbol{N}^\tau)$, where $\boldsymbol{N}^\tau$ is the realized anonymous preference decomposition from the independent period τ type-assignment process. Then $\mathcal{M}^{\beta,N^\tau}$ shapes the system for a sufficiently long time so that all population statistics are distributed according to the stationary measure μ^{β,N^τ}. Convergence to the stationary measure is guaranteed for positive noise levels. In particular, the distribution over behavioral matrices $\mathbf{Z} \in \mathcal{Z}(\boldsymbol{N}^\tau)$ is then governed by the temporary equilibrium measure $\gamma^{\beta,N^\tau}(\cdot|\boldsymbol{N}^\tau)$. Iterating this process over τ leads to a population growth process, where the time averages of the realized anonymous preference decompositions converges to the target distribution $\boldsymbol{m}$.[17] To fix this process, consider the following definition.

[17] See remark 5.2.1. In fact this is a simple consequence of the strong law of large numbers, due to the independence in the type assignment mechanism of nature.

Definition 5.7.1. *Let $\{\mathcal{I}^\tau\}_{\tau=0}^\infty$ be a population growth process with initial population $\mathcal{I}^0 = \{1, 2, \ldots, N\}$. Let $\boldsymbol{m}$ be a full support distribution on Θ. The population growth process is said to be good, if it satisfies the following criteria:*

Unit Growth: *For all $\tau \geq 1 : \mathcal{I}^\tau = \mathcal{I}^{\tau-1} \cup \{N + \tau\}$;*

Independent assignment: *At the τ-th step, nature assigns independently types to each player in $\mathcal{I}^\tau$ anew;*

Relaxation: *Before period $\tau + 1$ the co-evolutionary process $\mathcal{M}^{\beta,N^\tau}$ is close to its invariant distribution μ^{β,N^τ}.*

Calling $\hat{m}_r^\tau := N_r^\tau / N^\tau$ for all $\tau = 0, 1, \ldots,$ and $r = 1, 2, \ldots, K$, a good population growth process satisfies a law of large numbers in the sense that $\hat{m}_r^\tau \to m_r$ almost surely as $\tau \to \infty$.

In the following we will consider some good population growth process $\{\mathcal{I}^\tau\}_{\tau=0}^\infty$. In order to keep track of the temporal Bayesian strategies within each sub-population, define the matrices

$$
\begin{aligned}
\operatorname{diag}[\boldsymbol{N}^{-1}] &= \begin{bmatrix} 1/N_1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/N_K \end{bmatrix}, \\
\boldsymbol{\sigma}^N(\omega) = \mathbf{Z}(\omega) \operatorname{diag}[\boldsymbol{N}^{-1}] &= \begin{bmatrix} \sigma_1^N(1,\omega) & \sigma_2^N(1,\omega) & \ldots & \sigma_K^N(1,\omega) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1^N(n,\omega) & \sigma_2^N(n,\omega) & \ldots & \sigma_K^N(n,\omega) \end{bmatrix} \\
&= \left[\boldsymbol{\sigma}_1^N(\omega)^T, \ldots, \boldsymbol{\sigma}_K^N(\omega)^T\right].
\end{aligned}
$$

For every N and $\boldsymbol{N}$, $\boldsymbol{\sigma}^N$ is a column stochastic matrix, i.e. $\mathbf{1}^T \boldsymbol{\sigma}^N = \mathbf{1}^T$, living in the set $\mathcal{Z}(\boldsymbol{N}) \operatorname{diag}[\boldsymbol{N}^{-1}] = \prod_{r=1}^K \mathcal{X}_r^N(\boldsymbol{N}) \equiv \mathcal{X}^N(\boldsymbol{N})$. The invariant measure over behavior matrices (5.7.1) is, by definition, in a one-to-one correspondence with an invariant measure over finite population interim Bayesian strategy profiles (5.2.5). To see this, note that for any $\mathbf{Z} \in \mathcal{Z}(\boldsymbol{N})$, we have

$$
\begin{aligned}
\gamma^{\beta,N}(\mathbf{Z}|\boldsymbol{N}) &= \gamma^{\beta,N}(\boldsymbol{\sigma}^N \operatorname{diag}[\boldsymbol{N}]|\boldsymbol{N}) \\
&= \prod_{r=1}^K \frac{N_r!}{\prod_{s=1}^n (N_r \sigma_r^N(s))!} \prod_{s=1}^n \Phi_r^s(\boldsymbol{\sigma}^N \operatorname{diag}[\boldsymbol{N}], \beta, N)^{N_r \sigma_r^N(s)} \\
&=: \hat{\gamma}^{\beta,N}(\boldsymbol{\sigma}^N|\boldsymbol{N})
\end{aligned}
$$

with corresponding $\boldsymbol{\sigma}^N \in \mathcal{X}^N(\boldsymbol{N})$. Before presenting the main result of this section, we need to make sure that Bayesian strategies in fact converge in the large population limit. Therefore the following assumption is imposed.[18]

Assumption 5.7.1. *Let $\{\mathcal{I}^\tau\}_{\tau=0}^\infty$ be a good population growth process. The sequence of interim Bayesian strategy profiles $\boldsymbol{\sigma}^{N^\tau}$ converges component-wise to a limit Bayesian strategy $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ as $\tau \to \infty$, where $\boldsymbol{\Sigma}$ is the Cartesian product of K unit simplices of dimensions $n-1$.*

Next, let us write the interaction weights as

$$d_{rl}^{sv}(\beta, N) = \frac{\bar{d}_{rl}^{sv}(\beta)}{N}, \; \bar{d}_{rl}^{sv}(\beta) := \frac{2\exp(\pi(a_s, a_v)/\beta)}{\xi_{rl}^\beta}$$

and define the $n \times n$ symmetric matrix

$$\mathbf{D}_{rl}^\beta := \left[\bar{d}_{rl}^{sv}(\beta)\right]_{1 \le s,v \le n}.$$

Denote the standard inner product on $\mathbb{R}^n$ as $\langle \boldsymbol{a}, \mathbf{D}\boldsymbol{b}\rangle = \sum_{i,j} a_i d_{i,j} b_j$.

Theorem 5.7.1. *Let $\{N^\tau\}_{\tau=0}^\infty$ be a sequence of population sizes generated by a good population growth process. Under Assumption 5.7.1 we have*

$$\lim_{\tau\to\infty} \frac{1}{N^\tau} \log \hat{\gamma}^{\beta, N^\tau}(\boldsymbol{\sigma}^{N^\tau} | \boldsymbol{N}^\tau) = \sum_{r=1}^K m_r F_r(\boldsymbol{\sigma}, \beta), \tag{5.7.2}$$

where for all $1 \le r < l \le K$

$$F_r(\boldsymbol{\sigma}, \beta) := h(\boldsymbol{\sigma}_r) + \sum_{s=1}^n \frac{\theta_r(a_s)}{\beta}\sigma_r(s) + \sum_{l \ge r} m_l f_{rl}(\boldsymbol{\sigma}, \beta), \tag{5.7.3}$$

$$f_{rr}(\boldsymbol{\sigma}, \beta) := \frac{1}{2}\langle \boldsymbol{\sigma}_r, \mathbf{D}_{rr}^\beta \boldsymbol{\sigma}_r\rangle, \tag{5.7.4}$$

$$f_{rl}(\boldsymbol{\sigma}, \beta) := \langle \boldsymbol{\sigma}_r, \mathbf{D}_{rl}^\beta \boldsymbol{\sigma}_l\rangle, \tag{5.7.5}$$

$$h(\sigma_r) := -\sum_{s=1}^n \sigma_r(s) \log \sigma_r(s). \tag{5.7.6}$$

Proof. See Section 5.9. □

[18]We think that this is a mild assumption, and might even be redundant. However, we were not able to prove that Bayesian strategies do (almost) always converge. This remains for future research.

The proof uses the representation (5.7.1), and then we apply Stirling's formula repeatedly. It is a powerful result, telling us that the equilibrium distribution for strategy profiles puts almost all mass on the argmax of a single real-valued function $F(\boldsymbol{\sigma}, \beta) := \sum_{r=1}^{K} m_r F_r(\boldsymbol{\sigma}, \beta)$, for $N \to \infty$. The value of the functions $\{F_r\}_{r=1}^{K}$ depends on the complete Bayesian strategy profile $\boldsymbol{\sigma}$, as well as the interaction weights. The entropy function $h(\cdot)$ enters via the combinatorial terms, counting the number of ways the strategies can be distributed in the various preference groups, when this is done in a uniform random way.

5.8 Conclusion

We have presented a co-evolutionary model of networks and play where the agents' preferences are diverse. The main parameters of the model were the noise level and the size of the population. We call such class of games structured Bayesian interaction games, building on earlier work of Morris (1997; 2000) and Ely and Sandholm (2005), Sandholm (2007a). We gave a complete description of the long-run probabilistic law of the two dimensions of the model; the population profile of actions and the network of interaction. Further, we studied the behavior of the model by performing two different limit operations. First, we fixed the population size and asked the question how the invariant distribution behaves in the small noise limit. Such a limit emphasizes the importance of errors made by the individuals in their decisions. Low noise translates into a situation where agents make few errors, and the limiting outcome of such a process is conventionally called the stochastically stable set in the small noise limit. We gave sufficient conditions when this subset of states can be characterized by knowing the potential function alone, as if the model were formulated for a fixed interaction structure as in Blume (1993), Young (1998). Then, we fixed the noise level of the system at a positive value, and studied the behavior of the invariant measure as the population grows in magnitude. In this large population framework, individual errors are averaged out, and so we obtain a different way to formulate an equilibrium selection criterion.
There are many remaining questions waiting for an answer. In this paper we have just derived small noise or large population properties for the invariant measure over action Bayesian strategy profiles. We have not yet characterized the limiting behavior of the induced random graph ensemble. In particular, it would be interesting to explore the evolution of connected components in the society as one of the model parameters approaches its limit. A branching process approach would be helpful for this, and the literature on random

graphs (see e.g. Söderberg, 2002, Durrett, 2007, Bollobás et al., 2007) gives us a hint how to start. Moreover, the present model is very tractable simply because we made rather special assumptions. First, the game was assumed to be a potential game, a quite small class of games indeed. Second, the link creation process is not fully general due to the assumption on the players' rate functions (5.3.3). Future work should improve the model in these directions. The general framework outlined in chapter 2 of this thesis seems to be a promising first step.

5.9 Proofs of selected Lemmas, Propositions and Theorems

Proof of Lemma 5.6.2. We start with reformulating the interaction Hamiltonians as follows: for interactions within the same preference class $1 \leq r \leq K$, let us use the identity $e_{rr}(\omega) = \sum_{s=1}^{n} \sum_{v \geq s} e_{rr}^{sv}(\omega)$ for all $\omega \in \Omega^N$. Then we get

$$
\begin{aligned}
H_{rr}^{\beta,N}(\omega|\mathcal{T}) &= P_{rr}(\omega|\mathcal{T}) + \beta \log\left(\frac{2}{N\xi_{rr}^{\beta}}\right) e_{rr}(\omega) \\
&= \sum_{s=1}^{n} \sum_{i \in I_r(s,\omega)} \theta_r(a_s) + \beta \sum_{s=1, v \geq s}^{n} \sum_{i \in I_r(s,\omega)} \sum_{j \in I_r(v,\omega), j>i} \pi(a_s, a_v) g_{ij} \\
&+ \beta \log\left(\frac{2}{N\xi_{rr}^{\beta}}\right) e_{rr}(\omega) \\
&= \sum_{s=1}^{n} \theta_r(a_s)|I_r(s,\omega)| + \beta \sum_{s=1}^{n} \sum_{v \geq s} \left(\frac{\pi(a_s, a_v)}{\beta} + \log\left(\frac{2}{N\xi_{rr}^{\beta}}\right)\right) e_{rr}^{sv}(\omega) \\
&= \sum_{s=1}^{n} \theta_r(a_s)|I_r(s,\omega)| + \beta \sum_{s=1}^{n} \sum_{v \geq s} e_{rr}^{sv}(\omega) \log\left(d_{rr}^{sv}(\beta, N)\right)
\end{aligned}
$$

where we have defined the *interaction weight* $d_{rr}^{sv}(\beta, N) := \frac{2\exp(\pi(a_s,a_v)/\beta)}{N\xi_{rr}^{\beta}}$. Similarly, for interactions between players of type r and players of type l, we get

$$
\begin{aligned}
H_{rl}^{\beta,N}(\omega|\mathcal{T}) &= P_{rl}(\omega|\mathcal{T}) + \beta \log\left(\frac{2}{N\xi_{rl}^{\beta}}\right) e_{rl}(\omega) \\
&= \beta \sum_{s,v=1}^{n} \sum_{i\in I_r(s,\omega)} \sum_{j\in I_l(v,\omega)} \pi(a_s,a_v) g_{ij} + \beta \log\left(\frac{2}{N\xi_{rl}^{\beta}}\right) e_{rl}(\omega) \\
&= \beta \sum_{s,v=1}^{n} e_{rl}^{sv}(\omega) \log\left(d_{rl}^{sv}(\beta,N)\right).
\end{aligned}
$$

By eq. (5.4.4), the invariant measure takes the form

$$
\begin{aligned}
\rho^{\beta,N}(\omega|\mathcal{T}) &= \exp\left(\frac{H^{\beta,N}(\omega|\mathcal{T})}{\beta}\right) \\
&= \prod_{r=1}^{K}\prod_{l\geq r} \exp\left(\frac{H_{rl}^{\beta,N}(\omega|\mathcal{T})}{\beta}\right) \\
&= \prod_{r=1}^{K}\left[\prod_{s=1}^{n} \exp\left(\frac{\theta_r(a_s)|I_r(s,\omega)|}{\beta}\right) \prod_{v\geq s} d_{rr}^{sv}(\beta,N)^{e_{rr}^{sv}(\omega)}\right] \\
&\quad \times \left[\prod_{l>r}\prod_{s,v=1}^{n} d_{rl}^{sv}(\beta,N)^{e_{rl}^{sv}(\omega)}\right].
\end{aligned}
$$

Let us define for all $1\leq r\leq K,\ l>r$ and $\omega\in\Omega^N$ the following functions:

$$
\begin{aligned}
\phi_{rr}^{s}(\omega,\beta,N|\mathcal{T}) &:= \exp\left(\frac{\theta_r(a_s)|I_r(s,\omega)|}{\beta}\right) \prod_{v\geq s} d_{rr}^{sv}(\beta,N)^{e_{rr}^{sv}(\omega)}, \\
\phi_{rl}^{s}(\omega,\beta,N|\mathcal{T}) &:= \prod_{v=1}^{n} d_{rl}^{sv}(\beta,N)^{e_{rl}^{sv}(\omega)}, \\
\phi_{r}^{s}(\omega,\beta,N|\mathcal{T}) &:= \prod_{l\geq r} \phi_{rl}^{s}(\omega,\beta,N|\mathcal{T}),
\end{aligned}
$$

to arrive at the factorized representation

$$
\rho^{\beta,N}(\omega|\mathcal{T}) = \prod_{r=1}^{K}\prod_{s=1}^{n} \phi_{r}^{s}(\omega,\beta,N|\mathcal{T}).
$$

□

Proof of Proposition 5.6.1. On $\Omega^N(\mathbf{I})$ we know that the action profile is fixed, leaving only the network as free variable. Moreover, for a given subdivision $\mathbf{I}$, we can partition every graph g, corresponding to a population state $\omega \in \Omega^N(\mathbf{I})$, into many conditionally independent subgraphs. This high level of subgraph-decomposability makes the computation rather simple. Let $\omega \in \Omega^N(\mathbf{I})$, and define for all $1 \leq r \leq l \leq K$ and $1 \leq s, v \leq n$ the edge set

$$\mathcal{E}_{rl}^{sv}(\omega) := \mathcal{E}(I_r(s,\omega), I_l(v,\omega)).$$

For every $\omega \in \Omega^N(\mathbf{I})$, the cardinality of these sets is restricted to the range

$$e_{rl}^{sv}(\omega) := |\mathcal{E}_{rl}^{sv}(\omega)| \in \left\{0, 1, \dots, \frac{|I_r(s)|(|I_l(v)| - \delta_{r,l}\delta_{s,v})}{1 + \delta_{r,l}\delta_{s,v}}\right\}.$$

Since the edge set $\mathcal{E}_{rl}^{sv}(\omega)$ completely defines the subgraph describing the interactions between (r,s)-players and (l,v)-players, we see that $\mathcal{E}_{rl}(\omega)$ is the union over all these disjoint subgraphs. To be precise, we have for all $\omega \in \Omega^N(\mathbf{I})$ and $1 \leq r < l \leq K$, the following identities:

$$\mathcal{E}_{rr}(\omega) = \bigcup_{s=1}^{n} \bigcup_{v \geq s} \mathcal{E}_{rr}^{sv}(\omega),$$
$$\mathcal{E}_{rl}(\omega) = \bigcup_{s=1}^{n} \bigcup_{v=1}^{n} \mathcal{E}_{rl}^{sv}(\omega).$$

In order to aggregate over all states in $\Omega^N(\mathbf{I})$ we have to sum over all possible edge counts in these subgraphs. On $\Omega^N(\mathbf{I})$, the vertex set is $\mathcal{I} = \bigcup_{r=1}^{K} \bigcup_{s=1}^{n} I_r(s)$. On $\Omega^N(\mathbf{I})$ every $g \in \mathcal{G}[\mathcal{I}]$ is the union of its disjoint subgraphs

$$g = \left[\oplus_{r=1}^{K} \oplus_{v \geq s} g[I_r(s) \cup I_r(v)]\right] \oplus \left[\oplus_{l > r} \oplus_{s,v} g[I_r(s) \cup I_l(v)]\right].$$

It follows that

$$\rho^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T}) = \sum_{\omega \in \Omega^N(\mathbf{I})} \rho^{\beta,N}(\omega|\mathcal{T}) = \sum_{g \in \mathcal{G}[\mathcal{I}]} \rho^{\beta,N}((\alpha, g)|\mathcal{T}, \mathbf{I}).$$

In order to compute this sum, we can exploit the subgraph-decomposability of our model, and so sum over each subgraph separately. Start by taking the sum over edges $e_{11}^{11} \in \{0, 1, \dots, \frac{|I_1(1)|(|I_1(1)|-1)}{2}\}$. Call $E_{11}^{11}(\mathbf{I})$ the maximal element of this set. By the Factorization Lemma 5.6.2, we can partial out the factors affected by this summation as follows

$$\rho^{\beta,N}(\omega|\mathcal{T}, \mathbf{I}) = \exp\left(\frac{\theta_1(a_1)|I_1(1)|}{\beta}\right) d_{11}^{11}(\beta, N)^{e_{11}^{11}(\omega)} B.$$

B is a catch-all term, collecting all factors unaffected when summing over e_{11}^{11}. With this representation it is clear that we can now sum over all possible networks in the set $\mathcal{G}[I_1(1) \cup I_1(1)]$ without affecting the remainder term B. Further, observe that all graphs is this set with the same number of edges have the same weight. Hence, we can do the following summation exercise over the set of subgraphs $\mathcal{G}[I_1(1) \cup I_1(1)]$:

$$\begin{aligned}\sum_{g \in \mathcal{G}[I_1(1) \cup I_1(1)]} \rho^{\beta,N}((\alpha, g)|\mathcal{T}, \mathbf{I}) &= \exp\left(\frac{\theta_1(a_1)|I_1(1)|}{\beta}\right) B \sum_{k=0}^{E_{11}^{11}(\mathbf{I})} \binom{E_{11}^{11}(\mathbf{I})}{k} d_{11}^{11}(\beta, N)^k \\ &= \exp\left(\frac{\theta_1(a_1)|I_1(1)|}{\beta}\right) (1 + d_{11}^{11}(\beta, N))^{E_{11}^{11}} B \\ &= \rho_{(11),(11)}^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T}, \mathbf{I}).\end{aligned}$$

Store this expression and continue with aggregating over all graphs in the set $\mathcal{G}[I_1(1) \cup I_1(2)]$. In the same way as before, we arrive at

$$\begin{aligned}\rho_{(11),(12)}^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T}, \mathbf{I}) = \exp&\left(\frac{\theta_1(a_1)|I_1(1)|}{\beta}\right) \\ &\times (1 + d_{11}^{11}(\beta, N))^{E_{11}^{11}} (1 + d_{11}^{12}(\beta, N))^{E_{11}^{12}} B\end{aligned}$$

where B is again a catch-all term. Continuing in this fashion $(n-2)$ times, we arrive at the expression

$$\begin{aligned}\rho_{(11),(1n)}^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T}, \mathbf{I}) &= \exp\left(\frac{\theta_1(a_1)|I_1(1)|}{\beta}\right) \prod_{s=1}^{n} (1 + d_{11}^{1s}(\beta, N))^{E_{11}^{1s}(\mathbf{I})} B \\ &= \Phi_1^1(\mathbf{I}, \beta, N)^{|I_1(1)|} B.\end{aligned}$$

The next step is to sum over edges $e_{11}^{2v}, v \geq 2$, by following the same procedure. Iteration yields then the desired result. This summation operation terminates after $n \cdot K$ steps. $\square$

Proof of Theorem 5.6.1. For every $\omega \in \Omega^N$ we have

$$\mu^{\beta,N}(\omega|\mathcal{T}) = \frac{\rho^{\beta,N}(\omega|\mathcal{T})}{\sum_{\mathbf{I}} \rho^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T})},$$

where the sum extends over all partitions $\mathbf{I}$. From this it follows that

$$\mu^{\beta,N}(\omega|\mathcal{T}, \mathbf{I}) = \frac{\mu^{\beta,N}(\omega|\mathcal{T}) \mathbb{1}_{\Omega^N(\mathbf{I})}(\omega)}{\mu^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T})} = \frac{\rho^{\beta,N}(\omega|\mathcal{T}) \mathbb{1}_{\Omega^N(\mathbf{I})}(\omega)}{\rho^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T})},$$

and this extends to any subset $A \subseteq \Omega^N(\mathbf{I})$, by counting elements. Let

$$A := \{\omega \in \Omega^N(\mathbf{I}) | e^{sv}_{rl}(\omega) = \bar{e}^{sv}_{rl}\},$$

the subset of $\Omega^N(\mathbf{I})$ on which the number of connections between (r,l) and (s,v)-players equals the fixed number $\bar{e}^{sv}_{rl}$. By Proposition 5.6.1, we can compute the mass of this set as

$$\rho^{\beta,N}(A|\mathcal{T},\mathbf{I}) = d^{sv}_{rl}(\beta,N)^{\bar{e}^{sv}_{rl}} \frac{\rho^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T})}{(1+d^{sv}_{rl}(\beta,N))^{E^{sv}_{rl}(\mathbf{I})}}.$$

To see this is observe the following. $\rho^{\beta,N}(A|\mathcal{T},\mathbf{I})$ is the mass received by states in A. On A all interactions are unrestricted, except those between (r,l) and (s,v)-players. The mass of the set A is determined by summing over all possible subgraphs, conditional on $\mathbf{I}$, but holding the edge count of the subgraph $g[I_r(s) \cup I_l(v)]$ constant at $\bar{e}^{sv}_{rl}$. To get this mass from eq. (5.6.2) all one has to do is to replace the factor $(1+d^{sv}_{rl}(\beta,N))^{E^{sv}_{rl}(\mathbf{I})}$ with $d^{sv}_{rl}(\beta,N)^{\bar{e}^{sv}_{rl}}$. Hence

$$\begin{aligned}
\mu^{\beta,N}(A|\mathcal{T},\mathbf{I}) &= \frac{\rho^{\beta,N}(A|\mathcal{T})}{\rho^{\beta,N}(\Omega^N(\mathbf{I})|\mathcal{T})} \\
&= \frac{d^{sv}_{rl}(\beta,N)^{\bar{e}^{sv}_{rl}}}{(1+d^{sv}_{rl}(\beta,N))^{E^{sv}_{rl}(\mathbf{I})}} \qquad (\diamond)
\end{aligned}$$

Set $\frac{p^{sv}_{rl}(\beta,N)}{1-p^{sv}_{rl}(\beta,N)} = d^{sv}_{rl}(\beta,N)$, and solve this for $p^{sv}_{rl}(\beta,N)$. Substitute this into equation $(\diamond)$, by noting that $E^{sv}_{rl}(\mathbf{I}) = \frac{|I_r(s)|(|I_l(v)|-\delta_{rl}\delta_{vs})}{1+\delta_{rl}\delta_{vs}}$. We get

$$\mu^{\beta,N}(A|\mathcal{T},\mathbf{I}) = p^{sv}_{rl}(\beta,N)^{\bar{e}^{vs}_{rl}}(1-p^{vs}_{rl}(\beta,N))^{\frac{|I_r(s)|(|I_l(v)|-\delta_{rl}\delta_{vs})}{1+\delta_{rl}\delta_{vs}}-\bar{e}^{vs}_{rl}}.$$

It follows form this that if G^{sv}_{rl} is a graph in $\mathcal{G}[I_r(s) \cup I_l(s')]$ with $0 \leq k \leq E^{sv}_{rl}(\mathbf{I})$ edges, then its probability is

$$\mathbb{P}(\{G^{sv}_{rl}\}|\mathcal{T},\mathbf{I}) = p^{sv}_{rl}(\beta,N)^k(1-p^{vs}_{rl}(\beta,N))^{\frac{|I_r(s)|(|I_l(v)|-\delta_{rl}\delta_{vs})}{1+\delta_{rl}\delta_{vs}}-k}. \qquad (5.9.1)$$

This is the probability measure found by Erdös and Rényi (1960) to describe the probability space of a random graph where each edge exists with the probability $p^{sv}_{rl}(\beta,N)$. $\square$

Proof of Proposition 5.7.1. Start by writing

$$\begin{aligned}
\log\hat{\gamma}^{\beta,N^\tau}(\boldsymbol{\sigma}^{N^\tau}|\boldsymbol{N}^\tau) = \sum_{r=1}^{K}\Big\{&\log\left(\frac{N^\tau_r!}{\prod_{s=1}^n(N^\tau_r\sigma^{N^\tau}_r(s))!}\right) \\
&+ \sum_{s=1}^{n} N^\tau_r\sigma^{N^\tau}_r(s)\log\Phi^s_r(\boldsymbol{\sigma}^{N^\tau}\operatorname{diag}[\boldsymbol{N}^\tau],\beta,N^\tau)\Big\}.
\end{aligned}$$

We will first determine the limit as $\tau \to \infty$ for the combinatorial term by applying Stirling's Formula: $n! \cong \sqrt{2\pi n}(n/e)^n$ meaning $\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}(n/e)^n} = 1$. Under assumption 5.7.1 we get

$$\log\left(\frac{N_r^\tau!}{\prod_{s=1}^n (N_r^\tau \sigma_r^{N^\tau}(s))!}\right) = N_r^\tau \left[h(\boldsymbol{\sigma}_r^{N^\tau}) + o(1)\right]$$

For the second term in the measure, note that

$$\log \Phi_r^s = \log \Phi_{rr}^s + \sum_{l>r} \log \Phi_{rl}^s.$$

We have for all $r = 1, 2, \ldots, K$,

$$\log \Phi_{rr}^s(\boldsymbol{\sigma}^{N^\tau} \operatorname{diag}[\boldsymbol{N}^\tau], \beta, N^\tau) = \frac{\theta_r(a_s)}{\beta} + \sum_{v\geq s} \left(\frac{N_r^\tau \sigma_r^{N^\tau}(v) - \delta_{s,v}}{1+\delta_{s,v}}\right) \log\left(1 + d_{rr}^{sv}(\beta, N)\right),$$

and for all $l > r$

$$\log \Phi_{rl}^s(\boldsymbol{\sigma}^{N^\tau} \operatorname{diag}[\boldsymbol{N}^\tau], \beta, N^\tau) = \sum_{v=1}^n N_l^\tau \sigma_l^{N^\tau}(v) \log\left(1 + d_{rl}^{sv}(\beta, N)\right).$$

For fixed $\beta > 0$, we can write

$$\log\left(1 + d_{rl}^{sv}(\beta, N)\right) = \frac{1}{N}(\bar{d}_{rl}^{sv}(\beta) + o(1)).$$

Using this first-order approximation, we get for all $r = 1, 2, \ldots, K$ and $s = 1, 2, \ldots, n$

$$\begin{aligned}
&N_r^\tau \sigma_r^{N^\tau}(s) \log \Phi_{rr}^s(\boldsymbol{\sigma}^{N^\tau} \operatorname{diag}[\boldsymbol{N}^\tau], \beta, N^\tau) \\
&= N_r^\tau \Big[\frac{1}{\beta}\theta_r(a_s)\sigma_r^{N^\tau}(s) + \frac{\sigma_r^{N^\tau}(s)}{N} \sum_{v>s} N_r^\tau \sigma_r^{N^\tau}(v)(\bar{d}_{rr}^{sv}(\beta) + o(1)) \\
&+ \frac{\sigma_r^{N^\tau}(s)}{N}\left(\frac{N_r^\tau \sigma_r^{N^\tau}(s) - 1}{2}\right)(\bar{d}_{rr}^{ss} + o(1))\Big] \\
&= N_r^\tau \Big[\frac{1}{\beta}\theta_r(a_s)\sigma_r^{N^\tau}(s) + \hat{m}_r^\tau \frac{1}{2}(\sigma_r^{N^\tau}(s))^2 \bar{d}_{rr}^{ss}(\beta) + \hat{m}_r^\tau \sigma_r^{N^\tau} \sum_{v>s} \sigma_r^{N^\tau}(v) \bar{d}_{rr}^{sv}(\beta) + o(1)\Big]
\end{aligned}$$

It follows

$$\sum_{s=1}^n N_r^\tau \sigma_r^{N^\tau}(s) \log \Phi_{rr}^s(\boldsymbol{\sigma}^{N^\tau} \operatorname{diag}[\boldsymbol{N}^\tau], \beta, N^\tau) = N_r^\tau \left[\frac{1}{\beta}\sum_{s=1}^n \theta_r(s)\sigma_r^{N^\tau}(s) + \hat{m}_r^\tau f_{rr}(\boldsymbol{\sigma}^{N^\tau}, \beta) + o(1)\right]$$

Now, for $l > r$ we obtain analogously

$$\log \Phi_{rl}^{s}(\boldsymbol{\sigma}^{N^\tau} \operatorname{diag}[\boldsymbol{N}^\tau], \beta, N^\tau) = \sum_{v=1}^{n} \hat{m}_l^\tau \sigma_l^{N^\tau}(v)(\bar{d}_{rl}^{sv}(\beta) + o(1)),$$

so that

$$\sum_{s=1}^{n} N_r^\tau \sigma_r^{N^\tau}(s) \log \Phi_{rl}^{s}(\boldsymbol{\sigma}^{N^\tau} \operatorname{diag}[\boldsymbol{N}^\tau], \beta, N^\tau) = N_r^\tau \hat{m}_l^\tau (f_{rl}(\boldsymbol{\sigma}, \beta) + o(1)).$$

Putting the pieces together gives

$$\begin{aligned}
\log \hat{\gamma}^{\beta, N^\tau}(\boldsymbol{\sigma}^{N^\tau} | \boldsymbol{N}^\tau) &= \sum_{r=1}^{K} N_r^\tau \Big[h(\boldsymbol{\sigma}_r^{N^\tau}) + \frac{1}{\beta} \sum_{s=1}^{n} \theta_r(a_s) \sigma_r^{N^\tau}(s) + \sum_{l \geq r} \hat{m}_l^\tau f_{rl}(\boldsymbol{\sigma}^{N^\tau}, \beta) + o(1) \Big] \\
&= \sum_{r=1}^{K} N_r^\tau F_r(\boldsymbol{\sigma}^{N^\tau}, \beta)
\end{aligned}$$

Dividing by N^τ and letting $\tau \to \infty$ gives the desired result. $\square$

Chapter 6

Conclusion

In this dissertation we have discussed three models on the co-evolution of networks and play. Chapter 2 characterized the general mathematical framework for these kind of models, and we have shown that already with a minimal set of assumptions many interesting long-run characteristics of the system can be identified. In this article we encountered for first time the close connection between co-evolutionary models and the class of inhomogeneous random graphs. Subsequent chapters verified this finding in different scenarios, and it seems that this relationship is generic (in the sense of transition matrices satisfying the "axioms" of a co-evolutionary model of chapter 2. In general, the random graphs identified in this thesis are very close to classical Erdös-Rényi graphs. This results is promising and disappointing at the same time. It is promising, because it shows us that evolutionary games are a useful framework to study the dynamic evolution of networks and play in an integrated way. Further, a large class of evolutionary dynamics generates the same probabilistic ensemble of networks, but with different edge-success probabilities (recall Definition 1.2.1). It is disappointing, since the inhomogeneous random graph found here are, from the empirical point of view, only slightly better than an Erdös-Rényi graph.[1] The "problem" with this graph ensemble is that it misses any *local structure*, at least for sufficiently large population size. We see this already from the "toy model" presented in section 1.2. Compare Figures 1.3, 1.4 with the Erdös-Rényi graphs displayed in Figure 1.5, 1.6. At the first sight, it seems that our model possesses significant more structure than the Erdös-Rényi graphs, since we clearly observe a tendency to form cliques, "colored" by the action of the game. However, if we would consider only one of these components, the zoomed picture will essentially look as the networks of Figures 1.5 and 1.6, depending on the level of volatility. Empirically, even large massive networks display significant local structure, meaning significant clustering, short average distances, and center-periphery structures.[2] With the specific behavioral rules and simple games studied in chapters 4 and 5, we are not able to cope with these stylized facts. It is currently unknown if there exist behavioral rules which are able to generate "realistic" networks.

So why should one think that the results of this dissertation are useful? First, all the papers in this thesis are written in a mathematically rigorous way, standing in contrast to the bulk of (very interesting) results produced by physicists with the help of computer simulations and mean-field techniques. I have tried to convince the reader that a full study of the stochastic

[1] A nice survey on empirical observations on technological, biological and social networks can be found in Dorogovtsev and Mendes (2003) and Newman (2003).

[2] The first two properties are taken also as the definition of the famous small-world networks of Watts and Strogatz (1998).

processes is not completely out of reach, although it may be quite cumbersome. Second, I regard this thesis as a first step into a more general line of research, where tools of evolutionary game theory provide the mathematical language to "grow" random networks, evolving at a comparable time scale to the evolution of actions. This endows random graph theory with its necessary justification from behavioral principles. Therefore, I regard this co-evolutionary element as crucial and indispensable.

Second, even if the models presented in this thesis are not able to reproduce the stylized facts, inhomogeneous random graphs are still mathematically fascinating objects, and we have not yet studied their properties to a full extent. This remains to be done for future research. There is a rich literature which may be useful for this task, and all of them have been mentioned in the articles.

Third, a co-evolutionary model with noise, as defined in chapter 2, still allows for many possible specifications of behavioral rules. I firmly belief that one should be able to find behavioral rules, which may generate more intricate network topologies, having significant local clustering, or even scale-free degree distributions. The expected-degree model of Chung and Lu (see e.g. Chung and Lu, 2006), seems to be a promising starting point. It should also not be too difficult to introduce into the abstract behavioral rules of the general model in chapter 2 "local search protocols", such as the agents employ in Marsili et al. (2004). This modification will lead to a process which favors networks with significant clustering, yielding the desired local structure.

A real extension of the model could also be made by questioning the nature of connections in a network. The models in this thesis treated interaction as a 0-1-decision, i.e. a link is either on or off. Real world social networks show however different *intensities* in interactions. This calls for modeling the evolving networks as weighted graphs, and the weights of interaction are shaped by evolutionary forces. Weighted networks offer an even more refined picture of the overall interaction structure of the agents (see Schweitzer et al., 2009), but also make the mathematical analysis much more intricate. The problem is that, a weight can in principle be any real number, calling for the analysis of a Markov process with continuous state-spaces, and this requires much more mathematical sophistication than the models treated here.

All in all these are just some suggestions and ideas, and certainly many more avenues are open for future research in this fascinating and young area of science.

Bibliography

Alós-Ferrer, C., 1999. Dynamical systems with a continuum of randomly matched agents. Jounal of Economic Theory 86, Issue 2, 245–267.

Alós-Ferrer, C., Netzer, N., 2007. The logit-response dynamics, 2007, Forthcoming: Games and Economic Behavior.

Alós-Ferrer, C., Weidenholzer, S., 2008. Contagion and efficiency. Jounal of Economic Theory 143, Issue 1, 251–274.

Anderson, C., Wasserman, S., Crouch, B., 1999. A p^* primer: logit models for social networks. Social Networks 21, 37–66.

Anderson, S., de Palma, A., Thisse, J., 1992. Discrete Choice theory of Product differentiation. The MIT Press.

Bala, V., Goyal, S., 2000. A non-cooperative model of network formation. Econometrica 68, no.5, 1181–1229.

Barabási, A., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.

Beggs, A., 2005. Waiting times and equilibrium selection. Economic Theory 25, 599–628.

Benaïm, M., Weibull, J. W., 2003. Deterministic approximation of stochastic evolution in games. Econometrica 71 3, 873–903.

Bergin, J., Lipman, B., 1996. Evolution with state-dependent mutations. Econometrica 64, 943–956.

Binmore, K., Samuelson, L., 1997. Muddling through: Noisy equilibrium selection. Jounal of Economic Theory 74, 235–265.

Binmore, K., Samuelson, L., 1999. Evolutionary drift and equilibrium selection. Review of Economic Studies 66, 363–393.

Binmore, K., Samuelson, L., Vaughn, R., 1995. Musical chairs: Modeling noisy evolution. Games and Economic Behavior 11, 1–35.

Blume, L. E., 1993. Statistical mechanics of strategic interaction. Games and Economic Behavior 5, 387–424.

Blume, L. E., 1995. Evolutionary equilibrium with forward-looking players, mimeo, Cornell University.

Blume, L. E., 1997. Population games. In: Arthur, W., Durlauf, S., Lane, D. (Eds.), The Economy as an evolving complex system II. SFI Studies in the science of complexity, Vol. XXVII, Addison-Wesley, pp. 425–462.

Blume, L. E., 2003. How noise matters. Games and Economic Behavior 44, 251–271.

Bollobás, B., 1998. Modern Graph Theory. Springer: New York, Graduate Texts in Mathematics 184.

Bollobás, B., 2001. Random Graphs. Cambridge University Press, 2nd edition.

Bollobás, B., Janson, S., Riordan, O., 2007. The phase transition in inhomogeneous random graphs. Random Structures and Algorithms vol. 31, Issue 1, 3–122.

Boylan, R. T., 1992. Laws of large numbers for dynamical systems with randomly matched individuals. Journal of Economic Theory 57, 473–504.

Brémaud, P., 1998. Markov Chains - Gibbs Fields, Monte Carlo Simulation and Queues. Springer-Verlag.

Catoni, O., 1999. Simulated annealing algorithms and Markov chains with rare transitions. Séminaire de Probabilités XXXIII. Lecture Notes in Mathematics, Vol. 1079.

Catoni, O., 2001. Statistical Learning Theory and Stochastic Optimization. Lecture Notes in Mathematics, Vol. 1851.

Chung, F. R. K., Lu, L., 2006. Complex graphs and networks. CBMS regional Conference Series in Mathematics.

Currarini, S., Jackson, M. O., Pin, P., 2008. An economic model of friendship: Homophily, minorities and segregation, forthcoming: Econometrica.

Dorogovtsev, S. N., Mendes, J. F. F., 2003. Evolution of Networks - From Biological Nets to the Internet and WWW. Oxford University Press.

Durrett, R., 2007. Random Graph Dynamics. Cambridge Series in Statistical and Probabilistic Mathematics.

Ehrhardt, G., Marsili, M., Vega-Redondo, F., 2006a. Diffusion and growth in an evolving network. International Journal of Game Theory 34, 383–397.

Ehrhardt, G., Marsili, M., Vega-Redondo, F., 2006b. Phenomenological models of socio-economic network dynamics. Physical Review E 74, 036106.

Ehrhardt, G., Marsili, M., Vega-Redondo, F., 2008a. Emergence and resilience of social networks: a general theoretical framework. Annales d' économie et de statistique 86, 1–13.

Ehrhardt, G., Marsili, M., Vega-Redondo, F., 2008b. Networks emerging in a volatile world. EUI Working papers.

Ellison, G., 1993. Learning, local interaction, and coordination. Econometrica 61, no. 5, 1047–1071.

Ellison, G., 2000. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. Review of Economic Studies 67, 17–45.

Ely, J., Sandholm, W. H., 2005. Evolution in bayesian games I: Theory. Games and Economic Behavior 53, 83–109.

Erdös, P., Rényi, A., 1960. On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences 5, 17–61.

Foster, D., Young, H. P., 1990. Stochastic evolutionary game dynamics. Theoretical Population Biology 38, 219–232.

Freidlin, M., Wentzell, A., 1998. Random Perturbations of Dynamical Systems, 2nd Ed. Springer Verlag, New York.

Fudenberg, D., Levine, D. K., 1998. The theory of learning in games. The MIT Press, Cambridge Massachusetts.

Fudenberg, D., Tirole, J., 1991. Game Theory. The MIT Press, Cambridge Massachusetts.

Gilboa, I., Matsui, A., 1991. Social stability and social equilibrium. Econometrica 59 (3), 859–867.

Goyal, S., Vega-Redondo, F., 2005. Network formation and social coordination. Games and Economic Behavior 50, 178–207.

Grimmett, G., Stirzaker, D., 2001. Probability and Random Processes. Oxford University Press, USA, 3rd edition.

Harsanyi, J. C., 1973. Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. International Journal of Game Theory 2, 1–23.

Hofbauer, J., 1999. The spatially dominant equilibrium of a game. Annals of Operations Research 89, 233–251.

Hofbauer, J., Sandholm, W. H., 2002. On the global convergence of stochastic fictious play. Econometrica 70 (6), 2265–2294.

Hofbauer, J., Sandholm, W. H., 2007. Evolution in games with randomly disturbed payoffs. Journal of Economic Theory 132, 47–69.

Hofbauer, J., Sigmund, K., 1998. Evolutionary Games and Population Dynamics. Cambridge University Press.

Hojman, D. A., Szeidl, A., 2006. Endogenous networks, social games, and evolution. Games and Economic Behavior 55, Issue 1, 112–130.

Jackson, M. O., 2008. Social and Economic Networks. Princeton University Press.

Jackson, M. O., Watts, A., 2002. On the formation of interaction networks in social coordination games. Games and Economic Behavior 41, 265–291.

Jackson, M. O., Wolinsky, A., 1996. A strategic model of social and economic networks. Journal of Economic Theory 71, 44–74.

Jackson, M. O., Yariv, L., 2008. Diffusion, strategic interaction, and social structure. In: Benhabib, J., Bisin, A., Jackson, M. O. (Eds.), Handbook of Social Economics. Elsevier.

Kandori, M., Mailath, G., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61, 29–56.

Katz, M. L., Shapiro, C., 1994. Systems competition and network effects. The Journal of Economic Perspectives 8, no.2, 93–115.

López-Pintado, D., 2008. Diffusion in complex social networks. Games and Economic Behavior 62, 573–590.

Marsili, M., Vega-Redondo, F., Slanina, F., 2004. The rise and fall of the networked society: A formal model. PNAS 101, 1439–1442.

McKelvey, R., Palfrey, T., 1995. Quantal response equilibria for normal form games. Games and Economic Behavior 10, 6–38.

Monderer, D., Shapley, L., 1996. Potential games. Games and Economic Behavior 14, 124–143.

Morris, S., 1997. Interaction games: A unified analysis of incomplete information, local interaction and random matching. CARESS Working Paper 97-02.

Morris, S., 2000. Contagion. Review of Economic Studies 67, 57–78.

Morris, S., Ui, T., 2005. Generalized potentials and robust sets of equilibria. Journal of Economic Theory 124, 45–78.

Newman, M. E. J., 2002. Random graphs as models of networks. arXiv:cond-mat/0202208v1 [cond-mat.stat-mech].

Newman, M. E. J., 2003. The structure and function of complex networks. SIAM Review 45, no.2, 167–256.

Norris, J., 1997. Markov Chains. Cambridge Series in Statistical and Probabilistic Mathematics.

Park, J., Newman, M. E. J., 2004. The statistical mechanics of networks. Physical Review E 70, 066117–1–15.

Pastor-Satorras, R., Vespignani, A., 2001. Epidemic spreading in scale-free networks. Physical Review E 86, 14.

Powell, W. W., White, D. R., Koput, K. W., Owen-Smith, J., 2005. Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. American Journal of Sociology 110, no. 4, 1132–1205.

Rosenthal, R. W., 1973. On a class of games possessing pure-strategy nash equilibria. International Journal of Game Theory 2, 65–67.

Samuelson, L., 1997. Evolutionary Games and Equilibrium Selection. MIT Press, Cambridge Massachusetts.

Sandholm, W. H., 2007a. Evolution in bayesian games II: Stability of purified equilibria. Journal of Economic Theory 136, 641–667.

Sandholm, W. H., 2007b. Pigouvian pricing and stochastic evolutionary implementation. Journal of Economic Theory 132, 367–382.

Sandholm, W. H., 2009a. Decompositions and potentials for normal form games, mimeo, University of Wisconsin.

Sandholm, W. H., 2009b. Evolutionary dynamics and Population games. Forthcoming: The MIT Press.

Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., White, D. R., 2009. Economic networks: The new challenges. Science 325, 422–425.

Snijders, T. A. B., 2001. The statistical evaluation of social network dynamics. Sociological Methodology 31, 361–395.

Snijders, T. A. B., Pattison, P., Robins, G. L., Handcock, M. S., 2006. New specifications for exponential random graph models. Sociological Methodology 36, No. 1, 99–153.

Söderberg, B., 2002. General formalism for inhomogeneous random graphs. Physical Review E 66, 066121– 1– 6.

Strauss, D., 1986. On a general class of models for interaction. SIAM Review 28, no.4, 513–527.

Strauss, D., Frank, O., 1986. Markov graphs. Journal of the American Statistical Association 81, no. 395, 832–842.

Stroock, D., 2005. An Introduction to Markov Processes. Graduate Texts in Mathematics 230, Springer Berlin Heidelberg New York.

Ui, T., 2000. A shapley value representation of potential games. Games and Economic Behavior 31, 121–135.

van Damme, E., Weibull, J. W., 2002. Evolution in games with endogenous mistake probabilities. Jounal of Economic Theory 106, 296–315.

Vega-Redondo, F., 2006. Building up social capital in a changing world. Journal of Economic dynamics and control 30, 2305–2338.

Vega-Redondo, F., 2007. Complex Social Networks. Cambridge University Press, Econometric Society Monograph series.

Wasserman, S., Pattison, P., 1996. Logit models and logistic regressions for models of social networks: An introduction to p^*-models of social networks. Psychometrika 61, no.3, 401–425.

Watts, D. J., Strogatz, S. H., 1998. Collective dynamics of 'small world' networks. Nature 393, 440–442.

Weibull, J., 1995. Evolutionary Game Theory. The MIT Press Cambridge, Massachusetts.

Wilensky, U., 1999. NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Young, H., 1993. The evolution of conventions. Econometrica 61 (1), 57–84.

Young, H., 1998. Individual Strategy and Social Structure - An Evolutionary Theory of Institutions. Princeton University Press.

Young, H. P., 2003. The diffusion of innovations in social networks. In: Blume, L. E., Durlauf, S. N. (Eds.), The Economy as an Evolving Complex System III. Oxford University Press.

Abstract

This dissertation presents three interrelated papers on the co-evolution of networks and play. The general structure of these models combines three elementary events - action adjustment, link creation and link destruction - to one stochastic population dynamics, and focuses on the asymptotic properties of these processes.

Chapter 2 presents the general mathematical framework of a co-evolutionary model. The players are allowed to have arbitrary utility functions, defined on a common set of actions, and employ probabilistic behavioral rules in the above mentioned events. The class of admissible rules is characterized by irreducibility and a large deviations assumption. Beside these technical assumptions, not many behavioral assumptions are imposed. This generates a well-defined Markov chain, whose long-run properties can be studied analytically by making use of tree-characterization methods due to Freidlin and Wentzell (1998). We show how stochastically stable states can be detected in such co-evolutionary models, by defining suitable cost-functions. Then, under some further mild assumptions on the structure of the behavioral rules, we can prove an interesting connection between the derived ensemble of networks and inhomogeneous random graphs.

The models presented in chapters 4 and 5 particularize the general framework to the class of potential games and behavioral rules of the logit form. Under these assumptions, chapter 4 gives a full description of the induced ensemble of networks, and provides closed-form expressions for some statistics of this ensemble, such as the degree distribution. This article also presents a general concentration result of the long-run distribution of the co-evolutionary process on the set of potential maximizers, thereby generalizing earlier results from evolutionary game theory. The model presented in chapter 5 is more general by allowing the players to have idiosyncratic preferences. We introduce a new class of games, called structured Bayesian interaction games. This class of games combines ideas from so-called interaction games (Morris, 1997), and a very young evolutionary theory on Bayesian population games (Ely and Sandholm, 2005, Sandholm, 2007a).

Chapter 3 establishes a connection between the Markov chain constructed in the general framework of chapter 2 and the continuous-time Markov processes considered in the models of chapters 4 and 5. Chapter 6 closes the thesis with some thoughts and suggestions for future research.

Zusammenfassung

Diese Dissertation besteht aus drei inhaltlich zusammenhängenden Artikeln zu ko-evolutionären Dynamiken von Netzwerken und Strategien. Ein ko-evolutionärer Prozess besteht aus drei elementaren Ereignissen - Revision von Aktionen, Kreation eines links, Zerstörung eines links- welche zusammengefasst eine aggregierte Populationsdynamik definieren. Unser Augenmerk liegt in der Charakterisierung des langfristigen Verhaltens derartiger Prozesse.
Kapitel 2 beschreibt den abstrakten mathematischen Rahmen eines ko-evolutionären Modells. Die Spieler sind charakterisiert durch eine Nutzenfunktion auf einem gemeinsamen Raum von Aktionen und verwenden probabilistische Verhaltensregeln in den eingangs genannten elementaren Ereignissen. Zulässige Verhaltensregeln erfüllen eine Irreduzibilitätsannahme sowie ein Prinzip großer Abweichungen. Neben diesen technischen Annahmen werden strukturelle Annahmen an das Verhalten der Spieler auf ein Minimum reduziert. Dies generiert eine wohl definierte Markov-Kette, dessen langfristiges Verhalten durch Graphen-theoretische Methoden nach Freidlin und Wentzell (1998) studiert werden kann. Wir beschreiben eine allgemeine Methode mit der stochastisch-stabile Zustände identifiziert werden können. Unter weiteren schwachen Annahmen ist das induzierte Zufallsgraphenmodell vollständig charakterisierbar. Es zeigt sich eine bemerkenswerte Beziehung zwischen ko-evolutionären Modellen und dem Modell der inhomogenen Zufallsgraphen.
In Kapitel 4 und 5 werden diese Resultate verwendet um partikulre Modelle zu analysieren. Der Artikel "Potential games played in volatile environments" diskutiert ein Ko-evolutorisches Modell in der Klasse von Potentialspielen und Logit-Verhaltensregeln. Die invariante Verteilung und das generierte Zufallsgraphenmodell sind vollständig bestimmbar, und wir präsentieren einige Statistiken des Zufallsgraphenmodells in geschlossener Form, wie etwa die "degree distribution". Des Weiteren beweisen wir ein Konzentrationsresultat der invarianten Verteilung auf der Menge der Maxima der Potentialfunktion. Dies ist eine Verallgemeinerung wohl bekannte Resultate der Evolutionren Spieltheorie. Kapitel 5 erweitert das Modell von Kapitel 4 durch Heterogenität in den Präferenzen der Spieler. Wir definieren eine neue Klasse von Spielen, genannt "structured Bayesian interaction games". Diese Klasse von Spielen vereint Ideen von "interaction games" (Morris, 1997) und einer jungen evolutionären Literatur über Bayesianische Populationsspiele (Ely und Sandhom 2005, Sandholm, 2007a).
Kapitel 3 stellt eine Verbindung zwischen den Markov-Ketten von Kapitel 2, und den Markov-Prozessen der Kapitel 4 und 5 her. In Kapitel 6 diskutieren wir mögliche Erweiterungen der beschriebenen Modelle und skizzieren zuknftige Forschungsvorhaben.

Mathias Staudigl

University of Vienna
Department of Economics
A-1010 Wien
Hohenstaufengasse 9
Date of Birth: September 22, 1983
Citizenship: Austrian

Phone: (+43) 1-4277-37454
Email: mathias.staudigl@unvie.ac.at
Homepage: http://homepage.univie.ac.at/mathias.staudigl

Education

PhD Economics, University of Vienna, defense in February 2010.

Dissertation: Co-evolutionary dynamics of networks and play
Supervisors: Immanuel Bomze, Manfred Nermuth
External supervision: Fernando Vega-Redondo
Job-market paper: Potential games played in volatile environments

MSc (Mag.rer.soc.oec) Economics with honors, University of Vienna, 2007.

Thesis: Evolutionary dynamics and rationality
Supervisor: Ana-B. Ania Martinez

Fields of Interest

Game theory, Evolutionary game theory, Complex Networks, Random graph Dynamics, Stochastic processes

Teaching Experience

Teaching Assistant for Professor Immanuel Bomze, Linear Algebra for students in economics and statistics, Winter semester 2009/2010.

Teaching Assistant for Professor Immanuel Bomze, Analysis for students in economics and statistics, Summer semester 2009.

Tutor for Professor Konrad Podczeck, Advanced Microeconomics and Mathematical economics

Languages

German: Native speaker

English: Excellent written and spoken proficiency

Spanish: Very good written and spoken proficiency

French: Basic knowledge

Working Papers

Potential games played in volatile environments
(Submitted and currently under revision.)

Abstract: This papers studies the co-evolution of networks and play in the context of finite population potential games. Action revision, link creation and link destruction are combined in a continuous-time Markov process. We derive the unique invariant distribution of this process in closed form, as well as the marginal distribution over action profiles and the conditional distribution over networks. It is shown that the equilibrium interaction topology is a so-called inhomogeneous random graph, as defined in Bollobás et al. 2007 [Random Structures and Algorithms 31, 2007]. Furthermore, we are able to characterize the set of stochastically stable states by demonstrating a rather general concentration result of the invariant measure on the set of potential maximizers. This generalizes the well-known results due to Blume [Games and Economic Behavior 5, 1993] to models with endogenous interaction structures.

Co-Evolutionary dynamics and Bayesian interaction games

Abstract: We present a model on the co-evolution of networks and play for settings where agents' preferences are diverse. This leads to a definition of structured Bayesian interaction games, which is an adaption of Bayesian population games, introduced by Ely and Sandholm [Games and Economic Behavior 53, 2005], to the case of finite populations with general interaction architecture. We prove that under the logit dynamics of evolutionary game theory the co-evolutionary process is ergodic and we calculate its invariant distribution explicitly. We also derive a marginal distribution over Bayesian action profiles, as well as a probability measure over networks. We are able to characterize the long-run probabilistic ensemble of random graphs and show that it is an inhomogeneous random graph. We further perform two different limit operations to examine stochastically stable states in the small noise limit and in the large population limit, as suggested by Sandholm [Forthcoming: MIT Press, 2009].

On a general class of co-evolutionary dynamics

Abstract: This paper presents a unified framework to study the co-evolution of networks and behavior using the language of evolutionary game theory. The set-up is rich enough to encompass many recent models discussed in the literature. We completely characterize the invariant distribution of such processes and show how to calculate stochastically stable states by means of a tree-characterization algorithm. Moreover, specializing the process a bit further allows us to completely characterize the generated random graph ensemble. We demonstrate a rather deep and unexpected relationship between inhomogeneous random graphs and evolutionary game dynamics. This gives a new definition of an inhomogeneous random graph model, based on the behavioral rules of the players and the underlying game theoretic situation.

Conference Presentations

Stochastik Kolloquium (Department of Statistics and Decision support systems). Organized by Immanuel Bomze

Mini-Workshop Evolutionary Game Theory (Department of Mathematics, University of Vienna). Organized by Josef Hofbauer.

Doctoral Workshop on Game theory, University of Konstanz (Germany). Organized by Carlos Alós-Ferrer.

Economic theory seminar, University of Pavia (Italy). Invited by Stefano DeMichelis.

QED conference 2009 in Amsterdam.

Summer meeting of the Econometric Society 2009 in Barcelona.

Honors, & Awards

2005 Prize for Academic Excellence (Leistungsstipendium), University of Vienna

2006 Scholarship "ERASMUS". I had the pleasure to visit the University of Alicante, Spain, for the winter term.

2008 *Würdigungspreis des Bundesministeriums für Bildung, Wissenschaft und Kultur Österreich*; prize awarded by the Austrian federal ministry of education and culture to the 40 best undergraduate students in Austria.

References

Immanuel Bomze
Full Professor in Applied Mathematics and Statistics
University of Vienna
Department of Statistics and Decision support systems
`immanuel.bomze@univie.ac.at`

Manfred Nermuth
Full Professor of Economics
University of Vienna
Department of Economics
`manfred.nermuth@univie.ac.at`

Gerhard Sorger
Full Professor of Economics
University of Vienna
Department of Economics
`gerhard.sorger@univie.ac.at`

Last updated: December 17, 2009