Is L^2 Physics-Informed Loss Always Suitable for Training Physics-Informed Neural Network?

Chuwei Wang[1*], **Shanda Li**[2*], **Di He**[3], **Liwei Wang**[3,4†]
[1]School of Mathematical Sciences, Peking University
[2]Machine Learning Department, School of Computer Science, Carnegie Mellon University
[3]Key Lab. of Machine Perception (MoE), School of Intelligence Science and Technology, Peking University
[4] Center for Data Science, Peking University
chuwei.wang@pku.edu.cn, shandal@cs.cmu.edu
dihe@pku.edu.cn, wanglw@pku.edu.cn

Abstract

The Physics-Informed Neural Network (PINN) approach is a new and promising way to solve partial differential equations using deep learning. The L^2 Physics-Informed Loss is the de-facto standard in training Physics-Informed Neural Networks. In this paper, we challenge this common practice by investigating the relationship between the loss function and the approximation quality of the learned solution. In particular, we leverage the concept of stability in the literature of partial differential equation to study the asymptotic behavior of the learned solution as the loss approaches zero. With this concept, we study an important class of high-dimensional non-linear PDEs in optimal control, the Hamilton-Jacobi-Bellman (HJB) Equation, and prove that for general L^p Physics-Informed Loss, a wide class of HJB equation is stable only if p is sufficiently large. Therefore, the commonly used L^2 loss is not suitable for training PINN on those equations, while L^∞ loss is a better choice. Based on the theoretical insight, we develop a novel PINN training algorithm to minimize the L^∞ loss for HJB equations which is in a similar spirit to adversarial training. The effectiveness of the proposed algorithm is empirically demonstrated through experiments. Our code is released at https://github.com/LithiumDA/L_inf-PINN.

1 Introduction

Recently, with the explosive growth of available data and computational resources, there have been growing interests in developing machine learning approaches to solve partial differential equations (PDEs) [14, 13, 33, 28]. One seminal work in this direction is the Physics-Informed Neural Network (PINN) approach [28] which parameterizes the PDE's solution as a neural network. By defining differentiable loss functionals that measure how well the model fits the PDE and boundary conditions, the network parameters can be efficiently optimized using gradient-based approaches. L^2 distance is one of the most popularly used measures, which calculates the L^2 norm of the PDE and boundary residual on the domain and boundary, respectively. Previous works demonstrated that PINN could solve a wide range of PDE problems using the L^2 Physics-Informed Loss, such as Poisson equation, Burgers' equation, and Navier-Stokes equation [28, 6].

Although previous works empirically demonstrated promising results using L^2 Physics-Informed Loss, we argue the plausibility of using this loss for (high-dimensional) non-linear PDE problems.

*Equal contribution.

†Correspondence to: Liwei Wang <wanglw@pku.edu.cn>.

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

We know the trivial fact that the learned solution will equal the exact solution when its L^2 loss equals zero. However, the quality of a learned solution with a small but non-zero loss, which is a more realistic scenario in practice, remains unknown to have any approximation guarantees. In this work, we aim at answering a fundamental question:

Can we guarantee that a learned solution with a small Physics-Informed Loss always corresponds to a good approximator of the exact solution?

To thoroughly investigate the problem, we advocate analyzing the stability of PDE [8] in the PINN framework. Stability characterizes the asymptotic behavior of the distance between the learned solution and the exact solution when the Physics-Informed Loss approaches zero. If the PDE is not stable with respect to certain loss functions, we may not obtain good approximate solutions by minimizing the loss. To show the strength of the theory, we perform a comprehensive study on the stability of an important class of high-dimensional non-linear PDEs in optimal control, the Hamilton-Jacobi-Bellman (HJB) equation, which establishes a necessary and sufficient condition for a control's optimality with regard to the cost function. Interestingly, we prove that for general L^p Physics-Informed Loss, the HJB equation is stable only if p is sufficiently large. This finding suggests that the most widely used L^2 loss may not be suitable for training PINN on HJB equations as the learned solution can be arbitrarily distant from the exact solution. Empirical observation verifies the theoretical results.

We further show our theory can serve as a principled way to design loss functions for training PINN. For the high-dimensional HJB equation we target in the paper, the theoretical result suggests that L^∞ loss may be a better choice to learn approximate solutions. Motivated by this insight, we propose a new algorithm for training PINN, which adopts a min-max optimization procedure to minimize the L^∞ loss. Our approach resembles the well-known adversarial training framework. In each iteration, we first fix the network parameters and learn adversarial data points to approximate L^∞ loss, and then optimize the network parameters to minimize the loss. When the training finishes, the learned network will converge to a solution with small L^∞ losses and is close to the exact solution. We conduct experiments to demonstrate the effectiveness of the proposed algorithm. All empirical results show that our method can indeed learn accurate solutions for HJB equations and is much better than several baseline methods.

The contribution of the paper is summarized as follows.

- We make the first step towards theoretically studying the loss design in PINN, and formally introduce the concept of stability in the literature of PDE to characterize the quality of a learned solution with small but non-zero Physics-Informed Loss.
- We provide rigorous investigations on an important class of high-dimensional non-linear PDEs in optimal control, the HJB equation. Our results suggest that the widely used L^2 loss is not a suitable choice for training PINN on HJB equations.
- Based on the theoretical insight, we develop a novel PINN training algorithm to minimize the L^∞ loss for HJB equations. We empirically demonstrate that the proposed algorithm can significant improve the accuracy of PINN in solving the optimal control problems.

2 Related Works

Physics-Informed Neural Network approaches [33, 28] learn to find parametric solutions to satisfy equations and boundary conditions with gradient descent. There has been a notable scarcity of papers that rigorously justify why PINNs work. Important works include [31], which prove the convergence of PINN for second-order elliptic and parabolic equations. In [21], the authors study the statistical limit of learning a PDE solution from sampled observations for elliptic equations. In [29], the convergence of PINN with L^2 loss is established for Navier-Stokes equations. At the same time, several works observed different failure modes for training PINN in other PDE problems. In [17], researchers discover that PINN sometimes fails to learn accurate solutions to a class of convection and reaction equations. Their analysis shows that this may be attributed to the complicated loss landscape. [34] observed PINN failed to learn the Helmholtz equation due to the incommensurability between PDE and boundary losses.

In this work, we mainly experiment with the Hamilton-Jacobi-Bellman (HJB) equation in optimal control. Previously, there were several works aiming at solving the HJB equation using deep learning methods [12, 13, 25, 39, 24, 1, 27, 5]. [13] is among the first to leverage neural networks to solve HJB equations. In particular, [13] targets constructing an approximation to a solution value u at $T = 0$, which is further transformed into a backward stochastic differential equation and learned by neural networks. The main difference between [13] and ours is that [13] only learns the solution on a pre-defined time frame, while with our method, the obtained solution can be evaluated for any time frame. Recently, [12] tackled the offline reinforcement learning problem and developed a soft relaxation of the classical HJB equation, which can be learned using offline behavior data. The main difference between [12] and ours is that no additional data is required in our setting.

Stability is one of the most fundamental concepts in studying the well-posedness of PDE problems. Formally speaking, it characterizes the behavior of the solution to a PDE problem when a small perturbation modifies the operator, initial condition, boundary condition, or force term. We say the equation is stable if the solution of the perturbed PDE converges to the exact solution as the perturbations approach zero [8]. The problem regarding whether a PDE is stable has been intensively studied [8, 19, 11] in literature. There are also some works studying how (regularity) conditions affect stability. [20] and [7] investigate in which topology Couette Flow is asymptotic stable. [4] answers in which Sobolev space defocusing nonlinear Schrodinger equation, real Korteweg-de Vries (KdV), and modified KdV are locally well-posed. Our main focus is akin to the latter works but is settled in the machine learning framework.

3 Preliminary

In this section, we introduce basic background on Physics-Informed Neural Networks and stochastic optimal control problems. Without loss of generality, we formulate any partial differential equation as:

$$\begin{cases} \mathcal{L}u(x) = \varphi(x) & x \in \Omega \subset \mathbb{R}^n \\ \mathcal{B}u(x) = \psi(x) & x \in \partial\Omega, \end{cases} \tag{1}$$

where $\mathcal{L}$ is the partial differential operator and $\mathcal{B}$ is the boundary condition. We use x to denote the spatiotemporal-dependent variable, and use Ω and $\partial\Omega$ to denote the domain and boundary.

Physics-Informed Neural Networks (PINN) PINN [28] is a popular choice to learn the function $u(x)$ automatically by minimizing the loss function induced by the PDE (1). To be concrete, given $p \in (1, +\infty)$, we define the L^p Physics-Informed Loss as

$$\ell_{\Omega,p}(u) = \|\mathcal{L}u(x) - \varphi(x)\|^p_{L^p(\Omega)}, \tag{2}$$

$$\ell_{\partial\Omega,p}(u) = \|\mathcal{B}u(x) - \psi(x)\|^p_{L^p(\partial\Omega)}. \tag{3}$$

The loss term $\ell_{\Omega,p}(u)$ in Eq. (2) corresponds to the PDE residual, which evaluates how $u(x)$ fits the partial differential equation on Ω; and $\ell_{\partial\Omega,p}(u)$ in Eq. (3) corresponds to the boundary residual, which measures how well $u(x)$ satisfies the boundary condition on $\partial\Omega$. L^p denotes p-norm, where p is usually set to 2, leading to a "mean squared error" interpretation of the loss function [33, 28]. The goal is to find u^* that minimizes a linear combination of the two losses defined above. The function $u(x)$ is usually parameterized by neural network $u_\theta(x)$ with parameter $\theta \in \Theta$. To find θ^* efficiently, PINN approaches use gradient-based optimization methods. Note that computing the loss involves integrals over Ω and $\partial\Omega$. Thus, Monte Carlo methods are commonly used to approximate $\ell_{\Omega,p}(u)$ and $\ell_{\partial\Omega,p}(u)$ in practice.

Stochastic Optimal Control Stochastic control [9, 3] is an important sub-field in optimal control theory. In stochastic control, the state function $\{X_t\}_{0\le t\le T}$ is a stochastic process, where T is the time horizon of the control problem. The evolution of the state function is governed by the following stochastic differential equation:

$$\begin{cases} \mathrm{d}X_s = m(s, X_s)\mathrm{d}s + \sigma \mathrm{d}W_s & s \in [t, T] \\ X_t = x \end{cases}, \tag{4}$$

where $m : [t, T] \times \mathbb{R}^n \to \mathbb{R}^n$ is the control function and $\{W_s\}$ is a standard n-dimensional Brownian motion.

Given a control function m, its total cost is defined as $J_{x,t}(m) = \mathbb{E}\int_t^T r(X_s, m, s)\mathrm{d}s + g(X_T)$, where $r : \mathbb{R}^n \times \mathbb{R}^n \times [0, T] \to \mathbb{R}$ measures the cost rate during the process and $g : \mathbb{R}^n \to \mathbb{R}$ measures the final cost at the terminal state. The expectation is taken over the randomness of the trajectories.

We are interested in finding a control function that minimizes the total cost for a given initial state. Formally speaking, we define the *value function* of the control problem (4) as $u(x, t) = \min_{m\in\mathcal{M}} J_{x,t}(m)$, where $\mathcal{M}$ denotes the set of possible control functions that we take into consideration. It can be obtained that the value function will follow a particular partial differential equation as stated below.

Definition 3.1 ([38])**.** *The value function $u(x, t)$ is the unique solution to the following partial differential equation, which is called* ***Hamilton-Jacobi-Bellman Equation****:*

$$\begin{cases} \partial_t u(x,t) + \frac{1}{2}\sigma^2 \Delta u(x,t) + \min_{m\in\mathcal{M}} [r(x, m(t,x), t) + \nabla u \cdot m_t] = 0 \\ u(x,T) = g(x). \end{cases} \tag{5}$$

Hamilton-Jacobi-Bellman (HJB) equation establishes a necessary and sufficient condition for a control's optimality with regard to the cost functions. It is one of the most important high-dimensional PDEs [16] in optimal control with tremendous applications in physics [32], biology [18], and finance [26]. Many well-known equations, including Riccati equation, Linear–Quadratic–Gaussian control problem [36], Merton's portfolio problem [22] are special cases of HJB equation [37].

Conventionally, the solution to the HJB equation, i.e., the value function $u(x, t)$, can be computed using dynamic programming [2]. However, the computational complexity of dynamic programming will grow exponentially with the dimension of state function. Considering that the state function in many applications is high-dimensional, solving such HJB equations is notoriously difficult in practice using conventional solvers. As neural networks have shown impressive power in learning high-dimensional functions, it's natural to resort to neural-network-based approaches for solving high-dimensional HJB equations.

4 Failure Mode of PINN on High-Dimensional Stochastic Optimal Control

Note that $u(x)$ is the exact solution to the PDE (1) if and only if both loss terms $\ell_{\Omega,p}(u)$ and $\ell_{\partial\Omega,p}(u)$ are zero. However, in practice, we usually can only obtain small but non-zero loss values due to the randomness in the optimization procedure or the capacity of the neural network. In such cases, a natural question arises: whether a learned $u(x)$ with a small loss will correspond to a good approximator to the exact solution $u^*(x)$? Such a property is highly related to the concept *stability* in PDE literature, which can be defined as below in our learning scenario:

Definition 4.1. *Suppose Z_1, Z_2, Z_3 are three Banach spaces. We say a PDE defined as Eq. (1) is (Z_1, Z_2, Z_3)-stable, if $\|u^*(x) - u(x)\|_{Z_3} = O(\|\mathcal{L}u(x) - \varphi(x)\|_{Z_1} + \|\mathcal{B}u(x) - \psi(x)\|_{Z_2})$ as $\|\mathcal{L}u(x) - \varphi(x)\|_{Z_1}, \|\mathcal{B}u(x) - \psi(x)\|_{Z_2} \to 0$ for any function u.*

By definition, if a PDE is $(L^2(\Omega), L^2(\partial\Omega), Z)$-stable with a suitable Banach space Z, we can minimize the widely used L^2 Physics-Informed Losses $\|\mathcal{L}u(x) - \varphi(x)\|^2_{L^2(\Omega)}$ and $\|\mathcal{B}u(x) - \psi(x)\|^2_{L^2(\partial\Omega)}$, and the learned solution is guaranteed to be close to the exact solution when the loss terms approach zero. However, stability is not always an obvious property for PDEs. There are tremendous equations that are unstable, such as the inverse heat equation. Moreover, even if an equation is stable, it is possible that the equation is not $(L^2(\Omega), L^2(\partial\Omega), Z)$-stable, which suggests that the original L^2 Physics-Informed Loss might not be a good choice for solving it. We will show later that for control problems, some practical high-dimensional HJB equations are stable but not $(L^2(\Omega), L^2(\partial\Omega), Z)$-stable, and using L^2 Physics-Informed Loss will fail to find an approximated solution in practice.

We consider a class[3] of HJB equations in which the cost rate function is formulated as $r(x, m) = a_1|m_1|^{\alpha_1} + \cdots + a_n|m_n|^{\alpha_n} - \varphi(x, t)$. The corresponding Hamilton-Jacobi-Bellman equation can

[3]The form of cost function we investigate in the paper is representative in optimal control. For example, in financial markets, we often face power-law trading cost in optimal execution problems [10, 30]. The cost function in Linear–Quadratic–Gaussian control and Merton's portfolio model (constant relative risk aversion utility function in [22]) is also of this form. Therefore, we believe our theoretical analysis for this class of HJB equation is relevant for practical applications.

be reformulated as:

$$\begin{cases} \mathcal{L}_{\text{HJB}} u := \partial_t u(x,t) + \frac{1}{2}\sigma^2 \Delta u(x,t) - \sum_{i=1}^{n} A_i |\partial_{x_i} u|^{c_i} = \varphi(x,t) & (x,t) \in \mathbb{R}^n \times [0,T] \\ \mathcal{B}_{\text{HJB}} u := u(x,T) = g(x) & x \in \mathbb{R}^n \end{cases}, \quad (6)$$

where $A_i = (a_i\alpha_i)^{-\frac{1}{\alpha_i - 1}} - a_i(a_i\alpha_i)^{-\frac{\alpha_i}{\alpha_i - 1}} \in (0, +\infty)$ and $c_i = \frac{\alpha_i}{\alpha_i - 1} \in (1, \infty)$. See Appendix B for the detailed derivation. For a function $f : X \to \mathbb{R}$, where X is a measurable space, we denote by $\text{supp} f$ the support set of f, i.e. the closure of $\{x \in X : f(x) \neq 0\}$.

An important concept for analyzing PDEs is the Sobolev space, which is defined as follows:

Definition 4.2. *For $m \in \mathbb{N}$, $p \in [1, +\infty)$ and an open set $\Omega \subset \mathbb{R}^n$, the Sobolev space $W^{m,p}(\Omega)$ is defined as $\{f(x) \in L^p(\Omega) : D^\alpha f \in L^p(\Omega), \forall \alpha \in \mathbb{N}^n, |\alpha| \leq m\}$. The function space $W^{m,p}(\Omega)$ is equipped with Sobolev norm, which is defined as $\|f\|_{W^{m,p}(\Omega)} = \left(\sum_{|\alpha| \leq m} \|D^\alpha f\|^p_{L^p(\Omega)} \right)^{\frac{1}{p}}$.*

The definition above can be extended to functions defined on a spatiotemporal domain $Q \subseteq \mathbb{R}^n \times [0,T]$. With a slight abuse of notation, we define $W^{m,p}(Q) = \{f(x,t) \in L^p(Q) : D^\alpha f \in L^p(Q), \forall \alpha \in \mathbb{N}^n, |\alpha| \leq m\}$, where the differential D^α is only operated over spatial variable x. The norm $\|\cdot\|_{W^{m,p}(Q)}$ can also be defined accordingly.

Stability of the HJB Equation We present our main theoretical result which characterizes the stability of the HJB equation (Eq. (6)). In particular, we show that the HJB equation is $(L^p(\mathbb{R}^n \times [0,T]), L^q(\mathbb{R}^n), W^{1,r}(\mathbb{R}^n \times [0,T]))$-stable when p, q and r satisfies certain conditions. We take the Banach space Z_3 in Definition 4.1 as $W^{1,r}$ here because it captures the properties of both the value and the derivatives of a function, but L^p spaces do not. However, as could be seen from Appendix B, for optimal control problems, it is essential to obtain an accurate approximator for both the value and the gradient of the value function u (the solution of (Eq. (6)). Thus, it is appropriate to analyze the quality of the approximate solution in $W^{1,r}$ space.

Theorem 4.3. *For $p, q \geq 1$, let $r_0 = \frac{(n+2)q}{n+q}$. Assume the following inequalities hold for p, q and r_0:*

$$p \geq \max\left\{2, \left(1 - \frac{1}{\bar{c}}\right) n\right\};\ q > \frac{(\bar{c}-1)n^2}{(2-\bar{c})n + 2};\ \frac{1}{r_0} \geq \frac{1}{p} - \frac{1}{n}, \quad (7)$$

where $\bar{c} = \max_{1 \leq i \leq n} c_i$ in Eq. (6). Then for any $r \in [1, r_0)$ and any bounded open set $Q \subset \mathbb{R}^n \times [0,T]$, Eq. (6) is $(L^p(\mathbb{R}^n \times [0,T]), L^q(\mathbb{R}^n), W^{1,r}(Q))$-stable for $\bar{c} \leq 2$.

The proof of Theorem 4.3 can be found in Appendix C and an improved theorem with relaxed dependency on $\bar{c}$ can be found in Appendix E. Intuitively, Theorem 4.3 states that $(L^p, L^q, W^{1,r})$-stability of Eq. (6) can be achieved when $p, q = \Omega(n)$. We further show that this linear dependency on n cannot be relaxed in the following theorem:

Theorem 4.4. *There exists an instance of Eq. (6), whose exact solution is u^*, such that for any $\varepsilon > 0, A > 0, r \geq 1, m \in \mathbb{N}$ and $p \in \left[1, \frac{n}{4}\right]$, there exists a function $u \in C^\infty(\mathbb{R}^n \times (0,T])$ which satisfies the following conditions:*

- *$\|\mathcal{L}_{\text{HJB}} u - \varphi\|_{L^p(\mathbb{R}^n \times [0,T])} < \varepsilon$, $\mathcal{B}_{\text{HJB}} u = \mathcal{B}_{\text{HJB}} u^*$, and $\text{supp}(u - u^*)$ is compact, where $\mathcal{L}_{\text{HJB}}$ and $\mathcal{B}_{\text{HJB}}$ are defined in Eq. (6).*
- *$\|u - u^*\|_{W^{m,r}(\mathbb{R}^n \times [0,T])} > A$.*

The proof of Theorem 4.4 can be found in Appendix D.

Discussion. Theorem 4.3 and 4.4 together state that when the dimension of the state function n is large, the HJB equation in Eq. (6) cannot be $(L^p, L^q, W^{1,r})$-stable if p and q are small. Furthermore, since L^r=$W^{0,r}$ by definition, Theorem 4.4 also implies that Eq. (6) is not even (L^p, L^q, L^r)-stable. Therefore, for high-dimensional HJB problems, if we use classic L^2 Physics-Informed Loss for training PINN, the learned solution may be arbitrarily distant from u^* even if the loss is very small. Such theoretical results are verified in our empirical studies in Section 6.

More importantly, our theoretical results indicate that the design choice of the Physics-Informed Loss plays a significant role in solving PDEs using PINN. In this work, we shed light upon this problem using HJB equations. We believe the relationship between PDE's stability and the Physics-Informed Loss should be carefully investigated in the future, especially for high-dimensional non-linear PDEs whose stability are more complicated than low-dimensional and linear ones [8, 11, 19]. Given the above observations, we further propose a new algorithm for training PINN to solve HJB Equations, which will be presented in the subsequent sections.

5 Solving HJB Equations with Adversarial Training

The above results suggest that we should use a large value of p and q in the loss $\ell_{\Omega,p}(u)$ and $\ell_{\partial\Omega,q}(u)$ to guarantee a learned solution u is close to u^* for high-dimensional HJB problems. Note that L^p-norm and L^∞-norm behave similarly when p is large. We can substitute L^p-norm by L^∞-norm and directly optimize $\ell_{\Omega,\infty}(u)$ and $\ell_{\partial\Omega,\infty}(u)$. Overall, the training objective can be formulated as:

$$\min_u \ \ell_\infty(u) = \sup_{x\in\Omega} |\mathcal{L}u(x) - \varphi(x)| + \lambda \sup_{x\in\partial\Omega} |\mathcal{B}u(x) - \psi(x)|, \tag{8}$$

where $\lambda > 0$ is a hyper-parameter to trade off the two objectives.

It is straightforward to obtain that setting p and q to infinity satisfies the conditions in Theorem 4.3, and thus the quality of the learned solution enjoys theoretical guarantee. Furthermore, Eq. (8) can be regarded as a min-max optimization problem. The inner loop is a maximization problem to find data points on Ω and $\partial\Omega$ where u violates the PDE most, and the outer loop is a minimization problem to find u (i.e., the neural network parameters) that minimizes the loss on those points.

In deep learning, such a min-max optimization problem has been intensively studied, and adversarial training is one of the most effective learning approaches in many applications. We leverage adversarial training, and the detailed implementation is described in Algorithm 1. In each training step, the model parameters and data points are iteratively updated. We first fix the model u and randomly sample data points $x^{(1)}, \cdots, x^{(N_1)} \in \Omega$ and $\tilde{x}^{(1)}, \cdots, \tilde{x}^{(N_2)} \in \partial\Omega$, serving as a random initialization of the inner loop optimization. Then we perform gradient-based methods to obtain data points with large point-wise Physics-Informed Losses, which leads to the following inner-loop update rule:

$$x^{(k)} \leftarrow \mathrm{Project}_\Omega\left(x^{(k)} + \eta\, \mathrm{sign}\nabla_x \left(\mathcal{L}u_\theta(x^{(k)}) - \varphi(x^{(k)})\right)^2\right); \tag{9}$$

$$\tilde{x}^{(k)} \leftarrow \mathrm{Project}_{\partial\Omega}\left(\tilde{x}^{(k)} + \eta\, \mathrm{sign}\nabla_x \left(\mathcal{B}u_\theta(\tilde{x}^{(k)}) - \psi(\tilde{x}^{(k)})\right)^2\right), \tag{10}$$

Algorithm 1 L^∞ Training for Physics-Informed Neural Networks

Input: Target PDE (Eq. (1)); neural network u_θ; initial model parameters θ
Output: Learned PDE solution u_θ
Hyper-parameters: Number of total training iterations M; number of iterations and step size of inner loop K, η; weight for combining the two loss term λ

for $i = 1, \cdots, M$ **do**
  Sample $x^{(1)}, \cdots, x^{(N_1)} \in \Omega$ and $\tilde{x}^{(1)}, \cdots, \tilde{x}^{(N_2)} \in \partial\Omega$
  for $j = 1, \cdots, K$ **do**
    for $k = 1, \cdots, N_1$ **do**
      $x^{(k)} \leftarrow \mathrm{Project}_\Omega\left(x^{(k)} + \eta\, \mathrm{sign}\nabla_x \left(\mathcal{L}u_\theta(x^{(k)}) - \varphi(x^{(k)})\right)^2\right)$
    for $k = 1, \cdots, N_2$ **do**
      $\tilde{x}^{(k)} \leftarrow \mathrm{Project}_{\partial\Omega}\left(\tilde{x}^{(k)} + \eta\, \mathrm{sign}\nabla_x \left(\mathcal{B}u_\theta(\tilde{x}^{(k)}) - \psi(\tilde{x}^{(k)})\right)^2\right)$
  $g \leftarrow \nabla_\theta\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\left(\mathcal{L}u_\theta(x^{(i)}) - \varphi(x^{(i)})\right)^2 + \lambda\cdot\frac{1}{N_2}\sum_{i=1}^{N_2}\left(\mathcal{B}u_\theta(\tilde{x}^{(i)}) - \psi(\tilde{x}^{(i)})\right)^2\right)$
  $\theta \leftarrow \mathrm{Optimizer}\,(\theta, g)$
return u_θ

Table 1: **Experimental results of solving the 100/250-dimensional LQG control problems.** n denotes the dimensionality of the problem. Performances are measured by L^1, L^2, and $W^{1,1}$ relative error in $[0,1]^n \times [0,T]$. The best performances are indicated in **bold**.

Method	Relative error for $n = 100$			Relative error for $n = 250$		
	L^1	L^2	$W^{1,1}$	L^1	L^2	$W^{1,1}$
Original PINN [28]	3.47%	4.25%	11.31%	6.74%	7.67%	17.51%
Adaptive time sampling [35]	3.05%	3.67%	13.63%	7.18%	7.91%	18.38%
Learning rate annealing [34]	11.09%	11.82%	33.61%	6.94%	8.04%	18.47%
Curriculum regularization [17]	3.40%	3.91%	9.53%	6.72%	7.51%	17.52%
Adversarial training (ours)	**0.27%**	**0.33%**	**2.22%**	**0.95%**	**1.18%**	**4.38%**

where $\mathrm{Project}_{\Omega}(\cdot)$ and $\mathrm{Project}_{\partial\Omega}(\cdot)$ project the updated data points to the domain. When the inner-loop optimization finishes, we fix the generated data points and calculate the gradient g to the model parameter:

$$g \leftarrow \nabla_\theta \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \left(\mathcal{L} u_\theta(x^{(i)}) - \varphi(x^{(i)}) \right)^2 + \lambda \cdot \frac{1}{N_2} \sum_{i=1}^{N_2} \left(\mathcal{B} u_\theta(\tilde{x}^{(i)}) - \psi(\tilde{x}^{(i)}) \right)^2 \right), \tag{11}$$

then the model parameter can be updated using any first-order optimization methods. When the training finishes, the learned neural network will converge to a solution with small L^∞ losses and is guaranteed to be close to the exact solution.

6 Experiments

In this section, we conduct experiments to verify the effectiveness of our approach. Ablation studies on the design choices and hyper-parameters are then provided. Our codes are implemented based on `PyTorch` [23]. All the models are trained on one NVIDIA Tesla V100 GPU with 16GB memory. Due to space limitation, we only showcase our methods on the Linear Quadratic Gaussian control problem in the main body of the paper. More experimental results on other PDE problems can be found in Appendix G.

6.1 High Dimensional Linear Quadratic Gaussian Control Problem

We follow [13] to study the classical linear-quadratic Gaussian (LQG) control problem in n dimensions, a special case of the HJB equation:

$$\begin{cases} \partial_t u(x,t) + \Delta u(x,t) - \mu \|\nabla_x u(x,t)\|^2 = 0 & x \in \mathbb{R}^n, t \in [0,T] \\ u(x,T) = g(x) & x \in \mathbb{R}^n, \end{cases} \tag{12}$$

As is shown in [13], there is a unique solution to Eq. (12):

$$u(x,t) = -\frac{1}{\mu} \ln \left(\int_{\mathbb{R}^n} (2\pi)^{-n/2} \mathrm{e}^{-\|y\|^2/2} \cdot \mathrm{e}^{-\mu g(x - \sqrt{2(T-t)}y)} \mathrm{d}y \right), \tag{13}$$

We set $\mu = 1$, $T = 1$, and the terminal cost function $g(x) = \ln\left(\frac{1 + \|x\|^2}{2}\right)$.

Experimental Design The neural network used for training is a 4-layer MLP with 4096 neurons and $\tanh$ activation in each hidden layer. To train the models, we use Adam as the optimizer [15]. The learning rate is set to $7\mathrm{e}-4$ in the beginning and then decays linearly to zero during training. The total number of training iterations is set to 5000/10000 for the 100/250-dimensional problem. In each training iteration, we sample $N_1 = 100/50$ points from the domain $\mathbb{R}^n \times [0,T]$ and $N_2 = 100/50$ points from the boundary $\mathbb{R}^n \times \{T\}$ to obtain a mini-batch for the 100/250-dimensional problem. The number of inner-loop iterations K is set to 20, and the inner-loop step size η is set to 0.05 unless otherwise specified. Evaluations are performed on a hold-out validation set which is unseen during

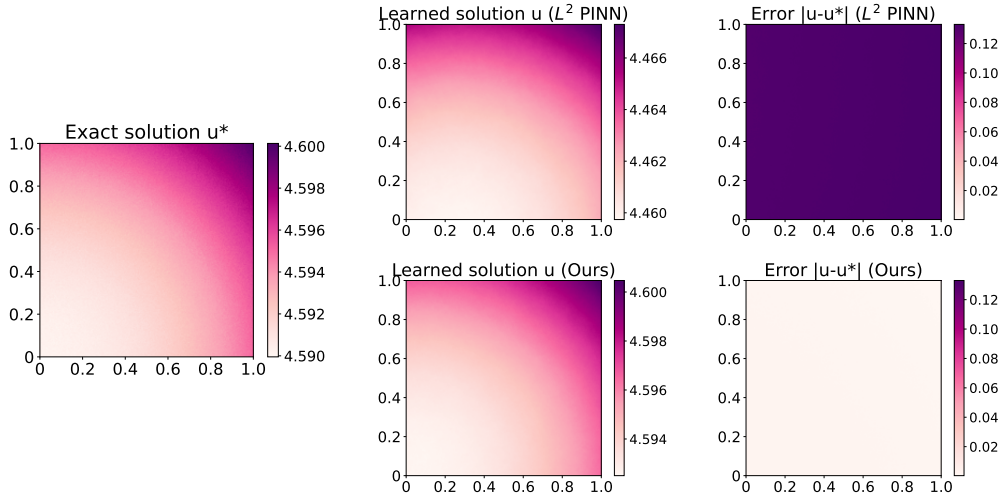


Figure 1: **Visualization for the solutions of Eq. (12)**. The left panel shows the exact solution u^*; the middle panel shows the learned solutions u of the original PINN method with L^2 loss and our method with adversarial training; the right panel show the point-wise absolute error $|u - u^*|$. Note that the solution is a high dimensional function, and we visualize its snapshot on a two-dimensional domain. Specifically, we visualize a bivariate function $u(x_1, x_2, 0, \cdots, 0; 0)$ for $x_1, x_2 \in [0, 1]$ with the horizontal axis and vertical axis corresponding to x_1 and x_2 respectively.

training. We use the L^1, L^2, and $W^{1,1}$ relative error in $[0, 1]^n \times [0, T]$ as evaluation metrics: L^1 and L^2 relative errors are popular evaluation metrics in literature. We additionally consider $W^{1,1}$ relative error since the gradient of the solution to HJB equations plays an important role in applications, and our theory indicates that Eq. (12) is $(L^\infty, L^\infty, W^{1,r})$ stable. More detailed descriptions of the experimental setting and evaluation metrics can be found in Appendix F.

We compare our method with a few strong baselines: 1) original PINN trained with L^2 Physics-Informed Loss [28]; 2) adaptive time sampling for PINN training proposed in [35]; 3) PINN with the learning rate annealing algorithm proposed in [34]; 4) curriculum PINN regularization proposed in [17]. The training recipes for the baseline methods, including the neural network architecture, the training iterations, the optimizer, and the learning rate, are the same as those of our method described above. It should be noted that although these approaches modifies the data sampler, training algorithms or the loss function, they all keep the L^2 norm of the PDE residual and boundary residual unchanged in the training objective.

Experimental Results The experimental results are summarized in Table 1. It's clear that the relative error of the model trained using the original PINN does not fit the solution well, e.g., the L^1 relative error is larger than 6% when $n = 250$. This empirical observation aligns well with our theoretical analysis, i.e., minimizing L^2 loss cannot guarantee the learned solution to be accurate. Advanced methods, e.g., curriculum PINN regularization [17], can improve the accuracy of the learned solution but with marginal improvement, which suggests that these methods do not address the key limitation of PINN in solving high-dimensional HJB Equations. By contrast, our proposed method significantly outperforms all the baseline methods in terms of both L^p relative error and Sobolev relative error, which indicates that both the values and the gradients of our learned solutions are more accurate than the baselines.

We also examine the quality of the learned solution $u(x, t)$ by visualization. As the solution is a high-dimensional function, we visualize its snapshot on a two-dimensional space. Specifically, we consider a bivariate function $u(x_1, x_2, 0, \cdots, 0; 0)$ and use a heatmap to show its function value given different x_1 and x_2. Figure 1 shows the ground truth u^*, the learned solutions u of original PINN and our method, and the point-wise absolute error $|u - u^*|$ for each methods. The two axises correspond x_1 and x_2, respectively. We can see that the point-wise error of the learned solution using our algorithm is less than $2\mathrm{e} - 2$ on average. In contrast, the point-wise error of the learned solution using original PINN method with L^2 loss is larger than $1.3\mathrm{e} - 1$ for most areas. Therefore,

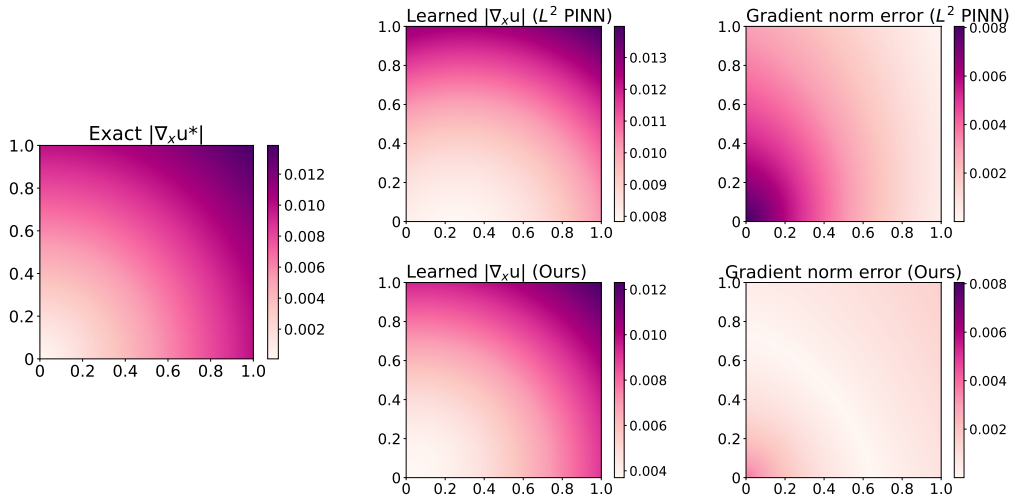


Figure 2: **Visualization for the gradient norm of the solutions of Eq. (12)**. The left panel shows the ground truth $|\nabla_x u^*|$; the middle panel shows the learned $|\nabla_x u|$ of the original PINN method with L^2 loss and our method with adversarial training; the right panel shows the point-wise absolute error $|\nabla_x(u - u^*)|$. We only visualize a snapshot on a two-dimensional domain. Specifically, we visualize a bivariate function $u(x_1, x_2, 0, \cdots, 0; 0)$ for $x_1, x_2 \in [0, 1]$ with the horizontal axis and vertical axis corresponding to x_1 and x_2 respectively.

the visualization of the solutions clearly illustrate that PINN learned more accurate solution using our proposed algorithm.

Furthermore, we visualize the gradient norm $|\nabla_x u|$ of the learned solution of both our method and the original PINN in Figure 2, to illustrate that not only can the learned solution of our method accurately approximate the exact solution, but also the *gradient* of the learned solution can accurately approximate the *gradient* of the exact solution. Again, since $|\nabla_x u|$ is a high dimensional function, we use a heatmap to show its function value given different x_1 and x_2, and set the other variables to 0. From the right panel of Figure 2, we can clearly see that the gradient of the learned solution of our method is much more accurate compared with that of the gradient of the learned solution using original PINN. The gradient norm error of the vanilla PINN approach is nearly $1\mathrm{e}-2$ in some areas shown in the visualization, while the error of our method is less than $1\mathrm{e}-3$ for most data points. This empirical observation aligns well with our theory, which states that Eq. (12) is $(L^\infty, L^\infty, W^{1,r})$ stable.

6.2 Ablation studies

We conduct ablation studies on the 100-dimensional LQG control problem (Eq. (12)) to ablate the main designs in our algorithm.

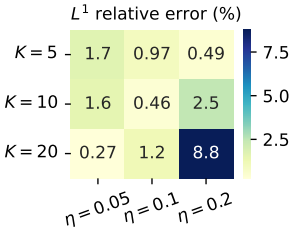
Adversarial training v.s. Directly optimizing L^p physic-informed loss. Our proposed algorithm introduces a min-max optimization procedure. One may have concerns that such an approach may be unnecessarily complicated, and directly minimizing L^p physic-informed loss with a large p would have the same effect. We use these two methods to solve Eq. (12), and compare their performance in the left panel of Table 2. It can be seen that directly minimizing L^p physic-informed loss does not lead to satisfactory results.

We point out that this observation does not contradict our theoretical analysis (Theorem 4.3). Theorem 4.3 focuses on the *approximation* ability, which indicates that a model with a small L^p loss can approximate the exact solution well. The empirical results in Table 2 demonstrate the *optimization* difficulty of learning such a model with L^p loss. By comparison, our proposed adversarial training method is more stable and leads to better performance. More detailed discussions are provided in Appendix H.

Table 2: **Experimental results for ablation studies.** The left panel compares PINN trained with L^p Physic Informed Loss and our method; the right panel compares PINN trained with partial or no adversarial training and our method. In the first two columns of the right panel, ✗ indicates using the original L^2 Physics-Informed Loss for the PDE/boundary residual loss term, while ✓ indicates using the proposed adversarial training method for the corresponding loss term. Performances are measured by L^1 relative error. Best performances are indicated in **bold**.

Method	Relative error
L^4 Loss	2.42%
L^8 Loss	53.55%
L^{16} Loss	113.24%
Ours	**0.27%**

Adversarial training Domain	Adversarial training Boundary	Relative error
✗	✗	3.47%
✗	✓	2.79%
✓	✓	**0.27%**



Adversarial training should be applied to both the PDE residual and the boundary residual. Our theoretical analysis suggests that we should use a large value of p and q in the loss $\ell_{\Omega,p}(u)$ and $\ell_{\partial\Omega,q}(u)$ to guarantee the quality of the learned solution u. Thus, in the proposed Algorithm 1, both the data points inside the domain and the data points on the boundary are learned in the inner-loop maximization. From the right panel of Table 2, we can see that when adversarial training is applied to one loss term, the performance is slightly improved, but its accuracy is still not satisfactory. When both loss terms use adversarial training, the solution is one order of magnitude more accurate, indicating that applying adversarial training to the whole loss function is essential.

Hyper-parameters K and η for the inner loop maximization. Our approach introduces additional hyper-parameters K (the number of inner-loop iterations) and η (the inner-loop step size). These two parameters control the accuracy of inner-loop maximization. We conduct ablation studies to examine the effects of different design choices. Specifically, we experiment with $K = 5, 10, 20$ and $\eta = 0.05, 0.1, 0.2$, and show the L^1 relative error in the right panel of Table 2. Typically we find that setting the product $K\eta = 1$ achieves the best performance. When $K\eta$ is fixed, our results suggest that using a larger K and a smaller η, i.e., more inner-loop iterations and smaller step sizes, will lead to better performance while being more time-consuming.

7 Conclusions

In this paper, we theoretically investigate the relationship between the loss function and the approximation quality of the learned solution using the concept of stability in the partial differential equation. We study an important class of high-dimensional non-linear PDEs in optimal control, the Hamilton-Jacobi-Bellman (HJB) equation, and prove that for general L^p Physics-Informed Loss, the HJB equation is stable only if p is sufficiently large. Such a theoretical finding reveals that the widely used L^2 loss is not suitable for training PINN on high-dimensional HJB equations, while L^∞ loss is a better choice. The theory also inspires us to develop a novel PINN training algorithm to minimize the L^∞ loss for HJB equations in a similar spirit to adversarial training. One limitation of this work is that we only work on the HJB Equation. Theoretical investigation of other important equations can be an exciting direction for future works. We believe this work provides important insights into the loss design in Physics-Informed deep learning.

Acknowledgements

We thank Weinan E, Bin Dong, Yiping Lu, Zhifei Zhang, and Yufan Chen for the helpful discussions.

This work is supported by National Science Foundation of China (NSFC62276005), The Major Key Project of PCL (PCL2021A12), Exploratory Research Project of Zhejiang Lab (No. 2022RC0AN02), and Project 2020BD006 supported by PKUBaidu Fund.

References

[1] Christian Beck, Sebastian Becker, Patrick Cheridito, Arnulf Jentzen, and Ariel Neufeld. Deep splitting method for parabolic pdes. *SIAM Journal on Scientific Computing*, 43(5):A3135–A3154, 2021.

[2] Richard Bellman. Dynamic programming. *Science*, 153(3731):34–37, 1966.

[3] Dimitri Bertsekas and Steven E Shreve. *Stochastic optimal control: the discrete-time case*, volume 5. Athena Scientific, 1996.

[4] Michael Christ, James Colliander, and Terrence Tao. Asymptotics, frequency modulation, and low regularity ill-posedness for canonical defocusing equations. *American journal of mathematics*, 125(6):1235–1293, 2003.

[5] Ashley Davey and Harry Zheng. Deep learning for constrained utility maximisation. *Methodology and Computing in Applied Probability*, 24(2):661–692, 2022.

[6] Tim De Ryck, Ameya D Jagtap, and Siddhartha Mishra. Error estimates for physics informed neural networks approximating the navier-stokes equations. *arXiv preprint arXiv:2203.09346*, 2022.

[7] Yu Deng and Nader Masmoudi. Long time instability of the couette flow in low gevrey spaces. *arXiv preprint arXiv:1803.01246*, 2018.

[8] Lawrence C Evans. Partial differential equations. *Graduate studies in mathematics*, 19(4):7, 1998.

[9] Wendell H Fleming and Raymond W Rishel. *Deterministic and stochastic optimal control*, volume 1. Springer Science & Business Media, 2012.

[10] Peter A Forsyth, J Shannon Kennedy, Shu Tong Tse, and Heath Windcliff. Optimal trade execution: a mean quadratic variation approach. *Journal of Economic dynamics and Control*, 36(12):1971–1991, 2012.

[11] David Gilbarg, Neil S Trudinger, David Gilbarg, and NS Trudinger. *Elliptic partial differential equations of second order*, volume 224. Springer, 1977.

[12] Igor Halperin. Distributional offline continuous-time reinforcement learning with neural physics-informed pdes (sciphy rl for doctr-l). *arXiv preprint arXiv:2104.01040*, 2021.

[13] Jiequn Han, Arnulf Jentzen, and E Weinan. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018.

[14] Yuehaw Khoo, Jianfeng Lu, and Lexing Ying. Solving parametric pde problems with artificial neural networks. *arXiv preprint arXiv:1707.03351*, 2017.

[15] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *ICLR (Poster)*, 2015.

[16] Donald E Kirk. *Optimal control theory: an introduction*. Courier Corporation, 2004.

[17] Aditi Krishnapriyan, Amir Gholami, Shandian Zhe, Robert Kirby, and Michael W Mahoney. Characterizing possible failure modes in physics-informed neural networks. *Advances in Neural Information Processing Systems*, 34, 2021.

[18] Weiwei Li, Emanuel Todorov, and Dan Liu. Inverse optimality design for biological movement systems. *IFAC Proceedings Volumes*, 44(1):9662–9667, 2011.

[19] Gary M Lieberman. *Second order parabolic differential equations*. World scientific, 1996.

[20] Zhiwu Lin and Chongchun Zeng. Inviscid dynamical structures near couette flow. *Archive for rational mechanics and analysis*, 200(3):1075–1097, 2011.

[21] Yiping Lu, Haoxuan Chen, Jianfeng Lu, Lexing Ying, and Jose Blanchet. Machine learning for elliptic pdes: Fast rate generalization bound, neural scaling law and minimax optimality. *arXiv preprint arXiv:2110.06897*, 2021.

[22] Robert C Merton. Optimum consumption and portfolio rules in a continuous-time model. In *Stochastic optimization models in finance*, pages 621–661. Elsevier, 1975.

[23] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32:8026–8037, 2019.

[24] Marcus Pereira, Ziyi Wang, Tianrong Chen, Emily Reed, and Evangelos Theodorou. Feynman-kac neural network architectures for stochastic control using second-order fbsde theory. In *Learning for Dynamics and Control*, pages 728–738. PMLR, 2020.

[25] Marcus Pereira, Ziyi Wang, Ioannis Exarchos, and Evangelos A Theodorou. Learning deep stochastic optimal control policies using forward-backward sdes. *arXiv preprint arXiv:1902.03986*, 2019.

[26] Huyên Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.

[27] Huyen Pham, Xavier Warin, and Maximilien Germain. Neural networks-based backward scheme for fully nonlinear pdes. *SN Partial Differential Equations and Applications*, 2(1):1–24, 2021.

[28] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[29] Tim De Ryck, Ameya D. Jagtap, and Siddhartha Mishra. Error estimates for physics informed neural networks approximating the navier-stokes equations. *CoRR*, abs/2203.09346, 2022.

[30] Alexander Schied and Torsten Schöneborn. Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. *Finance and Stochastics*, 13(2):181–204, 2009.

[31] Yeonjong Shin, Zhongqiang Zhang, and George Em Karniadakis. Error estimates of residual minimization using neural networks for linear pdes. *arXiv preprint arXiv:2010.08019*, 2020.

[32] Stanislaw Sieniutycz. Hamilton–jacobi–bellman framework for optimal control in multistage energy systems. *Physics Reports*, 326(4):165–258, 2000.

[33] Justin Sirignano and Konstantinos Spiliopoulos. Dgm: A deep learning algorithm for solving partial differential equations. *Journal of computational physics*, 375:1339–1364, 2018.

[34] Sifan Wang, Yujun Teng, and Paris Perdikaris. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(5):A3055–A3081, 2021.

[35] Colby L Wight and Jia Zhao. Solving allen-cahn and cahn-hilliard equations using the adaptive physics informed neural networks. *arXiv preprint arXiv:2007.04542*, 2020.

[36] Jan Willems. Least squares stationary optimal control and the algebraic riccati equation. *IEEE Transactions on automatic control*, 16(6):621–634, 1971.

[37] William M Wonham. On a matrix riccati equation of stochastic control. *SIAM Journal on Control*, 6(4):681–697, 1968.

[38] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.

[39] Yajie Yu, Bernhard Hientzsch, and Narayan Ganesan. Backward deep bsde methods and applications to nonlinear problems. *arXiv preprint arXiv:2006.07635*, 2020.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [No] We don't see any potential negative societal impacts of this work.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [Yes] The proofs are provided in the appendix.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes]
   (b) Did you mention the license of the assets? [Yes]
   (c) Did you include any new assets either in the supplemental material or as a URL? [No]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

A Notation and Auxiliary Results

This section gives an overview of the notations used in the paper and summarizes some basic results in partial differential equation and functional analysis.

A.1 Basic notations

For $n \in \mathbb{N}$, we denote $\{1, 2, ...n\}$ by $[n]$ for simplicity in the paper.

For two Banach spaces X, Y, $\mathscr{L}(X, Y)$ refers to the set of continuous linear operator mapping from X to Y.

For a mapping $f : X \to Y$, and $u, v \in X$, $df(u, v)$ denotes Gateaux differential of f at u in the direction of v, and $f'(u)$ denotes Fréchet derivative of f at u.

For $r > 0$, and $x_0 \in X$, where X is a Banach space equipped with norm $\|\cdot\|_X$, $B_r(x_0)$ refers to $\{x \in X : \|x - x_0\|_X < r\}$.

For a function $f : X \to \mathbb{R}$, where X is a measurable space. We denote by $\mathrm{supp} f$ the support set of f, i.e. the closure of $\{x \in X : f(x) \neq 0\}$.

For a measurable set $\Omega \subset \mathbb{R}^n$, define the parabolic region $Q_t = Q_t(\Omega)$ as $\Omega \times [0, t]$.

The parabolic boundary $\partial_p Q_t$ is then defined as $\Omega \times \{0\} \cup \partial\Omega \times [0, t]$.

For $R > 0$, the parabolic neighborhood of 0 (denoted by $Q(R)$) is defined as $\{(x, t) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0} : |x| < R,\ t < R^2\}$.

Its parabolic boundary $\partial_p(Q(R))$ is defined as $\{(x, t) :\ |x| < R,\ t = 0 \text{ or } |x| = R,\ t \in [0, R^2]\}$.

A.2 Multi-index notations

For $n \in \mathbb{N}$, we call an $n-$tuple of non-negative integers $\alpha \in \mathbb{N}^n$ a multi-index. We use the notation $|\alpha| = \Sigma_{i=1}^n \alpha_i,\ \alpha! = \Pi_{i=1}^n \alpha_i!$. For $x = (x_1, x_2, ...x_n) \in \mathbb{R}^n$, we denote by $x^\alpha = \Pi_{i=1}^n x_i^{\alpha_i}$ the corresponding multinomial. Given two multi-indices $\alpha, \beta \in \mathbb{N}^d$, we say $\alpha \leq \beta$ if and only if $\alpha_i \leq \beta_i,\ \forall i \in [n]$.

For an open set $\Omega \subset \mathbb{R}^n,\ T \in \mathbb{R}^+$ and a function $f(x) : \Omega \to \mathbb{R}$ or $f(x, t) : \Omega \times [0, T] \to \mathbb{R}$, we denote by

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} ... \partial x_n^{\alpha_n}} \tag{14}$$

the classical or weak derivative of f.

For $k \in \mathbb{R}^n$, we denote by $D^k f$ (or $\nabla^k f$) the vector whose components are $D^\alpha f$ for all $|\alpha| = k$, and we abbreviate $D^1 f$ as Df.

A.3 Norm notations

Let $n \in \mathbb{N}^*,\ m \in \mathbb{N},\ T \in \mathbb{R}^+$, and $\Omega \subset \mathbb{R}^n$, $Q \subset \mathbb{R}^n \times [0, T]$ be open sets.We denote by $L^p(\Omega)$ and $L^p(Q)$ the usual Lebesgue space.

The Sobolev space $W^{m,p}(\Omega)$ is defined as

$$\{f(x) \in L^p(\Omega) : D^\alpha f \in L^p(\Omega),\ \forall \alpha \in \mathbb{N}^n\ with\ |\alpha| \leq m\}. \tag{15}$$

And we define $W^{m,p}(Q)$ as

$$\{f(x) \in L^p(Q) : D^\alpha f \in L^p(Q),\ \forall \alpha \in \mathbb{N}^n\ with\ |\alpha| \leq m\}. \tag{16}$$

We define

$$\|f\|_{W^{m,p}(\Omega)} := \left(\sum_{|\alpha| \leq m} \|D^\alpha f\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}} \tag{17}$$

for $1 \leq p < \infty$, and

$$\|f\|_{W^{m,\infty}(\Omega)} := \max_{|\alpha| \leq m} \|D^\alpha f\|_{L^\infty(\Omega)} \tag{18}$$

for $p = \infty$.

We define $\|f\|_{W^{m,p}(Q)}$ for $p \in [1, \infty]$ similarly.

We will use simplified notations $\|f\|_p$ and $\|f\|_{m,p}$ for L^p-norm and $W^{m,p}-$norm when the domain is whole space ($\mathbb{R}^n$, $\mathbb{R}^n \times [0, T]$ or $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$) or when it is clear to the reader.

$W_0^{m,p}(\Omega)$ is defined as the completion of $C_0^\infty(\Omega)$ under $\|\cdot\|_{m,p}$ norm. In the similar way we define $W_0^{m,p}(Q_t)$.

A.4 Auxiliary results

In this section, we list out several fundamental yet important results in the field of PDE and functional analysis.

To begin with, we would like to recall three useful inequalities in real analysis.

Lemma A.1 (Young's convolution inequality)**.** *In $\mathbb{R}^n$, we define the convolution of two functions f and g as $(f * g)(x) := \int_{\mathbb{R}^n} f(y)g(x-y)dy$. Suppose $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$, and $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$ with $p, q, r \in [1, \infty]$, then$\|f * g\|_r \leq \|f\|_p \|g\|_q$.*

Lemma A.2 (Sobolev embedding theorem)**.** *Let Ω be an open set in $\mathbb{R}^n$, $p \in [1, \infty]$, and $m \leq k$ be a non-negative integer.*

(i) If $\frac{1}{p} - \frac{k}{n} > 0$, and set $q = \frac{np}{n-pk}$, then $W^{m,p} \subset W^{m-k,q}$ and the embedding is continuous, i.e. there exists a constant $c > 0$ such that $\|u\|_{m-k,q} \leq c\|u\|_{m,p}$, $\forall u \in W^{m,p}$.

(ii) If $\frac{1}{p} - \frac{k}{n} \leq 0$, then for any $q \in [1, \infty)$, $W^{m,p} \subset W^{m-k,q}$ and the embedding is continuous.

Lemma A.3 (A special case of Gagliardo-Nirenberg inequality)**.** *Ω is an open set in $\mathbb{R}^n$, Let $q \in [1, \infty]$ and $j, k \in \mathbb{N}$, and suppose $j \neq 0$ and*

$$\begin{cases} 1 < r < \infty \\ k - j - \frac{n}{r} \notin \mathbb{N} \\ \frac{j}{k} \leq \theta < 1. \end{cases}$$

If we set

$$\frac{1}{p} = \frac{j}{n} + \theta\left(\frac{1}{r} - \frac{k}{n}\right) + \frac{1-\theta}{q}, \tag{19}$$

then there exists a constant C independent of u such that

$$\|\nabla^j u\|_p \leq C\|\nabla^k u\|_r^\theta \|u\|_q^{1-\theta}, \ \forall u \in L^q(\Omega) \cap W^{k,r}(\Omega). \tag{20}$$

Next, we list out several basic results in second order parabolic equations, which will be of great use in Appendix C.

In the following lemmas, we denote by $\mathcal{L}_0$ the operator $\frac{\partial}{\partial t} - \Delta$.

Lemma A.4. *Suppose Ω is bounded, and $Q_T = \Omega \times [0, T]$.*

Let $u \in W^{2,2}(Q_T)$ be the solution to

$$\begin{cases} \mathcal{L}_0 u = f(x,t), & (x,t) \in Q_T \\ u = 0, & (x,t) \in \partial_p Q_T, \end{cases} \tag{21}$$

then for $2 \leq p < \infty$, if $f \in L^p(Q_T)$, we have $u \in W^{2,p}(Q_T)$ and there exists a constant C such that $\|u\|_{2,p} \leq C\|f\|_p$.

Proof. Since Ω is bounded, we can choose an $R_0 > 1$ such that $Q_T \subset Q(R_0)$. Set $\hat{u}(x,t) = u(x,t)\mathbf{1}_{Q_T}$ and $\hat{f}(x,t) = f(x,t)\mathbf{1}_{Q_T}$, which are extensions of u and f in $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$, respectively. The boundary condition in Eq. (21) implies that $\hat{u} \in W^{2,p}(\mathbb{R}^n \times \mathbb{R}_{\geq 0})$.

Furthermore, for any $R > R_0$, it holds that,

$$\begin{cases} \mathcal{L}_0 \hat{u} = \hat{f}(x,t), & (x,t) \in Q(R) \\ \hat{u} = 0, & (x,t) \in \partial_p Q(R). \end{cases}$$

From proposition 7.18 in [19], and in light of the fact that both $\hat{u}$ and $\hat{f}$ are supported on Q_T, we obtain that $\|D^2\hat{u}\|_p \leq C(\|\hat{f}\|_p + \frac{1}{R}\|D\hat{u}\|_p + \frac{1}{R^2}\|\hat{u}\|_p)$ holds for $\forall R > R_0$. Additionally, Poincaré inequality guarantees that there exists a constant $C' > 0$ depending only on Ω and n, such that $\|\hat{u}\|_{2,p} \leq C'\|D^2\hat{u}\|_p$.

Therefore we have

$$\frac{1}{C'}\|\hat{u}\|_{2,p} \leq C\left(\|\hat{f}\|_p + \frac{1}{R}\|\hat{u}\|_{2,p} + \frac{1}{R^2}\|\hat{u}\|_{2,p}\right) \leq C\left(\|\hat{f}\|_p + \frac{2}{R}\|\hat{u}\|_{2,p}\right) \tag{22}$$

$$\left(\frac{1}{C'} - \frac{2C}{R}\right)\|\hat{u}\|_{2,p} \leq C\|\hat{f}\|_p. \tag{23}$$

Let $R \to \infty$ and we derive $\|\hat{u}\|_{2,p} \leq CC'\|\hat{f}\|_p$. Since $\|\hat{u}\|_{W^{2,p}(\mathbb{R}^n \times \mathbb{R}_{\geq 0})} = \|u\|_{W^{2,p}(Q_T)}$ and $\|\hat{f}\|_{L^p(\mathbb{R}^n \times \mathbb{R}_{\geq 0})} = \|f\|_{L^p(Q_T)}$, we completes the proof. □

Lemma A.5. *Let u be the solution to*

$$\begin{cases} \mathcal{L}_0 u(x,t) = 0, & (x,t) \in \mathbb{R}^n \times [0,T] \\ u(x,0) = g(x), & x \in \mathbb{R}^n. \end{cases}$$

For any compact set $Q \subset \mathbb{R}^n \times [0,T]$ and $p \geq 1$, let $r < \frac{(n+2)p}{n+p}$, then

(i) there exists a constant C such that $\|u\|_{W^{1,r}(Q)} \leq C\|g\|_{L^p(\mathbb{R}^n)}$,

(ii) there exists a constant C' such that $\|u\|_{W^{2,r}(Q)} \leq C'\|g\|_{W^{1,p}(\mathbb{R}^n)}$.

Proof. u have the explicit form

$$u(x,t) = \int_{\mathbb{R}^n} \frac{a}{t^{\frac{n}{2}}} e^{-b\frac{|x-y|^2}{t}} g(y)dy, \tag{24}$$

where $a = (4\pi)^{-\frac{n}{2}}$ and $b = \frac{1}{4}$.

Note that u is the convolution between heat kernel $K(x,t) := \frac{a}{t^{\frac{n}{2}}} e^{-b\frac{|x|^2}{t}}$ and $g(x)$, with Lemma A.1, we derive

$$\|u(\cdot,t)\|_{r_0} \leq \|g\|_{p_0}\|K(\cdot,t)\|_{q_0} \tag{25}$$

for $p_0, q_0, r_0 \in [1,\infty]$ satisfying $\frac{1}{p_0} + \frac{1}{q_0} = \frac{1}{r_0} + 1$.

Due to the uniform convergence of (24), we have $\frac{\partial u}{\partial x_i} = (\frac{\partial}{\partial x_i}K(x,t)) * g(x)$ for all $i \in [n]$. Thus we have

$$\|\frac{\partial u(\cdot,t)}{\partial x_i}\|_{r'} \leq \|g\|_{p'}\|\frac{\partial}{\partial x_i}K(x,t)\|_{q'}. \tag{26}$$

for $p', q', r' \in [1,\infty]$ satisfying $\frac{1}{p'} + \frac{1}{q'} = \frac{1}{r'} + 1$.

It is enough to decide appropriate tuples of $(p_0, q_0, r_0), (p', q', r')$.

For $q \geq 1$,

$$\|K(\cdot,t)\|_q^q = \frac{a^q}{t^{\frac{nq}{2}}}\int_{\mathbb{R}^n} e^{-bq\frac{\|x\|^2}{t}}dx = \frac{a^q}{t^{\frac{nq-n}{2}}}\int_{\mathbb{R}^n} e^{-bq\|y\|^2}dy \qquad (y = \frac{1}{\sqrt{t}}x). \tag{27}$$

Thus $\|K(\cdot,t)\|_q = \frac{C}{t^{\frac{nq-n}{2q}}}$, where C is a constant.

As a result, for any $p,q,r \in [1,\infty]$ satisfying $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}+1$,

$$\|u\|^r_{L^r(\mathbb{R}^n\times[0,T])} = \int_0^T \|u(\cdot,t)\|^r_r dt \le \|g\|^r_p \int_0^T \|K(\cdot,t)\|^r_q dt = C^r \|g\|^r_p \int_0^T \frac{dt}{t^{\frac{nq-n}{2q}r}}. \tag{28}$$

Here we have $\|u\|_r \le C_1\|g\|_p$ for a constant $C_1 \iff$ the integral in the R.H.S. of Eq. (28) converges $\iff \frac{nq-n}{2q}r<1$. Then we could decide appropriate tuple (p_0,q_0,r_0) for (25): tuples that satisfy $\frac{n}{n+2}\frac{1}{p_0} < \frac{1}{r_0} \le \frac{1}{p_0}$. (The second inequality comes from the constraint $q_0 \in [1,\infty]$. It could be removed when these L^p-norms are calculated in a bounded domain).

We handle (p',q',r') in (26) with exactly the same method and find that tuples which satisfy $\frac{n}{n+2}\frac{1}{p'}+\frac{1}{n+2} < \frac{1}{r'} \le \frac{1}{p'}$ gives

$$\left\|\frac{\partial u}{\partial x_i}\right\|_{r'} \le C_2\|g\|_{p'}, \tag{29}$$

where C_2 is a constant.

Finally, note that for any bounded set Ω, and $1 \le q < p$, there is a constant C such that $\|v\|_{L^q(\Omega)} \le C\|v\|_{L^p(\Omega)}$ for all $v \in L^p(\Omega)$. Together with the inequalities (28) and (29), this means that for any compact set $Q \subset \mathbb{R}^n \times [0,T]$, $p \ge 1$, and $r < \frac{(n+2)p}{n+p}$,

$$\begin{aligned}
\|u\|_{W^{1,r}(Q)} &= (\|u\|^r_{L^r(Q)} + \sum_{i=1}^n \|\frac{\partial u}{\partial x_i}\|^r_{L^r(Q)})^{\frac{1}{r}} && (30)\\
&\le (C_3\|u\|^r_{L^{r_0}(Q)} + C_4\sum_{i=1}^n \|\frac{\partial u}{\partial x_i}\|^r_{L^{r'}(Q)})^{\frac{1}{r}} && (31)\\
&\le (C_3\|u\|^r_{L^{r_0}(\mathbb{R}^n\times[0,T])} + C_4\sum_{i=1}^n \|\frac{\partial u}{\partial x_i}\|^r_{L^{r'}(\mathbb{R}^n\times[0,T])})^{\frac{1}{r}} && (32)\\
&\le (C_5\|g\|^r_{L^p(\mathbb{R}^n)} + C_6\sum_{i=1}^n \|g\|^r_{L^p(\mathbb{R}^n)})^{\frac{1}{r}} && (33)\\
&= C_7\|g\|_{L^p(\mathbb{R}^n)}, && (34)
\end{aligned}$$

where $r_0, r' \in [p, \frac{(n+2)p}{n+p}) \cap [r,+\infty)$ and all C_i are constants.

This gives the first statement.

Next, we prove the second statement.

Note that $\frac{\partial^2 u}{\partial x_i x_j} = (\frac{\partial}{\partial x_i}K(x,t)) * \frac{\partial g(x)}{\partial x_j}$, $\forall i,j \in [n]$.

With the same argument for (29), we obtain that for any r'', p'' satisfying $\frac{n}{n+2}\frac{1}{p''}+\frac{1}{n+2} < \frac{1}{r''} \le \frac{1}{p''}$,

$$\left\|\frac{\partial^2 u}{\partial x_i x_j}\right\|_{L^{r''}(\mathbb{R}^n\times[0,T])} \le C\left\|\frac{\partial g(x)}{\partial x_j}\right\|_{L^{p''}(\mathbb{R}^n)}, \quad \forall i,j \in [n], \tag{35}$$

where C is a constant.

Therefore, for any compact set $Q \subset \mathbb{R}^n \times [0,T]$, $p \geq 1$, and $r < \frac{(n+2)p}{n+p}$,

$$\|u\|_{W^{2,r}(Q)} \tag{36}$$

$$= \left(\|u\|^r_{L^r(Q)} + \sum_{i=1}^n \left\| \frac{\partial u}{\partial x_i} \right\|^r_{L^r(Q)} + \sum_{i,j=1}^n \left\| \frac{\partial^2 u}{\partial x_i x_j} \right\|^r_{L^r(Q)} \right)^{\frac{1}{r}} \tag{37}$$

$$\leq \left(C_1 \|u\|^r_{L^{r_0}(Q)} + C_2 \sum_{i=1}^n \left\| \frac{\partial u}{\partial x_i} \right\|^r_{L^{r'}(Q)} + C_3 \sum_{i,j=1}^n \left\| \frac{\partial^2 u}{\partial x_i x_j} \right\|^r_{L^{r''}(Q)} \right)^{\frac{1}{r}} \tag{38}$$

$$\leq \left(C_1 \|u\|^r_{L^{r_0}(\mathbb{R}^n \times [0,T])} + C_2 \sum_{i=1}^n \left\| \frac{\partial u}{\partial x_i} \right\|^r_{L^{r'}(\mathbb{R}^n \times [0,T])} + C_3 \sum_{i,j=1}^n \left\| \frac{\partial^2 u}{\partial x_i x_j} \right\|^r_{L^{r''}(\mathbb{R}^n \times [0,T])} \right)^{\frac{1}{r}} \tag{39}$$

$$\leq \left(C_4 \|g\|^r_{L^p(\mathbb{R}^n)} + C_5 \sum_{i=1}^n \|g\|^r_{L^p(\mathbb{R}^n)} + C_6 \sum_{i,j=1}^n \left\| \frac{\partial g}{\partial x_j} \right\|^r_{L^p(\mathbb{R}^n)} \right)^{\frac{1}{r}} \tag{40}$$

$$\leq C_7 \|g\|_{W^{1,p}(\mathbb{R}^n)}, \tag{41}$$

where $r_0, r', r'' \in [p, \frac{(n+2)p}{n+p}) \cap [r, +\infty)$ and all C_i are constants.

This completes the proof. □

At last, we present it here a well-known result in functional analysis.

Lemma A.6 (Inverse function theorem in Banach space)**.** *Let X, Y be two Banach spaces, $V \subset X$ be an open set, and $g \in C^1(V, Y)$ be a mapping. Assume $x_0 \in V$, $y_0 = g(x_0)$ and the inverse of the Fréchet derivative $(g'(x_0))^{-1} \in \mathscr{L}(Y, X)$. Then there exists $r > 0$ and $s > 0$ such that $B_r(y_0) \subset g(V)$, $B_s(x_0) \subset V$ and $g : B_s(x_0) \to g(B_s(x_0))$ is a differmorphism.*

B Derivation of a Class of Hamilton-Jacobi-Bellman (HJB) Equations

For the sake of completeness of the paper, we give the derivation of a class of Hamilton-Jacobi-Bellman (HJB) Equations as below.

To start with, we derive the general form of HJB Equation in stochastic control problem.

In stochastic control, the state function $\{X_t\}_{0 \leq t \leq T}$ is a stochastic process, where T is the time horizon of the control problem. The evolution of the state function is governed by the following stochastic differential equation:

$$\begin{cases} \mathrm{d}X_s = m(s, X_s)\mathrm{d}s + \sigma \mathrm{d}W_s \quad s \in [t, T] \\ X_t = x \end{cases}, \tag{42}$$

where $m : [t, T] \times \mathbb{R}^n \to \mathbb{R}^n$ is the control function and $\{W_s\}$ is a standard n-dimensional Brownian motion.

Given a control function $m = (m_1(s, y), m_2(s, y), ...m_n(s, y))$, $s \in [t, T]$, $y \in \mathbb{R}^n$, its total cost is defined as $J_{x,t}(m) = \mathbb{E} \int_t^T r(X_s, m(s, X_s), s)\mathrm{d}s + g(X_T)$, where $r : \mathbb{R}^n \times \mathbb{R}^n \times [0, T] \to \mathbb{R}$ measures the cost rate during the process and $g : \mathbb{R}^n \to \mathbb{R}$ measures the final cost at the terminal state. The expectation is taken over the randomness of the trajectories.

We are interested in finding a control function that minimizes the total cost for a given initial state. Formally speaking, we define the *value function* of the control problem (42) as $u(x, t) = \min_{m \in \mathcal{M}} J_{x,t}(m)$, where $\mathcal{M}$ denotes the set of possible control functions that we take into consideration.

It is obvious that u satisfies $u(x, T) = g(x)$. In addition, according to dynamical programming principle, we have

$$u(x, t) = \min_{m \in \mathcal{M}} \mathbb{E}\left(\int_t^{t+h} r(X_s, m(s, X_s), s)ds + u(X_{t+h}, t + h)\right), \tag{43}$$

With Ito's formula, we derive

$$u(X_{t+h}, t+h) = u(x,t) + (\partial_t u + \frac{1}{2}\sigma^2 \Delta u)h + \nabla u \cdot (m(t,x)h + \sigma(W_{t+h} - W_t)) + o(h) \quad (44)$$

After taking expectation and some calculation, we derive from (44),

$$0 = (\partial_t u + \frac{1}{2}\sigma^2 \Delta u)h + \min_{m \in \mathcal{M}} \mathbb{E}(\int_t^{t+h} r(X_s, m(s, X_s), s)ds + \nabla u \cdot m(t,x)h) + o(h) \quad (45)$$

$$0 = \partial_t u(x,t) + \frac{1}{2}\sigma^2 \Delta u(x,t) + \min_{m \in \mathcal{M}} (r(x, m(t,x), t) + \nabla u \cdot m(t,x)) \quad (46)$$

Then we get HJB equation

$$\begin{cases} \partial_t u(x,t) + \frac{1}{2}\sigma^2 \Delta u(x,t) + \min_{m \in \mathcal{M}} (r(x, m(t,x), t) + \nabla u \cdot m(t,x)) = 0 \\ u(x,T) = g(x). \end{cases} \quad (47)$$

Next, we further simplify this equation in some special cases.

In practice, different components of the state have different meanings, and thus the effects of controlling corresponding components have different significance. Therefore, the cost function's dependence on each component of m_t takes a very different form.

Based on this argument, we consider the case when $r(x,y)$ takes the form

$$r(x,y,t) = \sum_{i=1}^{n} a_i |y_i|^{\alpha_i} - \varphi(x,t) \quad (48)$$

for some appropriate function φ and $a_i \geq 0, \alpha_i > 1$ (if $\alpha_i \leq 1$, the minimizing term might be $-\infty$),$\forall i \in [n]$.

Denote $m(t,x) = (m_1(t,x), m_2(t,x), ...m_n(t,x))$ as $y \in \mathbb{R}^n$, and $\frac{\partial u(x,t)}{\partial x_i}$ as $\partial_i u$, and suppose that $\mathcal{M}$ is so large that it includes the global minimizor of (43), then the third term in HJB equation (47) could be written as

$$\min_{y \in \mathbb{R}^n} (-\varphi(x,t) + \sum_{i=1}^{n} (a_i |y_i|^{\alpha_i} + y_i \partial_i u)) = \varphi(x,t) + \sum_{i=1}^{n} \min_{y_i \in \mathbb{R}} (a_i |y_i|^{\alpha_i} + y_i \partial_i u). \quad (49)$$

With some simple computation, we get

$$\min_{y_i \in \mathbb{R}} (a_i |y_i|^{\alpha_i} + y_i \partial_i u) = \left(\frac{a_i}{(a_i \alpha_i)^{\frac{\alpha_i}{\alpha_i - 1}}} - \frac{1}{(a_i \alpha_i)^{\frac{1}{\alpha_i - 1}}} \right) |\partial_i u|^{\frac{\alpha_i}{\alpha_i - 1}}. \quad (50)$$

As a result, HJB equation in this case is

$$\begin{cases} \partial_t u(x,t) + \frac{1}{2}\sigma^2 \Delta u(x,t) - \varphi(x,t) - \sum_{i=1}^{n} A_i |\partial_i u|^{c_i} = 0 \\ u(x,T) = g(x), \end{cases} \quad (51)$$

where $A_i = (a_i \alpha_i)^{-\frac{1}{\alpha_i - 1}} - a_i (a_i \alpha_i)^{-\frac{\alpha_i}{\alpha_i - 1}} \in (0, +\infty)$ and $c_i = \frac{\alpha_i}{\alpha_i - 1} \in (1, +\infty)$.

Remark B.1. *After taking the transform $v(x,t) := u(x, T-t)$, the equation above becomes*

$$\begin{cases} \partial_t v(x,t) - \frac{1}{2}\sigma^2 \Delta v(x,t) + \sum_{i=1}^{n} A_i |\partial_i v|^{c_i} = -\varphi(x, T-t) \\ v(x,0) = g(x). \end{cases} \quad (52)$$

We will study this equation in the rest of the paper instead.

Remark B.2. *The minimizer of (49) is $y_i^* = (\frac{|\partial_i u|}{a_i \alpha_i})^{\frac{1}{\alpha_i - 1}}$ and this gives the i-th component of the optimal control m^*. Based on the fact that both the value function u and the optimal control m^* are of interest in applications, it is necessary to study this equation in $W^{1,p}$ space.*

Moreover, in most cases, only a bounded domain $\Omega \subset \mathbb{R}^n$ is taken into consideration in both real applications and numerical experiments. Therefore, we study this equation in the space of $W^{1,p}(\Omega \times [0,T])$ for a bounded domain Ω, instead of $W^{1,p}(\mathbb{R}^n \times [0,T])$.

Remark B.3. *The form of cost function (48) we investigate in the paper is a generalization of the widely-used power-law cost (or utility) function, which is representative in optimal control. For example, in financial markets, we often face power-law trading cost in optimal execution problems [10, 30]. The cost function in Linear–Quadratic–Gaussian control and Merton's portfolio model (constant relative risk aversion utility function in [22]) is also of this form. Therefore, we believe our theoretical analysis for this class of HJB equation is relevant for practical applications.*

C Proof of Theorem 4.3

In this section, we give the proof of an equivalent statement of Theorem 4.3.

In light of remark B.1, it is equivalent to consider the stability property (as is defined in Definition 4.1) for the following equation:

$$\begin{cases} \partial_t u(x,t) - \frac{1}{2}\sigma^2 \Delta u(x,t) + \sum_{i=1}^n A_i |\partial_i u|^{c_i} = h(x,t) \quad (x,t) \in \mathbb{R}^n \times [0,T] \\ u(x,0) = g(x), \end{cases} \tag{53}$$

where $A_i > 0$, $c_i \in (1,\infty)$, and $h(x,t)$ corresponds to $-\varphi(x, T-t)$ in Eq. (52). Without loss of generality, we assume $\sigma = \sqrt{2}$ for simplicity in the discussion below.

We define operators $\mathcal{L}_0 := \frac{\partial}{\partial t} - \Delta$, $\tilde{\mathcal{L}}_{\mathrm{HJB}} u := \mathcal{L}_0 u + \sum_{i=1}^n A_i |\partial_i u|^{c_i}$ and $\tilde{\mathcal{B}}_{\mathrm{HJB}} u(x,t) := u(x,0)$ for clarity. We define $\bar{c}$ as $\max\limits_{i \in [n]} c_i$.

We start with the proof of some auxiliary results.

Lemma C.1. *For every $c > 1$, there exist $k \in \mathbb{N}$, $\{t_i\}_{i=1}^k$ satisfying $1 \le t_1 < t_2 < ... < t_k < c$, and k power functions $f_1, ...f_k$ whose orders are strictly less than c and no smaller than 0, such that*

$$(b+w)^c - b^c - w^c \le \sum_{i=1}^k f_i(b) w^{t_i}, \ \forall b, w \ge 0. \tag{54}$$

Proof. Obviously, the inequality holds when $c \in \mathbb{N}$ because of the binomial expansion. We will only consider the case when $c \notin \mathbb{N}$.

*Step 1.*We prove that for any $b, w \ge 0$ such that $\max\{b, x\} \ge 1$, $F_{b,w}(c) := (b+w)^c - b^c - w^c$ is monotone increasing.

Without loss of generality, suppose $b \ge w$. Then

$$F'(c) = (b+w)^c \ln(b+w) - b^c \ln b - w^c \ln w \tag{55}$$

$$= (b+\eta)^{c-1}(1 + c\ \ln(b+\eta))w - w^c \ln w \tag{56}$$

holds for an $\eta \in (0, w)$ (mean value theorem).

Thus

$$F'(c) \ge w^c(1 + c\ \ln(b+\eta)) - w^c \ln w > w^c(1 + \ln(b+\eta) - \ln w) > w^c > 0. \tag{57}$$

The inequalities rely on the assumption $b \ge 1$, which means $\ln(b+\eta) > 0$, and the first inequality comes from $b \ge w$.

This completes the proof in this step.

Step 2. We then construct $k, \{t_i\}_{i=1}^k, \{f_i\}_{i=1}^k$ stated in the lemma.

Set $n = \lceil c \rceil$. By virtue of the increasing property of $F(c)$ when $\max\{b, w\} \ge 1$, we get

$$F_{b,w}(c) \le F_{b,w}(n) = \sum_{i=1}^{n-1} \binom{n}{i} b^{n-i} \cdot w^i. \tag{58}$$

When b, w satisfies $\max\{b, w\} < 1$,

$$F_{b,w}(c) \le (b+w)^c - b^c = c(b+\eta)^{c-1} w < c 2^{c-1} w \tag{59}$$

holds for an $\eta \in (0, w)$. The equality is an application of mean value theorem, and the second inequality comes from the fact that $\eta \in (0, w)$.

To conclude, $(b+w)^c - b^c - w^c \leq c2^{c-1}w + \sum_{i=1}^{n-1} \binom{n}{i} b^{n-i} \cdot w^i,\ \forall b, w \geq 0$. Since $c \notin \mathbb{N}$, which means $n-1 < c$, this completes the proof. □

Lemma C.2. *Suppose u^* is the exact solution to*

$$\begin{cases} \tilde{\mathcal{L}}_{\text{HJB}} u = h & (x,t) \in \mathbb{R}^n \times [0,T] \\ \tilde{\mathcal{B}}_{\text{HJB}} u = g \end{cases}.$$

Fix a bounded open set $\Omega \subset \mathbb{R}^n$. Suppose u_1 satisfies $\tilde{\mathcal{B}}_{\text{HJB}} u^* = \tilde{\mathcal{B}}_{\text{HJB}} u_1$ and that $\text{supp}(u_1 - u^*) \subset Q_T(\Omega)$. Recall c_i are parameters in the operator $\tilde{\mathcal{L}}_{\text{HJB}}$. Let $p \in [2, \infty)$.

If $p \geq n \cdot \max_{i \in [n]} \frac{c_i - 1}{c_i} = (1 - \bar{c}^{-1})n$, then there exists $\delta_0 > 0$ such that, when $\|\tilde{\mathcal{L}}_{\text{HJB}} u_1 - h\|_p < \delta_0$, we have $\|u^* - u_1\|_{2,p} \leq C \|\tilde{\mathcal{L}}_{\text{HJB}} u_1 - h\|_p$ for a constant C independent of u_1.

Proof. Define $w = w_{u_1} := u_1 - u^*$, then $\text{supp}(w)$ is compact and $w(x, 0) = 0$. We further define $f = f_{u_1} := \tilde{\mathcal{L}}_{\text{HJB}} u_1 - h$. Since $u_1 = u^*$ in $\mathbb{R}^n \times [0,T] \backslash Q_T(\Omega)$ and $f = \tilde{\mathcal{L}}_{\text{HJB}} u_1 - \tilde{\mathcal{L}}_{\text{HJB}} u^*$, we have $\text{supp}(f) \subset Q_T(\Omega)$. The $W^{m,p}$ and L^p norm in the rest of the proof is defined on the domain $Q_T(\Omega)$.

Compute

$$f = \tilde{\mathcal{L}}_{\text{HJB}} u_1 - \tilde{\mathcal{L}}_{\text{HJB}} u^* = \mathcal{L}_0 w + \sum_{i=1}^{n} A_i |\partial_i (u^* + w)|^{c_i} - A_i |\partial_i u^*|^{c_i} \tag{60}$$

Thus, for any (x, t),

$$\begin{aligned} |\mathcal{L}_0 w(x,t)| &= |f - \sum_{i=1}^{n} (A_i |\partial_i (u^* + w)|^{c_i} - A_i |\partial_i u^*|^{c_i})| \Big|_{(x,t)} && (61) \\ &\leq |f(x,t)| + \sum_{i=1}^{n} A_i \|\partial_i (u^* + w)|^{c_i} - |\partial_i u^*|^{c_i}| \Big|_{(x,t)} && (62) \\ &\leq |f(x,t)| + \sum_{i=1}^{n} A_i ((|\partial_i u^*| + |\partial_i w|)^{c_i} - |\partial_i u^*|^{c_i}) \Big|_{(x,t)} && (63) \end{aligned}$$

where the second inequality could be derived from the fact that $(a+b)^c - a^c \geq a^c - (a-b)^c$ for $a \geq b \geq 0$ and $c \geq 1$.

For $i \in [n]$, apply Lemma C.1 for $c = c_i$ and we obtain k_i and $\{t_{ij}\}_{j=1}^{k_i}$, $\{f_{ij}\}_{j=1}^{k_i}$ satisfying corresponding properties.

We have

$$|\mathcal{L}_0 w(x,t)| \leq |f(x,t)| + \sum_{i=1}^{n} A_i (|\partial_i w|^{c_i} + \sum_{j=1}^{k_i} f_{ij}(|\partial_i u^*|) |\partial_i w|^{t_{ij}}) \Big|_{(x,t)}. \tag{64}$$

With Lemma A.4 and triangle inequality, we obtain

$$\|w\|_{2,p} \leq C \|\mathcal{L}_0 w\|_p \leq C(\|f\|_p + \sum_{i=1}^{n} A_i (\||\partial_i w|^{c_i}\|_p + \sum_{j=1}^{k_i} \|f_{ij}(|\partial_i u^*|) |\partial_i w|^{t_{ij}}\|_p)). \tag{65}$$

We will handle each term respectively.

Using Lemma A.2, we have

$$\||\partial_i w|^{c_i}\|_p = \|\partial_i w\|_{c_i p}^{c_i} \leq \|w\|_{1, c_i p}^{c_i} \leq \hat{C}_i \|w\|_{2, \frac{n c_i p}{n + c_i p}}^{c_i}. \tag{66}$$

for constants $\hat{C}_i$.

Using Lemma A.2 and Hölder inequality, we have

$$\||f_{ij}(|\partial_i u^*|)|\partial_i w|^{t_{ij}}\|_p \leq \|f_{ij}(|\partial_i u^*|)\|_\infty \||\partial_i w|^{t_{ij}}\|_p \tag{67}$$

$$= \|f_{ij}(|\partial_i u^*|)\|_\infty \|\partial_i w\|_{t_{ij}p}^{t_{ij}} \leq \tilde{C}_{ij}\|w\|_{2,\frac{nt_{ij}p}{n+t_{ij}p}}^{t_{ij}}. \tag{68}$$

for constants $\tilde{C}_{ij}$ (Since $\overline{Q_T(\Omega)}$ is compact, we can tell that $\|f_{ij}(|\partial_i u^*|)\|_\infty < \infty$ and thus $\tilde{C}_{ij}$ are well-defined).

When $p \geq n \cdot \max_{i\in[n]} \frac{c_i-1}{c_i}$, because of $t_{ij} < c_i$, we have $\frac{nc_ip}{n+c_ip} \leq p$, $\frac{nt_{ij}p}{n+t_{ij}p} \leq p$ for all i, j.

Note that Ω is bounded, so for $1 \leq q < p$,there is a constant C such that $\|v\|_{L^q(\Omega)} \leq \|v\|_{L^p(\Omega)}$ for all $v \in L^p(\Omega)$.

As a consequence, we can derive from Eq. (65,66,68) that $M := \|w\|_{2,p}$ satisfies the inequality

$$K_0 M - \sum_{i=1}^{n}(K_i M^{c_i} + \sum_{1\leq j\leq k_i, t_{ij}>1} K_{ij} M^{t_{ij}}) \leq \|f\|_p, \tag{69}$$

where all K_i and K_{ij} are positive constants depending only on p, n, u^* and Ω.

For clarity, We define the L.H.S. of (69) as a function F with variable M.

With the observations (i)$F(0) = 0$, (ii) $F'(0) > 0$, (iii)$F''(M) < 0$, (iv)$F(+\infty) = -\infty$, we could tell that $F(M)$ has a unique zero m_0 in $\mathbb{R}^+$. We could further tell that for any non-negative number $C \leq \max_{M\in[0,m_0]} F(M)$, solving $F(M) \leq C$ $(M \geq 0)$ derives $M \in [0, a] \cap [b, \infty]$ for some $0 < a < b$ depending on C and that $a \to 0, b \to m_0$ monotonously as C decreases to 0. Note that there exists $\delta > 0$ such that $a \leq \frac{2}{K_0}C$ for $\forall C \in [0, \delta]$. In order to prove $\|w\|_{2,p} = O(\|f\|_p)$, it suffices to show that $\|w\|_{2,p}$ (i.e. M in the discussion above) would not fall in the second interval providing C (or $\|f\|_p$, correspondingly) is sufficiently small. We will prove by contradiction.

Note that $\tilde{\mathcal{L}}_{\text{HJB}}$ is a continuous injection from $W_0^{2,p}(Q_T(\Omega))$ to $L^p(Q_T(\Omega))$, and that there exists $r_0^* > 0$ such that $\tilde{\mathcal{L}}_{\text{HJB}}$ is a differmorphism from $B_{r_0}(u^*)$ (in $W_0^{2,p}(Q_T(\Omega))$) to $\tilde{\mathcal{L}}_{\text{HJB}}(B_{r_0}(u^*)) \supset B_{r_1}(h)$ (in $L^p(Q_T(\Omega))$) for any $r_0 \in (0, r_0^*)$ and any $r_1 \in (0, r_1^*)$ with r_1^* depending on r_0 (this comes from an application of Lemma A.6).

Select $r_0 < \frac{m_0}{2}$ and determine the corresponding r_1^*. By the property of b, there exists $\delta_0 \in (0, \min\{\frac{r_1^*}{2}, \delta\})$ such that for any $C < \delta_0$, the corresponding b is larger than $\frac{3}{4}m_0$. If there exists $w \in W_0^{m,p}(Q_T(\Omega))$ satisfying $\|f\|_p = \|\tilde{\mathcal{L}}_{\text{HJB}}(u^* + w) - \tilde{\mathcal{L}}_{\text{HJB}}u^*\|_p := C < \delta_0$ while $\|w\|_{2,p} > b$ (b depends on C), then there will also be a $w' \in B_{r_0}(0)$ such that $\tilde{\mathcal{L}}_{\text{HJB}}(u^* + w') - \tilde{\mathcal{L}}_{\text{HJB}}u^* = f$. We could tell from the difference between their norm that $w \neq w'$. This contradicts the property of injection.

The proof is completed. □

Lemma C.3. *Suppose u^* follows Lemma C.2, and Ω is a fixed bounded open set in $\mathbb{R}^n$. Suppose u_1 satisfies $\tilde{\mathcal{L}}_{\text{HJB}}u^* = \tilde{\mathcal{L}}_{\text{HJB}}u_1$ and that $\text{supp}(u_1 - u^*) \subset Q_T(\Omega)$. Let $q \in [1, \infty)$.*

If $\bar{c} \leq 2$ and $q > \frac{(\bar{c}-1)n^2}{(2-\bar{c})n+2}$, there exists $\delta_0 > 0$ such that when $\|\tilde{\mathcal{B}}_{\text{HJB}}u_1 - \tilde{\mathcal{B}}_{\text{HJB}}u^*\|_q < \delta_0$, we have $\|u^* - u_1\|_{1,r} \leq C\|\tilde{\mathcal{B}}_{\text{HJB}}u_1 - \tilde{\mathcal{B}}_{\text{HJB}}u^*\|_q$ for a constant C independent of u_1, where $r < \frac{(n+2)q}{n+q}$.

Proof. Define $w = w_{u_1} := u_1 - u^*$ and $f = f_{u_1} := \tilde{\mathcal{B}}_{\text{HJB}}u_1 - \tilde{\mathcal{B}}_{\text{HJB}}u^*$. Let w_1 be the solution to

$$\begin{cases} \mathcal{L}_0 u = 0, & (x,t) \in \mathbb{R}^n \times [0,T] \\ \tilde{\mathcal{B}}_{\text{HJB}}u = f. \end{cases}$$

The $W^{m,p}$ and L^p norm in the rest of the proof is defined on the domain $Q_T(\Omega)$.

Since the conditions $\bar{c} \leq 2$ and $q > \frac{(\bar{c}-1)n^2}{(2-\bar{c})n+2}$ hold, we have $[(\bar{c}-1)n, \frac{(n+2)q}{n+q}) \neq \varnothing$. Thus we could choose $r' \in [(\bar{c}-1)n, \frac{(n+2)q}{n+q}) \cap [r, \infty)$. Since $Q_T(\Omega)$ is bounded, it suffices to bound $\|u_1 - u^*\|_{r'}$.

To start with, from Lemma A.5, we have $\|w_1\|_{1,r'} \leq C\|f\|_q$.

Then we bound the difference between w_1 and w. Define $v = w_1 - w$, then v satisfies

$$\begin{cases} \mathcal{L}_0 v = \sum_{i=1}^n A_i|\partial_i(u^* + w)|^{c_i} - A_i|\partial_i u^*|^{c_i} \\ \tilde{\mathcal{B}}_{\mathrm{HJB}} v = 0. \end{cases}$$

By Lemma A.2 and Lemma A.4 , we get $\|v\|_{1,r'} \leq C\|v\|_{2,\frac{nr'}{n+r'}} \leq C'\|\mathcal{L}_0 v\|_{\frac{nr'}{n+r'}}$.

Therefore we have $\|w\|_{1,r'} \leq \|w_1\|_{1,r'} + \|v\|_{1,r'} \leq C\|f\|_q + C'\|\mathcal{L}_0 v\|_{\frac{nr'}{n+r'}}$.

Next, we give an estimation for $\|\mathcal{L}_0 v\|_{\frac{nr'}{n+r'}}$.

Following from the proof in Lemma C.2, we obtain $\{k_i\}_{i=1}^n \subset \mathbb{N}$, $\{t_{ij}\}_{1\leq i\leq n, 1\leq j\leq k_i} \subset \mathbb{R}$, and power functions $\{f_{ij}\}_{1\leq i\leq n, 1\leq j\leq k_i}$.

With similar computations, we have

$$\|\mathcal{L}_0 v\|_{\frac{nr'}{n+r'}} \leq \sum_{i=1}^n A_i(\||\partial_i w|^{c_i}\|_{\frac{nr'}{n+r'}} + \sum_{j=1}^{k_i} \|f_{ij}(|\partial_i u^*|)|\partial_i w|^{t_{ij}}\|_{\frac{nr'}{n+r'}}) \tag{70}$$

$$\leq \sum_{i=1}^n A_i(\|\partial_i w\|^{c_i}_{\frac{c_i nr'}{n+r'}} + \sum_{j=1}^{k_i} \|f_{ij}(|\partial_i u^*|)\|_\infty \|\partial_i w\|^{t_{ij}}_{\frac{nt_{ij}r'}{n+r'}}). \tag{71}$$

Because of $r' \geq (\bar{c}-1)n$ and $t_{ij} < c_i$, we have $\frac{nc_i r'}{n+r'} \leq r'$, $\frac{nt_{ij}r'}{n+r'} \leq r'$ for all i, j.

Thus, due to the fact that $Q_T(\Omega)$ is bounded, all $\|\partial_i w\|_{\frac{c_i nr'}{n+r'}}$ and $\|\partial_i w\|_{\frac{nt_{ij}r'}{n+r'}}$ could be bounded by $C_{i,j}\|w\|_{1,r'}$, where $C_{i,j}$ are constants.

With similar methods applied in Lemma C.2, we could prove this lemma. □

Lemma C.4. *We denote by u^* the exact solution to*

$$\begin{cases} \tilde{\mathcal{L}}_{\mathrm{HJB}} u(x,t) = h(x,t) & (x,t) \in \mathbb{R}^n \times [0,T], \\ \tilde{\mathcal{B}}_{\mathrm{HJB}} u(x,t) = g(x) & x \in \mathbb{R}^n. \end{cases}$$

Fix Ω, which is an arbitrary bounded open set in $\mathbb{R}^n$. For two functions $\hat{f}_1(x,t)$, $\hat{f}_2(x)$, denote by u_1 the solution to

$$\begin{cases} \tilde{\mathcal{L}}_{\mathrm{HJB}} u(x,t) = h(x,t) + \hat{f}_1(x,t), \ in\ \mathbb{R}^n \times [0,T] \\ \tilde{\mathcal{B}}_{\mathrm{HJB}} u(x,t) = g(x) + \hat{f}_2(x), \ in\ \mathbb{R}^n. \end{cases}$$

For $p, q \geq 1$, let $r_0 = \frac{(n+2)q}{n+q}$. Assume the following inequalities hold for p, q and r_0:

$$p \geq \max\left\{2, \left(1 - \frac{1}{\bar{c}}\right)n\right\};\ q > \frac{(\bar{c}-1)n^2}{(2-\bar{c})n+2};\ \frac{1}{r_0} \geq \frac{1}{p} - \frac{1}{n}, \tag{72}$$

Further assume that $\bar{c} \leq 2$ and $\mathrm{supp}(u_1 - u^*) \subset Q_T(\Omega)$.

Then for $\forall r \in [1, r_0)$, there exists $\delta_0 > 0$ such that, when $\|\hat{f}_1\|_p < \delta_0$ and $\|\hat{f}_2\|_q < \delta_0$, $\|u_1 - u^*\|_{1,r} \leq C(\|\hat{f}_1\|_p + \|\hat{f}_2\|_q)$ for a constant C independent of u_1.

Proof. It is straight-forward to define u_2 as the solution to

$$\begin{cases} \tilde{\mathcal{L}}_{\mathrm{HJB}} u(x,t) = h(x,t) + \hat{f}_1(x,t) & (x,t) \in \mathbb{R}^n \times [0,T] \\ \tilde{\mathcal{B}}_{\mathrm{HJB}} u(x,t) = g(x) & x \in \mathbb{R}^n \end{cases}$$

and bound $\|u^* - u_2\|_{1,r}$, $\|u_2 - u_1\|_{1,r}$ respectively.

From Lemma C.2, there exists $\delta_1 > 0, C_1 > 0$ such that $\|\hat{f}_1\|_p < \delta_1$ implies $\|u^* - u_2\|_{2,p} \leq C_1\|\hat{f}_1\|_p$. And from Lemma C.3, there exists $\delta_2 > 0, C_2 > 0$ such that $\|\hat{f}_2\|_q < \delta_2$ implies $\|u_2 - u_1\|_{1,r} \leq C_2\|\hat{f}_2\|_q$. By virtue of the condition $\frac{1}{r} > \frac{1}{r_0} \geq \frac{1}{p} - \frac{1}{n}$, with Lemma A.2 and the fact that we are considering $\|u^* - u_2\|_{1,r},\ \|u_2 - u_1\|_{1,r}$ on a compact domain, providing $\|\hat{f}_1\| < \delta_1$ and $\|\hat{f}_2\| < \delta_2$, we derive

$$\|u^* - u_1\|_{1,r} \leq \|u^* - u_2\|_{1,r} + \|u_2 - u_1\|_{1,r} \tag{73}$$
$$\leq C\|u^* - u_2\|_{2,p} + \|u_2 - u_1\|_{1,r} \tag{74}$$
$$\leq CC_1\|\hat{f}_1\|_p + C_2\|\hat{f}_2\|_q, \tag{75}$$

where C is a constant.

This concludes the proof. □

Finally, we give the proof of an equivalent statement of Theorem 4.3.

Theorem C.5. *Let $\hat{f}_1, \hat{f}_2, u^*$ and u_1 follow from Lemma C.4. Let p, q, r_0 satisfy the conditions in Lemma C.4. Assume $\bar{c} \leq 2$. For any bounded open set $Q \subset \mathbb{R}^n \times [0, T]$, it holds that for any $r \in [1, r_0)$, there exists $\delta > 0$ and a constant C independent of $u_1,\ \hat{f}_1$ and $\hat{f}_2$, such that $\max\{\|\hat{f}_1\|_{L^p(\mathbb{R}^n\times[0,T])}, \|\hat{f}_2\|_{L^q(\mathbb{R}^n)}\} < \delta$ implies $\|u_1 - u^*\|_{W^{1,r}(Q)} \leq C(\|\hat{f}_1\|_{L^p(\mathbb{R}^n\times[0,T])} + \|\hat{f}_2\|_{L^q(\mathbb{R}^n)})$.*

Proof. Since Q is bounded, there exists $R > 0$ such that $Q \subset Q(R)$.

Let $\hat{u}_1$ be the constraint of u_1 in $Q(R)$. Construct an extension v of $\hat{u}_1$ to $\mathbb{R}^n \times \mathbb{R}_{\geq 0}$ such that

(i) $v = u^*$ in $(\mathbb{R}^n \times \mathbb{R}_{\geq 0})\backslash Q(2R)$,

(ii) $\|\tilde{f}_1\|_p \leq C'\|\hat{f}_1\|_{L^p(Q)}$ and $\|\tilde{f}_2\|_{\leq} C'\|\hat{f}_2\|_{L^q(B_R(0))}$ for a constant C' depending only on n, R, p, q, Q, where $\tilde{f}_1 := \tilde{\mathcal{L}}_{\mathrm{HJB}}v - \hat{f}_1,\ \tilde{f}_2 := \tilde{\mathcal{B}}_{\mathrm{HJB}}v - \hat{f}_2$.

Note that $\mathrm{supp}(\tilde{f}_1) \subset \overline{Q(2R)}$ and $\mathrm{supp}(\tilde{f}_2) \subset B_{2R}(0)$, the existence of v is obvious.

From Lemma C.4, there exists $C > 0$ and $\delta > 0$ such that $\|\tilde{f}_1\|_p < \delta$ and $\|\tilde{f}_2\|_q < \delta$ imply $\|v - u^*\|_{1,r} \leq C(\|\tilde{f}_1\|_p + \|\tilde{f}_2\|_q)$. Thus

$$\|u_1 - u^*\|_{W^{1,r}(Q)} = \|v - u^*\|_{W^{1,r}(Q)} \leq \|v - u^*\|_{1,r} \leq C(\|\tilde{f}_1\|_p + \|\tilde{f}_2\|_q) \tag{76}$$
$$\leq CC'(\|\hat{f}_1\|_{L^p(Q)} + \|\hat{f}_2\|_{L^q(B_R(0))}) \leq CC'(\|\hat{f}_1\|_p + \|\hat{f}_2\|_q). \tag{77}$$

The proof is completed. □

D Proof of Theorem 4.4

In this section, we give the proof of Theorem 4.4.

Based on remark B.1, it is equivalent to consider Eq. (53). We will show that the following equation satisfies the properties stated in Theorem 4.4,

$$\begin{cases} \partial_t u - \Delta u + |Du|^2 = 0 \ \ in\ \mathbb{R}^n \times [0, T] \\ u(x, 0) = g(x). \end{cases} \tag{78}$$

This equation is a special case for Eq. (53) with $A_i = 1,\ c_i = 2,\ \forall i \in [n]$ and $h(x, t) \equiv 0$.

Denote by u^* the exact solution to the equation above. The notations $\mathcal{L}_0, \tilde{\mathcal{L}}_{\mathrm{HJB}}, \tilde{\mathcal{B}}_{\mathrm{HJB}}$ in the following discussion have the same meaning as in section C.

We prove some auxiliary results first.

Lemma D.1. *For $p \in [2, 2n)$, and an open set or a parabolic region $\mathfrak{A}$, we denote the function space $W^{2,p}(\mathfrak{A})$ as X and $L^{\frac{np}{2n-p}}(\mathfrak{A})$ as Y. For any $u \in X$, we have $\|u - u'\|_X \geq A\sqrt{\|\tilde{\mathcal{L}}_{\mathrm{HJB}}u - \tilde{\mathcal{L}}_{\mathrm{HJB}}u'\|_Y + B} - C$ holds for $\forall u' \in X$, where A, B, C are positive constants depending on u.*

Proof. We divide the proof into two steps.

Step 1. We check that $\tilde{\mathcal{L}}_{\mathrm{HJB}}$ as an operator mapping from X to Y is Fréchet-differentiable.

For any $u, v \in X,\ t \in \mathbb{R}$, since $v \in X$, which means $|Dv|^2 \in W^{1,\frac{p}{2}} \subset Y$, we have

$$\|\tilde{\mathcal{L}}_{\mathrm{HJB}}(u + tv) - \tilde{\mathcal{L}}_{\mathrm{HJB}}u - t \cdot (\mathcal{L}_0 v + 2Du \cdot Dv)\|_Y = \|t^2|Dv|^2\|_Y = o(t) \ \ (t \to 0). \tag{79}$$

Therefore, $d\tilde{\mathcal{L}}_{\mathrm{HJB}}(u, v) = \mathcal{L}_0 v + 2Du \cdot Dv$ by definition.

Define operator $A(u) \in \mathscr{L}(X, Y)$ as $A(u)v = dL(u, v)$. For $u, u' \in X$ and any $v \in X$, note that

$$\|A(u')v - A(u)v\|_Y = 2\|D(u' - u) \cdot Dv\|_Y \leq C_0\|D(u' - u) \cdot Dv\|_{1,\frac{p}{2}} \tag{80}$$

$$\leq C_0\|D(u' - u)\|_{1,p}\|Dv\|_{1,p} \leq C_0\|u' - u\|_X\|v\|_X \tag{81}$$

for a constant C_0, where the second inequality comes from Cauchy-Schwarz inequality. Thus we have $\|A(u) - A(u')\|_{\mathscr{L}(X,Y)} \leq C_0\|u - u'\|_X$, which means A is continuous with regard to u. As a result, $\tilde{\mathcal{L}}_{\mathrm{HJB}}$ is Fréchet-differentiable and $\tilde{\mathcal{L}}'_{\mathrm{HJB}}(u) = A(u)$. Moreover, we derive $\|L'(u)\| \leq \|L'(0)\| + \|L'(0) - L'(u)\| \leq C_0\|u\|_X + C_1$.

Step 2. For any $u, u' \in X$, let $y = \tilde{\mathcal{L}}_{\mathrm{HJB}}u,\ y' = \tilde{\mathcal{L}}_{\mathrm{HJB}}u'$.

Define $u_\eta = (1 - \eta)u + \eta u'$ and $y_\eta = \tilde{\mathcal{L}}_{\mathrm{HJB}}u_\eta$ for $\eta \in [0, 1]$. Fix a number $m \in \mathbb{N}$.

From the property proved in step 1, for any $\eta \in [0, 1]$, there exists $r_\eta \in (0, \frac{1}{m})$ such that

$$\|\tilde{\mathcal{L}}_{\mathrm{HJB}}v - \tilde{\mathcal{L}}_{\mathrm{HJB}}u_\eta\| \leq 2\|L'(u_\eta)\| \cdot \|u_\eta - v\| \leq 2(C_0\|u_\eta\| + C_1)\|v - u_\eta\|, \ \ \forall v \in B_{r_\eta}(u_\eta). \tag{82}$$

Note that $\{B_{\frac{r_\eta}{2}}(u_\eta) : \eta \in [0, 1]\}$ is an open cover of $\{u_\eta : \eta \in [0, 1]\}$. Because of the compactness of $\{u_\eta : \eta \in [0, 1]\}$, we obtain an increasing finite sequence $\{\eta_i\}_{i=0}^N$ with $\eta_0 = 0, \eta_N = 1$ such that either $u_{\eta_i} \in B_{r_{\eta_{i-1}}}(u_{\eta_{i-1}})$ or $u_{\eta_{i-1}} \in B_{r_{\eta_i}}(u_{\eta_i})$ for $\forall i \in [N]$. This means that

$$\|y_{\eta_i} - y_{\eta_{i-1}}\| \leq 2(C_0\|u_{\eta_j}\| + C_1)\|u_{\eta_i} - u_{\eta_{i-1}}\|, \ j \in \{i - 1, i\} \tag{83}$$

holds for $\forall i \in [N]$.

Therefore

$$\|y' - y\| \leq \sum_{i=1}^N \|y_{\eta_i} - y_{\eta_{i-1}}\| \leq \sum_{i=1}^N 2(C_0\|u_{\eta_j}\| + C_1)\|u_{\eta_i} - u_{\eta_{i-1}}\| \tag{84}$$

Note that this inequality holds for every m. As $m \to \infty$, R.H.S. of (84) converges to

$$\|u - u'\| \int_0^1 2(C_0\|u + s(u' - u)\| + C_1)ds. \tag{85}$$

$$= \|u - u'\|(C_0\|u + \theta(u' - u)\| + C_1), \ \ (\theta \in (0, 1)) \tag{86}$$

$$\leq \|u - u'\|(C_0(\|u\| + \|u' - u\|) + C_1). \tag{87}$$

The equality comes from mean value theorem for integral, and the inequality comes from triangular inequality.

Combining Eq. (84) and (87) together completes the proof. □

Lemma D.2. *Let u_1 satisfies $\tilde{\mathcal{B}}_{\mathrm{HJB}}u^* = \tilde{\mathcal{B}}_{\mathrm{HJB}}u_1$ and that $\mathrm{supp}(u_1 - u^*)$ is compact in $\mathbb{R}^n \times \mathbb{R}^+$. Define $f(x, t) := \tilde{\mathcal{L}}_{\mathrm{HJB}}u_1$. Let $p \in [2, \infty), m \in \mathbb{N}$. If $p \geq \frac{n}{2}$ then there exists $\delta_0 > 0$ such that $\|f\|_{m,p} < \delta_0$ implies $\|u^* - u_1\|_{m+2,p} = O(\|f\|_{m,p})$.*

Proof. When $m = 0$, this statement is a direct consequence of Lemma C.2.

When $m > 0$, for every multi-index α with $|\alpha| \leq m$, operate D^α on both sides of $\tilde{\mathcal{L}}_{\text{HJB}} u_1 = f$ and $\tilde{\mathcal{L}}_{\text{HJB}} u^* = 0$. We then obtain

$$\mathcal{L}_0 D^\alpha u_1 + D^\alpha |Du_1|^2 = D^\alpha f \tag{88}$$

$$\mathcal{L}_0 D^\alpha u^* + D^\alpha |Du^*|^2 = 0. \tag{89}$$

Define $w := u_1 - u^*$, and compute the difference between (88) and (89), we get

$$\mathcal{L}_0 D^\alpha w = D^\alpha f - \sum_{i=1}^{n} (2D^\alpha(\partial_i u^* \partial_i w) + D^\alpha(\partial_i w)^2). \tag{90}$$

With similar methods used in Lemma C.2, we could bound $\|D^\alpha w\|_{2,p}$ with $\|D^\alpha f\|_p$, based on which we complete the proof. □

Finally, we show that Eq. (78) satisfies the properties stated in Theorem 4.4, which will conclude the proof for Theorem 4.4.

Theorem D.3. *For any $\varepsilon > 0, A > 0, r \geq 1, m \in \mathbb{N}$ and $p \in \left[1, \frac{n}{4}\right]$, there exists a function $u \in C^\infty(\mathbb{R}^n \times (0, T])$ which satisfies the following conditions:*

- $\|\tilde{\mathcal{L}}_{\text{HJB}} u\|_{L^p(\mathbb{R}^n \times [0,T])} < \varepsilon$, $\tilde{\mathcal{B}}_{\text{HJB}} u = \tilde{\mathcal{B}}_{\text{HJB}} u^*$, *and* $\text{supp}(u - u^*)$ *is compact.*
- $\|u - u^*\|_{W^{m,r}(\mathbb{R}^n \times [0,T])} > A$.

Proof. Since $L^1(\Omega)$ has the weakest topology in function spaces $W^{m,r}(\Omega)$ when Ω is bounded, it is enough to consider the case for $r = 1,\ m = 0$.

Set $p_0 = \frac{5}{9}n,\ p_1 = \frac{11}{9}n$ and $p_2 = \frac{11}{7}n$.

Step 1. We construct two families of functions $\{v_{a,c}\}, \{F_{a,c}\}$ as the basis of proof.

For any $a, c > 0$, define $f_{a,c}(x) = c|x|^{-0.7}, x \in \overline{B_1(0)} \backslash B_a(0)$ in $\mathbb{R}^n$. We could extend it to a C^∞ function $\tilde{f}_{a,c}(x)$ defined on $\mathbb{R}^n$ such that (i)$\|\tilde{f}_{a,c}\|_\infty < \|f_{a,c}\|_\infty + \min\{1, c\}$ and (ii)$\text{supp}(\tilde{f}_{a,c}) \subset B_{1.1}(0)$.

We could further construct a C^∞ function $\hat{f}_{a,c}(x,t)$ such that

(i) $\text{supp}(\hat{f}_{a,c}) \subset B_{1.1}(0) \times (0, T]$,

(ii) $\hat{f}_{a,c}(x,t) = \tilde{f}_{a,c}(x), \forall t \in [\frac{T}{2}, T]$,

(iii) $\|\hat{f}_{a,c}(x,t)\|_{L^\infty(\mathbb{R}^n) \times [0,T]} \leq \|\tilde{f}_{a,c}\|_{L^\infty(\mathbb{R}^n)}$.

Define $u_{a,c}$ as the solution to

$$\begin{cases} \tilde{\mathcal{L}}_{\text{HJB}} u = \hat{f}_{a,c} \ \ in\, \mathbb{R}^n \times [0, T] \\ \tilde{\mathcal{B}}_{\text{HJB}} u = g. \end{cases}$$

Select a function $w_{a,c} \in C^\infty(\mathbb{R}^n \times \mathbb{R})$ with compact support $\text{supp}(w_{a,c}) \subset \mathbb{R}^n \times (0, T]$, such that $\|u_{a,c} - u^* - w_{a,c}\| < \epsilon_c$ in $W^{3,4p_0}(\mathbb{R}^n \times [0, T]), W^{2,4p_1}(\mathbb{R}^n \times [0, T])$ and $W^{2,4p}(\mathbb{R}^n \times [0, T])$, where ϵ_c is an small value depending on c and is to be decided later.

We define $v_{a,c} = u^* + w_{a,c}$ and $F_{a,c} = \tilde{\mathcal{L}}_{\text{HJB}} v_{a,c}$.

Step 2. We show that$\{v_{a,c}\}$ and $\{F_{a,c}\}$ have following properties:

(i) $\text{supp}(v_{a,c} - u^*)$ is compact in $\mathbb{R}^n \times (0, T]$.

(ii) $\tilde{\mathcal{B}}_{\mathrm{HJB}} v_{a,c} = \tilde{\mathcal{B}}_{\mathrm{HJB}} u^*$.

(iii) $\mathrm{supp}(F_{a,c})$ is compact in $\mathbb{R}^n \times (0,T]$.

(iv) There exists a constant $M < \infty$ such that $\|F_{a,c}\|_q < cM$ and $\|F_{a,c}\|_{1,p_0} < cM$.

(v) For any $c > 0$, $\|F_{a,c}\|_{p_2} \to \infty$ as $a \to 0$.

(i) and (ii) comes directly from the construction of $v_{a,c}$.

Because $\mathrm{supp}(v_{a,c} - u^*)$ is close, for any $(x,t) \in (\mathbb{R}^n \times [0,T]) \backslash \mathrm{supp}(v_{a,c} - u^*)$, there exists $r > 0$ such that $(B_r(x,t) \cap (\mathbb{R}^n \times [0,T])) \subset (\mathbb{R}^n \times [0,T]) \backslash \mathrm{supp}(v_{a,c} - u^*)$, which means $v_{a,c} = u^*$ in $B_r(x,t) \cap (\mathbb{R}^n \times [0,T])$ and thus $\tilde{\mathcal{L}}_{\mathrm{HJB}} v_{a,c} = \tilde{\mathcal{L}}_{\mathrm{HJB}} u^* = 0$. This gives (iii).

Due to the fact that the function $|x|^{-0.7} \in L^p(B_2(0)) \cap W^{1,p_0}(B_2(0))$, there exists a constant $M < \infty$ such that $\|\hat{f}_{a,1}\|_p < M - 1$ and $\|\hat{f}_{a,1}\|_{1,p_0} < M - 1$ holds for any a and any construction of $\hat{f}_{a,1}$ based on $f_{a,1}$. Due to the linearity of norms, we derive $\|\hat{f}_{a,c}\|_p < c(M-1)$ and $\|\hat{f}_{a,c}\|_{1,p_0} < c(M-1)$.

It is easy to check that $\tilde{\mathcal{L}}_{\mathrm{HJB}}$ is a continuous mapping from $W^{3,4p_0}(\Omega)$ to $W^{1,p_0}(\Omega)$, from $W^{2,4p_1}(\Omega)$ to $L^{p_2}(\Omega)$ and from $W^{2,4p}(\Omega)$ to $L^p(\Omega)$ for any compact set $\Omega \subset \mathbb{R}^n \times [0,T]$. Therefore, $\|F_{a,c} - \hat{f}_{a,c}\|$ is small in $W^{1,p_0}(\mathbb{R}^n \times [0,T]),\ L^{p_2}(\mathbb{R}^n \times [0,T]),$ and $L^p(\mathbb{R}^n \times [0,T])$.

Since $\||x|^{-0.7}\|_{L^{p_2}(B_1(0))} = +\infty$, by the construction of $\hat{f}_{a,c}$ we have $\|\hat{f}_{a,c}\|_{p_2} \to +\infty$ as $a \to 0$.

As a result of the continuity of $\tilde{\mathcal{L}}_{\mathrm{HJB}}$, we could guarantee $\|F_{a,c}\|_p < cM$, $\|F_{a,c}\|_{1,p_0} < cM$ and $\|F_{a,c}\|_{p_2} > \frac{1}{2}\|\hat{f}_{a,c}\|_{p_2}$ by choosing ϵ_c sufficiently small previously. This gives (iv) and (v).

Step 3. We give an estimation for $\|v_{a,c} - u^*\|_1$, i.e., $\|w_{a,c}\|_1$).

We mention at the beginning of this part that all C_i appeared below are positive constants.

For any $\epsilon > 0$, set c to $\frac{1}{2M}\min\{\epsilon, \delta_0\}$, where δ_0 follows from an application of Lemma D.2 for the case $p = p_0$. Then for any $a > 0$, $\|F_{a,c}\|_p < \epsilon$ and we obtain $\|w_{a,c}\|_{3,p_0} \le C_0 \|F_{a,c}\|_{1,p_0}$.

In Lemma A.3, we choose $j = 2,\ k = 3,\ \theta = \frac{n+\frac{13}{11}}{n+\frac{6}{5}},\ r = p_0,\ q = 1,\ p = \frac{11}{9}n$ and derive $\|\nabla^2 w_{a,c}\|_{p_1} \le C_1 \|\nabla^3 w_{a,c}\|_{p_0}^{\theta} \|w_{a,c}\|_1^{1-\theta}$. Since $\|\nabla^3 w_{a,c}\|_{p_0} \le \|w_{a,c}\|_{3,p_0}$, we get

$$\|w_{a,c}\|_1 \ge \left(\frac{\|\nabla^2 w_{a,c}\|_{p_1}}{C_1 \|w_{a,c}\|_{3,p_0}^{\theta}}\right)^{\frac{1}{1-\theta}} \ge C_2 \left(\frac{\|\nabla^2 w_{a,c}\|_{p_1}}{\|F_{a,c}\|_{1,p_0}^{\theta}}\right)^{\frac{1}{1-\theta}} \ge \frac{C_2}{(C_3 M)^{\frac{1-\theta}{\theta}}} \|\nabla^2 w_{a,c}\|_{p_1}^{\frac{1}{1-\theta}}, \quad (91)$$

where the last inequality comes from property (iv) in Step 2.

By virtue of property (i) in Step 2, we have $\|\nabla^2 w_{a,c}\|_{p_1} \ge C_4 \|w_{a,c}\|_{2,p_1}$, which is an application of Poincaré inequality. Together with Lemma D.1,

$$\|w_{a,c}\|_1 \ge \frac{C_2}{(C_3 M)^{\frac{1-\theta}{\theta}}} \|\nabla^2 w_{a,c}\|_{p_1}^{\frac{1}{1-\theta}} \ge C_5 \|w_{a,c}\|_{2,p_1}^{\frac{1}{1-\theta}} \ge C_5 \left(C_6 \sqrt{\|F_{a,c}\|_{p_2}} - C_7\right)^{\frac{1}{1-\theta}}. \quad (92)$$

Since R.H.S. above goes to $+\infty$ as $a \to 0$ due to property (v), for any $A > 0$, there exists $a_0 > 0$ such that $\|w_{a_0,c}\|_1 > A$.

Setting $u = v_{a_0,c} = u^* + w_{a_0,c}$ completes the proof. □

E Improved Theorem 4.3

In this section, we give the stability result for Eq. (53) (Theorem E.3). Different from Theorem C.5, the constraint $\bar{c} \le 2$ is released here.

The notations $u^*,\ \bar{c},\ \mathcal{L}_0, \tilde{\mathcal{L}}_{\mathrm{HJB}}, \tilde{\mathcal{B}}_{\mathrm{HJB}}$ in the following discussion come from C.

The proof of Theorem E.3 is quite similar to that of Theorem C.5.

We begin with some auxiliary results.

Lemma E.1. *Suppose Ω is a fixed bounded open set in $\mathbb{R}^n$. Suppose u_1 satisfies $\tilde{\mathcal{L}}_{\mathrm{HJB}} u^* = \tilde{\mathcal{L}}_{\mathrm{HJB}} u_1$ and that* $\mathrm{supp}(u_1 - u^*) \subset Q_T(\Omega)$. *Let $q \in [1, \infty)$.*

If $q > \frac{(\bar{c}-1)n^2}{n+2\bar{c}}$, then there exists $\delta_0 > 0$ such that when $\|\tilde{\mathcal{B}}_{\mathrm{HJB}} u_1 - \tilde{\mathcal{B}}_{\mathrm{HJB}} u^*\|_{1,q} < \delta_0$, we have $\|u^* - u_1\|_{2,r} \leq C\|\tilde{\mathcal{B}}_{\mathrm{HJB}} u_1 - \tilde{\mathcal{B}}_{\mathrm{HJB}} u^*\|_{1,q}$ for a constant C independent of u_1, where $r < \frac{(n+2)q}{n+q}$.

Proof. The proof is almost the same as that for Lemma C.3.

Following its proof, we define $w,\ f,\ w_1$ and v similarly. And the $W^{m,p}$ and L^p norm in the rest of the proof will also be defined on the domain $Q_T(\Omega)$.

Since $q > \frac{(\bar{c}-1)n^2}{n+2\bar{c}}$, we have $[(1-\bar{c}^{-1})n, \frac{(n+2)q}{n+q}) \neq \varnothing$.

Thus we could choose $r' \in [(1-\bar{c}^{-1})n, \frac{(n+2)q}{n+q}) \cap [r, \infty)$.

Since $Q_T(\Omega)$ is bounded, it suffices to bound $\|u_1 - u^*\|_{r'}$.

From Lemma A.5, we have $\|w_1\|_{2,r'} \leq C\|f\|_{1,q}$. And from Lemma A.4 , we get $\|v\|_{2,r'} \leq C'\|\mathcal{L}_0 v\|_{r'}$.

Therefore we have $\|w\|_{2,r'} \leq \|w_1\|_{2,r'} + \|v\|_{2,r'} \leq C\|f\|_{1,q} + C'\|\mathcal{L}_0 v\|_{r'}$.

Next, we give an estimation for $\|\mathcal{L}_0 v\|_{r'}$.

Following from the proof in Lemma C.2, we obtain $\{k_i\}_{i=1}^n \subset \mathbb{N}$, $\{t_{ij}\}_{1\leq i\leq n, 1\leq j\leq k_i} \subset \mathbb{R}$, and power functions $\{f_{ij}\}_{1\leq i\leq n, 1\leq j\leq k_i}$.

With similar computations, we have

$$\|\mathcal{L}_0 v\|_{r'} \leq \sum_{i=1}^n A_i(\||\partial_i w|^{c_i}\|_{r'} + \sum_{j=1}^{k_i} \|f_{ij}(|\partial_i u^*|)|\partial_i w|^{t_{ij}}\|_{r'}) \tag{93}$$

$$\leq \sum_{i=1}^n A_i(\|\partial_i w\|_{c_i r'}^{c_i} + \sum_{j=1}^{k_i} \|f_{ij}(|\partial_i u^*|)\|_\infty \|\partial_i w\|_{t_{ij} r'}^{t_{ij}}). \tag{94}$$

Due to the fact that $Q_T(\Omega)$ is bounded, all $\|\partial_i w\|_{c_i r'}$ and $\|\partial_i w\|_{t_{ij} r'}$ could be bounded by $C_{i,j}\|w\|_{1,\bar{c}r'}$, where $C_{i,j}$ are constants.

Moreover, since $r' \geq (1-\bar{c}^{-1})n$, we could tell from Lemma A.2 that

$$\|w\|_{1,\bar{c}r'} \leq \hat{C}\|w\|_{2,\frac{n\bar{c}r'}{n+\bar{c}r'}} \leq \hat{C}'\|w\|_{2,r'}, \tag{95}$$

where $\hat{C}$ and $\hat{C}'$ are constants.

With similar methods applied in Lemma C.2, we could prove this lemma. □

Lemma E.2. *Fix Ω, which is an arbitrary bounded open set in $\mathbb{R}^n$. For two functions $\hat{f}_1(x,t),\ \hat{f}_2(x)$, denote by u_1 the solution to*

$$\begin{cases} \tilde{\mathcal{L}}_{\mathrm{HJB}} u(x,t) = h(x,t) + \hat{f}_1(x,t),\ in\ \mathbb{R}^n \times [0,T] \\ \tilde{\mathcal{B}}_{\mathrm{HJB}} u(x,t) = g(x) + \hat{f}_2(x),\ in\ \mathbb{R}^n. \end{cases}$$

For $p, q \geq 1$, let $r_0 = \frac{(n+2)q}{n+q}$. Assume the following inequalities hold for p, q and r_0:

$$p \geq \max\left\{2, \left(1 - \frac{1}{\bar{c}}\right)n\right\};\ q > \frac{(\bar{c}-1)n^2}{n+2\bar{c}}. \tag{96}$$

Further assume $\mathrm{supp}(u_1 - u^*) \subset Q_T(\Omega)$.

Then for $\forall r \in [1, \min\{r_0, p\})$, there exists $\delta_0 > 0$ such that, when $\|\hat{f}_1\|_p < \delta_0$ and $\|\hat{f}_2\|_q < \delta_0$, $\|u_1 - u^*\|_{2,r} \leq C(\|\hat{f}_1\|_p + \|\hat{f}_2\|_{1,q})$ for a constant C independent of u_1.

Proof. The proof follows as in Lemma C.4 by replacing the use of Lemma C.3 with Lemma E.1. □

Theorem E.3. *Let $\hat{f}_1, \hat{f}_2$ and u_1 follow from Lemma E.2. For $p, q, r \geq 1$, let $r_0 = \frac{(n+2)q}{n+q}$. Assume the following inequalities hold for p, q, r and r_0:*

$$p \geq \max\left\{2, \left(1 - \frac{1}{\bar{c}}\right) n\right\};\ q > \frac{(\bar{c}-1)n^2}{n+2\bar{c}};\ \frac{1}{r} > \frac{1}{\min\{r_0, p\}} - \frac{1}{n}. \tag{97}$$

Then for any bounded open set $Q \subset \mathbb{R}^n \times [0,T]$, there exists $\delta > 0$ and a constant C independent of $u_1,\ \hat{f}_1$ and $\hat{f}_2$, such that $\max\{\|\hat{f}_1\|_{L^p(\mathbb{R}^n\times[0,T])}, \|\hat{f}_2\|_{W^{1,q}(\mathbb{R}^n)}\} < \delta$ implies $\|u_1 - u^*\|_{W^{1,r}(Q)} \leq C(\|\hat{f}_1\|_{L^p(\mathbb{R}^n\times[0,T])} + \|\hat{f}_2\|_{W^{1,q}(\mathbb{R}^n)})$.

Proof. By replacing the use of Lemma C.4 in the proof for Theorem C.5 with Lemma E.2, we can bound $\|u_1 - u^*\|_{W^{2,r'}(Q)}$ with $\|\hat{f}_1\|_{L^p(\mathbb{R}^n\times[0,T])}$ and $\|\hat{f}_2\|_{W^{1,q}(\mathbb{R}^n)}$ for any $r' \in [1, \min\{r_0, p\})$.

We could further bound $\|u_1 - u^*\|_{W^{1,r}(Q)}$ with the help of Lemma A.2.

This concludes the proof. □

F Experimental Settings

Hyperparameters. The hyperparameters used in our experiment is described in Table 3.

Table 3: **Derailed experimental settings** of Section 6.

	$n = 100$	$n = 250$
Model Configuration		
Layers	4	
Hidden dimension	4096	
Activation	tanh	
Hyperparameters		
Toal iterations	5000	10000
Domain Batch Size N_1	100	50
Boundary Batch Size N_2	100	50
Inner Loop Iterations K	20	
Inner Loop Step Size η	0.05	
Learning Rate	$7\mathrm{e}-4$	
Learning Rate Decay	Linear	
Adam ε	$1\mathrm{e}-8$	
Adam(β_1, β_2)	(0,9, 0,999)	

Training data. In all the experiments, the training data is sampled *online*. Specifically, in each iteration, we sample N_1 i.i.d. data points, $(x^{(1)}, t^{(1)}), \cdots, (x^{(N_1)}, t^{(N_1)})$, from the domain $\mathbb{R}^n \times [0,T]$, and N_2 i.i.d. data points, $(\tilde{x}^{(1)}, T), \cdots, (\tilde{x}^{(N_2)}, T)$, from the boundary $\mathbb{R}^n \times \{T\}$, where $(x^{(i)}, t^{(i)}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_n) \times \mathcal{U}(0,1)$ and $\tilde{x}^{(j)} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_n)$.

Evaluation metrics. We use L^1, L^2, and $W^{1,1}$ relative errors to evaluate the quality of the learned solution.

L^1 and L^2 relative errors are two popular evaluation metrics, which are defined as

$$\frac{1}{S}\sum_{j=1}^{S} \frac{|u^*(x_j, t_j) - u_\theta(x_j, t)|^p}{|u^*(x_j)|^p},\ \ p = 1, 2, \tag{98}$$

where u_θ is the learned approximate solution, u^* is the exact solution and $\{x_j\}_{j=1}^S$ are S i.i.d. uniform samples from the domain $[0,1]^n \times [0,T]$.

Since the gradient of the solution to HJB equations plays an important role in applications, we also evaluate the solution using $W^{1,1}$ relative error, which is defined as

$$\frac{1}{S}\sum_{j=1}^{S}\frac{|u^*(x_j)-u_\theta(x_j)|+\sum_{i=1}^n|\partial_{x_i}u^*(x_j)-\partial_{x_i}u_\theta(x_j)|}{|u^*(x_j)|+\sum_{i=1}^n|\partial_{x_i}u^*(x_j)|}. \tag{99}$$

G More experiments and visualizations

G.1 More instance of HJB Equations

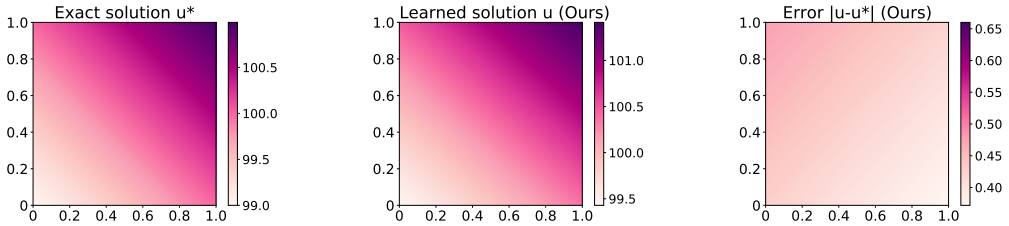


Figure 3: **Visualization for the solution snapshot of Eq. (100)**. c is set to 1.25.

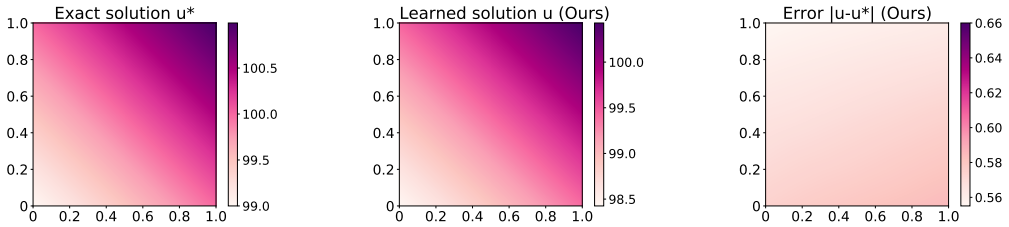


Figure 4: **Visualization for the solution snapshot of Eq. (100)**. c is set to 1.5.

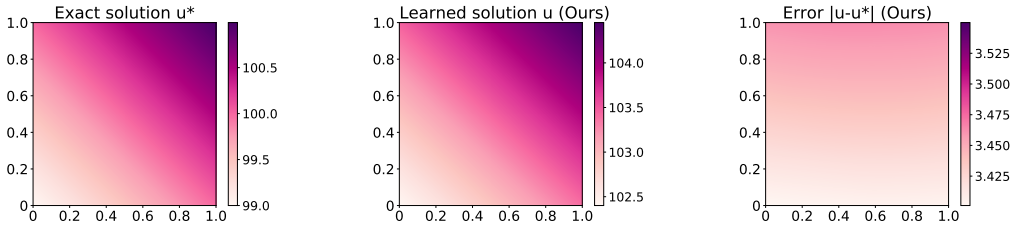


Figure 5: **Visualization for the solution snapshot of Eq. (100)**. c is set to 1.75.

To demonstrate the power of our method in solving general HJB Equations beyond classical LQG problems, we consider a more complicated HJB Equation as below:

$$\begin{cases}\partial_t u(x,t)+\Delta u(x,t)-\dfrac{1}{n}\displaystyle\sum_{i=1}^n|\partial_{x_i}u|^c=-2 & (x,t)\in\mathbb{R}^n\times[0,T]\\ u(x,T)=\displaystyle\sum_{i=1}^n x_i & x\in\mathbb{R}^n\end{cases}, \tag{100}$$

Eq. (100) has a unique solution $u(x,t)=x_1+\cdots+x_n+T-t$. We consider to solve Eq. (100) for different valued of c using our method. We choose $c=1.25, 1.5$ and 1.75 in the experiment. The neural network used for training is a 5-layer MLP with 4096 neurons and ReLU activation in each

Table 4: **Experimental results of solving the high dimensional HJB equations.** c is the parameter in Eq. (100). The dimensionality n is 100. Performances are measured by the L^1 relative error in the domain $[0,1]^n \times [0,T]$. Best performances are indicated in **bold**.

Method	$c = 1.25$	$c = 1.5$	$c = 1.75$
Original PINN [28]	1.11%	3.82%	2.73%
Adaptive time sampling [35]	1.18%	2.34%	7.94%
Learning rate annealing [34]	0.98%	1.13%	1.06%
Curriculum regularization [17]	6.27%	0.37%	3.51%
Adversarial training (ours)	**0.61%**	**0.15%**	**0.29%**

hidden layer. The training recipe, including the optimizer, learning rate, batch size, and the total iterations are the same as those in Appendix F. The number of inner-loop iterations K is set to 5, and the inner-loop step size η is searched from $\{2\mathrm{e}-1, 2\mathrm{e}-2, 2\mathrm{e}-3\}$.

Again, we examine the quality of the learned solution $u(x,t)$ by visualizing its snapshot on a two-dimensional space. Specifically, we consider the bivariate function $u(x_1, x_2, 1, 1, \cdots, 1; 0)$ and use a heatmap to show its function value given different x_1 and x_2. Figure 3-5 shows the ground truth u^*, the learned solutions u of our method, and the point-wise absolute error $|u - u^*|$ given different values of c.

From the above visualization, we can see that our method can solve Eq. (100) for different values of c effectively. Specifically, when $c = 1.25$ or 1.5, the point-wise absolute error is less than 0.5 for most of the area shown in the figures. When $c = 1.75$, the point-wise absolute error seems slightly larger, but it's still negligible compared with the scale of the learned solution. Thus, PINNs trained with our method fit the solution of Eq. (100) well, given different values of c.

We also compare our models with other baselines on these equations. The evaluation metric is L^1 relative error in the domain $[0,1]^n \times [0,1]$. The results are shown in Table 4. It's clear that our models outperform all the baselines on all these equations, showing the efficacy of our approach.

G.2 Tracing loss and error during the training

To give a more comprehensive comparison between original PINN and our method, we trace the loss and error during the training.

Table 5: **Error/loss-vs-time result of original PINN for Eq. (12)**.

Iteration	**1000**	**2000**	**3000**	**4000**	**5000**
L^2 Loss	0.098	0.088	0.070	0.584	0.041
L^1 Relative Error	6.18%	5.36%	3.86%	3.94%	3.47%
$W^{1,1}$ Relative Error	17.53%	17.67%	14.83%	14.40%	11.31%

Table 6: **Error/loss-vs-time result of our method for Eq. (12).**

Iteration	**1000**	**2000**	**3000**	**4000**	**5000**
L^∞ Loss	11.841	9.352	2.404	1.605	0.711
L^1 Relative Error	15.22%	4.26%	0.97%	1.10%	0.27%
$W^{1,1}$ Relative Error	21.91%	18.62%	5.14%	4.96%	2.22%

It is clear that for the original PINN approach, the L^2 loss drops very quickly during training, while its $W^{1,1}$ relative error remains high. This result indicates the optimization is successful in this experiment, and that the stability property of the PDE leads to the high test error. By contrast, our proposed training approach enables the test error goes down steadily during training, which aligns with the theoretical claims.

H Discussions on training with L^p loss

As is shown in the left panel of Table 2, directly optimizing L^p loss with large p fails to achieve a good approximator. This might seem to contradict our theoretical analysis in section 4. However, there is actually **no contradiction between our theorems and empirical results**. Theorem 4.3 focuses on the approximation ability, which indicates that if we have a model whose L^p loss is small, it will approximate the true solution well. The empirical results in Table 2 demonstrate the optimization difficulty of learning such a model.

Intuitively, we randomly sample points in each training iteration in the domain/boundary to calculate the loss. When p is large, most sampled points will hardly contribute to the loss, which leads to inefficiency and makes the training hard to converge. In Algorithm 1, we adversarially learn the points with large loss values, making all of them contribute to the model update (Step 8), significantly improving the model training.

Technically, directly applying Monte Carlo to compute L^p loss in experiments will lead to large variance estimations. For a function f,

$$\int |f|^p \mathrm{d}x = \frac{1}{N}\sum_{i=1}^{N} |f(X_i)|^p + \mathcal{O}\left(\sqrt{\frac{\mathrm{Var}|f(X)|^p}{N}}\right),$$

where $\{X_i\}_{i=1}^N$ are i.i.d. samplings in the domain.

Thus, $||f||_p$ suffers from an $\mathcal{O}((\mathrm{Var}|f(X)|^p/N)^{1/2p})$ error.

As $p \to \infty, \mathrm{Var}|f(X)|^p \sim ||f||_\infty^{2p}$. Therefore, the errors for estimating both Eq.(2,3) and the L^p norm of the residual are very large when p is large.