Journal of Computational Physics 491 (2023) 112346

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

High order computation of optimal transport, mean field planning, and potential mean field games

Guosheng Fu [a,1], Siting Liu [b,2], Stanley Osher [b,2], Wuchen Li [c,*,3]

[a] Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, IN 46556, USA
[b] Department of Mathematics, University of California, Los Angeles, Los Angeles, CA 90095, USA
[c] Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

ARTICLE INFO

Article history:
Received 3 February 2023
Received in revised form 3 June 2023
Accepted 2 July 2023
Available online 24 July 2023

Keywords:
High order computation
Optimal transport
Mean-field planning
Mean-field games

ABSTRACT

Mean-field games (MFGs) have shown strong modeling capabilities for large systems in various fields, driving growth in computational methods for mean-field game problems. However, high order methods have not been thoroughly investigated. In this work, we explore applying general high-order numerical schemes with finite element methods in the space-time domain for computing the optimal transport (OT), mean-field planning (MFP), and potential MFG problems. We conduct several experiments to validate the convergence rate of the high order method numerically. Those numerical experiments also demonstrate the efficiency and effectiveness of our approach.

1. Introduction

Proposed by Lasry and Lions [34] and independently by Caines, Huang, and Malhamé [30], the mean-field game models an infinite number of identical agents' interactions in a mean-field manner, characterizing the equilibrium state of the system. Thanks to its substantial descriptive ability, the MFG becomes an important approach to studying complex systems with large populations of interacting agents, such as crowd dynamics, financial markets, power systems, pandemics, etc. [19,20,33,6,7,32,41,40]. The mean-field planning is a class of MFGs where the distribution of agents at the terminal time is imposed [50]. On the other hand, the Benamou-Brenier dynamic formulation [10] of the optimal transport problem connects with the variational form of the potential mean-field games. It can be treated as a special case of the mean-field planning problem, which aims to find an efficient way of moving one probability distribution to another. Along with the empirical success of MFG and OT in modeling and real-world applications, the study of mean-field games is also expanding. From the PDE view, the mean-field game model can be described by a system of coupled partial differential equations: a forward-in-time Fokker-Planck (FP) equation governs the evolution of the population, and a backward-in-time Hamilton-Jacobi-Bellman (HJB) equation for the value function that characterizes the control problem. For a review of MFG theory, we refer to [35,27,16,26]. In particular, a concept of displacement convexities of mean-field games has been recently studied in [23,24].

* Corresponding author.
E-mail addresses: gfu@nd.edu (G. Fu), siting6@math.ucla.edu (S. Liu), sjo@math.ucla.edu (S. Osher), wuchen@mailbox.sc.edu (W. Li).

[1] G. Fu's work is supported in part by NSF DMS-2012031.
[2] S. Osher and S. Liu's work are supported in part by AFOSR MURI FP 9550-18-1-502, and ONR grants: N00014-20-1-2093, and N00014-20-1-2787.
[3] W. Li's work is supported by AFOSR MURI FP 9550-18-1-502, AFOSR YIP award No. FA 9550-23-1-0087, and NSF RTG: 2038080.

https://doi.org/10.1016/j.jcp.2023.112346
0021-9991/

With such a wide range of applications, computational methods play a crucial role since most MFG and OT problems do not have analytical solutions. While some recent computational approaches take advantage of machine learning methods and game theories [51,45,4,21,37,28,17,29], classical numerical methods are mostly developed discretization using finite difference schemes or semi-Lagrangian schemes. In [1], the MFG system is discretized using a finite difference scheme and then solved by Newton's method. Semi-Lagrangian methods are studied in [18]. As for MFGs and OTs that can be written in a variational form, optimization methods, such as augmented Lagrangian, primal-dual hybrid gradient, alternating direction method of multipliers, are applied to solve the discretized system [10,11,49,3,8,14,46]. For dynamic optimal transport, the convergence of the discretized problem to the continuous one has been studied in [25,38,48] for different discretization approaches. In [39], dynamical optimal transport is applied to interpolate probability distribution on a discrete surface. Recently, the computation of MFGs on manifolds has been investigated in [54]. For the survey of the numerical methods, we refer to [2,36]. Within the augmented Lagrangian framework, the (low-order) finite element discretization has been used frequently; see, e.g., [10,11,5,31].

Pioneering works on computational OT/MFG focus on first or second order methods; the general high order method is not well studied. Yet, high order methods generally have faster convergence rates in numerical analysis and provide more accurate solutions on a much coarse computational mesh than low order methods. Therefore, exploring high order computational methods for mean-field games and optimal transport problems is vital.

In this work, we propose a general high order numerical method for solving the optimal transportation problem and mean-field game (control) problems using the finite element method. More precisely,

1. We discretize the augmented Lagrangian formulations of the MFP and MFG systems using high-order space-time finite elements. Considering derivation information used in the saddle point formulation, we approximate the value function (dual variable) ϕ using high order H^1-conforming finite elements, while the density and momentum (primal variables) $\rho, \boldsymbol{m}$ are approximated via a high-order (discontinuous) integration rule space which only records values on the high-order (space-time) integration points. Our discrete saddle-point problem is then solved via the ALG2 algorithm, following [11]. To the best of our knowledge, this is the first time high-order schemes with more than second order accuracy being applied.
2. We present a series of comprehensive experiments to showcase the efficacy and efficiency of the proposed numerical algorithms. These experiments numerically validate the convergence rate of the algorithms as a function of mesh size and polynomial degree. In particular, we show a high-order method on a coarse mesh is more accurate than a low order method on a fine mesh with the same number of degrees of freedom. Furthermore, we apply the finite element scheme to a set of mean-field planning and mean-field game problems on non-rectangular domains (with obstacles) and computational graphics, demonstrating the validity and practicality of our method.

We would like to point out that there are two separate convergence issues related to our high-order algorithm. The first one is the convergence of the ALG2 optimization solver, while the second one is the convergence towards mesh refinement of the proposed FEM discretization. While per iteration cost is low, the convergence issue for ALG2 is well-known as it requires a lot of iterations to achieve high accuracy [13]. We present numerical results to support the high-order convergence towards mesh refinements of our FEM discretization for the optimal transport problems. The detailed theoretical/computational studies of these issues will be the subject of our future work.

This paper is organized as follows. Section 2 reviews the dynamic formulation of optimal transportation, mean-field planning, and mean-field games. Section 3 presents the high-order schemes we designed for computing the above problems and the companion algorithm. Section 4 demonstrates the effectiveness of the high-order method with numerical experiments. Finally, we make some conclusions and remarks in Section 5.

2. OT, MFP, and MFG

In this section, we briefly review the formulation of mean-field planning problems involving dynamic mean-field control and potential MFG problems (also referred as mean-field control problems) studied in this work.

2.1. Dynamic mean-field planning problems (MFP)

Consider the model on time interval $[0, 1]$ and space region $\Omega \subset \mathbb{R}^D$. Let ρ be the density of agents through $t \in [0, 1]$, $\boldsymbol{m}$ be the flux of the density which models strategies (control) of the agents, and $(\rho, \boldsymbol{m}) \in \mathcal{C}$:

$$\mathcal{C} := \left\{ \begin{array}{c} (\rho, \boldsymbol{m}) : \rho : [0,1] \times \Omega \to \mathbb{R}^+, \int_\Omega \rho(t, \boldsymbol{x}) \mathrm{d}\boldsymbol{x} = 1, \forall t \in [0,1], \\ \boldsymbol{m} : [0,1] \times \Omega \to \mathbb{R}^D \text{ is Lebesgue measurable}, \end{array} \right\}. \tag{2.1}$$

We are interested in ρ with given initial and terminal density ρ_0, ρ_1 and $(\rho, \boldsymbol{m})$ satisfying zero boundary flux and mass conservation law, which satisfies the constraint set $\mathcal{C}(\rho_0, \rho_1)$:

$$\mathcal{C}(\rho_0,\rho_1) := \mathcal{C} \cap \left\{ \begin{array}{l} (\rho,\boldsymbol{m}) : \partial_t \rho + \mathrm{div}_{\boldsymbol{x}} \boldsymbol{m} = 0, \\ \qquad \boldsymbol{m}\cdot\boldsymbol{n} = 0 \text{ for } \boldsymbol{x} \in \partial\Omega,\ \rho(0,\cdot) = \rho_0,\ \rho(1,\cdot) = \rho_1, \end{array} \right\}. \tag{2.2}$$

where the equation hold in the sense of distribution.

We denote $L : \mathbb{R}^+ \times \mathbb{R}^D \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ as the dynamic cost function and $A : \mathbb{R} \to \overline{\mathbb{R}}$ as a function modeling interaction cost. The goal of MFP is to minimize the total cost among all feasible $(\rho, \boldsymbol{m}) \in \mathcal{C}(\rho_0, \rho_1)$. Therefore, the problem can be formulated as

$$\inf_{(\rho,\boldsymbol{m})\in\mathcal{C}(\rho_0,\rho_1)} \int_0^1 \int_\Omega L(\rho(t,\boldsymbol{x}),\boldsymbol{m}(t,\boldsymbol{x})) + A(\rho(t,\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\mathrm{d}t. \tag{2.3}$$

It is clear to see $\mathcal{C}(\rho_0,\rho_1)$ is convex. In addition, the mass conservation law $\partial_t\rho + \mathrm{div}_{\boldsymbol{x}}\boldsymbol{m} = 0$ and zero flux boundary condition $\boldsymbol{m}\cdot\boldsymbol{n} = 0, \boldsymbol{x} \in \partial\Omega$ imply that $\mathcal{C}(\rho_0,\rho_1) \neq \emptyset$ if and only if $\int_\Omega \rho_0 = \int_\Omega \rho_1$. Once $\mathcal{C}(\rho_0,\rho_1)$ is non-empty, the existence and uniqueness of the optimizer depend on L and A. In this paper, we consider a typical dynamic cost function L by

$$L(\beta_0,\boldsymbol{\beta}_1) := \begin{cases} \frac{\|\boldsymbol{\beta}_1\|^2}{2\beta_0} & \text{if } \beta_0 > 0 \\ 0 & \text{if } \beta_0 = 0, \boldsymbol{\beta}_1 = \boldsymbol{0} \\ +\infty & \text{if } \beta_0 = 0, \boldsymbol{\beta}_1 \neq \boldsymbol{0}. \end{cases} \cdot \tag{2.4}$$

Various choices of the interaction function will be given in the numerics section. We note that our proposed numerical scheme works for more general dynamic cost L, like the ones used in [12]. However, this manuscript exclusively focuses on the quadratic case in (2.4).

If the interaction cost function $A = 0$, the MFP becomes the dynamic formulation of optimal transport problem:

$$(\mathrm{OT}) \quad \min_{\rho,\boldsymbol{m}\in\mathcal{C}(\rho_0,\rho_1)} \int_0^1 \int_\Omega L(\rho(t,\boldsymbol{x}),\boldsymbol{m}(t,\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\mathrm{d}t. \tag{2.5}$$

Since $\boldsymbol{m} = \rho\boldsymbol{v}$, this definition of L makes sure that $\boldsymbol{m} = \boldsymbol{0}$ wherever $\rho = 0$. OT can be viewed as a special case of MFP where masses move freely in Ω through $t \in [0,1]$.

To simplify notation, we denote an element in the set $\mathcal{C}$ as $\boldsymbol{\alpha} := (\alpha_0, \boldsymbol{\alpha}_1) \in \mathcal{C}$. Introducing the Lagrangian multiplier $\phi : [0,1]\times\Omega$ for the constraint (2.2), the MFP problem (2.3) can be reformulated as the following saddle-point problem:

$$\inf_{\boldsymbol{\alpha}} \sup_{\phi} F(\boldsymbol{\alpha}) - G(\phi) - \langle \boldsymbol{\alpha}, \nabla_{t,x}\phi \rangle, \tag{2.6a}$$

where

$$F(\boldsymbol{\alpha}) := \int_0^1 \int_\Omega L(\alpha_0(t,\boldsymbol{x}),\boldsymbol{\alpha}_1(t,\boldsymbol{x})) + A(\alpha_0(t,\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\mathrm{d}t, \tag{2.6b}$$

$$G(\phi) := \int_\Omega -\phi(1,\boldsymbol{x})\rho_1(\boldsymbol{x}) + \phi(0,\boldsymbol{x})\rho_0(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{2.6c}$$

$\nabla_{t,x} = (\partial_t, \mathrm{grad}_{\boldsymbol{x}})$ is the space-time gradient operator, and $\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle := \int_0^1 \int_\Omega \boldsymbol{\alpha}\cdot\boldsymbol{\beta}\,\mathrm{d}\boldsymbol{x}\mathrm{d}t$ is the space-time integral. The KKT system for this saddle-point system (with cost function L in (2.4)) is the following PDE system on the space-time domain $[0,1]\times\Omega$

$$\partial_t\rho + \mathrm{div}_{\boldsymbol{x}}\boldsymbol{m} = 0, \tag{2.7a}$$

$$\frac{\boldsymbol{m}}{\rho} - \mathrm{grad}_{\boldsymbol{x}}\phi = 0, \tag{2.7b}$$

$$\partial_t\phi + \frac{|\boldsymbol{m}|^2}{2\rho^2} = A'(\rho), \tag{2.7c}$$

with boundary conditions

$$\boldsymbol{m}\cdot\boldsymbol{n} = 0, \quad \text{on } [0,1]\times\partial\Omega, \tag{2.7d}$$

$$\rho(0,\boldsymbol{x}) = \rho_0(\boldsymbol{x}), \quad \rho(1,\boldsymbol{x}) = \rho_1(\boldsymbol{x}) \quad \text{on } \Omega. \tag{2.7e}$$

We would like to point out that the above equations (2.7a)–(2.7c) are the standard PDE formulation for MFG [35]. Moreover, equation (2.7b) and boundary condition (2.7d) implies the following boundary condition for the dual variable: $\rho(\text{grad}_x\phi \cdot \boldsymbol{n}) = 0$ on $[0,1]\times\partial\Omega$. Denoting $[L+A]^*(\boldsymbol{\alpha}^*)$ as the convex conjugate (Legendre transformation) of $L(\alpha_0,\boldsymbol{\alpha}_1)+A(\alpha_0)$ with L given in (2.4), i.e.,

$$\begin{aligned}
[L+A]^*(\boldsymbol{\alpha}^*) &= \sup_{\boldsymbol{\alpha}} \boldsymbol{\alpha}\cdot\boldsymbol{\alpha}^* - L(\alpha_0,\boldsymbol{\alpha}_1) - A(\alpha_0)\\
&= \sup_{\alpha_0} \alpha_0\cdot(\alpha_0^* + |\boldsymbol{\alpha}_1^*|^2) - \underbrace{L(\alpha_0,\alpha_0\boldsymbol{\alpha}_1^*)}_{=\frac{1}{2}\alpha_0|\boldsymbol{\alpha}_1^*|^2} - A(\alpha_0)\\
&= \sup_{\alpha_0} \alpha_0\cdot(\alpha_0^* + \frac{1}{2}|\boldsymbol{\alpha}_1^*|^2) - A(\alpha_0)\\
&= A^*(\alpha_0^* + \frac{1}{2}|\boldsymbol{\alpha}_1^*|^2),
\end{aligned}$$

where in the second equality, we used the optimality condition $\boldsymbol{\alpha}_1 = \alpha_0\boldsymbol{\alpha}_1^*$. By duality, we have

$$L(\alpha_0,\boldsymbol{\alpha}_1) + A(\alpha_0) = \sup_{\boldsymbol{\alpha}^*} \boldsymbol{\alpha}\cdot\boldsymbol{\alpha}^* - A^*(\alpha_0^* + \frac{1}{2}|\boldsymbol{\alpha}_1^*|^2). \tag{2.8}$$

Using the above relation, we have the following dual formulation of the saddle-point problem (2.6a):

$$\sup_{\boldsymbol{\alpha}}\inf_{\phi,\boldsymbol{\alpha}^*} F^*(\boldsymbol{\alpha}^*) + G(\phi) + \langle\boldsymbol{\alpha}, \nabla_{t,x}\phi - \boldsymbol{\alpha}^*\rangle, \tag{2.9}$$

where

$$F^*(\boldsymbol{\alpha}^*) = \int_0^1\int_\Omega A^*(\alpha_0^* + \frac{1}{2}|\boldsymbol{\alpha}_1^*|^2)\,\mathrm{d}\boldsymbol{x}\mathrm{d}t.$$

Introducing the augmented Lagrangian

$$L_r(\phi,\boldsymbol{\alpha},\boldsymbol{\alpha}^*) := F^*(\boldsymbol{\alpha}^*) + G(\phi) + \langle\boldsymbol{\alpha}, \nabla_{t,x}\phi - \boldsymbol{\alpha}^*\rangle + \frac{r}{2}\langle\nabla_{t,x}\phi - \boldsymbol{\alpha}^*, \nabla_{t,x}\phi - \boldsymbol{\alpha}^*\rangle,$$

where r is a positive parameter, it is clear that the corresponding saddle-point problem

$$\sup_{\boldsymbol{\alpha}}\inf_{\phi,\boldsymbol{\alpha}^*} L_r(\phi,\boldsymbol{\alpha},\boldsymbol{\alpha}^*) \tag{2.10}$$

has the same solution as (2.9).

2.2. Dynamic potential mean-field games (MFG)

For potential MFG, the terminal density ρ_1 is not explicitly provided but satisfies a given preference. The goal of MFG is to minimize the total cost among all feasible $(\rho,\boldsymbol{m})\in\mathcal{C}(\rho_0)$:

$$\inf_{(\rho,\boldsymbol{m})\in\mathcal{C}(\rho_0)} F((\rho,\boldsymbol{m})) + \underbrace{\int_\Omega \Gamma(\rho(1,\boldsymbol{x}))\mathrm{d}\boldsymbol{x}}_{:=R(\rho(1,\cdot))}, \tag{2.11}$$

where $\Gamma:\mathbb{R}\to\overline{\mathbb{R}}$ is the terminal cost, and the constraint set $\mathcal{C}(\rho_0)$ is similar to $\mathcal{C}(\rho_0,\rho_1)$:

$$\mathcal{C}(\rho_0) := \mathcal{C}\cap\left\{(\rho,\boldsymbol{m}) : \begin{array}{l}\partial_t\rho + \mathrm{div}_{\boldsymbol{x}}\boldsymbol{m} = 0,\\ \boldsymbol{m}\cdot\boldsymbol{n} = 0 \text{ for } \boldsymbol{x}\in\partial\Omega,\ \rho(0,\cdot) = \rho_0,\end{array}\right\}. \tag{2.12}$$

Similar to MFP, we reformulate the problem (2.11) into a saddle-point problem:

$$\inf_{\boldsymbol{\alpha},\rho_1}\sup_{\phi} F(\boldsymbol{\alpha}) + R(\rho_1) + (\rho_1,\phi(1,\cdot)) - (\rho_0,\phi(0,\cdot)) - \langle\boldsymbol{\alpha}, \nabla_{t,x}\phi\rangle, \tag{2.13}$$

in which $(\alpha,\beta) := \int_\Omega \alpha\beta\,\mathrm{d}\boldsymbol{x}$ is the spatial integration. Here the KKT system of the saddle-point problem (2.13) is simply the MFP system (2.7) with boundary condition (2.7e) replaced by the following:

$$\rho(0,\boldsymbol{x}) = \rho_0(\boldsymbol{x}),\quad \phi(1,\boldsymbol{x}) = -\Gamma'(\rho_1(\boldsymbol{x}))\quad \text{on } \Omega.$$

Introducing the dual variables $\boldsymbol{\alpha}^*$ and ρ_1^* for $\boldsymbol{\alpha}$ and ρ_1, respectively, we get the following equivalent saddle-point problem:

$$\sup_{\boldsymbol{\alpha},\rho_1}\inf_{\phi,\boldsymbol{\alpha}^*,\rho_1^*} F^*(\boldsymbol{\alpha}^*) + \langle \boldsymbol{\alpha}, \nabla_{t,x}\phi - \boldsymbol{\alpha}^*\rangle + R^*(\rho_1^*) - (\rho_1, \phi(1,\cdot) + \rho_1^*) + (\rho_0, \phi(0,\cdot)), \tag{2.14}$$

where $R^*(\rho_1^*) := \int_\Omega \Gamma^*(\rho_1(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}$, with Γ^* being the convex conjugate of Γ.

The augmented Lagrangian reformulation of (2.14) is the following:

$$\begin{aligned}\sup_{\boldsymbol{\alpha},\rho_1}\inf_{\phi,\boldsymbol{\alpha}^*,\rho_1^*} & F^*(\boldsymbol{\alpha}^*) + R^*(\rho_1^*) + (\rho_0, \phi(0,\cdot)) \\ & + \langle \boldsymbol{\alpha}, \nabla_{t,x}\phi - \boldsymbol{\alpha}^*\rangle + \frac{r_1}{2}\langle \nabla_{t,x}\phi - \boldsymbol{\alpha}^*, \nabla_{t,x}\phi - \boldsymbol{\alpha}^*\rangle \\ & - (\rho_1, \phi(1,\cdot) + \rho_1^*) + \frac{r_2}{2}(\phi(1,\cdot) + \rho_1^*, \phi(1,\cdot) + \rho_1^*),\end{aligned} \tag{2.15}$$

where r_1, r_2 are two positive parameters. Here we allow the augmented Lagrangian parameters r_1 and r_2 to be different. In the numerical experiments, we simply take $r_1 = r_2 = 1$.

Remark 2.1. Following the seminal works in [9,11], we propose our high-order schemes for MFP and MFG based on the augmented Lagrangian formulations (2.10). and (2.15). The discrete saddle-point problem is then solved using the ALG2 algorithm [22]. The major novelty of our scheme is the use of high-order space-time finite elements for the discretization of the variables in (2.10) and (2.15). This is the first time high-order schemes with more than second order accuracy being applied to such problems.

3. High-order schemes for OT, MFP and MFG

In this section, we discretize the augmented Lagrangian problems (2.10) and (2.15) using high-order space-time finite element spaces. We start with notation including the mesh and definition of finite element spaces to be used. We then formulate the discrete saddle-point problems using these finite element spaces, which is solved iteratively using the ALG2 algorithm [22]. Throughout this section, we restrict the discussion to $D = 2$ spatial dimensions.

Since space/time derivative information is needed for ϕ, we approximate it using (high-order) H^1-conforming finite elements. On the other hand, since no derivative information appears for $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}^*$, (and ρ_1 and ρ_1^* for MFG), it is natural to approximate these variables only on the (high-order) integration points.

3.1. The finite element spaces and notation

Let $\mathcal{I}_h = \{I_j\}_{j=1}^N$ be a triangulation of the time domain $[0,1]$ with $I_j = [x_{j-1}, x_j]$, and $0 = x_0 < x_1 < \cdots < x_N = 1$. Let $\mathcal{T}_h = \{T_\ell\}_{\ell=1}^M$ be a conforming triangulation of the spatial domain Ω, where we assume the element $T_\ell := \Phi_{T_\ell}(\widehat{T})$ is obtained from a polynomial mapping Φ_{T_ℓ} from the reference element $\widehat{T}$, which, is a unit triangle or unit square. We obtain the space-time mesh for $\Omega_T := [0,1]\times\Omega$ using tensor product of the spatial and temporal meshes:

$$\mathcal{I}_h \otimes \mathcal{T}_h := \{I_j \otimes T_\ell : \forall j \in \{1,\cdots,N\}, \text{ and } \ell \in \{1,\cdots,M\}\}.$$

We denote $\mathcal{P}^k(I)$ as the polynomial space of degree no greater than k on the interval I, and $\mathcal{P}^k(\widehat{T})$ as the polynomial space of degree no greater than k if $\widehat{T}$ is a unit triangle, or the tensor-product polynomial space of degree no greater than k in each direction if $\widehat{T}$ is a unit square for $k \geq 1$. The mapped polynomial space on a spatial physical element $T \in \mathcal{T}_h$ is denoted as

$$\mathcal{P}^k(T) := \{\widehat{v} \circ (\Phi_T)^{-1} : \forall \widehat{v} \in \mathcal{P}^k(\widehat{T}))\}.$$

We denote $\{\widehat{\boldsymbol{\xi}}_i\}_{i=1}^{N_k}$ as a set of quadrature points with positive weights $\{\widehat{\omega}_i\}_{i=1}^{N_k}$ that is accurate for polynomials of degree up to $2k+1$ on the reference element $\widehat{T}$, i.e.,

$$\int_{\widehat{T}} \widehat{f}\,\mathrm{d}\boldsymbol{x} = \sum_{i=1}^{N_k} \widehat{\omega}_i \widehat{f}(\widehat{\boldsymbol{\xi}}_i), \quad \forall \widehat{f} \in \mathcal{P}^{2k+1}(\widehat{T}). \tag{3.1}$$

Note that when $\widehat{T}$ is a reference square, we simply use the Gauss-Legendre quadrature rule with $N_k = (k+1)^2$, which is optimal. On the other hand, when $\widehat{T}$ is a reference triangle, the optimal choice of quadrature rule is more complicated; see, e.g., [55,53] and references cited therein. For example, the number N_k for $0 \leq k \leq 6$ of the symmetric quadrature

Table 1
Number of quadrature points N_k for the quadrature rule on a triangle that is accurate up to degree $2k+1$ for $0 \le k \le 6$.

	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
N_k on Triangle	1	6	7	15	19	28	37

rules on a triangle provided in [55] are given in Table 1. The integration points and weights on a physical element T_ℓ are simply obtained via mapping: $\{\boldsymbol{\xi}_i^\ell := \Phi_{T_\ell}(\widehat{\boldsymbol{\xi}}_i)\}_{i=1}^{N_k}$, and $\{\omega_i^\ell := |\mathrm{grad}_{\boldsymbol{x}} \Phi_{T_\ell}(\widehat{\boldsymbol{\xi}}_i)|\widehat{\omega}_i\}_{i=1}^{N_k}$. Moreover, we denote $\{\eta_i^j\}_{i=1}^{k+1}$ as the set of $(k+1)$ Gauss-Legendre quadrature points on the interval I_j with corresponding weights $\{\zeta_i^j\}_{i=1}^{k+1}$. To simplify the notation, we denote the set of physical integration points and weights

$$\Xi_h^k := \{\boldsymbol{\xi}_i^\ell : \ 1 \le i \le N_k, \ 1 \le \ell \le M\}, \tag{3.2a}$$
$$\Omega_h^k := \{\omega_i^\ell : \ 1 \le i \le N_k, \ 1 \le \ell \le M\}, \tag{3.2b}$$
$$H_h^k := \{\eta_i^j : \ 1 \le i \le k+1, \ 1 \le j \le N\}, \tag{3.2c}$$
$$Z_h^k := \{\zeta_i^j : \ 1 \le i \le k+1, \ 1 \le j \le N\}. \tag{3.2d}$$

Moreover, we denote $(\cdot,\cdot)_h$ as the discrete inner-product on the mesh $\mathcal{T}_h$ using the quadrature points Ξ_h^k and weights Ω_h^k:

$$(\alpha, \beta)_h := \sum_{\ell=1}^{M} \sum_{i=1}^{N_k} \alpha(\boldsymbol{\xi}_i^\ell)\beta(\boldsymbol{\xi}_i^\ell)\omega_i^\ell,$$

and $\langle\cdot,\cdot\rangle_h$ as the discrete inner-product on the space-time mesh $\mathcal{I}_h \otimes \mathcal{T}_h$ using the quadrature points Ξ_h^k, H_h^k and weights Ω_h^k, Z_h^k:

$$\langle\alpha, \beta\rangle_h := \sum_{\ell=1}^{M} \sum_{i_s=1}^{N_k} \sum_{j=1}^{N} \sum_{i_t=1}^{k+1} \alpha(\eta_{i_t}^j, \boldsymbol{\xi}_{i_s}^\ell)\beta(\eta_{i_t}^j, \boldsymbol{\xi}_{i_s}^\ell)\omega_{i_s}^\ell \zeta_{i_t}^\ell.$$

We are now ready to present our finite element spaces:

$$V_h^k := \{v \in H^1(\Omega_T) : \ v|_{I_j \times T_\ell} \in \mathcal{P}^k(I_j) \otimes \mathcal{P}^k(T_\ell) \ \forall j, \ell\}, \tag{3.3}$$
$$W_h^k := \{w \in L^2(\Omega_T) : \ w|_{I_j \times T_\ell} \in \mathcal{P}^k(I_j) \otimes W^k(T_\ell) \ \forall j, \ell\}, \tag{3.4}$$
$$M_h^k := \{\mu \in L^2(\Omega) : \quad \mu|_{T_\ell} \in W^k(T_\ell) \ \forall \ell\}, \tag{3.5}$$

where V_h^k is an H^1-conforming space on the space-time mesh $\mathcal{I}_h \otimes \mathcal{T}_h$, W_h^k is an L^2-conforming space on the space-time mesh $\mathcal{I}_h \otimes \mathcal{T}_h$, and M_h^k is an L^2-conforming space on the spacial mesh $\mathcal{T}_h$, in which the local space

$$W^k(T_\ell) := \mathcal{P}^k(T_\ell) \oplus \delta W_k(T_\ell),$$

is associated with the integration rule in (3.1) such that $\dim W^k(T_\ell) = N_k$, and the nodal conditions

$$\varphi_i^\ell(\boldsymbol{\xi}_j^\ell) = \delta_{ij}, \quad \forall 1 \le j \le N_k, \tag{3.6}$$

in which δ_{ij} is the Kronecker delta function determines a unique solution $\varphi_i^\ell \in W^k(T_\ell)$. This implies that $\{\varphi_i^\ell\}_{i=1}^{N_k}$ is a set of nodal bases for the space $W^k(T_\ell)$, i.e.,

$$W^k(T_\ell) = \mathrm{span}_{1 \le i \le N_k}\{\varphi_i^\ell\}. \tag{3.7}$$

When T^ℓ is mapped from a reference square, we have $N_k = (k+1)^2$, hence $W^k(T_\ell)$ is simply the (mapped) tensor product polynomial space $\mathcal{P}^k(T_\ell)$. Moreover, we emphasize that the explicit expression of the basis function ϕ_i^ℓ does not matter in our construction, as only their nodal degrees of freedom (DOFs) on the quadrature nodes will enter into the numerical integration. Furthermore, let $\{\psi_i^j(t)\}_{i=1}^{k+1}$ be the set of basis functions for $\mathcal{P}^k(I_j)$ corresponding to the Gauss-Legendre quadrature nodes $\{\eta_i^j\}_{i=1}^{k+1}$, i.e., $\psi_i^j \in \mathcal{P}^k(I_j)$ satisfies

$$\psi_i^j(\eta_l^j) = \delta_{il}, \quad \forall 1 \le l \le k+1.$$

With this notation by hand, we have

$$W_h^k = \mathrm{span}\left\{\psi_{i_t}^j(t)\varphi_{i_s}^\ell(\boldsymbol{x}) : \quad \begin{array}{l} 1 \le i_t \le k+1, 1 \le i_s \le N_k, \\ 1 \le j \le N, 1 \le \ell \le M \end{array}\right\} \tag{3.8}$$

and

$$M_h^k = \text{span}\left\{\varphi_{i_s}^{\ell}(\boldsymbol{x}):\ 1 \le i_s \le N_k, 1 \le \ell \le M\right\} \tag{3.9}$$

We approximate the dual variable ϕ using the H^1-conforming finite element space V_h^{k+1}, each components of $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ using the integration rule space W_h^k, and the variables ρ_1 and ρ_1^* (for MFG) using the integration rule space M_h^k.

3.2. High-order FEM for MFP and MFG

The discrete scheme for MFP (2.10) reads as follows: given a space-time mesh $\mathcal{I}_h \otimes \mathcal{T}_h$ and a polynomial degree $k \ge 0$, find $\boldsymbol{\alpha}_h, \boldsymbol{\alpha}_h^* \in [W_h^k]^3$, and $\phi_h \in V_h^{k+1}$ such that

$$\sup_{\boldsymbol{\alpha}_h \in [W_h^k]^3} \inf_{\phi_h \in V_h^{k+1}, \boldsymbol{\alpha}_h^* \in [W_h^k]^3} L_{r,h}(\phi_h, \boldsymbol{\alpha}_h, \boldsymbol{\alpha}_h^*), \tag{3.10}$$

where the discrete augmented Lagrangian is

$$\begin{aligned} L_{r,h} := {} & F_h^*(\boldsymbol{\alpha}_h^*) + G_h(\phi_h) + \langle \boldsymbol{\alpha}_h, \nabla_{t,x}\phi_h - \boldsymbol{\alpha}_h^* \rangle_h \\ & + \frac{r}{2}\langle \nabla_{t,x}\phi_h, \nabla_{t,x}\phi_h \rangle - r\langle \nabla_{t,x}\phi_h, \boldsymbol{\alpha}_h^* \rangle_h + \frac{r}{2}\langle \boldsymbol{\alpha}_h^*, \boldsymbol{\alpha}_h^* \rangle_h, \end{aligned} \tag{3.11}$$

in which

$$F_h^*(\boldsymbol{\alpha}_h^*) := \langle A^*(\alpha_{0,h}^* + \frac{1}{2}|\boldsymbol{\alpha}_{1,h}^*|^2), 1\rangle_h, \tag{3.12}$$

$$G_h^*(\phi_h) := -(\phi_h(1,\boldsymbol{x}), \rho_1(\boldsymbol{x}))_h + (\phi_h(0,\boldsymbol{x}), \rho_0(\boldsymbol{x}))_h. \tag{3.13}$$

Note that all terms in the discrete augmented Lagrangian (3.11) are integrated using numerical integration $(\cdot,\cdot)_h$ or $\langle \cdot,\cdot \rangle_h$, except the space-time Laplacian term in the second row of (3.11), which is integrated using exact integration $\langle \cdot,\cdot \rangle$ to avoid a singular matrix for the Laplacian.

Similarly, the discrete scheme for MFG (2.15) reads as follows: given a space-time mesh $\mathcal{I}_h \otimes \mathcal{T}_h$ and a polynomial degree $k \ge 0$, find $\boldsymbol{\alpha}_h, \boldsymbol{\alpha}_h^* \in [W_h^k]^3$, $\rho_{1,h}, \rho_{1,h}^* \in M_h^k$, and $\phi_h \in V_h^{k+1}$ such that

$$\sup_{\boldsymbol{\alpha}_h \in [W_h^k]^3, \rho_{1,h} \in M_h^k} \inf_{\phi_h \in V_h^{k+1}, \boldsymbol{\alpha}_h^* \in [W_h^k]^3, \rho_{1,h}^* \in M_h^k} L_{r,h}^{MFG}(\phi_h, \boldsymbol{\alpha}_h, \rho_{1,h}, \boldsymbol{\alpha}_h^*, \rho_{1,h}^*), \tag{3.14}$$

where the discrete augmented Lagrangian is

$$\begin{aligned} L_{r,h}^{MFG} = {} & F_h^*(\boldsymbol{\alpha}_h^*) + R_h^*(\rho_{1,h}^*) + (\rho_0, \phi_h(0,\cdot))_h \\ & + \langle \boldsymbol{\alpha}_h, \nabla_{t,x}\phi - \boldsymbol{\alpha}_h^* \rangle_h - (\rho_{1,h}, \phi_h(1,\cdot) + \rho_{1,h}^*)_h \\ & + \frac{r_1}{2}\langle \nabla_{t,x}\phi_h, \nabla_{t,x}\phi_h \rangle - r_1\langle \nabla_{t,x}\phi_h, \boldsymbol{\alpha}_h^* \rangle_h + \frac{r_1}{2}\langle \boldsymbol{\alpha}_h^*, \boldsymbol{\alpha}_h^* \rangle_h \\ & + \frac{r_2}{2}(\phi_h(1,\cdot), \phi_h(1,\cdot)) + r_2(\phi_h(1,\cdot), \rho_{1,h}^*)_h + \frac{r_2}{2}(\rho_{1,h}^*, \rho_{1,h}^*)_h, \end{aligned} \tag{3.15}$$

in which

$$R_h^*(\rho_{1,h}^*) := (\Gamma^*(\rho_{1,h}^*), 1)_h. \tag{3.16}$$

3.3. The ALG2 algorithm

The discrete saddle-point problems (3.10) and (3.14) can be solved efficiently using the ALG2 algorithm [22], where minimization of ϕ_h, $\boldsymbol{\alpha}_h^*$, and $\rho_{1,h}^*$ are decoupled. For simplicity, we only illustrate the main steps for the discrete MFG problem (3.14); see also [9,11]. One iteration of ALG2 contains the following three steps.

Step A: update ϕ_h

Minimize $L_{r,h}^{MFG}$ with respect to the first component by solving the elliptic problem: Find $\phi_h^{m+1} \in V_h^{k+1}$ such that it is the solution to

$$\inf_{\phi_h \in V_h^{k+1}} L_{r,h}^{MFG}(\phi_h, \boldsymbol{\alpha}_h^m, \rho_{1,h}^m, \boldsymbol{\alpha}_h^{*,m}, \rho_{1,h}^{*,m}).$$

This is simply a linear, constant-coefficient, space-time diffusion problem: Find $\phi_h^{m+1} \in V_h^{k+1}$ such that

$$r_1\langle\nabla_{t,x}\phi_h^{m+1},\nabla_{t,x}\psi_h\rangle + r_2(\phi_h^{m+1}(1,\cdot),\psi_h(1,\cdot)) \tag{3.17}$$
$$= \langle r_1\boldsymbol{\alpha}_h^{*,m} - \boldsymbol{\alpha}_h^m, \nabla_{t,x}\psi_h\rangle_h - (r_2\rho_{1,h}^{*,m} - \rho_{1,h}^m, \psi_h(1,\cdot))_h - (\rho_0,\psi_h(0,\cdot))_h,$$

for all $\psi_h \in V_h^{k+1}$.

Step B: update $\boldsymbol{\alpha}_h^*$ and $\rho_{1,h}^*$

Minimize $L_{r,h}^{MFG}$ with respect to the last two components by solving the nonlinear problem: Find $\boldsymbol{\alpha}_h^{*,m+1} \in [W_h^k]^3$ and $\rho_{1,h}^{*,m+1} \in M_h^k$ such that they are the solutions to

$$\inf_{\boldsymbol{\alpha}_h^*\in[W_h^k]^3,\rho_{1,h}^*\in M_h^k} L_{r,h}^{MFG}(\phi_h^{m+1},\boldsymbol{\alpha}_h^m,\rho_{1,h}^m,\boldsymbol{\alpha}_h^*,\rho_{1,h}^*).$$

Using the basis functions in (3.8) and (3.9), we write

$$\boldsymbol{\alpha}_h = \sum_{\ell=1}^{M}\sum_{i_s=1}^{N_k}\sum_{j=1}^{N}\sum_{i_t=1}^{k+1}\boldsymbol{a}_{\ell,i_s,j,i_t}\psi_{i_t}^j(t)\varphi_{i_s}^\ell(\boldsymbol{x}), \quad \rho_{1,h} = \sum_{\ell=1}^{M}\sum_{i_s=1}^{N_k}\mathsf{r}_{\ell,i_s}\varphi_{i_s}^\ell(\boldsymbol{x}),$$
$$\boldsymbol{\alpha}_h^* = \sum_{\ell=1}^{M}\sum_{i_s=1}^{N_k}\sum_{j=1}^{N}\sum_{i_t=1}^{k+1}\boldsymbol{a}_{\ell,i_s,j,i_t}^*\psi_{i_t}^j(t)\varphi_{i_s}^\ell(\boldsymbol{x}), \quad \rho_{1,h}^* = \sum_{\ell=1}^{M}\sum_{i_s=1}^{N_k}\mathsf{r}_{\ell,i_s}\varphi_{i_s}^\ell(\boldsymbol{x}),$$

with $\boldsymbol{a}_{\ell,i_s,j,i_t}, \boldsymbol{a}_{\ell,i_s,j,i_t}^*$, r_{ℓ,i_s} and r_{ℓ,i_s}^*.

By the choice of the numerical integration and the nodal bases for W_h^k and M_h^k, we observe that this optimization problem is decoupled for each DOF of $\boldsymbol{\alpha}_h^{*,m+1}$ and $\rho_{1,h}^*$, hence can be efficiently solved pointwisely: for each ℓ, i_s, j, i_t, find $\boldsymbol{a}_{\ell,\mathsf{i_s},j,\mathsf{i_t}}^{*,\mathsf{m+1}} \in \mathbb{R}^3$ such that it solves

$$\inf_{\boldsymbol{a}^*=(\mathsf{a}_0^*,\boldsymbol{a}_1^*)\in\mathbb{R}^3} A^*(\mathsf{a}_0^* + \frac{1}{2}|\boldsymbol{a}_1^*|^2) + \frac{r_1}{2}|\boldsymbol{a}^*|^2 - (\boldsymbol{a}_{\ell,i_s,j,i_t}^m + r_1\nabla_{t,x}\phi_h^{m+1}(\eta_{i_t}^j,\boldsymbol{\xi}_{i_s}^\ell))\cdot\boldsymbol{a}^*, \tag{3.18}$$

and find $\mathsf{r}_{\ell,\mathsf{i_s}}^{*,\mathsf{m+1}} \in \mathbb{R}$ such that it solves

$$\inf_{\mathsf{r}^*\in\mathbb{R}^+} \Gamma^*(\mathsf{r}^*) + \frac{r_2}{2}|\mathsf{r}^*|^2 - (\mathsf{r}_{\ell,i_s}^m - r_2\phi_h^{m+1}(1,\boldsymbol{\xi}_{i_s}^\ell))\cdot\mathsf{r}^*. \tag{3.19}$$

Note that in the optimization problem (3.18), we evaluate the space-time gradient $\nabla_{t,x}\phi_h^{m+1}$ on the quadrature points $(\eta_{i_t}^j,\boldsymbol{\xi}_{i_s}^\ell)$, similarly in (3.19), we evaluate the function ϕ_h^{m+1} on the quadrature points $(1,\boldsymbol{\xi}_{i_s}^\ell)$. Both optimization problems can be efficiently solved in parallel using the Newton's method.

Step C: update $\boldsymbol{\alpha}_h$ and $\rho_{1,h}$

This is a simple pointwise update for the DOFs of the Lagrange multipliers $\boldsymbol{\alpha}_h$ and $\rho_{1,h}$:

$$\boldsymbol{a}_{\ell,i_s,j,i_t}^{m+1} = \boldsymbol{a}_{\ell,i_s,j,i_t}^m + r_1(\nabla_{t,x}\phi_h^{m+1}(\eta_{i_t}^j,\boldsymbol{\xi}_{i_s}^\ell) - \boldsymbol{a}_{\ell,i_s,j,i_t}^{*,m+1}), \tag{3.20}$$
$$\mathsf{r}_{\ell,i_s}^{m+1} = \mathsf{r}_{\ell,i_s}^m - r_2(\phi_h^{m+1}(1,\boldsymbol{\xi}_{i_s}^\ell)) + \mathsf{r}_{\ell,i_s}^{*,m+1}). \tag{3.21}$$

We use the ℓ_∞-errors in the Lagrange multipliers

$$err_m^a := \max_{\ell,i_s,j,i_t}|\boldsymbol{a}_{\ell,i_s,j,i_t}^{m+1} - \boldsymbol{a}_{\ell,i_s,j,i_t}^m|, \tag{3.22}$$
$$err_m^r := \max_{\ell,i_s}|\mathsf{r}_{\ell,i_s}^{m+1} - \mathsf{r}_{\ell,i_s}^m|, \tag{3.23}$$

to monitor the convergence of the ALG2 algorithm.

It is not immediate clear whether the above ALG2 iteration preserves the positivity of density in the discrete level. Our next result shows that density on all quadrature points are guaranteed to be non-negative.

Proposition 3.1. *Let $\boldsymbol{a}_{\ell,i_s,j,i_t}^{m+1}$ and $\mathsf{r}_{\ell,i_s}^{m+1}$ be the solution to the ALG2 algorithm at Step C in (3.20) and (3.21). Then the first component of $\boldsymbol{a}_{\ell,i_s,j,i_t}^{m+1}$, which is density evaluated at the quadrature point $(\eta_{i_t}^j,\boldsymbol{\xi}_{i_s}^\ell)$, is non-negative. Also, $\mathsf{r}_{\ell,i_s}^{m+1}$, the terminal density on the quadrature point $(1,\boldsymbol{\xi}_{i_s}^\ell)$, is non-negative.*

Proof. We just prove the first result via a duality argument since the second result follows from the same procedure. Let $\boldsymbol{a}^{*,m+1}_{\ell,i_s,j,i_t}$ be the solution to the optimization problem (3.18), and let $\boldsymbol{a}^{m+1}_{\ell,i_s,j,i_t}$ be defined in (3.20). We claim that $\boldsymbol{a}^{m+1}_{\ell,i_s,j,i_t}$ is the optimizer to the following primal problem:

$$\inf_{\boldsymbol{a}=(\mathsf{a}_0,\boldsymbol{a}_1)\in\mathbb{R}^3} L(\mathsf{a}_0,\boldsymbol{a}_1)+A(\mathsf{a}_0)+\frac{1}{2r_1}|\boldsymbol{a}-\bar{\boldsymbol{a}}^{m+1}_{\ell,i_s,j,i_t}|^2, \tag{3.24}$$

where $\bar{\boldsymbol{a}}^{m+1}_{\ell,i_s,j,i_t}:=\nabla_{t,x}\phi_h^{m+1}(\eta^j_{i_t},\boldsymbol{\xi}^\ell_{i_s})+\frac{1}{r_1}\boldsymbol{a}^m_{\ell,i_s,j,i_t}$. By the choice of the dynamic cost L in (2.4), it is clear that the optimizer satisfies the non-negativity condition $\mathsf{a}_0\geq 0$, where equality holds only if $\boldsymbol{a}_1$ is also zero.

Let us now prove the claim. Using the definition of the bar-value and (3.20), we have

$$\boldsymbol{a}^{m+1}_{\ell,i_s,j,i_t}=r_1(\bar{\boldsymbol{a}}^{m+1}_{\ell,i_s,j,i_t}-\boldsymbol{a}^{*,m+1}_{\ell,i_s,j,i_t}) \tag{3.25}$$

Using the duality result (2.8), the solution $\boldsymbol{a}^{*,m+1}_{\ell,i_s,j,i_t}$ to the minimization problem is part of the saddle point solution of

$$\inf_{\boldsymbol{a}^*\in\mathbb{R}^3}\sup_{\boldsymbol{a}\in\mathbb{R}^3}\boldsymbol{a}^*\cdot\boldsymbol{a}-L(\mathsf{a}_0,\boldsymbol{a}_1)-A(\mathsf{a}_0)+\frac{r_1}{2}|\boldsymbol{a}^*-\bar{\boldsymbol{a}}^{m+1}_{\ell,i_s,j,i_t}|^2, \tag{3.26}$$

Taking the infinium with respect to $\boldsymbol{a}^*$ first, we get

$$\boldsymbol{a}+r_1(\boldsymbol{a}^*-\bar{\boldsymbol{a}}^{*,m+1}_{\ell,i_s,j,i_t})=0, \tag{3.27}$$

which implies

$$\boldsymbol{a}^*=\bar{\boldsymbol{a}}^{m+1}_{\ell,i_s,j,i_t}-\frac{\boldsymbol{a}}{r_1}.$$

Plugging this expression back to the (3.26) and simplifying, we find that $\boldsymbol{a}$ is the minimizer to the primal problem (3.24). Comparing (3.25) with (3.27) where $\boldsymbol{a}^*=\boldsymbol{a}^{*,m+1}_{\ell,i_s,j,i_t}$, it is now clear that the minimizer to (3.24) is nothing but the solution $\boldsymbol{a}^{m+1}_{\ell,i_s,j,i_t}$. This completes the proof. □

Remark 3.2. We specifically note that the use of the integration rule space W_h^k and numerical integration is crucial for the efficient implementation of Step B in the ALG2 algorithm, which leads to a pointwise update per integration point. If this space and numerical integration were not chosen carefully, additional unnecessary degrees of freedom coupling maybe introduced, which slows down the overall algorithm.

4. Numerical experiments

In this section, we conduct comprehensive experiments to show the efficiency and effectiveness of the proposed numerical algorithms. We restrict ourself to structured (hyper-)rectangular meshes. The case with unstructured meshes will be considered elsewhere. We first numerical verify the convergence of rate of the algorithm related to the mesh size and polynomial degree. Throughout, we take the augmented Lagrangian parameters to be $r=r_1=r_2=1$. Our numerical simulations are performed using the open-source finite-element software NGSolve [52], https://ngsolve.org/.

4.1. Convergence rates

We first consider OT problems with known exact solutions. Specifically, we take the domain $\Omega=\mathbb{R}^d$ with $d=1$ or $d=2$, cost $A(\rho)=0$ in (2.3) with initial and terminal densities:

$$\rho_0(\boldsymbol{x})=\exp(-50|\boldsymbol{x}-\boldsymbol{x}_0|^2),\quad \rho_1(\boldsymbol{x})=\exp(-50|\boldsymbol{x}-\boldsymbol{x}_1|^2),$$

where $\boldsymbol{x}_0=0.25,\boldsymbol{x}_1=0.75$ when spatial dimension $d=1$, and $\boldsymbol{x}_0=(0.25,0.25),\boldsymbol{x}_1=(0.75,0.75)$ when spatial dimension $d=2$. The exact solution is simply a traveling wave solution:

$$\begin{aligned}\rho_{ex}(t,\boldsymbol{x})&=\exp(-50|\boldsymbol{x}-(1+2t)\boldsymbol{x}_0|^2),\\ m_{ex,i}(t,\boldsymbol{x})&=0.5\exp(-50|\boldsymbol{x}-(1+2t)\boldsymbol{x}_0|^2),\quad \forall 1\leq i\leq d,\end{aligned}$$

where $\boldsymbol{m}_{ex}=(m_{ex,1},\cdots.,m_{ex,d})$. We truncate the domain Ω to be a unit box $[0,1]^d$, and replace the homogeneous boundary condition (2.7d) with a boundary source term

$$\boldsymbol{m}\cdot\boldsymbol{n}=\boldsymbol{m}_{ex}\cdot\boldsymbol{n},\quad \text{on } [0,1]\times\partial\Omega.$$

With this modification, the G-term in (2.6a) contains an additional boundary source term:

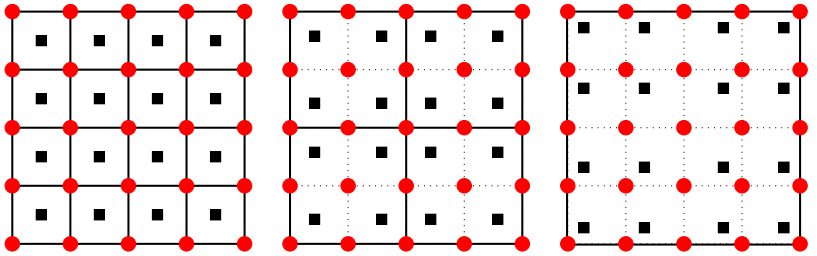


Fig. 1. Coarse mesh DOFs. Circles: DOFs for V_h^{k+1}; Squares: DOFs for W_h^k. The coarse mesh is 4×4 for $k=0$, 2×2 for $k=1$, and 1×1 for $k=3$.

Table 2
Convergence rates of scheme (3.10) applied to 1D OT problem.

k	mesh	L^2-err in ρ	order	L^2-err in $\boldsymbol{m}$	order	W_2^2 error	order
0	4^2	2.068E-01		1.097E-01		2.834E-03	
0	8^2	1.159E-01	0.84	5.985E-02	0.87	5.472E-04	2.37
0	16^2	6.007E-02	0.95	2.970E-02	1.01	5.788E-05	3.24
0	32^2	3.002E-02	1.00	1.497E-02	0.99	4.196E-06	3.79
1	2^2	1.868E-01		1.110E-01		1.127E-02	
1	4^2	7.496E-02	1.32	3.863E-02	1.52	4.625E-04	4.61
1	8^2	2.169E-02	1.78	1.077E-02	1.84	9.523E-06	5.60
1	16^2	5.683E-03	1.93	2.844E-03	1.92	1.611E-07	5.89
3	1^2	2.148E-01		1.301E-01		3.337E-02	
3	2^2	6.602E-02	1.70	3.548E-02	1.87	5.390E-04	5.95
3	4^2	7.234E-03	3.19	3.595E-03	3.30	5.044E-07	10.1
3	8^2	5.079E-04	3.83	2.542E-04	3.82	4.521E-09	6.80

Table 3
Convergence rates of scheme (3.10) applied to 2D OT problem.

k	mesh	L^2-err in ρ	order	L^2-err in $\boldsymbol{m}$	order	W_2^2 error	order
0	4^3	1.172E-01		8.385E-02		1.602E-03	
0	8^3	6.832E-02	0.78	4.879E-02	0.78	1.646E-04	3.28
0	16^3	3.559E-02	0.94	2.505E-02	0.96	2.693E-05	2.61
0	32^3	1.787E-02	0.99	1.262E-02	0.99	2.391E-06	3.49
1	2^3	1.113E-01		8.196E-02		7.008E-03	
1	4^3	4.540E-02	1.29	3.260E-02	1.33	3.882E-05	7.50
1	8^2	1.326E-02	1.78	9.354E-03	1.80	3.563E-06	3.45
1	16^2	3.474E-03	1.93	2.457E-03	1.93	3.854E-08	6.53
3	1^2	1.432E-01		1.109E-01		2.278E-02	
3	2^2	3.873E-02	1.89	2.804E-02	1.98	2.795E-04	6.35
3	4^2	4.353E-03	3.15	3.072E-03	3.19	2.004e-07	10.4
3	8^2	3.068E-04	3.83	2.170E-04	3.82	3.977E-09	5.65

$$G(\phi) := \int_\Omega -\phi(1,\boldsymbol{x})\rho_1(\boldsymbol{x}) + \phi(0,\boldsymbol{x})\rho_0(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_0^1\int_{\partial\Omega} \phi(t,\boldsymbol{x})\boldsymbol{m}_{ex}\cdot\boldsymbol{n}\,\mathrm{d}s\mathrm{d}t.$$

We apply the scheme (3.10) with polynomial degree $k=0,1,3$ on a sequence of uniform hypercubic meshes with $2^{s+2}/(k+1)$ cells in each direction for $s=0,1,2,3$. The total number of DOFs on the s-level meshes is the same for each polynomial degree, which is $2^{(s+2)(d+1)}$ for W_h^k, and $(2^{(s+2)}+1)^{d+1}$ for V_h^{k+1}. So their computational costs are similar. We apply the ALG2 algorithm to (3.10) with a stopping tolerance $err_m^a < 10^{-10}$ where err_m^a is given in (3.22). We take the parameter $r=1$. The DOFs on the coarsest meshes for $d=1$ are shown in Fig. 1.

We record the $L^2(\Omega_T)$-convergence rates of ρ_h and $\boldsymbol{m}_h$, along with the convergence rate of the distance

$$W_2^2 = \int_0^1\int_\Omega \frac{|\boldsymbol{m}|^2}{2\rho}\,\mathrm{d}\boldsymbol{x}\mathrm{d}t$$

in Table 2 for $d=1$, and Table 3 for $d=2$. We find that the convergence behavior for $d=1$ and $d=2$ are similar, in particular, (nearly) optimal L^2-convergence rates of $k+1$ are observed on the finest mesh for each case, and the average

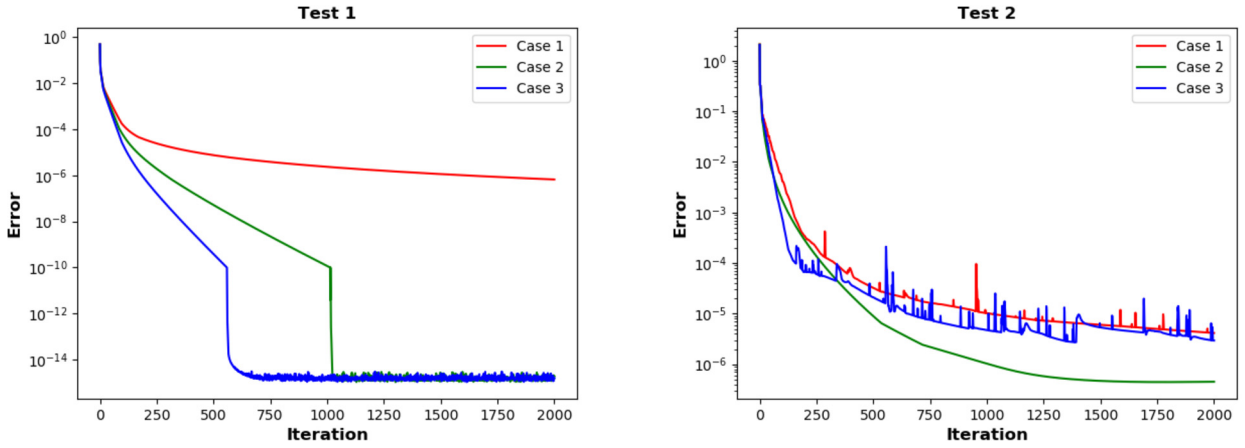


Fig. 2. Example 4.2. Evolution of the ℓ_∞-error err_m^a against number of ALG2 iterations for polynomial degree $k = 3$ on a 32×32 mesh.

Table 4
Example 4.2. Estimated space-time L^2 errors for different polynomial degrees and mesh resolutions. Here N_{DOF} is the total number of density DOFs.

	k	mesh	N_{DOF}	Case 1	Case 2	Case 3
Test 1	0	256×256	65536	2.53E-03	2.29E-03	2.16E-03
	1	32×32	4096	1.09E-03	7.74E-04	6.98E-04
	3	8×8	1024	4.68E-04	4.96E-04	4.45E-04
Test 2	0	256×256	65536	1.28E-02	1.10E-02	1.07E-02
	1	32×32	4096	1.48E-02	3.82E-03	8.57E-03
	3	16×16	4096	1.26E-02	1.07E-03	7.03E-03

convergence rates for the distance W_2^2 are between $2k + 2$ and $2k + 4$. Moreover, the advantage of higher order scheme is clearly observed on the fine meshes where the $k = 3$ case on the 8^{d+1} mesh produces L^2-errors that are 50 times smaller, and W_2^2 error that is three orders of magnitude smaller than the $k = 0$ case on the 32^{d+1} mesh, although the same number of DOFs are used.

4.2. Comparison of high-order and low-order schemes

The previous example demonstrated the advantage of high order methods over low order methods in terms of accuracy when the same number of DOFs are used for pure OT problems. In this example, we show that high order methods can achieve a similar accuracy compared with low order methods with a much less computational cost.

We consider one-dimensional MFP problems (2.3) with domain $\Omega = [0, 1]$. We use three interaction costs, namely, Case 1: $A(\rho) = 0$, Case 2: $A(\rho) = 0.01\rho\log(\rho)$, and Case 3: $A(\rho) = 0.01\rho^2$. Two sets of initial/terminal densities are considered:

$$\begin{cases} \text{Test 1}: \rho_0(x) = 1 + 0.5\sin(2\pi x),\ \rho_1(x) = 1 - 0.5\sin(2\pi x); \\ \text{Test 2}: \rho_0(x) = \max\{\pi\sin(2\pi x), 0\},\ \rho_1(x) = \max\{-\pi\sin(2\pi x), 0\}. \end{cases}$$

Note that the initial and terminal densities for Test 1 are everywhere positive and away from zero, but those for Test 2 are compactly supported with support $[0, 0.5]$ for the initial density and $[0.5, 1]$ for the terminal density. Since no analytic solutions are available for these problems, we compute the reference solution using the scheme (3.10) with polynomial degree $k = 3$ on a 32×32 rectangular mesh for the space-time domain $\Omega_T = [0, 1]^2$. We use a total of 2000 ALG2 iterations. The evolution of the error err_m^a in (3.22) against the number of iterations are recorded in Fig. 2. It is observed that the additional interaction cost $A(\rho)$ leads to faster convergence for ALG2 iterations, especially for Test 1 with positive initial/terminal densities.

We now use these obtained reference solutions to compute the space-time L^2-errors in density approximation for the schemes (3.10) with polynomial degrees $k = 0$, $k = 1$, and $k = 3$ using different mesh resolutions. The mesh resolution for the two tests is as follows. For Test 1, the mesh resolution is 256×256 for polynomial degree $k = 0$, 32×32 for polynomial degree $k = 1$, and 8×8 for polynomial degree $k = 3$. For Test 2, the mesh resolution is 256×256 for polynomial degree $k = 0$, 32×32 for polynomial degree $k = 1$, and 16×16 for polynomial degree $k = 3$. The errors are recorded in Table 4. From this table, we observe that the errors for the high-order scheme with $k = 3$ are always the smallest compared with those for $k = 0$ and $k = 1$, even using a much coarser mesh. In particular, for Test 1, the errors with $k = 3$ using 1024 density DOFs are around 5E-04, which are about 5 times smaller than those with $k = 0$ using 65536 density DOFs; and the errors for Test 2 with $k = 3$ using 4096 density DOFs are similar to the low order method with $k = 0$ using 65536 density DOFs.

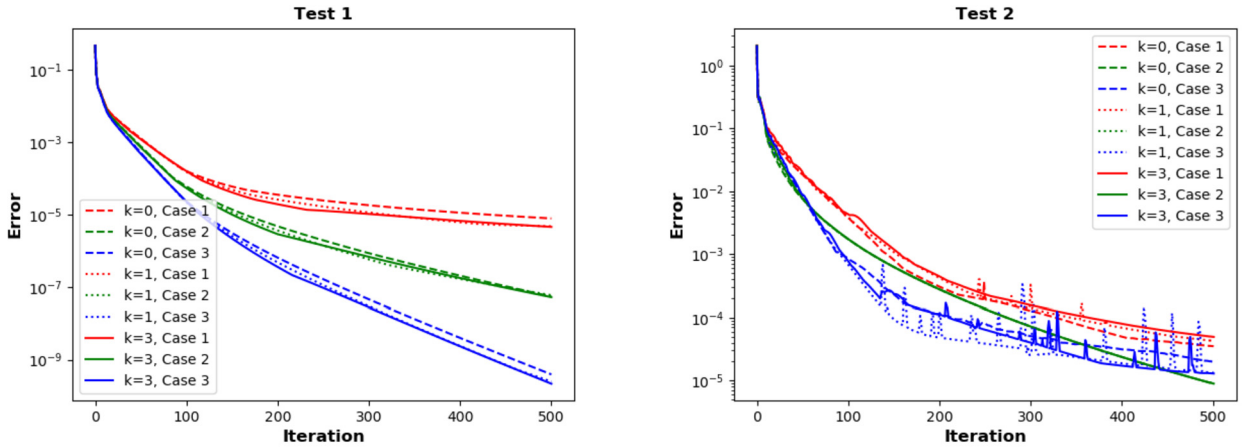


Fig. 3. Example 4.2. Evolution of the ℓ_∞-error err_m^a against the number of ALG2 iterations for different polynomial degrees and mesh resolutions.

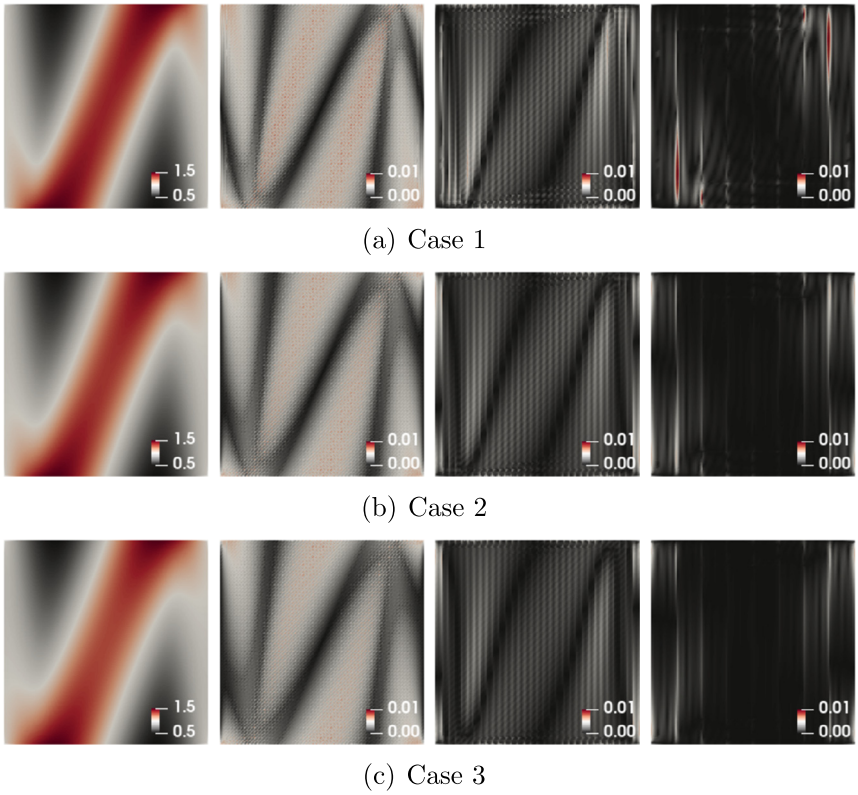
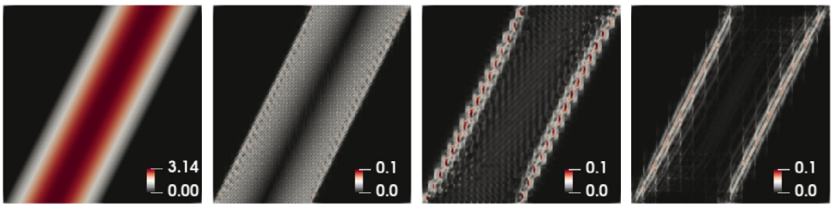


Fig. 4. Example 4.2. Space-time density contours and absolute errors for Test 1. From left to right: reference solution, error for $k = 0$, error for $k = 1$, and error for $k = 3$.

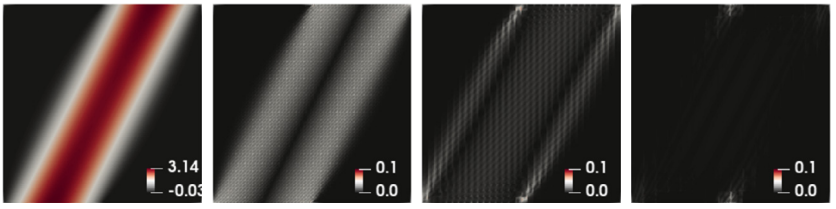
To generate these error tables, we apply 500 ALG2 iterations for all simulations. The convergence curves are shown in Fig. 3. Interestingly, the ALG2 iteration convergence is not sensitive to the mesh resolution nor the polynomial degree. Finally, we plot the error contour along with the reference solutions in Figs. 4–5. In particular, Table 4 and Fig. 5 show that high-order methods still have advantages over low-order methods even when the solution is not smooth and have moving boundaries.

4.3. MFP with obstacles

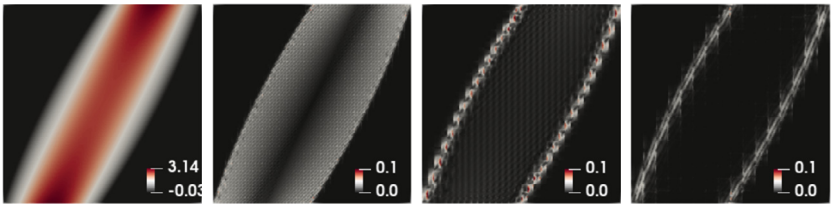
We consider a similar MFP problem used in [15], in which the spatial domain is a square excluding some obstacles that mass can not cross:



(a) Case 1



(b) Case 2



(c) Case 3

Fig. 5. Example 4.2. Space-time density contours and absolute errors for Test 2. From left to right: reference solution, error for $k=0$, error for $k=1$, and error for $k=3$.

$$\Omega = [-1,1]^2 \backslash \{\Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4\},$$

where the obstacles $\Omega_1 = [-0.2, 0.2] \times [-1.0, -0.7]$, $\Omega_2 = [-0.2, 0.2] \times [-0.5, -0.1]$, $\Omega_3 = [-0.2, 0.2] \times [0.1, 0.5]$, $\Omega_4 = [-0.2, 0.2] \times [0.7, 1.0]$. We take initial and terminal densities as two Gaussians

$$\rho_0(\boldsymbol{x}) = \frac{1}{2\pi\sigma^2} \exp(-\frac{1}{2\sigma^2}|\boldsymbol{x}-\boldsymbol{x}_0|^2), \quad \rho_1(\boldsymbol{x}) = \frac{1}{2\pi\sigma^2} \exp(-\frac{1}{2\sigma^2}|\boldsymbol{x}-\boldsymbol{x}_1|^2),$$

where the standard deviation $\sigma = 0.1$, and $\boldsymbol{x}_0 = (-0.65, 0)$, $\boldsymbol{x}_1 = (0.65, 0)$. We take the following 5 choices of interaction cost functions in the MFP problem (2.3), whose convex conjugate are also recorded for completeness:

$$\begin{cases}
\text{Case 1}: A(\rho) = \begin{cases} 0 & \text{if } \rho \geq 0 \\ +\infty & \text{else} \end{cases}, \quad A^*(\rho^*) = \begin{cases} 0 & \text{if } \rho^* \leq 0 \\ +\infty & \text{if } \rho^* > 0 \end{cases}, \\
\text{Case 2}: A(\rho) = \begin{cases} c\rho^2 & \text{if } \rho \geq 0 \\ +\infty & \text{else} \end{cases}, \quad A^*(\rho^*) = \begin{cases} 0 & \text{if } \rho^* \leq 0 \\ (\rho^*)^2/(4c) & \text{if } \rho^* > 0 \end{cases}, \\
\text{Case 3}: A(\rho) = c\rho\log(\rho), \quad A^*(\rho^*) = c\exp(\rho^*/c - 1), \\
\text{Case 4}: A(\rho) = \begin{cases} c/\rho & \text{if } \rho > 0 \\ +\infty & \text{else} \end{cases}, \quad A^*(\rho^*) = \begin{cases} -2\sqrt{-c\rho^*} & \text{if } \rho^* \leq 0 \\ +\infty & \text{if } \rho^* > 0 \end{cases}, \\
\text{Case 5}: A(\rho) = \begin{cases} 0 & \text{if } 0 \leq \rho \leq \rho_{\max} \\ +\infty & \text{else} \end{cases}, \quad A^*(\rho^*) = \begin{cases} 0 & \text{if } \rho^* \leq 0 \\ \rho_{\max}\rho^* & \text{if } \rho^* > 0 \end{cases}.
\end{cases}$$

where we take the scaling constant $c = 0.1$ in Cases 2–4, and maximum density $\rho_{\max} = \frac{1}{2\pi\sigma^2}$ in Case 5.

We apply the scheme (3.10) with polynomial degree $k=3$ on a structured hexahedral mesh obtained from tensor product of a uniform spatial rectangular mesh with mesh size $\Delta x = 0.1$ and uniform temporal mesh with $\Delta t = 0.1$. The spatial mesh for Ω is shown in Fig. 6. We apply 2000 ALG2 iterations and record the evolution of the ℓ_∞-error err_m^a from (3.22) in Fig. 7. As typical of the ALG2 algorithm, we observe a sharp decrease of the errors in the first few iterations followed by a much slower/flatter convergence for all five cases. It is also interesting to find the error curve is very oscillatory for Case 2, while those for Cases 1/3/4 are monotonically decreasing. Moreover, we observe the errors for Cases 2/3 after 2000 iterations

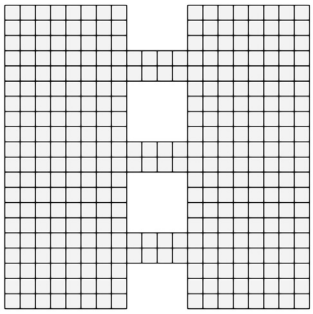


Fig. 6. A uniform rectangular mesh with $\Delta t = 0.1$ for the spatial domain Ω.

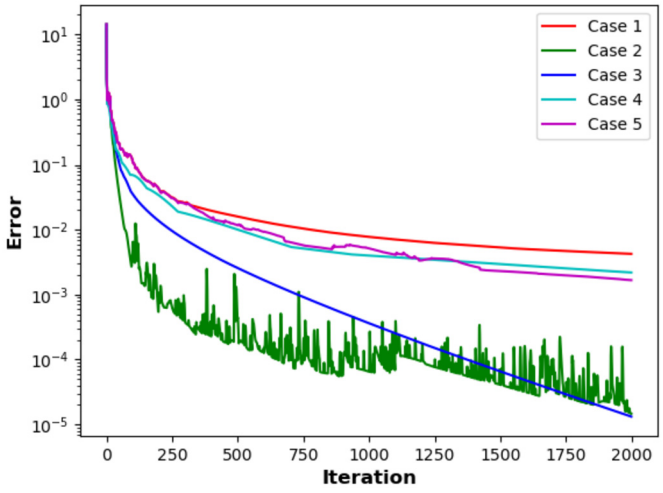


Fig. 7. Example 4.3. Evolution of the ℓ_∞-error err_m^a against number of ALG2 iterations.

are about two orders of magnitude smaller than those for Cases 1/4/5. This indicates the interaction cost $A(\rho)$ plays an important role in the convergence of the ALG2 algorithm.

Snapshots of the density contour at different times are shown in Fig. 8. The effects of different interaction cost functions on the density profile are clearly observed. It is observed from Fig. 8 that the initial ($t = 0$) and terminal ($t = 1$) density approximations are numerically consistent with the Gaussian initial/terminal data for all cases. However, the initial density approximation for Case 3 leads to some visible errors near the right half of the Gaussian bump. We note that the density approximations are defined on the space-time quadrature space W_h^k, hence polynomial interpolation is used to plot these snapshots. The MFP problem conserves the mass $\int_\Omega \rho(t,x)dx = \int_\Omega \rho_0(x)\mathrm{d}\boldsymbol{x}$ for all time $0 \le t \le 1$. But the proposed algorithm (3.10) does not preserve mass conservation. We plot the evolution of the total mass error $|\int_{\Omega_T} \rho_h^m(t,x) - \rho_0(x)\mathrm{d}\boldsymbol{x}\mathrm{d}t|$ for iterations $0 \le m \le 2000$ in Fig. 9. It is observed that the total mass error at the final iteration is reasonable, which is about 4×10^{-7} for Case 2, followed by 2×10^{-5} for Cases 1 and 5, then 2×10^{-4} for Case 3, and 0.01 for Case 4. It is not immediately clear why Case 3 leads to such a large mass conservation error.

4.4. MFG with obstacles

We consider a similar setting as in Example 4.3, where we consider a MFG problem with terminal cost

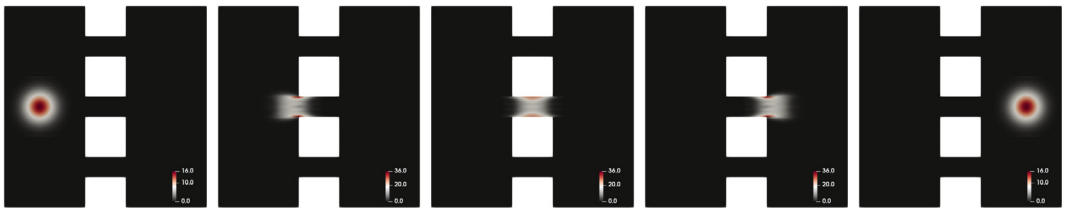
$$\Gamma(\rho) := \begin{cases} \frac{1}{2}(\rho - \rho_T)^2 & \text{if } \rho \ge 0, \\ +\infty & \text{otherwise,} \end{cases}$$

where the target density

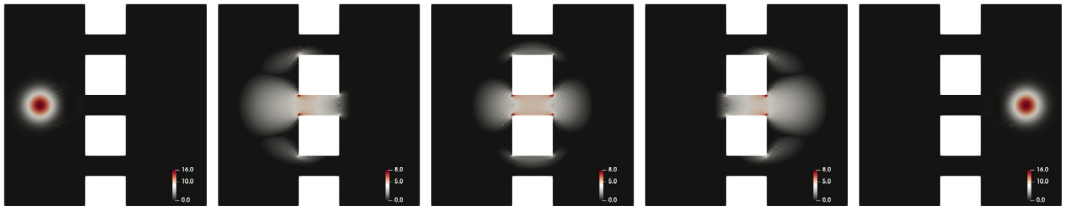
$$\rho_T := \frac{1}{2\pi\sigma^2}\left(\exp(-\frac{1}{2\sigma^2}|\boldsymbol{x} - (0.65, 0.3)|^2) + \exp(-\frac{1}{2\sigma^2}|\boldsymbol{x} - (0.65, -0.3)|^2)\right)$$

with $\sigma = 0.1$. We allow ρ_T and ρ_0 to have different total masses here.

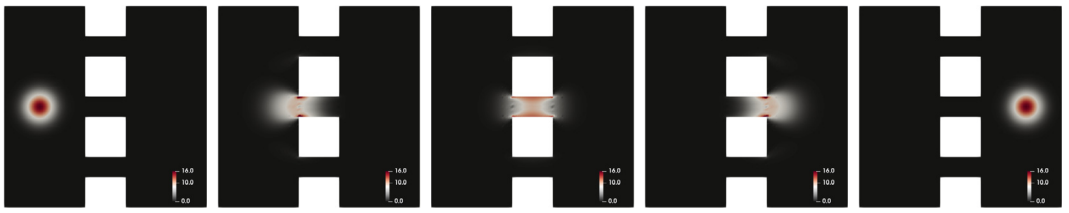
We apply the scheme (3.14) with polynomial degree $k = 3$ on the same mesh as in Example 4.3, and terminate the ALG2 algorithm when the ℓ_∞-error err_m^a is less than 0.01 (relative error is about 10^{-3}). The number of iterations needed for



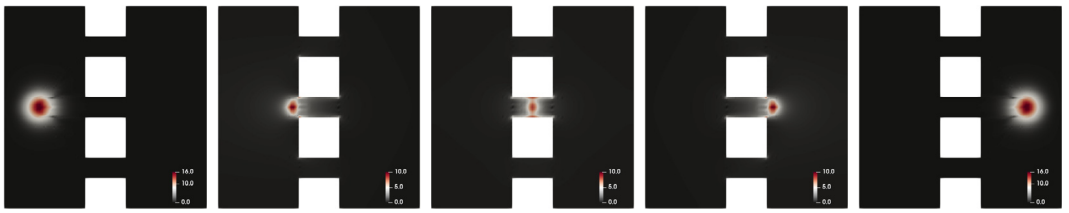
(a) Case 1



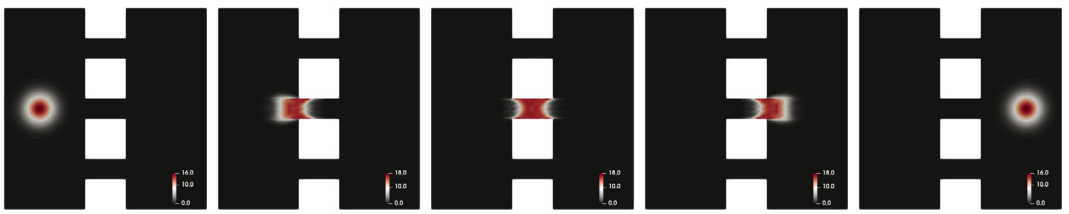
(b) Case 2



(c) Case 3



(d) Case 4



(e) Case 5

Fig. 8. Example 4.3. Snapshots of ρ at $t = 0.0, 0.3, 0.5, 0.7, 1.0$ (left to right).

Table 5
Example 4.4. Number of ALG2 iterations for each case.

	Case 1	Case 2	Case 3	Case 4	Case 5
iterations	3510	82	476	503	798

convergence for the 5 cases is recorded in Table 5, where we find Case 2 has the smallest number of iterations. Furthermore, we plot the convergence curve in Fig. 10 for 2000 ALG2 iterations, which has a similar pattern as Fig. 7 for Example 4.3 above.

Snapshots of the density contour at different times are shown in Fig. 11. The results are similar to those in Example 4.3, where different interaction cost function leads to different density evolution.

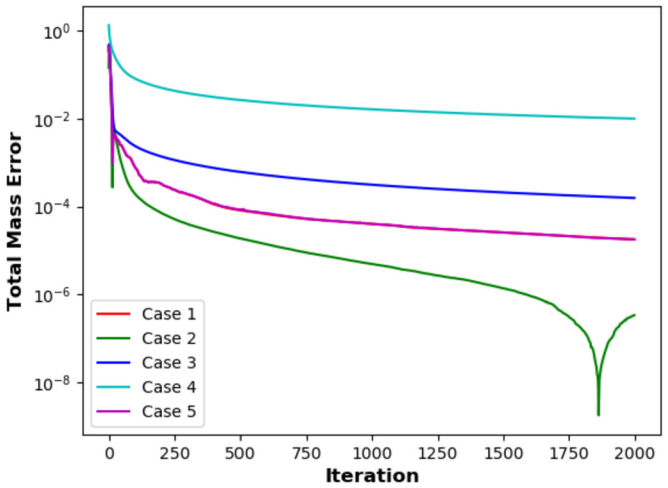


Fig. 9. Example 4.3. Evolution of the total mass error $|\int_{\Omega_T} \rho_h^m(t,x) - \rho_0(x)\mathrm{d}\boldsymbol{x}\mathrm{d}t|$ against number of ALG2 iterations.

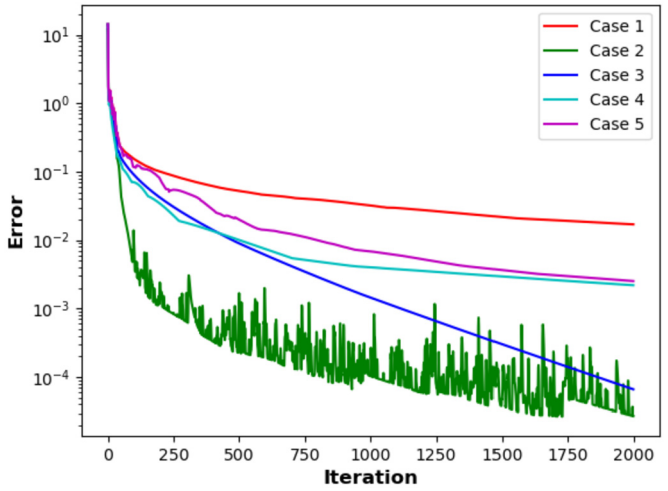


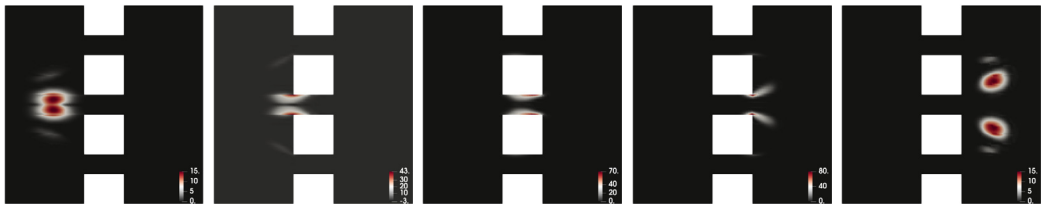
Fig. 10. Example 4.4. Evolution of the ℓ_∞-error err_m^a against number of ALG2 iterations.

4.5. MFP between mascot images

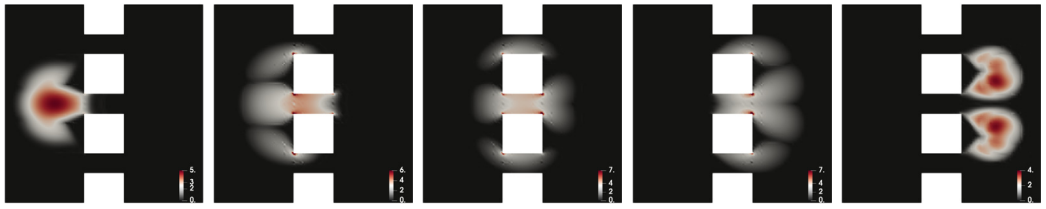
Our last example concerns with OT and MFP (2.3) between images. The initial or terminal densities are normalized images of athletics mascots from University of Notre Dame (Leprechaun), UCLA (Bruins), and University of South Carolina (Gamecocks); see Fig. 13. The spatial domain is a unit square $\Omega = [0,1] \times [0,1]$, and the initial/terminal densities are normalized to have unit mass.

We apply the scheme (3.10) with polynomial degree $k = 3$ on a structured hexahedral mesh of size $64 \times 64 \times 16$, where the time step size is $\Delta t = 1/16$. Three sets of initial/terminal density pairs are considered: (i) ND $\rightarrow$ UCLA where initial density is the ND image and terminal density is the UCLA image, (ii) UCLA$\rightarrow$ USC where initial density is the UCLA image and terminal density is the USC image, and (iii) USC $\rightarrow$ ND where initial density is the USC image and terminal density is the ND image. For each pair of data, we consider three choices of interaction cost, namely, Case 1: $A(\rho) = 0$ (OT), Case 2: $A(\rho) = 0.01\rho\log(\rho)$, and Case 3: $A(\rho) = 0.01/\rho$. The ALG2 algorithm is terminated when err_m^a is less than 0.001. The number of iterations needed for convergence is recorded in Table 6, and the evolution of err_m^a against number of iterations are shown in Fig. 12. We observe the algorithm converges much faster for Cases 2/3 than Case 1 due to additional interaction cost $A(\rho)$.

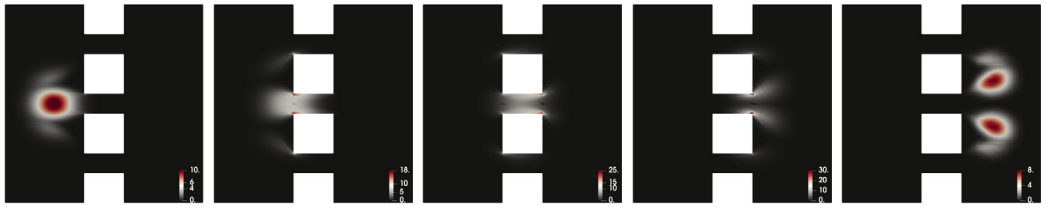
Snapshots of the density contour at different times are shown in Fig. 14 for (i) ND $\rightarrow$ UCLA, in Fig. 15 for (ii) UCLA $\rightarrow$ USC, and in Fig. 16 for (iii) USC$\rightarrow$ ND. We observe in these figures that Case 1 (OT) produces the most sharp results for the density evolution and that both interaction costs in Case 2/3 have a strong smoothing effect which blurs the density profile, where Case 3 with $A(\rho) = 0.01/\rho$ also leads to an everywhere positive density.



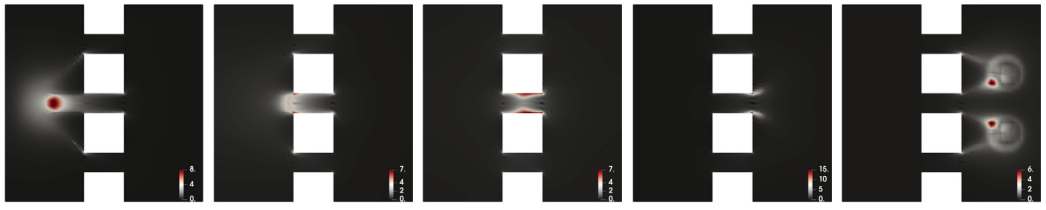
(a) Case 1



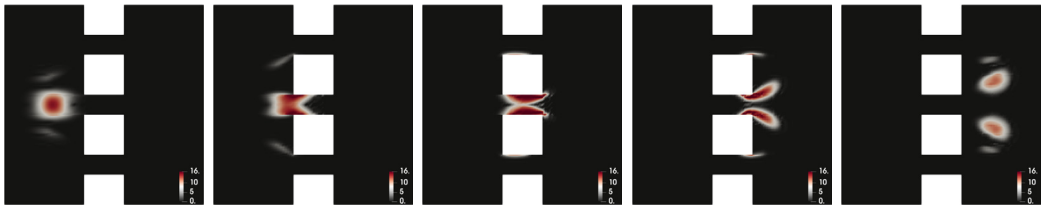
(b) Case 2



(c) Case 3



(d) Case 4



(e) Case 5

Fig. 11. Example 4.4. Snapshots of ρ at $t = 0.1, 0.3, 0.5, 0.7, 0.9$ (left to right).

Table 6
Example 4.5. Number of ALG2 iterations for each case.

	Case 1	Case 2	Case 3
ND→UCLA	2440	471	892
UCLA→USC	1511	211	244
USC→ND	3577	496	907

5. Conclusion

This paper applies high-order accurate finite element methods to compute optimal transport (OT) and mean field games (MFG). To our best knowledge, it is the first time to apply high order numerical methods in OT and MFGs. Despite that achieving convergence for high-order finite element methods is still an open problem, we verify the accuracy of algorithms

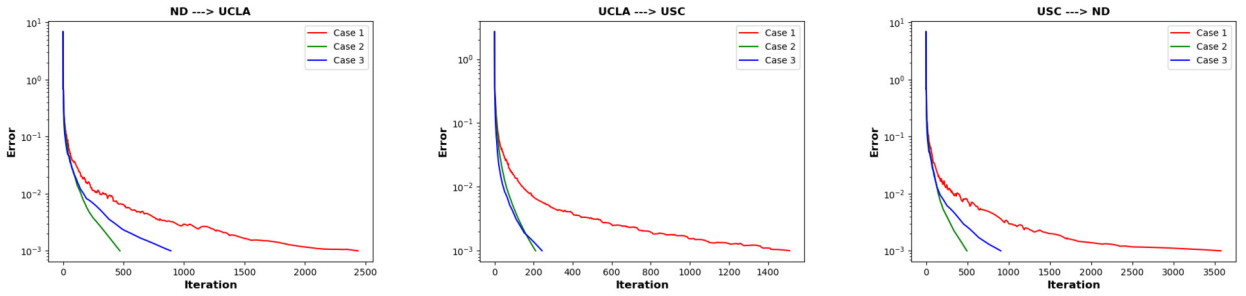
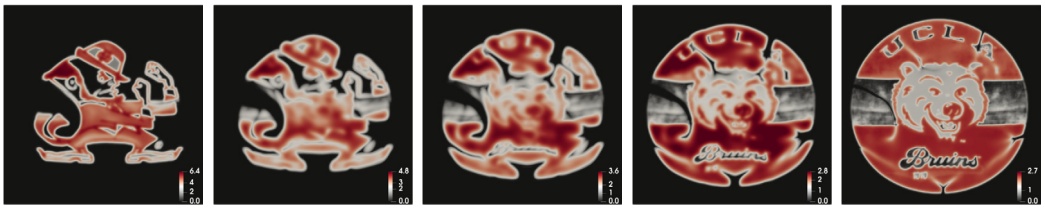


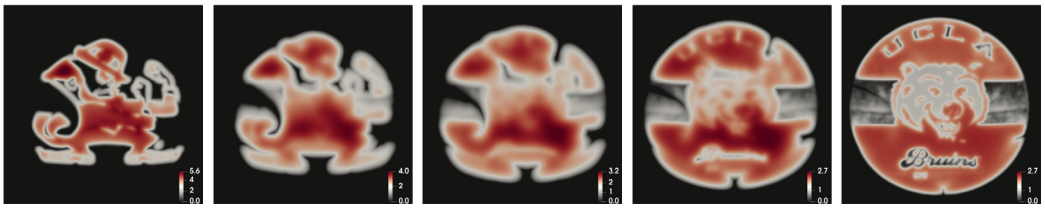
Fig. 12. Example 4.5. Evolution of the ℓ_∞-error err_m^a against number of ALG2 iterations.

(a) ND (Leprechaun) (b) UCLA (Bruins) (c) USC (Gamecocks)

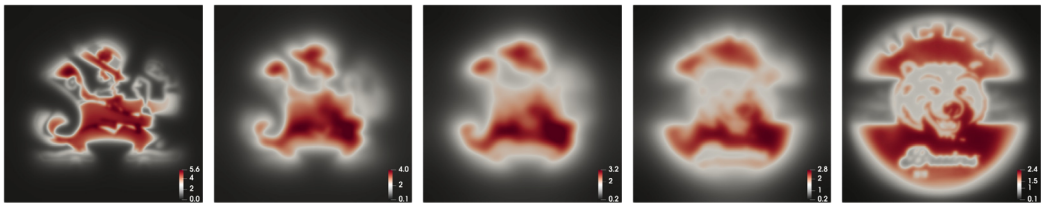
Fig. 13. Example 4.5. Initial/final densities.



(a) Case 1: $A(\rho) = 0$. ND $\rightarrow$ UCLA

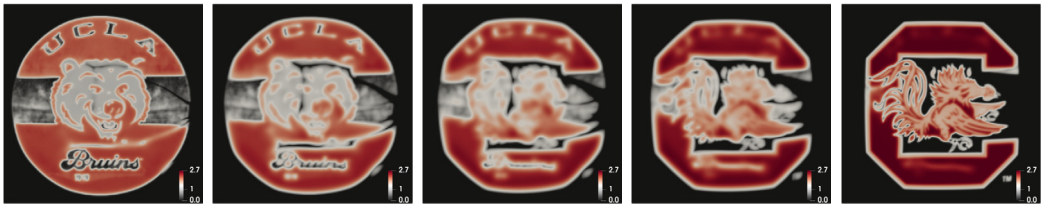


(b) Case 2: $A(\rho) = 0.01\rho\log(\rho)$. ND $\rightarrow$ UCLA

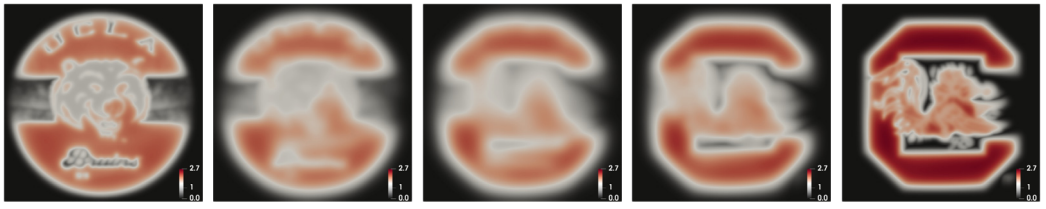


(c) Case 3: $A(\rho) = 0.01/\rho$. ND $\rightarrow$ UCLA

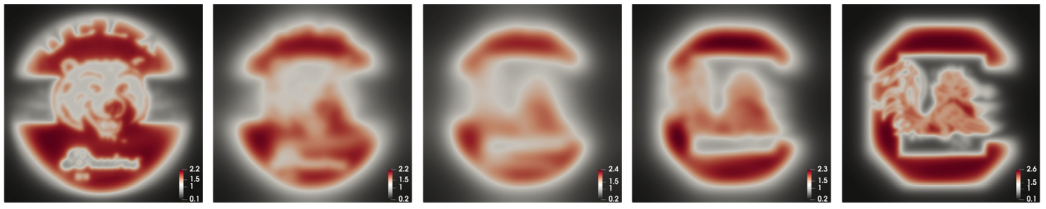
Fig. 14. Example 4.5. Initial density: ND. Terminal density: UCLA. Snapshots of ρ at $t = 0.1,0.3,0.5,0.7,0.9$ (left to right).



(a) Case 1: $A(\rho) = 0$. UCLA $\to$ USC

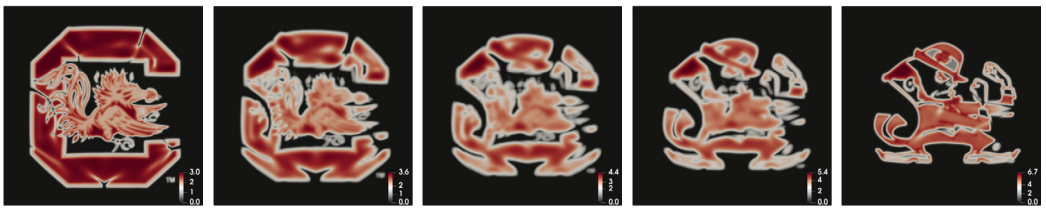


(b) Case 2: $A(\rho) = 0.01\rho\log(\rho)$. UCLA $\to$ USC

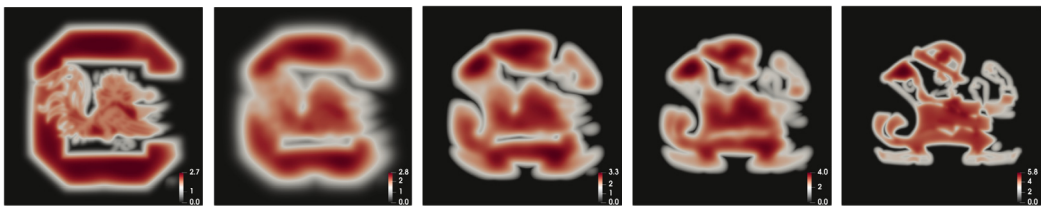


(c) Case 3: $A(\rho) = 0.01/\rho$. UCLA $\to$ USC

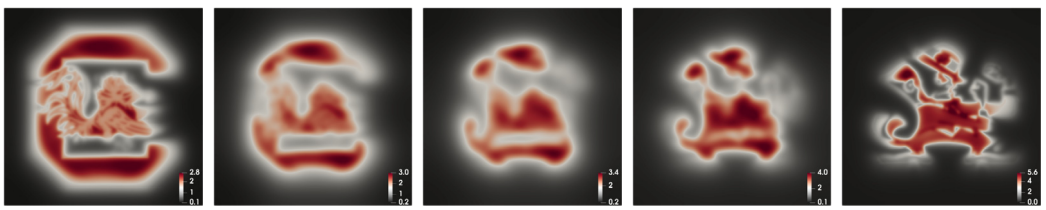
Fig. 15. Example 4.5. Initial density: UCLA. Terminal density: USC. Snapshots of ρ at $t = 0.1,0.3,0.5,0.7,0.9$ (left to right).



(a) Case 1: $A(\rho) = 0$. USC $\to$ ND



(b) Case 2: $A(\rho) = 0.01\rho\log(\rho)$. USC $\to$ ND



(c) Case 3: $A(\rho) = 0.01/\rho$. USC $\to$ ND

Fig. 16. Example 4.5. Initial density: USC. Terminal density: ND. Snapshots of ρ at $t = 0.1,0.3,0.5,0.7,0.9$ (left to right).

through numerical examples. In future works, we shall investigate the numerical property of high-order accuracy FEM methods in OT and MFG-related dynamics. We expect they might have potential applications in computational physics, social science, biology modeling, pandemics control, and computer vision. We also expect to apply high order FEM methods in generalized mean field control formalisms to compute time implicit schemes for fluid dynamics [42,44,43,47].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] Yves Achdou, Italo Capuzzo-Dolcetta, Mean field games: numerical methods, SIAM J. Numer. Anal. 48 (3) (2010) 1136–1162.
[2] Yves Achdou, Pierre Cardaliaguet, François Delarue, Alessio Porretta, Filippo Santambrogio, Yves Achdou, Mathieu Laurière, Mean field games and applications: numerical aspects, in: Mean Field Games, Cetraro, Italy, 2019, 2020, pp. 249–307.
[3] Yves Achdou, Mathieu Laurière, Mean field type control with congestion (II): an augmented Lagrangian method, Appl. Math. Optim. 74 (3) (2016) 535–578.
[4] Sudhanshu Agrawal, Wonjun Lee, Samy Wu Fung, Levon Nurbekyan, Random features for high-dimensional nonlocal mean-field games, J. Comput. Phys. 459 (2022) 111136.
[5] Roman Andreev, Preconditioning the augmented Lagrangian method for instationary mean field games with diffusion, SIAM J. Sci. Comput. 39 (6) (2017) A2763–A2783, MR 3731033.
[6] Alexander Aurell, Boualem Djehiche, Mean-field type modeling of nonlocal crowd aversion in pedestrian crowd dynamics, SIAM J. Control Optim. 56 (1) (2018) 434–455.
[7] Fabio Bagagiolo, Dario Bauso, Mean-field games and dynamic demand management in power grids, Dyn. Games Appl. 4 (2014) 155–176.
[8] J.-D. Benamou, G. Carlier, F. Santambrogio, Variational mean field games, in: Active Particles. vol. 1. Advances in Theory, Models, and Applications, in: Model. Simul. Sci. Eng. Technol., Birkhäuser/Springer, Cham, 2017, pp. 141–171, MR 3644590.
[9] Jean-David Benamou, Yann Brenier, A numerical method for the optimal time-continuous mass transport problem and related problems, Contemp. Math. 226 (1999) 1–12.
[10] Jean-David Benamou, Yann Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, Numer. Math. 84 (3) (2000) 375–393.
[11] Jean-David Benamou, Guillaume Carlier, Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations, J. Optim. Theory Appl. 167 (1) (2015) 1–26, MR 3395203.
[12] Jean-David Benamou, Guillaume Carlier, Maxime Laborde, An augmented Lagrangian approach to Wasserstein gradient flows and applications, ESAIM Proc. Surv. 54 (2016) 1–17.
[13] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al., Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (1) (2011) 1–122.
[14] L. Briceño Arias, D. Kalise, Z. Kobeissi, M. Laurière, Á. Mateos González, F.J. Silva, On the implementation of a primal-dual algorithm for second order time-dependent mean field games with local couplings, in: CEMRACS 2017—Numerical Methods for Stochastic Models: Control, Uncertainty Quantification, Mean-Field, in: ESAIM Proc. Surveys, vol. 65, EDP Sci., Les Ulis, 2019, pp. 330–348, MR 3968547.
[15] G. Buttazzo, C. Jimenez, E. Oudet, An optimization problem for mass transportation with congested dynamics, SIAM J. Control Optim. 48 (3) (2009) 1961–1976, MR 2516195.
[16] P. Cardaliaguet, Notes on mean field games, https://www.ceremade.dauphine.fr/~cardaliaguet/, 2013.
[17] Pierre Cardaliaguet, Saeed Hadikhanloo, Learning in mean field games: the fictitious play, ESAIM Control Optim. Calc. Var. 23 (2) (2017) 569–591.
[18] Elisabetta Carlini, Francisco J. Silva, A fully discrete semi-Lagrangian scheme for a first order mean field game problem, SIAM J. Numer. Anal. 52 (1) (2014) 45–67.
[19] René Carmona, Jean-Pierre Fouque, Li-Hsien Sun, Mean field games and systemic risk, Commun. Math. Sci. 13 (4) (2015) 911–933.
[20] Philippe Casgrain, Sebastian Jaimungal, Mean-field games with differing beliefs for algorithmic trading, Math. Finance 30 (3) (2020) 995–1034.
[21] Kai Cui, Heinz Koeppl, Approximately solving mean field games via entropy-regularized deep reinforcement learning, in: International Conference on Artificial Intelligence and Statistics, PMLR, 2021, pp. 1909–1917.
[22] Michel Fortin, Roland Glowinski, Augmented Lagrangian methods, in: Applications to the Numerical Solution of Boundary Value Problems, in: Studies in Mathematics and Its Applications, vol. 15, North-Holland Publishing Co., Amsterdam, 1983, translated from the French by B. Hunt and D.C. Spicer. MR 724072.
[23] Wilfrid Gangbo, Alpár R. Mészáros, Global well-posedness of master equations for deterministic displacement convex potential mean field games, Commun. Pure Appl. Math. 75 (12) (2022) 2685–2801 (en).
[24] Wilfrid Gangbo, Alpár R. Mészáros, Chenchen Mou, Jianfeng Zhang, Mean field games master equations with nonseparable Hamiltonians and displacement monotonicity, Ann. Probab. 50 (6) (2022) 2178–2217.
[25] Peter Gladbach, Eva Kopfer, Jan Maas, Scaling limits of discrete optimal transport, SIAM J. Math. Anal. 52 (3) (2020) 2759–2802.
[26] Diogo A. Gomes, J. Saúde, Mean field games models—a brief survey, Dyn. Games Appl. 4 (2) (2014) 110–154, MR 3195844.
[27] Olivier Guéant, Jean-Michel Lasry, Pierre-Louis Lions, Mean field games and applications, in: Paris-Princeton Lectures on Mathematical Finance 2010, in: Lecture Notes in Math., vol. 2003, Springer, Berlin, 2011, pp. 205–266, MR 2762362.
[28] Xin Guo, Anran Hu, Renyuan Xu, Junzi Zhang, A general framework for learning mean-field games, Math. Oper. Res. 48 (2) (2023) 656–686.
[29] Saeed Hadikhanloo, Francisco J. Silva, Finite mean field games: fictitious play and convergence to a first order continuous mean field game, J. Math. Pures Appl. 132 (2019) 369–397.
[30] Minyi Huang, Roland P. Malhamé, Peter E. Caines, et al., Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, Commun. Inf. Syst. 6 (3) (2006) 221–252.
[31] Noureddine Igbida, Van Thanh Nguyen, Augmented Lagrangian method for optimal partial transportation, IMA J. Numer. Anal. 38 (1) (2018) 156–183, MR 3800018.

[32] Arman C. Kizilkale, Rabih Salhab, Roland P. Malhamé, An integral control formulation of mean field game based large scale coordination of loads in smart grids, Automatica 100 (2019) 312–322.
[33] Aimé Lachapelle, Marie-Therese Wolfram, On a mean field game approach modeling congestion and aversion in pedestrian crowds, Transp. Res., Part B, Methodol. 45 (10) (2011) 1572–1589.
[34] Jean-Michel Lasry, Pierre-Louis Lions, Mean field games, Jpn. J. Math. 2 (1) (2007) 229–260.
[35] Jean-Michel Lasry, Pierre-Louis Lions, Mean field games, Jpn. J. Math. 2 (1) (2007) 229–260, MR 2295621.
[36] Mathieu Lauriere, Numerical methods for mean field games and mean field type control, in: Mean Field Games, vol. 78, 2021, p. 221.
[37] Mathieu Laurière, Sarah Perrin, Sertan Girgin, Paul Muller, Ayush Jain, Theophile Cabannes, Georgios Piliouras, Julien Pérolat, Romuald Élie, Olivier Pietquin, et al., Scalable deep reinforcement learning algorithms for mean field games, in: International Conference on Machine Learning, PMLR, 2022, pp. 12078–12095.
[38] Hugo Lavenant, Unconditional convergence for discretizations of dynamical optimal transport, Math. Comput. 90 (328) (2021) 739–786.
[39] Hugo Lavenant, Sebastian Claici, Edward Chien, Justin Solomon, Dynamical optimal transport on discrete surfaces, ACM Trans. Graph. (TOG) 37 (6) (2018) 1–16.
[40] Wonjun Lee, Siting Liu, Wuchen Li, Stanley Osher, Mean field control problems for vaccine distribution, Res. Math. Sci. 9 (3) (2022) 51.
[41] Wonjun Lee, Siting Liu, Hamidou Tembine, Wuchen Li, Stanley Osher, Controlling propagation of epidemics via mean-field control, SIAM J. Appl. Math. 81 (1) (2021) 190–207.
[42] Wuchen Li, Wonjun Lee, Stanley Osher, Computational mean-field information dynamics associated with reaction-diffusion equations, J. Comput. Phys. 466 (2022) 111409.
[43] Wuchen Li, Siting Liu, Stanley Osher, Controlling conservation laws II: compressible Navier–Stokes equations, J. Comput. Phys. 463 (2022) 111264.
[44] Wuchen Li, Siting Liu, Stanley Osher, Controlling conservation laws I: entropy–entropy flux, J. Comput. Phys. 480 (2023) 112019.
[45] Alex Tong Lin, Samy Wu Fung, Wuchen Li, Levon Nurbekyan, Stanley J. Osher, Alternating the population and control neural networks to solve high-dimensional stochastic mean-field games, Proc. Natl. Acad. Sci. 118 (31) (2021) e2024713118.
[46] Siting Liu, Matthew Jacobs, Wuchen Li, Levon Nurbekyan, Stanley J. Osher, Computational methods for first-order nonlocal mean field games with applications, SIAM J. Numer. Anal. 59 (5) (2021) 2639–2668.
[47] Siting Liu, Stanley Osher, Wuchen Li, Chi-Wang Shu, A primal-dual approach for solving conservation laws with implicit in time approximations, J. Comput. Phys. 472 (2023) 111654.
[48] Andrea Natale, Gabriele Todeschi, A mixed finite element discretization of dynamical optimal transport, J. Sci. Comput. 91 (2) (2022) 38.
[49] Nicolas Papadakis, Gabriel Peyré, Edouard Oudet, Optimal transport with proximal splitting, SIAM J. Imaging Sci. 7 (1) (2014) 212–238.
[50] Alessio Porretta, On the planning problem for a class of mean field games, C. R. Math. 351 (11–12) (2013) 457–462.
[51] Lars Ruthotto, Stanley J. Osher, Wuchen Li, Levon Nurbekyan, Samy Wu Fung, A machine learning framework for solving high-dimensional mean field game and mean field control problems, Proc. Natl. Acad. Sci. 117 (17) (2020) 9183–9193.
[52] J. Schöberl, C++11 Implementation of Finite Elements in NGSolve, ASC Report 30/2014, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2014.
[53] F.D. Witherden, P.E. Vincent, On the identification of symmetric quadrature rules for finite element methods, Comput. Math. Appl. 69 (10) (2015) 1232–1241, MR 3333661.
[54] Jiajia Yu, Rongjie Lai, Wuchen Li, Stanley Osher, Computational mean-field games on manifolds, arXiv:2206.01622, 2022, N/A.
[55] Linbo Zhang, Tao Cui, Hui Liu, A set of symmetric quadrature rules on triangles and tetrahedra, J. Comput. Math. 27 (1) (2009) 89–96, MR 2493559.