

# A Primal-dual hybrid gradient method for solving optimal control problems and the corresponding Hamilton-Jacobi PDEs

Tingwei Meng · Siting Liu · Wuchen Li · Stanley Osher

Received: date / Accepted: date

**Abstract** Optimal control problems are crucial in various domains, including path planning, robotics, and humanoid control, demonstrating their broad applicability. The connection between optimal control and Hamilton-Jacobi (HJ) partial differential equations (PDEs) underscores the need for solving HJ PDEs to address these control problems effectively. While numerous numerical methods exist for tackling HJ PDEs across different dimensions, this paper introduces an innovative optimization-based approach that reformulates optimal control problems and HJ PDEs into a saddle point problem using a Lagrange multiplier. Our method, based on the preconditioned primal-dual hybrid gradient (PDHG) method, offers a solution to HJ PDEs with first-order accuracy and numerical unconditional stability, enabling larger time steps and avoiding the limitations of explicit time discretization methods. Our approach has ability to handle a wide variety of Hamiltonian functions, including those

---

T. Meng, S. Liu and S. Osher are partially funded by Air Force Office of Scientific Research (AFOSR) MURI FA9550-18-502 and Office of Naval Research (ONR) N00014-20-1-2787. W. Li is supported by Air Force Office of Scientific Research (AFOSR) MURI FP 9550-18-502, Air Force Office of Scientific Research (AFOSR) YIP award No. FA9550-23-1-0087, National Science Foundation (NSF) DMS-2245097, and National Science Foundation (NSF) RTG: 2038080.

---

Tingwei Meng  
Mathematics Department, University of California, Los Angeles  
E-mail: tingwei@math.ucla.edu

Siting Liu  
Mathematics Department, University of California, Los Angeles  
E-mail: siting6@math.ucla.edu

Stanley Osher  
Mathematics Department, University of California, Los Angeles  
E-mail: sjo@math.ucla.edu

Wuchen Li  
Mathematics Department, University of South Carolina  
E-mail: wuchen@mailbox.sc.edu

that are non-smooth and dependent on time and space, through a simplified saddle point formulation that facilitates easy and parallelizable updates. Furthermore, our framework extends to viscous HJ PDEs and stochastic optimal control problems, showcasing its versatility. Through a series of numerical examples, we demonstrate the method's effectiveness in managing diverse Hamiltonians and achieving efficient parallel computation, highlighting its potential for wide-ranging applications in optimal control and beyond.

**Keywords** saddle point problems · optimal control problems · Hamilton-Jacobi PDEs · primal-dual hybrid gradient algorithms · time-implicit scheme

**Mathematics Subject Classification (2020)** 49L12 · 93E20 · 49M25

## 1 Introduction

Optimal control problems are pivotal in numerous practical applications, such as trajectory planning [21, 89, 49, 30, 86, 72], robot manipulator control [74, 58, 62, 39, 16], and humanoid robot control [36, 38, 66, 42, 37, 31], highlighting their significance across various fields. The relationship between optimal control and Hamilton-Jacobi (HJ) partial differential equations (PDEs) is well-established (see [6, 94]), underscoring the importance of solving HJ PDEs for addressing optimal control challenges. Various numerical strategies have been developed to tackle HJ PDEs and optimal control problems, ranging from high-order grid-based methods in lower-dimensional settings to innovative approaches designed to overcome the curse of dimensionality in higher dimensions. For lower-dimensional problems, advanced grid-based methods like essentially nonoscillatory (ENO) schemes [84], weighted ENO schemes [57], and the discontinuous Galerkin method [51] are frequently utilized. To tackle the complexity of higher dimensions, several sophisticated strategies have been introduced. These approaches encompass a diverse range of methodologies, including max-plus algebra methods [78, 1, 2, 34, 40, 44, 79, 80, 81], dynamic programming and reinforcement learning [3, 8], tensor decomposition techniques [33, 50, 91], sparse grids [9, 43, 61], model order reduction [4, 67], polynomial approximation [59, 60], optimization methods [22, 29, 64, 63, 26, 28, 93, 19, 20, 75, 17, 18], and neural network-based solutions [25, 5, 32, 56, 47, 53, 54, 69, 83, 87, 88, 90, 27, 23, 95, 96].

This paper introduces a novel optimization-based framework with time-implicit discretization to solve optimal control problems and HJ PDEs by reformulating them into a saddle point problem using a Lagrange multiplier. We propose an algorithm based on the preconditioned primal-dual hybrid gradient (PDHG) method [15] for finding the saddle point, effectively solving the HJ PDEs. Unlike other grid-based methods, our approach achieves first-order accuracy and numerical unconditional stability through implicit time discretization, allowing for larger time steps and avoiding the Courant–Friedrichs–Lewy (CFL) condition limitation of explicit time discretization methods.

Our method distinguishes itself within optimization-based approaches by its capability to handle a broad set of Hamiltonians, notably including those

that exhibit non-smooth behavior and dependencies on both time and space variables. The saddle point formulation simplifies iteration updates, eliminating the need for complex nonlinear inversion. The most intricate update step can be executed efficiently and in parallel through the proximal point operator. For scenarios frequently encountered that involve quadratic or 1-homogeneous Hamiltonians, possibly dependent on  $(x, t)$ , the proximal point operator can be straightforwardly implemented using explicit expressions. We provide a convergence analysis of the PDHG algorithm under certain conditions and establish connections with traditional numerical schemes for HJ PDEs. Furthermore, our approach is applicable to viscous HJ PDEs and corresponding stochastic optimal control problems, demonstrating its versatility and broad applicability.

The technique of converting PDE problems into saddle point problems and solving them with the PDHG method has been effectively applied to reaction-diffusion equations [76, 41, 14], conservation laws [77], and HJ PDEs [82]. While the reaction-diffusion and conservation law studies employ PDHG alongside integration by parts for solving initial value problems, this method is not readily applicable to HJ PDEs without modifications. The approach for HJ PDEs in [82] employs the Fenchel-Legendre transform within its saddle point formulation, leading to an objective function that is linear with respect to the HJ PDE solution  $\varphi$ , thereby avoiding nonlinear updates on  $\varphi$ . This method also facilitates the use of larger time steps thanks to implicit temporal Euler discretization. However, extracting optimal controls directly from HJ PDE solutions necessitates additional steps. Expanding upon this groundwork, our research specifically focuses on optimal control problems by proposing a dedicated saddle point formulation and algorithm. Unlike the previous work, our saddle point formulation is derived from the intrinsic link between optimal control problems and HJ PDEs, not the Fenchel-Legendre transform. This adjustment retains the advantages of the earlier method while making it more appropriate for optimal control issues. Our method's simplicity in the saddle point formulation allows for updates that are either explicitly defined or amenable to parallel computation, thus broadening its utility.

In addition, our proposed saddle point formulation establishes a connection with the primal-dual formulation of potential mean-field games (MFGs), a concept introduced in [70, 52] for modeling the equilibrium behavior of large populations of agents engaged in strategic interactions. This framework has been applied across a variety of fields [68, 46, 45, 13, 12, 65, 35, 73]. An MFG system is characterized by coupled PDEs, comprising a forward-evolving Fokker-Planck equation and a backward-evolving HJ equation. In our research, we utilize this interconnection to approach the solution of HJ PDEs. By drawing on this link, we can apply methods initially devised for MFGs to the task of solving HJ PDEs that have specific initial conditions. For example, the PDHG method, which has been used to solve potential MFGs through a saddle point formulation [85, 11, 10], is adapted in this work to solve optimal control problems and HJ PDEs.

To evaluate the effectiveness of our proposed methodology, we present a series of numerical examples in both one-dimensional and two-dimensional settings<sup>1</sup>. These examples demonstrate the method’s capability in handling diverse Hamiltonians that depend on the spatial variable. A notable feature of our method is that, during each iteration, the updates at individual points are decoupled, facilitating the application of parallel computing techniques to expedite the computation process. Additionally, we showcase an example where the Lagrangian is defined by an indicator function, imposing control constraints without introducing running costs. This scenario yields a Hamiltonian  $H(x, t, p)$  that is 1-homogeneous with respect to  $p$  for any given spatial variable  $x$  and temporal variable  $t$ . The numerical outcomes reveal the emergence of discontinuous feedback control functions and bang-bang control strategies. Such phenomena are also observed in the stochastic optimal control problem variants, where, despite the smoothness of the solution to the viscous HJ PDE, the control function remains non-smooth, with evident discontinuities in the sampled open-loop control trajectories.

This paper is structured as follows: Section 2 lays out the saddle point problem and introduces our algorithm tailored for optimal control problems and their related HJ PDEs. We delve into the formulation and algorithm within function spaces in Section 2.1 and proceed to discuss spatial and temporal discretization strategies for one-dimensional cases in Section 2.2.1 and two-dimensional cases in Section 2.2.2. Several numerical experiments are shown in Section 2.3. Section 3 proposes similar saddle point formulation and algorithm for stochastic optimal control problems and the corresponding viscous HJ PDEs, with further numerical experiments presented in Section 3.2. The paper wraps up with a summary and potential directions for future work in Section 4. The appendices provide in-depth technical insights: Appendix A examines our algorithm’s convergence analysis for specific scenarios, and Appendix B connects our methodology to traditional numerical solvers for HJ PDEs. Discretization details for the saddle point formulation and algorithm applied to viscous HJ PDEs and stochastic control problems are outlined in Appendix C.

## 2 Optimal control problems

The optimal control problems are in the following form

$$\min_{\alpha(\cdot)} \left\{ \int_t^T L(\gamma(s), s, \alpha(s)) ds + g(\gamma(T)) : \gamma(t) = x, \dot{\gamma}(s) = f(\gamma(s), s, \alpha(s)) \forall s \in (t, T) \right\}, \quad (1)$$

where  $\alpha: [t, T] \rightarrow \mathbb{R}^m$  is called control,  $\gamma: [t, T] \rightarrow \Omega$  is called state ( $\Omega \subseteq \mathbb{R}^n$  is the state space),  $f: \Omega \times [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  is the source term of ODE constraint,  $L: \Omega \times [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$  is called Lagrangian, and  $g: \Omega \rightarrow \mathbb{R}$  is a terminal cost. Assume these functions are all continuous and satisfy certain conditions such that the optimal control problem has a unique minimizer.

<sup>1</sup> Codes are available at <https://github.com/TingweiMeng/PDHG-optimal-control>.

Denote the optimal value in (1) by  $\phi(x, t)$ . Then the function  $\phi$  solves the following Hamilton-Jacobi (HJ) PDE

$$\begin{cases} \frac{\partial \phi(x, t)}{\partial t} + \inf_{\alpha \in \mathbb{R}^m} \{ \langle f(x, t, \alpha), \nabla_x \phi(x, t) \rangle + L(x, t, \alpha) \} = 0, & x \in \Omega, t \in [0, T], \\ \phi(x, T) = g(x), & x \in \Omega, \end{cases}$$

where the terminal condition is given by the terminal cost  $g$  in (1), and  $\Omega$  is the state space in the optimal control problem. Throughout this paper, we predominantly consider  $\Omega$  to be a rectangular domain subject to periodic boundary conditions. Nonetheless, our methodology is adaptable to alternative boundary conditions, as illustrated in Section 2.3.3. Utilizing a time reversal strategy and setting  $\varphi(x, t) = \phi(x, T - t)$ , we derive the initial value HJ PDE as follows:

$$\begin{cases} \frac{\partial \varphi(x, t)}{\partial t} + H(x, t, \nabla_x \varphi(x, t)) = 0, & x \in \Omega, t \in [0, T], \\ \varphi(x, 0) = g(x), & x \in \Omega, \end{cases} \quad (2)$$

where the Hamiltonian  $H$  is defined by

$$H(x, t, p) = \sup_{\alpha \in \mathbb{R}^m} \{ -\langle f(x, T - t, \alpha), p \rangle - L(x, T - t, \alpha) \}. \quad (3)$$

This Hamiltonian  $H$  is convex with respect to  $p$  for any  $x \in \Omega$  and  $t \in [0, T]$ . From the HJ PDE (2), we obtain the function  $\alpha: \Omega \times [0, T] \rightarrow \mathbb{R}^m$  by

$$\alpha(x, t) = \arg \max_{\alpha \in \mathbb{R}^m} \{ -\langle f(x, T - t, \alpha), \nabla_x \varphi(x, t) \rangle - L(x, T - t, \alpha) \}. \quad (4)$$

Following the application of time reversal, this function yields the feedback control function represented by  $\alpha(x, T - t)$ . The optimal trajectory  $\gamma^*$  and optimal open-loop control  $\alpha^*$  in the optimal control problem (1) are then computed by

$$\begin{cases} \dot{\gamma}^*(s) = f(\gamma^*(s), s, \alpha(\gamma^*(s), T - s)), & s \in (t, T), \\ \gamma^*(t) = x, \\ \alpha^*(s) = \alpha(\gamma^*(s), T - s), & s \in [t, T]. \end{cases} \quad (5)$$

For additional insights into the mathematical relationship between optimal control problems and HJ PDEs, we refer readers to [6].

## 2.1 Saddle point formulation

We introduce a saddle point formulation to solve the optimal control problem (1) and its associated initial value HJ PDE (2). Following this, we propose a numerical algorithm derived from the PDHG method [15, 92, 24] to effectively solve the saddle point problem. To simplify the notation, we will use  $f_{x,t}$  and  $L_{x,t}$  to denote the functions  $f(x, T - t, \cdot)$  and  $L(x, T - t, \cdot)$ , respectively.

We treat the HJ PDE given in (2) as a constraint within an optimization problem, where the objective function is defined as  $-c \int_{\Omega} \varphi(x, T) dx$ . This specific objective function is adopted based on insights from numerical experiments detailed in [82]. By employing the Lagrange multiplier  $\rho$ , we derive:

$$\begin{aligned}
& \min_{\varphi \text{ satisfying (2)}} -c \int_{\Omega} \varphi(x, T) dx \\
&= \min_{\varphi(x, 0)=g(x)} \max_{\rho} \int_0^T \int_{\Omega} \rho(x, t) \left( \frac{\partial \varphi(x, t)}{\partial t} + \sup_{\alpha \in \mathbb{R}^m} \{ -\langle f_{x,t}(\alpha), \nabla_x \varphi(x, t) \rangle - L_{x,t}(\alpha) \} \right) dx dt - c \int_{\Omega} \varphi(x, T) dx \\
&\geq \min_{\varphi(x, 0)=g(x)} \max_{\rho \geq 0, \alpha} \int_0^T \int_{\Omega} \rho(x, t) \left( \frac{\partial \varphi(x, t)}{\partial t} - \langle f_{x,t}(\alpha(x, t)), \nabla_x \varphi(x, t) \rangle - L_{x,t}(\alpha(x, t)) \right) dx dt - c \int_{\Omega} \varphi(x, T) dx \\
&= \min_{\varphi(x, 0)=g(x)} \max_{\rho \geq 0, \alpha} \int_0^T \int_{\Omega} -\frac{\partial \rho(x, t)}{\partial t} \varphi(x, t) + \nabla_x \cdot (f_{x,t}(\alpha(x, t)) \rho(x, t)) \varphi(x, t) - \rho(x, t) L_{x,t}(\alpha(x, t)) dx dt \\
&\quad - c \int_{\Omega} \varphi(x, T) dx + \int_{\Omega} (\rho(x, T) \varphi(x, T) - \rho(x, 0) g(x)) dx,
\end{aligned}$$

where the boundary term over  $\partial\Omega$  vanishes during integration by parts, owing to the periodic boundary condition imposed on  $\Omega$ . While this approach can be adapted to different boundary conditions, it necessitates a thorough evaluation of the boundary terms specific to each case. In this study, we introduce algorithms designed to solve HJ PDEs and their related optimal control problems through solving the subsequent saddle point problem

$$\min_{\varphi(x, 0)=g(x)} \max_{\rho \geq 0, \alpha} \int_0^T \int_{\Omega} \rho(x, t) \left( \frac{\partial \varphi(x, t)}{\partial t} - \langle f_{x,t}(\alpha(x, t)), \nabla_x \varphi(x, t) \rangle - L_{x,t}(\alpha(x, t)) \right) dx dt - c \int_{\Omega} \varphi(x, T) dx. \quad (6)$$

Assuming that the stationary point  $(\varphi, \rho, \alpha)$  satisfies  $\rho > 0$ , the first-order optimality condition is

$$\begin{cases} \partial_t \varphi(x, t) - \langle f_{x,t}(\alpha(x, t)), \nabla_x \varphi(x, t) \rangle - L_{x,t}(\alpha(x, t)) = 0, \\ \alpha(x, t) = \arg \min_{\alpha \in \mathbb{R}^m} \{ \langle f_{x,t}(\alpha), \nabla_x \varphi(x, t) \rangle + L_{x,t}(\alpha) \}, \\ \partial_t \rho(x, t) - \nabla_x \cdot (f_{x,t}(\alpha(x, t)) \rho(x, t)) = 0, \\ \varphi(x, 0) = g(x), \quad \rho(x, T) = c. \end{cases} \quad (7)$$

Upon applying the Hamiltonian  $H$ , as defined in (3), the resulting simplification leads to:

$$\begin{cases} \partial_t \varphi(x, t) + H(x, t, \nabla_x \varphi(x, t)) = 0, \\ \partial_t \rho(x, t) + \nabla_x \cdot (\nabla_p H(x, t, \nabla_x \varphi(x, t)) \rho(x, t)) = 0, \\ \varphi(x, 0) = g(x), \quad \rho(x, T) = c, \end{cases} \quad (8)$$

yielding a system containing the HJ PDE (2) with a continuity equation. Furthermore, the stationary point  $\alpha$  determines the function specified in (4), which facilitates the calculation of the optimal control. Thus, addressing the saddle point problem (6) enables the derivation of the HJ PDE solution  $\varphi$  along with the feedback optimal control function via  $\alpha$ . While theoretically, a discrepancy exists between the saddle point problem's stationary point  $(\varphi, \rho, \alpha)$  and the

HJ PDE solution if  $\rho$  takes zero values within certain regions of  $\Omega \times [0, T]$ , such instances were not encountered during our numerical experiments.

To solve the saddle point problem (6), we propose an iterative method, with the update at the  $\ell$ -th iteration described as follows:

$$\begin{cases} \rho^{\ell+1}(x, t) = (\rho^\ell(x, t) + \tau_\rho (\partial_t \tilde{\varphi}^\ell(x, t) - \langle f_{x,t}(\alpha^\ell(x, t)), \nabla_x \tilde{\varphi}^\ell(x, t) \rangle - L_{x,t}(\alpha^\ell(x, t))))_+, \\ \alpha^{\ell+1}(x, t) = \arg \min_{\alpha \in \mathbb{R}^m} \left\{ \langle f_{x,t}(\alpha), \nabla_x \tilde{\varphi}^\ell(x, t) \rangle + L_{x,t}(\alpha) + \frac{\rho^{\ell+1}(x, t)}{2\tau_\alpha} \|\alpha - \alpha^\ell(x, t)\|^2 \right\}, \\ \varphi^{\ell+1}(x, t) = \varphi^\ell(x, t) + \tau_\varphi (I - \Delta)^{-1} (\partial_t \rho^{\ell+1}(x, t) - \nabla_x \cdot (f_{x,t}(\alpha^{\ell+1}(x, t)) \rho^{\ell+1}(x, t))), \\ \tilde{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^\ell. \end{cases} \quad (9)$$

It is pertinent to highlight that both the saddle point problem and our algorithm exclusively engage with affine functions of  $\nabla_x \varphi$ , from both an analytical and numerical standpoint. This approach strategically circumvents the non-linear aspects inherent in HJ PDEs. The methodology we propose utilizes a preconditioned PDHG approach (for an extensive discussion on PDHG, the reader is referred to [15, 92, 24]). Further elaboration on this method is provided in the subsequent remark and detailed in Appendix A.

*Remark 1 (Convergence of the Proposed Algorithm)* In Appendix A, we delve into the specifics of the convergence of the PDHG algorithm when applied to (6) after a change of variable (see (28)). Specifically, our discussion is limited to scenarios where  $f_{x,t}$  is an affine function and  $L_{x,t}$  is non-negative, proper, convex, lower semi-continuous, and 1-coercive with  $L_{x,t}(0) = 0$ . While this argument can extend to broader cases, we refrain from addressing the most general scenarios as our current assumptions encompass several significant instances, including quadratic Lagrangians or indicator functions of a set that includes 0. Such instances are relevant to quadratic Hamiltonians and 1-homogeneous Hamiltonians. For illustrative examples, refer to Section 2.3.

In Remark A1, we clarify the connection between the update methods detailed in (28) and the updates we propose in (9). Our approach revises (28) by opting for a sequential update strategy for  $\rho$  and  $\alpha$ , as opposed to their concurrent modification. This strategy leverages the convergence principles outlined in Appendix A, which pertain to the joint updates in (28). Based on that, the sequential update scheme in (9) facilitates a more computationally streamlined and versatile algorithm. To address the errors introduced by the staggered updates for  $\rho$  and  $\alpha$ , our implementation includes multiple iterative updates for  $\rho$  and  $\alpha$  followed by a single update for  $\varphi$ .

*Remark 2 (Details on Implementation)* For the variable  $\alpha$ , we employ the penalty term  $\frac{\rho^{k+1}(x, t)}{2\tau_\alpha} \|\alpha - \alpha^k(x, t)\|^2$ , while for  $\varphi$ , the penalty is based on the  $H^1$ -norm, leading to the inclusion of the preconditioning operator  $(I - \Delta)^{-1}$  in  $\varphi$ 's update equation. These penalty terms are specifically selected to enhance the speed of the algorithm. The use of preconditioning methods is a well-established practice in the field, with references such as [55, 77, 82]

providing further context. For an in-depth discussion on the practical aspects of computing  $(I - \Delta)^{-1}$ , we direct the reader to [82].

In cases where the grid is extensively partitioned, leading to a high-dimensional optimization challenge, a reduction in dimensionality can be crucial for feasible numerical resolution. To this end, we implemented a strategy that divides the time interval, applying our algorithm within each segment sequentially. Additional information on this approach can be found in [77, Algorithm 3] and [82].

*Remark 3 (Relation to MFGs)* Upon establishing the stationary point  $(\varphi, \rho, \alpha)$  where  $\rho > 0$ , we derive the coupled PDE system as delineated in (8). Implementing time reversal transforms the equations for  $\varphi(x, T - t)$  and  $\tilde{\rho}(x, t) = \rho(x, T - t)$ , aligning them with the first-order optimality conditions of MFGs, which are expressed as follows:

$$\min_{\tilde{\alpha}} \left\{ \int_0^T \int_{\Omega} L(x, s, \tilde{\alpha}(x, s)) \tilde{\rho}(x, s) dx ds + \int_{\Omega} g(x) \tilde{\rho}(x, T) dx : \right. \\ \left. \begin{aligned} \partial_t \tilde{\rho}(x, s) + \nabla_x \cdot (f(x, s, \tilde{\alpha}(x, s)) \tilde{\rho}(x, s)) &= 0 \quad \forall x \in \Omega, s \in [0, T], \\ \tilde{\rho}(x, 0) &= c \quad \forall x \in \Omega \end{aligned} \right\}.$$

This linkage emerges as inherently logical, given that an optimal control problem is fundamentally intertwined with an MFG scenario characterized by an initial density condition resembling a Dirac mass. Through this association, numerical methods devised for MFGs can also be utilized in addressing HJ PDEs and associated optimal control problems.

*Remark 4 (Comparison with the Approach in [82])* While both the current work and [82] leverage saddle point formulations to tackle HJ PDEs, distinctions arise particularly in the context of HJ PDEs derived from optimal control problems. The following analysis addresses scenarios where  $f_{x,t}$  is a bijection for any  $(x, t)$  within  $\Omega \times [0, T]$ . Through implementing a change of variable  $v = -f(x, T - t, \alpha)$ , we obtain

$$\begin{aligned} H(x, t, p) &= \sup_{\alpha \in \mathbb{R}^m} \{ -\langle f(x, T - t, \alpha), p \rangle - L(x, T - t, \alpha) \} \\ &= \sup_{v \in \mathbb{R}^n} \{ \langle v, p \rangle - L(x, T - t, f^{-1}(x, T - t, -v)) \}, \end{aligned}$$

with the inversion  $f^{-1}$  applied solely concerning  $\alpha$ . Consequently,  $H(x, t, \cdot)$ 's Fenchel-Legendre transform emerges as the convex lower semi-continuous hull of  $v \mapsto L(x, T - t, f^{-1}(x, T - t, -v))$ .

Assuming the function  $v \mapsto L(x, T - t, f^{-1}(x, T - t, -v))$  is convex and lower semi-continuous, the saddle point formulation presented in [82] is:

$$\min_{\varphi(x,0)=g(x)} \max_{\rho \geq 0, v} \int_0^T \int_{\Omega} \rho(x, t) \left( \frac{\partial \varphi(x, t)}{\partial t} + \langle v(x, t), \nabla_x \varphi(x, t) \rangle - L(x, T - t, f^{-1}(x, T - t, -v(x, t))) \right) dx dt \\ - c \int_{\Omega} \varphi(x, T) dx. \quad (10)$$



Although both approaches employ saddle point methodologies, the specific formulations outlined in (6) and (10) lead to distinct algorithms. The strategy developed in this work is expressly designed for addressing optimal control problems, effectively providing solutions to HJ PDEs as well as the feedback optimal control functions crucial for determining optimal controls.

A key component of the above computation is the change of variable  $-v = f(x, T - t, \alpha)$ , a method also used in [71]. Unlike [71], which concentrates on analyzing a single trajectory in the context of optimal control, our proposed approach provides solutions to HJ PDEs and feedback controls over the entire domain. These solutions facilitate the computation of optimal controls for various initial conditions.

## 2.2 Discretization

This section is dedicated to discretizing the saddle point problem presented in (6) and outlining the associated algorithms for one-dimensional scenarios in Section 2.2.1 and for two-dimensional situations in Section 2.2.2. For each scenario, we start with upwind spatial discretization, an essential step for correctly approximating the viscous solutions of the HJ PDEs. Following this, implicit Euler temporal discretization is applied, enabling us to circumvent the restrictive CFL condition often encountered with explicit schemes.

### 2.2.1 One-dimensional problems

**Spatial discretization.** The first step in discretizing the saddle point problem involves partitioning the spatial domain  $\Omega$ . We define  $\{x_i\}_{i=1}^{n_x}$  as the evenly spaced grid points within  $\Omega$ , with  $\Delta x$  representing the distance between adjacent points. For a domain  $\Omega = [a, b]$  (where  $-\infty < a < b < \infty$ ) subject to periodic boundary conditions, we calculate the grid size as  $\Delta x = \frac{b-a}{n_x}$  and assign grid points such that  $x_i = a + (i-1)\Delta x$ . Our objective is to determine the values of the functions at these grid points, specifically  $\varphi_i(t) = \varphi(x_i, t)$  and  $\rho_i(t) = \rho(x_i, t)$ . In addressing the HJ PDE, the treatment of first-order spatial derivatives necessitates consideration of both positive and negative velocities, corresponding to scenarios where the function  $f$  assumes positive and negative values, respectively. To manage these cases, we introduce two variables,  $\alpha_1$  and  $\alpha_2$ , and aim to calculate their values at the grid points, denoted as  $\alpha_{1,i}(t)$  and  $\alpha_{2,i}(t)$ . Additionally, the numerical Lagrangian  $\hat{L}: \Omega \times [0, T] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$  is utilized as an approximation of  $L$ , and its determination will vary across different scenarios. The selection of  $\hat{L}$  for particular cases will be discussed subsequently.

For simplicity, we represent the functions  $f(x_i, T-t, \cdot)$  and  $\hat{L}(x_i, T-t, \cdot)$  as  $f_{i,t}$  and  $\hat{L}_{i,t}$  respectively. Upon applying first-order spatial discretization, the saddle point problem expressed in (6) (after division by  $\Delta x$ ) is reformulated

as:

$$\min_{\varphi_i(0)=g(x_i)} \max_{\rho \geq 0, \alpha} \int_0^T \sum_{i=1}^{n_x} \rho_i(t) \left( \dot{\varphi}_i(t) - f_{i,t}(\alpha_{1,i}(t))_+ (D_x^+ \varphi)_i(t) - f_{i,t}(\alpha_{2,i}(t))_- (D_x^- \varphi)_i(t) - \hat{L}_{i,t}(\alpha_{1,i}(t), \alpha_{2,i}(t)) \right) dt - c \sum_{i=1}^{n_x} \varphi_i(T). \quad (11)$$

The variables of this saddle point problem contain all semi-discretized functions, including  $\varphi_i(t)$ ,  $\rho_i(t)$ ,  $\alpha_{1,i}(t)$ , and  $\alpha_{2,i}(t)$ , across the time interval  $t \in [0, T]$  and for grid indices  $i = 1, \dots, n_x$ . Here,  $f_+ = \max\{f, 0\}$  and  $f_- = \min\{f, 0\}$  denote the positive and negative components of  $f$ , respectively. Meanwhile,  $(D_x^+ \varphi)_i = \frac{\varphi_{i+1} - \varphi_i}{\Delta x}$  and  $(D_x^- \varphi)_i = \frac{\varphi_i - \varphi_{i-1}}{\Delta x}$  represent the right and left finite differences, respectively. In this paper, for the sake of simplicity, we interchangeably use the notations  $(D\eta)_i$  and  $D(\eta_i)$  for any finite difference operator  $D$ , function  $\eta$ , and index  $i$ , regardless of whether they pertain to one-dimensional or two-dimensional scenarios.

Examining a stationary point  $(\varphi, \rho, \alpha)$  within this saddle point problem, and under the assumption that  $\rho_i(t) > 0$  for any  $t \in [0, T]$  and  $i = 1, \dots, n_x$ , we arrive at the first-order optimality conditions as follows:

$$\begin{cases} \dot{\varphi}_i(t) - f_{i,t}(\alpha_{1,i}(t))_+ (D_x^+ \varphi)_i(t) - f_{i,t}(\alpha_{2,i}(t))_- (D_x^- \varphi)_i(t) - \hat{L}_{i,t}(\alpha_{1,i}(t), \alpha_{2,i}(t)) = 0, \\ (\alpha_{1,i}(t), \alpha_{2,i}(t)) = \arg \min_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \{f_{i,t}(\alpha_1)_+ (D_x^+ \varphi)_i(t) + f_{i,t}(\alpha_2)_- (D_x^- \varphi)_i(t) + \hat{L}_{i,t}(\alpha_1, \alpha_2)\}, \\ \dot{\rho}_i(t) - D_x^-(f_{i,t}(\alpha_{1,i}(t))_+ \rho_i(t)) - D_x^+(f_{i,t}(\alpha_{2,i}(t))_- \rho_i(t)) = 0. \end{cases} \quad (12)$$

By integrating the first two equations, we obtain

$$\dot{\varphi}_i(t) + \sup_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \{-f_{i,t}(\alpha_1)_+ (D_x^+ \varphi)_i(t) - f_{i,t}(\alpha_2)_- (D_x^- \varphi)_i(t) - \hat{L}_{i,t}(\alpha_1, \alpha_2)\} = 0,$$

which provides a semi-discrete formulation for the HJ PDE (2). The corresponding numerical Hamiltonian  $\hat{H}$  is given by

$$\hat{H}(x, t, D_x^+ \varphi, D_x^- \varphi) = \sup_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \{-f_{x,t}(\alpha_1)_+ D_x^+ \varphi - f_{x,t}(\alpha_2)_- D_x^- \varphi - \hat{L}_{x,t}(\alpha_1, \alpha_2)\}. \quad (13)$$

This yields a semi-discrete approach to solving HJ PDEs and providing feedback optimal controls.

To obtain a valid numerical Hamiltonian, the following conditions must be met:

1. **Monotonicity:** Specifically,  $\hat{H}$  should be non-increasing with respect to  $D_x^+ \varphi$ , and is non-decreasing with respect to  $D_x^- \varphi$ . This condition is ensured by the formulation given in (13).
2. **Consistency:** This requires  $\hat{H}(x, t, p, p) = H(x, t, p)$  for any  $x \in \Omega$ ,  $t \in [0, T]$ , and  $p \in \mathbb{R}$ . The numerical Lagrangian  $\hat{L}$  must be chosen to ensure that the corresponding numerical Hamiltonian is consistent. In scenarios where  $f$  linearly depends on  $\alpha$  and  $L$  is a non-negative, convex, 1-coercive function of  $\alpha$  with  $L_{x,t}(0) = 0$  for every  $x$  and  $t$ , we can choose  $\hat{L}_{x,t}(\alpha_1, \alpha_2) = L_{x,t}(\alpha_1) + L_{x,t}(\alpha_2)$ . For more technical details, refer to Appendix B.1.

With this spatial discretization, the update for the  $\ell$ -th iteration is described as follows:

$$\left\{ \begin{array}{l} \rho_i^{\ell+1}(t) = \left( \rho_i^\ell(t) + \tau_\rho (\dot{\varphi}_i^\ell(t) - f_{i,t}(\alpha_{1,i}^\ell(t))_+ (D_x^+ \tilde{\varphi}^\ell)_i(t) - f_{i,t}(\alpha_{2,i}^\ell(t))_- (D_x^- \tilde{\varphi}^\ell)_i(t) - \hat{L}_{i,t}(\alpha_{1,i}^\ell(t), \alpha_{2,i}^\ell(t)) \right)_+ \\ (\alpha_{1,i}^{\ell+1}(t), \alpha_{2,i}^{\ell+1}(t)) = \arg \min_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \left\{ f_{i,t}(\alpha_1)_+ (D_x^+ \tilde{\varphi}^\ell)_i(t) + f_{i,t}(\alpha_2)_- (D_x^- \tilde{\varphi}^\ell)_i(t) + \hat{L}_{i,t}(\alpha_1, \alpha_2) \right. \\ \quad \left. + \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} (\|\alpha_1 - \alpha_{1,i}^\ell(t)\|^2 + \|\alpha_2 - \alpha_{2,i}^\ell(t)\|^2) \right\}. \\ \varphi_i^{\ell+1}(t) = \varphi_i^\ell(t) + \tau_\varphi (I - \partial_{tt} - D_{xx})^{-1} (\rho_i^{\ell+1}(t) - D_x^-(f_{i,t}(\alpha_{1,i}^{\ell+1}(t))_+ \rho_i^{\ell+1}(t)) - D_x^+(f_{i,t}(\alpha_{2,i}^{\ell+1}(t))_- \rho_i^{\ell+1}(t))). \\ \tilde{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^\ell, \end{array} \right. \quad (14)$$

where  $D_{xx}$  denotes the central difference operator for the second-order derivative, defined as  $(D_{xx}\varphi)_i(t) = \frac{\varphi_{i+1}(t) + \varphi_{i-1}(t) - 2\varphi_i(t)}{\Delta x^2}$ .

**Temporal discretization.** The next step involves discretizing the time domain  $[0, T]$ . We define the uniform temporal grid points as  $t_k$  with a grid size  $\Delta t$ , setting  $\Delta t = \frac{T}{n_t - 1}$  and assigning  $t_i = (i - 1)\Delta t$  for  $i = 1, \dots, n_t$ . For temporal discretization, particularly for  $\partial_t \varphi$ , we adopt an implicit Euler scheme, which circumvents the restrictive CFL condition encountered in explicit Euler discretization. Mirroring our approach in spatial discretization, we simplify notation by using  $f_{i,k}$  and  $\hat{L}_{i,k}$  for the functions  $f(x_i, T - t_k, \cdot)$  and  $\hat{L}(x_i, T - t_k, \cdot)$ , respectively. Our goal is to determine the values of functions at the grid points, namely  $\varphi_{i,k} = \varphi(x_i, t_k)$ ,  $\rho_{i,k} = \rho(x_i, t_k)$ ,  $\alpha_{1,i,k} = \alpha_1(x_i, t_k)$ , and  $\alpha_{2,i,k} = \alpha_2(x_i, t_k)$ . Upon dividing by  $\Delta t$ , the saddle point problem is reformulated as:

$$\min_{\varphi_{i,1}=g(x_i)} \max_{\rho \geq 0, \alpha} \sum_{k=2}^{n_t} \sum_{i=1}^{n_x} \rho_{i,k} \left( (D_t^- \varphi)_{i,k} - f_{i,k}(\alpha_{1,i,k})_+ (D_x^+ \varphi)_{i,k} - f_{i,k}(\alpha_{2,i,k})_- (D_x^- \varphi)_{i,k} - \hat{L}_{i,k}(\alpha_{1,i,k}, \alpha_{2,i,k}) \right) - \frac{c}{\Delta t} \sum_{i=1}^{n_x} \varphi_{i,n_t}, \quad (15)$$

targeting optimization over  $\varphi_{i,k}$ ,  $\rho_{i,k}$ ,  $\alpha_{1,i,k}$ , and  $\alpha_{2,i,k}$  for grid indices  $i = 1, \dots, n_x$  and time steps  $k = 2, \dots, n_t$ . Henceforth,  $D_t^-$  and  $D_t^+$  will represent the implicit and explicit Euler discretizations, respectively, defined as  $(D_t^- \varphi)_{i,k} = \frac{\varphi_{i,k} - \varphi_{i,k-1}}{\Delta t}$  and  $(D_t^+ \varphi)_{i,k} = \frac{\varphi_{i,k+1} - \varphi_{i,k}}{\Delta t}$ .

The proposed update for the  $\ell$ -th iteration is then formulated as follows

$$\left\{ \begin{array}{l} \rho_{i,k}^{\ell+1} = \left( \rho_{i,k}^\ell + \tau_\rho ((D_t^- \tilde{\varphi}^\ell)_{i,k} - f_{i,k}(\alpha_{1,i,k}^\ell)_+ (D_x^+ \tilde{\varphi}^\ell)_{i,k} - f_{i,k}(\alpha_{2,i,k}^\ell)_- (D_x^- \tilde{\varphi}^\ell)_{i,k} - \hat{L}_{i,k}(\alpha_{1,i,k}^\ell, \alpha_{2,i,k}^\ell)) \right)_+ \\ (\alpha_{1,i,k}^{\ell+1}, \alpha_{2,i,k}^{\ell+1}) = \arg \min_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \left\{ f_{i,k}(\alpha_1)_+ (D_x^+ \tilde{\varphi}^\ell)_{i,k} + f_{i,k}(\alpha_2)_- (D_x^- \tilde{\varphi}^\ell)_{i,k} + \hat{L}_{i,k}(\alpha_1, \alpha_2) \right. \\ \quad \left. + \frac{\rho_{i,k}^{\ell+1}}{2\tau_\alpha} (\|\alpha_1 - \alpha_{1,i,k}^\ell\|^2 + \|\alpha_2 - \alpha_{2,i,k}^\ell\|^2) \right\}. \\ \varphi_{i,k}^{\ell+1} = \varphi_{i,k}^\ell + \tau_\varphi (I - D_{tt} - D_{xx})^{-1} \left( (D_t^+ \rho^{\ell+1})_{i,k} - D_x^-(f_{i,k}(\alpha_{1,i,k}^{\ell+1})_+ \rho_{i,k}^{\ell+1}) - D_x^+(f_{i,k}(\alpha_{2,i,k}^{\ell+1})_- \rho_{i,k}^{\ell+1}) \right). \\ \tilde{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^\ell, \end{array} \right.$$

where  $D_{tt}$  represent the central difference operators for second-order derivative, defined as  $(D_{tt}\varphi)_{i,k} = \frac{\varphi_{i,k+1} + \varphi_{i,k-1} - 2\varphi_{i,k}}{\Delta t^2}$ . The Fast Fourier Transform (FFT) is utilized to numerically calculate  $(I - D_{tt} - D_{xx})^{-1}$ . For additional information regarding the numerical updates, we refer readers to [55, 77, 82].

Note that if the numerical Lagrangian is chosen to be  $L_{x,t}(\alpha_1) + L_{x,t}(\alpha_2)$ , then the updating step for  $\alpha$  can be computed in parallel for  $\alpha_1$  and  $\alpha_2$ .

### 2.2.2 Two-dimensional problems

The approach to discretizing two-dimensional problems (where  $n = 2$ ) mirrors that of one-dimensional cases. Initially, we focus on spatial discretization, establishing an upwind scheme. Subsequently, for time discretization, we employ an implicit Euler method, facilitating the use of a larger temporal grid size  $\Delta t$ .

**Spatial discretization.** In our analysis, we utilize regular grid structures within the domain  $\Omega$ . This study specifically focuses on rectangular domains  $\Omega$  with periodic boundary conditions, although our methodology is adaptable to various boundary conditions. Suppose  $\Omega = [a_1, b_1] \times [a_2, b_2]$ , with grid sizes defined as  $\Delta x = \frac{b_1 - a_1}{n_x}$  and  $\Delta y = \frac{b_2 - a_2}{n_y}$ . The grid points are denoted by  $x_{i,j} = (a_1 + (i-1)\Delta x, a_2 + (j-1)\Delta y)$ . As with the one-dimensional scenario, handling positive and negative velocities in two dimensions necessitates multiple  $\alpha$  variables. Given that  $f$  outputs values in  $\mathbb{R}^2$ , we require four  $\alpha$  variables, unlike the two needed for one-dimensional cases. The components of  $f$  are represented by  $f_1$  and  $f_2$  for the first and second components, respectively. For each component  $i = 1, 2$ , we introduce  $\alpha_{i1}$  and  $\alpha_{i2}$  to capture the positive and negative aspects of  $f_i$ . Our goal is to calculate the function values at the grid points; thus, we aim at computing  $\varphi_{i,j}(t) = \varphi(x_{i,j}, t)$ ,  $\rho_{i,j}(t) = \rho(x_{i,j}, t)$ , and  $\alpha_{IJ,i,j}(t) = \alpha_{IJ}(x_{i,j}, t)$  for  $I, J = 1, 2$ . For ease of reference, the functions  $f_1(x_{i,j}, T-t, \cdot)$ ,  $f_2(x_{i,j}, T-t, \cdot)$ , and  $\hat{L}(x_{i,j}, T-t, \cdot)$  are abbreviated as  $f_{1,i,j,t}$ ,  $f_{2,i,j,t}$ , and  $\hat{L}_{i,j,t}$ , respectively. Here,  $\hat{L}: \Omega \times [0, T] \times (\mathbb{R}^m)^4 \rightarrow \mathbb{R}$  denotes the numerical Lagrangian, which will be defined later, similar to the approach taken in one-dimensional scenarios. The formulation of the saddle point problem, divided by  $\Delta x \Delta y$ , is as follows:

$$\begin{aligned} \min_{\varphi} \max_{\rho \geq 0, \alpha} \int_0^T \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \rho_{i,j}(t) & \left( \dot{\varphi}_{i,j}(t) - f_{1,i,j,t}(\alpha_{11,i,j}(t))_+ (D_x^+ \varphi)_i(t) - f_{1,i,j,t}(\alpha_{12,i,j}(t))_- (D_x^- \varphi)_i(t) \right. \\ & - f_{2,i,j,t}(\alpha_{21,i,j}(t))_+ (D_y^+ \varphi)_i(t) - f_{2,i,j,t}(\alpha_{22,i,j}(t))_- (D_y^- \varphi)_i(t) \\ & \left. - \hat{L}_{i,j,t}(\alpha_{11,i,j}(t), \alpha_{12,i,j}(t), \alpha_{21,i,j}(t), \alpha_{22,i,j}(t)) \right) dt - c \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \varphi_{i,j}(T), \end{aligned} \quad (16)$$

where  $D_y^-$  and  $D_y^+$  represent the left and right finite difference schemes in the  $y$ -dimension, analogous to  $D_x^-$  and  $D_x^+$  in the  $x$ -dimension.

Consider a stationary point  $(\varphi, \rho, \alpha)$  in this saddle point problem. If we further assume  $\rho_{i,j}(t) > 0$  for any  $t \in [0, T]$ , then the first order optimality

condition is

$$\left\{ \begin{array}{l} \dot{\varphi}_{i,j}(t) - f_{1,i,j,t}(\alpha_{11,i,j}(t))_+(D_x^+ \varphi)_{i,j}(t) - f_{1,i,j,t}(\alpha_{12,i,j}(t))_-(D_x^- \varphi)_{i,j}(t) \\ - f_{2,i,j,t}(\alpha_{21,i,j}(t))_+(D_y^+ \varphi)_{i,j}(t) - f_{2,i,j,t}(\alpha_{22,i,j}(t))_-(D_y^- \varphi)_{i,j}(t) \\ - \hat{L}_{i,j,t}(\alpha_{11,i,j}(t), \alpha_{12,i,j}(t), \alpha_{21,i,j}(t), \alpha_{22,i,j}(t)) = 0, \\ (\alpha_{11,i,j}(t), \alpha_{12,i,j}(t), \alpha_{21,i,j}(t), \alpha_{22,i,j}(t)) = \arg \min_{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m} \{f_{1,i,j,t}(\alpha_{11})_+(D_x^+ \varphi)_{i,j}(t) \\ + f_{1,i,j,t}(\alpha_{12})_-(D_x^- \varphi)_{i,j}(t) + f_{2,i,j,t}(\alpha_{21})_+(D_y^+ \varphi)_{i,j}(t) \\ + f_{2,i,j,t}(\alpha_{22})_-(D_y^- \varphi)_{i,j}(t) + \hat{L}_{i,j,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})\}, \\ \dot{\rho}_{i,j}(t) - D_x^-(f_{1,i,j,t}(\alpha_{11,i,j}(t))_+ \rho_{i,j}(t)) - D_x^+(f_{1,i,j,t}(\alpha_{12,i,j}(t))_- \rho_{i,j}(t)) \\ - D_y^-(f_{2,i,j,t}(\alpha_{21,i,j}(t))_+ \rho_{i,j}(t)) - D_y^+(f_{2,i,j,t}(\alpha_{22,i,j}(t))_- \rho_{i,j}(t)) = 0. \end{array} \right. \quad (17)$$

Integrating the first two equations of this optimality condition yields:

$$\dot{\varphi}_{i,j}(t) + \sup_{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m} \{-f_{1,i,j,t}(\alpha_{11})_+(D_x^+ \varphi)_{i,j}(t) - f_{1,i,j,t}(\alpha_{12})_-(D_x^- \varphi)_{i,j}(t) - f_{2,i,j,t}(\alpha_{21})_+(D_y^+ \varphi)_{i,j}(t) \\ - f_{2,i,j,t}(\alpha_{22})_-(D_y^- \varphi)_{i,j}(t) - \hat{L}_{i,j,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})\} = 0.$$

This integration results in a semi-discrete formulation for the HJ PDE (2), with the numerical Hamiltonian being

$$\hat{H}(x, t, D_x^+ \varphi, D_x^- \varphi, D_y^+ \varphi, D_y^- \varphi) = \sup_{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m} \{-f_{1,x,t}(\alpha_{11})_+ D_x^+ \varphi - f_{1,x,t}(\alpha_{12})_- D_x^- \varphi \\ - f_{2,x,t}(\alpha_{21})_+ D_y^+ \varphi - f_{2,x,t}(\alpha_{22})_- D_y^- \varphi - \hat{L}_{x,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})\}. \quad (18)$$

Here,  $f_{1,x,t}$  and  $f_{2,x,t}$  refer to the first and second components of the function  $f_{x,t}$ , respectively. For the numerical Hamiltonian to be considered valid, it must meet the criteria outlined below:

1. **Monotonicity:** Specifically,  $\hat{H}$  should be non-increasing with respect to  $D_x^+ \varphi$  and  $D_y^+ \varphi$ , and should be non-decreasing with respect to  $D_x^- \varphi$  and  $D_y^- \varphi$ . This property is ensured by the formulation specified in (18).
2. **Consistency:** It is required that  $\hat{H}(x, t, p_1, p_1, p_2, p_2) = H(x, t, p)$  for every  $x \in \Omega$ ,  $t \in [0, T]$ , and  $p = (p_1, p_2) \in \mathbb{R}^2$ . This consistency must be considered when selecting  $\hat{L}$ . When  $f$  exhibits a linear relationship with  $\alpha$ ,  $L$  is non-negative and convex with 1-coerciveness over  $\alpha$  and satisfies  $L_{x,t}(0) = 0$ , and the Hamiltonian  $H$  is separable into  $H_{x,t}(p_1, p_2) = H_1(x, t, p_1) + H_2(x, t, p_2)$  for some functions  $H_1$  and  $H_2$  applicable across all  $x \in \Omega$ ,  $t \in [0, T]$ , and for any  $p_1, p_2 \in \mathbb{R}$ , then the selection for  $\hat{L}(x, t, \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) = L(x, t, \alpha_{11}) + L(x, t, \alpha_{12}) + L(x, t, \alpha_{21}) + L(x, t, \alpha_{22})$  is appropriate. For more technical details, refer to Appendix B.2.

Following this discretization approach, the update for the  $\ell$ -th iteration is as follows:

$$\left\{ \begin{array}{l} \rho_{i,j}^{\ell+1}(t) = \left( \rho_{i,j}^{\ell}(t) + \tau_{\rho} (\hat{\varphi}_{i,j}^{\ell}(t) - f_{1,i,j,t}(\alpha_{11,i,j}^{\ell}(t))_+ (D_x^+ \hat{\varphi}^{\ell})_{i,j}(t) \right. \\ \quad - f_{1,i,j,t}(\alpha_{12,i,j}^{\ell}(t))_- (D_x^- \hat{\varphi}^{\ell})_{i,j}(t) - f_{2,i,j,t}(\alpha_{21,i,j}^{\ell}(t))_+ (D_y^+ \hat{\varphi}^{\ell})_{i,j}(t) \\ \quad \left. - f_{2,i,j,t}(\alpha_{22,i,j}^{\ell}(t))_- (D_y^- \hat{\varphi}^{\ell})_{i,j}(t) - \hat{L}_{i,j,t}(\alpha_{1,i}^{\ell}(t), \alpha_{2,i}^{\ell}(t)) \right)_+ \\ (\alpha_{11,i,j}^{\ell+1}(t), \alpha_{12,i,j}^{\ell+1}(t), \alpha_{21,i,j}^{\ell+1}(t), \alpha_{22,i,j}^{\ell+1}(t)) = \arg \min_{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m} \left\{ f_{1,i,j,t}(\alpha_{11})_+ (D_x^+ \hat{\varphi}^{\ell})_{i,j}(t) \right. \\ \quad + f_{1,i,j,t}(\alpha_{12})_- (D_x^- \hat{\varphi}^{\ell})_{i,j}(t) + f_{2,i,j,t}(\alpha_{21})_+ (D_y^+ \hat{\varphi}^{\ell})_{i,j}(t) + f_{2,i,j,t}(\alpha_{22})_- (D_y^- \hat{\varphi}^{\ell})_{i,j}(t) \\ \quad + \hat{L}_{i,j,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) + \frac{\rho_{i,j}^{\ell+1}(t)}{2\tau_{\alpha}} (\|\alpha_{11} - \alpha_{11,i,j}^{\ell}(t)\|^2 + \|\alpha_{12} - \alpha_{12,i,j}^{\ell}(t)\|^2 \\ \quad \left. + \|\alpha_{21} - \alpha_{21,i,j}^{\ell}(t)\|^2 + \|\alpha_{22} - \alpha_{22,i,j}^{\ell}(t)\|^2) \right\}. \\ \varphi_{i,j}^{\ell+1}(t) = \varphi_{i,j}^{\ell}(t) + \tau_{\varphi} (I - \partial_{tt} - D_{xx} - D_{yy})^{-1} \left( \rho_{i,j}^{\ell+1}(t) - D_x^- (f_{1,i,j,t}(\alpha_{11,i,j}^{k+1}(t))_+ \rho_{i,j}^{\ell+1}(t)) \right. \\ \quad - D_x^+ (f_{1,i,j,t}(\alpha_{12,i,j}^{k+1}(t))_- \rho_{i,j}^{\ell+1}(t)) - D_y^- (f_{2,i,j,t}(\alpha_{21,i,j}^{k+1}(t))_+ \rho_{i,j}^{\ell+1}(t)) \\ \quad \left. - D_y^+ (f_{2,i,j,t}(\alpha_{22,i,j}^{k+1}(t))_- \rho_{i,j}^{\ell+1}(t)) \right). \\ \hat{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^{\ell}, \end{array} \right.$$

where  $D_{yy}$  denotes the central difference operators for the second-order derivative in the  $y$ -dimension, analogous to the  $D_{xx}$  and  $D_{tt}$  operators for the  $x$ -dimension and  $t$ -dimension, respectively.

**Temporal discretization.** We proceed to discretize the time domain, employing an implicit Euler method for  $\hat{\varphi}$  to facilitate the use of larger time steps. The time step size is defined as  $\Delta t = \frac{T}{n_t - 1}$ , with grid points marked by  $t_k = (k - 1)\Delta t$  for  $k = 1, \dots, n_t$ . The functions  $f_I(x_{i,j}, T - t_k, \cdot)$  (for  $I = 1, 2$ ) and  $\hat{L}(x_{i,j}, T - t_k, \cdot)$  are represented as  $f_{I,i,j,k}$  and  $\hat{L}_{i,j,k}$ , respectively. Our objective is to compute the values of these functions at the grid points, specifically  $\varphi_{i,j,k}$ ,  $\rho_{i,j,k}$ , and  $\alpha_{IJ,i,j,k}$  for  $I, J = 1, 2$ . Following this, the formulation of the saddle point challenge, normalized by  $\Delta t$ , is presented as:

$$\min_{\varphi_{i,j,1}=g(x_{i,j})} \max_{\rho \geq 0, \alpha} \sum_{k=2}^{n_t} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \rho_{i,j,k} \left( (D_t^- \varphi)_{i,j,k} - f_{1,i,j,k}(\alpha_{11,i,j,k})_+ (D_x^+ \varphi)_{i,j,k} - f_{1,i,j,k}(\alpha_{12,i,j,k})_- (D_x^- \varphi)_{i,j,k} \right. \\ \quad - f_{2,i,j,k}(\alpha_{21,i,j,k})_+ (D_y^+ \varphi)_{i,j,k} - f_{2,i,j,k}(\alpha_{22,i,j,k})_- (D_y^- \varphi)_{i,j,k} \\ \quad \left. - \hat{L}_{i,j,k}(\alpha_{11,i,j,k}, \alpha_{12,i,j,k}, \alpha_{21,i,j,k}, \alpha_{22,i,j,k}) \right) - \frac{c}{\Delta t} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \varphi_{i,j,n_t}.$$

The update process for the  $\ell$ -th iteration in the proposed algorithm is then established as

$$\left\{ \begin{array}{l} \rho_{i,j,k}^{\ell+1} = \left( \rho_{i,j,k}^{\ell} + \tau_{\rho} \left( (D_t^- \tilde{\varphi}^{\ell})_{i,j,k} - f_{1,i,j,k}(\alpha_{11,i,j,k}^{\ell}) + (D_x^+ \tilde{\varphi}^{\ell})_{i,j,k} - f_{1,i,j,k}(\alpha_{12,i,j,k}^{\ell}) - (D_x^- \tilde{\varphi}^{\ell})_{i,j,k} \right. \right. \\ \quad - f_{2,i,j,k}(\alpha_{21,i,j,k}^{\ell}) + (D_y^+ \tilde{\varphi}^{\ell})_{i,j,k} - f_{2,i,j,k}(\alpha_{22,i,j,k}^{\ell}) - (D_y^- \tilde{\varphi}^{\ell})_{i,j,k} \\ \quad \left. \left. - \hat{L}_{i,j,k}(\alpha_{11,i,j,k}^{\ell}, \alpha_{12,i,j,k}^{\ell}, \alpha_{21,i,j,k}^{\ell}, \alpha_{22,i,j,k}^{\ell}) \right) \right)_{+} \\ (\alpha_{11,i,j,k}^{\ell+1}, \alpha_{12,i,j,k}^{\ell+1}, \alpha_{21,i,j,k}^{\ell+1}, \alpha_{22,i,j,k}^{\ell+1}) = \underset{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m}{\operatorname{arg\,min}} \{ f_{1,i,j,k}(\alpha_{11}) + (D_x^+ \tilde{\varphi}^{\ell})_{i,j,k} \\ \quad + f_{1,i,j,k}(\alpha_{12}) - (D_x^- \tilde{\varphi}^{\ell})_{i,j,k} + f_{2,i,j,k}(\alpha_{21}) + (D_y^+ \tilde{\varphi}^{\ell})_{i,j,k} \\ \quad + f_{2,i,j,k}(\alpha_{22}) - (D_y^- \tilde{\varphi}^{\ell})_{i,j,k} + \hat{L}_{i,j,k}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) \\ \quad + \frac{\rho_{i,j,k}^{\ell+1}}{2\tau_{\alpha}} (\|\alpha_{11} - \alpha_{11,i,j,k}^{\ell}\|^2 + \|\alpha_{12} - \alpha_{12,i,j,k}^{\ell}\|^2 + \|\alpha_{21} - \alpha_{21,i,j,k}^{\ell}\|^2 + \|\alpha_{22} - \alpha_{22,i,j,k}^{\ell}\|^2) \}. \\ \varphi_{i,j,k}^{\ell+1} = \varphi_{i,j,k}^{\ell} + \tau_{\varphi} (I - D_{tt} - D_{xx} - D_{yy})^{-1} \left( (D_t^+ \rho^{\ell+1})_{i,j,k} - D_x^- \left( f_{1,i,j,k}(\alpha_{11,i,j,k}^{\ell+1}) + \rho_{i,j,k}^{\ell+1} \right) \right. \\ \quad - D_x^+ \left( f_{1,i,j,k}(\alpha_{12,i,j,k}^{\ell+1}) - \rho_{i,j,k}^{\ell+1} \right) - D_y^- \left( f_{2,i,j,k}(\alpha_{21,i,j,k}^{\ell+1}) + \rho_{i,j,k}^{\ell+1} \right) \\ \quad \left. - D_y^+ \left( f_{2,i,j,k}(\alpha_{22,i,j,k}^{\ell+1}) - \rho_{i,j,k}^{\ell+1} \right) \right). \\ \tilde{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^{\ell}. \end{array} \right.$$

It is important to note that selecting the numerical Lagrangian  $\hat{L}$  as  $L_{x,t}(\alpha_{11}) + L_{x,t}(\alpha_{12}) + L_{x,t}(\alpha_{21}) + L_{x,t}(\alpha_{22})$  enables parallel computation of the update steps for each of the variables  $\alpha_{11}$ ,  $\alpha_{12}$ ,  $\alpha_{21}$ , and  $\alpha_{22}$ .

### 2.3 Numerical examples

This section focuses on the optimal control problem as described in (1), characterized by the dynamics

$$f(x, t, \alpha) = A(x, t)\alpha + b(x, t),$$

where  $A(x, t)$  represents an  $n \times m$  matrix, and  $b(x, t)$  is a vector in  $\mathbb{R}^n$ . It is further assumed that the Lagrangian  $L$  exhibits convexity with respect to  $\alpha$ . The corresponding Hamiltonian  $H$  in (2) is detailed as follows

$$\begin{aligned} H(x, t, p) &= \sup_{\alpha \in \mathbb{R}^n} \{ -\langle A(x, T-t)\alpha + b(x, T-t), p \rangle - L(x, T-t, \alpha) \} \\ &= -\langle b(x, T-t), p \rangle + \sup_{\alpha \in \mathbb{R}^n} \{ \langle \alpha, -A(x, T-t)^T p \rangle - L(x, T-t, \alpha) \} \\ &= -\langle b(x, T-t), p \rangle + L^*(x, T-t, -A(x, T-t)^T p), \end{aligned}$$

with  $L^*(x, T-t, \cdot)$  denoting the Fenchel-Legendre transform of  $L(x, T-t, \cdot)$  for any  $x \in \Omega$  and  $t \in [0, T]$ .

In the cases where  $n = 1$ , the updates of  $\alpha$  in (14) are computed by

$$\begin{aligned}
\alpha_{1,i}^{\ell+1}(t) &= \arg \min_{f_{i,t}(\alpha) \geq 0} \left\{ (A(x, T-t)\alpha + b(x, T-t))(D_x^+ \bar{\varphi}^\ell)_i(t) + L_{i,t}(\alpha) + \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} \|\alpha - \alpha_{1,i}^\ell(t)\|^2 \right\} \\
&= \arg \min_{f_{i,t}(\alpha) \geq 0} \left\{ L_{i,t}(\alpha) + \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} \left\| \alpha - \alpha_{1,i}^\ell(t) + \frac{\tau_\alpha (D_x^+ \bar{\varphi}^\ell)_i(t)}{\rho_i^{\ell+1}(t)} A(x, T-t)^T \right\|^2 \right\}, \\
\alpha_{2,i}^{\ell+1}(t) &= \arg \min_{f_{i,t}(\alpha) \leq 0} \left\{ (A(x, T-t)\alpha + b(x, T-t))(D_x^- \bar{\varphi}^\ell)_i(t) + L_{i,t}(\alpha) + \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} \|\alpha - \alpha_{2,i}^\ell(t)\|^2 \right\} \\
&= \arg \min_{f_{i,t}(\alpha) \leq 0} \left\{ L_{i,t}(\alpha) + \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} \left\| \alpha - \alpha_{2,i}^\ell(t) + \frac{\tau_\alpha (D_x^- \bar{\varphi}^\ell)_i(t)}{\rho_i^{\ell+1}(t)} A(x, T-t)^T \right\|^2 \right\},
\end{aligned} \tag{19}$$

which are  $m$ -dimensional proximal point problems of the functions  $L_{x,t}$  and can be computed in parallel for different  $x, t$ . The two-dimensional  $\alpha$ -updates are similar as the one-dimensional case.

In each example, we begin by numerically solving for  $\varphi$ ,  $\alpha_1$ , and  $\alpha_2$  in the one-dimensional case, or  $\varphi$ ,  $\alpha_{11}$ ,  $\alpha_{12}$ ,  $\alpha_{21}$ , and  $\alpha_{22}$  in the two-dimensional case. Following this, applying the forward Euler method to (5), we obtain several optimal trajectories  $\gamma^*$  and the optimal open-loop controls  $\alpha^*$ , which provide the solutions to (1) with  $t = 0$  and different initial conditions  $x$ . For the purpose of demonstration, in one-dimensional examples, we illustrate the level sets of  $\varphi$  and  $\alpha_1 + \alpha_2$  within the spatial-temporal domain. We also plot the graphs of the functions  $\gamma^*$  and  $\alpha^*$ . In the case of two-dimensional examples, the level sets of  $\varphi(\cdot, t)$  and each component of  $\alpha_{11}(\cdot, t) + \alpha_{12}(\cdot, t) + \alpha_{21}(\cdot, t) + \alpha_{22}(\cdot, t)$  are depicted in the spatial domain for several times  $t \in [0, T]$ . We display the paths of  $\gamma^*$  and  $\alpha^*$  by plotting their values in the spatial domain.

### 2.3.1 Quadratic Hamiltonian with spatial dependent coefficients

In this section, we explore an optimal control problem characterized by dynamics  $f(x, t, \alpha)_I = -(|x_I - 1|^2 + 0.1)\alpha_I$  for  $I = 1, \dots, n$ , with a Lagrangian  $L(x, t, \alpha) = \frac{1}{2}|\alpha|^2$ . Accordingly, the Hamiltonian defined in (2) is expressed as  $H(x, t, p) = \frac{1}{2} \sum_{I=1}^n (|x_I - 1|^2 + 0.1)^2 p_I^2$ . The terminal cost function is  $g(x) = \sum_{I=1}^n \sin \pi x_I$ , over the spatial domain  $[0, 2]^n$  under periodic boundary conditions. The control dimension  $m$  is equal to the state dimension  $n$ .

We adopt the numerical Lagrangian  $\hat{L}$  as  $\hat{L}_{x,t}(\alpha_1, \alpha_2) = L_{x,t}(\alpha_1) + L_{x,t}(\alpha_2)$  in one-dimensional cases and  $\hat{L}_{x,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) = L_{x,t}(\alpha_{11}) + L_{x,t}(\alpha_{12}) + L_{x,t}(\alpha_{21}) + L_{x,t}(\alpha_{22})$  for two-dimensional scenarios. As discussed in Section B, these choices for the numerical Lagrangians yield consistent Hamiltonians in this example. For the one-dimensional scenario, the  $\alpha$  updates as outlined



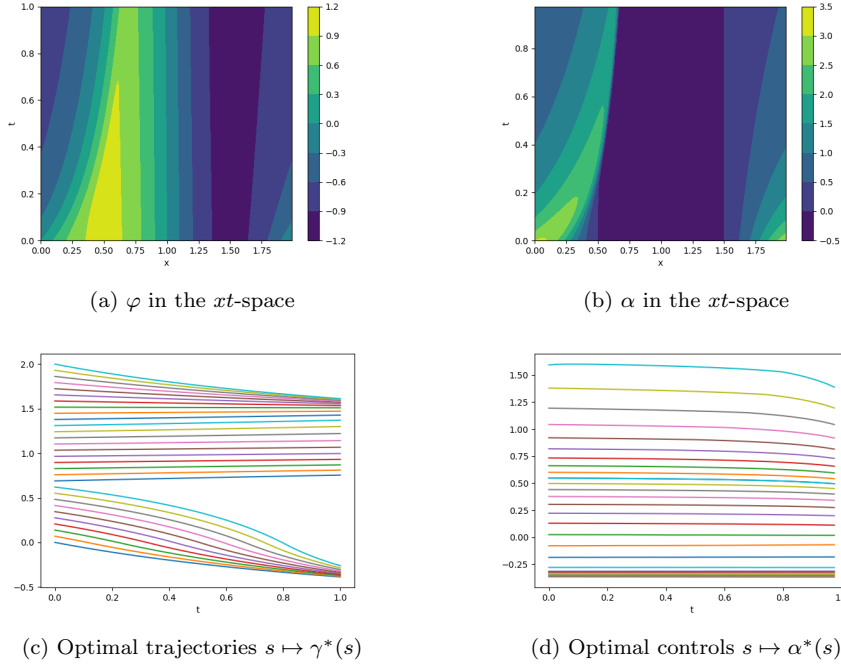


Fig. 1: Visualization of the solution for the one-dimensional scenario discussed in Section 2.3.1, using  $n_t = 41$  and  $n_x = 160$  grid points. Figures (a) and (b) showcase the level sets of the solution  $\varphi$  to the HJ PDE (2), along with the corresponding function  $\alpha$  from (4), which represents the time reversal of the feedback control function. Figures (c) and (d) depict several optimal paths  $s \mapsto \gamma^*(s)$  and their associated open-loop optimal controls  $s \mapsto \alpha^*(s)$ . These paths and control trajectories are the solutions to the optimal control problem (1), each beginning from a unique initial condition  $x$ .

in (19) are calculated by:

$$\begin{aligned}
\alpha_{1,i}^{\ell+1}(t) &= \arg \min_{\alpha \leq 0} \left\{ \frac{1}{2} \alpha^2 + \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} \left( \alpha - \alpha_{1,i}^\ell(t) - \frac{\tau_\alpha}{\rho_i^{\ell+1}(t)} (|x_i - 1|^2 + 0.1) (D_x^+ \tilde{\varphi}^\ell)_i(t) \right)^2 \right\}, \\
&= \min \left\{ 0, \frac{\rho_i^{\ell+1}(t) \alpha_{1,i}^\ell(t) + \tau_\alpha (|x_i - 1|^2 + 0.1) (D_x^+ \tilde{\varphi}^\ell)_i(t)}{\rho_i^{\ell+1}(t) + \tau_\alpha} \right\}, \\
\alpha_{2,i}^{\ell+1}(t) &= \arg \min_{\alpha \geq 0} \left\{ \frac{1}{2} \alpha^2 + \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} \left( \alpha - \alpha_{2,i}^\ell(t) - \frac{\tau_\alpha}{\rho_i^{\ell+1}(t)} (|x_i - 1|^2 + 0.1) (D_x^- \tilde{\varphi}^\ell)_i(t) \right)^2 \right\} \\
&= \max \left\{ 0, \frac{\rho_i^{\ell+1}(t) \alpha_{2,i}^\ell(t) + \tau_\alpha (|x_i - 1|^2 + 0.1) (D_x^- \tilde{\varphi}^\ell)_i(t)}{\rho_i^{\ell+1}(t) + \tau_\alpha} \right\}.
\end{aligned} \tag{20}$$

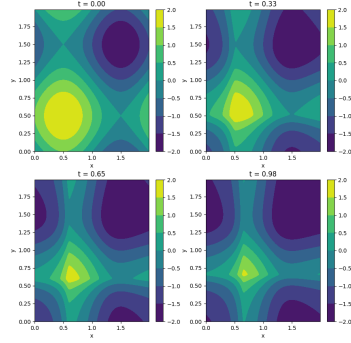
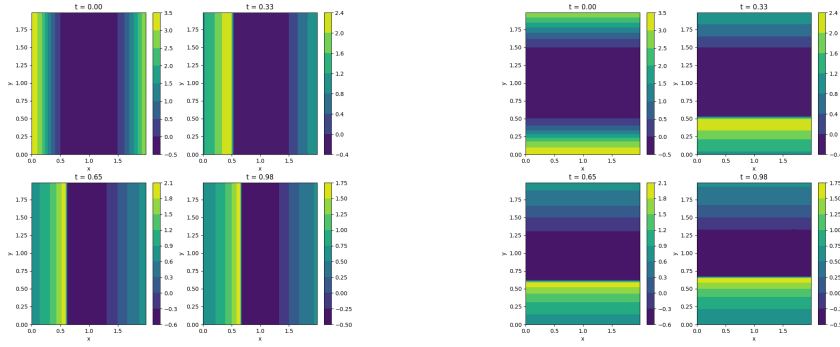
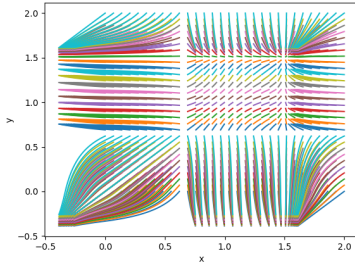
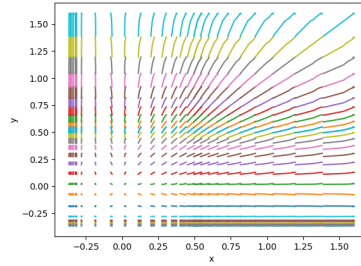
(a)  $(x, y) \mapsto \varphi(x, y, t)$  at different  $t$ (b) The first component of  $(x, y) \mapsto \alpha(x, y, t)$  at different  $t$ (c) The second component of  $(x, y) \mapsto \alpha(x, y, t)$  at different  $t$ (d) Optimal trajectories  $\gamma^*$ (e) Optimal controls  $\alpha^*$ 

Fig. 2: Depiction of the two-dimensional solution as discussed in Section 2.3.1, utilizing  $n_t = 41$  and  $n_x = n_y = 160$  grid points. Figure (a) illustrates the level sets of the solution  $\varphi(\cdot, t)$  to the HJ PDE (2) at different times  $t$ . Figures (b) and (c) show the first and second components, respectively, of the associated function  $\alpha(\cdot, t)$  from (4) at various times  $t$ , which depict the time reversal of the feedback control function. Figures (d) and (e) present several optimal trajectories  $s \mapsto \gamma^*(s)$  along with their corresponding open-loop optimal controls  $s \mapsto \alpha^*(s)$ . These trajectories and control strategies solve the optimal control problem specified in (1), starting from distinct initial conditions  $x$ . Notably, both  $\gamma^*$  and  $\alpha^*$  take values in  $\mathbb{R}^2$ . For visualization, they are plotted within the spatial domain, excluding the time dimension for clarity.

Two-dimensional scenarios follow a similar computation process. In our numerical experiment, we selected  $n_t = 41$  and  $n_x = n_y = 160$  as the number of grid points. Utilizing the implicit Euler method for time discretization allows us to circumvent the CFL condition, enabling the adoption of a larger time step  $\Delta t$ .

The solution of the one-dimensional problem ( $n = m = 1$ ) is illustrated in Figure 1, showcasing the level sets of our numerical solutions  $\varphi$  and  $\alpha_1 + \alpha_2$  in (a) and (b), respectively. Optimal trajectories and controls for evenly spaced initial conditions within  $\Omega$  are depicted in (c) and (d). The solution for the two-dimensional problem ( $n = m = 2$ ) follows a similar methodology and is depicted in Figure 2.

### 2.3.2 One-homogeneous Hamiltonian with spatial dependent coefficients

We maintain the same spatial domain  $\Omega$  and utilize the functions  $f$  and  $g$  as delineated in Section 2.3.1. The Lagrangian is set to  $L(x, t, \alpha) = \chi_{B_\infty}(\alpha)$ , where  $\chi$  signifies the indicator function, and  $B_\infty$  represents the unit ball in the  $\ell^\infty$  space, defined as  $B_\infty = \{\alpha \in \mathbb{R}^m : \|\alpha_I\|_\infty \leq 1\}$ . Recall that the indicator function  $\chi_C(x)$  associated with a set  $C$  is set to 0 when  $x$  falls within  $C$  and is assigned  $+\infty$  for values of  $x$  outside  $C$ . This choice of Lagrangian imposes a constraint on the control values. The Hamiltonian described in (2) becomes  $H(x, t, p) = \sum_{I=1}^n (|x_I - 1|^2 + 0.1)|p_I|$ , which is 1-homogeneous with respect to  $p$ . We select the numerical Lagrangian  $\hat{L}$  as  $\hat{L}_{x,t}(\alpha_1, \alpha_2) = L_{x,t}(\alpha_1) + L_{x,t}(\alpha_2)$  for the one-dimensional scenarios and  $\hat{L}_{x,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) = L_{x,t}(\alpha_{11}) + L_{x,t}(\alpha_{12}) + L_{x,t}(\alpha_{21}) + L_{x,t}(\alpha_{22})$  for two-dimensional cases. As discussed in Section B, these selections ensure consistency in the resulting Hamiltonians.

For one-dimensional scenarios, the  $\alpha$  updates detailed in (19) are computed as follows:

$$\begin{aligned} \alpha_{1,i}^{\ell+1}(t) &= \arg \min_{\alpha \in [-1,0]} \left\{ \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} \left( \alpha - \alpha_{1,i}^\ell(t) - \frac{\tau_\alpha}{\rho_i^{\ell+1}(t)} (|x_i - 1|^2 + 0.1)(D_x^+ \tilde{\varphi}^\ell)_i(t) \right)^2 \right\}, \\ &= \max \left\{ -1, \min \left\{ 0, \alpha_{1,i}^\ell(t) + \frac{\tau_\alpha}{\rho_i^{\ell+1}(t)} (|x_i - 1|^2 + 0.1)(D_x^+ \tilde{\varphi}^\ell)_i(t) \right\} \right\}, \\ \alpha_{2,i}^{\ell+1}(t) &= \arg \min_{\alpha \in [0,1]} \left\{ \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} \left( \alpha - \alpha_{2,i}^\ell(t) - \frac{\tau_\alpha}{\rho_i^{\ell+1}(t)} (|x_i - 1|^2 + 0.1)(D_x^- \tilde{\varphi}^\ell)_i(t) \right)^2 \right\} \\ &= \min \left\{ 1, \max \left\{ 0, \alpha_{2,i}^\ell(t) + \frac{\tau_\alpha}{\rho_i^{\ell+1}(t)} (|x_i - 1|^2 + 0.1)(D_x^- \tilde{\varphi}^\ell)_i(t) \right\} \right\}. \end{aligned}$$

The methodology for two-dimensional scenarios follows a similar approach. As with the previous example, we opted for  $n_t = 41$  and  $n_x = n_y = 160$  for the grid points. The implicit Euler method for time discretization demonstrates its benefit by allowing for a larger time step,  $\Delta t$ .

The solution for the one-dimensional case ( $n = m = 1$ ) is depicted in Figure 3. Initially, we employ the proposed method to calculate  $\varphi$ ,  $\alpha_1 + \alpha_2$

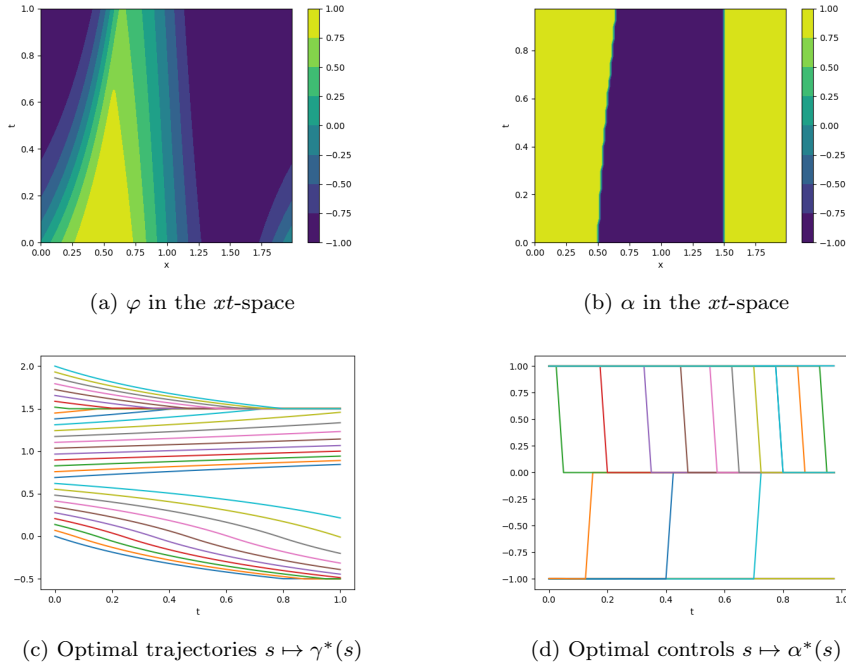


Fig. 3: Visualization of the solution for the one-dimensional scenario discussed in Section 2.3.2, using  $n_t = 41$  and  $n_x = 160$  grid points. Figures (a) and (b) showcase the level sets of the solution  $\varphi$  to the HJ PDE (2), along with the corresponding function  $\alpha$  from (4), which represents the time reversal of the feedback control function. Figures (c) and (d) depict several optimal paths  $s \mapsto \gamma^*(s)$  and their associated open-loop optimal controls  $s \mapsto \alpha^*(s)$ . These paths and control trajectories are the solutions to the optimal control problem (1), each beginning from a unique initial condition  $x$ .

at the grid points, as illustrated in Figures 3 (a) and (b). Subsequently, we determine the optimal trajectories  $\gamma^*$  and optimal open-loop controls  $\alpha^*$  in (1) for  $t = 0$  with varying initial conditions  $x$  uniformly distributed in  $\Omega$ . These are showcased in Figures 3 (c) and (d). The process for solving the two-dimensional problem ( $n = m = 2$ ) mirrors the one-dimensional case and is exhibited in Figure 4.

In contrast to the numerical results presented in Section 2.3.1, the control function  $\alpha$  in this scenario lacks smoothness, predominantly adopting values of  $-1$ ,  $0$ , or  $1$ . This phenomenon is evidenced by the discontinuities observed in Figures 3 (b) and 4 (b)-(c), as well as in the depicted paths of the optimal controls in Figures 3 (d) and 4 (e). The emergence of this distinct bang-bang control pattern is attributable to the control function's formulation as provided in (4). Specifically, in the context of this example, the control function is

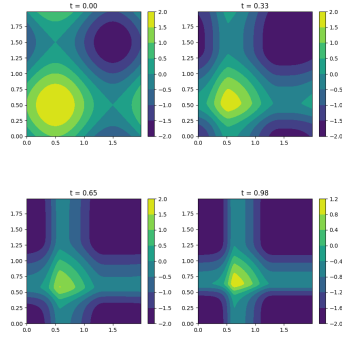
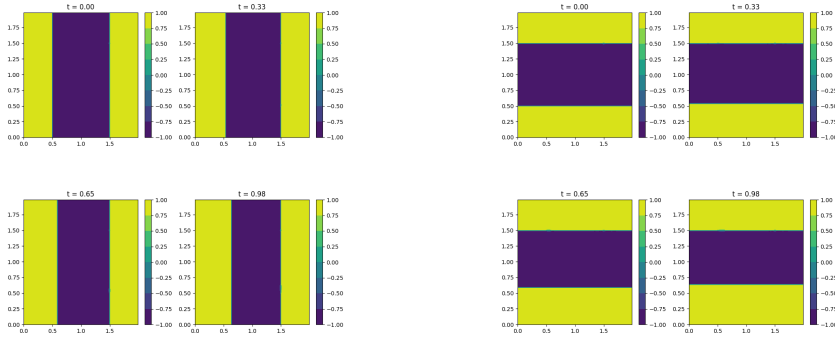
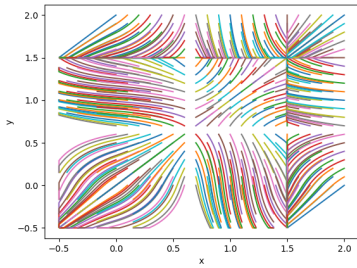
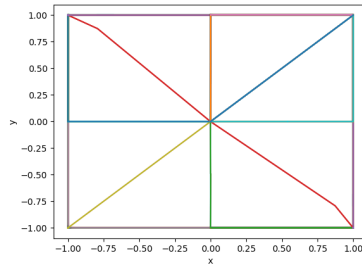
(a)  $(x, y) \mapsto \varphi(x, y, t)$  at different  $t$ (b) The first component of  $(x, y) \mapsto \alpha(x, y, t)$  at different  $t$ (c) The second component of  $(x, y) \mapsto \alpha(x, y, t)$  at different  $t$ (d) Optimal trajectories  $\gamma^*$ (e) Optimal controls  $\alpha^*$ 

Fig. 4: Depiction of the two-dimensional solution as discussed in Section 2.3.2, utilizing  $n_t = 41$  and  $n_x = n_y = 160$  grid points. Figure (a) illustrates the level sets of the solution  $\varphi(\cdot, t)$  to the HJ PDE (2) at different times  $t$ . Figures (b) and (c) show the first and second components, respectively, of the associated function  $\alpha(\cdot, t)$  from (4) at various times  $t$ , which depict the time reversal of the feedback control function. Figures (d) and (e) present several optimal trajectories  $s \mapsto \gamma^*(s)$  along with their corresponding open-loop optimal controls  $s \mapsto \alpha^*(s)$ . These trajectories and control strategies solve the optimal control problem specified in (1), starting from distinct initial conditions  $x$ . Notably, both  $\gamma^*$  and  $\alpha^*$  take values in  $\mathbb{R}^2$ . For visualization, they are plotted within the spatial domain, excluding the time dimension for clarity.

defined as

$$\alpha(x, t) \in \arg \min_{\alpha \in B_\infty} \sum_{I=1}^n -(|x_I - 1|^2 + 0.1) \partial_{x_I} \varphi(x, t) \alpha_I.$$

Consequently, when the condition  $(|x_I - 1|^2 + 0.1) \partial_{x_I} \varphi(x, t) > 1$  is satisfied, it results in  $\alpha_I(x, t) = 1$ . Conversely, if  $(|x_I - 1|^2 + 0.1) \partial_{x_I} \varphi(x, t) < -1$ , then  $\alpha_I(x, t)$  equals  $-1$ . This mechanism underpins the prevalence of these specific control values in the visual representations.

In this example, the optimal control problem is described as:

$$\min_{\alpha(\cdot)} \{g(\gamma(T)) : \gamma(t) = x, \dot{\gamma}(s) = f(\gamma(s), s, \alpha(s)), \alpha(s) \in B_\infty, \forall s \in (t, T)\}.$$

This setup lacks a running cost and instead imposes a constraint on control. This means that if a feasible control leads to a trajectory that starts at  $\gamma(t) = x$  and ends at a point  $\gamma(T)$  that minimizes  $g$  over  $\Omega$ , such a control is optimal. Consequently, it's possible to have multiple optimal controls. In our numerical results, we predominantly encounter controls exhibiting a bang-bang behavior, initially adopting values of 1 or  $-1$  and subsequently transitioning to zero. The corresponding trajectories initially move at maximum speed in the desired direction and then halt upon reaching a minimizer of  $g$ . It is an interesting question to characterize the controls selected by the saddle point problem in (6) and through our numerical algorithms.

### 2.3.3 Newton mechanics

This example delves into an optimal control problem underpinned by Newtonian mechanics. With the settings  $m = 1$ ,  $n = 2$ , and the dynamic equation  $f(x_1, x_2, t, \alpha) = [\alpha, x_1]^T$ , we derive the ODE constraints as  $\dot{\gamma}_1(s) = \alpha(s)$  and  $\dot{\gamma}_2(s) = \gamma_1(s)$ . From a physical standpoint,  $\gamma_2$  is interpreted as the position,  $\gamma_1$  as the velocity, and the control variable  $\alpha$  as the acceleration. The Lagrangian  $L$  is defined as the quadratic function  $L(x, t, \alpha) = \frac{1}{2}|\alpha|^2$ . The Hamiltonian in (2), is formulated as

$$H(x, t, p) = \sup_{\alpha \in \mathbb{R}} \{-\langle f(x, \alpha), p \rangle - L(x, t, \alpha)\} = \frac{1}{2}|p_1|^2 - x_1 p_2.$$

The terminal cost function  $g$  is expressed as

$$g(x_1, x_2) = \exp\left(-\frac{x_1^2}{2}\right) \sin(\pi x_2).$$

The domain for the position variable  $x_2$  is set to  $[-1, 1]$  with periodic boundary conditions, whereas the domain for the velocity variable  $x_1$  is  $[-2, 2]$  with Neumann boundary conditions.

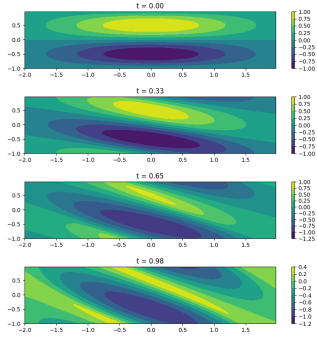
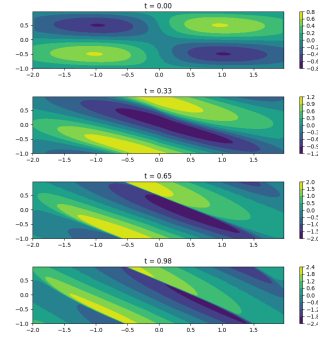
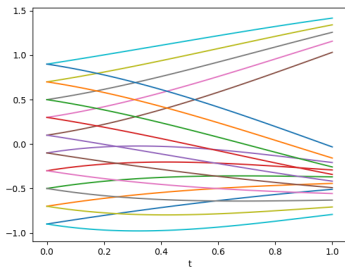
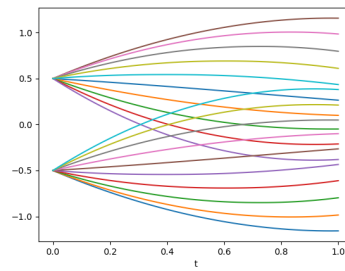
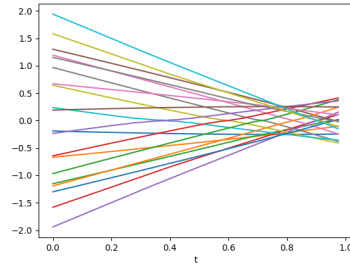
(a)  $(x, y) \mapsto \varphi(x, y, t)$  at different  $t$ (b)  $(x, y) \mapsto \alpha(x, y, t)$  at different  $t$ (c) Optimal positions  $\gamma_2^*$ (d) Optimal velocities  $\gamma_1^*$ (e) Optimal controls  $\alpha^*$ 

Fig. 5: Visualization of the solution for example discussed in Section 2.3.3, using  $n_t = 41$ ,  $n_x = 160$ , and  $n_y = 80$  grid points. Figures (a) and (b) showcase the level sets of the solution  $\varphi$  to the HJ PDE (2), along with the corresponding function  $\alpha$  from (4), which represents the time reversal of the feedback control function. Figures (c), (d), (e) depict the first and second components of several optimal paths  $s \mapsto \gamma^*(s)$  and their associated open-loop optimal controls  $s \mapsto \alpha^*(s)$ . These paths and control trajectories are the solutions to the optimal control problem (1), each beginning from a unique initial condition  $x$ .

Given that the function  $f_2$  does not depend on the control variable, we can disregard the variables  $\alpha_{21}$  and  $\alpha_{22}$  and select  $\hat{L}_{x,t}(\alpha_{11}, \alpha_{12}) = L_{x,t}(\alpha_{11}) + L_{x,t}(\alpha_{12})$ . Thus, the saddle point problem is formulated as

$$\min_{\varphi} \max_{\rho \geq 0, \alpha_{11}, \alpha_{12}} \int_0^T \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \rho_{i,j}(t) \left( \dot{\varphi}_{i,j}(t) - (\alpha_{11,i,j}(t))_+ (D_x^+ \varphi)_{i,j}(t) - (\alpha_{12,i,j}(t))_- (D_x^- \varphi)_{i,j}(t) - ((x_1)_{i,j})_+ (D_y^+ \varphi)_{i,j}(t) - ((x_1)_{i,j})_- (D_y^- \varphi)_{i,j}(t) - \frac{1}{2} \alpha_{11,i,j}(t)^2 - \frac{1}{2} \alpha_{12,i,j}(t)^2 \right) dt - c \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \varphi_{i,j}(T).$$

Updates for  $\varphi$  and  $\rho$  proceed as outlined in (14), with  $\alpha$  updates following the approach described in (19). Specifically, we have

$$\alpha_{11,i,j}^{\ell+1}(t) = \max \left\{ 0, \frac{\rho_{i,j}^{\ell+1}(t) \alpha_{11,i,j}^{\ell}(t) - \tau_{\alpha} (D_x^+ \tilde{\varphi}^{\ell})_{i,j}(t)}{\rho_{i,j}^{\ell+1}(t) + \tau_{\alpha}} \right\},$$

$$\alpha_{12,i,j}^{\ell+1}(t) = \min \left\{ 0, \frac{\rho_{i,j}^{\ell+1}(t) \alpha_{12,i,j}^{\ell}(t) - \tau_{\alpha} (D_x^- \tilde{\varphi}^{\ell})_{i,j}(t)}{\rho_{i,j}^{\ell+1}(t) + \tau_{\alpha}} \right\}.$$

In the numerical experiment, we chose  $n_t = 41$ ,  $n_x = 160$ , and  $n_y = 80$  as the grid specifications, circumventing the CFL condition through the application of the implicit Euler method for temporal discretization.

Figure 5 (a) and (b) display the level sets for the solution  $\varphi$  of the HJ PDE (2), along with the associated function  $\alpha$  described in (4), captured at various times. For the initial conditions  $x = (x_1, x_2)$  specified in the optimal control problem (1), we define the position variable  $x_2$  as points spanning the range  $[-1, 1]$ , and set the velocity  $x_1$  to either  $-0.5$  or  $0.5$ . The calculated optimal controls and trajectories are subsequently illustrated in Figure 5 (c), (d), and (e). Since the functions  $\gamma_1^*$ ,  $\gamma_2^*$ , and  $\alpha^*$  take values in  $\mathbb{R}$ , these three functions are plotted across the spatial-temporal dimensions.

### 3 Stochastic optimal control problems

The methods discussed in Section 2 are also applicable to stochastic optimal control problems. In this section, we consider the following stochastic optimal control problems:

$$\min_{\alpha} \left\{ \mathbb{E} \left[ \int_t^T L(X_s, s, \alpha_s) ds + g(X_T) \right] : X_t = x, dX_s = f(X_s, s, \alpha_s) ds + \sqrt{2\epsilon} dW_s \right\}, \quad (21)$$

where  $W_s$  represents Brownian motion in  $\mathbb{R}^n$ , the control  $\alpha_s$  is an adapted process, and  $f$ ,  $L$ , and  $g$  function as the drift, Lagrangian, and terminal cost, respectively, akin to their definitions in Section 2. The minimal value of problem (21) is represented by  $\phi(x, t)$ , which satisfies a specific viscous HJ PDE as follows:

$$\begin{cases} \frac{\partial \phi(x, t)}{\partial t} + \inf_{\alpha \in \mathbb{R}^m} \{ \langle f(x, t, \alpha), \nabla_x \phi(x, t) \rangle + L(x, t, \alpha) \} + \epsilon \Delta_x \phi(x, t) = 0, & x \in \Omega, t \in [0, T], \\ \phi(x, T) = g(x), & x \in \Omega. \end{cases}$$



Applying a time reversal technique, we arrive at a viscous HJ PDE with an initial condition as:

$$\begin{cases} \frac{\partial \varphi(x, t)}{\partial t} + \sup_{\alpha \in \mathbb{R}^m} \{-\langle f(x, t, \alpha), \nabla_x \varphi(x, t) \rangle - L(x, t, \alpha)\} = \epsilon \Delta_x \varphi(x, t), & x \in \Omega, t \in [0, T], \\ \varphi(x, 0) = g(x), & x \in \Omega. \end{cases} \quad (22)$$

We introduce the function  $\alpha: \Omega \times [0, T] \rightarrow \mathbb{R}^m$  defined as:

$$\alpha(x, t) = \arg \max_{\alpha \in \mathbb{R}^m} \{-\langle f(x, t, \alpha), \nabla_x \varphi(x, t) \rangle - L(x, t, \alpha)\}. \quad (23)$$

Upon reversing time, the function  $\alpha(x, T - t)$  serves as the feedback control function. Consequently, the optimal trajectories and controls are determined through the following computations:

$$\begin{cases} d\gamma_s^* = f(\gamma_s^*, s, \alpha(\gamma_s^*, T - s))ds + \sqrt{2\epsilon}dW_s, & s \in (t, T), \\ \gamma_t^* = x, \\ \alpha_s^* = \alpha(\gamma_s^*, T - s), & s \in [t, T]. \end{cases} \quad (24)$$

For an in-depth exploration of the linkage between stochastic optimal control problems and viscous HJ PDEs, refer to [94].

### 3.1 Saddle point formulation

To solve the viscous HJ PDE, we employ methodologies akin to those used for the first-order HJ PDE as outlined in Section 2. This approach involves framing the PDE as a constraint within an optimization problem, which has the objective function  $-c \int_{\Omega} \varphi(x, T)dx$ , and subsequently introducing a Lagrange multiplier  $\rho$ . Following computations akin to those for the first-order case yield the subsequent saddle point formula:

$$\min_{\varphi(x, 0)=g(x)} \max_{\rho \geq 0, \alpha} \int_0^T \int_{\Omega} \rho(x, t) \left( \frac{\partial \varphi(x, t)}{\partial t} - \langle f_{x,t}(\alpha(x, t)), \nabla_x \varphi(x, t) \rangle - L_{x,t}(\alpha(x, t)) - \epsilon \Delta_x \varphi(x, t) \right) dx dt - c \int_{\Omega} \varphi(x, T) dx. \quad (25)$$

For a stationary point  $(\varphi, \rho, \alpha)$ , where  $\rho(x, t) > 0$  across all  $x \in \Omega$  and  $t \in [0, T]$ , the first-order optimality conditions are given as follows:

$$\begin{cases} \partial_t \varphi(x, t) + H(x, t, \nabla_x \varphi(x, t)) = \epsilon \Delta_x \varphi(x, t), \\ \partial_t \rho(x, t) + \nabla_x \cdot (\nabla_p H(x, t, \nabla_x \varphi(x, t)) \rho(x, t)) + \epsilon \Delta_x \rho(x, t) = 0, \\ \varphi(x, 0) = g(x), \quad \rho(x, T) = c. \end{cases}$$

Echoing the insights from Remark 3, this saddle point problem similarly establishes a connection to an MFG problem. Specifically, with a time reversal,

the formulations for  $\varphi(x, T - t)$  and  $\tilde{\rho}(x, t) = \rho(x, T - t)$  align with the first-order optimality conditions of the following MFG problem:

$$\min_{\tilde{\alpha}} \left\{ \int_0^T \int_{\Omega} L(x, s, \tilde{\alpha}(x, s)) \tilde{\rho}(x, s) dx ds + \int_{\Omega} g(x) \tilde{\rho}(x, T) dx : \right. \\ \left. \partial_t \tilde{\rho}(x, s) + \nabla_x \cdot (f(x, s, \tilde{\alpha}(x, s)) \tilde{\rho}(x, s)) = \epsilon \Delta_x \tilde{\rho}(x, s) \quad \forall x \in \Omega, s \in [0, T], \right. \\ \left. \tilde{\rho}(x, 0) = c \quad \forall x \in \Omega \right\}.$$

Given that the constraint represents a Fokker-Planck equation associated with drifted Brownian motion, the resulting density is inherently positive throughout, thereby naturally fulfilling the  $\rho > 0$  assumption.

The process of discretizing the saddle point problem as described in (25) follows a similar approach to that outlined in Section 2.2, with an additional step to incorporate the diffusion term using a second-order centered difference scheme. The specifics of this approach are elaborated in Appendix C.

The primary distinction between the formulations in (25) and (6) lies in the inclusion of the diffusion term. This addition introduces challenges for the convergence analysis detailed in Appendix A, notably due to the coupling term  $-\int_0^T \int_{\Omega} \epsilon \rho \Delta_x \varphi dx dt$ , which lacks continuity for  $\rho \in L^2$  and  $\varphi \in H^1$ . This challenge is mitigated with the implementation of discretization. Through spatial discretization, a bilinear continuous operator emerges. However, its norm depends on the spatial grid dimensions  $\Delta x$  in one dimension or  $\Delta x \Delta y$  in two dimensions, leading to reduced step sizes as the granularity of spatial discretization increases. For illustration, let's consider the one-dimensional scenario, noting that the two-dimensional case follows a similar pattern. After discretization, the space for  $\varphi$  is represented as  $X = \mathbb{R}^{n_t \times n_x}$ , equipped with an inner product  $\langle \cdot, \cdot \rangle_X$  defined by

$$\langle \phi, \varphi \rangle_X = \sum_{i=1}^{n_x} \sum_{k=1}^{n_t} (\phi_{i,k} \varphi_{i,k} + (D_x^+ \phi)_{i,k} (D_x^+ \varphi)_{i,k}) + \sum_{i=1}^{n_x} \sum_{k=2}^{n_t} (D_t^- \phi)_{i,k} (D_t^- \varphi)_{i,k}.$$

The spaces for  $\rho$  and  $\alpha$  align with Euclidean spaces. Consequently, the discretization of the term  $-\int_0^T \int_{\Omega} \epsilon \rho \Delta_x \varphi dx dt$ , divided by  $\Delta x \Delta t$ , is presented as  $-\sum_{k=2}^{n_t} \sum_{j=1}^{n_x} \epsilon \rho_{i,k} (D_{xx} \varphi)_{i,k}$ , showcasing bilinearity and continuity for  $\varphi \in X$  and  $\rho \in \mathbb{R}^{(n_t-1) \times n_x}$ . However, the norm of this bilinear operator is inversely proportional to  $\Delta x$ , suggesting that the step sizes  $\tau_\varphi$ ,  $\tau_\rho$ , and  $\tau_\alpha$  in the algorithm depend on  $\Delta x$  and may lead to slower convergence compared to scenarios where  $\epsilon = 0$ . Exploring solutions to this slower convergence requires further research. Our experiments with  $H^2$ -preconditioning for  $\varphi$  or  $H^1$ -preconditioning for  $\rho$  did not result in faster convergence.

### 3.2 Numerical examples

This section presents numerical examples under settings akin to those in Section 2.3, with a diffusion coefficient  $\epsilon = 0.1$ . Given that the functions  $f$  and

$L$  remain the same with those previously described, the updates for  $\alpha$  remain unchanged. For each example, we initially deploy our proposed methodology to calculate the functions  $\varphi$  and  $\alpha$ . Subsequently, we employ the Euler–Maruyama method on (24) to generate samples of the optimal processes  $\gamma_s^*$  and controls  $\alpha_s^*$ .

### 3.2.1 Quadratic Hamiltonian with spatial dependent coefficients

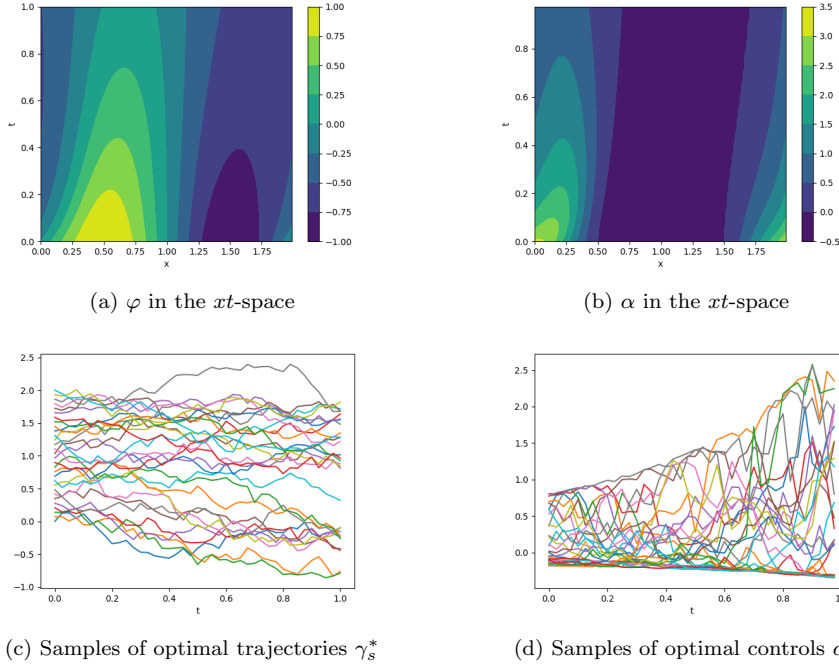


Fig. 6: Visualization of the solution for the one-dimensional scenario discussed in Section 3.2.1, using  $n_t = 41$  and  $n_x = 160$  grid points. Figures (a) and (b) showcase the level sets of the solution  $\varphi$  to the viscous HJ PDE (22), along with the corresponding function  $\alpha$  from (23), which represents the time reversal of the feedback control function. Figures (c) and (d) depict several samples of optimal paths  $\gamma_s^*$  and their associated open-loop optimal controls  $\alpha_s^*$ . These paths and control trajectories are the solutions to the stochastic optimal control problem (21), each beginning from a unique initial condition  $x$ .

We employ the same spatial domain  $\Omega$  and functions  $f, g, L$  as in Section 2.3.1. The solution of the one-dimensional case is depicted in Figure 6, while the solution for the two-dimensional problem is illustrated in Figure 7.

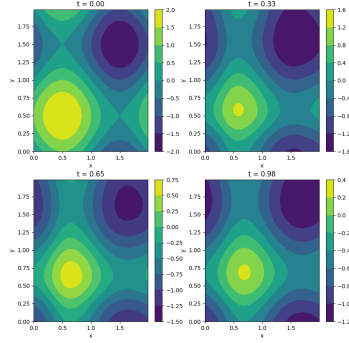
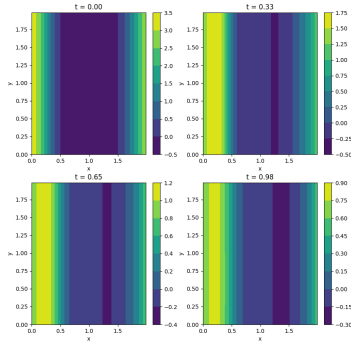
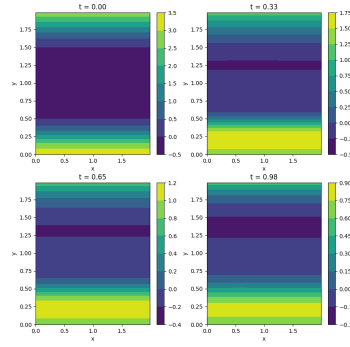
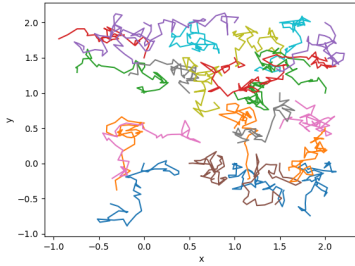
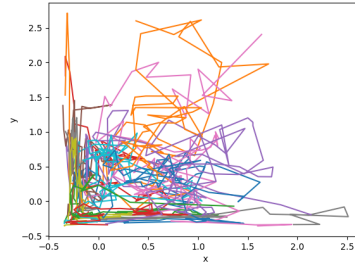
(a)  $(x, y) \mapsto \varphi(x, y, t)$  at different  $t$ (b) The first component of  $(x, y) \mapsto \alpha(x, y, t)$  at different  $t$ (c) The second component of  $(x, y) \mapsto \alpha(x, y, t)$  at different  $t$ (d) Samples of optimal trajectories  $\gamma_s^*$ (e) Samples of optimal controls  $\alpha_s^*$ 

Fig. 7: Depiction of the two-dimensional solution as discussed in Section 3.2.1, utilizing  $n_t = 41$  and  $n_x = n_y = 160$  grid points. Figure (a) illustrates the level sets of the solution  $\varphi(\cdot, t)$  to the viscous HJ PDE (22) at different times  $t$ . Figures (b) and (c) show the first and second components, respectively, of the associated function  $\alpha(\cdot, t)$  from (23) at various times  $t$ , which depict the time reversal of the feedback control function. Figures (d) and (e) present several samples of optimal trajectories  $\gamma_s^*$  along with their corresponding open-loop optimal controls  $\alpha_s^*$ . These trajectories and control strategies solve the stochastic optimal control problem specified in (21), starting from distinct initial conditions  $x$ . Notably, both  $\gamma_s^*$  and  $\alpha_s^*$  take values in  $\mathbb{R}^2$ . For visualization, they are plotted within the spatial domain, excluding the time dimension for clarity.

Upon examining the level set diagrams of  $\varphi$  and  $\alpha$ , it becomes evident that the numerical solutions exhibit greater smoothness compared to those presented in Section 2.3.1. Despite the stochastic nature of the optimal trajectories, which allows us to only display certain samples, there is a discernible pattern where the trajectories seem to cluster around the minimum value of  $g$  at  $t = T$ .

### 3.2.2 One-homogeneous Hamiltonian with spatial dependent coefficients

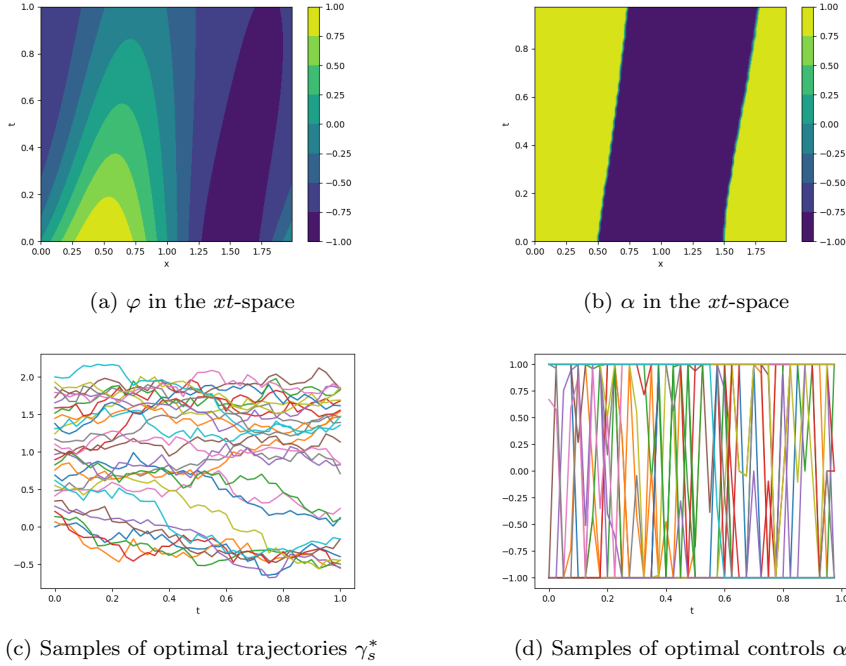


Fig. 8: Visualization of the solution for the one-dimensional scenario discussed in Section 3.2.2, using  $n_t = 41$  and  $n_x = 160$  grid points. Figures (a) and (b) showcase the level sets of the solution  $\varphi$  to the viscous HJ PDE (22), along with the corresponding function  $\alpha$  from (23), which represents the time reversal of the feedback control function. Figures (c) and (d) depict several samples of optimal paths  $\gamma_s^*$  and their associated open-loop optimal controls  $\alpha_s^*$ . These paths and control trajectories are the solutions to the stochastic optimal control problem (21), each beginning from a unique initial condition  $x$ .

This example follows a similar setup to that described in Section 2.3.2, with a diffusion parameter of  $\epsilon = 0.1$ . The outcomes for the one-dimensional scenario are depicted in Figure 8, and those for the two-dimensional scenario

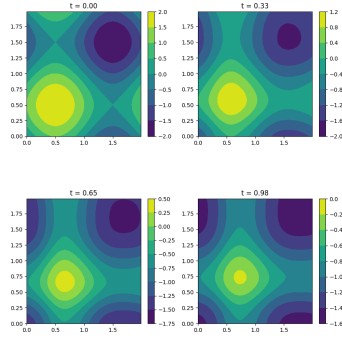
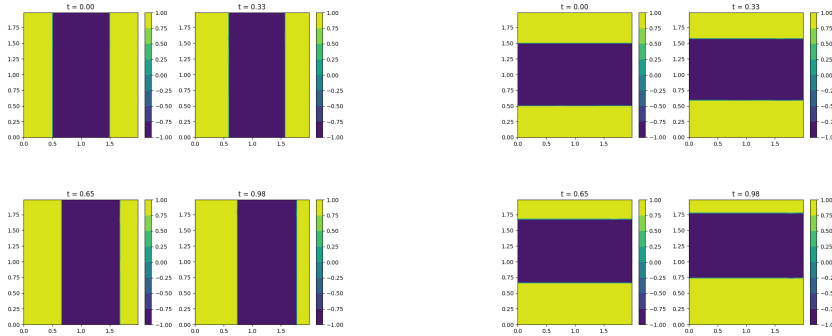
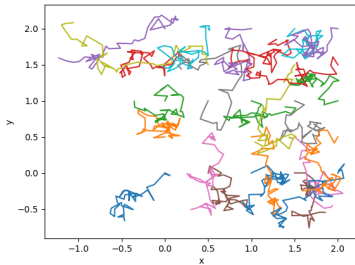
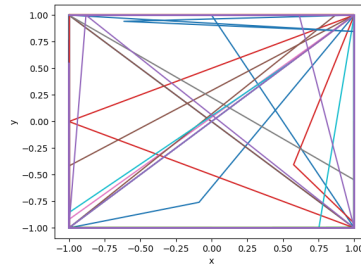
(a)  $(x, y) \mapsto \varphi(x, y, t)$  at different  $t$ (b) The first component of  $(x, y) \mapsto \alpha(x, y, t)$  at different  $t$ (c) The second component of  $(x, y) \mapsto \alpha(x, y, t)$  at different  $t$ (d) Samples of optimal trajectories  $\gamma_s^*$ (e) Samples of optimal controls  $\alpha_s^*$ 

Fig. 9: Depiction of the two-dimensional solution as discussed in Section 3.2.2, utilizing  $n_t = 41$  and  $n_x = n_y = 160$  grid points. Figure (a) illustrates the level sets of the solution  $\varphi(\cdot, t)$  to the viscous HJ PDE (22) at different times  $t$ . Figures (b) and (c) show the first and second components, respectively, of the associated function  $\alpha(\cdot, t)$  from (23) at various times  $t$ , which depict the time reversal of the feedback control function. Figures (d) and (e) present several samples of optimal trajectories  $\gamma_s^*$  along with their corresponding open-loop optimal controls  $\alpha_s^*$ . These trajectories and control strategies solve the stochastic optimal control problem specified in (21), starting from distinct initial conditions  $x$ . Notably, both  $\gamma_s^*$  and  $\alpha_s^*$  take values in  $\mathbb{R}^2$ . For visualization, they are plotted within the spatial domain, excluding the time dimension for clarity.

are presented in Figure 9. The solution  $\varphi$  exhibits greater smoothness compared to scenarios lacking a diffusion term. Nonetheless, the function  $\alpha$  still demonstrates discontinuities. This occurrence is attributable to  $\alpha$  employing the same expression (23) as found in the deterministic optimal control problem in Section 2, thus the rationale from Section 2.3.2 remains applicable. It should be noted that, due to the element of randomness, the trajectories for optimal control display a higher frequency of jumps in comparison to the deterministic configuration.

### 3.2.3 Newton mechanics

This section employs the same functions  $f$ ,  $L$ , and  $g$  detailed in Section 2.3.3. We present the numerical solution for a diffusion coefficient of  $\epsilon = 0.1$  in Figure 10.

## 4 Summary

In this study, we present a framework for solving optimal control problems and HJ PDEs through a novel saddle point formulation, which is solved using an algorithm derived from the preconditioned PDHG method. This approach extends to solve stochastic optimal control problems and associated viscous HJ PDEs. Through a series of numerical examples, we illustrate the efficacy of our algorithm, especially in handling Hamiltonians which depend on spatial and temporal variables. A key advantage of our method is the use of implicit time discretization, which allows for larger time steps than explicit methods, thereby enhancing computational speed. The core strength of our method lies in its straightforward saddle point formulation, linearly related to the solution  $\varphi$  of the HJ PDEs, a feature facilitated by leveraging the inherent relationship between optimal control problems and HJ PDEs. This simplicity enables the method to support explicit updates or to benefit from parallel computation, making it versatile and efficient.

Although our method is of first order, it shows potential as an initial step towards methods of higher accuracy, especially in applications that require small errors. Future work could explore integrating this method with higher-order schemes or employing it within machine learning frameworks, where its formulation could inspire novel loss function designs for solving optimal control problems and related HJ PDEs.

**Acknowledgements** We would like to thank Dr. Donggun Lee for his valuable discussions. Additionally, we are thankful to Dr. Levon Nurbekyan for his assistance with the analysis presented in Appendix A.

### Conflict of interest

The authors declare that they have no conflict of interest.

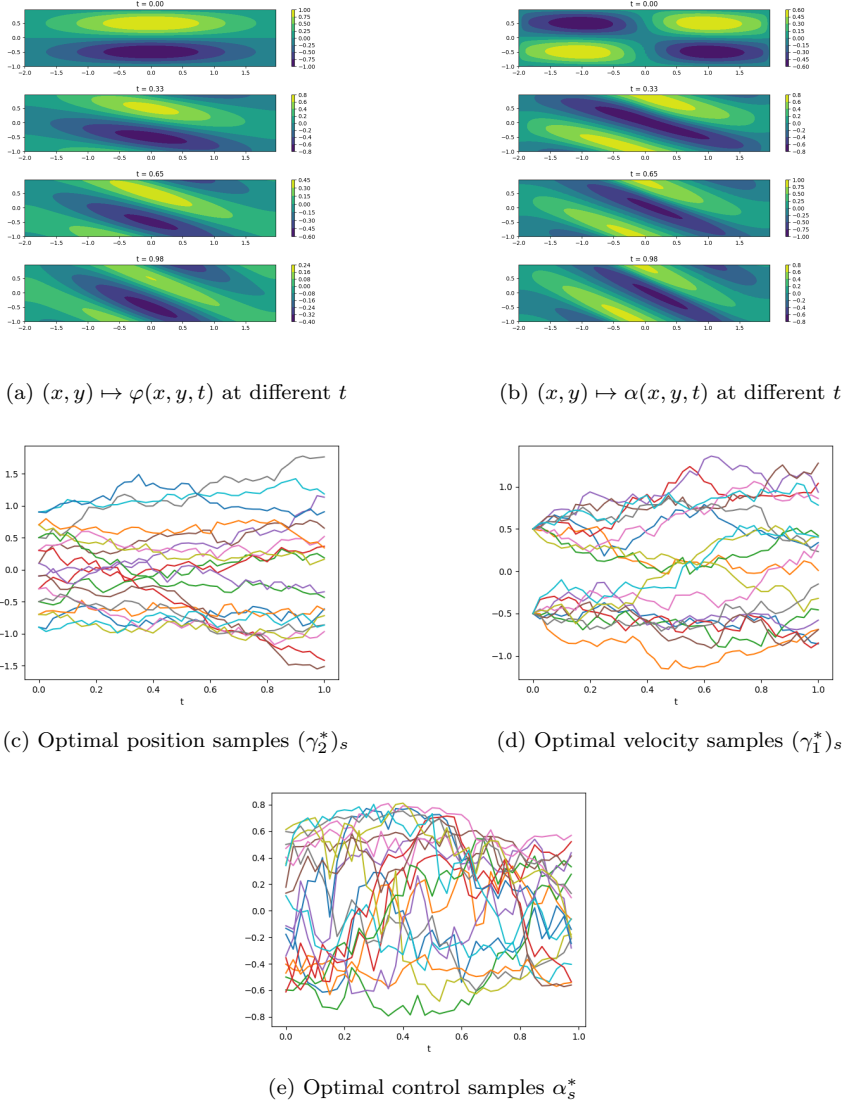


Fig. 10: Visualization of the solution for example discussed in Section 3.2.3, using  $n_t = 41$ ,  $n_x = 160$ , and  $n_y = 80$  grid points. Figures (a) and (b) showcase the level sets of the solution  $\varphi$  to the viscous HJ PDE (22), along with the corresponding function  $\alpha$  from (23), which represents the time reversal of the feedback control function. Figures (c), (d), (e) depict the first and second components of several samples of optimal paths  $\gamma_s^*$  and their associated open-loop optimal controls  $\alpha_s^*$ . These paths and control trajectories are the solutions to the stochastic optimal control problem (21), each beginning from a unique initial condition  $x$ .



## References

1. Akian, M., Bapat, R., Gaubert, S.: Max-plus algebra. *Handbook of linear algebra* **39** (2006)
2. Akian, M., Gaubert, S., Lakhoua, A.: The max-plus finite element method for solving deterministic optimal control problems: basic properties and convergence analysis. *SIAM Journal on Control and Optimization* **47**(2), 817–848 (2008)
3. Alla, A., Falcone, M., Saluzzi, L.: An efficient DP algorithm on a tree-structure for finite horizon optimal control problems. *SIAM Journal on Scientific Computing* **41**(4), A2384–A2406 (2019)
4. Alla, A., Falcone, M., Volkwein, S.: Error analysis for POD approximations of infinite horizon problems via the dynamic programming approach. *SIAM Journal on Control and Optimization* **55**(5), 3091–3115 (2017)
5. Bachouch, A., Huré, C., Langrené, N., Pham, H.: Deep neural networks algorithms for stochastic control problems on finite horizon: numerical applications. *arXiv preprint arXiv:1812.05916* (2018)
6. Bardi, M., Capuzzo-Dolcetta, I.: *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. *Systems & Control: Foundations & Applications*. Birkhäuser Boston, Inc., Boston, MA (1997). DOI 10.1007/978-0-8176-4755-1. With appendices by Maurizio Falcone and Pierpaolo Soravia
7. Bauschke, H.H., Combettes, P.L.: *Convex analysis and monotone operator theory in Hilbert spaces*. *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*. Springer, New York (2011). DOI 10.1007/978-1-4419-9467-7. URL <https://doi.org/10.1007/978-1-4419-9467-7>. With a foreword by Hedy Attouch
8. Bertsekas, D.P.: *Reinforcement learning and optimal control*. Athena Scientific, Belmont, Massachusetts (2019)
9. Bokanowski, O., Garcke, J., Griebel, M., Klompaker, I.: An adaptive sparse grid semi-Lagrangian scheme for first order Hamilton-Jacobi Bellman equations. *Journal of Scientific Computing* **55**(3), 575–605 (2013)
10. Briceno-Arias, L., Kalise, D., Kobeissi, Z., Lauriere, M., González, A.M., Silva, F.J.: On the implementation of a primal-dual algorithm for second order time-dependent mean field games with local couplings. *ESAIM: Proceedings and Surveys* **65**, 330–348 (2019)
11. Briceno-Arias, L.M., Kalise, D., Silva, F.J.: Proximal methods for stationary mean field games with local couplings. *SIAM Journal on Control and Optimization* **56**(2), 801–836 (2018)
12. Cardaliaguet, P., Lehalle, C.A.: Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics* **12**, 335–363 (2018)
13. Carmona, R., Fouque, J., Sun, L.: *Mean field games and systemic risk*. *Communications in Mathematical Sciences* (2015)
14. Carrillo, J.A., Wang, L., Wei, C.: Structure preserving primal dual methods for gradient flows with nonlinear mobility transport distances. *arXiv preprint arXiv:2303.16534* (2023)
15. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision* **40**, 120–145 (2011)
16. Chen, M., Hu, Q., Fisac, J.F., Akametalu, K., Mackin, C., Tomlin, C.J.: Reachability-based safety and goal satisfaction of unmanned aerial platoons on air highways. *Journal of Guidance, Control, and Dynamics* **40**(6), 1360–1373 (2017). DOI 10.2514/1.G000774. URL <https://doi.org/10.2514/1.G000774>
17. Chen, P., Darbon, J., Meng, T.: Hopf-type representation formulas and efficient algorithms for certain high-dimensional optimal control problems. *arXiv preprint arXiv:2110.02541* (2021)
18. Chen, P., Darbon, J., Meng, T.: Lax-Oleinik-type formulas and efficient algorithms for certain high-dimensional optimal control problems. *arXiv preprint arXiv:2109.14849* (2021)
19. Chow, Y.T., Darbon, J., Osher, S., Yin, W.: Algorithm for overcoming the curse of dimensionality for time-dependent non-convex Hamilton–Jacobi equations arising from optimal control and differential games problems. *Journal of Scientific Computing* **73**, 617–643 (2017)

20. Chow, Y.T., Darbon, J., Osher, S., Yin, W.: Algorithm for overcoming the curse of dimensionality for state-dependent Hamilton-Jacobi equations. *Journal of Computational Physics* **387**, 376–409 (2019)
21. Coupechoux, M., Darbon, J., Kélif, J., Sigelle, M.: Optimal trajectories of a uav base station using lagrangian mechanics. In: *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 626–631 (2019)
22. Darbon, J.: On convex finite-dimensional variational methods in imaging sciences and Hamilton–Jacobi equations. *SIAM Journal on Imaging Sciences* **8**(4), 2268–2293 (2015). DOI 10.1137/130944163
23. Darbon, J., Dower, P.M., Meng, T.: Neural network architectures using min-plus algebra for solving certain high-dimensional optimal control problems and Hamilton–Jacobi PDEs. *Mathematics of Control, Signals, and Systems* **35**(1), 1–44 (2023)
24. Darbon, J., Langlois, G.P.: Accelerated nonlinear primal-dual hybrid gradient methods with applications to supervised machine learning. *arXiv preprint arXiv:2109.12222* (2021)
25. Darbon, J., Langlois, G.P., Meng, T.: Overcoming the curse of dimensionality for some Hamilton–Jacobi partial differential equations via neural network architectures. *Res. Math. Sci.* **7**(3), 20 (2020). DOI 10.1007/s40687-020-00215-6. URL <https://doi.org/10.1007/s40687-020-00215-6>
26. Darbon, J., Meng, T.: On decomposition models in imaging sciences and multi-time Hamilton–Jacobi partial differential equations. *SIAM Journal on Imaging Sciences* **13**(2), 971–1014 (2020). DOI 10.1137/19M1266332. URL <https://doi.org/10.1137/19M1266332>
27. Darbon, J., Meng, T.: On some neural network architectures that can represent viscosity solutions of certain high dimensional hamilton–jacobi partial differential equations. *Journal of Computational Physics* **425**, 109907 (2021)
28. Darbon, J., Osher, S.: Algorithms for overcoming the curse of dimensionality for certain Hamilton–Jacobi equations arising in control theory and elsewhere. *Research in the Mathematical Sciences* **3**(1), 19 (2016). DOI 10.1186/s40687-016-0068-7
29. Darbon, J., Osher, S.J.: Splitting enables overcoming the curse of dimensionality. *Splitting Methods in Communication, Imaging, Science, and Engineering* pp. 427–432 (2016)
30. Delahaye, D., Puechmorel, S., Tsiotras, P., Feron, E.: Mathematical models for aircraft trajectory design: A survey. In: *Air Traffic Management and Systems*, pp. 205–247. Springer Japan, Tokyo (2014)
31. Denk, J., Schmidt, G.: Synthesis of a walking primitive database for a humanoid robot using optimal control techniques. In: *Proceedings of IEEE-RAS International Conference on Humanoid Robots*, pp. 319–326 (2001)
32. Djeridane, B., Lygeros, J.: Neural approximation of PDE solutions: An application to reachability computations. In: *Proceedings of the 45th IEEE Conference on Decision and Control*, pp. 3034–3039 (2006). DOI 10.1109/CDC.2006.377184
33. Dolgov, S., Kalise, D., Kunisch, K.: A tensor decomposition approach for high-dimensional Hamilton–Jacobi–Bellman equations. *arXiv preprint arXiv:1908.01533* (2019)
34. Dower, P.M., McEneaney, W.M., Zhang, H.: Max-plus fundamental solution semigroups for optimal control problems. In: *2015 Proceedings of the Conference on Control and its Applications*, pp. 368–375. SIAM (2015)
35. E, W., Han, J., Li, Q.: A mean-field optimal control formulation of deep learning. *Research in the Mathematical Sciences* **6**(1), 1–41 (2019)
36. El Khoury, A., Lamiraux, F., Taïx, M.: Optimal motion planning for humanoid robots. In: *2013 IEEE International Conference on Robotics and Automation*, pp. 3136–3141 (2013). DOI 10.1109/ICRA.2013.6631013
37. Fallon, M., Kuindersma, S., Karumanchi, S., Antone, M., Schneider, T., Dai, H., D’Arpino, C.P., Deits, R., DiCicco, M., Fourie, D., et al.: An architecture for online affordance-based perception and whole-body planning. *Journal of Field Robotics* **32**(2), 229–254 (2015)
38. Feng, S., Whitman, E., Xinjilefu, X., Atkeson, C.G.: Optimization based full body control for the atlas robot. In: *2014 IEEE-RAS International Conference on Humanoid Robots*, pp. 120–127 (2014). DOI 10.1109/HUMANOIDS.2014.7041347

39. Feng Lin, Brandt, R.D.: An optimal control approach to robust control of robot manipulators. *IEEE Transactions on Robotics and Automation* **14**(1), 69–77 (1998). DOI 10.1109/70.660845
40. Fleming, W., McEneaney, W.: A max-plus-based algorithm for a Hamilton–Jacobi–Bellman equation of nonlinear filtering. *SIAM Journal on Control and Optimization* **38**(3), 683–710 (2000). DOI 10.1137/S0363012998332433
41. Fu, G., Osher, S., Li, W.: High order spatial discretization for variational time implicit schemes: Wasserstein gradient flows and reaction-diffusion systems. *arXiv preprint arXiv:2303.08950* (2023)
42. Fujiwara, K., Kajita, S., Harada, K., Kaneko, K., Morisawa, M., Kanehiro, F., Nakaoka, S., Hirukawa, H.: An optimal planning of falling motions of a humanoid robot. In: 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 456–462 (2007). DOI 10.1109/IROS.2007.4399327
43. Garcke, J., Kröner, A.: Suboptimal feedback control of PDEs by solving HJB equations on adaptive sparse grids. *Journal of Scientific Computing* **70**(1), 1–28 (2017)
44. Gaubert, S., McEneaney, W., Qu, Z.: Curse of dimensionality reduction in max-plus based approximation methods: Theoretical estimates and improved pruning algorithms. In: 2011 50th IEEE Conference on Decision and Control and European Control Conference, pp. 1054–1061. IEEE (2011)
45. Gomes, D., Nurbekyan, L., Pimentel, E.: Economic models and mean-field games theory (2015)
46. Gomes, D.A., Saúde, J.: Mean field games models—a brief survey. *Dynamic Games and Applications* **4**, 110–154 (2014)
47. Han, J., Jentzen, A., E, W.: Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences* **115**(34), 8505–8510 (2018). DOI 10.1073/pnas.1718942115
48. Hiriart-Urruty, J.B., Lemaréchal, C.: Convex analysis and minimization algorithms. II, *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 306. Springer-Verlag, Berlin (1993). Advanced theory and bundle methods
49. Hofer, M., Muehlebach, M., D’Andrea, R.: Application of an approximate model predictive control scheme on an unmanned aerial vehicle. In: 2016 IEEE International Conference on Robotics and Automation (ICRA), pp. 2952–2957 (2016). DOI 10.1109/ICRA.2016.7487459
50. Horowitz, M.B., Damle, A., Burdick, J.W.: Linear Hamilton Jacobi Bellman equations in high dimensions. In: 53rd IEEE Conference on Decision and Control, pp. 5880–5887. IEEE (2014)
51. Hu, C., Shu, C.: A discontinuous Galerkin finite element method for Hamilton–Jacobi equations. *SIAM Journal on Scientific Computing* **21**(2), 666–690 (1999). DOI 10.1137/S1064827598337282
52. Huang, M., Malhamé, R.P., Caines, P.E.: Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle
53. Huré, C., Pham, H., Bachouch, A., Langrené, N.: Deep neural networks algorithms for stochastic control problems on finite horizon, part I: convergence analysis. *arXiv preprint arXiv:1812.04300* (2018)
54. Huré, C., Pham, H., Warin, X.: Some machine learning schemes for high-dimensional nonlinear PDEs. *arXiv preprint arXiv:1902.01599* (2019)
55. Jacobs, M., Léger, F., Li, W., Osher, S.: Solving large-scale optimization problems with a convergence rate independent of grid size. *SIAM Journal on Numerical Analysis* **57**(3), 1100–1123 (2019)
56. Jiang, F., Chou, G., Chen, M., Tomlin, C.J.: Using neural networks to compute approximate and guaranteed feasible Hamilton-Jacobi-Bellman PDE solutions. *arXiv preprint arXiv:1611.03158* (2016)
57. Jiang, G., Peng, D.: Weighted ENO schemes for Hamilton–Jacobi equations. *SIAM Journal on Scientific Computing* **21**(6), 2126–2143 (2000). DOI 10.1137/S106482759732455X
58. Jin, L., Li, S., Yu, J., He, J.: Robot manipulator control using neural networks: A survey. *Neurocomputing* **285**, 23 – 34 (2018). DOI <https://doi.org/10.1016/j.neucom.2018.01.002>. URL <http://www.sciencedirect.com/science/article/pii/S0925231218300158>

59. Kalise, D., Kundu, S., Kunisch, K.: Robust feedback control of nonlinear PDEs by numerical approximation of high-dimensional Hamilton-Jacobi-Isaacs equations. arXiv preprint arXiv:1905.06276 (2019)
60. Kalise, D., Kunisch, K.: Polynomial approximation of high-dimensional Hamilton-Jacobi-Bellman equations and applications to feedback control of semilinear parabolic PDEs. *SIAM Journal on Scientific Computing* **40**(2), A629–A652 (2018)
61. Kang, W., Wilcox, L.C.: Mitigating the curse of dimensionality: sparse grid characteristics method for optimal feedback control and HJB equations. *Computational Optimization and Applications* **68**(2), 289–315 (2017)
62. Kim, Y.H., Lewis, F.L., Dawson, D.M.: Intelligent optimal control of robotic manipulators using neural networks. *Automatica* **36**(9), 1355 – 1364 (2000). DOI [https://doi.org/10.1016/S0005-1098\(00\)00045-5](https://doi.org/10.1016/S0005-1098(00)00045-5). URL <http://www.sciencedirect.com/science/article/pii/S0005109800000455>
63. Kirchner, M.R., Hewer, G., Darbon, J., Osher, S.: A primal-dual method for optimal control and trajectory generation in high-dimensional systems. In: 2018 IEEE Conference on Control Technology and Applications (CCTA), pp. 1583–1590. IEEE (2018)
64. Kirchner, M.R., Mar, R., Hewer, G., Darbon, J., Osher, S., Chow, Y.T.: Time-optimal collaborative guidance using the generalized Hopf formula. *IEEE Control Systems Letters* **2**(2), 201–206 (2017)
65. Kizilkale, A.C., Salhab, R., Malhamé, R.P.: An integral control formulation of mean field game based large scale coordination of loads in smart grids. *Automatica* **100**, 312–322 (2019)
66. Kuindersma, S., Deits, R., Fallon, M., Valenzuela, A., Dai, H., Permenter, F., Koolen, T., Marion, P., Tedrake, R.: Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot. *Autonomous robots* **40**(3), 429–455 (2016)
67. Kunisch, K., Volkwein, S., Xie, L.: HJB-POD-based feedback design for the optimal control of evolution problems. *SIAM Journal on Applied Dynamical Systems* **3**(4), 701–722 (2004)
68. Lachapelle, A., Wolfram, M.T.: On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transportation research part B: methodological* **45**(10), 1572–1589 (2011)
69. Lambrianides, P., Gong, Q., Venturi, D.: A new scalable algorithm for computational optimal control under uncertainty. arXiv preprint arXiv:1909.07960 (2019)
70. Lasry, J.M., Lions, P.L.: Mean field games. *Japanese journal of mathematics* **2**(1), 229–260 (2007)
71. Lee, D., Deka, S.A., Tomlin, C.J.: Convexifying state-constrained optimal control problems. *IEEE Transactions on Automatic Control* (2022)
72. Lee, D., Tomlin, C.J.: A Hopf-Lax formula in Hamilton-Jacobi analysis of reach-avoid problems. *IEEE Control Systems Letters* **5**(3), 1055–1060 (2021). DOI 10.1109/LCSYS.2020.3009933
73. Lee, W., Liu, S., Tembine, H., Li, W., Osher, S.: Controlling propagation of epidemics via mean-field control. *SIAM Journal on Applied Mathematics* **81**(1), 190–207 (2021). DOI 10.1137/20M1342690
74. Lewis, F., Dawson, D., Abdallah, C.: Robot Manipulator Control: Theory and Practice. Control engineering. Marcel Dekker (2004). URL [https://books.google.com/books?id=BDS\\_PQAACAAJ](https://books.google.com/books?id=BDS_PQAACAAJ)
75. Lin, A.T., Chow, Y.T., Osher, S.: 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. arXiv preprint arXiv:1803.01215 (2018)
76. Liu, S., Liu, S., Osher, S., Li, W.: A first-order computational algorithm for reaction-diffusion type equations via primal-dual hybrid gradient method (2023)
77. Liu, S., Osher, S., Li, W., Shu, C.W.: A primal-dual approach for solving conservation laws with implicit in time approximations. *Journal of Computational Physics* **472**, 111654 (2023)
78. McEneaney, W.: Max-plus methods for nonlinear control and estimation. Springer Science & Business Media, Boston, MA (2006)

79. McEneaney, W.: A curse-of-dimensionality-free numerical method for solution of certain HJB PDEs. *SIAM Journal on Control and Optimization* **46**(4), 1239–1276 (2007). DOI 10.1137/040610830
80. McEneaney, W.M., Deshpande, A., Gaubert, S.: Curse-of-complexity attenuation in the curse-of-dimensionality-free method for HJB PDEs. In: 2008 American Control Conference, pp. 4684–4690. IEEE (2008)
81. McEneaney, W.M., Kluberg, L.J.: Convergence rate for a curse-of-dimensionality-free method for a class of HJB PDEs. *SIAM Journal on Control and Optimization* **48**(5), 3052–3079 (2009)
82. Meng, T., Hao, W., Liu, S., Osher, S.J., Li, W.: Primal-dual hybrid gradient algorithms for computing time-implicit Hamilton-Jacobi equations. *arXiv preprint arXiv:2310.01605* (2023)
83. Niarchos, K.N., Lygeros, J.: A neural approximation to continuous time reachability computations. In: Proceedings of the 45th IEEE Conference on Decision and Control, pp. 6313–6318 (2006). DOI 10.1109/CDC.2006.377358
84. Osher, S., Shu, C.: High-order essentially nonoscillatory schemes for Hamilton-Jacobi equations. *SIAM Journal on Numerical Analysis* **28**(4), 907–922 (1991). DOI 10.1137/0728049
85. Papadakis, N., Peyré, G., Oudet, E.: Optimal transport with proximal splitting. *SIAM Journal on Imaging Sciences* **7**(1), 212–238 (2014)
86. Parzani, C., Puechmorel, S.: On a Hamilton-Jacobi-Bellman approach for coordinated optimal aircraft trajectories planning. In: CCC 2017 36th Chinese Control Conference, Control Conference (CCC), 2017 36th Chinese, pp. ISBN: 978-1-5386-2918-5. IEEE, Dalian, China (2017). DOI 10.23919/ChiCC.2017.8027369. URL <https://hal-enac.archives-ouvertes.fr/hal-01340565>
87. Reisinger, C., Zhang, Y.: Rectified deep neural networks overcome the curse of dimensionality for nonsmooth value functions in zero-sum games of nonlinear stiff systems. *arXiv preprint arXiv:1903.06652* (2019)
88. Royo, V.R., Tomlin, C.: Recursive regression with neural networks: Approximating the HJI PDE solution. *arXiv preprint arXiv:1611.02739* (2016)
89. Rucco, A., Sujit, P.B., Aguiar, A.P., de Sousa, J.B., Pereira, F.L.: Optimal rendezvous trajectory for unmanned aerial-ground vehicles. *IEEE Transactions on Aerospace and Electronic Systems* **54**(2), 834–847 (2018). DOI 10.1109/TAES.2017.2767958
90. Sirignano, J., Spiliopoulos, K.: DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics* **375**, 1339 – 1364 (2018). DOI 10.1016/j.jcp.2018.08.029
91. Todorov, E.: Efficient computation of optimal actions. *Proceedings of the national academy of sciences* **106**(28), 11478–11483 (2009)
92. Valkonen, T.: A primal–dual hybrid gradient method for nonlinear operators with applications to MRI. *Inverse Problems* **30**(5), 055012 (2014)
93. Yegorov, I., Dower, P.M.: Perspectives on characteristics based curse-of-dimensionality-free numerical approaches for solving Hamilton–Jacobi equations. *Applied Mathematics & Optimization* pp. 1–49 (2017)
94. Yong, J., Zhou, X.Y.: *Stochastic controls, Applications of Mathematics (New York)*, vol. 43. Springer-Verlag, New York (1999). DOI 10.1007/978-1-4612-1466-3. URL <https://doi.org/10.1007/978-1-4612-1466-3>. Hamiltonian systems and HJB equations
95. Zhou, M., Han, J., Lu, J.: Actor-critic method for high dimensional static Hamilton–Jacobi–Bellman partial differential equations based on neural networks. *SIAM Journal on Scientific Computing* **43**(6), A4043–A4066 (2021)
96. Zhou, M., Lu, J.: A policy gradient framework for stochastic optimal control problems with global convergence guarantee. *arXiv preprint arXiv:2302.05816* (2023)

## A The convergence of the PDHG algorithms in Section 2.1

In this section, we show the convergence of the PDHG algorithm to a saddle point in certain cases.

Let  $\Omega \subset \mathbb{R}^n$  be a bounded rectangular domain with periodic boundary condition. We consider the case when  $f_{x,t}$  is an affine function for any  $x \in \Omega$  and  $t \in [0, T]$ , i.e.,  $f(x, t, \alpha) = A(x, t)\alpha + b(x, t)$  for continuous functions  $A: \Omega \times [0, T] \rightarrow \mathbb{R}^{n \times m}$  (where  $\mathbb{R}^{n \times m}$  denotes the set of matrices with  $n$  rows and  $m$  columns) and  $b: \Omega \times [0, T] \rightarrow \mathbb{R}^n$ . Assume  $L_{x,t}: \mathbb{R}^m \rightarrow [0, +\infty]$  is proper, convex, lower semi-continuous, 1-coercive, and satisfies  $L_{x,t}(0) = 0$  for any  $x \in \Omega$  and  $t \in [0, T]$ . To simplify the notation, we use  $A_{x,t}$  and  $b_{x,t}$  to denote  $A(x, T-t)$  and  $b(x, T-t)$ . Then, we have  $f_{x,t}(\alpha) = A_{x,t}\alpha + b_{x,t}$ . In this case, the saddle point problem (6) becomes

$$\min_{\varphi} \max_{\rho \geq 0, \alpha} \int_0^T \int_{\Omega} \rho(x, t) \left( \frac{\partial \varphi(x, t)}{\partial t} - \langle A_{x,t}\alpha(x, t) + b_{x,t}, \nabla_x \varphi(x, t) \rangle - L_{x,t}(\alpha(x, t)) \right) dx dt - c \int_{\Omega} \varphi(x, T) dx. \quad (26)$$

With a change of variable  $m(x, t) = \rho(x, t)\alpha(x, t)$ , the problem becomes

$$\min_{\varphi} \max_{\rho \geq 0, m} \int_0^T \int_{\Omega} \rho(x, t) \frac{\partial \varphi(x, t)}{\partial t} - m(x, t)^T A_{x,t}^T \nabla_x \varphi(x, t) - \rho(x, t) b_{x,t}^T \nabla_x \varphi(x, t) - \rho(x, t) L_{x,t} \left( \frac{m(x, t)}{\rho(x, t)} \right) dx dt - c \int_{\Omega} \varphi(x, T) dx, \quad (27)$$

where  $\rho L_{x,t}(\frac{m}{\rho})$  is defined to be  $\chi_{\{0\}}(m)$  if  $\rho$  is zero in order to get a lower semi-continuous function (by [48, Prop.X.1.2.1] and the assumption that  $L_{x,t}$  is 1-coercive which implies  $\text{dom } L_{x,t}^* = \mathbb{R}^m$ ). Recall that  $\chi_C$  denotes the indicator function of a set  $C$  which takes value 0 at points in  $C$  and  $+\infty$  otherwise. Let  $X = H^1(\Omega \times [0, T])$  and  $Y = L^2(\Omega \times [0, T]) \times L^2(\Omega \times [0, T]; \mathbb{R}^m)$ . Now, we prove the convergence of the PDHG algorithm applied to (27) for  $\varphi \in X$  and  $(\rho, m) \in Y$ .

Define the operators  $K: X \times Y \rightarrow \mathbb{R}$ ,  $G: X \rightarrow \mathbb{R} \cup \{+\infty\}$ , and  $\hat{F}: Y \rightarrow \mathbb{R} \cup \{+\infty\}$  as follows

$$\begin{aligned} K(\varphi, \rho, m) &= \int_0^T \int_{\Omega} \rho(x, t) \frac{\partial \varphi(x, t)}{\partial t} - m(x, t)^T A_{x,t}^T \nabla_x \varphi(x, t) - \rho(x, t) b_{x,t}^T \nabla_x \varphi(x, t) dx dt, \\ G(\varphi) &= -c \int_{\Omega} \varphi(x, T) dx + \chi_{\{\varphi: \varphi(\cdot, 0) = g\}}(\varphi), \\ \hat{F}(\rho, m) &= \int_0^T \int_{\Omega} \rho(x, t) L_{x,t} \left( \frac{m(x, t)}{\rho(x, t)} \right) + \chi_{[0, +\infty)}(\rho(x, t)) dx dt. \end{aligned}$$

It is straightforward to check that  $K$  is bilinear with respect to  $\varphi$  and  $(\rho, m)$ , and  $G, \hat{F}$  are convex and proper functions (recall that a function is proper means it is not constantly  $+\infty$ ).

Now, we prove that  $G$  is lower semi-continuous on  $X$ . Let  $\varphi_j$  be a sequence converging to  $\varphi$  in  $X$ . By trace theorem,  $\varphi_j(\cdot, T)$  converges to  $\varphi(\cdot, T)$  in  $L^2(\Omega)$ . Since  $\Omega$  is bounded, we have  $\lim_{j \rightarrow \infty} \int_{\Omega} \varphi_j(x, T) dx = \int_{\Omega} \varphi(x, T) dx$ . Similarly, by trace theorem,  $\varphi_j(\cdot, 0)$  converges to  $\varphi(\cdot, 0)$  in  $L^2(\Omega)$ , and hence there exists a subsequence of  $\varphi_j$ , denoted by  $\varphi_{j_i}$ , such that  $\varphi_{j_i}(\cdot, 0)$  converges to  $\varphi(\cdot, 0)$  almost everywhere. As a result,  $\liminf_{j \rightarrow \infty} \chi_{\{\varphi: \varphi(\cdot, 0) = g\}}(\varphi_j) \geq \chi_{\{\varphi: \varphi(\cdot, 0) = g\}}(\varphi)$  holds. Then, we conclude that  $G$  is lower semi-continuous.

We also prove that  $\hat{F}$  is lower semi-continuous on  $Y$ . Let  $(\rho_j, m_j)$  be a convergent sequence in  $Y$  with limit  $(\rho, m)$ . We need to prove  $\liminf_{j \rightarrow \infty} \hat{F}(\rho_j, m_j) \geq \hat{F}(\rho, m)$ . Without loss of generality, we can assume  $\rho_j$  is non-negative almost everywhere for all  $j$ . Since  $L$  is non-negative,  $\rho_j(x, t) L_{x,t}(\frac{m_j(x, t)}{\rho_j(x, t)})$  is non-negative almost everywhere for all  $j$ . Since  $(\rho_j, m_j)$  converges to  $(\rho, m)$  in  $L^2$ , there exists a subsequence  $(\rho_{j_i}, m_{j_i})$  converging to  $(\rho, m)$  pointwisely almost everywhere. According to the lower semi-continuity of  $L_{x,t}$  (and using the 1-coercivity of  $L_{x,t}$  in the case of  $\rho(x, t) = 0$ ), we obtain

$$\liminf_{i \rightarrow \infty} \rho_{j_i}(x, t) L_{x,t} \left( \frac{m_{j_i}(x, t)}{\rho_{j_i}(x, t)} \right) \geq \rho(x, t) L_{x,t} \left( \frac{m(x, t)}{\rho(x, t)} \right) \text{ a.e. } x \in \Omega, t \in [0, T].$$

Then, we get  $\liminf_{j \rightarrow \infty} \hat{F}(\rho_j, m_j) \geq \hat{F}(\rho, m)$  by Fatou's lemma. Define  $F$  as the Fenchel-Legendre transform of  $\hat{F}$ , i.e.,  $F = \hat{F}^*$ . Since  $\hat{F}$  is proper, convex, and lower semi-continuous on  $Y$ , by [7, Thm 13.37], we have  $\hat{F} = F^*$ .

Now, we prove the continuity of  $K$  as follows

$$\begin{aligned} |K(\varphi, \rho, m)| &\leq \max \left\{ 1, \sup_{x,t} \|b_{x,t}\|_\infty \right\} \|\rho\|_{L^2} \|\varphi\|_{H^1} + \sup_{x,t,i,j} |A_{x,t,i,j}| \|m\|_{L^2} \|\varphi\|_{H^1} \\ &\leq \left( \max \left\{ 1, \sup_{x,t} \|b_{x,t}\|_\infty \right\}^2 + \sup_{x,t,i,j} |A_{x,t,i,j}|^2 \right) \|(\rho, m)\|_{L^2} \|\varphi\|_{H^1}. \end{aligned}$$

Therefore,  $K$  is continuous with operator norm  $\|K\| \leq \max\{1, \sup_{x,t} \|b_{x,t}\|_\infty\}^2 + \sup_{x,t,i,j} |A_{x,t,i,j}|^2$ .

Then, the saddle point problem (27) is written as the standard form for PDHG algorithm:

$$\min_{\varphi \in X} \max_{(\rho, m) \in Y} K(\varphi, \rho, m) + G(\varphi) - F^*(\rho, m),$$

where  $K$  is bilinear and continuous with respect to  $\varphi$  and  $(\rho, m)$ ,  $G$  and  $F$  are both proper, convex and lower semi-continuous. The  $\ell$ -th iteration of the PDHG updates is given by

$$\begin{cases} (\rho^{\ell+1}, m^{\ell+1}) = \arg \max_{(\rho, m) \in Y} \left\{ K(\tilde{\varphi}^\ell, \rho, m) - F^*(\rho, m) - \frac{1}{2\tau_{\rho, m}} \left( \|\rho - \rho^\ell\|_{L^2}^2 + \|m - m^\ell\|_{L^2}^2 \right) \right\}, \\ \varphi^{\ell+1} = \arg \min_{\varphi \in X} \left\{ K(\varphi, \rho^{\ell+1}, m^{\ell+1}) + G(\varphi) + \frac{1}{2\tau_\varphi} \|\varphi - \varphi^\ell\|_{H^1}^2 \right\}, \\ \tilde{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^\ell. \end{cases} \quad (28)$$

Then, the convergence of these updates follows from [24] if the stepsizes  $\tau_\varphi$  and  $\tau_{\rho, m}$  are small enough. To be specific, assume the problem (27) has a saddle point, the stepsizes satisfy  $\tau_\varphi \tau_{\rho, m} < \frac{1}{\|K\|^2}$ , where  $\|K\|$  is estimated above, and let  $\varphi^j, \rho^j, m^j$  be the functions given by the algorithm (28). Then, the sequence  $\{\tilde{\varphi}^\ell, \tilde{\rho}^\ell, \tilde{m}^\ell\}_\ell$  has a subsequence which converges weakly to a saddle point of (27), where  $\tilde{\varphi}^\ell, \tilde{\rho}^\ell, \tilde{m}^\ell$  are the average functions defined by  $(\tilde{\varphi}^\ell, \tilde{\rho}^\ell, \tilde{m}^\ell) = \frac{1}{\ell} \sum_{j=1}^\ell (\varphi^j, \rho^j, m^j)$ .

**Remark A1 (Connection of the PDHG updates above with our proposed updates)**

First, we work on the details about the update for  $\varphi$  as follows. After simplifying the updates for  $\varphi$ , we get

$$\begin{aligned} \varphi^{\ell+1} &= \arg \min_{\substack{\varphi \in X \\ \varphi(\cdot, 0) = g}} \left\{ \int_0^T \int_\Omega \rho \frac{\partial \varphi}{\partial t} - m(x, t)^T A_{x,t}^T \nabla_x \varphi - \rho(x, t) b_{x,t}^T \nabla_x \varphi dx dt - c \int_\Omega \varphi(x, T) dx + \frac{1}{2\tau_\varphi} \|\varphi - \varphi^\ell\|_{H^1}^2 \right\} \\ &= \arg \min_{\substack{\varphi \in X \\ \varphi(\cdot, 0) = g}} \left\{ \int_0^T \int_\Omega \left( -\partial_t \rho^{\ell+1} + \nabla_x \cdot (A_{x,t} m^{\ell+1} + \rho^{\ell+1} b_{x,t}) \right) \varphi dx dt + \int_\Omega (\rho^{\ell+1}(x, T) - c) \varphi(x, T) dx + \frac{1}{2\tau_\varphi} \|\varphi - \varphi^\ell\|_{H^1}^2 \right\}. \end{aligned}$$

If  $\rho^{\ell+1}$  satisfies the terminal condition  $\rho^{\ell+1}(x, T) = c$  for all  $x \in \Omega$ , the first order optimality condition gives

$$0 = \tau_\varphi \left( -\partial_t \rho^{\ell+1} + \nabla_x \cdot (A_{x,t} m^{\ell+1} + \rho^{\ell+1} b_{x,t}) \right) + (I - \Delta)(\varphi^{\ell+1} - \varphi^\ell),$$

which implies

$$\varphi^{\ell+1} = \varphi^\ell + \tau_\varphi (I - \Delta)^{-1} \left( \partial_t \rho^{\ell+1} - \nabla_x \cdot (A_{x,t} m^{\ell+1}(x, t) + \rho^{\ell+1}(x, t) b_{x,t}) \right).$$

With a change of variable  $\alpha = \frac{m}{\rho}$ , this update on  $\varphi$  is the same as the third line in (9).

Then, we focus on the update for  $(\rho, m)$  (the first line in (28)), which, after the change of variable, becomes

$$\max_{\rho, \alpha} \left\{ \mathcal{L}(\tilde{\varphi}^\ell, \rho, \alpha) - \frac{1}{2\tau_{\rho, m}} \|\rho - \rho^\ell\|_{L^2}^2 - \frac{1}{2\tau_{\rho, m}} \|\rho \alpha - \rho^\ell \alpha^\ell\|_{L^2}^2 \right\},$$

where  $\mathcal{L}$  is the objective function in (26). To numerically solve this problem, we apply the alternative updates on  $\rho$  and  $\alpha$ . Fixing  $\alpha = \alpha^\ell$ , the update for  $\rho$  gives the first line in (9) with  $\tau_\rho = \tau_{\rho,m}$ . Then, fixing  $\rho = \rho^{\ell+1}$  and approximating  $\rho^\ell \alpha^\ell$  by  $\rho^{\ell+1} \alpha^\ell$ , the update for  $\alpha$  is

$$\begin{aligned} \alpha^{\ell+1} &= \arg \max_{\alpha} \left\{ \mathcal{L}(\tilde{\varphi}^\ell, \rho^{\ell+1}, \alpha) - \frac{1}{2\tau_{\rho,m}} \|\rho^{\ell+1} \alpha - \rho^{\ell+1} \alpha^\ell\|_{L^2}^2 \right\} \\ &= \arg \max_{\alpha} \int_0^T \int_{\Omega} -\rho^{\ell+1}(x,t) \langle A_{x,t} \alpha(x,t), \nabla_x \varphi^\ell(x,t) \rangle - \rho^{\ell+1}(x,t) L_{x,t}(\alpha(x,t)) \\ &\quad - \frac{\rho^{\ell+1}(x,t)^2}{2\tau_{\rho,m}} (\alpha(x,t) - \alpha^\ell(x,t))^2 dx dt. \end{aligned}$$

This optimization problem can be solved in parallel for any  $(x,t)$  by updating  $\alpha(x,t)$  as follows

$$\begin{aligned} \alpha^{\ell+1}(x,t) &= \arg \max_{\alpha \in \mathbb{R}^m} \left\{ -\rho^{\ell+1}(x,t) \langle A_{x,t} \alpha, \nabla_x \varphi^\ell(x,t) \rangle - \rho^{\ell+1}(x,t) L_{x,t}(\alpha) - \frac{\rho^{\ell+1}(x,t)^2}{2\tau_{\rho,m}} (\alpha - \alpha^\ell(x,t))^2 \right\} \\ &= \arg \min_{\alpha \in \mathbb{R}^m} \left\{ \langle A_{x,t} \alpha, \nabla_x \varphi^\ell(x,t) \rangle + L_{x,t}(\alpha) + \frac{\rho^{\ell+1}(x,t)}{2\tau_{\rho,m}} (\alpha - \alpha^\ell(x,t))^2 \right\}, \end{aligned}$$

which gives the second line in (9) with  $\tau_\alpha = \tau_{\rho,m}$ . This also provides an intuition for choosing the specific penalty for  $\alpha$ .

**Remark A2 (PDHG convergence for discretized problems)** Note that it is not easy to apply the analysis above to the discretized saddle point problem in Section 2.2. The main difficulty is, after spatial discretization, the function  $K$  is no longer bilinear and smooth, which requires PDHG analysis for non-linear and non-smooth coupling term. Moreover, how to analyze the convergence of the proposed algorithm for more general  $f$  and  $L$  (for instance, when  $f$  is not affine with respect to  $\alpha$ ) requires more study.

## B Consistency of the numerical Hamiltonian

In Appendix A, we explore how the numerical algorithm converges to a saddle point, necessitating a subsequent examination of how this saddle point relates to the solution of the HJ PDE. The linkage between the saddle point and the HJ PDE solution becomes apparent through the first-order optimality conditions outlined in (7), when the stationary point  $\rho$  is positive. Furthermore, by referencing the optimality conditions specified for discretized scenarios in (12) and (17), we derive a discretized formulation for solving HJ PDEs, alongside the associated numerical Hamiltonians in (13) and (18). This section aims to validate the consistency of the numerical Hamiltonians, particularly in scenarios where the function  $f$  exhibits linearity in relation to  $\alpha$ , and under a set of additional assumptions.

### B.1 One-dimensional cases

In this section, we consider the one-dimensional cases ( $n = 1$ ). Assume  $f$  is linear with respect to  $\alpha$ , i.e.,  $f(x,t,\alpha) = a(x,t)^T \alpha$  where  $a$  is a continuous function from  $\Omega \times [0, T]$  to  $\mathbb{R}^m$ . Assume  $L_{x,t}: \mathbb{R}^m \rightarrow \mathbb{R}$  is non-negative, convex, and 1-coercive with  $L_{x,t}(0) = 0$  for any  $x \in \Omega$  and  $t \in [0, T]$ . In this case, we define the numerical Lagrangian by  $\hat{L}_{x,t}(\alpha_1, \alpha_2) = L_{x,t}(\alpha_1) + L_{x,t}(\alpha_2)$ . We are going to show the consistency of the corresponding numerical Hamiltonian. The corresponding numerical Hamiltonian  $\hat{H}$  is

$$\begin{aligned} \hat{H}(x,t,p^+,p^-) &= \sup_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \{ -(a(x,t)^T \alpha_1)_+ p^+ - (a(x,t)^T \alpha_2)_- p^- - L_{x,t}(\alpha_1) - L_{x,t}(\alpha_2) \} \\ &= \sup_{\alpha_1 \in \mathbb{R}^m} \{ -(a(x,t)^T \alpha_1)_+ p^+ - L_{x,t}(\alpha_1) \} + \sup_{\alpha_2 \in \mathbb{R}^m} \{ -(a(x,t)^T \alpha_2)_- p^- - L_{x,t}(\alpha_2) \}. \end{aligned}$$



To show the consistency of the numerical Hamiltonian, we need to show that  $\hat{H}(x, t, p, p) = H(x, t, p)$  for any  $x \in \Omega$ ,  $t \in [0, T]$  and  $p \in \mathbb{R}$  where  $H$  is the Hamiltonian defined in (3).

Let  $x \in \Omega$ ,  $t \in [0, T]$  and  $p \in \mathbb{R}$  be arbitrary numbers. To simplify notations, we ignore the  $x, t$  dependency in  $a$  and  $L$ , and denote  $a(x, t)$  by  $a$  and  $L_{x,t}$  by  $L$ . Define  $h_+$  by  $h_+ = \sup_{\alpha \in \mathbb{R}^m} \{-a^T \alpha + p - L(\alpha)\}$  and  $h_-$  by  $h_- = \sup_{\alpha \in \mathbb{R}^m} \{-a^T \alpha + p - L(\alpha)\}$ . Then, we need to show  $h_+ + h_- = H(x, t, p)$ . Since  $0 = L(0) = \min_{\alpha \in \mathbb{R}^m} L(\alpha)$ , we have

$$h_+ = \sup_{\alpha: a^T \alpha \geq 0} \{-pa^T \alpha - L(\alpha)\}, \quad h_- = \sup_{\alpha: a^T \alpha \leq 0} \{-pa^T \alpha - L(\alpha)\}. \quad (29)$$

Since  $L$  is convex and 1-coercive, then the optimizer in (3) exists. Denote the set of optimizers by  $\mathcal{A}$ , which is a closed convex set. Define  $\mathcal{A}_+$  and  $\mathcal{A}_-$  by  $\mathcal{A}_+ = \{\alpha \in \mathcal{A}: a^T \alpha \geq 0\}$  and  $\mathcal{A}_- = \{\alpha \in \mathcal{A}: a^T \alpha \leq 0\}$ .

If  $\mathcal{A}_+$  is non-empty, then we have  $h_+ = H(x, t, p)$ . If  $\mathcal{A}_+$  is empty, we show that  $h_+$  equals zero. Denote the objective function in the minimization problem in (3) by  $F$ , i.e.,  $F(\alpha) = -pa^T \alpha - L(\alpha)$ . Note that  $F$  is also the objective functions in the optimization problems in (29). Since  $L$  is convex and 1-coercive, the minimizer in the first optimization problem in (29) exists. If there is a minimizer  $\bar{\alpha}$  satisfying  $a^T \bar{\alpha} > 0$ , then there is a neighborhood of  $\bar{\alpha}$ , denoted by  $\mathcal{N}$ , such that  $\bar{\alpha}$  is the minimizer of the function  $F$  in  $\mathcal{N}$ . In other words,  $\bar{\alpha}$  is a local minimizer of  $F$ , and hence it is also a global minimizer by convexity. Therefore, we have  $\bar{\alpha} \in \mathcal{A}_+$ , which contradicts to the assumption that  $\mathcal{A}_+$  is empty. As a result, we have  $h_+ = \sup_{\alpha: a^T \alpha = 0} \{-pa^T \alpha - L(\alpha)\} = 0$  in the case when  $\mathcal{A}_+$  is empty. Then, we obtain

$$h_+ = \begin{cases} H(x, t, p) & \mathcal{A}_+ \neq \emptyset, \\ 0 & \mathcal{A}_+ = \emptyset. \end{cases}$$

With a similar argument, we can get

$$h_- = \begin{cases} H(x, t, p) & \mathcal{A}_- \neq \emptyset, \\ 0 & \mathcal{A}_- = \emptyset. \end{cases}$$

If either  $\mathcal{A}_+$  or  $\mathcal{A}_-$  is empty, then  $h_+ + h_- = H(x, t, p)$  follows directly from these two formulas. If both  $\mathcal{A}_+$  and  $\mathcal{A}_-$  are non-empty, by convexity of  $\mathcal{A}$ , there exists  $\bar{\alpha} \in \mathcal{A}$  satisfying  $a^T \bar{\alpha} = 0$ , which implies  $H(x, t, p) = \sup_{\alpha: a^T \alpha = 0} \{-pa^T \alpha - L(\alpha)\} = 0$ , and hence  $h_+ + h_- = 2H(x, t, p) = H(x, t, p)$ . Therefore, we conclude that  $\hat{H}$  is consistent.

**Remark B1** *The above argument can be applied to more general cases. Let  $f$  be in the form of  $f(x, t, \alpha) = a(x, t)^T \alpha + b(x, t)$ , where  $a$  is a continuous function from  $\Omega \times [0, T]$  to  $\mathbb{R}^m$  and  $b$  is a continuous function from  $\Omega \times [0, T]$  to  $\mathbb{R}$ . Let  $L$  be a continuous function such that  $L_{x,t}$  is convex and 1-coercive for any  $x \in \Omega$  and  $t \in [0, T]$ . Further assume that  $\arg \min_{\alpha \in \mathbb{R}^m} L_{x,t}(\alpha) \cap \{\alpha \in \mathbb{R}^m: f(x, t, \alpha) = 0\}$  is non-empty for any  $x \in \Omega$  and  $t \in [0, T]$ . Then, define the numerical Lagrangian  $\hat{L}$  by  $\hat{L}_{x,t}(\alpha_1, \alpha_2) = L_{x,t}(\alpha_1) + L_{x,t}(\alpha_2) - \min_{\alpha \in \mathbb{R}^m} L_{x,t}(\alpha)$ . With a similar argument as above, we can prove that the corresponding numerical Hamiltonian defined by (13) is consistent.*

**Remark B2** *We can further remove the requirement in Remark B1 that the minimal value of  $L_{x,t}$  can be achieved on a root of  $f_{x,t}$ . In other words, we only need to assume the continuity of  $f$  and  $L$ , and assume that  $f$  is affine on  $\alpha$ , and  $L_{x,t}$  is convex and 1-coercive for any  $x \in \Omega$  and  $t \in [0, T]$ . Note that the 1-coercivity assumption of  $L_{x,t}$  is for the existence of the optimizer in (3). It can be replaced by any other assumption which guarantees the existence of the minimizers. The assumption that  $f$  is affine on  $\alpha$  guarantees the convexity of the optimization problem in (3) so that we can use convex analysis techniques in the proof. Under this general setup, we need to modify the saddle point problem (11) to*

$$\min_{\varphi} \max_{\substack{\rho \geq 0, \alpha \\ f_{i,t}(\alpha_{1,i}(t)) \geq 0 \forall i,t \\ f_{i,t}(\alpha_{2,i}(t)) \leq 0 \forall i,t}} \int_0^T \sum_{i=1}^{n_x} \rho_i(t) \left( \dot{\varphi}_i(t) - f_{i,t}(\alpha_{1,i}(t))(D_x^+ \varphi)_i(t) - f_{i,t}(\alpha_{2,i}(t))(D_x^- \varphi)_i(t) \right. \\ \left. - \hat{L}_{i,t}(\alpha_{1,i}(t), \alpha_{2,i}(t)) \right) dt - c \sum_{i=1}^{n_x} \varphi_i(T).$$

Then, we define the numerical Lagrangian by

$$\hat{L}_{x,t}(\alpha_1, \alpha_2) = L_{x,t}(\alpha_1) + L_{x,t}(\alpha_2) - \min_{\substack{\alpha \in \mathbb{R}^m \\ f(x,t,\alpha)=0}} L_{x,t}(\alpha).$$

The consistency of the corresponding numerical Hamiltonian defined by

$$\hat{H}(x, t, p^+, p^-) = \sup_{\substack{\alpha_1, \alpha_2 \in \mathbb{R}^m \\ f_{x,t}(\alpha_1) \geq 0 \\ f_{x,t}(\alpha_2) \leq 0}} \{-f_{x,t}(\alpha_1)p^+ - f_{x,t}(\alpha_2)p^- - \hat{L}_{x,t}(\alpha_1, \alpha_2)\}$$

can be proved similarly as above using the formulas

$$\sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{x,t}(\alpha) \geq 0}} \{-f_{x,t}(\alpha)p - L_{x,t}(\alpha)\} = \begin{cases} H(x, t, p) & \mathcal{A}_{+,x,t} \neq \emptyset, \\ -\min_{\substack{\alpha \in \mathbb{R}^m \\ f(x,t,\alpha)=0}} L_{x,t}(\alpha) & \mathcal{A}_{+,x,t} = \emptyset, \end{cases}$$

$$\sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{x,t}(\alpha) \leq 0}} \{-f_{x,t}(\alpha)p - L_{x,t}(\alpha)\} = \begin{cases} H(x, t, p) & \mathcal{A}_{-,x,t} \neq \emptyset, \\ -\min_{\substack{\alpha \in \mathbb{R}^m \\ f(x,t,\alpha)=0}} L_{x,t}(\alpha) & \mathcal{A}_{-,x,t} = \emptyset, \end{cases}$$

where  $\mathcal{A}_{+,x,t}$  ( $\mathcal{A}_{-,x,t}$  resp.) are the set containing the minimizers  $\bar{\alpha}$  in (3) which satisfy  $f_{x,t}(\bar{\alpha}) \geq 0$  ( $f_{x,t}(\bar{\alpha}) \leq 0$  resp.). Then, the implicit time discretization and PDHG updates can also be applied to this modified saddle point problem.

## B.2 Two-dimensional cases

In this section, we consider the case when the spatial domain is two-dimensional ( $n = 2$ ).

We assume that  $f$  is a continuous function in the form  $f(x, t, \alpha) = \begin{pmatrix} a_1^T(x, t) \\ a_2^T(x, t) \end{pmatrix} \alpha$ , where  $a_1$  and  $a_2$  are continuous functions from  $\Omega \times [0, T]$  to  $\mathbb{R}^m$ . Let  $L$  be a continuous non-negative function. Assume  $L_{x,t}: \mathbb{R}^m \rightarrow \mathbb{R}$  is convex and 1-coercive with  $L_{x,t}(0) = 0$  for any  $x \in \Omega$  and  $t \in [0, T]$ . Let  $H$  be the Hamiltonian defined by (3). Assume that  $H$  is in the form of  $H_{x,t}(p_x, p_y) = H_1(x, t, p_x) + H_2(x, t, p_y)$  for any  $x \in \Omega$ ,  $t \in [0, T]$ ,  $p_x, p_y \in \mathbb{R}$  for some functions  $H_1$  and  $H_2$ . In this case, we choose  $\hat{L}(x, t, \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) = L(x, t, \alpha_{11}) + L(x, t, \alpha_{12}) + L(x, t, \alpha_{21}) + L(x, t, \alpha_{22})$ . Now, we want to show that such defined numerical Lagrangian gives a consistent numerical Hamiltonian.

Let  $x \in \Omega$ ,  $t \in [0, T]$ , and  $p_x, p_y \in \mathbb{R}$  be arbitrary vectors and numbers. For simplicity of the notations, we ignore the  $(x, t)$  dependence in the functions  $a_1$ ,  $a_2$ , and  $L$ . Define  $h_{x,+}$ ,  $h_{x,-}$ ,  $h_{y,+}$ , and  $h_{y,-}$  by

$$h_{x,+} = \sup_{\alpha: a_1^T \alpha \geq 0} \{-p_x a_1^T \alpha - L(\alpha)\}, \quad h_{x,-} = \sup_{\alpha: a_1^T \alpha \leq 0} \{-p_x a_1^T \alpha - L(\alpha)\},$$

$$h_{y,+} = \sup_{\alpha: a_2^T \alpha \geq 0} \{-p_y a_2^T \alpha - L(\alpha)\}, \quad h_{y,-} = \sup_{\alpha: a_2^T \alpha \leq 0} \{-p_y a_2^T \alpha - L(\alpha)\}.$$

Since  $L_{x,t}$  is a non-negative function with  $L_{x,t}(0) = 0$ , the numerical Hamiltonian defined in (18) satisfies  $\hat{H}(x, t, p_x, p_x, p_y, p_y) = h_{x,+} + h_{x,-} + h_{y,+} + h_{y,-}$ . Then, it suffices to prove that  $h_{x,+} + h_{x,-} + h_{y,+} + h_{y,-} = H(x, t, p_x, p_y)$ . Denote the set of optimizers in (3) by  $\mathcal{A}(p)$ , which, according to the assumptions on  $f$  and  $L$ , is non-empty, closed, and convex for

any  $p \in \mathbb{R}^2$ . With a similar argument as in the one-dimensional case, we have

$$\begin{aligned} h_{x,+} &= \begin{cases} H(x, t, p_x, 0) & \mathcal{A}(p_x, 0) \cap \{\alpha \in \mathbb{R}^m : a_1^T \alpha \geq 0\} \neq \emptyset, \\ 0 & \mathcal{A}(p_x, 0) \cap \{\alpha \in \mathbb{R}^m : a_1^T \alpha \geq 0\} = \emptyset, \end{cases} \\ h_{x,-} &= \begin{cases} H(x, t, p_x, 0) & \mathcal{A}(p_x, 0) \cap \{\alpha \in \mathbb{R}^m : a_1^T \alpha \leq 0\} \neq \emptyset, \\ 0 & \mathcal{A}(p_x, 0) \cap \{\alpha \in \mathbb{R}^m : a_1^T \alpha \leq 0\} = \emptyset, \end{cases} \\ h_{y,+} &= \begin{cases} H(x, t, 0, p_y) & \mathcal{A}(0, p_y) \cap \{\alpha \in \mathbb{R}^m : a_2^T \alpha \geq 0\} \neq \emptyset, \\ 0 & \mathcal{A}(0, p_y) \cap \{\alpha \in \mathbb{R}^m : a_2^T \alpha \geq 0\} = \emptyset, \end{cases} \\ h_{y,-} &= \begin{cases} H(x, t, 0, p_y) & \mathcal{A}(0, p_y) \cap \{\alpha \in \mathbb{R}^m : a_2^T \alpha \leq 0\} \neq \emptyset, \\ 0 & \mathcal{A}(0, p_y) \cap \{\alpha \in \mathbb{R}^m : a_2^T \alpha \leq 0\} = \emptyset. \end{cases} \end{aligned}$$

If either  $\mathcal{A}(p_x, 0) \cap \{\alpha \in \mathbb{R}^m : a_1^T \alpha \geq 0\}$  or  $\mathcal{A}(p_x, 0) \cap \{\alpha \in \mathbb{R}^m : a_1^T \alpha \leq 0\}$  is empty, then we have  $h_{x,+} + h_{x,-} = H(x, t, p_x, 0)$ . Otherwise, similarly as in the one-dimensional case, there exists  $\bar{\alpha} \in \mathcal{A}(p_x, 0)$  satisfying  $a_1^T \bar{\alpha} = 0$ , which implies  $H(x, t, p_x, 0) = 0$ , and hence  $h_{x,+} + h_{x,-} = 2H(x, t, p_x, 0) = H(x, t, p_x, 0)$  holds. Similarly, we have  $h_{y,+} + h_{y,-} = H(x, t, 0, p_y)$ . Therefore, we get

$$\begin{aligned} h_{x,+} + h_{x,-} + h_{y,+} + h_{y,-} &= H(x, t, p_x, 0) + H(x, t, 0, p_y) \\ &= H_1(x, t, p_x) + H_2(x, t, 0) + H_1(x, t, 0) + H_2(x, t, p_y) \\ &= H(x, t, p_x, p_y) + H(x, t, 0, 0) \\ &= H(x, t, p_x, p_y) - \min_{\alpha \in \mathbb{R}^m} L_{x,t}(\alpha) = H(x, t, p_x, p_y), \end{aligned}$$

which ends the proof.

**Remark B3** Similarly as in the one-dimensional case, the assumptions can be further relaxed. Assume that  $f$  and  $L$  are both continuous,  $f_{x,t}$  is an affine function, and  $L_{x,t}$  is convex and 1-coercive for any  $x \in \Omega$  and  $t \in [0, T]$ . Note that the 1-coercivity assumption of  $L_{x,t}$  is for the existence of the optimizer in (3), and the assumption that  $f_{x,t}$  is affine makes the optimization problem in (3) a convex problem. Under this general setup, we modify the saddle point problem (16) to

$$\begin{aligned} \min_{\varphi_{i,j}(0)=g(x_{i,j})} \max_{\substack{\rho \geq 0, \alpha \\ f_{1,i,j,t}(\alpha_{11,i,j}(t)) \geq 0 \forall i,j,t \\ f_{1,i,j,t}(\alpha_{12,i,j}(t)) \leq 0 \forall i,j,t \\ f_{2,i,j,t}(\alpha_{21,i,j}(t)) \geq 0 \forall i,j,t \\ f_{2,i,j,t}(\alpha_{22,i,j}(t)) \leq 0 \forall i,j,t}} \int_0^T \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \rho_{i,j}(t) \left( \varphi_{i,j}(t) - f_{1,i,j,t}(\alpha_{11,i,j}(t))(D_x^+ \varphi)_{i,t}(t) - f_{1,i,j,t}(\alpha_{12,i,j}(t))(D_x^- \varphi)_{i,t}(t) \right. \\ \left. - f_{2,i,j,t}(\alpha_{21,i,j}(t))(D_y^+ \varphi)_{i,t}(t) - f_{2,i,j,t}(\alpha_{22,i,j}(t))(D_y^- \varphi)_{i,t}(t) \right. \\ \left. - \hat{L}_{i,j,t}(\alpha_{11,i,j}(t), \alpha_{12,i,j}(t), \alpha_{21,i,j}(t), \alpha_{22,i,j}(t)) \right) dt - c \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \varphi_{i,j}(T). \end{aligned}$$

We define the numerical Lagrangian by

$$\begin{aligned} \hat{L}_{x,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) &= L_{x,t}(\alpha_{11}) + L_{x,t}(\alpha_{12}) + L_{x,t}(\alpha_{21}) + L_{x,t}(\alpha_{22}) - \min_{\alpha \in \mathbb{R}^m} L_{x,t}(\alpha) \\ &= \min_{\substack{\alpha \in \mathbb{R}^m \\ f_1(x,t,\alpha)=0}} L_{x,t}(\alpha) - \min_{\substack{\alpha \in \mathbb{R}^m \\ f_2(x,t,\alpha)=0}} L_{x,t}(\alpha). \end{aligned}$$

The numerical Hamiltonian is then given by

$$\begin{aligned} \hat{H}(x, t, p_x^+, p_x^-, p_y^+, p_y^-) &= \sup_{\substack{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m \\ f_{1,x,t}(\alpha_{11}) \geq 0 \\ f_{1,x,t}(\alpha_{12}) \leq 0 \\ f_{2,x,t}(\alpha_{21}) \geq 0 \\ f_{2,x,t}(\alpha_{22}) \leq 0}} \{-f_{1,x,t}(\alpha_{11})p_x^+ - f_{1,x,t}(\alpha_{12})p_x^- - f_{2,x,t}(\alpha_{21})p_y^+ \\ &\quad - f_{2,x,t}(\alpha_{22})p_y^- - \hat{L}_{x,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})\}. \end{aligned}$$

The consistency of this numerical Hamiltonian follows from a similar argument. Then, the implicit time discretization and PDHG updates can be similarly applied to this modified saddle point problem.

**Remark B4** Note that the assumption in the two-dimensional cases is more restricted than the one-dimensional cases. To be more specific, we assume that there exists  $H_1$  and  $H_2$  such that  $H$  is in the form of  $H_{x,t}(p_x, p_y) = H_1(x, t, p_x) + H_2(x, t, p_y)$  for any  $x \in \Omega$ ,  $t \in [0, T]$ ,  $p_x, p_y \in \mathbb{R}$ . If we consider the numerical Lagrangian  $\hat{L}$  which can be written as  $\hat{L}_{x,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) = L_{11,x,t}(\alpha_{11}) + L_{12,x,t}(\alpha_{12}) + L_{21,x,t}(\alpha_{21}) + L_{22,x,t}(\alpha_{22})$  for some functions  $L_{11}$ ,  $L_{12}$ ,  $L_{21}$ , and  $L_{22}$ , then the corresponding numerical Hamiltonian is

$$\begin{aligned} \hat{H}(x, t, p_x^+, p_x^-, p_y^+, p_y^-) &= \sup_{\substack{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m \\ f_{1,x,t}(\alpha_{11}) \geq 0 \\ f_{1,x,t}(\alpha_{12}) \leq 0 \\ f_{2,x,t}(\alpha_{21}) \geq 0 \\ f_{2,x,t}(\alpha_{22}) \leq 0}} \{-f_{1,x,t}(\alpha_{11})p_x^+ - f_{1,x,t}(\alpha_{12})p_x^- - f_{2,x,t}(\alpha_{21})p_y^+ - f_{2,x,t}(\alpha_{22})p_y^- \\ &\quad - (L_{11,x,t}(\alpha_{11}) + L_{12,x,t}(\alpha_{12}) + L_{21,x,t}(\alpha_{21}) + L_{22,x,t}(\alpha_{22}))\} \\ &= \sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{1,x,t}(\alpha) \geq 0}} \{-f_{1,x,t}(\alpha)p_x^+ - L_{11,x,t}(\alpha)\} + \sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{1,x,t}(\alpha) \leq 0}} \{-f_{1,x,t}(\alpha)p_x^- - L_{12,x,t}(\alpha)\} \\ &\quad + \sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{2,x,t}(\alpha) \geq 0}} \{-f_{2,x,t}(\alpha)p_y^+ - L_{21,x,t}(\alpha)\} + \sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{2,x,t}(\alpha) \leq 0}} \{-f_{2,x,t}(\alpha)p_y^- - L_{22,x,t}(\alpha)\}. \end{aligned}$$

Then, if  $\hat{H}$  is consistent,  $H$  must be in the form of  $H_{x,t}(p_x, p_y) = H_1(x, t, p_x) + H_2(x, t, p_y)$  with

$$\begin{aligned} H_1(x, t, p_x) &= \sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{1,x,t}(\alpha) \geq 0}} \{-f_{1,x,t}(\alpha)p_x - L_{11,x,t}(\alpha)\} + \sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{1,x,t}(\alpha) \leq 0}} \{-f_{1,x,t}(\alpha)p_x - L_{12,x,t}(\alpha)\}, \\ H_2(x, t, p_y) &= \sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{2,x,t}(\alpha) \geq 0}} \{-f_{2,x,t}(\alpha)p_y - L_{21,x,t}(\alpha)\} + \sup_{\substack{\alpha \in \mathbb{R}^m \\ f_{2,x,t}(\alpha) \leq 0}} \{-f_{2,x,t}(\alpha)p_y - L_{22,x,t}(\alpha)\}. \end{aligned}$$

For more general  $H$ , we need to design the numerical Lagrangian  $\hat{L}$  in other forms. We also considered  $\hat{L}_{x,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) = L_{x,t}(\alpha_{11} + \alpha_{12} + \alpha_{21} + \alpha_{22})$ , but it has too many degrees of freedom on  $\alpha$ . To illustrate this intuition, we consider the case when we have  $m = n = 2$  and  $f(x, t, \alpha) = \alpha$ . Denote  $\alpha_{IJ} = (\alpha_{IJ,1}, \alpha_{IJ,2}) \in \mathbb{R}^2$  for any  $I, J = 1, 2$ . Then, the numerical Hamiltonian is

$$\hat{H}(x, t, p_x, p_x, p_y, p_y) = \sup_{\substack{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^2 \\ \alpha_{11,1} \geq 0 \\ \alpha_{12,1} \leq 0 \\ \alpha_{21,2} \geq 0 \\ \alpha_{22,2} \leq 0}} \{-\alpha_{11,1}p_x - \alpha_{12,1}p_x - \alpha_{21,2}p_y - \alpha_{22,2}p_y - L_{x,t}(\alpha_{11} + \alpha_{12} + \alpha_{21} + \alpha_{22})\}.$$

For any  $\alpha_{11,1}, \alpha_{21,2} \geq 0$  and  $\alpha_{12,1}, \alpha_{22,2} \leq 0$ , we can always choose  $\alpha_{11,2}, \alpha_{12,2}, \alpha_{21,1}, \alpha_{22,1}$  such that  $L_{x,t}(\alpha_{11} + \alpha_{12} + \alpha_{21} + \alpha_{22})$  achieves the minimal value of  $L_{x,t}$ . Therefore, the numerical Hamiltonian  $\hat{H}$  is infinity if  $p_x$  is non-zero or  $p_y$  is non-zero, which is not consistent. How to choose a numerical Lagrangian for more general cases is a possible interesting future direction.

## C Discretization for the viscous HJ PDEs

In this section, we provide details of the spatial and temporal discretization of the saddle point problem (25) for solving stochastic optimal control problems and the corresponding viscous HJ PDEs. We apply similar notation as in Section 2.2. Note that the difference between this part with Section 2.2 is merely on the diffusion term.

### C.1 One-dimensional problems

After upwind spatial discretization, the one-dimensional saddle point problem (divided by  $\Delta x$ ) becomes

$$\min_{\varphi} \max_{\rho \geq 0, \alpha} \int_0^T \sum_{i=1}^{n_x} \rho_i(t) \left( \dot{\varphi}_i(t) - f_{i,t}(\alpha_{1,i}(t))_+ (D_x^+ \varphi)_i(t) - f_{i,t}(\alpha_{2,i}(t))_- (D_x^- \varphi)_i(t) - \hat{L}_{i,t}(\alpha_{1,i}(t), \alpha_{2,i}(t)) - \epsilon(D_{xx} \varphi)_i(t) \right) dt - c \sum_{i=1}^{n_x} \varphi_i(T).$$

Consider a stationary point  $(\varphi, \rho, \alpha)$  in this saddle point problem. If we further assume  $\rho_i(t) > 0$  for any  $t \in [0, T]$ , then the first order optimality condition is

$$\begin{cases} \dot{\varphi}_i(t) - f_{i,t}(\alpha_{1,i}(t))_+ (D_x^+ \varphi)_i(t) - f_{i,t}(\alpha_{2,i}(t))_- (D_x^- \varphi)_i(t) - \hat{L}_{i,t}(\alpha_{1,i}(t), \alpha_{2,i}(t)) - \epsilon(D_{xx} \varphi)_i(t) = 0, \\ (\alpha_{1,i}(t), \alpha_{2,i}(t)) = \arg \min_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \{f_{i,t}(\alpha_1)_+ (D_x^+ \varphi)_i(t) + f_{i,t}(\alpha_2)_- (D_x^- \varphi)_i(t) + \hat{L}_{i,t}(\alpha_1, \alpha_2)\}, \\ \dot{\rho}_i(t) - D_x^- (f_{i,t}(\alpha_{1,i}(t))_+ \rho_i(t)) - D_x^+ (f_{i,t}(\alpha_{2,i}(t))_- \rho_i(t)) + \epsilon(D_{xx} \rho)_i(t) = 0. \end{cases}$$

Combining the first two lines in this optimality condition, we get

$$\dot{\varphi}_i(t) + \sup_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \{-f_{i,t}(\alpha_1)_+ (D_x^+ \varphi)_i(t) - f_{i,t}(\alpha_2)_- (D_x^- \varphi)_i(t) - \hat{L}_{i,t}(\alpha_1, \alpha_2)\} - \epsilon(D_{xx} \varphi)_i(t) = 0,$$

which gives a semi-discrete scheme for the HJ PDE (22) where the numerical Hamiltonian is defined by (13).

With this discretization, the  $\ell$ -th update becomes

$$\begin{cases} \rho_i^{\ell+1}(t) = \left( \rho_i^\ell(t) + \tau_\rho (\dot{\varphi}_i^\ell(t) - f_{i,t}(\alpha_{1,i}^\ell(t))_+ (D_x^+ \varphi^\ell)_i(t) - f_{i,t}(\alpha_{2,i}^\ell(t))_- (D_x^- \varphi^\ell)_i(t) - \hat{L}_{i,t}(\alpha_{1,i}^\ell(t), \alpha_{2,i}^\ell(t)) - \epsilon(D_{xx} \varphi^\ell)_i(t) \right)_+, \\ (\alpha_{1,i}^{\ell+1}(t), \alpha_{2,i}^{\ell+1}(t)) = \arg \min_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \left\{ f_{i,t}(\alpha_1)_+ (D_x^+ \varphi^\ell)_i(t) + f_{i,t}(\alpha_2)_- (D_x^- \varphi^\ell)_i(t) + \hat{L}_{i,t}(\alpha_1, \alpha_2) + \frac{\rho_i^{\ell+1}(t)}{2\tau_\alpha} (\|\alpha_1 - \alpha_{1,i}^\ell(t)\|^2 + \|\alpha_2 - \alpha_{2,i}^\ell(t)\|^2) \right\}, \\ \varphi_i^{\ell+1}(t) = \varphi_i^\ell(t) + \tau_\varphi (I - \partial_{tt} - D_{xx})^{-1} \left( \rho_i^{\ell+1}(t) - D_x^- (f_{i,t}(\alpha_{1,i}^{\ell+1}(t))_+ \rho_i^{\ell+1}(t)) - D_x^+ (f_{i,t}(\alpha_{2,i}^{\ell+1}(t))_- \rho_i^{\ell+1}(t)) + \epsilon(D_{xx} \rho^{\ell+1})_i(t) \right), \\ \dot{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^\ell. \end{cases}$$

For the time discretization, we apply implicit Euler scheme for  $\varphi$ , and then the saddle point problem (divided by  $\Delta t$ ) becomes

$$\min_{\varphi} \max_{\rho \geq 0, \alpha} \sum_{k=2}^{n_t} \sum_{i=1}^{n_x} \rho_{i,k} \left( (D_t^- \varphi)_{i,k} - f_{i,k}(\alpha_{1,i,k})_+ (D_x^+ \varphi)_{i,k} - f_{i,k}(\alpha_{2,i,k})_- (D_x^- \varphi)_{i,k} - \hat{L}_{i,k}(\alpha_{1,i,k}, \alpha_{2,i,k}) - \epsilon(D_{xx} \varphi)_{i,k} \right) - \frac{c}{\Delta t} \sum_{i=1}^{n_x} \varphi_{i,n_t}.$$

The corresponding algorithm in the  $\ell$ -th iteration becomes

$$\left\{ \begin{array}{l} \rho_{i,k}^{\ell+1} = \left( \rho_{i,k}^{\ell} + \tau_{\rho} \left( (D_t^- \tilde{\varphi}^{\ell})_{i,k} - f_{i,k}(\alpha_{1,i,k}^{\ell}) + (D_x^+ \tilde{\varphi}^{\ell})_{i,k} - f_{i,k}(\alpha_{2,i,k}^{\ell}) - (D_x^- \tilde{\varphi}^{\ell})_{i,k} \right. \right. \\ \quad \left. \left. - \hat{L}_{i,k}(\alpha_{1,i,k}^{\ell}, \alpha_{2,i,k}^{\ell}) - \epsilon(D_{xx} \tilde{\varphi}^{\ell})_{i,k} \right) \right)_+ \\ (\alpha_{1,i,k}^{\ell+1}, \alpha_{2,i,k}^{\ell+1}) = \arg \min_{\alpha_1, \alpha_2 \in \mathbb{R}^m} \left\{ f_{i,k}(\alpha_1) + (D_x^+ \tilde{\varphi}^{\ell})_{i,k} + f_{i,k}(\alpha_2) - (D_x^- \tilde{\varphi}^{\ell})_{i,k} + \hat{L}_{i,k}(\alpha_1, \alpha_2) \right. \\ \quad \left. + \frac{\rho_{i,k}^{\ell+1}}{2\tau_{\alpha}} \left( \|\alpha_1 - \alpha_{1,i,k}^{\ell}\|^2 + \|\alpha_2 - \alpha_{2,i,k}^{\ell}\|^2 \right) \right\}. \\ \varphi_{i,k}^{\ell+1} = \varphi_{i,k}^{\ell} + \tau_{\varphi} (I - D_{tt} - D_{xx})^{-1} \left( (D_t^+ \rho^{\ell+1})_{i,k} - D_x^- (f_{i,k}(\alpha_{1,i,k}^{\ell+1}) + \rho_{i,k}^{\ell+1}) \right. \\ \quad \left. - D_x^+ (f_{i,k}(\alpha_{2,i,k}^{\ell+1}) - \rho_{i,k}^{\ell+1}) + \epsilon(D_{xx} \rho^{\ell+1})_{i,k} \right). \\ \tilde{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^{\ell}. \end{array} \right.$$

It's worth mentioning that selecting the numerical Lagrangian as  $L_{x,t}(\alpha_1) + L_{x,t}(\alpha_2)$  allows for the update process of  $\alpha$  to be executed in parallel across the variables  $\alpha_1$  and  $\alpha_2$ .

## C.2 Two-dimensional problems

For two dimensional cases, with upwind spatial discretization, the saddle point problem (divided by  $\Delta x \Delta y$ ) becomes

$$\min_{\varphi_{i,j}(0)=g(x_{i,j})} \max_{\rho \geq 0, \alpha} \int_0^T \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \rho_{i,j}(t) \left( \dot{\varphi}_{i,j}(t) - f_{1,i,j,t}(\alpha_{11,i,j}(t)) + (D_x^+ \varphi)_{i,j}(t) - f_{1,i,j,t}(\alpha_{12,i,j}(t)) - (D_x^- \varphi)_{i,j}(t) \right. \\ \quad - f_{2,i,j,t}(\alpha_{21,i,j}(t)) + (D_y^+ \varphi)_{i,j}(t) - f_{2,i,j,t}(\alpha_{22,i,j}(t)) - (D_y^- \varphi)_{i,j}(t) - \epsilon(D_{xx} \varphi + D_{yy} \varphi)_{i,j}(t) \\ \quad \left. - \hat{L}_{i,j,t}(\alpha_{11,i,j}(t), \alpha_{12,i,j}(t), \alpha_{21,i,j}(t), \alpha_{22,i,j}(t)) \right) dt - c \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \varphi_{i,j}(T).$$

Consider a stationary point  $(\varphi, \rho, \alpha)$  in this saddle point problem. If we further assume  $\rho_{i,j}(t) > 0$  for any  $t \in [0, T]$ , then the first order optimality condition is

$$\left\{ \begin{array}{l} \dot{\varphi}_{i,j}(t) - f_{1,i,j,t}(\alpha_{11,i,j}(t)) + (D_x^+ \varphi)_{i,j}(t) - f_{1,i,j,t}(\alpha_{12,i,j}(t)) - (D_x^- \varphi)_{i,j}(t) \\ \quad - f_{2,i,j,t}(\alpha_{21,i,j}(t)) + (D_y^+ \varphi)_{i,j}(t) - f_{2,i,j,t}(\alpha_{22,i,j}(t)) - (D_y^- \varphi)_{i,j}(t) \\ \quad - \hat{L}_{i,j,t}(\alpha_{11,i,j}(t), \alpha_{12,i,j}(t), \alpha_{21,i,j}(t), \alpha_{22,i,j}(t)) - \epsilon(D_{xx} \varphi + D_{yy} \varphi)_{i,j}(t) = 0, \\ (\alpha_{11,i,j}(t), \alpha_{12,i,j}(t), \alpha_{21,i,j}(t), \alpha_{22,i,j}(t)) = \arg \min_{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m} \{ f_{1,i,j,t}(\alpha_{11}) + (D_x^+ \varphi)_{i,j}(t) \\ \quad + f_{1,i,j,t}(\alpha_{12}) - (D_x^- \varphi)_{i,j}(t) + f_{2,i,j,t}(\alpha_{21}) + (D_y^+ \varphi)_{i,j}(t) \\ \quad + f_{2,i,j,t}(\alpha_{22}) - (D_y^- \varphi)_{i,j}(t) + \hat{L}_{i,j,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) \}, \\ \dot{\rho}_{i,j}(t) - D_x^- (f_{1,i,j,t}(\alpha_{11,i,j}(t)) + \rho_{i,j}(t)) - D_x^+ (f_{1,i,j,t}(\alpha_{12,i,j}(t)) - \rho_{i,j}(t)) \\ \quad - D_y^- (f_{2,i,j,t}(\alpha_{21,i,j}(t)) + \rho_{i,j}(t)) - D_y^+ (f_{2,i,j,t}(\alpha_{22,i,j}(t)) - \rho_{i,j}(t)) + \epsilon(D_{xx} \rho + D_{yy} \rho)_{i,j}(t) = 0. \end{array} \right.$$

Combining the first two lines in this optimality condition, we get

$$\dot{\varphi}_{i,j}(t) + \sup_{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m} \{ -f_{1,i,j,t}(\alpha_{11}) + (D_x^+ \varphi)_{i,j}(t) - f_{1,i,j,t}(\alpha_{12}) - (D_x^- \varphi)_{i,j}(t) - f_{2,i,j,t}(\alpha_{21}) + (D_y^+ \varphi)_{i,j}(t) \\ \quad - f_{2,i,j,t}(\alpha_{22}) - (D_y^- \varphi)_{i,j}(t) - \hat{L}_{i,j,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) \} - \epsilon(D_{xx} \varphi + D_{yy} \varphi)_{i,j}(t) = 0,$$

which gives a semi-discrete scheme for the HJ PDE (22) whose numerical Hamiltonian is defined in (18).

With this discretization, the  $\ell$ -th update becomes

$$\left\{ \begin{array}{l} \rho_{i,j}^{\ell+1}(t) = \left( \rho_{i,j}^{\ell}(t) + \tau_{\rho} (\dot{\varphi}_{i,j}^{\ell}(t) - f_{1,i,j,t}(\alpha_{11,i,j}^{\ell}(t)) + (D_x^+ \varphi^{\ell})_{i,j}(t) \right. \\ \quad - f_{1,i,j,t}(\alpha_{12,i,j}^{\ell}(t)) - (D_x^- \varphi^{\ell})_{i,j}(t) - f_{2,i,j,t}(\alpha_{21,i,j}^{\ell}(t)) + (D_y^+ \varphi^{\ell})_{i,j}(t) \\ \quad \left. - f_{2,i,j,t}(\alpha_{22,i,j}^{\ell}(t)) - (D_y^- \varphi^{\ell})_{i,j}(t) - \hat{L}_{i,j,t}(\alpha_{1,i}^{\ell}(t), \alpha_{2,i}^{\ell}(t)) - \epsilon(D_{xx}\varphi^{\ell} + D_{yy}\varphi^{\ell})_{i,j}(t) \right)_+ \\ (\alpha_{11,i,j}^{\ell+1}(t), \alpha_{12,i,j}^{\ell+1}(t), \alpha_{21,i,j}^{\ell+1}(t), \alpha_{22,i,j}^{\ell+1}(t)) = \arg \min_{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m} \{ f_{1,i,j,t}(\alpha_{11}) + (D_x^+ \varphi^{\ell})_{i,j}(t) \\ \quad + f_{1,i,j,t}(\alpha_{12}) - (D_x^- \varphi^{\ell})_{i,j}(t) + f_{2,i,j,t}(\alpha_{21}) + (D_y^+ \varphi^{\ell})_{i,j}(t) \\ \quad + f_{2,i,j,t}(\alpha_{22}) - (D_y^- \varphi^{\ell})_{i,j}(t) + \hat{L}_{i,j,t}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) \\ \quad + \frac{\rho_{i,j}^{\ell+1}(t)}{2\tau_{\alpha}} (\|\alpha_{11} - \alpha_{11,i,j}^{\ell}(t)\|^2 + \|\alpha_{12} - \alpha_{12,i,j}^{\ell}(t)\|^2 + \|\alpha_{21} - \alpha_{21,i,j}^{\ell}(t)\|^2 + \|\alpha_{22} - \alpha_{22,i,j}^{\ell}(t)\|^2) \}. \\ \varphi_{i,j}^{\ell+1}(t) = \varphi_{i,j}^{\ell}(t) + \tau_{\varphi} (I - \partial_{tt} - D_{xx} - D_{yy})^{-1} (\dot{\rho}_{i,j}^{\ell+1}(t) - D_x^- (f_{1,i,j,t}(\alpha_{11,i,j}^{k+1}(t)) + \rho_{i,j}^{\ell+1}(t)) \\ \quad - D_x^+ (f_{1,i,j,t}(\alpha_{12,i,j}^{k+1}(t)) - \rho_{i,j}^{\ell+1}(t)) - D_y^- (f_{2,i,j,t}(\alpha_{21,i,j}^{k+1}(t)) + \rho_{i,j}^{\ell+1}(t)) \\ \quad - D_y^+ (f_{2,i,j,t}(\alpha_{22,i,j}^{k+1}(t)) - \rho_{i,j}^{\ell+1}(t)) + \epsilon(D_{xx}\rho^{\ell+1} + D_{yy}\rho^{\ell+1})_{i,j}(t)). \\ \tilde{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^{\ell}. \end{array} \right.$$

Then, we apply implicit Euler scheme for the time derivative of  $\varphi$ , and then the saddle point problem (divided by  $\Delta t$ ) becomes

$$\min_{\varphi} \max_{\rho \geq 0, \alpha} \sum_{k=2}^{n_x} \sum_{j=1}^{n_y} \rho_{i,j,k} \left( (D_t^- \varphi)_{i,j,k} - f_{1,i,j,k}(\alpha_{11,i,j,k}) + (D_x^+ \varphi)_{i,j,k} - f_{1,i,j,k}(\alpha_{12,i,j,k}) - (D_x^- \varphi)_{i,j,k} \right. \\ \quad \left. - f_{2,i,j,k}(\alpha_{21,i,j,k}) + (D_y^+ \varphi)_{i,j,k} - f_{2,i,j,k}(\alpha_{22,i,j,k}) - (D_y^- \varphi)_{i,j,k} - \epsilon(D_{xx}\varphi + D_{yy}\varphi)_{i,j,k} \right. \\ \quad \left. - \hat{L}_{i,j,k}(\alpha_{11,i,j,k}, \alpha_{12,i,j,k}, \alpha_{21,i,j,k}, \alpha_{22,i,j,k}) \right) - \frac{c}{\Delta t} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \varphi_{i,j,n_t}.$$

The corresponding algorithm in the  $\ell$ -th iteration becomes

$$\left\{ \begin{array}{l} \rho_{i,j,k}^{\ell+1} = \left( \rho_{i,j,k}^{\ell} + \tau_{\rho} ((D_t^- \varphi^{\ell})_{i,j,k} - f_{1,i,j,k}(\alpha_{11,i,j,k}) + (D_x^+ \varphi^{\ell})_{i,j,k} - f_{1,i,j,k}(\alpha_{12,i,j,k}) - (D_x^- \varphi^{\ell})_{i,j,k} \right. \\ \quad - f_{2,i,j,k}(\alpha_{21,i,j,k}) + (D_y^+ \varphi^{\ell})_{i,j,k} - f_{2,i,j,k}(\alpha_{22,i,j,k}) - (D_y^- \varphi^{\ell})_{i,j,k} \\ \quad \left. - \hat{L}_{i,j,k}(\alpha_{11,i,j,k}^{\ell}, \alpha_{12,i,j,k}^{\ell}, \alpha_{21,i,j,k}^{\ell}, \alpha_{22,i,j,k}^{\ell}) - \epsilon(D_{xx}\varphi^{\ell} + D_{yy}\varphi^{\ell})_{i,j,k} \right)_+ \\ (\alpha_{11,i,j,k}^{\ell+1}, \alpha_{12,i,j,k}^{\ell+1}, \alpha_{21,i,j,k}^{\ell+1}, \alpha_{22,i,j,k}^{\ell+1}) = \arg \min_{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{R}^m} \{ f_{1,i,j,k}(\alpha_{11}) + (D_x^+ \varphi^{\ell})_{i,j,k} \\ \quad + f_{1,i,j,k}(\alpha_{12}) - (D_x^- \varphi^{\ell})_{i,j,k} + f_{2,i,j,k}(\alpha_{21}) + (D_y^+ \varphi^{\ell})_{i,j,k} \\ \quad + f_{2,i,j,k}(\alpha_{22}) - (D_y^- \varphi^{\ell})_{i,j,k} + \hat{L}_{i,j,k}(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) \\ \quad + \frac{\rho_{i,j,k}^{\ell+1}}{2\tau_{\alpha}} (\|\alpha_{11} - \alpha_{11,i,j,k}^{\ell}(t)\|^2 + \|\alpha_{12} - \alpha_{12,i,j,k}^{\ell}(t)\|^2 + \|\alpha_{21} - \alpha_{21,i,j,k}^{\ell}(t)\|^2 + \|\alpha_{22} - \alpha_{22,i,j,k}^{\ell}(t)\|^2) \}. \\ \varphi_{i,j,k}^{\ell+1} = \varphi_{i,j,k}^{\ell} + \tau_{\varphi} (I - D_{tt} - D_{xx} - D_{yy})^{-1} ((D_t^+ \rho^{\ell+1})_{i,j,k} - D_x^- (f_{1,i,j,k}(\alpha_{11,i,j,k}^{k+1}) + \rho_{i,j,k}^{\ell+1}) \\ \quad - D_x^+ (f_{1,i,j,k}(\alpha_{12,i,j,k}^{k+1}) - \rho_{i,j,k}^{\ell+1}) - D_y^- (f_{2,i,j,k}(\alpha_{21,i,j,k}^{k+1}) + \rho_{i,j,k}^{\ell+1}) \\ \quad - D_y^+ (f_{2,i,j,k}(\alpha_{22,i,j,k}^{k+1}) - \rho_{i,j,k}^{\ell+1}) + \epsilon(D_{xx}\rho^{\ell+1} + D_{yy}\rho^{\ell+1})_{i,j,k}). \\ \tilde{\varphi}^{\ell+1} = 2\varphi^{\ell+1} - \varphi^{\ell}. \end{array} \right.$$

Just as with the one-dimensional scenario, by selecting the numerical Lagrangian as  $L_{x,t}(\alpha_{11}) + L_{x,t}(\alpha_{12}) + L_{x,t}(\alpha_{21}) + L_{x,t}(\alpha_{22})$ , the update process for  $\alpha$  can be carried out in parallel across the variables  $\alpha_{11}$ ,  $\alpha_{12}$ ,  $\alpha_{21}$ , and  $\alpha_{22}$ .