

Fourier approximation methods for first-order nonlocal mean-field games

Levon Nurbekyan and João Saúde

Abstract. In this note, we develop Fourier approximation methods for the solutions of first-order nonlocal mean-field games (MFG) systems. Using Fourier expansion techniques, we approximate a given MFG system by a simpler one that is equivalent to a convex optimization problem over a finite-dimensional subspace of continuous curves. Furthermore, we perform a time-discretization for this optimization problem and arrive at a finite-dimensional saddle point problem. Finally, we solve this saddle-point problem by a variant of a primal dual hybrid gradient method.

Mathematics Subject Classification (primary; secondary): 35Q91, 35Q93; 35A01

Keywords: Infinite-dimensional differential games, mean-field games, nonlocal interactions, Fourier expansions

1. Introduction

The mean-field game (MFG) framework [28], [29], [30], [31], [32] models systems with a huge number of small identical rational players (agents) that play non-cooperative differential games. In this framework, a generic player aims at minimizing a cost functional that takes the distribution of the whole population as a parameter. Consequently, the problem is to find a Nash equilibrium where a generic player cannot unilaterally improve his position. For a detailed account on MFG systems we refer the reader to [12], [15], [18], [19], [24], [25], [27], [34].

In this note, we introduce Fourier approximation techniques for first-order nonlocal MFG models. More precisely, we consider the system

$$\begin{cases} -\partial_t u + H(x, \nabla u) = F[x, m], \\ \partial_t m - \operatorname{div}(m \nabla_\rho H(x, \nabla u)) = 0, & (x, t) \in \mathbb{T}^d \times [0, 1], \\ m(x, 0) = M(x), \quad u(x, 1) = U(x), \quad x \in \mathbb{T}^d. \end{cases} \quad (1)$$

Here, $u : \mathbb{T}^d \times [0, 1] \rightarrow \mathbb{R}$ and $m : \mathbb{T}^d \times [0, 1] \rightarrow \mathbb{R}_+$ are the unknown functions. Furthermore, $H \in C^2(\mathbb{T}^d \times \mathbb{R}^d)$ is a Hamiltonian, and $F : \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d) \rightarrow \mathbb{R}$ is a nonlocal coupling term between the Hamilton–Jacobi and Fokker–Planck equations. Above, \mathbb{T}^d is the d -dimensional flat torus, and $\mathcal{P}(\mathbb{T}^d)$ is the space of Borel probability measures on \mathbb{T}^d . Next, $U \in C^2(\mathbb{T}^d)$ and $M \in L^\infty(\mathbb{T}^d) \cap \mathcal{P}(\mathbb{T}^d)$ (with a slight abuse of notation we identify the absolutely continuous measures with their densities) are terminal-initial conditions for u and m , respectively.

In (1), u represents the value function of a generic agent from a continuum population of players, whereas m represents the density of this population. Each agent aims at solving the optimization problem

$$u(x, t) = \inf_{\gamma \in H^1([t, 1]), \gamma(t)=x} \int_t^1 L(\gamma(s), \dot{\gamma}(s)) + F(\gamma(s), m(\cdot, s)) ds + U(\gamma(1)), \quad (2)$$

where L is the Legendre transform of H ; that is,

$$L(x, v) = \sup_p -v \cdot p - H(x, p), \quad (x, v) \in \mathbb{T}^d \times \mathbb{R}^d.$$

Hence, U is a terminal cost function. Since a generic agent is small and her actions on the population distribution can be neglected, we assume that m is fixed, but unknown, in (2). Consequently, u must solve a Hamilton–Jacobi equation; that is, the first PDE in (1) with terminal data U .

Furthermore, given u , optimal trajectories of agents are determined by

$$\dot{\gamma}(s) = -\nabla_p H(\gamma(s), \nabla u(\gamma(s), s)).$$

Therefore, m , being the population density, must satisfy the Fokker–Planck equation; that is, the second PDE in (1) with initial data M . Hence, M is the population density at time $t = 0$.

The existence, uniqueness and stability theories for (1) are well understood [14], [15], [32]. Here, we are specifically interested in approximation methods for the solutions of (1) that can be useful for numerical solution and modeling purposes.

Currently, there are number of effective approximation methods for solutions of MFG systems. We refer to [1], [2], [3], [4] for finite-difference schemes, [16], [17] for semi-Lagrangian methods, [6], [9], [10], [11], [13] for convex optimization techniques, [5], [26] for monotone flows, and [22] for infinite-dimensional Hamilton–Jacobi equations. Although general, the majority of methods above are particularly advantageous when F in (1) depends locally on m . The reason is that local F yield analytic pointwise formulas for infinite-dimensional operators involved in the algorithms. Instead, nonlocal F do not yield such formulas. Additionally,

fixed-grid methods suffer from dimensionality issues. Also, the number of inter-nodal couplings grows significantly for nonlocal F which leads to an increased complexity of such schemes. Hence, we are interested in developing approximation methods that specifically suit nonlocal F and are grid-free.

Our approach is based on a Fourier approximation of F and is inspired by the methods in [35]. Here, we use the classical trigonometric polynomials as an approximation basis. Nevertheless, our method is flexible and allows more general bases. For instance, one may consider (1) on different domains and boundary conditions and choose a basis accordingly.

Additionally, our approach yields a grid-free numerical approximation of u and m . More precisely, we directly recover the optimal trajectories of the agents rather than the values of u and m on a given grid. In particular, our methods may blend well with recently developed ideas for fast and curse-of-the-dimensionality-resistant solution approach for first-order Hamilton–Jacobi equations [21], [33], [36]. Hence, our techniques may lead to numerical schemes for nonlocal MFG that are efficient in high dimensions.

To avoid technicalities, we consider a linear F . More precisely, we assume that

$$F[x, m] = \int_{\mathbb{T}^d} K(x, y)m(y, t) dy, \quad x \in \mathbb{T}^d, m \in \mathcal{P}(\mathbb{T}^d),$$

where $K \in C^2(\mathbb{T}^d \times \mathbb{T}^d)$. Thus, here we deal with the system

$$\begin{cases} -\partial_t u + H(x, \nabla u) = \int_{\mathbb{T}^d} K(x, y)m(y, t) dy, \\ \partial_t m - \operatorname{div}(m \nabla_p H(x, \nabla u)) = 0, & (x, t) \in \mathbb{T}^d \times [0, 1], \\ m(x, 0) = M(x), \quad u(x, 1) = U(x), & x \in \mathbb{T}^d. \end{cases} \quad (3)$$

Our basic idea is to show that when K is a generalized polynomial in a given basis then (3) is equivalent to a fixed point problem, in a space of continuous curves, that has nice structural properties. In particular, when K is symmetric and positive semi-definite, (3) is equivalent to a convex optimization problem in the space of continuous curves.

Furthermore, we discuss how to construct generalized polynomial kernels that approximate a given K . Additionally, we observe that for translation invariant K the approximating kernels have a particularly simple structure. Consequently, for such K the aforementioned optimization problem is much simpler to solve.

The paper is organized as follows. In Section 2, we present standing assumptions and some preliminary results. In Section 3, we prove the equivalence of (3) to a fixed point problem over the space of continuous curves when K is a generalized polynomial. Next, in Section 4, we discuss approximation methods for a general kernel. Furthermore, in Section 5, we construct a discretization for the

optimization problem from Section 3 and devise a variant of a primal dual hybrid gradient algorithm for the discrete problem. Finally, in Section 6, we study several numerical examples.

2. Assumptions and preliminary results

We denote by \mathbb{T}^d the d -dimensional flat torus. Furthermore, throughout the paper, we assume that $H \in C^2(\mathbb{T}^d \times \mathbb{R}^d)$, and

$$\begin{aligned} \frac{1}{C}I_d &\leq \nabla_{pp}^2 H(x, p) \leq CI_d, \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d, \\ -C(1 + |p|^2) &\leq \nabla_x H(x, p) \cdot p, \quad \forall (x, p) \in \mathbb{T}^d \times \mathbb{R}^d, \end{aligned} \tag{4}$$

for some constant $C > 0$. Next, we assume that $M \in L^\infty(\mathbb{T}^d) \cap \mathcal{P}(\mathbb{T}^d)$, $U \in C^2(\mathbb{T}^d)$, $K \in C^2(\mathbb{T}^d \times \mathbb{T}^d)$, and

$$\|M\|_{L^\infty(\mathbb{T}^d)}, \|U\|_{C^2(\mathbb{T}^d)}, \|K\|_{C^2(\mathbb{T}^d \times \mathbb{T}^d)} \leq C. \tag{5}$$

Additionally, we suppose that K is positive semi-definite; that is,

$$\int_{\mathbb{T}^d \times \mathbb{T}^d} K(x, y) f(x) f(y) \, dx \, dy \geq 0, \quad \forall f \in L^\infty(\mathbb{T}^d). \tag{6}$$

We call K symmetric if

$$K(x, y) = K(y, x), \quad \forall x, y \in \mathbb{T}^d. \tag{7}$$

Next, we denote by $\mathcal{P}(\mathbb{T}^d)$ the space of Borel probability measures on \mathbb{T}^d . We equip $\mathcal{P}(\mathbb{T}^d)$ with the Monge–Kantorovich distance that is given by

$$\|m_2 - m_1\|_{MK} = \sup \left\{ \int_{\mathbb{T}^d} \phi(x) (m_2(x) - m_1(x)) \, dx \text{ s.t. } \|\phi\|_{\text{Lip}} \leq 1 \right\}. \tag{8}$$

In the rest of this section, we present some preliminary results and formulas. For the optimal control and related Hamilton–Jacobi equations theory we refer to [8], [23]. We begin by the definition of a solution for (3).

Definition 2.1. A pair (u, m) is a solution of (3) if $u \in W^{1,\infty}(\mathbb{T}^d \times [0, 1])$ is a viscosity solution of

$$\begin{cases} -\partial_t u + H(x, \nabla u) = \int_{\mathbb{T}^d} K(x, y) m(y, t) \, dy, & (x, t) \in \mathbb{T}^d \times [0, 1], \\ u(x, 1) = U(x), & x \in \mathbb{T}^d, \end{cases} \tag{9}$$

and $m \in L^\infty(\mathbb{T}^d \times [0, 1]) \cap C([0, 1]; \mathcal{P}(\mathbb{T}^d))$ is a distributional solution of

$$\begin{cases} \partial_t m - \operatorname{div}(m \nabla_p H(x, \nabla u)) = 0, & (x, t) \in \mathbb{T}^d \times [0, 1], \\ m(x, 0) = M(x), & x \in \mathbb{T}^d. \end{cases} \tag{10}$$

The following theorem [14], [15], [32] asserts that (3) is well-posed.

Theorem 2.2. (i) *Under assumptions (4) and (5), system (3) admits a solution (u, m) . Moreover, there exists a constant $C_1(C) > 0$ such that*

$$\nabla_{xx}^2 u, \|u\|_{W^{1,\infty}}, \|m\|_{L^\infty} \leq C_1, \tag{11}$$

for any solution (u, m) . Additionally, if (6) holds then (u, m) is unique.

(ii) *Solutions of (3) are stable with respect to variations of U, M and K in respective norms. Particularly, suppose that $\{K_r\}_{r=1}^\infty \subset C^2(\mathbb{T}^d \times \mathbb{T}^d)$ is such that*

$$\lim_{r \rightarrow \infty} \|K - K_r\|_{C^2(\mathbb{T}^d \times \mathbb{T}^d)} = 0, \tag{12}$$

and $\{(u_r, m_r)\}_{r=1}^\infty$ are solutions of (3) corresponding to kernel K_r . Then, the sequence $\{(u_r, m_r)\}_{r=1}^\infty$ is precompact in $C(\mathbb{T}^d \times [0, 1]) \times C([0, 1]; \mathcal{P}(\mathbb{T}^d))$ with all accumulation points being solutions of (3). Consequently, if (6) holds then

$$\begin{aligned} \lim_{r \rightarrow \infty} u_r(x, t) &= u(x, t), & \text{uniformly in } (x, t) \in \mathbb{T}^d \times [0, 1], \\ \lim_{r \rightarrow \infty} \|m_r(\cdot, t) - m(\cdot, t)\|_{MK} &= 0, & \text{uniformly in } t \in [0, 1], \end{aligned} \tag{13}$$

where (u, m) is the unique solution of (3).

Next, consider an arbitrary basis of smooth functions

$$\Phi = \{\phi_1, \phi_2, \dots, \phi_r\} \subset C^2(\mathbb{T}^d). \tag{14}$$

For $a = (a_1, a_2, \dots, a_r) \in C([0, 1]; \mathbb{R}^r)$ we denote by u_a the viscosity solution of

$$\begin{cases} -\partial_t u(x, t) + H(x, \nabla u(x, t)) = \sum_{i=1}^r a_i(t) \phi_i(x), & (x, t) \in \mathbb{T}^d \times [0, 1] \\ u(x, 1) = U(x), & x \in \mathbb{T}^d. \end{cases} \tag{15}$$

From the optimal control theory, we have that

$$u_a(x, t) = \inf_{\gamma \in H^1([t, 1], \gamma(t)=x)} \int_t^1 \left(L(\gamma(s), \dot{\gamma}(s)) + \sum_{i=1}^r a_i(s) \phi_i(\gamma(s)) \right) ds + U(\gamma(1)), \tag{16}$$

for all $(x, t) \in \mathbb{T}^d \times [0, 1]$, where

$$L(x, v) = \sup_{p \in \mathbb{R}^d} -v \cdot p - H(x, p). \tag{17}$$

Moreover, for all $(x, t) \in \mathbb{T}^d \times [0, 1]$ there exists $\gamma_{x,t,a} \in C^2([t, 1]; \mathbb{T}^d)$ such that

$$u_a(x, t) = \int_t^1 \left(L(\gamma_{x,t,a}(s), \dot{\gamma}_{x,t,a}(s)) + \sum_{i=1}^r a_i(s) \phi_i(\gamma_{x,t,a}(s)) \right) ds + U(\gamma_{x,t,a}(1)), \tag{18}$$

and

$$\begin{aligned} & \frac{d}{ds} \nabla_v L(\gamma_{x,t,a}(s), \dot{\gamma}_{x,t,a}(s)) \\ &= \nabla_x L(\gamma_{x,t,a}(s), \dot{\gamma}_{x,t,a}(s)) + \sum_{i=1}^r a_i(s) \nabla \phi_i(\gamma_{x,t,a}(s)), \quad s \in [t, 1]. \end{aligned} \tag{19}$$

Additionally,

$$\begin{aligned} -\nabla_v L(x, \dot{\gamma}_{x,t,a}(t)) &\in \nabla_x^+ u_a(x, t), \\ -\nabla_v L(\gamma_{x,t,a}(s), \dot{\gamma}_{x,t,a}(s)) &= \nabla_x u_a(\gamma_{x,t,a}(s), s), \quad s \in (t, 1], \\ -\dot{\gamma}_{x,t,a}(s) &= \nabla_p H(\gamma_{x,t,a}(s), \nabla_x u_a(\gamma_{x,t,a}(s), s)), \quad s \in (t, 1]. \end{aligned} \tag{20}$$

In fact, this previous equation is also sufficient for (18) to hold. For lighter notation, we denote $\gamma_{x,0,a}$ by $\gamma_{x,a}$.

In general, u_a is not everywhere differentiable. Nevertheless, u_a is semiconcave and hence $\nabla^+ u_a(x, t) \neq \emptyset$ for all (x, t) , and $\nabla^+ u_a(x, t) = \{\nabla u_a(x, t)\}$ for a.e. (x, t) . In fact, points (x, t) where u_a is not differentiable are precisely those for which (16) admits multiple minimizers. Thus, at points $x \in \mathbb{T}^d$ where $u_a(x, 0)$ is not differentiable we choose $\gamma_{x,a}$ in such a way that the map $(x, t) \mapsto \gamma_{x,a}(t)$ is Borel measurable.

Furthermore, we denote by m_a the distributional solution of

$$\begin{cases} \partial_t m - \operatorname{div}(m \nabla_p H(x, \nabla u_a)) = 0, & (x, t) \in \mathbb{T}^d \times [0, 1], \\ m(x, 0) = M(x), & x \in \mathbb{T}^d. \end{cases} \tag{21}$$

One can show that m_a is given by the push-forward of the measure M by the map $\gamma_{\cdot,a}(t)$; that is,

$$m_a(\cdot, t) = \gamma_{\cdot,a}(t) \# M. \tag{22}$$

We equip $C([0, 1]; \mathbb{R}^r)$ with the L^∞ norm

$$\|a\|_\infty = \max_i \sup_{t \in [0, 1]} |a_i(t)|.$$

Then, one has that

$$\lim_{n \rightarrow \infty} \|m_{a_n}(\cdot, t) - m_a(\cdot, t)\|_{MK} = 0, \quad \text{uniformly in } t \in [0, 1], \tag{23}$$

if $\lim_{n \rightarrow \infty} \|a_n - a\|_\infty = 0$. For a detailed discussion on m_a see Chapter 4 in [15].

Finally, we denote by

$$G(a) = \int_{\mathbb{T}^d} u_a(x, 0) M(x) dx, \quad a \in C([0, 1]; \mathbb{R}^r). \tag{24}$$

Our first theorem addresses the properties of G .

Theorem 2.3. *The functional $a \mapsto G(a)$ is concave and everywhere Fréchet differentiable. Moreover,*

$$\partial_{a_i} G = \int_{\mathbb{T}^d} \phi_i(x) m_a(x, \cdot) dx, \quad 1 \leq i \leq r. \tag{25}$$

Proof. We denote by

$$p(a) = \left(\int_{\mathbb{T}^d} \phi_i(x) m_a(x, \cdot) dx \right)_{i=1}^r, \quad a \in C([0, 1]; \mathbb{R}^r).$$

We prove that for every $a \in C([0, 1]; \mathbb{R}^r)$

$$0 \geq G(b) - G(a) - (b - a) \cdot p(a) \geq o(\|b - a\|_\infty).$$

We have that

$$\begin{aligned} & G(b) - G(a) - (b - a) \cdot p(a) \\ &= \int_{\mathbb{T}^d} M(x) dx \int_0^1 \left(L(\gamma_{x,b}(t), \dot{\gamma}_{x,b}(t)) + \sum_{i=1}^r b_i(t) \phi_i(\gamma_{x,b}(t)) \right) dt + U(\gamma_{x,b}(1)) \\ &\quad - \int_{\mathbb{T}^d} M(x) dx \int_0^1 \left(L(\gamma_{x,a}(t), \dot{\gamma}_{x,a}(t)) + \sum_{i=1}^r a_i(t) \phi_i(\gamma_{x,a}(t)) \right) dt + U(\gamma_{x,a}(1)) \\ &\quad - \sum_{i=1}^r \int_0^1 (b_i(t) - a_i(t)) dt \int_{\mathbb{T}^d} \phi_i(x) m_a(x, t) dx. \end{aligned}$$

From (22) we have that

$$\int_{\mathbb{T}^d} \phi_i(x) m_a(x, t) dx = \int_{\mathbb{T}^d} \phi_i(\gamma_{x,a}(t)) M(x) dx, \quad t \in [0, 1], 1 \leq i \leq r.$$

Hence,

$$\begin{aligned} G(b) - G(a) - (b - a) \cdot p(a) &= \int_{\mathbb{T}^d} M(x) dx \int_0^1 L(\gamma_{x,b}(t), \dot{\gamma}_{x,b}(t)) - L(\gamma_{x,a}(t), \dot{\gamma}_{x,a}(t)) dt \\ &\quad + \int_{\mathbb{T}^d} M(x) dx \int_0^1 \sum_{i=1}^r b_i(t) (\phi_i(\gamma_{x,b}(t)) - \phi_i(\gamma_{x,a}(t))) dt \\ &\quad + \int_{\mathbb{T}^d} M(x) (U(\gamma_{x,b}(1)) - U(\gamma_{x,a}(1))) dx. \end{aligned}$$

By definition, we have that

$$\begin{aligned} &\int_0^1 L(\gamma_{x,b}(t), \dot{\gamma}_{x,b}(t)) + \sum_{i=1}^r b_i(t) \phi_i(\gamma_{x,b}(t)) dt + U(\gamma_{x,b}(1)) \\ &\leq \int_0^1 L(\gamma_{x,a}(t), \dot{\gamma}_{x,a}(t)) + \sum_{i=1}^r b_i(t) \phi_i(\gamma_{x,a}(t)) dt + U(\gamma_{x,a}(1)), \quad \forall x \in \mathbb{T}^d. \end{aligned}$$

Hence,

$$G(b) - G(a) - (b - a) \cdot p(a) \leq 0, \quad \forall a, b \in C([0, 1]; \mathbb{T}^d).$$

This previous inequality yields the concavity of G . On the other hand, we have that

$$\begin{aligned} G(b) - G(a) - (b - a) \cdot p(a) &= \int_{\mathbb{T}^d} M(x) dx \int_0^1 L(\gamma_{x,b}(t), \dot{\gamma}_{x,b}(t)) - L(\gamma_{x,a}(t), \dot{\gamma}_{x,a}(t)) dt \\ &\quad + \int_{\mathbb{T}^d} M(x) dx \int_0^1 \sum_{i=1}^r a_i(t) (\phi_i(\gamma_{x,b}(t)) - \phi_i(\gamma_{x,a}(t))) dt \\ &\quad + \int_{\mathbb{T}^d} M(x) (U(\gamma_{x,b}(1)) - U(\gamma_{x,a}(1))) dx \\ &\quad + \int_{\mathbb{T}^d} M(x) dx \int_0^1 \sum_{i=1}^r (b_i(t) - a_i(t)) (\phi_i(\gamma_{x,b}(t)) - \phi_i(\gamma_{x,a}(t))) dt. \end{aligned}$$

Therefore, again by the definition of $\gamma_{x,a}$ and $\gamma_{x,b}$, we have that

$$\begin{aligned}
 & G(b) - G(a) - (b - a) \cdot p(a) \\
 & \geq \int_{\mathbb{T}^d} M(x) \, dx \int_0^1 \sum_{i=1}^r (b_i(t) - a_i(t)) (\phi_i(\gamma_{x,b}(t)) - \phi_i(\gamma_{x,a}(t))) \, dt \\
 & \geq -\|b - a\|_\infty \sum_{i=1}^r \left| \int_0^1 \int_{\mathbb{T}^d} \phi_i(\gamma_{x,b}(t)) M(x) \, dx - \int_{\mathbb{T}^d} \phi_i(\gamma_{x,a}(t)) M(x) \, dx \right| dt \\
 & = -\|b - a\|_\infty \sum_{i=1}^r \left| \int_0^1 \int_{\mathbb{T}^d} \phi_i(x) m_b(x, t) \, dx - \int_{\mathbb{T}^d} \phi_i(x) m_a(x, t) \, dx \right| dt \\
 & \geq -\|b - a\|_\infty \sum_{i=1}^r \text{Lip}(\phi_i) \int_0^1 \|m_b(\cdot, t) - m_a(\cdot, t)\|_{MK} \, dt.
 \end{aligned}$$

Hence, by (23) the proof is complete. □

3. The optimization problem

In this section, we assume that K is a generalized polynomial in the basis Φ ; that is,

$$K(x, y) = \sum_{i,j=1}^r k_{ij} \phi_i(x) \phi_j(y), \quad x, y \in \mathbb{T}^d. \tag{26}$$

where $\mathbf{K} = (k_{ij})_{i,j=1}^r \in M_{r,r}(\mathbb{R})$ is a matrix of coefficients. For such K , (3) takes form

$$\begin{cases} -\partial_t u + H(x, \nabla u) = \sum_{i=1}^r \phi_i(x) \sum_{j=1}^r k_{ij} \int_{\mathbb{T}^d} \phi_j(y) m(y, t) \, dy, \\ \partial_t m - \text{div}(m \nabla_p H(x, \nabla u)) = 0, \quad (x, t) \in \mathbb{T}^d \times [0, 1], \\ m(x, 0) = M(x), \quad u(x, 1) = U(x), \quad x \in \mathbb{T}^d. \end{cases} \tag{27}$$

Our main observation is the following theorem.

Theorem 3.1. (i) *A pair (u, m) is a solution of (27) if and only if $(u, m) = (u_{a^*}, m_{a^*})$ for some $a^* \in C([0, 1]; \mathbb{R}^r)$ such that*

$$a^* = \mathbf{K} \partial_a G(a^*). \tag{28}$$

- (ii) If \mathbf{K} is positive-definite then (28) is equivalent to finding a 0 of a monotone operator $a \mapsto \mathbf{K}^{-1}a - \partial_a G(a)$, $a \in C([0, 1]; \mathbb{R}^r)$.
- (iii) Additionally, if \mathbf{K} is symmetric, (28) is equivalent to the convex optimization problem

$$\begin{aligned} & \inf_{a \in C([0, 1]; \mathbb{R}^r)} \frac{1}{2} \langle \mathbf{K}^{-1}a, a \rangle - G(a) \\ &= \inf_{a \in C([0, 1]; \mathbb{R}^r)} \frac{1}{2} \langle \mathbf{K}^{-1}a, a \rangle - \int_{\mathbb{T}^d} u_a(x, 0)M(x) dx. \end{aligned} \tag{29}$$

Proof. Items (ii) and (iii) follow immediately from (i) by the concavity of G . Thus, we just prove (i).

By Theorem 2.2 (27) admits a solution (u, m) . Furthermore, define a^* as

$$a_i^*(t) = \sum_{j=1}^r k_{ij} \int_{\mathbb{T}^d} \phi_j(y)m(y, t) dy, \quad t \in [0, 1]. \tag{30}$$

Then $a^* \in C([0, 1]; \mathbb{R}^r)$, and by the definition of u_a and m_a we have that $(u, m) = (u_{a^*}, m_{a^*})$. Hence, by Theorem 2.3, we have that

$$\partial_{a_i} G(a^*) = \int_{\mathbb{T}^d} \phi_i(x)m(x, \cdot) dx, \quad 1 \leq i \leq r.$$

Consequently, from (30) obtain

$$a_i^* = \sum_{j=1}^r k_{ij} \partial_{a_j} G(a^*). \tag{31} \quad \square$$

Remark 3.2. The optimization problem (29) is equivalent to the optimal control of Hamilton–Jacobi PDE pointed out in [32] (equations (58)–(59) in Section 2.6). One can think of (29) as (58)–(59) of [32] written in Fourier coordinates.

4. Approximating the kernel

In this section, we show that one can construct suitable approximations for an arbitrary K . We begin by a simple lemma.

Lemma 4.1. *Suppose that K is given by (26). Then K is positive semi-definite if and only if $\mathbf{K} = (k_{ij})_{ij=1}^r$ is positive semi-definite.*

Proof. Fix an arbitrary $(\zeta_i)_{i=1}^r \in \mathbb{R}^r$. Then there exists a unique $(\lambda_i)_{i=1}^r \in \mathbb{R}^r$ such that

$$\zeta_i = \sum_{j=1}^r \lambda_j \int_{\mathbb{T}^d} \phi_i(x) \phi_j(x) dx, \quad 1 \leq i \leq r,$$

because $\{\phi_i\}$ are linearly independent. Therefore, for

$$f = \sum_{j=1}^r \lambda_j \phi_j$$

we have that

$$\zeta_i = \int_{\mathbb{T}^d} f(x) \phi_i(x) dx, \quad 1 \leq i \leq r.$$

Hence,

$$\int_{\mathbb{T}^d \times \mathbb{T}^d} K(x, y) f(x) f(y) dx dy = \sum_{i,j=1}^r k_{ij} \zeta_i \zeta_j,$$

that yields the proof. □

Now, we fix our basis to be the trigonometric one:

$$\phi_\alpha(x) = e^{2i\pi\alpha \cdot x}, \quad x \in \mathbb{T}^d, \alpha \in \mathbb{Z}^d. \tag{31}$$

Remark 4.2. Unlike in (14), here it is more practical to use multi-dimensional indexes to enumerate the trigonometric functions in higher dimensions. Additionally, it is more economical in terms of notation to use the complex-valued trigonometric functions. Nevertheless, our discussion is always about real valued K , and the reader can think of the end results as expansions in terms of $\{\cos(2\pi\alpha \cdot x), \sin(2\pi\alpha \cdot x)\}_{\alpha \in \mathbb{Z}^d}$.

For $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{Z}^d$, we denote by

$$|\alpha| = (|\alpha_1|, |\alpha_2|, \dots, |\alpha_d|),$$

and for $\alpha, r \in \mathbb{Z}^d$

$$\alpha \leq r \iff \alpha_j \leq r_j, \quad 1 \leq j \leq d.$$

For $r_1, r_2 \in \mathbb{N}_0^d$ we denote by

$$K_{r_1 r_2}(x, y) = \sum_{|\alpha| \leq r_1, |\beta| \leq r_2} \hat{K}_{\alpha\beta} e^{2i\pi(\alpha \cdot x + \beta \cdot y)}, \quad x, y \in \mathbb{T}^d,$$

where

$$\hat{K}_{\alpha\beta} = \int_{\mathbb{T}^d} K(x, y) e^{-2i\pi(\alpha \cdot x + \beta \cdot y)} dx dy, \quad \alpha, \beta \in \mathbb{Z}^d.$$

Furthermore, for $r_1, r_2 \in \mathbb{N}_0^d$ we denote by

$$\Sigma_{r_1 r_2}(x, y) = \frac{1}{\prod_{j=1}^d (1 + r_{1j})(1 + r_{2j})} \sum_{|\alpha| \leq r_1, |\beta| \leq r_2} K_{r_1 r_2}(x, y), \quad x, y \in \mathbb{T}^d.$$

Remark 4.3. The function $K_{r_1 r_2}$ is the rectangular partial Fourier sum of K . Correspondingly, $\Sigma_{r_1 r_2}$ is the rectangular Fejér average of K . Additionally, if K is real valued then $K_{r_1 r_2}$ and $\Sigma_{r_1 r_2}$ are real valued for any $r_1, r_2 \in \mathbb{N}_0^d$.

Proposition 4.4. *If K is positive semi-definite (symmetric) then, K_{rr} and Σ_{rr} are also positive semi-definite (symmetric) for all $r \in \mathbb{N}_0^d$. Moreover,*

$$\lim_{\min_j r_j \rightarrow \infty} \|\Sigma_{rr} - K\|_{C^2(\mathbb{T}^d \times \mathbb{T}^d)} = 0, \tag{32}$$

Additionally, if $K \in C^3(\mathbb{T}^d \times \mathbb{T}^d)$ then

$$\lim_{\min_j r_j \rightarrow \infty} \|K_{rr} - K\|_{C^2(\mathbb{T}^d \times \mathbb{T}^d)} = 0. \tag{33}$$

Proof. The convergence properties (32), (33) are classical results in Fourier analysis. Thus, we will just prove that K_{rr} and Σ_{rr} are positive semi-definite (symmetric). For that, we use the representation formulas

$$\begin{aligned} K_{rr}(x, y) &= \int_{\mathbb{T}^d \times \mathbb{T}^d} K(z, w) D_{rr}(x - z, y - w) dz dw, \\ \Sigma_{rr}(x, y) &= \int_{\mathbb{T}^d \times \mathbb{T}^d} K(z, w) F_{rr}(x - z, y - w) dz dw, \quad x, y \in \mathbb{T}^d, \end{aligned}$$

where D_{rr} and F_{rr} are, respectively, the $2d$ -dimensional rectangular Dirichlet and Fejér kernels. A crucial feature of D_{rr} and F_{rr} is that they are symmetric and decompose into lower dimensional kernels:

$$D_{rr}(z, w) = D_r(z)D_r(w), \quad F_{rr}(z, w) = F_r(z)F_r(w), \quad z, w \in \mathbb{T}^d,$$

where D_r and F_r are the corresponding d -dimensional kernels. In particular, K_{rr} , Σ_{rr} are symmetric if K is such. Furthermore, for an arbitrary $f \in L^\infty(\mathbb{T}^d)$ we have that

$$\begin{aligned} & \int_{\mathbb{T}^d \times \mathbb{T}^d} K_{rr}(x, y) f(x) f(y) \, dx \, dy \\ &= \int_{\mathbb{T}^d \times \mathbb{T}^d} K(z, w) \, dz \, dw \int_{\mathbb{T}^d \times \mathbb{T}^d} f(x) f(y) D_{rr}(x - z, y - w) \, dx \, dy \\ &= \int_{\mathbb{T}^d \times \mathbb{T}^d} K(z, w) \, dz \, dw \int_{\mathbb{T}^d \times \mathbb{T}^d} f(x) f(y) D_r(x - z) D_r(y - w) \, dx \, dy \\ &= \int_{\mathbb{T}^d \times \mathbb{T}^d} K(z, w) f_r(z) f_r(w) \, dz \, dw \geq 0. \end{aligned}$$

Thus, K_{rr} is positive semi-definite if K is such. The proof for Σ_{rr} is identical.

Remark 4.5. By Proposition 4.4, kernels K_{rr} , Σ_{rr} are positive semi-definite. Therefore, their coefficients matrices with respect to basis $\{\cos(2\pi\alpha \cdot x), \sin(2\pi\alpha \cdot x)\}$ are also positive semi-definite by Lemma 4.1. Nevertheless, to take full advantage of Theorem 3.1 one would need these matrices to be positive definite (invertible). To solve this problem one can add εI regularization term, where I is the identity matrix of the suitable dimension and $\varepsilon > 0$ is a small constant. However, as discussed below, this regularization is not necessary for translation invariant kernels.

Suppose that

$$K(x, y) = \eta(x - y), \quad x, y \in \mathbb{T}^d,$$

where η is a periodic function. Then, we have

$$\begin{aligned} & \int_{\mathbb{T}^d} K(x, y) \cos(2\pi\alpha \cdot y) \, dy \\ &= \int_{\mathbb{T}^d} \eta(x - y) \cos(2\pi\alpha \cdot y) \, dy = \int_{\mathbb{T}^d} \eta(y) \cos(2\pi\alpha \cdot (x - y)) \, dy \\ &= \cos(2\pi\alpha \cdot x) \int_{\mathbb{T}^d} \eta(y) \cos(2\pi\alpha \cdot y) \, dy \\ &\quad + \sin(2\pi\alpha \cdot x) \int_{\mathbb{T}^d} \eta(y) \sin(2\pi\alpha \cdot y) \, dy. \end{aligned} \tag{34}$$

Similarly, we obtain that

$$\begin{aligned}
& \int_{\mathbb{T}^d} K(x, y) \sin(2\pi\alpha \cdot y) dy \\
&= \int_{\mathbb{T}^d} \eta(x - y) \sin(2\pi\alpha \cdot y) dy = \int_{\mathbb{T}^d} \eta(y) \sin(2\pi\alpha \cdot (x - y)) dy \\
&= \sin(2\pi\alpha \cdot x) \int_{\mathbb{T}^d} \eta(y) \cos(2\pi\alpha \cdot y) dy \\
&\quad - \cos(2\pi\alpha \cdot x) \int_{\mathbb{T}^d} \eta(y) \sin(2\pi\alpha \cdot y) dy.
\end{aligned} \tag{35}$$

Therefore, we have that

$$\begin{aligned}
& \int_{\mathbb{T}^d} K(x, y) \cos(2\pi\alpha \cdot x) \cos(2\pi\alpha \cdot y) dx dy = \int_{\mathbb{T}^d} \eta(y) \cos(2\pi\alpha \cdot y) dy, \\
& \int_{\mathbb{T}^d} K(x, y) \sin(2\pi\alpha \cdot x) \cos(2\pi\alpha \cdot y) dx dy = \int_{\mathbb{T}^d} \eta(y) \sin(2\pi\alpha \cdot y) dy, \\
& \int_{\mathbb{T}^d} K(x, y) \cos(2\pi\alpha \cdot x) \sin(2\pi\alpha \cdot y) dx dy = - \int_{\mathbb{T}^d} \eta(y) \sin(2\pi\alpha \cdot y) dy, \\
& \int_{\mathbb{T}^d} K(x, y) \sin(2\pi\alpha \cdot x) \sin(2\pi\alpha \cdot y) dx dy = \int_{\mathbb{T}^d} \eta(y) \cos(2\pi\alpha \cdot y) dy.
\end{aligned}$$

Hence, the coefficients matrices of partial Fourier sums (and their linear combinations) of K consist of 2×2 blocks that correspond to expansion terms with a frequency $\alpha \in \mathbb{Z}^d$; that is,

$$\Delta_\alpha = \begin{pmatrix} \int_{\mathbb{T}^d} \eta(y) \cos(2\pi\alpha \cdot y) dy & \int_{\mathbb{T}^d} \eta(y) \sin(2\pi\alpha \cdot y) dy \\ - \int_{\mathbb{T}^d} \eta(y) \sin(2\pi\alpha \cdot y) dy & \int_{\mathbb{T}^d} \eta(y) \cos(2\pi\alpha \cdot y) dy \end{pmatrix}. \tag{36}$$

Thus, the coefficient matrix will be degenerate if $\det(\Delta_\alpha) = 0$ for some α . But we have that

$$\det(\Delta_\alpha) = \left(\int_{\mathbb{T}^d} \eta(y) \cos(2\pi\alpha \cdot y) dy \right)^2 + \left(\int_{\mathbb{T}^d} \eta(y) \sin(2\pi\alpha \cdot y) dy \right)^2.$$

Hence, $\det(\Delta_\alpha) = 0$ if and only if $\Delta_\alpha = 0$ or, equivalently, there are no expansion terms with frequency α . But then, we can simply ignore these terms in our basis and obtain a non-degenerate matrix.

Moreover, to invert the coefficients matrix one just has to invert the 2×2 blocks. Additionally, if K is symmetric; that is, $\eta(y) = \eta(-y)$, we get that

$$\int_{\mathbb{T}^d} \eta(y) \sin(2\pi\alpha \cdot y) dy = 0, \quad \forall \alpha \in \mathbb{Z}^d.$$

Hence, the coefficient matrices are simply diagonal. Therefore, we have proved the following proposition.

Proposition 4.6. *If K is translation invariant then all partial Fourier sums of K and their linear combinations, such as K_r and Σ_r , contain only $\cos(2\pi\alpha \cdot x) \cos(2\pi\alpha \cdot y)$, $\cos(2\pi\alpha \cdot x) \sin(2\pi\alpha \cdot y)$, $\sin(2\pi\alpha \cdot x) \cos(2\pi\alpha \cdot y)$, $\sin(2\pi\alpha \cdot x) \sin(2\pi\alpha \cdot y)$ expansion terms. Therefore, coefficient matrices of such approximations with respect to trigonometric basis consist of 2×2 blocks that are multiples of Δ_α in (36). Additionally, if K is symmetric these coefficient matrices are diagonal.*

Remark 4.7. In general, if $\{\phi_1, \phi_2, \dots, \phi_r, \dots\}$ is an orthonormal basis consisting of eigenfunctions of Hilbert–Schmidt integral operator $f(\cdot) \mapsto \int_{\mathbb{T}^d} K(\cdot, y)f(y) dy$; that is,

$$\int_{\mathbb{T}^d} K(x, y)\phi_\alpha(y) dx = \lambda_\alpha\phi_\alpha(x), \quad x \in \mathbb{T}^d, \alpha \in \mathbb{N},$$

for some $\{\lambda_\alpha\} \subset \mathbb{R}$. Then, one has that

$$k_{\alpha\beta} = \int_{\mathbb{T}^d} K(x, y)\phi_\alpha(x)\phi_\beta(y) dx dy = \lambda_\beta\delta_{\alpha\beta}.$$

Consequently, for arbitrary $I \subset \mathbb{N} \times \mathbb{N}$ we have that

$$K_I(x, y) = \sum_{(\alpha, \beta) \in I} k_{\alpha\beta}\phi_\alpha(x)\phi_\beta(y) = \sum_{(\alpha, \alpha) \in I} \lambda_\alpha\phi_\alpha(x)\phi_\alpha(y).$$

Therefore, all partial Fourier sums of K in basis $\{\phi_\alpha(x)\phi_\beta(y)\}$ contain only terms $\phi_\alpha(x)\phi_\alpha(y)$ and yield diagonal coefficient matrices consisting of corresponding eigenvalues of the Hilbert–Schmidt integral operator.

In general, it is not easy to calculate the eigenfunctions of a given Hilbert–Schmidt integral operator. Nevertheless, as we saw above, for translation invariant symmetric periodic K these eigenfunctions are precisely the trigonometric functions.

5. A numerical method

In this section we propose a numerical method to solve (3) for a symmetric and positive semi-definite K . We assume that an approximation K_r of the form (26) is already constructed with a symmetric and positive definite \mathbf{K} . Thus, we devise an algorithm for the solution of (27).

By Theorem 3.1 we have that (27) is equivalent to (29). Therefore, in what follows, we present a suitable discretization of (29). We rewrite latter as

$$\inf_{a \in C([0, 1]; \mathbb{R}^r)} S(a), \tag{37}$$

where

$$S(a) = \frac{1}{2} \langle \mathbf{J}a, a \rangle - G(a),$$

and $\mathbf{J} = \mathbf{K}^{-1}$.

5.1. Discretization of the u_a . We start with the discretization of u_a . For that, we discretize the representation formula (16). We can rewrite latter as

$$u_a(x, 0) = \inf_{\mathbf{b}} \int_0^1 L_a(\mathbf{x}(s), \mathbf{b}(s), s) ds + U(\mathbf{x}(1)), \tag{38}$$

where \mathbf{x} satisfies the following controlled ODE

$$\dot{\mathbf{x}}(s) = \mathbf{b}(s), \quad \mathbf{x}(0) = x, \quad s \in [0, 1]. \tag{39}$$

Recall that

$$L_a(x, b, s) = L(x, b) + \sum_{k=1}^r a_k(s) \phi_k(x), \quad (x, b, s) \in \mathbb{T}^d \times \mathbb{R}^d \times [0, 1].$$

We choose a uniform discretization of the time interval:

$$0 = s_0 < s_1 < s_2 < \dots < s_N = 1,$$

with a step size $h_t = \frac{1}{N}$, hence $s_i = ih_t = \frac{i}{N}$. We denote the values of \mathbf{x} and \mathbf{b} at time s_i by $\mathbf{x}(s_i) = x_i$, $\mathbf{b}(s_i) = b_i$. Using a backward Euler discretization of (39) we have

$$u_i = \frac{x_i - x_{i-1}}{h_t}, \quad i \in \{1, \dots, N\}.$$

Discretizing the integral (38) with a right point quadrature rule and using the above discretization we get

$$\begin{cases} [u_a](x, 0) = \inf_{\{x_i\}_0^N} h_t \sum_{i=1}^N L_a\left(x_i, \frac{x_i - x_{i-1}}{h_t}, s_i\right) + U(x_N), \\ \text{subject to: } x_0 = x. \end{cases} \tag{40}$$

5.2. Discretization of G . We start by discretizing the initial measure M using a convex combination of Dirac δ distributions. Denoting the discretized measure $[M]$, we have

$$[M] = \sum_{\alpha=1}^Q c_\alpha \delta_{y_\alpha}$$

or, in the distributional sense,

$$\int_{\mathbb{T}^d} \psi(y) d[M](y) = \sum_{\alpha=1}^Q c_\alpha \psi(y_\alpha), \quad \psi \in C(\mathbb{T}^d), \tag{41}$$

for some $\{y_\alpha\}_{\alpha=1}^Q \subset \mathbb{T}^d$ and $\{c_\alpha \geq 0\}_{\alpha=1}^Q$ such that $\sum_{\alpha=1}^Q c_\alpha = 1$. Then, G is discretized as follows

$$[G](a) = \sum_{\alpha=1}^Q c_\alpha [u_a](y_\alpha, 0). \tag{42}$$

5.3. Discretization of S . Now, we discretize (37). We first discretize a_k -s by taking their values at times s_i , that we denote by:

$$[a]_k = (a_k(s_0), \dots, a_k(s_N)) = (a_{k0}, \dots, a_{kN}), \quad k = 1, 2, \dots, r.$$

Recall that

$$\langle \mathbf{J}a, a \rangle = \sum_{k,l=1}^r \mathbf{J}_{kl} \int_0^1 a_k(s) a_l(s) ds.$$

We discretize this previous quadratic form by a simple right point quadrature rule.

$$[\langle \mathbf{J}a, a \rangle] = h_t \sum_{k,l=1}^r \mathbf{J}_{kl} \sum_{i=1}^N a_{ki} a_{li}.$$

So the discretization of S is

$$\begin{aligned} [S](a) &= \frac{h_t}{2} \sum_{k,l=1}^r \mathbf{J}_{kl} \sum_{i=1}^N a_{ki} a_{li} - [G](a) \\ &= \frac{h_t}{2} \sum_{k,l=1}^r \mathbf{J}_{kl} \sum_{i=1}^N a_{ki} a_{li} - \sum_{\alpha=1}^Q c_\alpha [u_a](y_\alpha, 0), \end{aligned} \tag{43}$$

where we used (42). Therefore, the discretization of (37) is

$$\begin{aligned} \inf_{\{a_{ki}\}} [S](a) = \inf_{\{a_{ki}\}} \sup_{\{x_{\alpha i}: x_{\alpha 0} = y_{\alpha}\}} & \frac{h_t}{2} \sum_{k,l=1}^r \mathbf{J}_{kl} \sum_{i=1}^N a_{ki} a_{li} - h_t \sum_{\alpha=1}^Q \sum_{i=1}^N c_{\alpha} L\left(x_{\alpha i}, \frac{x_{\alpha i} - x_{\alpha(i-1)}}{h_t}\right) \\ & - h_t \sum_{\alpha=1}^Q \sum_{i=1}^N \sum_{k=1}^r c_{\alpha} a_{ki} \phi_k(x_{\alpha i}) - \sum_{\alpha=1}^Q c_{\alpha} U(x_{\alpha N}). \end{aligned} \tag{44}$$

5.4. Primal-dual hybrid-gradient method. Now, we specify the Lagrangian to be quadratic and devise a primal-dual hybrid-gradient algorithm [20] to solve (37). More precisely, we assume that

$$L(x, b) = \frac{|b|^2}{2}, \quad (x, b) \in \mathbb{T}^d \times \mathbb{R}^d,$$

and therefore (44) becomes

$$\begin{aligned} \inf_{\{a_{ki}\}} [S](a) = \inf_{\{a_{ki}\}} \sup_{\{x_{\alpha i}: x_{\alpha 0} = y_{\alpha}\}} & \frac{h_t}{2} \sum_{k,l=1}^r \mathbf{J}_{kl} \sum_{i=1}^N a_{ki} a_{li} - \sum_{\alpha=1}^Q \sum_{i=1}^N c_{\alpha} \frac{|x_{\alpha i} - x_{\alpha(i-1)}|^2}{2h_t} \\ & - h_t \sum_{\alpha=1}^Q \sum_{i=1}^N \sum_{k=1}^r c_{\alpha} a_{ki} \phi_k(x_{\alpha i}) - \sum_{\alpha=1}^Q c_{\alpha} U(x_{\alpha N}). \end{aligned} \tag{45}$$

Now, we describe the algorithm. For each iteration time $v \geq 0$ we have three groups of variables: $a^v = \{a_{ki}^v\}_{k,i=1,1}^{r,N}$, $x^v = \{x_{\alpha i}^v\}_{\alpha,i=1,0}^{Q,N}$, and $z^v = \{z_{\alpha i}^v\}_{\alpha,i=1,0}^{Q,N}$. Furthermore, we fix $\lambda, \omega > 0$ that are proximal step parameters for variables a and x , respectively. Additionally, we take $0 \leq \theta \leq 1$.

Step 1. Given a^v, x^v, z^v the first step of the algorithm is to solve the proximal problem

$$\begin{aligned} \inf_{\{a_{ki}\}} & \frac{h_t}{2} \sum_{k,l=1}^r \mathbf{J}_{kl} \sum_{i=1}^N a_{ki} a_{li} - \sum_{\alpha=1}^Q \sum_{i=1}^N c_{\alpha} \frac{|z_{\alpha i}^v - z_{\alpha(i-1)}^v|^2}{2h_t} \\ & - h_t \sum_{\alpha=1}^Q \sum_{i=1}^N \sum_{k=1}^r c_{\alpha} a_{ki} \phi_k(z_{\alpha i}^v) - \sum_{\alpha=1}^Q c_{\alpha} U(z_{\alpha N}^v) + \frac{1}{2\lambda} \sum_{k=1}^r \sum_{i=1}^N (a_{ki} - a_{ki}^v)^2, \end{aligned}$$

that is equivalent to

$$\inf_{\{a_{ki}\}} \frac{h_t}{2} \sum_{k,l=1}^r \mathbf{J}_{kl} \sum_{i=1}^N a_{ki} a_{li} - h_t \sum_{\alpha=1}^Q \sum_{i=1}^N \sum_{k=1}^r c_{\alpha} a_{ki} \phi_k(z_{\alpha i}^v) + \frac{1}{2\lambda} \sum_{k=1}^r \sum_{i=1}^N (a_{ki} - a_{ki}^v)^2.$$

Thus, we obtain the following update of the a -variable.

$$\begin{pmatrix} a_{1i}^{v+1} \\ a_{2i}^{v+1} \\ \vdots \\ a_{ri}^{v+1} \end{pmatrix} = (\lambda h_t \mathbf{J} + \text{Id}_r)^{-1} \begin{pmatrix} a_{1i}^v + \lambda h_t \sum_{\alpha=1}^Q c_\alpha \phi_1(z_{\alpha i}^v) \\ a_{2i}^v + \lambda h_t \sum_{\alpha=1}^Q c_\alpha \phi_2(z_{\alpha i}^v) \\ \vdots \\ a_{ri}^v + \lambda h_t \sum_{\alpha=1}^Q c_\alpha \phi_r(z_{\alpha i}^v) \end{pmatrix}, \quad 1 \leq i \leq N. \quad (46)$$

Remark 5.1. Although the number of variables $\{a_{ki}\}_{k,i=1,1}^{r,N}$ is $r \times N$, the calculations of $\{a_{ki}\}$ for different i -s are mutually independent. Therefore, the only complexity is in the inversion of an $r \times r$ matrix $\lambda \sigma \mathbf{J} + \text{Id}_r$ that can be computed beforehand and used throughout the scheme. Moreover, as seen in Section 4, translation invariant symmetric kernels yield diagonal matrices that extremely simplify the calculations.

Step 2. Given a^{v+1}, x^v, z^v we update x -variable by solving the proximal problem

$$\begin{aligned} \inf_{\{x_{\alpha i}: x_{\alpha 0}=y_\alpha\}} & \sum_{\alpha=1}^Q \sum_{i=1}^N c_\alpha \frac{|x_{\alpha i} - x_{\alpha(i-1)}|^2}{2h_t} + h_t \sum_{\alpha=1}^Q \sum_{i=1}^N \sum_{k=1}^r c_\alpha a_{ki}^{v+1} \phi_k(x_{\alpha i}) \\ & + \sum_{\alpha=1}^Q c_\alpha U(x_{\alpha N}) + \frac{1}{2\omega} \sum_{\alpha=1}^Q \sum_{i=1}^N |x_{\alpha i} - x_{\alpha i}^v|^2. \end{aligned}$$

Solving this previous problem may be a costly operation. Hence, we just perform a one step gradient descent. Therefore, we obtain

$$\begin{aligned} x_{\alpha 1}^{v+1} &= x_{\alpha 1}^v - \frac{\omega c_\alpha}{h_t} (x_{\alpha 1} - y_\alpha) - \frac{\omega c_\alpha}{h_t} (x_{\alpha 1} - x_{\alpha 2}) - \omega c_\alpha h_t \sum_{k=1}^r a_{k1}^{v+1} \nabla \phi_k(x_{\alpha 1}), \\ x_{\alpha i}^{v+1} &= x_{\alpha i}^v - \frac{\omega c_\alpha}{h_t} (x_{\alpha i} - x_{\alpha(i-1)}) - \frac{\omega c_\alpha}{h_t} (x_{\alpha i} - x_{\alpha(i+1)}), \\ & - \omega c_\alpha h_t \sum_{k=1}^r a_{ki}^{v+1} \nabla \phi_k(x_{\alpha i}), \quad 1 \leq i \leq N - 1, \\ x_{\alpha N}^{v+1} &= x_{\alpha N}^v - \frac{\omega c_\alpha}{h_t} (x_{\alpha N} - x_{\alpha(N-1)}) - \omega c_\alpha \nabla U(x_{\alpha N}) \\ & - \omega c_\alpha h_t \sum_{k=1}^r a_{kN}^{v+1} \nabla \phi_k(x_{\alpha N}). \end{aligned} \quad (47)$$

Step 3. In the final step we update the z -variable by

$$z_{\alpha i}^{v+1} = x_{\alpha i}^{v+1} + \theta(x_{\alpha i}^{v+1} - x_{\alpha i}^v), \quad 1 \leq \alpha \leq Q, 1 \leq i \leq N. \quad (48)$$

Remark 5.2. Note that the updates for $\{x_{zi}\}$, $\{z_{zi}\}$ variables are mutually independent for different α -s. Therefore, our α -updates are parallel in time, and x , z -updates are parallel in space.

Remark 5.3. Strictly speaking, one cannot simply apply the primal-dual hybrid gradient method to (45) because the coupling between a and x is not bilinear, and there is no concavity in x . Nevertheless, our calculations always yield solid results. Therefore, there is a natural problem of rigorously understanding the convergence properties of the aforementioned algorithm. We plan to address this problem in our future work.

6. Numerical examples

In this section, we present several numerical experiments. We first look into one-dimensional case, in Section 6.1, and after we consider the two-dimensional case, in Section 6.2.

For our calculations, we choose the periodic Gaussian kernel that is given by

$$K_{\sigma,\mu}^d(x, y) = \prod_{i=1}^d K_{\sigma,\mu}^1(x_i, y_i), \quad x, y \in \mathbb{T}^d, \tag{49}$$

where

$$K_{\sigma,\mu}^1(x, y) = \frac{\mu}{\sqrt{2\pi(\frac{\sigma}{2})^2}} \sum_{k=-\infty}^{\infty} e^{-(x-y-k)^2/2(\sigma/2)^2}, \quad x, y \in \mathbb{T}, \tag{50}$$

and $\sigma, \mu > 0$ are given parameters. Here, σ models how spread is the kernel. The smaller σ the more weight agents assign to their immediate neighbors – this translates into crowd-aversion in the close neighborhood only. Furthermore, μ is the total weight of the agents. Therefore, μ measures how sensitive is a generic agent to the total population, the bigger the more averse is the agent to others. As we observe in the numerical experiments, the less σ and the larger μ the more separated are the agents. This phenomenon was also observed in [7].

Throughout the section we denote by

$$\phi_k(x) = \begin{cases} 1, & \text{if } k = 1, \\ \sqrt{2} \cos \pi(k - 1)x, & \text{if } k \text{ is odd, and } k > 0, \\ \sqrt{2} \sin \pi kx, & \text{if } k \text{ is even, } x \in \mathbb{T}. \end{cases} \tag{51}$$

Therefore, we have

$$\{\phi_1, \phi_2, \phi_3, \dots\} = \{1, \sqrt{2} \sin 2\pi x, \sqrt{2} \cos 2\pi x, \dots\}.$$

6.1. One-dimensional examples. For all simulations we use the same initial-terminal conditions

$$M(x) = \frac{1}{6} + \frac{5}{3} \sin^2 \pi x, \quad U(x) = 1 + \sin\left(4\pi x + \frac{\pi}{2}\right), \quad x \in \mathbb{T},$$

that are depicted in Figure 1.

We also use the same time and space discretization for all one dimensional experiments, and the same parameters for the numerical scheme. We discretize the time using a step size $\Delta t = \frac{1}{N}$. For the discretization of M we use

$$y_\alpha = \frac{\alpha}{Q+1}, \quad c_\alpha = \frac{M(y_\alpha)}{\sum_{\beta=1}^Q M(y_\beta)} \quad 1 \leq \alpha \leq Q.$$

We choose $N = 20$, $Q = 50$ and use eight basis functions, $r = 8$. Additionally, we set the numerical scheme parameters to $\lambda = 3$, $\omega = \frac{1}{12}$ and $\theta = 1$.

Remark 6.1. For the standard primal-dual hybrid gradient method, one must have $\omega\lambda < \frac{1}{A^2}$, where A is the norm of the bilinear-form matrix. As we mentioned in Remark 5.3, here we do not have a bilinear coupling between a and x . Thus, we estimate A by an upper bound on the (l_2, l_2) Lipschitz norm of the mapping

$$F_{ki}(x) = h_t \sum_{\alpha=1}^Q c_\alpha \phi_k(x_{\alpha i}), \quad 1 \leq l \leq r, 1 \leq i \leq N.$$

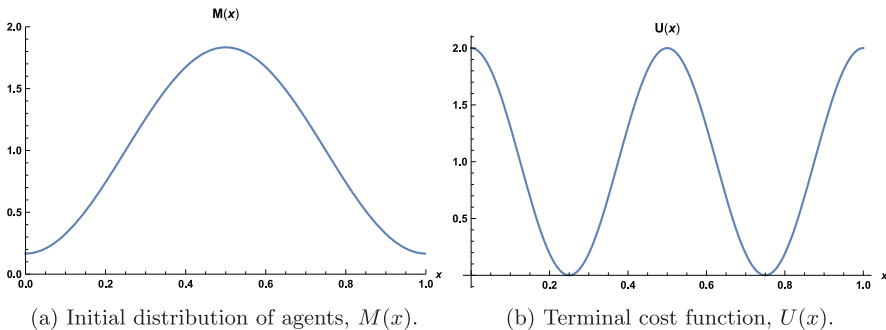


Figure 1. Initial distribution of agents, $M(x)$.

More precisely, we have that

$$\begin{aligned}
 \text{Lip}(F)^2 &= \sup_{\{x_{\beta j}\}} \sup_{\|w_{\beta j}\|_2 \leq 1} \sum_{k,i} \left(\sum_{\beta,j} \frac{\partial F_{ki}}{\partial x_{\beta j}} w_{\beta j} \right)^2 \\
 &= \sup_{\{x_{\beta j}\}} \sup_{\|w_{\beta j}\|_2 \leq 1} \sum_{k,i} \left(\sum_{\beta} h_t c_{\beta} \nabla \phi_k(x_{\beta i}) w_{\beta i} \right)^2 \\
 &\leq h_t^2 \sup_{\{x_{\beta j}\}} \sup_{\|w_{\beta j}\|_2 \leq 1} \sum_{k,i} \left(\sum_{\beta} c_{\beta}^2 \|\nabla \phi_k(x_{\beta i})\|_2^2 \cdot \sum_{\beta} w_{\beta i}^2 \right) \\
 &\leq h_t^2 \sup_{\|w_{\beta j}\|_2 \leq 1} \sum_{k,i} \text{Lip}(\phi_k)^2 \left(\sum_{\beta} c_{\beta}^2 \cdot \sum_{\beta} w_{\beta i}^2 \right) \\
 &= h_t^2 \sup_{\|w_{\beta j}\|_2 \leq 1} \sum_k \text{Lip}(\phi_k)^2 \sum_{\beta} c_{\beta}^2 \sum_{\beta,i} w_{\beta i}^2 \\
 &= h_t^2 \sum_k \text{Lip}(\phi_k)^2 \sum_{\beta} c_{\beta}^2.
 \end{aligned}$$

Thus, we take

$$A^2 = h_t^2 \sum_{k=1}^r \text{Lip}(\phi_k)^2 \sum_{\beta=1}^Q c_{\beta}^2.$$

The trigonometric expansion of $K_{\sigma,\mu}^1$ is given by

$$K_{\sigma,\mu}^1(x, y) = \mu \left(1 + 2 \sum_{n=1}^{\infty} e^{-(\pi n \sigma)^2 / 2} \cos 2\pi n(x - y) \right), \quad x, y \in \mathbb{T}, \tag{52}$$

or

$$K_{\sigma,\mu}^1(x, y) = \sum_{k=1}^{\infty} \mu e^{-(1/2)(\pi \sigma [k/2])^2} \phi_k(x) \phi_k(y), \quad x, y \in \mathbb{T}, \tag{53}$$

in our notation. Therefore, for a given r , the matrices \mathbf{K} , \mathbf{J} are given by

$$\begin{aligned}
 \mathbf{K} &= \text{diag}(\mu e^{-(1/2)(\pi \sigma [k/2])^2})_{k=1}^r, \\
 \mathbf{J} &= \text{diag}(\mu^{-1} e^{(1/2)(\pi \sigma [k/2])^2})_{k=1}^r.
 \end{aligned} \tag{54}$$

In Figure 2 we plot the Gaussian kernels we used, for $r = 8$ and different values of μ and σ . We see the influence of these values in Figure 3. In the first column of Figure 3 we compare the results regarding for different values of μ and σ .

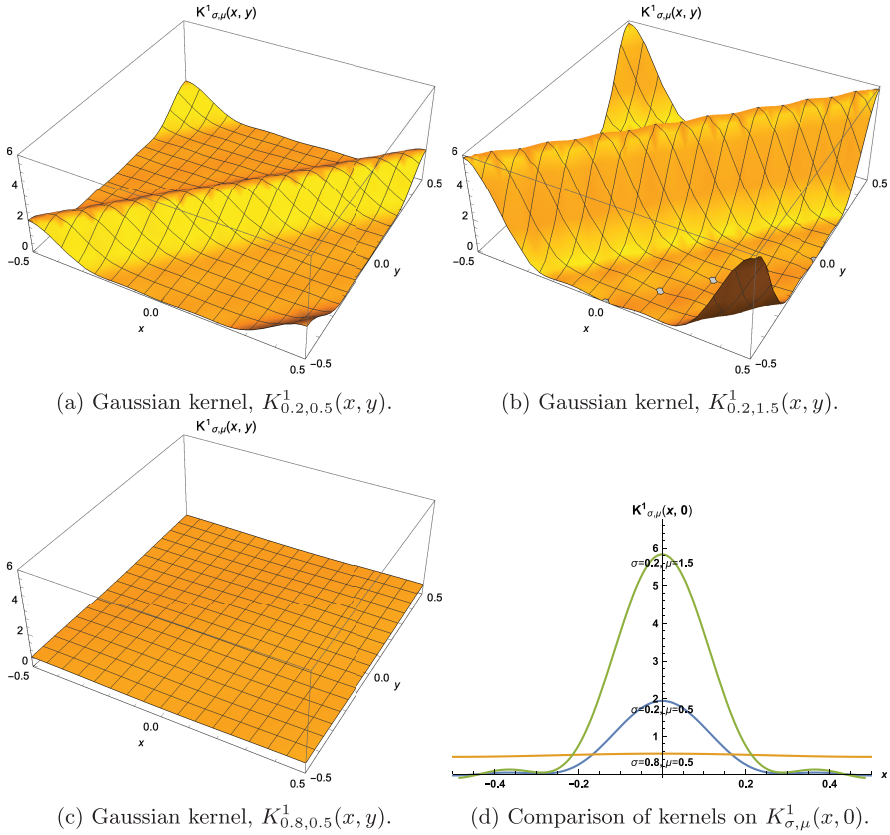


Figure 2. Plots of the three Gaussian kernels in (a)–(c), and a comparison of their sections in (d).

Comparing the first and the second columns of Figure 3, we see that the trajectories of the agents in the first column are closer than in the second one. This is due to the fact that $\mu = 0.5$ in the first kernel and $\mu = 1.5$ in the second one, hence the second kernel (higher value of μ) penalizes more high density of agents. Therefore, the agents spread out more before the final time when they converge to the points of low-cost near minima of the terminal cost function, U , see Figure 1(b).

In the last column the value of $\sigma = 0.8$ is higher, this means that agents are indifferent to the distances between them – they just feel the total mass. Hence, they minimize the travel distances from initial positions to low-cost locations of U ignoring the population density. In fact, in this case $K_{\sigma,\mu}^1 \approx \mu$, and therefore $\int_{\mathbb{T}} K_{\mu,\sigma}^1(x, y)m(y, t) dy \approx \mu$. Thus, in this case (3) approximates a decoupled system of Hamilton–Jacobi and Fokker–Planck equations. But the optimal trajec-

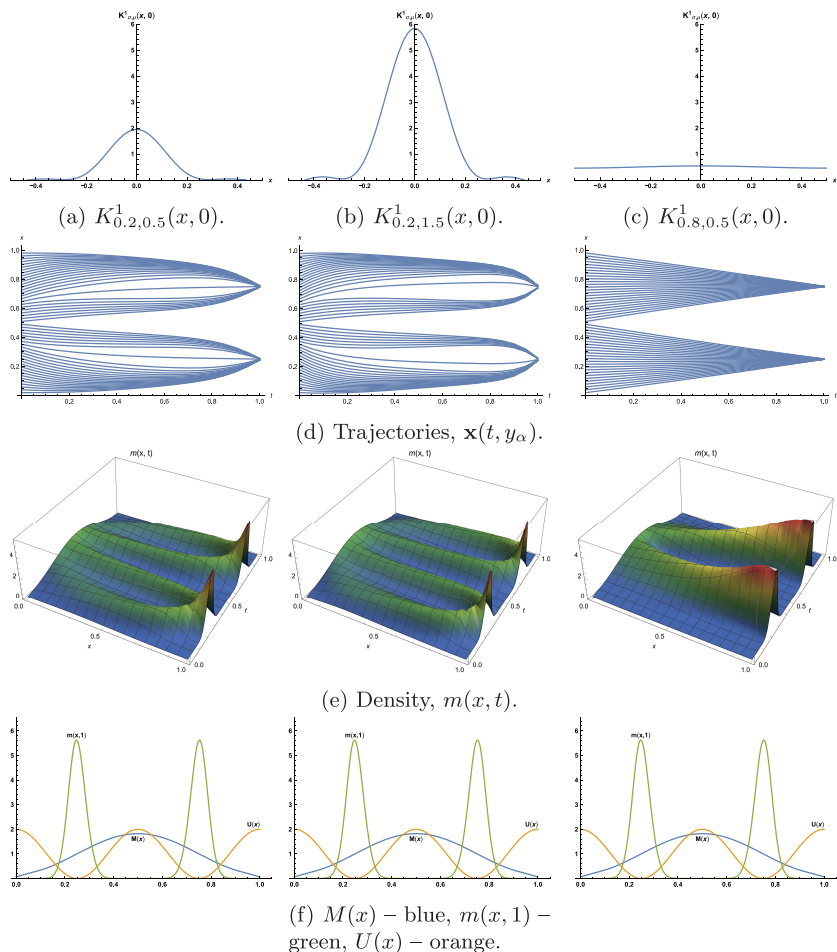
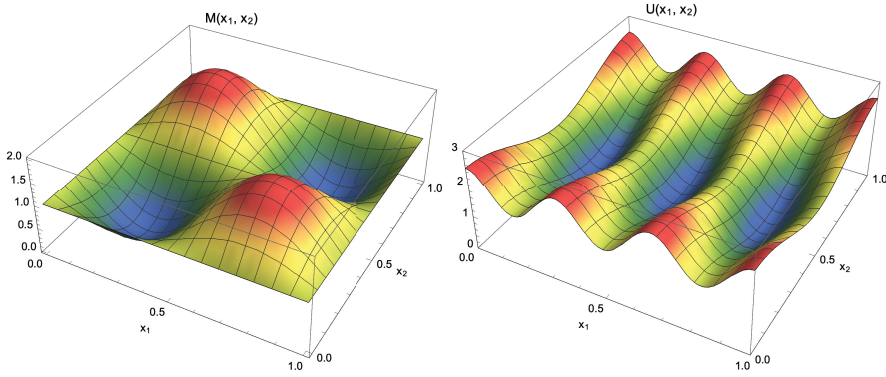


Figure 3. Simulations using Gaussian kernels with different parameters, $(\sigma, \mu) \in \{(0.2, 0.5), (0.2, 1.5), (0.8, 0.5)\}$, for each column. In the first row, we show a section of each kernel. In the second row, we plot the trajectories of the agents, $\{\mathbf{x}(t, y_\alpha)\}_{\alpha=1}^Q$, at time $t \in [0, 1]$ and initial positions $\{y_\alpha\}_{\alpha=1}^Q \subset \mathbb{T}$. In the third row, we plot the time evolution of the distribution of players, $m(t, x)$. Each plot of the last row displays the initial-terminal conditions, $M(x)$ and $U(x)$, and the final distribution, $m(x, 1)$.

trajectories of the decoupled system are straight lines by Hopf–Lax formula. As we can see in Figure 3(d), this fact is consistent with the straight-line trajectories that we obtain.

6.2. Two-dimensional examples. Here, we consider the case of two-dimensional state space. The initial distribution of players and the terminal cost function



(a) Initial distribution of agents, $M(x_1, x_2)$. (b) Terminal cost function, $U(x_1, x_2)$.

Figure 4. Initial-terminal conditions.

are given by

$$\begin{aligned}
 M(x_1, x_2) &= 1 + \frac{1}{2} \cos(\pi + 2\pi(x_1 - x_2)) + \frac{1}{2} \sin\left(\frac{\pi}{2} + 2\pi(x_1 + x_2)\right), \\
 U(x_1, x_2) &= \frac{3}{2} + \frac{1}{2} (\cos(6\pi x_1) + \cos(2\pi x_2)), \quad (x_1, x_2) \in \mathbb{T}^2,
 \end{aligned}$$

that are depicted in Figure 4.

The corresponding expansion of the kernel is given by

$$\begin{aligned}
 &K_{\sigma, \mu}^2(x_1, x_2; y_1, y_2) \\
 &= \sum_{k, k'=1}^{\infty} \mu^2 e^{-(\pi^2 \sigma^2 / 2)([k/2]^2 + [k'/2]^2)} \phi_k(x_1) \phi_k(y_1) \phi_{k'}(x_2) \phi_{k'}(y_2) \\
 &= \sum_{k, k'=1}^{\infty} \mu^2 e^{-(\pi^2 \sigma^2 / 2)([k/2]^2 + [k'/2]^2)} \phi_{k, k'}(x_1, x_2) \phi_{k, k'}(y_1, y_2), \tag{55}
 \end{aligned}$$

where

$$\phi_{k, k'}(x_1, x_2) = \phi_k(x_1) \phi_{k'}(x_2), \quad x_1, x_2 \in \mathbb{T}, \quad k, k' \in \mathbb{N}. \tag{56}$$

Thus, for a fixed r we take as a basis functions the set:

$$\{\phi_{1,1}, \phi_{1,2}, \dots, \phi_{1,r-1}, \phi_{2,1}, \dots, \phi_{2,r-2}, \dots, \phi_{r-1,1}\} = \{\psi_1, \psi_2, \dots, \psi_{r(r-1)/2}\}.$$

Therefore, we take all functions $\phi_{k,k'}$ such that $k + k' \leq r$ and order them in the lexicographic order. The corresponding matrices will be of size $\frac{r(r-1)}{2} \times \frac{r(r-1)}{2}$:

$$\begin{aligned} \mathbf{K} &= \text{diag}(\mu^2 e^{-(\pi^2 \sigma^2 / 2)([k/2]^2 + [k'/2]^2)})_{k+k' \leq r}, \\ \mathbf{J} &= \text{diag}(\mu^{-2} e^{(\pi^2 \sigma^2 / 2)([k/2]^2 + [k'/2]^2)})_{k+k' \leq r}, \end{aligned} \tag{57}$$

where the order is again lexicographic.

To compare the results, we use the same time and space discretization throughout all our 2-dimensional experiments, as well as the same parameters for the numerical scheme. We discretize the time using a step size $\Delta t = \frac{1}{N}$. For the discretization of M we use

$$y_{\alpha\alpha'} = \left(\frac{\alpha}{Q+1}, \frac{\alpha'}{Q+1} \right), \quad c_{\alpha\alpha'} = \frac{M(y_{\alpha\alpha'})}{\sum_{\beta, \beta'=1}^Q M(y_{\beta\beta'})}, \quad 1 \leq \alpha, \alpha' \leq Q.$$

We choose $N = 20$, $Q = 20$ and use eight basis functions, $r = 8$. Furthermore, we set the numerical scheme parameters to $\lambda = 1$, $\omega = \frac{1}{12}$ and $\theta = 1$.

In Figure 5, we plot the Gaussian kernels used in the simulations, with different values of μ and σ . We see that the bigger μ is the higher the peak of the kernel, see (a) and (b) in Figure 5. This means that each agent in (a) is more adverse of being in crowded areas than agents is (b), $\mu = 0.75$ and $\mu = 0.5$ respectively. For higher values of σ we see that the kernel becomes flat, compare (b) with (c) in Figure 5, for $\sigma = 0.1$ and $\sigma = 1$ respectively. As before, this means that the agents penalize others independent of mutual distances.

In Figure 6, we compare the simulation results using the same initial-terminal conditions, see Figure 4, but different kernel functions (plotted in the first row of Figure 6). In the last row of Figure 6 we have the final distribution of agents.

We see that for larger values of μ , left column compared with the middle one, the agents' concentration near low-cost regions of terminal cost, U , is less dense. We also see that when σ is bigger the agents become more indifferent to the

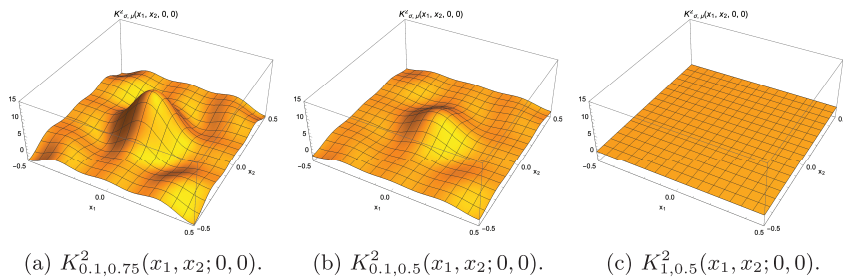


Figure 5. Plots of the Gaussian kernels for $(\sigma, \mu) \in \{(0.1, 0.75), (0.1, 0.5), (1, 0.5)\}$.

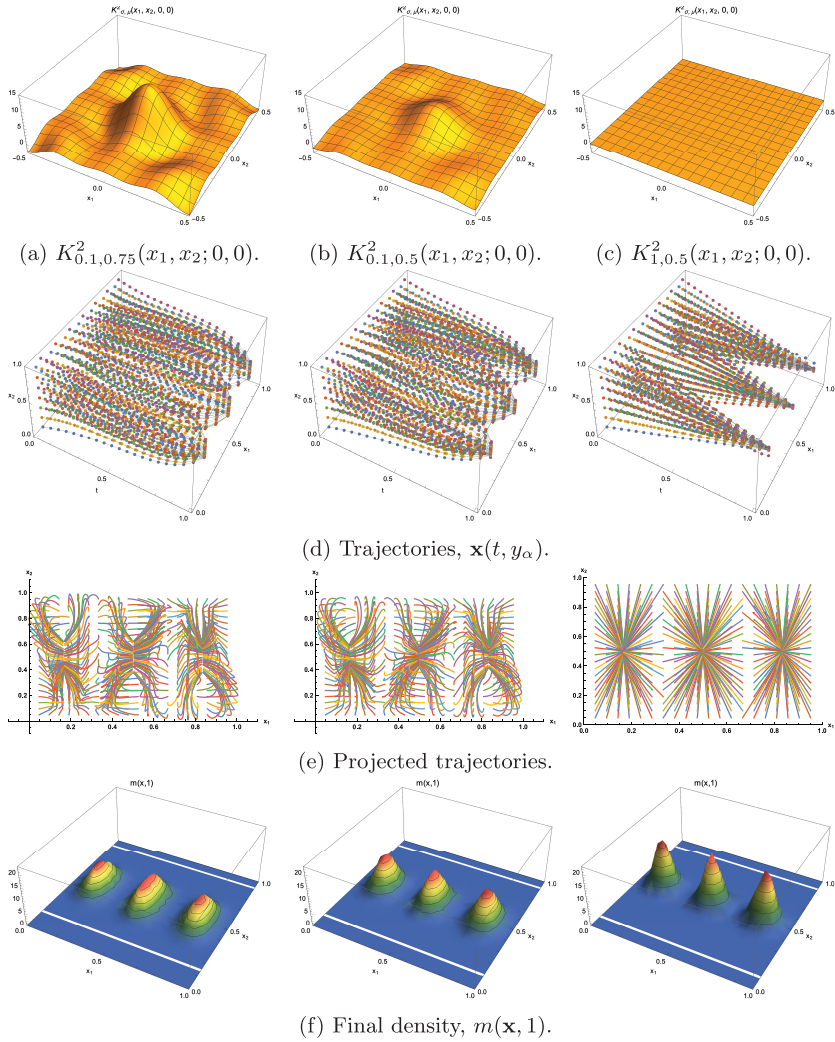


Figure 6. Simulations using Gaussian kernels with different parameters, $(\sigma, \mu) \in \{(0.1, 0.75), (0.1, 0.5), (1, 0.5)\}$, for each column. In the first row we show a section of each kernel. In the second row we show the trajectories of the agents, $\{\mathbf{x}(t, y_{\alpha\alpha'})\}_{\alpha, \alpha'=1}^Q$, $t \in [0, 1]$, with initial positions $\{y_{\alpha\alpha'}\}_{\alpha, \alpha'=1}^Q \in \mathbb{T}^2$. In the third row, we plot the 2D projection of the trajectories. And in the last row, we plot the final distribution of the agents, $m(\mathbf{x}, 1)$.

density of the crowd, and concentrate more densely near low-cost values of U – see the right column in Figure 6(f).

As in the 1-dimensional case, looking to the projected trajectories in the 2-dimensional plane we observe that for flat kernel agents follow straight lines from the initial positions to closest low-cost regions of the terminal cost function.

Acknowledgments. J. Saúde was partially supported by FCT/Portugal through the CMU-Portugal Program.

References

- [1] Y. Achdou. Finite difference methods for mean field games. In *Hamilton-Jacobi equations: approximations, numerical analysis and applications*, volume 2074 of *Lecture Notes in Math.*, pages 1–47. Springer, Heidelberg, 2013.
- [2] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games: convergence of a finite difference method. *SIAM J. Numer. Anal.*, 51(5):2585–2612, 2013.
- [3] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: numerical methods. *SIAM J. Numer. Anal.*, 48(3):1136–1162, 2010.
- [4] Y. Achdou and A. Porretta. Convergence of a finite difference scheme to weak solutions of the system of partial differential equations arising in mean field games. *SIAM J. Numer. Anal.*, 54(1):161–186, 2016.
- [5] N. Almulla, R. Ferreira, and D. A. Gomes. Two numerical approaches to stationary mean-field games. *Dyn. Games Appl.*, 7(4):657–682, 2017.
- [6] R. Andreev. Preconditioning the augmented Lagrangian method for instationary mean field games with diffusion. *SIAM J. Sci. Comput.*, 39(6):A2763–A2783, 2017.
- [7] A. Aurell and B. Djehiche. Mean-field type modeling of nonlocal crowd aversion in pedestrian crowd dynamics. *SIAM J. Control Optim.*, 56(1):434–455, 2018.
- [8] M. Bardi and I. 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.
- [9] J.-D. Benamou and G. Carlier. Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations. *J. Optim. Theory Appl.*, 167(1):1–26, 2015.
- [10] J.-D. Benamou, G. Carlier, S. Di Marino, and L. Nenna. An entropy minimization approach to second-order variational mean-field games. Preprint, 2018. [arXiv:1807.09078v1](https://arxiv.org/abs/1807.09078v1) [math.OC].
- [11] J.-D. Benamou, G. Carlier, and F. Santambrogio. Variational mean field games. In *Active particles. Vol. 1. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 141–171. Birkhäuser/Springer, Cham, 2017.

- [12] A. Bensoussan, J. Frehse, and P. Yam. *Mean field games and mean field type control theory*. SpringerBriefs in Mathematics. Springer, New York, 2013.
- [13] L. M. Briceño Arias, D. Kalise, and F. J. Silva. Proximal methods for stationary mean field games with local couplings. *SIAM J. Control Optim.*, 56(2):801–836, 2018.
- [14] P. Cardaliaguet. Long time average of first order mean field games and weak KAM theory. *Dyn. Games Appl.*, 3(4):473–488, 2013.
- [15] P. Cardaliaguet. Notes on mean field games, 2013. <https://www.ceremade.dauphine.fr/cardaliaguet/>.
- [16] 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.
- [17] E. Carlini and F. J. Silva. A semi-Lagrangian scheme for a degenerate second order mean field game system. *Discrete Contin. Dyn. Syst.*, 35(9):4269–4292, 2015.
- [18] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I*. Probability Theory and Stochastic Modelling. Springer, Cham, 2018.
- [19] A. Cesaroni and M. Cirant. Introduction to variational methods for viscous ergodic mean-field games with local coupling. 2017. Lecture notes in Springer INdAM Series, to appear.
- [20] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, May 2011.
- [21] Y. T. Chow, J. Darbon, S. Osher, and W. Yin. Algorithm for overcoming the curse of dimensionality for time-dependent non-convex Hamilton-Jacobi equations arising from optimal control and differential games problems. *J. Sci. Comput.*, 73(2–3):617–643, 2017.
- [22] Y. T. Chow, W. Li, S. Osher, and W. Yin. Algorithm for Hamilton-Jacobi equations in density space via a generalized Hopf formula. Preprint, 2018. [arXiv:1805.01636v1](https://arxiv.org/abs/1805.01636v1) [math.NA].
- [23] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1993.
- [24] D. A. Gomes, E. A. Pimentel, and V. Voskanyan. *Regularity theory for mean-field game systems*. SpringerBriefs in Mathematics. Springer, 2016.
- [25] D. A. Gomes and J. Saúde. Mean field games models – a brief survey. *Dynamic Games and Applications*, 4(2):110–154, Jun 2014.
- [26] D. A. Gomes and J. Saúde. Numerical methods for finite-state mean-field games satisfying a monotonicity condition. *Applied Mathematics & Optimization*, Aug 2018.
- [27] O. Guéant, J.-M. Lasry, and P.-L. Lions. *Mean Field Games and Applications*, pages 205–266. Springer, Berlin, Heidelberg, 2011.
- [28] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ -Nash equilibria. *IEEE Trans. Automat. Control*, 52(9):1560–1571, 2007.

- [29] 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.
- [30] 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.
- [31] 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.
- [32] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
- [33] A. T. Lin, Y. T. Chow, and S. Osher. A splitting method for overcoming the curse of dimensionality in Hamilton-Jacobi equations arising from nonlinear optimal control and differential games with applications to trajectory generation. Preprint, 2018. [arXiv:1803.01215v1](https://arxiv.org/abs/1803.01215v1) [math.OC].
- [34] P.-L. Lions. Collège de France course on mean-field games. 2007-2011. <https://www.college-de-france.fr/site/en-pierre-louis-lions/index.htm>.
- [35] L. Nurbekyan. One-dimensional, non-local, first-order stationary mean-field games with congestion: a Fourier approach. *Discrete Contin. Dyn. Syst. Ser. S*, 11(5):963–990, 2018.
- [36] I. Yegorov and P. M. Dower. Perspectives on characteristics based curse-of-dimensionality-free numerical approaches for solving Hamilton-Jacobi equations. *Applied Mathematics & Optimization*, Jul 2018.

Received November 14, 2018; revision received January 25, 2019

L. Nurbekyan, Department of Mathematics and Statistics, McGill University,
805 Sherbrooke St W, Montreal, H3A 0B9 QC, Canada

E-mail: levon.nurbekyan@mcgill.ca

J. Saúde, Department of Electrical and Computer Engineering, Carnegie Mellon
University, 5000 Forbes Avenue, Pittsburgh, USA

E-mail: jsaude@cmu.edu