A HIGH-ORDER SCHEME FOR MEAN FIELD GAMES

ELISA CALZOLA, ELISABETTA CARLINI, AND FRANCISCO J. SILVA

Abstract. In this paper we propose a high-order numerical scheme for time-dependent mean field games systems. The scheme, which is built by combining Lagrange-Galerkin and semi-Lagrangian techniques, is consistent and stable for large time steps compared with the space steps. We provide a convergence analysis for the exactly integrated Lagrange-Galerkin scheme applied to the Fokker-Planck equation, and we propose an implementable version with inexact integration. Finally, we validate the convergence rate of the proposed scheme through the numerical approximation of two mean field games systems.

1. INTRODUCTION

This work concerns the numerical approximation of Mean Field Games (MFGs), introduced simultaneously by Lasry-Lions in [32, 33, 34] and by Huang-Caines-Malhamé in [26]. MFGs characterize Nash equilibria of stochastic differential games with an infinite number of indistinguishable players. In some specific instances, the aforementioned equilibria are described by a system of parabolic Partial Differential Equations (PDEs) consisting of a Hamilton-Jacobi-Bellman (HJB) equation, with a terminal condition, coupled with a Fokker-Planck (FP) equation with an initial condition.

The numerical approximation of MFGs has been an active area of research over the last decade (see e.g. [2, 35] and the references therein). Let us mention, for instance, the articles [1] and [14] proposing a semi-implicit finite difference scheme and a Semi-Lagrangian (SL) type scheme, respectively. The scheme studied in [14], which allows for large time steps compared to space steps, has been extended in [15] to deal with nonlinear FP equations and in [16] to approximate MFGs with non-local diffusions. On the other hand, to the best of our knowledge, only few works deal with high-order numerical schemes for MFG systems. Let us mention [43] and [36], where the authors propose finite difference based second-order accurate methods, and the recent contribution [24], where high-order space-time finite elements are used to approximate variational MFGs.

The main purpose of this article is to provide a new high-order approximation scheme, meaning an order of convergence larger than two, for a class of second-order MFG systems with constant diffusion. The scheme combines a high-order Lagrange-Galerkin (LG) discretization for the FP equation with a high-order SL discretization for the HJB equation. The main novelty of our scheme lies in the discretization of the FP equation which, inspired by [40] and [15], is constructed by using SL techniques for the time discretization (see e.g. [10, 18]) and LG techniques for the space discretization (see e.g. [40, 5]). More precisely, the stochastic characteristic curves of the FP equation are approximated with a Crank-Nicolson method (see e.g. [28, 39]), as in high-order SL schemes for parabolic equations (see [7]), and the space variable is discretized by using a LG scheme with a symmetric Lagrangian basis of odd order. This last choice is inspired by the results in [20, 21], where the equivalence between SL and LG schemes has been studied, and where symmetric odd basis have shown a good behavior in terms of stability. The resulting scheme for the FP equation is explicit, conservative, consistent, stable, allows for large time steps compared with space steps, is convergent, and high-order accurate. When coupled with a high-order SL for the HJB equation, one obtains a high-order scheme for the MFG system which, because of its forward-backward structure, is not explicit and is solved by fixed-point iterations. We numerically show high-order accuracy of the scheme by considering two MFG systems. The first one is a linear-quadratic

"Sapienza", Università di Roma, Dipartimento di Matematica Guido Castelnuovo, 00185 Rome, Italy (calzola@mat.uniroma1.it, carlini@mat.uniroma1.it).

Institut de recherche XLIM-DMI, UMR 7252 CNRS, Faculté des Sciences et Techniques, Université de Limoges, 87060 Limoges, France (francisco.silva@unilim.fr).

MFG with non-local couplings (see e.g. [4]), for which we are able to compute its analytical solution, and the second one, taken from [43], is a MFG with local couplings (see e.g. [12]) and no explicit solution.

The article is organized as follows. In Section 2, we recall the MFG system we are interested in, as well as some basic results on FP equations. Section 3 introduces a new scheme for FP equations, based on SL techniques and LG approximations, and establish its main properties. In Section 4, we present a high-order SL scheme for HJB equations and couple it with the scheme for the FP equation studied in Section 3 to derive a new scheme for the MFG system. Finally, in Section 5 we provide an implementable version of the method, derived from the use of a cubic basis and Simpson's rule in the LG approximation. The paper concludes by showing the performance of the proposed scheme in two examples: a linear-quadratic MFG with non-local couplings and admitting an explicit solution, and a MFG with local couplings and without explicit solutions. In all the numerical examples, an order of accuracy between two and three is observed. Finally, we provide in the Appendix of this work the proof of some needed technical results.

2. Preliminary results

In the following, given a function $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ and $(t,x) \in (0,T) \times \mathbb{R}^d$, the notations $\nabla u(t,x)$ and $\Delta u(t,x)$ refer to the gradient and Laplacian of u with respect to the spatial variable x. Similarly, given $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, the notation Dv and $\mathrm{div}(v)$ refer to the Jacobian matrix and the divergence of v with respect to the space variable, respectively. We also denote by $(\mathcal{P}_1(\mathbb{R}^d), \mathbf{d})$ be the metric space of Borel probability measures on $\mathbb{R}^d$ with finite first order moment, endowed with the 1-Wasserstein distance $\mathbf{d}$ (see e.g. [3, Section 7.1] for the definition of $\mathbf{d}$).

We focus on the numerical approximation of the following time-dependent second-order MFG with non-local couplings (see [33, 34]):

$$
\begin{aligned}
-\partial_t v - \tfrac{\sigma^2}{2}\Delta v + H(x, \nabla v) &= F(x, m(t)) \ \text{ in } [0,T) \times \mathbb{R}^d, \\
\text{(MFG)} \qquad \partial_t m - \tfrac{\sigma^2}{2}\Delta m - \mathrm{div}\big(\partial_p H(x, \nabla v)m\big) &= 0 \ \text{ in } (0,T] \times \mathbb{R}^d, \\
v(T,\cdot) = G(\cdot, m(T)), \qquad m(0,\cdot) &= m_0^* \ \text{ in } \mathbb{R}^d,
\end{aligned}
$$

where $T > 0$, $\sigma \in \mathbb{R} \setminus \{0\}$, $\mathbb{R}^d \times \mathbb{R}^d \ni (x,p) \mapsto H(x,p) \in \mathbb{R}$ is convex and differentiable with respect to p, F, $G : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}$, and $m_0^* : \mathbb{R}^d \to \mathbb{R}$. Notice that (MFG) consists of a HJB equation, with a terminal condition, coupled with a FP equation with an initial condition. For the sake of simplicity, in what follows we will suppose that the *Hamiltonian* H is purely quadratic, i.e. $H(x,p) = |p|^2/2$ for all $x, p \in \mathbb{R}^d$ and we assume that:

(H1) m_0^* is nonnegative, Hölder continuous, has compact support, and $\int_{\mathbb{R}^d} m_0^*(x)\mathrm{d}x = 1$. .

(H2) F and G are bounded and Lipschitz continuous. Moreover, for every $\mu \in \mathcal{P}_1(\mathbb{R}^d)$, $F(\cdot, \mu)$ is of class C^2 and

$$
\sup_{x \in \mathbb{R}^d, \mu \in \mathcal{P}_1(\mathbb{R}^d)} \big\{ \|DF(x,\mu)\|_\infty + \|D^2 F(x,\mu)\|_\infty \big\} < \infty.
$$

Under **(H1)-(H2)** system (MFG) admits at least one classical solution (v^*, m^*) (see e.g. [11, Theorem 3.1]). Moreover, if the coupling terms F and G satisfy a monotonicity condition with respect to m, then the classical solution is unique (see [34, Theorem 2.4]).

In order to obtain a high-order scheme for (MFG), our first task will be to construct a high-order LG scheme for the following linear FP equation:

$$
\begin{aligned}
\text{(FP)} \qquad \partial_t m - \tfrac{\sigma^2}{2}\Delta m + \mathrm{div}\,(bm) &= 0 \ \text{ in } (0,T) \times \mathbb{R}^d, \\
m(0,\cdot) &= m_0^* \ \text{ in } \mathbb{R}^d,
\end{aligned}
$$

where $\sigma \in \mathbb{R} \setminus \{0\}$, $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, and $m_0^* : \mathbb{R}^d \to \mathbb{R}$. We will assume that:

(H3) $b \in C([0,T] \times \mathbb{R}^d)$, b is bounded and there exists $C_b > 0$ such that

$$
|b(t,x) - b(t,y)| \le C_b |x-y|, \quad \text{for } t \in [0,T] \text{ and } x, y \in \mathbb{R}^d.
$$

In the following result, proved in the Appendix, we summarize some properties of equation (FP).

Theorem 2.1. *Assume* **(H1)** *and* **(H3)**. *Then the following hold:*

(i) *Equation* (FP) *admits a unique classical solution* $m^* \in C^{1,2}([0,T] \times \mathbb{R}^d)$.

(ii) $m^* \geq 0$.

(iii) $\int_{\mathbb{R}^d} m^*(t,x)\mathrm{d}x = 1$ *for all* $t \in [0,T]$.

(iv) m^* *is the unique solution in* $L^2([0,T] \times \mathbb{R}^d)$ *to* (**FP**) *in the distributional sense.*

Let us recall the probabilistic interpretation of the solution m^* to (**FP**), which will be useful in order to construct a LG scheme. Let W be a d-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $Y_0 : \Omega \to \mathbb{R}^d$ be a random variable, independent of W, and whose distribution is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$, with density given by m_0^*. Given $(t,x) \in [0,T] \times \mathbb{R}^d$, we define $Y^{t,x}$ as the unique strong solution to the SDE:

$$(2.1) \qquad \begin{aligned} \mathrm{d}Y(s) &= b(s, Y(s))\mathrm{d}s + \sigma \mathrm{d}W(s) \quad \text{for } s \in (t,T), \\ Y(t) &= x. \end{aligned}$$

Denote by $\mathbb{E}(X)$ the expectation of a random variable $X : \Omega \to \mathbb{R}$. Under the asspumptions of Theorem 2.1, $Y^{0,Y_0}(t)$ is well defined for all $t \in [0,T]$ and its distribution is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$, with density given by $m^*(t,\cdot)$ (see e.g. [22]). From the $\mathbb{P}$-a.s. equality $Y^{0,Y_0}(s) = Y^{t,Y^{0,Y_0}(t)}(s)$ for every $0 \leq t \leq s \leq T$, we deduce that for every continuous and bounded function $\phi : \mathbb{R}^d \to \mathbb{R}$, we have

$$(2.2) \qquad \int_{\mathbb{R}^d} \phi(x)m^*(s,x)\mathrm{d}x = \int_{\mathbb{R}^d} \mathbb{E}\left(\phi(Y^{t,x}(s))\right)m^*(t,x)\mathrm{d}x.$$

3. A Lagrange-Galerkin type scheme for a Fokker-Planck equation

Let us focus on the numerical approximation of (**FP**). Notice that if b is differentiable with respect to the space variable, (**FP**) can be written as

$$\partial_t m - \frac{\sigma^2}{2}\Delta m + \langle b, \nabla m\rangle + \mathrm{div}(b)m = 0 \qquad \text{in } (0,T) \times \mathbb{R}^d,$$
$$m(0,\cdot) = m_0^* \qquad\qquad \text{in } \mathbb{R}^d.$$

Using this formulation, a second-order accurate semi-Lagrangian scheme can be derived to approximate m^* (see e.g. [7]). However, such a scheme is not conservative, i.e. the discrete solution does not satisfy the discrete analogous of Theorem 2.1(iii). The scheme that we consider, which will be built from (2.2), will allow us to preserve this property (see Theorem 3.1(ii) below).

Let us fix $N_{\Delta t} \in \mathbb{N}$, set $\mathcal{I}_{\Delta t} = \{0, \ldots, N_{\Delta t}\}$, $\mathcal{I}_{\Delta t}^* = \mathcal{I}_{\Delta t} \setminus \{N_{\Delta t}\}$, $\Delta t = T/N_{\Delta t}$, and $t_k = k\Delta t$ ($k \in \mathcal{I}_{\Delta t}$). Let $x \in \mathbb{R}^d$ and consider the sequence of random variables $(y_k)_{k=0}^{N_{\Delta t}}$ defined by $y_0 = x$ and, for every $k \in \mathcal{I}_{\Delta t}^*$, y_{k+1} is the unique solution to

$$(3.1) \qquad y = y_k + \frac{\Delta t}{2}\left(b(t_k, y_k) + b(t_{k+1}, y)\right) + \sqrt{\Delta t}\sigma\xi_k,$$

where $(\xi_k)_{k=0}^{N_{\Delta t}-1}$ is a sequence of i.i.d. $\mathbb{R}^d$-valued random variables with i.i.d. components such that, for every $k \in \mathcal{I}_{\Delta t}^*$,

$$(3.2) \qquad \mathbb{P}((\xi_k)_i = 0) = 2/3 \quad \text{and} \quad \mathbb{P}((\xi_k)_i = \pm\sqrt{3}) = 1/6 \quad \text{for all} \quad i = 1, \ldots, d.$$

Since b is Lipschitz continuous, the sequence $(y_k)_{k=0}^{N_{\Delta t}}$, called the Crank-Nicolson (CN) approximation of $Y^{0,x}$, is well-defined for Δt sufficiently small.

An interesting feature of the law of $(\xi_k)_i$ in (3.2) is that, provided that b is smooth enough, $(y_k)_{k=0}^{N_{\Delta t}}$ is a second order weak approximation of $Y^{0,x}$ (see e.g. [28, Section 15.4, equation (4.11)] and also [19, Section 2, Table 1]), i.e. for every $\phi : \mathbb{R}^d \to \mathbb{R}$ smooth enough and for every $k \in \mathcal{I}_{\Delta t}^*$, we have

$$(3.3) \qquad \left|\mathbb{E}\left(\phi(y_k)\right) - \mathbb{E}\left(\phi(Y^{0,x}(t_k))\right)\right| = O\left((\Delta t)^2\right).$$

Notice that this estimate is better than the one obtained by considering a classical random walk in $\mathbb{R}^d$, i.e. when distribution of $(\xi_k)_i$ is given by

$$(3.4) \qquad \mathbb{P}((\xi_k)_i = \pm 1) = 1/2 \quad \text{for all} \quad i = 1, \ldots, d,$$

for which it is known that second order accuracy does not hold (see Sections 5.1.A and 5.1.B in [29]).

In order to discretize (2.2), for every $k \in \mathcal{I}_{\Delta t}^*$ and $x \in \mathbb{R}^d$, denote by $y^{t_k,x}$ the one-step CN approximation of $Y^{t_k,x}(t_{k+1})$, given by the unique solution to (3.1). Let $\mathcal{I}_d = \{1, \ldots, 3^d\}$, define $\{e^\ell \,|\, \ell \in \mathcal{I}_d\} \subset \mathbb{R}^d$

as the set of possible values of ξ_k, set $\omega^\ell = \mathbb{P}(\xi_k = e^\ell)$, and denote by $y_k^\ell(x)$ the unique solution to (3.1) for $\xi_k = e^\ell$ ($\ell \in \mathcal{I}_d$). By setting $t = t_k$, $s = t_{k+1}$, and replacing $Y^{t_k, x}(t_{k+1})$ by $y^{t_k, x}$ in (2.2), we obtain the following semi-discrete scheme for (**FP**):

$$(3.5) \qquad \int_{\mathbb{R}^d} \phi(x) m_{k+1}(x) \mathrm{d}x = \sum_{\ell \in \mathcal{I}_d} \omega_\ell \int_{\mathbb{R}^d} \phi(y_k^\ell(x)) m_k(x) \mathrm{d}x \quad \text{for } \phi \text{ continuous and bounded, } k \in \mathcal{I}_{\Delta t},$$

with $m_0 = m_0^*$ and unknowns $\{m_k : \mathbb{R}^d \to \mathbb{R} \,|\, k \in \mathcal{I}_{\Delta t} \setminus \{0\}\}$. Note that the assumption that m_0 has a compact support implies the existence of $L_{\Delta t} = O(1/\sqrt{\Delta t})$ such that the solution $m_{\Delta t}$ to (3.5) satisfies

$$(3.6) \qquad \mathrm{supp}(m_{\Delta t, k}) \subset [-L_{\Delta t}, L_{\Delta t}]^d \quad \text{for } k \in \mathcal{I}_{\Delta t}.$$

In order to construct a space discretization of (3.5), and hence a fully-discrete scheme for (**FP**), we consider a symmetric Lagrangian basis of odd order. More precisely, let us fix $p \in \mathbb{N}$, set $q := 2p+1$, and let $\widehat{\beta} : \mathbb{R} \to \mathbb{R}$ be defined by

$$(3.7) \qquad (\forall \xi \in [0, \infty)) \quad \widehat{\beta}(\xi) = \begin{cases} \displaystyle\prod_{k \neq 0,\, k=-p}^{p+1} \frac{\xi - k}{-k} & \text{if } \xi \in [0, 1], \\[2ex] \displaystyle\prod_{k \neq 0,\, k=-p+1}^{p+2} \frac{\xi - k}{-k} & \text{if } \xi \in (1, 2], \\ \vdots & \\ \displaystyle\prod_{k=1}^{2p+1} \frac{\xi - k}{-k} & \text{if } \xi \in (p, p+1], \\[2ex] 0 & \text{if } \xi \in (p+1, \infty), \\[1ex] \widehat{\beta}(-\xi) & \text{if } \xi \in (-\infty, 0). \end{cases}$$

Following [20], for $\Delta x \in (0, \infty)$, we consider the symmetric Lagrange interpolation basis functions $\{\beta_i\}_{i \in \mathbb{Z}^d}$ defined as

$$(\forall z = (z_1, \dots z_d) \in \mathbb{R}^d, \ i = (i_1, \dots, i_d) \in \mathbb{Z}^d) \quad \beta_i(z) = \prod_{j=1}^{d} \widehat{\beta}\left(\frac{z_j}{\Delta x} - i_j \right).$$

For all $i \in \mathbb{Z}^d$, let us set $x_i = i\Delta x$. Notice that β_i has compact support, $\beta_i(x_j) = 1$ if $i = j$ and $\beta_i(x_j) = 0$ otherwise, and, for all $x \in \mathbb{R}^d$, $\sum_{j \in \mathbb{Z}^d} \beta_j(x) = 1$. Given $f \in W^{q+1, \infty}(\mathbb{R}^d)$, we define the interpolant $I[f] : \mathbb{R}^d \to \mathbb{R}$ by

$$(3.8) \qquad I[f](x) = \sum_{i \in \mathbb{Z}^d} f(x_i) \beta_i(x) \quad \text{for } x \in \mathbb{R}^d.$$

By [17, Theorem 16.1], the following estimate holds

$$(3.9) \qquad \sup_{x \in \mathbb{R}^d} |f(x) - I[f](x)| \leq C_I (\Delta x)^{q+1} \|D^{q+1} f\|_{L^\infty},$$

where $C_I > 0$ is independent of f and Δx. Notice that in the one dimensional case ($d = 1$), $I[f]$ restricted to a given interval (x_i, x_{i+1}) ($i \in \mathbb{Z}$) is the Lagrange interpolating polynomial of degree q constructed on the symmetric stencil $x_{i-(q-1)/2}, \dots, x_{i+1+(q-1)/2}$.

Let $L_{\Delta t} > 0$ be as in (3.6), let $N_{\Delta x} \in \mathbb{N}$, and set $\mathcal{I}_{\Delta x} = \{-N_{\Delta x}, \dots, N_{\Delta x}\}^d$. From now on, we assume that $\Delta x = L_{\Delta t}/N_{\Delta x}$, we set

$$\Delta = (\Delta t, \Delta x), \quad \text{and} \quad \mathcal{O}_\Delta = [-L_{\Delta t} - p\Delta x, L_{\Delta t} + p\Delta x]^d.$$

We look for an approximation m_Δ of the solution m^* to (**FP**) such that, for all $k \in \mathcal{I}_{\Delta t}$,

$$(3.10) \qquad m_\Delta(t_k, x) = \sum_{i \in \mathcal{I}_{\Delta x}} m_{k,i} \beta_i(x) \text{ for } x \in \mathcal{O}_\Delta, \quad m_\Delta(t_k, x) = 0 \text{ for } x \in \mathbb{R}^d \setminus \mathcal{O}_\Delta,$$

where $m_{k,i} \in \mathbb{R}$ ($k \in \mathcal{I}_{\Delta t}$, $i \in \mathcal{I}_{\Delta x}$) have to be determined. Notice that, by definition of $\mathcal{I}_{\Delta x}$, for all $k \in \mathcal{I}_{\Delta t}$ we have that $\operatorname{supp}\{m_\Delta(t_k,\cdot)\} \subset \mathcal{O}_\Delta$. Replacing m by m_Δ and taking $\phi = \beta_i$ ($i \in \mathcal{I}_{\Delta x}$) in (3.5) yields the following explicit iterative scheme for the unknowns $m_{k,i} \in \mathbb{R}$ ($k \in \mathcal{I}_{\Delta t}$, $i \in \mathcal{I}_{\Delta x}$)

$$(3.11) \qquad \begin{aligned} \sum_{j \in \mathcal{I}_{\Delta x}} m_{k+1,j} \int_{\mathcal{O}_\Delta} \beta_i(x)\beta_j(x)\mathrm{d}x &= \sum_{j \in \mathcal{I}_{\Delta x}} m_{k,j} \sum_{\ell \in \mathcal{I}_d} \omega_\ell \int_{\mathcal{O}_\Delta} \beta_i(y_k^\ell(x))\beta_j(x)\mathrm{d}x \\ &\qquad\qquad\qquad\qquad \text{for } k \in \mathcal{I}_{\Delta t}^*,\ i \in \mathcal{I}_{\Delta x}, \\ \sum_{j \in \mathcal{I}_{\Delta x}} m_{0,j} \int_{\mathcal{O}_\Delta} \beta_i(x)\beta_j(x)\mathrm{d}x &= \int_{\mathcal{O}_\Delta} m_0^*(x)\beta_i(x)\mathrm{d}x. \end{aligned}$$

Let A be the $(2N_{\Delta x}+1)^d \times (2N_{\Delta x}+1)^d$ real mass matrix with entries given by

$$(3.12) \qquad A_{i,j} = \int_{\mathcal{O}_\Delta} \beta_i(x)\beta_j(x)\mathrm{d}x \quad \text{for} \quad (i,j) \in \mathcal{I}_{\Delta x} \times \mathcal{I}_{\Delta x}.$$

For $k \in \mathcal{I}_{\Delta t}^*$ and $\ell \in \mathcal{I}_d$, let B_k^ℓ be the $(2N_{\Delta x}+1)^d \times (2N_{\Delta x}+1)^d$ real matrix with entries given by

$$(3.13) \qquad (B_k^\ell)_{i,j} = \int_{\mathcal{O}_\Delta} \beta_i(y_k^\ell(x))\beta_j(x)\mathrm{d}x \quad \text{for} \quad (i,j) \in \mathcal{I}_{\Delta x} \times \mathcal{I}_{\Delta x}.$$

Let $m_{0,\Delta x}$ be the $(2N_{\Delta x}+1)^d$ dimensional real vector with entries

$$(m_{0,\Delta x})_i = \int_{\mathcal{O}_\Delta} m_0^*(x)\beta_i(x)\mathrm{d}x \quad \text{for } i \in \mathcal{I}_{\Delta x}.$$

Calling $m_k = (m_{k,i})_{i \in \mathcal{I}_{\Delta x}}$, scheme (3.11) can be rewritten in the following matrix form: find m_k ($k \in \mathcal{I}_{\Delta t}$) such that

$$(3.14) \qquad \begin{aligned} Am_{k+1} &= \sum_{\ell \in \mathcal{I}_d} \omega_\ell B_k^\ell m_k \quad \text{for } k \in \mathcal{I}_{\Delta t}^*, \\ Am_0 &= m_{0,\Delta x}. \end{aligned}$$

3.1. **Properties of the space-time Lagrange-Galerkin scheme.** We show below some important properties of the scheme (3.11).

Theorem 3.1. *Assume* **(H1)**,**(H3)**. *Then for fixed Δ, there exists a unique solution* $(m_{k,i})_{k \in \mathcal{I}_{\Delta t}, i \in \mathcal{I}_{\Delta x}}$ *to (3.14) and, defining m_Δ as in (3.10), the following hold:*

(i)[Initial condition] $\|m_0^* - m_\Delta(0,\cdot)\|_{L^2} = O((\Delta x)^{q+1})$ *if* $m_0^* \in H^{q+1}(\mathbb{R}^d)$.

(ii)[Mass conservation] $\int_{\mathbb{R}^d} m_\Delta(t_k,x)\mathrm{d}x = 1$ *for* $k \in \mathcal{I}_{\Delta t}$.

(iii)[L^2-stability] *If $b(t,\cdot)$ is differentiable for all $t \in [0,T]$, then* $\max_{k \in \mathcal{I}_{\Delta t}} \|m_\Delta(t_k,\cdot)\|_{L^2}$ *is uniformly bounded with respect to Δ for Δt small enough.*

Proof. The well-posedness of (3.14) follows from the positive definiteness of A (see e.g. [44, Proposition 6.3.1]) and assertion (i) is proven in [44, Section 3.5]. In order to prove (ii), fix $k \in \mathcal{I}_{\Delta t}^*$ and sum over $i \in \mathbb{Z}^d$ in the first equation of (3.11) to obtain

$$\sum_{j \in \mathcal{I}_{\Delta x}} m_{k+1,j} \sum_{i \in \mathbb{Z}^d} \int_{\mathcal{O}_\Delta} \beta_j(x)\beta_i(x)\mathrm{d}x = \sum_{j \in \mathcal{I}_{\Delta x}} m_{k,j} \sum_{\ell \in \mathcal{I}_d} \omega_\ell \sum_{i \in \mathbb{Z}^d} \int_{\mathcal{O}_\Delta} \beta_j(x)\beta_i(y_k^\ell(x))\mathrm{d}x.$$

Recalling that, for every $y \in \mathbb{R}^d$, $\sum_{i \in \mathbb{Z}^d} \beta_i(y) = 1$, the cardinality $\{i \in \mathbb{Z}^d \,|\, \beta_i(y) \neq 0\}$ is bounded uniformly in y, and $\sum_{\ell \in \mathcal{I}_d} \omega_\ell = 1$, Fubini's theorem yields

$$(3.15) \qquad \begin{aligned} \int_{\mathcal{O}_\Delta} m_\Delta(t_{k+1},x)\mathrm{d}x &= \sum_{j \in \mathcal{I}_{\Delta x}} m_{k+1,j} \int_{\mathcal{O}_\Delta} \beta_j(x)\mathrm{d}x \\ &= \sum_{j \in \mathcal{I}_{\Delta x}} m_{k,j} \int_{\mathcal{O}_\Delta} \beta_j(x)\mathrm{d}x \\ &= \sum_{j \in \mathcal{I}_{\Delta x}} m_{0,j} \int_{\mathcal{O}_\Delta} \beta_j(x)\mathrm{d}x \\ &= \int_{\mathcal{O}_\Delta} m_\Delta(0,x)\mathrm{d}x. \end{aligned}$$

Analogously, using the second equation in (3.11) and summing over $i \in \mathbb{Z}^d$, we get that

$$(3.16) \qquad \int_{\mathcal{O}_\Delta} m_\Delta(0,x)\mathrm{d}x = \int_{\mathcal{O}_\Delta} m_0^*(x)\mathrm{d}x = 1.$$

Assertion (ii) follows from (3.15), (3.16), and (3.10). Finally, let us show assertion (iii). For $k = 0$, (iii) follows from Assumption **(H1)** and Theorem 3.1(i). For $k \in \mathcal{I}_{\Delta t}^*$, (3.11) implies that

$$(3.17) \qquad \begin{aligned} \|m_\Delta(t_{k+1},\cdot)\|_{L^2}^2 &= \sum_{\ell \in \mathcal{I}_d} \omega_\ell \sum_{i,j \in \mathcal{I}_{\Delta x}} m_{k+1,i} m_{k,j} \int_{\mathcal{O}_\Delta} \beta_i(x)\beta_j(y_k^\ell(x))\mathrm{d}x \\ &= \sum_{\ell \in \mathcal{I}_d} \omega_l \int_{\mathcal{O}_\Delta} m_\Delta(t_k, y_k^\ell(x)) m_\Delta(t_{k+1}, x)\mathrm{d}x, \end{aligned}$$

and hence, by the Cauchy-Schwarz inequality,

$$(3.18) \qquad \|m_\Delta(t_{k+1},\cdot)\|_{L^2} \leq \max_{\ell \in \mathcal{I}_d} \left(\int_{\mathcal{O}_\Delta} |m_\Delta(t_k, y_k^\ell(x))|^2 \mathrm{d}x \right)^{1/2}.$$

In order to estimate the right-hand-side above, fix $x \in \mathbb{R}^d$, $\ell \in \mathcal{I}_d$, and notice that

$$(3.19) \qquad Dy_k^\ell(x) = I_d + \frac{\Delta t}{2}\left(Db(t_k, x) + Db(t_{k+1}, y_k^\ell(x))Dy_k^\ell(x) \right),$$

where I_d denotes the $d \times d$ identity matrix. Since $Db(\cdot,\cdot)$ is bounded, there exists $\overline{\Delta t} > 0$ such that for all $k \in \mathcal{I}_{\Delta t}^*$ and $\Delta t \in [0, \overline{\Delta t}]$, y_k^ℓ is one-to-one, and, for all $z \in \mathbb{R}^d$, the matrix $I_d - \frac{\Delta t}{2} Db(t_{k+1}, z)$ is invertible. Therefore, by (3.19),

$$(3.20) \qquad Dy_k^\ell(x) = \left(I_d - \frac{\Delta t}{2} Db(t_{k+1}, y_k^\ell(x)) \right)^{-1} \left(I_d + \frac{\Delta t}{2} Db(t_k, x) \right),$$

from which we deduce that $Dy_k^\ell(x)$ is invertible. Then, by the change of variable formula, we get that

$$(3.21) \qquad \int_{\mathcal{O}_\Delta} |m_\Delta(t_k, y_k^\ell(x))|^2 \mathrm{d}x = \int_{y_k^\ell(\mathcal{O}_\Delta)} |m_\Delta(t_k, z)|^2 \big|\det\left(Dy_k^\ell((y_k^\ell)^{-1}(z)) \right)\big|^{-1} \mathrm{d}z.$$

On the other hand, by (3.20) and Jacobi's formula, for all $x \in \mathbb{R}^d$ we have

$$(3.22) \qquad \begin{aligned} \left[\det\left(Dy_k^\ell(x)\right)\right]^{-1} &= \frac{\det\left(I_d - \frac{\Delta t}{2} Db(t_{k+1}, y_k^\ell(x))\right)}{\det\left(I_d + \frac{\Delta t}{2} Db(t_k, x)\right)} \\ &= \frac{1 - \frac{\Delta t}{2}\mathrm{Tr}\left(Db(t_{k+1}, y_k^\ell(x))\right) + O((\Delta t)^2)}{1 + \frac{\Delta t}{2}\mathrm{Tr}\left(Db(t_k, x) + O((\Delta t)^2)\right)} \\ &= \frac{1 - \frac{\Delta t}{2}\mathrm{div}\left(b(t_{k+1}, y_k^\ell(x))\right) + O((\Delta t)^2)}{1 + \frac{\Delta t}{2}\mathrm{div}\left(b(t_k, x)\right) + O((\Delta t)^2)}. \end{aligned}$$

Thus, there exists a constant $C > 0$, independent of x, k, ℓ, and Δt, such that

$$(3.23) \qquad \big| \left[\det\left(Dy_k^\ell(x)\right)\right]^{-1} \big| \leq 1 + C\Delta t.$$

Combining the previous inequality with (3.21) yields

$$(3.24) \qquad \int_{\mathcal{O}_\Delta} |m_\Delta(t_k, y_k^\ell(x))|^2 \mathrm{d}x \leq (1 + C\Delta t)\|m_\Delta(t_k,\cdot)\|_{L^2}^2,$$

and hence, by (3.18),

$$\|m_\Delta(t_{k+1},\cdot)\|_{L^2} \leq (1 + C\Delta t)^{\frac{1}{2}}\|m_\Delta(t_k,\cdot)\|_{L^2}^2.$$

Thus,

$$\|m_\Delta(t_{k+1},\cdot)\|_{L^2} \leq \left(1 + \frac{CT}{N_{\Delta t}}\right)^{N_{\Delta t}/2} \|m_\Delta(0,\cdot)\|_{L^2} \leq e^{CT/2}\|m_\Delta(0,\cdot)\|_{L^2},$$

from which assertion (iii) follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 3.1. *Notice that Proposition 3.1(iii) and the Cauchy-Schwarz inequality imply that, for any compact set $K \subseteq \mathbb{R}^d$, there exists $C_K > 0$, independent of Δ for Δt small enough, such that*

$$\max_{k \in \mathcal{I}_{\Delta t}} \int_K |m_\Delta(t_k, x)| \mathrm{d}x \leq C_K.$$

In the following, we still denote by m_Δ its extension to $[0, T] \times \mathbb{R}^d$, defined as

$$(3.25) \qquad m_\Delta(t, x) = \frac{t - t_k}{\Delta t} m_\Delta(t_{k+1}, x) + \frac{t_{k+1} - t}{\Delta t} m_\Delta(t_k, x) \quad \text{if } (t, x) \in [t_k, t_{k+1}] \times \mathbb{R}^d \ (k \in \mathcal{I}_{\Delta t}^*).$$

Notice that (3.25) and Theorem 3.1(ii)-(iii) imply that

$$(3.26) \qquad \int_{\mathcal{O}_\Delta} m_\Delta(t, x)\mathrm{d}x = 1 \quad \text{for all } t \in [0, T] \quad \text{and} \quad \max_{t \in [0,T]} \|m_\Delta(t, \cdot)\|_{L^2} \leq C,$$

for some $C > 0$, independent of Δ for Δt small enough.

For $k \in \mathbb{N} \cup \{\infty\}$, we denote by $C_0^k(\mathbb{R}^d)$ the set of functions of class C^k with compact support.

Proposition 3.1. *Under **(H1)**-**(H3)**, the following hold:*

(i)[Equicontinuity] *Let $\phi \in C_0^{q+1}(\mathbb{R}^d)$. Then there exists $C_\phi > 0$ such that for all Δ, with Δt small enough and $(\Delta x)^{q+1} \leq \Delta t$, we have*

$$(3.27) \qquad \left| \int_{\mathbb{R}^d} \phi(x)m_\Delta(t, x)\mathrm{d}x - \int_{\mathbb{R}^d} \phi(x)m_\Delta(s, x)\mathrm{d}x \right| \leq C_\phi |t - s| \quad \text{for all } s, t \in [0, T].$$

(ii)[Consistency] *Assume that $b(t, \cdot) \in C^{q+1}(\mathbb{R}^d)$ for all $t \in [0, T]$ and let $\phi \in C_0^\infty(\mathbb{R}^d)$. Then for any $k \in \mathcal{I}_{\Delta t}^*$ and Δ, with Δt small enough and $(\Delta x)^{q+1} \leq \Delta t$, we have*
(3.28)

$$\int_{\mathbb{R}^d} \phi(x) \left(m_\Delta(t_{k+1}, x) - m_\Delta(t_k, x) \right) \mathrm{d}x = \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}^d} \left(\frac{\sigma^2}{2} \Delta\phi(x) + \langle b(s, x), \nabla\phi(x) \rangle \right) m_\Delta(s, x)\mathrm{d}x\mathrm{d}s$$

$$+ O\left((\Delta x)^{q+1} + (\Delta t)^2 + \Delta t\omega_\phi(\Delta t) \right),$$

where $\omega_\phi : [0, \infty) \to \mathbb{R}$ is a modulus of continuity of b on $[0, T] \times \mathrm{supp}(\phi)$.

The proof is given in the Appendix. Let us denote by $\mathcal{D}'(\mathbb{R}^d)$ the space of distributions, which we endow with the weak* topology. In the following, for every $\Delta = (\Delta t, \Delta x) \in (0, \infty)^2$ and $t \in [0, T]$, we identify $m_\Delta(t, \cdot)$ with the map

$$C_0^\infty(\mathbb{R}^d) \ni \phi \mapsto \int_{\mathbb{R}^d} \phi(x)m_\Delta(t, x)\mathrm{d}x \in \mathbb{R},$$

which, by Remark 3.1, is a regular distribution. For every Δ, let us denote, with a slight abuse of notation, m_Δ the map $[0, T] \ni t \mapsto m_\Delta(t, \cdot) \in \mathcal{D}'(\mathbb{R}^d)$. Notice that Proposition 3.1(i) implies that $m_\Delta \in C([0, T]; \mathcal{D}'(\mathbb{R}^d))$.

Lemma 3.1. *Suppose that **(H1)**(i),**(H2)** hold. Then there exists $\Delta t_0 > 0$ such that the family $\mathcal{M} = \{m_\Delta \mid \Delta t \leq \Delta t_0, \ (\Delta x)^{q+1} \leq \Delta t\}$ is relatively compact in $C([0, T]; \mathcal{D}'(\mathbb{R}^d))$.*

The proof is given in the Appendix. We conclude the section with the following convergence result for scheme (3.11).

Theorem 3.2. *Assume **(H1)** and **(H3)**, $m_0^* \in H^{q+1}(\mathbb{R}^d)$, b bounded and $b(t, \cdot) \in C^{q+1}(\mathbb{R}^d)$ for all $t \in [0, T]$. Consider a sequence $(\Delta_n)_{n \in \mathbb{N}} = ((\Delta t_n, \Delta x_n))_{n \in \mathbb{N}} \subseteq (0, \infty)^2$ such that, as $n \to \infty$, $\Delta_n \to (0, 0)$ and $(\Delta x_n)^{q+1}/\Delta t_n \to 0$. Setting $m^n := m_{\Delta_n}$, as $n \to \infty$ we have that $(m^n)_{n \in \mathbb{N}}$ converges to m^* in $C([0, T]; \mathcal{D}'(\mathbb{R}^d))$ and weakly in $L^2([0, T] \times \mathbb{R}^d)$, where m^* is the unique classical solution to **(FP)**.*

Proof. By Theorem 3.1(iii), the sequence $(m^n)_{n \in \mathbb{N}}$ is bounded in $L^2([0, T] \times \mathbb{R}^d)$. Thus, there exists $\widehat{m}$ in $L^2([0, T] \times \mathbb{R}^d)$ such that, as $n \to \infty$ and up to some subsequence, m^n converges weakly to $\widehat{m}$ in $L^2([0, T] \times \mathbb{R}^d)$.

Let us first show that for any $\phi \in C_0^\infty((0, T) \times \mathbb{R}^d)$, we have

$$(3.29) \qquad \int_0^T \int_{\mathbb{R}^d} \left[\partial_t \phi(t, x) - \frac{\sigma^2}{2} \Delta\phi(t, x) - \langle b(s, x), \nabla\phi(t, x) \rangle \right] \widehat{m}(t, x)\mathrm{d}x\mathrm{d}t = 0.$$

Let $\eta \in C_0^\infty([0,T])$, $\psi \in C_0^\infty(\mathbb{R}^d)$ and suppose that ϕ has the form $\phi = \eta\psi \in C_0^\infty([0,T] \times \mathbb{R}^d)$. Denote by $K \subset \mathbb{R}^d$ the support of ψ. By (3.25) and Proposition 3.1(i), we have

(3.30)
$$\int_0^T \int_{\mathbb{R}^d} \partial_t \phi(t,x) m^n(t,x) \mathrm{d}x \mathrm{d}t = \sum_{k=0}^{N_{\Delta t_n}-1} \int_{t_k}^{t_{k+1}} \int_K \partial_t \phi(t,x) m^n(t_k,x) \mathrm{d}x \mathrm{d}t$$
$$+ \sum_{k=0}^{N_{\Delta t_n}-1} \int_{t_k}^{t_{k+1}} \int_K \partial_t \phi(t,x)(m^n(t_{k+1},x) - m^n(t_k,x)) \frac{t-t_k}{\Delta t_n} \mathrm{d}x \mathrm{d}t$$
$$= \sum_{k=0}^{N_{\Delta t_n}-1} \int_{t_k}^{t_{k+1}} \int_K \partial_t \phi(t,x) m^n(t_k,x) \mathrm{d}x \mathrm{d}t + O(\Delta t_n).$$

On the other hand, by Remark 3.1 we have

(3.31)
$$\sum_{k=0}^{N_{\Delta t_n}-1} \int_{t_k}^{t_{k+1}} \int_K \partial_t \phi(t,x) m^n(t_k,x) \mathrm{d}x \mathrm{d}t = \sum_{k=0}^{N_{\Delta t_n}-1} \Delta t_n \int_K \partial_t \phi(t_k,x) m^n(t_k,x) \mathrm{d}x + O\left(\Delta t_n\right)$$
$$= \sum_{k=0}^{N_{\Delta t_n}-1} \Delta t_n \dot\eta(t_k) \int_K \psi(x) m^n(t_k,x) \mathrm{d}x + O(\Delta t_n)$$
$$= \sum_{k=0}^{N_{\Delta t_n}-1} (\eta(t_{k+1}) - \eta(t_k)) \int_K \psi(x) m^n(t_k,x) \mathrm{d}x + O(\Delta t_n)$$
$$= \sum_{k=0}^{N_{\Delta t_n}-2} \eta(t_{k+1}) \left(\int_K \psi(x)[m^n(t_k,x) - m^n(t_{k+1},x)] \mathrm{d}x \right)$$
$$+ O(\Delta t_n).$$

By (3.30), (3.31) and using that ψ vanishes outside K, we get

(3.32)
$$\int_0^T \int_{\mathbb{R}^d} \partial_t \phi(t,x) m^n(t,x) \mathrm{d}x \mathrm{d}t = \sum_{k=0}^{N_{\Delta t_n}-2} \eta(t_{k+1}) \left(\int_{\mathbb{R}^d} \psi(x)[m^n(t_k,x) - m^n(t_{k+1},x)] \mathrm{d}x \right)$$
$$+ O(\Delta t_n).$$

Using (3.32), Proposition 3.1(ii), and Remark 3.1, we have

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \phi(t,x) m^n(t,x) \mathrm{d}x \mathrm{d}t = \sum_{k=0}^{N_{\Delta t_n}-1} \eta(t_{k+1}) \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}^d} \left(\frac{\sigma^2}{2} \Delta\psi(x) + \langle b(s,x), \nabla\psi(x) \rangle \right) m^n(s,x) \mathrm{d}x \mathrm{d}s$$
$$+ O((\Delta x_n)^{q+1}/\Delta t_n + \Delta t_n + \omega(\Delta t_n))$$
$$= \int_0^T \int_{\mathbb{R}^d} \left(\frac{\sigma^2}{2} \Delta\phi(t,x) + \langle b(s,x), \nabla\phi(t,x) \rangle \right) m^n(t,x) \mathrm{d}x \mathrm{d}t$$
$$+ O((\Delta x_n)^{q+1}/\Delta t_n + \Delta t_n + \omega(\Delta t_n)),$$

where $\omega : [0,\infty) \to \mathbb{R}$ is a modulus of continuity of b on $[0,T] \times K$. Thus,

$$\int_0^T \int_{\mathbb{R}^d} \left[\partial_t \phi(t,x) - \frac{\sigma^2}{2} \Delta\phi(t,x) - \langle b(s,x), \nabla\phi(t,x) \rangle \right] m^n(t,x) \mathrm{d}x \mathrm{d}t = O((\Delta x_n)^{q+1}/\Delta t_n + \Delta t_n + \omega(\Delta t_n))$$

and hence, passing to the weak limit in $L^2([0,T] \times \mathbb{R}^d)$, we get

(3.33)
$$\int_0^T \int_{\mathbb{R}^d} \left[\partial_t \phi(t,x) - \frac{\sigma^2}{2} \Delta\phi(t,x) - \langle b(s,x), \nabla\phi(t,x) \rangle \right] \widehat{m}(t,x) \mathrm{d}x \mathrm{d}t = 0.$$

Since the vector space spanned by $\{\eta\psi \mid \eta \in C_0^\infty((0,T)), \ \psi \in C_0^\infty(\mathbb{R}^d)\}$ is dense in $C_0^{1,2}((0,T) \times \mathbb{R}^d)$ (as in [41, Corollary 1.6.2 of the Weierstrass Approximation Theorem]), we get that (3.29) holds for any $\phi \in C_0^{1,2}((0,T) \times \mathbb{R}^d)$.

Finally, let us show that for any $\phi \in C_0(\mathbb{R}^d)$

$$(3.34) \qquad \int_{\mathbb{R}^d} \phi(x)(\widehat{m}(t,x) - m_0^*(x))\mathrm{d}x \to 0 \quad \text{as } t \to 0^+.$$

By Lemma 3.1, we have that $\widehat{m} \in C([0,T]; \mathcal{D}'(\mathbb{R}^d))$. Moreover, by [22, Lemma 2.1], for any $t \in [0,T]$ and for every $\phi \in C_0(\mathbb{R}^d)$, it holds that

$$(3.35) \qquad \lim_{s \to t, \, s \in [0,T]} \int_{\mathbb{R}^d} \phi(x)\widehat{m}(s,x)\mathrm{d}x = \int_{\mathbb{R}^d} \phi(x)\widehat{m}(t,x)\mathrm{d}x.$$

Since Theorem 3.1(i) implies that $\widehat{m}(0, \cdot) = m_0^*(\cdot)$, (3.34) follows from (3.35) with $t = 0$. Thus, the result follows from (3.29), (3.34) and the uniqueness result in Theorem 2.1(iv). $\qquad \square$

Remark 3.2. *The convergence of the sequence $(m^n)_{n \in \mathbb{N}}$ to m^* in the previous theorem is rather weak. On the other hand, to the best of our knowledge this is the first convergence result of a high-order LG scheme for equation (FP). Notice that our proof does not depend on the smoothness of m^* recalled in Theorem 2.1(i), and it can be easily adapted to deal with equations whose second-order term are not uniformly elliptic (see e.g. [19, 15]).*

4. The scheme for MFG

To derive a high-order scheme that approximates a solution (v^*, m^*) of the MFG system, we are left to derive a high-order method for the HJB equation, which coupled with (3.11) will provide the desired discretization of system (MFG).

Given $\mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$, we consider the HJB equation:

$$(\textbf{HJB}) \qquad \begin{aligned} -\partial_t v - \tfrac{\sigma^2}{2}\Delta v + \tfrac{1}{2}|\nabla v|^2 &= F(x, \mu(t)) \quad \text{in } (0,T) \times \mathbb{R}^d, \\ v(T, \cdot) &= G(\cdot, \mu(T)) \quad \text{in } \mathbb{R}^d. \end{aligned}$$

Standard results for quasilinear parabolic equations (see e.g. [31, Chapter IV and V]) yield that (HJB) admits a unique classical solution $v[\mu]$. Moreover, using that $v[\mu]$ is the value function associated with a stochastic optimal control problem (see e.g. [23, Chapters IV and V]), it is easy to check that (H1)-(H2) imply the existence of $R > 0$ such that

$$|\nabla v[\mu](t,x)| \le R \quad \text{for all } t \in [0,T], \, x \in \mathbb{R}^d, \, \mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d)).$$

We now describe a variation of the scheme in [7] to deal with the nonlinearity of the Hamiltonian in (HJB) with respect to ∇v (see also [38, 42] for related constructions). For a given $\mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$, let us define $\{v_{k,i} \,|\, k \in \mathcal{I}_{\Delta t}, \, i \in \mathcal{I}_{\Delta x}\} \subset \mathbb{R}$ as the solution to

$$(4.1) \qquad \begin{aligned} v_{k,i} &= S[\mu](v_{\cdot,k+1}, k, i) \quad \text{for all } k \in \mathcal{I}_{\Delta t}^*, \, i \in \mathcal{I}_{\Delta x}, \\ v_{N_{\Delta t}, i} &= G(x_i, \mu(t_{N_{\Delta t}})) \quad \text{for all } i \in \mathcal{I}_{\Delta x}, \end{aligned}$$

where, for a given $f = \{f_i\}_{i \in \mathcal{I}_{\Delta x}} \subset \mathbb{R}$, $k \in \mathcal{I}_{\Delta t}^*$, and $i \in \mathcal{I}_{\Delta x}$,

$$(4.2) \qquad \begin{aligned} S[\mu](f, k, i) = \inf_{\alpha \in A} &\left[\sum_{\ell \in \mathcal{I}_d} \omega_\ell \left(I[f](x_i - \Delta t \alpha + \sqrt{\Delta t}\sigma e^\ell) + \frac{\Delta t}{2} F(x_i - \Delta t \alpha + \sqrt{\Delta t}\sigma e^\ell, \mu(t_{k+1})) \right) \right. \\ &\left. + \frac{\Delta t}{2}|\alpha|^2 \right] + \frac{\Delta t}{2} F(x_i, \mu(t_k)), \end{aligned}$$

with $A = \{\alpha \in \mathbb{R}^d \,|\, |\alpha| \le R\}$ and $I[f]$ being defined by (3.8). The following consistency result for $S[\mu]$ follows from (4.2) and (H2).

Proposition 4.1. *Let $(\Delta t_n, \Delta x_n)_{n \in \mathbb{N}} \subset (0, +\infty)^2$, $(k_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$, $(i_n)_{n \in \mathbb{N}} \subset \mathbb{Z}^d$, $(\mu_n)_{n \in \mathbb{N}} \subset C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$, and $\mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$. Assume that (H2) holds and, as $n \to \infty$, $(\Delta t_n, \Delta x_n) \to (0,0)$, $(\Delta x_n)^{q+1}/\Delta t_n \to 0$, $k_n \in \mathcal{I}_{\Delta t_n}$, $i_n \in \mathcal{I}_{\Delta x_n}$, $t_{k_n} \to t$, $x_{i_n} \to x$, and $\mu_n \to \mu$. Then, for every $\phi \in C_b^{1,3}([0,T] \times \mathbb{R}^d)$ satisfying $\|\nabla \phi\|_{L^\infty([0,T] \times \mathbb{R}^d)} \le R$, we have*

$$\lim_{n \to \infty} \frac{1}{\Delta t_n}[\phi(t_{k_n}, x_{i_n}) - S[\mu_n](\phi_{k_n+1}, k_n, i_n)] = -\partial_t \phi(t,x) - \frac{\sigma^2}{2}\Delta \phi(t,x) + \frac{1}{2}|\nabla \phi(t,x)|^2 - F(x, \mu(t)),$$

where $\phi_k = \{\phi(t_k, x_i)\}_{i \in \mathcal{I}_{\Delta x}}$.

The proof is given in the Appendix.

For $\mu \in C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$, let us define

$$(4.3) \qquad v_\Delta[\mu](t,x) := I[v_{[t/\Delta t]}](x) \quad \text{for all} \quad (t,x) \in [0,T] \times \mathcal{O}_\Delta,$$

where $v_{k,i}$ is given by (4.1). For $\ell \in \mathcal{I}_d$ and $k \in \mathcal{I}_{\Delta t}^*$, let $y_k^\ell[\mu](x)$ be the unique solution to

$$(4.4) \qquad y = x - \frac{\Delta t}{2}\left[Dv_\Delta[\mu](t_k,x) + Dv_\Delta[\mu](t_{k+1},x),y)\right] + \sqrt{\Delta t}\sigma e^\ell,$$

where $Dv_\Delta[\mu](t_k,x)$ represents a numerical gradient with respect to x of $v_\Delta[\mu](t_k,x)$, computed by a fourth-order finite difference approximation.

We propose the following scheme for (**MFG**): find $\{(v_{k,i}, m_{k,i}) \in \mathbb{R}^2 \mid k \in \mathcal{I}_{\Delta t}, \ i \in \mathcal{I}_{\Delta x}\}$ such that, for all $k \in \mathcal{I}_{\Delta t}^*$ and $i \in \mathcal{I}_{\Delta x}$,

$$v_{k,i} = S_\Delta[m_\Delta](v_{k+1},k,i),$$
$$v_{N_{\Delta t},i} = G(x_i,m),$$

$$(4.5) \qquad \sum_{j \in \mathcal{I}_{\Delta x}} m_{k+1,j} \int_{\mathcal{O}_\Delta} \beta_i(x)\beta_j(x)\mathrm{d}x = \sum_{j \in \mathcal{I}_{\Delta x}} m_{k,j} \sum_{\ell \in \mathcal{I}_d} \omega_\ell \int_{\mathcal{O}_\Delta} \beta_i(y_k^\ell[m_\Delta](x))\beta_j(x)\mathrm{d}x,$$

$$\sum_{j \in \mathcal{I}_{\Delta x}} m_{0,j} \int_{\mathcal{O}_\Delta} \beta_i(x)\beta_j(x)\mathrm{d}x = \int_{\mathcal{O}_\Delta} m_0^*(x)\beta_i(x)\mathrm{d}x.$$

In the examples considered in the following section, system (4.5) will be solved, heuristically and as in [13, 36, 16], by using fixed-point iterations.

5. Numerical results

In this section, we show the performance of the proposed scheme on two different problems: a MFG with non-local couplings and an explicit solution, and a MFG with local couplings and no explicit solutions. For each test, we measure the accuracy of the scheme by computing the following relative errors in the discrete uniform and L^2 norms

$$E_\infty = \frac{\max_{i \in \mathcal{I}_{\Delta x}} |h_\Delta(T,x_i) - h(T,x_i)|}{\max_{i \in \mathcal{I}_{\Delta x}} |h(T,x_i)|}, \quad E_2 = \left(\frac{\mathrm{Int}_{\mathcal{O}_\Delta}(|h_\Delta(T,x) - h(T,x)|^2)}{\mathrm{Int}_{\mathcal{O}_\Delta}(|h(T,x)|^2)}\right)^{1/2},$$

where $h = m$, v, $h_\Delta = m_\Delta$, v_Δ, and $\mathrm{Int}_{\mathcal{O}_\Delta}$ denotes the approximation of the Riemann integral on $\mathcal{O}_\Delta$ by using the Simpson's Rule. We denote by p_∞ and p_2 the rates of convergence for E_∞ and E_2, respectively.

Notice that, for the exactly integrated scheme (3.11), the local truncation error is given by the contributions of (3.3) and (3.9), which yield a global truncation error of order $(\Delta x)^{q+1}/\Delta t + (\Delta t)^2$. As in [19], we get that the order of consistency is maximized by taking $\Delta t = O((\Delta x)^{(q+1)/3})$. With respect to the space discretization step, the previous choice suggests an order of convergence given by $2(q+1)/3$. In all the simulations we take $q = 3$, which yields an heuristic optimal rate equal to $8/3$, and Simpson's Rule to approximate the integrals in (3.11). The convergence rate of the resulting scheme is illustrated numerically in the examples below. Indeed, the tables in the tests show rates of convergence p_∞ and p_2 greater than 2 in most of the cases. The positivity preservation of the discrete density is true only when linear basis functions are used. This property is not in general verified by our method, and it holds only asymptotically. In the first numerical test, we calculate the maximum value of the negative part of the approximate density in the space-time mesh and call this value positivity error. We will show that with the refinement of the mesh, the positivity error decreases until it cancels.

5.1. An implementable version of the scheme (3.11).
In order to obtain an implementable version of (3.14), an approximation of the integrals therein has to be introduced. For simplicity, we consider the one-dimensional case, we use Simpson's Rule on each element $[x_j, x_j + 2\Delta x]$ ($j = 2m$, $m \in \mathbb{Z}$) and cubic symmetric Lagrange interpolation basis functions β_j ($p = 1$ in (3.7)). Recalling that β_j has support in

$[x_{j-2}, x_{j+2}]$, letting $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ otherwise, the entries of the mass matrix A (see (3.12)) are approximated by

$$(5.1) \qquad \int_{\mathcal{O}_\Delta} \beta_i(x)\beta_j(x)\mathrm{d}x = \int_{x_{j-2}}^{x_j} \beta_i(x)\beta_j(x)\mathrm{d}x + \int_{x_j}^{x_{j+2}} \beta_i(x)\beta_j(x)\mathrm{d}x \simeq \frac{2\Delta x}{3}\delta_{i,j},$$

while the entries of B_k^ℓ (see (3.13)) are approximated by

$$(5.2) \qquad (B_k^\ell)_{i,j} = \int_{x_{j-2}}^{x_{j+2}} \beta_i(y_k^\ell(x))\beta_j(x)\mathrm{d}x \simeq \frac{2\Delta x}{3}\beta_i(y_k^\ell(x_j)).$$

We observe that, as usual in LG methods, the integrands in (5.1) and (5.2) have not the necessary regularity in order to guarantee the standard accuracy order of the quadrature rule. This can lead to fluctuations in the order of convergence, as can be observed in some instances of the numerical tests below. However, in those tests we will see that the aforementioned quadrature rule provides an overall order of convergence close to 8/3.

Using (5.1) and (5.2), the scheme (3.14) is approximated by

$$(5.3) \qquad \begin{aligned} m_{k+1} &= \sum_{\ell \in \mathcal{I}_d} \omega_\ell \widetilde{B}_k^\ell m_k \quad \text{for } k \in \mathcal{I}_{\Delta t}^*, \\ m_0 &= \widetilde{m}_0, \end{aligned}$$

where $\widetilde{B}_k^\ell$ is a $(2N_{\Delta x} + 1) \times (2N_{\Delta x} + 1)$ matrix with entries given by

$$(\widetilde{B}_k^\ell)_{i,j} = \beta_i(y_k^\ell(x_j))$$

and $\widetilde{m}_0$ is vector of length $2N_{\Delta x} + 1$ given by

$$\widetilde{m}_{0,i} = m_0^*(x_i) \quad \text{for } i \in \mathcal{I}_{\Delta x}.$$

Remark 5.1. *Applied to a linearization of equation (**HJB**), scheme (5.3) is the dual of the semi-Lagrangian scheme [19] when a Crank-Nicolson method is used to discretize the characteristic curves, together with a cubic symmetric Lagrange interpolation to reconstruct the values in the space variable. Moreover, scheme (5.3) is also a natural higher-order extension of the scheme proposed in [14, 15] to approximate second-order MFGs.*

5.2. **Fixed-point iterations.** In view of (5.3), it is natular to propose the following implementable version of the scheme for (**MFG**): find $(\mathcal{V}, \mathcal{M}) := \{(\overline{v}_{k,i}, \overline{m}_{k,i}) \in \mathbb{R}^2 \mid k \in \mathcal{I}_{\Delta t}, \ i \in \mathcal{I}_{\Delta x}\}$ such that, for all $k \in \mathcal{I}_{\Delta t}^*$ and $i \in \mathcal{I}_{\Delta x}$,

$$(5.4) \qquad \begin{aligned} v_{k,i} &= S_\Delta[m_\Delta](v_{k+1}, k, i), \\ v_{N_{\Delta t},i} &= G(x_i, m), \end{aligned}$$

$$(5.5) \qquad \begin{aligned} m_{k+1,i} &= \sum_{j \in \mathcal{I}_{\Delta x}} m_{k,j} \sum_{\ell \in \mathcal{I}_d} \omega_\ell \beta_i(y_k^\ell(x_j)), \\ m_{0,i} &= m_0^*(x_i). \end{aligned}$$

This system is heuristically solved by the fixed-point iterations described in **Algorithm 1**, which has as input data a damping (or relaxation) parameter $\omega \in [0, 1]$, an initial guess $\mathcal{M}_\circ$ for the density $\mathcal{M}$, and a tolerance parameter $\tau > 0$. The iterations are stopped as soon as the L^1-norm, approximated by the Simpson's Rule, of the difference between two consecutive approximations of $\mathcal{M}$ is less than τ.

5.3. **Non-local MFG with analytical solution.** Consider the non-local MFG system

$$(5.6) \qquad \begin{aligned} -\partial_t v - \tfrac{\sigma^2}{2}\Delta v + \tfrac{1}{2}|\nabla v|^2 &= \tfrac{1}{2}\left(x - \int_{\mathbb{R}^d} y m(t,y)\mathrm{d}y\right)^2 \quad \text{in } [0,T) \times \mathbb{R}^d, \\ \partial_t m - \tfrac{\sigma^2}{2}\Delta m - \mathrm{div}\,(\nabla v m) &= 0 \quad \text{in } (0,T] \times \mathbb{R}^d, \\ v(T, \cdot) = 0, \qquad m(0, \cdot) &= m_0^* \quad \text{in } \mathbb{R}^d, \end{aligned}$$

where m_0^* is the density of a Gaussian random vector with mean $\mu_0 \in \mathbb{R}^d$ and covariance matrix $\Sigma_0 \in \mathbb{R}^{d \times d}$. For simplicity, we will assume that Σ_0 is a diagonal matrix.

Algorithm 1: Fixed-point iterations

Input   : Initial guess $\mathcal{M}_\circ$, damping parameter ω, tolerance $\tau > 0$.

Output: Approximation $(\mathcal{V}, \mathcal{M})$ of the solution to (5.4)-(5.5).

1 Initialize $\widetilde{\mathcal{M}}^{(0)} = \mathcal{M}^{(0)} = \mathcal{M}_\circ$, $p = 0$,

2 do

3 $\quad$ compute $\mathcal{V}^{(p+1)}$ solution to (5.4) with m_Δ replaced by $\widetilde{\mathcal{M}}^{(p)}$,

4 $\quad$ compute $Dv_\Delta^{(p+1)}$ numerical gradient of $\mathcal{V}^{(p+1)}$,

5 $\quad$ compute $\mathcal{M}^{(p+1)}$ solution to (5.5) with y_k^ℓ solution to (4.4) obtained using $Dv_\Delta^{(p+1)}$,

6 $\quad$ compute $E^{(p+1)} = \text{Int}_{[0,T] \times \mathcal{O}_\Delta} |\mathcal{M}^{(p+1)} - \mathcal{M}^{(p)}|$,

7 $\quad$ let $\widetilde{\mathcal{M}}^{(p+1)} = \omega \widetilde{\mathcal{M}}^{(p)} + (1 - \omega)\mathcal{M}^{(p+1)}$,

8 $\quad$ set $p = p + 1$,

9 while $E^{(p+1)} > \tau$;

10 return $(\mathcal{V}^{(p+1)}, \mathcal{M}^{(p+1)})$.

In what follows, we compute explicitly the unique solution (v^*, m^*) to (5.6) (see e.g. [4]). Since v^* is the value function associated with a linear-quadratic optimal control problem, standard results (see e.g. [46, Chapter 6]) show that v^* has the form

$$(5.7) \qquad v^*(t,x) = \tfrac{1}{2}\langle \Pi(t)x, x\rangle + \langle s(t), x\rangle + c(t) \quad \text{for } (t,x) \in [0,T] \times \mathbb{R}^d,$$

where, setting $\overline{Y}(t) = \int_{\mathbb{R}^d} y m^*(t,y)\mathrm{d}y$ for all $t \in [0,T]$, Π, s, and c satisfy

$$(5.8) \qquad \begin{aligned} -\dot{\Pi}(t) &= -\Pi^2(t) + I_d & &\text{for } t \in (0,T), \\ -\dot{s}(t) &= -\Pi(t)s(t) - \overline{Y}(t) & &\text{for } t \in (0,T), \\ -\dot{c}(t) &= \tfrac{\sigma^2}{2}\text{Tr}(\Pi(t)) - \tfrac{1}{2}|s(t)|^2 + \tfrac{1}{2}\left|\overline{Y}(t)\right|^2 & &\text{for } t \in (0,T), \\ \Pi(T) &= 0, \quad s(T) = 0, \quad c(T) = 0. \end{aligned}$$

Notice that Π satisfies a Riccati equation whose analytical solution is given by

$$(5.9) \qquad \Pi(t) = \left(\frac{e^{2T-t} - e^t}{e^{2T-t} + e^t} \right) I_d \quad \text{for } t \in [0,T].$$

Since $\nabla v^*(t,x) = \Pi(t)x + s(t)$, the SDE underlying the FP equation in (5.6) (see (2.1)) is given by

$$\begin{aligned} \mathrm{d}Y(t) &= [-\Pi(t)Y(t) - s(t)]\,\mathrm{d}t + \sigma \mathrm{d}W(t) \quad \text{for } t \in (0,T), \\ Y(0) &= Y_0, \end{aligned}$$

where Y_0 is a Gaussian random variable, independent of the d-dimensional Brownian motion W, with mean μ_0 and covariance matrix Σ_0. Since

$$(5.10) \qquad Y(t) = Y_0 + \int_0^t [-\Pi(r)Y(r) - s(r)]\,\mathrm{d}r + \sigma W(t) \quad \text{for } t \in (0,T),$$

and the coordinates $Y_{0,i}$ ($i = 1, \ldots, d$) of Y_0 are independent Gaussian random variables with means $\mu_{0,i}$ and variance $(\Sigma_0)_{i,i}$, for every $t \in [0,T]$, $Y(t)$ is a vector of independent Gaussian random variables $Y_i(t)$ ($i = 1, \ldots, d$) with mean $\overline{Y}_i(t)$ and variance $(\Sigma(t))_{i,i} = \mathbb{E}\left(Y_i^2(t)\right) - \overline{Y}_i^2(t)$ to be determined. In other words,

$$(5.11) \qquad m^*(t,x) = \Pi_{i=1}^d m_i^*(t, x_i) \quad \text{for } t \in [0,T], \ x \in \mathbb{R}^d,$$

where, for every $t \in [0,T]$ and $i = 1, \ldots, d$, $m_i^*(t, \cdot)$ is a univariate Gaussian density with parameters $\overline{Y}_i(t)$ and variance $(\Sigma(t))_{i,i}$ In order to compute these parameters, notice that (5.10) implies that

$$\overline{Y}(t) = \mu_0 + \int_0^t \left(-\Pi(r)\overline{Y}(r) - s(r)\right)\mathrm{d}r \quad \text{for } t \in [0,T],$$

i.e. $\overline{Y}$ solves

(5.12)
$$\begin{aligned} \dot{\overline{Y}}(t) &= -\Pi(t)\overline{Y}(t) - s(t) \quad \text{for } t \in (0,T), \\ \overline{Y}(0) &= \mu_0. \end{aligned}$$

Thus, by (5.8) and (5.12), the couple $(\overline{Y}, s)$ solves the boundary value problem

$$\begin{aligned} \dot{\overline{Y}}(t) &= -\Pi(t)\overline{Y}(t) - s(t) && \text{for } t \in (0,T), \\ \dot{s}(t) &= \Pi(t)s(t) + \overline{Y}(t) && \text{for } t \in (0,T), \\ \overline{Y}(0) &= \mu_0, && s(T) = 0, \end{aligned}$$

whose unique solution is given by (see e.g. [25])

(5.13)
$$\overline{Y}(t) = \mu_0, \quad s(t) = -\Pi(t)\mu_0 \quad \text{for } t \in [0,T],$$

where we recall that Π is given by (5.9).

On the other hand, by (5.10) and Itô's lemma, for every $i = 1, \ldots, d$, we have

$$Y_i^2(t) = Y_{0,i}^2 - \int_0^t 2Y_i(r)\left[\Pi_{i,i}(r)Y_i(r) + s_i(r)\right]\mathrm{d}r + 2\sigma\int_0^t Y_i(r)\mathrm{d}W_i(r) + \sigma^2 t \quad \text{for } t \in [0,T].$$

Thus, denoting by $m_{0,i}^*$ the i the marginal of m_0^* $(i = 1, \ldots, d)$, (5.13) yields

$$\mathbb{E}\left(Y_i^2(t)\right) = \int_{\mathbb{R}} x^2 m_{0,i}^*(x)\,\mathrm{d}x - 2\int_0^t \left[\Pi_{i,i}(r)\mathbb{E}\left(Y_i^2(r)\right) + \mu_{0,i}s_i(r)\right]\mathrm{d}r + \sigma^2 t \quad \text{for } t \in [0,T].$$

In particular, $[0,T] \ni t \mapsto \mathbb{E}\left(Y_i^2(t)\right) \in \mathbb{R}$ is the unique solution to

$$\begin{aligned} \dot{M}(t) &= -2\Pi_{i,i}(t)M(t) - 2\mu_{0,i}s_i(t) + \sigma^2 && t \in (0,T), \\ M(0) &= \int_{\mathbb{R}} x^2 m_{0,i}^*(x)\,\mathrm{d}x, \end{aligned}$$

which, for all $t \in [0,T]$, is given by

(5.14)
$$\mathbb{E}\left(Y_i^2(t)\right) = \left(e^{2T-t} + e^t\right)^2 \left[\frac{2\int_{\mathbb{R}} x^2 m_{0,i}^*(x)\,\mathrm{d}x - 2\mu_{0,i}^2 + \sigma^2\left(e^{2T}+1\right)}{2(e^{2T}+1)^2} - \frac{\sigma^2}{2(e^{2T}+e^{2t})}\right] + \mu_{0,i}^2.$$

Thus, for all $i = 1, \ldots, d$ and $t \in [0,T]$,

(5.15)
$$(\Sigma(t))_{i,i} = \left(e^{2T-t} + e^t\right)^2 \left[\frac{2\int_{\mathbb{R}} x^2 m_{0,i}^*(x)\,\mathrm{d}x - 2\mu_{0,i}^2 + \sigma^2\left(e^{2T}+1\right)}{2(e^{2T}+1)^2} - \frac{\sigma^2}{2(e^{2T}+e^{2t})}\right].$$

Altogether, for all $t \in [0,T]$, $m^*(t, \cdot)$ is given by (5.11), where the parameters of the univariate Gaussian densities $m_i^*(t, \cdot)$ are given by (5.13) and (5.15), and the value function v^* is given by (5.7), with Π and s given by (5.9) and (5.13), respectively, and c, obtained by integrating the third equation of (5.8), is given by

$$c(t) = \tfrac{1}{2}\langle\Pi(t)\mu_0, \mu_0\rangle - \frac{\sigma^2 d}{2}\ln\left(\frac{2e^T}{e^{2T-t}+e^t}\right) \quad \text{for } t \in [0,T].$$

In this test, the assumption on the boundedness of b is not verified in all $\mathbb{R}^d$, however it is true in every bounded domain $\mathcal{O}_\Delta$. Let us now solve system (5.6) on a bounded domain in dimension $d = 1$. We choose $[0,T] \times \mathcal{O}_\Delta = [0, 0.25] \times (-2, 2)$, with Dirichlet boundary conditions on $\partial\mathcal{O}_\Delta$, the latter being equal to the exact solution of (5.6) for the HJB equation and homogeneous for the FP equation. The numerical approximation of the boundary conditions for the HJB equation is based on the technique proposed in [7], while for the FP equation we proceed as in the previous test. We consider two cases, one with $\sigma^2/2 = 0.05$ and the other one with $\sigma^2/2 = 0.005$. As input parameters for the fixed-point iterations, we set the initial guess for the density equal to the initial datum m_0^* at each time step, as damping parameter $\omega = 0$, and as tolerance $\tau = 10^{-9}$. Tables 1 and 2 show the errors and the convergence rates for the approximation of the HJB and FP equations, in the case where $\sigma^2/2 = 0.05$ and $\Delta t = (\Delta x)^{4/3}/4$. In order to show the advantages of the high-order scheme in this paper, we also solve system (5.6) with the low order numerical scheme proposed in [14]. We display in Table 2 (columns 6-9) the errors and convergence rates for the approximation of the FP equation. The comparison between the errors and the orders of the two schemes clearly shows the gain in accuracy achieved by the new scheme.

Tables 3 and 4 show the errors and the convergence rates for the approximation of the HJB and FP equations, in the case where $\sigma^2/2 = 0.005$ and $\Delta t = (\Delta x)^{4/3}/4$. The convergence rates tend to be close to the theoretical optimal rate $8/3$. Tables 5 and 6 show the errors and convergence rates for $v_\Delta(0, \cdot)$ and $m_\Delta(T, \cdot)$, which are calculated by taking $\Delta t = \Delta x/4$, and Tables 7 and 8 consider the case $\Delta t = (\Delta x)^2$. These tables show that the scheme is stable as the time steps change, however the convergence rate deteriorates slightly, especially for the approximation of the time-dependent density.

In Figure 1 we show the solution to (5.6) on $[0, T] \times \mathcal{O}_\Delta = [0, 0.25] \times (-2, 2)$ with $\sigma^2/2 = 0.005$, computed with $\Delta x = 1.25 \cdot 10^{-2}$ and $\Delta t = (\Delta x)^{4/3}/4$. Figure 2 displays a zoom of the initial density m_0^*, the exact solution $m^*(T, \cdot)$ and its approximation $m_\Delta(T, \cdot)$, computed with $\Delta x = 6.25 \cdot 10^{-3}$ and $\Delta t = (\Delta x)^{4/3}/4$.

Δx	Errors for the approximation of $v^*(0, \cdot)$			
	E_∞	E_2	p_∞	p_2
$2.00 \cdot 10^{-1}$	$6.20 \cdot 10^{-5}$	$7.40 \cdot 10^{-5}$	-	-
$1.00 \cdot 10^{-1}$	$1.09 \cdot 10^{-5}$	$1.43 \cdot 10^{-5}$	2.51	2.37
$5.00 \cdot 10^{-2}$	$2.13 \cdot 10^{-6}$	$3.41 \cdot 10^{-6}$	2.36	2.07
$2.50 \cdot 10^{-2}$	$5.42 \cdot 10^{-7}$	$1.00 \cdot 10^{-6}$	1.97	1.77

TABLE 1. Errors and convergence rates for the approximation of the value function of problem (5.6) with $d = 1$, $\sigma^2/2 = 0.05$, and $\Delta t = (\Delta x)^{4/3}/4$.
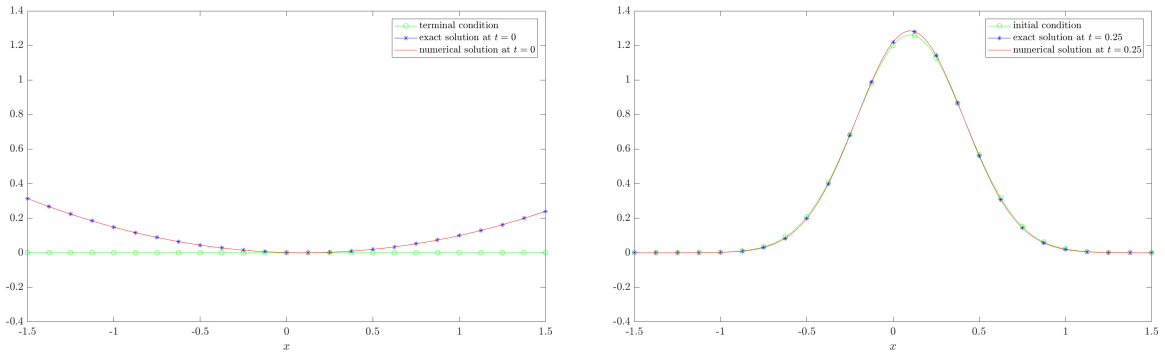

FIGURE 1. Solution to (5.6) on $[0, T] \times \mathcal{O}_\Delta = [0, 0.25] \times (-2, 2)$ with $\sigma^2/2 = 0.005$. On the left, we display the exact value function v^* at times $t = 0$, $t = 0.25$, and the numerical approximation v_Δ at time $t = 0$. On the right, we display the exact density m^* at times $t = 0$, $t = 0.25$, and the numerical approximation m_Δ at time $t = 0.25$.

5.4. **Mean field games with local couplings.** In this section, we approximate the solution of the second-order MFG system with local couplings studied in [43, Section 5.2]. Namely, we consider system

$$-\partial_t v - \frac{\sigma^2}{2}\Delta v + \frac{1}{2}|\nabla v|^2 = F(x, m) \quad \text{in } [0, T) \times \mathcal{O}_\Delta,$$

(5.16)
$$\partial_t m - \frac{\sigma^2}{2}\Delta m - \text{div}(\nabla v m) = 0 \quad \text{in } (0, T] \times \mathcal{O}_\Delta,$$

$$v(T, \cdot) = 0, \qquad m(0, \cdot) = m_0^* \quad \text{in } \mathbb{R}^d,$$

Δx	High-order scheme				Low-order scheme			
	E_∞	E_2	p_∞	p_2	E_∞	E_2	p_∞	p_2
$2.00 \cdot 10^{-1}$	$2.22 \cdot 10^{-2}$	$2.32 \cdot 10^{-2}$	-	-	$1.90 \cdot 10^{-1}$	$1.84 \cdot 10^{-1}$	-	-
$1.00 \cdot 10^{-1}$	$5.43 \cdot 10^{-3}$	$5.10 \cdot 10^{-3}$	2.03	2.19	$1.56 \cdot 10^{-1}$	$1.41 \cdot 10^{-1}$	0.28	0.38
$5.00 \cdot 10^{-2}$	$9.32 \cdot 10^{-4}$	$8.90 \cdot 10^{-4}$	2.54	2.52	$1.14 \cdot 10^{-1}$	$1.01 \cdot 10^{-1}$	0.45	0.48
$2.50 \cdot 10^{-2}$	$1.33 \cdot 10^{-4}$	$1.26 \cdot 10^{-4}$	2.81	2.82	$7.77 \cdot 10^{-2}$	$6.86 \cdot 10^{-2}$	0.55	0.56

TABLE 2. Errors and convergence rates for the approximation of the density of problem (5.6) with $d = 1$, $\sigma^2/2 = 0.05$, and $\Delta t = (\Delta x)^{4/3}/4$.

Δx	Errors for the approximation of $v^*(0,\cdot)$			
	E_∞	E_2	p_∞	p_2
$2.00 \cdot 10^{-1}$	$1.68 \cdot 10^{-4}$	$1.70 \cdot 10^{-4}$	-	-
$1.00 \cdot 10^{-1}$	$3.56 \cdot 10^{-5}$	$3.48 \cdot 10^{-5}$	2.24	2.29
$5.00 \cdot 10^{-2}$	$5.86 \cdot 10^{-6}$	$5.75 \cdot 10^{-6}$	2.60	2.60
$2.50 \cdot 10^{-2}$	$1.06 \cdot 10^{-6}$	$1.04 \cdot 10^{-6}$	2.47	2.47

TABLE 3. Errors and convergence rates for the approximation of the value function of problem (5.6) with $d = 1$, $\sigma^2/2 = 0.005$, and $\Delta t = (\Delta x)^{4/3}/4$.

Δx	Errors for the approximation of $m^*(T,\cdot)$				
	E_∞	E_2	p_∞	p_2	positivity error
$2.00 \cdot 10^{-1}$	$8.81 \cdot 10^{-3}$	$1.01 \cdot 10^{-2}$	-	-	$-3.51 \cdot 10^{-4}$
$1.00 \cdot 10^{-1}$	$3.06 \cdot 10^{-3}$	$2.53 \cdot 10^{-3}$	1.53	2.00	$-9.45 \cdot 10^{-9}$
$5.00 \cdot 10^{-2}$	$8.01 \cdot 10^{-4}$	$5.56 \cdot 10^{-4}$	1.93	2.19	0
$2.50 \cdot 10^{-2}$	$1.81 \cdot 10^{-4}$	$1.14 \cdot 10^{-4}$	2.15	2.29	0

TABLE 4. Errors and convergence rates for the approximation of the density of problem (5.6) with $d = 1$, $\sigma^2/2 = 0.005$, and $\Delta t = (\Delta x)^{4/3}/4$.
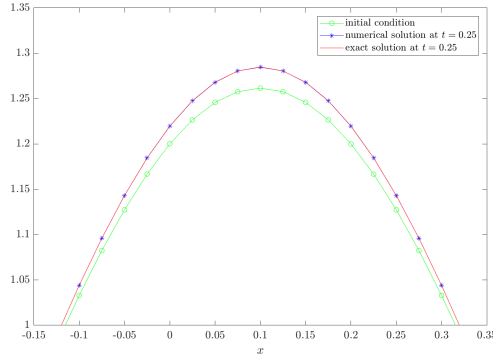

FIGURE 2. Zoom of the exact density m^* at times $t = 0$, $t = 0.25$, and of the numerical approximation m_Δ at time $t = 0.25$.

Δx	Errors for the approximation of $v^*(0,\cdot)$			
	E_∞	E_2	p_∞	p_2
$2.00 \cdot 10^{-1}$	$2.72 \cdot 10^{-4}$	$2.56 \cdot 10^{-4}$	-	-
$1.00 \cdot 10^{-1}$	$7.62 \cdot 10^{-5}$	$6.72 \cdot 10^{-5}$	1.84	1.93
$5.00 \cdot 10^{-2}$	$1.61 \cdot 10^{-5}$	$1.44 \cdot 10^{-5}$	2.24	2.22
$2.50 \cdot 10^{-2}$	$3.69 \cdot 10^{-6}$	$3.59 \cdot 10^{-6}$	2.13	2.00

TABLE 5. Errors and convergence rates for the approximation of the value function of problem (5.6) and $\sigma^2/2 = 0.005$, $\Delta t = \Delta x/4$.

with $T = 0.05$, $\mathcal{O}_\Delta = (0,1)$, homogeneous Neumann boundary conditions at $x = 0$ and $x = 1$, $\sigma^2/2 = 0.05$,

$$m_0^*(x) = \begin{cases} 4\sin^2(2\pi x - \frac{1}{4}) & \text{if } x \in [\frac{1}{4}, \frac{3}{4}], \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad F(x,m) = 3m_0^*(x) - \min(4,m).$$

Notice that the coupling term F depends on the density m in a pointwise (or local) manner. The homogeneous Neumann boundary conditions are approximated as in [9]. In this example, we do not have an explicit expression for (v^*, m^*).

Δx	Errors for the approximation of $m^*(T, \cdot)$				
	E_∞	E_2	p_∞	p_2	positivity error
$2.00 \cdot 10^{-1}$	$5.93 \cdot 10^{-3}$	$7.01 \cdot 10^{-3}$	-	-	$-8.18 \cdot 10^{-5}$
$1.00 \cdot 10^{-1}$	$2.63 \cdot 10^{-3}$	$2.17 \cdot 10^{-3}$	1.17	1.69	$-3.58 \cdot 10^{-10}$
$5.00 \cdot 10^{-2}$	$1.23 \cdot 10^{-3}$	$4.80 \cdot 10^{-4}$	1.10	2.18	0
$2.50 \cdot 10^{-2}$	$3.39 \cdot 10^{-4}$	$9.61 \cdot 10^{-5}$	1.86	2.32	0

TABLE 6. Errors and convergence rates for the approximation of the density of problem (5.6) and $\sigma^2/2 = 0.005$, $\Delta t = \Delta x/4$.

Δx	Errors for the approximation of $v^*(0, \cdot)$			
	E_∞	E_2	p_∞	p_2
$2.00 \cdot 10^{-1}$	$1.98 \cdot 10^{-4}$	$1.87 \cdot 10^{-4}$	-	-
$1.00 \cdot 10^{-1}$	$2.84 \cdot 10^{-5}$	$2.86 \cdot 10^{-5}$	2.80	2.71
$5.00 \cdot 10^{-2}$	$3.41 \cdot 10^{-6}$	$3.94 \cdot 10^{-5}$	3.06	2.86
$2.50 \cdot 10^{-2}$	$4.56 \cdot 10^{-7}$	$5.08 \cdot 10^{-6}$	2.90	2.96

TABLE 7. Errors and convergence rates for the approximation of the value function of problem (5.6) and $\sigma^2/2 = 0.005$, $\Delta t = (\Delta x)^2$.

Δx	Errors for the approximation of $m^*(T, \cdot)$				
	E_∞	E_2	p_∞	p_2	positivity error
$2.00 \cdot 10^{-1}$	$6.60 \cdot 10^{-3}$	$7.63 \cdot 10^{-3}$	-	-	$-1.01 \cdot 10^{-4}$
$1.00 \cdot 10^{-1}$	$3.11 \cdot 10^{-3}$	$2.60 \cdot 10^{-3}$	1.09	1.55	$-1.19 \cdot 10^{-8}$
$5.00 \cdot 10^{-2}$	$9.17 \cdot 10^{-4}$	$7.16 \cdot 10^{-4}$	1.76	1.86	0
$2.50 \cdot 10^{-2}$	$2.42 \cdot 10^{-4}$	$1.81 \cdot 10^{-4}$	1.92	1.98	0

TABLE 8. Errors and convergence rates for the approximation of the density of problem (5.6) and $\sigma^2/2 = 0.005$, $\Delta t = (\Delta x)^2$.

We consider two cases, one with $\sigma^2/2 = 0.05$ and the other one with $\sigma^2/2 = 0.005$. As input parameters for the fixed-point iterations, we set the initial guess for the density equal to the initial datum m_0^* at each time step, as damping parameter $\omega = 0.5$ and as tolerance $\tau = 10^{-9}$. In order to compute the errors and rates of convergence, we compare our approximations (v_Δ, m_Δ) with a reference solution, which is still denoted by (v^*, m^*), computed with $\Delta x = 6.67 \cdot 10^{-4}$ and $\Delta t = (\Delta x)^{3/2}/3$. In Tables 9 and 10, we show the errors and convergence rates for $v_\Delta(0, \cdot)$, $\partial_x v_\Delta(0, \cdot)$, and $m_\Delta(T, \cdot)$, which are computed by taking $\Delta t = (\Delta x)^{3/2}/3$ for different values of Δx. We observe an order of convergence greater than two in most of the cases. In order to show the main advantage of the proposed scheme over a low-order scheme, we compare the proposed scheme with a first-order semi-Lagrangian scheme for MFG, developed in [14]. Table 10, columns 4 to 8, shows that, using the low-order scheme, the density is approximated with a much lower accuracy and the convergence rate is also much lower. Finally, Figure 3 shows the approximated density m_Δ at time $t = T$ and the approximated value function v_Δ, together with its gradient Dv_Δ, at time $t = 0$. These approximations are computed with $\Delta x = 1.56 \cdot 10^{-3}$.

Δx	Errors for the approximation of $v^*(0, \cdot)$				Errors for the approximation of $\partial_x v^*(0, \cdot)$			
	E_∞	E_2	p_∞	p_2	E_∞	E_2	p_∞	p_2
$5.00 \cdot 10^{-2}$	$5.38 \cdot 10^{-2}$	$3.80 \cdot 10^{-2}$	-	-	$8.09 \cdot 10^{-2}$	$4.96 \cdot 10^{-2}$	-	-
$2.50 \cdot 10^{-2}$	$1.43 \cdot 10^{-2}$	$1.29 \cdot 10^{-2}$	1.91	1.55	$1.37 \cdot 10^{-2}$	$1.19 \cdot 10^{-2}$	2.53	2.05
$1.25 \cdot 10^{-2}$	$4.25 \cdot 10^{-3}$	$3.24 \cdot 10^{-3}$	1.75	1.99	$3.94 \cdot 10^{-3}$	$2.79 \cdot 10^{-3}$	1.80	2.09
$6.25 \cdot 10^{-3}$	$8.84 \cdot 10^{-4}$	$7.99 \cdot 10^{-4}$	2.27	2.02	$8.34 \cdot 10^{-4}$	$7.07 \cdot 10^{-4}$	2.23	1.98

TABLE 9. Errors and convergence rates for the approximation of $v^*(0, \cdot)$ and $\partial_x v^*(0, \cdot)$ of problem (5.16), with $\Delta t = (\Delta x)^{3/2}/3$

Δx	High-order scheme				Low-order scheme			
	E_∞	E_2	p_∞	p_2	E_∞	E_2	p_∞	p_2
$5.00 \cdot 10^{-2}$	$9.07 \cdot 10^{-2}$	$4.82 \cdot 10^{-2}$	-	-	$1.97 \cdot 10^{-0}$	$6.30 \cdot 10^{-1}$	-	-
$2.50 \cdot 10^{-2}$	$1.81 \cdot 10^{-2}$	$6.79 \cdot 10^{-3}$	2.32	2.82	$1.15 \cdot 10^{-0}$	$3.63 \cdot 10^{-1}$	0.77	0.79
$1.25 \cdot 10^{-2}$	$4.81 \cdot 10^{-3}$	$1.36 \cdot 10^{-3}$	1.91	2.32	$9.03 \cdot 10^{-1}$	$2.80 \cdot 10^{-1}$	0.35	0.48
$6.25 \cdot 10^{-3}$	$7.64 \cdot 10^{-4}$	$2.06 \cdot 10^{-4}$	2.65	2.72	$1.96 \cdot 10^{-1}$	$4.96 \cdot 10^{-2}$	0.48	0.51

TABLE 10. Errors and convergence rates for the approximation of $m^*(T, \cdot)$ of problem (5.16), with $\Delta t = (\Delta x)^{3/2}/3$
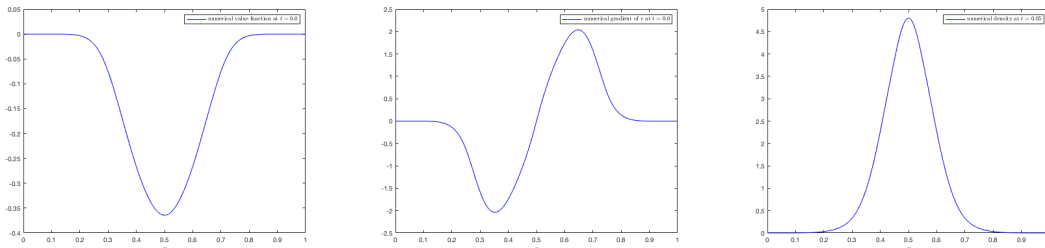

FIGURE 3. Approximated value function $v_\Delta(0, \cdot)$ (left), the derivate $Dv_\Delta[\mu](0, \cdot)$ (center), and approximated density $m_\Delta(T, \cdot)$ of problem (5.16)(right).

6. CONCLUSIONS AND FUTURE PERSPECTIVES

The main aim of the paper is to present a new and efficient high-order scheme to solve MFG systems with regular solutions. In order to do so, we have developed a new high-order scheme for the (**FP**) equation, based on Lagrange-Galerkin methods combined with a second-order weak approximation of the underlying stochastic characteristic curves. A convergence analysis has been provided in the distributional sense and with respect to the weak topology in L^2. We have then combined the new scheme for the (**FP**) equation with a high-order semi-Lagrangian scheme for the (**HJB**) equation to obtain a high-order scheme for the (**MFG**) system. We have shown the performance of the scheme by numerical simulations. We heuristically expect convergence rate 8/3, which is reached in some cases. The main advantage of the scheme, as usual for semi-Lagrangian schemes, is that during the fixed point iterations the equations (**HJB**) and (**FP**) are solved by schemes which are explicit and do not require the standard parabolic CFL condition $\Delta t = O((\Delta x)^2)$ in order to be stable. Recall that the CFL condition is required by standard explicit finite difference schemes to approximate parabolic PDEs. Instead, in Theorem 3.2 and in Proposition 4.1 the relation $(\Delta x)^{q+1}/\Delta t \to 0$ is assumed, which implies that larger time steps than $\Delta t = O((\Delta x)^2)$ are allowed. The restriction on the time step is only due to accuracy. In fact, as Tables 6 and 8 show, the accuracy may decrease for time steps that are far from the optimal one. Similar considerations were observed in [19], where high-order semi-Lagrangian schemes are applied to approximate linear parabolic PDEs. In addition, the scheme for the (**FP**) is conservative, which is not generally true for semi-Lagrangian type schemes applied to conservation laws (see e.g. [8]). The main drawbacks of the scheme are the loss of positivity for the discrete density, and the lack of a constant high-order convergence rate. Both drawbacks are due to the choice of standard cubic basis functions. Investigation on the use of a different class of basis function is an interesting point to be addressed in the future. Concerning the observed oscillations in the rate of convergence, we attribute them mostly to the lack of regularity of the integrands appearing in the schemes which yields, possibly, a lack of accuracy in the approximation of the integral terms. More regular basis functions may help to improve the quadrature error and then the overall truncation errors.

References

[1] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: numerical methods. *SIAM J. Numer. Anal.*, 48(3):1136–1162, 2010.

[2] Y. Achdou and M. Laurière. Mean field games and applications: numerical aspects. In *Mean field games*, volume 2281 of *Lecture Notes in Math.*, pages 249–307. Springer, Cham, 2020.

[3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Second edition. Lecture notes in Mathematics ETH Zürich. Birkhäuser Verlag, Bassel, 2008.

[4] A. Bensoussan, K. C. J. Sung, S. C. P. Yam, and S. P. Yung. Linear-quadratic mean field games. *J. Optim. Theory Appl.*, 169(2):496–529, 2016.

[5] R. Bermejo and L. Saavedra. Modified Lagrange-Galerkin methods of first and second order in time for convection-diffusion problems. *Numer. Math.*, 120(4):601–638, 2012.

[6] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov. *Fokker-Planck-Kolmogorov equations*, volume 207 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2015.

[7] L. Bonaventura, E. Calzola, E. Carlini, and R. Ferretti. Second order fully semi-Lagrangian discretizations of advection-diffusion-reaction systems. *J. Sci. Comput.*, 88(1):Paper No. 23, 29, 2021.

[8] L. Bonaventura and R. Ferretti. Semi-Lagrangian methods for parabolic problems in divergence form. *SIAM J. Sci. Comput.*, 36(5):A2458–A2477, 2014.

[9] E. Calzola, E. Carlini, X. Dupuis, and F. J. Silva. A semi-Lagrangian scheme for Hamilton-Jacobi-Bellman equations with oblique derivatives boundary conditions. *Numer. Math.*, 153(1):49–84, 2023.

[10] F. Camilli and M. Falcone. An approximation scheme for the optimal control of diffusion processes. *RAIRO Modél. Math. Anal. Numér.*, 29(1):97–122, 1995.

[11] P. Cardaliaguet. Notes on Mean Field Games: from P.-L. Lions' lectures at Collège de France. *Lecture Notes given at Tor Vergata*, 2010.

[12] P. Cardaliaguet and A. Porretta. An introduction to mean field game theory. In *Mean field games*, volume 2281 of *Lecture Notes in Math.*, pages 1–158. Springer, Cham, 2020.

[13] E. Carlini and F. J. Silva. A fully discrete semi-Lagrangian scheme for a first order mean field game problem. *SIAM J. Numer. Anal.*, 52(1):45–67, 2014.

[14] E. Carlini and F. J. Silva. A semi-Lagrangian scheme for a degenerate second order mean field game system. *Discrete and Continuous Dynamical Systems*, 35(9):4269–4292, 2015.

[15] E. Carlini and F. J. Silva. On the discretization of some nonlinear Fokker-Planck-Kolmogorov equations and applications. *SIAM J. Numer. Anal.*, 56(4):2148–2177, 2018.

[16] I. Chowdhury, O. Ersland, and E. R. Jakobsen. On numerical approximations of fractional and nonlocal mean field games. *Found. Comput. Math.*, 23(4):1381–1431, 2023.

[17] P. G. Ciarlet and J.-L. Lions, editors. *Handbook of numerical analysis. Vol. II.* Handbook of Numerical Analysis, II. North-Holland, Amsterdam, 1991. Finite element methods. Part 1.

[18] M. Falcone and R. Ferretti. *Semi-Lagrangian Approximation Schemes for Linear and Hamilton-Jacobi Equations.* MOS-SIAM Series on Optimization, 2013.

[19] R. Ferretti. A technique for high-order treatment of diffusion terms in semi-Lagrangian schemes. *Commun. Comput. Phys.*, 8(2):445–470, 2010.

[20] R. Ferretti. On the relationship between semi-Lagrangian and Lagrange-Galerkin schemes. *Numer. Math.*, 124(1):31–56, 2013.

[21] R. Ferretti and M. Mehrenberger. Stability of semi-Lagrangian schemes of arbitrary odd degree under constant and variable advection speed. *Math. Comp.*, 89(324):1783–1805, 2020.

[22] A. Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 253:109–153, 2008.

[23] 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.

[24] G. Fu, S. Liu, S. Osher, and W. Li. High order computation of optimal transport, mean field planning, and potential mean field games. *J. Comput. Phys.*, 491:Paper No. 112346, 21, 2023.

[25] M.T. Heath. *Scientific Computing: An Introductory Survey.* McGraw-Hill Education, 2005.

[26] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–251, 2006.

[27] J. L. Kelley. *General topology.* D. Van Nostrand Co., Inc., Toronto-New York-London, 1955.

[28] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York).* Springer-Verlag, Berlin, 1992.

[29] P. E. Kloeden, E. Platen, and Schurz H. *Numerical Solution of SDE Through Computer Experiments*, volume 23 of *Universitext.* Springer Berlin, Heidelberg, 1994.

[30] M. Krukowski. Arzelà-Ascoli's theorem in uniform spaces. *Discrete Contin. Dyn. Syst. Ser. B*, 23(1):283–294, 2018.

[31] O. A. Ladyvzenskaja, V. A. Solonnikov, and N. N. Uralceva. *Linear and quasilinear equations of parabolic type.* Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1968.

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

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

[34] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2:229–260, 2007.

[35] M. Laurière. Numerical methods for mean field games and mean field type control. In *Mean field games*, volume 78 of *Proc. Sympos. Appl. Math.*, pages 221–282. Amer. Math. Soc., Providence, RI, 2021.

[36] H. Li, Y. Fan, and L. Ying. A simple multiscale method for mean field games. *J. Comput. Phys.*, 439:Paper No. 110385, 18, 2021.

[37] R. Meise and D. Vogt. *Introduction to functional analysis*, volume 2 of *Oxford Graduate Texts in Mathematics*. The Clarendon Press, Oxford University Press, New York, 1997. Translated from the German by M. S. Ramanujan.

[38] G.N. Milstein and M.V. Tretyakov. Numerical algorithms for semilinear parabolic equations with small parameter based on approximation of stochastic equations. *Math. Comp.*, 69:237–567, 2000.

[39] G.N. Milstein and M.V. Tretyakov. *Stochastic numerics for mathematical physics*. Springer Science & Business Media, 2013.

[40] K. W. Morton, A. Priestley, and E. Süli. Stability of the Lagrange-Galerkin method with nonexact integration. *RAIRO Modél. Math. Anal. Numér.*, 22(4):625–653, 1988.

[41] R. Narasimhan. *Analysis on Real and Complex Manifolds*. Advanced studies in pure mathematics. Masson, 1973.

[42] A. Picarelli and C. Reisinger. Probabilistic error analysis for some approximation schemes to optimal control problems. *Systems Control Lett.*, 137:104619, 11, 2020.

[43] B. Popov and V. Tomov. Central schemes for mean field games. *Commun. Math. Sci.*, 13(8):2177–2194, 2015.

[44] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*. Springer Verlag, 1994.

[45] J. Voigt. *A course on topological vector spaces*. Compact Textbooks in Mathematics. Birkhäuser/Springer, Cham, 2020.

[46] J. Yong and X. Y. Zhou. *Stochastic controls*, volume 43 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1999. Hamiltonian systems and HJB equations.

7. Appendix: Proofs.

Proof of Theorem 2.1 We refer the reader to [6, Theorem 6.6.1] for the existence result in (i) as well as for the nonnegativity property in (ii). The uniqueness result in (i) and the mass conservation property in (iii) follow from [6, Theorem 9.3.6] and [6, Corollary 6.6.6], respectively. Finally, the proof of (iv) is given in [22, Theorem 4.3].   □

Proof of Proposition 3.1. In the proof of both assertions, we fix $\phi \in C_0^2(\mathbb{R}^d)$ and we will denote by C a positive real number which can depend on ϕ but not on Δ. We will also use the estimate

$$(7.1) \qquad \left| \sum_{\ell \in \mathcal{I}_d} \omega_\ell \phi(y_k^\ell(x)) - \left[\phi(x) + \Delta t \left(\frac{\sigma^2}{2} \Delta \phi(x) + \langle b(x, t_k), \nabla \phi(x) \rangle \right) \right] \right| \leq C(\Delta t)^2 \quad \text{for } x \in \mathbb{R}^d,$$

which follows from the definition of $y_k^\ell(x)$ and a Taylor expansion (see for instance [7]).

(i) Let us first show the assertion for $t = t_{k+1}$ and $s = t_k$ for some $k \in \mathcal{I}_{\Delta t}^*$. Set $\varepsilon := \phi - I[\phi]$ and fix $k \in \mathcal{I}_{\Delta t}^*$. Remark 3.1 yields the existence of $C > 0$ such that

$$(7.2) \qquad \left| \int_{\mathbb{R}^d} \phi(x) \left(m_\Delta(t_{k+1}, x) - m_\Delta(t_k, x) \right) \mathrm{d}x \right| \quad \leq \quad \left| \int_{\mathbb{R}^d} I[\phi](x) \left(m_\Delta(t_{k+1}, x) - m_\Delta(t_k, x) \right) \mathrm{d}x \right|$$
$$+ C \|\varepsilon\|_{L^\infty}.$$

Recalling that $\text{supp}\{m_\Delta(t_k, \cdot)\} \subset \mathcal{O}_\Delta$ and using the definition of the scheme in (3.11), we have that

(7.3)
$$
\begin{aligned}
\int_{\mathbb{R}^d} I[\phi](x)\, (m_\Delta(t_{k+1}, x) - m_\Delta(t_k, x))\, \mathrm{d}x &= \int_{\mathcal{O}_\Delta} \sum_{i \in \mathbb{Z}^d} \phi(x_i) \beta_i(x) \left(\sum_{j \in \mathcal{I}_{\Delta x}} (m_{k+1,j} - m_{k,j})\, \beta_j(x) \right) \mathrm{d}x \\
&= \sum_{i \in \mathbb{Z}^d} \phi(x_i) \left(\sum_{j \in \mathcal{I}_{\Delta x}} (m_{k+1,j} - m_{k,j}) \int_{\mathcal{O}_\Delta} \beta_i(x) \beta_j(x) \mathrm{d}x \right) \\
&= \sum_{i \in \mathbb{Z}^d} \phi(x_i) \Bigg[\sum_{\ell \in \mathcal{I}_d} \omega_\ell \sum_{j \in \mathcal{I}_{\Delta x}} m_{k,j} \bigg(\int_{\mathcal{O}_\Delta} \beta_i(y_k^\ell(x)) \beta_j(x) \mathrm{d}x \\
&\qquad\qquad\qquad\qquad\qquad\qquad - \int_{\mathcal{O}_\Delta} \beta_i(x) \beta_j(x) \mathrm{d}x \bigg) \Bigg] \\
&= \sum_{\ell \in \mathcal{I}_d} \omega_\ell \sum_{j \in \mathcal{I}_{\Delta x}} m_{k,j} \int_{\mathcal{O}_\Delta} \Big[I[\phi](y_k^\ell(x)) - I[\phi](x) \Big] \beta_j(x) \mathrm{d}x \\
&= \sum_{\ell \in \mathcal{I}_d} \omega_\ell \int_{\mathcal{O}_\Delta} \Big[I[\phi](y_k^\ell(x)) - I[\phi](x) \Big] m_\Delta(t_k, x)\, \mathrm{d}x.
\end{aligned}
$$

On the other hand, since ϕ has a compact support, there exists $C > 0$ such that

(7.4)
$$
\left\| \sum_{\ell \in \mathcal{I}_d} \omega_\ell \left(I[\phi](y_k^\ell(\cdot)) - \phi(y_k^\ell(\cdot)) \right) \right\|_{L^2} + \|\phi - I[\phi]\|_{L^2} \leq C \|\varepsilon\|_{L^\infty}
$$

and, by (7.1) and **(H2)**, there exists $C > 0$ such that

(7.5)
$$
\left\| \sum_{\ell \in \mathcal{I}_d} \omega_\ell \left(\phi(y_k^\ell(\cdot)) - \phi \right) \right\|_{L^2} \leq C \Delta t.
$$

Thus, by the triangular and the Cauchy-Schwarz inequalities, Theorem 3.1(iii), (7.2), (7.3), (7.4), and (7.5), we get the existence of $C > 0$ such that

$$
\left| \int_{\mathbb{R}^d} \phi(x)\, (m_\Delta(t_{k+1}, x) - m_\Delta(t_k, x))\, \mathrm{d}x \right| \leq C \left(\|\varepsilon\|_{L^\infty} + \Delta t \right).
$$

It follows from (3.9), and the condition $(\Delta x)^{q+1} \leq \Delta t$, the existence of $C > 0$ such that (3.27) holds for $t = t_{k+1}$ and $s = t_k$. Using this relation and the triangular inequality, we deduce that (3.27) holds for every $s = t_k$ and $t = t_m$ with $k, m \in \mathcal{I}_{\Delta t}$.

Now, let us fix $s, t \in [0, T]$ and assume, without loss of generality, that $t > s$. Let $k_1, k_2 \in \mathcal{I}_{\Delta t}^*$ be such that $s \in [t_{k_1}, t_{k_1+1}]$ and $t \in [t_{k_2}, t_{k_2+1}]$. If $k_1 = k_2$, then it follows from (3.25) that (3.27) holds with $C_\phi = C$. Otherwise, $k_2 \geq k_1 + 1$ and (3.25) yield

(7.6)
$$
\begin{aligned}
\left| \int_{\mathbb{R}^d} \phi(x)\, (m_\Delta(t_{k_1+1}, x) - m_\Delta(s, x))\, \mathrm{d}x \right| &\leq \frac{t_{k_1+1} - s}{\Delta t} \left| \int_{\mathbb{R}^d} \phi(x)\, (m_\Delta(t_{k_1+1}, x) - m_\Delta(t_{k_1}, x))\, \mathrm{d}x \right| \\
&\leq C(t_{k_1+1} - s).
\end{aligned}
$$

Similarly,

(7.7)
$$
\left| \int_{\mathbb{R}^d} \phi(x)\, (m_\Delta(t_{k_2}, x) - m_\Delta(t, x))\, \mathrm{d}x \right| \leq C(t - t_{k_2}).
$$

Altogether, it follows from the triangular inequality, (7.6), (7.7), and (3.27), with $t = t_{k_2}$ and $s = t_{k_1+1}$, that (3.27) holds with $C_\phi = C$.

(ii) By (3.9), Remark 3.1, and the definition of the scheme (3.11), for each $k \in \mathcal{I}_{\Delta t}^*$ we have

$$
\begin{aligned}
\int_{\mathbb{R}^d} \phi(x) m_\Delta(t_{k+1}, x) \mathrm{d}x &= \int_{\mathbb{R}^d} I[\phi](x) m_\Delta(t_{k+1}, x) \mathrm{d}x + O((\Delta x)^{q+1}) \\
&= \sum_{i \in \mathbb{Z}^d} \phi(x_i) \sum_{j \in \mathcal{I}_{\Delta x_n}} m_{k+1,j} \int_{\mathbb{R}^d} \beta_i(x) \beta_j(x) \mathrm{d}x + O((\Delta x)^{q+1}) \\
&= \sum_{i \in \mathbb{Z}^d} \phi(x_i) \sum_{j \in \mathcal{I}_{\Delta x}} m_{k,j} \sum_{\ell \in \mathcal{I}_d} \omega_\ell \int_{\mathbb{R}^d} \beta_i(y_k^\ell(x)) \beta_j(x) \mathrm{d}x + O((\Delta x)^{q+1}) \\
&= \sum_{j \in \mathcal{I}_{\Delta x_n}} m_{k,j} \sum_{\ell \in \mathcal{I}_d} \omega_\ell \int_{\mathbb{R}^d} I[\phi](y_k^\ell(x)) \beta_j(x) \mathrm{d}x + O((\Delta x)^{q+1}) \\
&= \int_{\mathbb{R}^d} \left(\sum_{\ell \in \mathcal{I}_d} \omega_\ell \phi(y_k^\ell(x)) \right) m_\Delta(t_k, x) \mathrm{d}x + O((\Delta x)^{q+1}).
\end{aligned}
\tag{7.8}
$$

Using (7.1) and Remark 3.1, we obtain
$$
\begin{aligned}
\int_{\mathbb{R}^d} \phi(x) \left(m_\Delta(t_{k+1}, x) - m_\Delta(t_k, x) \right) \mathrm{d}x &= \Delta t \int_{\mathbb{R}^d} \left(\frac{\sigma^2}{2} \Delta \phi(x) + \langle b(t_k, x), \nabla\phi(x) \rangle \right) m_\Delta(t_k, x) \mathrm{d}x \\
&\quad + O((\Delta x)^{q+1} + (\Delta t)^2).
\end{aligned}
\tag{7.9}
$$

Notice that, for any $s \in [t_k, t_{k+1}]$, Remark 3.1 implies that

$$
\int_{\mathbb{R}^d} \langle b(t_k, x), \nabla\phi(x) \rangle m_\Delta(t_k, x) \mathrm{d}x = \int_{\mathbb{R}^d} \langle b(s, x), \nabla\phi(x) \rangle m_\Delta(t_k, x) \mathrm{d}x + O(\omega_\phi(\Delta t)).
\tag{7.10}
$$

By (3.25) and the fact that $b(s, \cdot) \in C^{q+1}(\mathbb{R}^d)$, together with assertion (i), we have

$$
\left| \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}^d} \left(\frac{\sigma^2}{2} \Delta\phi(x) + \langle b(s, x), \nabla\phi(x) \rangle \right) (m_\Delta(s, x) - m_\Delta(t_k, x)) \mathrm{d}x \mathrm{d}s \right| = O((\Delta t)^2).
\tag{7.11}
$$

Thus, (3.28) follows from (7.10), (7.11), and (7.9). $\qquad\square$

Proof of Lemma 3.1. In view of the Arzelà-Ascoli theorem [27, Chapter 7, Theorem 18] (see also [30, Section 4]) and Proposition 3.1(i), it suffices to show that the family $\mathcal{M}$ is pointwise relatively compact. Let us consider the absolutely convex set $U_0 := \{\phi \in C_0^\infty(\mathbb{R}^d) \mid \|\phi\|_{L^\infty} < 1, \text{ supp}\, \phi \subseteq \overline{B}(0,1)\}$. This set is a neighborhood of 0 in the standard topology of $C_0^\infty(\mathbb{R}^d)$ (see e.g. [45, Chapter 10]) and, for any $t \in [0, T]$,

$$
\sup_{\phi \in U_0} \left| \int_{\mathbb{R}^d} m_\Delta(t, x) \phi(x) \mathrm{d}x \right| = \sup_{\phi \in U_0} \left| \int_{\overline{B}(0,1)} m_\Delta(t, x) \phi(x) \mathrm{d}x \right| \leq \|m_\Delta(t, \cdot)\|_{L^1(\overline{B}(0,1))} \leq r,
$$

where $r := \sup\{\|m_\Delta(t, \cdot)\|_{L^1(\overline{B}(0,1))} \,|\, \Delta \in (0, \infty)^2\}$ belongs to $[0, +\infty)$ by (3.26). This proves that $\{m_\Delta(t, \cdot) \,|\, \Delta \in (0, \infty)^2\} \subset \{T \in \mathcal{D}'(\mathbb{R}^d) \,|\, \sup_{\phi \in U_0} |T(\phi)| \leq r\}$ which, by the Banach-Alaoglu-Bourbaki theorem (see e.g. [37, Theorem 23.5]), is a compact subset of $\mathcal{D}'(\mathbb{R}^d)$. $\qquad\square$

Proof of Proposition 4.1. Let $\Delta t > 0$, $\Delta x > 0$, and $\alpha \in A$. In the computations below, the big O terms are uniform with respect to $\alpha \in A$. Let us apply (7.1) to $\phi(t_{k+1}, \cdot)$, with $b(t, x) = -\alpha$, to obtain

$$
\begin{aligned}
\sum_{\ell \in \mathcal{I}_d} \omega_\ell \phi\left(t_{k+1}, x_i - \Delta t \alpha + \sqrt{\Delta t} \sigma e^\ell\right) &= \phi(t_{k+1}, x_i) + \Delta t \left(\frac{\sigma^2}{2} \Delta\phi(t_{k+1}, x_i) - \langle \nabla\phi(t_{k+1}, x_i), \alpha \rangle \right) \\
&\quad + O\left((\Delta t)^2\right).
\end{aligned}
\tag{7.12}
$$

By (**H2**) and using the first-order Taylor expansion of $F(\cdot, \mu(t_k))$ around x_i, we get

$$
\begin{aligned}
\frac{1}{2} \left(\sum_{\ell \in \mathcal{I}_d} \omega_\ell F(x_i - \Delta t \alpha + \sqrt{\Delta t} \sigma e^\ell, \mu(t_{k+1})) + F(x_i, \mu(t_k)) \right) &= F(x_i, \mu(t_{k+1})) \\
&\quad + O(\Delta t + \mathbf{d}(\mu(t_{k+1}), \mu(t_k))).
\end{aligned}
\tag{7.13}
$$

Thus, by (4.2), (7.12), (7.13), and (3.9), we obtain

$$
\begin{aligned}
S[\mu](\phi_{k+1}, k, i) &= \phi(t_{k+1}, x_i) - \Delta t \sup_{\alpha \in A} \left[\langle \nabla \phi(t_{k+1}, x_i), \alpha \rangle - \frac{|\alpha|^2}{2} \right] + \Delta t \frac{\sigma^2}{2} \Delta \phi(t_{k+1}, x_i) \\
&\quad + \Delta t F(x_i, \mu(t_{k+1})) + O\left((\Delta t)^2 + (\Delta x)^{q+1} + \Delta t \mathbf{d}(\mu(t_{k+1}), \mu(t_k))\right) \\
&= \phi(t_{k+1}, x_i) - \frac{\Delta t}{2} |\nabla \phi(t_{k+1}, x_i)|^2 + \Delta t \frac{\sigma^2}{2} \Delta \phi(t_{k+1}, x_i) \\
&\quad + \Delta t F(x_i, \mu(t_{k+1})) + O\left((\Delta t)^2 + (\Delta x)^{q+1} + \Delta t \mathbf{d}(\mu(t_{k+1}), \mu(t_k))\right).
\end{aligned}
$$

Finally, we get

$$
\begin{aligned}
\frac{1}{\Delta t} \left[\phi(t_k, x_i) - S_\Delta[\mu](\phi_{k+1}, k, i) \right] = &-\partial_t \phi(t_{k+1}, x_i) - \frac{\sigma^2}{2} \Delta \phi(t_{k+1}, x_i) + \frac{1}{2} |\nabla \phi(t_{k+1}, x_i)|^2 - F(x_i, \mu(t_{k+1})) \\
&+ O\left(\Delta t + \frac{(\Delta x)^{q+1}}{\Delta t} + \mathbf{d}(\mu(t_{k+1}), \mu(t_k)) \right),
\end{aligned}
$$

from which the result follows. $\qquad\square$