Università degli Studi di Padova

Dipartimento Matematica
Dipartimento di Matematica "Tullio Levi-Civita"

Doctoral programme in Mathematical Sciences
Mathematics curriculum
XXXVII cycle

On nonlinear systems of PDEs arising in the theory of large population differential games

Coordinator: Prof. Giovanni Colombo
Supervisor: Prof. Marco Cirant

PhD candidate: Davide Francesco Redaelli

Thesis written with the financial contribution of Fondazione Cassa di Risparmio di Padova e Rovigo

Abstract

This thesis is concerned with the study of stochastic differential games with many players, under structural hypotheses that differ from the classic ones of Mean Field Game theory. We focus on Nash equilibria and the systems of partial differential equations that describe them, within two main settings, namely *games with sparse interactions* and *Generalised Mean Field Games*.

In the first part of the thesis, we deal with network games with interactions between players governed by sparse graphs. We introduce the concept of *unimportance of distant players* and provide two precise declinations of it, one for open-loop and one for closed-loop games. Related implications are also investigated.

The main character of the second part is the Nash system, of parabolic equations of Hamilton–Jacobi–Bellman type, describing closed-loop equilibria. We make use of structural assumptions inspired by the unimportance of distant players to prove short-time existence and uniqueness for a class of Nash systems in infinitely many dimensions.

Afterwards, we enter the framework of Generalised Mean Field Games and, for some N-player nonsymmetric Nash systems under hypotheses of *semimonotonicity*, we prove certain a priori estimates historically known to be both hard to obtain and crucial for a rigorous derivation of the Master Equation directly from of the Nash system as N diverges. Making use of such estimates in this bottom-up approach to the large population limit of the Nash system, we conclude by proving that in our context suitable generalisations of both the Mean Field system and a weak form of the Master Equation can be obtained.

Contents

Notations

$\mathcal{A}_{R,L}$	see Definition 2.3.1.
$C^0(X)$	the space of all *bounded and continuous* functions on X.
$C^0_0([0,+\infty))$	the space of all functions in $f \in C^0([0,+\infty))$ such that $\lim_{r\to+\infty} f(r) = 0$.
$C^\gamma(X;Y)$	given $\gamma \in (0,1)$, the space of all *bounded and γ-Hölder continuous* Y-valued functions on X. $Y = \mathbb{R}$ is implied. If the domain involves $\mathbb{R}^\omega$, see Section 3.2 instead.
$C^k_{(\mathrm{loc})}(X)$	for $k \in \mathbb{N} \cup \{\infty\}$, the space of all infinitely many times differentiable functions on X, with (locally) bounded derivatives, zeroth-order included.
$\mathcal{X}(X;\mathcal{Y}(Y))$	given two Banach spaces, $\mathcal{X}(X)$ of functions on X, and $\mathcal{Y}(Y)$ of functions on Y, the corresponding *Bochner space* of functions on $X \times Y$.
$C^{1,2}_{\mathrm{loc}}(\mathbb{R}^n_T)$	the space of all *continuous* functions on $\mathbb{R}^n_T$ (with typical variables (t,x)) with *continuous derivatives* ∂_t, D and D^2. Derivatives are understood to be defined on $(0,T) \times \mathbb{R}^n$.
$C^{1,2}(\mathbb{R}^n_T)$	the space of all *bounded and continuous* functions on $\mathbb{R}^n_T$ (with typical variables (t,x)) with *bounded and continuous derivatives* ∂_t, D and D^2. Derivatives are understood to be defined on $(0,T) \times \mathbb{R}^n$.
$C^{m+\delta}_{(\mathrm{loc})}(X)$	given $m \in \mathbb{N}$ and $\delta \in (0,1)$, the classic high-order *Hölder space* on X; cf. [**57**]. As for the subscript loc, $u \in C^{m+\delta}_{\mathrm{loc}}(\mathbb{R}^d)$ if and only if $u \in C^{m+\delta}(\Omega)$ for any $\Omega \Subset \mathbb{R}^d$. If the domain involves $\mathbb{R}^\omega$, see Section 3.2 instead.
$C^{\frac{\delta}{2},\delta}_{(\mathrm{loc})}(X), C^{1+\frac{\delta}{2},2+\delta}_{(\mathrm{loc})}(X)$	given $\delta \in (0,1)$, the classic *parabolic Hölder spaces*; cf. [**57**]. As for the subscript loc, see the previous notation. If the domain involves $\mathbb{R}^\omega$, see Section 3.2 instead.
$C^{m+\gamma}_\beta$, $\\|\cdot\\|_{m+\gamma;\beta}$, ...	see Definitions 3.2.8 and 3.2.9.
∂_t	the *time derivative* (with respect to the typical variable t).
D , or D_x , or ∂_x	the *gradient*, or *space derivative* (with respect to the typical variable x).
D_j , or ∂_{x^j}	the derivative with respect to the typical j-th coordinate x^j.
$D^2, D^2_{ij}, \partial^2_{xy}, D^3, D^3_{ijk}$	spatial derivatives of second and third order, in analogy with the notations for the first order.

Δ, or Δ_x	the *Laplacian* (with respect to the typical variable x).
D^α	given $\alpha = (\alpha_0, \dots, \alpha_{N-1}) \in \mathbb{N}^N$, the *multi-index* notation for derivatives, namely $D_1^{\alpha_1} \cdots D_N^{\alpha_N}$.
$\frac{\delta}{\delta m}$	the *flat derivative* (with respect to the typical probability measure variable m); cf. [**23**, Definition 2.2.1].
D_m	the *intrinsic derivative* (with respect to the typical probability measure variable m); cf. [**23**, Definition 2.2.2].
$\mathcal{D}$	see (4.2.4).
$\mathbb{D}_r(z)$	the open complex disc of radius r about z. We omit the centre when it is 0 and the radius when it is 1; hence, e.g., $\mathbb{D}_r = \mathbb{D}_r(0) = r\mathbb{D}$.
$\deg^- v,\ \deg^+ v$	the *indegree* and the *outdegree* of the vertex v.
$\deg v$	abridged notation for $\deg^- v$.
$\operatorname{diag} x$	given $x = (x^i)_{i \in [\![n]\!]} \in \mathbb{R}^n$, the $n \times n$ diagonal matrix having x^i as the entry (i,i).
div, or div_x	the *divergence* (with respect to the typical variable x).
δ_x	the *Dirac delta* distribution centred at x.
$\mathbb{E}$	the probabilistic *expectation*.
$v \sim v'$	$(v, v') \in E$.
$f\big\|_y^x$	given a function f, the difference $f(x) - f(y)$.
$A \geqslant 0$	given an $n \times n$ real matrix $A = (A_{ij})_{i,j \in [\![n]\!]}$ (not necessarily symmetric), $\sum_{i,j \in [\![n]\!]} A_{ij} \xi^i \xi^j \geqslant 0$ for all $\xi \in \mathbb{R}^n$.
I_n	the $n \times n$ *identity matrix*.
I	the matrix I_{Nd}.
id_X	the *identity map* on X.
J_n	the $n \times n$ *matrix of ones*.
J	the matrix $J_N \otimes I_d$.
$\mathcal{L}$	see (4.2.6).
$\mathcal{L}_X$	given a random variable X, on a probability space with probability $\mathbb{P}$, the *law* of X, namely $X_\sharp \mathbb{P}$.
$L^p(X; Y)$	for $p \in (0, \infty]$ and Y Banach, the space of Y-valued measurable functions f on a measure space $(X, \mathfrak{M}, \mu)$ such that $\\|f\\|_p := \int_X \\|f\\|_Y^p \,\mathrm{d}\mu < \infty$ (for $p \neq \infty$) or $\\|f\\|_\infty := \operatorname{ess\,sup}_X \\|f\\|_Y < \infty$ (for $p = \infty$). $Y = \mathbb{R}$, with Euclidean metric, is implied.
$\ell^p(X; Y)$	$L^p(X; Y)$ when X is countable, $\mathfrak{M}$ is the discrete σ-algebra and μ is the counting measure.
$\lesssim$	less than or equal to, up to a positive multiplicative constant.
$\mathrm{Lip}_{\mathrm{loc}}(X)$	the space of all *locally Lipschitz continuous* functions on X.
m_x	given a vector $x = (x^i)_{i \in [\![n]\!]}$, the *empirical measure* $\frac{1}{n} \sum_{i \in [\![n]\!]} \delta_{x^i}$.
$\mathbb{N}$	the set of all *natural numbers*, including 0.

$\lvert\cdot\rvert$	the *Euclidean norm* when applied to a vector, the *Frobenius norm* for matrices and, more generally, the *full tensor contraction* for tensors (and according to this, inner products between tensors are understood as tensor contractions); for example, $\lvert D^3_{ijk}u^j\rvert^2 = \sum_{h_i,h_j,h_k\in[\![d]\!]}\lvert D^3_{x^{ih_i}x^{jh_j}x^{kh_k}}u^j\rvert^2$ if $x^i, x^j, x^k \in \mathbb{R}^d$.
$\lvert\cdot\rvert_p$	the p-norm $\lVert\cdot\rVert_p$ on $\ell^p(\mathbb{N};\mathbb{R}^d)$, with $\mathbb{R}^d$ equipped with the Euclidean metric; that is, $\lvert x\rvert_p^p := \sum_{j\in\mathbb{N}}\lvert x^j\rvert^p$. The same notation is used for the p-norm of $x\in(\mathbb{R}^d)^n$, $n\in\mathbb{N}$, given the natural embedding $x\mapsto(x,0,\dots)$ into $\ell^p(\mathbb{N};\mathbb{R}^d)$.
$\lVert\cdot\rVert_X$	the norm on the space X.
$\lVert f\rVert_\infty$	given a function f defined on X, its *supremum norm*, namely $\sup_{x\in X}\lvert f(x)\rvert$.
$\lVert f\rVert_{\infty;I}$	given a function f defined on $\mathbb{R}\times X$ and an interval $I\subset\mathbb{R}$, $\sup_{t\in I}\lVert f(t,\cdot)\rVert_\infty$.
$[f]_\gamma$	the *Hölder seminorm* of order γ of the function f. For functions defined on $\mathbb{R}^\omega$, see Section 3.2.
$\lVert x\rVert^i$	the *i-th weighted norm* of the vector $x\in X^N$, defined as $\left(\lvert x^i\rvert^2+\frac{1}{N}\sum_{j\in[\![N]\!]\setminus\{i\}}\lvert x^j\rvert^2\right)^{\frac{1}{2}}$.
$\lVert\cdot\rVert_{\infty;\beta,\alpha}$, $[\cdot]_{\gamma;\beta,\alpha}$	see Definition 3.2.3.
1_n	the *all-ones vector* in $\mathbb{R}^n$.
$x\cdot y$, or $\langle x,y\rangle_X$	the *Euclidean inner product* between the vectors $x,y\in X$. The subscript X can be implied if it is clear from the context.
$\otimes$	the *outer product* between vectors, or the *Kronecker product* between matrices, or the (tensor) *product* between measures.
$\mathcal{P}(X)$	the space of *probability measures* on X.
$\mathcal{P}_p(X)$	the *Wasserstein space* of order $p\in[1,\infty)$; cf. [**78**, Definition 6.4].
P_v	given a vector $v\neq 0$ in a Hilbert space, the projection onto the direction of v, namely $\frac{v}{\lvert v\rvert}\otimes\frac{v}{\lvert v\rvert}$.
Q	see below (4.2.4).
$\Pi(\mu,\nu)$	the set of all *couplings* of the measures (μ,ν).
π_{X^i}	given a product space $X=\prod_{i\in I}X^i$, the projection of X onto X^i.
$\mathbb{R}^n_T$	the set $[0,T]\times\mathbb{R}^n$.
$\mathbb{R}^\omega$	the space $\mathbb{R}^\mathbb{N}$ endowed with the *product topology*.
$[\![n]\!]$	given $n\in\mathbb{Z}_+$, the set $\{0,\dots,n-1\}$.
$[m]_n$	given $m\in\mathbb{Z}$ and $n\in\mathbb{Z}_+$, the unique natural number $r\in[\![n]\!]$ such that $r=m \bmod n$.
$\mathcal{Q}_r$	see (2.2.28).
$\mathbb{R}_+$	the set of all *positive real* numbers.
$\#S$	the *cardinality* (or *counting measure*) of the set S.

$f_\sharp\mu$	the *push-forward* (or *image*) *measure* of the measure μ via the map f.
S_n	the group of all *permutations* on $[\![n]\!]$.
x^σ	given a vector $x = (x^i)_{i\in[\![n]\!]}$ and a permutation $\sigma \in S_n$, the vector $(x^{\sigma(i)})_{i\in[\![n]\!]}$.
$\operatorname{spt}\mu$	the *support* of the measure μ.
$f \star g$	given two functions f, g defined on a measure space (X, μ), the *convolution* between f, g, namely $\int_X f(y)g(\cdot - y)\,\mu(\mathrm{d}y)$.
$x \star_n y$	given two vectors $x = (x^i)_{i\in[\![n]\!]}, y = (y^i)_{i\in[\![n]\!]} \in \mathbb{R}^n$, the *cyclic discrete convolution* of order n between x, y, namely $\sum_{i\in[\![n]\!]} x^i y^{[\cdot - i]_n}$.
$X \Subset Y$	$X \subseteq Y$, bounded and with closure also contained in Y.
$\mathscr{S}(X)$	the space of all *real symmetric linear operators* on X.
$\mathscr{S}(n)$	the space of all *real symmetric* $n \times n$ *matrices.* $\mathscr{S}(\mathbb{R}^n)$ and $\mathscr{S}(n)$ are identified through the canonical basis on $\mathbb{R}^n$.
$\operatorname{Tr} A$	the *trace* of the matrix A.
A^T	the *transpose* of the matrix A.
$\triangle_h$	if h is a function, see (4.2.7); if $h \in \mathbb{Z}_+$, it is the *diagonal* $\{(x, y) \in \mathbb{R}^h \times \mathbb{R}^h : x = y\}$.
$x \vee y$, $x \wedge y$	$\max\{x, y\}$, $\min\{x, y\}$.
$(x : y)$	given $x = (x^i)_{i\in[\![n]\!]} \in X^n$, $y = (y^i)_{i\in[\![n]\!]} \in Y^n$, the vector $((x^i, y^i))_{i\in[\![n]\!]} \in (X \times Y)^n$.
x^{-i}	given a vector x, the vector obtained by *removing* the i-th coordinate from x.
(x^{-i}, y)	given two vectors x, y, the vector obtained from x by *replacing* its i-th coordinate with y. — The order is important, in the sense that we specify first the vector x and then the replacement y; both the following *not equivalent* notations will occur: given a function h, $h(x^{-i}, y) = h(x^1, \dots, x^{i-1}, y, x^{i+1}, \dots, x^N)$, but $h(y, x^{-i}) = h(y, x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^N)$.
$\hat{x}$	given a vector x, see Section 5.1.
W_p	the *Wasserstein distance* of order $p \in [1, \infty)$; cf. [**78**, Definition 6.1].
$W_p^{1,2}$	the standard *parabolic Sobolev space* with exponent $p \in [1, \infty)$; cf. [**58**, Chapter 2, Section 2].
$\mathbb{Z}_+$, $\mathbb{Z}_-$	the sets of all *positive integers* and of all *negative integers.*

Introduction

Contents

This thesis contains new results on stochastic differential games (henceforth simply *games*) from the viewpoint of partial differential equations (PDEs). The *trait d'union* between the different works collected herein is the interest in studying games under assumptions outside or beyond the Mean Field Game (MFG) canon, with the intent of identifying substantial dissimilarities with or attainable generalisations of that renowned standard.

The introduction of MFG theory, independently by Lasry and Lions [**65**] and Huang, Caines and Malhamé [**51**], has been the most influential contribution to the rapid development of the study of games with many players in the last two decades. The theory provides an effective paradigm to deal with dynamic games with many players that are both indistinguishable from one another and individually negligible as the number of players grows; we refer to the book [**25**] and the lecture notes [**24**] for a recent account of it.

If the foundational assumptions of MFGs are abandoned, or at least eased, one enters the broader framework of network games. In particular, the last few years have witnessed an increasing interest in the understanding of the large population limits of Nash equilibria among players whose interactions are described by graphs.

Broadly speaking, when the number of interactions is large, the MFG description, or a suitable generalisation based on the notion of *graphon*, effectively characterises large population limits: see for instance [**3, 8, 19, 37, 63**] and references therein.

On the other hand, when the underlying network structure is sparse, few results are available in the literature; for instance, the reader can have a look at [**61**] and references therein. While in the dense regime one expects all the players to have an individual negligible influence on a given player, because their running costs involve cumulative functions of a number of variables that diverges with the number of players, in sparse regimes, only most of the players should have a small impact, just because they are far enough with respect to the graph distance. This mechanism of independency between distant players, in the sense of an underlying graph, has been first observed in [**62**], by means of correlation estimates.

The first part of the thesis concerns the study of such *games with sparse interactions*, with particular interest in the equations describing Nash equilibria in the limit of infinitely many players. We highlight two crucial features in contrast to those of MFGs and deriving from the aforementioned *unimportance of distant players*.

As a first illustrative setting, we consider open-loop equilibria. Under the hypotheses of MFG theory, one exploits the averaging structure of the costs to reduce the prospective infinitely many equations describing Nash equilibria in the large population limit to the familiar two-equation Mean Field system, whereas in games with sparse interactions we have no hope to be able to exploit any aggregate effect coming into play as the number of player grows. On the contrary, we shall show that, if one wishes to find, with a small error, the optimal trajectory of a given player, a reduction of the complexity of the system to be studied is possible by going in the opposite direction; that is, by neglecting distant players and thus considering a new game governed by just a subgraph of the original one. This is presented in Chapter 1.

Considering instead closed-loop equilibria, the canonical object describing them is the Nash system, a notoriously strongly coupled, and thus intricate, system of parabolic equations of Hamilton–Jacobi–Bellman (HJB) type. As described in [**23**] expanding ideas of [**68**], within the MFG scenario the natural limit of the Nash system is provided by the fabled Master Equation, a parabolic equation set in the space of probability measures. On the other hand, under hypotheses of sparse interactions, we have quantified the mutual influence of players by estimating the derivatives of the solutions to the Nash system, consequently showing that a formally-written infinite-dimensional Nash system can be well-posed. This is introduced in Chapter 2.

The second part of the thesis is focused on the study of the Nash system.

Local existence and uniqueness for a rather general infinite-dimensional Nash system are proven in Chapter 3, adapting ideas developed in the previous chapter within the linear-quadratic (LQ) framework. Contextually, we also prove a general linear result, providing a priori estimates, stable with respect to the dimension, for transport-diffusion equations whose drifts (and their derivatives) enjoy appropriate decay properties.

In Chapter 4, again inspired by results in the LQ setting, we exit the realm of sparse interactions and recover some structure of Mean Field type: under *Generalised Mean Field* hypotheses (also referred to as *Mean-Field-like*), for a class of nonsymmetric Nash systems with *semimonotonicity* we prove a priori estimates enjoying the stability with respect to the number of equations first prospected in [**68**] for the standard symmetric case and expected to allow to pass the system to the infinite population limit and obtain the Master Equation.

Finally, Chapter 5 is devoted to the use of our a priori estimates to deal with such a passage to the limit. We obtain both a weak form of the Master Equation and the Mean Field system expected in our generalised MFG context.

Exordia of the chapters

Chapter 1 (*d'après* [**30**])**.** We consider a graph $\Gamma := (V, E)$ with $\#V = N \in \mathbb{N} \cup \{\infty\}$, which can be directed but is assumed to have no loops. We recall that this means that the existence of the edge $(v, v') \in E$ does not imply that of $(v', v) \in E$, while $(v, v) \notin E$ for any vertex $v \in V$.

The indegree and the outdegree of a vertex are respectively defined as $\deg^- v := \#\{v' \in V : v' \sim v\}$ and $\deg^+ v := \#\{v' \in V : v \sim v'\}$, where $v' \sim v$ if and only if $(v', v) \in E$; however, in the following, since the former will appear quite often, we will use the lighted notation $\deg v := \deg^- v$. We shall assume that $\deg v > 0$ for all $v \in V$ and, if V is countable, $\sup_{v \in V} \deg v < \infty$.

The main object of our study is a system of N HJB and N Fokker–Planck–Kolmogorov (FPK) equations in $\mathbb{R}^d_T := [0, T] \times \mathbb{R}^d$, which we call the *graph-game system*, of the form

$$
\text{(I)} \qquad \begin{cases} -\partial_t u^v - \Delta u^v + H^v(t, Du^v) = F^v(t, x) \\ \partial_t m^v - \Delta m^v - \operatorname{div}(\partial_p H^v(t, Du^v) m^v) = 0 \qquad v \in V\,. \\ u^v(T, \cdot) = G^v, \;\; m^v(0) = m_0^v\,, \end{cases}
$$

The unknowns are the (value) functions u^v and the measure flows m^v; the Hamiltonians $H^v(t, p)$, the data (or *cost functions*) $F^v(t, x)$ and $G^v(x)$, and the initial distributions m_0^v are given. In particular, we consider

$$
F^v(t, x) := \int_{(\mathbb{R}^d)^{\deg v}} f^v(t, x, y)\, \mu_t^v(\mathrm{d}y)\,, \quad G^v(x) := \int_{(\mathbb{R}^d)^{\deg v}} g^v(x, y)\, \mu_T^v(\mathrm{d}y)\,,
$$

for some functions $f^v : [0, T] \times \mathbb{R}^d \times (\mathbb{R}^d)^{\deg v} \to \mathbb{R}$ and $g^v : \mathbb{R}^d \times (\mathbb{R}^d)^{\deg v} \to \mathbb{R}$, where $\mu_t^v \in \mathcal{P}((\mathbb{R}^d)^{\deg v})$ is the product measure having $m_t^{v'}$, $v' \sim v$, as its marginals.

Such a graph-game system naturally arises to describe open-loop Nash equilibria for a differential game with interactions between the players governed by the graph Γ. Indeed, consider an N-player game where players are labelled with $v \in V$

and the state of the player v is governed by the following $\mathbb{R}^d$-valued SDE on $[0,T]$:

$$\mathrm{d}X_t^v = b^v(t,\alpha_t^v)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^v\,, \tag{II}$$

with given $X_0^v = Z^v$. The B^v are independent Brownian motions, defined on a filtered probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$, $\mathbb{F}$ being the completion of the natural filtration of the Brownian motion; the controls α^v are progressively measurable processes with values in $\mathbb{R}^d$; the drift b^v is a given map $[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$. Player v aims at minimising the cost functional

$$J^v(\alpha) := \mathbb{E}\biggl[\int_s^T \Bigl(L^v(t,\alpha_t^v) + f^v\bigl(t,X_t^v,(X_t^{v'})_{v'\sim v}\bigr)\Bigr)\,\mathrm{d}t + g^v\bigl(X_T^v,(X_T^{v'})_{v'\sim v}\bigr)\biggr]\,,$$

where $L^v\colon [0,T]\times A\to\mathbb{R}$ is given. Note that such a cost depends (directly) only on the behaviour of the neighbours of player v; that is, players that are adjacent to it according to the graph Γ. Also note right away that if $\deg v = 0$, then player v is by no means connected to the other, so it will be playing alone an optimisation problem; nevertheless, such players will be of no interest in our study of the game, whence our assumption that $\deg v > 0$.

The Hamiltonian of player v is defined as $H^v(t,p) := -\inf_{a\in\mathbb{R}^d}\bigl(p\cdot b^v(t,a) + L^v(t,a)\bigr)$. Hence, if for any (t,p) there exists a unique $a^v(t,p)\in\mathbb{R}^d$ such that $H^v(t,p) = -p\cdot b^v(t,a^v(t,p)) - L^v(t,a^v(t,p))$, and that the map a^v is continuous (for example, this is the case if $p\cdot b^v(t,a)+L^v(t,a)$ is continuous in all variables, and is C^2, strictly convex and coercive in a), then the envelope theorem yields $\partial_p H^v(t,p) = -b^v(t,a^v(t,p))$. Under these circumstances, the choice $\alpha_t^v = \alpha_t^{*v} := -\partial_p H^v(t,Du^v(t,X_t^v))$ in (II), with u^v solving system (I), is expected to provide the open-loop equilibria of the game (see Remark 1.1.1 for further discussion). In particular, one has $m_0^v = Z^v{}_\sharp\mathbb{P}$ and $\mu_t^v := (X_t^{v'})_{v'\sim v\sharp}\mathbb{P}$, when X^v solves (II) with optimal drift α^{*v}.

If Γ is complete (that is, $E = V\times V\setminus\{(v,v)\,:\,v\in V\}$) and f^v and g^v depend on the average of $(X^v)_{v'\sim v}$, then we are in the setting of MFG theory; in particular, the MFG paradigm presumes that we have an increasing sequence of graphs $(\Gamma^N)_{N\in\mathbb{N}}$ (that is, Γ^N is a subgraph of Γ^{N+1}) with $\min_{v\in V^N}\deg v = N\to\infty$ as $N\to\infty$ (and actually much more that that). On the contrary, here we are interested in addressing the antipodal problem when $\sup_{N\in N}\sup_{v\in V^N}\deg v<\infty$.

The crucial difference with this dense regime is the loss of the negligibility assumption, and thus basically the loss of a fruitful ground to have a Law of Large Numbers: if the costs of the players do not depend on cumulative functions of a diverging number of variables, one is led to expect that most of the players should have a small influence on a given one, with no cumulative effect arising from their interactions in the large population limit. In other words, given a player, we expect players that far from it with respect to the graph distance to be asymptotically *unimportant* in determining the optimal distribution of that player.

Here we focus on open-loop games governed by fairly general sparse graphs, with the goal of showing the *unimportance of distant players* under the following perspective: in order to play almost optimally, a given player can neglect all players that are far enough from it and assume to be taking part in a new *reduced* game where interactions are modelled according to a subgraph of the original one.

Our main result will be Theorem 1.2.8, yet we informally state here an immediate consequence of it (precisely formalised as Corollary 1.2.9), which we believe to be illustrative of the meaning of the preceding paragraph.

THEOREM I. *Let (m, u) solve (I) with infinite-dimensional underlying graph $\Gamma = (V, E)$ and $m_0 \in \mathcal{P}_2(\mathbb{R}^d)^V$. Under suitable structural assumptions, for any $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that, for any $v \in V$, one has*

$$\sup_{t\in[0,T]} W_2(m_t^v, \hat{m}_t^v) < \varepsilon$$

if $\hat{m}$ solves (I) with graph $\Gamma_{v;k}$ (which is the minimal subgraph of Γ in which v is k-stable with respect to Γ, according to Definition 1.2.2).

The subgraph $\Gamma_{v;k}$ is basically the smallest subgraph of Γ obtained by removing all vertices that are more than k edges apart from v.

The idea of the proof is rather simple: by means of a synchronous coupling, we show that (objects very close to) the Wasserstein distances of two distributions solving the graph-game system with two different underlying graphs obey a recursive inequality; then, analysing such inequality, we are able to provide quantitative decay of that "distance" with respect to k, which is the radius of the neighbourhood of a vertex we wish, or need, to keep when performing our dimensional reduction.

We have implicitly taken into account that not only we are losing the negligibility assumptions of MFG theory, also we are making no symmetry assumptions; therefore, the localisation provided by $\Gamma_{v;k}$ depends indeed on v, in general. It clearly follows that in order to determine approximate optimal trajectories for all the players, one needs to localise around each of them.

On the other hand, if Γ is vertex-transitive and the players have identically distributed initial states (that is, m_0^v is independent of $v \in V$), then system (I) is invariant under automorphisms of Γ (as observed and exploited in [**62**]), which implies that any player is representative for the game equilibrium. In this setting, the localisation $\Gamma_{v;k}$ is independent of v, so our result says that, instead of solving a system of infinitely many HJB and FPK equations, one can only solve (making a small error) one system of N HJB and N FPK equations, with N depending on Γ and k.

Our results mentioned so far are genuine a priori estimates, valid for solutions $(u, m) = ((u^v, m^v))_{v\in V}$ to system (I) with $Du^v \in C^{1,2}_{\mathrm{loc}}(\mathbb{R}^d_T)$ and sublinear (that is, $|Du^v(t,x)| \lesssim 1 + |x|$, with implied constant independent of $(t,x) \in \mathbb{R}^d_T$), and $m^v \in C^0([0,T];\mathcal{P}_2(\mathbb{R}^d))$.

For the sake of a more complete discussion, in Section 1.3, we will prove the existence of unbounded solutions fitting the aforementioned requirements, in a couple of cases compatible with our main assumptions, which we are going to present here below. Let us also point out that those will be general existence results which do not rely on the information provided by the unimportance of distant players, and thus Section 1.3 is independent of Section 1.2. A different approach to the well-posedness of a graph-game Nash system is instead adopted in Chapter 3, where the expected behaviour of the solution according to the results in Chapter 2 is exploited as a starting point to prove existence.

Chapter 2 (*d'après* [**32**])**.** We consider an N-player game, indexed by $i \in [\![N]\!]$, where the state X^i of the i-th player evolves according to the $\mathbb{R}^d$-valued SDE

$$\mathrm{d}X^i_t = \alpha^i_t \,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B^i_t\,.$$

Its cost, in the fixed time horizon $[0,T]$, is given by

$$J^i(\alpha) = \frac{1}{2}\,\mathbb{E}\bigg[\int_0^T \big(|\alpha^i|^2 + \langle F^i X, X\rangle\big)\,\mathrm{d}t + \langle G^i X_T, X_T\rangle\bigg]\,,$$

for some $F^i = f^i \otimes I_d$ and $G^i = g^i \otimes I_d$, with $f^i \in C^0([0,T];\mathscr{S}(N))$ and $g^i \in \mathscr{S}(N)$. The B^i are independent $\mathbb{R}^d$-valued Brownian motions, and the α^i are closed-loop controls in feedback form, that is $\alpha^i = \alpha^i(t,X_t)$.

It is known (see for instance [**25**, Section 2.1.4], or [**46**]) that the value functions of the players, $u^i = u^i(t,x)$ with $i \in [\![N]\!]$, $t \in [0,T]$ and $x = (x^0,\dots,x^{N-1}) \in (\mathbb{R}^d)^N$, solve the Nash system of Hamilton–Jacobi PDEs

$$\text{(III)}\qquad \begin{cases} -\partial_t u^i - \Delta u^i + \frac{1}{2}|D_i u^i|^2 + \sum_{j\neq i} D_j u^j D_j u^i = \bar{F}^i \\ u^i(T,\cdot) = \bar{G}^i \end{cases}$$

where $\bar{F}^i = \frac{1}{2}\langle F^i\cdot,\cdot\rangle$, $\bar{G}^i = \langle G^i\cdot,\cdot\rangle$ and $i \in [\![N]\!]$. The equilibrium feedbacks are then given by $\alpha^i = -D_i u^i$. Since the dynamic of each player is linear, and the costs are quadratic in the state and control variables, it is well-known that the previous system can be recast into a system of ODEs of Riccati type, by making the ansatz that u^i are quadratic functions of the states. Here, we look for conditions on the solvability of such system, and we focus in particular on properties that are *stable* as the number of players goes to infinity, with the aim of addressing the limit problem with infinitely many players, whenever possible. The Laplacian appears in (III) by the presence of the independent noises B^i, but it will not actually play any relevant role in our analysis. We keep it since the purely deterministic system is not known to be well posed, beyond the linear quadratic setting.

Our framework is somehow similar to that addressed in [**62**]. The linear-quadratic setting is considered, and under some symmetry assumptions on the underlying graph, Nash equilibria are computed explicitly (exploiting also the running costs $\bar{F}^i$ being identically zero, and $\bar{G}^i$ having a specific structure). Then, a probabilistic information on the covariance of any two players' equilibrium state processes is derived. A main goal in our work is to derive an analogous information in more analytic terms; that is, we wish to quantify the influence of the j-th player on the i-th one by estimating $D_j u^i$, with the perspective of developing some ideas that could be applied also beyond the linear-quadratic framework.

We now describe our results more in detail. The first part of the chapter is devoted to the analysis of a sort of sparse regime with a special structure, namely *shift-invariance*, where basically f^i and g^i coincide with f^{i-1} and g^{i-1}, respectively, after the permutation of variables $x^i \mapsto x^{i-1}$. Most importantly, we assume players to have *nearsighted interactions*, that means

$$|f^i_{hk}|, |g^i_{hk}| \lesssim \beta_{h-i}\beta_{k-i} \tag{IV}$$

for all h, k, where $(\beta_k)_k$ is a suitable sequence in $\ell^1(\mathbb{Z})$. Since β_{k-i} decays as $|k-i| \to \infty$, this means that f^i_{hk}, which quantifies the influence of the h-th and k-th player on the i-th one, decreases as $|h-i|$ and $|k-i|$ increase. A prototypical example is given by the bi-directed chain, where f^i_{hk} is a sparse matrix with zero entries except for $|h-i| \leqslant 1$ and $|k-i| \leqslant 1$, so that the i-th player cost depends only on the $(i+1)$-th and the $(i-1)$-th state. See also [**42**] for results on models that build upon a similar structure.

The first statement we get, which sums up Theorems 2.2.11 and 2.2.12, is the following.

Theorem ii. *There exists $T^* > 0$ such that if $T \leqslant T^*$ then for any $N \in \mathbb{N} \cup \{\infty\}$ there exists a smooth solution to system* (2.1.2) *such that, for any i, j and $m \in \mathbb{N}$,*

$$\left\| \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^m D_{hk} u^i \right\|_\infty \lesssim \beta_{h-i}\beta_{k-i}.$$

Note in particular that we get existence of a smooth classical solution to the infinite-dimensional Nash system (iii) with $N = \infty$, $i \in \mathbb{Z}$. We stress that the key issue in the analysis of our problem is that, despite the cost functions f^i, g^i may depend on very few variables, the system itself is strongly coupled by the transport terms $\sum_{j \neq i} D_j u^j D_j u^i$, which become in fact series when $N = \infty$. The closed-loop structure of equilibria forces the equilibrium feedbacks $\alpha^j = -D_j u^j$ to be strongly "nonlocal", that is, they depend on the full vector state x. Hence, decay estimates on $D_j u^i$ which are stable as N increases are crucial to pass to the limit $N \to \infty$. These are obtained here by a careful choice of β: below, we will use the terminology *self-controlled for the discrete convolution*, or briefly *c-self-controlled* (where the "c" stands for "convolution"; see Definition 2.2.3), that fits well with the structure of cyclic discrete convolution appearing in our problem. From the game perspective, the equilibrium feedbacks α^j turn out to be almost "local", in the sense that the influence of "far" players is still negligible in the sense explained above. In other words, despite the strong coupling given by the full information structure of closed-loop equilibria, the property of *unimportance of distant players* is observed (see Remark 2.2.13 for further discussions).

While the shift-invariance condition can actually be dropped (see Section 2.2.3), one of the main restrictions of the previous result is that it guarantees short-time existence only. Note that, even with a finite number of players N, Riccati systems may fail to have long time solutions in general; therefore we look for further conditions on f^i and g^i such that existence holds for any time horizon T, and independently on N. To achieve this goal, we strengthen the previous assumption on nearsighted interactions, that now become of *strongly gathering* type, and require further *directionality* conditions. Section 2.2.4 contains precise definition of this notion and examples. The main existence result is stated in Theorem 2.2.24. Here, we exploit the possibility to relate a solution to system (III) to a flow of generating functions, which works well when $N = \infty$ but has no clear adaption to the finite N setting.

Within the special setting of systems with cost of strongly gathering type and directionality, we are able to push further our analysis and study the long time limit $T \to \infty$, that is, we show that the value functions u^i converge to solutions of the ergodic problem as the time horizon goes to infinity. To complete this program, estimates on solutions of the Riccati system are obtained at the level of the game with $N = \infty$ players, uniformly with respect to T. Theorem 2.5.1 is the main result of this second part of the chapter, which we conclude by showing that equilibria of the infinite-dimensional game provide ε-Nash closed-loop equilibria of suitable N-player games, see Section 2.3.

Chapter 3 (*d'après* [**77**])**.** We consider the following system of infinitely many backward parabolic differential equations of HJB type, which we refer to as an *infinite-dimensional Nash system*:

$$\text{(V)} \qquad \begin{cases} -\partial_t u^i - \sum_{jk} A^{jk}(t,x) D^2_{jk} u^i + H^i(t,x,\mathcal{D}u) + \sum_{j \neq i} \partial_{p^j} H^j(t,x,\mathcal{D}u) D_j u^i = 0 \\ u^i(T,\cdot) = G^i\,, \qquad i \in \mathbb{N}\,, \end{cases}$$

where $\mathcal{D}u := (D_j u^j)_{j\in\mathbb{N}}$, set in $\mathbb{R}^\omega_T := [0,T] \times \mathbb{R}^\omega$, where the space $\mathbb{R}^\omega$ is $\mathbb{R}^\mathbb{N}$ equipped with the product topology.

The unknowns are the *value functions* $u^i \colon \mathbb{R}^\omega_T \to \mathbb{R}$; the data are the *horizon* T, the *diffusion* $A = (A^{jk})_{j,k\in\mathbb{N}} \colon [0,T] \to \mathbb{R}^{\omega\times\omega}$, the *Hamiltonians* $H^i \colon \mathbb{R}^\omega_T \times \mathbb{R}^\omega \to \mathbb{R}$ and the *terminal costs* $G^i \colon \mathbb{R}^\omega \to \mathbb{R}$.

The notation $\mathbb{R}^{\omega\times\omega}$ indicates infinite-dimensional square matrices, or equivalently the space $(\mathbb{R}^\omega)^\omega$, in analogy with the notation $\mathbb{R}^{N\times N}$ for $N \times N$ matrices.

The typical element of $[0,T]\times\mathbb{R}^\omega\times\mathbb{R}^\omega$ is denoted by (t,x,p), with coordinates x^j and p^j, $j\in\mathbb{N}$.

Derivatives are understood in the sense of Gateaux. We denote by ∂_t the derivative with respect to t, by D_j the derivative with respect x^j (that is, with respect to x along e_j) and by ∂_{p^j} the derivative with respect to p^j.

Finite-dimensional systems like (V) (that is, with $i\in[\![N]\!]$ and thus set in $\mathbb{R}^N_T := [0,T]\times\mathbb{R}^N$) naturally arise to describe closed-loop Nash equilibria in stochastic differential N-player games, and they are notoriously hard to study due to the strong coupling Hamiltonian terms. Let us briefly recall how this happens, referring the reader to [**25**, **46**] for further details.

Consider a game in which the vector $X_t=(X^i_t)_{i\in[\![N]\!]}$ of the states of the players obeys the following $\mathbb{R}^N$-valued SDE on $[0,T]$:

$$\mathrm{d}X_t = B(t,X_t,\alpha_t)\,\mathrm{d}t + \Sigma(t,X_t)\,\mathrm{d}B_t\,,$$

where $B=(B^i)_{i\in[\![N]\!]}\colon[0,T]\times\mathbb{R}^N\times\mathbb{R}^N\to\mathbb{R}^N$ is the given vector of the *drifts*, $\alpha=(\alpha^i)_{i\in[\![N]\!]}\colon[0,T]\to\mathbb{R}^N$ is the vector of closed-loop controls in feedback form (that is, $\alpha^i_t=\phi^i(t,X_t)$ for some $\phi^i\colon[0,T]\times\mathbb{R}^N\to\mathbb{R}$), $\Sigma\colon[0,T]\times\mathbb{R}^N\to\mathbb{R}^{N\times M}$ is the *volatility* matrix and B is an $\mathbb{R}^M$-valued Brownian motion. Suppose that each player i aims at minimising a cost of the form

$$J^i(\alpha)=\mathbb{E}\bigg[\int_0^T F^i(t,X_t,\alpha_t)\,\mathrm{d}t + G^i(X_T)\bigg]$$

and that there exists a unique $\alpha^*(t,x,q)$ (with $q\in(\mathbb{R}^N)^N$) such that

$$\alpha^{*i}(t,x,q)\in\arg\min_{\alpha^i}\Big(B\big(t,x,(\alpha^i,\alpha^{*,-i}(t,x,q))\big)\cdot q^i + F^i\big(t,x,(\alpha^i,\alpha^{*,-i}(t,x,q))\big)\Big)$$

for each $i\in[\![N]\!]$. Then one defines the Hamiltonians

$$\tilde H^i(t,x,q) := -B(t,x,\alpha^*(t,x,q)))\cdot q^i - F^i(t,x,\alpha^*(t,x,q))\,, \tag{VI}$$

and the optimal controls (in the sense of Nash equilibrium) for the game are expected to be given by $\alpha^*(t,x,Du)$, where the so-called value functions $u^i\colon[0,T]\times\mathbb{R}^N\to\mathbb{R}$ solve the Nash system

$$-\partial_t u^i - \mathrm{Tr}\big(A(t,x)D^2u^i\big) + \tilde H^i(t,x,Du) = 0\,, \tag{VII}$$

with $A=\frac12\Sigma\Sigma^\intercal$ and terminal condition $u^i(T,\cdot)=G^i$. Now, by the envelope theorem (see, e.g., [**24**, Lemma 1.1]) one has $B(t,x,\alpha^*(t,x,q)) = -\partial_{q^i}\tilde H^i(t,x,q)$; therefore, if in addition B^i is independent of α^{-i}, then we can consider α^{*i} (and thus also H^i) that is independent of q^{jk} for $j\neq k$. In this situation one gets to a system of the form (V) by letting $H^i(t,x,(q^{jj})_j) := -B^i(t,x,\alpha^{*i}(t,x,(q^{jj})_j))\,q^{ii} - F^i(t,x,\alpha^*(t,x,(q^{jj})_j))$.

Note that in this setting we are considering each state X^i_t to be one-dimensional, while in general one could consider $X^i_t\in\mathbb{R}^d$. Our choice is only made for the sake of a simpler notation, as the willing reader can check that our results can be easily extended to the d-dimensional case, where one has $(\mathbb{R}^d)^\omega$ instead of $\mathbb{R}^\omega$.

In general, system (VII) is not well-defined for $i \in \mathbb{N}$, and thus in $[0,T] \times \mathbb{R}^{\mathbb{N}}$; when formally writing its infinite-dimensional analogue, two series appear which are not guaranteed to converge: one coming from the trace and another one carried by the Hamiltonian due to the scalar product in its definition (VI). The most notable example that such a system cannot be taken to the limit "as it is" is offered by MFG theory, where the natural limit of the Nash system is provided by the Master Equation, a parabolic equation set in the space $[0,T] \times \mathbb{R} \times \mathcal{P}(\mathbb{R})$, where $\mathcal{P}$ stands for probability measures; in other words, exploiting the structural hypotheses of MFGs and letting $N \to \infty$ one passes from $\mathbb{R}^N$ to $\mathbb{R} \times \mathcal{P}(\mathbb{R})$ instead of $\mathbb{R}^{\mathbb{N}}$.

In *network games* one drops the general symmetry and negligibility assumptions of MFGs, which basically lead to assume that the functions B, F and G above depend on X and i only through X^i and the mean $\frac{1}{N-1}\sum_{j \neq i} X^j$; instead, it is assumed that interactions are governed by a graph $\Gamma = (V,E)$, with $\#V = N$, in such a way that players are labelled by $v \in V$ and, for example, $G^v(X) = G^v(X^v, \sum_{(v',v)\in E} w(v,v')X^{v'})$ for some weights w. Nevertheless, when, as one lets $N \to \infty$, the corresponding sequence of graphs consists of *dense* graphs converging (in an appropriate sense) to a limit object which is a *graphon*, it is possible to find conditions that allow to pass to the limit in a spirit analogous to that of MFGs (basically exploiting suitable generalisations of the Law of Large Numbers), thus entering the theory of Graphon Mean Field Games (GMFGs); see, e.g., [**3, 8, 19, 63**] and the references therein.

Relatively few results are available, instead, for *sparse* graphs; see, e.g., [**61**] and references therein. Studying large population games governed by such graphs is intrinsically more difficult due to the general lack of any overall effect that could come into play as the number of players increases. In (G)MFGs (most of) the other players have an asymptotically negligible impact on a given one, which is though influenced by the mean distribution of them altogether, thus making their average not negligible in the large population limit. One the other hand, if for instance the sequence of graphs has eventually constant maximum degree, then a given player is not affected at all (at least directly) by most of the others. What one expects in this case is that mutually distant players (in the sense of the graph) should have an unimportant impact on one another, so that a characteristic property of such games should be the possibility for a player to neglect all those too far from it and still play almost optimally. This situation is indeed antipodal to that of (G)MFGs, since a natural way to obtain ε-Nash equilibria for games with many players turns out to be cutting some of them out of the game instead of supposing them to be infinitely many in order to exploit the aforementioned typical effects occurring in (G)MFG theory.

The above claims have been investigated and proved to be true in the previous chapter, in peculiar LQ settings for closed-loop equilibria, as well as in Chapter 1, in more general sparse-graph-based open-loop games. In particular, we recall that the way of depicting the *unimportance of distant players* proposed in Chapter 2 is showing that, if the graphs have vertices $V = [\![N]\!]$, then $D_j u^i \to 0$ as $N, |i-j| \to \infty$ (u being the solution to the corresponding Nash system), whereas in (G)MFGs one has $D_j u^i \to 0$ for any $j \neq i$. Indeed, note that the derivative of the i-th value function along the direction j vanishing is interpreted as the i-th player's optimal behaviour disregarding the j-th player, and, more in general, a small absolute value of $D_j u^i$ is associated to a small impact of the j-th player's behaviour on the i-th player's optimal strategy.

As a byproduct, we showed that there exist structural conditions under which $D_j u^i$ vanish fast enough for the series appearing in the formally-written infinite-dimensional Nash system to be convergent. In other words, a consequence of the unimportance of distant players is the possibility to preserve the formal structure of the Nash system in the limit $N \to \infty$. This is done in Chapter 2 in a LQ framework, so the purpose of this third chapter is to make use of the takeaways of that case study in order to formulate suitable assumptions to prove short-time existence and uniqueness for more general infinite-dimensional Nash systems.

Our main result of this chapter is Theorem 3.4.2; we informally state it here for the convenience of the reader and of the following discussion.

THEOREM III. *Let $\beta \in \ell^{\frac{1}{2}}(\mathbb{Z})$ be positive, even and c-self-controlled. Under suitable regularity and decay assumptions (involving β) on the data, there exists $T^* > 0$ such that if $T < T^*$ the Nash system* (V) *has a unique classical solution u.*

The notion of classical solution to (V) is specified in Definition 3.2.15; in particular, it entails that the derivatives $D_j u^i$ (as well as higher-order ones) decay, as j diverges, fast enough for the series appearing in the equations to be summable.

The regularity and decay assumptions are formalised in the definitions of convenient weighted Hölder spaces, in Section 3.2. We point out that we assume, for instance, that

$$|D^2_{hk} G^i| \lesssim \beta^{i-h} \wedge \beta^{i-k} \wedge \sqrt{\beta^{i-h}\beta^{i-k}}\,, \tag{VIII}$$

which is a weaker form of (IV), seemingly more manageable in a non-LQ context.

The strategy of the proof consists in obtaining such a solution as the limit of solutions to finite-dimensional Nash systems, rather than working directly in an infinite-dimensional setting. This "bottom-up" approach anticipates that pursued in Chapters 4 and 5, where a weak formulation of the Master Equation is deduced from a detailed analysis of N-dimensional Nash systems in a generalised MFG framework. Herein just as therein, the key step consists in proving certain a priori estimates for the N-dimensional system that are stable with respect to N, and in this work we will need more precise index-wise a priori estimates in order to have a sufficiently fine control on all partial derivatives of the solutions (because we wish to exploit bounds like that in (VIII), provided by the c-self controlled sequence).

The pivotal result we prove to this end is stated as Proposition 3.3.2, which we believe to be of interest in itself as well; it provides, for solutions to linear drift-diffusion equations on an arbitrary – finite – time horizon, a priori estimates that are stable with respect to the dimension, provided that the data – drift included – fulfil bounds like (VIII).

Finally, we point out that the above theorem only providing existence for sufficiently small time horizons (while under our other assumptions uniqueness always holds; see Theorem 3.4.1) is to be considered a natural consequence of our assumptions only concerning the regularity and the decay of the data. Indeed, in MFG theory it is well-known that having long-time well-posedness of the Master Equation (which is the limit object corresponding to our infinite-dimensional Nash system in the symmetric and dense setting) requires additional structural assumptions of *monotonicity*, the most influential ones being the Lasry–Lions monotonicity and the displacement monotonicity; the reader can have a look at [**1, 5, 23, 29, 47, 49, 73**], and references therein.

Outside the MF context, it is not clear how an analogous assumption should be formulated. A sufficient condition in order to deal with long horizons in our sparse setting is presented in Chapter 2, Section 2.2.4, yet it is so tailored to the LQ structure that adapting it to other frameworks seems unviable. Summing up, proving long-time well-posedness for the problem we are considering should require the identification of some correct notion of "monotonicity" in this sparse setting, but this is way beyond what we are able to do at the moment.

We conclude these introductory lines by stressing that, despite the strong relationship between the infinite-dimensional Nash system and network games, this work deals with system (V) from a purely PDE viewpoint, as the reader could have realised by this point. To the best of our knowledge, our approach is novel among those adopted so far to study infinite-dimensional (or *abstract*) HJB equations (we refer, e.g., to [**16, 17, 20, 34, 70, 71**], and references therein), for multiple reasons. First, as already noted, we mainly work on finite-dimensional problems, deducing estimates that are stable with respect to the dimension. Second, our method does not make use of any spectral theoretic or Hilbert-space-related tool, rather it relies on particular decay properties of the data G and H. Third, our ambient space $\mathbb{R}^\omega$ is not even Banach, though we note that our strategy could be set in $\ell^\infty(\mathbb{N})$, which is the most natural Banach space for our Nash system; in that case, we would be considering A to be $\ell^\infty(\mathbb{N}^2)$-valued, so we should expect to work with solutions such that D^2u^i (in the Fréchet sense) belongs to $\ell^1(\mathbb{N}^2)$, which is indeed consistent with our decay assumptions.

Chapter 4 (*d'après* [**31, 32**])**.** We open the fourth chapter by considering the system of Riccati equations arising from the Nash system in the LQ framework, introduced in Chapter 2, but within the *dense regime*, and without any symmetry assumption (like shift-invariance). Our goal is to deduce some estimates on the Riccati system that do not depend on the number of players, under the *Mean-Field-like* condition

$$\sup_i \Big(N \sum_{\substack{h,k \\ k\neq i}} |f^i_{hk}|^2 + N \sum_{k\neq i} |f^k_{ki}|^2 + |f^i_{ii}|^2 \Big) \lesssim 1\,,$$

and the same for g^i; see Remark 4.1.2 for a basic explanation of the connection between this condition and the standard Mean Field setting. It is crucial to observe that in this dense setting there is no hope to have a limit system of the same form. Indeed, in the sparse case we had an ℓ^1 control of coefficients independent of N (given by β), so that the series $\sum_{j\neq i} D_j u^j D_j u^i$ carries to the limit by dominated convergence, which now cannot be performed. In fact, at least in the symmetric case, where the costs are of the form $F^i(x) = V(x^i, \frac{1}{N-1}\sum_{j\neq i}\delta_{x^j})$, the correct limiting object is the so-called Master Equation, which has been the object of the seminal work [**23**]. The convergence of u^i to a limit function defined on the space of probability measures has been shown under monotonicity assumptions of f^i, and its regularity is obtained as a consequence of the stability properties of the MFG system, which characterises the limit model.

Here, we wish to deduce the estimates on u^i that allow for a passage to the limit $N\to\infty$ directly on the Riccati system (that is, on the Nash system). For short time horizon, one can basically reproduce the same approach of Chapter 2 (cf. Theorem 2.2.15). To achieve stability for arbitrary values of T, we impose a bound from below on the matrices $(f^i_{ij})_{ij}$ and $(g^i_{ij})_{ij}$:

$$(f^i_{ij})_{ij},\ (g^i_{ij})_{ij} \geqslant -\kappa I, \quad \kappa > 0. \tag{IX}$$

If κ is small enough, then the existence of a solution to the Nash system is guaranteed for large T, by a mechanism that produces an analogous bound from below on the matrix $(D_{ij}u^i)_{ij}$. Recalling that the equilibrium feedbacks are given by $\alpha^j = -D_j u^j$, this is equivalent to the one-sided Lipschitz condition

$$\sum_j \langle \alpha^j(t,x) - \alpha^j(t,y), x^j - y^j\rangle \leqslant \delta |x-y|^2. \tag{X}$$

Our main estimate, which pave the way for a convergence result in the $N\to\infty$ limit, is contained in Theorem 4.1.1. Interestingly, the monotonicity (or mild non-monotonicity) condition (IX) generates the nice structural property (X) of the equilibrium drift vector $(-D_j u^j)_j$; it is worth observing that, when $\kappa = 0$ and in the symmetric MFG case, (IX) is equivalent to the *displacement monotonicity condition* used in [**47**] to get well-posedness of the Master Equation.

Further related references are [**6, 41**], where the convergence problem of open-loop equilibria in symmetric N-person linear-quadratic games is addressed, or [**38**], which deals with the selection problem for models without uniqueness. See also [**7**] for a convergence result in a finite state model. LQ MFGs are studied in several works, see for instance [**13, 14, 48, 66, 67**] and references therein. Further contributions on the convergence problem are contained in [**22, 39, 43, 55, 59**], while recent results on the structure of many players cooperative equilibria can be found in [**27, 35, 54**].

Treasuring the intuitions tested in such LQ case study, we wish to derive analogous bounds on u^i in a larger Mean-Field-like setting, again by working directly on the Nash system, and without making use of limit objects such as the Master Equation. This is the content of the main part of this chapter, in which we consider the following semilinear (backward) parabolic system for the unknowns $u^i = u^i_N : [0,T] \times (\mathbb{R}^d)^N \to \mathbb{R}$, $i \in [\![N]\!]$,

$$\text{(XI)} \qquad \begin{cases} -\partial_t u^i - \operatorname{Tr}((\sigma \mathsf{I} + \beta \mathsf{J}) D^2 u^i) + \dfrac{1}{2}|D_i u^i|^2 + \displaystyle\sum_{j \in [\![N]\!] \setminus \{i\}} D_j u^j \cdot D_j u^i = f^i \\ u^i|_{t=T} = g^i \end{cases}$$

stated in $[0,T] \times (\mathbb{R}^d)^N \ni (t,x) = (t, x^0, \dots, x^{N-1})$. As said, our main goal will be deriving estimates on u^i and their space/time derivatives that are *stable* with respect to the number of equations; that is, uniformly in N.

The data are the maps $f^i, g^i : (\mathbb{R}^d)^N \to \mathbb{R}$, the parameters $\sigma > 0$ and $\beta \geqslant 0$ and the horizon $T \geqslant 0$. The symbol $\mathsf{I} = I_{Nd}$ denotes the Nd-dimensional identity matrix and $\mathsf{J} = J_N \otimes I_d$, where $J_N = 1_N \otimes 1_N$ is the N-dimensional matrix of ones. Hence, (XI) is a compact formulation of

$$\begin{cases} -\partial_t u^i(t,x) - \sigma \displaystyle\sum_{j \in [\![N]\!]} \Delta_{x^j} u^i(t,x) - \beta \sum_{j,k \in [\![N]\!]} \operatorname{Tr} D^2_{x^j x^k} u^i(t,x) \\ \qquad\qquad + \dfrac{1}{2}|D_{x_i} u^i(t,x)|^2 + \displaystyle\sum_{j \in [\![N]\!] \setminus \{i\}} D_{x^j} u^j(t,x) \cdot D_{x^j} u^i(t,x) = f^i(t,x) \\ u^i(T,x) = g^i(x). \end{cases}$$

System (XI) describes Markovian Nash equilibria in N-player differential games, in particular it characterises the value function u^i of the i-th agent for each $i \in [\![N]\!]$. In our setting, agents control via feedbacks $\alpha^i = \alpha^i(t,x)$ their own states, which are driven by the following $\mathbb{R}^d$-valued SDEs on $[0,T]$:

$$\mathrm{d}X^i_t = \alpha^i(t, X_t)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}B^i_t + \sqrt{2\beta}\,\mathrm{d}W_t, \quad i \in \{1, \dots, N\},$$

where the B^i_t's and W_t are d-dimensional independent Brownian motions. The Brownian motions B^i_t correspond to the individual noises, while W_t is the so-called common noise, as it is the same for all the equations. The i-th agents aims at minimising the following cost functional

$$\alpha^i \mapsto \mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}|\alpha^i(s,X_s)|^2 + f^i(X_s)\Big)\,\mathrm{d}s + g^i(X_T)\Big].$$

It is known that the choice $\alpha^{*,i}(t,x) = -D_i u^i(t,x)$ characterises Nash equilibria, see for instance [**25, 46**]. Moreover, since one expects uniqueness of solutions to (XI) by its (uniformly) parabolic structure, such equilibria are unique.

For *fixed* N, existence and uniqueness results for semilinear systems of the form (XI) (or close to it) are now classical, see for example [**9, 10, 26, 46, 53, 64**]. Nevertheless, the methods that are involved seem to be strongly sensitive to the dimension N. The main reason is that one typically wants to employ parabolicity to get existence, uniqueness and regularity, but its smoothing effect may deteriorate as N increases (a simple and concrete example can be found, e.g., [**36**, Example 2.13] in the setting of mean field control), while the strong transport term $\sum_j D_j u^j \cdot D_j u^i$ may take over. In fact, it is now widely accepted that (XI) should be regarded as a nonlinear transport system for large N. On top of that, the quadratic behaviour in the gradients is critical: for some quadratic systems, one may even have that solutions fail to be Hölder continuous [**44**] (at least in the elliptic case). For the sake of self-containedness, we included an existence theorem for (XI) (for any fixed N) in the appendix to this chapter (Section 4.6.3).

If interactions are assumed to be *symmetric* (that is, $f^i(x) = f(x^i, x^{-i})$ with f being symmetric in the variable x^{-i}, and the same for g^i), one expects, borrowing intuitions from Statistical Mechanics, that a simplified limit model could be derived in the limit $N \to \infty$. In this direction, MFG theory, originated by the works [**51, 65**], has seen during the last fifteen years extensive development in many fields of mathematics and applied sciences. Within this framework, the problem can be embedded into the space of probability measures $\mathcal{P}(\mathbb{R}^d)$: one may write

$$f^i(x) = F(x^i, m_{x^{-i}}),$$

where $m_{x^{-i}} := \frac{1}{N-1}\sum_{i\neq j}\delta_{x^j}$ is the empirical measure and $F : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$, and the same for g^i (everything that will be said on f^i in this introduction will hold for g^i also). Note that not only one is assuming symmetry, but also that interactions between players are "small", since a variation in x^j produces a variation in F (and then in f^i) of order $1/N$ if $j \neq i$. Nevertheless, this smallness is compensated by the number N of interactions. Ideally, since each u^i is also symmetric in x^{-i} one may hope in the convergence of $u^i(t,x)$ to $U(t, x^i, m_{x^{-i}})$, where U is a function over $[0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)$ that should be characterised by being the solution of the formal *limit* of (XI). This limit object is called the Master Equation, and it is a nonlinear transport equation over the set of probability measures. The function $U(t,x,m)$ can be regarded as the value function of a player with state x at time t, observing a population distribution m. This approach to the limit of N-player symmetric games has been discussed for the first time by P.-L. Lions during his lectures at the Collège de France [68]. Since then, the derivation of the estimates needed to complete the program in a general framework remained an open problem.

Such a convergence result is of primary importance. Not only does it justify the use of the limit object (broadly speaking, the MFG theory) to approximate Nash equilibria of the N-players problem, but it carries fundamental implications. For instance, the optimal control in the limit for each agent reads as $-D_x U(t, x^i, m)$, so it has *decentralised* structure, or, in other words, it is *open-loop* in nature. This fact reflects into the possibility of characterising Nash equilibria in the limit by a simple (compared to the Nash system (XI)) backward-forward PDE system that goes under the name of MFG system.

The convergence problem has been a matter of extensive research, and the fundamental contribution from the PDE perspective has been given by [23]. Interestingly, since the estimates for the Nash system (XI) seemed to be out of reach, the authors proposed a different approach: from the stability properties of the MFG system, one may produce smooth solutions $U(t, x, m)$ to the Master equation. These can be in turn projected over empirical measures; the $v^i(x) = U(t, x^i, m_{x^{-i}})$ are then shown to be "almost" solutions to the Nash system. Comparing the v^i with the solutions u^i, one observes a discrepancy that vanishes as $N \to \infty$.

This program has been carried out under the so-called *Lasry-Lions* monotonicity assumption

$$\text{(XII)} \qquad \int_{\mathbb{R}^d} \big(F(x, m_1) - F(x, m_2)\big)(m_1 - m_2)(\mathrm{d}x) \geqslant 0 \qquad \forall\, m_1, m_2 \in \mathcal{P}(\mathbb{R}^d),$$

which is crucial to get uniqueness and stability at the level of the MFG system. Later on, other conditions guaranteeing uniqueness of solutions have been formulated [1, 5, 29, 47, 49, 73]. Among them, we focus our attention now on the *displacement* monotonicity assumption

$$\text{(XIII)} \qquad \int_{\mathbb{R}^d \times \mathbb{R}^d} \big(D_x F(x, m_1) - D_x F(y, m_2)\big) \cdot (x - y)\, \mu(\mathrm{d}x, \mathrm{d}y) \geqslant 0$$

for all m_1, m_2 and μ having m_1 and m_2 as first and second marginal, respectively. This has been used successfully to obtain the well-posedness of the Master Equation in [47], and later in [72] to deduce uniqueness of solutions to the MFG system. To the best of our knowledge, the convergence problem in this framework has not been addressed, at least in the closed-loop case (for a recent result in the open-loop framework, see [55]). An important observation is that the monotonicity assumptions (XII) and (XIII) both *propagate* at the level of the (limit) value function $U(t, x, m)$; that is, (XII) or (XIII) hold for $U(t, \cdot, \cdot)$ in place of $F(\cdot, \cdot)$ for any $t \in [0, T]$.

To attack the convergence problem *directly* from the Nash system (that is, without employing smooth solutions of the Master Equation), one has to look for estimates of the following form:

$$\text{(XIV)}\qquad |D_{x^i}u^i|, |D_{x^i}(D_{x^i}u^i)| \lesssim 1, \qquad |D_{x_j}u^i|, |D_{x^j}(D_{x^i}u^i)| \lesssim \frac{1}{N} \qquad \forall\, i, j \neq i.$$

Roughly speaking, if such bounds are uniform in N, a compactness argument allows to define (up to subsequences) a limit value function U and a limit drift D_xU, which is the optimal feedback of a typical player. These will be Lipschitz continuous in the (x, m) variables. Our strategy to derive these estimates adapts that originated within the displacement monotone setting and used in the introductory LQ case. It can be described as follows.

As we already mentioned, under the monotonicity condition (XIII) one knows that $U(t, \cdot, \cdot)$ also satisfies (XIII) for any $t \in [0, T]$, see for example [**72**, eq. (4.7)]. Hence, by evaluating such inequality on empirical measures, and assuming that $u^i(t, x)$ is close to $U(t, x^i, m_{x^{-i}})$, one observes that for all $t \in [0, T]$

$$\text{(XV)}\qquad \sum_{i\in[\![N]\!]} \big(D_iu^i(t,x) - D_iu^i(t,y)\big) \cdot (x^i - y^i) \geqslant -c_N|x-y|^2,$$

where c_N vanishes as $N \to \infty$ (see Remark 4.2.5). We recall that $\alpha^{*,i}(t, x) = -D_iu^i(t, x)$ is the optimal drift for the i-th player. Therefore, the previous inequality reads as a *one-sided Lipschitz condition* (or dissipativity) on the *global* drift vector α^*. This property turns out to be crucial in deriving estimates on the derivatives of u^i. For example, one can deduce gradient bounds that are independent of the dimension N for u^i, which solves in fact

$$-\partial_t u^i - \mathrm{Tr}((\sigma\mathsf{I} + \beta\mathsf{J})D^2u^i) + \sum_{1\leqslant j\leqslant N} (-\alpha^{*,j}) \cdot D_ju^i = f^i + \frac{1}{2}|D_iu^i|^2,$$

by using doubling of variables methods (see Lemma 4.3.3). This observation on the structure of α^* suggests to pursue the following program:

(1) assume first that the vector (D_iu^i) satisfies the one-sided Lipschitz condition (XV) with some $c_N = M > 0$ (so, uniformly in N), and, with this condition in force, prove the desired estimates on the derivatives of u^i;
(2) employ the estimates (that will depend on the value of M, but not on N), along with the full structure of the Nash system to show that in fact the vector (D_iu^i) satisfies (XV) with $c_N = M/2$, at least for N large enough.

This shows that the Nash system enforces one-sided Lipschitz estimates on the optimal drift, or, in other words, (XV) defines (with suitable $c_N = M$) an invariant set for the equilibrium controls. Practically, it allows to show that solutions u^i satisfy the one-sided Lipschitz condition for some constant that does not depend on N, implying the desired estimates on u^i.

We will first show that this strategy works in a framework which resembles the one of displacement monotonicity in the symmetric case. In particular, we will prove that such an M exists, and it has to be smaller and smaller as T increases.

We stress that we are *not* going to assume any symmetry on the data, but only the fact that derivatives of f^i, g^i with respect to x^j behave as in the MF case. This is why we will use the terminology *Mean-Field-like* costs, see **(MF)** at p. 82. Similarly, we will not assume (XIII), rather we formulate the condition

$$\sum_{i\in[\![N]\!]} (D_i f^i(x) - D_i f^i(y)) \cdot (x^i - y^i) \geqslant -C|x-y|^2 \quad \forall\, x, y \in (\mathbb{R}^d)^N\,,$$

which reads like (XIII) once it is specialised to the MF setting and $C = 0$. Since we are going to allow for a negative right-hand side, we will refer to this case as the *$\mathscr{D}$-semimonotone* one, see **(DS)** at p. 83.

Let us now come back to the Lasry-Lions monotonicity assumption (XII), and discuss how the previous program can be adapted to this case. It is convenient to look at the one-sided Lipschitz condition (XV) via the second-order formulation (cf. Remark 4.2.3)

$$\text{(XVI)} \qquad \sum_{i,j\in[\![N]\!]} D^2_{ji} u^i(t,x)\xi^i \cdot \xi^j \geqslant -c_N|\xi|^2 \qquad \forall\, \xi \in (\mathbb{R}^d)^N.$$

It turns out that, under Lasry-Lions monotonicity, the counterpart of (XVI) is (as before by projecting onto empirical measures the Lasry-Lions monotonicity of U)

$$\begin{aligned}\sum_{\substack{i,j\in[\![N]\!]\\ i\neq j}} D^2_{ji} u^i(t,x)\xi^i \cdot \xi^j &= \sum_{i,j\in[\![N]\!]} D^2_{ji} u^i(t,x)\xi^i \cdot \xi^j - \sum_{i\in[\![N]\!]} D^2_{ii} u^i(t,x)\xi^i \cdot \xi^i \\ &\geqslant -c_N|\xi|^2\,.\end{aligned}$$

Therefore, we have now an *off-diagonal* one-sided Lipschitz information on the drift. To get the full one-sided Lipschitz control, it is necessary to bound from below the term $\sum_i D^2_{ii} u^i(t,x)\xi^i \cdot \xi^i$, which encodes a sort of “global” semiconvexity of the set of value functions u^i. This introduces an additional step in our analysis, that is handled as before. Under Lasry-Lions monotonicity, we show that the Nash system enforces not only the one-sided Lipschitz estimates on the optimal drift α^*, but also this kind of global semiconvexity. In other words, in step (1) we assume that both $\sum_{i,j} D_{ij} u^i(t,x)\xi^i \cdot \xi^j$ and $\sum_i D^2_{ii} u^i(t,x)\xi^i \cdot \xi^i$ are bounded below, and in step (2) we verify that these bounds are improved. To do so, as before we do not require symmetry, but the Mean-Field-like assumption **(MF)** and the global (with respect to i) condition

$$\sum_{i\in[\![N]\!]} \big(f^i(x) - f^i(x^{-i}, y^i) - f^i(y^{-i}, x^i) + f^i(y)\big) \geqslant -\kappa|x-y|^2 \quad \forall\, x, y \in (\mathbb{R}^d)^N,$$

$(x^{-i}, y) = (x^0, \dots, x^{i-1}, y, x^{i+1}, \dots, x^{N-1})$, as the monotonicity condition. This inquality will be called the *$\mathscr{L}$-semimonotone* assumption (see **(LS)** at p. 83), and becomes (XII) once it is specialised to the MF setting and $\kappa = 0$.

Our main result on the estimates for u^i is stated in Theorem 4.2.7: assuming Mean-Field-like costs and either $\mathscr{D}$-semimonotonicity or $\mathscr{L}$-semimonotonicity, the desired (a priori) estimates follow for any fixed $T > 0$, sufficiently small semimonotonicity constants and sufficiently large $N \in \mathbb{N}$. Note that we are not exactly obtaining (XIV); regarding second-order derivatives, we get a bound of the form $\sup_i \sum_{j\neq i} |D_j D_i u^i|^2 \lesssim 1/N$. This is still enough, in the symmetric case described below, to produce in the limit $N \to \infty$ a Lipschitz function (with respect to the Wasserstein W_2 distance).

A few comments on the two steps (1) and (2) are now in order. Step (2) is achieved, in the $\mathscr{D}$-semimonotone case, in Section 4.4. We evaluate the quantity $\sum_i \big(D_i u^i(t,x) - D_i u^i(t,y)\big) \cdot (x^i - y^i)$ along optimal trajectories, see Proposition 4.4.1. One observes some nice properties (signs…) that appear also when operating at the limit, with an additional term, which is proven to be of order $1/N$ provided that derivatives of u^i are suitably controlled up to the third order. A similar strategy is developed in Section 4.5 for the $\mathscr{L}$-monotone case, see in particular Proposition 4.5.3. We speculate that the two concepts of semimonotonicity explored here might be two special cases of a more general framework.

The estimates under the one-sided Lipschitz condition are the core of step (1) presented above, which are developed in Section 4.3. Mainly two techniques are employed there: the method of doubling variables and the Bernstein method. Both are well suited to exploit the a priori one-sided assumption on $D_i u^i$ (the former uses (XV), while the latter works well in the form (XVI)). To reach the third-order derivatives, several steps are necessary: sometimes we address each equation separately (Lemma 4.3.3, Proposition 4.3.8), sometimes we need to proceed by estimating derivatives of u^i for all i at the same time (Proposition 4.3.6, Proposition 4.3.7, Proposition 4.3.9, Proposition 4.3.11).

Our main result is the following, which consists in the merging of Theorems 4.4.2 and 4.5.5.

THEOREM IV. *Assume Mean-Field-like and semimonotone interactions (that is, assume* ***(MF)****, and either* ***(DS)*** *or* ***(LS)****). If, given $T > 0$, the semimonotonicity constants are sufficiently small (or vice versa), and $N \in \mathbb{N}$ is sufficiently large, any solution u to the Nash system* (4.2.1) *on $[0,T] \times (\mathbb{R}^d)^N$ satisfies*

$$\sup_{i\in[\![N]\!]} \Bigg(\sup_{j\in[\![N]\!]\setminus\{i\}} \|D_j u^i\|_\infty + \Bigg\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D^2_{ij} u^j\big|^2 \Bigg\|_\infty + \sum_{j\in[\![N]\!]\setminus\{i\}} \|D(D_j u^i)\|^2_\infty \Bigg) \lesssim \frac{1}{N}$$

and

$$\sup_{i\in[\![N]\!]} \|Du^i\|_\infty + \sup_{i\in[\![N]\!]} \sum_{j\in[\![N]\!]} \|D(D_i u^j)\|^2_\infty + \sum_{\substack{i,j\in[\![N]\!] \\ j\neq i}} \|D(D_{ij} u^j)\|^2_\infty \lesssim 1,$$

where the implied constants are independent of N. In addition, u shares the same type of semimonotonicity of the data.

The “vice versa” part in the statement could be not straightforwardly clear, so let us clarify that it means that the relationship between T and the semimonotonicity constants can be rephrased as follows: it is sufficient that the semimonotonicity constants are less than some threshold, which depends on T in such a way that it vanishes as $T \to \infty$ and explodes as $T \to 0$ (at suitable rates). In other words, Theorem IV holds both for fixed time and small semimonotonicity constants, and for fixed semimonotonicity constants and small time.

What is developed in this chapter is restricted to quadratic Hamiltonians. We expect that this method can be adapted to more general Hamiltonians $H^i(t, x, D_i u^i)$ with suitable semimonotone structure, though there are nontrivial technical issues that need to be overcome. Most importantly, while one still verifies the one-sided Lipschitz condition on $D_i u^i$ in step (2), this does not propagate immediately to the optimal drift $\alpha^{*,i}(t, x) = -D_p H^i(t, x, D_i u^i)$, since a composition of semimonotone maps needs not be semimonotone in general.

Chapter 5 (*d'après* [**31**])**.** With the estimates obtained in the previous chapter, we address the convergence problem of $N \to \infty$. In the full symmetric case, where all the players react in the *same way* with respect to the empirical measure of the population, we are able to obtain the classical MFG limit (Proposition 5.3.1), while the convergence problem in full generality appears to be at this stage too complicated. We put ourselves in a sort of intermediate setting: each player observes the empirical measure as in the standard MFG theory, but players can react differently. More precisely, we assume that the costs f^i (and g^i) are taken from a pool of possible costs; that is,

$$f^i(x) = F(\lambda^i, x^i, m_{x^{-i}}),$$

where the parameter λ^i belongs to $[0, 1]$. If λ^i varies in a finite subset of $[0, 1]$ for all N, one may think of a multi-population MFGs as in [**12, 28**], but λ^i can actually vary in the continuum $[0, 1]$. To obtain a convergence result, we assume Lipschitz regularity of F in the λ variable. This setting is somehow similar to the one of *Graphon* MFGs [**3, 19, 63**], with the difference that typically the cost functionals are player-independent in Graphon games, but may depend on weighted empirical measures, and weights are player-dependent. Here, costs are player-dependent, but they see the standard empirical measure of the others.

Our convergence argument is in the spirit of [**63**]. If one thinks of any player i in terms of its label-state couple (λ^i, X^i_t), it is natural to work in the limit with probabilities μ over $[0, 1] \times \mathbb{R}^d$. More precisely, we are able to define a value function $U(\lambda, t, x, \mu)$ from limits of the Nash system. Furthermore, U generates a solution of a generalised MFG system, that is a continuum of classical MFG systems that are parametrised by λ. This is the second main result of this work, and it is stated in Theorem V.

For simplicity, we proceed in this part without common noise; that is, $\beta = 0$ in (XI). A compactness argument allows to define U from limits of u^i (Theorem 5.2.6), and the convergence is strong enough to pass to the limit the notion of Nash equilibrium (via a propagation of chaos argument; see Proposition 5.3.1 and Theorem 5.3.4); this in turn yields the MFG system. We first show how to get convergence in the usual MFG setting; that is, for $\lambda^i = \lambda^{N,i} = \lambda$ for all N, i. In the multi-population (but finite) case, the argument is similar, even though it is technically more involved. Finally, the general case is obtained by approximating via a finite (and increasing) number of populations.

Our main contribution contained in this chapter consists in identifying a limit function and showing that it satisfies a certain representation formula. This in turn implies that a solution to a generalised Mean Field system can be constructed. These results are contained in Theorems 5.2.6, 5.3.4 and 5.4.2 and we summarise them here below, omitting the more technical integral characterisation provided by the second theorem. The novel assumptions **(S)** and **(LP)** can be found at p. 115 and p. 116, respectively.

In the upcoming statement, we use the notation $\underline{x} = (x, \hat{x}) \in \mathbb{R}^d \times (\mathbb{R}^d)^{N-1}$ and $\underline{\lambda} = (\lambda, \hat{\lambda}) \in \Lambda \times \Lambda^{N-1}$, with $\Lambda := [0,1]$. The map $\mathfrak{u}_N \colon \Lambda^N \times [0,T] \times (\mathbb{R}^d)^N \to \mathbb{R}$ is defined as $\mathfrak{u}_N(\underline{\lambda}, t, \underline{x}) := u^0_{\underline{\lambda}}(t, \underline{x})$, where, in order to specify the choice of $\underline{\lambda} \in \Lambda^N$ with which a Nash system is built, $(u^i_{\underline{\lambda}})_{i \in [\![N]\!]}$ denotes the solution to (XI) on $[0,T] \times (\mathbb{R}^d)^N$ where $h^i = h_N(\underline{\lambda}^i, \cdot)$ for $h \in \{f, g\}$, $i \in [\![N]\!]$. As discussed in Remark 5.2.1, the function $\mathfrak{u}_N$ is representative of any solution of an N-dimensional Nash system of the form we are considering.

Theorem V. *Let assumptions **(MF)**, **(S)** and **(LP)** be in force. Assume also that one between **(DS)** and **(LS)** holds, and that the corresponding semimonotonicity constants are such that the thesis of Theorem IV holds.*

Then, there exists a Lipschitz continuous map

$$U \colon \Lambda \times [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\Lambda \times \mathbb{R}^d) \to \mathbb{R}$$

with bounded derivative $\partial_x U$ which is Lipschitz continuous on $\Lambda \times \mathbb{R}^d \times \mathcal{P}_2(\Lambda \times \mathbb{R}^d)$ and $\frac{1}{3}$-Hölder continuous on $[0,T]$, such that, up to a subsequence,

$$\sup \left| \partial_x^k \mathfrak{u}_N(\underline{\lambda}, t, \underline{x}) - \partial_x^k U(\lambda, t, x, m_{(\hat{\lambda}:\hat{x})}) \right| \xrightarrow[N \to \infty]{} 0\,, \quad k \in \{0, 1\}\,,$$

whenever the supremum is taken over a set of the form

$$\left\{ (\underline{\lambda}, t, \underline{x}) \in \Lambda^N \times [0,T] \times (\mathbb{R}^d)^N : \ |x| \leqslant R,\ m_{\hat{x}} \in \mathcal{K} \right\}$$

for some $R > 0$ and some compact set $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R}^d)$.

Furthermore, given $\mu \in \mathcal{P}_2(\Lambda \times \mathbb{R}^d)$ with continuous disintegration with respect to the projection onto Λ, the function $u^\lambda(t, x) := U^\lambda(t, x, \mu_t)$ solves the following *(generalised) Mean Field system on* $(0,T) \times \Lambda \times \mathbb{R}^d$*:*

$$\begin{cases} -\partial_t u^\lambda - \Delta_x u^\lambda + \dfrac{1}{2}|D_x u^\lambda|^2 = f^\lambda(x, \pi_{\mathbb{R}^d\sharp}\mu_t) \\ \partial_t \mu - \Delta_x \mu - \mathrm{div}_x(Du^\lambda \mu) = 0 \\ u^\lambda(T,\cdot) = g^\lambda(\cdot, \pi_{\mathbb{R}^d\sharp}\mu_T), \quad \mu_0 = \mu\,, \end{cases}$$

where both the Hamilton–Jacobi and the Fokker–Planck equations are satisfied in the classical sense.

Note that, since compactness arguments are involved, convergence is obtained *up to subsequences*. Clearly, under displacement or Lasry–Lions monotonicity and in the classical homogeneous MFG setting one has that limit objects (MFG equilibria, or solutions to the MFG system, or solutions $U(t,x,m)$ of the Master Equation, ...) are unique, hence convergence is a posteriori along the full sequence. Nevertheless, we prefer not to stress this point here, as in fact we are requiring the milder assumptions of *semimonotonicity*, and we allow for a certain degree of heterogeneity among agents. While we are not aware of uniqueness theorems in such generality, we believe that they should be true, as consequence of the fact that $U(\lambda,t,x,\mu)$ and $D_xU(\lambda,t,x,\mu)$ are Lipschitz continuous.

In this chapter, we propose a new PDE approach to the convergence problem in large population stochastic differential games, which allows to treat in a unified way two scenarios of monotonicity. In addition, by employing the notion of semi-monotonicity of Chapter 4, we can consider at the same time the large time horizon T case, where one needs almost monotonicity, and the short time T case, where (almost) no monotonicity is needed, as in [**22**].

Besides the aforementioned approach via the Master Equation, let us mention that different probabilistic methods have been developed to tackle the convergence problem in MFGs, within the framework of closed-loop equilibria. In [**76**], propagation of chaos for BSDEs arguments are employed. In [**39**], by making use of the notion of measure-valued MFG equilibria, convergence is studied for a general class of MFGs of controls; remarkably, no uniqueness of MF equilibria is assumed here, as in the important works [**59, 60**], where semi-Markov equilibria are obtained by means of compactness arguments.

By developing further the analytic approach proposed here, we plan to extend our convergence arguments in the following two directions, for which, to the best of our knowledge, no results are yet available. First, the $\mathscr{D}$-monotone setting appears to be strong enough to guarantee estimates that are not only independent of N, but also on T, thus allowing to study the infinite-horizon (or ergodic) problem. Second, we aim at analysing the vanishing viscosity limit, where the idiosyncratic noise $\sigma = \sigma(N) > 0$ vanishes as $N \to \infty$. Note that the estimates obtained here are stable with respect to the common noise strength β, but σ needs to be bounded away from 0 uniformly in N.

Acknowledgements

My sincere thanks to my supervisor Prof. Marco Cirant, for his calmness, his openness and his attention; for his encouragement, his help, his guidance.

My appreciation for the reviewers' accurate suggestions and comments, which helped improve the quality of this work.

My gratitude and affection to all those who helped me see flashes of light from within the darkness.

Part 1

The unimportance of distant players in games with sparse interactions

CHAPTER 1

Open-loop games

Contents

1.1. Setting and main assumptions

Consider a graph $\Gamma := (V, E)$ with $\#V = N \in \mathbb{N} \cup \{\infty\}$, which can be directed but is assumed to have no loops. In the following, since the indegree of a vertex will appear quite often, we will use the lighted notation $\deg v := \deg^- v$. We shall assume that $\deg v > 0$ for all $v \in V$ and, if V is countable, $\sup_{v \in V} \deg v < \infty$.

The main object of our study is a system of N HJB and N FPK equations in $\mathbb{R}^d_T$, which we call the *graph-game system*, of the form

$$\begin{cases} -\partial_t u^v - \Delta u^v + H^v(t, Du^v) = F^v(t,x) \\ \partial_t m^v - \Delta m^v - \operatorname{div}(\partial_p H^v(t, Du^v) m^v) = 0 \qquad v \in V\,. \\ u^v(T, \cdot\,) = G^v, \;\; m^v(0) = m_0^v\,, \end{cases} \tag{1.1.1}$$

The unknowns are the value functions u^v and the measure flows m^v; the Hamiltonians $H^v(t,p)$, the cost functions $F^v(t,x)$ and $G^v(x)$, and the initial distributions $m_0^v \in \mathcal{P}_2(\mathbb{R}^d)$ are given. In particular, we consider

$$F^v(t,x) := \int_{(\mathbb{R}^d)^{\deg v}} f^v(t,x,\,\cdot\,)\,\mathrm{d}\mu_t^v\,, \quad G^v(x) := \int_{(\mathbb{R}^d)^{\deg v}} g^v(x,\,\cdot\,)\,\mathrm{d}\mu_T^v\,, \tag{1.1.2}$$

for some functions $f^v \colon [0,T] \times \mathbb{R}^d \times (\mathbb{R}^d)^{\deg v} \to \mathbb{R}$ and $g^v \colon \mathbb{R}^d \times (\mathbb{R}^d)^{\deg v} \to \mathbb{R}$, where $\mu_t^v \in \mathcal{P}((\mathbb{R}^d)^{\deg v})$ is the product measure having $m_t^{v'}$, $v' \sim v$, as its marginals.

Remark 1.1.1. In the introduction we claimed that system (1.1.1) is expected to provide open-loop Nash equilibria of a certain stochastic differential game. Such a statement could seem unclear if one has in mind some well-known general definitions of open-loop equilibria, such as those in [**25**, Chapter 2], so let us point out the correct interpretation of our terminology. By a standard duality argument between the HJB and the FPK equations of (1.1.1) (and the relationship between H^v and L^v discussed in the Introduction) one obtains that, for each $v \in V$,

$$\int_{\mathbb{R}^d} u^v(0,\cdot)\,\mathrm{d}m_0^v \leqslant \int_{\mathbb{R}^d}\int_0^T \big(L^v(t,\alpha^v(t,\cdot)) + F^v(t,\cdot)\big)\,\mathrm{d}m_t^v\mathrm{d}t + \int_{\mathbb{R}^d} G^v\,\mathrm{d}m_T^v$$

for any $\alpha^v\colon \mathbb{R}^d_T \to \mathbb{R}$, with equality if $\alpha^v(t,x) = -\partial_p H^v(t, Du^v(t,x))$. With the notation used in the Introduction, this is equivalent to $J^v(\alpha^*) \leqslant J^v(\alpha^{*,-v},\alpha^v)$ for any α^v adapted to B^v, for each $v \in V$. Hence, in other words, by an open-loop game we precisely mean that the controls α_t^v, at any given time $t \in [0,T]$, can depend only on the initial state $X_0 = (X_0^v)_{v\in V}$ and the noise B_t^v.

Our main structural assumptions on the data are the following, which are supposed to hold for all $v \in V$, $t \in [0,T]$, $x \in \mathbb{R}^d$, $y, y' \in (\mathbb{R}^d)^{\deg v}$ and $p, p' \in \mathbb{R}^d$:

(L) there exists a constant $\ell_g^v \geqslant 0$ such that

$$\big|\partial_x g^v(x,y) - \partial_x g^v(x,y')\big| \leqslant \ell_g^v |y-y'|\,,$$

and analogously for $f^v(t,\cdot)$ in lieu of g^v, with constant $\ell_f^v \geqslant 0$;

(C) there exists a function $\mathfrak{K}_g^v\colon [0,\infty) \to \mathbb{R}$, with

$$\mathfrak{L}_g^v := -\inf_{\varrho\geqslant 0} \mathfrak{K}_g^v(\varrho)\varrho^2 < +\infty\,,$$

such that

$$\big(\partial_x g^v(x,y) - \partial_x g^v(x',y)\big)\cdot(x-x') \geqslant \mathfrak{K}_g^v(|x-x'|)|x-x'|^2\,, \tag{1.1.3}$$

and analogously for $f^v(t,\cdot)$ in lieu of g^v, with function $\mathfrak{K}_f^v\colon [0,+\infty) \to \mathbb{R}$ and corresponding limit $\mathfrak{L}_f^v < +\infty$;

(H) the Hamiltonian has the form

$$H^v(t,p) = \frac{1}{2}\,\theta^v(t)|p|^2 + \tilde{H}^v(t,p)$$

with $\theta^v\colon [0,T] \to \mathbb{R}_+$ and $|\partial_p \tilde{H}^v| \leqslant C^v$ for some $C^v \geqslant 0$, also it satisfies $\lambda^v \leqslant \partial^2_{pp} H^v \leqslant \Lambda^v$ for some positive constants λ^v and Λ^v.

We can note that the requirements in condition **(H)** imply that $\lambda^v - \theta^v(t) \leqslant \partial^2_{pp}\tilde{H}^v(t,\cdot) \leqslant \Lambda^v - \theta^v(t)$, and thus $\lambda^v \leqslant \theta^v \leqslant \Lambda^v$, otherwise $\partial_p \tilde{H}^v$ could not be bounded.

In assumption **(C)** we do not make any sign assumption on $\mathfrak{K}_g^v$ and $\mathfrak{K}_f^v$, nor we give any lower bound for them. We will give one in the hypotheses of Proposition 1.2.3, showing that we can allow them to be negative as long as they vanish at a suitable rate as $T \to \infty$. On the other hand, the requirement $\mathfrak{L}_g^v < +\infty$ implies that $\liminf_{\varrho\to+\infty} \mathfrak{K}_g^v(\varrho) \geqslant 0$, meaning that while we can allow some concavity near the origin, we will be needing "convexity at infinity". Note that a function $\mathfrak{K}$ fulfilling an inequality like (1.1.3) is, for instance, the convexity profile used in [**33**] to model a suitable notion of weak semiconvexity for the quadratic HJB equation.

It is worth noting also that assumption **(C)** is genuinely needed to deal with data with bounded second-order derivatives but unbounded (sublinear) gradients. Let us anticipate that it will be used (in the proof of Proposition 1.2.3) basically to guarantee that m^v has bounded second moment; on the other hand, since the other conditions on f and g (namely, assumption **(L)** and the semiconvexity bounds (1.2.3) in the hypotheses of Proposition 1.2.3) are compatible with the requirement of f and g having bounded derivatives, our following results can be adapted to such data as well. Indeed, in this case assumption **(C)** can be dropped and the needed bound on the second moments is provided by standard estimates on the HJB and FPK equations. Additionally, the existence of solutions to system (1.1.1) is well-known if f and g are themselves bounded (for example, by classic general results on quasi-linear parabolic equations; see, e.g., [**64**]).

1.2. Unimportance of distant players

Throughout this section, even if not specified, solution $(u, m) = ((u^v, m^v))_{v \in V}$ to system (1.1.1) are understood to have sublinear $Du^v \in C^{1,2}_{\mathrm{loc}}(\mathbb{R}^d_T)$ and $m^v \in C^0([0,T]; \mathcal{P}_2(\mathbb{R}^d))$.

We start with a useful standard result on FPK equations.

Proposition 1.2.1. *Let $b \in \mathrm{Lip}_{\mathrm{loc}}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ be sublinear. Then the FPK equation*

$$\partial_t m - \Delta m + \operatorname{div}(bm) = 0 \tag{1.2.1}$$

with $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$ has a unique solution $m \in C^{\frac{1}{2}}([0,T]; \mathcal{P}_2(\mathbb{R}^d))$.

Recall that the space $C^{\frac{1}{2}}([0,T]; \mathcal{P}_2(\mathbb{R}^d))$ of all $\frac{1}{2}$-Hölder continuous $\mathcal{P}_2(\mathbb{R}^d)$-valued functions on $[0,T]$ is defined as that of all $m \in C^0([0,T]; \mathcal{P}_2(\mathbb{R}^d))$ such that

$$\sup_{\substack{s,t \in [0,T] \\ s \neq t}} \frac{W_2(m_t, m_s)}{|t-s|^{\frac{1}{2}}} < +\infty.$$

Proof of Proposition 1.2.1. Let Z be a random variable with law m_0 and consider the following SDE: $\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t$, with initial condition $X_0 = Z$. By [**4**, Theorems 9.1, 9.4 and 9.5], it has a unique solution in law, with distribution $m_t \in C^{\frac{1}{2}}([0,T]; \mathcal{P}_2(\mathbb{R}^d))$. By Itô's formula (see, e.g., [**4**, Theorem 8.3]), for any function $\varphi \in C^{1,2}((0,T) \times \mathbb{R}^d) \cap C^0([0,T] \times \mathbb{R}^d)$ with derivatives that are bounded in $t \in [0,T]$ and subquadratic in $x \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} \varphi(t, \cdot)\,\mathrm{d}m_t = \int_{\mathbb{R}^d} \varphi(0, \cdot)\,\mathrm{d}m_0 + \int_0^t \int_{\mathbb{R}^d} (\partial_t + \Delta + b \cdot D)\,\varphi(s, \cdot)\,\mathrm{d}m_s\,\mathrm{d}s \tag{1.2.2}$$

for all $t \in [0,T]$. In particular, m is the (unique, by, e.g., [**18**, Theorem 9.4.6]) solution to (1.2.1). ∎

Our key estimate (Proposition 1.2.3) essentially provides a recursive inequality for the (Wasserstein) distance of two distributions solving system (1.1.1), with possibly different underlying graphs. Before stating it, let us give the following definition. Recall that the (quasi-)distance $d(v,v')$ on a graph Γ between $v,v'\in V$ is given by the length of the shortest directed path of arcs from v to v'; hence, for example $v\sim v'$ if and only if $d(v,v')=1$.

Definition 1.2.2. Given $\Gamma=(V,E)\leqslant\bar\Gamma=(\bar V,\bar E)$ (that is, $V\subseteq\bar V$ and $E\subseteq\bar E$) and a positive integer k, we say that a vertex $v\in V$ is *k-stable* with respect to $\bar\Gamma$ if the subgraphs of Γ and $\bar\Gamma$, respectively, containing v and all vertices v' such that $d(v',v)\leqslant k$ coincide.

Also, given Γ and $v\in V$, we will denote by $\Gamma_{v;k}$ the minimal subgraph of Γ containing v such that v is k-stable with respect to Γ.

For example, $v\in V$ is 1-stable if and only if $(v',v)\in\bar E$ implies $(v',v)\in E$ for all $v'\in\bar V$.

Proposition 1.2.3. *Assume* ***(L)***, ***(C)*** *and* ***(H)***, *with*

$$\inf\mathfrak{K}^v_g+T\inf\mathfrak{K}^v_f>-\frac{\lambda^v}{6T(\Lambda^v)^2}\,. \tag{1.2.3}$$

Let $(u^v,m^v)_{v\in V}$ and $(\bar u^v,\bar m^v)_{v\in V}$ be solutions to (1.1.1) *with underlying graph Γ and $\bar\Gamma$, respectively. Let ν^v_t be the coupling of $(m^v_t,\bar m^v_t)$ such that $\nu^v_0=(\mathrm{id}_{\mathbb{R}^d},\mathrm{id}_{\mathbb{R}^d})_\sharp m^v_0$. Define*

$$\mathfrak{D}^v_T:=\sup_{t\in[0,T]}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^2\,\nu^v_t(\mathrm{d}x,\mathrm{d}y).$$

Suppose that $\Gamma\leqslant\bar\Gamma$ and $v\in V$ is 1-stable. Then one has

$$\mathfrak{D}^v_T\leqslant\Big(\delta^v_T\sum_{v'\sim v}\mathfrak{D}^{v'}_T\Big)\wedge K^v_T\,, \tag{1.2.4}$$

with

$$\delta^v_T:=\frac{2\big(\ell^v_f+\frac{\ell^v_g}{T}\big)^2}{\frac{\lambda^v}{6T(\Lambda^v)^2}+\inf\mathfrak{K}^v_f+\frac{1}{T}\inf\mathfrak{K}^v_g} \tag{1.2.5}$$

and

$$K^v_T:=\bigg(2\big(\mathfrak{L}^v_g+T\mathfrak{L}^v_f\big)\int_0^T\theta^v+(C^v)^2T\bigg)e^T\,. \tag{1.2.6}$$

Remark 1.2.4. For $g^v(\,\cdot\,,y)$ of class C^2, one has $\partial^2_{xx}g^v(x,y)\geqslant\mathfrak{K}^v_g(0)$ for all $(x,y)\in\mathbb{R}^d\times(\mathbb{R}^d)^{\deg v}$, as an immediate consequence of the fundamental theorem of calculus. Also, we can replace $\mathfrak{K}^v_g$ in assumption **(C)** with

$$\tilde{\mathfrak{K}}^v_g(\varrho):=\begin{cases}\mathfrak{K}^v_g(0) & \text{if } \mathfrak{K}^v_g(\varrho)_-\geqslant\mathfrak{K}^v_g(0)_-\\ \mathfrak{K}^v_g(\varrho) & \text{otherwise}\,;\end{cases}$$

in this case, $\inf\mathfrak{K}^v_g=\mathfrak{K}^v_g(0)$, so condition (1.2.3) can rightfully be considered as a lower bound on the semiconvexity constants of the data f and g.

REMARK 1.2.5. In the convex case $\mathfrak{K}_f^v > 0$ and $\mathfrak{K}_g^v \geqslant 0$, from (1.2.5) one sees that $\lim_{T\to+\infty} \delta_T^v = 2\ell_f^v / \inf \mathfrak{K}_f^v$ is finite.

PROOF OF PROPOSITION 1.2.3. It is not difficult to see (for example, by Itô's formula for the synchronous coupling of two processes with laws m^v and $\bar{m}^v$) that ν^v solves in the sense of distributions in $(0,T)\times\mathbb{R}^d\times\mathbb{R}^d$ the PDE

$$\partial_t \nu^v - \Delta_{(x,y)}\nu^v - 2\sum_{j\in[\![d]\!]} \partial^2_{x^j y^j}\nu^v \\ - \operatorname{div}_x(\partial_p H^v(t, Du^v(t,x))\nu^v) - \operatorname{div}_y(\partial_p H^v(t, D\bar{u}^v(t,y))\nu^v) = 0\,,$$

and that we can test the above equation by any function $\varphi \in C^{1,2}((0,T)\times(\mathbb{R}^d)^2)\cap C^0([0,T]\times(\mathbb{R}^d)^2)$ with derivatives that are bounded in $t\in[0,T]$ and subquadratic in $(x,y)\in(\mathbb{R}^d)^2$; cf., e.g., [**4**, Chapter 8].

Let us now define $\mathfrak{d}_t^v := \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\,\nu_t^v(\mathrm{d}x,\mathrm{d}y)$. Testing the above equation by $|x-y|^2$ one gets

$$\mathfrak{d}_t^v = -2\int_0^t \int_{(\mathbb{R}^d)^2} \big(\partial_p H^v(s, Du^v(s,x)) - \partial_p H^v(s, D\bar{u}^v(s,y))\big) \\ \cdot (x-y)\,\nu_s^v(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}s\,, \tag{1.2.7}$$

where we have also used that $\mathfrak{d}_0^v = 0$ because ν_0^v is supported on the diagonal $\triangle_d := \{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d : x = y\}$. By Hölder's and Young's inequalities one obtains

$$\mathfrak{d}_t^v \leqslant T\int_0^t \int_{(\mathbb{R}^d)^2} \big|\partial_p H^v(s, Du^v(s,x)) \\ - \partial_p H^v(s, D\bar{u}^v(s,y))\big|^2\,\nu_s^v(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}s + \frac{1}{T}\,\mathfrak{D}_t^v\,, \tag{1.2.8}$$

whence, recalling the Lipschitz continuity of $\partial_p H^v$ and using Gronwall's lemma,

$$\mathfrak{d}_t^v \leqslant 3T(\Lambda^v)^2 \int_0^t \int_{(\mathbb{R}^d)^2} \big|Du^v(s,x) - D\bar{u}^v(s,y)\big|^2\,\nu_s^v(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}s\,. \tag{1.2.9}$$

Now define

$$\varphi^v(t,x,y) := (Du^v(t,x) - D\bar{u}^v(t,y))\cdot(x-y) \tag{1.2.10}$$

and note that

$$- \partial_t\varphi^v - L\varphi^v \\ = \big(DF^v(t,x) - DF^v(t,y)\big)\cdot(x-y) \\ + \big(\partial_p H^v(t,x,Du^v(t,x)) - \partial_p H^v(t,y,D\bar{u}^v(t,y))\big)\cdot\big(Du^v(t,x) - D\bar{u}^v(t,y)\big)$$

with

$$L := \Delta_{(x,y)} + 2\sum_{j\in[\![d]\!]} \partial^2_{x^j y^j} - \partial_p H^v(t, Du^v(t,x))\cdot D_x - \partial_p H^v(t, D\bar{u}^v(t,y))\cdot D_y\,;$$

here we used that v is 1-stable to have that F^v is the same with respect to both Γ and $\bar\Gamma$, and the same will be true for G^v here below. Testing the equation of ν^v by φ^v one gets

$$\begin{aligned}(1.2.11)\qquad \int_0^T\int_{(\mathbb{R}^d)^2}\big(\partial_p H^v(t,Du^v(t,x)) - \partial_p H^v(t,D\bar u^v(t,y))\big)&\\ \cdot\,(Du^v(t,x)-D\bar u^v(t,y))\,\nu^v_t(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t&\\ = -\int_{(\mathbb{R}^d)^2}\big(DG^v(x)-DG^v(y)\big)\cdot(x-y)\,\nu^v_T(\mathrm{d}x,\mathrm{d}y)&\\ -\int_0^T\int_{(\mathbb{R}^d)^2}(DF^v(t,x)-DF^v(t,y))\cdot(x-y)\,\nu^v_t(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t\,;&\end{aligned}$$

here we have also used that $\int_{\mathbb{R}^d\times\mathbb{R}^d}\varphi(0,x,y)\,\nu^v_0(\mathrm{d}x,\mathrm{d}y)=0$, again because ν^v_0 is supported on $\triangle_d$ and $\varphi(0,\cdot)|_{\triangle_d}=0$. By the strong convexity of H^v and (1.2.9), the first line of (1.2.11) is bounded below by $\frac{\lambda^v}{6T(\Lambda^v)^2}\,\mathfrak{D}^v_T$. On the other hand, note that we can write

$$DF^v(t,x)-DF^v(t,y)=\int_{((\mathbb{R}^d)^{\deg v})^2}\big(\partial_x f^v(t,x,z)-\partial_x f^v(t,y,w)\big)\,\eta^v_t(\mathrm{d}z,\mathrm{d}w)$$

for any coupling η^v_t of μ^v_t and $\bar\mu^v_t$, the latter being defined in the obvious manner; then, by our structural assumptions on f^v along with Hölder's and Young's inequalities, we have, for any $\varepsilon^v_f>0$,

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}^d}(DF^v(t,x)-DF^v(t,y))\cdot(x-y)\,\nu^v_t(\mathrm{d}x,\mathrm{d}y)&\\ \geqslant(c^v_f-\varepsilon^v_f\ell^v_f)\,\mathfrak{d}^v_t-\frac{\ell^v_f}{\varepsilon^v_f}\int_{((\mathbb{R}^d)^{\deg v})^2}|z-w|^2\,\eta^v_t(\mathrm{d}z,\mathrm{d}w)\,,&\end{aligned}$$

where $c^v_f:=\inf\mathfrak{K}^v_f$. Also note now that, since η^v_t can be an arbitrary coupling of μ^v_t and $\bar\mu^v_t$, we can choose it in such a way that

$$\int_{((\mathbb{R}^d)^{\deg v})^2}|z-w|^2\,\eta^v_t(\mathrm{d}z,\mathrm{d}w)=\sum_{v'\sim v}\mathfrak{d}^{v'}_t.$$

Analogous considerations and an akin estimate hold for the term in (1.2.11) involving G^v, with parameter ε^v_g. Plugging all estimates together, with

$$\varepsilon^v_f=\varepsilon^v_g=\frac12\,\frac{\frac{\lambda^v}{6(T\lambda^v)^2}+c^v_f+\frac{c^v_g}{T}}{\ell^v_f+\frac{\ell^v_g}{T}}$$

we obtain the first upper bound in (1.2.4). To get the second one, namely $\mathfrak{D}^v_T \leqslant K^v_T$, it suffices to estimate

$$\begin{aligned}\int_{(\mathbb{R}^d)^2} &\big(DF^v(t,x) - DF^v(t,y)\big)\cdot(x-y)\,\nu^v_t(\mathrm{d}x,\mathrm{d}y)\\ &= \int_{(\mathbb{R}^d)^2}\int_{(\mathbb{R}^d)^{\deg v}} \big(\partial_x f^v(t,x,\cdot) - \partial_x f^v(t,y,\cdot)\big)\cdot(x-y)\,\mathrm{d}\mu^v_t\,\nu^v_t(\mathrm{d}x,\mathrm{d}y)\\ &\geqslant \int_{(\mathbb{R}^d)^2} \mathfrak{K}^v_f(|x-y|)|x-y|^2\,\nu^v_t(\mathrm{d}x,\mathrm{d}y) \geqslant -\mathfrak{L}^v_f\,,\end{aligned}$$

and analogously for G. Indeed, testing the equation of ν^v by φ^v over (t,T) and using the convexity of H^v as well as the estimate here above, one deduces that

$$-\int_{(\mathbb{R}^d)^2}\big(Du^v(t,x) - D\bar{u}^v(t,y)\big)\cdot(x-y)\,\nu^v_t(\mathrm{d}x,\mathrm{d}y) \leqslant \mathfrak{L}^v_g + T\mathfrak{L}^v_f\,.$$

Using the structure of H^v, this allows to deduce from (1.2.7) that

$$\mathfrak{d}^v_t \leqslant 2\big(\mathfrak{L}^v_g + T\mathfrak{L}^v_f\big)\int_0^t \theta^v + (C^v)^2 t + \int_0^t \mathfrak{d}^v_s\,\mathrm{d}s\,,$$

so Gronwall's lemma yields the desired bound. ∎

In the following, to lighten the notation, we will be implying the subscript T in the quantities appearing in inequality (1.2.4); that is, we will be writing $\mathfrak{D}^v$, δ^v, and K^v in lieu of $\mathfrak{D}^v_T$, δ^v_T, and K^v_T, respectively.

An immediate consequence of (1.2.4) in the case $\Gamma = \bar{\Gamma}$ is a uniqueness result for system (1.1.1).

COROLLARY 1.2.6. *In addition to the assumptions of Proposition 1.2.3, suppose that either*

$$\mathfrak{K}^v_g, \mathfrak{K}^v_f \geqslant 0 \quad \textit{and} \quad C^v = 0$$

or

$$\delta \sup_{v\in V} \deg^- v < 1\,,$$

where $\delta := \sup_{v\in V}\delta^v$. Then, if (u,m) and $(\bar{u},\bar{m})$ are solutions to (1.1.1), one has $m = \bar{m}$ and $u^v = \bar{u}^v$ m^v-almost everywhere for all $v \in V$.

PROOF. In the former regime we have $\mathfrak{L}^v_g = \mathfrak{L}^v_f = 0$ and thus $K^v = 0$ for all $v \in V$, see (1.2.6). In the latter one, summing (1.2.4) (with $\Gamma = \bar{\Gamma}$) over $v \in V$, we have $\sum_{v\in V}\mathfrak{D}^v \leqslant \delta\sum_{v\in V}\deg^- v\,\mathfrak{D}^v$, yielding $\mathfrak{D}^v = 0$ for all $v \in V$. Hence, in either case, $m = \bar{m}$. Note that this also implies that $\nu_t = (\mathrm{id}_{\mathbb{R}^d}, \mathrm{id}_{\mathbb{R}^d})_\sharp m_t$. Now, using the strong convexity of H^v to control the left-hand side of (1.2.11) from below and the estimates on the right-hand side obtained in the previous proof, we see that this implies that $Du^v = D\bar{u}^v$ m^v-almost everywhere, for all $v \in V$. Then testing the FPK equation by $|u^v - \bar{u}^v|^2$ one obtains that $u^v = \bar{u}^v$ m^v-a.e., for all $v \in V$. ∎

The above corollary states that system (1.1.1) has at most one solution, at least in two regimes: either if $\mathfrak{K}_g^v, \mathfrak{K}_f^v \geqslant 0$ and $C^v = 0$ for all $v \in V$ (that is, f and g are convex with respect to $x \in \mathbb{R}^d$ and $\tilde{H}$ is constant) or if δ is small enough with respect to the maximum outdegree of a vertex of Γ. Note that, looking at (1.2.5), we can say that we are in the latter regime if T is small, or so are the Lipschitz constants ℓ_f^v, ℓ_g^v and Λ^v, or f and g have large convexity constants, or the strong convexity constant λ^v is large.

Besides highlighting cases in which we have uniqueness, the importance of Proposition 1.2.3 for our purposes resides in the fact that it can be iterated to obtain some interesting information, as stated in upcoming Proposition 1.2.7. We will use the notations $v \sim_k v'$ if and only if $d(v, v') = k$, and $\deg_k v := \#\{v' \in V : v' \sim_k v\}$.

Proposition 1.2.7. *Let $v \in V$ be k-stable. Suppose that $\delta^v < 1$ and*

$$\gamma_j^v := \sup_{v' \sim_j v} \delta^{v'} < \frac{1}{3} \tag{1.2.12}$$

for all $j \in \{1, \dots, k\}$. Define by recursion $\alpha_1^v := \delta^v$ and

$$\alpha_{j+1}^v := \frac{\gamma_j^v}{1 - \gamma_j^v(1 + \alpha_j^v)}\,.$$

for $j \geqslant 1$. Then $\alpha_j^v \in (0,1)$ for all $j \in \{1, \dots, k\}$ and

$$\mathfrak{D}^v \leqslant \alpha^{v,(k)} \sum_{v' \sim_k v} \mathfrak{D}^{v'}\,, \tag{1.2.13}$$

with $\alpha^{v,(k)} := \prod_{j=1}^k \alpha_j^v$.

Proof. The entire proof is by induction. We start by showing that α_j^v is well-defined as a real number in $(0,1)$, for $j \in \{1, \dots, k\}$. This is true for $j = 1$, so suppose it is true for some $j \in \{2, \dots, k-1\}$. We have $1 - \gamma_j^v(1 + \alpha_j^v) > 0$, which implies that $\alpha_{j+1}^v > 0$; then we also have that $\alpha_{j+1}^v < 1$ if and only if $\gamma_j^v < (2 + \alpha_j^v)^{-1}$, which is true since $2 + \alpha_j^v < 3$. We now prove that

$$\sum_{v' \sim_j v} \mathfrak{D}^{v'} \leqslant \alpha_{j+1}^v \sum_{v' \sim_{j+1} v} \mathfrak{D}^{v'} \quad \forall\, j < k. \tag{1.2.14}$$

It is clearly true for $j = 0$, because it reduces to (1.2.4); then, since if $j < k$ then all v such that $v' \sim_j v$ are 1-stable, from (1.2.4) we obtain

$$\begin{aligned}\sum_{v' \sim_j v} \mathfrak{D}^{v'} \leqslant \sum_{v' \sim_j v} \delta^{v'} \sum_{v'' \sim v'} \mathfrak{D}^{v''} &\leqslant \gamma_j^v \sum_{|l-j| \leqslant 1} \sum_{v' \sim_l v} \mathfrak{D}^{v'} \\ &\leqslant \gamma_j^v \Big(\sum_{v' \sim_{j+1} v} \mathfrak{D}^{v'} + (1 + \alpha_j^v) \sum_{v' \sim_j v} \mathfrak{D}^{v'} \Big)\,,\end{aligned}$$

so (1.2.14) is proved. At this point (1.2.13) follows by iteratively using (1.2.14) in (1.2.4). ∎

We can now exploit this result to prove the unimportance of distant players. As we explained in the Introduction, under suitable conditions, given a player v of a game with underlying graph Γ, it is possible to consider an analogous game with a subgraph $\hat{\Gamma} \leqslant \Gamma$ as the underlying graph in such a way that the respective optimal distributions m^v and $\hat{m}^v$ are arbitrarily close (in the Wasserstein sense). Such a subgraph does not contain players that are too far from player v (in the sense of the graph), so this means that that player can neglect distant players and still play almost optimally.

THEOREM 1.2.8. *Let m solve (1.1.1) with graph $\Gamma = (V, E)$.*[1] *Let $v \in V$ and $\varepsilon > 0$. With the notation of Proposition 1.2.7, suppose that $\delta^v < 1$ and there exists $k < \#V$ such that*

$$\max_{1 \leqslant j \leqslant k} \gamma_j^v < \frac{1}{3} \qquad \textit{and} \qquad \alpha^{v,(k)} \sum_{v' \sim_k v} K^{v'} < \varepsilon^2 \,,$$

where $K^v = K_T^v$ is that given in (1.2.6). Then, given $\hat{m}$ solving (1.1.1) with graph $\Gamma_{v;k}$, one has

$$\|W_2(m^v, \hat{m}^v)\|_\infty < \varepsilon \,. \tag{1.2.15}$$

PROOF. Combine Propositions 1.2.3 and 1.2.7 and note that, by its definition, $\mathfrak{D}^v \geqslant \|W_2(m^v, \hat{m}^v)\|_\infty^2$. ∎

This theorem provides a fairly explicit insight on the interplay between the structural parameters related to the graph and the data in order for the localisation of the game described above to be possible. The following corollary presents the main implication of this result in the interesting case of Γ being infinite-dimensional (that is, having a countable vertex set V).

COROLLARY 1.2.9. *Let m solve (1.1.1) with an infinite-dimensional underlying graph $\Gamma = (V, E)$ having $\sup_{v \in V} \deg v < \infty$, and assume that all the structural constants in assumptions **(L)**, **(C)** and **(H)** are independent of $v \in V$ (thus, we will omit the superscript v in the following). Suppose that $\delta < 1$, $\sup_{j \geqslant 1} \gamma_j < \frac{1}{3}$. Then $\alpha := \sup_{j \geqslant 1} \alpha_j < 1$ and, if $\alpha^k \deg_k v \to 0$ as $k \to \infty$, for any $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that, for any $v \in V$, (1.2.15) holds for $\hat{m}$ solving (1.1.1) with graph $\Gamma_{v;k}$.*

PROOF. As $\sup_{j \geqslant 1} \gamma_j < \frac{1}{3}$, one easily sees that $\alpha < 1$. Then $\alpha^{(k)} \sum_{v' \sim_k v} K^{v'} = K \alpha^k \deg_k v < \varepsilon^2$ for k large enough (independent of v). ∎

We give a couple of basic examples, in the LQ setting, to illustrate this last corollary.

[1] That is, there exists u such that (u, m) is a solution to (1.1.1) with graph Γ.

EXAMPLE 1.2.10. Consider $H^v(t,p) = \frac{1}{2}\theta^v(t)|p|^2$,

$$f^v(t,x,y) = \frac{1}{2}\,\mathfrak{f}^v(t)(x,y)\cdot(x,y) \quad \text{and} \quad g^v(x,y) = \frac{1}{2}\,\mathfrak{g}^v(x,y)\cdot(x,y)$$

for all $t \in [0,T]$ and $(x,y) \in \mathbb{R}^d \times (\mathbb{R}^d)^{\deg v}$, where $\mathfrak{f}^v \colon [0,T] \to \mathscr{S}(\mathbb{R}^d \times (\mathbb{R}^d)^{\deg v})$ is a matrix-valued function and $\mathfrak{g}^v \in \mathscr{S}(\mathbb{R}^d \times (\mathbb{R}^d)^{\deg v})$ is a matrix. We refer to Section 1.3.2 for general considerations on the properties needed to be fulfilled by θ, $\mathfrak{f}$ and $\mathfrak{g}$ in order for our structural assumptions to be satisfied.

Take Γ to be the *bi-directed chain*, given by $V = \mathbb{Z}$ and $E = \{(i, i\pm 1) : i \in \mathbb{Z}\}$, which is the natural limit of a sequence of cycle graphs of length $N \to \infty$. Clearly $\deg_k v = 2$ for all $k \in \mathbb{N}$. Let $\mathfrak{g} = 0$ and

$$\mathfrak{f}^i \equiv \begin{pmatrix} a+b & -a & -b \\ -a & a & 0 \\ -b & 0 & b \end{pmatrix} \otimes I_d$$

for some constants $a, b \in \mathbb{R}$, so that $f^i(t,x,y) = f^i(t,x,y_+,y_-) = \frac{a}{2}|x-y_+|^2 + \frac{b}{2}|x-y_-|^2$. In this case, we have $\ell^i_g = 0$ and $\ell^i_f = \frac{1}{2}\sqrt{a^2+b^2}$ for all $i \in \mathbb{Z}$, and one sees that, if $a+b > 0$,

$$\gamma_j = \delta < \frac{a^2+b^2}{2(a+b)} \qquad \forall\, j \geqslant 1\,.$$

Hence, a sufficient condition for Corollary 1.2.9 to apply is that (a,b) is a point inside the circle of centre $(\frac{1}{3}, \frac{1}{3})$ and radius $\frac{\sqrt{2}}{3}$.

EXAMPLE 1.2.11. Consider Γ being an infinite *r-regular graph*, for some $r > 2$. Assume it is undirected; nevertheless very similar considerations hold for the directed case as well. Let $\mathfrak{g} = 0$ and

$$\mathfrak{f}^v \equiv \begin{pmatrix} \sum_{1\leqslant j\leqslant r} a^j & -a^1 & -a^2 & \cdots & -a^r \\ -a^1 & a^1 & 0 & \cdots & 0 \\ -a^2 & 0 & a^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a^r & 0 & 0 & \cdots & a^r \end{pmatrix} \otimes I_d$$

for constants a^j such that $\sum_{1\leqslant j\leqslant r} a^j > 0$. The situation is analogous to that of Example 1.2.10, with $\gamma_j = \delta < \frac{1}{3}$ provided that the vector $a = (a^j)_{1\leqslant j\leqslant r} \in \mathbb{R}^r$ lies in the ball of centre $\frac{1}{3}1_r$ and radius $\frac{\sqrt{r}}{3}$. The main difference with the above case is that $\deg_k v$ has in general exponential growth: indeed, we have the universal bound $\deg_k v \leqslant (r-1)^k$ (corresponding to the case without closed paths), or, for example, in the natural setting of Γ being a rectangular lattice one has $\deg_k v = 2^{k+1}$.

As $\delta < \frac{1}{3}$, we have $\alpha_j > \alpha_{j+1}$ if and only if $\alpha_j < \frac{1-\delta+\sqrt{1-2\delta-3\delta^2}}{2\delta}$, this latter number being greater than δ; therefore, $\alpha_j < \frac{1}{3}$ for all $j \geqslant 1$, so that we can estimate $\alpha^k < 3^{-k}$. This shows that for the rectangular lattice the condition $\alpha^k \deg_k v \to 0$ as $k \to \infty$ is fulfilled without any further assumptions. In general, one can write sharper estimates on $\alpha = \alpha(\delta)$ and show that Corollary 1.2.9 applies provided that δ is less than some $\bar{\delta}(r)$; in other words, the vector a needs to be inside the ball of centre $\bar{\delta}(r)1_r$ and radius $\frac{\sqrt{r}}{\bar{\delta}(r)}$.

1.3. Existence of solutions to the graph-game system

In this last section, we consider two settings in which we prove existence of solutions to the graph-game system (1.1.1) belonging to the class for which the a priori estimates of the previous section hold.

The first one deals with data with bounded derivatives, hence sublinear. As foreshadowed in Section 1.1, this case is compatible with our main structural assumptions once **(C)** is dropped and T is finite, so that (1.2.3) allows f and g to be semiconvex.

The second one is linear-quadratic, thus allowing for the most explicit computations in a genuinely quadratic setting. We are not able to deal with a general subquadratic framework, for which, to the best of our knowledge, the problem of (determining conditions for) existence of classical solutions with bounded second-order derivatives is still open. We mention, for example, the results of [**11**] or [**52**] on HJB equations with quadratic growth, which though cannot provide sufficiently strong second-order estimates.

1.3.1. Data with bounded derivatives. We start with a standard a priori estimate on the derivative of the solution to an HJB equation. It will be used in the proof of the following intermediate result (Proposition 1.3.2), stating existence of sublinear solutions to such HJB equation with data having bounded derivatives.

LEMMA 1.3.1. *Let $u \in C^{1,2}(\mathbb{R}^d_T)$ with $Du \in C^{1,2}(\mathbb{R}^d_T)$ solve*

$$\begin{cases} -\partial_t u - \Delta u + H(t, Du) = F & \text{in } \mathbb{R}^d_T \\ u(T, \cdot) = G \end{cases} \tag{1.3.1}$$

where $|DF| \leqslant c_F$, $|DG| \leqslant c_G$ and $(t,p) \mapsto H(t,p)$ continuous in t and differentiable in p. Then $\|Du\|_\infty \leqslant C$ with C depending only on T, c_F and c_G.

PROOF. Fix $(\tau, x) \in \mathbb{R}^d_T$ and let m solve

$$\begin{cases} \partial_t m - \Delta m - \operatorname{div}(\partial_p H(t, Du(t,x))m) = 0 \\ m(\tau) = \delta_x \,. \end{cases} \tag{1.3.2}$$

Testing the FPK equation by $|Du|^2$, we get

$$\begin{aligned} &\int_{\mathbb{R}^d} |Du(t,\cdot)|^2 \,\mathrm{d}m(t) + \int_t^T \int_{\mathbb{R}^d} |D^2 u(s,\cdot)|^2 \,\mathrm{d}m(s)\mathrm{d}s \\ &\qquad = \int_{\mathbb{R}^d} |DG|^2 \,\mathrm{d}m(T) + 2\int_t^T \int_{\mathbb{R}^d} DF(s,\cdot)\cdot Du(s,\cdot)\,\mathrm{d}m(s,\mathrm{d}y)\mathrm{d}s \end{aligned} \tag{1.3.3}$$

for all $t \in [\tau, T]$. Let now, for the sake of shortness,

$$\mathcal{V}(t) := \sup_{s\in[t,T]} \int_{\mathbb{R}^d} |Du(s,\cdot)|^2 \,\mathrm{d}m(s)\,;$$

using Young's inequality on the right-hand side of the identity above, we have

$$\mathcal{V}(t) \leqslant c_G^2 + Tc_F^2 + (T-t)\mathcal{V}(t) \qquad \forall\, t \in [\tau, T]\,.$$

For $T - t \leqslant \frac{1}{2}$ this yields $\mathcal{V}(t) \leqslant 2(c_G^2 + Tc_F^2)$; then we can iterate this argument (with different final time) finitely many times to get $\mathcal{V}(\tau) \leqslant c_T(c_G^2 + Tc_F^2)$, where c_T is a constant depending only on T. In particular, $\|Du\|_\infty^2 \leqslant c_T(c_G^2 + Tc_F^2)$ by the arbitrariness of (τ, x). ∎

Proposition 1.3.2. *For some $\delta \in (0,1)$, let $F \in C^{\frac{\delta}{2},\delta}_{\mathrm{loc}}(\mathbb{R}^d_T)$ with $DF \in C^{\frac{\delta}{2},\delta}(\mathbb{R}^d)^d$ and $G \in C^{3+\delta}_{\mathrm{loc}}(\mathbb{R}^d_T)$ with $DG \in C^{2+\delta}(\mathbb{R}^d)^d$. Suppose that $F(t,\cdot)$ is both semiconcave and semiconvex uniformly in t. Let $H \in C^{\frac{\delta}{2}}([0,T]; C^2_{\mathrm{loc}}(\mathbb{R}^d))$ with bounded second-order derivative $\partial^2_{pp}H$. Then there exists a unique solution $u \in C^{1+\delta,2+\delta}_{\mathrm{loc}}(\mathbb{R}^d_T)$ with $Du \in C^{1+\delta,2+\delta}(\mathbb{R}^d_T)$ to (1.3.1).*

Proof. Suppose that D^2F is of class $C^{\frac{\delta}{2},\delta}$ and $G \in C^{4+\delta}_{\mathrm{loc}}(\mathbb{R}^d)$ with $DG \in C^{3+\delta}(\mathbb{R}^d)^d$. Let $\psi := (1 + |\cdot|^2)^{-\frac{1}{2}}$ and, for each $n \in \mathbb{N}$, let $\psi_n := \psi(\frac{\cdot}{n})$. Define $F_n := F\psi_n$ and $G_n := G\psi_n$. Let $u_n \in C^{1+\delta,2+\delta}(\mathbb{R}^d_T)$ with Du_n and D^2u_n of class $C^{1+\delta,2+\delta}$ solve (1.3.1) with F_n and G_n in lieu of F and G, respectively. In the following, c_F, c_ψ, c_G and c_H denote constants that respectively depend only on F, ψ, G and H. Observe that we have

$$|DF_n| \leqslant |FD\psi| + |\psi DF| \leqslant c_F(1 + c_\psi) \tag{1.3.4}$$

and an akin estimate holds for G_n as well. By Lemma 1.3.1, $\|Du_n\|_\infty$ is controlled by a constant depending only on T and the right-hand sides of (1.3.4) and its analogue for DG_n; in particular such a norm is bounded uniformly with respect to n. Fix now $(\tau, x) \in \mathbb{R}^d_T$ and test (1.3.2), with $u = u_n$, by $|u_n|^2$; we obtain

$$\begin{aligned}\int_{\mathbb{R}^d} &|u_n(t,\cdot)|^2 \,\mathrm{d}m(t) \\ &= \int_{\mathbb{R}^d} |G_n|^2 \,\mathrm{d}m(T) + 2\int_t^T \int_{\mathbb{R}^d} F(s,\cdot)u_n(s,\cdot)\,\mathrm{d}m(s)\mathrm{d}s \\ &\quad + 2\int_t^T \int_{\mathbb{R}^d} \big(\partial_p H(s, Du_n(s,\cdot)) \cdot Du_n(s,\cdot) - H(s, Du_n(s,\cdot))\big)\, u_n(s,\cdot)\,\mathrm{d}m(s)\mathrm{d}s\end{aligned}$$

for any $t \in [\tau, T]$. From this identity, using that if $|\partial^2_{pp}H| \leqslant \Lambda$ then

$$\big|\partial_p H(s, Du_n(s,\cdot)) \cdot Du_n(s,\cdot) - H(s, Du_n(s,\cdot))\big| \leqslant |H(s,0)| + \Lambda |Du_n(s,\cdot)|^2$$

along with Young's inequality, we deduce that

$$\begin{aligned}\mathcal{U}(t) \leqslant (1 + c_\psi)(c_G + Tc_F)\Big(1 + \sup_{s\in[\tau,T]} \int_{\mathbb{R}^d} |\cdot|^2 \,\mathrm{d}m(s)\Big) + Tc_H(1 + \|Du_n\|_\infty^4) \\ + 2(T-t)\,\mathcal{U}(t)\,,\end{aligned}$$

with $\mathcal{U}(t) := \sup_{s\in[t,T]} \int_{\mathbb{R}^d} |u_n(s,\cdot)|^2 \,\mathrm{d}m(s)$. Arguing as in the proof of Lemma 1.3.1, we deduce that there exists a constant C depending only on T, F, G, ψ and H such that

$$|u_n(\tau,x)|^2 \leqslant C\biggl(1 + \sup_{s\in[\tau,T]} \int_{\mathbb{R}^d} |\cdot|^2 \,\mathrm{d}m(s)\biggr).$$

Testing now the same FPK equation by $|\cdot|^2$ it is easy to see that Gronwall's lemma ensures that the right-hand side above is controlled by $C'(1+|x|^2)$, for some constant C' which depends only on the same parameters as C. We conclude that $\|\psi u_n\|_\infty$ is bounded uniformly in n.

To estimate the second-order derivatives, we test the FPK equation by $|D^2 u_n|^2$. After easy computations, we obtain

$$\begin{aligned}\int_{\mathbb{R}^d} &|D^2 u_n(t,\cdot)|^2 \,\mathrm{d}m(t) \\ &= \int_{\mathbb{R}^d} |D^2 G_n|^2 \,\mathrm{d}m(T) + 2\int_t^T \int_{\mathbb{R}^d} D^2 F(s,\cdot) : D^2 u_n(s,\cdot) \,\mathrm{d}m(s)\mathrm{d}s \\ &\quad + 2\int_t^T \int_{\mathbb{R}^d} \partial^2_{pp} H(s, Du_n(s,\cdot)) D^2 u_n(s,\cdot)^2 : D^2 u_n(s,\cdot) \,\mathrm{d}m(s)\mathrm{d}s\,,\end{aligned}$$

where $A : B := \mathrm{Tr}(AB)$. Note that by (1.3.3) we can deduce that

$$\int_t^T \int_{\mathbb{R}^d} |D^2 u_n(s,\cdot)|^2 \,\mathrm{d}m(s)\mathrm{d}s \leqslant C$$

for some constant C which does not depend on n or (τ,x). Therefore we have

$$\begin{aligned}\int_{\mathbb{R}^d} &|D^2 u_n(t,\cdot)|^2 \,\mathrm{d}m(t) \\ &\leqslant c_G + Tc_F + (T-t)\sup_{s\in[t,T]} \|D^2 u_n\|_\infty^2 + 2\Lambda^2 C^2 + \frac{1}{2}\sup_{s\in[t,T]} \|D^2 u_n\|_\infty^2\end{aligned}$$

for any m solving the FPK equation (that is, regardless the initial time and distribution); so, choosing $m(t) = \delta_z$ with $z \in \arg\max_{\mathbb{R}^d} |D^2 u_n(t,\cdot)|$, we have

$$\begin{aligned}&\|D^2 u_n(t,\cdot)\|_\infty^2 \\ &\leqslant c_G + Tc_F + (T-t)\sup_{s\in[t,T]} \|D^2 u_n(s,\cdot)\|_\infty^2 + 2\Lambda^2 C^2 + \frac{1}{2}\sup_{s\in[t,T]} \|D^2 u_n(s,\cdot)\|_\infty^2\,,\end{aligned}$$

whence

$$\mathcal{W}(t) \leqslant 2(c_G + Tc_F + 2\Lambda^2 C^2) + 2(T-t)\mathcal{W}(t)\,,$$

with $\mathcal{W}(t) := \sup_{s\in[t,T]} \|D^2 u_n(s,\cdot)\|_\infty^2$. We deduce from the above inequality that $\|D^2 u_n\|_\infty$ is bounded uniformly in n, again arguing as in the proof of Lemma 1.3.1.

At this point, by the equation of u_n we have a uniform bound on $\|\partial_t u_n\|$, while by a standard Hölder estimate for the heat equation (see, e.g., Lemma 4.6.2) we have $Du_n \in C^{\frac{1}{3},\frac{2}{3}}(\mathbb{R}^d_T)^d$ uniformly in n. By the Ascoli–Arzelà theorem there exists $u \in C^{\frac{1}{3}}([0,T]; C^1_{\mathrm{loc}}(\mathbb{R}^d))$ with $Du \in C^{\frac{1}{3},\frac{2}{3}}(\mathbb{R}^d_T)^d \cap C^0([0,T]; \mathrm{Lip}(\mathbb{R}^d))^d$ such that, up to a subsequence, $u_n \to u$ locally uniformly and $Du_n \to Du$ uniformly (as $n \to \infty$).

Then we can pass to the limit Duhamel's formula for the heat equation and say that $v = u$ solves in the mild sense

$$\begin{cases} -\partial_t v - \Delta v = F - H(t, Du) & \text{in } \mathbb{R}^d_T \\ v(T, \cdot) = G\,. \end{cases}$$

It is now standard to prove that u is the unique classical solution to (1.3.1), and the desired regularity follows from classic Hölder estimates for parabolic equations.

For less regular data than we supposed at the beginning of the proof, it suffices to consider smooth approximations f_n and g_n of F_n and G_n, respectively, such that $(\psi f_n, Df_n) \to (\psi F, DF)$ uniformly and the semi-convexity/concavity constants are preserved (and the same for G) and note that the estimates on u_n used to pass to the limit as $n \to \infty$ are stable under such an approximation. ∎

We can now state our first main existence theorem.

THEOREM 1.3.3. *Let $\Gamma = (V, E)$ be a graph with $D := \sup_{v \in V} \deg v < \infty$. For each $v \in V$ let $f^v \colon \mathbb{R}^d_T \times (\mathbb{R}^d)^{\deg v} \to \mathbb{R}$, $g^v \colon \mathbb{R}^d \times (\mathbb{R}^d)^{\deg v} \to \mathbb{R}$ and $H^v \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be given. Assume the following, for each $(t, x, y) \in \mathbb{R}^d_T \times (\mathbb{R}^d)^{\deg v}$:*

- ***(L)** is in force;*
- *there exists $\delta \in (0,1)$ and constants C_f, C_g and C_H such that*

$$\|f^v(\cdot, x, y)\|_{C^{\frac{\delta}{2}}([0,T])} \leqslant C_f(1+|y|^2)^{\frac12}(1+|x|^2)^{\frac12}\,,$$
$$\|\partial_x f^v(\cdot, y)\|_{C^{\frac{\delta}{2},\delta}(\mathbb{R}^d_T)} \leqslant C_f(1+|y|^2)^{\frac12}\,,$$
$$\|\partial_x g^v(\cdot, y)\|_{C^{2+\delta}(\mathbb{R}^d)} \leqslant C_g(1+|y|^2)^{\frac12}\,,$$

  and $H^v \in C^{\frac{\delta}{2}}([0,T]; C^2_{\mathrm{loc}}(\mathbb{R}^d))$ with $|\partial^2_{pp} H^v| \leqslant C_H$;
- *$\partial_x f^v(t, \cdot, y)$ is Lipschitz continuous and there exist (positive) constants κ^v_f and K^v_f such that*

$$-\kappa^v_f(1+|y|^2)^{\frac12} I_d \leqslant \partial^2_{xx} f^v(t,x,y) \leqslant K^v_f(1+|y|^2)^{\frac12} I_d$$

  in the sense of distributions.

Let $m^v_0 \in \mathcal{P}_2(\mathbb{R}^d)$ with $\sup_{v\in V} \int_{\mathbb{R}^d} |\cdot|^2 \,\mathrm{d}\mu^v_0 \leqslant \bar{M}$ for some finite constant $\bar{M}$. Then, for any $T > 0$, there exists a solution (u, m) to system (1.1.1), with u^v having the growth and the regularity guaranteed by Proposition 1.3.2 and $m^v \in C^{\frac12}([0,T]; \mathcal{P}_2(\mathbb{R}^d))$, uniformly in v.

PROOF. For some constant $K > 0$ to be determined, define

$$\mathcal{X}^v := \left\{ \mu^v \in C^0([0,T]; \mathcal{P}_1((\mathbb{R}^d)^{\deg v})) : \sup_{[0,T]} \int_{(\mathbb{R}^d)^{\deg v}} |\cdot|^2 \,\mathrm{d}\mu^v + [\mu^v]_{\frac12} \leqslant K \right\},$$

where $[\mu^v]_{\frac12} := \sup_{s \neq t} |s-t|^{-\frac12} W_1(\mu^v_s, \mu^v_t)$; note that $\mathcal{X}^v$ is a convex and compact subset of $C^0([0,T]; \mathcal{P}_1((\mathbb{R}^d)^{\deg v}))$ (this follows, e.g., from [**24**, Lemma 1.5] and the

Ascoli–Arzelà theorem). Given $\mu^v \in \mathcal{X}^v$, let

$$F^v(t,x;\mu^v_t) := \int_{(\mathbb{R}^d)^{\deg v}} f^v(t,x,y)\,\mu^v_t(\mathrm{d}y)\,, \quad G^v(x;\mu^v_T) := \int_{(\mathbb{R}^d)^{\deg v}} g^v(x,y)\,\mu^v_T(\mathrm{d}y)\,.$$

Observe that

$$\|DF^v(\,\cdot\,;\mu^v)\|_{C^{\frac{\delta}{2},\delta}(\mathbb{R}^d_T)} \leqslant C_f \sup_{[0,T]} \int_{\mathbb{R}^d} \Psi\,\mathrm{d}\mu^v + \ell^v_f K T^{\frac{1-\delta}{2}}\,,$$

where $\Psi := (1+|\cdot|^2)^{\frac12}$ and the second term in the right-hand side comes only from the Hölder seminorm $\sup_{x\in\mathbb{R}^d}[DF^v(\,\cdot\,,x;\mu^v_\cdot)]_{\frac{\delta}{2}}$; similarly,

$$\|D^2F^v(\,\cdot\,;\mu^v)\|_{L^\infty(\mathbb{R}^d_T)} \leqslant \sqrt{d}\,(\kappa^v_f \vee K^v_f) \sup_{[0,T]} \int_{\mathbb{R}^d} \Psi\,\mathrm{d}\mu^v$$

and an analogous estimate holds for $\|DG^v(\,\cdot\,;\mu^v)\|_{C^{2+\delta}(\mathbb{R}^d)}$ as well.

For each $v \in V$, let u^v be the solution to

$$\begin{cases} -\partial_t u^v - \Delta u^v + H^v(t,Du^v) = F^v(t,x;\mu^v_t) & \text{in } \mathbb{R}^d_T \\ u^v(T,\,\cdot\,) = G^v(\,\cdot\,;\mu^v_T) \end{cases}$$

provided by Proposition 1.3.2; note that the strategy of the proof of Lemma 1.3.1 applies to estimate

$$\|Du^v\|^2_\infty \leqslant c_T(C^2_g + TC^2_f) \sup_{[0,T]} \int_{\mathbb{R}^d} \Psi^2\,\mathrm{d}\mu^v \leqslant c_T(C^2_g + TC^2_f)(1+K) \tag{1.3.5}$$

Then let $\bar\mu^v \in C^{\frac12}([0,T];\mathcal{P}_2((\mathbb{R}^d)^{\deg v})$ be the solution to

$$\begin{cases} \partial_t \bar\mu^v - \Delta\bar\mu^v - \sum_{v'\sim v} \operatorname{div}_{y^{v'}}(\partial_p H^{v'}(t,Du^{v'}(t,y^{v'}))\bar\mu^v) = 0 & \text{in } [0,T]\times(\mathbb{R}^d)^{\deg v} \\ \bar\mu^v(0) = \mu^v_0 \in \mathcal{P}_2(\mathbb{R}^d)\,, \end{cases}$$

where we denoted by $y^{v'}$, $v' \sim v$, the coordinates of $y \in (\mathbb{R}^d)^{\deg v}$. It is easy to see that the marginals of $\bar\mu^v$ solve the FPK equations of (1.1.1) corresponding to vertexes $v' \sim v$, provided that μ^v_0 is a suitable measure with marginals $m^{v'}_0$, $v' \sim v$ (recall the definition (1.1.2) of F and G); on the other hand, such features of μ^v_0 are actually enough to have the only piece of information we need about it, namely that $\int_{\mathbb{R}^d} |\cdot|^2\,\mathrm{d}\mu^v_0 \leqslant D\bar M$.

By (1.3.5) and standard Hölder estimates for the FPK equation (cf., e.g., [**24**, Lemma 1.6]),

$$[\bar\mu]_{\frac12} \leqslant 1 + \big(T\hat C_H D\big(1 + c_T(C^2_g + TC^2_f)(1+K)\big)\big)^{\frac12}\,,$$

where $\hat C_H := \|\partial_p H(\,\cdot\,,0)\|^2_\infty + C^2_H$; hence, for $K > 3$ and $T\hat C_H D\big(1 + c_T(C^2_g + TC^2_f)(1+K) < 1$ we have $[\bar\mu]_{\frac12} \leqslant K$. Testing instead the FPK equation of $\bar\mu^v$ by $|\cdot|^2$, Gronwall's lemma yields

$$\sup_{[0,T]} \int_{\mathbb{R}^d} |\cdot|^2\,\mathrm{d}\bar\mu^v \leqslant \bigg(\int_{\mathbb{R}^d} |\cdot|^2\,\mathrm{d}\mu^v_0 + T\Big(2d\deg v + \sum_{v'\sim v}\|Du^{v'}\|^2_\infty\Big)\bigg)e^T\,, \tag{1.3.6}$$

which along with (1.3.5) gives

$$\sup_{[0,T]} \int_{\mathbb{R}^d} |\cdot|^2 \,\mathrm{d}\bar{\mu}^v \leqslant D\Big(\bar{M} + T\big(2d + \hat{C}_H\big(1 + c_T(C_g^2 + TC_f^2)(1+K)\big)\Big)e^T .$$

Hence, for instance, for $K > 2(1 + eD\bar{M})$ and $T < 1$ so small that also $4edDT < 1$ and $2eT\hat{C}_H D\big(1 + c_T(C_g^2 + TC_f^2)(1+K) < 1$, we have $\sup_{[0,T]} \int_{\mathbb{R}^d} |\cdot|^2 \,\mathrm{d}\bar{\mu}^v \leqslant K$.

We deduce that for K large (depending only on $\bar{M}$ and D) and T sufficiently small (depending only on D, C_g, C_f and $\hat{C}_H$) the map $\Phi\colon \mu \mapsto \bar{\mu}$ is well-defined from $\mathcal{X} := \prod_{v\in V} \mathcal{X}^v$ to itself. Since $\mathcal{X}$ is a convex and compact (by Tychonoff's theorem), Schauder's fixed point theorem guarantees that Φ has a fixed point $\mu = (\mu^v)_{v\in V}$. This proves that, for small T (say $T \leqslant \bar{T}$), system (1.1.1) has a solution, with (uniformly in v) $m^v \in C^{\frac{1}{2}}([0,T];\mathcal{P}_2(\mathbb{R}^d))$ and u^v having the growth and the regularity guaranteed by Proposition 1.3.2. Nevertheless, since $\bar{T}$ does not depend on K, we can iterate such an existence argument finitely many times (enlarging K and keeping the horizon's length constantly equal to $\bar{T}$) to obtain existence on any arbitrary (finite) horizon $[0,T]$. ∎

1.3.2. Linear-Quadratic data. As in Example 1.2.10, suppose that, for all $v \in V$, $H^v(t,p) = \frac{1}{2}\theta^v(t)|p|^2$ and there exist a matrix-valued function $\mathfrak{f}^v\colon [0,T] \to \mathscr{S}(\mathbb{R}^d \times (\mathbb{R}^d)^{\deg v})$ and a matrix $\mathfrak{g}^v \in \mathscr{S}(\mathbb{R}^d \times (\mathbb{R}^d)^{\deg v})$ such that

$$f^v(t,x,y) = \frac{1}{2}\,\mathfrak{f}^v(t)(x,y)\cdot(x,y) \quad \text{and} \quad g^v(x,y) = \frac{1}{2}\,\mathfrak{g}^v(x,y)\cdot(x,y)$$

for all $t \in [0,T]$ and $(x,y) \in \mathbb{R}^d \times (\mathbb{R}^d)^{\deg v}$.

It will be useful to denote the coordinates of $y \in (\mathbb{R}^d)^{\deg v}$ by $y^{v'}$, $v' \sim v$, and to consider $\mathfrak{g}^v$ (as well as $\mathfrak{f}^v(t)$) as a block matrix with entries $\mathfrak{g}^v_{iv'} \in (\mathbb{R}^d)^*$, with $i \in \{1,\dots,d\}$ and $v' \sim v$. Also, we introduce some notation for four super-blocks partitioning $\mathfrak{g}^v$ which will come into play:

$$\mathfrak{g}^v_{pp} := (\mathfrak{g}^v_{ij})_{i,j\in\{1,\dots,d\}}, \qquad \mathfrak{g}^v_{nn} := (\mathfrak{g}^v_{v'v''})_{v',v''\sim v},$$
$$\mathfrak{g}^v_{pn} = (\mathfrak{g}^v_{np})^\mathsf{T} := (\mathfrak{g}^v_{iv'})_{i\in\{1,\dots,d\},\, v'\sim v}\,,$$

p and n standing for "player" and "neighbours", respectively; similar notation will be used for $\mathfrak{f}^v(t)$ as well.

Let us have a look at how our main structural assumptions appear in this setting. Assumption **(L)** is essentially a control on the mixed derivatives $\partial^2_{xy} f^v$ and $\partial^2_{xy} g^v$, so with this notation it reads

$$|\mathfrak{f}^v_{np}(t)| \leqslant \ell^v_f, \qquad |\mathfrak{g}^v_{np}| \leqslant \ell^v_g$$

for all $t \in [0,T]$. Assumption **(C)**, as we have already mentioned, basically forces f^v and g^v to be convex in $x \in \mathbb{R}^d$ at large distances, and so in all of $\mathbb{R}^d$ in the LQ setting; that is, it reads

$$\mathfrak{f}^v_{pp}(t),\ \mathfrak{g}^v_{pp} \geqslant 0$$

for all $t \in [0,T]$. Finally, assumption **(H)** is satisfied if $\theta^v([0,T]) \subseteq [\lambda^v, \Lambda^v] \subset (0,+\infty)$.

As for the definitions of F^v and G^v (recall (1.1.2)), letting $\beta(\mu) := \int_{\mathcal{X}} y\,\mu(\mathrm{d}y)$ for $\mu \in \mathcal{P}_1(\mathcal{X})$, we have

$$F^v(t,x) = \frac{1}{2}\,\mathfrak{f}^v_{pp}(t)x \cdot x + \frac{1}{2}\big(\mathfrak{f}^v_{pn}(t) + \mathfrak{f}^v_{np}(t)^\top\big)\beta(\mu^v_t) \cdot x \\ + \frac{1}{2}\int_{(\mathbb{R}^d)^{\deg v}} \mathfrak{f}^v_{nn}(t)y \cdot y\,\mu^v_t(\mathrm{d}y)$$

and

$$G^v(x) = \frac{1}{2}\,\mathfrak{g}^v_{pp}x \cdot x + \frac{1}{2}\big(\mathfrak{g}^v_{pn} + \mathfrak{g}^{v\,\top}_{np}\big)\beta(\mu^v_T) \cdot x + \frac{1}{2}\int_{(\mathbb{R}^d)^{\deg v}} \mathfrak{g}^v_{nn}y \cdot y\,\mu^v_T(\mathrm{d}y)\,.$$

Then we make the ansatz that

$$u^v(t,x) = \frac{1}{2}\,\mathfrak{u}^v(t)x \cdot x + \mathfrak{v}^v(t) \cdot x + \mathfrak{h}^v(t)\,, \quad (t,x) \in \mathbb{R}^d_T\,, \tag{1.3.7}$$

for some $\mathfrak{u}^v \colon [0,T] \to \mathscr{S}(\mathbb{R}^d)$, $\mathfrak{v} \colon [0,T] \to \mathbb{R}^d$ and $\mathfrak{h}^v \colon [0,T] \to \mathbb{R}$. The HJB equation in (1.1.1) reads

$$\frac{1}{2}\big(-\dot{\mathfrak{u}}^v + \theta^v(\mathfrak{u}^v)^2 - \mathfrak{f}^v_{pp}\big)x \cdot x + \Big(-\dot{\mathfrak{v}}^v + \theta^v\mathfrak{u}^v\mathfrak{v}^v - \frac{1}{2}\big(\mathfrak{f}^v_{pn} + \mathfrak{f}^{v\,\top}_{np}\big)\beta(\mu^v)\Big) \cdot x \\ - \dot{\mathfrak{h}}^v - \operatorname{Tr}\mathfrak{u}^v + \frac{1}{2}\,\theta|\mathfrak{v}^v|^2 - \frac{1}{2}\,\beta^v_{\mathfrak{f}}(\mu^v) = 0\,,$$

where the dot over a symbol indicates the time derivative and we have set

$$\beta^v_{\mathfrak{f}}(\mu^v) := \int_{(\mathbb{R}^d)^{\deg v}} \mathfrak{f}^v_{nn}(t)y \cdot y\,\mu^v_t(\mathrm{d}y)\,,$$

thus yielding

$$\begin{cases} -\dot{\mathfrak{u}}^v + \theta^v(\mathfrak{u}^v)^2 - \mathfrak{f}^v_{pp} = 0 \\ -\dot{\mathfrak{v}}^v + \theta^v\mathfrak{u}^v\mathfrak{v}^v - \frac{1}{2}\big(\mathfrak{f}^v_{pn} + \mathfrak{f}^{v\,\top}_{np}\big)\beta(\mu^v) = 0 \\ -\dot{\mathfrak{h}}^v - \operatorname{Tr}\mathfrak{u}^v + \frac{1}{2}\,\theta|\mathfrak{v}^v|^2 - \frac{1}{2}\,\beta^v_{\mathfrak{f}}(\mu^v) = 0\,. \end{cases} \tag{1.3.8}$$

Note that, having fixed μ^v, the existence of u^v of the form (1.3.7) is equivalent to the solvability of the first equation in (1.3.8) (which is of Riccati type and completely by its own), since the system is decoupled and the other two equations are linear. For example, by standard methods one can explicitly solve the Riccati equation, with terminal condition $\mathfrak{u}^v(T) = \mathfrak{g}^v_{pp}$, if $(\theta^v)^{-1}\mathfrak{f}^v$ is constant and positive definite.

As for the FPK equation in (1.1.1), it turns out to be driven by the drift $-\theta^v(t)(\mathfrak{u}^v(t)x + \mathfrak{v}^v(t))$, hence the corresponding SDE

$$\mathrm{d}X^v_t = -\theta^v(t)(\mathfrak{u}^v(t)X^v_t + \mathfrak{v}^v(t))\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B^v_t$$

is linear. According to our setting, μ^v_t is the law of $Y^v := (X^{v'}_t)_{v' \sim v}$, which solves

$$\mathrm{d}Y^v_t = -\big(\operatorname{diag}(\theta^{v'}\mathfrak{u}^{v'})_{v' \sim v}Y^v_t + (\mathfrak{v}^{v'})_{v' \sim v}\big)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}(B^{v'}_t)_{v' \sim v}\,.$$

It follows that $\beta^v = \beta(\mu^v)$ solves the equation

$$\dot\beta^v + \Delta^v \beta^v + V^v = 0\,, \tag{1.3.9}$$

with $\Delta^v := \operatorname{diag}(\theta^{v'}\mathfrak{u}^{v'})_{v'\sim v}$ and $V^v := (\mathfrak{v}^{v'})_{v'\sim v}$.

Therefore, the strategy to find a solution to the graph-game system is the following. We start by solving the Riccati equation for $\mathfrak{u}^v$, for each $v \in V$. Then, equation (1.3.9) gives us the relation

$$\beta(\mu^v_t) = e^{-\int_0^t \Delta^v}\left(\beta(\mu^v_0) - \int_0^t \sum_{v'\sim v} e^{\int_0^s \theta^{v'}\mathfrak{u}^{v'}}\mathfrak{v}^{v'}(s)\,\mathrm{d}s\right) \tag{1.3.10}$$

At this point it all amounts to proving that there exists a fixed point for the map Φ that associates to $\bar{\mathfrak{v}} \in C^0([0,T];\mathbb{R}^d)^{\#V}$ the vector $(\mathfrak{v}^v)_{v\in V}$ with $\mathfrak{v}^v$ solving the second equation of (1.3.8) where $\beta(\mu^v)$ is given by (1.3.10) with $\mathfrak{v} = \bar{\mathfrak{v}}$. Indeed, once $\mathfrak{v}^v$ is also given for all $v \in V$, we can compute $\beta^v_{\mathfrak{f}}(\mu^v)$ and the third equation of (1.3.8) readily gives $\mathfrak{h}^v$.

Let us admit that, as $\mathfrak{g}^v_{pp}, \mathfrak{f}^v_{pp} \geqslant 0$, we have $\mathfrak{u}^v \geqslant 0$ on $[0,T]$. Standard estimates for linear equations yield

$$|\Phi(\mathfrak{v})^v(t) - \Phi(\bar{\mathfrak{v}})^v(t)| \leqslant (T-t)T\|\mathfrak{f}^v_{np}\|_\infty \sqrt{\deg v}\,\sup_{v'\sim v}|\mathfrak{v}^{v'} - \bar{\mathfrak{v}}^{v'}|\,,$$

so the contraction theorem ensure the existence of a unique fixed point (and thus, a solution), provided that T is small enough with respect to $\mathfrak{f}$ and the graph Γ only. For arbitrary T, we conclude by iterating this argument finitely many times on time horizon with smaller fixed length.

We have thus arrived at the following result.

Theorem 1.3.4. *Let $\Gamma = (V,E)$ be a graph, with $\sup_{v\in V}\deg v < \infty$. For each $v \in V$, let $H^v(t,p) = \frac{1}{2}\theta^v(t)|p|^2$ be continuous and, for all $t \in [0,T]$ and $(x,y) \in \mathbb{R}^d \times (\mathbb{R}^d)^{\deg v}$, let*

$$f^v(t,x,y) = \frac{1}{2}\,\mathfrak{f}^v(t)(x,y)\cdot(x,y) \quad \textit{and} \quad g^v(x,y) = \frac{1}{2}\,\mathfrak{g}^v(x,y)\cdot(x,y)$$

for some $\mathfrak{f}^v \in C^0([0,T];\mathscr{S}(\mathbb{R}^d \times (\mathbb{R}^d)^{\deg v}))$ and $\mathfrak{g}^v \in \mathscr{S}(\mathbb{R}^d \times (\mathbb{R}^d)^{\deg v})$ such that $\partial^2_{xx}f^v, \partial^2_{xx}g^v \geqslant 0$ and $\sup_{v\in V}(|\partial^2_{xy}f^v| + |\partial^2_{xy}g^v|) < \infty$. Suppose that, for each $v \in V$, there exists a solution to

$$\begin{cases} -\dot{\mathfrak{u}}^v + \theta^v(\mathfrak{u}^v)^2 = \partial^2_{xx}f^v & \text{in } [0,T) \\ \mathfrak{u}^v(T) = \partial^2_{xx}g^v\,. \end{cases}$$

Then there exists a unique classical solution to the graph-game system (1.1.1) with u^v of the form (1.3.7).

Proof. Standing all the preceding discussion, we only need to prove that $\mathfrak{u}^v \geqslant 0$ and that the Riccati equation has a unique solution. For $\tau \in [0,T)$ and any $\zeta \in \mathbb{R}^d$, let $\xi\colon [0,T] \to \mathbb{R}^d$ solve $\dot\xi = -\frac{1}{2}\theta^v\mathfrak{u}^v\xi$ with $\xi(\tau) = \zeta$. Then $\frac{\mathrm{d}}{\mathrm{d}t}(\mathfrak{u}^v\xi\cdot\xi) = -\mathfrak{f}^v_{pp}\xi\cdot\xi \leqslant 0$, yielding $\mathfrak{u}^v\zeta\cdot\zeta \geqslant \mathfrak{g}^v_{pp}\xi(T)\cdot\xi(T) \geqslant 0$, and the first conclusion follows by the arbitrariness of ζ. Now, given two solution $\mathfrak{u}^v$ and $\bar{\mathfrak{u}}^v$, we have $\frac{\mathrm{d}}{\mathrm{d}t}|\mathfrak{u}^v - \bar{\mathfrak{u}}^v|^2 = \theta((\mathfrak{u}^v)^2 - (\bar{\mathfrak{u}}^v)^2) : (\mathfrak{u}^v - \bar{\mathfrak{u}}^v) \geqslant 0$ because $\mathfrak{u}^v, \bar{\mathfrak{u}}^v \geqslant 0$, and uniqueness follows. ∎

CHAPTER 2

Linear-quadratic closed-loop games

Contents

2.1. The setting: shift-invariant games

Consider a stochastic differential game with N players, indexed by $i \in [\![N]\!]$, where the state X^i of the i-th player evolves according to the $\mathbb{R}^d$-valued SDE

$$\mathrm{d}X^i_t = \alpha^i_t \,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B^i_t\,.$$

Its cost, in the fixed time horizon $[0,T]$, is given by

$$J^i(\alpha) = \frac{1}{2}\,\mathbb{E}\int_0^T \big(|\alpha^i|^2 + \langle F^i X, X\rangle\big)\,\mathrm{d}t + \langle G^i X_T, X_T\rangle\,, \tag{2.1.1}$$

for some $F^i = f^i \otimes I_d$ and $G^i = g^i \otimes I_d$, with $f^i \in C^0([0,T];\mathscr{S}(N))$ and $g^i \in \mathscr{S}(N)$. The B^i's are independent $\mathbb{R}^d$-valued Brownian motions, and the α^i's are closed-loop controls in feedback form, that is $\alpha^i = \alpha^i(t, X_t)$.

It is known (see for instance [**25**, Section 2.1.4], or [**46**]) that the value functions of the players, $u^i = u^i(t,x)$ with $i \in [\![N]\!]$, $t \in [0,T]$ and $x = (x^0, \dots, x^{N-1}) \in (\mathbb{R}^d)^N$ solve the so-called Nash system of Hamilton–Jacobi PDEs

$$\begin{cases} -\partial_t u^i - \Delta u^i + \frac{1}{2}|D_i u^i|^2 + \sum_{j\neq i} D_j u^j D_j u^i = \bar{F}^i \\ u^i(T,\cdot) = \bar{G}^i \end{cases} \tag{2.1.2}$$

where $\bar{F}^i = \frac{1}{2}\langle F^i\cdot,\cdot\rangle$, $\bar{G}^i = \langle G^i\cdot,\cdot\rangle$ and $i \in [\![N]\!]$. The equilibrium feedbacks are then given by $\alpha^i = -D_i u^i$. Since the dynamic of each player is linear, and the costs are quadratic in the state and control variables, it is well-known that the previous system can be recast into a system of ODEs of Riccati type, by making the ansatz that u^i are quadratic functions of the states. Here, we look for conditions on the solvability of such system, and we focus in particular on properties that are *stable* as the number of players goes to infinity, with the aim of addressing the limit problem with infinitely many players, whenever possible. The Laplacian appears in (2.1.2) by the presence of the independent noises B^i, but it will not actually play any relevant role in our analysis.

For $a \in \mathbb{Z}$ and $b \in \mathbb{Z}_+$, we will use the notation $[a]_b$ to identify the unique natural number $r \in [\![b]\!]$ such that $r = a \bmod b$. The entries of a $N \times N$ matrix will be indexed over $[\![N]\!]$. We will denote by $L^{(N)} \in \mathscr{S}(N)$ the lower shift matrix $\bmod N$, defined by

$$L^{(N)}_{hk} = \delta_{h,[k+1]_N} \quad \forall\, h,k \in [\![N]\!]\,,$$

where δ is the Kronecker symbol.

Our main assumption, on the structure of the running cost, is that f^i and g^i are shift-invariant, in the sense of the following definition.

DEFINITION 2.1.1. A collection of matrices $(M^i)_{i\in[\![N]\!]} \subset \mathscr{S}(N)$ is *shift-invariant* if

$$M^{[i+1]_N} = L^{(N)} M^i L^{(N)\mathsf{T}} \quad \forall\, i \in [\![N]\!]. \tag{2.1.3}$$

Note that this is equivalent to requiring that for all $i \in [\![N]\!]$ and $x \in \mathbb{R}^N$,

$$\begin{aligned}&\langle M^i(x^0,x^1,\dots,x^{N-1}),(x^0,x^1,\dots,x^{N-1})\rangle \\ &\qquad = \langle M^{[i+1]_N}(x^1,\dots,x^{N-1},x^0),(x^1,\dots,x^{N-1},x^0)\rangle\,.\end{aligned}$$

EXAMPLE 2.1.2. A very basic shift-invariant case is that with $g^0 = 0$ and

$$f^0 = w \otimes w, \tag{2.1.4}$$

with

$$w = \frac{1}{\ell-1}\,(\ell-1, \underbrace{-1,\dots,-1}_{\ell-1\text{ times}}, \underbrace{0,\dots,0}_{N-\ell\text{ times}}), \quad \ell \leqslant N;$$

that is, by (2.1.3),

$$\langle F^i X_t, X_t\rangle = \left| X^i_t - \frac{1}{\ell-1}\sum_{j=[i+1]_N}^{[i+\ell-1]_N} X^j_t\right|^2. \tag{2.1.5}$$

When $\ell = N$, the cost is actually of Mean Field type, that is, it penalises the deviation of the private state X^i from the mean of the vector X.

Here f^0 induces an underlying directed circulant graph structure G_ℓ to the problem; indeed, by assumption (2.1.3)

$$A = (f^i_{ij})_{i,j\in[\![N]\!]} - \operatorname{diag}(f^i_{ii})_{i\in[\![N]\!]} = (f^i_{ij})_{i,j\in[\![N]\!]} - I_N$$

can be considered as the asymmetric and circulant adjacency matrix of G_ℓ, so that (2.1.5) reads

$$\langle F^i X_t, X_t\rangle = \left| X_t^i - \frac{1}{\#\{j : (i,j) \in G_\ell\}} \sum_{j:\ (i,j)\in G_\ell} X_t^j \right|^2. \tag{2.1.6}$$

The same is true in the more general case when

$$w = (1, -w_1, \dots, -w_{\ell-1}), \quad \text{with } w_j \geqslant 0 \ \forall j, \ \sum_{j=1}^{\ell-1} w_j = 1; \tag{2.1.7}$$

here the w_j's are regarded as normalised weights, and we have

$$\langle F^i X_t, X_t\rangle = \left| X_t^i - \sum_{j:\ (i,j)\in G_\ell} w_j X_t^j \right|^2,$$

which generalises (2.1.6).

Note finally that a sort of "directionality" is encoded in the above examples, that is, each player i is affected by the "following" ones $j > i$ in the chain. This is not yet important at the current stage, namely we may allow for

$$w = \frac{1}{\ell - 1}\,(\ell - 1, \underbrace{-1, \dots, -1}_{\ell-1 \text{ times}}, \underbrace{0, \dots, 0}_{N-\ell-m \text{ times}}, \underbrace{-1, \dots, -1}_{m \text{ times}}), \quad \ell + m \leqslant N; \tag{2.1.8}$$

It is only from Section 2.2.4 that $m = 0$ will be required.

2.2. The evolutive infinite-dimensional Nash system

The main object of our study is the Nash system

$$\begin{cases} -\partial_t u^i - \Delta u^i + \frac{1}{2}|D_i u^i|^2 + \sum_{j\neq i} D_j u^j D_j u^i = \bar{F}^i & \text{on } (0,T) \times (\mathbb{R}^d)^N \\ u^i(T, \cdot) = \bar{G}^i \end{cases} \tag{2.2.1}$$

When $N = \infty$, we need to be a bit careful about the notion of solution. In this case, $x \in \mathcal{X} = \ell^\infty(\mathbb{Z}; \mathbb{R}^d)$. Then, we mean that (2.2.1) admits a classical solution in the following sense.

Definition 2.2.1. A sequence of $\mathbb{R}$-valued functions $(u^i)_{i\in\mathbb{Z}}$ defined on $[0,T] \times \mathcal{X}$ is a classical solution to the Nash system (2.2.1) on $[0,T] \times \mathcal{X}$ if the following hold:

- **(S1)** each u^i is of class C^1 with respect to $t \in (0,T)$ and C^2 with respect to $x \in \mathcal{X}$, in the Fréchet sense;
- **(S2)** for each $i \in \mathbb{N}$, the Laplacian series $\Delta u^i = \sum_j \Delta_j u^i$ and the series $\sum_{j\neq i} D_j u^j D_j u^i$ uniformly converge on all bounded subsets of $[0,T] \times \mathcal{X}$;
- **(S3)** system (2.2.1) is satisfied pointwise for all $(t,x) \in (0,T) \times \mathcal{X}$;
- **(S4)** $u^i(T, \cdot) = \bar{G}^i$ for all $i \in \mathbb{N}$.

REMARK 2.2.2. Here we made a choice of Banach space, namely $\mathcal{X} = \ell^\infty(\mathbb{Z};\mathbb{R}^d)$. This seems a natural choice for a twofold reason. First, it is the largest ℓ^p space contained in the limit set $(\mathbb{R}^d)^\infty \cong (\mathbb{R}^d)^\mathbb{Z}$, so it seems to provide a quite general setting for a Nash system for infinitely many players. Second, in a linear-quadratic setting as the one we are considering, it is fairly clear that (see the upcoming discussion about the standard ansatz $u^i(t,x) = \frac{1}{2}\langle (c^i(t)\otimes I_d)x, x\rangle_{\mathbb{R}^{Nd}} + \eta^i(t)$) the uniform convergence of Δu^i when $N = \infty$ is strictly related to $(c^i_{jj})_{j\in\mathbb{Z}}$ being summable; therefore the variable x is expected to live in the dual space of ℓ^1, that is ℓ^∞. This is even more evident if one considers the case (and the results we present can indeed be adapted to it) of a more general diffusion such as the one arising from common noise, $\sum_j \Delta_j u^i + \beta \sum_{jk} \operatorname{Tr}(D^2_{jk}u^i)$.

As it is customary in linear-quadratic N-player games, we look for solutions of the form

$$u^i(t,x) = \frac{1}{2}\langle (c^i(t)\otimes I_d)x, x\rangle_{\mathbb{R}^{Nd}} + \eta^i(t), \tag{2.2.2}$$

for some functions $c^i\colon [0,T] \to \mathscr{S}(N)$ such that $c^i(T) = g^i$ and $\eta^i\colon [0,T]\to\mathbb{R}$ which vanish at T. We have $D_j u^i(t,x) = e^{j\mathsf{T}}(c^i(t)\otimes I_d)x$, where $e^j = e_j \otimes I_d$, $\{e_j\}_{j=0}^{N-1}$ being the canonical basis of $\mathbb{R}^N$. Hence, from (2.2.1) we obtain

$$x^\mathsf{T}\bigg(-\frac{1}{2}\,\dot c^i\otimes I_d + \frac{1}{2}(c^i\otimes I_d)e^i e^{i\mathsf{T}}(c^i\otimes I_d) + \sum_{j\in[\![N]\!]\setminus\{i\}} (c^j\otimes I_d)e^j e^{j\mathsf{T}}(c^i\otimes I_d) - \frac{1}{2}F^i\bigg)x = \operatorname{Tr}(c^i\otimes I_d) + \dot\eta^i;$$

that is,

$$x^\mathsf{T}\bigg(\bigg(-\frac{1}{2}\dot c^i + \frac{1}{2}c^i e_i {e_i}^\mathsf{T} c^i + \sum_{j\in[\![N]\!]\setminus\{i\}} c^j e_j {e_j}^\mathsf{T} c^i - \frac{1}{2}f^i\bigg)\otimes I_d\bigg)x = d\operatorname{Tr} c^i + \dot\eta^i.$$

As this must hold for all $x\in\mathbb{R}^{Nd}$, it follows that

$$\dot\eta^i = -d\sum_{j\in[\![N]\!]} c^i_{jj},$$

and

$$-\dot c^i + c^i e_i {e_i}^\mathsf{T} c^i + \sum_{j\in[\![N]\!]\setminus\{i\}} \big(c^j e_j {e_j}^\mathsf{T} c^i + c^i e_j {e_j}^\mathsf{T} c^j\big) = f^i, \tag{2.2.3}$$

as $x^\mathsf{T} A x = 0$ for all x if and only if $A + A^\mathsf{T} = 0$ for any square matrix A.

Now, given the shift invariance of f^i and g^i, one expects a solution to (2.2.3) to enjoy the same property, hence we look for a solution such that

$$c^i = \big(L^{(N)}\big)^i c\,\big(L^{(N)\mathsf{T}}\big)^i, \tag{2.2.4}$$

for some $c\colon [0,T]\to\mathscr{S}(N)$. Clearly, this makes η^i independent of i, as we will have $\eta^i = \eta := \int_\cdot^T \operatorname{Tr} c$. By plugging (2.2.4) into (2.2.3) and letting $i = 0$ one obtains the following system of ODEs for the entries of c:

$$\begin{cases} -\dot{c}_{hk} - c_{0h}c_{0k} + \sum_{j\in[\![N]\!]} \left(c_{0,[h-j]_N}c_{jk} + c_{0,[k-j]_N}c_{jh}\right) = f_{hk} \\ c_{hk}(T) = g_{hk}\,, \end{cases} \tag{2.2.5}$$

where $f := f^0$ and $g := g^0$.

As we are interested in the limit problem of infinitely many players, which we expect to be indexed by $\mathbb{Z}$ since we have an undirected structure, it is convenient to shift the indices in such a way that $i = 0$ "stays in the middle"; that is, we let $i = -N', \dots, N''$, instead of $i = 0, \dots, N$, where, for example, $N' = N'' = (N-1)/2$ if N is odd, and $N' = N/2 = N'' + 1$ if N is even. Therefore, we rewrite system (2.2.5) as

$$\begin{cases} -\dot{c}_{hk} + c_{0h}c_{0k} + \sum_{j\neq 0} (c_{0,h-j}c_{jk} + c_{hj}c_{0,k-j}) = f_{hk} \\ c_{hk}(T) = g_{hk}\,, \end{cases} \qquad i = -N', \dots, N''\,, \tag{2.2.6}$$

where all indices are understood mod N and between $-N'$ and N''. Now it is immediate to identify a convenient limit system by letting $N \to \infty$.

2.2.1. Self-controlled sequences for the discrete convolution.

We can recognise a structure of a cyclic discrete convolution in the sums in (2.2.5); that is,

$$\sum_{j=0}^{N-1} c_{0,[h-j]_N}c_{jk} = (c_{\cdot 0} \star_N c_{\cdot k})_h, \qquad \sum_{j=0}^{N-1} c_{hj}c_{0,[k-j]_N} = (c_{h\cdot} \star_N c_{0\cdot})_k. \tag{2.2.7}$$

We wish to exploit this fact in order to prove existence of the solution to system (2.2.6) for small T, for any (possibly infinite) N.

Our main tool will be the following.

DEFINITION 2.2.3. A nonnegative sequence $\beta \in \ell^2(\mathbb{Z})$ is said to be convolution-self-controlled, or *c-self-controlled*, if $\beta \star \beta \lesssim \beta$; that is,

$$\sum_{j\in\mathbb{Z}} \beta_j\beta_{i-j} \leqslant C\beta_i \quad \forall\, i \in \mathbb{Z}, \tag{2.2.8}$$

for some constant $C > 0$ independent of i.

We will mainly consider positive c-self-controlled sequences in $\ell^1(\mathbb{Z})$ that are even and "weakly decreasing", in the sense that $\beta_i = \beta_{-i}$ and there exists $K > 0$ such that $\beta_j \leqslant K\beta_i$, for all $i, j \in \mathbb{N}$ with $j \geqslant i$. Such sequences, which we will refer to as *regular*, indeed exist, as shown by the following result.

LEMMA 2.2.4. *For any $\varepsilon > 0$, there exists a positive sequence $\beta \in \ell^1(\mathbb{Z})$, with $\|\beta\|_2 < \varepsilon$ and such that $\beta \star \beta \leqslant 4\beta$. In particular, one can choose β of the form*

$$\beta_i = \beta_i(\alpha) := \frac{2\alpha}{\alpha^2 + i^2}\left(1 - (-)^i e^{-\alpha\pi}\right), \quad i \in \mathbb{Z},$$

for some $\alpha = \alpha(\varepsilon) > 0$*, so that* β *is even and* $\beta_j \leqslant \coth(\frac{\alpha\pi}{2})\beta_i$ *for all* $i, j \in \mathbb{N}$ *with* $j \geqslant i$*.*

PROOF. Recall that the Fourier coefficients of a function $f \in L^2((-\pi,\pi))$, given by

$$\hat{f}_j := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ijx}\mathrm{d}x, \quad j \in \mathbb{Z},$$

satisfy the following property: if $f, g \in L^2((-\pi,\pi))$, then $\widehat{fg}_j = (\hat{f} \star \hat{g})_j$ for all $j \in \mathbb{Z}$. Let now be $f_\alpha := e^{-\alpha|\cdot|}$, for any $\alpha > 0$. It is elementary to compute, for each $j \in \mathbb{Z}$,

$$\widehat{f_\alpha}_j = \frac{2\alpha}{\alpha^2 + j^2}\left(1 - (-)^j e^{-\alpha\pi}\right) > 0, \tag{2.2.9}$$

whence $(\widehat{f_\alpha} \star \widehat{f_\alpha})_j = \widehat{f^2_\alpha}_j = \widehat{f_{2\alpha}}_j \leqslant 4\widehat{f_\alpha}_j$, the last inequality being straightforward to check using the explicit expression (2.2.9). Also, by Parseval's identity $\|\widehat{f_\alpha}\|^2_{\ell^2(\mathbb{Z})} = \|f_\alpha\|^2_{L^2(-\pi,\pi)} = \alpha^{-1}(1 - e^{-2\alpha\pi}) \to 0$ as $\alpha \to +\infty$, so that $\beta = \widehat{f_\alpha}$ has the desired properties for any choice of $\alpha = \alpha(\varepsilon)$ sufficiently large. Finally, note that $\beta_j \leqslant \coth(\frac{\alpha\pi}{2})\beta_i$ for all $i, j \in \mathbb{N}$ with $j \geqslant i$. ■

REMARK 2.2.5 (*Variations on a c-self-controlled sequence).* Clearly any positive multiple of a c-self-controlled sequence stays self-controlled. This allows to have self-controlled sequences of arbitrarily large ℓ^∞-norm, although with a larger constant C in (2.2.8). On the other hand, one can also build c-self-controlled sequences which decay exponentially faster; indeed, if β is c-self-controlled, then for any $\gamma > 0$ so is the sequence defined by setting $\tilde{\beta}_i := \beta_i e^{-\gamma|i|}$, with the same implied constant C. This is easily proven as follows. Suppose that $i \leqslant 0$; then

$$\begin{aligned}
(\tilde{\beta} \star \tilde{\beta})_i &= e^{\gamma i}\sum_{j>0}\beta_j\beta_{i-j}e^{-2\gamma j} + e^{\gamma i}\sum_{i\leqslant j\leqslant 0}\beta_j\beta_{i-j} + e^{-\gamma i}\sum_{j<i}\beta_j\beta_{i-j}e^{2\gamma j} \\
&\leqslant e^{\gamma i}\left(\sum_{j>0}\beta_j\beta_{i-j} + \sum_{i\leqslant j\leqslant 0}\beta_j\beta_{i-j}\right) + e^{-\gamma i}\sum_{j<i}\beta_j\beta_{i-j}e^{2\gamma i} \\
&= e^{\gamma i}\sum_{j\in\mathbb{Z}}\beta_j\beta_{i-j} \\
&\leqslant C\beta_i e^{\gamma i}.
\end{aligned}$$

The case $i \geqslant 0$ is analogous.

The next result will be useful to deal with convolution of the form (2.2.7). It essentially states that c-self-controllability is preserved by suitable perturbations, which include all perturbations with compact support.

LEMMA 2.2.6. *Let β be c-self-controlled and let $\theta = (\theta_{hk})_{h,k\in\mathbb{Z}}$ be a nonnegative sequence such that $\theta \lesssim \beta \otimes \beta$.*[1] *Let d be the sequence given by $d := \beta\otimes\beta + \theta$. Then*

$$(d_{\cdot 0} \star d_{\cdot k})_h \lesssim \beta_h \beta_k \quad \forall\, h,k \in \mathbb{Z}.$$

PROOF. It suffices to compute

$$(d_{\cdot 0} \star d_{\cdot k})_h = \beta_0(\beta\star\beta)_h\beta_k + (\theta_{\cdot 0}\star\beta)_h\beta_k + \beta_0(\beta\star\theta_{\cdot k})_h + (\theta_{\cdot 0}\star\theta_{\cdot k})_h \lesssim \beta_h\beta_k. \quad \blacksquare$$

REMARK 2.2.7. A straightforward implication of the above inequality is that $(d_{\cdot 0}\star d_{\cdot k})_h \lesssim d_{hk}$.

2.2.2. Short-time existence for nearsighted interactions. We are now ready for our first existence and uniqueness results (Theorems 2.2.11 and 2.2.12 below), which is a direct consequence of the following proposition.

By a game with *nearsighted* interactions we are meaning that $|f| \vee |g| \leqslant \theta$ pointwise (that is, index-wise) for some $\theta \in \ell^1(\mathbb{Z}^2)$ satisfying the hypotheses of Lemma 2.2.6; said differently, we are meaning that $|f|\vee|g| \lesssim \beta\otimes\beta$ for some regular c-self-controlled β.

REMARK 2.2.8. Any compactly supported sequence $h\colon \mathbb{Z}^2 \to [0,+\infty)$ is nearsighted. Indeed, given a positive c-self-controlled sequence β, define

$$\tilde\beta := \max_{(i,j)\in\operatorname{spt}\beta} \frac{h_{ij}}{\beta_i\beta_j}\cdot\beta\,;$$

then $h \lesssim \tilde\beta\otimes\tilde\beta$.

PROPOSITION 2.2.9. *Let $N\in\mathbb{N}$ and β be a regular c-self-controlled sequence. Let $f \in C^0([0,T];\mathscr{S}(N))$ and $g\in\mathscr{S}(N)$ satisfy $|f|\vee|g| \lesssim \beta\otimes\beta$. Define $\mathcal{C}_N := C^0([0,T])^{2N+1}$ and write $c = (c_{ij})_{i,j=-N}^N \in \mathcal{C}_N$. For $d = \beta\otimes\beta + |g| \vee \sup_{[0,T]}|f|$, set*

$$\mathcal{K}_N := \prod_{i,j=-N}^{N} \big\{ w \in C^0([0,T]) :\ \|w\|_\infty \leqslant 2d_{ij} \big\}. \tag{2.2.10}$$

Let $J_N\colon \mathcal{K}_N \to \mathcal{C}_N$ be the map given, for each $i,j=-N,\dots,N$, by

$$J_N(c)_{ij}(t) := g_{ij} + \int_t^T \Big(f_{ij} + c_{0i}c_{0j} - (c_{\cdot 0}\star_{2N+1} c_{\cdot j})_i - (c_{i\cdot}\star_{2N+1} c_{0\cdot})_j \Big). \tag{2.2.11}$$

Then there exist $T^*>0$, depending on β but independent of N, such that

$$T\leqslant T^* \quad\Longrightarrow\quad J_N(\mathcal{K}_N) \subseteq \mathcal{K}_N.$$

[1] We mean $(\beta\otimes\beta)_{hk} = \beta_h\beta_k$ for all $h,k\in\mathbb{Z}$.

Proof. Let $c \in \mathcal{K}_N$. If $i \geqslant 0$,

$$\begin{aligned}
\|(c_{\cdot 0} \star_N c_{\cdot j})_i\|_\infty &= \Big\| \sum_{k=-N+i}^{N} c_{0,i-k} c_{jk} + \sum_{k=-N}^{-N+i-1} c_{0,i-k-2N-1} c_{jk} \Big\|_\infty \\
&\leqslant 4 \sum_{k=-N+i}^{N} d_{0,i-k} d_{jk} + 4 \sum_{k=-N}^{-N+i-1} d_{0,i-k-2N-1} d_{jk} \\
&\leqslant 4 (d_{\cdot 0} \star d_{\cdot j})_i + 4 (d_{\cdot 0} \star d_{\cdot j})_{i-2N-1} \\
&\lesssim (\beta_i + \beta_{i-2N-1}) \beta_j \,,
\end{aligned}$$

where the last estimate comes from Lemma 2.2.6. As β is regular and $|i-2N-1| = 2N+1-i > i$, we have that $\beta_{i-2N-1} \lesssim \beta_i$, and thus $\|(c_{\cdot 0} \star_N c_{\cdot j})_i\|_\infty \lesssim d_{ij}$, with an implied constant which does not depend on N. The same holds for $-N \leqslant i < 0$ by a symmetrical argument. Analogously, $\|(c_{i \cdot} \star_N c_{0 \cdot})_j\|_\infty \lesssim d_{ij}$ and clearly $\|c_{0i} c_{0j}\|_\infty \leqslant 4 d_{0i} d_{0j} \leqslant 4 (d_{\cdot 0} \star d_{\cdot j})_i \lesssim d_{ij}$. Therefore,

$$\|J_N(c)_{ij}\|_\infty \leqslant |g_{ij}| + CT d_{ij} \leqslant (1+CT) d_{ij}, \tag{2.2.12}$$

where the constant C depends only on β. It follows that for $T > 0$ small enough, depending on β, one has $\|J_N(c)_{ij}\|_\infty < 2 d_{ij}$ for all $i, j = -N, \ldots, N$. ∎

Remark 2.2.10. We stated and proved Proposition 2.2.9 for an odd number of players. This is just a matter of having expressions that look more symmetrical, yet there is no preference about the parity of the number of players, so that the above result holds, *mutatis mutandis*, also if the number of players is even. It is also clear that, with a very much analogous proof, the thesis of Proposition 2.2.9 also holds for $N = \infty$, where one defines, in a natural way,

$$\mathcal{K}_\infty := \prod_{i,j \in \mathbb{Z}} \big\{ w \in C^0([0,T]) : \ \|w\|_\infty \leqslant 2 d_{ij} \big\}$$

and

$$J_\infty(c)_{ij}(t) := g_{ij} + \int_t^T \big(f_{ij} + c_{0i} c_{0j} - (c_{\cdot 0} \star c_{\cdot j})_i - (c_{i \cdot} \star c_{0 \cdot})_j \big).$$

Theorem 2.2.11. *Under the hypotheses of Proposition 2.2.9, there exists $T^* > 0$ such that if $T \leqslant T^*$ then for any $N \in \mathbb{N} \cup \{\infty\}$ there exists a unique smooth solution to system (2.2.6) such that, for any $i, j \in -N', \ldots, N''$ and $m \in \mathbb{N}$,*

$$\Big\| \Big(\frac{\mathrm{d}}{\mathrm{d}t} \Big)^m c_{ij} \Big\|_\infty \lesssim \beta_i \beta_j, \tag{2.2.13}$$

where the implied constants depend only on T^*, f, g and m.

Proof. A fixed point of the map J_N defined in (2.2.11) is a solution. We deal with the case of J_N with $N = \infty$, as the case with $N \in \mathbb{N}$ can be included therein. Note that $\mathcal{K}_\infty$ can be considered as a closed ball of the Banach space

$\ell^\infty_d(\mathbb{Z}^2; C^0([0,T]))$; that is, the space of functions from $\mathbb{Z}^2$ to $C([0,T])$ with finite norm

$$\|\!| \cdot \|\!|_\infty := \sup_{i,j\in\mathbb{Z}} d_{ij}^{-1} \|\cdot_{ij}\|_\infty.$$

We prove that the map J_∞ is a contraction on $\mathcal{K}_\infty$, provided that T is sufficiently small with respect to d. The conclusion will follow from Proposition 2.2.9, Remark 2.2.10 and the contraction mapping theorem; then, once one have a continuous solution, by the structure of equations (2.2.6), one bootstraps its regularity up to C^∞, while estimate (2.2.13) for $m > 1$ follows by induction differentiating (2.2.6) and estimating as in the proof of Proposition 2.2.9. Let now $c, \bar{c} \in \mathcal{K}_\infty$. We have, for i, j fixed,

$$\begin{aligned} (2.2.14) \quad & \big\| J_\infty(\bar{c})_{ij} - J_\infty(c)_{ij} \big\|_\infty \\ & \leqslant T\Big(\big\| (\bar{c}_{\cdot 0} \star_N \bar{c}_{\cdot j})_i - (c_{\cdot 0} \star_N c_{\cdot j})_i \big\|_\infty \\ & \quad + \big\| c_{0i}\bar{c}_{0j} - c_{0i}c_{0j} - \big((\bar{c}_{i\cdot} \star_N \bar{c}_{0\cdot})_j - (c_{i\cdot} \star_N c_{0\cdot})_j \big) \big\|_\infty \Big). \end{aligned}$$

We have

$$\begin{aligned} & \big\| (\bar{c}_{\cdot 0} \star \bar{c}_{\cdot j})_i - (c_{\cdot 0} \star c_{\cdot j})_i \big\|_\infty \\ & \leqslant \sum_{k\in\mathbb{Z}} \Big(\|\bar{c}_{k0}\|_\infty \|\bar{c}_{i-k,j} - c_{i-k,j}\|_\infty + \|\bar{c}_{k0} - c_{k0}\|_\infty \|c_{i-k,j}\|_\infty \Big) \\ & \lesssim d_{ij} \|\!| \bar{c} - c \|\!| \,; \end{aligned}$$

that is, $\|\!| (\bar{c}_{\cdot 0} \star \bar{c}_{\cdot j})_i - (c_{\cdot 0} \star c_{\cdot j})_i \|\!| \lesssim \|\!| \bar{c} - c \|\!|$ and analogously for the second term in (2.2.14). Hence $\|\!| J_\infty(\bar{c}) - J_\infty(c) \|\!| \lesssim T \|\!| \bar{c} - c \|\!|$. ∎

Theorem 2.2.12. *Suppose that f^i and g^i are shift-invariant and there exists a regular c-self-controlled β such that $|f^0| \vee |g^0| \leqslant C(\beta \otimes \beta)$, $C > 0$. There exists $T^* > 0$, depending only on C and β, such that if $T \leqslant T^*$ then there exists a smooth classical solution to the infinite-dimensional Nash system (2.2.1) with $i \in \mathbb{Z}$ on $[0,T] \times \mathcal{X}$. Furthermore, such a solution is unique in the class of functions of the form (2.2.2).*

Proof. Let c be the solution given by Theorem 2.2.11. For $x = (x^i)_{i\in\mathbb{Z}} \in \ell^\infty(\mathbb{Z};\mathbb{R}^d)$, define

$$U(t,x) = \frac{1}{2} \sum_{i,j\in\mathbb{Z}} c_{ij}(t) x^i \cdot x^j + \int_t^T \sum_{i\in\mathbb{Z}} c_{ii}(s)\,\mathrm{d}s,$$

where we denoted by $\cdot$ the standard scalar product on $\mathbb{R}^d$. U is well-defined for $x \in \ell^\infty(\mathbb{Z};\mathbb{R}^d)$, and continuous in t, because the series normally converge thanks to estimate (2.2.13); for the same reason, also

$$t \mapsto \partial_t^k U(t,x) = \frac{1}{2} \sum_{i,j\in\mathbb{Z}} \Big(\frac{\mathrm{d}}{\mathrm{d}t}\Big)^k c_{ij}(t) x^i \cdot x^j - \sum_{i\in\mathbb{Z}} \Big(\frac{\mathrm{d}}{\mathrm{d}t}\Big)^{k-1} c_{ii}(t), \quad k \in \mathbb{N},$$

are well-defined and continuous. Finally, for $h \in \ell^\infty(\mathbb{Z};\mathbb{R}^d)$, note that (omitting the dependence on t)

$$U(x+h) - U(x) = \sum_{i,j\in\mathbb{Z}} c_{ij} x^i h^j + \frac{1}{2}\sum_{i,j\in\mathbb{Z}} c_{ij} h^i \cdot h^j,$$

thus $U(t,\cdot)$ is infinitely many times Fréchet-differentiable in $\ell^\infty(\mathbb{Z};\mathbb{R}^d)$. Define now $u = (u^i)_{i\in\mathbb{Z}}$ by setting

$$u^0 := U, \qquad u^{i+1}(t,x) := u^i(t,\sigma x), \quad i \in \mathbb{Z},$$

where $(\sigma x)_j := x_{j-1}$ for $j \in \mathbb{Z}$. We have

$$D_j u^i(t,x) = D_j[u(t,\sigma^i x)] = D_{j-i} u(t,x) = \sum_{k\in\mathbb{Z}} c_{j-i,k}(t) x^k, \quad i,j \in \mathbb{Z},$$

hence $\sum_{j\in\mathbb{Z}} D_j u^j D_j u^i$ locally uniformly converges by estimate (2.2.13). Hence, by construction, u solves (2.2.1) in the desired sense. ∎

Remark 2.2.13 (*Unimportance of distant players).* What essentially allows the Nash system to have a solution in infinite dimensions is what we call an *unimportance of distant players*, in the sense that the farther a player is from a given one (say the 0-th player) the smaller the impact it has on it is. This is seen in the fact that, on any bounded $\mathcal{B} \subset \mathcal{X}$,

$$\|D_j U\|_{\infty,[0,T]\times\mathcal{B}} = \Big\| \sum_{i\in\mathbb{Z}} c_{ij} x^i \Big\|_{\infty,[0,T]\times\mathcal{B}} \lesssim \beta_j$$

is infinitesimal as $|j| \to \infty$.

2.2.3. Beyond shift-invariance. We have made the shift-invariance hypothesis to reduce our system of infinitely many equations for c to one equation. Nevertheless, the reader should be aware that the above results can be adapted to a more general setting.

One can suppose that we only have a shift-invariant control on the data; that is

$$|f^i_{hk}| \vee |g^i_{hk}| \lesssim \beta_{h-i}\beta_{k-i}. \tag{2.2.15}$$

In this case, the most natural limit of (2.2.3) is indexed over $\mathbb{N}$ and it suffices to replace $\mathcal{K}_\infty$ with

$$\tilde{\mathcal{K}}_\infty := \big\{ w = (w^i_{hk})_{i,h,k\in\mathbb{N}} \in \ell^\infty(\mathbb{N}^3; C^0([0,T])) : \\ \|w^i_{hk}\|_\infty \leqslant 2 d_{|h-i|,|k-i|} \ \forall\, i,h,k \in \mathbb{N} \big\},$$

which is a closed ball in $\ell^\infty_{\tilde{d}}(\mathbb{N}^3; C^0([0,T]))$, letting $\tilde{d}^i_{hk} := d_{|h-i|,|k-i|}$. Then, for instance, one obtains the following result.

THEOREM 2.2.14. *Assume* (2.2.15). *There exists* $T^* > 0$ *such that if* $T \leqslant T^*$ *then for any* $N \in \mathbb{N} \cup \{\infty\}$ *there exists a unique smooth solution to system* (2.2.3) *such that, for any* $i, h, k \in [\![N]\!]$ *and* $m \in \mathbb{N}$,

$$\left\| \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^m c^i_{hk} \right\|_\infty \lesssim \beta_{|h-i|}\beta_{|k-i|},$$

where the implied constants depend only on T^*, f, g *and* m.

Also, one can fix the dimension $N \geqslant 2$ and consider $\beta^N \in (\mathbb{R}_+)^{2N+1}$ given by

$$\beta^N_j := \begin{cases} 1 & \text{if } j = 0 \\ \dfrac{1}{N-1} & \text{if } |j| \in \{1, \dots, N-1\}, \end{cases}$$

which is c-self-controlled in the sense that for $|j| \in [\![N]\!]$

$$(\beta^N \star \beta^N)_j = \sum_{|k| \in [\![N]\!]} \beta^N_k \beta^N_{j-k} = \beta^N_j + \frac{1}{N-1} \sum_{|k| \in [\![N]\!] \setminus \{0\}} \beta^N_{j-k} \leqslant \beta^N_j + \frac{3}{N-1} \leqslant 4\beta^N_j.$$

In this case one can look for a solution to (2.2.3) starting with assumption (2.2.15) with $\beta = \beta^N$; that is,

$$|f^i_{hk}| \vee |g^i_{hk}| \lesssim \begin{cases} 1 & \text{if } h = i = k \\ N^{-1} & \text{if } h \neq i = k \text{ or vice versa} \\ N^{-2} & \text{if } h \neq i \text{ and } k \neq i. \end{cases} \tag{2.2.16}$$

What we get is the following statement.

THEOREM 2.2.15. *Let* $N \in \mathbb{N}$, $N \geqslant 2$ *and assume* (2.2.16). *There exists* $T^* > 0$ *such that if* $T \leqslant T^*$ *then there exists a unique smooth solution to system* (2.2.3) *such that, for any* $i, h, k \in [\![N]\!]$ *and* $m \in \mathbb{N}$,

$$\left\| \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^m c^i_{hk} \right\|_\infty \lesssim \begin{cases} 1 & \text{if } h = i = k \\ N^{-1} & \text{if } h \neq i = k \text{ or vice versa} \\ N^{-2} & \text{if } h \neq i \text{ and } k \neq i. \end{cases}$$

where the implied constants depend only on T^*, f, g *and* m.

Notice that this can be regarded as a result in a nonsymmetric Mean-Field-like setting (which we will consider in Chapter 4), as assumption (2.2.16) is consistent with the fact that we expect the j-th derivative of a *Mean-Field-like* cost for the i-th player to scale by a factor of N^{-1} whenever $j \neq i$.

2.2.4. Long-time existence for shift-invariant directed strongly gathering interactions. It is clear that the previous construction follows the standard Cauchy-Lipschitz local (in time) existence argument, and the existence (and uniqueness) of a solution can be as usual continued up to a maximal time T^*, that is when the quantity $\max_{i,j} \beta_i^{-1}\beta_j^{-1}|c^0_{ij}|$ blows up. So far, we cannot exclude in general that such blow up time T^* is finite.

Upcoming Definition 2.2.16 introduces an additional assumption on the running and terminal costs which will allow us to prove long-time existence of a solution to the infinite-dimensional Nash system. Given $m \leqslant n$, we will identify $M \in \mathscr{S}(m) \subset \mathscr{S}(n)$ by considering $\mathbb{R}^m \simeq \mathbb{R}^m \times \{0\} \subset \mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n$ and then extending $M = 0$ on $\mathbb{R}^{n-m}$; that is, identify $M \in \mathscr{S}(m)$ with $\left(\begin{smallmatrix} M & 0 \\ 0 & 0 \end{smallmatrix}\right) \in \mathscr{S}(n)$. Also, given $M \in \mathscr{S}(m)$ and $M' \in \mathscr{S}(n)$ we will say that $M = M'$ if M' equals the above-mentioned extension of M over $\mathbb{R}^n$.

Definition 2.2.16. Let $\mathscr{M} = (M^{(N)})_{N\in\mathbb{N}}$ be a sequence of matrices, with $M^{(N)} \in \mathscr{S}(N)$. We say that $\mathscr{M}$ is *directed* if there exists $\ell \in \mathbb{N}$ such that $M^{(N)} = M^{(\ell)} \in \mathscr{S}(\ell)$ for all N large enough. Given $\varrho > 1$, we say that $M \in \mathscr{S}(\ell)$ is *ϱ-strongly gathering* if the polynomial

$$\mu(z,w) = \sum_{h,k\in[\![\ell]\!]} M_{hk} z^h w^k, \quad z,w \in \mathbb{C}$$

is such that $\mu(z,0) \notin (-\infty,0)$ if $z \in \varrho\bar{\mathbb{D}}$.

Remark 2.2.17. We anticipate here the main idea behind this definition. We wish to consider the generating function of the coefficients c_{hk} in (2.2.6) with $N = \infty$; that is, formally, $\Xi(t,z,w) := \sum_{h,k\in\mathbb{Z}} c_{hk}(t) z^h w^k$. This is a priori singular on $zw = 0$, nevertheless we are going to show that if $((f_{hk})_{|h|,|k|\leqslant N})_{N\in\mathbb{N}}$ and $((g_{hk})_{|h|,|k|\leqslant N})_{N\in\mathbb{N}}$ are directed, then one can assume that $c_{hk} = 0$ if $h \wedge k < 0$, so that Ξ is analytic and satisfies a "functional Riccati equation" (see (2.2.23) below).

At this point the strong gathering condition, which will be put on f (see upcoming Assumptions ($\star$)), has a twofold utility. First, it ensures that the functional Riccati equation has a solution which is defined in a neighborhood of $(z,w) = (0,0)$ and has real coefficients c_{hk}; this is basically due to the fact that the principal branch of the square root of the function $\mu(\cdot,0)$ associated to f is well-defined. Second, as we require $\varrho > 1$ (and not just $\varrho > 0$), it ensures that the c_{hk}'s will be summable by standard properties of the derivatives of an analytic function.

We also point out that $\varrho > 1$ is not necessary in order for summability to be satisfied, while $\varrho = 1$ is not a priori sufficient. This makes the latter a limiting condition that we have found as interesting as difficult to study in the generality we would have wished for; we have devoted to it a brief discussion in Section 2.6.

Remark 2.2.18. The reader could find a first sight quite strange that the crucial condition in the notion of strong gathering regards $\mu(\cdot,0)$, which sees only the first column $(M_{h0})_{h\in[[\ell]]}$ of the matrix M. From a technical point of view, we claimed in Remark 2.2.17 that it is all we need to solve the functional Riccati equation (2.2.23). From an interpretative point of view, instead, recalling that we will require f to be strongly gathering, it is interesting to note that the coefficients f_{h0} quantify the interaction of the 0-th player with the others; in the shift-invariant setting the 0-th player is basically our reference player for the game, hence we are somehow saying that the solvability of the Nash system is related to a condition which sees only the "direct" interactions between the reference player and the others, and not the "indirect" influence of the interactions between pairs of other players.

REMARK 2.2.19. The term *directed* is related to the fact that the i-th player's cost is affected by the states of the *following* players $j > i$ in the chain, and this fact might not be immediately clear from the previous definition, which just requires the matrices $M^{(N)}$ to be extensions of a fixed matrix, not depending on N. One may then have a look to the matrices of the form $M = w \otimes w$ in (the end of) Example 2.1.2. In particular, $w = w^{(N)}$ in (2.1.8) gives rise to a directed family only when $m = 0$.

Moreover, in the situations described in Example 2.1.2, the associated sequence of matrices as the dimension N diverges is not strongly gathering. Indeed, even though ℓ stays bounded, one has that μ is a polynomial with $\mu(1, 0) = 0$. In fact, as said in Remark 2.2.17, those situations can be seen as limit settings corresponding to taking $\varrho = 1$ and we will comment on this case later on in Section 2.6.

The validity of the strong gathering assumption as it is, can be achieved in different ways; two basic settings are given in the following examples.

EXAMPLE 2.2.20. If we want to stick to a matrix f^0 of the form (2.1.4), in order to have $(f^0(N))_{N\in\mathbb{N}}$ (where N is the dimension) directed with strongly gathering limit one can require that ℓ is independent of N large and

$$\sum_{j=1}^{\ell-1} w_j = 1 - \varepsilon, \quad \varepsilon > 0,$$

so that $\mu(z, 0) \geqslant \varepsilon$ if $|z| \leqslant 1$. Put it differently, it suffices to consider

$$w = (\nu, -w_1, \dots, -w_{\ell-1}), \quad \text{with } w_j \geqslant 0 \ \forall j, \ \sum_{j=1}^{\ell-1} w_j = 1, \text{ and } \nu > 1;$$

this means that we are considering an underlying graph where each node is directly connected with itself as well, and such link has a negative weight. As in this model a positive weight is associated to the tendency, in order to reduce their cost, of each player to get closer to their neighbours, a negative connection with themself is to be interpreted as a drift towards self-annihilation. More prosaically, this means that the state of each player will also tend to the common position $0 \in \mathbb{R}^d$.

Regarding this example, it is also worth pointing out that the common attractive position $0 \in \mathbb{R}^d$ cannot be any other arbitrary point, in the sense that the structure of problem is not invariant under translation of the coordinates. This is due to the fact that we are considering a graph whose nodes have outdegree different from 1.

EXAMPLE 2.2.21. Another setting with directedness and strong gathering is that of

$$q^0 = \nu E^0 + w \otimes w,$$

where w can be given by (2.1.7) with ℓ independent of N, $\nu > 0$ and $E^0 x = x^0$ for all $x = (x^0, \dots, x^{N-1}) \in (\mathbb{R}^d)^N$. Here

$$\langle F^i X_t, X_t \rangle = \nu |X_t^i|^2 + \Big| X_t^i - \sum_{j\colon (i,j) \in G_\ell} w_j X_t^j \Big|^2,$$

and with a translation of the coordinates we can lead to this situation also costs like

$$J^i(\alpha) = \frac{1}{2} \mathbb{E} \int_0^T \bigg(|\alpha^i|^2 + \nu |X_t^i - y|^2 + \Big| X_t^i - \sum_{j\colon (i,j) \in G_\ell} w_j X_t^j \Big|^2 \bigg),$$

for any given $y \in \mathbb{R}^d$. This example, more genuinely than the previous one, shows that the strong gathering assumption entails that we are giving some sort of preference about where the players should aggregate. This will strengthen the attractive structure yielded by a graph satisfying only assumption (2.1.3) (Example 2.1.2), thus providing more stability to our game.

From this section on in this part, the following assumptions will be in force, declined according to suitable choices of the interval I which will be specified (compare, for instance, the statements of Theorems 2.2.24 and 2.5.1).

ASSUMPTIONS ($\star$). The matrices f^i and g^i are shift-invariant and $f^0 = f^0(N)$ and $g^0 = g^0(N)$ are directed with limits $f, g \in \mathscr{S}(\ell)$ for some $\ell \in \mathbb{N}$. The matrix f is ϱ-strongly gathering for some $\varrho > 0$ and g is *compatible on* I with f for some interval $I \subset \mathbb{R}$, in the following sense: given

$$\phi(z, w) := \sum_{h,k=0}^{\ell-1} f_{hk} z^h w^k, \qquad \Psi(z, w) := \sum_{h,k=0}^{\ell-1} g_{hk} z^h w^k$$

and setting

$$\xi := \sqrt{\varphi(\cdot, 0)},^{2} \qquad \psi := \Psi(\cdot, 0),$$

we have

$$\inf_{t \in I} |\psi \tanh(t\xi) + \xi| > 0 \quad \text{on } \varrho\mathbb{D}. \tag{2.2.17}$$

REMARK 2.2.22. Note that for any $t \in \mathbb{R}$, one has that $\psi \tanh(t\xi) + \xi$ is bounded on D, as $\xi(z)t$ can be a singular point only if $\xi(z)^2 = \varphi(z, 0) < 0$, which contradicts that $z \in D$. Also note that condition (2.2.17) holds with $I = \mathbb{R}$ if, e.g., $g = 0$ or $|\psi| \geqslant |\xi|$.

[2] The symbol $\sqrt{\cdot}$ denotes the principal branch of the square root function.

Considering system (2.2.6), we see that for any N sufficiently large with respect to ℓ we have $f_{hk} = 0 = g_{hk}$ if either h or k is negative; hence the limit system as $N \to \infty$ will be given by

$$-\dot{c}_{hk} - c_{0h}c_{0k} + \sum_{j\in\mathbb{Z}} \big(c_{0,h-j}c_{jk} + c_{0,k-j}c_{jh}\big) = f_{hk}, \quad c_{hk}(T) = g_{hk}, \tag{2.2.18}$$

where f_{hk} and g_{hk} extend to $h, k \in \mathbb{Z}^2$ by letting $f_{hk} = 0 = g_{hk}$ if $(h,k) \notin [\![\ell]\!]^2$.[3] Given this system, another reduction is possible. Since $f_{hk} = 0 = g_{hk}$ whenever $h \wedge k < 0$, the assumption that $c_{hk} = 0$ if $h \wedge k < 0$ is not a priori incompatible with the structure of system (2.2.19). Therefore we will look for a solution with this additional property; in this way, the coefficients which are not a priori null, $(c_{hk})_{h,k\in\mathbb{N}}$, will define in a natural way a symmetric operator $c(t) \otimes I_d$ on $\mathcal{X}$ which vanishes on $\ell^\infty(\mathbb{Z}_-;\mathbb{R}^d)$, so that it can also be seen as a trivial extension to $\mathcal{X}$ of a symmetric operator on $\ell^\infty(\mathbb{N};\mathbb{R}^d)$. This reduces the system of ODEs in (2.2.18) to

$$-\dot{c}_{hk} - c_{0h}c_{0k} + \sum_{j=0}^{h} c_{0,h-j}c_{jk} + \sum_{j=0}^{k} c_{0,k-j}c_{jh} = f_{hk}, \quad c_{hk}(T) = g_{hk}, \tag{2.2.19}$$

which will be the object of our following study, and will eventually provide a solution with the particular form presented in the upcoming definition.

Definition 2.2.23. A *quadratic shift-invariant directed (QSD) solution* to (2.2.1) is a classical solution of the form

$$u^i(t,x) = \frac{1}{2}\sum_{h,k\in\mathbb{N}} c_{hk}(t)\langle x^{h+i}, x^{k+i}\rangle_{\mathbb{R}^d} + \int_t^T \operatorname{Tr} c(s)\,\mathrm{d}s, \tag{2.2.20}$$

for some $c\colon [0,T] \to \ell^1(\mathbb{N}^2) \subset \ell^1(\mathbb{Z}^2)$.[4]

Theorem 2.2.24. *Under Assumptions* $(\star)$ *with* $[0,T] \subseteq I$*, there exists a unique QSD solution to* (2.2.1) *on* $[0,T] \times \mathcal{X}$*.*

The following lemmata, in whose statement the hypotheses of Theorem 2.2.24 will be implied, provide the steps of our proof of Theorem 2.2.24.

Lemma 2.2.25. *There exists a unique sequence* $(c_{hk})_{h,k\in\mathbb{N}} \subset C([0,T]) \cap C^\infty((0,T))$*, with* $c_{hk} = c_{kh}$*, which solves the infinite dimensional system* (2.2.19) *on* $[0,T]$*.*

Proof. We perform the change of variables $t \mapsto T - t$ and prove that there exists a unique solution on $[0,T]$ with initial condition $c(0) = g$ to the forward system

$$\dot{c}_{hk} - c_{0h}c_{0k} + \sum_{j=0}^{h} c_{0,h-j}c_{jk} + \sum_{j=0}^{k} c_{0,k-j}c_{jh} = f_{hk} \tag{2.2.21}$$

[3] Note that the limit operator q is exactly the matrix $f \in \mathscr{S}(\ell)$ seen as embedded into the space of symmetric linear operators on $\mathbb{R}^{\mathbb{Z}}$.

[4] A natural immersion $\imath\colon \ell^1(\mathbb{N}^2) \hookrightarrow \ell^1(\mathbb{Z}^2)$ is given by $\imath(x)^i = x^i$ if $i \in \mathbb{N}$ and $\imath(x)^i = 0$ otherwise.

which smoothly extends to $\mathbb{R}$. Then, the solution to (2.2.19) on $[0,T]$ with $c(T)=0$ will be given by the restriction to $[0,T]$ of $\hat{c}(T-\cdot)$, where $\hat{c}$ is the unique solution to (2.2.21) on $\mathbb{R}$ with $\hat{c}(0)=g$. Note that $\hat{c}_{00}$ is the solution to the Riccati equation $\dot{\hat{c}}_{00}+\hat{c}_{00}^2=f_{00}$, where $f_{00}=\phi(0,0)>0$ by strong gathering, hence

$$\hat{c}_{00}(t)=\nu\,\frac{\nu\sinh(\nu t)+\mathfrak{g}\cosh(\nu t)}{\mathfrak{g}\sinh(\nu t)+\nu\cosh(\nu t)},\quad \mathfrak{g}:=g_{00},\ \nu:=\sqrt{f_{00}}\,.$$

All the other $\hat{c}_{hk}$'s satisfy first-order ODEs with coefficient which are second-order polynomials depending only on f_{hk} and $\hat{c}_{h'k'}$ with $(h',k')\prec(h,k)$, where $\prec$ denotes the strict Pareto preference.[5] Therefore, existence and uniqueness of the solution to the infinite system may be proved by induction. Indeed, suppose $(\hat{c}_{0k'})_{0\leqslant k'<k}\subset C^\infty(\mathbb{R})$ are given, and note that

$$\dot{\hat{c}}_{0k}+2\hat{c}_{00}\hat{c}_{0k}=f_{0k}-\sum_{j=1}^{k-1}\hat{c}_{0,k-j}\hat{c}_{0j};$$

then $\hat{c}_{0k}$ is unique and smooth as well. This proves the existence and uniqueness of $\hat{c}_{0k}$ for all $k\in\mathbb{N}$; then, looking at this argument as the base step for a new induction over h, one proves analogously the existence and uniqueness of $\hat{c}_{hk}$ for all $h\in\mathbb{N}$ and any $k\in\mathbb{N}$. Finally, $c_{hk}=c_{kh}$ since equations (2.2.5) are invariant with respect to the swap $(h,k)\mapsto(k,h)$. ∎

The arguments below show that the coefficients c_{hk} can be thought as derivatives of a generating function $\hat{\Xi}$, which will play a fundamental role in the long-time analysis.

LEMMA 2.2.26. *The solution to* (2.2.19) *on* $[0,T]$ *is given by*

$$c_{hk}(t)=\frac{1}{h!\,k!}\left.\frac{\partial^{h+k}}{\partial z^h\partial w^k}\right|_{(0,0)}\hat{\Xi}(T-t),\quad \forall\, t\in[0,T] \tag{2.2.22}$$

for some function $\hat{\Xi}\colon I\times\varrho\mathbb{D}^2\to\mathbb{C}$, *of class* C^∞ *with respect to* $t\in I$ *and analytic in* $\varrho\mathbb{D}^2$, *such that* $\hat{\Xi}(\cdot,z,w)=\hat{\Xi}(\cdot,w,z)$ *and* $\hat{\Xi}(0,\cdot,\cdot)=\Psi$.

PROOF. Suppose that, for all $z,w\in D$ fixed, there exists a solution $\hat{\Xi}$ on I to

$$\partial_t\hat{\Xi}(t,z,w)-\hat{\Xi}(t,z,0)\hat{\Xi}(t,0,w)+\big(\hat{\Xi}(t,z,0)+\hat{\Xi}(t,0,w)\big)\hat{\Xi}(t,z,w)=\phi(z,w), \tag{2.2.23}$$

with the desired properties of smoothness, invariant with respect the swap $(z,w)\mapsto(w,z)$ and such that $\hat{\Xi}(0,z,w)=\Psi(z,w)$. Then by taking the derivatives $\partial_z^h\partial_w^k|_{(0,0)}$ one recovers equation (2.2.21), and thus the coefficients given by (2.2.22) satisfy (2.2.19) on $[0,T]$.[6] To see that (2.2.23) admits such a solution, note that $\hat{\Xi}(t,z,0)$

[5] That is, $(h',k')\prec(h,k)$ if and only if $h'\leqslant h$ and $k'\leqslant k$ with at least one strict inequality.

[6] In other words, we are saying that $\partial_z^h\partial_w^k\hat{\Xi}(t,0,0)=h!\,k!\,\hat{c}_{hk}(t)$, where the coefficients $\hat{c}_{hk}$ are those in the proof of Lemma 2.2.25.

solves the Riccati equation $\partial_t\hat{\Xi}(t,z,0)+\hat{\Xi}(t,z,0)^2=\phi(z,0)$; hence,

$$\hat{\Xi}(t,z,0)=\xi(z)\,\frac{\xi(z)\sinh(\xi(z)t)+\psi(z)\cosh(\xi(z)t)}{\psi(z)\sinh(\xi(z)t)+\xi(z)\cosh(\xi(z)t)},\quad \psi:=\Psi(\cdot,0).$$

Note that letting

$$\mathcal{E}(t,z;\zeta_1,\zeta_2):=\zeta_1(z)\sinh(\xi(z)t)+\zeta_2(z)\cosh(\xi(z)t)$$

one can write

$$\hat{\Xi}(t,z,0)=\xi(z)\,\frac{\mathcal{E}(t,z;\xi,\psi)}{\mathcal{E}(t,z;\psi,\xi)}=\frac{\partial}{\partial t}\log\mathcal{E}(t,z;\psi,\xi).$$

For any $t\in I$ fixed, this function is well-defined for $z\in\varrho\mathbb{D}$ and analytic therein by the compatibility assumption (2.2.17). At this point, (2.2.23) becomes a first-order ODE in $t\in I$, for all $z,w\in\varrho\mathbb{D}$ fixed, whose solution is given by

(2.2.24)
$$\hat{\Xi}(t,z,w)=\frac{\xi(z)\xi(w)}{\mathcal{E}(t,z;\psi,\xi)\mathcal{E}(t,w;\psi,\xi)}\Bigg(\Psi(z,w)+\int_0^t\Big(\mathcal{E}(t,z;\xi,\psi)\mathcal{E}(t,w;\xi,\psi)+\frac{\phi(z,w)}{\xi(z)\xi(w)}\mathcal{E}(t,z;\psi,\xi)\mathcal{E}(t,w;\psi,\xi)\Big)\Bigg).$$

Note that $\hat{\Xi}(t,\cdot,\cdot)$ is well-defined for $z,w\in\varrho\mathbb{D}$ by the same argument we applied to $\hat{\Xi}(t,\cdot,0)$. Also, it is trivial that $\hat{\Xi}(\cdot,z,w)\in C^\infty(I)$, and by differentiating under the integral sign one proves the analyticity of $\hat{\Xi}(t,\cdot,\cdot)$. ∎

Remark 2.2.27. Let $\tilde{\mathcal{E}}(t,z;\zeta_1,\zeta_2):=\mathcal{E}(t,z;\zeta_2,\zeta_1)$. As, omitting the dependence on ζ_1 and ζ_2 in $\mathcal{E}$, we have $\frac{\partial}{\partial t}\mathcal{E}(t,z)=\xi(z)\tilde{\mathcal{E}}(t,z)$, it is easy to see that

$$\frac{\partial}{\partial t}\big(\tilde{\mathcal{E}}(t,z)\mathcal{E}(t,w)\pm\mathcal{E}(t,z)\tilde{\mathcal{E}}(t,w)\big)=(\xi(z)\pm\xi(w))\big(\mathcal{E}(t,z)\mathcal{E}(t,w)\pm\tilde{\mathcal{E}}(t,z)\tilde{\mathcal{E}}(t,w)\big)$$

and thus

$$\int_0^t\mathcal{E}(s,z)\mathcal{E}(s,w)\,\mathrm{d}s=\frac{1}{2}\Big(\frac{\tilde{\mathcal{E}}(\cdot,z)\mathcal{E}(\cdot,w)+\mathcal{E}(\cdot,z)\tilde{\mathcal{E}}(\cdot,w)}{\xi(z)+\xi(w)}+\frac{\tilde{\mathcal{E}}(\cdot,z)\mathcal{E}(\cdot,w)-\mathcal{E}(\cdot,z)\tilde{\mathcal{E}}(\cdot,w)}{\xi(z)-\xi(w)}\Big)\Big|_0^t.$$

Letting $(\zeta_1,\zeta_2)\in\{(\xi,\psi),(\psi,\xi)\}$ one can then compute the integral in (2.2.24). Set

(2.2.25)
$$\sigma^\pm(t,z,w)=\frac{\mathcal{F}(t,z)+\mathcal{F}(t,w)}{\xi(z)+\xi(w)}\pm\frac{\mathcal{F}(t,z)-\mathcal{F}(t,w)}{\xi(z)-\xi(w)},\quad \mathcal{F}(t,\cdot):=\frac{\mathcal{E}(t,\cdot;\xi,\psi)}{\tilde{\mathcal{E}}(t,\cdot;\xi,\psi)};$$

then

$$2\,\hat{\Xi}(t,z,w)=\tilde{\Psi}(z,w)+\phi(z,w)\sigma^+(t,z,w)+\xi(z)\xi(w)\sigma^-(t,z,w),\tag{2.2.26}$$

where

$$\tilde{\Psi}(z,w)=\Psi(z,w)-\phi(z,w)\sigma^+(0,z,w)+\xi(z)\xi(w)\sigma^-(0,z,w).$$

LEMMA 2.2.28. *Let $c = (c_{hk})_{h,k\in\mathbb{N}}$ be the solution to system (2.2.19) on $[0,T]$. Then, for any $r \in (1,\varrho)$,*

$$\|c_{hk}\|_{\infty;[0,T]} \leqslant \frac{K(r,T)}{r^{h+k}} \quad \forall\, h,k \in \mathbb{N}, \tag{2.2.27}$$

for some constant $K(r,T)$ depending only on T, r, f and g. In particular, $c \in C^0([0,T];\ell^1(\mathbb{N}^2))$ and the same is true for $\dot{c}$.[7]

PROOF. By the strong gathering assumption, the function $\hat{\Xi}(t,\cdot,\cdot)$ given in Lemma 2.2.26 is analytic in a neighbourhood of

$$\mathcal{Q}_r := \overline{\mathbb{D}}_r \times \overline{\mathbb{D}}_r\,, \tag{2.2.28}$$

with $r \in (1,\varrho)$. Then, by Cauchy's theorem on derivatives,

$$\|c_{hk}\|_{\infty;[0,T]} \leqslant \frac{1}{r^{h+k}} \max_{[0,T]\times\partial\mathcal{Q}_r} |\hat{\Xi}|. \tag{2.2.29}$$

This proves that $c \in C^0([0,T];\ell^1(\mathbb{N}^2))$, and further regularity is easily proven by induction by exploiting system (2.2.19). ∎

LEMMA 2.2.29. *Let c be the solution to (2.2.19) on $[0,T]$. The functions u^i given by (2.2.20) are well-defined for $(t,x) \in [0,T]\times\mathcal{X}$, differentiable with respect to t, and twice Fréchet-differentiable with respect to x.*

PROOF. Since a suitable shift of coordinates in $\mathcal{X}$ transforms $u^i(t,\cdot)$ into $u^0(t,\cdot)$, it is sufficient to prove the result for u^0. The existence of a solution follows from Lemma 2.2.28, and the differentiability with respect to t follows from the same lemma and the fundamental theorem of calculus. The differentiability with respect to x follows again from Lemma 2.2.28, as it is trivial to see that for any $h \in \mathcal{X}$ small one has

$$u^0(t,x+h) - u^0(t,x) - \langle (c(t)\otimes I_d)x, h\rangle_{\mathcal{X}} - \frac{1}{2}\langle (c(t)\otimes I_d)h,h\rangle_{\mathcal{X}} = 0,$$

where, with an obvious notation, $\langle (c(t)\otimes I_d)y,z\rangle_{\mathcal{X}} = \sum_{h,k\in\mathbb{N}} c_{hk}(t)\langle y^h, z^k\rangle_{\mathbb{R}^d}$. ∎

LEMMA 2.2.30. *Let u^i be given by (2.2.20), where c is the solution to (2.2.19) on $[0,T]$. Let $\mathcal{B}_R$ be the closed ball of radius R in $\mathcal{X}$. Then, for any $r \in (1,\varrho)$,*

$$\sum_{j\in\mathbb{Z}\setminus\{i\}} \left\|D_j u^j D_j u^i\right\|_{\infty;[0,T]\times\mathcal{B}_R} \leqslant \frac{1}{r-1}\left(\frac{RKr}{r-1}\right)^2,$$

where $K = K(r,T)$ is the constant appearing in Lemma 2.2.28.

PROOF. We have $D_j u^i(t,x) = 0$ if $j < i$, and $D_j u^i(t,x) = \sum_{h\in\mathbb{N}} c_{j-i,h}(t)x^h$ if $j \geqslant i$. Then, for $x \in \mathcal{B}_R$, and $r \in (1,\varrho)$ fixed, estimate (2.2.27) yields

$$\left|D_j u^j(t,x) D_j u^i(t,x)\right| \leqslant R^2 \sum_{h,k\in\mathbb{N}} |c_{0h}(t)||c_{j-i,k}(t)| \leqslant \left(\frac{RKr}{r-1}\right)^2 \frac{1}{r^{j-i}},$$

[7] And for all higher order derivatives.

for all $t \in [0,T]$. The thesis now follows by computing $\sum_{j>i} r^{i-j}$. ∎

Since, by construction, choosing u^i as in (2.2.20) satisfying **(S1)** and **(S2)** in Definition 2.2.1 yields a classical solution, the proof of Theorem 2.2.24 is complete.

2.3. Almost-optimal controls for the N-player game

We did not prove long-time existence of a solution to the Nash system for the N-player game, with $N > \ell$ finite; nevertheless, in Corollary 2.3.4 below we show that on any fixed horizon $[0,T]$ the infinite-dimensional optimal control for the i-th player, given by $\bar{\alpha}^{*i}(t, X_t) := \sum_{j \geqslant i} c_{0,j-i}(t) X_t^j$, if suitably "projected" onto $(\mathbb{R}^d)^N$, provides an ε-Nash equilibrium for the N-player game, with $\varepsilon \to 0$ as $N \to \infty$. In claiming so, we consider classes of admissible controls in the sense of the following definition.

Definition 2.3.1. Let $R, L \geqslant 0$. A control $\alpha \colon [0,T] \times (\mathbb{R}^d)^N \to (\mathbb{R}^d)^N$ in feedback form belongs to $\mathcal{A}_{R,L}$ if, for all $t \in [0,T]$, $x, y \in (\mathbb{R}^d)^N$,

$$|\alpha(t,0)| \leqslant R, \qquad |\alpha(t,x) - \alpha(t,y)| \leqslant L|x-y|.$$

The Lipschitz constant L is said to be *admissible* if $L \geqslant \|c_{0\cdot}\|_{C^0([0,T];\ell^1(\mathbb{N}))}$.

Remark 2.3.2. For such controls, it is known (cf., e.g., [4, Theorems 9.1 and 9.2]) that

$$\begin{cases} \mathrm{d}X_t = \alpha(t, X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t & t \in [0,T] \\ X_0 = x_0 \in (\mathbb{R}^d)^N \end{cases} \tag{2.3.1}$$

has a unique solution, satisfying $\mathbb{E}\sup_{[0,T]} |X|^2 \leqslant C$ where C is a locally bounded function of R, L and T, directly proportional to $1 + |x_0|^2$.

Theorem 2.3.3. *Consider the N-player game on $[0,T]$ with time evolution of the state of the players given by (2.3.1) and costs given by (2.1.1) with $f^0, g^0 \geqslant 0$. Let Assumptions ($\star$) be in force with $[0,T] \subseteq I$. Let c solve (2.2.19) and define the control α^* by*

$$-\alpha^{*i}(t, X_t^0, \dots, X_t^{N-1}) := \sum_{j=0}^{N-i-1} c_{0j}(t) X_t^{j+i} + \sum_{j=N-i}^{N-1} c_{0j}(t) X_t^{j+i-N}, \qquad i \in [\![N]\!].$$

Then, for any $R \geqslant 0$ and admissible L, for any $i \in [\![N]\!]$ and any $(\alpha^{*,-i}, \psi) \in \mathcal{A}_{R,L}$,[8]

$$J^i(\alpha^*) \leqslant J^i((\alpha^{*,-i}, \psi)) + \hat{C}(\delta^M + (\delta^{-M} + N)\delta^N) \qquad \forall\, \delta \in (\varrho^{-1}, 1),\ \forall M \geqslant \ell,$$

where ℓ is the dimension appearing in Assumptions ($\star$) and the constant $\hat{C}$ is a locally bounded function of R, L, T and δ, directly proportional to $1 + |x_0|^2$.

[8] With this notation we mean that all components are those of α^* but the i-th one, which is a suitable $\mathbb{R}^d$-valued function $\psi = \psi(t,x)$. Note that $\alpha^* \in \mathcal{A}_{R,L}$ for any R and admissible L.

Proof. Since f^i and g^i are shift-invariant and α^* is linear in the state variable with $D\alpha^*$ circulant, without loss of generality we can prove the thesis for $i = 0$. We denote by X^* the solution to (2.3.1) when $\alpha = \alpha^*$ and by X the solution to (2.3.1) when $\alpha = (\psi, \alpha^{*1}, \dots, \alpha^{*,N-1}) =: \hat{\alpha}^*$. We wish to estimate from above the quantity

$$J^0(\alpha^*) - J^0(\hat{\alpha}^*) = \frac{1}{2}\,\mathbb{E}\bigg[\int_0^T \big(|\alpha^{*0}(t, X_t^*)|^2 - |\psi(t, X_t)|^2 + F(X_t^*) - F(X_t)\big)\,\mathrm{d}t + G(X_T^*) - G(X_T)\bigg], \tag{2.3.2}$$

where $F := \langle F^0 \cdot, \cdot\rangle$ and $G := \langle G^0 \cdot, \cdot\rangle$. Let u be the QSD solution to (2.2.1) on $[0,T] \times \mathcal{X}$. Since G is convex and $DG = Du^0(T, \cdot)$,

$$\begin{aligned}
G(X_T^*) - G(X_T) &\leqslant \langle Du^0(T, X_T^*), X_T^* - X_T\rangle_{\mathbb{R}^{dN}} \\
&= \sum_{j=0}^{N-1}\sum_{k=0}^{\ell-1} c_{jk}(T)\langle X_T^{*k}, X_T^{*j} - X_T^j\rangle_{\mathbb{R}^d} \\
&= \sum_{j=0}^{N-1}\sum_{k=0}^{\ell-1}\bigg(\int_0^T \dot{c}_{jk}(t)\langle X_t^{*k}, X_t^{*j} - X_t^j\rangle_{\mathbb{R}^d}\,\mathrm{d}t \\
&\qquad + c_{jk}(t)\langle \mathrm{d}X_t^{*k}, X_t^{*j} - X_t^j\rangle_{\mathbb{R}^d} + c_{jk}(t)\langle X_t^{*k}, \mathrm{d}(X_t^{*j} - X_t^j)\rangle_{\mathbb{R}^d}\bigg).
\end{aligned}$$

Note that we can replace ℓ with any $M \geqslant \ell$ because $c_{jk}(T) = g_{jk}(T) = 0$ if $k > \ell$. As c solves (2.2.19) we obtain

$$\begin{aligned}
G(X_T^*) - G(X_T) \leqslant \sum_{j=0}^{N-1}\sum_{k=0}^{M-1}\bigg(&\int_0^T \sum_{h=1}^{j} c_{0,j-h}(t)c_{hk}(t)\langle X_t^{*k}, X_t^{*j} - X_t^j\rangle_{\mathbb{R}^d}\,\mathrm{d}t \\
&+ \int_0^T \sum_{h=0}^{k} c_{0,k-h}(t)c_{hj}(t)\langle X_t^{*k}, X_t^{*j} - X_t^j\rangle_{\mathbb{R}^d}\,\mathrm{d}t \\
&- f_{jk}\int_0^T \langle X_t^{*k}, X_t^{*j} - X_t^j\rangle_{\mathbb{R}^d}\,\mathrm{d}t \\
&+ \int_0^T c_{jk}(t)\langle \alpha^{*k}(t, X_t^*), X_t^{*j} - X_t^j\rangle_{\mathbb{R}^d}\,\mathrm{d}t \\
&+ \int_0^T c_{jk}(t)\langle X_t^{*k}, \alpha^{*j}(t, X_t^*) - \hat{\alpha}^{*j}(t, X_t)\rangle_{\mathbb{R}^d}\,\mathrm{d}t\bigg) + Z_T\,,
\end{aligned}$$

where $(Z_t)_{0\leqslant t\leqslant T}$ is a martingale starting from 0. Straightforward computations show that, omitting the dependence on t,

$$-\sum_{k=0}^{M-1} c_{jk}\alpha^{*k}(X^*) = \sum_{k=0}^{M-1}\sum_{h=0}^{k} c_{0,k-h}c_{hj}X^{*k} + \sum_{h=N-M+1}^{N-1}\sum_{k=N-h}^{M-1} c_{0h}c_{jk}X^{*,h+k-N}$$

and

$$-\sum_{j=0}^{N-1} c_{jk}(\alpha^{*j}(X^*) - \hat{\alpha}^{*j}(X)) = -c_{0k}(\alpha^{*0}(X^*) - \psi(X)) + \sum_{j=0}^{N-1}\sum_{h=1}^{j} c_{hk}c_{0,j-h}(X^{*j} - X^j) + \sum_{h=N}^{2N-2}\sum_{j=h-N+1}^{N-1} c_{jk}c_{0,h-j}(X^{*,h-N} - X^{h-N})\,;$$

therefore,

$$G(X_T^*) - G(X_T) \leqslant -\sum_{k=0}^{M-1}\int_0^T c_{0k}(t)\langle X_t^{*k}, \alpha^{*0}(X^*) - \psi(X)\rangle_{\mathbb{R}^d}\,\mathrm{d}t - \sum_{j=0}^{N-1}\sum_{k=0}^{M-1} f_{jk}\int_0^T \langle X_t^{*k}, X_t^{*j} - X_t^j\rangle_{\mathbb{R}^d}\,\mathrm{d}t - \int_0^T \mathcal{E}_t\,\mathrm{d}t\,, \tag{2.3.3}$$

where we have set

$$\mathcal{E} := \sum_{j=0}^{N-1}\sum_{h=N-M+1}^{N-1}\sum_{k=N-h}^{M-1} c_{0h}c_{jk}\langle X^{*,h+k-N}, X^{*j} - X^j\rangle_{\mathbb{R}^d} + \sum_{k=0}^{M-1}\sum_{h=N}^{2N-2}\sum_{j=h-N+1}^{N-1} c_{jk}c_{0,h-j}\langle X^{*k}, X^{*,h-N} - X^{h-N}\rangle_{\mathbb{R}^d}\,.$$

Note now that by the convexity of $\frac{1}{2}|\cdot|^2$, omitting the dependence on t,

$$\frac{1}{2}|\alpha^{*0}(X^*)|^2 - \frac{1}{2}|\psi(X)|^2 - \langle \alpha^{*0}(X^*), \alpha^{*0}(X^*) - \psi(X)\rangle_{\mathbb{R}^d} \leqslant 0 \tag{2.3.4}$$

and by the convexity of F

$$F(X^*) - F(X) - \langle DF(X^*), X^* - X\rangle_{\mathbb{R}^d} \leqslant 0. \tag{2.3.5}$$

Using (2.3.3), (2.3.4) and (2.3.5) in (2.3.2) we obatin

$$J^0(\alpha^*) - J^0(\hat{\alpha}^*) \leqslant \mathbb{E}\Big[\sum_{k=\ell}^{N-1}\int_0^T c_{0k}(t)\langle X_t^{*k}, \alpha^{*0}(X^*) - \psi(X)\rangle_{\mathbb{R}^d}\,\mathrm{d}t - \int_0^T \mathcal{E}_t\,\mathrm{d}t\Big]\,. \tag{2.3.6}$$

As, whenever $\{Y, Z\} \subseteq \{X^*, X\}$,

$$\sup_{j,k\in[\![N]\!]} \mathbb{E}\int_0^T |\langle Y_t^k, Z_t^j\rangle_{\mathbb{R}^d}|\,\mathrm{d}t \leqslant TC\,,$$

where $C = C(|x_0|, R, L, T)$ is the constant appearing in Remark 2.3.2, we have

$$\left| \mathbb{E} \int_0^T \mathcal{E}_t \,\mathrm{d}t \right| \leqslant 2TC \Big(\|c\|_{C^0([0,T];\ell^1(\mathbb{N}^2))} \sum_{h \geqslant N-M} \|c_{0h}\|_{\infty;[0,T]} + \sum_{h \geqslant N} \sum_{j=0}^{N-1} \|c_{j\cdot}\|_{C^0([0,T];\ell^1(\mathbb{N}))} \|c_{0,h-j}\|_{\infty;[0,T]} \Big),$$

so that by Lemma 2.2.28, for any $\delta \in (\varrho^{-1}, 1)$, there exists $K > 0$ such that

$$\left| \mathbb{E} \int_0^T \mathcal{E}_t \,\mathrm{d}t \right| \leqslant \tilde{C}(\delta^{-M} + N)\delta^N, \quad \tilde{C} := \frac{2TCK^2}{(1-\delta)^3}\,. \tag{2.3.7}$$

On the other hand,

$$\sup_k \mathbb{E} \int_0^T |\langle X_t^{*k}, \alpha^{*0}(X^*) - \psi(X) \rangle_{\mathbb{R}^d}| \leqslant 2R\sqrt{TC} + 2LC\,, \tag{2.3.8}$$

where we used again that $\alpha^*, \hat{\alpha}^* \in \mathcal{A}_{R,L}$. Therefore, from (2.3.6), (2.3.7), (2.3.8) and Lemma 2.2.28 we obtain

$$J^0(\alpha^*) - J^0(\hat{\alpha}^*) \leqslant \hat{C}(\delta^M + (\delta^{-M} + N)\delta^N)\,, \quad \hat{C} := 2(R\sqrt{TC} + LC)\frac{K}{1-\delta} \vee \tilde{C}\,.$$

This concludes the proof. ∎

Corollary 2.3.4. *Let $\Omega \subset \mathbb{R}^d$ bounded, and assume the same as in Theorem 2.3.3, but with $x_0 \in \Omega^N$. Let $R \geqslant 0$ and L be admissible (in the sense of Definition 2.3.1); consider as admissible those controls belonging to $\mathcal{A}_{R,L}$. Then, for any $\varepsilon > 0$ there exists $N_0 = N_0(\varepsilon, R, L, \varrho, \ell, \Omega)$ such that if $N \geqslant N_0$ then the control α^* provides an ε-Nash equilibrium of the game.*

Proof. It suffices to require that

$$\hat{C}(\delta^M + (\delta^{-M} + N)\delta^N) \leqslant \varepsilon \quad \forall\, N \geqslant N_0, \tag{2.3.9}$$

where, for instance, one sets $\delta = \frac{1}{2}(1 + \varrho^{-1})$. Choose $M = M(N)$ such that $M \to \infty$ and $M = o(N)$ as $N \to \infty$; for example, $M = \lfloor \sqrt{N} \rfloor$ for all $N > \ell^2$. Since $x_0 \in \Omega^N$ there exists a constant $\hat{C}'$, independent of N, such that $\hat{C} \leqslant N\hat{C}'$. Then the conclusion follows from the fact that the left-hand side of (2.3.9) goes to 0 as $N \to \infty$. ∎

2.4. The ergodic Nash system

Consider now the related ergodic problem, with costs

$$\bar{J}^i(\alpha) = \liminf_{T \to +\infty} \frac{1}{2T} \mathbb{E} \int_0^T \big(|\alpha^i|^2 + \langle F^i X, X \rangle \big),$$

where the dynamics and the assumptions on F^i are the same as before. The corresponding Nash system reads

$$\lambda^i - \Delta v^i + \frac{1}{2}|D_i v^i|^2 + \sum_{j \neq i} D_j v^j D_j v^i = \bar{F}^i \quad \text{in } \mathcal{Y}, \tag{2.4.1}$$

where $\mathcal{Y}$ is either $(\mathbb{R}^d)^N$, for the N-player game, or $\mathcal{X}$, for the limit game with infinitely many players. In the latter case we give notions of classical solution and QSD solution which are analogous as those in the previous section.

DEFINITION 2.4.1. A sequence of pairs $((\lambda^i, v^i))_{i\in\mathbb{Z}}$ of real numbers λ^i and $\mathbb{R}$-valued functions v^i defined on $\mathcal{X}$ is a classical solution to the ergodic Nash system (2.4.1) on $\mathcal{X}$ if the following hold:

- **(E1)** each v^i is of class C^2 with respect to $x \in \mathcal{X}$, in the Fréchet sense;
- **(E2)** for each $i \in \mathbb{N}$, the series $\sum_{j\neq i} D_j v^j D_j v^i$ uniformly converges on all bounded subsets of $\mathcal{X}$;
- **(E3)** system (2.4.1) is satisfied pointwise for all $x \in \mathcal{X}$;

A *QSD ergodic (QSDE) solution* will be a classical solution to (2.4.1) with

$$\lambda^i \equiv \lambda = \operatorname{Tr} \bar{c}, \qquad v^i(x) = \frac{1}{2} \sum_{h,k\in\mathbb{N}} \bar{c}_{hk} \langle x^{h+i}, x^{k+i} \rangle_{\mathbb{R}^d}, \tag{2.4.2}$$

for some $\bar{c} \in \ell^1(\mathbb{N}^2)$.

REMARK 2.4.2. By the structure of (2.4.1), it is clear that if $((\lambda^i, v^i))_{i\in\mathbb{Z}}$ is a classical solution, then so will be $((\lambda^i, v^i + \mu^i))_{i\in\mathbb{Z}}$ for any choice of real numbers μ^i. We will prove that there exists a special choice of $\mu \in \mathbb{R}$, and of $\bar{c} \in \ell^1(\mathbb{N}^2)$, such that the solution $((\lambda, v^i + \mu))_{i\in\mathbb{Z}}$, with λ and v^i given by (2.4.2), is in a precise sense the limit of the QSD solution as $T \to +\infty$ (see Theorem 2.5.1).

Arguing as in the previous section, the coefficients $\bar{c}_{hk}$ of a QSDE solution are given by the solutions of the following system:

$$- c_{0h} c_{0k} + \sum_{j=0}^{h} c_{0,h-j} c_{jk} + \sum_{j=0}^{k} c_{0,k-j} c_{jh} = f_{hk}. \tag{2.4.3}$$

It is immediate to see that if c solves (2.4.3), then so does $-c$, hence we cannot have a unique solution to this limit system.

LEMMA 2.4.3. *There are exactly two sequences $(c^{\pm}_{hk})_{h,k\in\mathbb{N}}$ which solve (2.4.3). Such sequences are one the opposite of the other; that is, $c^- = -c^+$.*

PROOF. We have $c_{00}^2 = f_{00} > 0$, hence $c_{00} \in \{\pm\sqrt{f_{00}}\}$. Once the sign of c_{00} is chosen, all other c_{hk} can be uniquely determined by induction on h, k. ∎

Both solutions can be represented also in this case by a generating function.

LEMMA 2.4.4. *The two solutions to (2.4.3) are given by*

$$c^{\pm}_{hk} = \pm \frac{1}{h!\,k!} \left. \frac{\partial^{h+k}}{\partial z^h \partial w^k} \right|_{(0,0)} \Xi, \tag{2.4.4}$$

for some analytic function $\overline{\Xi}\colon \varrho\mathbb{D}^2 \to \mathbb{C}$ such that $\overline{\Xi}(z,w) = \overline{\Xi}(w,z)$.

Proof. A function $\overline{\Xi}$ is the desired generating function if it solves the equation

$$-\overline{\Xi}(z,0)\,\overline{\Xi}(0,w)+\big(\overline{\Xi}(z,0)+\overline{\Xi}(0,w)\big)\,\overline{\Xi}(z,w)=\phi(z,w),$$

for $z,w\in\varrho\mathbb{D}$. It follows that $\overline{\Xi}(z,0)^2=\phi(z,0)$, so choose $\overline{\Xi}(\cdot,0)=\xi$; then

$$\overline{\Xi}(z,w)=\frac{\phi(z,w)+\xi(z)\xi(w)}{\xi(z)+\xi(w)}. \tag{2.4.5}$$

Note that the real part of the denominator in (2.4.5) can vanish only if $|z|,|w|=\varrho$, hence $\overline{\Xi}$ is well-defined and analytic for $(z,w)\in\varrho\mathbb{D}^2$. ∎

This is sufficient to prove the existence of exactly two opposite QSDE solutions, and thus that the limiting ergodic Nash system in $\mathcal{X}$ is soluble. We state this as a theorem.

Theorem 2.4.5. *There exist exactly two QSDE solutions to (2.4.1) on $\mathcal{X}$. Such solutions are one the opposite of the other.*

Proof. Argue as in the first part of the proof of Lemma 2.2.28 to say that $c^+\in\ell^1(\mathbb{N}^2)$, then argue as in the proof of Lemma 2.2.29 to build the QSDE solution determined by choosing $\bar c=c^+$. Finally note that the solution determined by the choice $\bar c=c^-$ is the opposite function. ∎

2.5. Long-time asymptotics

As expected by KAM theory and ergodic control, we are going to prove now that, up to a constant, the QSDE solution corresponding to c^+ (that is, the solution with $c_{00}>0$) describes the long-time asymptotics of the QSD solution as $T\to+\infty$, while the opposite solution should be regarded as the result of considering the limit $T\to-\infty$ instead. To highlight the dependence of the QSD solution on T, we will write it as u^i_T; on the other hand, since by the shift-invariance property it will suffice to show the convergence of u^0_T to the QSDE solution v^0, we will omit the superscript 0.

Theorem 2.5.1. *Let Assumptions ($\star$) be in force with $[0,+\infty)\subseteq I$. Let u_T be the value function of the 0-th player corresponding to the QSD solution to the Nash system on $[0,T]\times\mathcal{X}$; let v be the value function of the 0-th player corresponding to the QSDE solution on $\mathcal{X}$ determined by $\bar c=c^+$. Let $\lambda:=\operatorname{Tr}\bar c$. Then, for any $t\geqslant 0$, as $t<T\to+\infty$, the following limits hold, locally uniformly in both $x\in\mathcal{X}$ and t:*

$$\frac{u_T(t,x)}{T-t}\to\lambda \tag{2.5.1}$$

and there exists a constant $\mu\in\mathbb{R}$ such that

$$u_T(t,x)-\lambda(T-t)\to v(x)+\mu. \tag{2.5.2}$$

The proof is based on the following result, which is strictly related to a refinement of Lemma 2.2.28 (cf. Remark 2.5.3 below) and is due to the possibility of explicitly compute the integral in formula (2.2.24) as showed in Remark 2.2.27.

LEMMA 2.5.2. *Under Assumptions* ($\star$) *with* $[0,+\infty) \subseteq I$, *there exists a nonnegative function* $\gamma \in C_0^0([0,+\infty)) \cap L^1([0,+\infty))$, *depending only on* r, f *and* g, *such that*

$$\sup_{(z,w)\in\mathcal{Q}_r}\left|\sigma^{\pm}(t,z,w) - \frac{2}{\xi(z)+\xi(w)}\right| \leqslant \gamma(t)$$

for all $t \geqslant 0$, *where* $\sigma^{\pm}$ *are defined as in* (2.2.25).

PROOF. By the continuity of ξ, given $r' \in (r,\varrho)$,[9] there exists $\varepsilon > 0$ such that

$$\Re\xi \geqslant \varepsilon \quad \text{on } \overline{\mathbb{D}}_{r'} \tag{2.5.3}$$

whence

$$\inf_{k\in\mathbb{Z}} |\xi(w)t - \big(k+\tfrac{1}{2}\big)i\pi| \geqslant \mathfrak{d}(t) \qquad \forall\, w \in \overline{\mathbb{D}}_{r'},\ t\in[0,+\infty), \tag{2.5.4}$$

for some function $\mathfrak{d}$ which is uniformly positive on $[0,+\infty)$.[10] By (2.5.4) there exists a uniformly positive function $\mathfrak{f}\colon [0,+\infty) \to \mathbb{R}_+$, which depends only on r and f, such that $|\cosh(\cdot\, t)| \geqslant \mathfrak{f}(t)$ on $\xi(\overline{\mathbb{D}}_r)$; also, by (2.5.3) we can suppose that $\mathfrak{f}$ be asymptotic to $\frac{1}{2}e^{3\varepsilon|\cdot|}$ at ∞. Since, with $\mathcal{F}$ defined as in (2.2.25),

$$\frac{\partial}{\partial\xi}\mathcal{F}(t,z;\xi(z),\psi(z)) = \frac{1}{\cosh(\xi(z)t)}\,\frac{(\xi(z)^2-\psi(z)^2)t - \frac{\psi(z)}{\cosh(\xi(z)t)}}{(\psi(z)\tanh(\xi(z)t)+\xi(z))^2},$$

we obtain, also using (2.2.17),

$$\left|\frac{\mathcal{F}(t,z)-\mathcal{F}(t,w)}{\xi(z)-\xi(w)}\right| \lesssim \frac{t}{\mathfrak{f}(t)^2} \quad \forall\,(t,z,w)\in[0,+\infty)\times\mathcal{Q}_r, \tag{2.5.5}$$

where the implied constant depends only on r, f and g. At this point, it is easy to see that the desired conclusion follows. ∎

REMARK 2.5.3. This proof shows that if $[0,+\infty) \subseteq I$, then the constant K appearing in Lemma 2.2.28 is in fact independent of T, since $\sup_{\mathbb{R}\times\mathcal{Q}_r}|\hat{\Xi}|$ is finite. In particular, $c \in C^0([0,T];\ell^1(\mathbb{N}^2))$ along with its derivatives, and their the norms are bounded uniformly with respect to $T>0$.

PROOF OF THEOREM 2.5.1. Fix $r\in(1,\varrho)$. By comparing expressions (2.2.26) and (2.4.5) one sees that

$$\sup_{|z|,|w|\leqslant r} |\Xi(t,z,w) - \overline{\Xi}(z,w)| \lesssim \gamma(T-t), \tag{2.5.6}$$

[9] E.g., $r' = \frac{1}{2}(r+\varrho)$.

[10] For instance, $\mathfrak{d}(t) = \big(t \vee \frac{1}{3}\|\xi\|^{-2}_{\infty;\overline{\mathbb{D}}_{r'}}\big)\varepsilon$.

where γ is the function given in Lemma 2.5.2, and the implied constant depends only on the L^∞-norms of ϕ and ξ on $\mathbb{D}_r$ and $\mathcal{Q}_r$, respectively. By Lemma 2.2.26 and Cauchy's theorem on derivatives,

$$|c^T_{hk}(t) - \bar{c}_{hk}| \leqslant \frac{1}{r^{h+k}} \sup_{|z|,|w|\leqslant r} |\Xi(t,z,w) - \overline{\Xi}(z,w)| \qquad \forall\, h,k \in \mathbb{N}; \tag{2.5.7}$$

where we have used the superscript T to stress the fact that $c_{hk}(t) = c^T_{hk}(t)$ depends on the horizon T. Plugging (2.5.6) into (2.5.7) yields

$$\sum_{h,k\in\mathbb{N}} |c^T_{hk}(t) - \bar{c}_{hk}| \lesssim \gamma(T-t) \tag{2.5.8}$$

as well as

$$|\operatorname{Tr} c^T(t) - \lambda\,| \lesssim \gamma(T-t), \tag{2.5.9}$$

where the implied constants depend only on r and f. As γ is integrable on $[0,+\infty)$, by (2.5.9) so is $\operatorname{Tr}\hat{c} - \lambda$, where we use the notation $\hat{c} = c(T-\cdot)$ introduced in the proof of Lemma 2.2.25; thus by the dominated convergence theorem there exists $\mu \in \mathbb{R}$ such that

$$\int_t^T \operatorname{Tr} c^T - (T-t)\lambda \to \mu \quad \text{as } T \to +\infty,$$

locally uniformly in t. Along with (2.5.8), this proves (2.5.1) and (2.5.2). ∎

Remark 2.5.4. The argument of the previous proof also applies to the case when $t = sT$, with $s \in [0,1]$. In this case, we can give the following estimate of the rate of convergence of (2.5.1): for any $r \in (1, \varrho)$,

$$\sup_{\|x\|_{\mathcal{X}} \leqslant L,\ s\in[0,1]} \left| \frac{u_T(sT,x)}{T} - (1-s)\lambda \right| \lesssim_{L,r} \frac{1}{T}.$$

The implied constant can be computed quite explicitly, by retracing the proofs above; we confine ourselves to pointing out that it depends only on L, r and f, and that it explodes as $L \to \infty$ or $r \to 1$.

2.6. Digression on a delicate limit case

We have noted in Example 2.1.2 that the case of a cost designed according to an underlying directed circulant graph structure is limit among those satisfying our assumptions, in the sense that Definition 2.2.16 holds with $\varrho = 1$.

Results like Lemmata 2.2.25, 2.2.26, 2.2.28, 2.4.3 and 2.4.4 continue to hold, but the other methods used in the previous sections are not refined enough to successfully prove all the previous theorems for those limit cases, even though, along with Remark 2.5.3, they are sufficient in order to establish ℓ^∞-stability at the level of the system for c; that is, convergence in $\ell^\infty(\mathbb{N}^2)$ of the solution to (2.2.19) to a solution of (2.4.3).

On the other hand, having a closer look at the simplest limit case, which is the directed chain given by the choice $g = 0$, $f^0_{00} = f^0_{11} = 1 = -f^0_{01} = -f^0_{10}$ and $f^0_{hk} = 0$

for all other h, k, we note that we are also able to compute QSDE solutions quite easily, thanks to formula (2.4.5). In this case we have $\phi(z, 0) = 1 - z$, and we find the expansion

$$\overline{\Xi}(z,w) = 1 + \sum_{j\geqslant 1}(-)^j\binom{\frac{1}{2}}{j}\left(z^j + w^j\right) + \sum_{j\geqslant 2}(-)^j\binom{\frac{3}{2}}{j}\sum_{\substack{h,k\geqslant 1\\ h+k=j}} z^h w^k,$$

yielding

$$c^{\pm}_{hk} = \pm(-)^{h+k}\binom{\frac{3}{2}-\delta_{0,hk}}{h+k}. \tag{2.6.1}$$

As by Stirling's formula $|c^{\pm}_{hk}| \asymp (h+k)^{\delta_{0,hk}-\frac{5}{2}}$, one easily sees that these coefficients well-define a QSDE solution, hence the limit ergodic Nash system has a solution.

One can also note that the coefficients enjoy the property that

$$\bar{c}_{hk} = \bar{c}_{0,h+k} - \bar{c}_{0,h+k-1} \quad \text{if } hk \neq 0\,, \tag{2.6.2}$$

where $\bar{c}$ is either c^+ or c^-; this can be seen from (2.6.1) or proved by induction using system (2.4.3). In fact, the same can be done for system (2.2.19), so that property (2.6.2) also holds for the[11] solution of (2.2.19) on $[0, T]$, for any fixed $T > 0$. Therefore, the information about the coefficients of a prospective QSD or QSDE solution is encoded in the functions of one complex variable $\Xi_0(t, \cdot) := \Xi(t, \cdot, 0)$ and $\overline{\Xi}_0 = \overline{\Xi}(\cdot, 0)$, namely

$$\Xi_0(t,z) = \sqrt{1-z}\,\tanh\left(\sqrt{1-z}\,(T-t)\right) \quad \text{and} \quad \overline{\Xi}_0(z) = \sqrt{1-z}\,.$$

This peculiar fact is specific of the directed chain. It could be useful as it allows to conclude existence of a solution to the infinite-dimensional evolutive Nash system provided that the sequence of functions $(c_{0k})_{k\in\mathbb{N}}$ is monotone,[12] yet this would not still be enough to deal with convergence to an ergodic solution.

Another property is instead shared by all problems having f of the form (2.1.4): the polynomial ϕ factors as $\phi(z, w) = \xi^2(z)\xi^2(w)$; this makes Ξ and $\overline{\Xi}$ functions of $(\xi(z), \xi(w))$, possibly helping in the analysis of the aforementioned limit cases.

[11] Note that it is indeed unique, as Lemma 2.2.25 is still true with the same proof.

[12] As it would seem by computing the first functions of the sequence. . .

Part 2

Nash systems. Stable a priori estimates and their consequences

CHAPTER 3

Nash systems in infinitely many dimensions

Contents

3.1. The problem

We consider the following system of infinitely many (backward) parabolic differential equations of Hamilton–Jacobi–Bellman (HJB) type, which we refer to as an *infinite-dimensional Nash system*:

$$
\begin{cases}
-\partial_t u^i - \displaystyle\sum_{jk} A^{jk}(t,x) D^2_{jk} u^i \\
\qquad\qquad + H^i(t,x,\mathcal{D}u) + \displaystyle\sum_{j\neq i} \partial_{p^j} H^j(t,x,\mathcal{D}u) D_j u^i = 0 \qquad i \in \mathbb{N}\,, \\
u^i(T,\cdot) = G^i
\end{cases}
\tag{3.1.1}
$$

where $\mathcal{D}u := (D_j u^j)_{j\in\mathbb{N}}$, set in $\mathbb{R}^\omega_T := [0,T]\times\mathbb{R}^\omega$, where the space $\mathbb{R}^\omega$ is $\mathbb{R}^\mathbb{N}$ equipped with the product topology.

The unknowns are the *value functions* $u^i\colon \mathbb{R}^\omega_T \to \mathbb{R}$; the data are the *horizon* T, the *diffusion* $A = (A^{jk})_{j,k\in\mathbb{N}}\colon [0,T] \to \mathbb{R}^{\omega\times\omega}$, the *Hamiltonians* $H^i\colon \mathbb{R}^\omega_T \times \mathbb{R}^\omega \to \mathbb{R}$ and the *terminal costs* $G^i\colon \mathbb{R}^\omega \to \mathbb{R}$.

The notation $\mathbb{R}^{\omega\times\omega}$ indicates infinite-dimensional square matrices, or equivalently the space $(\mathbb{R}^\omega)^\omega$, in analogy with the notation $\mathbb{R}^{N\times N}$ for $N\times N$ matrices. The typical element of $[0,T]\times\mathbb{R}^\omega\times\mathbb{R}^\omega$ is denoted by (t,x,p), with coordinates x^j and p^j, $j\in\mathbb{N}$. Also, we will use the notation $\mathbb{R}^\omega_x$ and $\mathbb{R}^\omega_p$ when we need to specify which copy of $\mathbb{R}^\omega$ we are considering, while $\mathbb{R}^\omega_T$ will always mean $[0,T]\times\mathbb{R}^\omega_x$.

Derivatives are understood in the sense of Gateaux; recall that given k vectors $v^1,\dots,v^k \in \mathbb{R}^\omega$, the k-th Gateaux derivative of V along them is recursively defined by

$$
D^k V(x; v^1,\dots,v^k) := \frac{\partial}{\partial s}\bigg|_0 D^{k-1} V(x + s v^k; v^1,\dots,v^{k-1})\,.
$$

We denote by ∂_t the derivative with respect to t, by D_j the derivative with respect x^j (that is, with respect to x along e_j) and by ∂_{p^j} the derivative with respect to p^j.

The strategy in Chapter 2 to prove existence for the LQ infinite-dimensional Nash system is based on the existence of a positive *c-self-controlled* sequence $\beta \in \ell^1(\mathbb{Z})$ providing bounds for the derivatives on the data; that is, for instance,

$$|D^2_{hk} G^i| \lesssim \beta^{i-h} \beta^{i-k} \,,$$

with β satisfying the peculiar property to control its discrete self-convolution:

$$(\beta \star \beta)^i := \sum_{j \in \mathbb{Z}} \beta^{i-j} \beta^j \lesssim \beta^i \qquad \forall\, i \in \mathbb{Z}\,.$$

Thinking of the simpler form of (3.1.1) where the equations are given by

$$-\partial_t u^i - \Delta u^i + \frac{1}{2} |D_i u^i|^2 + \sum_{j \neq i} D_j u^j \cdot D_j u^i = 0\,,$$

the above hypotheses allow to obtain nice a priori estimates by efficiently handling the strongly coupling terms $\sum_{j \neq i} D_j u^j \cdot D_j u^i$.

Our main result (Theorem 3.4.2) is a local existence and uniqueness theorem for (3.1.1) under suitable decay assumptions inspired by those described here above; we refer to the Introduction for a more extensive discussion. Contextually, we also prove (see Proposition 3.3.2) a priori estimates, stable with respect to the dimension, for general linear transport-diffusion equations whose drifts (and their derivatives) enjoy appropriate decay properties.

3.2. Weighted Hölder spaces of functions on $\mathbb{R}^\omega$

We start with some useful definitions and related comments. At first sight, the reader might find the spaces defined in this section a bit odd, nevertheless we point out that they will turn out to be appropriate for dealing with decays governed by a c-self controlled sequence.

DEFINITION 3.2.1. For $V \colon \mathbb{R}^\omega \to \mathbb{R}$ and $\gamma \in [0,1]$ we define

$$[V]_\gamma := \sup_{j \in \mathbb{N}} \sup_{x^{-j} \in \mathbb{R}^\omega} \sup_{y^j \neq z^j} \frac{\big| V(x)|_{x^j = y^j} - V(x)|_{x^j = z^j} \big|}{|y^j - z^j|^\gamma}\,.$$

REMARK 3.2.2. We included also $\gamma = 0$, for which $[V]_0 \leqslant 2\|V\|_\infty$, where $\|\cdot\|_\infty$ is the usual sup-norm.

DEFINITION 3.2.3. Given $\beta \in (\mathbb{R}_+)^{\mathbb{N}}$ and $\alpha \in \mathbb{N}^{\mathbb{N}}$ with $|\alpha| = \sum_{i \in \mathbb{N}} \alpha^i < \infty$, for $V \colon \mathbb{R}^\omega \to \mathbb{R}$ and $\gamma \in [0,1)$ we define

$$\|V\|_{\infty;\beta,\alpha} := \frac{\|V\|_\infty}{\beta^\alpha} \quad \text{and} \quad [V]_{\gamma;\beta,\alpha} := \frac{[V]_\gamma}{\beta^\alpha}\,,$$

where β^α is recursively defined as follows: if $|\alpha| = 0$, $\beta^\alpha := 1$, and for $|\alpha| \geqslant 1$,

$$\beta^\alpha := \Big(\prod_{i\in\mathbb{N}} (\beta^i)^{\alpha^i}\Big)^{\frac{1}{|\alpha|}} \wedge \min \beta^{\alpha'},$$

the minimum being taken over all $\alpha' \leqslant \alpha$ (according to the lexicographic order) with $|\alpha'| = |\alpha| - 1$.

EXAMPLE 3.2.4. We will essentially use the above definition of β^α for $|\alpha| \leqslant 2$. For example, note that if $\alpha^j = 1 = \alpha^k$ and $\alpha^i = 0$ for all $i \notin \{j,k\}$, then $\beta^\alpha = \beta^j \wedge \beta^k \wedge \sqrt{\beta^j\beta^k}$.

REMARK 3.2.5. One has that $\|V\|_{\infty;\beta,0} = \|V\|_\infty$ and $[V]_{\gamma;\beta,0} = [V]_\gamma$ are in fact independent of β. Similarly, with 1 denoting the all-ones sequence, $\|V\|_{\infty;1,\alpha} = \|V\|_\infty$ and $[V]_{\gamma;1,\alpha} = [V]_\gamma$.

REMARK 3.2.6. Since β^α is non-increasing with respect to α, one has $\|V\|_{\infty;\beta,\alpha} \leqslant \|V\|_{\infty;\beta,\alpha'}$ if $\alpha \leqslant \alpha'$.

REMARK 3.2.7. Let $\alpha \in \mathbb{N}^\mathbb{N}$ and suppose that $\|D^{\alpha'}V\|_{\infty;\beta,\alpha'}$ is finite for all $j \in \mathbb{N}$ and $\alpha' \geqslant \alpha$ with $|\alpha'| = |\alpha|+1$, where we are using the standard multi-index notation for derivatives (that is, $D^\alpha V = D_0^{\alpha^0} D_1^{\alpha^1} \cdots D_\ell^{\alpha^\ell} V$, with $\ell := \max\{i \in \mathbb{N} :\ \alpha^i \neq 0\}$). Then

$$\begin{aligned}[D^\alpha V]_{\gamma;\beta,\alpha} &\leqslant \frac{\sup_j \|D_j D^\alpha V\|_\infty \vee 2\|D^\alpha V\|_\infty}{\beta^\alpha} \\ &\leqslant \sup_{\substack{\alpha'\geqslant\alpha \\ |\alpha'|=|\alpha|+1}} \|D^{\alpha'}V\|_{\infty;\beta,\alpha} \ \vee\ 2\|D^\alpha V\|_{\infty;\beta,\alpha}\,.\end{aligned}$$

DEFINITION 3.2.8. Given $m \in \mathbb{N}$, $\gamma \in [0,1)$ and $\beta \in (\mathbb{R}_+)^\mathbb{N}$, the space $C^{m+\gamma}_\beta(\mathbb{R}^\omega)$ is that of all functions $V\colon \mathbb{R}^\omega \to \mathbb{R}$ such that, for all $\alpha \in \mathbb{N}^\mathbb{N}$ with $|\alpha| \leqslant m$, the derivative $D^\alpha V$ exists and it is continuous, with finite norm

$$\|V\|_{m+\gamma;\beta} := \sum_{k\leqslant m} \sup_{|\alpha|=k} \|D^\alpha V\|_{\infty;\beta,\alpha} + \sup_{|\alpha|=m} [D^\alpha V]_{\gamma;\beta,\alpha}.$$

For $\gamma = 1$, we denote by $C^{m,1}_\beta(\mathbb{R}^\omega)$ the space of all functions as above, with finite norm

$$\|V\|_{m,1;\beta} := \sum_{k\leqslant m} \sup_{|\alpha|=k} \|D^\alpha V\|_{\infty;\beta,\alpha} + \sup_{|\alpha|=m} [D^\alpha V]_{1;\beta,\alpha}.$$

We will also benefit of a slight variant of the above spaces, which is given as follows.

DEFINITION 3.2.9. Given $m \in \mathbb{N} \setminus \{0\}$, $\gamma \in [0,1)$ and $\beta \in (\mathbb{R}_+)^\mathbb{N}$, the space $C^{m+\gamma-}_\beta(\mathbb{R}^\omega)$ is that of all functions $V \in C^{m-1+\gamma}_\beta(\mathbb{R}^\omega)$ such that, for all $\alpha \in \mathbb{N}^\mathbb{N}$ with $|\alpha| = m$, the derivative $D^\alpha V$ exists and it is continuous, with finite norm

$$\|V\|_{m+\gamma-;\beta} := \|V\|_{m-1+\gamma;\beta} + \sup_{|\alpha|=m} \sup_{\substack{\alpha'\leqslant\alpha \\ |\alpha'|=m-1}} \big(\|D^\alpha V\|_{\infty;\beta,\alpha'} + [D^\alpha V]_{\gamma;\beta,\alpha'}\big).$$

One also defines $C^{m-,1}_\beta(\mathbb{R}^\omega)$ in an analogous manner.

REMARK 3.2.10. By Remark 3.2.6 (and the fundamental theorem of calculus) one has that $C^{m+\gamma}_\beta(\mathbb{R}^\omega)$ is a closed subspace of $C^{m+\gamma-}_\beta(\mathbb{R}^\omega)$, and by Remark 3.2.7

$$\sum_{k\leqslant m-1}\sup_{|\alpha|=k}\|D^\alpha V\|_{\infty;\beta,\alpha}+\sup_{|\alpha|=m}\sup_{\substack{\alpha'\leqslant\alpha\\|\alpha'|=m-1}}\big(\|D^\alpha V\|_{\infty;\beta,\alpha'}+[D^\alpha V]_{\gamma;\beta,\alpha'}\big)$$

is an equivalent norm on $C^{m+\gamma-}_\beta(\mathbb{R}^\omega)$.

In the following, in light of Remark 3.2.5, for $\gamma\in[0,1)$ we will simply write $C^\gamma(\mathbb{R}^\omega)$ in lieu of $C^\gamma_\beta(\mathbb{R}^\omega)$ and omit the subscripts β and α in the seminorms; similarly $C^{0,1}(\mathbb{R}^\omega):=C^{0,1}_\beta(\mathbb{R}^\omega)$. We will also omit the β when it is the all-ones sequence; that is, $C^{m+\gamma}(\mathbb{R}^\omega):=C^{m+\gamma}_1(\mathbb{R}^\omega)$. The same conventions are understood for the spaces given in Definition 3.2.9 as well.

LEMMA 3.2.11. *The spaces* $C^{m+\gamma-}_\beta(\mathbb{R}^\omega)$ *and* $C^{m+\gamma}_\beta(\mathbb{R}^\omega)$ *are Banach.*

PROOF. We prove that $C^\gamma(\mathbb{R}^\omega)$ is Banach, and then that so is $C^{m+\gamma-}_\beta(\mathbb{R}^\omega)$ for $m\geqslant 1$. This will give also the completeness of $C^{m+\gamma}_\beta(\mathbb{R}^\omega)$ due to Remark 3.2.10.

If $(V^N)_{N\in\mathbb{N}}\subset C^\gamma(\mathbb{R}^\omega)$ is Cauchy, by the Ascoli–Arzelà theorem there exists a subsequence $V^{N_n}\to V\in C^0(\mathbb{R}^\omega)$ as $n\to\infty$ in the topology of compact convergence. We have, for any $x,y\in\mathbb{R}^\omega$ with $x^i=y^i$ for all $i\neq j$,

$$\begin{aligned}\big|V(x)-V^N(x)-\big(V(y)-V^N(y)\big)\big|&\\&\leqslant\limsup_{n\to\infty}\big|V^{N_n}(x)-V^{N_n}(y)-\big(V^N(x)-V^N(y)\big)\big|\\&\leqslant|x^j-y^j|^\gamma\limsup_{n\to\infty}[V^{N_n}-V^N]_\gamma\,,\end{aligned}$$

and $\|V-V^N\|_\infty\leqslant\limsup_{n\to\infty}\|V^{N_n}-V^N\|_\infty$. Therefore $V^N\to V$ in $C^\gamma(\mathbb{R}^\omega)$.

Fix now a bijection $\hat\alpha\colon\mathbb{N}\to\{\alpha\in\mathbb{N}^\mathbb{N}:\ |\alpha|\leqslant m\}$ and let $\mathcal{Y}$ be the spaces of sequences of functions $W\in\ell^\infty(\mathbb{N};C^\gamma(\mathbb{R}^\omega))$ with finite norm

$$\|W\|_{\mathcal{Y}}:=\sum_{k\leqslant m}\sup_{i\in N\colon|\hat\alpha(i)|=k}\frac{\|W^i\|_\infty+[W^i]_\gamma}{w(\hat\alpha(i))}\,,$$

where

$$w(\hat\alpha(i)):=\begin{cases}\beta^{\hat\alpha(i)} & \text{if } |\hat\alpha(i)|<m\\ \min\limits_{\substack{\alpha\leqslant\hat\alpha(i)\\|\alpha'|=m-1}}\beta^\alpha & \text{if } |\hat\alpha(i)|=m\,;\end{cases}$$

$\mathcal{Y}$ is easily seen to be a Banach space. Then consider the linear map $\imath\colon C^{m+\gamma}_\beta(\mathbb{R}^\omega)\to\mathcal{Y}$ given by $\imath(V)^i:=D^{\hat\alpha(i)}V$, for $V\in C^{m+\gamma}_\beta(\mathbb{R}^\omega)$ and $i\in\mathbb{N}$ By Remark 3.2.7, $\|V\|_{m+\gamma-;\beta,\alpha}\leqslant\|\imath(V)\|_{\mathcal{Y}}\leqslant 3\|V\|_{m+\gamma-;\beta}$. Therefore, if $(V^N)_{N\in\mathbb{N}}\subset C^{m+\gamma-}_\beta(\mathbb{R}^\omega)$ is Cauchy, then $\imath(V^N)\to W$ in $\mathcal{Y}$. By compact convergence, if $|\alpha(j)|=|\alpha(i)|+1$ with $\alpha(j)^k=\alpha(i)^k+1$, then $W^i(x)-W^i(x)|_{x^k=0}=\int_0^{x^k}W^j(x)|_{x^k=y}\,\mathrm{d}y$. Exploiting the fundamental theorem of calculus, one can prove by induction that this implies that $W^i = D^{\alpha(i)} W^{\alpha^{-1}(0)}$; therefore, $W = \imath(W^{\alpha^{-1}(0)})$. By the above equivalence of norms and the linearity of $\imath$, we conclude that $V^N \to W^{\alpha^{-1}(0)}$ in $C^{m+\gamma-}_\beta(\mathbb{R}^\omega)$. ■

In a similar fashion, we define parabolic (Hölder) spaces on $\mathbb{R}^\omega_T$.

Definition 3.2.12. For $V\colon \mathbb{R}^\omega_T \to \mathbb{R}$ and $\gamma \in [0,1)$ we define

$$[V]_{\frac{\gamma}{2},\gamma} := \sup_{x\in\mathbb{R}^\omega} [V(\,\cdot\,,x)]_{\frac{\gamma}{2}} + \sup_{t\in[0,T]} [V(t,\,\cdot\,)]_\gamma$$

$$= \sup_{s\neq t} \frac{\|V(t,\,\cdot\,) - V(s,\,\cdot\,)\|_\infty}{|t-s|^{\frac{\gamma}{2}}} + \sup_{t\in[0,T]} [V(t,\,\cdot\,)]_\gamma \,.$$

Then, given $\beta \in (\mathbb{R}_+)^\mathbb{N}$ and $\alpha \in \mathbb{N}^\mathbb{N}$ with $|\alpha| < \infty$, for $V\colon \mathbb{R}^\omega_T \to \mathbb{R}$ we define

$$[V]_{\frac{\gamma}{2},\gamma;\beta,\alpha} := \sup_{x\in\mathbb{R}^\omega} \frac{[V(\,\cdot\,,x)]_{\frac{\gamma}{2}}}{(\beta^\alpha)^{\frac{1}{2}}} + \sup_{t\in[0,T]} \frac{[V(t,\,\cdot\,)]_\gamma}{\beta^\alpha}\,,$$

where β^α is defined as in Definition 3.2.3.

Definition 3.2.13. Given $\gamma \in [0,1)$, the space $C^{\frac{\gamma}{2},\gamma}(\mathbb{R}^\omega_T)$ is that of all continuous functions $V\colon \mathbb{R}^\omega_T \to \mathbb{R}$ with finite norm

$$\|V\|_{\frac{\gamma}{2},\gamma} := \|V\|_\infty + [V]_{\frac{\gamma}{2},\gamma}\,.$$

Given also $\beta \in (\mathbb{R}_+)^\mathbb{N}$, the space $C^{1,2}_\beta(\mathbb{R}^\omega_T)$ is that of all functions $V\colon \mathbb{R}^\omega_T \to \mathbb{R}$ such that the derivatives $\partial_t V$ and $D^\alpha V$, for all $\alpha \in \mathbb{N}^\mathbb{N}$ with $|\alpha| \leqslant 2$ exists and are continuous on $(0,T)\times\mathbb{R}^\omega$, with finite norm

$$\|V\|_{1,2;\beta} := \|\partial_t V\|_\infty + \sum_{k\leqslant 2} \sup_{|\alpha|=k} \|D^\alpha V\|_{\infty;\beta,\alpha}\,.$$

With a proof very much similar to that of Lemma 3.2.11 (which we thus omit), we can also show the following fact.

Lemma 3.2.14. *The spaces $C^{\frac{\gamma}{2},\gamma}(\mathbb{R}^\omega_T)$ and $C^{1,2}_\beta(\mathbb{R}^\omega_T)$ are Banach.*

We can now set the notion of classical solution to the infinite-dimensional Nash system. The space $C^0(\mathbb{R}^\omega_T)$ appearing below is defined in the usual way as the space of all continuous functions from $\mathbb{R}^\omega_T$ to $\mathbb{R}$ with bounded norm $\|\cdot\|_\infty$; alternatively, it can be understood as the space $C^{0,0}(\mathbb{R}^\omega_T)$ defined above, as the two respective norms are equivalent.

Definition 3.2.15. We will say that $u = (u^i)_{i\in\mathbb{N}}$ is a classical solution to the Nash system (3.1.1) if the following happens:

- for each $i \in \mathbb{N}$, $u^i \in C^{1,2}_{\beta_i}(\mathbb{R}^\omega_T)$ for some $\beta_i \in (\mathbb{R}_+)^\mathbb{N}$, and $\sup_{i\in\mathbb{N}}\|u^i\|_{1,2;\beta_i} < \infty$;
- the series appearing in the equations converge in $C^0(\mathbb{R}^\omega_T)$ uniformly in i;
- the equations are satisfied pointwise in $\mathbb{R}^\omega_T$.

3.3. A priori estimates on linear parabolic equations

We begin with an important general estimate on the derivatives of a solution to a Fokker–Planck–Kolmogorov equation. We state this result under slightly more general hypotheses than the ones under which we will use it, for the convenience of possible further applications.

LEMMA 3.3.1. *Let $N \in \mathbb{N}$. Let $A\colon \mathbb{R}\times\mathbb{R}^N \to \mathscr{S}(N)$ and $B\colon \mathbb{R}\times\mathbb{R}^N \to \mathbb{R}^N$ be bounded and continuous, with D_iA^{hk}, $D^2_{ij}A^{hk}$ and D_iB^j bounded for all $i,j,h,k \in [\![N]\!]$. Suppose that $A \geqslant \lambda I$ for some $\lambda > 0$ and $D_kA^{ij} = 0$ if $i \neq k \neq j$. Let $\rho = \rho_\varepsilon \in C^{1,2}((s,T)\times\mathbb{R}^N) \cap C^0([s,T)\times\mathbb{R}^N)$ solve*

$$\begin{cases} \partial_t\rho - \sum_{jk} D^2_{ij}(A^{ij}\rho) - \operatorname{div}(B\rho) = 0 & \text{in } (s,T)\times\mathbb{R}^N \\ \rho(s) = \delta_y \star \eta_\varepsilon, \end{cases} \tag{3.3.1}$$

for $T > 0$ and some $(s,y) \in [0,T)\times\mathbb{R}^N$, with $(\eta_\varepsilon)_{\varepsilon>0}$ being an approximation of the identity as $\varepsilon \to 0$. Then there exists $T^* > 0$, depending only on

$$\lambda, \quad \sup_i \Big\| \sum_j |D_iB^j|^2 \Big\|_\infty,$$

$$\sup_i \Big\| \sum_j D_iA^{ij} \Big\|_\infty, \quad \sup_i \Big\| \sum_j D^2_{ij}A^{ij} \Big\|_\infty, \quad \text{and} \quad \sup_i \Big\| \sum_j \Big| \sum_k D^2_{ik}A^{jk} \Big|^2 \Big\|_\infty,$$

and $C > 0$, depending only on the above quantities and T^*, such that, if $T < T^*$, then

$$\lim_{\varepsilon\to 0} \sup_k \int_s^t \int_{\mathbb{R}^N} |D_k\rho_\varepsilon| \leqslant C\sqrt{t-s} \qquad \forall\, t \in [s,T]. \tag{3.3.2}$$

PROOF. By the results in [**45**, Chapter 1], we know that (3.3.2) holds for some constant C and any $T > 0$. We want to make explicit the dependence of C on the data, for T near 0. Let $\tau \in (s,T)$ and consider a sequence of smooth functions $\psi^{\tau,k}_{\varepsilon,n} \to \operatorname{sgn} D_k\rho_\varepsilon(\tau)$ in $L^1_{\mathrm{loc}}(\mathbb{R}^N)$ as $n \to \infty$, with $\|\psi^{\tau,k}_{\varepsilon,n}\|_\infty = 1$ for all n. Since $\int_{\mathbb{R}^N} D_k\rho_\varepsilon(\tau)\psi^{\tau,k}_{\varepsilon,n}(\tau) \to \int_{\mathbb{R}^N} |D_k\rho_\varepsilon(\tau)|$ as $n \to \infty$, for any sequence $\varepsilon_n \to 0$ we can consider $\ell_n \to \infty$ such that

$$\int_{\mathbb{R}^N} \big| D_k\rho_n(\tau)\psi^{\tau,k}_n(\tau) - |D_k\rho_n(\tau)| \big| \to 0 \quad \text{as } n \to \infty,$$

where $\rho_n := \rho_{\varepsilon_n}$ and $\psi^{\tau,k}_n := \psi_{\varepsilon_n,\ell_n}$. Let now $w = w^{\tau,k}_n$ solving

$$-\partial_t w - \operatorname{Tr}(AD^2w) + \langle B, Dw\rangle = 0 \quad \text{on } (s,\tau)\times\mathbb{R}^N, \tag{3.3.3}$$

with terminal condition $w|_{t=\tau} = \psi_n = \psi^{\tau,k}_n$. We first notice that testing the equation of ρ_ε by $\frac{1}{2}w^2$ over (s,τ) we get

$$(w(s,\cdot)^2 \star \eta_\varepsilon)(y) + 2\int_s^\tau \int_{\mathbb{R}^N} \langle ADw, Dw\rangle \rho_\varepsilon \leqslant 1. \tag{3.3.4}$$

On the other hand, $v(t,x) := \frac{\tau-t}{2}|D_k w(t,x)|^2$ satisfies

$$\begin{aligned}&-\partial_t v - \operatorname{Tr}(AD^2v) + \langle B, Dv\rangle \\ &\qquad\leqslant \frac{1}{2}|D_k w|^2 - (\tau-t)\langle B_k, Dw\rangle D_k w + (\tau-t)\operatorname{Tr}(D_kAD^2w)D_kw,\end{aligned}$$

so testing by ρ_ε over (s,τ) one obtains

$$\begin{aligned}\int_{\mathbb{R}^N} v(s,\cdot)\eta_\varepsilon(y-\cdot) &\leqslant \frac{1}{2}\int_s^\tau\int_{\mathbb{R}^N}|D_kw|^2\rho_\varepsilon - \int_s^\tau\int_{\mathbb{R}^N}(\tau-t)\langle D_kB, Dw\rangle D_kw\rho_\varepsilon \\ &\quad + \int_s^\tau\int_{\mathbb{R}^N}(\tau-t)\operatorname{Tr}(D_kAD^2w)D_kw\rho_\varepsilon\,. \end{aligned} \tag{3.3.5}$$

Using estimate (3.3.4) and the ellipticity assumption $A \geqslant \lambda I$, along with Young's and the Cauchy–Schwarz inequalities, we estimate

$$\frac{1}{2}\int_s^\tau\int_{\mathbb{R}^N}|D_kw|^2\rho_\varepsilon - \int_s^\tau\int_{\mathbb{R}^N}(\tau-t)\langle D_kB, Dw\rangle D_kw\rho_\varepsilon \leqslant \frac{1+\tau(1+C_B)}{4\lambda},$$

with $C_B := \|D_kB\|^2_{2,\infty}$. On the other hand, using our assumptions on A and integrating by parts,

$$\begin{aligned}\int_s^\tau\int_{\mathbb{R}^N}(\tau-t)\operatorname{Tr}(D_kAD^2w)D_kw\rho_\varepsilon &= -\frac{1}{2}\int_s^\tau\int_{\mathbb{R}^N}(\tau-t)\sum_j D^2_{jk}A^{jk}|D_kw|^2\rho_\varepsilon \\ &\quad -\frac{1}{2}\int_s^\tau\int_{\mathbb{R}^N}(\tau-t)\sum_j D_kA^{jk}|D_kw|^2D_j\rho_\varepsilon\,,\end{aligned}$$

whence

$$\begin{aligned}&\int_s^\tau\int_{\mathbb{R}^N}(\tau-t)\operatorname{Tr}(D_kAD^2w)D_kw\rho_\varepsilon \\ &\qquad\leqslant \frac{TC_A}{4\lambda} + C'_A\sup_{t\in[s,\tau]}\|v(t,\cdot)\|_\infty \sup_j\int_s^\tau\int_{\mathbb{R}^N}|D_j\rho_\varepsilon|\,,\end{aligned}$$

with $C_A := \sup_k\big\|\sum_j D^2_{jk}A^{jk}\big\|_\infty$ and $C'_A := \sup_k\big\|\sum_j D_kA^{jk}\big\|_\infty$. Let now $y=z$ with $z = z(s,N,\tau,n)$ such that $|D_kw(s,z)|^2 \geqslant \|D_kw(s,\cdot)\|^2_\infty - 1$ for all $k \in [\![N]\!]$. We will write $\rho_\varepsilon = \rho^z_\varepsilon$ to stress that $\rho^z_\varepsilon(s) = \delta_z \star \eta_\varepsilon$. Then for $\varepsilon_n > 0$ so small that $\int_{\mathbb{R}^N} v(s,\cdot)\eta_{\varepsilon_n}(z-\cdot) \geqslant v(s,z)-1$, from the previous estimates we have

$$\begin{aligned}\|v(s,\cdot)\|_\infty &\leqslant 1 + \frac{T}{2} + \frac{1+T(1+C_B+C_A)}{4\lambda} \\ &\quad + C'_A\sup_{t\in[s,\tau]}\|v(t,\cdot)\|_\infty\sup_j\int_s^\tau\int_{\mathbb{R}^N}|D_j\rho^z_n|\,.\end{aligned} \tag{3.3.6}$$

By continuity, for any τ close enough to s (say $s < \tau < \hat\tau = \hat\tau(s,n,N) \leqslant T$) one has

$$\sup_j\int_s^\tau\int_{\mathbb{R}^N}|D_j\rho^z_n| \leqslant \frac{1}{2C'_A},$$

so (3.3.6) yields

$$\|v(s,\cdot)\|_\infty \leqslant 2 + T + \frac{1 + T(1 + C_B + C_A)}{2\lambda}.$$

This implies that for all $\tau \leqslant \bar\tau := \min_{t\in[s,T]} \hat\tau(t,n,N)$ one has

$$\sup_k \|D_k w_n^{\tau,k}(t,\cdot)\|_\infty \leqslant \frac{\bar C}{\sqrt{\tau - t}} \quad \forall\, t \in [s,\tau], \tag{3.3.7}$$

with $\bar C = \bar C(\lambda, T, C_B, C_A)$ being the square root of the constant above. We want to show that we can choose $\bar\tau \equiv T$, provided that T is small enough. Testing the equation for of w by $D_k\rho_n$ over (s,τ) we get

$$-\int_{\mathbb{R}^N} D_k\rho_n(\tau)\psi_n(\tau) = \int_{\mathbb{R}^N} D_k w(s,\cdot)\rho_n(s) + \int_s^\tau \int_{\mathbb{R}^N} \langle D_k B, Dw\rangle \rho_n + \sum_{ij}\int_s^\tau \int_{\mathbb{R}^N} D^2_{ik}A^{ij} D_j w \rho_n + \sum_{ij}\int_s^\tau\int_{\mathbb{R}^N} D_k A^{ij} D_j w D_i\rho_n,$$

whence, by the Cauchy–Schwarz and Young's inequalities, and estimate (3.3.4),

$$\int_{\mathbb{R}^N} D_k\rho_n(\tau)\psi_n(\tau) \leqslant \|D_k w(s,\cdot)\|_\infty + T\Big(\frac12 C_B + C''_A\Big) + \frac{1}{2\lambda} + C'_A \sup_i \int_s^\tau \sup_j \|D_j w(t,\cdot)\|_\infty \int_{\mathbb{R}^N} |D_i\rho_n|(t)\,\mathrm{d}t\,, \tag{3.3.8}$$

where $C''_A := \sup_k \big\|\sum_i D^2_{ik}A^{i\cdot}\big\|^2_{2,\infty}$. Let $n_0 = n_0(\tau, N) \in \mathbb{N}$ be such that

$$\int_{\mathbb{R}^N} D_k\rho_n(\tau)\psi_n^{\tau,k}(\tau) \geqslant \int_{\mathbb{R}^N} |D_k\rho_n(\tau)| - 1$$

for all $n \geqslant n_0$ and for all $k \in [\![N]\!]$; for those n, by Gronwall's lemma,

$$\sup_k \int_{\mathbb{R}^N} |D_k\rho_n|(\tau) \leqslant \big(1 + \sup_k \|D_k w_n^{\tau,k}(s,\cdot)\|_\infty + \hat C\big) \exp\Big\{ C'_A \int_s^\tau \sup_j \|D_j w_n^{\tau,k}(t,\cdot)\|_\infty \,\mathrm{d}t \Big\},$$

where $\hat C := T\big(\frac12 C_B + C''_A\big) + \frac{1}{2\lambda}$. For $\tau \leqslant \bar\tau$ as above, we obtain

$$\sup_k \int_s^\tau \int_{\mathbb{R}^N} |D_k\rho_n| \leqslant \sqrt{\tau - s}\,\big(\sqrt{T}(1+\hat C) + \bar C\big) e^{C'_A \bar C T} \leqslant \sqrt{T}\big(\sqrt{T}(1+\hat C) + \bar C\big) e^{C'_A \bar C T}\,, \tag{3.3.9}$$

which holds for any $\rho_n = \rho_n^y$; in particular, if $\sqrt{T}\big(\sqrt{T}(1+\hat C)+\bar C\big)e^{C'_A\bar C T} \leqslant \frac{1}{2C'_A}$ and we let $y = z$ as above, by a continuity argument we see that we can choose $\hat\tau \equiv T$, and thus $\bar\tau = T$. We deduce that estimate (3.3.9) in fact holds for all $\tau \in [s,T]$ and we can let $n \to \infty$ to get (3.3.2). ∎

The following result provides a crucial decay estimates for the derivatives of a solution to a linear transport-diffusion equation whose differential operator is the adjoint of the one of the FPK equation.

PROPOSITION 3.3.2. *Let $N \in \mathbb{N}$. Consider A and B as in Lemma 3.3.1, and let $G \in C^3(\mathbb{R}^N)$ and $F \in C^0([0,T]; C^2(\mathbb{R}^N))$. Suppose that $D_k A^{ij} = 0$ unless $i = j = k$. Let $\beta \in \ell^1(\mathbb{Z}; \mathbb{R}_+)$ be even and such that $\beta \star \beta \leqslant c\beta$; that is,*

$$\sum_{j \in \mathbb{Z}} \beta^j \beta^{i-j} \leqslant c\beta^i \quad \forall\, i \in \mathbb{Z}.$$

Suppose there exist constants $c_B, c_F, c_G \geqslant 0$ such that, for all $i, j, k, l \in [\![N]\!]$,

$$\|D_j B^i\|_\infty \leqslant c_B \beta^{j-i}\,, \qquad \|D_{jk} B^i\|_\infty \leqslant c_B(\beta^{j-i} \wedge \sqrt{\beta^{j-i}\beta^{k-i}}\,)\,, \tag{3.3.10}$$

$$\|D_j F\|_\infty \leqslant c_F \beta^j, \qquad \|D^2_{jk} F\|_\infty \leqslant c_F(\beta^j \wedge \sqrt{\beta^j \beta^k}\,)\,, \tag{3.3.11}$$

$$\|D_j G\|_\infty \leqslant c_G \beta^j\,, \qquad \|D^2_{jk} G\|_\infty \vee \|D^3_{jkl} G\|_\infty \leqslant c_G(\beta^j \wedge \sqrt{\beta^j \beta^k}\,)\,. \tag{3.3.12}$$

Then a classical solution $w \in C^0([0,T]; C^3(\mathbb{R}^N))$ to

$$\begin{cases} -\partial_t w - \mathrm{Tr}(AD^2 w) + \langle B, Dw \rangle = F & \text{on } (0,T) \times \mathbb{R}^N \\ w|_{t=T} = G \end{cases} \tag{3.3.13}$$

satisfies

$$\|D_j w\|_\infty + \|D^2_{jk} w\|_\infty + \|D^3_{jkl} w\|_\infty \leqslant K(\beta^j \wedge \sqrt{\beta^j \beta^k}\,) \tag{3.3.14}$$

for all $j, k, l \in \{0, \dots, N-1\}$, where the constant K depends only on T, c, c_B, c_F, c_G, $\|\beta\|_1$ and $c_A := \max_{1 \leqslant \ell \leqslant 3} \sup_k \|(D_k)^\ell A^{kk}\|_\infty$. More precisely, one can choose $K = c_G + c_F \tilde{K}(T)$ with $\tilde{K}(T)$ vanishing as $T \to 0$.

Note that recalling Example 3.2.4 the above bounds (3.3.10), (3.3.11) (3.3.12) and (3.3.14) can be expressed in terms of the norms of suitable spaces among those introduced Section 3.2. This allows a more compact notation, which will be used in the following Section 3.4; nevertheless, here we agreed to write the estimates in a more explicit form, for the benefit of the reader who needs to get used to the kind of controls we wish to eventually have on the derivatives of the data and the solution of the Nash system.

PROOF OF PROPOSITION 3.3.2. The following computations are performed assuming w to be smooth; nevertheless, by means of a standard approximation argument, one can prove that the estimates we get hold for w as regular as in the statement of this proposition.

First-order estimates. Testing the equation of $D_k w$ by ρ_ε solving (3.3.1) and letting $\varepsilon \to 0$, after easy computations one gets

$$\begin{aligned} &\|D_k w(s,\cdot)\|_\infty \\ &\quad \leqslant \|D_k G\|_\infty + (T-s)\bigg(c_B \sup_i \frac{\|D_i w\|_{\infty;[s,T]}}{\beta^i} \sum_{j \in [\![N]\!]} \beta^{k-j}\beta^j + \|D_k F\|_{\infty;[s,T]}\bigg) \\ &\qquad\qquad\qquad + \int_s^T \bigg| \int_{\mathbb{R}^N} D_k A^{kk} D^2_{kk} w \rho \bigg|\,; \end{aligned}$$

here $\|\cdot\|_{\infty;[s,T]} := \sup_{[s,T]\times\mathbb{R}^N} |\cdot|$ and $\rho(t) := \lim_{\varepsilon\to 0}\rho_\varepsilon(t) \in C^0(\mathbb{R}^d)$, and we have used (3.3.10). The last integral is obtained exploiting our assumption on A and which is controlled by

$$(T-s)C_A \sup_{ij} \frac{\|D^2_{ij}w\|_{\infty;[s,T]}}{\beta^i \wedge \sqrt{\beta^i\beta^j}}\, \beta^k,$$

with $C_A := \sup_k \|D_k A^{kk}\|_\infty$. Then, if $T-s$ small enough, using that $\beta\star\beta \leqslant c\beta$ we deduce from the above estimates that

$$\sup_k \frac{\|D_k w\|_{\infty;[s,T]}}{\beta^k} \leqslant c_1(T-s)\bigg(1 + \sup_{ij} \frac{\|D^2_{ij}w\|_{\infty;[s,T]}}{\beta^i \wedge \sqrt{\beta^i\beta^j}}\bigg), \tag{3.3.15}$$

where $c_1(T-s)$ is an increasing function of $T-s$ that can be explicitly written using the parameters $c, C_A, c_B, c_F, c_G, \|\beta\|_1$.

Second order estimates. Testing the equation of $D^2_{jk}w$ by ρ one gets

$$\begin{aligned}\|D^2_{jk}w(s,\cdot)\|_\infty \leqslant \|D^2_{jk}G\|_\infty &+ (T-s)\|D^2_{jk}F\|_{\infty;[s,T]} \\ &+ \int_s^T\int_{\mathbb{R}^N} \big|D^2_{jk}\langle B, Dw\rangle - \langle B, D^2_{jk}Dw\rangle\big|\rho \\ &+ \int_s^T \bigg|\int_{\mathbb{R}^N} \big(D^2_{jk}A^{jj}D^2_{jj}w + D_jA^{jj}D^3_{jjk}w \\ &\qquad\qquad + D_kA^{kk}D^3_{jkk}w\big)\rho\bigg|,\end{aligned} \tag{3.3.16}$$

where in particular $D^2_{jk}A^{jj} = 0$ if $j \neq k$. It easy to see that the last integral is controlled by

$$3(T-s)C'_A\Big(\sup_{il} \frac{\|D^2_{il}w\|_{\infty;[s,T]}}{\beta^i \wedge \sqrt{\beta^i\beta^l}} + \sup_{ilm} \frac{\|D^3_{ilm}w\|_{\infty;[s,T]}}{\beta^i \wedge \sqrt{\beta^i\beta^l}}\Big)(\beta^j \wedge \sqrt{\beta^j\beta^k}\,),$$

with $C'_A := C_A + \sup_k \|D_{kk}A^{kk}\|_\infty$. On the other hand, using (3.3.10) and the property $\beta\star\beta \leqslant c\beta$, one can estimates in two ways, according to whether one uses that $|D^2_{jk}B^i| \leqslant c_B\beta^j$ or $|D^2_{jk}B^i| \leqslant c_B\sqrt{\beta^j\beta^k}$: we have either

$$\begin{aligned}&\int_{\mathbb{R}^N} \big|D^2_{jk}\langle B, Dw\rangle - \langle B, D^2_{jk}Dw\rangle\big|\rho \\ &\quad\leqslant c_B\bigg(\sup_i \frac{\|D_i w\|_{\infty;[s,T]}}{\beta^i} \cdot \sum_i \beta^{j-i}\beta^i \\ &\qquad\qquad + \sup_{ij} \frac{\|D^2_{ij}w\|_{\infty;[s,T]}}{\beta^i} \cdot \Big(\|\beta\|_1\beta^j + \sum_i \beta^{j-i}\beta^i\Big)\bigg) \\ &\quad\leqslant c_B\bigg(c\sup_i \frac{\|D_i w\|_{\infty;[s,T]}}{\beta^i} + (\|\beta\|_1 + c)\sup_{ij} \frac{\|D^2_{ij}w\|_{\infty;[s,T]}}{\beta^i}\bigg)\beta^j,\end{aligned}$$

or, exploiting the Cauchy-Schwarz inequality,

$$
\begin{aligned}
&\int_{\mathbb{R}^N} \big|D^2_{jk}\langle B, Dw\rangle - \langle B, D^2_{jk}Dw\rangle\big|\rho \\
&\qquad \leqslant c_B \sup_i \frac{\|D_i w\|_{\infty;[s,T]}}{\beta^i}\cdot \sum_i (\beta^{j-i})^{\frac12}(\beta^{k-i})^{\frac12}\beta^i \\
&\qquad\quad + c_B \sup_{ij}\frac{\|D^2_{ij}w\|_{\infty;[s,T]}}{\sqrt{\beta^i\beta^j}}\cdot\Big(\sum_i \beta^{k-i}(\beta^j)^{\frac12}(\beta^i)^{\frac12} + \sum_i \beta^{j-i}(\beta^k)^{\frac12}(\beta^i)^{\frac12}\Big) \\
&\qquad \leqslant \bigg(cc_B \sup_i \frac{\|D_i w\|_{\infty;[s,T]}}{\beta^i} + 2c_B\sqrt{c\|\beta\|_1}\sup_{ij}\frac{\|D^2_{ij}w\|_{\infty;[s,T]}}{\sqrt{\beta^i\beta^j}}\bigg)\sqrt{\beta^j\beta^k}\,.
\end{aligned}
$$

Therefore, also using (3.3.15), we deduce that, for small $T-s$,

$$
\sup_{jk}\frac{\|D^2_{jk}w\|_{\infty;[s,T]}}{\beta^j\wedge\sqrt{\beta^j\beta^k}} \leqslant c_2(T-s)\bigg(1+\sup_{ijk}\frac{\|D^2_{ijk}w\|_{\infty;[s,T]}}{\beta^i\wedge\sqrt{\beta^i\beta^j}}\bigg), \tag{3.3.17}
$$

where $c_2(T-s)$ is an increasing function of $T-s$ depending also on the parameter C'_A and those listed for c_1.

Third-order estimates. Testing the equation of $D^3_{jkl}w$ by ρ_ε one gets

$$
\begin{aligned}
&\|D^3_{jkl}w(s,\cdot)\|_\infty \\
&\qquad \leqslant \|D^3_{jkl}G\|_\infty + \lim_{\varepsilon\to 0}\bigg(\bigg|\int_s^T\int_{\mathbb{R}^N}\big(D^3_{jkl}\langle B, Dw\rangle - \langle B, D^3_{jkl}Dw\rangle\big)\rho_\varepsilon\bigg| \\
&\qquad\qquad + \bigg|\int_s^T\int_{\mathbb{R}^N} D^3_{jkl}F\rho_\varepsilon\bigg| + \mathcal{A}_\varepsilon\bigg),
\end{aligned}
$$

where $\mathcal{A}_\varepsilon$ collects all terms involving derivatives of A. By analogous estimates as above, it is not difficult to see that all terms in $\mathcal{A}$ where no fourth-order derivatives of w appear can be controlled by

$$
(T-s)C''_A\|\beta\|_\infty\Big(\sup_{il}\frac{\|D^2_{il}w\|_{\infty;[s,T]}}{\beta^i\wedge\sqrt{\beta^i\beta^l}} + \sup_{ilm}\frac{\|D^3_{ilm}w\|_{\infty;[s,T]}}{\beta^i\wedge\sqrt{\beta^i\beta^l}}\Big)(\beta^j\wedge\sqrt{\beta^j\beta^k}\,),
$$

with $C''_A := C'_A + \sup_k\|D_{kkk}A^{kk}\|_\infty$. This excludes three terms, which are of the form

$$
\begin{aligned}
&\int_s^T\bigg|\int_{\mathbb{R}^N} D_jA^{jj}D^4_{jjkl}w\rho_\varepsilon\bigg| \\
&\qquad\qquad \leqslant \int_s^T\bigg|\int_{\mathbb{R}^N} D^2_{jl}A^{jj}D^3_{jjk}w\rho_\varepsilon\bigg| + \int_s^T\bigg|\int_{\mathbb{R}^N} D_jA^{jj}D^3_{jjk}wD_l\rho_\varepsilon\bigg|
\end{aligned}
$$

for any permutation of j,k,l, and thus, using also Lemma 3.3.1, as $\varepsilon\to 0$ they are controlled by

$$
\sqrt{T-s}\,\big(C'_A + \sqrt{T-s}\,C''_AC\big)\sup_{ilm}\frac{\|D^3_{ilm}w\|_{\infty;[s,T]}}{\beta^i\wedge\sqrt{\beta^i\beta^l}}(\beta^j\wedge\sqrt{\beta^j\beta^k}\,).
$$

On the other hand, we have

$$\begin{aligned}(3.3.18)\qquad &\left|\int_{\mathbb{R}^N} \big(D^2_{jkl}\langle B, Dw\rangle - \langle B, D^3_{jkl}Dw\rangle\big)\rho_\varepsilon(t,\cdot)\right| \\ &\qquad \leqslant \sum \big(\|\langle D_j B, D^2_{kl}Dw\rangle\|_{\infty;[s,T]} + \|\langle D^2_{jk}B, D_l Dw\rangle\|_{\infty;[s,T]}\big) \\ &\qquad\quad - \|\langle D^2_{jk}B, D_l Dw\rangle\|_{\infty;[s,T]} + \|\langle D^2_{jk}B, Dw\rangle\|_{\infty;[s,T]} \int_{\mathbb{R}^N} |D_l\rho_\varepsilon(t,\cdot)|\,,\end{aligned}$$

where the sums are over cyclic permutations of the indices j,k,l. The basic idea to derive (3.3.18) is using integration by parts to move a derivative from B to ρ_ε when otherwise the order of the derivative applied to B would be greater than 2. Then using (3.3.10), the property $\beta\star\beta \leqslant c\beta$ and Lemma 3.3.1, one can show that the integral of the right-hand side of (3.3.18), as $\varepsilon\to 0$, is controlled either by

$$\begin{aligned}\sqrt{T-s}\,c_B\Bigg(&\sqrt{T-s}\,(\|\beta\|_1+2c)\sup_{ilm}\frac{\|D^3_{ilm}w\|_{\infty;[s,T]}}{\beta^i} \\ &+\sqrt{T-s}\,(\|\beta\|_1+c)\sup_{il}\frac{\|D^2_{il}w\|_{\infty;[s,T]}}{\beta^i} + Cc\sup_i\frac{\|D_i w\|_{\infty;[s,T]}}{\beta^i}\Bigg)\beta^j,\end{aligned}$$

or, exploiting the Cauchy–Schwarz inequality, by

$$\begin{aligned}\sqrt{T-s}\,c_B\Bigg(&3\sqrt{(T-s)c\|\beta\|_1}\sup_{ilm}\frac{\|D^3_{ilm}w\|_{\infty;[s,T]}}{\sqrt{\beta^i\beta^l}} \\ &+2\sqrt{T-s}\,c\sup_{il}\frac{\|D^2_{il}w\|_{\infty;[s,T]}}{\beta^i} + Cc\sup_i\frac{\|D_i w\|_{\infty;[s,T]}}{\beta^i}\Bigg)\sqrt{\beta^j\beta^k}.\end{aligned}$$

Finally, integrating by parts and using (3.3.2) and (3.3.11),

$$\lim_{\varepsilon\to 0}\left|\int_s^T\int_{\mathbb{R}^N} D^3_{jkl}F\rho_\varepsilon\right| \leqslant Cc_F\sqrt{T-s}\,(\beta^j\wedge\sqrt{\beta^j\beta^k}\,).$$

Collecting the estimates we have proved and also using (3.3.15) and (3.3.17) we can now obtain that

$$\sup_{ijk}\frac{\|D^3_{ijk}w\|_{\infty;[s,T]}}{\beta^i\wedge\sqrt{\beta^i\beta^j}} \leqslant c_3(T-s), \tag{3.3.19}$$

for some c_3 increasing in $T-s$ and depending only on C''_A, C and the same parameters already listed above.

Plugging (3.3.19) back into (3.3.17) and the resulting estimate into (3.3.15), we conclude that there exists $\mathfrak{c}(T-s)$ such that

$$\sup_i\frac{\|D_i w\|_{\infty;[s,T]}}{\beta^i} + \sup_{ij}\frac{\|D^2_{ij}w\|_{\infty;[s,T]}}{\beta^i\wedge\sqrt{\beta^i\beta^j}} + \sup_{ijk}\frac{\|D^3_{ijk}w\|_{\infty;[s,T]}}{\beta^i\wedge\sqrt{\beta^i\beta^j}} \leqslant \mathfrak{c}(T-s),$$

provided that $T-s$ is small enough, depending on $c, C''_A, c_B, c_F, c_G, \|\beta\|_1, C$. It is now standard to iterate these estimates a finite number of times and get the desired ones on the whole interval $[0,T]$. ∎

Using the bounds we have obtained, it is also possible to deduce a Lipschitz estimate in time of the space derivatives up to the second-order, as we state here below.

COROLLARY 3.3.3. *Let the assumptions of Proposition 3.3.2 be in force; also, suppose that* $\beta \in \ell^{\frac{1}{2}}(\mathbb{Z})$. *Then, for all* $j, k \in [\![N]\!]$, *one has* $w \in C^{0,1}([0,T]; C^2(\mathbb{R}^N))$, *with*

$$\sup_{jk}\left(\|\partial_t w\|_\infty + \frac{\|\partial_t D_j w\|_\infty}{\sqrt{\beta^j}} + \sup_{s\neq t} \frac{\|D^2_{jk}w(s,\cdot) - D^2_{jk}w(t,\cdot)\|_\infty}{\sqrt{\beta^j}\,|s-t|} \right) \leqslant K',$$

where K' *depends only on* $\sup_j \|A^{jj}\|_\infty$, $\sup_j \|B^j\|_\infty$, $\|\beta\|_{\frac{1}{2}}$, K *and the parameters thereof. Furthermore, one can choose* $K' = CK$ *with* C *linearly depending on* c_B.

PROOF. The differentiability in time of Dw comes directly from the equation. We prove now only the Lipschitz estimate for the second-order derivatives; the other ones are obtained in an analogous manner. Testing over (s,t) the equation of $D^2_{jk}w$ by ρ_ε (solving (3.3.1), with arbitrary $y \in \mathbb{R}^N$) and letting $\varepsilon \to 0$, the analogue of (3.3.16) one obtains, after using all estimates on the data and w, is

$$\left| D^2_{jk}w(s,y) - \int_{\mathbb{R}^N} D^2_{jk}w(t,\cdot)\,\mathrm{d}\rho(t) \right| \lesssim (t-s)(\beta^j \wedge \sqrt{\beta^j\beta^k}),$$

with implied constant (also below) depending only on K and the parameters thereof. Therefore,

$$\begin{aligned} \left|D^2_{jk}w(s,y) - D^2_{jk}w(t,y)\right| &\lesssim (t-s)(\beta^j \wedge \sqrt{\beta^j\beta^k}) \\ &\qquad + \left| \int_{\mathbb{R}^N} D^2_{jk}w(t,\cdot)\,\mathrm{d}\rho(t) - D^2_{jk}w(t,y) \right| \\ &\lesssim \left((t-s) + \sum_{1\leqslant l\leqslant N} \sqrt{\beta^l} \int_{\mathbb{R}^N} |\cdot^{\,l} - y^l|\,\mathrm{d}\rho(t) \right)\sqrt{\beta^j}. \end{aligned}$$

As A and B are bounded, we can test (3.3.1) by $|\cdot^{\,l} - y^l|^2$ to get

$$\int_{\mathbb{R}^N} |\cdot^{\,l} - y^l|^2\,\mathrm{d}\rho(t) = \int_s^t \int_{\mathbb{R}^N} A^{ll}\,\mathrm{d}\rho - 2\int_s^t \int_{\mathbb{R}^N} B^\ell(\cdot^{\,l} - y^l)\,\mathrm{d}\rho,$$

so that Young's inequality and Gronwall's lemma yield

$$\int_{\mathbb{R}^N} |\cdot^{\,l} - y^l|^2\,\mathrm{d}\rho(t) \leqslant (t-s)(\|A^{ll}\|_\infty + \|B^l\|^2_\infty)e^T.$$

Using Hölder's inequality, we conclude that $\left|D^2_{jk}w(s,y) - D^2_{jk}w(t,y)\right| \lesssim (t-s)\sqrt{\beta^j}$, with implied constant depending also on $\sup_j \|A^{jj}\|_\infty$, $\sup_j \|B^j\|_\infty$ and $\|\beta\|_{\frac{1}{2}}$. ∎

3.4. Existence and uniqueness for the Nash system

We first prove that, under suitable bounds on the Hamiltonian, the infinite-dimensional Nash system can have at most one classical solution, according to Definition 3.2.15.

THEOREM 3.4.1. *Let H be such that $\partial_{p^i} H^i(t,\cdot)$ is twice differentiable in $\mathbb{R}^\omega \times \mathbb{R}^\omega$, with*

$$\sup_{ik} \sum_j \big(|\partial_{p^j} H^i| + |D_j \partial_{p^j} H^i| + |\partial^2_{p^j p^k} H^i| + |D_j \partial^2_{p^j p^k} H^i| + |\partial^3_{p^j p^k p^i} H^i| \big) < \infty$$

uniformly for $(t,x,p) \in [0,T] \times \mathbb{R}^\omega \times \Omega$ with $\Omega \subset \mathbb{R}^\omega$ bounded with respect to the ∞-norm.[1] *Then there exists at most one classical solution to the Nash system (3.1.1).*

PROOF. Let u and v be two classical solutions as in the statement. Let $w = u - v$, so that it solves

$$\begin{aligned} &-\partial_t w^i - \sum_{jk} A^{jk} D^2_{jk} w^i + \Big(\int_0^1 \partial_p H^i(\mathcal{D}(su + (1-s)v))\,\mathrm{d}s \Big) \cdot \mathcal{D}w \\ &\quad + \sum_{j \neq i} \partial_{p^j} H^j(\mathcal{D}u) D_j w^i + \sum_{j \neq i} \Big(\int_0^1 \partial^2_{pp^j} H^j(\mathcal{D}(su + (1-s)v))\,\mathrm{d}s \Big) \cdot \mathcal{D}w\, D_j v^i = 0 \end{aligned}$$

for each $i \in \mathbb{N}$, with terminal condition $w|_{t=T} = 0$. Let now ρ^i solve (3.3.1) where

$$B^j = \begin{cases} \partial_{p^j} H^j(\mathcal{D}u) & \text{if } j \neq i \\ \displaystyle\int_0^1 \partial_{p^i} H^i(\mathcal{D}(su + (1-s)v))\,\mathrm{d}s & \text{if } j = i; \end{cases}$$

for the sake of simplicity we will do the computation assuming that $\rho(s) = \delta_y$, nevertheless what follows is to be understood in the limit as $\varepsilon \to 0$. Testing the equation for w^i by ρ^i over (s,T) we get, with the notation $x_{-N} := (x^j)_{j \geqslant N}$,

$$\begin{aligned} &w^i(s,y,x_{-N}) \\ &= \sum_{j \neq i} \int_s^T \int_{\mathbb{R}^N} \Big(\Big(\int_0^1 \partial_{p^j} H^i(\mathcal{D}(s'u + (1-s')v))\,\mathrm{d}s' \Big) D_j w^j \Big)\Big|_{(t,(z,x_{-N}))} \rho^i(t,z)\,\mathrm{d}z\mathrm{d}t \\ &\quad + \sum_{j \neq i} \int_s^T \int_{\mathbb{R}^N} \Big(\sum_k \Big(\int_0^1 \partial^2_{p^k p^j} H^j(*)\,\mathrm{d}s' \Big) D_k w^k D_j v^i \Big)\Big|_{(t,(z,x_{-N}))} \rho^i(t,z)\,\mathrm{d}z\mathrm{d}t \\ &\quad + \sum_{j \geqslant N} \int_s^T \int_{\mathbb{R}^N} \Big(2 \sum_k A^{jk} D^2_{jk} w^i + \partial_{p^j} H^j(\mathcal{D}u) D_j w^i \Big)\Big|_{(t,(z,x_{-N}))} \rho^i(t,z)\,\mathrm{d}z\mathrm{d}t, \end{aligned}$$

[1] That is, Ω is such that $\sup_\Omega |\cdot|_\infty := \sup_{p\in\Omega} \sup_i |p^i| < \infty$. Later, we will refer to such sets as ∞-*bounded* sets.

where for the sake of brevity we used $*$ as a placeholder for $\mathcal{D}(s'u+(1-s')v)$. As long as $j<N$, integrating by parts one finds

$$
\begin{aligned}
&-\int_{\mathbb{R}^N}\Big(\Big(\int_0^1 \partial_{p^j}H^i(*)\,\mathrm{d}s'\Big)D_j w^j\Big)\Big|_{(t,(z,x_{-N}))}\rho^i(t,z)\,\mathrm{d}z\mathrm{d}t\\
&=\int_{\mathbb{R}^N}\Big(\Big(\int_0^1 \partial_{p^j}H^i(*)\,\mathrm{d}s'\Big)w^j\Big)\Big|_{(t,(z,x_{-N}))}D_j\rho^i(t,z)\,\mathrm{d}z\mathrm{d}t\\
&\quad+\int_{\mathbb{R}^N}\Big(\Big(\int_0^1 D_j\partial_{p^j}H^i(*)\,\mathrm{d}s'\Big)w^j\Big)\Big|_{(t,(z,x_{-N}))}\rho^i(t,z)\,\mathrm{d}z\mathrm{d}t\\
&\quad+\int_{\mathbb{R}^N}\Big(\sum_\ell\Big(\int_0^1 \partial^2_{p^\ell p^j}H^i(*)\,\mathrm{d}s'\Big)D^2_{j\ell}(su^\ell+(1-s)v^\ell)\,w^j\Big)\Big|_{(t,(z,x_{-N}))}\rho^i(t,z)\,\mathrm{d}z\mathrm{d}t\,;
\end{aligned}
$$

analogously,

$$
\begin{aligned}
&-\int_{\mathbb{R}^N}\Big(\Big(\int_0^1 \partial^2_{p^k p^j}H^j(*)\,\mathrm{d}s'\Big)D_k w^k D_j v^i\Big)\Big|_{(t,(z,x_{-N}))}\rho^i(t,z)\,\mathrm{d}z\mathrm{d}t\\
&\quad=\int_{\mathbb{R}^N}\Big(\Big(\int_0^1 \partial^2_{p^k p^j}H^j(*)\,\mathrm{d}s'\Big)w^k D^2_{jk}v^i\Big)\Big|_{(t,(z,x_{-N}))}\rho^i(t,z)\,\mathrm{d}z\\
&\qquad+\int_{\mathbb{R}^N}\Big(\Big(\int_0^1 \partial^2_{p^k p^j}H^j(*)\,\mathrm{d}s'\Big)w^k D_j v^i\Big)\Big|_{(t,(z,x_{-N}))}D_k\rho^i(t,z)\,\mathrm{d}z\\
&\qquad+\int_{\mathbb{R}^N}\Big(\Big(\int_0^1 D_k\partial^2_{p^k p^j}H^j(*)\,\mathrm{d}s'\Big)w^k D_j v^i\Big)\Big|_{(t,(z,x_{-N}))}\rho^i(t,z)\,\mathrm{d}z\\
&\qquad+\int_{\mathbb{R}^N}\Big(\sum_\ell\Big(\int_0^1 \partial^3_{p^\ell p^k p^j}H^j(*)\,\mathrm{d}s'\Big)\\
&\qquad\qquad D^2_{k\ell}(su^\ell+(1-s)v^\ell)\,w^k D_j v^i\Big)\Big|_{(t,(z,x_{-N}))}\rho^i(t,z)\,\mathrm{d}z\,.
\end{aligned}
$$

Therefore, using Lemma 3.3.1 and the properties required to classical solutions, we deduce that for $T-s<T^*$

$$\sup_i\|w^i\|_{\infty;[s,T]}\lesssim\sqrt{T-s}\,\Big(\sup_i\|w^i\|_{\infty;[s,T]}+\epsilon_N\Big)$$

with $\epsilon_N\to0$ as $N\to\infty$ and implied constant independent of N. It follows that if s is sufficiently close to T one has $w^i=0$ for all i (that is, $u=v$) on $[s,T]\times\mathbb{R}^\omega$. Finally define $\hat{s}:=\min\{s\in[0,T]:\ u=v \text{ on } [s,T]\times\mathbb{R}^\omega\}$; if $\hat{s}>0$ then by the argument above with $\hat{s}$ in lieu of T we get a contradiction, thus $\hat{s}=0$ and the proof is complete. ∎

To prove existence, we solve a version of the Nash system (3.1.1) reduced onto $\mathbb{R}^N_T:=[0,T]\times\mathbb{R}^N$ and then pass to the limit as $N\to\infty$ to obtain a solution to the infinite-dimensional system, thanks to the stability provided by our estimates of Proposition 3.3.2. This is going to be done by observing that letting $B_i(t,x):=$

$\hat{B}_i(t,x,\mathcal{D}(t,x))$, with

$$\hat{B}_i^j(t,x,p) := \begin{cases} \partial_{p^j} H^j(t,x,p) & \text{if } j \neq i \\ \displaystyle\int_0^1 \partial_{p^i} H^i(t,x,p^{-i},sp^i)\,\mathrm{d}s & \text{if } j = i\,, \end{cases} \tag{3.4.1}$$

and

$$F^i(t,x) := H^i(t,x,\mathcal{D}u^{-i}(t,x),0), \tag{3.4.2}$$

the equations in (3.1.1) take the form $-\partial_t u^i - \mathrm{Tr}(AD^2u^i) + \langle B_i, Du^i\rangle = F^i$, in such a way that they can be regarded as a system of linear transport-diffusion equations in the context of a fixed point argument.

When performing the reduction from $\mathbb{R}^\omega_T$ to $\mathbb{R}^N_T$, we will identify $\mathbb{R}^N \subset \mathbb{R}^\omega$ via $(x^0,\dots,x^{N-1}) \mapsto (x^0,\dots,x^{N-1},0,\dots)$. Given $V\colon \mathbb{R}^\omega \to \mathbb{R}$ we will consider the projection $V_N(y) := V(y)$ for all $y \in \mathbb{R}^N$, while for $A\colon \mathbb{R}^\omega \to \ell^\infty(\mathbb{N}^2)$ the notation A_N will mean that we have also restricted the infinite-dimensional matrix to $\mathbb{R}^N$; that is, $A_N(y) := A(y)|_{[\![N]\!]} = (A(y)^{ij})_{i,j\in[\![N]\!]}$ for all $y \in \mathbb{R}^N$. Finally, the spaces defined in Section 3.2 are readily adapted to functions on $\mathbb{R}^N$, simply by replacing $\mathbb{R}^\omega$ with $\mathbb{R}^N$.

In the upcoming statement, which is our main theorem, we will make use of the notations $\mathbb{R}^\omega_x$ and $\mathbb{R}^\omega_p$ presented in the opening section; $\mathbb{R}^\omega_T$ will always mean $[0,T]\times\mathbb{R}^\omega_x$.

THEOREM 3.4.2. *Let β be as in Corollary 3.3.3; for any $i \in \mathbb{N}$, define $\beta_i := \beta^{i-\cdot} \in \ell^{\frac{1}{2}}(\mathbb{Z})$. Assume that the following hypotheses are fulfilled, uniformly in the parameters $i,j,k \in \mathbb{N}$:*

- *$A\colon \mathbb{R}^\omega_T \to \ell^\infty(\mathbb{N}^2)$ is such that*
$$A^{ij} = A^{ji} \in C^{\frac{\gamma}{2},\gamma}(\mathbb{R}^\omega_T) \quad \text{for some } \gamma \in (0,1)\,,$$
with $D_k A^{ij} = 0$ unless $i = j = k$, and
$$D_k A^{kk} \in C^{\frac{\gamma}{2},\gamma}(\mathbb{R}^\omega_T) \cap C^0([0,T];C^2(\mathbb{R}^\omega))\,;$$
- *$G^i \in C^{3+\gamma-}_{\beta_i}(\mathbb{R}^\omega)$;*
- *for any ∞-bounded $\Omega \subset \mathbb{R}^\omega_p$,*
$$\partial_{p^i} H^i \in C^0([0,T];C^2_{\beta_i}(\mathbb{R}^\omega_x\times\Omega)) \cap C^0(\mathbb{R}^\omega_T;C^{3-}_{\beta_i}(\Omega))\,;$$
also,
$$D_j\partial_{p^i} H^i,\ \partial_{p^j}\partial_{p^i} H^i \in C^0(\Omega;C^{\frac{\gamma}{2},\gamma}(\mathbb{R}^\omega_T))\,.$$

Then there exists $T^* > 0$ such that if $T < T^*$ the Nash system (3.1.1) has a unique classical solution u (with β_i defined as above), which also belongs to $\ell^\infty(\mathbb{N};C^0([0,T];C^{2,1}_{\beta_i}(\mathbb{R}^\omega)) \cap C^{0,1}([0,T];C^{2-}_{\sqrt{\beta_i}}(\mathbb{R}^\omega)))$.

PROOF. We only need to prove existence, since uniqueness will then follow because the hypotheses of Theorem 3.4.1 are satisfied.

Let $N \in \mathbb{N}$ and let $\mathcal{Y}$ be the closed subset of

$$\big(C^0([0,T]; C^{2,1}_{\beta_i}(\mathbb{R}^N)) \cap C^{0,1}([0,T]; C^{2-}_{\sqrt{\beta_i}}(\mathbb{R}^N))\big)^N$$

of all u such that

$$\sup_i \|u^i\|_{C^0([0,T];C^{3-}_{\beta_i}(\mathbb{R}^N))} \leqslant R \quad \text{and} \quad \sup_i \|u^i\|_{C^{0,1}([0,T];C^{2-}_{\sqrt{\beta_i}}(\mathbb{R}^N))} \leqslant R', \tag{3.4.3}$$

for some $R, R' > 0$ to be determined. We will denote by $\|u\|$ the norm on $(C^0([0,T]; C^{2,1}_{\beta_i}(\mathbb{R}^N)) \cap C^{0,1}([0,T]; C^{2-}_{\sqrt{\beta_i}}(\mathbb{R}^N)))^N$, defined as the sum of the two norms above. Given $u \in \mathcal{Y}$, let $\mathcal{S}$ be the map that associates to u the solution w to

$$\begin{cases} -\partial_t w^i - \displaystyle\sum_{j,k \in [\![N]\!]} A^{jk}_N D^2_{jk} w^i + \sum_{j \in [\![N]\!]} B^j_{i,N} D_j w^i = F^i_N & \text{in } \mathbb{R}^N_T \\ w^i|_{t=T} = G^i_N \end{cases} \qquad i \in [\![N]\!],$$

with B and F being defined as in (3.4.1) and (3.4.2); here we have used the notation with the subscript N that we have previously introduced. By [**57**, Theorem 8.12.1], w and Dw are of class $C^{1+\frac{\gamma}{2},2+\gamma}$. Also, it is not difficult to see that there exists a constant κ_R which depends only on β, c_H and R such that

$$\sup_{ih}\big(\|B^h_i\|_{C^0([0,T];C^2_{\beta_h}(\mathbb{R}^N))} + \|F^i\|_{C^0([0,T];C^2_{\beta_i}(\mathbb{R}^N))}\big) \leqslant \kappa_R\,,$$

so, by Proposition 3.3.2 and Corollary 3.3.3 one can choose $R > \sup_i \|G^i\|_{3+\gamma-;\beta_i}$, T sufficiently small and R' large enough in such a way that, for all $N \in \mathbb{N}$, $\mathcal{S}$ is well-defined with values in $\mathcal{Y}$.

Consider now $u, v \in \mathcal{Y}$ and denote by $B_{i,N}[u]$ and $B_{i,N}[v]$ the corresponding drifts;[2] then, letting $\bar{w} = \mathcal{S}(u) - \mathcal{S}(v)$ and $\tilde{F}^i_N = \langle B_{i,N}[v] - B_{i,N}[u], D\mathcal{S}(v)^i\rangle$ we have

$$\begin{cases} -\partial_t \bar{w}^i - \displaystyle\sum_{jk} A^{jk}_N D^2_{jk} \bar{w}^i + \langle B_{i,N}[u], D\bar{w}^i\rangle = \tilde{F}^i_N \\ \bar{w}|_{t=T} = 0. \end{cases}$$

Note that

$$\tilde{F}^i = \sum_j \Big(\int_0^1 \partial_p \hat{B}^j_{i,N}[su + (1-s)v]\,\mathrm{d}s\Big) \cdot \mathcal{D}(u-v)\, D_j \mathcal{S}(v)^i, \tag{3.4.4}$$

hence, letting $c_{\tilde{F}} := \sup_i \|\tilde{F}^i\|_{C^0([0,T];C^2_{\beta_i}(\mathbb{R}^N))}$, we have $c_{\tilde{F}} \leqslant \kappa'_R \|u - v\|$, with κ'_R independent of N. In addition, by Proposition 3.3.2,

$$\sup_i \|\bar{w}^i\|_{C^0([0,T];C^{3-}_{\beta_i}(\mathbb{R}^N))} \leqslant c_{\tilde{F}} \tilde{K}(T)\,.$$

[2]That is, $B_{i,N}[u](t,x)$ is a short notation for $\hat{B}_{i,N}(t,x,\mathcal{D}u(t,x))$, according to (3.4.1), and analogously for v.

Then, also using Corollary 3.3.3, we deduce that there exists $\bar{T} > 0$, independent of N, such that if $T \leqslant \bar{T}$ then $|||\mathcal{S}(u) - \mathcal{S}(v)||| \leqslant \frac{1}{2}|||u - v|||$; hence the contraction theorem yields the existence of a unique fixed point for $\mathcal{S}$ in $\mathcal{Y}$, which is by its definition the solution u_N to the Nash system on $\mathbb{R}^N_T$ with data A_N, H_N and G_N.

For any $\delta \in (0,1)$, by the Ascoli–Arzelà theorem and a standard diagonal argument, up to a subsequence which is common to all $i \in \mathbb{N}$, each $(u^i_N)_{N \geqslant i}$ compactly converges in $C^0([0,T]; C^{2+\delta}_{\beta_i}(\mathbb{R}^\omega)) \cap C^\delta([0,T]; C^{2-}_{\sqrt{\beta_i}}(\mathbb{R}^\omega))$ to a function u^i enjoying bounds (3.4.3).

Since $\beta \in \ell^{\frac{1}{2}}(\mathbb{Z})$, by the dominated convergence theorem

$$\sum_{j,k<N} A^{jk}_N D^2_{jk} u^i_N \longrightarrow \sum_{j,k \in \mathbb{N}} A^{jk} D^2_{jk} u^i \,,$$

$$\sum_{\substack{j<N \\ j \neq i}} \partial_{p^i} H^j_N(\mathcal{D}u_N) D_j u^i_N \longrightarrow \sum_{\substack{j \in \mathbb{N} \\ j \neq i}} \partial_{p^i} H^j(\mathcal{D}u) D_j u^i$$

compactly in $C^0(\mathbb{R}^\omega_T)$ as $N \to \infty$. As also $H^i_N(\mathcal{D}u_N) \to H^i(\mathcal{D}u)$ compactly in $C^0(\mathbb{R}^\omega_T)$, by the equation we have that so $\partial_t u^i_N$ compactly converges in $C^0(\mathbb{R}^\omega_T)$ as well. In particular, for any $x \in \mathbb{R}^\omega_T$, $\partial_t u^i(\cdot, x)^i_N$ uniformly converges to some $v^i(\cdot, x)$ which is thus $\partial_t u^i(\cdot, x)$ by the fundamental theorem of calculus. We conclude that u solves (3.1.1) pointwise. Furthermore, by the decay estimates on the derivatives we see that the series appearing in the equations converge in $C^0(\mathbb{R}^\omega_T)$. To complete the proof, we need to show that such a convergence is uniform in i.

Note that

$$\sum_{0 \leqslant j,k \leqslant N} \|D^2_{jk} u^i\|_\infty \lesssim \bigg(\sum_{0 \leqslant j \leqslant N} \gamma^{j-i} \bigg)^2,$$

where we have set $\gamma^i := \sqrt{\beta^i}$ and the implied constant is independent of $i \in \mathbb{N}$. The latter sum equals

$$\begin{cases} \displaystyle\sum_{i-N \leqslant k \leqslant i} \gamma^k & \text{if } i \geqslant N \\ \displaystyle\gamma^0 + \; 2 \sum_{1 \leqslant k \leqslant i \wedge (N-i)} \gamma^k \; + \sum_{i \wedge (N-i) < k \leqslant i \vee (N-i)} \gamma^k & \text{if } i < N. \end{cases}$$

Since $(\gamma^k)_{k \geqslant 0}$ is summable, $\sum_{i-N \leqslant k \leqslant i} \gamma^k \to 0$ as $i \to \infty$, thus there exists $\bar{\imath} = \bar{\imath}(N) \geqslant N$ such that $\sum_{i-N \leqslant k \leqslant i} \gamma^k \leqslant \sum_{0 \leqslant k \leqslant \bar{\imath}} \gamma^k$ for all $i \geqslant N$; then one easily sees that

$$\sum_{0 \leqslant j,k \leqslant N} \|D^2_{jk} u^i\|_\infty \lesssim \bigg(\sum_{0 \leqslant k \leqslant \bar{\imath}} \gamma^k \bigg)^2 \quad \forall\, i \in \mathbb{N}.$$

Similarly, there exists $\tilde{\imath} = \tilde{\imath}(N) \geqslant N$ such that

$$\sum_{0 \leqslant j \leqslant N} \|D_j u^j D_j u^i\|_\infty \lesssim \beta^0 \sum_{0 \leqslant j \leqslant N} \beta^{j-i} \lesssim \sum_{0 \leqslant k \leqslant \tilde{\imath}(N)} \beta^k \quad \forall\, i \in \mathbb{N}.$$

Then the series appearing in the Nash system converge in $C^0(\mathbb{R}^\omega_T)$ uniformly in $i \in \mathbb{N}$. ∎

CHAPTER 4

Generalised Mean Field Games. A priori estimates on nonsymmetric Nash systems with semimonotonicity

Contents

4.1. *Amuse-bouche.* Linear-quadratic Mean-Field-like games

Consider the Nash system (2.2.3) without the shift-invariance assumption. Instead, make a *Mean-Field-like* assumption. Given $(a^i_{jk})_{i,j,k} \in \mathscr{S}(N)^N$ we will use the notation $B(a)$ for the matrix $B(a)_{hk} = a^h_{hk}$. Given the i-th vector e^i of the canonical basis of $\mathbb{R}^N$, we will write $E^i = e^i(e^i)^\mathsf{T}$. The indices will range over $[\![N]\!]$.

Note that system (2.2.3) can be rewritten in a forward form as

$$
\begin{cases} \dot{c}^i - (c^i)^\mathsf{T} E^i c^i + B(c)^\mathsf{T} c^i + c^i B(c) = f^i \\ c^i(0) = g^i \end{cases} \qquad i \in [\![N]\!], \tag{4.1.1}
$$

where $c^i(t), f^i(t), g^i \in \mathscr{S}(N)$ for all $i \in [\![N]\!]$ and $t \in [0,T]$.

THEOREM 4.1.1. *Let $T > 0$ and let $f \in L^1((0,T); \mathscr{S}(N)^N)$ and $g \in \mathscr{S}(N)^N$ satisfy*

$$\sup_i \Big(N \sum_{\substack{h,k \\ k \neq i}} |\bullet^i_{hk}|^2 + N \sum_{\substack{k \\ k\neq i}} |\bullet^k_{ki}|^2 + |\bullet^i_{ii}|^2 \Big) \leqslant \kappa_\bullet, \quad \kappa_\bullet > 0, \quad \text{for } \bullet \in \{f, g\}. \tag{4.1.2}$$

Suppose that $B(f) \geqslant -K_f I$ and $B(g) \geqslant -K_g I$ for some constants K_f and K_g such that

$$K_g < \sup_{M \in \mathbb{R}} M e^{-2MT} = \frac{1}{2eT}, \qquad K_f < \sup_{M\in\mathbb{R}} \frac{M(Me^{-MT} - K_g)}{T\big(1 \vee \frac{e^{2MT}-1}{2MT}\big)}. \tag{4.1.3}$$

Then there exists $N_0 \in \mathbb{N}$ such that if $N > N_0$ there exists a unique absolutely continuous solution c to (4.1.1) on $[0,T)$, which satisfies

$$\sup_i \Big(N \sum_{\substack{h,k \\ k \neq i}} |c^i_{hk}|^2 + N \sum_{\substack{k \\ k\neq i}} |c^k_{ki}|^2 + |c^i_{ii}|^2 \Big) \leqslant C \tag{4.1.4}$$

for some constant C which is independent of N.

Note that (4.1.3) is automatically satisfied when $K_f, K_g \leqslant 0$ for any $T > 0$. Otherwise, for fixed K_f and K_g, it poses a restriction on the size of T (or, similarly, a restriction on the size of $(K_f)_+$ and $(K_g)_+$ once T is fixed). Back to the value functions u^i, the previous estimate reads as follows:

$$\sup_i \Big(N \sum_{\substack{h,k \\ k \neq i}} \|D^2_{hk} u^i\|^2_\infty + N \sum_{\substack{k \\ k\neq i}} \|D^2_{ki} u^k\|^2_\infty + \|D^2_{ii} u^i\|^2_\infty \Big) \leqslant C;$$

in particular,

$$\sup_i \sum_{j \colon j \neq i} \|D_j \alpha^i\|^2_\infty = \sup_i \sum_{j\colon j\neq i} \|D^2_{ij} u^i\|^2_\infty \leqslant \frac{C}{N}.$$

REMARK 4.1.2. We use the terminology *Mean-Field-like* since any f^i such that

$$\sup_i |f^i_{ii}| + N \sup_{\substack{i,j \\ j \neq i}} |f^i_{ij}| + N^2 \sup_{\substack{i,j,k \\ j\neq i,\, k \neq i}} |f^i_{jk}| \leqslant C$$

satisfies (4.1.2), with κ_f depending on C (and not on N). In turn, the previous inequality is satisfied when $f^i(x) = V^i(x^i, (N-1)^{-1} \sum_{j\neq i} \delta_{x^j})$, where V^i is a smooth enough function defined over $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ (see for instance [**23**, Proposition 6.1.1]).

The proof of the theorem is based on the following lemmata.

LEMMA 4.1.3. *Let c be an absolutely continuous solution to (4.1.1) with $B(c) \geqslant -MI$ on $[0,T)$ for some $M \in \mathbb{R}$. Assume (4.1.2). Then the following estimates* *hold on* $[0,T)$*:*

$$\sum_{\substack{h,k\\k\neq i}} |c^i_{hk}|^2 \leqslant \kappa_0 N^{-1}, \qquad \kappa_0(t) := (\kappa_g + \kappa_f t) e^{(1+4M_+)t}, \tag{4.1.5}$$

$$\sup_k \sum_{i\neq k} |c^i_{ik}|^2 \leqslant \kappa_1 N^{-1}, \qquad \kappa_1(t) := 2\kappa_0(t) e^{2\int_0^t \sqrt{\kappa_0}}, \tag{4.1.6}$$

$$\sup_i |c^i_{ii}|^2 \leqslant \kappa_2, \qquad \kappa_2(t) = (\kappa_g + (\kappa_f + \kappa_0\kappa_1 N^{-2})t) e^{(2+M_+)t}, \tag{4.1.7}$$

where $M_+ = M \vee 0$*. As a consequence,* c *continuously extends on* $[0,T]$*.*

Proof. Multiplying the equation for c^i_{hk} by c^i_{hk} and summing over h and $k \neq i$ we have

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} \sum_{\substack{h,k\\k\neq i}} |c^i_{hk}|^2 &= \sum_{\substack{h,k\\k\neq i}} f^i_{hk} c^i_{hk} - \sum_{\substack{j,h,k\\j,k\neq i}} c^i_{hk} c^i_{jh} c^j_{jk} - \sum_{\substack{j,h,k\\k\neq i}} c^i_{hk} c^i_{jk} c^j_{jh} \\ &= \sum_{\substack{h,k\\k\neq i}} f^i_{hk} c^i_{hk} - \mathrm{Tr}(\hat{c}^{i\mathsf{T}} B(c) \hat{c}^i) - \mathrm{Tr}(\tilde{c}^{i\mathsf{T}} B(c) \tilde{c}^i),\end{aligned}$$

where we have set $\hat{c}^i_{hk} := c^i_{hk}(1-\delta_{hi})$ and $\tilde{c}^i_{hk} = c^i_{hk}(1-\delta_{ki})$. It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{\substack{h,k\\k\neq i}} |c^i_{hk}|^2 \leqslant \kappa_f N^{-1} + (1+4M_+) \sum_{\substack{h,k\\k\neq i}} |c^i_{hk}|^2,$$

thus, by Gronwall's inequality, (4.1.5) is proved. Multiplying the equation for c^i_{ik} by c^i_{ik} and summing over $i \neq k$ we obtain

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} \sum_{i\neq k} |c^i_{ik}|^2 &= \sum_{i\neq k} f^i_{ik} c^i_{ik} - \sum_{i\neq k} B(c)_{kk} |c^i_{ik}|^2 - \sum_{i,j\neq k} c^i_{ik} c^i_{ji} c^j_{jk} - \sum_{\substack{i\neq k\\j\neq i}} c^i_{ik} c^i_{jk} c^j_{ji} \\ &\leqslant \sum_{i\neq k} f^i_{ik} c^i_{ik} - \sum_{i\neq k} B(c)_{kk} |c^i_{ik}|^2 - \mathrm{Tr}(\hat{B}(c)^{\mathsf{T}} B(c) \hat{B}(c)) - \sum_{\substack{i\neq k\\j\neq i}} c^i_{ik} c^i_{jk} c^j_{ji},\end{aligned}$$

where $\hat{B}(c)_{hk} = B(c)_{hk}$ if $h = k$ and it is null otherwise. By (4.1.5)

$$\Bigg| \sum_{\substack{i\neq k\\j\neq i}} c^i_{ik} c^i_{jk} c^j_{ji} \Bigg| \leqslant \sup_\ell \sum_{i\neq \ell} |c^i_{i\ell}|^2 \Bigg(\sum_{\substack{i\neq k\\j\neq i}} |c^i_{jk}|^2 \Bigg)^{\frac{1}{2}} \leqslant \sqrt{\kappa_0} \sup_\ell \sum_{i\neq\ell} |c^i_{i\ell}|^2.$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{i\neq k} |c^i_{ik}|^2 \leqslant \kappa_f N^{-1} + (1+4M_+) \sum_{i\neq k} |c^i_{ik}|^2 + 2\sqrt{\kappa_0} \sup_\ell \sum_{i\neq\ell} |c^i_{i\ell}|^2;$$

that is,

$$\sup_k \sum_{i\neq k} |c^i_{ik}(t)|^2 \leqslant (\kappa_g + \kappa_f t) N^{-1} + \int_0^t (1 + 4M_+ + 2\sqrt{\kappa_0}\,) \sup_k \sum_{i\neq k} |c^i_{ik}|^2.$$

By Gronwall's inequality one finds (4.1.6). Consider at this point that c^i_{ii} solves

$$\dot c^i_{ii} + |c^i_{ii}|^2 + 2\sum_{j\neq i} c^j_{ji} c^i_{ij} = f^i_{ii}, \tag{4.1.8}$$

where by (4.1.5) and (4.1.6)

$$\Big|\sum_{j\neq i} c^j_{ji} c^i_{ij}\Big|^2 \leqslant \kappa_0\kappa_1 N^{-2}.$$

Therefore, multiplying equation (4.1.8) by c^i_{ii} one obtains

$$\frac{\mathrm{d}}{\mathrm{d}t}|c^i_{ii}|^2 \leqslant \kappa_f + (2+M_+)|c^i_{ii}|^2 + \kappa_0\kappa_1 N^{-2}$$

and thus (4.1.7) by Gronwall's inequality. ∎

Lemma 4.1.4. *Under the hypotheses of Theorem 4.1.1, let c be as in Lemma 4.1.3. Suppose that*

$$Me^{-2MT} > K_g, \qquad K_f < \frac{M(Me^{-MT}-K_g)}{T\big(1\vee \frac{e^{2MT}-1}{2MT}\big)}\,.$$

Then there exists $N_*(T)$ such that $B(c)(T) > -MI$ provided that $N > N_*(T)$. Furthermore, the map $T\mapsto N_*(T)$ is continuous on $[0,+\infty)$.

Proof. Consider that $B(c)$ solves the equation

$$\dot B(c) + B(c)^2 = B(f) - D. \qquad \text{where} \quad D_{ik} = \sum_{j\neq i} c^i_{kj}c^j_{ji}. \tag{4.1.9}$$

By estimates (4.1.5) and (4.1.6),

$$\|D\|_2^2 = \sum_{i,k}\Big(\sum_{j\neq i} c^i_{kj}c^j_{ji}\Big)^2 \leqslant \sup_i \sum_{j\neq i}|c^j_{ji}|^2\cdot \sum_{\substack{i,j,k\\ j\neq i}} |c^i_{jk}|^2 \leqslant \kappa_0\kappa_1 N^{-1}.$$

Let now ξ solve the linear equation $\dot\xi = B(c)^\mathsf{T}\xi$ on $[0,T)$ with terminal condition $\xi(T)=\zeta$, for some arbitrary $\zeta\in\mathbb{S}^{N-1}$; note that since $\frac{\mathrm{d}}{\mathrm{d}t}|\xi|^2 \geqslant -4M|\xi|^2$ we have

$$1\wedge e^{2M(T-\cdot)} \leqslant |\xi| \leqslant 1\vee e^{2M(T-\cdot)} \quad \text{on } [0,T].$$

By (4.1.9) we get

$$\begin{aligned}(\xi^\mathsf{T}B(c)\xi)^{\dot{}} = \xi^\mathsf{T}(B(f)-D+B(c)B(c)^\mathsf{T})\xi \geqslant \xi^\mathsf{T}(B(f)-D)\xi \\ \geqslant -(K_f + \sqrt{\kappa_0\kappa_1}N^{-\frac12})|\xi|^2\,.\end{aligned}$$

If $K_f<0$ and N is large enough, then $B(c)(T) > -\nu K_g I$, where

$$\nu := \begin{cases} e^{2MT} & \text{if } MK_g \geqslant 0 \\ 1 & \text{if } MK_g \leqslant 0, \end{cases}$$

thus it is easily seen that $B(c)(T) > -MI$. If $K_f\geqslant 0$, then

$$B(c)(T) \geqslant -\Big(\nu K_g + \int_0^T (K_f+\sqrt{\kappa_0\kappa_1}N^{-\frac12})\,(1\vee e^{2M(T-\cdot)})\Big)I. \tag{4.1.10}$$

It follows that in order to have $B(c)(T) > -MI$ it suffices that

$$\nu K_g + TK_f(1 \vee h(MT)) + N^{-\frac{1}{2}} T \sqrt{\kappa_0(T)\kappa_1(T)}\, h(MT) < M,$$

where $h(z) := (e^{2z} - 1)/(2z)$. This is guaranteed by our assumptions on K_g, K_f and M (by which $\nu K_g + TK_f(1 \vee h(MT)) < M$), provided that N is large enough. The continuity of $N_*(T)$ is easily seen as one can write it explicitly using the estimates above. ∎

PROOF OF THEOREM 4.1.1. Fix M according to Lemma 4.1.4 (note that this implies that $M > K_g$). By the Cauchy–Lipschitz theorem there exists $\tau > 0$ such that (4.1.1) has a unique absolutely continuous solution on $[0, \tau)$. By taking τ smaller if necessary, we may suppose that by continuity $B(c) > -MI$ on $[0, \tau)$. Then

$$\bar{\tau} := \sup\{\tau > 0 : \text{(4.1.1) has a unique absolutely continuous solution with } B(c) > -MI \text{ on } [0, \tau)\}$$

is well-defined. Seeking for a contradiction, suppose that $\bar{\tau} < T$. Let $N_0 := \max_{[0,T]} N_*$, where N_* is given by Lemma 4.1.4. By Lemma 4.1.3, c continuously extends on $[0, \bar{\tau}]$, with $B(c)(\bar{\tau}) > -MI$ as guaranteed by Lemma 4.1.4, thanks to our choice of N_0. By the Cauchy–Lipschitz theorem one can extend the solution on $[0, \tau')$ for some $\tau' > \bar{\tau}$ and by continuity we may suppose that $B(c) > -MI$ on $[0, \tau')$. This contradicts the maximality of $\bar{\tau}$, thus $\bar{\tau} \geqslant T$. Finally, estimate (4.1.4) follows from (4.1.5), (4.1.6) and (4.1.7). This concludes the proof. ∎

We can now move on to the main course.

4.2. Main setting and assumptions

Consider the following semilinear (backward) parabolic system for the unknowns $u^i = u^i_N : [0, T] \times (\mathbb{R}^d)^N \to \mathbb{R}$, $i \in [\![N]\!]$,

$$\begin{cases} -\partial_t u^i - \operatorname{Tr}((\sigma \mathsf{I} + \beta \mathsf{J}) D^2 u^i) + \dfrac{1}{2}|D_i u^i|^2 + \displaystyle\sum_{j \in [\![N]\!] \setminus \{i\}} D_j u^j \cdot D_j u^i = f^i \\ u^i|_{t=T} = g^i \end{cases} \tag{4.2.1}$$

stated in $[0, T] \times (\mathbb{R}^d)^N \ni (t, x) = (t, x^0, \ldots, x^{N-1})$. The main goal of this chapter is to derive estimates on u^i and their space/time derivatives that are *stable* with respect to the number of equations; that is, uniformly in N.

The data are the maps $f^i, g^i : (\mathbb{R}^d)^N \to \mathbb{R}$, the parameters $\sigma > 0$ and $\beta \geqslant 0$ and the horizon $T \geqslant 0$. Recall that $\mathsf{I} = I_{Nd}$ and $\mathsf{J} = J_N \otimes I_d$.

System (4.2.1) describes Markovian Nash equilibria in N-player differential games, in particular it characterises the value function u^i of the i-th agent for each $i \in [\![N]\!]$. In our setting, agents control via feedbacks $\alpha^i = \alpha^i(t, x)$ their own states, which are driven by the following $\mathbb{R}^d$-valued SDEs on $[0, T]$:

$$\mathrm{d}X^i_t = \alpha^i(t, X_t)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}B^i_t + \sqrt{2\beta}\,\mathrm{d}W_t, \quad i \in \{1, \ldots, N\}, \tag{4.2.2}$$

where the B_t^i's and W_t are d-dimensional independent Brownian motions. The Brownian motions B_t^i correspond to the individual noises, while W_t is the so-called common noise, as it is the same for all the equations. The i-th agents aims at mimimising the following cost functional

$$\alpha^i \mapsto \mathbb{E}\bigg[\int_0^T \Big(\frac{1}{2}|\alpha^i(s, X_s)|^2 + f^i(X_s)\Big)\,\mathrm{d}s + g^i(X_T)\bigg].$$

It is known that the choice $\alpha^{*,i}(t,x) = -D_i u^i(t,x)$ characterises Nash equilibria, see for instance [**25, 46**]. Moreover, since one expects uniqueness of solutions to (4.2.1) by its (uniformly) parabolic structure, such equilibria are unique.

We will suppose that all data f^i and g^i are sufficiently smooth (locally C^2 and C^4, respectively) with bounded derivatives. Solutions to the Nash system are assumed to be classical, locally C^1 in time and C^4 in space, with bounded derivatives. Moreover, we require second order derivatives in space to be uniformly continuous. Note that we are not assuming that f^i, g^i are globally bounded here: they may have linear growth.

4.2.1. Two notions of (semi)monotonicity. Along with the hypothesis of *Mean-Field-like* interactions (which will be presented later in Section 4.2.2), our main structural assumption on the system will be of *semimonotonicity* of the data, according to either one of the following definitions.

Definition 4.2.1. Let $h \in C^1((\mathbb{R}^d)^N; \mathbb{R}^N)$. For $M \geqslant 0$, we say that h is *M-$\mathscr{D}$-semimonotone* if

$$\sum_{i \in [\![N]\!]} (D_i h^i(x) - D_i h^i(y)) \cdot (x^i - y^i) \geqslant -M|x-y|^2 \tag{4.2.3}$$

for all $x, y \in (\mathbb{R}^d)^N$, and M is called the *$\mathscr{D}$-semimonotonicity constant*. We simply say that h is *$\mathscr{D}$-semimonotone* if (4.2.3) holds for some $M \geqslant 0$. If (4.2.3) holds with $M = 0$, we say that h is *$\mathscr{D}$-monotone*.

In order to abridge the notation in what follows we shall write $h|_y^x = h(x) - h(y)$ and

$$\mathcal{D}[h](x,y) := \sum_{i \in [\![N]\!]} D_i h^i\big|_y^x \cdot (x^i - y^i), \tag{4.2.4}$$

so that, inequality (4.2.3) will appear as $\mathcal{D}[h] \geqslant -MQ$ on $(\mathbb{R}^d)^{2N} \simeq (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$, where $Q = Q_{\mathsf{A}}$ is the quadratic form induced on $(\mathbb{R}^d)^{2N}$ by the matrix $\mathsf{A} := \left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right) \otimes \mathsf{I} \in \mathscr{S}(2Nd)$.

Definition 4.2.2. Let $h\colon (\mathbb{R}^d)^N \to \mathbb{R}^N$. For $\kappa \geqslant 0$, we say that h is *κ-$\mathscr{L}$-semimonotone* if

$$\sum_{i \in [\![N]\!]} \big(h^i(x) - h^i(x^{-i}, y^i) - h^i(y^{-i}, x^i) + h^i(y)\big) \geqslant -\kappa|x-y|^2 \tag{4.2.5}$$

for all $x, y \in (\mathbb{R}^d)^N$, and κ is called the *$\mathscr{L}$-semimonotonicity constant*. We simply say that h is *$\mathscr{L}$-semimonotone* if (4.2.5) holds for some $\kappa \geqslant 0$. If (4.2.5) holds with $\kappa = 0$, we also say that f is *$\mathscr{L}$-monotone*.

For the sake of brevity, we define

$$\mathcal{L}[h](x,y) := \sum_{i \in [\![N]\!]} \big(h^i(x) - h^i(x^{-i}, y^i) - h^i(y^{-i}, x^i) + h^i(y) \big), \tag{4.2.6}$$

so that inequality (4.2.5) will appear as $\mathcal{L}[h] \geqslant -\kappa Q$ on $(\mathbb{R}^d)^{2N}$.

REMARK 4.2.3 (*Second-order characterisations of semimonotonicity, and comparison)*. If h is of class C^2, then exploiting the mean value theorem and the fundamental theorem of calculus one easily shows that

$$h \text{ is } M\text{-}\mathscr{D}\text{-semimonotone (that is, } \mathcal{D}[h] \geqslant -MQ \text{ on } (\mathbb{R}^d)^{2N}) \iff (D^2_{ij} h^i)_{i,j \in [\![N]\!]} \geqslant -M\mathsf{I}\,.$$

Analogously, applying the mean value theorem twice to (4.2.5),

$$\mathcal{L}[h](x,y) = \int_0^1 \int_0^1 \sum_{\substack{i,j \in [\![N]\!] \\ j \neq i}} D^2_{ij} h^i (z^{-i}_{s'}, z^i_s)(x^j - y^j) \cdot (x^i - y^i) \,\mathrm{d}s \,\mathrm{d}s'\,,$$

where we have set $z_s := sx + (1-s)y$; thus, one easily shows that

$$h \text{ is } \kappa\text{-}\mathscr{L}\text{-semimonotone (that is, } \mathcal{L}[h] \geqslant -\kappa Q \text{ on } (\mathbb{R}^d)^{2N}) \iff (D^2_{ij} h^i)_{i,j \in [\![N]\!]} - \operatorname{diag}(D^2_{ii} h^i)_{i \in [\![N]\!]} \geqslant -\kappa\mathsf{I}\,.$$

At the level of second-order derivatives, one then observes that the two notions of semimonotonicity differ by the diagonal term

$$\triangle_h := \operatorname{diag}(D^2_{ii} h^i)_{i \in [\![N]\!]}. \tag{4.2.7}$$

Therefore, under some unilateral control of such diagonal, they can be somehow compared: if $D^2_{ii} h^i \geqslant -\gamma I_d$ on $\mathbb{R}^d$, for some constant $\gamma > 0$ and all i, $\mathscr{L}$-semimonotonicity implies $\mathscr{D}$-semimonotonicity

$$\begin{cases} \mathcal{L}[h] \geqslant -\kappa Q & \text{on } (\mathbb{R}^d)^{2N} \\ \triangle_h \geqslant -\gamma \mathsf{I} & \text{on } (\mathbb{R}^d)^N \end{cases} \implies \mathcal{D}[h] \geqslant -(\kappa + \gamma) Q \ \text{ on } (\mathbb{R}^d)^{2N}.$$

On the other hand, if $D^2_{ii} h^i \leqslant \gamma I_d$ on $(\mathbb{R}^d)^N$, $\mathscr{D}$-semimonotonicity implies $\mathscr{L}$-semimonotonicity

$$\begin{cases} \mathcal{D}[h] \geqslant -MQ & \text{on } (\mathbb{R}^d)^{2N} \\ \triangle_h \leqslant \gamma \mathsf{I} & \text{on } (\mathbb{R}^d)^N \end{cases} \implies \mathcal{L}[h] \geqslant -(M + \gamma) Q \ \text{ on } (\mathbb{R}^d)^{2N}.$$

Nevertheless, for a given h, it is clear that the two constants involved in the definitions may differ substantially, hence one should not regard the two notions as comparable at all, especially when κ or M are small, that will be likely in our analysis. For instance, in the limit case that h be $\mathscr{D}$*-monotone*, it will be false that h is also $\mathscr{L}$*-monotone*, and vice versa.

REMARK 4.2.4. Having a look at the second-order characterisation above, one notes that the notion of $\mathscr{L}$-monotonicity is rather strict, in the sense that only functions with a very particular structure can be $\mathscr{L}$-monotone. In fact, for h to be $\mathscr{L}$-monotone one needs the *zero-diagonal* matrix $(D^2_{ij}h^i)_{i,j\in[\![N]\!]} - \triangle_h$ to be positive semidefinite (in the generalised sense that its symmetrisation is positive semidefinite), which happens only if such a matrix is the null one. Hence this forces each h^i to have the form $h^i(x) = h^i_0(x^i) + h^i_1(x^{-i})$, for some $h^i_0 \colon \mathbb{R}^d \to \mathbb{R}$ and $h^i_1 \colon (\mathbb{R}^d)^{N-1} \to \mathbb{R}$.

REMARK 4.2.5 (*Relationship with monotonicity in the Mean Field setting*). Consider the case $f^i(x) = F(x^i, m_x)$ for all $i \in [\![N]\!]$, where $F : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ is smooth enough for the following computations to be admissible.

- Suppose that F is *displacement monotone*; that is,

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \big(D_xF(x,m_1) - D_xF(y,m_2)\big)\cdot(x-y)\,\mu(\mathrm{d}x,\mathrm{d}y) \geqslant 0 \tag{4.2.8}$$

for every $m_1, m_2 \in \mathcal{P}(\mathbb{R}^d)$ and $\mu \in \mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d)$ having m_1 and m_2 as first and second marginal, respectively. We show that f is “almost” $\mathscr{D}$-monotone. Indeed, for every i,

$$D_if^i(x) = D_xF(x^i,m_x) + \frac{1}{N}\,D_mF(x^i,m_x,x^i),$$

hence

$$\begin{aligned}
&\sum_{i\in[\![N]\!]} (D_if^i(x) - D_if^i(y))\cdot(x^i-y^i)\\
&\qquad = N\int_{(\mathbb{R}^d)^2} (D_xF(x,m_x) - D_xF(y,m_y))\cdot(x-y)\,\frac{1}{N}\sum_{i\in[\![N]\!]}\delta_{(x^i,y^i)}(\mathrm{d}x,\mathrm{d}y)\\
&\qquad\qquad + \frac{1}{N}\sum_{i\in[\![N]\!]}(D_mF(x^i,m_x,x^i) - D_mF(y^i,m_y,y^i))\cdot(x^i-y^i).
\end{aligned}$$

The former term on the right-hand side is nonnegative by the displacement monotonicity of F. The latter one can be estimated in absolute value, assuming that D_mF is globally L-Lipschitz with respect to the $(|\cdot|, W_1, |\cdot|)$ distance, by

$$\frac{L}{N}\,W_1(m_x,m_y)\sum_{i\in[\![N]\!]}|x^i-y^i| + \frac{2L}{N}\sum_{i\in[\![N]\!]}|x^i-y^i|^2 \leqslant \frac{3L}{N}\sum_{i\in[\![N]\!]}|x^i-y^i|^2.$$

Therefore, f is $\frac{3L}{N}$-$\mathscr{D}$-semimonotone. More generally, if the right-hand side of (4.2.8) is not zero, but $-M\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^2\,\mu(\mathrm{d}x,\mathrm{d}y)$, then F is just displacement semimonotone, as in [47, Definition 2.7]. In such case, f becomes $\big(\frac{3L}{N}+M\big)$-$\mathscr{D}$-semimonotone.

• Suppose now that F is *monotone* in the Lasry–Lions sense; that is, for all $m_1, m_2 \in \mathcal{P}(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \big(F(z,m_1) - F(z,m_2)\big)(m_1 - m_2)(\mathrm{d}z) \geqslant 0. \tag{4.2.9}$$

Then,

$$\begin{aligned}\sum_{i\in[\![N]\!]} \big(f^i(x) - f^i(x^{-i},y^i) - f^i(y^{-i},x^i) + f^i(y)\big) \\ = N\int_{\mathbb{R}^d} \big(F(z,m_x) - F(z,m_y)\big)(m_x - m_y)(\mathrm{d}z) \\ + \sum_{i\in[\![N]\!]} \Big(F(y^i,m_x) - F\Big(y^i, m_x + \frac{\delta_{y^i} - \delta_{x^i}}{N}\Big) \\ - F\Big(x^i, m_y + \frac{\delta_{x^i} - \delta_{y^i}}{N}\Big) + F(x^i,m_y)\Big).\end{aligned}$$

While the former term on the right-hand side is nonnegative by the monotonicity of F, the latter one can be written as

$$\begin{aligned}\frac{1}{N}\sum_{i\in[\![N]\!]}\int_0^1 \Big(\frac{\delta F}{\delta m}(y^i,m_s,\cdot)\Big|_{y^i}^{x^i} - \frac{\delta F}{\delta m}(x^i,m_s,\cdot)\Big|_{y^i}^{x^i}\Big)\,\mathrm{d}s \\ + \frac{1}{N}\sum_{i\in[\![N]\!]}\int_0^1 \Big(\frac{\delta F}{\delta m}(x^i,m_s,\cdot)\Big|_{y^i}^{x^i} - \frac{\delta F}{\delta m}(x^i,\hat m_s,\cdot)\Big|_{y^i}^{x^i}\Big)\,\mathrm{d}s\end{aligned}$$

where $\frac{\delta}{\delta m}$ denotes the flat derivative (see, e.g., [**23**, Definition 2.2.1]) and we set $m_s = m_x + \frac{1-s}{N}(\delta_{y^i} - \delta_{x^i})$ and $\hat m_s = m_y + \frac{1-s}{N}(\delta_{x^i} - \delta_{y^i})$. Using repeatedly the mean value theorem, these two sums can be bounded, respectively, by

$$\frac{\|D_x D_m F\|_\infty}{N}\sum_{i\in[\![N]\!]} |x^i - y^i|^2$$

and

$$\begin{aligned}\left|\frac{1}{N}\sum_{i\in[\![N]\!]}\Big(\int_0^1\int_0^1\int_0^1\int_{\mathbb{R}^d} \frac{\delta(D_m F)}{\delta m}(*)\,(m_s - \hat m_s)(\mathrm{d}z)\,\mathrm{d}s''\mathrm{d}s'\mathrm{d}s\Big)\cdot(x^i - y^i)\right| \\ \leqslant \|D^2_{mm}F\|_\infty\bigg(\Big(\frac{1}{N}\sum_{i\in[\![N]\!]}|x^i - y^i|\Big)^2 + \frac{2}{N^2}\sum_{i\in[\![N]\!]}|x^i - y^i|^2\bigg),\end{aligned}$$

where $* = (x^i, s'm_s + (1-s')\hat m_s, s''x^i + (1-s'')y^i, z)$. Therefore, f is κ-$\mathscr{L}$-semimonotone with $\kappa = O(N^{-1})$. More generally, if the right-hand side of (4.2.9) is not zero, but $-MW_2(m_1,m_2)^2$, then f becomes $\big(M + O(N^{-1})\big)$-$\mathscr{L}$-semimonotone.

4.2.2. *Mean-Field-like* and *semimonotone* interactions. As prefigured, we work in a regime of *Mean-Field-like interactions*, assuming the following:

(MF) for $h \in \{f, g\}$, there exists $L_h > 0$, independent of N, such that

$$\sup_{i \in [\![N]\!]} \left(\|D_i h^i\|_\infty + \|D^2_{ii} h^i\|_\infty \right) \leqslant L_h,$$

$$\sup_{i \in [\![N]\!]} \Big(\sup_{k \in [\![N]\!] \setminus \{i\}} \|D_k h^i\|_\infty + \sum_{k \in [\![N]\!] \setminus \{i\}} \|D^2_{ik} h^k\|^2_\infty \Big) \leqslant \frac{L_h^2}{N},$$

and in addition

$$\sup_{k \in [\![N]\!]} \sum_{i \in [\![N]\!]} \|D(D_k g^i)\|^2_\infty + \sum_{\substack{k,i \in [\![N]\!] \\ i \neq k}} \|D(D^2_{ki} g^i)\|^2_\infty \leqslant L_g^2,$$

$$\sup_{i \in [\![N]\!]} \sum_{k \in [\![N]\!] \setminus \{i\}} \|D(D_k g^i)\|^2_\infty \leqslant \frac{L_g^2}{N}.$$

We stress the fact that these bounds are of Mean Field type in that they provide the weight with respect to N which one expects in the Mean Field setting, as described in the remark below. Yet, we are not asking for any symmetry with respect to the variables of h; that is, we are essentially dropping the classic indistinguishability (or exchangeability) hypothesis of Mean Field Games.

Remark 4.2.6. Consider $f^i(x) = F^i(x^i, m_{x^{-i}})$; recall that $m_{x^{-i}} = \frac{1}{N-1} \sum_{j \neq i} \delta_{x^j}$. Suppose that $F^i \colon \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}$ is Lipschitz continuous with Lipschitz constant $L > 0$. Since for any $v \in \mathbb{R}^d$ and $k \neq i$

$$\begin{aligned} &|f^i(x^0, \dots, x^{k-1}, x^k + v, x^{k+1}, \dots) - f^i(x)| \\ &\qquad \leqslant L\, W_1 \Big(\frac{1}{N-1} \sum_{j \notin \{i,k\}} \delta_{x^j} + \frac{1}{N-1} \delta_{x^k + v}, \frac{1}{N-1} \sum_{j \neq i} \delta_{x^j} \Big) \leqslant \frac{L}{N+1} |v| , \end{aligned}$$

one easily checks that the assumptions on f^i in **(MF)** hold, independently of N, provided that

$$F^i \text{ and } D_x F^i \text{ are Lipschitz continuous w.r.t. the } (|\cdot|, W_1) \text{ distance,} \tag{4.2.10}$$

since in such case $\|D_k F^i\|_\infty$ and $\|D^2_{ki} F^i\|_\infty$ remain bounded proportionally to $1/N$ uniformly with respect to $k \neq i$ and N.

On the other hand, the assumptions on g^i in **(MF)** are a bit more restrictive. This is because several estimates below exploit the regularisation effect of the diffusion, hence bounds on $D^\alpha u^i$ will depend on $D^{\alpha-1} f^i$ and on $D^\alpha g^i$. If $g^i(x) = G^i(x^i, \frac{1}{N-1} \sum_{j \neq i} \delta_{x^j})$, where $G^i \in C^1(\mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d))$, then, whenever $j \neq i$,

$$D_j g^i(x) = \frac{1}{N-1} D_m G^i \Big(x^i, \frac{1}{N-1} \sum_{j \neq i} \delta_{x^j}, x^j \Big),$$

see [**23**, Remark 6.1.2]. Therefore, for **(MF)** to hold one should require (4.2.10) with G in lieu of F and that

$$D_m G^i \text{ and } D_x D_m G^i \text{ are Lipschitz continuous w.r.t. the } (|\cdot|, W_1, |\cdot|) \text{ distance.}$$

The a priori estimates are derived on the N-dimensional Nash system under one of the following two additional structural assumptions of *semimonotonicity* of f and g: for $h \in \{f, g\}$,

(DS) h is M_h-$\mathscr{D}$-semimonotone, for some constant $M_h > 0$ independent of N;

(LS) h is κ_h-$\mathscr{L}$-semimonotone, for some constant $\kappa_h \geqslant 0$ independent of N;

Our main result has been presented in the Introduction. We reproduce it here for the convenience of the reader.

THEOREM 4.2.7. *Assume Mean-Field-like and semimonotone interactions; that is, assume* ***(MF)****, and either* ***(DS)*** *or* ***(LS)****. If, given $T > 0$, the semimonotonicity constants are sufficiently small (or vice versa), and $N \in \mathbb{N}$ is sufficiently large, any solution u to the Nash system* (4.2.1) *on $[0, T] \times (\mathbb{R}^d)^N$ satisfies*

$$\sup_{i \in [\![N]\!]} \left(\sup_{j \in [\![N]\!] \setminus \{i\}} \|D_j u^i\|_\infty + \left\| \sum_{j \in [\![N]\!] \setminus \{i\}} \left|D^2_{ij} u^j\right|^2 \right\|_\infty + \sum_{j \in [\![N]\!] \setminus \{i\}} \|D(D_j u^i)\|^2_\infty \right) \lesssim \frac{1}{N}$$

and

$$\sup_{i \in [\![N]\!]} \|Du^i\|_\infty + \sup_{i \in [\![N]\!]} \sum_{j \in [\![N]\!]} \|D(D_i u^j)\|^2_\infty + \sum_{\substack{i,j \in [\![N]\!] \\ j \neq i}} \|D(D_{ij} u^j)\|^2_\infty \lesssim 1,$$

where the implied constants are independent of N. In addition, u shares the same type of semimonotonicity of the data.

As discussed in the introduction, we proceed as follows. First, if the solution u to the Nash system is semimonotone, then the desired estimates on the derivatives hold; then, the semimonotonicity of the data provide semimonotonicity u near the final time T in such a way that, as a consequence of the aforementioned estimates, the semimonotonicity of u "propagates" up to time 0, provided that N is large enough.

This is done for $\mathscr{D}$-semimonotone data first. Then, making use of the relationship highlighted in Remark 4.2.3 between the two notions of semimonotonicity, with little additional effort the case of $\mathscr{L}$-semimonotone data is covered as well.

4.3. Estimates under the one-sided Lipschitz condition on the drift

For the following computations, we assume that u is of class C^1 in time and of class C^4 in space. Throughout the section, we will always assume a one-sided Lipschitz condition on the first-order (drift) term, as in (4.3.2) below. Such term will be often the one appearing in (4.2.1), that is $b^i = D_i u^i$, in which case the one-sided condition is equivalent to the $\mathscr{D}$-semimonotonicity of u.

4.3.1. A Lipschitz estimate for linear equations. The first ingredient is a Lipschitz estimate for solutions of linear equations with semimonotone drifts. Here, we implement the method of doubling variables (also known as Ishii–Lions method, or coupling method in the probabilistic community) as in [75]. The crucial features of the estimate are that the Lipschitz constant is universal with respect to N, and its dependence on the Lipschitz norm of the final condition, the supremum norm of the right-hand side and t is explicit. We note that the technique is well-known, but we are aware of very few examples of statements that are dimensional independent (e.g., [2]). Having a nondegenerate diffusion is crucial here, while the size of β does not play any role.

LEMMA 4.3.1. *Let $v\colon [0,T]\times(\mathbb{R}^d)^N\to\mathbb{R}$ be a bounded classical solution to*

$$\begin{cases} -\partial_t v - \operatorname{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2v) + \displaystyle\sum_{1\leqslant j\leqslant N} b^j\cdot D_j v = F \\ v|_{t=T} = G, \end{cases} \tag{4.3.1}$$

where F and G are continuous and bounded, G having bounded derivative, $b\colon[0,T]\times(\mathbb{R}^d)^N\to(\mathbb{R}^d)^N$ is continuous and bounded and for some $\tau\in[0,T]$ and $M\geqslant 0$,

$$\sum_{i\in[\![N]\!]}(b^i(t,x)-b^i(t,y))\cdot(x^i-y^i)\geqslant -M|x-y|^2\quad \forall t\in(\tau,T],\ x,y\in(\mathbb{R}^d)^N. \tag{4.3.2}$$

Then

$$\|Dv(t,\cdot)\|_\infty\leqslant\big(4\sigma^{-\frac12}\|F\|_\infty(T-t)^{\frac12}e^{M(T-t)}+\|DG\|_\infty\big)e^{M(T-t)}\quad\forall\, t\in[\tau,T].$$

Note that the rate $\|F\|_\infty\sqrt{(T-t)/\sigma}$ is sharp and coincides with the one of the heat equation $-(\partial_t+\sigma\Delta)v=F$, that can be obtained, for instance, using Duhamel's formula.

PROOF. Let $w\colon[0,T]\times(\mathbb{R}^d)^N\times(\mathbb{R}^d)^N\to\mathbb{R}$ be given by

$$w(t,x,y):=v(t,\cdot)|_y^x-\psi(t,|x-y|)-\|DG\|_\infty e^{M(T-t)}|x-y|\,,$$

where

$$\psi(t,r):=2\sigma^{-\frac12}\|F\|_\infty e^{2M(T-t)}\big(r\sqrt{T-t}+\sigma^{\frac12}(T-t)(1-e^{-\frac{r}{\sqrt{\sigma(T-t)}}})\big),$$

Computations show that for $x\neq y$

$$\begin{aligned} &D^2_{(x,y)}w(t,x,y)\\ &\quad=\operatorname{diag}\big(D^2v(t,x),-D^2v(t,y)\big)-\psi_{rr}(t,|x-y|)\begin{pmatrix}\mathsf{P}_{x-y} & -\mathsf{P}_{x-y}\\ -\mathsf{P}_{x-y} & \mathsf{P}_{x-y}\end{pmatrix}\\ &\qquad-\frac{1}{|x-y|}\big(\psi_r(t,|x-y|)+\|DG\|_\infty e^{M(T-t)}\big)\begin{pmatrix}\mathsf{I}-\mathsf{P}_{x-y} & -\mathsf{I}+\mathsf{P}_{x-y}\\ -\mathsf{I}+\mathsf{P}_{x-y} & \mathsf{I}-\mathsf{P}_{x-y}\end{pmatrix}, \end{aligned}$$

where $\mathsf{P}_e := |e|^{-2} e \otimes e$ for $e \in (\mathbb{R}^d)^N \setminus \{0\}$. We now want to show that w is a subsolution of the parabolic operator $-\partial_t - L$ on $(0,T) \times \big((\mathbb{R}^d)^N \times (\mathbb{R}^d)^N \setminus \{x = y\}\big)$, where

$$L = \operatorname{Tr}\big((J_2 \otimes (\sigma \mathsf{I} + \beta \mathsf{J})) D^2_{(x,y)}\big) - 4\sigma \operatorname{Tr}\big(\mathsf{P}_{x-y} D^2_{xy}\big) - b(t,x) \cdot D_x - b(t,y) \cdot D_y.$$

Note that the operator is indeed (degenerate) parabolic, since its diffusion matrix

$$\begin{pmatrix} \sigma \mathsf{I} + \beta \mathsf{J} & \sigma \mathsf{I} + \beta \mathsf{J} - 2\sigma \mathsf{P}_{x-y} \\ \sigma \mathsf{I} + \beta \mathsf{J} - 2\sigma \mathsf{P}_{x-y} & \sigma \mathsf{I} + \beta \mathsf{J} \end{pmatrix}$$

is nonnegative. We have

$$\operatorname{Tr}\big((J_2 \otimes (\sigma \mathsf{I} + \beta \mathsf{J})) D^2_{(x,y)} w(t,x,y)\big) = \operatorname{Tr}((\sigma \mathsf{I} + \beta \mathsf{J}) D^2 v(t,\cdot))\big|_y^x$$

and

$$-4\sigma \operatorname{Tr}\big(\mathsf{P}_{x-y} D^2_{xy} w(t,x,y)\big) = -4\sigma \psi_{rr}(t, |x-y|),$$

where we used the notation $\psi_r = \frac{\partial}{\partial r}\psi$. Also, we compute

$$D_x w(t,x,y) = Dv(t,x) - \frac{1}{|x-y|}\big(\psi_r(t,|x-y|) + \|DG\|_\infty e^{M(T-t)}\big)(x-y),$$
$$D_y w(t,x,y) = -Dv(t,y) + \frac{1}{|x-y|}\big(\psi_r(t,|x-y|) + \|DG\|_\infty e^{M(T-t)}\big)(x-y)$$

and, according to (4.2.1),

$$\begin{aligned} \partial_t w(t,x,y) = & -\operatorname{Tr}\big((\sigma \mathsf{I} + \beta \mathsf{J}) D^2 v(t,\cdot)\big|_y^x\big) + \sum_{1 \leqslant j \leqslant N} b^j(t,\cdot) D_j v(t,\cdot)\big|_y^x - F\big|_y^x \\ & - \psi_t(t,|x-y|) + M\|DG\|_\infty e^{M(T-t)}|x-y|, \end{aligned}$$

where $\psi_t = \frac{\partial}{\partial t}\psi$. It follows that

$$\begin{aligned} &(\partial_t + L) w(t,x,y) \\ &\quad = -4\sigma\psi_{rr}(t,|x-y|) - F\big|_y^x - \psi_t(t,|x-y|) + M\|DG\|_\infty e^{M(T-t)}|x-y| \\ &\qquad + \frac{1}{|x-y|}\big(\psi_r(t,|x-y|)(|x-y|) \\ &\qquad\qquad + \|DG\|_\infty e^{M(T-t)}\big) \sum_{1 \leqslant j \leqslant N} b^j(t,\cdot)\big|_y^x \cdot (x^j - y^j)\,; \end{aligned}$$

hence by the hypothesis (4.3.2) on b,

$$\begin{aligned} &(\partial_t + L) w(t,x,y) \\ &\quad \geqslant -4\sigma\psi_{rr}(t,|x-y|) - M\psi_r(t,|x-y|)|x-y| - \psi_t(t,|x-y|) - 2\|F\|_\infty\,. \end{aligned}$$

Now, direct computations show that, with $r = |x-y|$, $s = T-t$ and $y = r/\sqrt{T-t}$,

$$\begin{aligned} &-4\sigma\psi_{rr}(t,r) - M\psi_r(t,r)r - \psi_t(t,r) \\ &\quad = 2\sigma^{-\frac{1}{2}}\|F\|_\infty e^{2Ms}\Big(\sigma^{\frac{1}{2}} + 3\sigma^{\frac{1}{2}} e^{-y\sigma^{-\frac{1}{2}}} + 2Ms\sigma^{\frac{1}{2}}(1 - e^{-y\sigma^{-\frac{1}{2}}}) \\ &\qquad\qquad + \frac{y}{2}(1 + 2Ms)(1 - e^{-y\sigma^{-\frac{1}{2}}})\Big) \\ &\quad > 2\|F\|_\infty. \end{aligned}$$

Since $w(t,\cdot)|_{x=y} = 0$ for all t, by the maximum principle[1]

$$w \leqslant \Big(\max_{(\mathbb{R}^d)^N \times (\mathbb{R}^d)^N} w(T,\cdot) \Big)_+ = 0 \quad \text{on } [\tau, T] \times (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N.$$

Being the choice of t, x and y arbitrary, the desired conclusion follows by observing that $\psi(t,r) \leqslant 4\sigma^{-\frac{1}{2}} \|F\|_\infty e^{2M(T-t)} r\sqrt{T-t}$. ■

REMARK 4.3.2. The previous estimate can be proven also for viscosity solutions, using standard methods. Unbounded solutions (with controlled growth) can be also considered.

The following estimates will make use of Lemma 4.3.1 with $b^j = D_j u^j$. They will be all derived under the assumption that u be $\mathscr{D}$-semimonotone on $(\mathbb{R}^d)^{2N}$ uniformly with respect to $t \in (\tau, T]$; that is, according to the introduced notation, $\mathcal{D}[u] \geqslant -MQ$ on $(\tau, T] \times (\mathbb{R}^d)^{2N}$.

4.3.2. First-order derivatives of the value functions. The next result is a Lipschitz estimate for the value function, it is obtained by doubling variables. We show that the Hamilton–Jacobi equations preserve some weighted Lipschitz seminorm on $(\mathbb{R}^d)^N$. Note that, contrarily to the previous lemma, possible regularisation effects from the diffusion are not exploited below.

Recall that, for any $x \in (\mathbb{R}^d)^N$, we have the *i-th weighted norm*

$$\|x\|^i := \Big(|x^i|^2 + \frac{1}{N} \sum_{j \in [\![N]\!] \setminus \{i\}} |x^j|^2 \Big)^{\frac{1}{2}}.$$

LEMMA 4.3.3 (Weighted Lipschitz continuity of the value functions). *Let $\tau \in [0,T)$. Suppose that $\mathcal{D}[u] \geqslant -MQ$ on $(\tau, T] \times (\mathbb{R}^d)^{2N}$ for some $M > 0$ and that, for $h \in \{f, g\}$, there exist $\widetilde{L}_h > 0$ such that*

$$|h^i(x) - h^i(y)| \leqslant \widetilde{L}_h \|x-y\|^i \qquad \forall\, x, y \in (\mathbb{R}^d)^N,\ i \in [\![N]\!]. \tag{4.3.3}$$

Then

$$|u^i(t,x) - u^i(t,y)| \leqslant c_1 \|x-y\|^i \qquad \forall\, t \in [\tau, T],\ x, y \in (\mathbb{R}^d)^N, \tag{4.3.4}$$

where the constant c_1 depends only on T, M, $\widetilde{L}_g$ and $\widetilde{L}_f$.

---

[1] Here one needs to apply the maximum principle for classical subsolutions of parabolic equations on unbounded domains. The proof of this is rather classical, and it is based on the fact that interior maxima of w cannot exist. Since w may not have a maximum on the unbounded set $(\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$, it should be additively perturbed for instance by $-\varepsilon(e^{K(T-t)}(1 + |x|^2 + |y|^2) + (T-\tau)^{-1})$, with K large enough. Being now w bounded (as well as b), then the conclusion follows by letting $\varepsilon \to 0$. We do not provide here further details, but we mention that identical computations appear in the proof of the next Lemma 4.3.3. Alternatively, see Section 4.6.3 for a more probability-oriented proof.

PROOF. For $\varepsilon > 0$, let w^i be the function

$$w^i(t,x,y) := u^i(t,\cdot)|_y^x - \psi(t)\|x-y\|^i - \varepsilon\varphi(t,x,y) - \frac{\varepsilon}{t-\tau},$$

defined for $t \in (\tau, T]$, $x, y \in (\mathbb{R}^d)^N$, with

$$\psi(t) := e^{M(T-t)}\widetilde{L}_g + \frac{e^{M(T-t)}-1}{M}\widetilde{L}_f,$$
$$\varphi(t,x,y) = e^{K(T-t)}(1+|x|^2+|y|^2),$$

for some constant $K \geqslant 0$ to be determined. Since u^i is bounded, w^i attains its maximum at some point $(\bar{t},\bar{x},\bar{y}) \in (\tau,T] \times (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$. Suppose that $\bar{t} \neq T$ and $\bar{x} \neq \bar{y}$. Computations show that for $x \neq y$ (for simplicity, one may follow the computations below with $\varepsilon = 0$, being the corresponding terms merely necessary to guarantee the existence of a maximum for w, and therefore of perturbative nature)

$$D^2_{(x,y)}w^i(t,x,y) = \begin{pmatrix} D^2u^i(t,x) - 2\varepsilon e^{K(T-t)}\mathsf{I} & 0 \\ 0 & -D^2u^i(t,y) - 2\varepsilon e^{K(T-t)}\mathsf{I} \end{pmatrix}$$
$$- \frac{\psi(t)}{\|x-y\|^i}\begin{pmatrix} \mathsf{I}^i - \mathsf{P}^i_{x-y} & -\mathsf{I}^i + \mathsf{P}^i_{x-y} \\ -\mathsf{I}^i + \mathsf{P}^i_{x-y} & \mathsf{I}^i - \mathsf{P}^i_{x-y} \end{pmatrix},$$

where I^i is the block diagonal matrix given by

$$(\mathsf{I}^i)_{jj} = \begin{cases} I_d & \text{if } j = i \\ N^{-1}I_d & \text{if } j \neq i \end{cases}$$

and we have set, for $e \in (\mathbb{R}^d)^N \setminus \{0\}$,

$$\mathsf{P}^i_e := \frac{N_i e}{\|e\|^i} \otimes \frac{N_i e}{\|e\|^i}, \qquad (N_i e)^j = \begin{cases} e^i & \text{if } j = i \\ N^{-1}e^j & \text{if } j \neq i. \end{cases}$$

Then, as $J_2 \otimes (\sigma\mathsf{I} + \beta\mathsf{J}) \geqslant 0$ and $\mathrm{Tr}(D^2w^i(\bar{t},\bar{x},\bar{y})) \leqslant 0$, we have

$$\begin{aligned} 0 &\geqslant \mathrm{Tr}\big((J_2 \otimes (\sigma\mathsf{I} + \beta\mathsf{J}))D^2_{(x,y)}w^i(\bar{t},\bar{x},\bar{y})\big) \\ &= \mathrm{Tr}((\sigma\mathsf{I} + \beta\mathsf{J})D^2u^i(\bar{t},\cdot))\big|_{\bar{y}}^{\bar{x}} - 4(\sigma+\beta)Nd\varepsilon e^{K(T-t)}. \end{aligned} \tag{4.3.5}$$

Also, since $(D_x, D_y, \partial_t)w^i(\bar{t},\bar{x},\bar{y}) = 0$, we have

$$0 = Du^i(\bar{t},\bar{x}) - \frac{\psi(\bar{t})}{\|\bar{x}-\bar{y}\|^i}N_i(\bar{x}-\bar{y}) - 2\varepsilon e^{K(T-t)}\bar{x}, \tag{4.3.6}$$

$$0 = -Du^i(\bar{t},\bar{y}) + \frac{\psi(\bar{t})}{\|\bar{x}-\bar{y}\|^i}N_i(\bar{x}-\bar{y}) - 2\varepsilon e^{K(T-t)}\bar{y} \tag{4.3.7}$$

and, using (4.2.1) when computing $\partial_t w^i$,

$$
\begin{aligned}
(4.3.8) \qquad 0 = &- \operatorname{Tr}\big((\sigma \mathsf{I} + \beta \mathsf{J}) D^2 u^i(\bar t, \cdot)\big|_{\bar y}^{\bar x}\big) + \sum_{j \in [\![N]\!] \setminus \{i\}} D_j u^j(\bar t, \cdot) \cdot D_j u^i(\bar t, \cdot)\big|_{\bar y}^{\bar x} \\
&+ \frac{1}{2} |D_i u^i(\bar t, \cdot)|^2 \big|_{\bar y}^{\bar x} - f^i\big|_{\bar y}^{\bar x} - \psi'(\bar t) \|\bar x - \bar y\|^i \\
&+ \varepsilon K e^{K(T-t)} (1 + |\bar x|^2 + |\bar y|^2) + \frac{\varepsilon}{(\bar t - \tau)^2} .
\end{aligned}
$$

Exploiting relations (4.3.6) and (4.3.7) one sees that

$$
\begin{aligned}
(4.3.9) \qquad &\sum_{j \in [\![N]\!] \setminus \{i\}} D_j u^j(\bar t, \cdot) \cdot D_j u^i(\bar t, \cdot)\big|_{\bar y}^{\bar x} + \frac{1}{2} |D_i u^i(\bar t, \cdot)|^2 \big|_{\bar y}^{\bar x} \\
&= \frac{1}{N} \frac{\psi(\bar t)}{\|\bar x - \bar y\|^i} \sum_{j \in [\![N]\!] \setminus \{i\}} D_j u^j(\bar t, \cdot)\big|_{\bar y}^{\bar x} \cdot (\bar x^j - \bar y^j) \\
&\quad + 4 \varepsilon e^{K(T-\bar t)} \frac{\psi(\bar t)}{\|\bar x - \bar y\|^i} (\bar x^i - \bar y^i) \cdot (\bar x^i + \bar y^i) + 4 \varepsilon^2 e^{2K(T-\bar t)} (|\bar x^i|^2 - |\bar y^i|^2) \\
&\quad + 2 \varepsilon e^{K(T-\bar t)} \sum_{j \in [\![N]\!] \setminus \{i\}} \big(D_j u^j(\bar t, \bar x) \cdot \bar x^j + D_j u^j(\bar t, \bar y) \cdot \bar y^i\big) .
\end{aligned}
$$

The terms on the right-hand side of identity (4.3.9) can be estimated as follows: since $\mathcal{D}[u](\bar t) \geqslant -MQ$ and $\sqrt{N}\, \|x - y\|^i \geqslant |x - y|$, using relations (4.3.6) and (4.3.7) along with the Cauchy–Schwarz inequality,

$$
\begin{aligned}
(4.3.10) \quad &\frac{1}{N} \frac{\psi(\bar t)}{\|\bar x - \bar y\|^i} \sum_{j \in [\![N]\!] \setminus \{i\}} D_j u^j(\bar t, \cdot)\big|_{\bar y}^{\bar x} \cdot (\bar x^j - \bar y^j) \\
&= \frac{1}{N} \frac{\psi(\bar t)}{\|\bar x - \bar y\|^i} \sum_{j \in [\![N]\!]} D_j u^j(\bar t, \cdot)\big|_{\bar y}^{\bar x} \cdot (\bar x^j - \bar y^j) \\
&\qquad\qquad - \frac{1}{N} \frac{\psi(\bar t)}{\|\bar x - \bar y\|^i} D_i u^i(\bar t, \cdot)\big|_{\bar y}^{\bar x} \cdot (\bar x^i - \bar y^i) \\
&\geqslant -M \psi(\bar t) \|\bar x - \bar y\|^i - \frac{1}{N} \frac{\psi(\bar t)}{\|\bar x - \bar y\|^i} 2 \varepsilon e^{K(T-\bar t)} (\bar x^i + \bar y^i) \cdot (\bar x^i - \bar y^i) \\
&\geqslant -M \psi(\bar t) \|\bar x - \bar y\|^i - \frac{2\varepsilon}{\sqrt{N}} \psi(\bar t) \varphi(\bar t, \bar x, \bar y);
\end{aligned}
$$

since $\|x - y\|^i \leqslant |x - y|$, by the Cauchy–Schwarz inequality

$$
\begin{aligned}
(4.3.11) \quad &4 \varepsilon e^{K(T-\bar t)} \frac{\psi(\bar t)}{\|\bar x - \bar y\|^i} (\bar x^i - \bar y^i) \cdot (\bar x^i + \bar y^i) + 4 \varepsilon^2 e^{2K(T-\bar t)} (|\bar x^i|^2 - |\bar y^i|^2) \\
&\geqslant -4 \varepsilon e^{K(T-\bar t)} \psi(\bar t) (|\bar x^i| + |\bar y^i|) - 4 \varepsilon^2 e^{2K(T-\bar t)} (|\bar x|^2 + |\bar y|^2) \\
&\geqslant -4 \varepsilon \Big(\psi(\bar t) + \varepsilon e^{K(T-\bar t)}\Big) \varphi(\bar t, \bar x, \bar y);
\end{aligned}
$$

finally, by the Cauchy–Schwarz and Young's inequalities,

$$
\begin{aligned}
(4.3.12) \quad &2 \varepsilon e^{K(T-\bar t)} \sum_{j \in [\![N]\!] \setminus \{i\}} \big(D_j u^j(\bar t, \bar x) \cdot \bar x^j + D_j u^j(\bar t, \bar y) \cdot \bar y^i\big) \\
&\qquad \geqslant -\varepsilon \Big(2N \sup_{j \in [\![N]\!]} \|D_j u^j\|_\infty^2 + 1\Big) \varphi(\bar t, \bar x, \bar y).
\end{aligned}
$$

Plugging (4.3.9) into (4.3.8) and exploiting inequalities (4.3.5), (4.3.10), (4.3.11) and (4.3.12) as well as (4.3.3), we get

$$0 < \big(M\psi(\bar t) + \widetilde{L}_f + \psi'(\bar t)\big)\|\bar x - \bar y\|^i + \varepsilon\big(4(\sigma+\beta)Nd + 4\psi(\bar t) + 2\varepsilon e^{K(T-\bar t)} + 2NC_N + 1 - K\big)\varphi(\bar t, \bar x, \bar y)\,, \tag{4.3.13}$$

where $C_N := \sup_{1\leqslant j\leqslant N}\|D_j u^j\|_\infty^2$. At this point, choose

$$K = 4(\sigma+\beta)Nd + 4\psi(0) + 2NC_N + 3\,;$$

thus for all $\varepsilon < e^{-KT}$ one has

$$4(\sigma+\beta)Nd + 4\psi(\bar t) + 2\varepsilon e^{K(T-\bar t)} + 2NC_N + 1 - K < 0\,.$$

On the other hand, by the definition of ψ we have $M\psi + \widetilde{L}_f + \psi' = 0$, thus the above inequality (4.3.13) cannot hold. We conclude that one must have $\bar t = T$ or $\bar x = \bar y$, for any sufficiently small ε. Letting $\varepsilon \to 0$, we deduce that

$$u^i(t,\cdot)|_{\bar y}^{\bar x} - \psi(t)\|\bar x - \bar y\|^i \leqslant (g^i|_{\bar y}^{\bar x} - \widetilde{L}_g\|\bar x - \bar y\|^i)_+ = 0\,,$$

where we have used (4.3.3) to obtain the last equality. The desired conclusion follows with $c_1 = \psi(0)$. ∎

Remark 4.3.4. The previous result can be obtained for solutions to Hamilton–Jacobi equations where the transport term $\sum_{j\in[\![N]\!]\setminus\{i\}} D_j u^j \cdot D_j u^i$ is replaced by $\sum_{j\in[\![N]\!]} b^j \cdot D_j u^i$, with b satisfying (4.3.2).

The previous proposition immediately provides an estimate for the *skew* first-order derivatives; that is, for those derivatives of the i-th value functions with respect to x^j with $j \neq i$.

Proposition 4.3.5 (Estimate on skew first-order derivatives). *Let $\tau \in [0,T)$. Suppose that $\mathcal{D}[u] \geqslant -MQ$ on $(\tau, T] \times (\mathbb{R}^d)^{2N}$. Then*

$$\sup_{i\in[\![N]\!]}\Big\|\sum_{j\in[\![N]\!]\setminus\{i\}} |D_j u^i(t,\cdot)|^2\Big\|_\infty \leqslant \frac{c_1^2}{N} \quad \forall\, t \in [\tau, T],$$

where c_1 is the constant given in Lemma 4.3.3 (that applies with $\widetilde{L}_f = \sqrt{2}L_f$ and $\widetilde{L}_g = \sqrt{2}L_g$, where L_f and L_g are the constants appearing in **(MF)***).*

Proof. Note first that by the mean value theorem and **(MF)** we have

$$|f^i(x) - f^i(y)| \leqslant \Big\|\Big(|D_i f^i|^2 + N\sum_{j\neq i}|D_j f^i|^2\Big)^{\frac12}\Big\|_\infty \|x-y\|^i \leqslant \sqrt{2}\,L_f\|x-y\|^i.$$

Let now $\xi \in (\mathbb{R}^d)^N$ be such that $|\xi| = 1$ and $\xi^i = 0$. Plug $y = x + h\xi$ into (4.3.4) and let $h \to 0$ to obtain $|Du^i(t,x)\cdot\xi| \leqslant c_1/\sqrt{N}$. Taking now the supremum over ξ yields $\big(\sum_{j\neq i}|D_j u^i(t,x)|^2\big)^{\frac12} \leqslant c_1/\sqrt{N}$. ∎

4.3.3. Second-order derivatives of the value functions. Although all constants contained in the following statements will depend also on the dimension d, we shall omit to specify it.

We now show that certain sums of second-order derivatives over the indices of the players are controlled uniformly with respect to N. This is a preliminary step to achieve more precise control on second-order derivatives.

PROPOSITION 4.3.6. *Let $\tau \in [0,T)$. Suppose that $\mathcal{D}[u] \geqslant -MQ$ on $(\tau,T] \times (\mathbb{R}^d)^{2N}$. Then*

$$\sup_{k\in[\![N]\!]} \sum_{i\in[\![N]\!]} \|D(D_k u^i)(t,\cdot)\|_\infty^2 \leqslant C_2 \quad \forall\, t\in[\tau,T]\,,$$

where the constant C_2 depends only on σ, T, M, L_g and L_f.

PROOF. Let $k \in [\![N]\!]$, $\ell \in [\![d]\!]$ and $s \in [0, T-\tau]$. Note that $v = v^{k\ell} = D_{x^{k\ell}} u^i$ solves (4.3.1) on $[T-s,T] \times (\mathbb{R}^d)^N$ with $b^j = D_j u^j$, $F = D_{x^{k\ell}} f^i - \sum_{j\in[\![N]\!]\setminus\{i\}} D_j(D_{x^{k\ell}} u^j) D_j u^i$ and $G = D_{x^{k\ell}} g^i$. By Lemma 4.3.1,

$$\begin{aligned}\|D_{x^{k\ell}} D u^i\|_\infty &\leqslant C_\sigma \|D_{x^{k\ell}} f^i\|_\infty s^{\frac12} e^{2Ms} + \|D(D_{x^{k\ell}} g^i)\|_\infty e^{Ms} \\ &\quad + C_\sigma s^{\frac12} e^{2Ms} \Big\| \sum_{j\in[\![N]\!]\setminus\{i\}} D_j(D_{x^{k\ell}} u^j) D_j u^i \Big\|_\infty,\end{aligned} \tag{4.3.14}$$

where the L^∞-norm are understood to be computed on $[T-s,T]\times(\mathbb{R}^d)^N$ or $(\mathbb{R}^d)^N$. By Proposition 4.3.5 and the Cauchy–Schwarz inequality we have

$$\Big\| \sum_{j\in[\![N]\!]\setminus\{i\}} D_j(D_{x^{k\ell}} u^j) D_j u^i \Big\|_\infty^2 \leqslant \frac{c_1^2}{N} \Big\| \sum_{j\in[\![N]\!]} \big|D^2_{jk} u^j\big|^2 \Big\|_\infty ;$$

therefore from (4.3.14) and Young's inequality we get

$$\begin{aligned}\|D(D_k u^i)\|_\infty^2 &\leqslant 3dC_\sigma^2 \|D_k f^i\|_\infty^2 e^{4Ms} s + 3d \|D(D_k g^i)\|_\infty^2 e^{2Ms} \\ &\quad + \frac{3dC_\sigma^2 c_1^2 e^{4Ms} s}{N} \Big\| \sum_{j\in[\![N]\!]} \big|D^2_{jk} u^j\big|^2 \Big\|_\infty .\end{aligned} \tag{4.3.15}$$

Since by **(MF)** we have $\sup_{1\leqslant k\leqslant N} \sum_{1\leqslant i\leqslant N} \|D_k f^i\|_\infty^2 \leqslant 2L_f^2$, summing (4.3.15) over i one deduces that

$$\begin{aligned}\sum_{i\in[\![N]\!]} \|D(D_k u^i)\|_\infty^2 &\leqslant 6dC_\sigma^2 e^{4Ms} s L_f^2 + 3de^{2Ms} \sum_{i\in[\![N]\!]} \|D_k D g^i\|_\infty^2 \\ &\quad + 3dC_\sigma^2 c_1^2 e^{4Ms} s \sum_{i\in[\![N]\!]} \|D(D_k u^i)\|_\infty^2 .\end{aligned}$$

If $3dC_\sigma^2 c_1^2 e^{4Ms} s \leqslant \frac12$ this yields

$$\sum_{i\in[\![N]\!]} \|D(D_k u^i)\|_\infty^2 \leqslant \frac{2L_f^2}{c_1^2} + 6de^{2Ms} \sum_{i\in[\![N]\!]} \|D(D_k g^i)\|_\infty^2 . \tag{4.3.16}$$

Considering that $g^i = u^i(T,\cdot)$, by iterating estimate (4.3.16) on the intervals $[T-\ell s, T-(l-1)s]$ with $u^i(T-(l-1)s,\cdot)$ in lieu of g^i for any positive integer $l \leqslant l^* := \lfloor (T-\tau)/s \rfloor$ and then on the interval $[\tau, T-l^*s]$, one can prove by induction that

$$\sum_{i\in[\![N]\!]} \|D(D_k u^i)\|_\infty^2$$
$$\leqslant \sum_{1\leqslant l\leqslant l^*+1} \bigg(\frac{2L_f^2}{c_1^2} \sum_{m\in[\![l]\!]} (6de^{2Ms})^m + (6de^{2Ms})^\ell \sum_{i\in[\![N]\!]} \|D(D_k g^i)\|_\infty^2 \bigg),$$

where the L^∞-norms are understood to be computed on $[\tau,T]\times(\mathbb{R}^d)^N$. We can estimate $sl^* \leqslant T$ and choose $s = (6dC_\sigma^2 c_1^2 e^{4MT})^{-1} \wedge (T-\tau)$, hence we conclude using that **(MF)** includes $\sup_{k\in[\![N]\!]} \sum_{i\in[\![N]\!]} \|D(D_k g^i)\|_\infty^2 \leqslant L_g^2$. ■

Exploiting the estimates deduced so far, Bernstein's method yields the first estimate on the skew second-order derivatives. It provides a control when letting one derivative be "in the direction" of the player to which the value function corresponds while the other derivative is skew and then summing the squares of the second-order derivatives over the players. In this sense, as the sum touches all value functions, we will refer to it as a *transversal* estimate.

PROPOSITION 4.3.7 (Transversal estimate on skew second-order derivatives). *Let $\tau \in [0,T)$. Suppose that $\mathcal{D}[u] \geqslant -MQ$ on $(\tau,T]\times(\mathbb{R}^d)^{2N}$. Then*

$$\sup_{i\in[\![N]\!]} \bigg\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D_{ij}^2 u^j(t,\cdot)\big|^2 \bigg\|_\infty \leqslant \frac{C_3}{N} \quad \forall\, t\in[\tau,T].$$

where the constant C_3 depends only on σ, T, M, L_g and L_f.

PROOF. Let

$$w^i := \frac{1}{2}\bigg(\sum_{\ell\in[\![N]\!]\setminus\{i\}} |D_{i\ell}^2 u^\ell|^2 + \frac{1}{N}\,|D_{ii}^2 u^i|^2 \bigg).$$

Direct computations exploiting (4.2.1) show that

$$\begin{aligned}
&-\partial_t w^i - \mathrm{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2 w^i) + \sum_{j\in[\![N]\!]} D_j u^j\cdot D_j w^i + \beta \sum_{\ell\in[\![N]\!]\setminus\{i\}} \bigg| \sum_{j\in[\![N]\!]} D_{ij\ell}^3 u^\ell \bigg|^2 \\
&+ \frac{\beta}{N}\bigg| \sum_{j\in[\![N]\!]} D_{iij}^3 u^i \bigg|^2 + \sigma \sum_{\substack{j,\ell\in[\![N]\!]\\ \ell\neq i}} \big|D_{ij\ell}^3 u^\ell\big|^2 + \frac{\sigma}{N} \sum_{j\in[\![N]\!]} \big|D_{iij}^3 u^i\big|^2 \\
&+ \sum_{\substack{j,\ell\in[\![N]\!]\\ \ell\neq i}} D_{j\ell}^2 u^\ell D_{ij}^2 u^j D_{i\ell}^2 u^\ell + \sum_{\substack{j,\ell\in[\![N]\!]\\ j\neq\ell\neq i}} \big(D_{ij}^2 u^\ell D_{j\ell}^2 u^j D_{i\ell}^2 u^\ell + D_j u^\ell D_{ij\ell}^3 u^j D_{i\ell}^2 u^\ell \big) \\
&+ \frac{1}{N} \sum_{j\in[\![N]\!]} D_{ij}^2 u^i D_{ij}^2 u^j D_{ii}^2 u^i + \frac{1}{N} \sum_{j\in[\![N]\!]\setminus\{i\}} \big(D_{ij}^2 u^i D_{ij}^2 u^j D_{ii}^2 u^i + D_j u^i D_{iij}^3 u^j D_{ii}^2 u^i \big) \\
&= \sum_{\ell\in[\![N]\!]\setminus\{i\}} D_{i\ell}^2 f^\ell D_{i\ell}^2 u^\ell + \frac{1}{N} D_{ii}^2 f^i D_{ii}^2 u^i,
\end{aligned}$$

where according to the fact that the we are considering the Frobenius inner product on $\mathbb{R}^d$, all products between tensors are to be understood in the sense of tensor contraction over the coordinates of repeated derivative indices; that is, for example,

$$D^2_{j\ell}u^\ell D^2_{ij}u^j D^2_{i\ell}u^\ell = \mathrm{Tr}(D^2_{\ell j}u^\ell D^2_{ji}u^j D^2_{i\ell}u^\ell)\,,$$
$$D_j u^\ell D^3_{ij\ell}u^j D^2_{i\ell}u^\ell = \sum_{k_1,k_2,k_3\in[\![d]\!]} D_{x^{jk_2}}u^\ell D^3_{x^{ik_1}x^{jk_2}x^{\ell k_3}}u^j D^2_{x^{ik_1}x^{\ell k_3}}u^\ell\,,$$
$$\Big|\sum_{j\in[\![N]\!]} D^3_{iij}u^i\Big|^2 = \sum_{k_1,k_2,k_3\in[\![d]\!]}\Big|\sum_{j\in[\![N]\!]} D^3_{x^{ik_1}x^{ik_2}x^{jk_3}}u^i\Big|^2\,.$$

We now estimate the terms above: since, according to Remark 4.2.3, $\mathcal{D}[u] \geqslant -MQ$ is equivalent to $(D^2_{j\ell}h^\ell)_{j,\ell\in[\![N]\!]} \geqslant -M\mathsf{I}$,

$$\begin{aligned}\sum_{\substack{j,\ell\in[\![N]\!]\\ \ell\neq i}} D^2_{j\ell}u^\ell D^2_{ij}u^j D^2_{i\ell}u^\ell &= \sum_{\substack{j,\ell\in[\![N]\!]\\ j,\ell\neq i}} D^2_{j\ell}u^\ell D^2_{ij}u^j D^2_{i\ell}u^\ell + \sum_{\ell\in[\![N]\!]\setminus\{i\}} D^2_{i\ell}u^\ell D^2_{ii}u^i D^2_{i\ell}u^\ell\\ &\geqslant -2M\sum_{\ell\in[\![N]\!]\setminus\{i\}}|D^2_{i\ell}u^\ell|^2 \geqslant -4Mw^i\,;\end{aligned}$$

by the Cauchy–Schwarz inequality and Proposition 4.3.6,

$$\begin{aligned}\sum_{\substack{j,\ell\in[\![N]\!]\\ j\neq\ell\neq i}}\big|D^2_{ij}u^\ell D^2_{j\ell}u^j D^2_{i\ell}u^\ell\big| &\leqslant \sup_{i\in[\![N]\!]}\sum_{\ell\in[\![N]\!]\setminus\{i\}}|D^2_{i\ell}u^\ell|^2\Bigg(\sum_{\substack{j,\ell\in[\![N]\!]\\ j\neq\ell\neq i}}|D^2_{ij}u^\ell|^2\Bigg)^{\frac12}\\ &\leqslant C_2^{\frac12}\sup_{i\in[\![N]\!]} w^i\,;\end{aligned}$$

by the Cauchy–Schwarz inequality, Proposition 4.3.5 and Young's inequality,

$$\begin{aligned}&\sum_{\substack{j,\ell\in[\![N]\!]\\ j\neq\ell\neq i}}\big|D_j u^\ell D^3_{ij\ell}u^j D^2_{i\ell}u^\ell\big|\\ &\quad\leqslant\Bigg(\sup_{\ell\in[\![N]\!]}\sum_{j\in[\![N]\!]\setminus\{i\}}|D_j u^\ell|^2\Bigg)^{\frac12}\Bigg(\sum_{\substack{j,\ell\in[\![N]\!]\\ j\neq\ell\neq i}}|D^3_{ij\ell}u^j|^2\Bigg)^{\frac12}\Bigg(\sum_{\ell\in[\![N]\!]\setminus\{i\}}|D^2_{i\ell}u^\ell|^2\Bigg)^{\frac12}\\ &\quad\leqslant\frac{c_1}{\sqrt N}\Bigg(\sum_{\substack{j,\ell\in[\![N]\!]\\ \ell\neq i}}|D^3_{ij\ell}u^\ell|^2+\sum_{j\in[\![N]\!]}|D^3_{iij}u^i|^2\Bigg)^{\frac12}\Bigg(\sum_{\ell\in[\![N]\!]\setminus\{i\}}|D^2_{i\ell}u^\ell|^2\Bigg)^{\frac12}\\ &\quad\leqslant\frac{\sigma}{2N}\sum_{\substack{j,\ell\in[\![N]\!]\\ \ell\neq i}}|D^3_{ij\ell}u^\ell|^2+\frac{\sigma}{2N}\sum_{j\in[\![N]\!]}|D^3_{iij}u^i|^2+\frac{c_1^2}{\sigma}\,w^i;\end{aligned}$$

by the Cauchy–Schwarz inequality, Young's inequality and Proposition 4.3.6,

$$\begin{aligned}\sum_{j\in[\![N]\!]} \big|D^2_{ij}u^i D^2_{ij}u^j D^2_{ii}u^i\big| &\leqslant |D^2_{ii}u^i|\Big(\sum_{j\in[\![N]\!]\setminus\{i\}} |D^2_{ij}u^i|^2\Big)^{\frac12}\Big(\sum_{j\in[\![N]\!]\setminus\{i\}} |D^2_{ij}u^j|^2\Big)^{\frac12} + |D^2_{ii}u^i|^3 \\ &\leqslant 2C_2^{\frac32} + C_2^{\frac12} w^i;\end{aligned}$$

by the Cauchy–Schwarz inequality, Young's inequality and Proposition 4.3.5,

$$\begin{aligned}\sum_{j\in[\![N]\!]\setminus\{i\}} \big|D_j u^i D^3_{iij}u^j D^2_{ii}u^i\big| &\leqslant |D^2_{ii}u^i|\Big(\sum_{j\in[\![N]\!]\setminus\{i\}} |D_j u^i|^2\Big)^{\frac12}\Big(\sum_{j\in[\![N]\!]\setminus\{i\}} |D^3_{iij}u^j|^2\Big)^{\frac12} \\ &\leqslant \frac{c_1^2}{\sigma} w^i + \frac{\sigma}{2}\sum_{j\in[\![N]\!]\setminus\{i\}} |D^3_{iij}u^j|^2;\end{aligned}$$

by Young's inequality,

$$\sum_{\ell\in[\![N]\!]\setminus\{i\}} D^2_{i\ell}f^\ell D^2_{i\ell}u^\ell + \frac1N D^2_{ii}f^i D^2_{ii}u^i \leqslant \frac12\Big(\sum_{\ell\in[\![N]\!]\setminus\{i\}} |D^2_{i\ell}f^\ell|^2 + \frac1N |D^2_{ii}f^i|^2\Big) + w^i.$$

Therefore, we get

$$\begin{aligned}&-\partial_t w^i - \mathrm{Tr}((\sigma \mathsf{I} + \beta \mathsf{J})D^2 w^i) + \sum_{j\in[\![N]\!]} D_j u^j\cdot D_j w^i \\ &\qquad\leqslant \Big(1 + 4M + \frac{2c_1^2}{\sigma} + C_2^{\frac12}\Big)w^i + C_2^{\frac12}\sup_{\ell\in[\![N]\!]} w^\ell + \frac{2C_2^{\frac32}}{N} \\ &\qquad\quad + \frac12\Big(\sum_{\ell\in[\![N]\!]\setminus\{i\}} |D^2_{i\ell}f^\ell|^2 + \frac1N |D^2_{ii}f^i|^2\Big).\end{aligned}$$

By (**MF**), $\sum_{\ell\in[\![N]\!]\setminus\{i\}} |D^2_{i\ell}f^\ell|^2 + \frac1N |D^2_{ii}f^i|^2 \leqslant 2L_f^2/N$, thus

$$-\partial_t w^i - \mathrm{Tr}((\sigma \mathsf{I} + \beta \mathsf{J})D^2 w^i) + \sum_{j\in[\![N]\!]} D_j u^j\cdot D_j w^i \leqslant C\Big(\sup_{\ell\in[\![N]\!]} w^\ell + \frac1N\Big),$$

where the constant C depends only on σ, T, M, L_g and L_f. By the maximum principle we get

$$\|w^i(t,\cdot)\|_\infty \leqslant \|w^i(T,\cdot)\|_\infty + (T-t)C\Big(\sup_{\ell\in[\![N]\!]} \|w^\ell\|_{\infty;[T-s,T]} + \frac1N\Big).$$

for any $s\in(0,T-\tau]$ and $t\in[T-s,T]$; here we used the notation $\|\cdot\|_{\infty;I}$ for the standard norm of $L^\infty(I\times(\mathbb{R}^d)^N)$. For $s\leqslant(2C)^{-1}$ we obtain

$$\sup_{i\in[\![N]\!]} \|w^i\|_{\infty;[T-s,T]} \leqslant 2\sup_{i\in[\![N]\!]} \|w^i(T,\cdot)\|_\infty + \frac1N; \tag{4.3.17}$$

letting $s=(2C)^{-1}\wedge(T-\tau)$ and iterating estimate (4.3.17) on the intervals $[T-\ell s, T-(\ell-1)s]$ for all positive integers $\ell\leqslant\ell^* := \lfloor (T-\tau)/s\rfloor$ and then on $[\tau, T-\ell^* s]$, we get by induction

$$\sup_{i\in[\![N]\!]} \|w^i\|_{\infty;[\tau,T]} \leqslant 2(2^{\ell^*+1}-1)\Big(2\sup_{i\in[\![N]\!]}\|w^i(T,\cdot)\|_\infty + \frac{1}{N}\Big) - \frac{\ell^*+1}{N}.$$

The conclusion follows recalling that $u^i(T,\cdot)=g^i$, and thus $\sup_{i\in[\![N]\!]}\|w^i(T,\cdot)\| \leqslant 2L_g^2/N$ by **(MF)**. ∎

In order to proceed, we need further control, for fixed i, on second-order derivatives $D_{jk}u^i$ of the value function as j,k vary. We will refer to it as a *horizontal* estimate. Differently from the one obtained in Proposition 4.3.6, we sum over $k\neq i$, thus expecting a decay of order $1/N$. To achieve this, we need more precise control on first-order derivatives: in Proposition 4.3.5, a cumulative information on Du^i was stated; here below, we show *index-wise* bounds, in the sense that no sum over the direction x^j is involved.

PROPOSITION 4.3.8 (Index-wise estimate on skew first-order derivatives). *Let $\tau\in[0,T)$. Suppose that $\mathcal{D}[u]\geqslant -MQ$ on $(\tau,T]\times(\mathbb{R}^d)^{2N}$. Then*

$$\sup_{\substack{i,k\in[\![N]\!]\\ k\neq i}} \|D_k u^i(t,\cdot)\|_\infty \leqslant \frac{C_1}{N} \quad \forall\, t\in[\tau,T],$$

where the constant C_1 depends only on σ, T, M, L_g and L_f.

PROOF. Let $i,k\in[\![N]\!]$, $i\neq k$, and $\ell\in[\![d]\!]$. Note that $v=v^{k\ell}=D_{x^{k\ell}}u^i$ solves the linear problem

$$\begin{cases} -\partial_t v - \operatorname{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2v) + \displaystyle\sum_{j\in[\![N]\!]} D_ju^j\cdot D_jv \\ \qquad\qquad\qquad\qquad = D_{x^{k\ell}}f^i - \displaystyle\sum_{j\in[\![N]\!]\setminus\{i\}} D_j(D_{x^{k\ell}}u^j)\cdot D_ju^i \\ v|_{t=T} = D_{x^{k\ell}}g^i\,. \end{cases}$$

Recall that by **(MF)**, $\|D_{x^{k\ell}}f^i\|_\infty\leqslant L_f/N$. For any $s\in(0,T-\tau]$ and $t\in[T-s,T]$, we control the second term of the right-hand side of the previous equation using Proposition 4.3.5, Proposition 4.3.7 and Proposition 4.3.6 as follows:

$$\begin{aligned} &\Bigg|\sum_{\substack{j\in[\![N]\!]\\ k\neq j\neq i}} D_j(D_{x^{k\ell}}u^j)\cdot D_ju^i + D_k(D_{x^{k\ell}}u^k)\cdot D_ku^i\Bigg| \\ &\qquad\leqslant \Big(\sum_{j\neq k}|D^2_{jk}u^j|^2\Big)^{\frac12}\Big(\sum_{j\neq i}|D_ju^i|^2\Big)^{\frac12} + |D^2_{kk}u^k|\,|D_ku^i| \\ &\qquad\leqslant \frac{c_1C_3}{N} + C_2\|D_ku^i\|_{\infty;[T-s,T]}. \end{aligned}$$

Therefore, by the maximum principle we have

$$\|v\|_{\infty;[T-s,T]} \leqslant s\,\frac{c_1C_3+L_f}{N} + sC_2\|D_ku^i\|_{\infty;[T-s,T]} + \|D_kg^i\|_\infty.$$

Since the previous estimate holds for all $\ell \in [\![d]\!]$, for $s \leqslant (2\sqrt{d}\,C_2)^{-1}$ we conclude that

$$\|D_k u^i\|_{\infty;[T-s,T]} \leqslant \frac{c_1 C_3 + L_f}{C_2 N} + 2\sqrt{d}\,\|D_k g^i\|_\infty .$$

By iterating the previous inequality as in the previous proof, and using **(MF)** again to control $\|D_{x^k} g^i\|_\infty \leqslant L_g/N$ we obtain the assertion on the time interval $[T-\tau, T]$. ∎

We can now get the horizontal bounds on skew second-order derivatives.

PROPOSITION 4.3.9 (Horizontal estimate on skew second-order derivatives). *Let $\tau \in [0,T)$. Suppose that $\mathcal{D}[u] \geqslant -MQ$ on $(\tau,T] \times (\mathbb{R}^d)^{2N}$. Then*

$$\sup_{i\in[\![N]\!]} \sum_{k\in[\![N]\!]\setminus\{i\}} \|D(D_k u^i)(t,\cdot)\|_\infty^2 \leqslant \frac{C_4}{N} \quad \forall\, t \in [\tau,T]\,,$$

where the constant C_4 depends only on σ, T, M, L_g and L_f.

PROOF. From inequality (4.3.14) we have

$$\begin{aligned}\|D_{x^{k\ell}} D u^i\|_\infty \leqslant{}& C_\sigma \|D_{x^{k\ell}} f^i\|_\infty s^{\frac12} e^{2Ms} + \|D(D_{x^{k\ell}} g^i)\|_\infty e^{Ms} \\ &+ C_\sigma s^{\frac12} e^{2Ms} \Bigg\| \sum_{\substack{j\in[\![N]\!]\\ k\neq j\neq i}} D_j(D_{x^{k\ell}} u^j) D_j u^i \Bigg\|_\infty \\ &+ C_\sigma s^{\frac12} e^{2Ms} \|D_k(D_{x^{k\ell}} u^k) D_k u^i\|_\infty \,. \end{aligned} \tag{4.3.18}$$

By the Cauchy–Schwarz inequality, Proposition 4.3.5 and Proposition 4.3.7,

$$\Bigg\| \sum_{\substack{j\in[\![N]\!]\\ k\neq j\neq i}} D_j(D_{x^{k\ell}} u^j) D_j u^i \Bigg\|_\infty^2 \leqslant \frac{c_1^2}{N} \Bigg\| \sum_{j\in[\![N]\!]\setminus\{k\}} \big|D^2_{jk} u^j\big|^2 \Bigg\|_\infty \leqslant \frac{c_1^2 C_3}{N^2};$$

therefore from (4.3.18) and Young's inequality we get

$$\begin{aligned}\|D(D_k u^i)\|_\infty^2 \lesssim{}& C_\sigma^2 \|D_k f^i\|_\infty^2 e^{4Ms} s + \|D(D_k g^i)\|_\infty^2 e^{2Ms} + \frac{C_\sigma^2 c_1^2 C_3 e^{4Ms} s}{N^2} \\ &+ C_\sigma^2 e^{4Ms} s \|D^2_{kk} u^k\|_\infty^2 \|D_k u^i\|_\infty^2, \end{aligned} \tag{4.3.19}$$

where the implied constant depends only on d. By **(MF)**, $\sum_{k\in[\![N]\!]\setminus\{i\}} \|D_k f^i\|_\infty^2 \leqslant L_f^2/N$ and $\sum_{k\in[\![N]\!]\setminus\{i\}} \|D(D_k g^i)\|_\infty^2 \leqslant L_g^2/N$; by Proposition 4.3.6 and Proposition 4.3.8

$$\sum_{k\in[\![N]\!]\setminus\{i\}} \|D^2_{kk} u^k\|_\infty^2 \|D_k u^i\|_\infty^2 \leqslant \frac{C_1 C_2}{N}\,. \tag{4.3.20}$$

The desired conclusion now follows, summing (4.3.19) over $k \neq i$. ∎

REMARK 4.3.10. Proposition 4.3.8 is clearly stronger than Proposition 4.3.5, nevertheless we cited them both in the previous proof in order to highlight when the weaker Proposition 4.3.5 was sufficient and, consequently, that the stronger Proposition 4.3.8 is only needed to get estimate (4.3.20).

4.3.4. Third-order derivatives of the value functions. Finally, we prove the most basic third-order version of the transversal estimate on skew-second order derivatives.

PROPOSITION 4.3.11 (Transversal estimate on third-order derivatives). *Let $\tau \in [0,T)$. Suppose that $\mathcal{D}[u] \geqslant -MQ$ on $(\tau,T] \times (\mathbb{R}^d)^{2N}$. Then*

$$\sum_{\substack{i,j \in [\![N]\!] \\ j \neq i}} \|D(D_{ij}u^j)(t,\cdot)\|_\infty^2 \leqslant C_5 \quad \forall\, t \in [\tau,T]\,,$$

where the constant C_5 depends only on σ, T, M, L_g and L_f.

PROOF. Proceed as in the proof of Proposition 4.3.6. Let $h,k \in [\![N]\!]$, $\ell,m \in [\![d]\!]$ and $s \in [0,T-\tau]$. Note that $v = D_{x^{k\ell}x^{hm}}u^i$ solves problem (4.3.1) on $[T-s,T] \times (\mathbb{R}^d)^N$ with $b^j = D_j u^j$,

$$\begin{aligned} F = D^2_{x^{k\ell}x^{hm}} f^i &- \sum_{j \in [\![N]\!] \setminus \{i\}} D_j(D^2_{x^{k\ell}x^{hm}}u^j) D_j u^i - \sum_{j \in [\![N]\!] \setminus \{i\}} D_j(D_{x^{k\ell}}u^j) D_j(D_{x^{hm}}u^i) \\ &- \sum_{j \in [\![N]\!] \setminus \{i\}} D_j(D_{x^{hm}}u^j) D_j(D_{x^{k\ell}}u^i) - D_i(D_{x^{hm}}u^i) D_i(D_{x^{k\ell}}u^i) \end{aligned}$$

and $G = D^2_{x^{k\ell}x^{hm}} g^i$. By Lemma 4.3.1,

$$\begin{aligned} &\|D(D^2_{x^{k\ell}x^{hm}}u^i)\|_\infty \\ &\quad \lesssim \|D^2_{x^{k\ell}x^{hm}} Dg^i\|_\infty e^{Ms} + C_\sigma \Bigg(\Bigg\| \sum_{j \in [\![N]\!] \setminus \{i\}} D_j(D^2_{x^{k\ell}x^{hm}}u^j) D_j u^i \Bigg\|_\infty \\ &\qquad + \Bigg\| \sum_{j \in [\![N]\!] \setminus \{i\}} D_j(D_{x^{k\ell}}u^j) D_j(D_{x^{hm}}u^i) \Bigg\|_\infty \\ &\qquad + \Bigg\| \sum_{j \in [\![N]\!] \setminus \{i\}} D_j(D_{x^{hm}}u^j) D_j(D_{x^{k\ell}}u^i) \Bigg\|_\infty \\ &\qquad + \|D_i(D_{x^{hm}}u^i) D_i(D_{x^{k\ell}}u^i)\|_\infty s^{\frac12} e^{2Ms} + \|D^2_{x^{k\ell}x^{hm}} f^i\|_\infty \Bigg) s^{\frac12} e^{2Ms}\,, \end{aligned} \tag{4.3.21}$$

where the implied constant is a number and the L^∞-norm are understood to be computed on $[T-s,T] \times (\mathbb{R}^d)^N$ or $(\mathbb{R}^d)^N$. By Proposition 4.3.5 and the Cauchy–Schwarz inequality we have

$$\Bigg\| \sum_{j \in [\![N]\!] \setminus \{i\}} D_j(D^2_{x^{k\ell}x^{hm}}u^j) D_j u^i \Bigg\|_\infty^2 \leqslant \frac{c_1^2}{N} \Bigg\| \sum_{j \in [\![N]\!] \setminus \{i\}} \big|D^3_{hjk}u^j\big|^2 \Bigg\|_\infty ;$$

therefore, choosing $k = i$ in (4.3.21), and applying Young's inequality we get

$$\begin{aligned}(4.3.22)\quad & \|D(D^2_{hi}u^i)\|^2_\infty \\ & \lesssim \|D^2_{hi}Dg^i\|^2_\infty e^{2Ms} + \Big\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D^2_{ji}u^j\big|^2 \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D^2_{hj}u^i\big|^2 \Big\|_\infty se^{4Ms} \\ & + \Big\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D^2_{hj}u^j\big|^2 \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D^2_{ji}u^i\big|^2 \Big\|_\infty se^{4Ms} + \|D^2_{hi}f^i\|^2_\infty se^{4Ms} \\ & + \|D_{hi}u^i\|^2_\infty \|D_{ii}u^i\|^2_\infty se^{4Ms} + \frac{c_1^2 se^{4Ms}}{N} \sum_{j\in[\![N]\!]\setminus\{i\}} \|D^3_{hij}u^j\|^2_\infty\,,\end{aligned}$$

where the implied constant depends only on C_σ and d. By assumption (**MF**), $\sum_{h,i\in[\![N]\!],\,i\neq h}\|D^2_{hi}f^i\|^2_\infty \leqslant L_f^2$; by Propositions 4.3.7 and 4.3.9,

$$\sup_{i\in[\![N]\!]} \Big\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D^2_{ij}u^j\big|^2 \Big\|_\infty \sum_{\substack{h,i\in[\![N]\!]\\ h\neq i}} \Big\| \sum_{j\in[\![N]\!]} \big|D^2_{hj}u^i\big|^2 \Big\|_\infty \leqslant \frac{C_3C_4}{N};$$

by Propositions 4.3.6 and 4.3.7,

$$\sum_{\substack{h,i\in[\![N]\!]\\ i\neq h}} \|D_{hi}u^i\|^2_\infty \|D_{ii}u^i\|^2_\infty \leqslant C_2C_3.$$

Therefore, summing inequality (4.3.22) over h and $k = i \neq h$ yields

$$\begin{aligned}&\sum_{\substack{h,i\in[\![N]\!]\\ i\neq h}} \|D(D^2_{hi}u^i)\|^2_\infty \\ &\qquad \lesssim se^{4Ms} + e^{2Ms} \sum_{\substack{h,i\in[\![N]\!]\\ i\neq h}} \|D(D^2_{hi}g^i)\|^2_\infty + \frac{1}{N} \sum_{\substack{h,i,j\in[\![N]\!]\\ j\neq i}} \|D^3_{hij}u^j\|^2_\infty se^{4Ms}\,,\end{aligned}$$

where the implied constant (here as well as below) depends only on σ, T, M, L_f and L_g. Exploiting now the fact that

$$\sum_{\substack{h,i,j\in[\![N]\!]\\ j\neq i}} \|D^3_{hij}u^j\|^2_\infty \leqslant N \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} \Big\| \sum_{h\in[\![N]\!]} \big|D^3_{hij}u^j\big|^2 \Big\|_\infty = N \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} \|D(D_{ij}u^j)\|^2_\infty\,,$$

we obtain, for small s,

$$\sum_{\substack{h,i\in[\![N]\!]\\ i\neq h}} \|D(D^2_{hi}u^i)\|^2_\infty \lesssim 1 + \sum_{\substack{h,i\in[\![N]\!]\\ i\neq h}} \|D(D^2_{hi}g^i)\|^2_\infty = 1 + \sum_{\substack{h,i\in[\![N]\!]\\ i\neq h}} \|D(D^2_{hi}u^i)(T,\cdot)\|^2_\infty\,.$$

Using that $\sum_{h,i\in[\![N]\!],\,i\neq h}\|D(D^2_{hi}g^i)\|^2_\infty \leqslant L_g^2$ due to assumption (**MF**), one concludes by arguing as in the proof of Proposition 4.3.6.

∎

4.4. The $\mathscr{D}$-semimonotone case

We prove in this section Theorem 4.2.7 in the $\mathscr{D}$-semimonotone case. The following result shows the crucial interplay between the previous estimates and the $\mathscr{D}$-semimonotonicity of the value functions. It basically shows that if there exists a left temporal open neighbourhood $(\tau, T]$ of T in which u is M-$\mathscr{D}$-semimonotone for some suitable M, then exploiting the estimates on the skew derivatives one can show that such a semimonotonicity actually holds with constant $\frac{1}{2}M$ on the whole $[\tau, T] \times (\mathbb{R}^d)^N$, provided that N is large.

This implies that if the matrix $(D_{ij}u^j)_{i,j\in[\![N]\!]} + M\mathsf{I}$ is positive definite near T, then it cannot degenerate at any time in $[0, T]$. In this sense, we can consider this fact as a (backward) *propagation* of the semimonotonicity. Such an argument will in turn imply that all estimates on the derivatives themselves propagate, thus small-time existence of a $\mathscr{D}$-semimonotone solution to the Nash system is sufficient to prove global existence on $[0, T]$ of a solution which satisfies all the above estimates (see Theorem 4.4.2 below).

Proposition 4.4.1 (Improvement of $\mathscr{D}$-semimonotonicity). *Assume **(MF)**. There exist positive constants M^*, M_f^* and M_g^* (with $M_g^* < M^*$) depending on T, and a natural number N_* depending on σ, T, L_g and L_f such that for any $\tau \in [0, T)$ one has*

$$\left.\begin{array}{ll} \mathcal{D}[u] \geqslant -M^*Q & \text{on } (\tau, T] \times (\mathbb{R}^d)^{2N} \\ \mathcal{D}[f] \geqslant -M_f^*Q & \text{on } (\mathbb{R}^d)^{2N} \\ \mathcal{D}[g] \geqslant -M_g^*Q & \text{on } (\mathbb{R}^d)^{2N} \\ N \geqslant N_* & \end{array}\right\} \implies \mathcal{D}[u(\tau, \cdot)] \geqslant -\frac{1}{2}M^*Q \quad \text{on } (\mathbb{R}^d)^{2N}.$$

Looking at (4.4.7) below, one can in fact choose $M_g^* = (12eT)^{-1}$, $M_f^* = (12eT^2)^{-1}$ and $M^* = (2T)^{-1}$.

The proof involves again the method of doubling variables. We could proceed with a PDE approach as before, but we prefer to follow a "dynamic" approach; that is, we argue along optimal trajectories. This is probably closer in the spirit to what is usually done in the MFG theory.

More precisely, we use a synchronous coupling between two solutions of (4.2.2), for agents playing optimally. For the benefit of the reader we briefly recall what one means by such a coupling. Let $\bar{t} \in [0, T)$ and consider the $(\mathbb{R}^d)^m$-valued SDE

$$\mathrm{d}Z_t = b(t, Z_t)\,\mathrm{d}t + \Sigma(t, Z_t)\,\mathrm{d}\tilde{B}_t, \quad t \in [\bar{t}, T], \tag{4.4.1}$$

where Σ is an $md \times \ell d$ matrix and $\tilde{B}_t$ is an ℓd-dimensional Brownian motion. Assume that the drift b and the diffusion coefficient Σ are locally Lipschitz continuous and have sublinear growth with respect to the state variable, uniformly with respect to time. If Z solves (4.4.1) with $Z_{\bar{t}} = \bar{z} \in (\mathbb{R}^d)^m$, then the following *Dynkin's formula* holds for any $h \in C^{1,2}([\bar{t},T]\times(\mathbb{R}^d)^m)$ such that $\partial_t^\ell D^\gamma h(t,\cdot)$ has polynomial growth for all $t\in[\bar{t},T]$ and $2\ell+|\gamma|\leqslant 2$:

$$\mathbb{E}\big[h(T,Z_T)\big] = h(\bar{t},\bar{z}) + \int_{\bar{t}}^{T} \mathbb{E}\big[\partial_t h(t,Z_t) + Lh(t,Z_t)\big]\,\mathrm{d}t, \tag{4.4.2}$$

where $L = \frac{1}{2}\operatorname{Tr}\big(\Sigma\Sigma^{\mathsf{T}} D_z^2\big) + b\cdot D_z$ is the differential generator of the process Z; see, e.g., [**4**, Chapter 9]. We are interested in the particular case where $m=2N$ and Z is the following coupling of two solution of the $(\mathbb{R}^d)^N$-valued SDE

$$\mathrm{d}X_t = \alpha(t,X_t)\,\mathrm{d}t + \sigma(t,X_t)\,\mathrm{d}\tilde{B}_t, \quad t\in[\bar{t},T], \tag{4.4.3}$$

where σ is an $Nd\times \ell d$ matrix and $\tilde{B}_t$ is a ℓd-dimensional Brownian motion. Given X and Y which both solve (4.4.3), with $X_{\bar{t}}=\bar{x}$ and $Y_{\bar{t}}=\bar{y}$, respectively, we say that $Z=(X,Y)^{\mathsf{T}}$ with $Z_{\bar{t}}\sim\mu$ having marginals $\delta_{\bar{x}}$ and $\delta_{\bar{y}}$ is a *synchronous coupling* of X and Y. It solves equation (4.4.1) with $b(t,x,y)=(\alpha(t,x),\alpha(t,y))^{\mathsf{T}}$ and $\Sigma(t,x,y)=(\sigma(t,x),\sigma(t,y))^{\mathsf{T}}$, thus

$$b\cdot D_z = \alpha(t,x)\cdot D_x + \alpha(t,y)\cdot D_y, \quad \Sigma\Sigma^{\mathsf{T}} = \begin{pmatrix} \sigma(t,x)\sigma(t,x)^{\mathsf{T}} & \sigma(t,x)\sigma(t,y)^{\mathsf{T}} \\ \sigma(t,y)\sigma(t,x)^{\mathsf{T}} & \sigma(t,y)\sigma(t,y)^{\mathsf{T}} \end{pmatrix}.$$

In particular, when (4.4.3) coincides with (4.2.2) in the equilibrium, which happens for $\alpha=\alpha^* := (-D_i u^i)_{i\in[\![N]\!]}$, $\tilde{B}=(B,W)^{\mathsf{T}}$ and $\sigma\equiv\sqrt{2}\big(\sqrt{\sigma}\,\mathsf{I}\mid\sqrt{\beta}\,1_N\otimes I_d\big)$, one has $\Sigma\Sigma^{\mathsf{T}} = 2(J_2\otimes(\sigma\mathsf{I}+\beta\mathsf{J}))$ and thus

$$L = \operatorname{Tr}\big((J_2\otimes(\sigma\mathsf{I}+\beta\mathsf{J}))D^2_{(x,y)}\big) + \alpha^*(t,x)\cdot D_x + \alpha^*(t,y)\cdot D_y. \tag{4.4.4}$$

Clearly, analogous considerations hold if $Z_{\bar{t}}$ is any random variable with finite $\mathbb{E}\big[|Z_{\bar{t}}|^k\big]$ for some $k\in\mathbb{N}$; in this case, (4.4.2) holds with $h(\bar{t},\bar{z})$ replaced by $\mathbb{E}\big[h(\bar{t},Z_{\bar{t}})\big]$ and provided that $\partial_t^\ell D^\gamma h(t,\cdot)\lesssim 1+|\cdot|^k$.

PROOF OF PROPOSITION 4.4.1. Let $w^i(t,x,y) := D_i u^i(t,\cdot)\big|_y^x\cdot(x^i-y^i)$. We want to use formula (4.4.2) where $\bar{t}=\tau$, $h=w^i$ and $Z=(X,Y)$ is a synchronous coupling of two solutions to (4.2.2); that is, formula (4.4.2) with L given by (4.4.4). We have

$$\begin{aligned} Lw^i(t,x,y) &= D_i \operatorname{Tr}\big((\sigma\mathsf{I}+\beta\mathsf{J})D^2u^i(t,\cdot)\big|_y^x\big)\cdot(x^i-y^i) \\ &\qquad - \sum_{j\in[\![N]\!]} D^2_{ij}u^i(t,\cdot)D_ju^j(t,\cdot)\big|_y^x\cdot(x^i-y^i) - \big|D_iu^i(t,\cdot)\big|_y^x\big|^2, \end{aligned}$$

and, according to (4.2.1),

$$\begin{aligned} \partial_t w^i(t,x,y) &= -D_i \operatorname{Tr}\big((\sigma\mathsf{I}+\beta\mathsf{J})D^2u^i(t,\cdot)\big|_y^x\big)\cdot(x^i-y^i) \\ &\quad + \sum_{j\in[\![N]\!]} D^2_{ij}u^i(t,\cdot)D_ju^j(t,\cdot)\big|_y^x\cdot(x^i-y^i) \\ &\quad + \sum_{j\in[\![N]\!]\setminus\{i\}} D^2_{ij}u^j(t,\cdot)D_ju^i(t,\cdot)\big|_y^x\cdot(x^i-y^i) - D_if^i\big|_y^x\cdot(x^i-y^i)\,. \end{aligned}$$

Then equality (4.4.2) gives

$$\mathbb{E}[w^i(T,X_T,Y_T)] \leqslant w^i(\tau,\bar{x},\bar{y}) - \int_\tau^T \mathbb{E}\Big[D_i f^i\big|_{Y_t}^{X_t} \cdot (X_t^i - Y_t^i)\Big] + \int_\tau^T \mathbb{E}\,\mathcal{E}_t^i \,\mathrm{d}t,$$

where

$$\mathcal{E}_t^i := \sum_{j\in[\![N]\!]\setminus\{i\}} D_{ij}^2 u^j(t,\cdot) D_j u^i(t,\cdot)\big|_{Y_t}^{X_t} \cdot (X_t^i - Y_t^i)\,.$$

Letting $w := \sum_{i\in[\![N]\!]} w^i$, it follows that

$$w(\tau,\bar{x},\bar{y}) \geqslant -M_g^*\,\mathbb{E}\big[|X_T - Y_T|^2\big] - M_f^* \int_\tau^T \mathbb{E}\big[|X_t - Y_t|^2\big]\mathrm{d}t - \int_\tau^T \mathbb{E}\Big[\sum_{i\in[\![N]\!]} \mathcal{E}_t^i\Big]\,\mathrm{d}t\,. \tag{4.4.5}$$

By the mean value theorem

$$D_{ij}^2 u^j(t,\cdot) D_j u^i(t,\cdot)\big|_{Y_t}^{X_t} = \bigg(\int_0^1 D\big(D_{ij}^2 u^j(t,Z_t(s)) D_j u^i(t,Z_t(s))\big)\,\mathrm{d}s\bigg)(X_t - Y_t)\,,$$

with $Z_t(s) = sX_t + (1-s)Y_t$; then note that

$$\begin{aligned}
&\sum_{i\in[\![N]\!]} \mathcal{E}_t^i \\
&\quad = \int_0^1 \sum_{i,k\in[\![N]\!]} \sum_{j\in[\![N]\!]\setminus\{i\}} \big(D_{ijk}^3 u^j(t,Z_t(s)) D_j u^i(t,Z_t(s))\big)(X_t^k - Y_t^k)\cdot(X_t^i - Y_t^i)\,\mathrm{d}s \\
&\quad + \int_0^1 \sum_{i,k\in[\![N]\!]} \sum_{j\in[\![N]\!]\setminus\{i\}} \big(D_{ij}^2 u^j(t,Z_t(s))) D_{jk}^2 u^i(t,Z_t(s))\big)(X_t^k - Y_t^k)\cdot(X_t^i - Y_t^i)\,\mathrm{d}s\,,
\end{aligned}$$

where

$$\begin{aligned}
&\bigg|\int_0^1 \sum_{i,k\in[\![N]\!]} \sum_{j\in[\![N]\!]\setminus\{i\}} \big(D_{ijk}^3 u^j(t,Z_t(s)) D_j u^i(t,Z_t(s))\big)(X_t^k - Y_t^k)\cdot(X_t^i - Y_t^i)\,\mathrm{d}s\bigg| \\
&\qquad \leqslant \bigg\| \sum_{i,k\in[\![N]\!]} \bigg| \sum_{j\in[\![N]\!]\setminus\{i\}} D_{ijk}^3 u^j D_j u^i \bigg|^2 \bigg\|_\infty^{\frac12} |X_t - Y_t|^2 \\
&\qquad \leqslant \bigg(\sum_{i\in[\![N]\!]} \bigg\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D(D_{ij}^2 u^j)\big|^2 \bigg\|_\infty \bigg\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D_j u^i\big|^2 \bigg\|_\infty \bigg)^{\frac12} |X_t - Y_t|^2
\end{aligned}$$

and similarly

$$\begin{aligned}
&\bigg|\int_0^1 \sum_{i,k\in[\![N]\!]} \sum_{j\in[\![N]\!]\setminus\{i\}} \big(D_{ij}^2 u^j(t,Z_t(s))) D_{jk}^2 u^i(t,Z_t(s))\big)(X_t^k - Y_t^k)\cdot(X_t^i - Y_t^i)\,\mathrm{d}s\bigg| \\
&\qquad \leqslant \bigg(\sum_{i\in[\![N]\!]} \bigg\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D_{ij}^2 u^j\big|^2 \bigg\|_\infty \bigg\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D(D_j u^i)\big|^2 \bigg\|_\infty \bigg)^{\frac12} |X_t - Y_t|^2.
\end{aligned}$$

Propositions 4.3.5 and 4.3.11 ensure that

$$\sum_{i\in[\![N]\!]} \Big\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D(D^2_{ij}u^j)\big|^2 \Big\|_\infty \Big\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D_j u^i\big|^2 \Big\|_\infty \leqslant \frac{c_1^2 C_5}{N}$$

and Propositions 4.3.7 and 4.3.9 give

$$\sum_{i\in[\![N]\!]} \Big\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D^2_{ij}u^j\big|^2 \Big\|_\infty \Big\| \sum_{j\in[\![N]\!]\setminus\{i\}} \big|D(D_j u^i)\big|^2 \Big\|_\infty \leqslant \frac{C_3 C_4}{N}\,.$$

Therefore we have

$$\int_\tau^T \mathbb{E}\Big[\sum_{i\in[\![N]\!]} \mathcal{E}^i_t\Big] \leqslant \frac{\overline{C}}{\sqrt{N}} \sup_{t\in[\tau,T]} \mathbb{E}\big[|X_t - Y_t|^2\big]$$

for some constant $\overline{C}$ which depends only σ, T, M^*, L_f and L_g. On the other hand, by the equation of $X - Y$,

$$\begin{aligned} \mathrm{d}|X_t - Y_t|^2 &= -2 \sum_{i\in[\![N]\!]} (D_i u^i(t,X_t) - D_i u^i(t,Y_t)) \cdot (X^i_t - Y^i_t)\,\mathrm{d}t \\ &\leqslant 2M^*|X_t - Y_t|^2\,\mathrm{d}t\,, \end{aligned}$$

where the inequality comes from the $\mathscr{D}$-semimonotonicity of u, so that Gronwall's lemma yields

$$\mathbb{E}\big[|X_t - Y_t|^2\big] \leqslant |\bar{x} - \bar{y}|^2 e^{2M^*(t-\tau)} \quad \forall\, t \in [\tau, T]. \tag{4.4.6}$$

Hence, from (4.4.5) we have obtained

$$w(\tau,\bar{x},\bar{y}) \geqslant -e^{2M^*T}\Big(M^*_g + TM^*_f + \frac{\overline{C}}{\sqrt{N}}\Big)|\bar{x} - \bar{y}|^2.$$

Fix now $M^* = (2T)^{-1}$; choosing M^*_g and M^*_f small enough and N_* large enough so that

$$M^*_g + TM^*_f + \frac{\overline{C}}{\sqrt{N_*}} \leqslant \frac{M^*}{2e^{2M^*T}} = \frac{1}{4eT} \tag{4.4.7}$$

we see that $w(\tau,\bar{x},\bar{y}) \geqslant -\frac{1}{2}M^*|\bar{x} - \bar{y}|^2$ holds for any $N \geqslant N_*$. The conclusion follows by the arbitrariness of $\bar{x}, \bar{y} \in (\mathbb{R}^d)^N$. ∎

As anticipated, such a non-degeneration of semimonotonicity of the solution is the key ingredient for the propagation of all derivative estimates, and of semimonotonicity, over the whole time horizon $[0,T]$. This constitutes the first half of Theorem 4.2.7.

THEOREM 4.4.2 (Estimates on the Nash system with $\mathscr{D}$-semimonotone data). *Assume that f and g satisfy assumptions **(MF)** and **(DS)**. Let $T > 0$. There exist positive constants M^*_f and M^*_g depending on T (in such a way that $M^*_f, M^*_g \to 0$ as $T \to +\infty$ and $M^*_f, M^*_g \to +\infty$ as $T \to 0$), and a natural number N^* depending* *only on σ, T, L_f and L_g, such that if $M_g \leqslant M_g^*$, $M_f \leqslant M_f^*$ and $N \geqslant N^*$ then any solution u to (4.2.1) on $[0,T] \times (\mathbb{R}^d)^N$ satisfies*

$$\sup_{i\in[\![N]\!]} \left(\sup_{j\in[\![N]\!]\setminus\{i\}} \|D_j u^i\|_\infty + \Bigg\| \sum_{j\in[\![N]\!]\setminus\{i\}} \left|D^2_{ij} u^j\right|^2 \Bigg\|_\infty + \sum_{j\in[\![N]\!]\setminus\{i\}} \|D(D_j u^i)\|^2_\infty \right) \lesssim \frac{1}{N}$$

and

$$\sup_{i\in[\![N]\!]} \|Du^i\|_\infty + \sup_{i\in[\![N]\!]} \sum_{j\in[\![N]\!]} \|D(D_i u^j)\|^2_\infty + \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} \|D(D_{ij} u^j)\|^2_\infty \lesssim 1,$$

where the implied constants depend only on σ, T, L_f and L_g. In addition, there exists $M^* > 0$ (depending only on T) such that u is M^*-$\mathscr{D}$-semimonotone on $[0,T]$.

Proof. Let M_f^*, M_g^*, M^* and N^* be given by Proposition 4.4.1. By assumption **(DS)**, $\mathcal{D}[u(T,\cdot)] + M^*Q \geqslant \varepsilon Q$ on $(\mathbb{R}^d)^{2N}$, with $\varepsilon = M^* - M_g^* > 0$; that is, by Remark 4.2.3, $(D^2_{ij} u^i(T,\cdot))_{i,j\in[\![N]\!]} + M^*\mathsf{I} \geqslant \varepsilon\mathsf{I}$ on $(\mathbb{R}^d)^N$. Since $D^2_{ij}u^i$ are assumed to be uniformly continuous, there exists $\tau \in [0,T)$ (a priori dependent of N) such that $(D^2_{ij} u^i(t,\cdot))_{i,j\in[\![N]\!]} + M^*\mathsf{I} \geqslant \frac{\varepsilon}{2}\mathsf{I}$ on $(\mathbb{R}^d)^N$ for all $t \in (\tau, T]$. Therefore,

$$\mathcal{T} := \big\{ s \in [0,T) : \ u \text{ extends to a solution on } (s,T] \times (\mathbb{R}^d)^N \\ \text{and } \mathcal{D}[u] > -M^*Q \text{ therein} \big\} \neq \emptyset$$

and $\tau^* := \inf \mathcal{T} \in [0,\tau]$. Seeking for a contradiction, suppose that $\tau^* > 0$. Then $\mathcal{D}[u] \geqslant -M^*Q$ on $[\tau^*,T] \times (\mathbb{R}^d)^{2N}$ and by Proposition 4.4.1 we have $\mathcal{D}[u(\tau^*,\cdot)] \geqslant -\frac{1}{2}M^*Q$ on $(\mathbb{R}^d)^{2N}$. Repeating the same argument as at the beginning of the proof, there exists $\tau' \in \mathcal{T}$, $\tau' < \tau^*$, thus contradicting the definition of τ^*. This proves the $\mathscr{D}$-semimonotonicity of u on $[0,T]$, hence all estimates on the derivatives are given by Propositions 4.3.6 to 4.3.9 and 4.3.11. ∎

4.5. The $\mathscr{L}$-semimonotone case

We prove in this section Theorem 4.2.7 in the $\mathscr{L}$-semimonotone case. Our observations in Remark 4.2.3 motivate the following result, which comprehends all estimates on derivatives we have made in the previous section, under the assumption of $\mathscr{L}$-semimonotonicity and a lower bound on the diagonal of the Jacobian matrix of the vector of optimal controls.

Proposition 4.5.1 (Estimates on the derivatives). *Let $\tau \in [0,T)$. Suppose that $\mathcal{L}[u] \geqslant -\kappa Q$ on $(\tau,T] \times (\mathbb{R}^d)^{2N}$ and there exists $\gamma > 0$ (independent of N), such that*

$$\triangle_{u(t,\cdot)} \geqslant -\gamma\mathsf{I} \quad \text{on } (\mathbb{R}^d)^N \quad \forall\, t \in (\tau,T]. \tag{4.5.1}$$

Then there exists constants K_i, $i \in \{1, \dots, 5\}$, depending only on σ, T, κ, γ, L_g and L_f, such that the following estimates hold for all $t \in [\tau, T]$:

$$\sup_{\substack{i,j\in[\![N]\!]\\ j\neq i}} \|D_j u^i(t,\cdot)\|_\infty \leqslant \frac{K_1}{N}; \tag{4.5.2}$$

$$\sup_{i\in[\![N]\!]} \sum_{j\in[\![N]\!]} \|D(D_i u^j)(t,\cdot)\|_\infty^2 \leqslant K_2; \tag{4.5.3}$$

$$\sup_{i\in[\![N]\!]} \Bigg\| \sum_{j\in[\![N]\!]\setminus\{i\}} \left|D^2_{ij} u^j(t,\cdot)\right|^2 \Bigg\|_\infty \leqslant \frac{K_3}{N}; \tag{4.5.4}$$

$$\sup_{i\in[\![N]\!]} \sum_{j\in[\![N]\!]\setminus\{i\}} \|D(D_j u^i)(t,\cdot)\|_\infty^2 \leqslant \frac{K_4}{N}; \tag{4.5.5}$$

$$\sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} \|D(D_{ij} u^j)(t,\cdot)\|_\infty^2 \leqslant K_5. \tag{4.5.6}$$

More precisely, each K_i is given by the constant C_i appearing in Propositions 4.3.6 to 4.3.9 and 4.3.11 with M replaced by $\kappa + \gamma$.

Proof. As noted in Remark 4.2.3, we have $\mathcal{D}[u] \geqslant -(\kappa+\gamma)Q$ on $(\tau, T] \times (\mathbb{R}^d)^{2N}$. Then the estimates follows from Propositions 4.3.6 to 4.3.9 and 4.3.11. ∎

With this piece of information one can show that the pair formed by $\mathscr{L}$-semimonotonicity and diagonal lower bound (4.5.1) does not degenerate, just like displacement semiconvexity did not (cf. Proposition 4.4.1).

Proposition 4.5.2 (Improvement of $\mathscr{L}$-monotonicity and diagonal lower bound). *There exist nonnegative constants κ_f^*, κ_g^*, κ^* and γ^* and a natural number N_*, all depending only on σ, T, L_g and L_f, such that $\kappa_g^* < \kappa^*$, $L_g < \gamma^*$ and, for any $\tau \in [0, T)$, one has*

$$\left.\begin{array}{ll} \mathcal{L}[u] \geqslant -\kappa^* Q & \text{on } (\tau,T]\times(\mathbb{R}^d)^{2N} \\ \triangle_u \geqslant -\gamma^* \mathsf{I} & \text{on } (\tau,T]\times(\mathbb{R}^d)^{N} \\ \mathcal{L}[f] \geqslant -\kappa_f^* Q & \text{on } (\mathbb{R}^d)^{2N} \\ \mathcal{L}[g] \geqslant -\kappa_g^* Q & \text{on } (\mathbb{R}^d)^{2N} \\ N \geqslant N_* & \end{array}\right\} \implies \begin{cases} \mathcal{L}[u(\tau,\cdot)] \geqslant -\frac{1}{2}\kappa^* Q & \text{on } (\mathbb{R}^d)^{2N} \\ \triangle_u(\tau,\cdot) \geqslant -\frac{1}{2}\gamma^* \mathsf{I} & \text{on } (\mathbb{R}^d)^{N}. \end{cases}$$

We are going to split the proof of this result in two propositions. We will prove that the $\mathscr{L}$-semimonotonicity improves, provided that one has the diagonal lower bound (see Proposition 4.5.3). Then, symmetrically, that such a lower bound improves, provided that one knows that u is $\mathscr{L}$-semimonotone (see Proposition 4.5.4).

Proposition 4.5.3 (Improvement of $\mathscr{L}$-semimonotonicity). *Let $\gamma > 0$. Then there exist a nonnegative constant κ^* depending on T (but not on γ), and constants κ_f^*,*

κ_g^* *(with* $\kappa_g^* < \kappa$*) and a natural number* N'_**, all depending only on* σ*,* T*,* γ*,* L_g *and* L_f*, such that for any* $\tau \in [0,T)$ *one has*

$$\left.\begin{aligned} &\triangle_{u(t,\cdot)} \geqslant -\gamma \mathsf{I} && \text{on } (\tau,T] \times (\mathbb{R}^d)^N \\ &\mathcal{L}[u] \geqslant -\kappa^* Q && \text{on } (\tau,T] \times (\mathbb{R}^d)^{2N} \\ &\mathcal{L}[f] \geqslant -\kappa_f^* Q && \text{on } (\mathbb{R}^d)^{2N} \\ &\mathcal{L}[g] \geqslant -\kappa_g^* Q && \text{on } (\mathbb{R}^d)^{2N} \\ &N \geqslant N'_* \end{aligned}\right\} \implies \mathcal{L}[u(\tau,\cdot)] \geqslant -\frac{1}{2}\kappa^* Q \quad \text{on } (\mathbb{R}^d)^{2N}.$$

Looking at the proof below, one can in fact choose $\kappa_g^* = (12Te^{2\gamma T+1})^{-1}$, $\kappa_f^* = (12T^2e^{2\gamma T+1})^{-1}$ and $\kappa^* = (2T)^{-1}$, therefore κ_h^* is actually independent of L_g and L_f.

Proof of Proposition 4.5.3. Let $w(t,x,y) := \mathcal{L}[u(t,\cdot)](x,y)$. We want to use formula (4.4.2) where $\bar{t} = \tau$, $h = w$ and $Z = (X,Y)$ is a synchronous coupling of two solutions to (4.2.2); that is, formula (4.4.2) with L given by (4.4.4). Omitting the dependence on t, we have

$$\begin{aligned} &Lw(x,y) \\ &\quad = \operatorname{Tr}\big((\sigma\mathsf{I} + \beta\mathsf{J})\mathcal{L}[D^2u](x,y)\big) - \sum_{j\in[\![N]\!]} \big(D_ju^j(x)\cdot D_{x^j} + D_ju^j(y)\cdot D_{y^j}\big)\mathcal{L}[u](x,y)\,, \end{aligned}$$

and, according to (4.2.1),

$$\begin{aligned} \partial_t w(x,y) = &-\operatorname{Tr}\big((\sigma\mathsf{I} + \beta\mathsf{J})\mathcal{L}[D^2u](x,y)\big) + \frac{1}{2}\mathcal{L}[(|D_iu^i|^2)_{i\in[\![N]\!]}](x,y) \\ &+ \sum_{j\in[\![N]\!]\setminus\{i\}} \mathcal{L}[D_ju^j\cdot D_ju](x,y) - \mathcal{L}[f](x,y). \end{aligned}$$

Note that

$$\begin{aligned} &\sum_{j\in[\![N]\!]} \big(D_ju^j(x)\cdot D_{x^j} + D_ju^j(y)\cdot D_{y^j}\big)\mathcal{L}[u](x,y) \\ &\qquad = \sum_{i,j\in[\![N]\!]} D_ju^j(x)\cdot D_ju^i(x) - \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_ju^j(x)\cdot D_ju^i(x^{-i},y^i) \\ &\qquad\quad - \sum_{i\in[\![N]\!]} D_iu^i(x)\cdot D_iu^i(y^{-i},x^i) - \sum_{i\in[\![N]\!]} D_iu^i(y)\cdot D_iu^i(x^{-i},y^i) \\ &\qquad\quad - \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_ju^j(y)\cdot D_ju^i(y^{-i},x^i) + \sum_{i,j\in[\![N]\!]} D_ju^j(y)\cdot D_ju^i(y) \end{aligned}$$

and

$$\begin{aligned}\sum_{j\in[\![N]\!]\setminus\{i\}} &\mathcal{L}[D_j u^j \cdot D_j u](x,y) \\ &= \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j(x)\cdot D_j u^i(x) - \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j(x^{-i},y^i)\cdot D_j u^i(x^{-i},y^i) \\ &\quad - \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j(y^{-i},x^i)\cdot D_j u^i(y^{-i},x^i) + \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j(y)\cdot D_j u^i(y)\,.\end{aligned}$$

Then easy computations show that

$$\begin{aligned}&(\partial_t + L)w(x,y) \\ &= \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j\big|_{(x^{-i},y^i)}^{x}\cdot D_j u^i(x^{-i},y^i) + \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j\big|_{(y^{-i},x^i)}^{y}\cdot D_j u^i(y^{-i},x^i) \\ &\quad - \sum_{i\in[\![N]\!]} \big|D_i u^i(x) - D_i u^i(y^{-i},x^i)\big|^2 - \sum_{i\in[\![N]\!]} \big|D_i u^i(y) - D_i u^i(x^{-i},y^i)\big|^2 \\ &\quad - \mathcal{L}[f](x,y)\,.\end{aligned}$$

Exploiting the mean value theorem,

$$\begin{aligned}&\sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j\big|_{(x^{-i},y^i)}^{x}\cdot D_j u^i(x^{-i},y^i) + \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j\big|_{(y^{-i},x^i)}^{y}\cdot D_j u^i(y^{-i},x^i) \\ &= \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} \Big(D_j u^j\big|_{(x^{-i},y^i)}^{x} + D_j u^j\big|_{(y^{-i},x^i)}^{y}\Big)\cdot D_j u^i(x^{-i},y^i) \\ &\qquad\qquad + \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j\big|_{(y^{-i},x^i)}^{y}\cdot D_j u^i\big|_{(x^{-i},y^i)}^{(y^{-i},x^i)} \\ &= \sum_{\substack{i,k\in[\![N]\!]\\ k\neq i}} \int_0^1\!\!\int_0^1 \sum_{j\in[\![N]\!]\setminus\{i\}} D^3_{ijk}u^j(z^{-i}_{s'},z^i_s) D_j u^i(x^{-i},y^i)(x^i-y^i)\cdot(x^k-y^k)\,\mathrm{d}s\,\mathrm{d}s' \\ &\quad + \sum_{i,k\in[\![N]\!]} \int_0^1\!\!\int_0^1 \sum_{j\in[\![N]\!]\setminus\{i\}} D^2_{ij}u^j(y^{-i},z^i_s) D^2_{jk}u^i(z^{-i}_s,z^i_{1-s})(x^i-y^i)\cdot(x^k-y^k)\,\mathrm{d}s\,\mathrm{d}s',\end{aligned}$$

where $z_s := y + s(x-y)$ and products between tensors are understood in the sense of tensor contraction as discussed in the proof of Proposition 4.3.7. By the Cauchy–Schwarz inequality and assumption **(LS)**,

(4.5.7) $(\partial_t + L)w(x,y)$

$$\begin{aligned}&\leqslant \Bigg(\kappa_f + \Bigg(\sum_{i\in[\![N]\!]}\Big\|\sum_{\substack{j,k\in[\![N]\!]\\ j,k\neq i}} \big|D^3_{ijk}u^j\big|^2\Big\|_\infty \Big\|\sum_{j\in[\![N]\!]\setminus\{i\}} \big|D_j u^i\big|^2\Big\|_\infty\Bigg)^{\frac12} \\ &\quad + \Bigg(\sum_{i\in[\![N]\!]}\Big\|\sum_{j\in[\![N]\!]\setminus\{i\}} \big|D^2_{ij}u^j\big|^2\Big\|_\infty \Big\|\sum_{j\in[\![N]\!]\setminus\{i\}} \big|D(D_j u^i)\big|^2\Big\|_\infty\Bigg)^{\frac12}\Bigg)|x-y|^2\,.\end{aligned}$$

We can now proceed with the proof as we did for the one of Proposition 4.4.1. By formula (4.4.2), estimate (4.5.7), Proposition 4.5.1 and estimate (4.4.6), we obtain

$$w(\tau,\bar x,\bar y) \geqslant -e^{2(\kappa^*+\gamma)T}\Big(\kappa_g^* + T\kappa_f^* + \frac{K}{\sqrt{N}}\Big)|\bar x-\bar y|^2,$$

where the constant K depend only on σ, T, $\kappa^*+\gamma$, L_g and L_f. Then we conclude by arguing as in the proof of Proposition 4.4.1: pick $\kappa^* = (2T)^{-1}$, κ_g^* and κ_f^* small enough and N'_* large enough so that

$$\kappa_g^* + T\kappa_f^* + \frac{K}{\sqrt{N'_*}} \leqslant \frac{\kappa^*}{2e^{2(\kappa^*+\gamma)T}} = \frac{1}{4Te^{2\gamma T+1}}. \qquad \blacksquare$$

PROPOSITION 4.5.4 (Improvement of the diagonal lower bound). *Let $\kappa>0$. Then there exists a constant γ^* depending only on σ, T, L_g and L_f (and not on κ) and a natural number N''_* depending only on σ, T, κ, L_g and L_f, such that for any $\tau\in[0,T)$ one has*

$$\left.\begin{array}{ll}\mathcal{L}[u]\geqslant -\kappa Q & \text{on } (\tau,T]\times(\mathbb{R}^d)^N\\ \triangle_u \geqslant -\gamma^*\mathsf{I} & \text{on } (\tau,T]\times(\mathbb{R}^d)^N\\ N\geqslant N''_* & \end{array}\right\} \implies \triangle_{u(\tau,\,\cdot\,)} \geqslant -\frac12\gamma^*\mathsf{I} \quad \text{on } (\mathbb{R}^d)^N.$$

The explicit constant γ^* can be found below, see (4.5.12).

PROOF OF PROPOSITION 4.5.4. We are going to use an argument "along optimal trajectories", but without doubling variables/coupling. Consider $v = D_i u^i$; it satisfies the $\mathbb{R}^d$-valued equation[2]

$$(4.5.8)\quad -\partial_t v - \mathrm{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2 v) + \sum_{j\in[\![N]\!]} D_j u^j\cdot D_j v = D_i f^i - \sum_{j\in[\![N]\!]\setminus\{i\}} D^2_{ij}u^j D_j u^i\,.$$

Fix $x\in(\mathbb{R}^d)^N$, $s\in[\tau,T)$ and let ρ be a flow of probability measures on $(\mathbb{R}^d)^N$ solving

$$(4.5.9)\qquad \begin{cases}\partial_t\rho - \mathrm{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2\rho) + \mathrm{div}(B\rho) = 0 & \text{on } (s,T)\times(\mathbb{R}^d)^N\\ \rho(s) = \delta_x,\end{cases}$$

where $B = (-D_j u^j)_{1\leqslant j\leqslant N}$. Duality between equations (4.5.8) and (4.5.9) gives

$$D_i u^i(s,x) = \int_{(\mathbb{R}^d)^N} D_i g^i\,\mathrm{d}\rho(T) + \int_s^T\int_{(\mathbb{R}^d)^N} D_i f^i\,\mathrm{d}\rho - \int_s^T\int_{(\mathbb{R}^d)^N}\sum_{j\in[\![N]\!]\setminus\{i\}} D^2_{ij}u^j D_j u^i\,\mathrm{d}\rho\,,$$

[2] We point out that we use the natural notation $\mathrm{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2 v)$ to denote the vector with coordinates

$$\sum_{j,k\in[\![N]\!],\ a,b\in[\![d]\!]} (\sigma\mathsf{I}+\beta\mathsf{J})_{dj+a,dk+b} D^3_{x^{kb}x^{ja}x^{i\ell}} v^i\,,\quad \ell\in[\![d]\!]\,.$$

hence by assumption **(MF)**, the Cauchy–Schwarz inequality, estimates (4.5.2) and (4.5.4) and the arbitrariness of $s \in [\tau, T)$,

$$\|D_i u^i(t,\cdot)\|_\infty \leqslant L_g + TL_f + KN^{-1} \quad \forall\, t \in [\tau, T], \tag{4.5.10}$$

where the constant K depends only on σ, T, κ, γ^* (that will be chosen below), L_g and L_f.

Letting $w = \frac{1}{2}|D_i u^i|^2$, which satisfies

$$\begin{aligned} &- \partial_t w - \mathrm{Tr}((\sigma \mathsf{I} + \beta \mathsf{J})D^2 w) + \sum_{j\in[\![N]\!]} D_j u^j \cdot D_j w \\ &\quad + \sigma \sum_{j\in[\![N]\!]} |D^2_{ij} u^i|^2 + \beta \Big| \sum_{j\in[\![N]\!]} D^2_{ij} u^i \Big|^2 = D_i u^i \cdot D_i f^i - \sum_{j\in[\![N]\!]\setminus\{i\}} D^2_{ij} u^j D_j u^i \cdot D_i u^i\,, \end{aligned}$$

again by duality with (4.5.9) with $s = \tau$ one obtains

$$\begin{aligned} \frac{1}{2}|D_i u^i(\tau,x)|^2 + \sigma \int_\tau^T \int_{(\mathbb{R}^d)^N} \sum_{j\in[\![N]\!]} |D^2_{ij} u^i|^2 \,\mathrm{d}\rho + \beta \int_\tau^T \int_{(\mathbb{R}^d)^N} \Big| \sum_{j\in[\![N]\!]} D^2_{ij} u^i \Big|^2 \mathrm{d}\rho& \\ = \int_{(\mathbb{R}^d)^N} \frac{1}{2}|D_i g^i|^2 \,\mathrm{d}\rho(T) + \int_\tau^T \int_{(\mathbb{R}^d)^N} D_i u^i \cdot D_i f^i \,\mathrm{d}\rho& \\ - \int_\tau^T \int_{(\mathbb{R}^d)^N} \sum_{j\in[\![N]\!]\setminus\{i\}} D^2_{ij} u^j D_j u^i \cdot D_i u^i \,\mathrm{d}\rho\,.& \end{aligned}$$

Therefore, using assumption **(MF)**, the Cauchy–Schwarz inequality and estimates (4.5.2), (4.5.4) and (4.5.10), one obtains in particular

$$\int_\tau^T \int_{(\mathbb{R}^d)^N} |D^2_{ii} u^i|^2 \,\mathrm{d}\rho \leqslant \frac{L_g^2}{2\sigma} + \frac{TL_f}{\sigma}(L_g + TL_f) + K' N^{-1}, \tag{4.5.11}$$

where the constant K' depends only on σ, T, κ, γ^*, L_g and L_f.

As a last step, $V = D^2_{ii} u^i$ solves the following $\mathscr{S}(d)$-valued equation

$$\begin{aligned} &- \partial_t V - \mathrm{Tr}((\sigma \mathsf{I} + \beta \mathsf{J})D^2 V) + \sum_{j\in[\![N]\!]} D_j u^j D_j V + V^2 \\ &\quad = D^2_{ii} f^i - \sum_{j\in[\![N]\!]\setminus\{i\}} D^2_{ij} u^i D^2_{ji} u^j - \sum_{j\in[\![N]\!]\setminus\{i\}} D^2_{ij} u^j D^2_{ji} u^i - \sum_{j\in[\![N]\!]\setminus\{i\}} D^3_{iij} u^j D_j u^i\,; \end{aligned}$$

thus duality with equation (4.5.9) with $s = \tau$ gives

$$\begin{aligned} D^2_{ii} u^i(\tau,x) + \int_\tau^T \int_{(\mathbb{R}^d)^N} (D^2_{ii} u^i)^2 \,\mathrm{d}\rho - \int_{(\mathbb{R}^d)^N} D_{ii} g^i \,\mathrm{d}\rho(T) - \int_\tau^T \int_{(\mathbb{R}^d)^N} D^2_{ii} f^i \,\mathrm{d}\rho& \\ = \int_\tau^T \int_{(\mathbb{R}^d)^N} \sum_{\substack{1\leqslant j\leqslant N \\ j\neq i}} \big(D^2_{ij} u^i D^2_{ji} u^j + D^2_{ij} u^j D^2_{ji} u^i + D^3_{iij} u^j D_j u^i \big) \,\mathrm{d}\rho\,.& \end{aligned}$$

By assumption **(MF)**, the Cauchy–Schwarz inequality, the estimates in Proposition 4.5.1 and estimate (4.5.11), we get

$$D^2_{ii}u^i(\tau,x) \geqslant -\Big(L_g + L_f + \frac{L_g^2}{2\sigma} + \frac{TL_f}{\sigma}(L_g + TL_f) + \frac{K'}{N} + \frac{K''}{\sqrt{N}}\Big)I_d,$$

where the constant K'' depends only on σ, T, κ, γ^*, L_g and L_f. At this point one sees that if

$$\gamma^* = 4\Big(L_g + L_f + \frac{L_g^2}{2\sigma} + \frac{TL_f}{\sigma}(L_g + TL_f)\Big) \tag{4.5.12}$$

then for any N large enough so that $K'N^{-1} + K''N^{-\frac{1}{2}} \leqslant \frac{1}{4}\gamma^*$ one has $D_{ii}u^i(\tau,x) \geqslant -\frac{1}{2}\gamma^* I_d$, uniformly with respect to $i \in [\![N]\!]$. The conclusion follows from the arbitrariness of $x \in (\mathbb{R}^d)^N$. ∎

For the benefit of the reader, we make explicit how Proposition 4.5.2 can be now obtained.

Proof of Proposition 4.5.2. Let κ^* be as in Proposition 4.5.3 and γ^* be as in Proposition 4.5.4 (note that these two do not depend on each other). Then, Proposition 4.5.3 applies with $\gamma = \gamma^*$, yielding the improvement of $\mathscr{L}$-semimonotonicity for $N \geqslant N'_*$. On the other hand, Proposition 4.5.4 applies with $\kappa = \kappa^*$, so that the diagonal lower bound improves for $N \geqslant N''_*$. Taking $N_* = \max\{N'_*, N''_*\}$ gives the assertion. ∎

Theorem 4.5.5 (Estimates on the Nash system with $\mathscr{L}$-semimonotone data). *Assume that f and g satisfy assumptions **(MF)** and **(LS)**. Let $T > 0$. There exist nonnegative constants κ_f^* and κ_g^* and a natural number N^*, all depending only on σ, T, L_f and L_g (in such a way that $\kappa_f^*, \kappa_g^* \to 0$ as $T \to +\infty$ and $\kappa_f^*, \kappa_g^* \to +\infty$ as $T \to 0$), such that if $\kappa_g \leqslant \kappa_g^*$, $\kappa_f \leqslant \kappa_f^*$ and $N \geqslant N^*$ then any solution u to (4.2.1) on $[0,T] \times (\mathbb{R}^d)^N$ satisfies*

$$\sup_{i\in[\![N]\!]}\bigg(\sup_{j\in[\![N]\!]\setminus\{i\}} \|D_j u^i\|_\infty + \bigg\|\sum_{j\in[\![N]\!]\setminus\{i\}} \big|D^2_{ij}u^j\big|^2\bigg\|_\infty + \sum_{j\in[\![N]\!]\setminus\{i\}} \|D(D_j u^i)\|^2_\infty\bigg) \lesssim \frac{1}{N}$$

and

$$\sup_{i\in[\![N]\!]} \|Du^i\|^2_\infty + \sup_{i\in[\![N]\!]} \sum_{j\in[\![N]\!]} \|D(D_i u^j)\|^2_\infty + \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} \|D(D_{ij}u^j)\|^2_\infty \lesssim 1,$$

where the implied constants depend only on σ, T, L_f and L_g. In addition, there exists $\kappa^* > 0$ (depending only on T) such that u is κ^*-$\mathscr{L}$-semimonotone on $[0,T]$.

Proof. The argument is very much analogous to the one of Theorem 4.4.2, so we omit it. ∎

4.6. Appendix

4.6.1. An existence and uniqueness theorem for the Nash system. For the sake of self-containedness, we present an existence theorem for (4.2.1).

THEOREM 4.6.1. *Let $f^i \in C^{\gamma+\alpha}((\mathbb{R}^d)^N)$ and $g^i \in C^{2+\gamma+\alpha}((\mathbb{R}^d)^N)$, for some integer $\gamma \geqslant 1$ and $\alpha \in (0,1)$. Then for any $T > 0$ there exists a unique bounded solution to the Nash system (4.2.1) which is of class $C^{1+\frac{\gamma}{2}+\frac{\alpha}{2},2+\gamma+\alpha}$.*

PROOF. *Step 1.* For $\tau \in [0,T)$, let $\mathcal{X} := \mathcal{X}_\tau = C([\tau,T]; C_b^1((\mathbb{R}^d)^N))^N$, with norm $\|v\|_{\mathcal{X}} := \sup_{i\in[\![N]\!]}(\|v^i\|_\infty + \|Dv^i\|_\infty)$. For $v \in \mathcal{X}$, let $u^i = \mathcal{S}(v)^i$ be the solution of the parabolic equation

$$\begin{cases} -\partial_t u^i - \operatorname{Tr}((\sigma \mathsf{I} + \beta \mathsf{J})D^2 u^i) = f^i - \dfrac{1}{2}|D_i v^i|^2 - \displaystyle\sum_{\substack{1\leqslant j\leqslant N\\ j\neq i}} D_j v^j \cdot D_j v^i \\ u^i|_{t=T} = g^i. \end{cases}$$

By [69, Theorem 5.1.2], $\mathcal{S}(v) \in \mathcal{X}$ (solutions actually enjoy further Hölder regularity). We have the following standard representation formula: let Φ be the fundamental solution to the diffusion equation $\partial_t u - \operatorname{Tr}((\sigma\mathsf{I} + \beta\mathsf{J})D^2 u) = 0$ on $[0,T] \times (\mathbb{R}^d)^N$; that is,

$$\begin{aligned} \Phi(t,x) &:= \frac{1}{\sqrt{\det(4\pi(\sigma\mathsf{I} + \beta\mathsf{J})t)}}\, e^{-\frac{(\sigma\mathsf{I}+\beta\mathsf{J})^{-1}x\cdot x}{4t}} \\ &= \frac{1}{\sqrt{1+Nd\beta/\sigma}}\,\frac{1}{(4\pi\sigma t)^{\frac{Nd}{2}}}\, e^{-\frac{\left(\mathsf{I}-\frac{\beta}{\sigma+Nd\beta}\mathsf{J}\right)x\cdot x}{4\sigma t}}\,. \end{aligned}$$

Then for $v \in \mathcal{X}$ and $i \in \{1,\dots,N\}$, we have

$$\mathcal{S}(v)^i(t,x) := (\Phi(T-t,\cdot) \star g^i)(x) + \int_t^T (\Phi(T-s,\cdot) \star (f^i - w^i(s-t,\cdot))(x)\,\mathrm{d}s,$$

where $\star$ is the convolution operator over $(\mathbb{R}^d)^N$ and $w^i := \frac{1}{2}|D_i v^i|^2 + \sum_{j\neq i} D_j v^j \cdot D_j v^i$. Then $\|\mathcal{S}(v)^i\|_\infty \lesssim \|g\|_{\mathcal{X}} + (T-\tau)(\|f\|_{\mathcal{X}} + \|v\|_{\mathcal{X}}^2)$, where the implied constant depends only on N and d. Also, for $j \in [\![N]\!]$ and $k \in [\![d]\!]$

$$\begin{aligned} D_{x^{jk}}\mathcal{S}(v)^i(t,x) &= (\Phi(T-t,\cdot) \star D_{x^{jk}} g^i)(x) + \int_t^T (\Phi(T-s,\cdot) \star D_{x^{jk}} f^i)(x)\,\mathrm{d}s \\ &\quad - \int_t^T (D_{x^{jk}}\Phi(T-s,\cdot) \star w^i(s-t,\cdot))(x)\,\mathrm{d}s\,; \end{aligned}$$

then, using that $\int_{(\mathbb{R}^d)^N} |D_{x^{jk}}\Phi(t,y)|\,\mathrm{d}y \leqslant Nd/\sqrt{\pi\sigma t}$, one has $\|D\mathcal{S}(v)^i\|_\infty \lesssim \|g\|_{\mathcal{X}} + (T-\tau)\|f\|_{\mathcal{X}} + \sqrt{T-\tau}\,\|v\|_{\mathcal{X}}^2$, the implied constant depending only on σ, N and d. Furthermore, given $v, \bar{v} \in \mathcal{X}$ similar computations show that $\|\mathcal{S}(v) - \mathcal{S}(\bar{v})\|_{\mathcal{X}} \lesssim (1 + \sqrt{T-\tau})\sqrt{T-\tau}\max\{\|v\|_{\mathcal{X}}, \|\bar{v}\|_{\mathcal{X}}\}\|v - \bar{v}\|_{\mathcal{X}}$, with implied constant depending only on σ, N and d. Letting $\mathcal{B}_c := \{v \in \mathcal{X} : \ \|v\|_{\mathcal{X}} \leqslant c\}$, we deduce that $\mathcal{S}\colon \mathcal{B}_c \to \mathcal{B}_c$ is a contraction for any $c = c(\|g\|_{\mathcal{X}}, \sigma, N, d)$ sufficiently large and $\tau = \tau(c, \sigma, N, d, \beta, \|f\|_{\mathcal{X}})$ such that $T - \tau$ is sufficiently small. It follows by the

Banach–Caccioppoli theorem that $\mathcal{S}$ has a unique fixed point $u \in \mathcal{B}_c$, which by the definition of $\mathcal{S}$ solves system (4.2.1) in the mild sense. Then note that by standard parabolic Schauder estimates we can bootstrap up to the desired regularity.

Step 2. In order to see that such a solution extends to the whole time horizon $[0,T]$, we consider the following a priori estimates, that hold for classical solutions that are C^3 in space and C^1 in time (therefore they will apply to the short-time solutions obtained in the previous step, since $\gamma \geqslant 1$). Recall the representation formula for the solution of the Nash system: for any $(s,z) \in [\tau,T] \times (\mathbb{R}^d)^N$,

$$u^i(s,z) = \int_s^T \int_{(\mathbb{R}^d)^N} \Big(\frac{1}{2}|D_i u^i(t,x)|^2 + f^i(x)\Big)\rho(t,\mathrm{d}x)\mathrm{d}t + \int_{(\mathbb{R}^d)^N} g^i(x)\,\rho(T,\mathrm{d}x)\,, \tag{4.6.1}$$

where $\rho = \rho^{s,z}$ solves $\partial_t \rho - \mathrm{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2\rho) + \sum_{j\in[\![N]\!]} \mathrm{div}_{x^j}(D_j u^j \rho) = 0$ on $[s,T]\times(\mathbb{R}^d)^N$ with $\rho|_{t=s} = \delta_z$. This is easily obtained by testing the equation of ρ by u^i. In particular,

$$u^i(s,z) \geqslant \int_s^T \int_{(\mathbb{R}^d)^N} f^i(y)\,\rho(t,\mathrm{d}x)\mathrm{d}t + \int_{(\mathbb{R}^d)^N} g^i(x)\,\rho(T,\mathrm{d}x).$$

On the other hand, since $-\partial_t u^i - \mathrm{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2u^i) + \sum_{j\in[\![N]\!]\setminus\{i\}} D_j u^j \cdot D_j u^i \leqslant f^i$, if $\hat\rho^i$ solves instead $\partial_t\hat\rho^i - \mathrm{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2\hat\rho^i) + \sum_{j\in[\![N]\!]\setminus\{i\}} \mathrm{div}_{x^j}(D_j u^j \hat\rho^i) = 0$ then

$$u^i(s,z) \leqslant \int_s^T \int_{(\mathbb{R}^d)^N} f^i(x)\,\hat\rho^i(t,\mathrm{d}x)\mathrm{d}t + \int_{(\mathbb{R}^d)^N} g^i(x)\,\hat\rho^i(T,\mathrm{d}x).$$

Then $\sup_{t\in[\tau,T]}\|u^i(t,\cdot)\|_\infty \leqslant (T+1)\max\{\|f^i\|_\infty, \|g^i\|_\infty\}$.

We now want to control the gradient of u^i. As a preliminary step, we show that the quantity

$$\mathfrak{c} = \mathfrak{c}(s,z) := \int_s^T \int_{(\mathbb{R}^d)^N} \sum_{i,j\in[\![N]\!]} |D_j u^i(t,\cdot)|^2\,\mathrm{d}\rho(t)\mathrm{d}t,$$

can be bounded uniformly in (s,z). Note indeed that $v^i := \frac{1}{2}|u^i|^2$ satisfies

$$-\partial_t v^i - \mathrm{Tr}((\sigma\mathsf{I}+\beta\mathsf{J})D^2v^i) + \sigma\sum_{j\in[\![N]\!]}|D_j u^i|^2 + \beta\Big|\sum_{j\in[\![N]\!]} D_j u^i\Big|^2 + \frac{1}{2}D_i u^i D_i v^i + \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} D_j u^j D_j v^i = f^i u^i\,,$$

so that by testing the equation of ρ by v^i one gets

$$\sigma \int_\tau^T \int_{(\mathbb{R}^d)^N} \sum_{1\leqslant j\leqslant N} |D_j u^i(t,\cdot)|^2 \,\mathrm{d}\rho(t)\mathrm{d}t$$
$$\leqslant \int_{(\mathbb{R}^d)^N} \frac{1}{2}|g^i|^2 \,\mathrm{d}\rho(T) + \int_\tau^T \int_{(\mathbb{R}^d)^N} \frac{1}{2}|D_i u^i(t,\cdot)|^2 u^i(t,\cdot)\,\mathrm{d}\rho(t)\mathrm{d}t$$
$$+ \int_\tau^T \int_{(\mathbb{R}^d)^N} f^i u^i(t,\cdot)\,\mathrm{d}\rho(t)\mathrm{d}t\,.$$

Then, letting $A := \max\{\|f^i\|_\infty, \|g^i\|_\infty\}$ and noting that (4.6.1) yields

$$\int_\tau^T \int_{(\mathbb{R}^d)^N} \frac{1}{2}|D_i u^i(t,\cdot)|^2 \,\mathrm{d}\rho(t)\mathrm{d}t \leqslant 2(T+1)A,$$

we obtain

$$\mathfrak{c} \leqslant \bar{\mathfrak{c}} := \Big(\frac{1}{2} + T(T+1)(2T+3)\Big)\frac{NA^2}{\sigma}. \tag{4.6.2}$$

Consider now that $\tilde{w} := \sum_{i,j\in[\![N]\!]} |D_j u^i|^2$ solves

$$-\partial_t\tilde{w} - \mathrm{Tr}((\sigma\mathsf{I} + \beta\mathsf{J})D^2\tilde{w}) + \sigma \sum_{i,j,k\in[\![N]\!]} |D^2_{jk}u^i|^2 + \beta\Big|\sum_{i,j,k\in[\![N]\!]} D^2_{jk}u^i\Big|^2$$
$$+ \sum_{k\in[\![N]\!]} D_k u^k D_k \tilde{w} = \sum_{i,j\in[\![N]\!]} D_j f^i D_j u^i - \sum_{\substack{i,j,k\in[\![N]\!]\\ k\neq i}} D^2_{kj}u^k D_k u^i D_j u^i\,.$$

By testing the equation of ρ by w, using Young and Hölder inequality, we deduce that

$$\|\tilde{w}(s,\cdot)\|_\infty \leqslant \frac{1}{2}\sum_{i,j\in[\![N]\!]} \int_{(\mathbb{R}^d)^N} |D_j g^i|^2 \,\mathrm{d}\rho(T) + \frac{1}{2}\sum_{i,j\in[\![N]\!]} \int_s^T \int_{(\mathbb{R}^d)^N} |D_j f^i|^2 \,\mathrm{d}\rho(t)\mathrm{d}t$$
$$+ \int_s^T \left(1 + \frac{1}{\sigma}\int_{(\mathbb{R}^d)^N} \sum_{\substack{i,j\in[\![N]\!]\\ j\neq i}} |D_j u^i(t,\cdot)|^2 \,\mathrm{d}\rho(t)\right)\|\tilde{w}(t,\cdot)\|_\infty \,\mathrm{d}t\,.$$

(For the sake of accuracy, we point out that one should consider $\rho^{s,z}$ such that $\tilde{w}(s,z) \geqslant \|\tilde{w}(s,\cdot)\|_\infty - \varepsilon$ for $\varepsilon > 0$, and let $\varepsilon \to 0$ in the following computations.) Whence, for

$$s \geqslant s^* := T - (2(T + \sigma^{-1}\bar{\mathfrak{c}}))^{-1} \tag{4.6.3}$$

one has $\sup_{t\in[s^*\vee 0,T]}\|\tilde{w}(t,\cdot)\|_\infty \leqslant C$, where C depends only on $\max_{ij}\|D_j f^i\|_\infty$, $\max_{ij}\|D_j g^i\|_\infty$, σ, N and T. Then by iterating the estimate a finite number of times κ, one obtains $\sup_{t\in[\tau,T]}\|\tilde{w}(t,\cdot)\|_\infty \leqslant 2^\kappa C$.

This is sufficient to conclude that $\|u\|_{\mathcal{X}_\tau} \leqslant C'$ for some constant C' independent of τ. Define at this point

$$\tau_* := \inf\{\tau \in [0,T] :\ (4.2.1) \text{ has a unique solution belonging to } \mathcal{X}_\tau\}\,.$$

We proved that $\tau_* < T$; seeking for a contradiction, suppose that $\tau_* > 0$. Let $\varepsilon \in (0, T-\tau_*)$ to be announced and consider a solution $u \in \mathcal{X}_{\tau_*+\varepsilon}$. We have proved that $\|u\|_{\mathcal{X}_{\tau_*+\varepsilon}} \leqslant C'$. Then one can redo the initial fixed point argument with $u(\tau_*+\varepsilon,\cdot)$ in lieu of g and conclude that u extends to the horizon $[\tau', T]$, provided that $\tau_* + \varepsilon - \tau' \leqslant \delta$ for some $\delta > 0$ which is independent of τ_* and ε. So one can fix $\delta_0 = \frac{\tau_*}{2} \wedge \delta$, $\varepsilon = \frac{1}{2}(\delta_0 \wedge (T-\tau_*))$ and $\tau' = \tau_* - \frac{\delta_0}{2} \in (0, \tau_*)$, thus contradicting the definition of τ_* ∎

4.6.2. An estimate on the space-time regularity for the heat equation. The Lemma below shows how to transfer weighted Lipschitz regularity in space of solution to heat equations to Hölder regularity in time, following a standard idea proposed in [**56**]. We state it for forward equations for simplicity. Note again that Hölder continuity is independent of N.

LEMMA 4.6.2. *Let $f : [0,T] \times (\mathbb{R}^d)^N \to \mathbb{R}$, $\sigma > 0$, $\beta \geqslant 0$, and u be a classical solution of*

$$\partial_t u - \mathrm{Tr}((\sigma \mathsf{I} + \beta \mathsf{J}) D^2 u) = f \qquad \text{in } (0,T) \times (\mathbb{R}^d)^N \ .$$

Assume also that

$$|u(t,x) - u(t,y)| \leqslant c_1 \|x-y\|^i \qquad \forall\, t \in [0,T],\ x, y \in (\mathbb{R}^d)^N,$$

and $\|f\|_\infty, \|u\|_\infty \leqslant c_2$ for some $c_1, c_2 > 0$ and $i \in [\![N]\!]$. Then,

$$|u(\tau,z) - u(s,z)| \leqslant (c_1 + 4(\sigma+\beta)dc_2)|\tau - s|^{\frac{1}{3}} + c_2|\tau - s|$$

for all $\tau, s \in [0,T]$ and $z \in (\mathbb{R}^d)^N$.

PROOF. Assume first that $s = 0$, $z = 0$ and $u(0,0) = 0$. Let $\rho > 0$ and

$$v(t,x) = c_1\rho + \frac{4(\sigma+\beta)dc_2}{\rho^2}\, t + c_2 t + \frac{c_2}{\rho^2}(\|x\|^i)^2.$$

Since

$$\partial_t v - \mathrm{Tr}((\sigma \mathsf{I} + \beta \mathsf{J}) D^2 v) - f = \frac{4(\sigma+\beta)dc_2}{\rho^2} + c_2 - \frac{(\sigma+\beta)dc_2}{\rho^2}\Big(4 - \frac{2}{N}\Big) - f \geqslant 0$$

everywhere, and

$$\begin{aligned} v(0,x) &\geqslant c_1\rho \geqslant c_1\|x\|^i \geqslant u(0,x) \qquad \text{on } \{\|x\|^i \leqslant \rho\}, \\ v(t,x) &\geqslant \frac{c_2}{\rho^2}(\|x\|^i)^2 \geqslant u(t,x) \qquad \text{on } [0,T] \times \{\|x\|^i = \rho\}, \end{aligned}$$

by the comparison principle we get $u(t,x) \leqslant v(t,x)$ on $[0,T] \times \{\|x\|^i \leqslant \rho\}$. For $\tau > 0$, pick then $\rho = \tau^{\frac{1}{3}}$, so that the previous inequality yields

$$u(\tau, 0) \leqslant c_1 \tau^{\frac{1}{3}} + 4(\sigma+\beta)dc_2\tau^{\frac{1}{3}} + c_2\tau.$$

Using $-v$ as a subsolution yields the analogous bound from below. It is now straightforward to conclude the estimate for general s, z. ∎

4.6.3. An alternative proof of Lemma 4.3.1. Since we used explicitly the more probability-oriented technique of synchronous coupling to prove Propositions 4.4.1 and 4.5.3, and we did not give all the details about the application of the maximum principle in the proof of Lemma 4.3.1, we take the occasion to briefly write here an alternative proof of that lemma, based on *coupling by reflection*. This "bypasses" (or "hides") the use of the maximum principle by relying on Itô's formula. In a similar manner, the proof of Lemma 4.3.3 can also be carried out by exploiting a synchronous coupling.

For the benefit of the reader, let us recall the basis we need of coupling by reflection. Assume that the diffusion coefficient σ in (4.4.3) is constant and invertible. Consider X and Y which solve, respectively, (4.4.3) with $X_{\bar t} = \bar x$ and

$$\begin{cases} \mathrm{d}Y_t = \mathbb{1}_{t<\theta}\big(\alpha(t,Y_t)\,\mathrm{d}t + \sigma(\mathsf{I} - 2\mathsf{P}_{\sigma^{-1}(X_t-Y_t)})\,\mathrm{d}\tilde B_t\big) + \mathbb{1}_{t\geqslant\theta}\,\mathrm{d}X_t, \quad t\in[\bar t,T] \\ Y_{\bar t} = \bar y, \end{cases}$$

where $\theta := \inf\{t \geqslant \bar t : X_t = Y_t\}$ is the first hitting time for $Z = (X,Y)^\mathsf{T}$ of the diagonal of $(\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$, and $\mathbb{1}_{t<\theta}$ and $\mathbb{1}_{t\geqslant\theta}$ are shorthands for the *indicator functions* of the intervals $(-\infty,\theta)$ and $[\theta,+\infty)$, respectively, computed at t. Also, recall that, according to the notation we adopted,

$$\mathsf{P}_{\sigma^{-1}(X_t-Y_t)} = \frac{\sigma^{-1}(X_t-Y_t)}{|\sigma^{-1}(X_t-Y_t)|} \otimes \frac{\sigma^{-1}(X_t-Y_t)}{|\sigma^{-1}(X_t-Y_t)|}\,.$$

If $Z_{\bar t} \sim \mu$ having marginals $\delta_{\bar x}$ and $\delta_{\bar y}$, then it is called a *coupling by reflection* of two solution to (4.4.3) starting at $\bar x$ and $\bar y$ (cf., e.g., [**40**]). If $T < \theta$, Z solves equation (4.4.1) with

$$b(t,x,y) = (\alpha(t,x),\alpha(t,y))^\mathsf{T}, \qquad \Sigma(t,x,y) = (\sigma, \sigma(\mathsf{I} - 2\mathsf{P}_{\sigma^{-1}(x-y)}))^\mathsf{T}$$

thus

$$b\cdot D_z = \alpha(t,x)\cdot D_x + \alpha(t,y)\cdot D_y$$

and

$$\begin{aligned} \Sigma\Sigma^\mathsf{T} &= J_2 \otimes \sigma\sigma^\mathsf{T} - 2\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\sigma\mathsf{P}_{\sigma^{-1}(x-y)}\sigma^\mathsf{T} \\ &= J_2 \otimes \sigma\sigma^\mathsf{T} - 2\,\frac{|x-y|^2}{|\sigma^{-1}(x-y)|^2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\mathsf{P}_{x-y}\,. \end{aligned}$$

In order to adapt this setting to that of system (4.2.2) where a common noise is present, we add a variable a consider, in lieu of (4.4.3), the $(\mathbb{R}^d)^{N+1}$-valued SDE $\mathrm{d}\hat X_t = \hat\alpha(t,\hat X_t)\,\mathrm{d}t + \hat\sigma\,\mathrm{d}\tilde B_t$, with

$$\hat\alpha = (\alpha^*, 0), \qquad \tilde B = (B,W)^\mathsf{T}, \qquad \hat\sigma = \sqrt{2}\left(\begin{array}{c|c} \mathsf{I} & \sqrt{\beta}\,1_N \otimes I_d \\ \hline 0 & 1 \end{array}\right).$$

Letting $\hat{x} = (x, x')$, $\hat{y} = (y, y')$ and considering the starting points $\bar{\hat{x}} = (\bar{x}, 0)$ and $\bar{\hat{y}} = (\bar{x}, 0)$, the corresponding coupling by reflection $\hat{Z}$ is such that

$$\begin{cases} \mathrm{d}(X' - Y')_t = 2\, \dfrac{(X' - Y')_t}{|\hat{\sigma}^{-1}(\hat{X}_t - \hat{Y}_t)|}\, \mathrm{d}\displaystyle\int_0^t \frac{\sigma^{-1}(\hat{X}_t - \hat{Y}_t)}{|\sigma^{-1}(\hat{X}_t - \hat{Y}_t)|} \otimes \mathrm{d}\tilde{B}_s, \quad t \in [\bar{t}, T] \\ (X' - Y')_{\bar{t}} = 0, \end{cases}$$

hence $\hat{Z}_t = ((X_t, 0), (Y_t, 0))$ for all $t \in [\bar{t}, T]$. We deduce that for $h = h(t, x, y)$ formula (4.4.2) reads

$$\mathbb{E}[h(T, X_{T\wedge\theta}, Y_{T\wedge\theta})] = h(\bar{t}, \bar{x}, \bar{y}) + \int_{\bar{t}}^{T\wedge\theta} \mathbb{E}[\partial_t h(t, X_t, Y_t) + Lh(t, X_t, Y_t)]\, \mathrm{d}t \tag{4.6.4}$$

with

$$\begin{aligned} L = \mathrm{Tr}\big((J_2 \otimes (\mathsf{I} + \beta \mathsf{J}))D^2_{(x,y)}\big) - \frac{4|x-y|^2}{|x-y|^2_\beta}\, \mathrm{Tr}\big(\mathsf{P}_{x-y} D^2_{xy}\big) \\ + \alpha^*(t,x) \cdot D_x + \alpha^*(t,y) \cdot D_y\,, \end{aligned} \tag{4.6.5}$$

where $|\cdot|_\beta$ is the norm given by the metric induced by the matrix $\mathsf{I} + \beta\mathsf{J}$. We can refer to such an operator L as the differential generator of the coupling by reflection of two solution to (4.2.2).

Alternative proof of Lemma 4.3.1. Let w and ψ be as in the former proof of Lemma 4.3.1. We are going to use formula (4.4.2) where $h = w$ and $Z = (X, Y)$ is a coupling by reflection of two solutions to (4.2.2); that is, formula (4.6.4) with L given by (4.6.5). The same computations as in the former proof of Lemma 4.3.1, along with the fact that $|\cdot|_\beta \geqslant |\cdot|$, yield

$$\begin{aligned} &(\partial_t + L)w(t,x,y) \\ &\quad \geqslant -4\psi(t)\varphi''(|x-y|) - M\psi(t)\varphi'(|x-y|)|x-y| - \psi'(t)\varphi(|x-y|) - 2\|F\|_\infty \\ &\quad > 0\,. \end{aligned}$$

Then formula (4.6.4) gives

$$w(\bar{t}, \bar{x}, \bar{y}) \leqslant \mathbb{E}[w(T\wedge\theta, X_{T\wedge\theta}, Y_{T\wedge\theta})] \leqslant \big(g|_y^x - \|DG\|_\infty |x-y|\big)_+ = 0$$

and the desired conclusion follows. ∎

CHAPTER 5

Generalised Mean Field Games. Infinite-population limit of the Nash system

Contents

5.1. Generalised Mean Field assumptions

In this last chapter, we show that the estimates of Theorem 4.2.7 allow to attack the convergence problem in the $N \to \infty$ limit. To be sure that, for any fixed N, solutions to the Nash system exist, we assume that f^i and g^i are of class $C^{2+\alpha}((\mathbb{R}^d)^N)$ and $C^{4+\alpha}((\mathbb{R}^d)^N)$ respectively, for every i and N. By Theorem 4.6.1, solutions u^i indeed exists and are unique, and belong to the class $C^{2+\frac{\alpha}{2},4+\alpha}$. We have then enough regularity to apply the a priori estimates obtained in the previous chapter.

We work, for the sake of simplicity, without common noise; that is, we set $\beta = 0$ in (4.2.2) and thus in (4.2.1). Then we set $\sigma = 1$.

Besides this reduction, we let $\Lambda := [0, 1]$ and we suppose that, for $h \in \{f, g\}$, the following symmetry assumption is fulfilled:

(S) in the N-dimensional Nash system, $h^i(x) = h_N(\lambda^i_N, x^i, x^{-i})$, for some *bounded* sequence of maps $h_N : \Lambda \times \mathbb{R}^d \times (\mathbb{R}^d)^{N-1} \to \mathbb{R}$, $N \in \mathbb{N}$, that are *symmetric* on $(\mathbb{R}^d)^{N-1}$, and for some $\lambda^i_N \in \Lambda$.

More explicitly, the assumption of symmetry means that for any $\sigma \in S_{N-1}$ (that is, σ which is a permutation on $[\![N-1]\!]$)[1] one has

$$h_N(\lambda, y, z^\sigma) = h_N(\lambda, y, z), \qquad z^\sigma := (z^{\sigma(0)}, \dots, z^{\sigma(N-2)}),$$

[1] Note that we are using now σ to denote both the diffusion coefficient in the Nash system and a permutation. Nevertheless, we believe this should not cause any confusion, also because, as we have said, in this part of the work the former is set to be 1 (and thus it will not even explicitly appear).

for all $\lambda \in \Lambda$, $y \in \mathbb{R}^d$ and $z \in (\mathbb{R}^d)^{N-1}$.

REMARK 5.1.1. The additional parameters λ_N^i in Assumption **(S)** provide a convenient way of modelling a situation where, although each player observes the empirical measure as in the standard MFG theory, different players can react differently. This generalises standard symmetric MFG interactions, which, on the other hand, are recovered by letting $\lambda_N^i = \lambda \in \Lambda$ for all N, i, so that $u^i(x) = u^1(x^i, x^{-i})$.

REMARK 5.1.2. A bounded and continuous function $h_N : (\mathbb{R}^d)^N \to \mathbb{R}$ that is symmetric on its domain can be seen as a "finite-dimensional projection" of the function $\mathfrak{h}_N : \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}$. Indeed, suppose that h_N has a modulus of continuity ω with respect to the Euclidean distance (that is, $|h_N(x) - h_N(y)| \leqslant \omega(|x-y|)$), and define

$$\mathfrak{h}_N(m) := \inf_{x \in (\mathbb{R}^d)^N} \{h_N(x) + \tilde{\omega}(W_p(m_x, m))\},$$

with $\tilde{\omega} := \omega(N\,\cdot)$. It is known (cf. [**74**, Theorem 2.1]) that there exists $\sigma \in S_N$ such that

$$W_p(m_x, m_y) = N^{-\frac{1}{p}} \Bigg(\sum_{i \in [\![N]\!]} \big|x^i - y^{\sigma(i)}\big|^p \Bigg)^{\frac{1}{p}} = N^{-\frac{1}{p}} |x - y^\sigma|_p \quad \forall\, x, y \in (\mathbb{R}^d)^N,$$

hence $|h_N(x) - h_N(y)| \leqslant \tilde{\omega}(W_p(m_x, m_y))$, by the symmetry of h_N and Hölder's inequality; then one readily checks that $h_N(x) = \mathfrak{h}_N(m_x)$.

With this additional structure on the data, we require *Lipschitz continuous* dependence on the *parameter* λ: for $h \in \{f, g\}$, $h = h(\lambda, y, z)$ as above,

(LP) there exist a $L_\Lambda > 0$, independent of N, such that

$$\|\partial_y^k h_N(\lambda, \cdot) - \partial_y^k h_N(\lambda', \cdot)\|_\infty \leqslant L_\Lambda |\lambda - \lambda'| \qquad \forall\, \lambda, \lambda' \in \Lambda,$$

for $k = 0$ if $h = f$ and $k \in \{0, 1\}$ if $h = g$.

Since we have introduced new hypotheses to the pool from Chapter 4, let us briefly discuss their implications in the MF scenario and their relationship with the assumptions of Theorem 4.2.7, which we wish to continue to hold.

EXAMPLE 5.1.3. Let $F, G : \Lambda \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be bounded, and consider $f_N(\lambda, y, z) = F(\lambda, y, m_z)$ and $g_N(\lambda, y, z) = G(\lambda, y, m_z)$, so that

$$f^i(x) = F(\lambda_N^i, x^i, m_{x^{-i}}), \qquad g^i(x) = G(\lambda_N^i, x^i, m_{x^{-i}});$$

that is, players' costs are taken from a pool of admissible symmetric costs, according to some label λ_N^i.

While **(S)** clearly holds and so does **(MF)** provided that the properties highlighted in Remark 4.2.6 are satisfied for any fixed λ (uniformly), **(LP)** requires Lipschitz continuity with respect to λ uniformly in the other variables. Summing up, F needs to satisfy

F and $D_x F$ are Lipschitz continuous with respect to the $(|\cdot|, |\cdot|, W_1)$ distance and so does G, for which we also require that $D_m G(\lambda, \cdot)$ and $D_x D_m G(\lambda, \cdot)$ are Lipschitz continuous with respect to the $(|\cdot|, W_1, |\cdot|)$ distance, uniformly in λ.

Finally, one can check the semimonotonicity assumption **(DS)** or **(LS)** by revising Remark 4.2.5, as follows. For any fixed $T > 0$, Theorem 4.2.7 gives some smallness condition on the semimonotonicity constant M_0.

- Suppose that, for any $\nu \in \mathcal{P}(\Lambda \times (\mathbb{R}^d)^2)$,

$$\int_\Lambda \int_{\mathbb{R}^d \times \mathbb{R}^d} \big(D_x F(\lambda, x, \pi^{(1)}_{\mathbb{R}^d \sharp}\nu) - D_x F(\lambda, y, \pi^{(2)}_{\mathbb{R}^d \sharp}\nu) + M(x - y)\big) \cdot (x - y)\, \nu(\mathrm{d}\lambda, \mathrm{d}x, \mathrm{d}y) \geqslant 0\,,$$

where $\pi^{(i)}_{\mathbb{R}^d} \colon \Lambda \times (\mathbb{R}^d)^2 \to \mathbb{R}^d$ is the projection onto the i-th copy of $\mathbb{R}^d$ ($i = 1, 2$). If $D_m F(\lambda, \cdot)$ is L-Lipschitz with respect to the $(|\cdot|, W_1, |\cdot|)$ distance, uniformly in λ, then f_N is $(M + 3L/N)$-$\mathscr{D}$-semimonotone. Therefore, if $M < M_0$, Theorem 4.2.7 applies (by requiring also the same inequality on G of course).

- Suppose that, for any $\mu_1, \mu_2 \in \mathcal{P}_2(\Lambda \times \mathbb{R}^d)$,

$$\int_{\Lambda \times \mathbb{R}^d} \big(F(\lambda, z, \pi_{\mathbb{R}^d \sharp}\mu_1) - F(\lambda, z, \pi_{\mathbb{R}^d \sharp}\mu_2)\big)(\mu_1 - \mu_2)(\mathrm{d}\lambda, \mathrm{d}z) \geqslant -M W_2(\pi_{\mathbb{R}^d \sharp}\mu_1, \pi_{\mathbb{R}^d \sharp}\mu_2)^2\,.$$

Then, arguing as in Remark 4.2.5, f_N is $(M + O(N^{-1}))$-$\mathscr{L}$-semimonotone. Hence, if $D_x D_m F$, $D^2_{mm} F$ are bounded uniformly in λ and $M < M_0$, Theorem 4.2.7 applies.

5.2. A limit value function

For the benefit of the following discussion, we explicitly recall that the estimates on the Nash system we obtained imply

$$\text{(5.2.1a)} \qquad \sup_{i \in [\![N]\!]} \|D_i u^i\|_\infty \lesssim 1\,, \qquad \sup_{i \in [\![N]\!]} \sup_{j \in [\![N]\!] \setminus \{i\}} \|D_j u^i\|_\infty \lesssim \frac{1}{N}\,,$$

$$\text{(5.2.1b)} \qquad \sup_{i \in [\![N]\!]} \|D^2_{ii} u^i\|_\infty \lesssim 1\,, \qquad \sup_{i \in [\![N]\!]} \sum_{j \in [\![N]\!] \setminus \{i\}} \|D_j D_i u^i\|^2_\infty \lesssim \frac{1}{N}\,.$$

In order to specify the choice of $\lambda \in \Lambda^N$ with which a Nash system is built, we use the notation u_λ to denote the solution to (4.2.1) on $[0, T] \times (\mathbb{R}^d)^N$ where $h^i = h_N(\lambda^i, \cdot)$ for $h \in \{f, g\}$, $i \in [\![N]\!]$. Then, we focus on the map

$$\mathfrak{u}_N \colon (\lambda, t, x) \mapsto u^0_\lambda(t, x),$$

which is representative of any solution of an N-dimensional Nash system of the form we are considering, as the following remark shows.

Remark 5.2.1. A solution u_λ is recovered by noticing that $u^i_\lambda(t, x) = \mathfrak{u}_N(\lambda^\sigma, t, x^\sigma)$ for any $\sigma \in S_N$ such that $\sigma(0) = i$, as by construction $\mathfrak{u}_N$ is symmetric on $(\Lambda \times (\mathbb{R}^d))^{N-1}$, in the sense that it is invariant under permutations of the pairs (λ^j, x^j), $j > 0$. Similarly, the optimal controls $D_i u^i$ are recovered from $D_0 \mathfrak{u}$. Heuristically, this means that switching two players does not affect the value functions, but the two corresponding labels need to be switched correspondingly in the Nash system.

As the first coordinate of $x \in (\mathbb{R}^d)^N$ and of $\lambda \in \Lambda^N$ have a distinct role for $\mathfrak{u}$ with respect to the other coordinates, a switch in our notation is now convenient: hereinafter, the typical element of $(\mathbb{R}^d)^N$ will be denoted by $\underline{x}$, and we will write $\underline{x} = (x, \hat{x}) \in \mathbb{R}^d \times (\mathbb{R}^d)^{N-1}$, thus using the notation x as a shortening for $\underline{x}^0$ (or, more generally, as a typical element of $\mathbb{R}^d$); analogously, we will have $\underline{\lambda} = (\lambda, \hat{\lambda}) \in \Lambda \times \Lambda^{N-1}$. We will write ∂_x for the derivative with respect to $\underline{x}^0$ (or x), and D_j for the derivative with respect to $\underline{x}^j = \hat{x}^{j-1} \in \mathbb{R}^d$, with $j \in \{1, \dots, N-1\}$.

We begin by showing that assumption **(LP)** reflects in a Lipschitz dependence of $\mathfrak{u}^N$ and $\partial_x \mathfrak{u}^N$ with respect to $\underline{\lambda} \in \Lambda^N$. We recall the notation

$$|\hat{\lambda} - \hat{\lambda}'|_1 := \sum_{i \in [\![N-1]\!]} |\hat{\lambda}^i - \hat{\lambda}'^i| = \sum_{j \in [\![N]\!] \setminus \{0\}} |\underline{\lambda}^j - \underline{\lambda}'^j| \,.$$

LEMMA 5.2.2. *For any $\underline{\lambda}, \underline{\lambda}' \in \Lambda^N$, one has*

$$\|\mathfrak{u}_N(\underline{\lambda}, \cdot) - \mathfrak{u}_N(\underline{\lambda}', \cdot)\|_\infty + \|\partial_x \mathfrak{u}_N(\underline{\lambda}, \cdot) - \partial_x \mathfrak{u}_N(\underline{\lambda}', \cdot)\|_\infty \lesssim |\lambda - \lambda'| + \frac{|\hat{\lambda} - \hat{\lambda}'|_1}{N} \,,$$

where the implied constant depends only on T, L_g, L_f and L_Λ.

PROOF. All implied constants appearing in this proof depend only on T, L_g and L_f. Let $\underline{\lambda}, \underline{\lambda}' \in \Lambda^N$. Consider $w^i := u^i_{\underline{\lambda}} - u^i_{\underline{\lambda}'}$ that solves the following equation in $(0,T) \times (\mathbb{R}^d)^N$:

$$\begin{aligned}(5.2.2) \qquad -\partial_t w^i - \Delta w^i + \sum_{j \in [\![N]\!]} D_j u^j_{\underline{\lambda}} \cdot D_j w^i \\ &= f_N\Big|^{(\lambda^i, \cdot)}_{(\lambda'^i, \cdot)} - \sum_{j \in [\![N]\!] \setminus \{i\}} D_j w^j \cdot D_j u^i_{\underline{\lambda}'} + \frac{1}{2}|D_i w^i|^2 \\ &\lesssim |\lambda^i - \lambda'^i| + \frac{1}{N} \sum_{j \in [\![N]\!] \setminus \{i\}} |D_j w^j| + |D_i w^i| \,,\end{aligned}$$

where the inequality is obtained by exploiting assumption **(LP)** and estimates (5.2.1a). By Lemma 4.3.1 we deduce that, for any $\tau \in [0,T]$,

$$\begin{aligned}(5.2.3) \qquad \sup_{t \in [\tau, T]} \|D_i w^i(t, \cdot)\|_\infty \lesssim (T-\tau)^{\frac{1}{2}} \Bigg(&|\lambda^i - \lambda'^i| + \sup_{t \in [\tau,T]} \|D_i w^i(t,\cdot)\|_\infty \\ &+ \frac{1}{N} \sum_{j \in [\![N]\!] \setminus \{i\}} \sup_{t \in [\tau,T]} \|D_j w^j(t,\cdot)\|_\infty \Bigg) .\end{aligned}$$

Since by summing over i we get

$$\sum_{i \in [\![N]\!]} \sup_{t \in [\tau,T]} \|D_i w^i(t,\cdot)\|_\infty \lesssim (T-\tau)^{\frac{1}{2}} \Bigg(\sum_{i \in [\![N]\!]} |\lambda^i - \lambda'^i| + \sum_{i \in [\![N]\!]} \sup_{t \in [\tau,T]} \|D_i w^i(t,\cdot)\|_\infty \Bigg),$$

possibly by iterating this estimate a finite number of times we obtain

$$\sum_{i \in [\![N]\!]} \sup_{t \in [0,T]} \|D_i w^i(t,\cdot)\|_\infty \lesssim \sum_{i \in [\![N]\!]} |\lambda^i - \lambda'^i| \,;$$

then plugging this back into (5.2.3) (and again possibly iterating the estimate) we have

$$\sup_{t\in[0,T]} \|D_i w^i(t,\cdot)\|_\infty \lesssim |\lambda^i - \lambda'^i| + \frac{1}{N}\sum_{i\in[\![N]\!]} |\lambda^i - \lambda'^i|,$$

which is equivalent to $\|\partial_x \mathfrak{u}_N(\underline{\lambda},\cdot) - \partial_x \mathfrak{u}_N(\underline{\lambda}',\cdot)\|_\infty \lesssim |\lambda - \lambda'| + N^{-1}|\lambda - \hat{\lambda}'|_1$. Now by the maximum principle one deduces that $\|w^1\|_\infty \lesssim |\lambda - \lambda'| + N^{-1}|\lambda - \hat{\lambda}'|_1$ as well, thus concluding the proof. ∎

REMARK 5.2.3. As a consequence of Lemma 5.2.2, by Rademacher's theorem $\underline{\lambda} \mapsto \partial_x \mathfrak{u}^N(\underline{\lambda}, t, \underline{x})$ is $\mathscr{L}^N$-a.e. differentiable on Λ^N for each $(t,x) \in [0,T]\times(\mathbb{R}^d)^N$, and we have $\sup_{j\in[\![N]\!]\setminus\{0\}} \|\partial_{\underline{\lambda}^j}\partial_x \mathfrak{u}\|_\infty \lesssim N^{-1}$, where the implied constant depends only on T, L_g, L_f and L_Λ. The same can be said for $\mathfrak{u}$ (clearly, with differentiability everywhere).

The following result will be instrumental in our study of the limit of the (solution to the) Nash system as $N\to\infty$; it adapts [**21**, Theorem 2.1].

LEMMA 5.2.4. *Let (Γ, d) be a metric space and let $(E, \|\cdot\|)$ be a finite-dimensional normed vector space. Let $(h_N)_{N\in\mathbb{N}}$ be a sequence of functions such that $h_N\colon \Gamma\times E^N \to \mathbb{R}$ is symmetric on E^N and there exist $C>0$ and concave moduli of continuity ω_Γ, ω_p, all independent of N, such that $\|h_N\|_\infty \leqslant C$ and, for some $p\geqslant 1$,*

$$|h_N(\gamma,v) - h_N(\gamma',v')| \leqslant \omega_\Gamma(d(\gamma,\gamma')) + \omega_p(W_p(m_v, m_{v'})) \tag{5.2.4}$$

for all $\gamma,\gamma'\in\Gamma$ and $v,v'\in E^N$. Then there exists $h\in C^0(\Gamma\times\mathcal{P}_p(E))$ such that, up to a subsequence, for any compact sets $H\subset\Gamma$ and $\mathcal{K}\subset\mathcal{P}_p(E)$,

$$\lim_{N\to\infty}\sup_{\gamma\in H}\sup_{v\in E^N\colon m_v\in\mathcal{K}} |h_N(\gamma,v) - h(\gamma,m_v)| = 0.$$

Furthermore, h has the same moduli of continuity with respect to $\gamma\in\Gamma$ and $m\in\mathcal{P}_p(E)$.

PROOF. Define $\mathfrak{h}_N\colon \Gamma\times\mathcal{P}_p(E)\to\mathbb{R}$ by

$$\mathfrak{h}_N(\gamma,m) := \inf_{v\in E^N}\big\{h_N(\gamma,v) + \omega_p(W_p(m_v,m))\big\}.$$

This is well-defined since h_N is bounded. By (5.2.4), $\mathfrak{h}_N(\gamma,m_v) = h_N(\gamma,v)$. Given $\gamma,\gamma'\in\Gamma$, $m,m'\in\mathcal{P}_p(E)$ and $v\in E^N$ which is ε-optimal in the definition of $\mathfrak{h}_N(\gamma',m')$,

$$\begin{aligned}
\mathfrak{h}_N(\gamma,m) &\leqslant h_N(\gamma,v) + \omega_p(W_p(m_v,m))\\
&\leqslant \mathfrak{h}_N(\gamma',m') + \omega_\Gamma(d(\gamma,\gamma')) + \omega_p(W_p(m_v,m)) - \omega_p(W_p(m_v,m')) + \varepsilon\\
&\leqslant \mathfrak{h}_N(\gamma',m') + \omega_\Gamma(d(\gamma,\gamma')) + \omega_p(W_p(m,m')) + \varepsilon.
\end{aligned}$$

We have used that ω_p is subadditive because it is concave and $\omega_p(0)=0$. Therefore, $(\mathfrak{h}_N)_{N\in\mathbb{N}}$ is a sequence of equicontinuous and equibounded functions, so a generalisation of the Ascoli–Arzelà theorem guarantees the existence of $h\in C^0(\Gamma\times\mathcal{P}_p(E))$

such that, up to a subsequence, $\mathfrak{h}_N \to h$ in the topology of compact convergence, that is our assertion. ∎

The importance of this lemma lies in the fact mentioned in [**21**, Remark 2.3] that if $E = \mathbb{R}^N$ and $\sup_{i\in[\![N]\!]}\|D_i h\|_\infty \lesssim N^{-1}$, then $x \mapsto h(\gamma, x)$ is Lipschitz continuous with respect to the 1-Wasserstein distance of the empirical measures in the sense of (5.2.4). Hence we will be able to apply this lemma to f_N and g_N thanks to the symmetry **(S)** and assumptions **(MF)**–**(LP)** as well as to $\mathfrak{u}$ and $\partial_x\mathfrak{u}$ thanks to Remark 5.2.3 and our estimates on the derivatives.

We formalise a generalisation of the above-mentioned fact in the following lemma, where we denote by $|\cdot|_p$ the standard p-norm on E^N, with E being as in Lemma 5.2.4.

LEMMA 5.2.5. *Let $h\colon E^N \to \mathbb{R}$ be symmetric and such that*

$$\||Dh|_q\|_\infty \leqslant CN^{-\frac{1}{p}}\,, \quad \text{with } \frac{1}{p}+\frac{1}{q}=1\,,$$

for some constants $C>0$ and $q\in(1,\infty]$ independent of N. Then

$$|h(v)-h(v')| \leqslant CW_p(m_v, m_{v'})$$

for all $v, v' \in E^N$.

PROOF. Recall that $W_p(m_v, m_{v'}) = N^{-\frac{1}{p}}|v - v'^{\sigma}|_p$ for some $\sigma \in S_N$. By the symmetry of h and Hölder's inequality, $|h(v)-h(v')| \leqslant \||Dh|_q\|_\infty |v-v'^{\sigma}|_p$. The conclusion follows. ∎

We are now ready to prove the existence of a limit of $\mathfrak{u}_N$ as $N\to\infty$.

THEOREM 5.2.6. *Let assumptions **(MF)**, **(S)** and **(LP)** be in force. Assume also that one between **(DS)** and **(LS)** holds, with the corresponding semimonotonicity constants being such that the thesis of Theorem 4.2.7 holds.*

Then there exists a map

$$U\colon \Lambda\times[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\Lambda\times\mathbb{R}^d)\to\mathbb{R}\,,$$

with bounded derivative $\partial_x U$, such that U and $\partial_x U$ are Lipschitz continuous on $\Lambda\times\mathbb{R}^d\times\mathcal{P}_2(\Lambda\times\mathbb{R}^d)$ and $\frac{1}{3}$-Hölder continuous on $[0,T]$, and, up to a subsequence,

$$\sup\big|\partial_x^k\mathfrak{u}_N(\underline{\lambda},t,\underline{x}) - \partial_x^k U(\lambda,t,x,m_{(\hat\lambda:\hat x)})\big| \xrightarrow[N\to\infty]{} 0\,, \qquad k\in\{0,1\}\,, \tag{5.2.5}$$

whenever the supremum is taken over any set of the form

$$\big\{(\underline{\lambda},t,\underline{x})\in\Lambda^N\times[0,T]\times(\mathbb{R}^d)^N :\ |x|\leqslant R,\ m_{\hat x}\in\mathcal{K}\big\}$$

for some $R>0$ and some compact set $\mathcal{K}\subset\mathcal{P}_2(\mathbb{R}^d)$.

PROOF. For the local uniform convergence of a subsequence $\mathfrak{u}_N$, we apply Lemma 5.2.4 with $\gamma=(t,\lambda,x)$ and $v=(\hat\lambda:\hat x)$. Indeed, first note that the $\mathfrak{u}_N$ are globally bounded, uniformly in N, by a standard comparison argument (recall that

f^i and g^i are now assumed to be bounded uniformly). The second inequality in (5.2.1a) and the estimate of Lemma 5.2.2 yield uniform Lipschitz continuity with respect to $(\hat\lambda : \hat x)$ in W_2 distance, in the sense provided by Lemma 5.2.5. Lipschitz continuity in (λ, x) is also a consequence of Lemma 5.2.2, and the first estimate in (5.2.1a). Finally, since the right-hand side of the equation for $\mathfrak{u}_N$

$$-\partial_t \mathfrak{u}_N - \Delta \mathfrak{u}_N = -\frac{1}{2}|\partial_x \mathfrak{u}_N|^2 - \sum_{j \in [\![N]\!] \setminus \{0\}} D_j u^j \cdot D_j \mathfrak{u}_N + f_N$$

is uniformly bounded by (5.2.1a), and $\mathfrak{u}_N$ is uniformly Lipschitz in the $\underline{x}$-variable with respect to the weighted norm $\|\cdot\|^0$, again by (5.2.1a), the desired uniform Hölder regularity in t follows by Lemma 4.6.2.

The local uniform convergence of $\partial_x \mathfrak{u}_N$ is analogous. We argue as before, employing estimates in (5.2.1b) instead of (5.2.1a). Time regularity is obtained from the equation

$$\begin{aligned} &-\partial_t(\partial_x \mathfrak{u}_N) - \Delta(\partial_x \mathfrak{u}_N) \\ &\qquad = -\sum_{j \in [\![N]\!]} D_j u^j \cdot D_j(\partial_x \mathfrak{u}_N) - \sum_{j \in [\![N]\!] \setminus \{0\}} (D^2_{j0} u^j) \cdot D_j \mathfrak{u}_N + \partial_x f_N\,, \end{aligned}$$

whose right-hand side can be controlled using (5.2.1a), (5.2.1b). ∎

Remark 5.2.7. Regularity of U is actually a bit better. First, (5.2.1a) yields Lipschitz continuity in the measure variable with respect to the W_1 distance. Moreover, looking at the equation for $\mathfrak{u}_N$ one observes that $\|\partial_t \mathfrak{u}_N\|_\infty \lesssim 1$, that gives uniform Lipschitz continuity in the variable t.

Remark 5.2.8. In Example 5.1.3 we discussed the implications of our assumption when taking the data h_N (that is, f_N and g_N) as projections over empirical measures of given $\mathfrak{h}_N : \Lambda \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}$, and in Remark 5.1.2 we argued that this can be considered a natural choice.

Furthermore, our standing assumptions force in fact also the convergence of h_N to functions defined on probability measures, at least up to subsequences. Indeed, assumption **(MF)** guarantees that $\|D_j h_N(\lambda, \cdot)\|_\infty \leqslant L_h N^{-1}$ for all $j > 0$. Then by Lemma 5.2.5 we can apply Lemma 5.2.4 and say that there exists a map $h \in \mathrm{Lip}(\Lambda \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d))$ such that, up to a subsequence,

$$\sup_{\lambda \in \Lambda,\, |x| \leqslant R,\, m_{\hat x} \in \mathcal{K}} |\partial_x^k h_N(\lambda, \underline{x}) - \partial_x^k h(\lambda, x, m_{\hat x})| \xrightarrow[N \to \infty]{} 0, \tag{5.2.6}$$

for any $R > 0$ and $\mathcal{K} \subset \mathcal{P}_1(\mathbb{R}^d)$ compact, and for k as in **(LP)**.

5.3. Characterisation of the limit

Henceforth, we will work with the three limit functions f, g and U defined in Remark 5.2.8 and Theorem 5.2.6. We will adopt the notation $f^\lambda = f(\lambda, \cdot)$ and analogously for g^λ and U^λ.

Our aim is to give a characterisation of U. As a first step and a starter, the most basic result we can prove regards the standard symmetric setting of MFGs; that is, when $\boldsymbol{\lambda} = \lambda 1_N$ for some $\lambda \in \Lambda$, so that all f^i (and all g^i) coincide up to a permutation of the coordinates of $(\mathbb{R}^d)^N$.

PROPOSITION 5.3.1. *For all $\tau \in [0,T]$, $\mathfrak{m}, \bar{\mathfrak{m}} \in \mathcal{P}_2(\mathbb{R}^d)$ and $\lambda \in \Lambda$,*

$$\int_{\mathbb{R}^d} U^\lambda(\tau, \cdot, \delta_\lambda \otimes \mathfrak{m})\, \mathrm{d}\bar{\mathfrak{m}} = \int_\tau^T \int_{\mathbb{R}^d} \Big(\frac{1}{2} \big|\partial_x U^\lambda(s, \cdot, \delta_\lambda \otimes m_s)\big|^2 + f^\lambda(\cdot, m_s) \Big) \mathrm{d}\bar{m}_s \mathrm{d}s + \int_{\mathbb{R}^d} g^\lambda(\cdot, m_T)\, \mathrm{d}\bar{m}_T\,, \tag{5.3.1}$$

where m and $\bar{m}$ solve, respectively,

$$\begin{cases} \partial_t m - \Delta m - \operatorname{div}(\partial_x U^\lambda(t, x, \delta_\lambda \otimes m) m) = 0 \\ m|_{t=\tau} = \mathfrak{m} \end{cases}$$

and

$$\begin{cases} \partial_t \bar{m} - \Delta \bar{m} - \operatorname{div}(\partial_x U^\lambda(t, x, \delta_\lambda \otimes m) \bar{m}) = 0 \\ \bar{m}|_{t=\tau} = \bar{\mathfrak{m}}. \end{cases}$$

for $t \in [\tau, T]$ and $x \in \mathbb{R}^d$.

In order to abridge the notation, given $\mathfrak{m} \in \mathcal{P}_2(\mathbb{R}^d)$ we will write $U^\lambda(t, x, \mathfrak{m})$ in lieu of $U^\lambda(t, x, \delta_\lambda \otimes \mathfrak{m})$. Also, throughout the proofs we will imply all constants which depend only on T, L_g, L_f and L_Λ. We state the basic propagation of chaos results needed for the proof in the following two lemmata.

LEMMA 5.3.2. *Let $\lambda \in \Lambda$, $\bar{m}, m_0 \in \mathcal{P}_q(\mathbb{R}^d)$, $q > 4$, and let $Z \sim \bar{m} \otimes m_0^{\otimes(N-1)}$; then consider X^N and Y^N solving, respectively,*

$$\begin{cases} \mathrm{d}X^{N,i} = -\partial_x U^\lambda(t, X_t^{N,i}, m_{X_t^N})\, \mathrm{d}t + \sqrt{2}\, \mathrm{d}B_t^i, \quad i \in [\![N]\!] \\ X_0^N = Z \end{cases}$$

and

$$\begin{cases} \mathrm{d}Y^{N,i} = -\partial_x U^\lambda(t, Y_t^{N,i}, \mathcal{L}_{Y_t^{N,i}})\, \mathrm{d}t + \sqrt{2}\, \mathrm{d}B_t^i, \quad i \in [\![N]\!] \setminus \{0\} \\ \mathrm{d}Y^{N,0} = -\partial_x U^\lambda(t, Y_t^{N,0}, \mathcal{L}_{Y_t^{N,1}})\, \mathrm{d}t + \sqrt{2}\, \mathrm{d}B_t^0 \\ Y_0^N = Z. \end{cases}$$

Let $m_t := \mathcal{L}_{Y_t^{N,i}}$ if $i \geqslant 1$ (note that $\mathcal{L}_{Y_t^{N,i}}$ is independent of $i \geqslant 1$) and $\bar{m}_t := \mathcal{L}_{Y_t^{N,0}}$. Then

$$\lim_{N\to\infty} \sup_{i \in [\![N]\!]} \mathbb{E}\Big[\sup_{t\in[0,T]} \big|X_t^{N,i} - Y_t^{N,i}\big|^2 \Big] = 0 \tag{5.3.2}$$

and

$$\lim_{N\to\infty} \mathbb{E}\Big[\sup_{t\in[0,T]} W_2(m_{X_t^N}, m_t) \Big] = 0. \tag{5.3.3}$$

PROOF. Subtracting the equations of X^N and Y^N and using the Lipschitz continuity of $\partial_x U$, we have

$$|X_t^{N,i} - Y_t^{N,i}| \lesssim \int_0^t \big(|X^{N,i} - Y^{N,i}| + W_2(m_{X^N}, m_{Y^N}) + W_2(m_{Y^N}, m)\big);$$

therefore, since due to the boundedness of $\partial_x U$ one estimates[2]

$$\mathbb{E}\big[\sup_{t\in[0,T]} W_2(m_{Y_t^{N,-0}}, m_{Y_t^N})^2\big] \lesssim \frac{1}{N},$$

we obtain

$$|X_t^{N,i} - Y_t^{N,i}|^2 \lesssim \int_0^t \Big(|X^{N,i} - Y^{N,i}|^2 + \frac{1}{N}\sum_{1\leqslant j\leqslant N} |X^{N,j} - Y^{N,j}|^2 + W_2(m_{Y^{N,-0}}, m)\Big) + \frac{1}{N}. \tag{5.3.4}$$

Averaging over i,

$$\frac{1}{N}\sum_{i\in[\![N]\!]} |X_t^{N,i} - Y_t^{N,i}|^2 \lesssim \frac{1}{N} + \int_0^t \Big(\frac{1}{N}\sum_{i\in[\![N]\!]} |X^{N,i} - Y^{N,i}|^2 + W_2(m_{Y^{N,-0}}, m)\Big),$$

thus by Gronwall's lemma

$$\frac{1}{N}\sum_{i\in[\![N]\!]} |X_t^{N,i} - Y_t^{N,i}|^2 \lesssim \frac{1}{N} + \int_0^t W_2(m_{Y^{N,-0}}, m)$$

and plugging this back into (5.3.4) one concludes that

$$\sup_{i\in[\![N]\!]} \mathbb{E}\Big[\sup_{t\in[0,T]} \big|X_t^{N,i} - Y_t^{N,i}\big|^2\Big] \lesssim \frac{1}{N} + \int_0^T \mathbb{E}\big[W_2(m_{Y^{N,-1}}, m)^2\big],$$

where the right-hand side goes to 0 as $N \to \infty$ by [**25**, Theorem 5.8]. Then note that by the triangle inequality one deduces (5.3.3) as well. ∎

LEMMA 5.3.3. *Let X^N be as in Lemma 5.3.2 and consider $\tilde{X}^N$ solving*

$$\begin{cases} \mathrm{d}\tilde{X}_t^{N,i} = -\partial_x \mathfrak{u}_N(\lambda 1_N, t, \tilde{X}_t^{N,i}, \tilde{X}_t^{N,-i})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i \in [\![N]\!] \\ \tilde{X}_0^N = Z. \end{cases}$$

Then

$$\lim_{N\to\infty} \sup_{i\in[\![N]\!]} \mathbb{E}\Big[\sup_{t\in[0,T]} |\tilde{X}_t^{N,i} - X_t^{N,i}|^2\Big] = 0. \tag{5.3.5}$$

[2]According to the notation adopted for empirical measures, we are writing $m_{Y^{N,-0}} = \frac{1}{N-1}\sum_{j\in[\![N]\!]\setminus\{0\}} \delta_{Y^{N,j}}$; note that in the literature the same empirical measure is often denoted by $m_Y^{N,0}$.

PROOF. Note that, since $\mathbb{E}\big[\sup_{t\in[0,T]} W_2(m_{\tilde X_t^{N,-i}}, m_{\tilde X_t^N})^2\big] \lesssim N^{-1}$,

$$
\begin{aligned}
(5.3.6)\quad & |\tilde X_s^{N,i} - X_s^{N,i}|^2 \\
&\lesssim \int_0^s \Big(\big|\partial_x \mathfrak{u}_N(\lambda 1_N, t, \tilde X_t^{N,i}, \tilde X_t^{N,-i}) - \partial_x \mathfrak{u}_N(\lambda 1_N, t, X_t^{N,i}, \tilde X_t^{N,-i})\big|^2 \\
&\quad + \big|\partial_x \mathfrak{u}_N(\lambda 1_N, t, X_t^{N,i}, \tilde X_t^{N,-i}) - \partial_x U^\lambda(t, X_t^{N,i}, m_{\tilde X_t^N})\big|^2 \\
&\quad + \big|\partial_x U^\lambda(t, X_t^{N,i}, m_{\tilde X_t^N}) - \partial_x U^\lambda(t, X_t^{N,i}, m_{X_t^N})\big|^2 \Big)\, \mathrm{d}t \\
&\lesssim \int_0^s \big(|\tilde X_t^{N,i} - X_t^{N,i}|^2 + \delta_t^{N,i} + W_2(m_{\tilde X_t^N}, m_{X_t^N})^2 \big)\, \mathrm{d}t
\end{aligned}
$$

where we have set

$$
\delta_t^{N,i} := \big|\partial_x \mathfrak{u}_N(\lambda 1_N, t, X_t^{N,i}, \tilde X_t^{N,-i}) - \partial_x U^\lambda(t, X_t^{N,i}, m_{\tilde X_t^{N,-i}})\big|^2 + \frac{1}{N}.
$$

Taking the averages,

$$
\frac{1}{N}\sum_{i\in[\![N]\!]} |\tilde X_s^{N,i} - X_s^{N,i}|^2 \lesssim \frac{1}{N}\sum_{i\in[\![N]\!]} \int_0^T \delta_t^{N,i}\, \mathrm{d}t + \int_0^s \frac{1}{N}\sum_{i\in[\![N]\!]} |\tilde X_t^{N,i} - X_t^{N,i}|^2\, \mathrm{d}t\,,
$$

so that Gronwall's lemma yields

$$
\sup_{t\in[0,T]} \frac{1}{N}\sum_{i\in[\![N]\!]} |\tilde X_t^{N,i} - X_t^{N,i}|^2 \lesssim \frac{1}{N}\sum_{i\in[\![N]\!]} \int_0^T \delta_t^{N,i}\, \mathrm{d}t
$$

and, plugging this back into (5.3.6),

$$
(5.3.7)\qquad \sup_{i\in[\![N]\!]} \mathbb{E}\Big[\sup_{t\in[0,T]} |\tilde X_t^{N,i} - X_t^{N,i}|^2\Big] \lesssim \frac{1}{N}\sum_{i\in[\![N]\!]} \int_0^T \mathbb{E}\,\delta_t^{N,i}\, \mathrm{d}t \xrightarrow[N\to\infty]{} 0\,;
$$

in order to justify the convergence to 0 of the right-hand side of (5.3.7), we proceed as follows. First note that, since $\partial_x \mathfrak{u}^N, D_x U$ are bounded uniformly and $\bar m, m_0 \in \mathcal{P}_q(\mathbb{R}^d)$, $q > 4$, it is standard that $\mathbb{E}|\tilde X_t^{N,i}|^q$ and $\mathbb{E}|X_t^{N,i}|^q$ are uniformly bounded (and thus, so is $N^{-1}\mathbb{E}\sum_{1\leqslant j\leqslant N} |\tilde X_t^{N,j}|^q$). Then, for $x \in (\mathbb{R}^d)^N$, let $M_q(x) := N^{-1}\sum_{1\leqslant j\leqslant N} |x^j|^q$; since

$$
\int_0^T \mathbb{E}\,\delta_t^{N,i} = \int_0^T \mathbb{E}\big[1_{\{M_q(\tilde X_t^N) + |X_t^{N,i}| \leqslant R\}} \delta_t^{N,i}\big] + \int_0^T \mathbb{E}\big[1_{\{M_q(\tilde X_t^N) + |X_t^{N,i}| > R\}} \delta_t^{N,i}\big]\,,
$$

the second term can be made arbitrarily small provided that R is large, by the aforementioned bounds on q-th moments, while the first can be made small by choosing N large enough ($\delta_t^{N,i}$ is bounded uniformly). Indeed, it is sufficient to apply the uniform convergence (5.2.5) on the set $\mathcal{K}$ of all probability measures with q-th moment bounded by R,[3] which is compact in $\mathcal{P}_2(\mathbb{R}^d)$. Note that the previous limit is uniform in i, since K and the bounds on the moments are independent of i. ∎

[3]here we mean that $\mathcal{K} := \{\nu \in \mathcal{P}_2(\mathbb{R}^d) :\ \int_{\mathbb{R}^d} |x|^q\, \nu(\mathrm{d}x) \leqslant R\}$.

PROOF OF PROPOSITION 5.3.1. We have

$$\Big|\big|\partial_x U^\lambda(t,\cdot)\big|^2\Big|^{(Y_t^{N,i},m_t)}_{(X_t^{N,i},m_{X_t^N})}\Big| + \Big|f^\lambda\Big|^{(Y_t^{N,i},m_t)}_{(X_t^{N,i},m_{X_t^N})}\Big| + \Big|g^\lambda\Big|^{(Y_t^{N,i},m_t)}_{(X_t^{N,i},m_{X_t^N})}\Big| \lesssim \big|X_t^{N,i} - Y_t^{N,i}\big| + W_2(m_{X_t^N}, m_t)\,,$$

the expectation of the right-hand side going to 0 as $N \to \infty$ (uniformly in $t \in [0,T]$) by Lemma 5.3.2. Then we deduce that

$$\mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}\big|\partial_x U^\lambda(t, X_t^{N,0}, m_{X_t^N})\big|^2 + f^\lambda(X_t^{N,0}, m_{X_t^N})\Big)\mathrm{d}t + g^\lambda(X_T^{N,0}, m_{X_T^N})\Big] \xrightarrow[N\to\infty]{} \int_0^T\int_{\mathbb{R}^d}\Big(\frac{1}{2}\big|\partial_x U^\lambda(t,\cdot,m_t)\big|^2 + f^\lambda(\cdot,m_t)\Big)\mathrm{d}\bar{m}_t\mathrm{d}t + \int_{\mathbb{R}^d} g^\lambda(\cdot,m_T)\,\mathrm{d}\bar{m}_T\,. \tag{5.3.8}$$

Now, by the representation formula for the value functions solving of the Nash system (see, e.g., [**25**, Section 2.1.4]),

$$\mathbb{E}\,\mathfrak{u}_N(\lambda 1_N, 0, Z) = \mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}\big|\partial_x \mathfrak{u}_N(\lambda 1_N, t, \tilde{X}_t^N)\big|^2 + f_N(\lambda, \tilde{X}_t^N)\Big)\mathrm{d}t + g_N(\lambda, \tilde{X}_T^N)\Big]\,, \tag{5.3.9}$$

where $\tilde{X}^N$ is the vector of optimal trajectories for the N-player game; that is, $\tilde{X}^N$ is as in Lemma 5.3.3. Then, combining (5.3.8), (5.3.9) and (5.3.5) one deduces that

$$\mathbb{E}\,\mathfrak{u}_N(\lambda 1_N, 0, Z) \xrightarrow[N\to\infty]{} \int_0^T\int_{\mathbb{R}^d}\Big(\frac{1}{2}\big|\partial_x U^\lambda(t,\cdot,m_t)\big|^2 + f^\lambda(\cdot,m_t)\Big)\mathrm{d}\bar{m}_t\mathrm{d}t + \int_{\mathbb{R}^d} g^\lambda(\cdot,m_T)\,\mathrm{d}\bar{m}_T\,. \tag{5.3.10}$$

Finally, since $\mathbb{E}\,W_2(m_{Z^{-1}}, m_0) \to 0$ by [**25**, Theorem 5.8], from (5.2.5) and (5.3.10) we obtain

$$\int_{\mathbb{R}^d} U^\lambda(0,\cdot,m_0)\,\mathrm{d}\bar{m} = \int_0^T\int_{\mathbb{R}^d}\Big(\frac{1}{2}\big|\partial_x U^\lambda(t,\cdot,m_t)\big|^2 + f^\lambda(\cdot,m_t)\Big)\mathrm{d}\bar{m}_t\mathrm{d}t + \int_{\mathbb{R}^d} g^\lambda(\cdot,m_T)\,\mathrm{d}\bar{m}_T\,.$$

As $\bar{m}, m_0$ can be chosen in a dense subset of $\mathcal{P}_2(\mathbb{R}^d)$ and the choice of $\tau = 0$ as initial time is arbitrary, (5.3.1) follows by standard stability of the Fokker-Planck equation. ∎

We finally extend the characterisation (5.3.1) to any measure $\mu \in \mathcal{P}_2(\Lambda \times \mathbb{R}^d)$, with a continuous disintegration with respect to the projection $\pi_\Lambda \colon \Lambda \times \mathbb{R}^d \to \Lambda$. This will be the main result of this section. The extension is rather technical, but the main point is performing the following reduction: the continuity in λ (of the disintegration and of the data) allows to approximate the problem with a simpler one, where the first marginal of μ is supported on a finite subset of Λ. This is a sort of problem with finitely many populations, where we apply the previous convergence argument; in other words, we are going to perform a sort of multiple "localised" propagation of chaos.

THEOREM 5.3.4. *Let $\mu \in \mathcal{P}_2(\Lambda \times \mathbb{R}^d)$ admit a disintegration $\mathfrak{m}\colon K \to \mathcal{P}_2(\mathbb{R}^d)$ with respect to π_Λ which is continuous on some compact subset $K \supseteq \operatorname{spt}(\pi_{\Lambda\sharp}\mu)$. Let $\bar{\mathfrak{m}} \in \mathcal{P}_2(\mathbb{R}^d)^\Lambda$. Then for all $\tau \in [0,T]$ and $\lambda \in K$,*

$$\begin{aligned}(5.3.11)\quad \int_{\mathbb{R}^d} U^\lambda(\tau,\cdot,\mu)\,\mathrm{d}\bar{\mathfrak{m}}^\lambda &= \int_\tau^T \int_{\mathbb{R}^d} \Big(\frac{1}{2}\big|\partial_x U^\lambda(s,\cdot,\mu_s)\big|^2 + f^\lambda(\cdot,\pi_{\mathbb{R}^d\sharp}\mu_s)\Big)\,\mathrm{d}\bar{m}^\lambda_s \mathrm{d}s \\ &\quad + \int_{\mathbb{R}^d} g^\lambda(\cdot,\pi_{\mathbb{R}^d\sharp}\mu_T)\,\mathrm{d}\bar{m}^\lambda_T\,,\end{aligned}$$

where $\mu \in C^0([\tau,T];\mathcal{P}_2(\Lambda\times\mathbb{R}^d))$ and $\bar{m} \in C^0([\tau,T];\mathcal{P}_2(\mathbb{R}^d))^\Lambda$ solve, respectively,

$$(5.3.12)\quad \begin{cases}\partial_t m^\lambda - \Delta m^\lambda - \operatorname{div}(\partial_x U^\lambda(t,x,\mu_t) m^\lambda) = 0 \\ \mu_t = \int_\Lambda m^\lambda_t\, \pi_{\Lambda\sharp}\mu(\mathrm{d}\lambda) \\ m^\lambda_\tau = \mathfrak{m}^\lambda\end{cases}$$

and

$$(5.3.13)\quad \begin{cases}\partial_t \bar{m}^\lambda - \Delta \bar{m}^\lambda - \operatorname{div}(\partial_x U^\lambda(t,x,\mu_t) \bar{m}^\lambda) = 0 \\ \bar{m}^\lambda_\tau = \bar{\mathfrak{m}}^\lambda\end{cases}$$

for $t \in [\tau,T]$, $x \in \mathbb{R}^d$ and $\lambda \in \Lambda$.

REMARK 5.3.5. Since $\partial_x U^\lambda$ and $\operatorname{div}(\partial_x U^\lambda)$ are globally bounded, $\bar{m}^\lambda$ is locally continuous as a function with values in $L^1(\mathbb{R}^d)$, it solves the Fokker-Planck equation in the strong sense, and it belongs locally to $W^{2,1}_p$ for every p, see for example [**18**]. Regularity extends up to time τ provided that $\bar{\mathfrak{m}}^\lambda$ is smooth enough.

Problem (5.3.13) is simply an uncoupled system of (possibly uncountably many) Fokker–Planck equations. On the other hand, before proceeding with the proof of this theorem we show that problem (5.3.12) is well-posed, as a result of the following lemma. Also note that it can be written as

$$(5.3.14)\quad \begin{cases}\partial_t \mu - \Delta_x \mu - \operatorname{div}_x(\partial_x U^\lambda(t,x,\mu_t)\mu) = 0 \quad \text{on } [\tau,T]\times\Lambda\times\mathbb{R}^d \\ \mu_\tau = \mu,\end{cases}$$

so that μ solves an equation of Fokker–Planck type with no derivatives with respect to λ.

LEMMA 5.3.6. *Let $B\colon \Lambda \times [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\Lambda\times\mathbb{R}^d) \to \mathbb{R}^d$ be bounded Lipschitz continuous on $\Lambda \times \mathbb{R}^d \times \mathcal{P}_2(\Lambda\times\mathbb{R}^d)$ uniformly on $[0,T]$. Let $\mu \in \mathcal{P}_2(\Lambda\times\mathbb{R}^d)$ and* *let $\rho := \pi_{\Lambda\sharp}\mu \in \mathcal{P}(\Lambda)$ be its first marginal. Then problem*[4]

$$\begin{cases} \partial_t m^\lambda - \Delta m^\lambda + \operatorname{div}(B^\lambda(t,x,\mu_t)m^\lambda) = 0 \quad \text{on } [0,T]\times\mathbb{R}^d, \quad \forall\,\lambda\in\Lambda \\ \mu_t = \int_\Lambda m_t^\lambda \rho(\mathrm{d}\lambda) \quad \forall\, t\in[0,T] \\ \mu_0 = \mu \end{cases} \tag{5.3.15}$$

has a unique solution $\mu \in C^{\frac12}([0,T];\mathcal{P}_2(\Lambda\times\mathbb{R}^d))$, whose disintegration m with respect to ρ is unique as well if one fixes the disintegration of μ.

Furthermore, if μ and $\tilde\mu$ are the solutions relative to data μ and $\tilde\mu$ with respective disintegrations given by $\mathfrak{m}$ and $\tilde{\mathfrak{m}}$, respectively, one has

$$\sup_{t\in[0,T]} W_2(\mu_t,\tilde\mu_t) \lesssim W_2(\mu,\tilde\mu), \tag{5.3.16}$$

$$\sup_{t\in[0,T]} W_2(m_t^\lambda,\tilde m_t^{\lambda'}) \lesssim W_2(\mathfrak{m}^\lambda,\tilde{\mathfrak{m}}^{\lambda'}) + W_2(\mu,\tilde\mu) \quad \forall\,\lambda,\lambda'\in\Lambda, \tag{5.3.17}$$

where the implied constants depend only on T and $\|B\|_{\mathrm{Lip}}$.

Proof. For some $A > 0$ to be determined, let $\mathfrak{C}$ be the set of all $\mu \in C^0([0,T];\mathcal{P}_2(\Lambda\times\mathbb{R}^d))$ such that $[\mu]_{0,\frac12} := \sup_{s,t\in[0,T],\, s\neq t} |s-t|^{-\frac12} W_2(\mu_s,\mu_t) \leqslant A$. Note that $\mathfrak{C}$ a closed convex subset $C^0([0,T];\mathcal{P}_2(\Lambda\times\mathbb{R}^d))$. Given $\mu\in\mathfrak{C}$ and $\lambda\in\Lambda$, let $m^\lambda \in C^0([0,T];\mathcal{P}_2(\mathbb{R}^d))$ be the unique solution to $\partial_t m^\lambda - \Delta m^\lambda + \operatorname{div}(B^\lambda(t,x,\mu_t)m^\lambda) = 0$ with $m_0^\lambda = \mathfrak{m}^\lambda$, where $(\mathfrak{m}^\lambda)_{\lambda\in\Lambda}$ is a fixed disintegration of μ with respect to ρ. Let $\Phi(\mu) := \int_\Lambda m^\lambda \rho(\mathrm{d}\lambda)$. It is standard (cf. [**24**, Lemma 1.6]) to show that $[m^\lambda]_{0,\frac12}$ is bounded by a constant which increases with $\|B^\lambda\|_\infty$. In addition, if $\pi_{s,t}^\lambda$ is an optimal coupling of $(m_s^\lambda, m_t^\lambda)$, then $\int_{\Lambda\times\Lambda} \pi_{s,t}^\lambda((\mathrm{id}_\Lambda,\mathrm{id}_\Lambda)_\sharp\rho)(\mathrm{d}\lambda,\mathrm{d}\lambda')$ is a coupling of $(\tilde\mu_s,\tilde\mu_t)$, showing that $W_2(\Phi(\mu)_s,\Phi(\mu)_t)^2 \leqslant \int_\Lambda W_2(m_s^\lambda,m_t^\lambda)^2\rho(\mathrm{d}\lambda)$. Therefore, since $\sup_{\lambda\in\Lambda}\|B^\lambda\|_\infty$ is finite, if A is sufficiently large then the map $\Phi\colon \mathfrak{C}\to\mathfrak{C}$ is well-defined.

The following estimates will prove that Φ is a contraction for T small, and thus (by iteration) the existence of a unique solution to (5.3.15) with arbitrary T. Furthermore, a posteriori, they will also yield the desired estimates (5.3.16) and (5.3.17).

Let $\mu,\tilde\mu\in\mathfrak{C}$ and let two initial data $\mu,\tilde\mu\in\mathcal{P}(\Lambda\times\mathbb{R}^d)$ be given; let $\rho = \pi_{\Lambda\sharp}\mu$, $\tilde\rho = \pi_{\Lambda\sharp}\tilde\mu$ and $(\mathfrak{m}^\lambda)_{\lambda\in\Lambda}$ (resp. $(\tilde{\mathfrak{m}}^\lambda)_{\lambda\in\Lambda}$) be a disintegration of μ (resp. $\tilde\mu$) with respect to ρ (resp. $\tilde\rho$). Let $\nu\in C^0([0,T];\mathcal{P}_2((\Lambda\times\mathbb{R}^d)^2)$ satisfy

$$\partial_t\nu - \Delta_x\nu - \Delta_y\nu + 2\sum_{1\leqslant j\leqslant d}\partial^2_{x^jy^j}\nu + \operatorname{div}_x(B^\lambda(t,x,\mu_t)\nu) + \operatorname{div}_y(B^{\lambda'}(t,y,\tilde\mu_t)\nu) = 0\,,$$

so that if $\nu_0\in\Pi(\mu,\tilde\mu)$, then ν_t is a coupling of $(\Phi(\mu)_t,\Phi(\tilde\mu)_t)$. Testing by $|\lambda-\lambda'|^2 + |x-y|^2$ and assuming that ν_0 is optimal for the 2-Wasserstein distance, we

[4]With an abuse of notation we write $\int_\Lambda m_t^\lambda\rho(\mathrm{d}\lambda)$ instead of $\int_\Lambda \hat m_t^\lambda\rho(\mathrm{d}\lambda)$ with $\hat m_t^\lambda(E) := m^\lambda(\pi_{\mathbb{R}^d}(E\cap\pi_\Lambda^{-1}(\{\lambda\})))$ for all Borel sets $E\subseteq\Lambda\times\mathbb{R}^d$. We will continue to identify $\hat m^\lambda$ with m^λ.

obtain

$$\int_{(\Lambda\times\mathbb{R}^d)^2} (|\lambda-\lambda'|^2+|x-y|^2)\,\nu_t(\mathrm{d}\lambda,\mathrm{d}x,\mathrm{d}\lambda',\mathrm{d}y)$$
$$= W_2(\mu,\tilde\mu)^2 + 2\int_0^t \int_{(\Lambda\times\mathbb{R}^d)^2} \big(B^\lambda(t,x,\mu_t) - B^{\lambda'}(t,y,\tilde\mu_t)\big) \cdot (x-y)\,\mathrm{d}\nu_s(\mathrm{d}\lambda,\mathrm{d}x,\mathrm{d}\lambda',\mathrm{d}y)\,\mathrm{d}s\,.$$

By the Cauchy–Schwarz and Young's inequalities,

$$\int_{(\Lambda\times\mathbb{R}^d)^2} (|\lambda-\lambda'|^2+|x-y|^2)\,\nu_t(\mathrm{d}\lambda,\mathrm{d}x,\mathrm{d}\lambda',\mathrm{d}y)$$
$$\leqslant W_2(\mu,\tilde\mu)^2 + C\int_0^t W_2(\mu_s,\tilde\mu_s)^2\,\mathrm{d}s$$
$$+ C\int_0^t \int_{(\Lambda\times\mathbb{R}^d)^2} (|\lambda-\lambda'|^2+|x-y|^2)\,\nu_s(\mathrm{d}\lambda,\mathrm{d}x,\mathrm{d}\lambda',\mathrm{d}y)\,\mathrm{d}s\,,$$

where C depends only on $\|B\|_{\mathrm{Lip}}$. Apply Gronwall's lemma to deduce

$$W_2(\Phi(\mu)_t,\Phi(\tilde\mu)_t)^2 \leqslant \bigg(W_2(\mu,\tilde\mu)^2 + C\int_0^t W_2(\mu_s,\tilde\mu_s)^2\,\mathrm{d}s\bigg)e^{Ct}, \tag{5.3.18}$$

which implies that Φ is a contraction on $\mathfrak{C}\subset C^0([0,T];\mathcal{P}_2(\Lambda\times\mathbb{R}^d))$ if $CTe^{CT}<1$. Since this smallness condition on T depends only on $\|B\|_{\mathrm{Lip}}$, we can iterate finitely many times the short-time uniqueness provided by the contraction theorem in order to obtain uniqueness on an arbitrary horizon.

Now $\Phi(\mu)=\mu$ and $\Phi(\tilde\mu)=\tilde\mu$ in (5.3.18) above, then we can apply Gronwall's lemma once again to get (5.3.16). With a similar argument (which considers a coupling of $(m_t^\lambda,\tilde m_t^{\lambda'})$, where m^λ is a solution to (5.3.15) given $m_0^\lambda=\mathfrak{m}^\lambda$ and analogously for $\tilde m^{\lambda'}$) one can also deduce that

$$W_2(m_t^\lambda,\tilde m_t^{\lambda'})^2 \lesssim W_2(\mathfrak{m}^\lambda,\tilde{\mathfrak{m}}^{\lambda'})^2 + \int_0^t W_2(\mu_s,\tilde\mu_s)^2\,\mathrm{d}s.$$

Then, using (5.3.16), one has (5.3.17). ∎

For the benefit of the reader, we also collect the crucial steps of the proof of Theorem 5.3.4 in some lemmata, in order to isolate the propagation of chaos and the approximation arguments. We begin with the latter.

Lemma 5.3.7. *Suppose that Theorem 5.3.4 holds for a sequence $(\mu_k)_{k\in\mathbb{N}}\subset\mathcal{P}_2(\Lambda\times\mathbb{R}^d)$ such that $\pi_{\Lambda\sharp}\mu_k$ is finitely supported and weakly converges to $\rho\in\mathcal{P}(\Lambda)$. Then Theorem 5.3.4 holds for any μ as in its statement such that $\pi_{\Lambda\sharp}\mu=\rho$.*

Proof. Given $\mu\in\mathcal{P}_2(\Lambda\times\mathbb{R}^d)$ as in the statement of Theorem 5.3.4 and $\rho:=\pi_{\Lambda\sharp}\mu$, let $(\rho_k)_{k\in\mathbb{N}}\subset\mathcal{P}(K)$ be a sequence of finitely-supported probability measures such that $\rho_k\to\rho$ weakly. Since W_2 metrises the weak convergence of probability measures on K (see, e.g., [**78**, Corollary 6.13]), the existence of such a sequence comes from the Banach–Alaoglu and the Krein–Milman theorems. Moreover, note that one can in particular choose ρ_k supported only on rational numbers.

Consider the continuous disintegration $\mathfrak{m}\colon K \to \mathcal{P}_2(\mathbb{R}^d)$ of μ with respect to π_Λ and define $\mu_k := \int_\Lambda \mathfrak{m}^\lambda \rho_k(\mathrm{d}\lambda)$; then, let m_k be the corresponding solution to (5.3.12) (note that $\pi_{\Lambda\sharp}(\mu_k)_t = \rho_k$ for all t). Thanks to the continuity of the disintegration $\mathfrak{m}$, the map $\lambda \mapsto \int_{\mathbb{R}^d} F(\lambda,x)\mathfrak{m}^\lambda(\mathrm{d}x)$ is continuous, so, for any $F \in C(\Lambda\times\mathbb{R}^d)$ with sub-quadratic growth with respect to $x\in\mathbb{R}^d$, by the weak convergence $\rho_k \to \rho$ we have

$$\int_{\Lambda\times\mathbb{R}^d} F(\lambda,x)\mathfrak{m}^\lambda(\mathrm{d}x)\rho_k(\mathrm{d}\lambda) \to \int_{\Lambda\times\mathbb{R}^d} F(\lambda,x)\mathfrak{m}^\lambda(\mathrm{d}x)\rho(\mathrm{d}\lambda);$$

that is, $\mu_k \to \mu$ in $\mathcal{P}_2(\Lambda\times\mathbb{R}^d)$. Therefore, by (5.3.16), $\mu_k\to\mu$ in $C^0([\tau,T];\mathcal{P}_2(\Lambda\times\mathbb{R}^d))$, which implies that $\pi_{\mathbb{R}^d\sharp}\mu_k \to \pi_{\mathbb{R}^d\sharp}\mu$ in $C^0([\tau,T];\mathcal{P}_2(\mathbb{R}^d))$ as well.

Let now $\lambda\in K$ and assume that $\lambda \in \bigcap_{k\in\mathbb{N}}\operatorname{spt}(\rho_k)$. Standard stability of the Fokker–Planck equation gives the convergence $\bar m_k^\lambda \to \bar m^\lambda$ in $C^0([\tau,T];\mathcal{P}(\mathbb{R}^d))$, so we have all the ingredients to pass the identity (5.3.11) (that we are assuming to hold when $\pi_{\Lambda\sharp}\mu$ is finitely supported) to the limit $k\to\infty$. On the other hand, if $\lambda\notin\bigcap_{k\in\mathbb{N}}\operatorname{spt}(\rho_k)$ one can consider, in place of ρ_k, measures $\tilde\rho_k := \frac{N-1}{N}\rho_k + \frac{1}{N}\delta_\lambda$, so that still $\tilde\rho_k\to\rho$ weakly and the above argument can be still performed. ∎

LEMMA 5.3.8. *Let $\ell\in\mathbb{Z}_+$ and $\lambda\in\Lambda^\ell$. Let $(\theta_j)_{j\in[\![\ell]\!]}\subset\mathbb{R}_+$ satisfy $\sum_{j\in[\![\ell]\!]}\theta_j = 1$. Define*

$$U^{\lambda^i}(t,x,\mathfrak{m}) := U^{\lambda^i}\Big(t,x,\sum_{j\in[\![\ell]\!]}\theta_j\delta_{\lambda^j}\otimes\mathfrak{m}^j\Big) \tag{5.3.19}$$

for any $i\in[\![\ell]\!]$, $t\in[0,T]$, $x\in\mathbb{R}^d$ and $\mathfrak{m}\in\mathcal{P}_2(\mathbb{R}^d)^\ell$. Let $m_0\in\mathcal{P}_q(\mathbb{R}^d)^\ell$, $q>4$, and consider $m^\lambda := (m^{\lambda^j})_{j\in[\![\ell]\!]}$ solving

$$\begin{cases}\partial_t m^{\lambda^j} - \Delta m^{\lambda^j} - \operatorname{div}(\partial_x U^{\lambda^j}(t,x,m_t^\lambda)m^{\lambda^j}) = 0\,, & j\in[\![\ell]\!]\\ m^\lambda|_{t=\tau} = m_0\,. \end{cases} \tag{5.3.20}$$

Fix $\bar{\mathfrak{m}}\in\mathcal{P}_2(\mathbb{R}^d)^\ell$. Let $(I_j)_{j\in[\![\ell]\!]}$ be a partition of $\mathbb{N}$ and let Z be an $(\mathbb{R}^d)^N$-valued random variable such that

$$Z^i \sim \begin{cases}\bar{\mathfrak{m}}^j & \text{if } i = \min I_j^N \ \text{(for some } j\in[\![\ell]\!])\\ m_0^j & \text{if } i\in I_j^N\setminus\bigcup_{k\in[\![\ell]\!]}\{\min I_k^N\},\end{cases}$$

where $I_j^N := I_j\cap[\![N]\!]$. Define $\lambda_N\in\Lambda^N$ by letting $\lambda_N^i := \lambda^j$ if $i\in I_j^N$. Let $X^{I^N_\cdot}$ and $\tilde X^N$ be the solutions of

$$\mathrm{d}X^{I_j^N,i} = -\partial_x U^{\lambda^j}\big(t, X^{I_j^N,i}, (m_t^{\lambda^{-j}}, m_{X_t^{I_j^N}})\big)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i\,, \qquad i\in I_j^N,\ j\in[\![\ell]\!]\,,[5]$$

and

$$\mathrm{d}\tilde X^{N,i} = -\partial_x \mathfrak{u}_N(\lambda_N^i,\lambda_N^{-i},t,\tilde X_t^{N,i},\tilde X_t^{N,-i})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \qquad i\in[\![N]\!]\,,$$

[5]For the sake of clarity, we remark that, according to the notation we are adopting, the vector $(m^{\lambda^{-j}},m)$ is that with all coordinates given by m^{λ^i} but the j-th one, which is m, hence (recall

respectively, with $X_0^{I^N_\cdot} = Z = \tilde X_0^N$. *If*

$$\lim_{N\to\infty} \frac{\# I_j^N}{N} = \theta_j \quad \forall\, j \in [\![\ell]\!]\,,$$

then

$$\lim_{N\to\infty} \sup_{j\in[\![\ell]\!]} \sup_{i\in I_j^N} \mathbb{E}\Big[\sup_{t\in[0,T]} \big|\tilde X_t^{N,i} - X_t^{I_j^N,i}\big|^2\Big] = 0\,. \tag{5.3.21}$$

Proof. We first note that, since m_t^λ is a given flow of measures, each agent $i \in I_j^N$ is driven by the same drift

$$V^j\big(t, X^{I_j^N,i}, m_{X_t^{I_j^N}}\big) := -\partial_x U^{\lambda^j}\big(t, X^{I_j^N,i}, (m_t^{\lambda^{-j}}, m_{X_t^{I_j^N}})\big),$$

we can invoke Lemma 5.3.2 to have, in particular,

$$\lim_{N\to\infty} \mathbb{E}\Big[\sup_{t\in[0,T]} W_2(m_{X_t^{I_j^N}}, m_t^{\lambda^j})\Big] = 0; \tag{5.3.22}$$

this will be needed afterwards.

Now, if $i \in I_j^N$ we have

$$\begin{aligned}
\big|\tilde X_s^{N,i} - X_s^{I_j^N,i}\big|^2 &\\
\leqslant 2\int_0^s \Big(&\big|\partial_x \mathfrak{u}^N(\lambda^j, \lambda_N^{-i}, t, \tilde X_t^{N,i}, \tilde X_t^{N,-i}) \\
&\qquad - \partial_x \mathfrak{u}^N(\lambda^j, \lambda_N^{-i}, t, X_t^{I_j^N,i}, \tilde X_t^{N,-i})\big|^2 \\
&+ \big|\partial_x \mathfrak{u}^N(\lambda^j, \lambda_N^{-i}, t, X_t^{I_j^N,i}, \tilde X_t^{N,-i}) - \partial_x U(\lambda^j, t, X_t^{I_j^N,i}, m_{(\lambda_N:\tilde X_t^N)})\big|^2 \\
&+ \big|\partial_x U(\lambda^j, t, X_t^{I_j^N,i}, m_{(\lambda_N:\tilde X_t^N)}) \\
&\qquad - \partial_x U^{\lambda^j}\big(t, X_t^{I_j^N,i}, (m_t^{\lambda^{-j}}, m_{X_t^{I_j^N}})\big)\big|^2\Big)\,\mathrm{d}t\,;
\end{aligned} \tag{5.3.23}$$

note that, according to the notation we are adopting,

$$m_{(\lambda_N:\tilde X_t^N)} = \frac{1}{N}\sum_{k\in[\![N]\!]} \delta_{\lambda_N^k} \otimes \delta_{\tilde X_t^{N,k}}\,.$$

By the Lipschitz continuity of $\partial_x U$ (throughout the proof we will imply constants if they depend only on T, L_g, L_f and L_Λ),

$$\begin{aligned}
\big|\partial_x U(\lambda^j, t, X_t^{I_j^N,i}, m_{(\lambda_N:\tilde X_t^N)}) - \partial_x U^{\lambda^j}\big(t, X_t^{I_j^N,i}, (m_t^{\lambda^{-j}}, m_{X_t^{I_j^N}})\big)\big|^2 &\\
\lesssim W_2\Big(m_{(\lambda_N:\tilde X_t^N)},\ \theta_j\delta_{\lambda^j}\otimes m_{X_t^{I_j^N}} + \sum_{k\in[\![\ell]\!]\setminus\{j\}} \theta_k\delta_{\lambda^k}\otimes m_t^{\lambda^k}\Big)^2&;
\end{aligned} \tag{5.3.24}$$

---

also notation (5.3.19))

$$\partial_x U^{\lambda^j}(t, x, (m_t^{\lambda^{-j}}, m)) = \partial_x U^{\lambda^j}\Big(t, x, \theta_j\delta_{\lambda^j}\otimes m + \sum_{k\in[\![\ell]\!]\setminus\{j\}} \theta_k\delta_{\lambda^k}\otimes m_t^{\lambda^k}\Big).$$

on the other hand, recalling the definition of λ_N and then controlling the Wasserstein distance with the total variation (cf., e.g., [78, Theorem 6.15]), one easily gets

$$W_2\Big(m_{(\lambda_N:\tilde{X}_t^N)}, \sum_{k\in[\![\ell]\!]} \theta_k \delta_{\lambda^k} \otimes \frac{1}{\#I_k^N} \sum_{i\in I_k^N} \delta_{\tilde{X}_t^{N,i}}\Big)^2 \leqslant 2 \sup_{i\in[\![N]\!]} |\tilde{X}_t^{N,i}|^2 \sum_{k\in[\![\ell]\!]} \Big| \frac{\#I_k^N}{N} - \theta_k \Big| =: \gamma_t^N \,,$$

where $\mathbb{E}\big[\sup_{[0,T]} \gamma^N\big] \to 0$ as $N \to \infty$ thanks to the uniform boundedness of $\partial_x \mathfrak{u}_N$. Then applying the triangle inequality to the right-hand side of (5.3.24) we see that it is controlled by

$$W_2\Big(\sum_{k\in[\![\ell]\!]} \theta_k \delta_{\lambda^k} \otimes \frac{1}{\#I_k^N} \sum_{i\in I_k^N} \delta_{\tilde{X}_t^{N,i}} \,, \theta_j \delta_{\lambda^j} \otimes m_{X_t^{I_j^N}} + \sum_{k\in[\![\ell]\!]\setminus\{j\}} \theta_k \delta_{\lambda^k} \otimes m_t^{\lambda^k} \Big)^2 + \gamma_t^N \,. \tag{5.3.25}$$

Now, by the convexity of W_2^2 (cf., e.g., [78, Theorem 4.8]), the first term in (5.3.25) is in turn controlled by

$$\theta_j W_2\Big(\frac{1}{\#I_j^N} \sum_{i\in I_j^N} \delta_{\tilde{X}_t^{N,i}},\ m_{X_t^{I_j^N}} \Big)^2 + \sum_{\substack{1\leqslant k\leqslant \ell \\ k\neq j}} \theta_k W_2\Big(\frac{1}{\#I_k^N} \sum_{i\in I_k^N} \delta_{\tilde{X}_t^{N,i}},\ m_t^{\lambda^k} \Big)^2 \leqslant 2 \sum_{k\in[\![\ell]\!]} \frac{1}{\#I_k^N} \sum_{i\in I_k^N} \big|\tilde{X}_t^{N,i} - X_t^{I_k^N,i}\big|^2 + 2 \sum_{k\in[\![\ell]\!]} W_2(m_{X_t^{I_k^N}}, m_t^{\lambda^k})^2$$

Therefore, from (5.3.23) we have

$$\big|\tilde{X}_s^{N,i} - X_s^{I_j^N,i}\big|^2 \lesssim \int_0^s \bigg(\sum_{k\in[\![\ell]\!]} \frac{1}{\#I_k^N} \sum_{i\in I_k^N} \big|\tilde{X}_t^{N,i} - X_t^{I_k^N,i}\big|^2 + \delta_t^{N,j} \bigg) \mathrm{d}t \,, \tag{5.3.26}$$

where

$$\delta_t^{N,j} := \big|\partial_x \mathfrak{u}_N(\lambda^j, \lambda_N^{-i}, t, X_t^{I_j^N,i}, \tilde{X}_t^{N,-i}) - \partial_x U(\lambda^j, t, X_t^{I_j^N,i}, m_{(\lambda^N,\tilde{X}_t^N)})\big|^2 + \gamma_t^N + \sum_{k\in[\![\ell]\!]} W_2(m_{X_t^{I_k^N}}, m_t^{\lambda^k})^2$$

is such that $\mathbb{E}\big[\sup_{[0,T]} \delta_t^{N,j}\big] \to 0$ as $N \to \infty$ (cf. Theorem 5.2.6, property (5.3.22) and see the last argument in the proof Lemma 5.3.2). Averaging (5.3.26) over $i \in I_j$ and summing over $j \in [\![\ell]\!]$ we have

$$\sum_{1\leqslant j\leqslant \ell} \frac{1}{\#I_j^N} \sum_{i\in I_j^N} \big|\tilde{X}_s^{N,i} - X_s^{I_j^N,i}\big|^2 \lesssim \int_0^s \bigg(\ell \sum_{j\in[\![\ell]\!]} \frac{1}{\#I_j^N} \sum_{i\in I_k^N} \big|\tilde{X}_t^{N,i} - X_t^{I_j^N,i}\big|^2 + \sum_{j\in[\![\ell]\!]} \delta_t^{N,j} \bigg) \mathrm{d}t \,,$$

whence

$$\sum_{j\in[\![\ell]\!]}\frac{1}{\#I_j^N}\sum_{i\in I_j^N}\mathbb{E}\Big[\sup_{t\in[0,T]}\big|\tilde X_t^{N,i}-X_t^{I_j^N,i}\big|^2\Big]\to 0\quad\text{as } N\to\infty.$$

by Gronwall's lemma; then, plugging this back into (5.3.26), we obtain (5.3.21). ∎

Proof of Theorem 5.3.4. By Lemma 5.3.7 it suffices to prove the theorem by assuming that the first marginal is of the form $\rho=\sum_{j\in[\![\ell]\!]}\theta_j\delta_{\lambda^j}\in\mathcal{P}(\Lambda)$ for some (fixed) $\ell\in\mathbb{Z}_+$, $\theta_j\in\mathbb{Q}_+$ and $\lambda^j\in\Lambda$.

We now want to construct N-player games, where each player is suitably labelled with λ^j, $j\in[\![\ell]\!]$. The distribution of labels in the limit $N\to\infty$ must be ρ. It is elementary to construct a partition $(I_j)_{j\in[\![\ell]\!]}$ of $\mathbb{N}$ such that $\mathfrak{d}(I_j)=\theta_j$, where $\mathfrak{d}$ denotes the natural density;[6] indeed, it suffices to consider the minimum natural number D such that $D\theta_j\in\mathbb{N}$ for all $j\in[\![\ell]\!]$, take a partition $\{\Theta_j\}_{j\in[\![\ell]\!]}$ of $[\![D]\!]$ such that $\#\Theta_j=D\theta_j$ and set $I_j:=\Theta_j+D\mathbb{N}$. So we can define $\bar\lambda\in\ell^\infty(\mathbb{N};\Lambda)$ by letting $\bar\lambda^i:=\lambda^j$ if $i\in I_j$ and then $\lambda_N\in\Lambda^N$ for any $N\in\mathbb{N}$ by taking the first N coordinates of $\bar\lambda$.

With the notation of Lemma 5.3.8, we want argue as in the proof of Proposition 5.3.1, with the replacements

$$[\![N]\!]\rightsquigarrow I_j^N\quad\text{and}\quad \partial_xU^\lambda(t,x,m)\rightsquigarrow\partial_xU^{\lambda^j}(t,x,(m_t^{\lambda^{-j}},m))\,.$$

This crucial intermediate step is an "intra-population" propagation of chaos. As noted at the beginning of the proof of Lemma 5.3.8, we can exploit Lemma 5.3.2 and argue exactly as in the proof of Proposition 5.3.1 (note that $\#I_j^N\to\infty$ as $N\to\infty$) to obtain, for $i=\min I_j^N$,

$$\begin{aligned}(5.3.27)\quad \mathbb{E}\bigg[\int_0^T\Big(&\frac12\big|\partial_xU^{\lambda^j}\big(t,X_t^{I_j^N,i},(m_t^{\lambda^{-j}},m_{X_t^{I_j^N}})\big)\big|^2\\&+f^{\lambda^j}\big(X_t^{I_j^N,i},(m_t^{\lambda^{-j}},m_{X_t^{I_j^N}})\big)\Big)\mathrm{d}t+g^{\lambda^j}\big(X_T^{I_j^N,i},(m_T^{\lambda^{-j}},m_{X_T^{I_j^N}})\big)\bigg]\\
\xrightarrow[N\to\infty]{}\int_0^T\int_{\mathbb{R}^d}\Big(&\frac12\big|\partial_xU^{\lambda^j}(t,\cdot,m_t^\lambda)\big|^2\\&+f^{\lambda^j}(\cdot,m_t^\lambda)\Big)\mathrm{d}\bar m_t^{\lambda^j}\,\mathrm{d}t+\int_{\mathbb{R}^d}g^{\lambda^j}(\cdot,m_T^\lambda)\,\mathrm{d}\bar m_T^{\lambda^j}\,,\end{aligned}$$

where m^λ and $\bar m^\lambda$ solve, respectively, (5.3.20) and

$$\begin{cases}\partial_t\bar m^{\lambda^j}-\Delta\bar m^{\lambda^j}-\operatorname{div}(\partial_xU^{\lambda^j}(t,x,m_t^\lambda)\bar m^{\lambda^j})=0\,,\quad j\in[\![\ell]\!]\\ \bar m^\lambda|_{t=\tau}=\bar{\mathfrak{m}}\,,\end{cases}$$

both with initial time $\tau=0$.

[6] Recall that the natural density of $S\subset\mathbb{N}$ is defined as $\mathfrak{d}(S):=\lim_{N\to\infty}N^{-1}\#(S\cap[\![N]\!])$, if the limit exists.

At this point consider the representation formula for the value functions solving the Nash system constructed above,

$$
\begin{aligned}
(5.3.28)\quad & \mathbb{E}\,\mathfrak{u}_N(\lambda_N, 0, Z) \\
&= \mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}\big|\partial_x \mathfrak{u}_N(\lambda_N, t, \tilde{X}^N_t)\big|^2 + f_N(\bar{\lambda}^0, \tilde{X}^N_t)\Big)\mathrm{d}t + g_N(\bar{\lambda}^0, \tilde{X}^N_T)\Big],
\end{aligned}
$$

where $\tilde{X}^N$ is as in Lemma 5.3.8; in order to pass to the limit this identity, we need to compare $\tilde{X}^N$ and $X^{I^N_\cdot}$, which we have thus already done in Lemma 5.3.8. Having fixed $j \in [\![\ell]\!]$, note that we can assume without loss of generality that $0 \in I_j$ (otherwise, just take a different partition $\{\Theta_j\}_{j\in[\![\ell]\!]}$ at the beginning of Step 2); therefore, $\bar{\lambda}^0 = \lambda^j$, and one deduces (5.3.11) from (5.3.27), passing (5.3.28) to the limit. In order to motivate this last claim, for the sake of completeness, we prove that, for $h \in \{f, g\}$, one has

$$
\lim_{N\to\infty} \mathbb{E}\Big[\sup_{t\in[0,T]} \big|h^{\lambda^j}\big(X^{I^N_j,i}_t, (m^{\lambda^{-j}}_t, m_{X^{I^N_j}_t})\big) - h_N(\lambda^j, \tilde{X}^N_t)\big|\Big] = 0
$$

(the other analogous convergence that we will need, involving $\partial_x \mathfrak{u}_N$ and U, can be proved in a similar manner). Under our latest assumption, $X^{I^N_j,i} = X^{I^N_j,0}$, so, by the Lipschitz continuity of h and the triangle inequality,

$$
\begin{aligned}
&\big|h^{\lambda^j}\big(X^{I^N_j,i}_t, (m^{\lambda^{-j}}_t, m_{X^{I^N_j}_t})\big) - h_N(\lambda^j, \tilde{X}^N_t)\big|^2 \\
&\qquad\qquad \lesssim \big|X^{I^N_j,0}_t - \tilde{X}^{N,0}_t\big|^2 + W_2\Big(\theta_j m_{X^{I^N_j}_t} + \sum_{k\in[\![\ell]\!]\setminus\{j\}} \theta_k m^{\lambda^k}_t,\, m_{\tilde{X}^N_t}\Big)^2 \\
&\qquad\qquad\quad + W_2(m_{\tilde{X}^N_t}, m_{\tilde{X}^{N,-0}_t})^2 + \big|h^{\lambda^j}(\tilde{X}^{N,1}_t, m_{\tilde{X}^{N,-0}_t}) - h_N(\lambda^j, \tilde{X}^N_t)\big|^2;
\end{aligned}
$$

it is straightforward to show that, after taking the supremum over $t \in [0, T]$ and then the expectation, the third term in the left-hand side vanished as $N \to \infty$, while the first one and the last one go to 0 as well thanks to (5.3.21) and Remark 5.2.8, respectively. As for the second term, by the triangle inequality and convexity of W_2^2,

$$
\begin{aligned}
&W_1\Big(\theta_j m_{X^{I^N_j}_t} + \sum_{k\in[\![\ell]\!]\setminus\{j\}} \theta_k m^{\lambda^k}_t,\, m_{\tilde{X}^N_t}\Big) \\
&\quad \leqslant \theta_j W_2\Big(m_{X^{I^N_j}_t}, \frac{1}{\# I^N_j}\sum_{i\in I^N_j} \delta_{\tilde{X}^{N,i}_t}\Big)^2 + \sum_{k\in[\![\ell]\!]\setminus\{j\}} \theta_k W_2\Big(m^{\lambda^k}_t, \frac{1}{\# I^N_k}\sum_{i\in I^N_k} \delta_{\tilde{X}^{N,i}_t}\Big)^2 \\
&\qquad + W_2\Big(\sum_{k\in[\![\ell]\!]} \frac{\theta_k}{\# I^N_k}\sum_{i\in I^N_k} \delta_{\tilde{X}^{N,i}_t},\, m_{\tilde{X}^N_t}\Big)^2;
\end{aligned}
$$

by using the estimates obtained in the previous lemmata, $\mathbb{E}\sup_{t\in[0,T]}$ of the right-hand side can be shown to vanish as $N \to \infty$. ■

5.4. A continuum of MFG systems

Some considerations about the characterisation provided by (5.3.11)–(5.3.12)–(5.3.13) are now in order.

REMARK 5.4.1. If U^λ is sufficiently regular on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\Lambda \times \mathbb{R}^d)$ (e.g., according to Definition 2.4.1 in [**23**] of a classical solution to the first-order Master Equation, so, in particular it is differentiable with respect to μ), one can differentiate $t \mapsto \int_{\mathbb{R}^d} U^\lambda(t, \cdot, \mu_t)\, \mathrm{d}m_t^\lambda$ with $m_\tau^\lambda = \delta_x$ and use (5.3.11), (5.3.12), (5.3.13) to get, after standard computations,

$$\begin{aligned}
&- \partial_t U^\lambda(t,x,\mu) - \Delta_x U^\lambda(t,x,\mu) - \int_{\Lambda\times\mathbb{R}^d} \Delta_y \frac{\delta}{\delta\mu} U^\lambda(t,x,\mu,\lambda',y)\, \mu(\mathrm{d}\lambda', \mathrm{d}y) \\
&\quad + \frac{1}{2}\big|\partial_x U^\lambda(t,x,\mu)\big|^2 + \int_{\Lambda\times\mathbb{R}^d} D_y \frac{\delta}{\delta\mu} U^\lambda(t,x,\mu,\lambda',y) \cdot \partial_y U^{\lambda'}(t,y,\mu)\, \mu(\mathrm{d}\lambda', \mathrm{d}y) \\
&\qquad = f^\lambda(x, \pi_{\mathbb{R}^d\sharp}\mu).
\end{aligned}$$

This is exactly the form one would expect a Master Equation associated to our non-symmetric game to exhibit. In this sense, the limit game (5.3.11)–(5.3.12)–(5.3.13) can be considered as a *weak formulation of the Master Equation for a nonsymmetric game* parametrised by $\lambda \in \Lambda$. Since U^λ is Lipschitz continuous with respect to the measure variable, this formulation is very close to the one proposed in [**15**]. Moreover, it brings naturally a continuum of standard MFG systems, as described below.

THEOREM 5.4.2. *With the notation of Theorem 5.3.4, define*

$$u^\lambda(t,x) := U^\lambda(t,x,\mu_t)\,.$$

Then u solves the following (generalised) Mean Field system on $(0,T)\times\Lambda\times\mathbb{R}^d$:

$$\begin{cases}
-\partial_t u^\lambda - \Delta_x u^\lambda + \frac{1}{2}|D_x u^\lambda|^2 = f^\lambda(x, \pi_{\mathbb{R}^d\sharp}\mu_t) \\
\partial_t \mu - \Delta_x \mu - \mathrm{div}_x(D u^\lambda \mu) = 0 \\
u^\lambda(T,\cdot) = g^\lambda(\cdot, \pi_{\mathbb{R}^d\sharp}\mu_T), \quad \mu_0 = \mu\,,
\end{cases} \tag{5.4.1}$$

where both the Hamilton–Jacobi and the Fokker–Planck equations are satisfied in the classical sense.

PROOF. Fix $\lambda \in \Lambda$. Note that, by (5.3.11), for every τ and smooth $\bar{\mathfrak{m}}^\lambda$,

$$\begin{aligned}
\int_{\mathbb{R}^d} u^\lambda(\tau,\cdot)\, \mathrm{d}\bar{\mathfrak{m}}^\lambda = \int_\tau^T \int_{\mathbb{R}^d} \Big(\frac{1}{2}\big|D_x u^\lambda(s,\cdot)\big|^2 + f^\lambda(\cdot, \pi_{\mathbb{R}^d\sharp}\mu_s) \Big)\, \mathrm{d}\bar{m}_s^\lambda \mathrm{d}s \\
+ \int_{\mathbb{R}^d} g^\lambda(\cdot, \pi_{\mathbb{R}^d\sharp}\mu_T)\, \mathrm{d}\bar{m}_T^\lambda\,.
\end{aligned} \tag{5.4.2}$$

Let v be solution of the linear backward Cauchy problem

$$\begin{cases}
-\partial_t v - \Delta_x v + D_x u^\lambda \cdot D_x v = f^\lambda(x, \pi_{\mathbb{R}^d\sharp}\mu_t) + \frac{1}{2}\big|D_x u^\lambda\big|^2 \\
v(T,\cdot) = g^\lambda(\cdot, \pi_{\mathbb{R}^d\sharp}\mu_T).
\end{cases}$$

By standard linear theory (see, e.g., [**64**]), v solves the equation a.e., and since coefficients are in L^∞, it is globally bounded and locally in $W^{2,1}_p$ for every $p \in (1,\infty)$. By duality between the equation for v and the equation for $\bar{m}^\lambda$, for every $\tau \in [0,T]$ we have

$$\int_{\mathbb{R}^d} v(\tau,\cdot)\,\mathrm{d}\bar{\mathfrak{m}}^\lambda = \int_\tau^T \int_{\mathbb{R}^d} \Big(\frac{1}{2}\big|D_x u^\lambda(s,\cdot)\big|^2 + f^\lambda(\,\cdot\,,\pi_{\mathbb{R}^d\sharp}\mu_s)\Big)\,\mathrm{d}\bar{m}^\lambda_s \mathrm{d}s + \int_{\mathbb{R}^d} g^\lambda(\,\cdot\,,\pi_{\mathbb{R}^d\sharp}\mu_T)\,\mathrm{d}\bar{m}^\lambda_T\,.$$

Since the previous equation has the same right-hand side of (5.4.2), by the arbitrariness of $\bar{\mathfrak{m}}^\lambda$ and τ, we obtain $v \equiv u^\lambda$. Plugging back $D_x v \equiv D_x u^\lambda$ into the equation for v we obtain the assertion. That u^λ is a classical solution follows by standard parabolic regularity. ∎

Remark 5.4.3. If f and g are constant in λ, then (5.4.1) is the classic Mean Field system (this case corresponds to the situation in Proposition 5.3.1). We also mention that, as a byproduct, under assumptions of monotonicity, one can then deduce uniqueness of the limit function U by classic uniqueness results for the Mean Field system.

Remark 5.4.4. The above result says, within the setting of the previous example, that the Mean Field system arises from a sequence of N-player games built by choosing certain λ_N, provided that the labels λ^i_N are "sufficiently well distributed", in the sense that their empirical measures weakly converge to some $\rho \in \mathcal{P}(\Lambda)$ as $N \to \infty$. Limit properties of those games are then captured by the Mean Field system when one chooses $\mu = \int_\Lambda m^\lambda \rho(\mathrm{d}\lambda)$, being m^λ the distribution of λ-labelled players.

On the other hand, by a theorem of Hedrlín (see [**50**]), for each Mean Field system of that kind there exists a suitable choice of λ_N in order for it to arise from the respective sequence of N-player games.

Bibliography

[1] Ahuja, S. Wellposedness of Mean Field Games with common noise under a weak monotonicity condition. *SIAM Journal on Control and Optimization 54*, 1 (Jan 2016), 30–48.

[2] Andrews, B., and Clutterbuck, J. Time-interior gradient estimates for quasilinear parabolic equations. *Indiana Univ. Math. J. 58*, 1 (2009), 351–380.

[3] Aurell, A., Carmona, R., and Laurière, M. Stochastic graphon games: II. The linear-quadratic case. *Appl. Math. Optim. 85*, 3 (2022), Article No. 39, 33.

[4] Baldi, P. *Stochastic calculus.* Universitext. Springer, Cham, 2017. An introduction through theory and exercises.

[5] Bansil, M., and Mészáros, A. R. Hidden monotonicity and canonical transformations for mean field games and master equations, 2024. arXiv:2403.05426.

[6] Bardi, M. Explicit solutions of some linear-quadratic mean field games. *Netw. Heterog. Media 7*, 2 (2012), 243–261.

[7] Bayraktar, E., Cecchin, A., Cohen, A., and Delarue, F. Finite state mean field games with Wright-Fisher common noise as limits of N-player weighted games. *Math. Oper. Res. 47*, 4 (2022), 2840–2890.

[8] Bayraktar, E., Wu, R., and Zhang, X. Propagation of chaos of forward-backward stochastic differential equations with graphon interactions. *Appl. Math. Optim. 88*, 1 (2023), Article No. 25, 44.

[9] Bensoussan, A., and Frehse, J. C^α-regularity results for quasilinear parabolic systems. *Comment. Math. Univ. Carolin. 31*, 3 (1990), 453–474.

[10] Bensoussan, A., and Frehse, J. Smooth solutions of systems of quasilinear parabolic equations. *ESAIM: Control, Optimisation and Calculus of Variations 8* (2002), 169–193.

[11] Bensoussan, A., Frehse, J., Peng, S., and Yam, S. C. P. *Parabolic Equations with Quadratic Growth in* $\mathbb{R}^n$. Springer International Publishing, Cham, 2019, pp. 91–110.

[12] Bensoussan, A., Huang, T., and Laurière, M. Mean field control and mean field game models with several populations. *Minimax Theory Appl. 3*, 2 (2018), 173–209.

[13] Bensoussan, A., Li, B., and Yam, S. C. P. Linear quadratic extended mean field games and control problems, 2023. arXiv:2311.05176.

[14] Bensoussan, A., Sung, K. C. J., Yam, S. C. P., and Yung, S. P. Linear-quadratic mean field games. *J. Optim. Theory Appl. 169*, 2 (2016), 496–529.

[15] Bertucci, C., Lasry, J.-M., and Lions, P.-L. On Lipschitz solutions of mean field games master equations, 2023. arXiv:2302.05218.

[16] Bertucci, C., Lasry, J.-M., and Lions, P.-L. A singular infinite dimensional Hamilton–Jacobi–Bellman equation arising from a storage problem, 2023. arXiv:2210.02780.

[17] Bianca, C., and Dogbe, C. Regularization and propagation in a Hamilton–Jacobi–Bellman-type equation in infinite-dimensional Hilbert space. *Symmetry 16*, 8 (2024).

[18] Bogachev, V. I., Krylov, N. V., Röckner, M., and Shaposhnikov, S. V. *Fokker-Planck-Kolmogorov equations*, vol. 207 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2015.

[19] Caines, P. E., and Huang, M. Graphon mean field games and their equations. *SIAM J. Control Optim. 59*, 6 (2021), 4373–4399.

[20] Calvia, A., Cappa, G., Gozzi, F., and Priola, E. HJB equations and stochastic control on half-spaces of Hilbert spaces. *J. Optim. Theory Appl. 198* (2023), 710–744.

[21] Cardaliaguet, P. Notes on mean field games. *Lecture notes* (2012). From P.-L. Lions' lectures at Collège de France.

[22] Cardaliaguet, P., Cirant, M., and Porretta, A. Splitting methods and short time existence for the master equations in mean field games. *J. Eur. Math. Soc. (JEMS) 25*, 5 (2023), 1823–1918.

[23] Cardaliaguet, P., Delarue, F., Lasry, J.-M., and Lions, P.-L. *The master equation and the convergence problem in mean field games*, vol. 201 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2019.

[24] Cardaliaguet, P., and Porretta, A. *An introduction to mean field game theory*, vol. 2281 of *Lecture Notes in Math.* Springer, Cham, 2020, pp. 1–158.

[25] Carmona, R., and Delarue, F. *Probabilistic theory of mean field games with applications. I*, vol. 83 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. Mean field FBSDEs, control, and games.

[26] Carmona, R., and Delarue, F. *Probabilistic theory of mean field games with applications. II*, vol. 84 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. Mean field games with common noise and master equations.

[27] Cecchin, A. Finite state N-agent and mean field control problems. *ESAIM Control Optim. Calc. Var. 27* (2021), Article No. 31, 33.

[28] Cirant, M. Multi-population mean field games systems with Neumann boundary conditions. *J. Math. Pures Appl. (9) 103*, 5 (2015), 1294–1315.

[29] Cirant, M., and Porretta, A. Long time behavior and turnpike solutions in mildly non-monotone mean field games. *ESAIM Control Optim. Calc. Var. 27* (2021), Article No. 86, 40.

[30] Cirant, M., and Redaelli, D. F. On the unimportance of distant players in graph-based open-loop differential games, 2024. Preprint.

[31] Cirant, M., and Redaelli, D. F. A priori estimates and large population limits for some nonsymmetric Nash systems with semimonotonicity, 2024. arXiv:2406.10822.

[32] Cirant, M., and Redaelli, D. F. Some remarks on Linear-Quadratic closed-loop games with many players. *Dyn. Games Appl.* (2024).

[33] Conforti, G. Weak semiconvexity estimates for Schrödinger potentials and logarithmic Sobolev inequality for Schrödinger bridges. *Probab. Theory Relat. Fields 189* (2024), 1045–1071.

[34] Da Prato, G., and Zabczyk, J. *Second Order Partial Differential Equations in Hilbert Spaces*. London Mathematical Society Lecture Note Series. Cambridge University Press, 2002.

[35] Daudin, S., Delarue, F., and Jackson, J. On the optimal rate for the convergence problem in Mean Field Control, 2023. arXiv:2305.08423.

[36] Daudin, S., Delarue, F., and Jackson, J. On the optimal rate for the convergence problem in mean field control. *Journal of Functional Analysis 287*, 12 (2024), 110660.

[37] Delarue, F. Mean field games: a toy model on an Erdös-Renyi graph. In *Journées MAS 2016 de la SMAI—Phénomènes complexes et hétérogènes*, vol. 60 of *ESAIM Proc. Surveys*. EDP Sci., Les Ulis, 2017, pp. 1–26.

[38] Delarue, F., and Foguen Tchuendom, R. Selection of equilibria in a linear quadratic mean-field game. *Stochastic Process. Appl. 130*, 2 (2020), 1000–1040.

[39] Djete, M. F. Large population games with interactions through controls and common noise: convergence results and equivalence between open-loop and closed-loop controls. *ESAIM Control Optim. Calc. Var. 29* (2023), Article No. 39, 42.

[40] Eberle, A. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields 166*, 3-4 (2016), 851–886.

[41] Feleqi, E. The derivation of ergodic mean field game equations for several populations of players. *Dyn. Games Appl. 3*, 4 (2013), 523–536.

[42] Feng, Y., Fouque, J.-P., and Ichiba, T. Linear-quadratic stochastic differential games on directed chain networks, 2020. arXiv:2003.08840.

[43] Fischer, M. On the connection between symmetric N-player games and mean field games. *Ann. Appl. Probab. 27*, 2 (2017), 757–810.

[44] Frehse, J. Remarks on diagonal elliptic systems. In *Partial differential equations and calculus of variations*, vol. 1357 of *Lecture Notes in Math.* Springer, Berlin, 1988, pp. 198–210.

[45] Friedman, A. *Partial Differential Equations of Parabolic Type.* Prentice-Hall, 1964.

[46] Friedman, A. Stochastic differential games. *J. Differential Equations 11* (1972), 79–108.

[47] Gangbo, W., Mészáros, A. R., Mou, C., and Zhang, J. Mean field games master equations with nonseparable Hamiltonians and displacement monotonicity. *Ann. Probab. 50*, 6 (2022), 2178–2217.

[48] Graber, P. J. Linear quadratic mean field type control and mean field games with common noise, with application to production of an exhaustible resource. *Appl. Math. Optim. 74*, 3 (2016), 459–486.

[49] Graber, P. J., and Mészáros, A. R. On monotonicity conditions for mean field games. *J. Funct. Anal. 285*, 9 (2023), Article No. 110095, 45.

[50] Hedrlín, Z. On integration in compact metric spaces. *Commentationes Mathematicae Universitatis Carolinae 002*, 4 (1961), 17–19.

[51] Huang, M., Malhamé, R. P., and Caines, P. E. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst. 6*, 3 (2006), 221–251.

[52] Ito, K. Existence of solutions to the Hamilton–Jacobi–Bellman equation under quadratic growth conditions. *Journal of Differential Equations 176*, 1 (2001), 1–28.

[53] Jackson, J. On quasilinear parabolic systems and FBSDEs of quadratic growth. *Ann. Appl. Probab. 34*, 1A (2024), 357–387.

[54] Jackson, J., and Lacker, D. Approximately optimal distributed stochastic controls beyond the mean field setting, 2023. arXiv:2301.02901.

[55] Jackson, J., and Tangpi, L. Quantitative convergence for displacement monotone mean field games with controlled volatility, 2023. arXiv:2304.04543.

[56] Kružkov, S. N. Results on the nature of the continuity of solutions of parabolic equations, and certain applications thereof. *Mat. Zametki 6* (1969), 97–108.

[57] Krylov, N. V. *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, vol. 12 of *Graduate Studies in Mathematics*. American Mathematical Society, 1996.

[58] Krylov, N. V. *Lectures on Elliptic and Parabolic Equations in Sobolev Spaces*, vol. 96 of *Graduate Studies in Mathematics*. American Mathematical Society, 2008.

[59] Lacker, D. On the convergence of closed-loop Nash equilibria to the mean field game limit. *Ann. Appl. Probab. 30*, 4 (2020), 1693–1761.

[60] Lacker, D., and Le Flem, L. Closed-loop convergence for mean field games with common noise. *Ann. Appl. Probab. 33*, 4 (2023), 2681–2733.

[61] Lacker, D., Ramanan, K., and Wu, R. Local weak convergence for sparse networks of interacting processes. *The Annals of Applied Probability 33*, 2 (2023), 843–888.

[62] Lacker, D., and Soret, A. A case study on stochastic games on large graphs in mean field and sparse regimes. *Math. Oper. Res. 47*, 2 (2022), 1530–1565.

[63] Lacker, D., and Soret, A. A label-state formulation of stochastic graphon games and approximate equilibria on large networks. *Math. Oper. Res. 48*, 4 (2023), 1987–2018.

[64] Ladyzenskaja, O. A., Solonnikov, V. A., and Ural'ceva, N. N. *Linear and quasilinear equations of parabolic type*. Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1968.

[65] Lasry, J.-M., and Lions, P.-L. Mean field games. *Jpn. J. Math. 2*, 1 (2007), 229–260.

[66] Li, M., Mou, C., Wu, Z., and Zhou, C. Linear-quadratic mean field games of controls with non-monotone data. *Trans. Amer. Math. Soc. 376*, 6 (2023), 4105–4143.

[67] Li, X., Shi, J., and Yong, J. Mean-field linear-quadratic stochastic differential games in an infinite horizon. *ESAIM Control Optim. Calc. Var. 27* (2021), Article No. 81, 40.

[68] Lions, P.-L. Cours du Collège de France. Available at https://www.college-de-france.fr/fr/agenda/cours/jeux-champ-moyen-suite-2, 2008.

[69] Lunardi, A. *Analytic semigroups and optimal regularity in parabolic problems*, vol. 16 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Verlag, Basel, 1995.

[70] Masiero, F. HJB equations in infinite dimensions under weak regularizing properties. *J. Evol. Equ. 16* (2016), 789–824.

[71] Masiero, F., and Richou, A. HJB equations in infinite dimensions with locally Lipschitz Hamiltonian and unbounded terminal condition. *Journal of Differential Equations 257*, 6 (2014), 1989–2034.

[72] Mészáros, A. R., and Mou, C. Mean field games systems under displacement monotonicity. *SIAM J. Math. Anal. 56*, 1 (2024), 529–553.

[73] Mou, C., and Zhang, J. Mean field game master equations with anti-monotonicity conditions, 2022. arXiv:2201.10762.

[74] Peyré, G., and Cuturi, M. Computational optimal transport. *Foundations and Trends in Machine Learning 11*, 5-6 (2019), 355–607.

[75] Porretta, A., and Priola, E. Global Lipschitz regularizing effects for linear and nonlinear parabolic equations. *J. Math. Pures Appl. (9) 100*, 5 (2013), 633–686.

[76] Possamaï, D., and Tangpi, L. Non-asymptotic convergence rates for mean-field games: weak formulation and McKean-Vlasov BSDEs, 2021. arXiv:2105.00484.

[77] Redaelli, D. F. Short-time existence and uniqueness for some infinite-dimensional Nash systems, 2024. arXiv:2411.18356.

[78] Villani, C. *Optimal transport*, vol. 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.