Finite Difference Methods for Mean Field Games Systems

Simone Cacace and Fabio Camilli

Abstract We discuss convergence results for a class of finite difference schemes approximating Mean Field Games systems either on the torus or a network. We also propose a quasi-Newton method for the computation of discrete solutions, based on a least squares formulation of the problem. Several numerical experiments are carried out including the case with two or more competing populations.

Keywords Mean field games · Networks · Numerical methods · Finite difference · Newton-like methods

1 Introduction

In this paper we describe a class of finite difference methods for the approximation of the stationary Mean Field Games (MFG in short) system

$$\begin{cases} -\nu\Delta u + H(x, Du) + \lambda = V[m] & x \in \mathcal{T}, \\ \nu\Delta m + \operatorname{div}\left(m\, \frac{\partial H}{\partial p}(x, Du)\right) = 0 & x \in \mathcal{T}, \\ \int_{\mathcal{T}} u(x)dx = 0, \int_{\mathcal{T}} m(x)dx = 1,\ m \geq 0\,, \end{cases} \tag{1.1}$$

where $\mathcal{T}$ can be either the unit torus $\mathbb{T}^d = [0,1]^d$ or a network Γ. The system consists in a couple of PDEs, respectively a Hamilton-Jacobi-Bellman equation and a Fokker-Planck equation plus normalization conditions on both u and m. The unknowns are the value function u, the density m and the ergodic constant λ and

S. Cacace
Dip. di Matematica e Fisica, Università degli Studi Roma Tre, Rome, Italy
e-mail: cacace@mat.uniroma3.it

F. Camilli (✉)
Dip. di Scienze di Base e Applicate per l'Ingegneria, "Sapienza" Università di Roma, Rome, Italy
e-mail: camilli@sbai.uniroma1.it

P. Cardaliaguet et al. (eds.), *PDE Models for Multi-Agent Phenomena*,
Springer INdAM Series 28, https://doi.org/10.1007/978-3-030-01947-1_2

the system also involves the scalar Hamiltonian $H(x, p)$ and the potential V (for a general presentation of the theory of Mean Field Games we refer [12, 15]). The results are based on the papers in [1, 3, 4, 6, 7] and include existence, uniqueness and regularity of the approximate solution, convergence of the scheme and efficient resolution of the discrete problem.

After the introduction of the MFG theory, an important research activity has been pursued for the approximation of the different types of MFG models and several papers have been devoted to this topic. Besides the finite difference method we describe in this chapter, we mention among the others: the semi-Lagrangian scheme proposed in [9]; the optimization algorithm connected with the optimal control interpretation of the MFG system in [14]; the monotone scheme in [13] which exploits the equivalence between the MFG system and a linear system in the case of a quadratic Hamiltonian; the gradient-flow method based on the variational characterization of certain MFG systems in [5].

A numerical method for MFG systems has to face several difficulties: the system is strongly coupled in both the equations, i.e. via the potential term V in the Hamilton-Jacobi-Bellman equation and via the drift term $\frac{\partial H}{\partial p}(x, Du)$ in the Fokker-Planck equation; in the stationary case the system is formally overdetermined, involving three unknowns (u, m, λ) and two equations, while in the evolutive case it has forward-backward structure with respect to the time variable; the approximation of the Hamilton-Jacobi-Bellman equation presents the typical curse of dimensionality issue complicated furthermore by the coupled structure; the constraint $m \geq 0$ may be difficult to impose for algorithms based on Gradient and Newton methods; moreover, in order to obtain convergence and error estimates, a numerical method for MFG systems should reproduce at a discrete level some main properties of the continuous problem: for example, it is well-known that the Fokker-Planck equation in MFG systems is the adjoint equation associated to the linearization of the Hamilton-Jacobi-Bellman equation and, indeed, this relation is usually employed to get several properties of the solution to the problem.

The numerical method introduced in [1] and described in Sect. 2 is designed to reproduce at the discrete level the same adjoint structure of the continuous system. Discretizing the Hamilton-Jacobi-Bellman equation via standard finite differences, then the approximation of the Fokker-Planck equation is obtained by means of the weak formulation of the linearization of the first equation. The adjoint structure of the discrete problem allows one to obtain, as in the continuous case, several properties of the discrete solution, such as existence, uniqueness and regularity. Moreover, since the continuous and the approximate problems have the same adjoint structure, convergence of the scheme and error estimates are obtained by substituting the continuous solution in the discrete problem and estimating the truncation error (see [2, 4]).

In the recent times, there has been an increasing interest in the study of differential models on networks, and in [8] we extend the MFG theory to this framework. In [7], we consider the numerical approximation of the network problem and, following the approach in [1], we find an approximation of the transition conditions at the vertices which preserves, at the discrete level, the adjoint structure

of the continuous problem. Also in this case, employing the similarity between the continuous and the approximate problems, we are able to prove the convergence of the scheme. The scheme for the network problem is described in Sect. 3.

Since MFG theory introduces an effective and efficient methodology for handling a wide variety of applications in different fields, it is particularly relevant to design efficient solvers for the discrete problem. Section 4 is devoted to a new method proposed in [6] which allows to compute solutions of (1.1) avoiding costly large-time and ergodic approximations usually employed in this framework. Indeed, once an effective discretizion of (1.1) is introduced, the discrete problem is solved directly, by interpreting the ergodic constant λ as an unknown of the problem and computing the solution of the overdeterminated system by a Newton-like method for inconsistent nonlinear systems. A large collection of numerical tests in dimensions one and two shows the performance of the proposed method, both in terms of accuracy and computational time.

2 A Finite Difference Scheme for Mean Field Games on the Torus

In this section we consider system (1.1) on the torus $\mathbb{T}^d$, i.e. with periodic boundary conditions. The Hamiltonian $H(x, p)$ is assumed to be convex w.r.t. p and regular w.r.t. x and p. The potential term V may be either a local operator, i.e. $V[m(\cdot)](x) = F(m(x))$ for some regular function F; or a non local operator which continuously maps the set of probability measures on $\mathbb{T}^d$ to a bounded subset of the Lipschitz functions on $\mathbb{T}^d$.

To simplify the notations, we assume that the dimension of the state space is $d = 2$, but the scheme can be easily generalized to any dimension. Hence, let $\mathbb{T}^2_h$ be a uniform grid on the two-dimensional torus with step h (assuming that $N_h = 1/h$ is an integer) and denote by $x_{i,j}$ a typical grid node in $\mathbb{T}^2_h$. The values of u and m at $x_{i,j}$ are approximated, respectively by $U_{i,j}$ and $M_{i,j}$. For a grid function U, we consider the finite difference operators

$$(D_1^+ U)_{i,j} = \frac{U_{i+1,j} - U_{i,j}}{h}, \qquad (D_2^+ U)_{i,j} = \frac{U_{i,j+1} - U_{i,j}}{h},$$

and define

$$[D_h U]_{i,j} = \left((D_1^+ U)_{i,j}, (D_1^+ U)_{i-1,j}, (D_2^+ U)_{i,j}, (D_2^+ U)_{i,j-1}\right)^T,$$
$$(\Delta_h U)_{i,j} = -\frac{4U_{i,j} - U_{i+1,j} - U_{i-1,j} - U_{i,j+1} - U_{i,j-1}}{h^2},$$

where T denotes the transposition operator. We approximate $H(\cdot, \nabla u)(x_{i,j})$ by $g(x_{i,j}, [D_h U]_{i,j})$, where the numerical Hamiltonian is a function $g : \mathbb{T}^2 \times \mathbb{R}^4 \to \mathbb{R}$, $(x, q_1, q_2, q_3, q_4) \to g(x, q_1, q_2, q_3, q_4)$ satisfying

(**G1**) *monotonicity*: g is non increasing w.r.t. q_1, q_3 and nondecreasing w.r.t. to q_2, q_4.
(**G2**) *consistency:* $g(x, q_1, q_1, q_2, q_2) = H(x, q)\ \forall x \in \mathbb{T}^2, \forall q = (q_1, q_2) \in \mathbb{R}^2$.
(**G3**) *differentiability*: g is of class C^1.
(**G4**) *convexity*: for all $x \in \mathbb{T}^2$, $(q_1, q_2, q_3, q_4) \to g(x, q_1, q_2, q_3, q_4)$ is convex.

Numerical Hamiltonians fulfilling these requirements are provided by Lax-Friedrichs or Godunov type schemes, see [16].

The operator $V[m](x_{i,j})$ is approximated by $V_h[M]_{i,j}$. We assume that $V_h[M]$ can be computed in practice. For example, if $V[m]$ is defined as the solution w of the equation $\Delta^2 w + w = m$ in $\mathbb{T}^2$, (Δ^2 being the bi-laplacian), then one can define $V_h[M]$ as the solution W of the discrete problem $\Delta_h^2 W + W = M$ in $\mathbb{T}_h^2$. If V is a local operator, i.e. $V[m](x) = F(m(x))$, then $V_h[M]_{i,j} = F(M_{i,j})$.

For a generic pair of grid functions U, V we define the scalar product $(U, V)_2 = h^2 \sum_{0 \le i,j < N_h} U_{i,j} V_{i,j}$ and we consider the compact and convex set

$$\mathcal{K}_h = \{M = (M_{i,j})_{0 \le i,j < N_h} : (M, \overline{1})_2 = 1; \quad M_{i,j} \ge 0\},$$

where $\overline{1}$ denotes the N_h^2-tuple with all components equal to 1. Note that $\mathcal{K}_h$ can be viewed as the set of the discrete probability measures on $\mathbb{T}_h^2$.

We make the following assumptions on the potential term, V being local or not:

(**V1**) V_h is continuous.
(**V2**) V_h is monotone, i.e.

$$\left(V_h[M] - V_h[\tilde{M}], M - \tilde{M}\right)_2 \le 0 \Rightarrow V_h[M] = V_h[\tilde{M}].$$

If $V[m](x) = F(m(x))$, the function F being continuous from $\mathbb{R}^+$ to $\mathbb{R}$, then V_h is continuous on the set of nonnegative grid functions.

If V is a nonlocal operator, we assume that the discrete operator V_h satisfies the following additional properties:

(**V3**) There exists a constant C independent of h such that, for every grid function $M \in \mathcal{K}_h$, it holds

$$\|V_h[M]\|_\infty \le C, \qquad \left|(V_h[M])_{i,j} - (V_h[M])_{k,\ell}\right| \le C d(x_{i,j}, x_{k,\ell})$$

where $d(x, y)$ is the distance between the two points x and y in the torus $\mathbb{T}^2$.

(**V4**) There exists a continuous, bounded function $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\omega(0) = 0$ and such that, for all $m \in \mathcal{K} := \{m \in L^1(\mathbb{T}^2) : \int_{\mathbb{T}^2} m dx = 1,\ m \geq 0\}$ and for all $M \in \mathcal{K}_h$,

$$\| V[m] - V_h[M] \|_{L^\infty(\mathbb{T}^2_h)} \leq \omega\left(\|m - \mathcal{I}_h M\|_{L^1(\mathbb{T}^2)}\right), \tag{2.1}$$

where $\mathcal{I}_h M$ is the piecewise constant function taking the value $M_{i,j}$ in the square $\left\{|x - x_{i,j}|_\infty \leq h/2\right\}$.

Remark 2.1 If $m \in \mathcal{K}$ and $\mathcal{P}_h m$ is the grid function whose value at $x_{i,j}$ is

$$\int_{|x - x_{i,j}|_\infty \leq h/2} m(x) dx,$$

then $\mathcal{P}_h m \in \mathcal{K}_h$ and (2.1) implies the convergence of the approximation to $V[m]$, i.e.

$$\lim_{h \to 0} \sup_{m \in \mathcal{K}} \| V[m] - V_h[\mathcal{P}_h m] \|_{L^\infty(\mathbb{T}^2_h)} = 0.$$

To approximate the Hamilton-Jacobi-Bellman equation in (1.1), we consider the scheme

$$-\nu(\Delta_h U)_{i,j} + g(x_{i,j}, [D_h U]_{i,j}) + \Lambda = (V_h[M])_{i,j}, \tag{2.2}$$

with $\Lambda \in \mathbb{R}$ and subject to the normalization condition $(U, \overline{1})_2 = 0$.

In order to approximate the Fokker-Planck equation in (1.1), we consider the linearization of the Hamilton-Jacobi-Bellman equation at u in the direction w

$$-\nu \Delta w + \frac{\partial H}{\partial p}(x, Du)\, Dw = 0.$$

Note that the weak formulation of previous equation involves the term

$$-\int_{\mathbb{T}^2} \operatorname{div}\left(m \frac{\partial H}{\partial p}(x, \nabla u)\right) w\, dx.$$

which, by periodicity, yields

$$\int_{\mathbb{T}^2} m \frac{\partial H}{\partial p}(x, \nabla u) \cdot \nabla w\, dx$$

for any test function w, and it can be approximated by

$$h^2 \sum_{i,j} M_{i,j} \nabla_q g(x_{i,j}, [D_h U]_{i,j}) \cdot [D_h W]_{i,j}. \tag{2.3}$$

By discrete integration by parts on $\mathbb{T}_h^2$, the sum in (2.3) is readily rewritten as

$$h^2 \sum_{i,j} \mathcal{T}_{i,j}(U, M) W_{i,j},$$

where the operator $\mathcal{T}$ is defined as follows:

$$\mathcal{T}_{i,j}(U, M) = \frac{1}{h}\Big[\begin{array}{l} M_{i,j} \frac{\partial g}{\partial q_1}(x_{i,j}, [D_h U]_{i,j}) - M_{i-1,j} \frac{\partial g}{\partial q_1}(x_{i-1,j}, [D_h U]_{i-1,j}) \\ + M_{i+1,j} \frac{\partial g}{\partial q_2}(x_{i+1,j}, [D_h U]_{i+1,j}) - M_{i,j} \frac{\partial g}{\partial q_2}(x_{i,j}, [D_h U]_{i,j}) \end{array} \Big] +$$

$$\frac{1}{h}\Big[\begin{array}{l} M_{i,j} \frac{\partial g}{\partial q_3}(x_{i,j}, [D_h U]_{i,j}) - M_{i,j-1} \frac{\partial g}{\partial q_3}(x_{i,j-1}, [D_h U]_{i,j-1}) \\ + M_{i,j+1} \frac{\partial g}{\partial q_4}(x_{i,j+1}, [D_h U]_{i,j+1}) - M_{i,j} \frac{\partial g}{\partial q_4}(x_{i,j}, [D_h U]_{i,j}) \end{array} \Big].$$

In conclusion the second equation in (1.1) is approximated by

$$\nu(\Delta_h M)_{i,j} + \mathcal{T}_{i,j}(U, M) = 0, \tag{2.4}$$

subject to the normalization conditions $(M, \overline{1})_2 = 1$, $M \geq 0$. As for the continuous problem, the operator $M \mapsto \big(- \nu(\Delta_h M)_{i,j} - \mathcal{T}_{i,j}(U, M)\big)_{i,j}$ is the adjoint of the linearized version of the operator $u \mapsto \big(- \nu(\Delta_h U)_{i,j} + g(x_{i,j}, [D_h U]_{i,j})\big)_{i,j}$. This is a crucial property in view of the uniqueness and the convergence of the scheme.

Summarizing the finite difference scheme for the system (1.1) is: for all $0 \leq i, j < N_h$

$$\begin{cases} -\nu(\Delta_h U)_{i,j} + g(x_{i,j}, [D_h U]_{i,j}) = (V_h[M])_{i,j}), \\ \nu(\Delta_h M)_{i,j} + \mathcal{T}_{i,j}(U, M) = 0, \\ (U, 1)_2 = 0, \ (M, 1)_2 = 1, \ M \geq 0. \end{cases} \tag{2.5}$$

The following theorem is proved in [1].

Theorem 2.1 *If the numerical Hamiltonian g satisfies **(G1)–(G3)** and the potential V satisfies **(V1)**, then (2.5) has a solution (U, M, Λ). Moreover if g also satisfies **(G4)** and V_h also satisfies **(V2)**, then the solution is unique.*

In the previous result ν can also vanish, hence the deterministic case is included. The proof of existence of a solution to (2.5) is based on a fixed point argument and careful estimates of the Lipschitz norm of the solution of (2.2), while uniqueness is proved with a duality argument similar to the one in [15] for the continuous problem. The next proposition gives a regularity result for the solution (2.5) with an estimate of the norm uniform in h.

Proposition 2.1 *Under the same assumptions of Theorem 2.1, assume moreover that* $\nu > 0$ *and*

$$\left|\frac{\partial g}{\partial x}(x,(q_1,q_2,q_3,q_4))\right| \leq C(1+|q_1|+|q_2|+|q_3|+|q_4|). \tag{2.6}$$

Then there exists a constant C *independent of* h *such that*

$$\|U\|_\infty + \|D_h U\|_\infty \leq C. \tag{2.7}$$

We now focus on the convergence of the scheme (2.5). In the rest of this section we make the following additional assumptions

- $\nu > 0$;
- the Hamiltonian is of the form

$$H(x,p) = \mathcal{H}(x) + |p|^\beta \tag{2.8}$$

  with the function $\mathcal{H} \in C^1(\mathbb{T}^2)$ and $\beta > 1$;
- the system (1.1) admits a unique classical solution (u, m, λ).

To approximate the Hamiltonian in (2.8) we consider a numerical Hamiltonian of the form

$$g(x,q) = \mathcal{H}(x) + G(q_1^-, q_2^+, q_3^-, q_4^+), \tag{2.9}$$

where, for a real number r, $r^+ = \max(r,0)$, $r^- = \max(-r,0)$ and $G : \mathbb{R}^4 \to \mathbb{R}_+$ is given by

$$G(p) = |p|^\beta = (p_1^2 + p_2^2 + p_3^2 + p_4^2)^{\frac{\beta}{2}}.$$

Assumptions **(G1)–(G4)** are satisfied by the numerical Hamiltonian in (2.9), hence Theorem 2.1 guarantees existence and uniqueness of the solution. In the following we denote by u^h (resp. m^h) the piecewise bilinear function in $C(\mathbb{T}^2)$ obtained by interpolating the values $U^h_{i,j}$ (resp $M^h_{i,j}$) of the solution (U^h, M^h, Λ^h) to (2.5) at the nodes of the space grid. For the convergence analysis we distinguish the cases of a nonlocal potential and the case of a local one.

2.1 Convergence for V Nonlocal Operator

We assume that V is monotone, nonlocal and smooth. In this case it is known that there exists a unique classical solution (u, m, λ) of (1.1) such that $m > 0$ [15]. Note that since g in (2.9) verifies condition (2.6), the regularity estimate (2.7) holds and U is uniformly Lipschitz continuous.

Theorem 2.2 *Consider the numerical Hamiltonian given by (2.9) and a discrete potential V_h such that **(V1)–(V4)** hold.*

***The case** $\beta \geq 2$: As h goes to 0, the functions u^h converge to u in $W^{1,\beta}(\mathbb{T}^2)$, the functions m^h converge to m in $H^1(\mathbb{T}^2)$, and Λ^h tends to λ.*

***The case** $\beta \in (1,2)$: As h goes to 0, the functions u^h converge to u in $W^{1,2}(\mathbb{T}^2)$, the functions m^h converge to m in $L^2(\mathbb{T}^2)$, and Λ^h tends to λ.*

2.2 Convergence for V Local Operator

If V is a local operator, i.e. $V[m](x) = F(m(x))$, existence of a classical solution to (1.1) for any $\beta > 1$ holds, for example, if F is non decreasing and satisfies

$$mF(m) \geq \delta |F(m)|^\gamma - C_1, \qquad \forall m \geq 0 \tag{2.10}$$

for some constant $C_1 > 0$ and $\gamma > 2$ (being 2 the dimension of the space). In the local case, there are no a priori Lipschitz estimates on U such as (2.7). Since these estimates are used several times in the proof of Theorem 2.2, in this case additional difficulties arise and further assumptions are need.

Theorem 2.3 *Consider the numerical Hamiltonian given by (2.9) and a local operator V defined by a continuous function $F : \mathbb{R}^+ \to \mathbb{R}$ satisfying (2.10) and*

$$F'(m) \geq \underline{\delta} \min(m^{\eta_1}, m^{-\eta_2})$$

for $\underline{\delta} > 0$, $\eta_1 > 0$ and $0 < \eta_2 < 1$. As h goes to 0, the functions u^h converge to u in $W^{1,\beta}(\mathbb{T}^2)$, the functions m^h converges to m in $L^{2-\eta_2}(\mathbb{T}^2)$, and Λ^h tends to λ.

The proofs of Theorems 2.2 and 2.3 are rather technical and require several accurate estimates, hence we skip the details here. We only point out that a key ingredient in the proofs is the fundamental identity given in the next lemma (see [4]).

Lemma 2.1 *Let A, B be two grid functions, (U, M, Λ) a solution of (2.5) and $(\tilde{U}, \tilde{M}, \tilde{\Lambda})$ a solution of the perturbed system*

$$\begin{cases} -\nu(\Delta_h \tilde{U})_{i,j} + g(x_{i,j}, [D_h \tilde{U}]_{i,j}) + \tilde{\Lambda} = (V_h[\tilde{M}])_{i,j} + A_{i,j}, \\ \nu(\Delta_h \tilde{M})_{i,j} + \mathcal{T}_{i,j}(\tilde{U}, \tilde{M}) = B_{i,j}, \\ (\tilde{U}, \overline{1})_2 = 0, \ (\tilde{M}, \overline{1})_2 = 1, \ \tilde{M} \geq 0. \end{cases} \tag{2.11}$$

Then the following identity holds

$$\begin{aligned} &\mathcal{G}(M, U, \tilde{U}) + \mathcal{G}(\tilde{M}, \tilde{U}, U) + (V_h[M] - V_h[\tilde{M}], M - \tilde{M})_2 \\ =&(A, M - \tilde{M})_2 + (B, U - \tilde{U})_2, \end{aligned} \tag{2.12}$$

where $\mathcal{G}$ is the nonlinear functional acting on grid functions defined by

$$\mathcal{G}(M, U, \tilde{U}) = \sum_{i,j} M_{i,j}\Big[g(x_{i,j}, [D\tilde{U}]_{i,j}) - g(x_{i,j}, [DU]_{i,j}) \\ - g_q(x_{i,j}, [DU]_{i,j}) \cdot ([D\tilde{U}]_{i,j} - [DU]_{i,j})\Big].$$

The identity (2.12) holds for a general numerical Hamiltonian g and it employs the crucial property that the second equation in (2.5) is the adjoint of the linearized version of the first equation of the system, as already observed. Moreover, if **(G4)** and **(V2)** hold, then the first line of (2.12) is made of three nonnegative terms, hence uniqueness for (2.5) is a straightforward consequence of this identity.

While Theorems 2.2 and 2.3 rely on the existence of a classical solution to (1.1), the convergence analysis has been extended in [2], where the existence of a weak solution of the MFG system is proved via a compactness argument on solutions of the discrete problem.

3 Mean Field Games on Networks

In this section we consider stationary Mean Field Games defined on a network. We first describe a formal derivation of the MFG system in terms of Pareto equilibria for dynamic games defined on a network with a large number of (indistinguishable) players. In this way we deduce the correct transition conditions at the vertices of the network which allow to prove existence and uniqueness of the classical solution to the problem. Hence we propose a finite difference scheme for the MFG system based on the approach of Sect. 2 and a correct approximation of the transition conditions at the vertices.

3.1 Networks and Functional Spaces

We start by describing the constitutive elements of the problem and the main assumptions. A network $\Gamma = (\mathcal{V}, \mathcal{E})$ is a finite collection of points $\mathcal{V} := \{v_i\}_{i\in I}$ in $\mathbb{R}^d$ connected by continuous, non self-intersecting edges $\mathcal{E} := \{e_j\}_{j\in J}$, respectively indexed by two finite sets I and J. Each edge $e_j \in \mathcal{E}$ is parametrized by a smooth function $\pi_j : [0, l_j] \to \mathbb{R}^d$, $l_j > 0$. Given $v_i \in \mathcal{V}$, we denote by $Inc_i := \{j \in J : v_i \in e_j\}$ the set of edges branching out from v_i and by $d_{v_i} := |Inc_i|$ the degree of v_i. A vertex v_i is said a boundary vertex if $d_{v_i} = 1$, otherwise it is said a transition vertex. For simplicity, we assume that the set of boundary vertices is empty. For a function $u : \Gamma \to \mathbb{R}$ we denote by $u_j : [0, l_j] \to \mathbb{R}$ the restriction of u to e_j, i.e. $u(x) = u_j(y)$ for $x \in e_j$, $y = \pi_j^{-1}(x)$, and by $\partial_j u(v_i)$ the oriented derivative of u

at v_i along the arc e_j defined by

$$\partial_j u(v_i) = \begin{cases} \lim_{h\to 0^+} (u_j(h) - u_j(0))/h, & \text{if } v_i = \pi_j(0); \\ \lim_{h\to 0^+} (u_j(l_j - h) - u_j(l_j))/h, & \text{if } v_i = \pi_j(l_j). \end{cases}$$

The integral of a function u on Γ is defined by

$$\int_\Gamma u(x)dx := \sum_{j\in J} \int_0^{l_j} u_j(r)dr.$$

The space $C^k(\Gamma)$, $k \in \mathbb{N}$, consists of all the continuous functions $u : \Gamma \to \mathbb{R}$ such that $u_j \in C^k([0, l_j])$ for $j \in J$ and $\|u\|_{C^k} = \max_{\beta\le k} \|\partial^\beta u\|_{L^\infty} < \infty$. Observe that no continuity condition at the vertices is prescribed for the derivatives of a function $u \in C^k(\Gamma)$.

3.2 A Formal Derivation of the MFG System on a Network

We first show that the transition conditions at the vertices can be deduced in a natural way by the formulation of the differential game associated to the MFG system on the network. Consider a population of agents, distributed at time $t = 0$ according to a probability measure m_0 on Γ; each agent moves on the network Γ and its dynamics inside the edge e_j is governed by the stochastic differential equation

$$dX_s = -\gamma_s\, ds + \sqrt{2\nu_j}\, dW_s,$$

where X_s is the state variable, γ is the control, $\nu_j > 0$ is a diffusion coefficient and W_t is a 1-dimensional Brownian motion. When the agent reaches a vertex $v_i \in \mathcal{V}$, it almost surely spends zero time at v_i and enters in one of the incident edges, say e_j with $j \in Inc_i$, with probability β_{ij} where

$$\beta_{ij} > 0, \quad \sum_{j\in Inc_i} \beta_{ij} = 1.$$

(see [11] for a rigorous definition of stochastic processes on networks). The cost criterion is given by

$$\liminf_{T\to\infty} \mathbb{E}_x \left[\frac{1}{T} \int_0^T \{L(X_t, \gamma_t) + V[m(X_t)]\} dt \right]$$

where m represents the distribution of the overall population of players, L is the Lagrangian and V is an the potential. A formal application of the dynamic

programming principle implies that the value function u of the previous control problem satisfies

$$\begin{cases} -\nu_j \partial^2 u_j + H_j(x, \partial u_j) + \lambda = V_j[m], \ x \in e_j, \ j \in J \\ \sum_{j \in Inc_i} \alpha_{ij} \nu_j \partial u_j(v_i) = 0 \qquad\qquad v_i \in \mathcal{V}, \\ u_j(v_i) = u_k(v_i), \qquad\qquad j, k \in Inc_i, \ v_i \in \mathcal{V}, \end{cases} \tag{3.1}$$

where $\alpha_{ij} := \beta_{ij} \nu_j^{-1}$, λ is the ergodic cost and the Hamiltonian is given on the edge e_j by the Fenchel transformation

$$H_j(x, p) = \sup_{\gamma} \big[-\gamma \cdot p - L_j(x, \gamma) \big].$$

Note that the differential equation inside e_j is defined in terms of the coordinate parametrizing the edge. The second equation in (3.1) is known as the Kirchhoff transition condition and it is consequence of the assumption on the behavior of X_t at the vertices (see [11]). Finally, the third line equation in (3.1) is a constraint prescribing the continuity of u at transition vertices.

In order to derive the equation satisfied by the distribution m of the agents, we follow a standard duality argument. Consider the linearization of Hamilton-Jacobi-Bellman equation at u in the direction w

$$\begin{cases} -\nu_j \partial^2 w_j + \partial_p H_j(x, \partial u_j) \partial w_j = 0, \ x \in e_j, \ j \in J \\ \sum_{j \in Inc_i} \nu_j \alpha_{ij} \partial w_j(v_i) = 0 \qquad\qquad v_i \in \mathcal{V} \\ w_j(v_i) = w_k(v_i), \qquad\qquad j, k \in Inc_i, \ v_i \in \mathcal{V}. \end{cases} \tag{3.2}$$

Writing the weak formulation of (3.2) for a test function m, integrating by parts along each edge and regrouping the boundary terms corresponding to the same vertex v_i, we get

$$\begin{aligned} 0 &= \sum_{j \in J} \int_{e_j} \big(-\nu_j \partial^2 w_j + \partial_p H_j(x, \partial u_j) \partial w_j \big) m_j \, dx \\ &= \sum_{j \in J} \int_{e_j} \big[-\nu_j \partial^2 m_j - \partial(m_j \partial_p H_j(x, \partial u_j)) \big] w_j \, dx \\ &\quad - \sum_{v_i \in \mathcal{V}} \left[\sum_{j \in Inc_i} \nu_j m_j(v_i) \partial w_j(v_i) - \big(\nu_j \partial m_j(v_i) + \partial_p H(v_i, \partial u_j) m_j(v_i) \big) w_j(v_i) \right]. \end{aligned}$$

By the previous identity we obtain that m satisfies inside each edge e_j the adjoint equation

$$\nu_j \partial^2 m_j + \partial(m_j \, \partial_p H_j(x, \partial u_j)) = 0.$$

Moreover, recalling the Kirchhoff transition condition for w, the first one of the terms computed at the transition vertices vanishes if

$$\frac{m_j(v_i)}{\alpha_{ij}} = \frac{m_k(v_i)}{\alpha_{ik}}, \qquad j, k \in Inc_i, \; v_i \in \mathcal{V}. \tag{3.3}$$

The vanishing of the other term for each $v_i \in \mathcal{V}$, namely

$$\sum_{j \in Inc_i} \nu_j \partial m_j(v_i) + \partial_p H_j(v_i, \partial u_j) m_j(v_i) = 0, \tag{3.4}$$

gives the transition condition for m at the vertices $v_i \in \mathcal{V}$. Note that (3.4) corresponds to the conservation of the total flux of the density m at v_i.

We restrict for simplicity to the case in which all the coefficients in the transition condition for u are equal, i.e. $\alpha_{ij} = \alpha_{ik} \; \forall i \in I, \; j, k \in Inc_i$ and therefore (3.3) reduces to the continuity of m at the vertices. Summarizing we get the following MFG system

$$\begin{cases} -\nu_j \partial^2 u_j + H_j(x, \partial u_j) + \lambda = V_j[m], & x \in e_j, \; j \in J \\ \nu_j \partial^2 m_j + \partial(m_j \, \partial_p H_j(x, \partial u_j)) = 0 & x \in e_j, \; j \in J \\ \displaystyle\sum_{j \in Inc_i} \nu_j \partial u_j(v_i) = 0 & v_i \in \mathcal{V} \\ \displaystyle\sum_{j \in Inc_i} [\nu_j \partial m_j(v_i) + \partial_p H_j(v_i, \partial u_j) m_j(v_i)] = 0 & v_i \in \mathcal{V} \\ u_j(v_i) = u_k(v_i), \; m_j(v_i) = m_k(v_i) & j, k \in Inc_i, \; v_i \in \mathcal{V} \\ \displaystyle\int_\Gamma u(x)dx = 0, \; \int_\Gamma m(x)dx = 1, \; m \geq 0 \end{cases} \tag{3.5}$$

where the ergodic constant $\lambda \in \mathbb{R}$ is also an unknown of the problem. The transition conditions for u and m (continuity and either Kirchhoff condition or conservation of total flux, respectively) give d_{v_i} linear conditions for each function at a vertex $v_i \in \mathcal{V}$, hence they uniquely determine the values $u_j(v_i)$ and $m_j(v_i)$, $j \in Inc_i$.

For the stationary system (3.5) we have the following existence and uniqueness result [8] in the case of a local coupling $V[m](x) = V(m(x))$, $x \in \Gamma$.

Theorem 3.1 *Assume that* $H = \{H_j\}_{j\in J}$, $H_j : [0, l_j] \times \mathbb{R} \to \mathbb{R}$, $\nu = \{\nu_j\}_{j\in J}$, $\nu_j \in \mathbb{R}$, *and* $V = \{V_j\}_{j\in J}$, $V_j : \mathbb{R} \to \mathbb{R}$, *satisfy*

$H_j \in C^2([0, l_j] \times \mathbb{R})$;

$H_j(x, \cdot)$ *is convex in* p *for each* $x \in [0, l_j]$;

$\delta|p|^2 - C \le H_j(x, p) \le C|p|^2 + C$ *for* $(x, p) \in [0, l_j] \times \mathbb{R}$ *and some* $\delta, C > 0$,

$\nu_0 := \inf_{j\in J} \nu_j > 0$,

$V_j[m](x) = V_j(m(x))$ *with* $V_j \in C^1([0, +\infty))$ *and bounded.*

Then, there exists a solution $(u, m, \lambda) \in C^2(\Gamma) \times C^2(\Gamma) \times \mathbb{R}$ *to* (3.5). *Moreover if*

$$\int_\Gamma (V(m_1) - V(m_2))(m_1 - m_2)dx \le 0 \Rightarrow m_1 = m_2,$$

then the solution is unique.

3.3 A Finite Difference Scheme for Mean Field Games on Networks

The differential equations in (3.5) are defined in terms of derivatives with respect to the coordinate $y = \pi_j^{-1}(x) \in [0, l_j]$ parametrizing the arc e_j. Hence the approximation scheme for the MFG system is obtained by discretizing this local coordinate.

Given a discretization step $h = \{h_j\}_{j\in J}$, we consider an uniform partition $y_{j,k} = kh_j$, $k = 0, \ldots, N_j^h$, of the interval $[0, l_j]$ parametrizing the edge e_j (we assume that $N_j^h = l_j/h_j$ is an integer). We obtain a spatial grid on Γ by setting

$$\mathcal{G}_h = \{x_{j,k} = \pi_j(y_{j,k}),\ j \in J,\ k = 0, \ldots, N_j^h\}.$$

In the notation $x_{j,k}$, the index j refers to the arc e_j, whereas the index k refers to the grid point on e_j. We set

$$|h| = \max_{j\in J}\{h_j\}, \quad N^h = \#(I) + \sum_{j\in J}(N_j^h - 1),$$

i.e. N^h is the total number of the points of $\mathcal{G}_h$ having identified, for each $i \in I$, the $\#(\mathrm{Inc}_i)$ grid points corresponding to the same vertex v_i. We make a partition of Inc_i

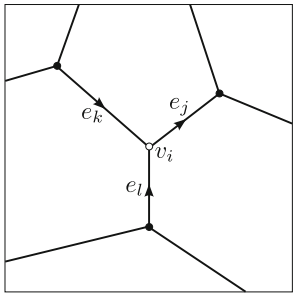

Fig. 1 Incident edges to the vertex v_i: $\mathrm{Inc}_i^+ = \{j\}$, $\mathrm{Inc}_i^- = \{k, l\}$

into the subsets

$$\mathrm{Inc}_i^+ = \{j \in \mathrm{Inc}_i : v_i = \pi_j(0)\}, \quad \mathrm{Inc}_i^- = \{j \in \mathrm{Inc}_i : v_i = \pi_j(N_j^h h_j)\},$$

as shown in Fig. 1.

For a grid function $U : \mathcal{G}_h \to \mathbb{R}$ we denote by $U_{j,k}$ its value at the grid point $x_{j,k}$. We say that a grid function $U : \mathcal{G}_h \to \mathbb{R}$ is said to be continuous at v_i if

$$U_{j,\ell} = U_{k,m} := U_i \quad \text{if } v_i = \pi_j(\ell h_j) = \pi_k(m h_k),\ j, k \in \mathrm{Inc}_i,\ \ell \in \{0, N_j^h\},\ \text{and } m \in \{0, N_k^h\},$$

i.e., the value of U at the vertex v_i is independent of the incident edge e_j, $j \in \mathrm{Inc}_i$. A grid function is continuous if it is continuous at v_i, for each $i \in I$.

Given a generic pair of grid functions $U, W : \mathcal{G}_h \to \mathbb{R}$, we define the scalar product

$$(U, W)_2 = \sum_{j\in J} \sum_{k=1}^{N_j^h - 1} h_j U_{j,k} W_{j,k} + \sum_{i\in I} \left(\sum_{j\in \mathrm{Inc}_i^+} \frac{h_j}{2} U_{j,0} W_{j,0} + \sum_{j\in \mathrm{Inc}_i^-} \frac{h_j}{2} U_{j,N_j^h} W_{j,N_j^h} \right).$$

and we introduce the compact and convex set

$$\mathcal{K}_h = \{(M_{j,k})_{j\in J,\, 0\le k\le N_j^h} : M \text{ is continuous}, M_{j,k} \ge 0,\ (M, 1)_2 = 1\}.$$

We finally define the following finite difference operators

$$(D^+U)_{j,k} = \frac{U_{j,k+1} - U_{j,k}}{h_j},$$

$$[D_h U]_{j,k} = \left((D^+U)_{j,k}, (D^+U)_{j,k-1}\right)^T,$$

$$(D_h^2 U)_{j,k} = \frac{U_{j,k-1} - 2U_{j,k} + U_{j,k+1}}{h_j^2}.$$

In order to approximate the Hamiltonian, $H = \{H_j\}_{j\in J}$, $H_j : [0, l_j] \times \mathbb{R} \to \mathbb{R}$, $j \in J$, we consider a numerical Hamiltonian $g = \{g_j\}_{j\in J}$, $g_j : [0, l_j] \times \mathbb{R}^2 \to \mathbb{R}$, $(x, q_1, q_2) \to g_j(x, q_1, q_2)$ satisfying

(**G1**) *monotonicity: g_j is non increasing w.r.t. q_1 and nondecreasing w.r.t. q_2;*
(**G2**) *consistency: $g_j(x, q, q) = H_j(x, q)\ \forall x \in [0, l_j], \forall q \in \mathbb{R}$;*
(**G3**) *differentiability: g_j is of class C^1;*
(**G4**) *convexity: for all $x \in e_j$, $(q_1, q_2) \mapsto g_j(x, q_1, q_2)$ is convex.*

The operator $V[m](x_{j,k})$ is approximated by $V_h[M]_{j,k} = V[\mathcal{I}_h m](x_{j,k})$ where $\mathcal{I}_h m$ is the piecewise constant function taking the value $M_{j,k}$ in the interval $\{|y - y_{j,k}| \le h_j/2\}$, $k = 1, \dots, N_j^h - 1$, $j \in J$ (at the vertices only the half interval contained in $[0, l_j]$ is considered). In particular, if V is a local operator, i.e. $V[m](x) = F(m(x))$, then we set $V_h[M]_{j,k} = F(M_{j,k})$. We assume that

(**V1**) *V_h is continuous and maps $\mathcal{K}_h$ on a bounded set of grid functions.*
(**V2**) *V_h is monotone, i.e. $(V_h[M] - V_h[\bar{M}], M - \bar{M})_2 \le 0 \Rightarrow M = \bar{M}$.*

For the discretization of the differential equations in (3.5) inside the edge, we follow the same approach in [1] and we refer to Sect. 2 for motivations and explanations. We just recall that the approximation of the transport operator in the Fokker-Planck equation comes from the discretization of the quantity

$$\int_{e_j} m \frac{\partial H_j}{\partial p}(x, \partial u)\partial w \, dx$$

for a test function w, which is related to the weak formulation of the equation on the network. At the internal grid points we consider the finite difference system

$$\begin{cases} -\nu_j (D_h^2 U)_{j,k} + g_j(x_{j,k}, [D_h U]_{j,k}) + \Lambda = V_h[M]_{j,k}, & k = 1, \dots, N_j^h - 1,\ j \in J \\ \nu_j (D_h^2 M)_{j,k} + \mathcal{B}^h(U, M)_{j,k} = 0, & k = 1, \dots, N_j^h - 1,\ j \in J \\ M \in \mathcal{K}_h, \qquad (U, 1)_2 = 0, \end{cases} \tag{3.6}$$

where U, M are grid functions and $\Lambda \in \mathbb{R}$. The transport operator $\mathcal{B}^h$ is defined for $j \in J$ and $k = 1$ by

$$\mathcal{B}^h(U, M)_{j,k} = \begin{array}{l} \frac{1}{h_j}\Big[M_{j,k} \frac{\partial g_j}{\partial q_1}(x_{j,k}, [D_h U]_{j,k}) + \\ M_{j,k+1} \frac{\partial g_j}{\partial q_2}(x_{j,k+1}, [D_h U]_{j,k+1}) - M_{j,k} \frac{\partial g_j}{\partial q_2}(x_{j,k}, [D_h U]_{j,k}) \Big]; \end{array}$$

for $k = 2, \dots, N_j^h - 2$ by

$$\mathcal{B}^h(U, M)_{j,k} = \begin{array}{l} \frac{1}{h_j}\Big[M_{j,k} \frac{\partial g_j}{\partial q_1}(x_{j,k}, [D_h U]_{j,k}) - M_{j,k-1} \frac{\partial g_j}{\partial q_1}(x_{j,k-1}, [D_h U]_{j,k-1}) \\ + M_{j,k+1} \frac{\partial g_j}{\partial q_2}(x_{j,k+1}, [D_h U]_{j,k+1}) - M_{j,k} \frac{\partial g_j}{\partial q_2}(x_{j,k}, [D_h U]_{j,k}) \Big]; \end{array}$$

for $k = N_j^h - 1$ by

$$\mathcal{B}^h(U,M)_{j,k} = \begin{array}{l} \frac{1}{h_j}\Big[M_{j,k}\frac{\partial g_j}{\partial q_1}(x_{j,k},[D_hU]_{j,k}) - M_{j,k-1}\frac{\partial g_j}{\partial q_1}(x_{j,k-1},[D_hU]_{j,k-1}) - \\ M_{j,k}\frac{\partial g_j}{\partial q_2}(x_{j,k},[D_hU]_{j,k})\Big]. \end{array}$$

For the approximation of the transition conditions in (3.5), we use a standard first order discretization of the normal derivative of u and m. In particular, we employ forward or backward finite differences depending on whether the vertex is, respectively, the initial or terminal point in the parametrization of the edge. The flux term in the Kirchhoff condition for m is approximated in a upwind fashion taking always into account the orientation of the edge. Moreover we impose the continuity at the vertices of U and M at the vertices so that the full set of discrete transition conditions is given by

$$\begin{cases} \mathcal{S}^h(U, V_h[M] - \Lambda)_i = 0, & i \in I, \\ \mathcal{T}^h(M,U)_i = 0 & i \in I, \\ U, M \text{ continuous at } v_i, & i \in I, \end{cases} \tag{3.7}$$

where,for every triple of grid functions U,V, M, the operators $\mathcal{S}^h : \mathcal{V} \to \mathbb{R}$ and $\mathcal{T}^h : \mathcal{V} \to \mathbb{R}$ are defined by

$$\mathcal{S}^h(U,V)_i = \sum_{j\in \mathrm{Inc}_i^+}\Big[v_j(D^+U)_{j,0} + \frac{h_j}{2}V_{j,0}\Big] - \sum_{j\in \mathrm{Inc}_i^-}\Big[v_j(D^+U)_{j,N_j^h-1} - \frac{h_j}{2}V_{j,N_j^h}\Big],$$

$$\begin{aligned} \mathcal{T}^h(M,U)_i = &\sum_{j\in \mathrm{Inc}_i^+}\Big[v_j(D^+M)_{j,0} + M_{j,1}\frac{\partial g}{\partial q_2}(x_{j,1},[D_hU]_{j,1})\Big] \\ &- \sum_{j\in \mathrm{Inc}_i^-}\Big[v_j(D^+M)_{j,N_j^h-1} + M_{j,N_j^h-1}\frac{\partial g}{\partial q_1}(x_{j,N_j^h-1},[D_hU]_{j,N_j^h-1})\Big] = 0. \end{aligned}$$

We observe that in the approximation of the Kirchhoff condition appears an additional term $\frac{h_j}{2}((V_h[M]) - \Lambda)$, vanishing for $h \to 0$. This term is necessary to obtain a fundamental identity analogous to (2.12), suitable for the network problem. Summarizing, the approximation scheme for the stationary problem (3.5) is given by (3.6)–(3.7).

3.4 Existence, Uniqueness and Convergence of the Numerical Scheme

Concerning existence and uniqueness of a solution to (3.6)–(3.7) we have the following result.

Theorem 3.2 *Assume that g satisfies* ***(G1)–(G3)****, V satisfies* ***(V1)****. Then the problem* (3.6)–(3.7) *has at least a solution* (U, M, Λ)*. If moreover g satisfies* ***(G4)*** *and V satisfies* ***(V2)****, then the solution is unique.*

Existence is proved as in the continuous case by a fixed point argument. For the uniqueness we rely on a fundamental identity similar to (2.12) which is also a crucial tool to prove the convergence of the scheme.

We describe a convergence result for the scheme (3.6)–(3.7) in the reference case

$$H(x, p) = |p|^{\beta} + f(x), \tag{3.8}$$

where $\beta \geq 2$ and $f : \Gamma \to \mathbb{R}$ is a continuous function. We consider a numerical Hamiltonian of the form

$$g(x, p) = G(p_1^-, p_2^+) + f(x) \tag{3.9}$$

where $G(p_1, p_2) = (p_1^2 + p_2^2)^{\beta/2}$ and $p^{\pm}$ denote the positive and negative part of $p \in \mathbb{R}$. We observe that g satisfies assumptions **(G1)–(G4)**.

Theorem 3.3 *Assume* (3.8)*, V is a local* C^1 *potential and g of the form* (3.9)*. Let* (u, m, λ) *be the unique solution of* (3.5) *and let* u^h *(resp.* m^h*) be the piecewise linear function on* Γ *obtained by interpolating the values* $U^h_{j,k}$ *(resp* $M^h_{j,k}$*) of the solution* (U^h, M^h, Λ^h) *to* (3.6) *and* (3.7) *at the nodes of the network grid. Then*

$$\lim_{|h|\to 0} \|u^h - u\|_{\infty} + \|m^h - m\|_{\infty} + |\Lambda^h - \lambda| = 0.$$

4 A Quasi-Newton Method for Stationary Mean Field Games

This section is devoted to the actual implementation and test of a numerical solver for stationary MFG systems, both in the Euclidean and Network cases introduced in (1.1) and (3.5) respectively. The main issue from an implementation point of view is that these systems are strongly coupled and, more important, they involve the ergodic constant λ as an additional unknown. A standard way in the literature to overcome this issue is a regularization technique, which is an effective tool both for theoretical and numerical results. The ergodic constant λ is replaced by the zero order term δu_δ (where $\delta > 0$ is a small parameter) or by the time derivative $\frac{\partial u}{\partial t}$ (associated to an initial datum), yielding to well posed problems. Indeed, it can

be proved that both $-\delta u_\delta$ and $-\frac{u(\cdot,t)}{t}$ converge uniformly to λ as $\delta \to 0$ and $t \to \infty$ respectively, whereas u_δ and $u(\cdot,t) - \lambda t$ converge to a solution of the original stationary equation. Unfortunately, this procedure introduces an additional approximation in the computation, affecting the accuracy of the corresponding numerical solutions and the computational time to reach convergence.

Here we review a new method that we introduced in [7] for stationary MFG systems on networks and extended in [6] to very general homogenization problems for Hamilton-Jacobi equations. The main novelty is that the discrete stationary MFG system is solved *directly* without any further (small δ or long time t) approximation, treating the ergodic constant λ as it is, an additional unknown.

To avoid cumbersome notations and focus only on the main idea, we keep the discussion at an abstract level. In particular, we no further distinguish between the Euclidean case (1.1) and the Network case (3.5), we only assume that, after the discretization, we end up with a generic lattice of N nodes. We collect all the unknowns in a single vector $X = (U, M, \Lambda)$, whose length turns out to be $2N + 1$. On the other hand, we recast the $2N$ equations in the system (including the transition conditions in the network case) plus the 2 normalization conditions as functions of X with zero right hand sides, obtaining a nonlinear map $F : \mathbb{R}^{2N+1} \to \mathbb{R}^{2N+2}$. The problem is then reduced to

$$\text{Find } X \in \mathbb{R}^{2N+1} \text{ such that } F(X) = 0 \in \mathbb{R}^{2N+2}\,. \tag{4.1}$$

Note that a zero of F exists and it is unique under the assumptions discussed in the previous sections, but the problem (4.1) is formally *overdetermined*, adopting, with a slight abuse, a terminology usually devoted to linear systems. Hence, the solution to (4.1) should be meant in a *least-squares* sense, namely as a solution of the following optimization problem (where we denote by $\|\cdot\|_2$ the Euclidean norm in $\mathbb{R}^{2N+2}$):

$$\min_{X\in\mathbb{R}^{2N+1}} \frac{1}{2}\|F(X)\|_2^2\,.$$

Assuming smoothness of F and defining $\mathcal{F}(X) := \frac{1}{2}\|F(X)\|_2^2$, the classical Newton method for finding critical points of $\mathcal{F}$ is given by

$$\mathcal{H}_{\mathcal{F}}(X^k)(X^{k+1} - X^k) = -\nabla\mathcal{F}(X^k) \qquad k \geq 0\,.$$

Computing the gradient $\nabla\mathcal{F}$ and the Hessian $\mathcal{H}_{\mathcal{F}}$ of $\mathcal{F}$ we have

$$\nabla\mathcal{F}(X) = J_F(X)^T F(X)\,,$$

$$\mathcal{H}_{\mathcal{F}}(X) = J_F(X)^T J_F(X) + \sum_{i=1}^{2N+2} \frac{\partial^2 F_i}{\partial X^2}(X) F_i(X)\,,$$

where the second order term is given by

$$\left(\frac{\partial^2 F_i}{\partial^2 X}(X)\right)_{k,\ell} = \frac{\partial^2 F_i}{\partial X_k \partial X_\ell}(X)\,.$$

Since a solution to the MFG system corresponds to a zero minimum of $\mathcal{F}(X)$, we expect $F(X^k)$ to be small for X^k close enough to a solution. Hence we approximate $\mathcal{H}_{\mathcal{F}}(X^k) \simeq J_F(X^k)^T J_F(X^k)$ and obtain the so called Gauss-Newton method, which requires only Fréchet differentiability of F:

$$J_F(X^k)^T J_F(X^k)\left(X^{k+1} - X^k\right) = -J_F(X^k)^T F(X^k) \qquad k \geq 0\,.$$

From a computational point of view, the presence of the transposed Jacobian restores the square size of the system, but it also squares its condition number, a crucial point for the approximation as N increases.

We proceed in an alternative way, by simply observing that, for $\delta := X^{k+1} - X^k$, the Gauss-Newton step above is just the optimality condition for the following linear least-squares problem:

$$\min_{\delta \in \mathbb{R}^{2N+1}} \frac{1}{2}\|J_F(X^k)\delta + F(X^k)\|_2^2\,, \tag{4.2}$$

which is in turn easily and efficiently solved by means of the QR *factorization* of J_F. Indeed, let $n = 2N + 1$ and suppose that $J_F(X^k) = QR$, where Q is a $(n+1) \times (n+1)$ orthogonal matrix (i.e. $Q^{-1} = Q^T$) and R is a $(n+1) \times n$ matrix of the form $R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$, with R_1 of size $n \times n$ and upper triangular. Writing $Q = (Q_1 \quad Q_2)$ with Q_1 of size $(n+1) \times n$ and Q_2 of size $(n+1) \times 1$, we get

$$\|J_F(X^k)\delta + F(X^k)\|_2^2 = \|Q^T\left(J_F(X^k)\delta + F(X^k)\right)\|_2^2 = \|Q^T QR\delta + Q^T F(X^k)\|_2^2 =$$

$$= \left\| \begin{pmatrix} R_1\delta \\ 0 \end{pmatrix} + \begin{pmatrix} Q_1^T F(X^k) \\ Q_2^T F(X^k) \end{pmatrix} \right\|_2^2 = \|R_1\delta + Q_1^T F(X^k)\|_2^2 + \|Q_2^T F(X^k)\|_2^2$$

which is finally minimized by getting rid of the first of the two latter terms, i.e. solving the square triangular $n \times n$ linear system $R_1\delta = -Q_1^T F(X^k)$ via back substitution.

Summarizing, we propose the following algorithm:

GIVEN AN INITIAL GUESS X AND A TOLERANCE $\varepsilon > 0$,
REPEAT

1. ASSEMBLE $F(X)$ AND $J_F(X)$
2. SOLVE THE LINEAR SYSTEM $J_F(X)\delta = -F(X)$ IN THE LEAST-SQUARES SENSE, USING THE QR FACTORIZATION OF $J_F(X)$

3. UPDATE $X \leftarrow X + \delta$

UNTIL $\|\delta\|_2^2 < \varepsilon$ AND/OR $\|F(X)\|_2^2 < \varepsilon$

We refer the interested reader to [6] and [7] for implementation details, performance tests of the proposed algorithm and a comparison with existing methods. In the remaining sections we present some simulations in different settings, showing the versatility of the new method to catch interesting features of the corresponding problems.

4.1 MFG in Euclidean Spaces

We consider a MFG system in dimension two, with an eikonal Hamiltonian, a cost function f and a local potential V, namely

$$\begin{cases} -\nu\Delta u + |Du|^2 + f(x) + \lambda = V(m) & x \in \mathbb{T}^2 \\ \nu\Delta m + 2\,\mathrm{div}(m\,Du) = 0 & x \in \mathbb{T}^2 \\ \int_{\mathbb{T}^2} u(x)dx = 0, \int_{\mathbb{T}^2} m(x)dx = 1\,, \end{cases}$$

with $f(x) = \sin(2\pi x_1) + \cos(4\pi x_1) + \sin(2\pi x_2)$ and $V(m(x)) = m^2(x)$. We discretize the torus $\mathbb{T}^2$ with $N = 2500$ uniformly distributed nodes, so that the size of the system is 5002×5001, corresponding to 2500 degrees of freedom for U, 2500 for M and 1 for Λ. We choose $U \equiv 0$, $M \equiv 1$ and $\Lambda = 0$ as initial guess for the Newton's method and we set to $\varepsilon = 10^{-6}$ the tolerance for the stopping criterion of the algorithm.

In the first test we set the diffusion coefficient $\nu = 1$. In Fig. 2, we show the surfaces and the level sets of the computed pair of solutions (U, M).

The convergence is fast, just five iterations in 8.06 s and we get $\Lambda = 0.9784$. Moreover, we observe the typical "dual" behavior of the solutions, namely the fact that the local maxima of the mass distribution M correspond to the local minima of the value function U and vice versa. Note that, due to the high diffusion, the mass density is well distributed on the whole domain.

On the contrary, we can push the diffusion close to the deterministic limit, repeating the test with $\nu = 0.01$ to enhance concentration. We reach convergence in 10.72 s with 21 iterations and we get $\Lambda = 1.1878$. In Fig. 3, we show the surfaces and the level sets of the computed pair of solutions (U, M).

We clearly see how the supports of U and M are almost disjoint and that the mass distribution tries to occupy all the region corresponding to the minimum of the value function.

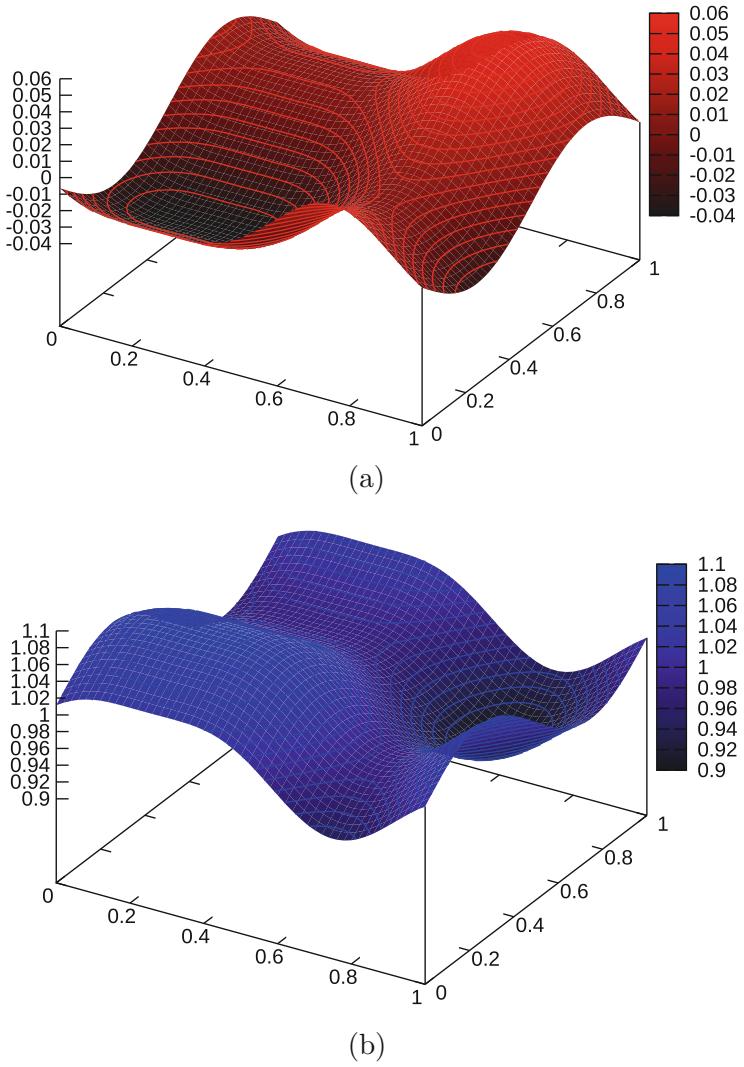

Fig. 2 Surfaces and level sets of the solutions U (**a**) and M (**b**)

4.2 Multi-Population MFG in Euclidean Spaces

This is a generalization of (1.1) to the case of P competing populations, each one described by a MFG-system, coupled via a potential term (see [15]). Here we consider the setting recently studied in [10] for problems with Neumann boundary conditions, and we present the case in dimension one and two of an eikonal

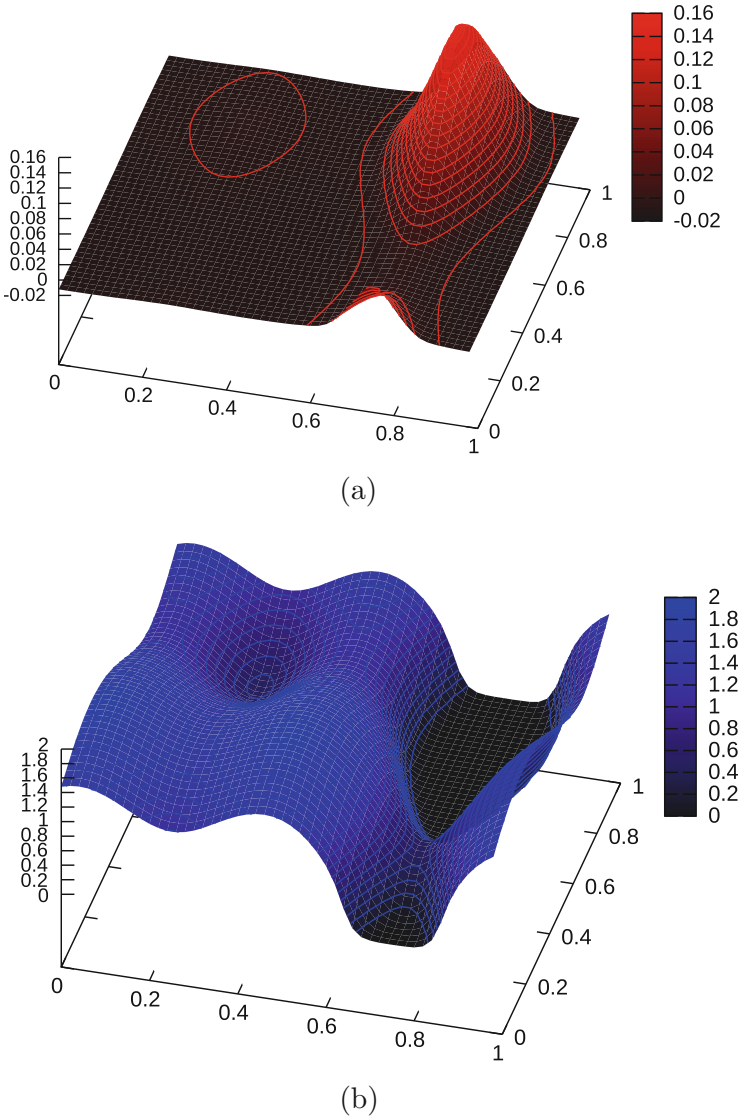

Fig. 3 Surfaces and level sets of the solutions U (**a**) and M (**b**)

Hamiltonian with a linear local potential, namely the problem

$$\begin{cases} -\nu\Delta u_i + |Du_i|^2 + \lambda_i = V_i(m) & \text{in } \Omega\,, \quad i = 1,\dots,P \\ \nu\Delta m_i + 2\mathrm{div}(m_i\, Du_i) = 0 & \text{in } \Omega\,, \quad i = 1,\dots,P \\ \partial_n u_i = 0\,, \quad \partial_n m_i = 0 & \text{on } \partial\Omega\,, \quad i = 1,\dots,P \\ \int_\Omega u_i(x)dx = 0\,, \quad \int_\Omega m_i(x)dx = 1\,, & i = 1,\dots,P\,, \end{cases}$$

where $\Omega = [0, 1]$ or $\Omega = [0, 1]^2$, the value function $u(x) = (u_1(x), \dots, u_P(x))$ and the mass distribution $m(x) = (m_1(x), \dots, m_P(x))$ are vector functions and $\lambda = (\lambda_1, \dots, \lambda_P) \in \mathbb{R}^P$ is a P-tuple of ergodic constants. Moreover, for $i = 1, \dots, P$, the linear local potential V_i takes the form

$$V_i(m(x)) = \sum_{j=1}^{P} \theta_{ij} m_j(x) \,,$$

for some given weights $\theta_{ij} \in \mathbb{R}$, or in matrix notation

$$V = (V_1, \dots, V_P) \,, \qquad \Theta = (\theta_{ij})_{i,j=1,\dots,P} \,, \qquad V(m) = \Theta m \,. \tag{4.3}$$

Existence and uniqueness of a solution (u, m, λ) can be proved under suitable monotonicity assumptions on V (see [10] for details).

Note that this problem is even more *overdetermined* than the previous one. Indeed, discretizing Ω with a uniform grid of N nodes, we end up with $P(2N + 2)$ equations in the $P(2N + 1)$ unknowns (U, M, Λ).

In the special case (4.3) uniqueness is guaranteed assuming that Θ is positive semi-definite and the solution is explicitly given, for $i = 1, \dots, P$, by $u_i \equiv 0$, $m_i \equiv 1$ and $\lambda_i = \sum_{j=1}^{P} \theta_{ij}$. By dropping this condition, the trivial solution is still found, but we expect to observe other more interesting solutions.

We start with some experiments in dimension one, in the case of $P = 2$ populations. We choose the coupling matrix (not positive semi-definite)

$$\Theta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

so that the potential for each population only depends on the other population. Moreover, we discretize the interval $\Omega = [0, 1]$ with $N = 100$ uniformly distributed nodes, we set to $\nu = 0.05$ the diffusion coefficient and to $\varepsilon = 10^{-6}$ the tolerance for the stopping criterion of the algorithm. To avoid the trivial solution, we choose non constant initial guesses, such as piecewise constant pairs with zero mean for U and piecewise constant pairs with mass one for M. Figure 4 shows four computed solutions. In the top panels we show the mass distribution $M = (M_1, M_2)$, while in the bottom panels the corresponding value function $U = (U_1, U_2)$. Segregation of the two populations is expected (see [10]) and clearly visible. This phenomenon can be enhanced by reducing the diffusion coefficient, as shown in Fig. 5, where $\nu = 10^{-4}$, close to the deterministic limit.

We finally consider the more complex and suggestive two dimensional case. We discretize the square $\Omega = [0, 1]^2$ with 25×25 uniformly distributed nodes and we push the diffusion ν up to 10^{-6}, in order to observe segregation among the populations. Moreover, we choose the interaction matrix as before, with all the entries equal to 1 except for the diagonal, which is set to 0.

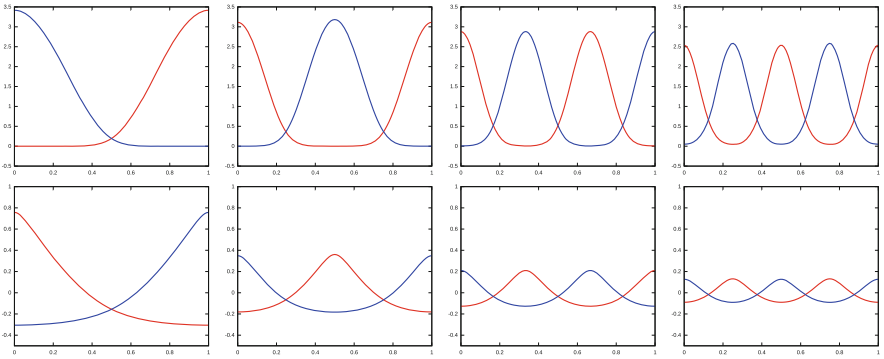

Fig. 4 Two-population MFG solutions ($\nu = 0.05$): mass distribution $M = (M_1, M_2)$ (top panels) and corresponding value function $U = (U_1, U_2)$ (bottom panels)

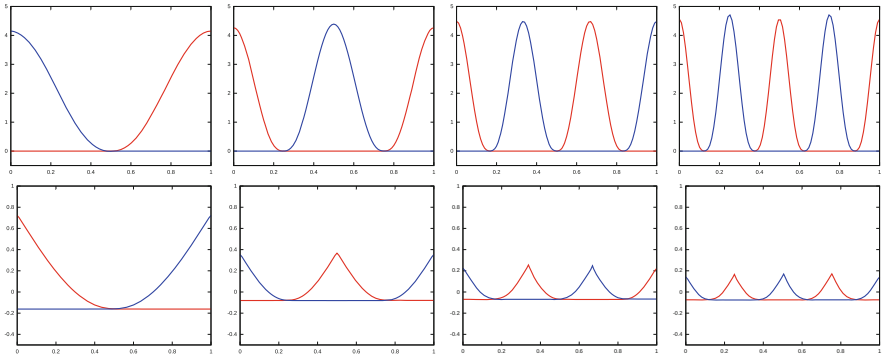

Fig. 5 Two-population MFG solutions ($\nu = 10^{-4}$): mass distribution $M = (M_1, M_2)$ (top panels) and corresponding value function $U = (U_1, U_2)$ (bottom panels)

Figure 6 shows a rich collection of solutions, corresponding to $P = 2$ (top panels), $P = 3$ (middle panels) and $P = 4$ (bottom panels) populations for different initial guesses of the Newton's method. We clearly see how the populations compete to share out all the domain.

4.3 MFG on Networks

Here we show the ability of the proposed method to handle problems on quite complex structures. To this end, we consider a MFG system on a network without boundary, in the special case of an eikonal Hamiltonian, a cost function f and a local potential V, i.e.

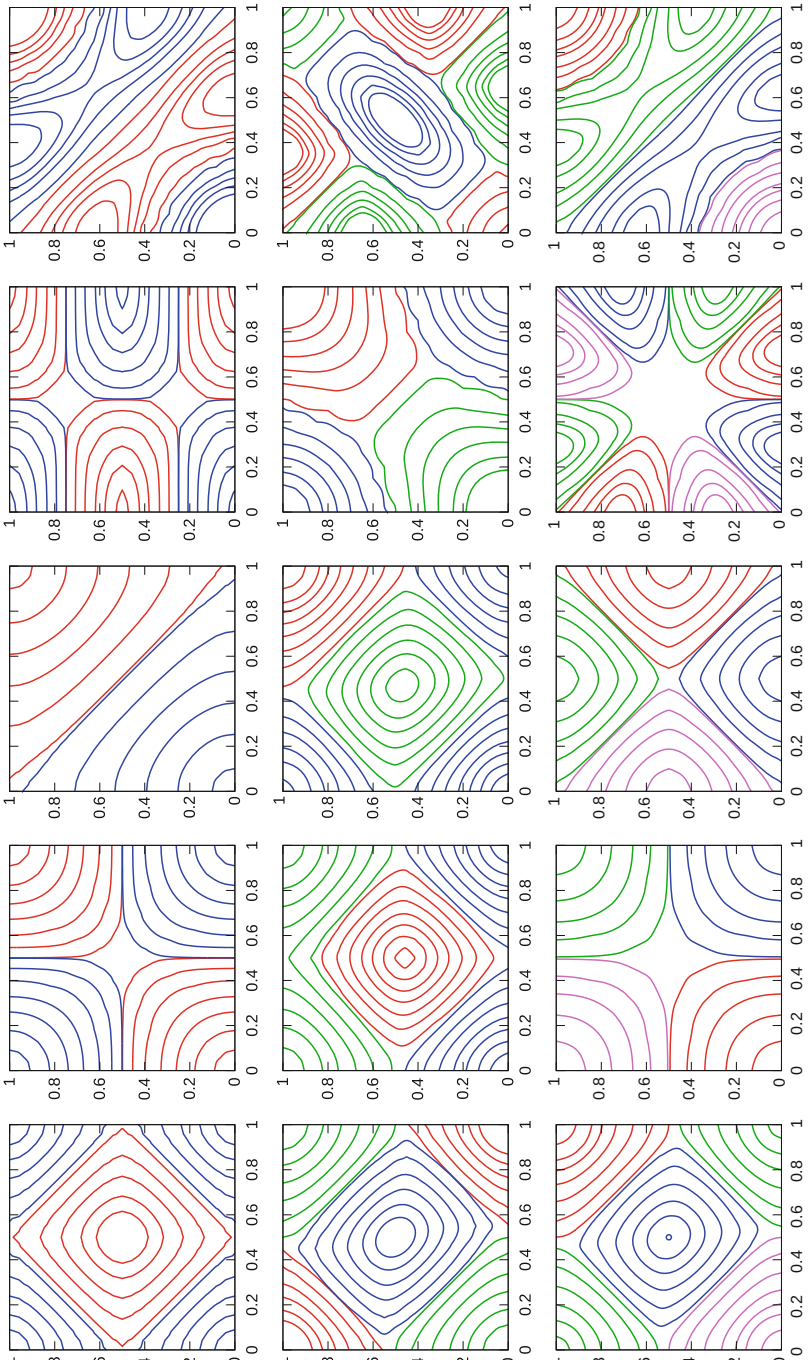

Fig. 6 Multi-population MFG mass distributions ($P = 2, 3, 4$)

$$
\begin{cases}
-\nu\partial^2 u + |\partial u|^2 + f + \lambda = V[m] & \text{in } \Gamma \\
\nu\partial^2 m + 2\partial(m\,\partial u) = 0 & \text{in } \Gamma \\
\displaystyle\sum_{j\in Inc_i} \nu_j \partial u_j(v_i) = 0 & v_i \in \mathcal{V} \\
\displaystyle\sum_{j\in Inc_i} [\nu_j \partial m_j(v_i) + 2\partial u_j(v_i) m_j(v_i)] = 0 & v_i \in \mathcal{V} \\
u_j(v_i) = u_k(v_i),\ m_j(v_i) = m_k(v_i) & j,k \in Inc_i,\ v_i \in \mathcal{V} \\
\displaystyle\int_\Gamma u(x)dx = 0, \quad \int_\Gamma m(x)dx = 1\,,
\end{cases}
$$

where $V[m] = m^2$, $\nu = 0.1$, the network Γ is shown in Fig. 7a and the cost f is the restriction to Γ of the function $\min\{|x - (3.5, 2.5)|, 1\}$ for $x \in \mathbb{R}^2$, shown in Fig. 7b.

The network consists in 26 vertices and 44 edges, each one uniformly discretized with 50 nodes, yielding a system with 4365 degrees of freedom. We set to $\varepsilon = 10^{-6}$ the tolerance for the stopping criterion of the algorithm. Note that the cost f is maximal ($\equiv$ 1) outside of the ball centered in the point (3.5, 2.5) with radius 1 where, as in the Euclidean case, we expect the value function u to attain its minimum and the mass m to be well distributed. This is what is observed in Fig. 8, showing the pair of computed solutions. The algorithm reaches convergence in 14 s after 15 iterations.

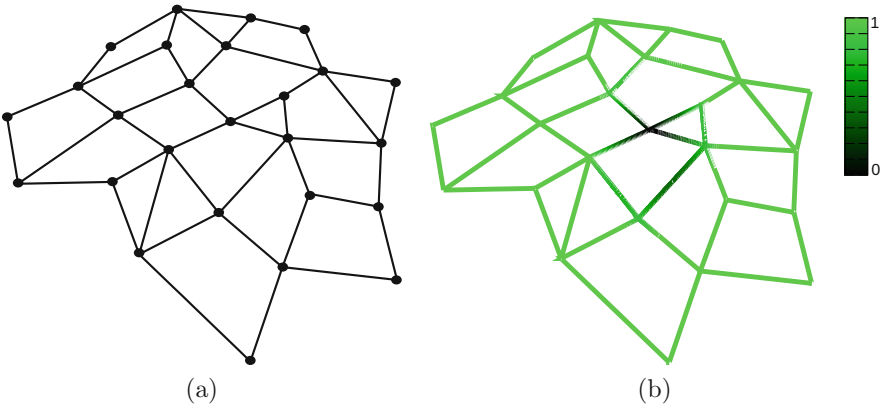

Fig. 7 The network Γ (**a**) and the cost function f (**b**)

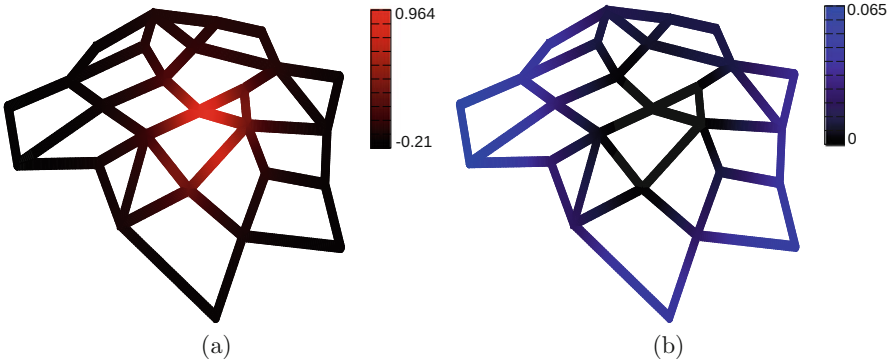

Fig. 8 The value function u (**a**) and the mass distribution m (**b**)

References

1. Achdou, Y., Capuzzo Dolcetta, I.: Mean field games: numerical methods. SIAM J. Numer. Anal. **48**(3), 1136–1162 (2010)
2. Achdou, Y., Porretta, A.: Convergence of a finite difference scheme to weak solutions of the system of partial differential equations arising in mean field games. SIAM J. Numer. Anal. **54**(1), 161–186 (2016)
3. Achdou, Y., Camilli, F., Capuzzo Dolcetta, I.: Mean field games: numerical methods for the planning problem. SIAM J. Control Optim. **50**, 77–109 (2012)
4. Achdou, Y., Camilli, F., Capuzzo Dolcetta, I.: Mean field games: convergence of a finite difference method. SIAM J. Numer. Anal. 51, 2585–2612 (2013)
5. Almulla, N., Ferreira, R., Gomes, D.: Two numerical approaches to stationary mean-field games. Dyn. Games Appl. **7**, 657–682 (2016). https://doi.org/10.1007/s13235-016-0203-5
6. Cacace, S., Camilli, F.: A generalized Newton method for homogenization of Hamilton-Jacobi equations. SIAM J. Sci. Comput. **38**, A3589–A3617 (2016)
7. Cacace, S., Camilli, F., Marchi, C.: A numerical method for mean field games on networks. ESAIM Math. Model. Numer. Anal. **51**, 63–88 (2017)
8. Camilli, F., Marchi, C.: Stationary mean field games on networks. SIAM J. Control Optim. **54**, 1085–1103 (2016)
9. Carlini, C., Silva, F.: A fully discrete semi-Lagrangian scheme for a first order mean field games problem. SIAM J. Numer. Anal. **52**, 45–67 (2014)
10. Cirant, M.: Multi-population mean field games systems with Neumann boundary conditions. J. Math. Pures Appl. **103**(5), 1294–1315 (2015)
11. Freidlin, M., Wentzell, A.: Diffusion processes on graphs and the averaging principle. Ann. Probab. **21**(4), 2215–2245 (1993)
12. Gomes, D., Saude, J.: Mean field games - a brief survey. Dyn. Games Appl. **4**(2), 110–154 (2014)
13. Guéant, O.: Mean field games equations with quadratic Hamiltonian: a specific approach. Math. Models Methods Appl. Sci. **22**(9), 37 pp., 1250022 (2012)
14. Lachapelle, A., Salomon, J., Turinici, G.: Computation of mean field equilibria in economics. Math. Models Methods Appl. Sci. **20**, 567–588 (2010)
15. Lasry, J.-M., Lions, P.-L.: Mean field games. Jpn. J. Math. **2**(1), 229–260 (2007)
16. LeVeque, R.J.: Numerical Methods for Conservation Laws. Lectures in Mathematics ETH Zürich (Birkhäuser Verlag, Basel, 1990)