Technical Notes of Mean Field Games Planning and Optimal Transport Problem

Peilin Yang

Abstract

This technique note is for learning the knowledge of Mean fields games and optimal transport problem. Mean fields games describe the asymptotic behavior of differential games in which the number of players tends to $+\infty$. Here we consider a numerical method for the optimal planning problem, i.e. the problem in which the positions of a very large number of identical rational agents, with common value function, evolve from a given initial spatial density to a desired target density at the final horizon time.

1   Introduction

Mean field type models describing the asymptotic behavior of stochastic differential game problems (Nash equilibria) as the number of players tends to $+\infty$ have recently been introduced by J-M. Lasry and P-L. Lions [13, 14, 15]. In the periodical setting, a typical such model comprises the following system of evolution partial differential equations for the unknown scalar functions $u = u(t, x)$ and $m = m(t, x)$

$$\frac{\partial u}{\partial t} - \nu \Delta u + H(x, \nabla u) = V[m], \quad \text{in } (0, T) \times \mathbb{T}^d$$
$$\frac{\partial m}{\partial t} + \nu \Delta m + \text{div}\left(m\frac{\partial H}{\partial p}(x, \nabla u)\right) = 0, \quad \text{in } (0, T) \times \mathbb{T}^d$$

with the initial and terminal conditions

$$m(0, x) = m_0(x), \quad m(T, x) = m_T(x), \quad \text{in } \mathbb{T}^d$$

given two probability densities m_0 and m_T. Since the function m represents a probability density it is natural to supplement (1)-(3) with the further conditions

$$\int_{\mathbb{T}^d} m(t, x)dx = 1, \quad m > 0$$

We denote by $\mathbb{T}^d = [0, 1]^d$ the d-dimensional unit torus, by ν a nonnegative constant and by Δ, ∇ and div, respectively, the Laplace, the gradient and the divergence operator acting on the x variable. The other operators occurring in the system are the scalar Hamiltonian $H(x, p)$, typically convex in the gradient variable, and the operator V associating a real valued function $V[m]$ on $\mathbb{T}^d$ to a probability density m. Precise assumptions on H and V as well as some examples will be specified below.

Consider the important special case when the Hamiltonian is of the form

$$H(x, \nabla u) = \sup_{\gamma}[\gamma \cdot \nabla u - L(x, \gamma)]$$

In this case, if u and m solve the system above, then Dynamic Programming arguments, see Bardi-Capuzzo Dolcetta [4], Fleming- Soner [9], show that the solution u of the forward in time Hamilton-Jacobi-Bellman equation (1) is the value function of an optimal control problem for the controlled dynamics defined on $\mathbb{T}^d$ by

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

(here s is the physical time and $t = T - s$ is the time to the horizon and (W_s) is a Brownian motion), and running cost density $L(X_s, \gamma_s) + V[m_s](X_s)$ depending on the position X_s, the control γ_s and the probability density m_s. On the other hand, (2) is a backward Fokker-Planck equation with velocity field $\frac{\partial H}{\partial p}(x, \nabla u)$ depending on the value function itself. In this context, conditions (3) represent the requirement that the positions of a very large number of identical rational agents whose dynamics is given by

$$dX_s = -\frac{\partial H}{\partial p}(X_s, \nabla u(s, X_s)) ds + \sqrt{2\nu} dW_s$$

evolve from a given spatial density m_T at $s = 0 \Leftrightarrow t = T$ to a desired target density m_0 at $s = T \Leftrightarrow t = 0$. Problem (1),(2),(3),(4) (we will occasionally refer to it as MFGP) has been introduced by P-L. Lions in his lectures at Collège de France (2009-2010), see [17], as a mean field planning problem.

Note that in MFGP, the initial time condition differs from the one studied in the seminal papers [14, 15], where the initial and final time conditions are indeed

$$u(0,x) = V_0[m|_{t=0}](x). \quad m(T,x) = m_T(x), \quad \text{in } \mathbb{T}^d$$

and $V_0[m|_{t=0}](X_T)$ is an additional terminal cost term (terminal with respect to the s variable). The mean field planning problem can be viewed as an inverse problem for the terminal cost $u|_{t=0}$ appearing in the direct mean field game system (1),(2),(5),(4) in order to drive the density of players from m_T to m_0.

Whereas existence (and uniqueness) results for (1),(2),(5),(4) are available under fairly general assumptions, see [14, 15], much less is known concerning MFGP. Indeed, as far as we know, P-L. Lions has proved existence for MFGP in mainly two cases:

1. $\nu = 0$ (deterministic case), H is a smooth and strictly convex Hamiltonian such that $\lim_{|p| \to \infty} \frac{H(x,p)}{|p|} = +\infty$, $V[m](x) = F(m(x))$ where F is a smooth and strictly increasing function, m_0 and m_T are smooth functions bounded away from 0.

2. $\nu > 0, H(p) = c|p|^2$ or $H(p)$ is close to $c|p|^2$, $V[m](x) = F(m(x))$ where F is a smooth, bounded and non-decreasing function, m_0 and m_T are smooth functions bounded away from 0 but existence is still an open question when $\nu > 0$ and the Hamiltonian is more general. P-L. Lions has also proved that if H is sublinear with respect to p and if $m_0 \neq m_T$, then there are no solutions to MFGP if T is small enough. Therefore, existence may only result from combined nonlinear effects.

We shall also consider a penalized version of system MFGP, namely

$$\begin{cases} \frac{\partial u^\epsilon}{\partial t} - \nu \Delta u^\epsilon + H(x, \nabla u^\epsilon) = V[m^\epsilon], & \text{in } (0,T) \times \mathbb{T}^d, \\ \frac{\partial m^\epsilon}{\partial t} + \nu \Delta m^\epsilon + \text{div}\left(m^\epsilon \frac{\partial H}{\partial p}(x, \nabla u^\epsilon)\right) = 0, & \text{in } (0,T) \times \mathbb{T}^d, \\ \int_{\mathbb{T}^d} m^\epsilon(t,x) dx = 1, \quad m^\epsilon > 0, \\ u^\epsilon(0,x) = \frac{1}{\epsilon}(m^\epsilon(0,x) - m_0(x)), \quad m^\epsilon(T,x) = m_T(x), & \text{in } \mathbb{T}^d \end{cases}$$

where ϵ is a small positive parameter. Problem (6) is clearly of the form (1),(2),(5),(4), which is more easily handled as we have already seen.

It is also worthwhile to observe that the planning problem described above can be seen as a generalization of the simpler system, (with in particular $V = 0, \nu = 0$),

$$\frac{\partial u}{\partial t} + \frac{1}{2}|\nabla u|^2 = 0, \quad \frac{\partial m}{\partial t} + \mathrm{div}(m\nabla u) = 0,$$
$$m(0, x) = m_0(x), \quad m(T, x) = m_T(x)$$

which was introduced by Benamou and Brenier [5], see also [18], as a fluid mechanics formulation of the Monge-Kantorovich mass transfer problem. In [5], a numerical method for the solution of $(7), (8)$ is proposed on the basis of a reformulation of the problem as the system of optimality conditions for a suitably constructed primal-dual pair of convex optimal control problems for the transport equation

$$\frac{\partial m}{\partial t} + \mathrm{div}(m\gamma) = 0$$

the velocity field $\gamma(x, t)$ playing here the role of a distributed control. Similarly, the mean field games models can also be reformulated as an optimal control problems for a density driven by a Fokker-Planck equation, see [14, 15].

An important research activity is currently going on about approximation procedures of different types of mean field games models, see [12] for a numerical method based on the reformulation of the model as an optimal control problem for the Fokker-Planck equation with an application in economics, and [10] for a very recent work on discrete time, finite state space mean field games. Also, in [1], the authors have proposed and analyzed the convergence as the discretization step tends to 0 of finite difference methods basically relying on monotone approximations of the Hamiltonian and on a suitable weak formulation of the Fokker-Planck equation, both in the stationary case and in the non-stationary one. Finally, applications of the theory of mean fields games to economics are considered in [11, 16].

The focus of the present paper is on numerical methods for the solutions of the general planning problems MFGPand MFGPP.

In Section 2, we propose finite differences semi-implicit schemes for the approximations of MFGPand MFGPP. Section 3 is dedicated to the optimal control formulation of those discrete schemes, following ideas in [5], [15], [18]. Existence and uniqueness of solutions of the discrete control problems are discussed, together with some convergence analysis on a penalized version of the semi-implicit scheme. In Section 4 we discuss a strategy based on the Newton algorithm for solving the discrete nonlinear system arising from the schemes introduced in the previous sections. Finally, in Section 5 we describe a few numerical experiments, which show that the proposed methods are successful even in some cases not covered by the theory. The schemes are also robust in the deterministic limit of the models, that is when the parameter ν is very close to 0.

2   Finite differences schemes

In this section we construct finite differences discretization schemes for system $(1), (2), (3), (4)$ and its penalized version (6). For simplicity of notations, we will always consider the case $d = 2$ although our approach and results hold for general d.

Let $\mathbb{T}_h^2$ be a uniform grid on the two-dimensional torus with mesh step h (assuming that $1/h$ is an integer N_h) and denote by x_{ij} a typical point in $\mathbb{T}_h^2$. Let N_T be a positive integer and $\Delta t = T/N_T, t_n = n\Delta t, n = 0, \dots, N_T$. The values of u and m at $(x_{i,j}, t_n)$ are approximated, respectively by $U_{i,j}^n$ and $M_{i,j}^n$.

We first discuss the approximations of the nonlinear operators in (1). The operator $V[m](x_{i,j})$ is approximated by

$$(V_h[M])_{i,j} = V[m_h](x_{i,j})$$

3

where m_h is the piecewise constant function taking the value $M_{i,j}$ in the square $|x - x_{i,j}|_\infty \leq h/2$. We assume that $V[m_h]$ can be computed in practice. In particular, if V is a local operator, i.e. $V[m](x) = V(m(x))$, then $(V_h[M])_{i,j} = V(M_{i,j})$.

We introduce next the finite difference operators

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

and define

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

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

In order to approximate the Hamiltonian H in equation (1), we consider a numerical Hamiltonian $g : \mathbb{T}^2 \times \mathbb{R}^4 \to \mathbb{R}$, $(x, q_1, q_2, q_3, q_4) \to g(x, q_1, q_2, q_3, q_4)$ satisfying the following conditions:

($\mathbf{G_1}$) monotonicity: g is nonincreasing with respect to q_1 and q_3 and nondecreasing with respect to q_2 and q_4.

($\mathbf{G_2}$) consistency: $g(x, q_1, q_1, q_2, q_2) = H(x, q)$, $\quad \forall x \in \mathbb{T}^2, \forall q = (q_1, q_2) \in \mathbb{R}^2$.

($\mathbf{G_3}$) differentiability: g is of class $\mathcal{C}^1$.

We will sometimes make further assumptions on g :

($\mathbf{G_4}$) convexity : $(q_1, q_2, q_3, q_4) \to g(x, q_1, q_2, q_3, q_4)$ is convex.

($\mathbf{G_5}$) coercivity:

$$\lim_{q_1 \to -\infty} \frac{g(x, q_1, q_2, q_3, q_4)}{|q_1|} = +\infty \text{ uniformly w.r.t. } x, q_2, q_3, q_4,$$

$$\lim_{q_2 \to +\infty} \frac{g(x, q_1, q_2, q_3, q_4)}{q_2} = +\infty \text{ uniformly w.r.t. } x, q_1, q_3, q_4,$$

$$\lim_{q_3 \to -\infty} \frac{g(x, q_1, q_2, q_3, q_4)}{|q_3|} = +\infty \text{ uniformly w.r.t. } x, q_1, q_2, q_4,$$

$$\lim_{q_4 \to +\infty} \frac{g(x, q_1, q_2, q_3, q_4)}{|q_4|} = +\infty \text{ uniformly w.r.t. } x, q_1, q_2, q_3.$$

This coercivity property implies that

$$\lim_{\|[D_h U]\|_\infty \to \infty} \frac{\max_{i,j} g\left(x_{i,j}, [D_h U]_{i,j}\right)}{\|[D_h U]\|_\infty} = +\infty$$

Standard examples of numerical Hamiltonians fulfilling these requirements are provided by Lax-Friedrichs or Godunov type schemes, see [1].

2.1 A scheme for the planning problem

The approximation of equation (1) is given by the semi-implicit scheme:

$$\frac{U_{i,j}^{n+1} - U_{i,j}^n}{\Delta t} - \nu \left(\Delta_h U^{n+1}\right)_{i,j} + g\left(x_{i,j}, \left[D_h U^{n+1}\right]_{i,j}\right) = (V_h[M^n])_{i,j}$$

Note that we could as well consider fully implicit or fully explicit schemes; the present choice will prove convenient because it is well suited to an optimal control formulation.

In order to approximate equation (2), it is convenient to consider its weak formulation which involves in particular the term

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

which by periodicity turns out be equal to

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

for any test function w. This term will be approximated by

$$h^2 \sum_{i,j} M_{i,j} \nabla_q g \left(x_{i,j}, [D_h U]_{i,j} \right) \cdot [D_h W]_{i,j} = h^2 \sum_{i,j} \mathcal{B}_{i,j}(U, M) W_{i,j},$$

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

$$\mathcal{B}_{i,j}(U, M) = \frac{1}{h} \left(\begin{pmatrix} M_{i,j} \frac{\partial g}{\partial q_1} \left(x_{i,j}, [D_h U]_{i,j} \right) - M_{i-1,j} \frac{\partial g}{\partial q_1} \left(x_{i-1,j}, [D_h U]_{i-1,j} \right) \\ + M_{i+1,j} \frac{\partial g}{\partial q_2} \left(x_{i+1,j}, [D_h U]_{i+1,j} \right) - M_{i,j} \frac{\partial g}{\partial q_2} \left(x_{i,j}, [D_h U]_{i,j} \right) \end{pmatrix} + \begin{pmatrix} M_{i,j} \frac{\partial g}{\partial q_3} \left(x_{i,j}, [D_h U]_{i,j} \right) - M_{i,j-1} \frac{\partial g}{\partial q_3} \left(x_{i,j-1}, [D_h U]_{i,j-1} \right) \\ + M_{i,j+1} \frac{\partial g}{\partial q_4} \left(x_{i,j+1}, [D_h U]_{i,j+1} \right) - M_{i,j} \frac{\partial g}{\partial q_4} \left(x_{i,j}, [D_h U]_{i,j} \right) \end{pmatrix} \right).$$

The discrete version of equation (2) is chosen as the following scheme:

$$\frac{M_{i,j}^{n+1} - M_{i,j}^n}{\Delta t} + \nu (\Delta_h M^n)_{i,j} + \mathcal{B}_{i,j} \left(U^{n+1}, M^n \right) = 0$$

Remark 1 It is important to realize that the operator $M \mapsto -\nu (\Delta_h M)_{i,j} - \mathcal{B}_{i,j}(U, M)$ is the adjoint of the linearization of the operator $U \mapsto -\nu (\Delta_h U)_{i,j} + g \left(x_{i,j}, [D_h U]_{i,j} \right)$. Finally, we introduce the compact and convex set

$$\mathcal{K} = \left\{ (M_{i,j})_{0 \leq i,j < N_T} : h^2 \sum_{i,j} M_{i,j} = 1, M_{i,j} \geq 0 \right\}$$

which can be viewed as the set of the discrete probability measures.

The fully discrete scheme for system (1),(2),(3),(4) is therefore the following:

$$\begin{cases} \frac{U_{i,j}^{n+1} - U_{i,j}^n}{\Delta t} - \nu (\Delta_h U^{n+1})_{i,j} + g \left(x_{i,j}, [D_h U^{n+1}]_{i,j} \right) = (V_h [M^n])_{i,j} \\ \frac{M_{i,j}^{n+1} - M_{i,j}^n}{\Delta t} + \nu (\Delta_h M^n)_{i,j} + \mathcal{B}_{i,j} (U^{n+1}, M^n) = 0, & n = 0, \dots N_T - 1, 0 \leq i,j < N_h \\ M^n \in \mathcal{K}, & n = 0, \dots N_T \\ M_{i,j}^{N_T} = (m_T)_{i,j}, \quad M_{i,j}^0 = (m_0)_{i,j}, & 0 \leq i,j < N_h \end{cases}$$

where

$$(m_T)_{i,j} = \frac{1}{h^2} \int_{|x - x_{i,j}|_\infty \leq h/2} m_T, \quad \text{and} \quad (m_0)_{i,j} = \frac{1}{h^2} \int_{|x - x_{i,j}|_\infty \leq h/2} m_0.$$

2.2 A scheme for the penalized planning problem

Here we introduce a discrete version of system MFGPP, see (6) in the Introduction, namely

$$\frac{U_{i,j}^{\epsilon,n+1} - U_{i,j}^{\epsilon,n}}{\Delta t} - \nu \left(\Delta_h U^{\epsilon,n+1}\right)_{i,j} + g\left(x_{i,j}, \left[D_h U^{\epsilon,n+1}\right]_{i,j}\right) = \left(V_h\left[M^n\right]\right)_{i,j},$$

$$\frac{M_{i,j}^{\epsilon,n+1} - M_{i,j}^{\epsilon,n}}{\Delta t} + \nu \left(\Delta_h M^{\epsilon,n}\right)_{i,j} + \mathcal{B}_{i,j}\left(U^{\epsilon,n+1}, M^{\epsilon,n}\right) = 0,$$

$$M^{\epsilon,n} \in \mathcal{K},$$

for $n = 0, \dots N_T - 1$ and $0 \le i, j < N_h$, with the final time and initial time conditions

$$U_{i,j}^{\epsilon,0} = \frac{1}{\epsilon}\left(M_{i,j}^{\epsilon,0} - (m_0)_{i,j}\right), \quad M_{i,j}^{\epsilon,N_T} = (m_T)_{i,j}, \quad \forall 0 \le i, j < N_h$$

Observe that the previous problem is a penalized approximation of (18). It can be proved that for $\epsilon > 0$, (20) − (23) has a unique solution (for the proof we refer to [1], Theorem 6).

3 Optimal control formulation of the discrete problems

3.1 The discrete planning problem

In this section we introduce an optimal control problem whose optimality conditions are interpreted as the semi-implicit scheme (18). In this way, using a convex duality argument based on the Fenchel-Rockafellar Theorem, we prove the existence of a solution to such a system, see Theorem 1 below.

In what follows, we assume that ν is nonnegative and that

$$\left(V_h[M]\right)_{i,j} = V\left(M_{i,j}\right), \text{ and that } V = W' \text{ where } W : \mathbb{R} \to \mathbb{R}$$
$$\text{is a strictly convex, coercive } C^2 \text{ function.}$$

It follows that the image of the interval $(0, +\infty)$ by V is some interval $\mathcal{J}_V = (\lambda, +\infty)$. We also assume g satisfies $(\mathbf{G_1}) - (\mathbf{G_5})$.

If χ denotes the indicator function of the set $\{m \ge 0\}$, the Legendre-Fenchel transform of $W + \chi$ is defined by

$$(W + \chi)^*(\alpha) = \sup_m [\alpha m - W(m) - \chi(m)]$$

It is clear that $(W + \chi)^*$ is convex, continuous and non decreasing. If $\alpha \in \mathcal{J}_V$ then $(W + \chi)^*(\alpha) = \alpha V^{-1}(\alpha) - W\left(V^{-1}(\alpha)\right)$. If $\alpha \notin \mathcal{J}_V$ then $(W + \chi)^*(\alpha) = -W(0)$.

Consider now the convex functional Θ on $\mathbb{R}^{N_T \times N_h^2} \times \mathbb{R}^{4N_T \times N_h^2}$:

$$\Theta(\alpha, \beta) = \sum_{n=1}^{N_T} \sum_{i,j} (W + \chi)^* \left(\alpha_{i,j}^n + g\left(x_{i,j}, [\beta^n]_{i,j}\right)\right)$$

where $\alpha = \left(\alpha_{i,j}^n\right), \beta = \left([\beta^n]_{i,j}\right)$ and $[\beta^n]_{i,j} = \left(\beta_{i,j}^{1,n}, \beta_{i,j}^{2,n}, \beta_{i,j}^{3,n}, \beta_{i,j}^{4,n}\right), 1 \le n \le N_T, 1 \le i, j \le N_h$.
The Legendre-Fenchel transform of Θ is defined by

$$\Theta^*(M, Z) = \sup_{\alpha, \beta} \left(\sum_{n=1}^{N_T} \sum_{i,j} M_{i,j}^{n-1} \alpha_{i,j}^n + \left\langle \left[Z^{n-1}\right]_{i,j}, [\beta^n]_{i,j}\right\rangle - (W + \chi)^* \left(\alpha_{i,j}^n + g\left(x_{i,j}, [\beta^n]_{i,j}\right)\right)\right),$$

where $M = \left(M_{i,j}^n\right), Z = \left([Z^n]_{i,j}\right)$ and $[Z^n]_{i,j} = \left(Z_{i,j}^{1,n}, Z_{i,j}^{2,n}, Z_{i,j}^{3,n}, Z_{i,j}^{4,n}\right), 0 \le n < N_T, 1 \le i, j \le N_h$ and $\langle [Z], [\beta]\rangle = \sum_{k=1}^{4} \beta^k Z^k$.

Remark 2 Note that in our definition, for $n = 1, \dots, N_T$, the dual variable of $\alpha_{i,j}^n$ is $M_{i,j}^{n-1}$, and the dual variable of $[\beta^n]_{i,j}$ is $\left[Z^{n-1}\right]_{i,j}$. This lag in the time index n will prove convenient for our purpose.

Let us introduce the minimization problem

where

$$\text{div}_h \left(Z^{n-1} \right)_{i,j} = \left(D_1^+ Z^{1,n-1} \right)_{i-1,j} + \left(D_1^+ Z^{2,n-1} \right)_{i,j} + \left(D_2^+ Z^{3,n-1} \right)_{i,j-1} + \left(D_2^+ Z^{4,n-1} \right)_{i,j}$$

The above minimization problem is an optimal control problem for a discrete density driven by a discrete Fokker-Planck equation. The data $(m_0)_{i,j}$, $(m_T)_{i,j} \in \mathcal{K}$ are discrete probability densities.

We are going to prove next that if the initial datum satisfies $(m_0)_{i,j} > 0$ for all i,j, then the optimal control problem above has at least a solution (M, Z), that there exists a solution (α, β) of the dual problem and that the optimality conditions at the saddle point coincide with the discrete scheme (18). The argument is based on convex duality and the Fenchel-Rockafellar theorem. Let us introduce for this purpose the functionals $\mathcal{F}, \Lambda, \Sigma$ by setting

$$\mathcal{F}(\Psi) = \frac{1}{\Delta t} \left(\sum_{i,j} (m_0)_{i,j} \, \Psi_{i,j}^0 - \sum_{i,j} m_{T,i,j} \Psi_{i,j}^{N_T} \right)$$

$$(\alpha, \beta) = \Lambda(\Psi) \Leftrightarrow \begin{cases} \alpha_{i,j}^{n+1} = \dfrac{\Psi_{i,j}^{n+1} - \Psi_{i,j}^{n}}{\Delta t} - \nu \left(\Delta_h \Psi^{n+1} \right)_{i,j}, \\[2mm] \left[\beta^{n+1} \right]_{i,j} = \left[D_h \Psi^{n+1} \right]_{i,j}, \quad 0 \le n < N_T, \end{cases}$$

and, finally,

$$\Sigma(\alpha, \beta) = \begin{cases} \mathcal{F}(\Psi) & \text{if } \exists \Psi \quad \text{s.t.} \quad (\alpha, \beta) = \Lambda(\Psi) \text{ and } \sum_{i,j} \Psi_{i,j}^0 = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Lemma 1 The functional Θ is convex and continuous. The functional Σ is convex and lower semicontinuous. Moreover, the following constraints qualification property holds: there exists (α, β) such that $\Sigma(\alpha, \beta) < +\infty$ (and of course $\Theta(\alpha, \beta) < +\infty$).

Proof. Convexity and continuity/semicontinuity are straightforward to check. For the constraint qualification it is enough to solve

$$\frac{U_{i,j}^{n+1} - U_{i,j}^{n}}{\Delta t} - \nu \left(\Delta_h U^{n+1} \right)_{i,j} + g \left(x_{i,j}, \left[D_h U^{n+1} \right]_{i,j} \right) = r_{i,j}^{n+1}$$

where $r_{i,j}^{n+1} \in \mathcal{J}_V$ (recall (24)) for all i,j,n, with an initial datum $U_{i,j}^0$ such that $\sum_{i,j} U_{i,j}^0 = 0$. Then, take (α, β) be such that

$$\begin{cases} \alpha_{i,j}^{n+1} = \dfrac{U_{i,j}^{n+1} - U_{i,j}^{n}}{\Delta t} - \nu \left(\Delta_h U^{n+1} \right)_{i,j}, \\[2mm] \left[\beta^{n+1} \right]_{i,j} = \left[D_h U^{n+1} \right]_{i,j}, \quad 0 \le n < N_T - 1, \end{cases}$$

Thus

$$\Sigma(\alpha, \beta) = -\frac{1}{\Delta t} \sum_{i,j} (m_T)_{,i,j} \, U_{i,j}^{N_T} + \frac{1}{\Delta t} \sum_{i,j} (m_0)_{,i,j} \, U_{i,j}^0 < +\infty$$

Lemma 2 The functionals Θ^* and Σ^* are convex and lower semicontinuous. Moreover,

$$\Theta^*(M, Z) = \sum_{n=1}^{N_T} \sum_{i,j} (W + \chi) \left(M_{i,j}^{n-1} \right) + \sup_{\beta} \left\{ \sum_{n=1}^{N_T} \sum_{i,j} \left\langle \left[Z_{i,j}^{n-1} \right], [\beta^n]_{i,j} \right\rangle - M_{i,j}^{n-1} g \left(x_{i,j}, [\beta^n]_{i,j} \right) \right\}$$

and

$$\Sigma^*(M,Z) = \sup_{\Psi} \left\{ \begin{array}{l} \frac{1}{\Delta t} \sum_{i,j} \left((m_T)_{i,j} + M_{i,j}^{N_T} \right) \Psi_{i,j}^{N_T} - \frac{1}{\Delta t} \sum_{i,j} \left((m_0)_{i,j} + M_{i,j}^0 \right) \Psi_{i,j}^0 \\ + \sum_{n=0}^{N_T-1} \sum_{i,j} \Psi_{i,j}^{n+1} \left(\frac{M_{i,j}^n - M_{i,j}^{n+1}}{\Delta t} - \nu (\Delta_h M^n)_{i,j} - \mathrm{div}_h (Z^n)_{i,j} \right) \end{array} \right\}.$$

Proof. Convexity and semi-continuity are a direct consequence of the previous lemma and the properties of the Legendre-Fenchel transform. Adding and subtracting a same term in (25), we get

$$\Theta^*(M,Z) = \sup_{\alpha,\beta} \left\{ \begin{array}{l} \sum_{n=1}^{N_T} \sum_{i,j} -M_{i,j}^{n-1} g\left(x_{i,j}, [\beta^n]_{i,j} \right) + \left\langle [\beta^n]_{i,j}, [Z^{n-1}]_{i,j} \right\rangle \\ + \sum_{n=1}^{N_T} \sum_{i,j} M_{i,j}^{n-1} \left(\alpha_{i,j}^n + g\left(x_{i,j}, [\beta^n]_{i,j} \right) \right) - (W + \chi)^* \left(\alpha_{i,j}^n + g\left(x_{i,j}, [\beta^n]_{i,j} \right) \right) \end{array} \right\}.$$

A simple computation shows that this can be written as

$$\sup_{\gamma,\beta} \left\{ \sum_{n=1}^{N_T} \sum_{i,j} -M_{i,j}^{n-1} g\left(x_{i,j}, [\beta^n]_{i,j} \right) + \left\langle [\beta^n]_{i,j}, \left[Z^{n-1} \right]_{i,j} \right\rangle + M_{i,j}^{n-1} \gamma_{i,j}^n - (W + \chi)^* \left(\gamma_{i,j}^n \right) \right\}$$

and the formula for Θ^* in the statement follows. As for Σ^*, observe that

$$\Sigma^*(M,Z) = \sup_{\alpha,\beta} \left(\sum_{n=0}^{N_T-1} \sum_{i,j} M_{i,j}^n \alpha_{i,j}^{n+1} + \left\langle [Z^n]_{i,j}, \left[\beta^{n+1} \right]_{i,j} \right\rangle - \Sigma(\alpha,\beta) \right)$$

Thus, taking the definition of Σ and Λ into account,

$$\Sigma^*(M,Z) = \sup_{\Psi} \left\{ \begin{array}{l} \frac{1}{\Delta t} \sum_{i,j} (m_T)_{i,j} \Psi_{i,j}^{N_T} - \frac{1}{\Delta t} \sum_{i,j} (m_0)_{i,j} \Psi_{i,j}^0 \\ + \sum_{n=0}^{N_T-1} \sum_{i,j} M_{i,j}^n \left(\frac{\Psi_{i,j}^{n+1} - \Psi_{i,j}^n}{\Delta t} - \nu (\Delta_h \Psi^{n+1})_{i,j} \right) + \left\langle [Z^n]_{i,j}, [D_h \Psi^{n+1}]_{i,j} \right\rangle \end{array} \right\},$$

and the claimed formula for Σ^* easily follows by a discrete integration by part.

Using Lemma 2, it easy to realize that the optimal control problem (26) can be equivalently formulated as the unconstrained minimization problem

$$\min_{M,Z} \Theta^*(M,Z) + \Sigma^*(-M,-Z)$$

The qualification condition is fulfilled for this problem also:

Lemma 3 Assume that $(m_0)_{i,j} > 0$ for all i,j. Then there exists (M,Z) such that

$$\left\{ \begin{array}{l} \Theta^*(M,Z) < +\infty \\ \Sigma^*(-M,-Z) < +\infty \\ \Theta^* \text{ is continuous in a neighborhood of } M, Z. \end{array} \right.$$

Proof. Take $M_{i,j}^n = \frac{n}{N_T} (m_T)_{i,j} + \left(1 - \frac{n}{N_T} \right) (m_0)_{i,j}$, and choose ϕ^n such that

$$\Delta_h \phi^n = \frac{1}{\Delta t} \left(M^{n+1} - M^n \right) + \nu \Delta_h M^n, \quad n = 0, \ldots, N_T - 1.$$

Since ϕ^n is unique up to the addition of a constant, one can always choose the constant in such a way that $\phi^n < \underline{\phi} < 0$, where $\underline{\phi}$ is a fixed negative number.

Set then

$$Z_{i,j}^{1,n} = \frac{\phi_{i,j}^n}{h}, \quad Z_{i,j}^{2,n} = -\frac{\phi_{i,j}^n}{h}, \quad Z_{i,j}^{3,n} = \frac{\phi_{i,j}^n}{h}, \quad Z_{i,j}^{4,n} = -\frac{\phi_{i,j}^n}{h}$$

We have

$$\operatorname{div}_h (Z^n)_{i,j} = - (\Delta_h \phi^n)_{i,j}$$

Therefore

$$\begin{cases} \dfrac{M_{i,j}^{n+1} - M_{i,j}^n}{\Delta t} + \nu \left(\Delta_h M^n\right)_{i,j} + \operatorname{div}_h (Z^n)_{i,j} = 0, \quad 0 \le n < N_T \\ \qquad\qquad\qquad M_{i,j}^{N_T} = (m_T)_{i,j}, \\ \qquad\qquad\qquad M_{i,j}^0 = (m_0)_{i,j}, \\ \qquad\qquad\qquad M_{i,j} \ge 0. \end{cases}$$

Observe that the assumption $(m_0)_{i,j} > 0$ implies $M^n \ge \underline{m} > 0$ for all $n < N_T$. Using Lemma 2 , this implies $\Sigma^*(-M, -Z) = 0$. Also, taking the definition of Z into account,

$$\Theta^*(M, Z) = \sum_{n=1}^{N_T} \sum_{i,j} W\left(M_{i,j}^{n-1}\right) + \sup_{\beta} \left\{ \sum_{n=1}^{N_T} \sum_{i,j} \left(\frac{\phi_{i,j}^{n-1}}{h} \left(\beta_{i,j}^{1,n} - \beta_{i,j}^{2,n} + \beta_{i,j}^{3,n} - \beta_{i,j}^{4,n} \right) \right. \right.$$
$$- M_{i,j}^{n-1} g\left(x_{i,j}, [\beta^n]_{i,j} \right) \right).$$

Since $\phi < 0$ and $M^n > \underline{m} > 0, n = 0, \dots, N_T - 1$, from the coercivity $(\mathbf{G_5})$ of g we deduce that $\Theta^*(M, Z)$ is finite and Θ^* is continuous in a neighborhood of (M, Z).

The next result gives sufficient conditions for the existence of a solution of the discrete system (18).

Theorem 1 Assume that

- g satisfies $(\mathbf{G_1}) - (\mathbf{G_5})$

- V satisfies (24).

- $(m_0)_{i,j}, (m_T)_{i,j} \in \mathcal{K}$ with $(m_0)_{i,j} > 0, \forall i, j$

- either $\nu > 0$ or $\left(\nu = 0 \text{ and } (m_T)_{i,j} > 0, \forall i, j\right)$.

Then the saddle point problem:

$$\min_{M,Z} \Theta^*(M, Z) + \Sigma^*(-M, -Z) = - \min_{\alpha, \beta} (\Theta(\alpha, \beta) + \Sigma(\alpha, \beta))$$

has a solution $(M, Z), (\alpha, \beta)$ and there exists U such that $(\alpha, \beta) = \Lambda(U)$. Moreover, (M, Z) and U satisfy the optimality conditions of (31)

$$-\Lambda^*(M, Z) \in \partial \mathcal{F}(U),$$
$$\Lambda(U) \in \partial \Theta^*(M, Z),$$

which are equivalent to the discrete system (18).

Proof. By applying the Fenchel-Rockafellar Duality Theorem to Θ and Σ (see for example [2, 8, 3, 18]) and using Lemma 1 , there exists a solution (M, Z) of the problem

$$\Theta^*(M, Z) + \Sigma^*(-M, -Z) = \inf_{M,Z} \left(\Theta^*(M, Z) + \Sigma^*(-M, -Z)\right) = - \inf_{\alpha, \beta}(\Theta(\alpha, \beta) + \Sigma(\alpha, \beta)).$$

By applying the Fenchel-Rockafellar Duality Theorem to $(M, Z) \mapsto \Theta^*(M, Z)$ and $(M, Z) \mapsto \Sigma^*(-M, -Z)$, and using Lemmas 2 and 3 , we deduce that there exist (α, β) such that

$$\Theta(\alpha,\beta) + \Sigma(\alpha,\beta) = \inf_{\alpha,\beta}(\Theta(\alpha,\beta) + \Sigma(\alpha,\beta)) = -\inf_{M,Z}(\Theta^*(M,Z) + \Sigma^*(-M,-Z))$$

We have thus proved the existence of a solution of the saddle point problem (31). By the optimality conditions, see [3] Theorem 2.4 page 205, we get

$$\begin{aligned}
-\Lambda^*(M,Z) &\in \partial\mathcal{F}(\Psi),\\
(\alpha,\beta) = \Lambda(U) &\in \partial\Theta^*(M,Z)
\end{aligned}$$

Recalling the definition of $\mathcal{F}$, (36) is seen to be in fact equivalent to

$$\begin{cases}
\dfrac{M_{i,j}^{n+1} - M_{i,j}^n}{\Delta t} + \nu\left(\Delta_h M^n\right)_{i,j} + \mathrm{div}_h\left(Z^n\right)_{i,j} = 0, \quad 0 \le n < N_T,\\[2mm]
M_{i,j}^{N_T} = (m_T)_{i,j},\\[2mm]
M_{i,j}^0 = (m_0)_{i,j}.
\end{cases}$$

On the other hand, it is easy to see that (37) is equivalent to

$$\Theta^*(M,Z) = \sum_{n=1}^{N_T}\sum_{i,j}\left(M_{i,j}^{n-1}\alpha_{i,j}^n + \left\langle\left[Z^{n-1}\right]_{i,j}, [\beta^n]_{i,j}\right\rangle - (W+\chi)^*\left(\alpha_{i,j}^n + g\left(x_{i,j},[\beta^n]_{i,j}\right)\right)\right).$$

Introducing $\gamma_{i,j}^n = \alpha_{i,j}^n + g\left(x_{i,j},[\beta^n]_{i,j}\right), n = 1,\dots,N_T$, the latter equation is equivalent to

$$Z_{i,j}^{k,n} = M_{i,j}^n \frac{\partial g}{\partial q_k}\left(x_{i,j},\left[\beta^{n+1}\right]_{i,j}\right), \quad k = 1,\dots,4,$$

$$0 = \sum_{n=1}^{N_T}\sum_{i,j}\left(M_{i,j}^{n-1}\gamma_{i,j}^n - (W+\chi)^*\left(\gamma_{i,j}^n\right) - (W+\chi)\left(M_{i,j}^{n-1}\right)\right).$$

Equation (40) is equivalent to

$$M_{i,j}^n \ge 0,$$

$$\gamma_{i,j}^{n+1} = \frac{U_{i,j}^{n+1} - U_{i,j}^n}{n+1} - \nu\left(\Delta_h U^{n+1}\right)_{i,j} + g\left(x_{i,j},\left[D_h U^{n+1}\right]_{i,j}\right) = W'\left(M_{i,j}^n\right) \quad \text{if } M_{i,j}^n > 0,$$

$$\gamma_{i,j}^{n+1} = \frac{U_{i,j}^n - U_{i,j}}{\Delta t} - \nu\left(\Delta_h U^{n+1}\right)_{i,j} + g\left(x_{i,j},\left[D_h U^{n+1}\right]_{i,j}\right) \le W'\left(M_{i,j}^n\right) \quad \text{if } M_{i,j}^n = 0,$$

for $0 \le n < N_T$

From (38) and (39), we deduce

$$\frac{M_{i,j}^{n+1} - M_{i,j}^n}{\Delta t} + \nu\left(\Delta_h M^n\right)_{i,j} + \mathcal{B}_{i,j}\left(U^{n+1}, M^n\right) = 0, \quad 0 \le i,j < N_h, 0 \le n < N_T$$

The fact that $M^{N_T} \in \mathcal{K}$ and (42) imply that $h^2\sum_{i,j} M_{i,j}^n = 1$ for all $n, 0 \le n < N_T$. Finally $M^n \in \mathcal{K}$ because of (42)

Finally, let us prove that $M^n > 0$ for all $0 \le n < N_T$. Indeed, assume that the minimum of $M_{i,j}^n$ is 0 and is reached at $n_0 < N_T, i_0, j_0$. Equation (42) for $n = n_0, i = i_0$ and $j = j_0$ can be written

10

$$0 =$$

$$\frac{1}{\Delta t}M_{i_0,j_0}^{n_0+1} + \frac{\nu}{h^2}\left(M_{i_0+1,j_0}^{n_0} + M_{i_0-1,j_0}^{n_0} + M_{i_0,j_0+1}^{n_0} + M_{i_0,j_0-1}^{n_0}\right) +$$

$$\frac{1}{h}\left(\begin{array}{c} -M_{i_0-1,j_0}^{n_0}\frac{\partial g}{\partial q_1}\left(x_{i_0-1,j_0}, [D_hU^{n_0+1}]_{i_0-1,j_0}\right) + M_{i_0+1,j_0}^{n_0}\frac{\partial g}{\partial q_2}\left(x_{i_0+1,j_0}, [D_hU^{n_0+1}]_{i_0+1,j_0}\right) \\ -M_{i_0,j_0-1}^{n_0}\frac{\partial g}{\partial q_3}\left(x_{i_0,j_0-1}, [D_hU^{n_0+1}]_{i_0,j_0-1}\right) + M_{i_0,j_0+1}^{n_0}\frac{\partial g}{\partial q_4}\left(x_{i_0,j_0+1}, [D_hU^{n_0+1}]_{i_0,j_0+1}\right) \end{array}\right)$$

If $\nu > 0$, then the nonnegativity of M and the monotonicity of g imply that $M_{i_0\pm1,j_0}^{n_0} = M_{i_0,j_0\pm1}^{n_0} = 0$. We can therefore repeat the argument for the triplets of indices $(n_0, i_0 \pm 1, j_0)$ and $(n_0, i_0, j_0\pm 1)$. Repeating the argument as many times as necessary, we finally obtain that $M^{n_0} = 0$, which is impossible since $M^{n_0} \in \mathcal{K}$.

If $\nu = 0$ and $M^{N_T} > 0$, a similar argument gives that $M_{i_0,j_0}^{n_0+1} = 0$. After a finite number of steps, we get that $M_{i_0,j_0}^{N_T} = 0$, which is in contradiction with the hypothesis.

As a consequence, (38), (39) and (41) can be written:

$$\frac{U_{i,j}^{n+1} - U_{i,j}^n}{\Delta t} - \nu\left(\Delta_h U^{n+1}\right)_{i,j} + g\left(x_{i,j}, \left[D_h U^{n+1}\right]_{i,j}\right) = V\left(M_{i,j}^n\right),$$

$$\frac{M_{i,j}^{n+1} - M_{i,j}^n}{\Delta t} + \nu\left(\Delta_h M^n\right)_{i,j} + \mathcal{B}_{i,j}\left(U^{n+1}, M^n\right) = 0,$$

for $n = 0, \ldots N_T - 1$ and $0 \leq i, j < N_h$, with

$$M_{i,j}^{N_T} = (m_T)_{i,j}, \quad M_{i,j}^0 = (m_0)_{i,j}, \quad 0 \leq i, j < N_h,$$

and

$$M^n \in \mathcal{K}, \quad 0 \leq n \leq N_T$$

Recognizing that equations (43) to (46) comprise indeed the semi-implicit finite difference scheme (18), the proof is complete.

System (18) also enjoys some uniqueness property:

Proposition 1 Under the same assumptions as in Theorem 1, if $\left(U_{i,j}^n, M_{i,j}^n\right)_{n,i,j}$ and $\left(\widetilde{U}_{i,j}^n, \widetilde{M}_{i,j}^n\right)_{n,i,j}$ are solutions of system (18), then

$$M_{i,j}^n = \widetilde{M}_{i,j}^n \text{ for all } n = 0, \ldots, N_T; \text{ and for all } (i, j).$$

Moreover if the numerical Hamiltonian g is strictly convex and $\sum_{ij} U_{ij}^0 = \sum_{ij} \widetilde{U}_{ij}^0 = 0$, then

$$U_{i,j}^n = \widetilde{U}_{i,j}^n \text{ for all } n = 0, \ldots, N_T; \text{ and for all } (i, j).$$

Proof. The proof resembles very much the proofs of uniqueness in [14] and in [1]. We give it for completeness.

Multiplying (43) satisfied by $\left(\widetilde{U}_{i,j}^n, \widetilde{M}_{i,j}^n\right)_{n,i,j}$ by $M_{i,j}^n - \widetilde{M}_{i,j}^n$, doing the same thing with (43) satisfied by $\left(U_{i,j}^n, M_{i,j}^n\right)_{n,i,j}$ and subtracting, then summing the results for all $n = 0, \ldots, N_T - 1$ and all (i, j), we obtain

$$\sum_{n=0}^{N_T-1}\frac{\left(\left(U^{n+1} - \widetilde{U}^{n+1}\right) - \left(U^n - \widetilde{U}^n\right), M^n - \widetilde{M}^n\right)_2}{\Delta t} - \nu\left(\Delta_h\left(U^{n+1} - \widetilde{U}^{n+1}\right), M^n - \widetilde{M}^n\right)_2$$

$$+ \sum_{n=0}^{N_T-1}\sum_{i,j}\left(g\left(x_{i,j}, \left[D_h U^{n+1}\right]_{i,j}\right) - g\left(x_{i,j}, \left[D_h\widetilde{U}^{n+1}\right]_{i,j}\right)\right)\left(M_{i,j}^n - \widetilde{M}_{i,j}^n\right)$$

$$= \sum_{n=0}^{N_T-1}\left(V_h\left[M^n\right] - V_h\left[\widetilde{M}^n\right], M^n - \widetilde{M}^n\right)_2$$

11

where $(X, Y)_2 = \sum_{i,j} X_{i,j} Y_{i,j}$. Similarly, subtracting the equation (44) satisfied by $\left(\tilde{U}_{i,j}^n, \tilde{M}_{i,j}^n \right)_{n,i,j}$ from the same equation satisfied by $\left(U_{i,j}^n, M_{i,j}^n \right)_{n,i,j}$, and multiplying the result by $U_{i,j}^{n+1} - \tilde{U}_{i,j}^{n+1}$, sum for all $n = 0, \ldots, N_T - 1$ and all (i, j) leads to

$$\frac{1}{\Delta t} \sum_{n=0}^{N_T - 1} \left(\left(\left(M^{n+1} - M^n \right) - \left(\tilde{M}^{n+1} - \tilde{M}^n \right), \left(U^{n+1} - \tilde{U}^{n+1} \right) \right)_2 + \nu \left(\left(M^n - \tilde{M}^n \right), \Delta_h \left(U^{n+1} - \tilde{U}^{n+1} \right) \right)_2 \right.$$

$$- \sum_{n=0}^{N_T - 1} \sum_{i,j} M_{i,j}^n \left[D_h \left(U^{n+1} - \tilde{U}^{n+1} \right) \right]_{i,j} \cdot \nabla_q g \left(x_{i,j}, \left[D_h U^{n+1} \right]_{i,j} \right)$$

$$+ \sum_{n=0}^{N_T - 1} \sum_{i,j} \tilde{M}_{i,j}^n \left[D_h \left(U^{n+1} - \tilde{U}^{n+1} \right) \right]_{i,j} \cdot \nabla_q g \left(x_{i,j}, \left[D_h \tilde{U}^{n+1} \right]_{i,j} \right) = 0.$$

Adding (47) and (48) and leads to

$$\sum_{n=0}^{N_T - 1} \sum_{i,j} M_{i,j}^n \left(g \left(x_{i,j}, \left[D_h \tilde{U}^{n+1} \right]_{i,j} \right) - g \left(x_{i,j}, \left[D_h U^{n+1} \right]_{i,j} \right) - \left[D_h \left(\tilde{U}^{n+1} - U^{n+1} \right) \right]_{i,j} \cdot \nabla_q g \left(x_{i,j}, \left[D_h U^{n+1} \right]_{i,j} \right) \right)$$

$$+ \sum_{n=0}^{N_T - 1} \sum_{i,j} \tilde{M}_{i,j}^n \left(g \left(x_{i,j}, \left[D_h U^{n+1} \right]_{i,j} \right) - g \left(x_{i,j}, \left[D_h \tilde{U}^{n+1} \right]_{i,j} \right) - \left[D_h \left(U^{n+1} - \tilde{U}^{n+1} \right) \right]_{i,j} \cdot \nabla_q g \left(x_{i,j}, \left[D_h \tilde{U}^{n+1} \right]_{i,j} \right) \right)$$

$$+ \sum_{n=0}^{N_T - 1} \left(V_h \left[M^n \right] - V_h \left[\tilde{M}^n \right], M^n - \tilde{M}^n \right)_2 = 0$$

because M and $\tilde{M}$ satisfy the same final and terminal conditions. The three terms in the sum above being nonnegative, they must be zero. The strict monotonicity of V implies that $\tilde{M}^n = M^n$ for all $n = 1, \ldots, N_T - 1$.

If g is strictly convex, we also get that $[D_h U^n] = \left[D_h \tilde{U}^n \right]$ for all $n = 0, \ldots, N_T$. Hence $\Delta_h U^n = \Delta_h \tilde{U}^n$ and substituting in the equation satisfied by U and $\tilde{U}$ we get

$$U_{i,j}^{n+1} - U_{i,j}^n = \tilde{U}_{i,j}^{n+1} - \tilde{U}_{i,j}^n \quad \text{for all } n = 0, \ldots, N_T - 1$$

Now it is clear that if $\tilde{U}_{ij}^0 = U_{ij}^0$ then by (49) $\tilde{U}_{ij}^n = U_{ij}^n$, for all n. Hence, there exists $\bar{i}, \bar{j}$ such that $U_{\bar{i},\bar{j}}^0 - \tilde{U}_{\bar{i},\bar{j}}^0 \geq \delta > 0$. By $(D_1^+ U)_{ij} = (D_1^+ \tilde{U})_{ij}$, we get that $U_{ij}^0 - \tilde{U}_{ij}^0 \geq \delta > 0$ for any i, j and this gives a contradiction to $\sum_{i,j} U_{ij}^0 = \sum_{i,j} \tilde{U}_{ij}^0 = 0$.

3.2 The discrete penalized planning problem

We discuss briefly the optimal control interpretation of the discrete penalized planning problem (20)-(23) and prove some estimates on the ϵ-dependence of its solutions.

Theorem 2 Under the same assumptions as in Theorem 1, the minimization problem

$$\text{Minimize } \Theta^*(M, Z) + \frac{1}{2\epsilon \Delta t} \sum_{i,j} \left(M_{i,j}^0 - (m_0)_{i,j} \right)^2 \text{ subject to}$$

$$\begin{cases} \dfrac{M_{i,j}^{n+1} - M_{i,j}^n}{\Delta t} + \nu \left(\Delta_h M^n \right)_{i,j} + \text{div}_h \left(Z^n \right)_{i,j} = 0, \quad 0 \leq n < N_T, \\ M_{i,j}^{N_T} = (m_T)_{i,j}, \end{cases}$$

where Θ^* is given by (25) has a solution (M^ϵ, Z^ϵ). Moreover, there exists a solution U^ϵ of the dual problem (which we do not write) and

$$Z_{i,j}^{\epsilon,k,n} = M_{i,j}^{\epsilon,n} \frac{\partial g}{\partial q_k}\left(x_{i,j}, \left[D_h U^{\epsilon,n+1}\right]_{i,j}\right)$$

The pair (M^ϵ, U^ϵ) is the solution of the discrete penalized system (20)-(23).

The existence and uniqueness of (20)-(23) have already been discussed in §2.2. The remaining part of the proof of Theorem 2 is almost the same as the proof of Theorem 1: we skip it for brevity.

Proposition 2 Under the same assumptions as in Theorem 1 the following estimate holds

$$\max_{i,j}\left|M_{i,j}^{\epsilon,0} - (m_0)_{i,j}\right| \le C\epsilon^{\frac{1}{2}}$$

for a constant C which may depend on h and Δt but not on ϵ.

Proof. For a solution (U, M) of (43)-(46), (M, Z) with $Z_{i,j}^{k,n} = M_{i,j}^n \frac{\partial g}{\partial q_k}\left(x_{i,j}, \left[D_h U^{n+1}\right]_{i,j}\right)$ is suboptimal for (50). Therefore

$$\Theta^*\left(M^\epsilon, Z^\epsilon\right) + \frac{1}{2\epsilon\Delta t}\sum_{i,j}\left(M_{i,j}^{\epsilon,0} - (m_0)_{i,j}\right)^2 \le \Theta^*(M, Z).$$

But

$$\Theta^*\left(M^\epsilon, Z^\epsilon\right)$$
$$= \sum_{n=1}^{N_T}\sum_{i,j} W\left(M_{i,j}^{\epsilon,n-1}\right) + \sup_\beta \sum_{n=1}^{N_T}\sum_{i,j} M_{i,j}^{\epsilon,n-1}\left(\left\langle\nabla_q g\left(x_{i,j}, [D_h U^{\epsilon,n}]_{i,j}\right), [\beta^n]_{i,j}\right\rangle - g\left(x_{i,j}, [\beta^n]_{i,j}\right)\right)$$
$$\ge \sum_{n=1}^{N_T}\sum_{i,j} W\left(M_{i,j}^{\epsilon,n-1}\right) - \sum_{n=1}^{N_T}\sum_{i,j} M_{i,j}^{\epsilon,n-1} g\left(x_{i,j}, 0\right) \ge C$$

where C does not depend on ϵ. From (52) and (53),

$$\frac{1}{2\epsilon\Delta t}\sum_{i,j}\left(M_{i,j}^{\epsilon,0} - (m_0)_{i,j}\right)^2 \le C$$

where C does not depend on ϵ. This yields the desired result. Proposition 3 With the same assumptions as in Theorem 1, the solution of the penalized discrete system (20)-(23) satisfies

$$\max_{n,i,j}\left|U_{i,j}^{\epsilon,n}\right| \le C$$

for a constant C which may depend of h and Δt but not of ϵ.

Proof. Throughout the proof, C denotes a positive constant that may vary and depend on h and Δt but is independent of ϵ.

Proposition 2 implies that

$$\min_{i,j} M^{\epsilon,0} \ge \alpha > 0$$

for ϵ small enough and α independent of ϵ, because $\min_{i,j}(m_0)_{i,j} > 0$.

Let $\mathcal{N}$ be the vector space of the grid functions with zero average:

$$\mathcal{N} = \left\{(W_{i,j})_{0\le i,j<N_h} : \sum_{i,j} W_{i,j} = 0\right\}$$

13

Consider the functional

$$J : \mathcal{N} \to \mathbb{R}, \quad W \mapsto - \sum_{i,j} W_{i,j} \left(\frac{M_{i,j}^{\epsilon,1} - M_{i,j}^{\epsilon,0}}{\Delta t} + \nu \left(\Delta_h M^{\epsilon,0} \right)_{i,j} \right) + \sum_{i,j} M_{i,j}^{\epsilon,0} g \left(x_{i,j}, [D_h W]_{i,j} \right)$$

We know that there exists a positive constant c such that $\| W \|_\infty \leq c \, \| [D_h W] \|_\infty$, $\forall W \in \mathcal{N}$. This observation, (55) and (13) imply that

$$\lim_{\| [D_h W] \|_\infty \to \infty} J(W) = +\infty$$

Since the functional J is continuous, (58) implies the existence of a minimum $\bar{U}$. It is also possible to check that there exists a constant C which only depends on α, h and Δt such that the minimum points satisfy $\| \bar{U} \|_\infty \leq C$. From $(\mathbf{G}_3)$ and $(\mathbf{G}_4)$, the Euler equations

$$- \sum_{i,j} W_{i,j} \left(\frac{M_{i,j}^{\epsilon,1} - M_{i,j}^{\epsilon,0}}{\Delta t} + \nu \left(\Delta_h M^{\epsilon,0} \right)_{i,j} \right) + \sum_{i,j} M_{i,j}^{\epsilon,0} \nabla_q g \left(x_{i,j}, [D_h \bar{U}]_{i,j} \right) \cdot [D_h W]_{i,j} = 0, \quad \forall W \in \mathcal{N}$$

characterize the points of minimum. They can be written

$$\sum_{i,j} W_{i,j} \left(\frac{M_{i,j}^{\epsilon,1} - M_{i,j}^{\epsilon,0}}{\Delta t} + \nu \left(\Delta_h M^{\epsilon,0} \right)_{i,j} + \mathcal{B}_{i,j} \left(\bar{U}, M^{\epsilon,0} \right) \right) = 0, \quad \forall W \in \mathcal{N}$$

On the other hand,

$$\sum_{i,j} \left(\frac{M_{i,j}^{\epsilon,1} - M_{i,j}^{\epsilon,0}}{\Delta t} + \nu \left(\Delta_h M^{\epsilon,0} \right)_{i,j} + \mathcal{B}_{i,j} \left(\bar{U}, M^{\epsilon,0} \right) \right) = 0$$

because $M^{\epsilon,1} \in \mathcal{K}$ and $M^{\epsilon,0} \in \mathcal{K}$. Therefore $\bar{U}$ is a minimum of J if and only if $\bar{U}$ is a solution of

$$\frac{M_{i,j}^{\epsilon,1} - M_{i,j}^{\epsilon,0}}{\Delta t} + \nu \left(\Delta_h M^{\epsilon,0} \right)_{i,j} + \mathcal{B}_{i,j} \left(\bar{U}, M^{\epsilon,0} \right) = 0$$

This proves that for the constant C introduced above, the solutions of (59) in $\mathcal{N}$ satisfy $\| \bar{U} \|_\infty \leq C$. This implies that $\| U^{\epsilon,1} - \langle U^{\epsilon,1} \rangle \|_\infty \leq C$.

On the other hand, from (20) for $n = 0$ and the fact that $U^{\epsilon,0} \in \mathcal{N}$, we find that $| \langle U^{\epsilon,1} \rangle | \leq C$ and therefore $\| U^{\epsilon,1} \|_\infty \leq C$. Then, again from (20) for $n = 0$, we obtain that $\| U^{\epsilon,0} \|_\infty \leq C$. The estimate for $U^{\epsilon,n}, n > 1$ follows as usual.

Using (23), an easy consequence of Proposition 3 is the following improvement of Proposition 2

Corollary 1 With the same assumptions as in Theorem 1, the solution of the penalized discrete system (20)-(23) satisfies

$$\max_{i,j} \left| M_{i,j}^{\epsilon,0} - (m_0)_{i,j} \right| \leq C\epsilon$$

for a constant C which may depend of h and Δt but not of ϵ.

Also since $\mathcal{K}^{N_T+1}$ is compact, there exists a family of grid functions (M^n) and a subsequence still noted ϵ such that $\lim_{\epsilon \to 0} \max_n \| M^{\epsilon,n} - M^n \|_\infty = 0$. We aim at passing to the limit as $\epsilon \to 0$.

Proposition 4 With the same assumptions as in Theorem 1. For a subsequence still called ϵ, let $(U^{\epsilon,n}, M^{\epsilon,n})$ be a solution of (20)-(23) and (M^n) be a family of grid functions in $\mathcal{K}$ such that $\lim_{\epsilon \to 0} \max_n \| M^{\epsilon,n} - M^n \|_\infty = 0$. There exists a family of grid functions (U^n) such that up to a further extraction of a subsequence, $\lim_{\epsilon \to 0} \max_n \| U^{\epsilon,n} - U^n \|_\infty = 0$ and $(U^n, M^n)_n$ is a solution of (43)-(46).

Moreover, if the numerical Hamiltonian g is strictly convex, all the sequence $(U^{\epsilon,n}, M^{\epsilon,n})$ converges to the unique solution of (43)-(46).

Proof. From (60), we deduce that $M^0_{i,j} = (m_0)_{i,j}$. From the uniform bound of $U^{\epsilon,n}$, it is possible to find a family of grid functions (U^n) such that up to the extraction of a subsequence $\lim_{\epsilon \to 0} \max_n \|U^{\epsilon,n} - U^n\|_\infty = 0$ and by passing to the limit, we find that $(U^n, M^n)_n$ is a solution of (43)-(46).

If g is strictly convex, any converging subsequence converges to the unique solution of (43)(46) and this gives the convergence of all the sequence.

3.3 Remarks on the optimal control formulation of the original pde's system

Here, we are going to introduce an optimal control problem whose optimality conditions give the system (1)-(3), at least in the case where the operator V is the gradient of some functional W. The optimal control problem is similar to the one proposed in [14, 15] for the so-called finite horizon mean field game problem, i.e. system (1)-(2) with the terminal conditions

$$u(0,x) = V_0[m](0,x), \quad m(T,x) = m_T(x).$$

As mentioned in the introduction, this approach is reminiscent of the simpler computational fluid mechanics formulation of the Monge-Kantorovith mass transfer problem proposed by Benamou and Brenier [5, 6, 7]. We refer to [18] for a rigorous presentation of the ideas in [5] in the framework of convex duality.

For simplicity, we take $\nu = 1$. The optimal control problem is to minimize with respect to admissible controls α the functional

$$\int_0^T \int_{\mathbb{T}^d} [m(t,x)L(x,\alpha(t,x)) + W(m)]dtdx$$

subject to the constraints

$$\begin{cases} \partial_t m + \Delta m + \operatorname{div}(m\alpha) = 0, & \text{in } (0,T) \times \mathbb{T}^d, \\ m \geq 0, & \text{in } (0,T) \times \mathbb{T}^d, \\ m(T,x) = m_T(x) & \text{in } \mathbb{T}^d, \\ m(0,x) = m_0(x) & \text{in } \mathbb{T}^d \end{cases}$$

where $L : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is the Legendre-Fenchel transform of the of the convex Hamiltonian H and $W' = V$. We assume that

$(\mathbf{H_1})$ $L : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$, is $\mathcal{C}^1$, and coercive in α, uniformly with respect to x,

$(\mathbf{H_3})$ $W : \mathbb{R} \to \mathbb{R}$ is a strictly convex coercive and $\mathcal{C}^2$ function (the image of $(0,+\infty)$ by $V = W'$ is an interval $(\lambda,+\infty)$ which we call $\mathcal{J}_V$).

$(\mathbf{H_2})$ m_T and m_0 are probability densities on $\mathbb{T}^d$

The Hamiltonian

$$H(x,p) = \sup_{\alpha \in \mathbb{R}^d} (p \cdot \alpha - L(x,\alpha))$$

is convex and LSC with respect to its second argument. We assume for simplicity that H is of class $\mathcal{C}^1$

Let us proceed first in an informal way. Following [5], we begin by modifying the problem in such a way that the constraints become linear; for this purpose, we introduce

$$\Phi(x,m,z) = \sup_{p \in \mathbb{R}^d} (p \cdot z - mH(x,p)) \quad \text{for} \quad x \in \mathbb{T}^d, m \geq 0, z \in \mathbb{R}^d$$

It is easy to check that Φ is convex and lower semicontinuous with respect to (m, z) and that

$$
\begin{cases}
m > 0 \Rightarrow \Phi(x, m, z) = mL\left(x, \dfrac{z}{m}\right), \\
z = 0 \Rightarrow \Phi(x, m, z) \leq mL(x, 0), \\
m = 0, z \neq 0 \Rightarrow \Phi(x, m, z) = +\infty.
\end{cases}
$$

At least formally, the new problem is to minimize over vector fields $z = z(t, x)$ the convex functional

$$
\int_0^T \int_{\mathbb{T}^d} [\Phi(x, m(t, x), z(t, x)) + W(m(t, x))] dt dx
$$

subject to

$$
\begin{cases}
\partial_t m + \Delta m + \operatorname{div} z = 0, & \text{in } (0, T) \times \mathbb{T}^d, \\
m \geq 0, & \text{in } (0, T) \times \mathbb{T}^d, \\
m(T, x) = m_T(x), & \text{in } \mathbb{T}^d, \\
m(0, x) = m_0(x), & \text{in } \mathbb{T}^d.
\end{cases}
$$

Note that the set defined by the constraints is convex and nonempty at least if $m_0 > 0$ and if m_0 and m_T are smooth enough. Indeed, just take $\hat{m}(t, x) = \frac{t}{T} m_T(x) + \frac{T-t}{T} m_0(x)$ and solve

$$
-\Delta \zeta = \partial_t \hat{m} + \Delta \hat{m}, \quad \langle \zeta \rangle = 0, \quad \zeta(T) = \zeta_T
$$

and take $z(t) = \nabla \zeta(t)$.

The definition of the latter minimization problem is only formal. To do things in a more rigorous way, let us proceed as follows.

We call χ the indicator function of the set $\{m \geq 0\} : \chi(m) = 0$ if $m \geq 0$ and $\chi(m) = +\infty$ if $m < 0$. χ is convex and LSC. we introduce the Legendre-Fenchel transform of the function $W + \chi$:

$$
(W + \chi)^*(\alpha) = \sup_m (\alpha m - W(m) - \chi(m)).
$$

It is clear that $(W + \chi)^*$ is convex, continuous and non decreasing. If $\alpha \in \mathcal{J}_V$ then $(W + \chi)^*(\alpha) = \alpha V^{-1}(\alpha) - W\left(V^{-1}(\alpha)\right)$

Consider the convex and continuous functional Θ :

$$
\Theta : \mathcal{C}\left([0, T] \times \mathbb{T}^d\right) \times \left(\mathcal{C}\left([0, T] \times \mathbb{T}^d\right)\right)^d \to \mathbb{R} \cup \{+\infty\}
$$

$$
(\alpha, \beta) \mapsto \int_0^T \int_{\mathbb{T}^d} (W + \chi)^*(\alpha + H(\beta)).
$$

Let $\mathcal{M} = \mathcal{M}\left([0, T] \times \mathbb{T}^d\right)$ denote the set of Radon measures. The Legendre-Fenchel transform of Θ is

$$
\Theta^* : \mathcal{M} \times (\mathcal{M})^d \to \mathbb{R} \cup \{+\infty\}
$$

$$
(m, z) \mapsto \sup_{\alpha, \beta} \left(\langle m, \alpha \rangle + \langle z, \beta \rangle - \int_0^T \int_{\mathbb{T}^d} (W + \chi)^*(\alpha + H(\beta)) \right).
$$

We have

$$
\Theta^*(m, z) = \sup_{\alpha, \beta} \left(-\langle m, H(\beta) \rangle + \langle z, \beta \rangle + \langle m, \alpha + H(\beta) \rangle - \int_0^T \int_{\mathbb{T}^d} (W + \chi)^*(\alpha + H(\beta)) \right)
$$

$$
= \sup_{\beta} (-\langle m, H(\beta) \rangle + \langle z, \beta \rangle) + \sup_{\gamma} \left(\langle m, \gamma \rangle - \int_0^T \int_{\mathbb{T}^d} (W + \chi)^*(\gamma) \right)
$$

Focus on the supremum w.r.t. γ :

$$\sup_{\gamma} \left(\langle m, \gamma \rangle - \int_0^T \int_{\mathbb{T}^d} (W + \chi)^*(\gamma) \right) = \sup_{\gamma} \left(\langle m, \gamma \rangle - \int_0^T \int_{\mathbb{T}^d} \sup_{y \in \mathbb{R}} (y\gamma(t,x) - W(y) - \chi(y)) \right).$$

If $dm = \frac{dm}{d\mu} d\mu + dm_s$, where μ is the Lebesgue measure on $[0,T] \times \mathbb{T}^d$, the quantity above reads

$$\sup_{\gamma} \left(\langle m_s, \gamma \rangle + \int_0^T \int_{\mathbb{T}^d} \inf_{y \in \mathbb{R}} \left(\left(\frac{dm}{d\mu}(t,x) - y \right) \gamma(t,x) + W(y) + \chi(y) \right) \right).$$

The supremum is $+\infty$ if $m_s \neq 0$ or if $\frac{dm}{d\mu}$ is non positive. In the opposite case, the supremum is $\widetilde{W}(m) = \sup_{\gamma \in \mathcal{C}} \int_0^T \int_{\mathbb{T}^d} \left(\frac{dm}{d\mu}(t,x)\gamma(t,x) - W^*(\gamma(t,x)) \right) d\mu$, which coincides with $W(m) = \int_0^T \int_{\mathbb{T}^d} W\left(\frac{dm}{d\mu}(t,x) \right) d\mu$ if $\frac{dm}{d\mu}(t,x)$ is continuous.

Therefore

$$\Theta^*(m,z) = \sup_{\beta} \{ \langle z, \beta \rangle - \langle m, H(\beta) \rangle + (\widetilde{W} + \tilde{\chi})(m) \}$$

where $\tilde{\chi}$ is the indicator function of the set of nonnegative measures on $[0,T] \times \mathbb{T}^d$ with a L^1 density w.r.t the Lebesgue measure. Thus the problem of minimizing Θ^* is the rigorous version of the minimization problem above. The construction of Θ^* shows that Θ^* is convex and LSC on $\mathcal{M}\left([0,T] \times \mathbb{T}^d\right) \times \left(\mathcal{M}\left([0,T] \times \mathbb{T}^d\right)\right)^d$

On the other hand, consider the functional $\Sigma : \mathcal{C}\left([0,T] \times \mathbb{T}^d\right) \times \left(\mathcal{C}\left([0,T] \times \mathbb{T}^d\right)\right)^d \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\Sigma(\alpha, \beta) = \begin{cases} \mathcal{F}(\phi) \text{ if } (\alpha, \beta) = \Lambda(\phi), \\ +\infty \text{ in the opposite case} \end{cases}$$

where $\mathcal{F}$ is the linear form on $\left\{ \phi \in \mathcal{C}^{1,2}\left([0,T] \times \mathbb{T}^d\right) : \int_{\mathbb{T}^d} \phi(t=0) = 0 \right\}$ defined by

$$\mathcal{F}(\phi) = - \langle m_T, \phi(t=T) \rangle + \langle m_0, \phi(t=0) \rangle$$

and where Λ is the linear operator from $\left\{ \phi \in \mathcal{C}^{1,2}\left([0,T] \times \mathbb{T}^d\right) : \int_{\mathbb{T}^d} \phi(t=0) = 0 \right\}$ to $\mathcal{C}([0,T] \times \mathbb{T}^d) \times \left(\mathcal{C}\left([0,T] \times \mathbb{T}^d\right)\right)^d$ defined by

$$(\alpha, \beta) = \Lambda(\phi) \Leftrightarrow \begin{cases} \alpha = \partial_t \phi - \Delta\phi \\ \beta = \nabla\phi \end{cases}$$

It is clear that Σ is convex. The Legendre-Fenchel transform of Σ is

$$\Sigma^*(m,z) = \sup_{\alpha, \beta} (\langle m, \alpha \rangle + \langle z, \beta \rangle - \Sigma(\alpha, \beta))$$

Thus

$$\Sigma^*(m,z) = \sup_{\phi} \left(- \langle m_T, \phi(t=T) \rangle + \langle m_0, \phi(t=0) \rangle + \langle m, \partial_t\phi - \Delta\phi \rangle + \langle z, \nabla\phi \rangle \right)$$

and Σ^* is a convex and LSC functional.

Moreover, there exists $(\tilde{\alpha}, \tilde{\beta}) \in \mathcal{C}\left([0,T] \times \mathbb{T}^d\right) \times \left(\mathcal{C}\left([0,T] \times \mathbb{T}^d\right)\right)^d$ such that $\Sigma(\alpha, \beta) < +\infty$ (and of course $\Theta(\alpha, \beta) < +\infty$). Indeed it is enough to solve

$$\partial_t\phi - \Delta\phi + H(x, \nabla\phi) = g$$

17

where g is a smooth function taking its values in $\mathcal{J}_V$, with a smooth initial data ϕ_0. It can be proved that ϕ is smooth. Then one can take $\tilde{\alpha} = \partial_t \phi - \Delta \phi, \tilde{\beta} = \nabla \phi$. We have $\tilde{\alpha} + H(\tilde{\beta}) = g$, thus $\Theta(\alpha, \beta) = (\alpha + H(\beta))V^{-1}(\alpha + H(\beta)) - W\left(V^{-1}(\alpha + H(\beta))\right)$ for (α, β) close to $(\tilde{\alpha}, \tilde{\beta})$. On the other hand, $\Sigma(\tilde{\alpha}, \tilde{\beta}) = -\langle m_T, \phi(t = T)\rangle + \langle m_0, \phi(t = 0)\rangle < +\infty$.

We can apply Fenchel-Rockafellar duality theorem, see for example $[2, 8, 3, 18]$:

$$\inf_{\alpha, \beta}(\Theta + \Sigma) = -\min_{m, z}\left(\Theta^*(m, z) + \Sigma^*(-m, -z)\right)$$

From [3] (theorem 2.4 page 205), the problem of finding (α, β) minimizing $\Theta + \Sigma$ and a solution of the dual problem, i.e. finding (m, z) minimizing $\Theta^*() +, \Sigma^*(-, -)$ is equivalent to solving the system of optimality conditions: $(\alpha, \beta) = \Lambda(\phi)$ with

$$\begin{aligned}-\Lambda^*(m, z) &\in \partial \mathcal{F}(\phi),\\ \Lambda(\phi) &\in \partial \Theta^*(m, z).\end{aligned}$$

Condition (66) is equivalent to the weak formulation of

$$\begin{aligned}\partial_t m + \Delta m + \mathrm{div}(z) &= 0,\\ m(t = T) &= m_T,\\ m(t = 0) &= m_0.\end{aligned}$$

With $\gamma = \alpha + H(\beta)$, condition (67) is equivalent to

$$z = m\frac{dH}{dq}(\nabla \phi),$$

$$(\widetilde{W} + \widetilde{\chi})(m) + \int_0^T \int_{\mathbb{T}^d}(W + \chi)^*(\gamma) - \langle m, \gamma\rangle = 0$$

The latter equation implies that m is a positive measure absolutely continuous w.r.t. the Lebesgue measure and that a.e.,

$$\gamma = \partial_t \phi - \Delta \phi + H(x, \nabla \phi) \in \partial W(m)$$

If $m > 0$ and m is continuous, we find (1)-(3).

Remark 3 It is worthwhile to observe that the present duality argument is not a proof of the existence of a solution of (1)-(3): we have only proved that if there exists a minimum of $\Theta + \Sigma$ such that $m > 0$ and m is continuous, then the optimality conditions yield (1)-(3).

4 Solving the systems of nonlinear equations

4.1 The Newton method for the penalized problem

4.1.1 Description of the method

In §2.2, we have seen that solutions of the systems of nonlinear equations (43)-(46) arising from a semi-implicit scheme for (1)-(3) can be found by passing to the limit in the penalized problem (20)-(23). It is therefore important to propose a solution procedure for the penalized problem (20)-(23). We propose to use the Newton method.

The system can be seen as a forward discrete HJB equation for U with a Cauchy condition at $t = 0$ coupled with a backward discrete Fokker-Planck equation for M with a Cauchy condition at final time. This structure prohibits the use of a straightforward time-marching solution procedure.

Call $\mathcal{U}$ and $\mathcal{M}$ the vectors of $\mathbb{R}^{(N^T+1)N_h^2}$ such that $\mathcal{U}_{kN_h^2+iN_h+j} = U_{i,j}^k$ and $\mathcal{M}_{kN_h^2+iN_h+j} = M_{i,j}^k$. The system of nonlinear equations can be written

$$\begin{cases} \mathcal{F}_U(\mathcal{U}, \mathcal{M}) = 0 \\ \mathcal{F}_M(\mathcal{U}, \mathcal{M}) = 0 \end{cases}$$

with

- $\mathcal{F}_U(\mathcal{U}, \mathcal{M}) = 0 \Leftrightarrow (20) \forall n, 0 \leq n < N_T, \forall i, j \quad$ and $U_{i,j}^0 = (V_{0,h}(M^0))_{i,j}, \forall i, j$

$$\mathcal{F}_M(\mathcal{U}, \mathcal{M}) = 0 \Leftrightarrow \begin{cases} \dfrac{M_{i,j}^n - M_{i,j}^{n+1}}{\Delta t} - \nu (\Delta_h M^n)_{i,j} - \mathcal{B}_{i,j}\left(U^{n+1}, M^n\right) = 0, 0 \leq n < N_T, \\ M^{N_T} \text{ given} \end{cases}$$

In order to discuss the Newton method for solving (69), we use the following notation

$$A_{U,U}(\mathcal{U}, \mathcal{M}) = D_{\mathcal{U}} \mathcal{F}_U(\mathcal{U}, \mathcal{M}), \quad A_{U,M}(\mathcal{U}, \mathcal{M}) = D_{\mathcal{M}} \mathcal{F}_U(\mathcal{U}, \mathcal{M}),$$
$$A_{M,U}(\mathcal{U}, \mathcal{M}) = D_{\mathcal{U}} \mathcal{F}_M(\mathcal{U}, \mathcal{M}), \quad A_{M,M}(\mathcal{U}, \mathcal{M}) = D_{\mathcal{M}} \mathcal{F}_M(\mathcal{U}, \mathcal{M}).$$

The matrices $A_{UU}(\mathcal{U}, \mathcal{M})$ and $A_{UM}(\mathcal{U}, \mathcal{M})$ have the form

$$A_{UU} = \begin{pmatrix} I & 0 & \cdots & \cdots & \cdots & 0 \\ -\frac{1}{\Delta t}I & D_1 & \ddots & & & \vdots \\ 0 & -\frac{1}{\Delta t}I & D_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & -\frac{1}{\Delta t}I & D_{N_T-1} & 0 \\ 0 & \cdots & \cdots & 0 & -\frac{1}{\Delta t}I & D_{N_T} \end{pmatrix} \quad A_{UM} = \begin{pmatrix} E_0 & 0 & \cdots & \cdots & 0 \\ E_1 & 0 & & & \vdots \\ 0 & E_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & E_{N_T} & 0 \end{pmatrix}.$$

The blocks of $A_{UU}(\mathcal{U}, \mathcal{M})$ are sparse. The block D_n corresponds to the discrete operator $(Z_{i,j}) \mapsto \left(\frac{1}{\Delta t}Z_{i,j} - \nu(\Delta_h Z)_{i,j} + [D_h Z$
From the regularity and monotonicity assumptions on g, D_n is a M-matrix, thus A_{UU} is invertible.

The blocks of $A_{UM}(\mathcal{U}, \mathcal{M})$ are sparse too with the assumptions used in §2, but they may be dense if V and V_0 are nonlocal operators.

From Remark 1, the matrices $A_{MM}(\mathcal{U}, \mathcal{M})$ and $A_{MU}(\mathcal{U}, \mathcal{M})$ have the form

$$A_{MM} = \begin{pmatrix} D_1^T & -\frac{1}{\Delta t}I & 0 & \cdots & 0 \\ 0 & D_2^T & -\frac{1}{\Delta t}I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & D_{N_T}^T & -\frac{1}{\Delta t}I \\ 0 & \cdots & \cdots & 0 & I \end{pmatrix} \quad A_{MU} = \begin{pmatrix} 0 & \tilde{E}_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \tilde{E}_2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \tilde{E}_{N_T} \\ 0 & \cdots & & \cdots & 0 \end{pmatrix}.$$

Note that

$$\mathcal{V}^T \tilde{E}_n \mathcal{W} = \sum_{i,j} M_{i,j}^{n-1} [D_h V]_{i,j} \cdot D_{q,q}^2 g\left(x_{i,j}, [D_h U^n]_{i,j}\right)[D_h W]_{i,j}$$

From the convexity of g, we see that the block $\tilde{E}_n$ is symmetric and positive semi-definite if M^{n-1} is a nonnegative grid functions.

In §4.1.2, we show that if V_h is strictly monotone and if the iterate produced by the Newton method satisfies $\mathcal{M} \geq 0$, then the Jacobian matrix

$$\begin{pmatrix} A_{U,U} & A_{U,M} \\ A_{M,U} & A_{M,M} \end{pmatrix}$$

is invertible. The positivity of $\mathcal{M}$ is not guaranteed though, but if the initial guess is close enough to a solution $(\widehat{U}, \widehat{\mathcal{M}})$ with $\widehat{\mathcal{M}} > 0$, then the iterates $\mathcal{M}$ will stay positive.

Assuming the invertibility of the matrix, the most time consuming part of the procedure lies in solving the system of linear equations

$$\begin{pmatrix} A_{U,U} & A_{U,M} \\ A_{M,U} & A_{M,M} \end{pmatrix} \begin{pmatrix} U \\ \mathcal{M} \end{pmatrix} = \begin{pmatrix} G_U \\ G_M \end{pmatrix}$$

The chosen algorithm is as follows:

1. solve first

$$A_{U,U} \widetilde{U} = G_U$$

This is done by observing that $\widetilde{U}^0 = G_U^0$ and sequentially solving

$$D_k \widetilde{U}^k = -L_k \widetilde{U}^{k-1} + G_{U,k}$$

i.e. marching in time in the forward direction. We know that (76) has a unique solution if g satisfies the regularity and monotonicity assumptions mentioned above. The systems (76) can be solved with efficient direct solvers (in our tests we have used the library SuperLu).

2. Introducing $\overline{U} = U - \widetilde{U}$, the vector $(\overline{U}, \mathcal{M})^T$ satisfies

$$\begin{pmatrix} A_{U,U} & A_{U,M} \\ A_{M,U} & A_{M,M} \end{pmatrix} \begin{pmatrix} \overline{U} \\ \mathcal{M} \end{pmatrix} = \begin{pmatrix} 0 \\ G_M - A_{M,U} \widetilde{U} \end{pmatrix}$$

which implies

$$\left(A_{M,M} - A_{M,U} A_{U,U}^{-1} A_{U,M} \right) \mathcal{M} = G_M - A_{M,U} \widetilde{U}$$

The system (78) is solved by means of an iterative method, for example, the BiCGstab algorithm as in the numerical tests described below; it only requires an implementation of the matrix-vector product with the matrix $A_{M,M} - A_{M,U} A_{U,U}^{-1} A_{U,M}$. Of course, this matrix is not assembled: the matrix-vector product involves matrix-vector products with the matrices $A_{M,M}$, $A_{M,U}$ and $A_{U,M}$ and solving a linear system of the form (75), similar to that appearing in the first step.

Once (78) is solved, $\overline{U}$ is obtained by solving the discrete forward linearized HJB equation

$$A_{U,U} \overline{U} = -A_{U,M} \mathcal{M}$$

by the same method as for (75).

4.1.2 Well posedness of the systems of linear equations arising in the Newton method

In the paragraph below, we prove that the systems of linear equations arising in the Newton method is well posed if the current iterate (produced by the Newton method) satisfies $\mathcal{M} \geq 0$. Consider the system

$$
\begin{pmatrix} A_{U,U} & A_{U,M} \\ A_{M,U} & A_{M,M} \end{pmatrix} \begin{pmatrix} \mathcal{V} \\ \mathcal{Q} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
$$

It is easy to see that (80) implies $Q^{N_T} = 0$. Call $\bar{A}_{UU}$ the submatrix of A_{UU}

$$
\bar{A}_{UU}(\mathcal{U}, \mathcal{M}) = \begin{pmatrix} D_1 & 0 & \cdots & \cdots & 0 \\ -\frac{1}{\Delta t} I & D_2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\frac{1}{\Delta t} I & D_{N_T} \end{pmatrix}
$$

and $\bar{A}_{MM} = \bar{A}_{UU}^T$. Introduce also the block-diagonal matrices $\bar{A}_{UM}(\mathcal{U}, \mathcal{M}) = \text{Block} - \text{Diag}\left(E_1, \ldots, E_{N_T}\right)$ and $\bar{A}_{MU}(\mathcal{U}, \mathcal{M}) = \text{Block} - \text{Diag}\left(\tilde{E}_1, \ldots, \tilde{E}_{N_T}\right)$. We obtain

$$
\bar{A}_{UU} \begin{pmatrix} V^1 \\ \vdots \\ V^{N_T} \end{pmatrix} + \bar{A}_{UM} \begin{pmatrix} Q^0 \\ \vdots \\ Q^{N_T-1} \end{pmatrix} + \frac{1}{\Delta t} \begin{pmatrix} E_0 Q^0 \\ 0 \\ \vdots \end{pmatrix} = 0,
$$

$$
\bar{A}_{MU} \begin{pmatrix} V^1 \\ \vdots \\ V^{N_T} \end{pmatrix} + \bar{A}_{UU}^T \begin{pmatrix} Q^0 \\ \vdots \\ Q^{N_T-1} \end{pmatrix} = 0.
$$

This implies that

$$
\begin{pmatrix} Q^0 \\ \vdots \\ Q^{N_T-1} \end{pmatrix}^T \bar{A}_{UM} \begin{pmatrix} Q^0 \\ \vdots \\ Q^{N_T-1} \end{pmatrix} + \frac{1}{\Delta t} \left(Q^0\right)^T E_0 Q^0 - \begin{pmatrix} V^1 \\ \vdots \\ V^{N_T} \end{pmatrix}^T \bar{A}_{MU} \begin{pmatrix} V^1 \\ \vdots \\ V^{N_T} \end{pmatrix} = 0
$$

From $(\mathbf{G_4})$, the matrix $\bar{A}_{MU}$ is symmetric and positive semi-definite if $\mathcal{M}$ has non negative coordinates, i.e. if the current iterate produced by the Newton method satisfies $m \geq 0$. Assuming that the operators V_h and $V_{0,h}$ are strictly monotone, the blocks E_n are negative definite. Therefore, if $\mathcal{M} \geq 0$, we deduce from (83) that $Q^n = 0, n = 0, \ldots N_T - 1$. Using this in the first set of equations of (82) yields that $V^n = 0, n = 1, \ldots N_T$ since the blocks D_n are invertible, and finally $V^0 = 0$ because $V^0 + E_0 Q^0 = 0$.

4.2 A solution procedure for the penalized problem

As explained above, the Newton method described above may break down if in the Newton loop, the approximation of m_h takes negative values. A similar phenomenon was observed by Benamou, Brenier and Guittet [6,7] when they studied a somewhat similar but simpler penalty method (using conjugate gradient iterations instead of Newton) for computing a mixed L^2 Wasserstein distance between two probability densities. This is of course a strong drawback of the method. However, breakdown does not happen if the initial guess is close enough to a positive solution. Therefore, it is important to find good initial guesses for the Newton method. For this reason, the penalty parameter ϵ in (20)-(23) is gradually decreased to a small value, say $1/100$, starting from a moderately

small value, say 1 . The fixed point found by the Newton method for a given value of ϵ is used as an initial guess for the next and smaller value of ϵ. Doing so, we have avoided breakdowns.

Another sensible way to avoid breakdowns would probably be to use multigrid/Newton algorithms, but this has not been tested yet.

5 References

[1] Y. Achdou and I. Capuzzo Dolcetta. Mean field games: Numerical methods. submitted, http://hal.archives-ouvertes.fr/hal-00392074.

[2] J.-P. Aubin. Applied functional analysis. Pure and Applied Mathematics (New York). Wiley-Interscience, New York, second edition, 2000. With exercises by Bernard Cornet and Jean-Michel Lasry, Translated from the French by Carole Labrousse.

[3] V. Barbu and Th. Precupanu. Convexity and optimization in Banach spaces, volume 10 of Mathematics and its Applications (East European Series). D. Reidel Publishing Co., Dordrecht, romanian edition, 1986.

[4] Martino Bardi and Italo Capuzzo-Dolcetta. Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.

[5] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the MongeKantorovich mass transfer problem. Numer. Math., 84(3):375-393, 2000.

[6] J.-D. Benamou and Y. Brenier. Mixed L^2-Wasserstein optimal mapping between prescribed density functions. J. Optim. Theory Appl., 111(2):255-271, 2001.

[7] J.-D. Benamou, Y. Brenier, and K. Guittet. The Monge-Kantorovitch mass transfer and its computational fluid mechanics formulation. Internat. J. Numer. Methods Fluids, 40(12):21-30, 2002. ICFD Conference on Numerical Methods for Fluid Dynamics (Oxford, 2001)

[8] H. Brezis. Analyse fonctionnelle. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].

[9] W. H. Fleming and H. M. Soner. Controlled Markov processes and viscosity solutions, volume 25 of Stochastic Modelling and Applied Probability. Springer, New York, second edition, 2006.

[10] D.A. Gomes, J. Mohr, and R.R. Souza. Discrete time, finite state space mean field games. Journal de Mathématiques Pures et Appliquées, 93(3):308 - 328, 2010.

[11] O. Guéant. Mean field games and applications to economics. PhD thesis, Université ParisDauphine, 2009.

[12] A. Lachapelle, J. Salomon, and G. Turinici. A monotonic algorithm for a mean field games model in economics. Technical report, CEREMADE, U. Paris Dauphine, 2009.

[13] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. C. R. Math. Acad. Sci. Paris, 343(9):619-625, 2006.

[14] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. C. R. Math. Acad. Sci. Paris, 343(10):679-684, 2006.

[15] J.-M. Lasry and P.-L. Lions. Mean field games. Jpn. J. Math., 2(1):229-260, 2007.

[16] J-M. Lasry, P-L. Lions, and O. Guéant. Application of mean field games to growth theory. Technical report, INRIA a CCSD electronic archive server based on P.A.O.L [http://hal.inria.fr/oai/oai.php] (France), 2008.

[17] P.L.Lions. Cours du Collège de France. 2009.

[18] C. Villani. Topics in optimal transportation, volume 58 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2003.