Chapter 1
Evolutionary game of coalition building under external pressure

Alekos Cecchin and Vassili N. Kolokoltsov

Abstract We study the fragmentation-coagulation, or merging and splitting, model as introduced in [16], where N small players can form coalitions to resist to the pressure exerted by the principal. It is a Markov chain in continuous time and the players have a common reward to optimize. We study the behavior as N grows and show that the problem converges to a (one player) deterministic optimization problem in continuous time, in the infinite dimensional state space ℓ^1. We apply the method developed in [8], adapting it to our different framework. We use tools involving dynamics in ℓ^1, generators of Markov processes, martingale problems and coupling of Markov chains.

Key words: Fragmentation-coagulation, merging and splitting, evolutionary coalition formation, Markov decision process, major agent, mean field limit.

1.1 Introduction

In this paper we study dynamic optimization problems on Markov decision processes composed of a large number of interacting agents, in particular we investigate the so called fragmentation-coagulation, or merging and splitting, model. Our aim is to analyze its limit as the number of players tends to infinity.

Following Kolokoltsov [16] we describe a model which is a Markov chain in continuous time. A natural reaction of the society of small players to the pressure exerted by the principal can be executed by forming stable groups that can con-

Alekos Cecchin
Department of Mathematics, University of Padua, Via Trieste 63, Padova, Italy,
e-mail: acecchin@math.unipd.it

Vassili N.Kolokoltsov
Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK
e-mail: v.kolokoltsov@warwick.ac.uk

front this pressure in an effective manner (but possibly imposing certain obligatory regulations for the members of the group). Analysis of such possibility leads one naturally to models of mean field enhanced coagulation processes under external pressure. The major player can change her strategy only in discrete deterministic time.

Coagulation fragmentation processes are well studied in statistical physics, see e. g. [19]. In particular, general mass exchange processes, that in our social environment become general coalition forming processes preserving the total number of participants, were analyzed in [13] and [14] with their law of large number limits for discrete and general state spaces. In the same way problems in economics, like merging banks or firms on the market, were studied in [21] and [22]. While an application to scientific citation networks or the network of internet links is discussed in [17]. Some simple situation of nonlinear Markov games on a finite state space was analyzed in [15], proving the convergence of Nash-equilibria for finite games to equilibria of a limiting deterministic differential game.

Very recently, several authors have studied games of coalition formation. A notion of core equilibrium in proposed in [9], and found via a fixed point method. An application to contracts and networks is analyzed in [10]. A study of the incentives offered by the government to municipalities to merge into larger groups is provided in [24]. Players preferences over winning coalitions are derived by applying strongly monotonic power indices on the game in [12], where the author also investigate whether there are core stable coalitions . An application of systems of coalition formation to the climate change problem is discussed in [7], where also numerical simulation are performed.

Here we are interested in the response of such systems to external parameters that may be set by the principal who has her own agenda. Thus we add to the analysis a major player fitting the model to a more general framework. There are two main difficulties in studying this model. Firstly the total number of coalitions is not constant in time, as they can merge or split. Secondly the dynamics both of the system of small players and of the limiting system are supposed to lie on the infinite dimensional space ℓ^1, which can be viewed also as a space of measures, instead of a fixed $\mathbb{R}^d$. In fact the dimension of the state space for the system of coalitions grows as the number N of small players tends to infinity. If the system is in the state $x \in \ell^1$ then $x_k = hn_k$ where n_k is the number of coalitions of size k and h is a suitable parameter depending on N, for instance the inverse of the initial number of coalitions.

Our main result is to show that this problem converges, as N grows, to a one player deterministic optimization problem in continuous time, in the infinite dimensional state space ℓ^1, the so called mean field limit. We prove convergence of the value functions and provide also an approximated optimal policy for the system of small players. Such optimal policy is usually found by using dynamic programming for the finite horizon case, but this approach suffer from the curse of dimensionality, which makes them impractical when the state space is too large. Solving the HJB equation for the limiting system numerically is sometimes rather easy. It provides a deterministic optimal policy whose reward is remarkably close to the optimal reward.

We apply the method developed by Gast, Gaujal and Le Boudec in [8], where the authors obtained the same kind of results , but in a different setting. They consider discrete time Markov chains as prelimit systems whose state space is finite and fixed. Their proofs are in line with classic mean field arguments and use stochastic approximation techniques. Moreover their approach is algorithmic: they construct two intermediate systems: one with a finite number of objects controlled by a limit policy and one with a limit system controlled by a stochastic policy induced by the finite system.

Several papers in the literature are concerned with the problem of mixing the limiting behavior of a large number of objects with optimization. In [5], the value function of the Markov decision process (MDP) is approximated by a linearly parametrized class of functions and a fluid approximation of the MDP is used. It is shown that a solution of the HJB equation is a value function for a modification of the original MDP problem. In [23], the curse of dimensionality of dynamic programming is circumvented by approximating the value function by linear regression. In [8] they use instead a mean field limit approximation and prove asymptotic optimality in N of limit policy. Actually, most of the papers dealing with mean field limits of optimization problems over large systems are set in a game theory framework, leading to the concept of mean field games, introduced by Lasry and Lions [18] and P.E. Caines, M. Huang and R.P. Malhamé [2].

Notice finally that in this paper we analyze only a preliminary step for a full game setting with major and minor players, namely the response of the minor players to the action of the major one. The full analysis (not developed here) would include the reaction of the major on the behavior of the minor players and the search for the corresponding equilibrium. However, this development does not seem to present serious difficulty, since our analysis reduces it effectively to a two-player game: the major and the pool of small players.

Contribution and structure of the paper

In [16] Kolokoltsov shows the convergence of the optimization problems related to the system of small players to an optimization problem in discrete time for the limiting system (this will be similar to theorem 3). His proof is based on an argument that is focused on the generators of the Markov chains and shows their convergence. In this paper we want to show the convergence to an optimization problem in continuous time, so we apply the ideas from [8] where they used a completely different argument for the proof, focusing on trajectories and constructing two auxiliary systems.

In section 2 we describe properly the fragmentation coagulation model starting from [16], the limiting system and the related optimization problems. We define the state space where all the dynamics considered lie, which is a compact set $S \subset \ell^1$, and in the end we state the assumptions we need to obtain the convergence. In section 3 we present our main results and define the two auxiliary systems. Then we show

how to construct an approximated optimal policy starting from an optimal action function for the mean field limit. Moreover we consider a class of applications in which a particular choice of the rate functions allows to reduce the limiting problem to an optimization problem in one dimension, providing an explicit solution in a simplified case and a more effective numerical scheme in general. Finally in section 4 we complete the proofs, showing that the general requirements for convergence expressed in [8] can fit to our model, with some modification. We use theorems about semigroups and generators of Markov processes and related martingale problems. Moreover we apply the notion of coupling of Markov chains and also a particular Markovian coupling.

1.2 Model and assumptions

1.2.1 The space $B_+(L,R]$

We denote, as usual, the space of measures

$$\ell^1(\mathbb{N}) := \left\{ x = (x_1, x_2, \dots) : x_k \in \mathbb{R}, \quad \sum_k |x_k| < \infty \right\}. \tag{1.1}$$

Denote by ℓ^1_+ the space of positive measures on $\mathbb{N}$: $\ell^1_+(\mathbb{N}) := \{x \in \ell^1 : x_k \geq 0\}$. The usual norm in ℓ^1 is $||x||_{\ell^1} := \sum_k |x_k|$.

Let $L : \mathbb{N} \longrightarrow \mathbb{R}$ be the identity function, which means $L(k) = k$. We define a new norm $||x||_{\ell^1(L)} := \sum_k L(k)|x_k|$, so that we can consider the subset of ℓ^1

$$\ell^1(L) := \left\{ x \in \ell^1 : ||x||_{\ell^1(L)} < \infty \right\}$$

which is a Banach space equipped with this norm.

Let us denote by $B(L,R)$ the ball of radius R in $\ell^1(L)$, centered in 0, and $\ell^1_+(L) := \ell^1(L) \cap \ell^1_+$, $B_+(L,R) := B(L,R) \cap \ell^1_+$.

Lemma 1. *The set $B_+(L,R)$ is relatively compact in the norm topology of ℓ^1.*

Proof. By *Prohorov*'s compactness criterion a family of measures is relatively compact in the weak topology if and only if it is tight. We have $\sum_k kx_k < R$ for any $x \in B_+(L,R)$. Thus for any $n \in \mathbb{N}$ and any $x \in B_+(L,R)$

$$n \sum_{k \geq n} x_k \leq \sum_{k \geq n} kx_k < \sum_k kx_k < R$$

which gives $\sum_{k \geq n} x_k \leq \frac{R}{n}$ for any $n \in \mathbb{N}$ and any $x \in B_+(L,R)$. So for any $\varepsilon > 0$ there exists $n \in N$ such that

$$x(\mathbb{N} \setminus [0,n]) = \sum_{k \geq n} x_k < \varepsilon$$

for all $x \in B_+(L,R)$, which means that the tightness condition is satisfied.

By *Schur*'s theorem, any weakly convergent sequence in ℓ^1 is actually convergent in the norm of ℓ^1. Therefore the set $B_+(L,R)$ is relatively compact in the topology of ℓ^1.

We denote by $B_+(L,R]$ the closure of $B_+(L,R)$ in the norm topology of ℓ^1, which is compact in ℓ^1, although not in the topology of $\ell^1(L)$. The set

$$S := B_+(L,R]$$

will be the state space for the dynamics considered.

We will assume that the functions defined on S have some regularity. Let Z be a closed convex subset of a normed space Y and $f : Z \longrightarrow Y$ a function. Recall that the *directional derivative* $Df(x) : Y \longrightarrow Y$ of f in the point $x \in Z$ is a linear form that calculated in a vector $\xi \in Y$ is defined as $Df(x).\xi := \lim_{t\to 0} \frac{f(x+t\xi)-f(x)}{t}$. The second order derivative $D^2f(x)$ is a bilinear form defined as $D^2f(x).[\xi,\eta] := D(Df(x).\xi).\eta$. Thus the norms of the derivatives in Y are defined as norms of linear maps:

$$||Df(x)||_Y := \sup_{||\xi||=1} ||Df(x).\xi||_Y, \tag{1.2}$$

$$||D^2f(x)||_Y := \sup_{||\xi||=||\eta||=1} ||D^2f(x).[\xi,\eta]||_Y. \tag{1.3}$$

We say that $f \in \mathscr{C}^1(Z)$ if the function $(x,\xi) \mapsto Df(x).\xi$ is continuous from $Z \times Y$ to Y. Similarly, $f \in \mathscr{C}^2(Z)$ if the function $(x,\xi,\eta) \mapsto D^2f(x).[\xi,\eta]$ is continuous from $Z \times Y^2$ to Y. These are subsets of $\mathscr{C}(Z)$ and Banach spaces under the norms

$$||f||_{\mathscr{C}^1(Z)} := \sup_{x\in Z} \{||Df(x)||_Y + ||f(x)||_Y\} \tag{1.4}$$

$$||f||_{\mathscr{C}^2(Z)} := \sup_{x\in Z} \left\{||D^2f(x)||_Y + ||f(x)||_Y\right\}. \tag{1.5}$$

We will use these definitions for the sets $Z = S$, which is convex and compact, and Y to be either ℓ^1 or $\ell^1(L)$.

1.2.2 System of small players

We describe a so called *fragmentation-coagulation*, or *merging and splitting*, model in which there are N indistinguishable small players that form coalitions to resist to the pressure exerted by a major player, following [16].

The state space is

$$\mathbb{N}^{fin} := \{n = (n_1, n_2, \dots) : \text{there is only a finite number of non zero entries}\} \tag{1.6}$$

where $n_k \in \mathbb{N}$ denotes the total number of coalitions of size k, so the total number of small players is $N = \sum_k kn_k$ and the total number of coalitions is $\sum_k n_k$. The dynamics will be better described in the rescaled space

$$h\mathbb{N}^{fin} = \{x = hn = (x_1, x_2, \dots)\} \tag{1.7}$$

where h can be taken, for instance, as the inverse of the initial number of coalitions. We want to study the limit as $h \to 0$. All this $h\mathbb{N}^{fin}$ spaces, as h changes, can be viewed as subspaces of the space ℓ^1. The total number of players is conserved: this motivates the choice of $L(k) = k$ in the previous section, since $||x||_{\ell^1(L)} = hN$ for any state x; we will return to this in 2.5.1.

The dynamics evolves in continuous time as a Markov chain. It is described as follows:

- to any randomly chosen pair of coalitions of size i and j is attached a random exponential clock of parameter $hC_{ij}(x,b)$ so that they merge if it rings;
- to any randomly chosen coalition of size i is attached an exponential clock of parameter $F_{ij}(x,b)$ such that, if it rings, the coalition splits into two coalitions of size j and $i-j$.

Here the functions C and F may depend on the whole composition x and b is a control parameter which lies in a compact metric space (E,d).

The minimum of all these exponential random variables is an exponential random variable with the parameter

$$s = s(x,b) := \sum_{i,j} n_i n_j hC_{ij}(x,b) + \sum_i n_i F_{ij}(x,b). \tag{1.8}$$

When this minimum clock rings, the system goes from the state n to either $n - e_i - e_j + e_{i+j}$ (two coalitions merge) or $n - e_i + e_j + e_{i-j}$ (a coalition splits). The sequence $(e_i)_{i=1}^{\infty}$ denotes the standard basis in $\mathbb{R}^{\infty}$. The first case happens if the minimum holds for the clock of parameter $hC_{ij}(x,b)$, thus with probability given by $\frac{hC_{ij}(x,b)n_i n_j}{s(x,b)}$. While the second case happens with probability $\frac{F_{ij}(x,b)n_i}{s(x,b)}$.

Hence the infinitesimal generator of this Markov chain on the space $\mathbb{N}^{fin}$ is

$$\begin{aligned} \Lambda_{b,n} G(n) = s(x,b) \sum_{i,j} \frac{hC_{ij}(x,b)n_i n_j}{s(x,b)} \left[G(n - e_i - e_j + e_{i+j}) - G(n)\right] \\ + s(x,b) \sum_i \sum_{i<j} \frac{F_{ij}(x,b)n_i}{s(x,b)} \left[G(n - e_i + e_j + e_{i-j}) - G(n)\right]. \end{aligned} \tag{1.9}$$

Equation (1.9) can be equivalently presented as the infinitesimal generator

$$\Lambda_{b,h}G(x) = \frac{1}{h}\sum_{i,j} C_{ij}(x,b)x_i x_j \left[G(x - he_i - he_j + he_{i+j}) - G(x)\right] \quad (1.10)$$
$$+ \frac{1}{h}\sum_{i}\sum_{j<i} F_{ij}(x,b)x_i \left[G(x - he_i + he_j + he_{i-j}) - G(x)\right]$$

of the Markov chain describing the system of small players on the space $h\mathbb{N}^{fin} \subset \ell^1(\mathbb{N})$, for every $G \in \mathscr{C}(S)$.

Notation 1 *$X^h(t,x,b)$ is the state (in S) at time t of the Markov Chain given by this generator (1.10) which is in x at $t = 0$, under the control parameter b given by the major player.*

The process $X^h(t,x,b)$ describes the evolution of the coalitions of small players, which will be also called the system with N agents.

1.2.3 Limiting system

The limiting deterministic evolution, the *mean field limit*, is described by the so called *Smoluchovski equation*. For every x in the compact subset $S \subset \ell^1$ the ODE for the component i is

$$\dot{x}_i = f_i(x,b) := \sum_{j<i} C_{j,i-j}(x,b)x_j x_{i-j} - 2\sum_{j} C_{ij}(x,b)x_i x_j \quad (1.11)$$
$$+ 2\sum_{j>i} F_{ji}(x,b)x_j - \sum_{j<i} F_{ij}(x,b)x_i.$$

Notation 2 *$X(t,x,b)$ is the* flow *at time t of the ODE*

$$\dot{x} = f(x,b) \quad (1.12)$$

starting in x at $t = 0$ under the control parameter $b \in E$, where f is given by (1.11). In integral form

$$X(t,x,b) = x + \int_0^t f(X(s,x,b),b)ds. \quad (1.13)$$

We view the dynamics given by a deterministic ODE as a Markov process. The semigroup is

$$U_t G(x) = G(X(t,x)) \quad (1.14)$$

for every $G \in \mathscr{C}(S)$, and its generator is given by

$$\Lambda G(x) := \sum_i f_i(x)\frac{\partial G}{\partial x_i}(x), \quad (1.15)$$

for any $G \in \mathscr{C}^1(S)$. The first order partial differential operator defined in (1.15) has characteristics which solve equation (1.12).

So, for the limiting ODE given by (1.11), the corresponding infinitesimal generator given by (1.15) is

$$\Lambda_b G(x) = \sum_{i,j} C_{ij}(x,b) x_i x_j \left[\frac{\partial G}{\partial x_{i+j}} - \frac{\partial G}{\partial x_i} - \frac{\partial G}{\partial x_j} \right] + \sum_i \sum_{j<i} F_{ij}(x,b) x_i \left[\frac{\partial G}{\partial x_{i-j}} + \frac{\partial G}{\partial x_j} - \frac{\partial G}{\partial x_i} \right] \tag{1.16}$$

for any $G \in \mathscr{C}^1(S)$ and $b \in E$.

We can thus deduce that pointwise convergence of the generators holds. Namely, for the generators of the Markov chains defined by (1.10) and the generator of the deterministic limit defined by (1.16), we obtain

$$\lim_{h \to 0} \Lambda_{b,h} G(x) = \Lambda_b G(x) \tag{1.17}$$

for every $G \in \mathscr{C}^1(S)$, $x \in S$ and every $b \in E$.

We show moreover the convergence in law, for any fixed parameter $b \in E$, of the processes X^h to X in the Skorokhod space $D([0,T],S)$ of cadlag functions, which is the right space where to study these processes. The convergence is then also in probability, as the limit is deterministic, and hence a constant in the Skorokhod space.

Proposition 1. *Let all the functions C_{ij} and F_{ij} be in $\mathscr{C}^1(S)$. Suppose that the initial points $x(h)$ converge in ℓ^1 to x_0, as $h \to 0$. Then the processes $X^h(\cdot, x(h), b)$ converges in law on the Skorokhod space $D([0,T],S)$ to $X(\cdot, x_0, b)$, as $h \to 0$, for any $b \in E$; whereas the processes are defined in notations 1 and 2.*

Proof. Let $b \in E$ be fixed. The set $\mathscr{C}^1(S)$ is a dense linear subspace of $\mathscr{C}(S)$ and under the assumption of smooth C_{ij} and F_{ij} it is invariant under the limiting semigroup (U_t) defined in (1.14), since the function f defined in (1.11) turns out to be in $\mathscr{C}^1(S)$. So by ([11], proposition 17.9) the set $\mathscr{C}^1(S)$ is a core for the generator Λ_b defined in (1.16).

Then expanding G in Taylor series we have that

$$\lim_{h \to 0} \Lambda_{b,h} G = \Lambda_b G \tag{1.18}$$

uniformly for every $G \in \mathscr{C}^1(S)$. Thus the claim follows from ([11], theorem 17.25) which characterizes the convergence of processes in $D([0,T],S)$.

1.2.4 Controlled systems

Here we deal with the system of small players under some control, i.e. a strategy given by the major player. We assume that this major player focuses in finite horizon time $n\tau$ and can update her strategy only in discrete times

$$k\tau \qquad k = 0, 1, \ldots, n-1.$$

$\tau > 0$ and $n \in \mathbb{N}$ are fixed, and in each time step the controller can change the control parameter b regarding what has happened in the time interval.

The starting point of the Markov chain is $x_0 = x(h) \in h\mathbb{Z}^{fin}$, with the control parameter b_0. After the first time step the Markov chain is in the state $x_1 = X_h(\tau, x_0, b_0)$. Now the major player can change the parameter, so it becomes $b_1 = b_1(x_1)$ that may depend on the current state of the system. She repeats the same procedure at each time step and therefore, in the end, we get what is called a *policy*.

Notation 3 *A* policy *is a sequence of decision rules*

$$\pi = (\pi_0, \pi_1, \ldots, \pi_{n-1}) \tag{1.19}$$

that specify the action of the mayor player at each time step, with

$$\pi_k = \pi_k(x_k) \tag{1.20}$$

and

$$x_k = X^h(\tau, x_{k-1}, \pi_{k-1}). \tag{1.21}$$

Let $X^h_\pi(t, x_0)$ denote the state of the system at time t when the controller applies policy π, starting from the initial point x_0. To shorten the notation we shall sometimes write $X^h_\pi(t)$ instead of $X^h_\pi(t, x_0)$. It is called the controlled system of small players.

Equation (1.21) can be also written as $x_k = X^h_\pi(k\tau)$. At each time step $k\tau$ the controller has an *instantaneous reward* $B(x_k, b_k)$ and in the end she has a *final reward* $V_0(x_n)$. Our goal is to find a strategy that maximizes

$$\begin{aligned} V^h_{\pi,n}(x(h)) :&= E[\tau B(x_0, \pi_0) + \ldots + \tau B(x_{n-1}, \pi_{n-1}) + V_0(x_n)] \\ &= E\left[\sum_{k=0}^{n-1} \tau B(X^h_\pi(k\tau), \pi(X^h_\pi(k\tau)) + V_0(X^h_\pi(n\tau)) \middle| X^h_\pi(0) = x_0\right] \end{aligned} \tag{1.22}$$

where B and V_0 are given continuous functions. It is called the *value* for the system with N players. The maximum over all possible policies is then the *optimal value* for the system with N agents

$$V^h_n(x(h)) := \sup_\pi V^h_{\pi,n}(x(h)). \tag{1.23}$$

We may want to find this optimum value via the usual *dynamic programming* method. First of all we define the *Shapley operator*

$$S[h]V(x) := \sup_{b \in E} [\tau B(x, b) + E(V(X_h(\tau, x, b)))] \tag{1.24}$$

and then by backward recurrence

$$V_k^h = S[h]V_{k-1}, \tag{1.25}$$

hence we get

$$V_n^h = S[h]^n V_0. \tag{1.26}$$

However this procedure might be unfeasible to calculate practically when the number of players increases. So we will consider the optimum of the limit and then study how close these optima are.

1.2.4.1 Controlled limiting system

We want to study the mean field limit system, given by equation (1.11), in a classical control theory setup. Recall that (E,d) is a compact metric space, the one where the parameter b lies.

Notation 4 *We define an* action function *to be a piecewise Lipschitz function from finite horizon time to E*

$$\alpha : [0,T] \to E.$$

We note that an action function is different from a policy, because the latter depends on the state of the system at each step, while the former does not. Thus in this context we rewrite equation (1.12) where f is defined in (1.11) as

$$\dot{x} = f(x,\alpha), \tag{1.27}$$

meaning $\dot{x}(t) = f(x(t),\alpha(t))$ for every $t \geq 0$, considering hence $b = \alpha(t)$, i.e. the control parameter is a function of the time.

Notation 5 $X(t,x,\alpha)$ *is the* flux *of the ODE (1.27), i.e. the solution at time t that is in x at* $t = 0$ *under the control parameter* $b = \alpha = \alpha(s)$. *In integral form*

$$X(t,x_0,\alpha) = x_0 + \int_0^t f(X(s,x_0,\alpha),\alpha(s))ds. \tag{1.28}$$

We are in a finite horizon time T and now we want to maximize

$$v_\alpha(x) := \int_0^T B(X(s,x,\alpha),\alpha(s))ds + V_0(X(T,x,\alpha)) \tag{1.29}$$

where B and V_0 are the same as in (1.22). This is the *value* of the limiting system. The *optimal value* is then

$$v(x) = \sup_\alpha v_\alpha(x). \tag{1.30}$$

Our aim is to study how and under what assumptions we have the convergence of the optimum of the system of small players (1.23) to this optimum (1.30). In fact we want both h and τ tend to 0. To achieve this goal we need further auxiliary systems, to get also the convergence for every policy and every action function.

1.2.5 *Stability of* S

We show that the state space $S := B_+(L,R]$ is stable for all the dynamics considered. We need some regularity for the functions involved in the model. We require that all the functions $C_{ij}(x,b)$ and $F_{ij}(x,b)$ are positive and in $\mathscr{C}^2(S)$ in the variable x, for any b, i.e. twice continuously differentiable on the compact subspace $S \subset \ell^1$, in the topology of ℓ^1. Since S is convex we can take the directional derivatives in every direction, so we have

$$C := \sup_{i,j} C_{ij}(x,b) < \infty, \qquad F = \sup_i \sum_{j<i} F_{ij}(x,b) < \infty \tag{1.31}$$

$$C(1) := \sup_{i,j,k} \left| \frac{\partial C_{i,j}}{\partial x_k}(x,b) \right| < \infty, \qquad F(1) := \sup_{i,k} \sum_{j<i} \left| \frac{\partial F_{i,j}}{\partial x_k}(x,b) \right| < \infty \tag{1.32}$$

$$C(2) := \sup_{i,j,k,l} \left| \frac{\partial^2 C_{i,j}}{\partial x_k \partial x_l}(x,b) \right| < \infty, \qquad F(2) := \sup_{i,k,l} \sum_{j<i} \left| \frac{\partial^2 F_{i,j}}{\partial x_k \partial x_l}(x,b) \right| < \infty \tag{1.33}$$

These constants are all finite as S is compact.

Let us recall that L is the identity, i.e. $L(k) = k$. Therefore, using the above equalities in (1.11) we get that also $f : S \to \ell^1$ is twice continuously differentiable (in $\mathscr{C}^2(S)$) as a map both in ℓ^1 and in $\ell^1(L)$ with the following bounds

$$||f(x)||_{\ell^1} \le 3C||x||^2_{\ell^1} + 3F||x||_{\ell^1} \tag{1.34}$$
$$||f(x)||_{\ell^1(L)} \le 3(C||x||_{\ell^1} + 3F)||x||_{\ell^1(L)} \tag{1.35}$$
$$\|Df(x)\|_{\ell^1} \le 6C||x||_{\ell^1} + 3F + 3[C(1)||x||_{\ell^1} + F(1)]||x||_{\ell^1} \tag{1.36}$$
$$\|Df(x)\|_{\ell^1(L)} \le 8C||x||_{\ell^1(L)} + 3F + 3[2C(1)||x||_{\ell^1} + F(1)]||x||_{\ell^1(L)} \tag{1.37}$$
$$||D^2 f(x)||_{\ell^1} \le 6[C + F(1) + [C(1) + F(2)]||x||_{\ell^1} + C(2)||x||^2_{\ell^1}] \tag{1.38}$$
$$||D^2 f(x)||_{\ell^1(L)} \le 9[C + F(1) + [C(1) + F(2)]||x||_{\ell^1(L)} + C(2)||x||^2_{\ell^1(L)}]. \tag{1.39}$$

We recall now some fact about ODEs in Banach space of measures. In the Markovian dynamics of the system of small players every state represents the number of coalitions of different sizes, which is of course positive. Hence we are interested in an evolution $f : S \longrightarrow \ell^1$ for the dynamic (1.12) $\dot{x} = f(x)$ that preserves positivity, i.e. such that for any initial point $x \in \ell^1_+$ the solution $X(t,x)$ belongs to ℓ^1_+ for any $t \ge 0$. We say that f must be *conditionally positive*, in the following sense:

Definition 1. A function $f : \ell^1 \longrightarrow \ell^1$ is said to be *conditionally positive* if for any $x \in \ell^1_+$ with $x_k = 0$ one has $f_k(x) \geq 0$.

Further, we need the following definitions.

Definition 2. A function $f : \ell^1_+ \longrightarrow \ell^1$ is called *L-subcritical* if

$$\sum_k L(k) f_k(x) \leq 0. \tag{1.40}$$

As a motivation, we observe that $\frac{d}{dt}||x||_{\ell^1(L)} \leq 0$ if f is L-subcritical and $\dot{x} = f(x)$.

Definition 3. A function $f : \ell^1_+ \longrightarrow \ell^1$ is said to satisfy the *Lyapunov condition* if

$$\sum_k L(k) f_k(x) \leq a \sum_k L(k) x_k + b \tag{1.41}$$

for some constant a and b, for all $x \in \ell^1_+$.

The main result concerning the dynamics in ℓ^1 is the following lemma.

Lemma 2. *Assume that the function f is conditionally positive, satisfies the Lyapunov condition and is Lypschitz continuous in the norm of $\ell^1(L)$ on any bounded set of $\ell^1(L)_+$. Then for any $x \in \ell^1(L)_+$ the Cauchy problem (1.12) has a unique global (defined for all times) solution $X(t,x)$ in $\ell^1_+(L)$. Moreover*

$$X(t,x) \in B_+(L, e^{at}(||x||_{\ell^1(L)} + bt)). \tag{1.42}$$

In particular if f is L-subcritical then any $B_+(L,R)$ is invariant.

Proof. By local Lipschitz continuity and conditional positivity, evolution (1.12) is locally well-posed and preserves positivity. Moreover, by the Lyapunov condition,

$$(L, X(t,x)) \leq (L,x) + \int_0^t [a(L, X(s,x)) + b] ds$$

where $(L,x) = \sum_k L(k) x_k$ is the duality between functions and measures. So by Gronwall's lemma and the preservation of positivity

$$0 \leq (L, X(t,x)) \leq e^{at}[(L,x) + bt],$$

implying that the solution can be extended to all times with required bounds.

We have found that, if the assumptions of the lemma are satisfied, the set $B_+(L,x_0]$ is invariant under an L-subcritical evolution f and the ODE has a unique global solution, starting from x_0.

Let us check that the assumptions of lemma 2 are satisfied for f defined in (1.11). The function f is conditionally positive because its domain is S, which is a subset of ℓ^1_+, and the functions $C_{ij}(x,b)$ and $F_{ij}(x,b)$ are positive. Further, if we consider the function $G \in \mathscr{C}^1(S)$ defined by $G(x) = ||x||_{\ell^1(L)} = \sum_k kx_k$ and apply the generator (1.16) to this function then we have $\sum_k kf_k(x) = 0$, since the derivatives of G are $\frac{\partial G}{\partial x_k}(x) = k$. This implies that f is L-subcritical.

Considering equation (1.37) and thanks to the boundedness of S in $\ell^1(L)$ we obtain that f is Lipschitz continuous as a map in $\ell^1(L)$. So all the assumptions of lemma 2 are satisfied, showing the well posedness of the problem and the invariance of S. We would like this set to be invariant also for the system of small players.

1.2.5.1 State space for the small players

For any h, $X^h(t,x)$ is a continuous-time Markov chains on $h\mathbb{N}^{fin} \cap \ell^1_+$. Let us say that $X^h(t,x)$ are *L-non-increasing*, if any jump of $X^h(t,x)$ cannot increase L. In this case a trajectory $X^h(t,x)$ stays forever in $B_+(L,R)$ whenever the initial point $x \in B_+(L,R)$. Moreover X^h is *L-subcritical* in the sense that its generator $\Lambda_{h,b}$ satisfies the inequality $\Lambda_{h,b}(L) \leq 0$.

The Markov chains $X^h(t,x(h))$ are L-non increasing and have bounded generators. In fact the state space of the coalitions of small players is actually finite for any fixed h. Indeed we recall that if $X^h(t,x)$ is in the state x then x_k is h times the number of coalitions of size k. The total number of small players is fixed $N = N(h)$ for any h, so $x_k = 0$ for any $k \geq N$ and $x_k \leq N/h$ for any $k \leq N$, meaning that the state space is finite.

The total number of small players N is of course constant. So the norm in $\ell^1(L)$ of the states x of the Markov chain $X^h(t)$ is conserved:

$$||x||_{\ell^1(L)} = \sum_{k=1}^{N} kx_k = h\sum_{k=1}^{N} kn_k = hN(h). \tag{1.43}$$

Hence if the initial point $x(h)$ is in $S = B_+(L,R]$ then any state is in S, i.e. the set $B_+(L,R]$ is invariant for the dynamics of X^h.

Notation 6 *$S(h)$ is the finite state space of the Markov chain X^h, the system of $N = N(h)$ small players. It is a subset of the compact $B_+(L,R]$ in ℓ^1 and a subset of $\mathbb{R}^N$ and of the set $h\mathbb{N}^{fin}$. Denote by $M(h)$ the number of elements of $S(h)$*

$$S(h) := h\mathbb{N}^{fin} \cap B_+(L,R].$$

So we can define $S := B_+(L,R]$ for a suitable R such that this set contains all the initial data $x(h)$ and x_0. Such an R exists because we will consider $\lim_{h\to 0} x(h) = x_0$ and then the sequence is bounded. S is the compact set in ℓ^1 invariant for all the dynamics considered. Thanks to (1.43) we have

$$N(h) \leq \frac{R}{h} \tag{1.44}$$

for any h. Further $\lim_{h\to 0} M(h) = \lim_{h\to 0} N(h) = +\infty$ and the finite spaces $S(h)$ are decreasing, i.e. if $h > l$ then $S(h) \subset S(l)$, and $\bigcup_{h>0} S(h) \subseteq S$. Moreover

$$S(h) \subset S \subset \ell^1(L) \subset \ell^1 \subset \ell^2. \tag{1.45}$$

1.2.6 Assumptions in the model

Let us summarize the assumptions we make on our model. Recall that in the system of N small players the controller acts at time steps $k\tau$ for $k = 0, 1, \ldots, n-1$.

- **(H1)** $S = B_+(L,R]$ is the state space of all the dynamics considered, and the initial states lie in S;
- **(H2)** The functions $C_{ij}(x,b)$ and $F_{ij}(x,b)$ are positive and in $\mathscr{C}^2(S)$ in the variable x, for any $b \in E$, and Lipschitz continuous in the variable b, for any $x \in S$;
- **(H3)** The rewards $B(x,b)$ and $V_0(x)$ are Lipschitz continuous in x in the ℓ^2-norm, uniformly in b, and B is bounded;
- **(H4)** The time step $\tau = \tau(h)$ depends on h, as well as $N = N(h)$ does, and
$$\lim_{h\to 0} \tau(h) = \lim_{h\to 0} \tau(h)\sqrt{N(h)} = 0;$$
- **(H5)** The horizon T is fixed and the number of steps is
$$n(h) := \left\lfloor \frac{T}{\tau(h)} \right\rfloor \tag{1.46}$$
for any h, which tends to infinity, as h tends to 0;
- **(H6)** The rescaling parameter h of the model is such that
$$\lim_{h\to 0} h(N(h))^2 = \lim_{h\to 0} \tau(h)(N(h))^2 = 0. \tag{1.47}$$

Equivalently, we can think of studying the limit as N tends to infinity. So the parameter $h = h(N)$ has to satisfy the latter conditions and it represents the rescaling parameter for the system of N players.

1.3 Mean field convergence

In this section we present our main results. We follow the ideas in [8], hence we firstly introduce the two auxiliary systems.

1.3.0.1 First auxiliary system

This is a system with N agents controlled by an action function borrowed from the mean field limit. More precisely, let α be an action function that specifies the action to be taken at time t. Although α has been defined for the limiting system, it can also be used for the system with N players. In this case, the action function α can be seen as a policy that does not depend on the state of the system.

At step k, the controller applies action

$$\alpha_k := \alpha(k\tau),$$

so (1.20) gives a policy $(\alpha_0, \ldots, \alpha_{n-1})$ as in (1.19), but independent of the state of the system.

By abuse of notation, we denote by $X^h_\alpha(t)$ as in (3) the state of the system at time t when applying the policy derived from the action function α as explained above. In what follows policies will always be denoted by π and action functions by α. Here (1.21) becomes

$$x_k = X^h(\tau, x_{k-1}, \alpha_{k-1}) = X^h_\alpha(k\tau),$$

starting from initial point x_0 with control parameter $\alpha_0 = \alpha(0)$.

The value for this system, similarly to (1.22), is defined by

$$\begin{aligned} V^h_{\alpha,n}(x_0) :&= E[\tau B(x_0, \alpha_0) + \ldots + \tau B(x_{n-1}, \alpha_{n-1}) + V_0(x_n)] \\ &= E\left[\sum_{k=0}^{n-1} \tau B(X^h_\alpha(k\tau), \alpha(k\tau)) + V_0(X^h_\alpha(n\tau)) \middle| X^h_\alpha(0) = x_0\right]. \end{aligned} \tag{1.48}$$

1.3.0.2 Second auxiliary system

The method of proof uses a second auxiliary system in which trajectories are considered. This is a limiting system controlled by an action function derived from the policy of the original system with N agents.

Consider the system with N players under policy π. The stochastic process $X^h_\pi = X^h_{\pi,n(h)}$ is defined on some probability space Ω. To every $\omega \in \Omega$ there corresponds a trajectory $X^h_\pi(\omega)$, and for every $\omega \in \Omega$ we define the piecewise constant action function $A^h_\pi(\omega)$, as explained in the following

Notation 7

$$A^h_\pi(\omega) : [0, T] \to E$$

is an action function such that

- *this random function is constant on each interval* $[k\tau, (k+1)\tau[$ *for any* $k = 0, 1, \ldots, n-1$*;*

- $A^h_\pi(\omega)(k\tau) := \pi_k = \pi_k(X^h_\pi(k\tau))$ *is the action taken by the major player of the system with N agents at time slot* $k\tau$*, under policy* π*.*

Recall from notation 5 that for any $x_0 \in S \subset \ell^1(\mathbb{N})$ and any action function α, $X(x_0,\alpha)$ is the solution of the ODE (1.27). For every ω, $X(t,x_0,A^h_\pi(\omega))$ is the solution of the limiting system with action function $A^h_\pi(\omega)$, as in (1.28), i.e.

$$X(t,x_0,A^h_\pi(\omega)) = x_0 + \int_0^t f(X(s,x_0,A^h_\pi(\omega)),A^h_\pi(\omega)(s))ds. \tag{1.49}$$

The value function for this system is as in (1.29).

When ω is fixed, $X(t,x_0,A^h_\pi(\omega))$ is a continuous time deterministic process corresponding to one trajectory $X^h_\pi(\omega)$. When considering all possible realizations of X^h_π, $X(t,x_0,A^h_\pi)$ is a random, continuous time function *coupled* to X^h_π, i.e. a stochastic process. Its randomness comes only from the action term A^h_π, in the ODE (1.27). In the following we omit the dependence on ω in our writing. A^h_π and X^h_π will always designate the processes corresponding to the same ω.

1.3.1 Main results

The main result establishes the convergence of the optimization problem for the system with N players to the optimization problem for the mean field limit, through their value functions.

Theorem 1. *Under assumptions (H1)-(H6), if* $\lim_{h\to 0} x(h) = x_0$ *almost surely, respectively in probability, then*

$$\lim_{h\to 0} V^h(x(h)) = v(x_0) \tag{1.50}$$

almost surely, respectively in probability, where V^h *and* v *are the optimal values defined in (1.23) and (1.30).*

The second result states that an optimal action function for the mean field limit provides an asymptotically optimal strategy for the system with N agents. Let us denote by $(\hat{X}^h_\alpha(t))_{t\geq 0}$ the continuous time process which is the affine interpolation of $X^h_\alpha(t)$ (the first auxiliary system) in the points $k\tau$ and similarly by $\hat{X}^h_\pi(t)$ the affine interpolation of $X^h_\pi(t)$ under policy π.

Theorem 2. *Under assumptions (H1)-(H6), let* α *be a piecewise Lipschitz continuous action function on* $[0,T]$*, of Lipschitz constant* K_α*, and with at most p discontinuity points. Then there exist functions J,* I'_0 *and* B' *satisfying*

$$\lim_{h\to 0} I'_0(h,\alpha) = \lim_{h\to 0} J(h,T) = 0, \quad \lim_{\substack{h\to 0\\ \delta\to 0}} B'(h,\delta) = 0$$

such that for all $\varepsilon > 0$

$$P\left\{\sup_{0<t<T}\left\|\hat{X}^h_\alpha(t) - X(t,x_0,\alpha)\right\| > \left[\left\|X^h(0) - x_0\right\| + I_0'(h,\alpha)T + \varepsilon\right] e^{L_1 T}\right\} \leq \frac{J(h,T)}{\varepsilon^2} \tag{1.51}$$

and

$$\left|V^h_\alpha(X^h(0)) - v_\alpha(x_0)\right| \leq B'\left(h, \left\|X^h(0) - x_0\right\|\right). \tag{1.52}$$

Inequality (1.52) implies also that if

$$\lim_{h\to 0} X^h(0) = x_0 \tag{1.53}$$

almost surely, respectively in probability, then

$$\lim_{h\to 0} V^h_\alpha(X^h(0)) = v_\alpha(x_0) \tag{1.54}$$

almost surely, respectively in probability.

The following corollary combines Theorems 1 and 2. It states that an optimal action function for the limiting system is asymptotically optimal for the system of small players.

Corollary 1. *If* α_* *is an optimal action function for the limiting system and if* $\lim_{h\to 0} X^h(0) = x_0$ *almost surely, respectively in probability, then*

$$\lim_{h\to 0}\left|V^h_{\alpha_*}(X^h(0)) - V^h(X^h(0))\right| = 0 \tag{1.55}$$

almost surely, respectively in probability.

Proof. The assumption says that there exists an action function α_* that maximizes v, i.e. $v(x_0) = v_{\alpha_*}(x_0) = \max_\alpha v_\alpha(x_0)$. Hence we have

$$\left|V^h_{\alpha_*}(X^h(0)) - V^h(X^h(0))\right| \leq \left|V^h_{\alpha_*}(X^h(0)) - v_{\alpha^*}(x_0)\right| + \left|V^h(X^h(0)) - v(x_0)\right|.$$

So, if (1.53) holds, the first modulus goes to 0 by (1.54) and the second by (1.50). Therefore (1.55) is given combining Theorems 1 and 2.

1.3.1.1 Auxiliary results

In order to prove the main theorems we need two auxiliary results.

Theorem 3. *Under assumptions (H1)-(H6), there exist functions* I_0 *and* J *satisfying* $\lim_{h\to 0} I_0(h,\alpha) = \lim_{h\to 0} J(h,T) = 0$ *such that for any* $\varepsilon > 0$, $h > 0$ *and any policy* π

$$P\left\{\sup_{0<t<T}\left\|\hat{X}^h_\pi(t)-X(t,x_0,A^h_\pi)\right\|>\left[\left\|X^h(0)-x_0\right\|+I_0(h)T+\varepsilon\right]e^{L_1T}\right\}\le\frac{J(h,T)}{\varepsilon^2}. \tag{1.56}$$

If (1.53) holds this theorem shows the convergence in probability of the controlled system with N agents, with explicit bounds.

The second statement deals with the convergence of the value for the controlled system of small players to the value of the second auxiliary system. Let π be a policy and A^h_π be the sequence of actions corresponding to a trajectory of X^h_π, as in notation 7. Equation (1.29) defines the value for the deterministic limit, whereas α is an action function. When applying the random action function A^h_π, this defines a random variable $v_{A^h_\pi}(x_0)$. A consequence of Theorem 3 is the convergence of $V^h_\pi(X^h(0))$ to the expectation of this random variable.

Theorem 4. *Let A^h_π be the random action function associated with X^h_π as in notation 7. Under assumptions (H1)-(H6), there exist a function B satisfying*

$$\lim_{\substack{h\to 0\\ \delta\to 0}} B(h,\delta)=0 \tag{1.57}$$

such that

$$\left|V^h_\pi(X^h(0))-E[v_{A^h_\pi}(x_0)]\right|\le B(h,\left\|X^h(0)-x_0\right\|). \tag{1.58}$$

This implies that if (1.53) holds almost surely, respectively in probability, then

$$\lim_{h\to 0}\left|V^h_\pi(X^h(0))-E[v_{A^h_\pi}(x_0)]\right|=0 \tag{1.59}$$

almost surely, respectively in probability.

The proofs of the main results use the two auxiliary systems. The first auxiliary system provides a strategy for the system with N agents derived from an action function of the mean field limit. It can not do better than the optimal value of the system of small players and it is close to the optimal value of the mean field limit. Therefore the optimal value for the system with N players is lower bounded by the optimal value for the mean field limit.

The second auxiliary system is used in the opposite direction: it shows that for large N the two optimal values are the same.

1.3.2 Requirements for convergence

Let us denote by $||x||$ the ℓ^2-norm of x and define the *drift* of the model as

$$F^h(x,b):=E(X^h(\tau,x,b)-x). \tag{1.60}$$

Due to Theorems 2-6 in [8], in order to prove Theorems 1-4 it is sufficient to show that

- **(A1)** There exist some non random functions $I_1(h)$ and $I_2(h)$ such that

$$\lim_{h\to 0} I_1(h) = \lim_{h\to 0} I_2(h) = 0$$

and that for all x and all policies π the number coalitions $\Delta_\pi^h(k)$ that perform a transition between time step $k\tau$ and $(k+1)\tau$ satisfies

$$E(\Delta_\pi^h(k)|X_\pi^h(k\tau) = x) \le \frac{1}{h} I_1(h) \tag{1.61}$$

$$E(\Delta_\pi^h(k)^2|X_\pi^h(k\tau) = x) \le \frac{1}{h^2} \tau I_2(h); \tag{1.62}$$

- **(A2)** There exist a function $I_0(h)$ such that $\lim_{h\to 0} I_0(h) = 0$ and

$$\left\| \frac{F^h(x,b)}{\tau(h)} - f(x,b) \right\| \le I_0(h) \tag{1.63}$$

for every $x \in S$ and $b \in E$ where the function f is the one in (1.11), and moreover f is defined on $S \times E$ and there exists a constant L_2 such that

$$|f(x,b)| \le L_2; \tag{1.64}$$

- **(A3)** There exist constants L_1, K and K_B such that for all x, $y \in S$ and a, $b \in E$

$$||F^h(x,b) - F^h(y,b)|| \le L_1 ||x-y|| I(h) \tag{1.65}$$

$$||f(x,b) - f(y,a)|| \le K(||x-y|| + d(a,b)) \tag{1.66}$$

$$||B(x,b) - B(y,b)|| \le K_B ||x-y|| \tag{1.67}$$

$$||V_0(x) - V_0(y)|| \le K_B ||x-y|| \tag{1.68}$$

and the reward is bounded:

$$\sup_{\substack{x\in S\\ b\in E}} \max\{|B(x,b)|, |V_0(x,b)|\} =: ||B||_\infty. \tag{1.69}$$

We will show in the next section that if our assumptions (H1)-(H3) are satisfied then (A1)-(A3) hold.

Let us now fix the functions appearing in (A1), (A2) and (A3):

$$K := 6CR + 3F + 3R[C(1)R + F(1)], \tag{1.70}$$
$$L_2 := 3CR^2 + 3FR, \tag{1.71}$$
$$I_0(h) := \sqrt{N(h)}\frac{\tau}{2}(R_1 + hR_2), \tag{1.72}$$
$$I_1(h) := \tau(CR^2 + F), \tag{1.73}$$
$$I_2(h) := (CR^2 + F)[\tau(CR^2 + F) + h], \tag{1.74}$$

where

$$R_1 := 3(CR^2 + FR)(6CR + 3F + 3C(1)R^2 + F(1)R),$$
$$R_2 := 54(CR^2 + FR)(C + F(1) + C(1)R + F(1)R + C(2)R^2).$$

Clearly $\lim_{h\to 0} I_0(h) = \lim_{h\to 0} I_1(h) = \lim_{h\to 0} I_2(h) = 0$ if (H4) holds. L_1 actually depends on h

$$L_1 = L_1(h) := Ke^{M_2\sqrt{N(h)}\tau}, \tag{1.75}$$

where $M_2 := 3(C(1)R^2 + 2CR + F(1)R + F)$. Hence L_1 is tends to the constant K by (H4), as h tends to 0.

Let us define the functions I_0', J, B, B' appearing in the statements of Theorems 2-4 by the following equations

$$I_0'(h,\alpha) := I_0(h) + \tau K e^{(K-L_1)T} \cdot \left[\frac{K_\alpha}{2} + 2\left(1 + \min\left\{\frac{1}{\tau}, p\right\}\right)||\alpha||_\infty\right], \tag{1.76}$$
$$J(h,T) := 8T\left\{L_1^2\left[I_2(h)\tau^2 + I_1(h)^2(T+\tau)\right] + N(h)^2\left[2I_2(h) + \tau L_2^2\right]\right\}, \tag{1.77}$$
$$B(h,\delta) := \tau||B||_\infty + K_B\sqrt{2}I_1(h) + K_B(\delta + I_0(h)T)\left(e^{L_1T} + \frac{e^{L_1T} - 1}{L_1}\right) \tag{1.78}$$
$$+ \frac{3}{2^{\frac{1}{3}}}\left[e^{L_1T} + \frac{e^{L_1T} - 1 + \frac{\tau}{2}}{L_1}\right]^{\frac{2}{3}} \cdot K_B^{\frac{2}{3}}||B||_\infty^{\frac{1}{3}} J(h,T)^{\frac{1}{3}}(T+1)^{\frac{2}{3}} \tag{1.79}$$

and $B'(h,\delta)$ has the same expression as $B(h,\delta)$ replacing $I_0(h)$ by $I_0'(h,\alpha)$. From (H4) and (H6) follow that $\lim_{h\to 0} J(h,T) = \lim_{h\to 0} I_0'(h,\alpha) = 0$ and $\lim_{\substack{h\to 0\\ \delta\to 0}} B'(h,\delta) = \lim_{\substack{h\to 0\\ \delta\to 0}} B(h,\delta) = 0$.

1.3.3 Constructing an optimal policy

By means of corollary 1, an optimal action function for the mean field limit is asymptotically optimal for the system of small players. This provides a way for constructing an asymptotically optimal policy.

We denote by $u(x,t)$ the optimal cost over horizon $[t,T]$ for the limiting system. Under our hypothesis, the following proposition holds.

Proposition 2. *The value function $u(x,t)$ is the unique, bounded and uniformly continuous, viscosity solution in $S \times [0,T] \subset \ell^2 \times [0,T]$ of the Hamilton-Jacobi-Bellman equation*

$$-\frac{\partial u(x,t)}{\partial t} - \max_{b \in E}\{\nabla u(x,t) \cdot f(x,b) + B(x,b)\} = 0 \tag{1.80}$$

which satisfies the terminal condition $u(x,T) = V_0(x)$.

Let us recall that the definition of viscosity solution in a Hilbert space, as ℓ^2, does not differ from the usual one. Further, under our assumptions, the Hamiltonian defined as $H(x,p) := \max_{b\in E}\{p \cdot f(x,b) + B(x,b)\}$ for any $(x,p) \in S \times \ell^2$ is such that

$$|H(x,p) - H(y,p)| \leq C||x-y||(1+||p||)$$
$$|H(x,p) - H(x,q)| \leq C||p-q||,$$

where C is a constant. Therefore existence and uniqueness of bounded and uniformly continuous viscosity solutions to (1.80) are implied by Theorem 5.1 in [6].

We state the algorithm presented in [8] for constructing an asymptotically optimal policy for the system of small players X^h via an optimal action function for the mean field limit X:

- Let $X^h(0)$ be the initial condition of the limiting system. Solve the Hamilton-Jacobi-Bellman equation (1.80) on $[0, \tau n(h)]$. Assume this provides an optimal control function α_*;
- Construct a policy π for the system of small players: the action to be taken under state $X^h(k\tau)$ at step k is
$$\pi_k(X^h(k\tau)) := \alpha_*(k\tau).$$

The asymptotic optimality of the related value is ensured by corollary 1. The policy π constructed above is static in the sense that it does not depend on the state $X^h(k\tau)$ but only on the initial state $X^h(0)$. The deterministic estimation of $X^h(k\tau)$ is provided by the differential equation.

The algorithm described above uses an optimal action function for the limiting system which may not exist, as we observed above. A sufficient condition for its existence is that the set $f(x,E) \times B(x,E)$ is convex for all $x \in S$; a proof of this fact can be found in [1].

However the existence of an optimal action function is actually not necessary in the algorithm described above. Indeed, if there is no optimal control of the HJB equation (1.80), one can replace the optimal α_* used in the algorithm by an action function which is h-optimal. This still provides an asymptotically optimal policy.

1.3.4 Example

We show here that in a class of applications the limiting problem can be reduced to an optimization problem in one dimension. This provides a computationally much more effective scheme in order to obtain an asymptotically optimal policy for the system of small players, in view of the algorithm presented above.

Firstly, we study a particular case in which the mean field limit admits an explicit solution. So let us consider $E = [0,1]$ and a particular shape of the functions C_{ij} and F_{ij}, in which the there is no dependence on x:

$$C_{ij}(x,b) = b \tag{1.81}$$

$$F_{ij}(x,b) = \frac{1-b}{i-1}. \tag{1.82}$$

In particular if b is 1 then only merging is possible. While if it is 0 only splitting is possible.

Let $m(x) := \sum_i x_i$ be the norm of $x \in \ell^1$. With these rates the limiting evolution f in (1.11) can be studied considering only the dynamics of the norm. Evaluating the limiting generator (1.16) on $G = m$ we obtain

$$\sum_i f_i(x,b) = -bm^2 + (1-b)m. \tag{1.83}$$

Hence the evolution of m is described by

$$\begin{cases} \dot{m} = -bm^2 + (1-b)m \\ m(0) = m_0, \end{cases} \tag{1.84}$$

which is in fact an ODE on $\mathbb{R}_+$.

If b is constant then the explicit solution to (1.84) is

$$m(t,m_0,b) = \frac{m_0(1-b)e^{(1-b)t}}{1-b-m_0 b + m_0 e^{(1-b)t}}. \tag{1.85}$$

if $b \neq 1$. In particular, if $b = 0$, $m(t,m_0,0) = m_0 e^t$. While if $b = 1$

$$m(t,m_0,1) = \frac{m_0}{1+tm_0}. \tag{1.86}$$

Observe that $m(t,m_0,b)$ is a continuous and strictly decreasing function of $b \in [0,1]$, for any value of $0 < t \leq T$ and $m_0 > 0$.

We assume that there is no instantaneous reward and that the final reward $V_0 = V_0(m)$ is a strictly concave function of m which has a unique maximum in m^*. The value function V in this case can be computed explicitely, that is

$$V(t,m) = \begin{cases} V_0(me^{T-t}) & \text{if } m < m^* e^{-(T-t)} \\ V_0(m^*) & \text{if } m^* e^{-(T-t)} \leq m \leq \frac{m^*}{1-(T-t)m^*} \\ V_0\left(\frac{m}{1+(T-t)m}\right) & \text{if } m > \frac{m^*}{1-(T-t)m^*}. \end{cases} \tag{1.87}$$

If $m^* e^{-(T-t)} \leq m_0 \leq \frac{m^*}{1-(T-t)m^*}$ denote by $b^*(T, m_0)$ the unique value of $b \in [0,1]$ such that $m(T, m_0, b) = m^*$. Therefore, if $m(0) = m_0$, an optimal action function for the mean field limiting problem is to choose the constant value

$$\alpha^*(t, m_0) = \alpha^*(m_0) := \begin{cases} 1 & \text{if } m_0 < m^* e^{-(T-t)} \\ b^*(T, m_0) & \text{if } m^* e^{-(T-t)} \leq m_0 \leq \frac{m^*}{1-(T-t)m^*} \\ 0 & \text{if } m_0 > \frac{m^*}{1-(T-t)m^*}. \end{cases} \tag{1.88}$$

This example can be generalized to the case in which the rates C_{ij} and F_{ij} depend also on the norm of x. Namely, let

$$C_{ij}(x, b) = b f_C(m(x)) \tag{1.89}$$

$$F_{ij}(x, b) = \frac{1-b}{i-1} f_B(m(x)), \tag{1.90}$$

where f_C and f_B are some Lipschitz continuous and non-negative functions. The dynamics (1.84) for m becomes

$$\begin{cases} \dot{m} = -b f_C(m) m^2 + (1-b) f_B(m) m \\ m(0) = m_0. \end{cases} \tag{1.91}$$

Also in this case, the mean field limiting problem reduces to an optimization problem in one dimension. Thus there are numerical schemes which provide an $\frac{1}{N}$-optimal action function in feedback form (see e.g. [1]). These are much more efficient than trying to solve the Bellman equation (1.26). In fact the prelimit problem is allowed to be tackled when N is lower than a few tens: see [20].

1.4 Proofs of convergence

In this section we show that (A1)-(A3) hold in our fragmentation-coagulation model. So (H1)-(H6) hold and recall that the compact state space $S \subset \ell^1$ is stable for all the dynamics considered.

1.4.1 Lipschitz continuity of the limit

We verify that (1.66) and (1.64) are satisfied for the function f in (1.11), if (H2) holds. We actually do not need f to be in $\mathscr{C}^2(S)$, but only in $\mathscr{C}^1(S)$. We use the fact that

$$||x|| = ||x||_{\ell^2} \leq ||x||_{\ell^1} \leq ||x||_{\ell^1(L)} \leq R \tag{1.92}$$

for any $x \in S$. So equation (1.34) gives

$$||f(x)|| \leq 3CR^2 + 3FR$$

for all $x \in S$, which is the assumption (1.64) with bound

$$L_2 := 3CR^2 + 3FR. \tag{1.93}$$

By (1.92) and its definition (1.2) we get also

$$||Df(x)||_{\ell^2} \leq ||Df(x)||_{\ell^1} \quad \forall x \in S$$

which, together with equation (1.36), yields

$$||Df(x)||_{\ell^2} \leq 6CR + 3F + 3R[C(1)R + F(1)]. \tag{1.94}$$

Applying then the mean value theorem in the convex space S, for every x and y in S there exist a point z in the segment $[x,y]$ such that

$$f(y) - f(x) = Df(z).(y-x).$$

Thus taking the ℓ^2-norm and using the fact that Df is a bounded linear map we get

$$||f(y) - f(x)|| \leq ||Df(z)||_{\ell^2} ||y-x||$$

for all x and y in S. According to (1.94) we find that the limiting function is Lipschitz continuous in x, for the ℓ^2-norm with

$$K = 6CR + 3F + 3R[C(1)R + F(1)] \tag{1.95}$$

as a Lipschitz constant.

Clearly (H3) implies (1.67), (1.68) and (1.69).

1.4.2 Convergence of the drift

Here we prove (1.63). Recall that the drift of the model is

$$F^h(x,b) := E[X^h(\tau,x,b) - x].$$

The semigroup U_t^h of the Markov process X^h is defined by

$$U_t^h G(x,b) := E[G(X^h(t,x,b))]$$

for any $G \in \mathscr{C}(S,\mathbb{R})$ and $t \geq 0$. U_0 is the identity and U_t satisfies the Kolmogorov differential equation

$$\frac{\partial}{\partial t} U_t^h G(x,b) = \Lambda_{h,b} U_t^h G(x,b) = U_t^h \Lambda_{h,b} G(x,b), \tag{1.96}$$

where $\Lambda_{h,b}$ is the infinitesimal generator defined in (1.10).

We apply U_t^h to the projection on the k-th coordinate G_k. This function is in $\mathscr{C}(S,\mathbb{R})$ and

$$F_k^h(x,t) := G_k\left(E[X^h(\tau,x,b)-x]\right) = EG_k(X^h(\tau,x,b)) - x_k = U_\tau^h G_k(x,b) - G_k(x). \tag{1.97}$$

Let us denote $u_k(t,x,b) := U_t^h G_k(x,b)$. The Taylor formula applied to u_k in the variable t gives

$$u_k(\tau,x,b) = u_k(0,x,b) + \tau\frac{\partial u_k}{\partial t}(0,x,b) + \frac{\tau^2}{2}\frac{\partial^2 u_k}{\partial t^2}(s,x,b) \tag{1.98}$$

for a fixed $s \in]0,\tau[$ and for every x and b. We have

$$u_k(0,x,b) = G_k(x) = x_k \tag{1.99}$$

and according to (1.96)

$$\frac{\partial u_k}{\partial t}(0,x,b) = \Lambda_{h,b} u_k(0,x,b) = \Lambda_{h,b} G_k(x) \tag{1.100}$$

for every x and b.

The generator (1.10) calculated on the projection, using $G_k(e_i) = \delta_{ik}$, gives

$$\begin{aligned}
\Lambda_{b,h} G_k(x) &= \frac{1}{h}\sum_{i,j} C_{ij}(x,b)x_i x_j \left[G_k(x - he_i - he_j + he_{i+j}) - G_k(x)\right] \\
&+ \frac{1}{h}\sum_i \sum_{i<j} F_{ij}(x,b)x_i \left[G_k(x - he_i + he_j + he_{i-j}) - G_k(x)\right] \\
&= \sum_{i,j} C_{ij}(x,b)x_i x_j \left[\delta_{i+j,k} - \delta_{ik} - \delta_{jk}\right] \\
&+ \sum_i \sum_{i<j} F_{ij}(x,b)x_i \left[\delta_{jk} + \delta_{i-j,k} - \delta_{ik}\right] \\
&= \sum_{i<k} C_{i,k-i}(x,b)x_i x_{k-i} - \sum_j C_{kj}(x,b)x_k x_j - \sum_i C_{ik}(x,b)x_i x_k \\
&+ \sum_{i>k} F_{ik}(x,b)x_i + \sum_{i>k} F_{i,i-k}(x,b)x_i - \sum_{j<k} F_{kj}(x,b)x_k
\end{aligned}$$

and, observing that $C_{ij} = C_{ji}$ and $F_{ij} = F_{i,i-j}$, the latter expression is exactly $f_k(x,b)$ defined in (1.11), whence

$$\Lambda_{b,h} G_k(x) = f_k(x,b) \tag{1.101}$$

for every x and b.

We need also an estimate of the latter term in (1.98). Equation (1.96) yields

$$\frac{\partial^2 u_k}{\partial t^2}(s,x,b)\frac{\partial}{\partial t}U_s^h \Lambda_{b,h} G_k(x) = U_s^h \Lambda_{b,h}\Lambda_{b,h} G_k(x) = U_s^h \Lambda_{b,h} f_k(x,b).$$

We use then the fact that the semigroup U_s^h is, for any s, a contraction in the space $\mathscr{C}(S,\mathbb{R})$ equipped with the sup-norm $||G||_\infty := \sup_{x\in S}|G(x)|$. Thus the latter term in (1.98) is bounded by

$$\left|\frac{\partial^2 u_k}{\partial t^2}(s,x,b)\right| \le \left\|U_s^h \Lambda_{b,h} f_k\right\|_\infty \le \left\|\Lambda_{b,h} f_k\right\|_\infty \tag{1.102}$$

for any $x\in S$, $b\in E$ and $s>0$.

The estimate for the norm of $\Lambda_{b,h} f_k$ is found applying again the Taylor formula to f_k:

$$\begin{aligned}
\Lambda_{b,h} f_k(x) &= \frac{1}{h}\sum_{i,j} C_{ij}(x,b)x_i x_j\left[f_k(x-he_i-he_j+he_{i+j})-f_k(x)\right]\\
&+\frac{1}{h}\sum_i\sum_{j<i} F_{ij}(x,b)x_i\left[f_k(x-he_i+he_j+he_{i-j})-f_k(x)\right]\\
&=\sum_{i,j} C_{ij}(x,b)x_i x_j\sum_l \frac{\partial f_k}{\delta x_l}(x)\left(\delta_{i+j,l}-\delta_{il}-\delta_{jl}\right)\\
&+h\sum_{i,j} C_{ij}(x,b)x_i x_j\sum_l\sum_m \frac{\partial}{\delta x_l}\frac{\partial}{\delta x_m} f_k(y)\left(\delta_{i+j,l}-\delta_{il}-\delta_{jl}\right)\left(\delta_{i+j,m}-\delta_{im}-\delta_{jm}\right)\\
&+\sum_i\sum_{j<i} F_{ij}(x,b)x_i\sum_l \frac{\partial f_k}{\delta x_l}(x)\left(\delta_{jl}+\delta_{i-j,l}-\delta_{il}\right)\\
&+h\sum_i\sum_{j<i} F_{ij}(x,b)x_i\sum_l\sum_m \frac{\partial}{\delta x_l}\frac{\partial}{\delta x_m} f_k(z)\left(\delta_{jl}+\delta_{i-j,l}-\delta_{il}\right)\left(\delta_{jm}+\delta_{i-j,m}-\delta_{im}\right)
\end{aligned}$$

for certain fixed points y and z in S. This gives, using the estimates (1.31), (1.36) and (1.38) and the definitions of the norms of the derivatives Df and $D^2 f$,

$$\begin{aligned}
\left|\Lambda_{b,h} f_k(x)\right| &\le C||x||_{\ell^1}^2\left(3||Df(x)||_{\ell^1}+9h||D^2 f(y)||_{\ell^1}\right)\\
&+F||x||_{\ell^1}\left(3||Df(x)||_{\ell^1}+9h||D^2 f(z)||_{\ell^1}\right)\\
&\le C||x||_{\ell^1}^2[3\,(6C||x||_{\ell^1}+3F+3[C(1)||x||_{\ell^1}+F(1)]||x||_{\ell^1})\\
&+54h\left(C+F(1)+[C(1)+F(2)]||y||_{\ell^1}+C(2)||y||_{\ell^1}^2\right)]\\
&+F||x||_{\ell^1}[3\,(6C||x||_{\ell^1}+3F+3[C(1)||x||_{\ell^1}+F(1)]||x||_{\ell^1})\\
&+54h\left(C+F(1)+[C(1)+F(2)]||z||_{\ell^1}+C(2)||z||_{\ell^1}^2\right)]\\
&\le (CR^2+FR)[3(6CR+3F+3C(1)R^2+F(1)R)\\
&+54h(C+F(1)+C(1)R+F(1)R+C(2)R^2)]
\end{aligned}$$

for every $x\in S$ and $b\in E$ and $k=1,\dots,N(h)$.

Therefore

$$\left\|\Lambda_{b,h} f_k\right\|_\infty \le R_1 + hR_2 \tag{1.103}$$

for every $x \in S$ and $b \in E$ and $k = 1, \dots, N(h)$, where $R_1 := 3(CR^2 + FR)(6CR + 3F + 3C(1)R^2 + F(1)R)$ and $R_2 := 54(CR^2 + FR)(C + F(1) + C(1)R + F(1)R + C(2)R^2)$.

Equations (1.97) and (1.98), by means of (1.99), (1.100) and (1.101), lead to

$$F_k^h(x,b) - \tau f_k(x,b) = \frac{\tau^2}{2} \frac{\partial^2 u_k}{\partial t^2}(s,x,b)$$

and this applying (1.102) and (1.103) yields

$$\left| F_k^h(x,b) - \tau f_k(x,b) \right| \le \frac{\tau^2}{2}(R_1 + hR_2)$$

for any $k = 1, \dots, N(h)$. Considering the ℓ^2-norm we have hence

$$\left\| \frac{F^h(x,b)}{\tau} - f(x,b) \right\| \le \sqrt{N(h)} \frac{\tau}{2}(R_1 + hR_2)$$

which is (1.63), whereas I_0 is given by (1.72).

1.4.3 Lipschitz continuity of the drift

Here we verify that the drift F^h of the model is Lipschitz continuous and we also find a constant for which it is bounded in ℓ^2. We use two tools. The first is the fact that the process

$$M_G(t) := G(X(t)) - G(X(0)) - \int_0^t \Lambda(X(s))ds \tag{1.104}$$

are martingales with respect to the filtration generated by the Markov process $X(t)_{t \ge 0}$, for every G in the domain of its generator. While the second tool is the notion of *coupling* for Markov chains.

We want to study the behavior of the two Markov chains $X^h(t,x,b)$ and $X^h(t,y,b)$ for $x \neq y$ and link them in some sense in the product space. So we could define a coupling of these stochastic processes in terms of their distributions in the space of paths $D([0,T],S)$, the space of cadlag functions, for fixed initial points. However, for given marginal Markov processes, the resulting coupled process may not be Markovian. So we introduce the following fundamental definition, which can be found for instance in [3].

Definition 4. Given two Markov processes with semigroups $U_j(t)$ and generators Λ_j, or transition probabilities $P_j(t, x_1, \cdot)$, on $(E_j, \mathscr{E}_j)$, $j = 1, 2$, a *Markovian coupling* is a Markov process with semigroup $\tilde{U}(t)$ and generator $\tilde{\Lambda}$, or transition probability $\tilde{P}(t; x_1, x_2, \cdot)$, on the product space $(E_1 \times E_2, \mathscr{E}_1 \otimes \mathscr{E}_2)$ having the marginality:

•

$$\tilde{P}(t;x_1,x_2,A_1\times E_2)=P_1(t,x_1,A_1)\quad \forall t\geq 0, x_1\in E_1, A_1\in\mathscr{E}_1,$$
$$\tilde{P}(t;x_1,x_2,E_1\times A_2)=P_2(t,x_2,A_2)\quad \forall t\geq 0, x_2\in E_2, A_2\in\mathscr{E}_2.$$

Or equivalently

$$\tilde{U}(t)G(x_1,x_2)=U_1(t)G(x_1),\quad \forall t\geq 0, x_1\in E_1, G\in B(\mathscr{E}_1),$$
$$\tilde{U}(t)G(x_1,x_2)=U_2(t)G(x_2),\quad \forall t\geq 0, x_2\in E_2, G\in B(\mathscr{E}_2),$$

where $B(\mathscr{E})$ is the set of all bounded $\mathscr{E}$-measurable functions. Here on the left hand side G is regarded as a bivariate function, although it depends on only one variable.

-
$$\tilde{\Lambda}G(x_1,x_2)=G(x_1)\quad \forall x_1\in E_1, G\in B(\mathscr{E}_1),$$
$$\tilde{\Lambda}G(x_1,x_2)=G(x_2)\quad \forall x_2\in E_2, G\in B(\mathscr{E}_2),$$

where G on the left hand side is regarded as above.

The main result concerning coupling of Markov chains, whose proof can be found in [3], is that the two parts given in the definition are equivalent under certain conditions, for instance if the two Markov chains take value in a finite state space. A Markovian coupling always exists: the simplest is the *independent coupling*. Consider a finite set $E=E_1=E_2$, in the notations of the above definition. The generator of the independent coupling is defines by

$$\tilde{\Lambda}G(x_1,x_2):=\Lambda_1 G(\cdot,x_2)(x_1)+\Lambda_1 G(x_1,\cdot)(x_2)$$

for any $G\in\mathscr{C}(E^2)$.

Hence, for fixed h and b (which will be omitted in the following writing) we consider *any* coupling operator $\widetilde{\Lambda}_h$ of the Markov chain $X^h(t)$ and itself, which gives the semigroup $\tilde{U}^h_t$. The coupled process will be called $(X^h(t),Y^h(t))$, although we consider the same process, to avoid misunderstandings. Then by (1.104) the process

$$M_g(t):=g(X^h(t),Y^h(t))-g(X^h(0),Y^h(0))-\int_0^t\tilde{\Lambda}g(X^h(s),Y^h(s))ds \qquad (1.105)$$

is a martingale for any $g\in\mathscr{C}(S(h)\times S(h))$.

We let

$$E^x[G(X^h(t))]:=E[G(X^h(t))|X(0)=x]=E[G(X^h(t,x))]=U^h_tG(x)$$

and

$$\tilde{E}^{x,y}[g(X^h(t),Y^h(t))]:=\tilde{U}^h_t g(x,y)=\tilde{E}\left[g(X^h(t),Y^h(t))|(X^h(0),Y^h(0))=(x,y)\right]$$

for any x and y in $S(h)$. Therefore the martingale M_g defined in (1.105) leads to

$$\tilde{E}^{x,y}[g(X^h(t),Y^h(t))] - g(x,y) = \tilde{E}^{x,y}\left[\int_0^t \tilde{\Lambda} g(X^h(s),Y^h(s))ds\right] \tag{1.106}$$

for any $t \geq 0$ and $g \in \mathscr{C}(S(h) \times S(h))$.

Considering $g_k(x,y) := x_k - y_k$ and $G_k(x) = x$, by means of definition 4 and identity (1.101), we have

$$\tilde{\Lambda} g_k(x,y) = \Lambda G_k(x) - \Lambda G_k(y) = f_k(x) - f_k(y)$$

and

$$\tilde{E}^{x,y}[g_k(X^h(\tau),Y^h(\tau))] - g_k(x,y) = E^x[X_k^h(\tau))] - x - E^y[X_k^h(\tau))] + y = F_k^h(x) - F_k^h(y)$$

by definition (1.60) of the drift. Hence equation (1.106) yields

$$F^h(x) - F^h(y) = \tilde{E}^{x,y}\left[\int_0^\tau \Big(f(X^h(s)) - f(Y^h(s))\Big)\, ds\right]$$

which, taking the ℓ^2-norm and applying Fubini's theorem and the Lipschitz continuity of f (1.66), becomes

$$||F^h(x) - F^h(y)|| \leq K \int_0^\tau \tilde{E}^{x,y} \left\| X^h(s) - Y^h(s) \right\| ds \tag{1.107}$$

for any coupling operator $\tilde{\Lambda}_h$, where K is defined in (1.95).

Now we need the following lemma, which is stated for instance in [4]:

Lemma 3. *Let U_t be a strongly continuous semigroup on E with generator Λ whose domain is $\mathscr{D}$, $\alpha \in \mathbb{R}$ be a constant and $f \in \mathscr{D}$. Then $U_t f \leq e^{\alpha t} f$ if and only if $\Lambda f \leq \alpha f$.*

Proof. ($\Leftarrow$) Let $\Lambda f \leq \alpha f$, then by Kolmogorov's equation

$$\frac{d}{dt} U_t f = \Lambda U_t f \leq \alpha U_t f.$$

Thus by Gronwall's lemma

$$U_t f \leq e^{\alpha t} U_0 f = e^{\alpha t} f.$$

($\Rightarrow$) Let $U_t f \leq e^{\alpha t} f$, then by definition of the generator

$$\Lambda f = \lim_{t \to 0} \frac{U_t f - f}{t} \leq \lim_{t \to 0} \frac{e^{\alpha t} - 1}{t} f = \alpha f.$$

Hence we want to apply this lemma to the coupling operator $\tilde{\Lambda}_h$ and the function $\rho(x,y) = ||x - y||$, in order to obtain an upper bound of the right hand side in (1.107). So we have to choose a particular coupling for which there exists $\alpha \in \mathbb{R}$ such that the condition

$$\tilde{\Lambda}_h \rho \leq \alpha \rho \tag{1.108}$$

is satisfied.

We use the so called *coupling of marching soldiers*, introduced by Chen in 1986 and whose description can be found for instance in [3]. This gives a Markov chain in $S(h) \times S(h)$ such that, if it is in (x,y), it jumps to

- $(x+z, y+z)$ at rate $\min\{q(x,x+z), q(y,y+z)\}$,
- $(x+z, y)$ at rate $[q(x,x+z) - q(y,y+z)]^+$,
- $(x, y+z)$ at rate $[q(x,x+z) - q(y,y+z)]^-$,

for any $z \in S(h)$, where $q(x,y)$ are the rates, i.e. the elements of the Q-matrix, of the Markov chain X^h. It is a Markov coupling, i.e. it satisfies definition 4 part 2, since $\min\{a,b\} + (a-b)^+ = a$ for any $a, b \geq 0$.

Set $h_{ij} = -he_i - he_j + he_{i+j}$ and $h^{ij} = -he_i + he_j + he_{i-j}$. Thus the generator of the marching coupling $\tilde{\Lambda}$ of the Markov chain X^h, whose generator is defined in (1.10), is given by

$$\begin{aligned}
\tilde{\Lambda}_h g(x,y) = & \frac{1}{h} \sum_{i,j} \min\{C_{ij}(x) x_i x_j, C_{ij}(y) y_i y_j\} \left[g(x+h_{ij}, y+h_{ij}) - g(x,y)\right] \qquad (1.109)\\
& + \frac{1}{h} \sum_{i,j} [C_{ij}(x) x_i x_j - C_{ij}(y) y_i y_j]^+ \left[g(x+h_{ij}, y) - g(x,y)\right] \\
& + \frac{1}{h} \sum_{i,j} [C_{ij}(x) x_i x_j - C_{ij}(y) y_i y_j]^- \left[g(x, y+h_{ij}) - g(x,y)\right] \\
& + \frac{1}{h} \sum_{i} \sum_{j<i} \min\{F_{ij}(x) x_i, F_{ij}(y) y_i\} \left[g(x+h^{ij}, y+h^{ij}) - g(x,y)\right] \\
& + \frac{1}{h} \sum_{i} \sum_{j<i} [F_{ij}(x) x_i - F_{ij}(y) y_i]^+ \left[g(x+h^{ij}, y) - g(x,y)\right] \\
& + \frac{1}{h} \sum_{i} \sum_{j<i} [F_{ij}(x) x_i - F_{ij}(y) y_i]^- \left[g(x, y+h^{ij}) - g(x,y)\right]
\end{aligned}$$

for any $g \in \mathscr{C}(S(h) \times S(h))$. Hence, calculating this generator (1.109) in the distance function $\rho(x,y) = ||x-y||$, we obtain

$$\tilde{\Lambda}_h \rho(x,y) \leq 3 \sum_{i,j} \left| C_{ij}(x) x_i x_j - C_{ij}(y) y_i y_j \right| + 3 \sum_{i} \sum_{j<i} \left| F_{ij}(x) x_i - F_{ij}(y) y_i \right|, \tag{1.110}$$

using the identity $a^+ + a^- = |a|$ for any real number a and the upper bound $||x-y-z|| - ||x-y|| \leq ||z|| \leq 3h$ whereas z can be either $-he_i - he_j + he_{i+j}$ or $-he_i + he_j + he_{i-j}$.

In order to get an estimate of the above equation, we shall consider the two functions $u, v : S \to \mathbb{R}^{N \times N} \cong \mathbb{R}^{N^2}$, where $N = N(h)$, defined by

$$\begin{aligned}
u_{ij}(x) &:= C_{ij}(x) x_i x_j \qquad (1.111)\\
v_{ij}(x) &:= F_{ij}(x) x_i \mathbb{I}_{]0,+\infty[}(i-j). \qquad (1.112)
\end{aligned}$$

So the right hand side in (1.110) is equal to

$$3||u(x)-u(y)||_{\ell^1}+3||v(x)-v(y)||_{\ell^1}.$$

The derivatives of u and v are given by

$$\frac{\partial u_{ij}}{\partial x_k}(x)=\frac{\partial C_{ij}(x)}{\partial x_k}x_i x_j+C_{ij}(x)\delta_{ik}x_j+C_{ij}(x)x_i\delta_{jk}$$

and

$$\frac{\partial v_{ij}}{\partial x_k}(x)=\left(\frac{\partial F_{ij}(x)}{\partial x_k}x_i+F_{ij}(x)\delta_{ik}\right)\mathbb{I}_{]0,+\infty[}(i-j).$$

Thus we apply the mean value theorem to get

$$\begin{aligned}||u(x)-u(y)||_{\ell^1}&+||v(x)-v(y)||_{\ell^1}\\&=\left\|\frac{\partial u}{\partial x}(z).\xi\right\|_{\ell^1}+\left\|\frac{\partial v}{\partial x}(w).\xi\right\|_{\ell^1}\\&\leq\sum_{ijk}\left|\frac{\partial u_{ij}}{\partial x_k}(z)\right||\xi_k|+\sum_{ik}\sum_{j<i}\left|\frac{\partial v_{ij}}{\partial x_k}(w)\right||\xi_k|\\&\leq C(1)||z||_{\ell^1}^2||\xi||_{\ell^1}+2C||z||_{\ell^1}||\xi||_{\ell^1}+F(1)||w||_{\ell^1}||\xi||_{\ell^1}+F||\xi||_{\ell^1}\end{aligned}$$

for any $x,y\in S(h)$ and for certain z and w in S, where $\xi=x-y$. The latter inequality follows from (1.31) and (1.32).

Therefore (1.110) becomes

$$\tilde{\Lambda}_h\rho(x,y)\leq 3(C(1)R^2+2CR+F(1)R+F)||x-y||_{\ell^1}$$

for any x and y in $S(h)$. If we use the estimate $||x-y||_{\ell^1}\leq\sqrt{N}||x-y||$ for any x and y in $S(h)$ then

$$\tilde{\Lambda}_h\rho(x,y)\leq M_2\sqrt{N(h)}||x-y||, \tag{1.113}$$

where $M_2:=3(C(1)R^2+2CR+F(1)R+F)$ is constant, which says that (1.108) holds with $\alpha:=M_2\sqrt{N(h)}>0$.

Thus we can apply lemma 3 to the marching coupling, so that

$$\tilde{E}^{x,y}\left\|X^h(s)-Y^h(s)\right\|\leq e^{\alpha s}||x-y||$$

for any $s\in[0,\tau]$ and $x\neq y\in S(h)$. Hence (1.107) becomes

$$||F^h(x)-F^h(y)||\leq K\int_0^\tau e^{\alpha s}||x-y||ds\leq K\tau e^{\alpha\tau}||x-y||,$$

which is (1.65) where L_1 is the function defined in (1.75):

$$||F^h(x)-F^h(y)||\leq K\tau e^{M_2\tau\sqrt{N(h)}}||x-y||.$$

1.4.3.1 Boundness of the drift

Applying (1.104) simply to the process X^h, without considering couplings, we obtain

$$F^h(x,b) = E\left[\int_0^\tau f(X^h(s,x,b))ds\right] \tag{1.114}$$

Hence from (1.114), by means of (1.64), we get also an upper bound for the drift:

$$||F^h(x,b)|| \le L_2\tau \tag{1.115}$$

for every $x \in S(h)$ and $b \in E$, where L_2 is defined in (1.93).

1.4.4 Bounds for Δ

We find an estimate for $E(\Delta^h_\pi(k)|X^h_\pi = x)$ where $\Delta^h_\pi(k)$ is the number of coalitions that perform a transition between time step $k\tau$ and $(k+1)\tau$. Because of Markovianity the above expectation is independent of k, so we can suppose $k = 0$. Hence we consider $E(\Delta^h|X^h(0) = x_0)$, where Δ^h is the number of coalitions that change their state between 0 and τ.

If the system is in x_0 in $t = 0$ there is an exponential clock of parameter $s(x_0,b)$ such that, when it clicks, the system changes its state, say it goes in x_1. Now there is an other exponential clock of parameter $s(x_1,b)$ such that, when it clicks, the system changes its state, say it goes in x_2. We repeat this procedure until we arrive at time τ. Note that Δ^h is then less or equal than the number of clicks that we get from 0 to τ.

Thus to estimate Δ^h we take an upper bound of $s(x,b)$ defined in (1.8)

$$s(x,b) = \sum_{i,j} n_i n_j hC_{ij}(x,b) + \sum_i n_i F_{ij}(x,b),$$

for any $x \in S(h)$. Using assumption (1.31) on C_{ij} and F_{ij} we have

$$s(x,b) \le Ch\sum_{i,j} n_i n_j + F\sum_i n_i = Ch\left(\sum_i n_i\right)^2 + F\sum_i n_i$$

for any x and b. According to (1.44) $\sum_i n_i \le N(h) \le \frac{R}{h}$, which gives

$$s(x,b) \le \frac{1}{h}(CR^2 + FR).$$

Hence for any $x \in S(h)$ and $b \in E$, $s(x,b)$ is bounded by a constant s^h that depends only on h:

$$s^h := \frac{1}{h}(CR^2 + F). \tag{1.116}$$

For a constant s the number of occurrences of clicks is known to be a Poisson process of intensity s. Thus for a constant s the number of clicks from 0 to τ is a random variable X with Poisson distribution of parameter $s\tau$. This implies in particulat that $E(X) = s\tau$ and $E(X^2) = s\tau(1+s\tau)$.

Hence the expectations for Δ^h are bounded by

$$E(\Delta^h|X^h(0) = x_0) \leq s^h\tau$$

and

$$E((\Delta^h)^2|X^h(0) = x_0) \leq s^h\tau(1+s^h\tau)$$

for any $x_0 \in S(h)$. Using (1.116) we find

$$E(\Delta^h|X^h(0) = x_0) \leq \frac{1}{h}\tau(CR^2+F)$$

and

$$\begin{aligned} E((\Delta^h)^2|X^h(0) = x_0) &\leq \frac{1}{h}\tau(CR^2+F)[1+\frac{1}{h}(CR^2+F)\tau] \\ &= \frac{1}{h^2}\tau(CR^2+F)[\tau(CR^2+F)+h]. \end{aligned}$$

Therefore (1.61) and (1.62) hold with the bounds specified in (1.73) and (1.74).

Acknowledgements Work (partially) supported by the PhD programme in Mathematical Sciences, Dipartimento di Matematica, Università di Padova (Italy) and Progetto Dottorati - Fondazione Cassa di Risparmio di Padova e Rovigo.

References

1. Bardi, M., Capuzzo Dolcetta, I.: Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Birkhauser (1997)
2. Caines, P.E., Huang, M., Malhamé, R.P.: Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. Commun. Inf. Syst. 6(3), 221-252 (2006)
3. Chen, M.F.: From Markov Chains to non equilibrium particle systems. World scientific (1992)
4. Chen, M.F.: Eigenvalues, Inequalities and ergodic theory. Springer (2005)
5. Chen, W., Huang, D., Kulkarni, A.A, Unnikrishnan, J., Zhu, Q., Mehta, P., Meyn, S., Wierman, A.: Approximate dynamic programming using fluid and diffusion approximations with applications to power management. In: Proceedings of the 48th IEEE Conference on Decision and Control held jointly with 28th Chinese Control Conference, CDC/CCC 2009, Shangai, September 2009, pp. 3575-3580 (2009)
6. Crandall, M.G., Lions, P.L.: Hamilton-Jacobi equations in infinite dimensions. II. Existence of viscosity solutions. J. Funct. Anal. 65, 360-405 (1986)
7. Finus, M., Rundshagen, B., Eyckmans, J.: Simulating a sequential coalition formation process for the climate change problem: first come, but second served? Ann. Oper. Res. 220, 5-23 (2014)

8. Gast, N., Gaujal, B., Le Boudec, J.Y.: Mean Field for Markov Decision Processes: from discrete to continuous optimization. IEEE Trans. Autom. Control 57, 2266-2280 (2012)
9. Inal, H.: Core of coalition formation games and fixed-point methods. Soc. Choice Welfare 45(4), 745-763 (2015)
10. Jouida, J.B., Sihem, K., Krichen, S., Klibi, W.: Coalition-formation problem for sourcing contract design in supply networks. Eur. J. Oper. Res. 257(2), 539-558 (2017)
11. Kallenberg, O.: Foundations of Modern Probability. 2nd ed. Springer (2002)
12. Karos. D.: Coalition formation in general apex games under monotonic power indices. Games Econom. Behav. 87, 239-252 (2014)
13. Kolokoltsov, V.N.: Hydrodynamic limit of coagulation-fragmentation type models of k-nary interacting particles. J. Stat. Phys. 115(5/6), 1621-1653 (2004)
14. Kolokoltsov, V.N.: Kinetic equations for the pure jump models of k-nary interacting particle systems. Markov Process. Relat. 12, 95-138 (2006)
15. Kolokoltsov, V.N.: Nonlinear Markov games on a finite state space (mean-field and binary interactions). Int. J. Stat. Probab. 1(1), 77-91(2012). doi:10.5539/ijsp.v1n1p77
16. Kolokoltsov, V.N.: The evolutionary game of pressure (or interference), resistance and collaboration. Math. Oper. Res. ISSN 0364-765X (2016, in press). doi:10.1287/moor.2016.0838
17. Krapivsky, P.L., Redner, S.: Organization of growing random networks. Phys. Rev. E. 63, 066123 (2001)
18. Lasry, J.M., Lions, P.L.: Mean field games. Jpn. J. Math. 2(1), 229-260 (2007)
19. Norris, J.: Cluster Coagulation. Comm. Math. Phys. 209, 407-435 (2000)
20. Papadimitriou, C.H., Tsitsiklis, J.N.: The complexity of optimal queuing network control. Math. Oper. Res. 24(2), 292-305 (1999)
21. Pushkin, D.O., Aref, H.: Bank mergers as scale-free coagulation. Physica A 336, 571-584 (2004)
22. Saichev, A., Malvergne, Y., Sornette, D.: Theory of Zipf's Law and Beyond. Springer, Berlin (2010)
23. Tsitsiklis, J.N., Van Roy, B.: An analysis of temporal-difference learning with function approximation. IEEE Tran. Autom. Control 42(5), 674-690 (1997)
24. Weese, E.; Political mergers as coalition formation: an analysis of the Heisei municipal amalgamations. Quant. Econ. 6(2), 257-307 (2015)