

CONVERGENCE RATE OF VANISHING VISCOSITY APPROXIMATIONS TO MEAN FIELD GAMES WITH NON-SEPARABLE HAMILTONIANS

WINSTON YU, QIANG DU, AND WENPIN TANG

ABSTRACT. This paper studies the vanishing viscosity approximation to mean field games (MFGs) in \mathbb{R}^d with a nonlocal and possibly non-separable Hamiltonian. We prove that the value function converges at a rate of $\mathcal{O}(\beta)$, where β^2 is the diffusivity constant, which matches the classical convergence rate of vanishing viscosity for Hamilton-Jacobi (HJ) equations. The same rate is also obtained for the approximation of the distribution of players as well as for the gradient of the value function. The proof is a combination of probabilistic and analytical arguments by first analyzing the forward-backward stochastic differential equation associated with the MFG, and then applying a general stability result for HJ equations. Applications of our result to N -player games, mean field control, and policy iteration for solving MFGs are also presented.

Key words: convergence rate, Fokker-Planck equation, forward-backward stochastic differential equation, mean field control, mean field games, non-separable Hamiltonian, nonlocal coupling, policy iteration, vanishing viscosity approximation.

1. INTRODUCTION

Mean field games (MFGs) were simultaneously proposed by Lasry and Lions in [40, 41, 42], and by Huang, Malhame, and Caines in [36], for the purpose of modeling a game with a large number of players whose decisions are influenced by the distribution of the other players. Due to the large number of players, each player is assumed to have an infinitesimal influence on all of the other players. Moreover, by assuming that all players act rationally (meaning that they each solve an optimization problem of some cost functional and act as though all other players are also playing rationally), the system is said to be in a Nash equilibrium. Now suppose that the agents are playing on the state space \mathbb{R}^d . Then one of the most common formulations of MFGs is as a system of coupled partial differential equations (PDEs), of which the first is a Hamilton-Jacobi (HJ) equation solved by $u^\beta : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, and the second is a Fokker-Planck equation solved by a flow ρ^β of probability measures on \mathbb{R}^d :

$$\begin{cases} -\partial_t u^\beta + H(x, -\nabla u^\beta, \rho_t^\beta) = \frac{\beta^2}{2} \Delta u & \text{on } [0, T] \times \mathbb{R}^d, \\ \partial_t \rho_t^\beta + \operatorname{div}_x \{ \rho_t^\beta \nabla_p H(x, -\nabla u^\beta, \rho_t^\beta) \} = \frac{\beta^2}{2} \Delta \rho_t^\beta & \text{on } [0, T] \times \mathbb{R}^d, \\ u^\beta(T, x) = g(x, \rho_T^\beta), \quad \rho_0^\beta(x) = m_0(x) & \text{on } \mathbb{R}^d, \end{cases} \quad (1.1)$$

where $\beta \geq 0$ is the idiosyncratic noise intensity¹, $H(x, p, \mu)$ is the (non-separable) Hamiltonian, $g(x, \mu)$ is the terminal cost, and m_0 is the initial distribution. When $\beta > 0$, the system is of second order. The system obtained by sending $\beta \rightarrow 0$ is called the *vanishing viscosity limit*,

Date: May 24, 2025.

¹The coefficient $\beta^2/2$ is also called the diffusivity constant.

which is of first order. Refer to Section 1.1 for a literature review on the well-posedness of MFGs (1.1).

Second-order MFGs are widely used to model complex systems in economics [1, 15, 48] and engineering [26, 35]. Recently, there has been a surge of interest in first-order MFG models, where the state evolves according to deterministic dynamics. Examples include the traffic flow of autonomous vehicles [34], and Proof-of-Stake cryptocurrency mining [53, Section 5]. As prior works have noted, such as in [34], traditional methods (e.g., Newton's method) for solving first-order MFGs tend to be numerically unstable. Various approaches [13, 14, 45] have recently been proposed to address such problems in solving first-order MFGs with a separable Hamiltonian (see (1.2) below), but with no quantitative guarantees. An obvious approach, as already suggested in [2], is to add a small second-order perturbation that corresponds to the addition of a small idiosyncratic noise, and then to solve the resulting second-order MFG². However, the price of gaining numerical stability from the perturbation is to introduce a source of error depending on the noise intensity β . Denoting $u = u^0$ for the value function of the first-order MFG, one would expect from the classical theory of HJ equations (see e.g., [22, Section IV]) that as $\beta \rightarrow 0$, $u^\beta \rightarrow u$ in some topology. But the classical theory of viscosity solutions does not provide a convergence rate with respect to β for MFGs, on account of the coupling with the Fokker-Planck equation.

The purpose of this paper is to provide a quantitative convergence rate of vanishing viscosity limits of the MFG (1.1) with a general (non-separable) Hamiltonian. Previous work [56] studied the convergence rate of vanishing viscosity limits of MFGs with a separable Hamiltonian on the torus \mathbb{T}^d ; see the end of Sections 1.1 and Remark 2.8 for discussions. Now let us briefly describe our results: we prove under suitable conditions on the model parameters that for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$,

$$\|u^\beta - u\|_{L^\infty([0,T] \times \mathcal{K})} \leq C_{\mathcal{K}}\beta,$$

for some constant $C_{\mathcal{K}}$ growing at most linearly in the diameter of \mathcal{K} . In other words, $u^\beta \rightarrow u$ at a rate of $O(\beta)$ in the topology of uniform convergence on compact sets (Theorem 4.2), which matches with the convergence rate from the classical viscosity theory of HJ equations. As intermediary steps, we prove that ρ_t^β and ∇u^β converge at the same rate of $O(\beta)$ in the 1-Wasserstein and $L^\infty([0, T] \times \mathbb{R}^d)$ metrics, respectively (Corollary 3.3 and Theorem 3.2). Our analysis does not specifically require the Lasry-Lions or displacement monotonicity condition; see Remark 2.8 for a discussion. We also show how to apply our result to varied problems, such as particle system approximations and policy iteration for solving MFGs, in Section 5.

Here we give a quick outline of the proof, which is a combination of probabilistic and analytical arguments. From the classical stability theory of PDEs, one may suspect that the difference between u^β and u is controlled by the difference in the coefficients, i.e., $H(\cdot, \cdot, \rho_t^\beta) - H(\cdot, \cdot, \rho_t)$, $g(\cdot, \rho_T^\beta) - g(\cdot, \rho_T)$, and β . However, the dependence on the measure in the first four terms complicates the analysis, because ρ^β and ρ satisfy their own PDEs that depend on u^β and u , respectively. We avoid this issue by instead analyzing the forward-backward stochastic differential equation (FBSDE) system (defined in (2.3)) associated with the MFG system (1.1). Our analysis consists of three steps:

²This approach is also reminiscent of the Lax-Friedrichs approximation scheme to first-order equations, where numerical viscosity is $\beta = \sqrt{\frac{2(\Delta x)^2}{\Delta t}}$ with $(\Delta x, \Delta t)$ as the space-time discretization (see e.g., [19, 20, 52]).

- (1) First we control the L^1 difference between ρ_t^β and ρ_t in terms of β and $\|\nabla u^\beta - \nabla u\|_\infty$ (Lemma 3.1).
- (2) Using the FBSDE representation, we prove that $\|\nabla u^\beta - \nabla u\|_\infty = \mathcal{O}(\beta)$ (Theorem 3.2). This implies the convergence of ρ_t^β to ρ_t at a rate of $\mathcal{O}(\beta)$ in the L^1 , and hence, W_1 metric (Corollary 3.3).
- (3) We obtain a rate of $\mathcal{O}(\beta)$ for the convergence of $H(\cdot, \cdot, \rho_t^\beta)$ and $g(\cdot, \rho_T^\beta)$ by Corollary 3.3 (and some assumptions on H and g). Finally, the convergence rate of u^β follows from a general stability result (Theorem 4.1).

Organization of the paper: The remainder of this paper is organized as follows. Section 1.1 provides a literature review on MFGs, and compares the result in this paper with prior work. In Section 2, we formally define the problem and collect some assumptions for our result. Section 3 proves the convergence rate of ρ_t^β and ∇u^β , and Section 4 proves the convergence rate of u^β . In Section 5, we give several applications of our result. Examples and numerical experiments are presented in Section 6. Concluding remarks are summarized in Section 7.

1.1. Literature Review. The well-posedness of MFGs has been studied extensively, particularly for the case of a separable Hamiltonian, in which the momentum and the measure arguments are additively separated:

$$H(x, p, \mu) = H_0(x, p) - f(x, \mu). \quad (1.2)$$

Here H can either be a local or nonlocal function of the measure argument. In the separable, local case and when $\beta = 0$, [10, 11, 29] are some of the major works proving the well-posedness. In the separable, nonlocal case and when $\beta = 0$, [12] ensures well-posedness of the MFG system, provided that the Lasry-Lions monotonicity condition holds.

The separable case is a strong structural assumption, upon which much of the previous literature relied. However, many applications may go beyond this assumption. In the economic model proposed by [1, 48], despite the model of the agent being relatively simple to formulate, such an agent corresponds to a MFG whose Hamiltonian is not separable. Another example is provided by [53], where the Hamiltonian has a term containing the product of the price process (which is a function of the player distribution) and the action of the player. Finally, [3] is one of the works that models mean field games with congestion, where players are penalized based on the density of other players at the current position; as a result of congestion penalizing movement, the Hamiltonian some function of the momentum divided by another function of the density. In order to make sense of the MFGs arising in these applications, the well-posedness of MFGs with a non-separable Hamiltonian must first be established. A breakthrough was made by [31], and later [47] for their proposal of a new condition, called displacement monotonicity, under which well-posedness of the MFG system (1.1) can be proved for all $\beta \geq 0$. To the best of our knowledge, the work that proves well-posedness under the least restrictive regularity assumptions on H and g is [5], whose main assumption, other than displacement monotonicity, is the uniform boundedness of the second derivatives of H and g .

Some earlier works take a probabilistic approach to MFGs as well. [17] considers a probabilistic formulation of the MFG, where the volatility is uncontrolled. Their Remark 7.12 is similar to our FBSDE system (2.3), though our equation of the adjoint process is for the gradient of the value function, not for the value function itself. The later work of [38] also studies MFGs

from a probabilistic perspective, and they prove the existence, though not its uniqueness, of a MFG solution where the volatility is controlled. See [16] for further developments in this direction.

Comparison to previous work: The only previous work addressing the convergence rate of vanishing viscosity approximations to MFGs is [56]. Its main assumption is that of a separable Hamiltonian, which (along with the terminal cost function) satisfies the Lasry-Lions monotonicity condition. In contrast, we do not assume any monotonicity condition until one is needed for the well-posedness of the MFG system (1.1). Moreover, our result addresses the case of a non-separable Hamiltonian which is nonlocal in the measure argument. We prove that the convergence rate for $\{u^\beta\}_{\beta>0}$ is $\mathcal{O}(\beta)$ in $L^\infty(\mathcal{K})$ for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$, which improves upon their rate $\mathcal{O}(\beta^{1/2})$ in $L^1(\mathbb{T}^d)$. This (partially) solves Problem 4(a) in [56] for MFGs with possibly non-separable and nonlocal Hamiltonians.

2. NOTATIONS, ASSUMPTIONS AND PROBLEM FORMULATION

2.1. Notations. For a metric space X , let $\mathcal{C}^k(X)$ be the space of functions mapping X to \mathbb{R} , which are k -times differentiable, and whose k -th order derivatives are continuous. If $X = [0, T] \times \mathbb{R}^n$, then for $f : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$, $\nabla f(t, x)$ refers to $\nabla_x f(t, x) = [\partial_i f(t, x)]_{i=1}^n$, and $\nabla^2 f(t, x)$ refers to $\nabla_{xx}^2 f(t, x) = [\partial_{ij} f(t, x)]_{i,j=1}^n$. In particular, ∇f and $\nabla^2 f$ do not include the partial derivatives with respect to t .

For $p \in [1, \infty]$ and a measure space (X, \mathcal{B}, μ) , $L^p(X, \mathcal{B}, \mu)$ is the space of real-valued functions whose p -th power is integrable with respect to μ , i.e., all $f : X \rightarrow \mathbb{R}$ such that $\|f\|_p^p = \int_X |f|^p d\mu < \infty$. If $p = \infty$, then $L^\infty(X, \mathcal{B}, \mu)$ is the space of functions f such that there exists a constant $C > 0$ satisfying $\mu(|f| > C) = 0$; the infimum of all such constants is denoted by $\|f\|_\infty = \|f\|_{L^\infty}$. If we only write L^p , then μ is assumed to be the Lebesgue measure on X , and \mathcal{B} is the Borel sigma-field on X . We will also make use of L^∞ spaces of functions mapping \mathbb{R}^n to \mathbb{R}^m , denoted by $L^\infty(\mathbb{R}^n; \mathbb{R}^m)$, consisting of functions $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that there exists a C with $|f| \leq C$ almost everywhere with respect to the Lebesgue measure on \mathbb{R}^n ($|\cdot|$ is the Euclidean norm on \mathbb{R}^n). If \mathbb{R}^m is replaced by $\mathbb{R}^{m \times m}$, the space of m by m matrices, then $|\cdot|$ is replaced by the operator norm $\|\cdot\|_\infty$ on matrices.

Now let $(X, \mathcal{B}, \mathbb{P})$ be a probability space. For a random variable ξ , $\text{Law}(\xi)$ is the law of ξ with respect to \mathbb{P} . For $p \in [1, \infty]$, we write $\mathcal{P}_p(X)$ for the space of probability measures μ with finite p -th moment, i.e., $\int_X |x|^p d\mu(x) < \infty$. On $\mathcal{P}_p(X)$, we define the p -Wasserstein distance W_p :

$$W_p(\mu, \nu) = \inf \left\{ \int_{X \times X} |x - y|^p d\pi(x, y) : \pi \in \mathcal{P}(X \times X) \text{ has marginals } \mu \text{ and } \nu \right\}^{1/p}.$$

We also introduce the Wasserstein gradient of a function $U : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$. For a more extensive introduction, one source is the textbook [16]. For $\mu \in \mathcal{P}_2(\mathbb{R}^n)$, the Wasserstein gradient of U at μ is denoted by $\nabla_\mu U(\mu, \cdot)$, and is an element of the closure (with respect to $L^2(\mathbb{R}^n, \mu)$) of gradients of $\mathcal{C}^\infty(\mathbb{R}^n)$ functions. The gradient $\nabla_\mu U$ is characterized by the property that for all L^2 random variables ξ and η in \mathbb{R}^n ,

$$U(\text{Law}(\xi + \eta)) - U(\text{Law}(\xi)) = \mathbb{E}[\nabla_\mu U(\text{Law}(\xi), \xi) \cdot \eta] + o(\mathbb{E}[|\eta|^2]^{1/2}).$$

Finally, for $a, b > 0$, the symbol $a = \mathcal{O}(b)$ means that a/b is bounded from above, as some problem parameter tends to 0 or ∞ . Similarly, $a \asymp b$ means that a/b is bounded from below and from above, as some problem parameter tends to 0 or ∞ .

2.2. Assumptions. Let $H : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ be the Hamiltonian, and $g : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ be the terminal cost function. We need the following Lipschitz and regularity conditions on H and g .

Assumption 2.1 (Lipschitz property of H). *H is Lipschitz in the measure argument with respect to W_1 , and we denote the Lipschitz constant by $\|\nabla_\mu H\|_\infty$: for all $\mu^1, \mu^2 \in \mathcal{P}_2(\mathbb{R}^d)$ and $x, p \in \mathbb{R}^d$,*

$$|H(x, p, \mu^1) - H(x, p, \mu^2)| \leq \|\nabla_\mu H\|_\infty W_1(\mu^1, \mu^2).$$

Assumption 2.2 (Regularity of H). *The derivatives $\nabla_{xx}^2 H$, $\nabla_{xp}^2 H$, $\nabla_{x\mu}^2 H$, $\nabla_{pp}^2 H$, and $\nabla_{p\mu}^2 H$ exist. Moreover, despite not specifying a measure on $\mathcal{P}_2(\mathbb{R}^d)$, we say that $\nabla_{xx}^2 H$, $\nabla_{xp}^2 H$, and $\nabla_{pp}^2 H$ have finite L^∞ norms on their respective domains in the sense that:*

$$\begin{aligned} \|\nabla_{xx}^2 H\|_\infty &:= \sup_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} \|\nabla_{xx}^2 H(\cdot, \cdot, \mu)\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d \times \mathbb{R}^d)} < \infty, \\ \|\nabla_{xp}^2 H\|_\infty &:= \sup_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} \|\nabla_{xp}^2 H(\cdot, \cdot, \mu)\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d \times \mathbb{R}^d)} < \infty, \\ \|\nabla_{pp}^2 H\|_\infty &:= \sup_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} \|\nabla_{pp}^2 H(\cdot, \cdot, \mu)\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d \times \mathbb{R}^d)} < \infty. \end{aligned}$$

Finally, $\nabla_x H$ and $\nabla_p H$ are Lipschitz in the measure argument with Lipschitz constants $\|\nabla_{x\mu}^2 H\|_\infty$ and $\|\nabla_{p\mu}^2 H\|_\infty$: for all $\mu^1, \mu^2 \in \mathcal{P}_2(\mathbb{R}^d)$ and $x, p \in \mathbb{R}^d$,

$$\begin{aligned} |\nabla_x H(x, p, \mu^1) - \nabla_x H(x, p, \mu^2)| &\leq \|\nabla_{x\mu}^2 H\|_\infty W_1(\mu^1, \mu^2), \\ |\nabla_p H(x, p, \mu^1) - \nabla_p H(x, p, \mu^2)| &\leq \|\nabla_{p\mu}^2 H\|_\infty W_1(\mu^1, \mu^2). \end{aligned}$$

Assumption 2.3 (Convexity of H in p). [47, (2.7)] *H is uniformly convex in p : there exists some $c_0 > 0$ such that $\nabla_{pp}^2 H \succcurlyeq c_0 I_d$. Also, for each $p \in \mathbb{R}^d$, there exists a constant $C(p)$ such that for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $|\nabla_p H(0, p, \mu)| \leq C(p)$.*

Remark 2.4. *Let us present two classes of Hamiltonians that satisfy Assumption 2.1–2.3.*

(1) *Let $\Gamma > 0$, F and γ be real-valued functions, and $m(\mu)$ be the mean of μ . Let*

$$H(x, p, \mu) = \Gamma|p|^2 + \gamma(p)F(x, m(\mu)). \quad (2.1)$$

Then Assumption 2.1–2.3 are satisfied if $\gamma, \gamma', \gamma'', F$, and ∇F are bounded, and $\Gamma > 0$ is sufficiently large.

(2) *Let $\Gamma > 0$, $\gamma, G : \mathbb{R}^d \rightarrow \mathbb{R}$, and $F : \mathbb{R} \rightarrow \mathbb{R}$, $\varphi \in \mathcal{C}_{bc}^1(\mathbb{R}^d; \mathbb{R})$ (i.e., φ and $\nabla\varphi$ are bounded and are compactly supported), and $\eta > 0$. Let*

$$H(x, p, \mu) = \Gamma|p|^2 + \frac{\gamma(p)}{|\eta + F((\varphi * \mu)(x))|} + G(x). \quad (2.2)$$

This example is motivated by non-separable Hamiltonians from MFGs modeling congestion [3]. The idea behind the convolution of μ with φ is to transform the Hamiltonian that is usually encountered in MFGs with congestion from a local one into a nonlocal one. From a modeling perspective, it should be interpreted as agents not only taking into

account the density of agents at their current position, but also the density of agents in some compact set around the agent's position. Then Assumption 2.1–2.3 are satisfied if $\gamma, \gamma', \gamma''$ are bounded, F is Lipschitz and nonnegative, G has a bounded Hessian, and $\Gamma > 0$ is sufficiently large.

Assumption 2.5 (Lipschitz property of g). g is Lipschitz in the measure argument with respect to W_1 , and we denote the Lipschitz constant by $\|\nabla_{\mu}g\|_{\infty}$: for all $\mu^1, \mu^2 \in \mathcal{P}_2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$,

$$|g(x, \mu^1) - g(x, \mu^2)| \leq \|\nabla_{\mu}g\|_{\infty} W_1(\mu^1, \mu^2).$$

Assumption 2.6 (Regularity of g). The derivatives $\nabla_{xx}^2 g$ and $\nabla_{x\mu}^2 g$ exist. Moreover, despite not specifying a measure on $\mathcal{P}_2(\mathbb{R}^d)$, we say that $\nabla_{xx}^2 g$ has finite L^{∞} norm in the sense that:

$$\|\nabla_{xx}^2 g\|_{\infty} := \sup_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} \|\nabla_{xx}^2 g(\cdot, \mu)\|_{L^{\infty}(\mathbb{R}^d, \mathbb{R}^{d \times d})} < \infty.$$

Finally, $\nabla_x g$ is Lipschitz in the measure argument with respect to W_1 , and we denote its Lipschitz constant by $\|\nabla_{x\mu}^2 g\|_{\infty}$: for all $\mu^1, \mu^2 \in \mathcal{P}_2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$,

$$|\nabla_x g(x, \mu^1) - \nabla_x g(x, \mu^2)| \leq \|\nabla_{x\mu}^2 g\|_{\infty} W_1(\mu^1, \mu^2).$$

The above assumptions largely agree with those in [5] (namely, Assumptions 2.6(1) and 2.7(1) therein). To reiterate, [5] is, to the best of our knowledge, the work with the least restrictive regularity assumptions that guarantee well-posedness of the master equation. However, there is a notable exception: when applying the PDE stability result to MFGs in Corollary 4.2, we found it quite convenient to assume the boundedness of $\nabla_{\mu}H$ and $\nabla_{\mu}g$, which [5] did not assume.

Assumption 2.7. For all $\beta \geq 0$, the MFG (1.1) is well-posed with solution $(u^{\beta}, \rho^{\beta})$ in the sense of Definition 2.9. This assumption will be in force for the rest of the paper, even when we do not say so explicitly.

As discussed in [47, Section 3], Assumption 2.7 can be satisfied by MFGs that do not necessarily possess displacement monotone H or g . Indeed, we allow H and g that satisfy any conditions (that do not contradict Assumptions 2.1–2.6) sufficient to guarantee well-posedness. See e.g., [32, 49] for more sufficient conditions for well-posedness that are beyond the Lasry-Lions and displacement monotonicity conditions.

Remark 2.8 (Comparison with [56]). We compare our assumptions to those of [56] in greater detail. In terms of regularity, their condition (H1') on the \mathcal{C}^2 norm of the coupling says that the coupling must have bounded first and second derivatives, which is slightly stronger than ours (we allow for $\nabla_x H$ to be unbounded). Their condition (H2'), that the second-order derivatives of H are locally bounded, is of course not as strong as uniform boundedness. Their condition (H3') is stronger than ours, which requires that the \mathcal{C}^2 norm of the terminal cost is bounded, while we only require that g be Lipschitz in the measure argument with respect to W_1 . Finally, (H4') and (H4'') do not seem to be directly comparable to the regularity of Wasserstein derivatives. However, observe that if the condition of $\varepsilon^{1/2}$ (in H4' and H4'') were replaced with ε , then they would have achieved the same convergence rate as we did.

Moreover, we did not find necessary to assume the monotonicity conditions that [56] imposed on their Hamiltonian and coupling. For instance, we could impose any of the four conditions

in [32] that guarantee well-posedness of the MFG system, while [56] did require the Lasry-Lions monotonicity condition for their proof³. Most importantly, [56] relied on the separable Hamiltonian structure, and their proof technique seems difficult to adapt to the non-separable case due to the dual equation technique [44].

2.3. Problem Formulation. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, \mathbb{P})$ be a filtered probability space for some time horizon $T > 0$, where the filtration is generated by a standard d -dimensional Brownian motion $\{B_t\}_{t=0}^T$. A representative agent takes a path $\{X_s^{\alpha, \beta}\}_{s=t}^T$ through \mathbb{R}^d , which satisfies the SDE:

$$\begin{cases} dX_s^{\alpha, \beta} = \alpha_s(X_s^{\alpha, \beta}, \rho_s^\beta) ds + \beta dB_s & \text{for } s \in (t, T], \\ X_t^{\alpha, \beta} = x, \end{cases}$$

for some initial position $x \in \mathbb{R}^d$, initial time $t \in [0, T]$, and adapted stochastic process $\{\alpha_s\}_{s=t}^T$ (referred to as the control). Its goal is to minimize the following cost functional J :

$$J(t, x; \alpha) = \mathbb{E} \left[\int_t^T L(X_s^{\alpha, \beta}, \alpha_s, \rho_s^\beta) ds + g(X_T^{\alpha, \beta}, \rho_T^\beta) \middle| X_t^{\alpha, \beta} = x \right],$$

over all adapted processes α . Here the Lagrangian L is a running cost function that is the Legendre transform of H in the second variable, g is a terminal cost function, ρ_s^β is the distribution of players at time s , and $\beta \geq 0$ is the idiosyncratic noise intensity faced by each player. Let $u^\beta(t, x) = \inf_\alpha J(t, x; \alpha)$. From classical optimal control, if all players play optimally in the sense of a Nash equilibrium, then at time s and position x , each player chooses the action $\alpha : [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$, defined by

$$\alpha_s(x, \rho) = \nabla_p H(x, -\nabla u^\beta(s, x), \rho).$$

Since α is already determined for each agent, we drop α from the superscripts, and we use X^β to refer to the stochastic process representing the path of the agent, and we use ρ_s^β to refer to the law of X_s^β . Moreover, the pair (u^β, ρ^β) solves the coupled PDEs (1.1) in the following sense:

Definition 2.9 (Definition 3.2 in [47]). *We say that (u^β, ρ^β) is a solution pair to the MFG system 1.1 if*

- (i) *the Lipschitz constant of u^β restricted to any compact set is finite, u^β solves the HJB equation in the viscosity sense, and $\nabla_{xx}^2 u^\beta \in L^\infty([0, T] \times \mathbb{R}^d; \mathbb{R}^{d \times d})$.*
- (ii) *$\rho^\beta : [0, T] \rightarrow (\mathcal{P}_1(\mathbb{R}^d), W_1)$ is continuous, and ρ^β solves the Fokker-Planck equation in the distributional sense: for all test functions $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^d)$, the following equation holds:*

$$\int_0^T \int_{\mathbb{R}^d} \left[-\partial_t \varphi + \langle \nabla \varphi, \nabla_p H(x, -\nabla u^\beta(t, x), \rho_t^\beta) \rangle \right] d\rho_t^\beta(x) dt = \frac{\beta^2}{2} \int_0^T \int_{\mathbb{R}^d} \Delta \varphi d\rho_t^\beta(x) dt,$$

and

$$\int_{\mathbb{R}^d} \varphi(0, x) dm_0(x) = \int_{\mathbb{R}^d} \varphi(0, x) d\rho_0^\beta(x).$$

³The Lasry-Lions monotonicity condition is needed in their equations (6.10) and (6.11).

It is worth noting that from [47, Lemma 3.4], there is a semi-concavity estimate for u^β which is uniform in β , i.e., $K := \sup_{\beta > 0} \|\nabla_{xx}^2 u^\beta\|_{L^\infty([0, T] \times \mathbb{R}^d; \mathbb{R}^{d \times d})} < \infty$. According to [47, Theorem 4.1], under Assumption 2.3, the MFG (1.1) has an FBSDE representation:

$$\begin{cases} dX_t^\beta = \nabla_p H(X_t^\beta, Y_t^\beta, \rho_t^\beta) dt + \beta dB_t, \\ dY_t^\beta = -\nabla_x H(X_t^\beta, Y_t^\beta, \rho_t^\beta) dt + \beta Z_t^\beta dB_t, \\ X_0^\beta \sim m_0, \quad Y_T^\beta = -\nabla_x g(X_T^\beta, \rho_T^\beta), \end{cases} \quad (2.3)$$

which has a strong solution. The first SDE describes the state dynamics of an agent playing in a Nash equilibrium. The key to our analysis is the fact that $\nabla u^\beta : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a decoupling field for the FBSDE (2.3) in the sense that:

$$Y_t^\beta = -\nabla u^\beta(t, X_t^\beta) \quad \text{for almost every } t \in [0, T].$$

The meaning of ∇u^β to the FBSDE is that the SDE for Y^β is solved by the gradient of the value function evaluated along the trajectory of a typical agent playing in a Nash equilibrium. Instead of relying on PDE methods such as the dual equation as [56] did, we use a probabilistic approach that hinges on stability properties of the FBSDE (2.3) to derive our main result.

3. CONVERGENCE OF ρ^β AND ∇u^β

The difficulty of analyzing MFGs with a non-separable Hamiltonian is that without the assumption of separability (1.2), we can no longer analyze each equation separately. However, as discussed previously, we can use the convenient property of the FBSDE (2.3) having $-\nabla u^\beta$ as a decoupling field in the sense of Equation (2.3). By substituting Equation (2.3) into the SDE for X^β , we can analyze it separately from the SDE for Y^β . Furthermore, because $\rho_t^\beta = \text{Law}(X_t^\beta)$ for all $t \in [0, T]$, we can get a convergence rate for ρ^β to ρ in L^1 , and hence, in W_1 .

Before we state the results, consider the following FBSDE system for $t_0 < T$ and $x \in \mathbb{R}^d$:

$$\begin{cases} dX_t^\beta = \nabla_p H(X_t^\beta, Y_t^\beta, \rho_t^\beta) dt + \beta dB_t, \\ dY_t^\beta = -\nabla_x H(X_t^\beta, Y_t^\beta, \rho_t^\beta) dt + \beta Z_t^\beta dB_t, \\ X_{t_0}^\beta = x, \quad Y_T^\beta = -\nabla_x g(X_T^\beta, \rho_T^\beta). \end{cases} \quad (3.1)$$

Other than the initial condition, its only difference compared to (2.3) is that the initial time t_0 is not necessarily 0. To lighten notation, when $\beta = 0$, we denote by $X_t = X_t^0$ the solution to the FBSDE (2.3) or (3.1).

Lemma 3.1. *For $\beta \geq 0$, $0 \leq t_0 < T$, and $\xi \in L^2(\Omega, \mathcal{F}_{t_0}, \mathbb{P})$, consider the uncoupled state variable dynamics:*

$$\begin{cases} dX_t^\beta = \nabla_p H(X_t^\beta, -\nabla u^\beta(t, X_t^\beta), \rho_t^\beta) dt + \beta dB_t, \\ X_{t_0}^\beta \sim \xi. \end{cases} \quad (3.2)$$

Under Assumptions 2.2, 2.3, 2.6, and 2.7, there exists a constant $C = C(H, g, T)$, where the dependence on H and g are only through the L^∞ norms of their second-order derivatives, such that for all $T' \in [t_0, T]$,

$$\sup_{t \in [t_0, T']} \mathbb{E}[|X_t^\beta - X_t|] \leq C \left\{ \int_{t_0}^{T'} \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_\infty ds + \beta \right\}. \quad (3.3)$$

Proof. Note that the equation (3.2) is the SDE satisfied by a solution to the X component of the FBSDE (3.1), the existence of which is guaranteed by [47, Theorem 4.1]. Using Assumptions 2.2 and 2.6, we have that for all $s \in [t_0, T']$:

$$\begin{aligned} & \mathbb{E}[|\nabla_p H(X_s^\beta, -\nabla u^\beta(s, X_s^\beta), \rho_s^\beta) - \nabla_p H(X_s, -\nabla u(s, X_s), \rho_s)|] \\ & \leq \mathbb{E}[\|\nabla_{xp}^2 H\|_\infty |X_s^\beta - X_s| + \|\nabla_{pp}^2 H\|_\infty (|\nabla u^\beta(s, X_s^\beta) - \nabla u^\beta(s, X_s)| \\ & \quad + \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_\infty)] + \|\nabla_{p\mu}^2 H\|_\infty W_1(\rho_s^\beta, \rho_s) \\ & \leq (\|\nabla_{xp}^2 H\|_\infty + \|\nabla_{p\mu}^2 H\|_\infty) \mathbb{E}[|X_s^\beta - X_s|] + \|\nabla_{pp}^2 H\|_\infty (K \mathbb{E}[|X_s^\beta - X_s|] \\ & \quad + \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_\infty). \end{aligned}$$

As a result, there exists some (explicitly computable) constant C depending only on T and the Lipschitz constants of the gradients of H and g , such that for all $T' \leq T$:

$$\begin{aligned} & \mathbb{E}[|X_{T'}^\beta - X_{T'}|] \\ & \leq C \mathbb{E} \left\{ \int_{t_0}^{T'} |\nabla_p H(X_s^\beta, -\nabla u^\beta(s, X_s^\beta), \rho_s^\beta) - \nabla_p H(X_s, -\nabla u(s, X_s), \rho_s)| ds + \beta |B_{T'}| \right\} \\ & \leq C \left\{ \int_{t_0}^{T'} \mathbb{E}[|X_s^\beta - X_s|] + \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_\infty ds + \beta \mathbb{E}[|B_{T'}|] \right\}. \end{aligned}$$

By Grönwall's inequality, there exists another constant C such that

$$\mathbb{E}[|X_{T'}^\beta - X_{T'}|] \leq C \left\{ \int_{t_0}^{T'} \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_\infty ds + \beta \mathbb{E}[|B_{T'}|] \right\}.$$

Using the fact that the right-hand side is non-decreasing in T' , absorbing $\mathbb{E}[|B_{T'}|]$ into C , and then taking the supremum on both sides over $t \in [t_0, T']$ yields the conclusion. \square

The previous result hints that in order to quantify the convergence of ρ_t^β to ρ_t , it suffices to control the convergence of $\{\nabla u^\beta(t, \cdot)\}_{\beta > 0}$ to $\nabla u(t, \cdot)$ in $L^\infty(\mathbb{R}^d)$. In the following theorem, we obtain a convergence rate of $O(\beta)$ for ∇u^β to ∇u by exploiting the decoupling field ∇u^β .

Theorem 3.2. *Let u^β and u be (possibly non-unique) value function solutions to the MFG system (1.1) with $\beta > 0$ and $\beta = 0$, respectively. Suppose that Assumptions 2.2, 2.3, 2.6, and 2.7 hold. Then there exists a constant $C = C(H, g, T)$, where the dependence on H and g is only through the L^∞ bounds on the second-order derivatives of H and g , such that*

$$\|\nabla u^\beta - \nabla u\|_{L^\infty([0, T] \times \mathbb{R}^d)} \leq C\beta.$$

Proof. Denote by $(X^\beta, Y^\beta, Z^\beta)$ and (X, Y, Z) the solutions to the FBSDE (3.1) with the initial condition $X_{t_0}^\beta = X_{t_0} = x$, $t_0 \in (0, T)$, and $\beta > 0$ and $\beta = 0$ respectively. We have $Y_t^\beta = -\nabla u^\beta(t, X_t^\beta)$ and $Y_t = -\nabla u(t, X_t)$ for almost every t . By [47, (4.13)], and using the uniform convexity of H in p from Assumption 2.3, we have that for $\beta \geq 0$, $-\nabla u^\beta(t_0, x) = \mathbb{E}[\nabla_x g(X_T^\beta, \rho_T^\beta)] - \int_{t_0}^T \mathbb{E}[\nabla_x H(X_t^\beta, Y_t^\beta, \rho_t^\beta)] dt$. By the triangle inequality, we get:

$$\begin{aligned} & |\nabla u^\beta(t_0, x) - \nabla u(t_0, x)| \leq \mathbb{E}[|\nabla_x g(X_T^\beta, \rho_T^\beta) - \nabla_x g(X_T, \rho_T)|] \\ & \quad + \int_{t_0}^T \mathbb{E}[|\nabla_x H(X_t^\beta, -\nabla u^\beta(t, X_t^\beta), \rho_t^\beta) - \nabla_x H(X_t, -\nabla u^\beta(t, X_t), \rho_t)|] dt. \end{aligned} \tag{3.4}$$

Using Lemma 3.1 and Assumption 2.6, we can bound the first term in (3.4) by

$$\begin{aligned}
\mathbb{E}[\|\nabla_x g(X_T^\beta, \rho_T^\beta) - \nabla_x g(X_T, \rho_T)\|] &\leq \|\nabla_{xx}^2 g\|_\infty \mathbb{E}[|X_T^\beta - X_T|] + \|\nabla_{x\mu}^2 g\|_\infty W_1(\rho_T^\beta, \rho_T) \\
&\leq (\|\nabla_{xx}^2 g\|_\infty + \|\nabla_{x\mu}^2 g\|_\infty) \mathbb{E}[|X_T^\beta - X_T|] \\
&\leq C \left\{ \int_{t_0}^{T'} \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_\infty ds + \beta \right\}.
\end{aligned} \tag{3.5}$$

By Assumption 2.2, we can bound the second term in (3.4):

$$\begin{aligned}
&\int_{t_0}^T \mathbb{E}[\|\nabla_x H(X_t^\beta, -\nabla u^\beta(t, X_t^\beta), \rho_t^\beta) - \nabla_x H(X_t, -\nabla u^\beta(t, X_t), \rho_t)\|] dt \\
&\leq \int_{t_0}^T \|\nabla_{xx}^2 H\|_\infty \mathbb{E}[|X_t^\beta - X_t|] + \|\nabla_{x\mu}^2 H\|_\infty W_1(\rho_t^\beta, \rho_t) \\
&\quad + \|\nabla_{xp}^2 H\|_\infty \mathbb{E}[|\nabla u^\beta(t, X_t^\beta) - \nabla u^\beta(t, X_t)| + |\nabla u^\beta(t, X_t) - \nabla u(t, X_t)|] dt \\
&\leq \int_{t_0}^T C \mathbb{E}[|X_t^\beta - X_t|] + \|\nabla_{xp}^2 H\|_\infty (\|\nabla_{xx}^2 u^\beta\|_\infty \mathbb{E}[|X_t^\beta - X_t|] + \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_\infty) dt.
\end{aligned} \tag{3.6}$$

By applying Lemma 3.1 and $K := \sup_{\beta>0} \{\|\nabla_{xx}^2 u^\beta\|_\infty\} < \infty$, we can continue from the last line in (3.6) to get:

$$\begin{aligned}
&\int_{t_0}^T \mathbb{E}[\|\nabla_x H(X_t^\beta, -\nabla u^\beta(t, X_t^\beta), \rho_t^\beta) - \nabla_x H(X_t, -\nabla u^\beta(t, X_t), \rho_t)\|] dt \\
&\leq \int_{t_0}^T (C + K \|\nabla_{xp}^2 H\|_\infty) \mathbb{E}[|X_t^\beta - X_t|] + \|\nabla_{xp}^2 H\|_\infty \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_\infty dt \\
&\leq C \left\{ (T - t_0) \sup_{t \in [t_0, T]} \mathbb{E}[|X_t^\beta - X_t|] + \int_{t_0}^T \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_\infty dt \right\} \\
&\leq C \left\{ \int_{t_0}^T \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_\infty ds + \beta + \int_{t_0}^T \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_\infty dt \right\} \\
&\leq C \left\{ \int_{t_0}^T \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_\infty ds + \beta \right\}.
\end{aligned} \tag{3.7}$$

In the above computations, the value of C may change from line to line but only depends on T and the Lipschitz constants of the relevant gradients of H and g . Therefore, after taking the supremum over all $x \in \mathbb{R}^d$ and combining (3.5) and (3.7), the equation (3.4) is bounded by

$$\|\nabla u^\beta(t_0, \cdot) - \nabla u(t_0, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq C \left\{ \int_{t_0}^T \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(\mathbb{R}^d)} ds + \beta \right\}.$$

Then we apply Grönwall's inequality to find that for a possibly different C still only depending on H , g , and T ,

$$\|\nabla u^\beta(t_0, \cdot) - \nabla u(t_0, \cdot)\|_{L^\infty(\mathbb{R}^d)} \leq C\beta.$$

Since C does not depend on t_0 , we conclude that $\|\nabla u^\beta - \nabla u\|_{L^\infty([0, T] \times \mathbb{R}^d)} \leq C\beta$. \square

Corollary 3.3. *Suppose that Assumptions 2.2, 2.3, 2.6, and 2.7 hold. Then, there exists a constant $C = C(H, g, T)$, again only depending on the second derivatives of H and g , such that for any $t \in [0, T]$, we have:*

$$W_1(\rho_t^\beta, \rho_t) \leq \mathbb{E}[|X_t^\beta - X_t|] \leq C\beta. \quad (3.8)$$

Proof. Note that

$$\mathbb{E}[|X_t^\beta - X_t|] \leq C \left\{ \int_{t_0}^t \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_\infty ds + \beta \right\} \leq C(T\beta + \beta) \leq C\beta,$$

where we use Lemma 3.1 for the first inequality, and Theorem 3.2 for the second inequality. This proves the result. \square

Remark 3.4. *The same convergence rate can be proved in W_2 . The proof would first use stability properties of the SDE and Gronwall's inequality to obtain for all $\beta \geq \nu > 0$ and $t \in [0, T]$,*

$$W_2(\rho_t^\beta, \rho_t^\nu) \leq C|\beta - \nu| \quad \text{for a constant } C = C(H, g, T).$$

Then, for some sequence $\{\nu_n\}_{n=0}^\infty$ such that $\nu_n \rightarrow 0$, we would like to take the limit on both sides as $n \rightarrow \infty$ to argue that

$$W_2(\rho_t^\beta, \rho_t) = \lim_{n \rightarrow \infty} W_2(\rho_t^\beta, \rho_t^{\nu_n}) \leq C\beta.$$

However, the first equality is not immediately clear. In order to prove it, we leave as an exercise to the reader to prove that $\{X_t^\nu\}_{\nu \in (0, \beta]}$ is tight. By Prokhorov's theorem, there exists a sequence $\{\nu_{n_k}\}_{k=0}^\infty$ such that $X_t^{\nu_{n_k}} \rightarrow X_t$ in distribution, as $k \rightarrow \infty$. We would also expect that X_t^ν has uniformly integrable second moments, uniform in both $\nu \in (0, \beta]$ and $t \in [0, T]$. By [4, Proposition 7.1.5], these two facts are equivalent to the convergence of $\rho_t^{\nu_{n_k}}$ to ρ_t in W_2 . Thus, taking the limit as $k \rightarrow \infty$ yields the desired equality.

4. CONVERGENCE OF u^β

We first present a general stability result concerning HJ equations with respect to the supremum norm on compact sets. Similar results exist in the literature but are not directly applicable in our scenario ⁴. To informally describe our theorem: if two HJ PDEs are well-posed with solutions u^1 and u^2 , Hamiltonians H^1 and H^2 , terminal cost functions g^1 and g^2 , and viscosity parameters ν^1 and ν^2 , then the maximum difference between u^1 and u^2 on any compact set in \mathbb{R}^d is bounded by the difference in the coefficients (H^i, g^i, ν^i) , on that compact set. Our proof is inspired by that of [56, Lemma 6.3], but we apply their ideas to derive stability results for the Hamiltonian and the terminal cost function, not just for the viscosity parameter.

⁴Closely related results are [51, Propositions 1.4 and 2.1]. However, the former proposition assumed that the solution to their HJ equation was bounded and that it lacked a second-order term, and the latter assumed that the Hamiltonian was bounded in space and time. Another potentially applicable result was [56, Lemma 6.3]. But on account of their formulation not permitting for non-separable Hamiltonians, we also cannot directly apply this lemma.

Theorem 4.1. *Let $i = 1, 2$, $T > 0$, and $\nu^1, \nu^2 \geq 0$. Let $H^i : \mathbb{R}^d \times \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ and $g^i : \mathbb{R}^d \rightarrow \mathbb{R}$ be continuous. Suppose that H is uniformly convex in the p variable, in the sense of Assumption 2.3. If the HJ equation*

$$\begin{cases} -\partial_t u^i + H^i(x, -\nabla u^i, t) = \nu^i \Delta u^i \\ u^i(T, x) = g^i(x) \end{cases} \quad (4.1)$$

is well-posed with solution $u^i : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, then for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$, there exists a constant $C_{\mathcal{K}} = C(\mathcal{K}, H^1, H^2, g^1, g^2, T)$ such that

$$\|u^1 - u^2\|_{L^\infty([0, T] \times \mathcal{K})} \leq C_{\mathcal{K}} \left\{ \|H^1 - H^2\|_{L^\infty(\mathcal{K} \times \mathbb{R}^d \times [0, T])} + \|g^1 - g^2\|_{L^\infty(\mathcal{K})} + |\nu^1 - \nu^2|^{1/2} \right\}, \quad (4.2)$$

and $C_{\mathcal{K}}$ grows at most linearly in $\text{diam}(\mathcal{K}) := \sup\{|x - y| : x, y \in \mathcal{K}\}$.

Proof. Before anything else, we note that the uniform convexity of H implies that the Hamilton-Jacobi equation is satisfied in the classical sense, so the differentiation of u in the rest of the proof is well-defined.

Step 1: First, assume $\nu = \nu^1 = \nu^2$. Fix a compact set $\mathcal{K} \subseteq \mathbb{R}^d$, and define:

$$3\sigma = \sup_{(t, x) \in [0, T] \times \mathcal{K}} \{u^1(t, x) - u^2(t, x)\}.$$

Since \mathcal{K} is compact, this difference is finite and is achieved at some $(t_0, x_0) \in [0, T] \times \mathcal{K}$.

Step 2: Define the quadratic penalty

$$\varphi(t, x) = \frac{1}{2}(|t|^2 + |x|^2),$$

and the penalized difference $\Phi_\alpha : [0, T]^2 \times \mathcal{K}^2 \mapsto \mathbb{R}$ for some $\alpha > 0$:

$$\Phi_\alpha(t, x, s, y) = u^1(t, x) - u^2(s, y) - \sigma \frac{2T - s - t}{T} - \alpha \varphi(t - s, x - y).$$

Since $\Phi_\alpha \in \mathcal{C}(0, T]^2 \times \mathcal{K}^2)$, there exists $(t_\alpha, s_\alpha, x_\alpha, y_\alpha)$ such that

$$\Phi_\alpha(t_\alpha, s_\alpha, x_\alpha, y_\alpha) = \max_{t, s, x, y} \Phi_\alpha(t, x, s, y).$$

Step 3: Suppose that $t_\alpha, s_\alpha \in [0, T)$, i.e., the maximum of the penalized difference is not achieved at the terminal condition. Define the test function $\psi : [0, T] \times \mathcal{K} \rightarrow \mathbb{R}$ by

$$\psi(t, x) = -\sigma \frac{t}{T} + \alpha \varphi(t - s_\alpha, x - y_\alpha),$$

so that by definition of $(t_\alpha, s_\alpha, x_\alpha, y_\alpha)$, the function $(t, x) \mapsto u^1(t, x) - \psi(t, x)$ is maximized at (t_α, x_α) . So by the definition of a viscosity solution, we have that

$$\begin{aligned} & \frac{\sigma}{T} - \alpha \partial_t \varphi(t_\alpha - s_\alpha, x_\alpha - y_\alpha) + H^1(x_\alpha, -\nabla_x \psi(t_\alpha, x_\alpha), t_\alpha) \\ &= -\partial_t \psi(t_\alpha, x_\alpha) + H^1(x_\alpha, -\alpha \nabla_x \varphi(t_\alpha - s_\alpha, x_\alpha - y_\alpha), t_\alpha) \\ &\leq \nu \Delta_x \psi(t_\alpha, x_\alpha) = \nu \alpha \Delta_x \varphi(t_\alpha - s_\alpha, x_\alpha - y_\alpha). \end{aligned}$$

Define another test function $\phi : [0, T] \times \mathcal{K} \rightarrow \mathbb{R}$ by

$$\phi(s, y) = \sigma \frac{s}{T} - \alpha \varphi(t_\alpha - s, x_\alpha - y),$$

so that the function $(s, y) \mapsto u^2(s, y) - \phi(s, y)$ is minimized at (s_α, y_α) . Again, by the definition of a viscosity solution,

$$\begin{aligned} -\frac{\sigma}{T} + \alpha \partial_s \phi(t_\alpha - s_\alpha, x_\alpha - y_\alpha) + H^2(y_\alpha, -\alpha \nabla_y \phi(t_\alpha - s_\alpha, x_\alpha - y_\alpha), s_\alpha) \\ = -\partial_s \phi(s_\alpha, y_\alpha) + H^2(y_\alpha, -\nabla_y \phi(s_\alpha, y_\alpha), s_\alpha) \\ \geq \nu \Delta_y \phi(s_\alpha, y_\alpha) = \nu \alpha \Delta_y \phi(t_\alpha - s_\alpha, x_\alpha - y_\alpha). \end{aligned}$$

Combining the two inequalities and using the facts that $\partial_s \varphi = \partial_t \varphi$, $\nabla_x \varphi = \nabla_y \varphi$, and $\Delta_x \varphi = \Delta_y \varphi$, and writing $\partial_t \varphi = \partial_t \varphi(t_\alpha - s_\alpha, x_\alpha - y_\alpha)$ (and similarly for $\nabla \varphi$), we obtain:

$$\sigma \leq \frac{T}{2} \left\{ H^1(x_\alpha, -\alpha \nabla \varphi, t_\alpha) - H^2(y_\alpha, -\alpha \nabla \varphi, s_\alpha) \right\} + T \alpha \partial_t \varphi(t_\alpha - s_\alpha, x_\alpha - y_\alpha).$$

Using the triangle inequality, we have:

$$\begin{aligned} \sigma \leq \frac{T}{2} \left\{ |H^1(x_\alpha, -\alpha \nabla \varphi, t_\alpha) - H^1(y_\alpha, -\alpha \nabla \varphi, t_\alpha)| \right. \\ + |H^1(y_\alpha, -\alpha \nabla \varphi, t_\alpha) - H^1(y_\alpha, -\alpha \nabla \varphi, s_\alpha)| \\ + |H^1(y_\alpha, -\alpha \nabla \varphi, s_\alpha) - H^2(y_\alpha, -\alpha \nabla \varphi, s_\alpha)| \left. \right\} \\ + T \alpha |t_\alpha - s_\alpha|. \end{aligned} \tag{4.3}$$

Sending $\alpha \rightarrow \infty$, using Proposition 3.7 in [21] (with $\mathcal{O} = \mathcal{K}$, $\Phi(t, s, x, y) = u^1(t, x) - u^2(s, y)$, $\Psi(t, s, x, y) = |t - s| + |x - y|$), and using the uniform continuity of H^1 and H^2 when restricted to \mathcal{K} , we conclude that the first, second, and fourth terms in (4.3) converge to zero. The third term can be simplified to:

$$\begin{aligned} |H^1(y_\alpha, -\alpha \nabla \varphi, s_\alpha) - H^2(y_\alpha, -\alpha \nabla \varphi, s_\alpha)| &= |H^1(y_\alpha, -\alpha |x_\alpha - y_\alpha|, s_\alpha) - H^2(y_\alpha, -\alpha |x_\alpha - y_\alpha|, s_\alpha)| \\ &\leq \|H^1 - H^2\|_{L^\infty(\mathcal{K} \times \mathbb{R}^d \times [0, T])}. \end{aligned}$$

We therefore conclude that $\sigma \leq \frac{T}{2} \|H^1 - H^2\|_{L^\infty(\mathcal{K} \times \mathbb{R}^d \times [0, T])}$.

Step 4: Suppose that one of $t_\alpha, s_\alpha = T$ (without loss of generality, let $t_\alpha = T$). Then we have the following lower bound for Φ_α :

$$\begin{aligned} \Phi_\alpha(t_\alpha, s_\alpha, x_\alpha, y_\alpha) &\geq \Phi_\alpha(t_0, t_0, x_0, x_0) \\ &= u^1(t_0, x_0) - u^2(t_0, x_0) - \frac{\sigma}{T}(2T - 2t_0) - \alpha \varphi(0, 0) \\ &= 3\sigma - 2\sigma + 2\sigma(t_0/T) \geq \sigma. \end{aligned}$$

It follows that

$$\begin{aligned} \sigma &\leq \Phi_\alpha(t_\alpha, s_\alpha, x_\alpha, y_\alpha) \\ &= u^1(t_\alpha, x_\alpha) - u^2(s_\alpha, y_\alpha) - \frac{\sigma}{T}(2T - s_\alpha - t_\alpha) - \alpha \varphi(t_\alpha - s_\alpha, x_\alpha - y_\alpha) \\ &= u^1(T, x_\alpha) - u^1(T, y_\alpha) + u^1(T, y_\alpha) - u^2(T, y_\alpha) + u^2(T, y_\alpha) - u^2(s_\alpha, y_\alpha) \\ &\leq |g^1(x_\alpha) - g^1(y_\alpha)| + |g^1(x_\alpha) - g^2(y_\alpha)| + |g^2(y_\alpha) - u^2(s_\alpha, y_\alpha)| \\ &\leq |g^1(x_\alpha) - g^1(y_\alpha)| + \|g^1 - g^2\|_{L^\infty(\mathcal{K})} + |g^2(y_\alpha) - u^2(s_\alpha, y_\alpha)| \longrightarrow \|g^1 - g^2\|_{L^\infty(\mathcal{K})}, \end{aligned} \tag{4.4}$$

where the limit is taken as $\alpha \rightarrow \infty$, and follows from the uniform continuity of $\partial_t u^2$ and g^1 on the compact set \mathcal{K} . Combining the results from Step 3 and 4 yields

$$\sigma \leq \frac{T}{2} \|H^1 - H^2\|_{L^\infty(\mathcal{K} \times \mathbb{R}^d \times [0, T])} + \|g^1 - g^2\|_{L^\infty(\mathcal{K})}.$$

Step 5: Now, we allow $\nu^1 \neq \nu^2$. Define for $i = 1, 2$:

$$\begin{cases} -\partial_t v^i + H^2(x, -\nabla v^i, t) = \nu^i \Delta v^i, \\ v^i(T, x) = g^2(x). \end{cases}$$

By modifying the proof of [56, Lemma 6.3], and by taking $\Omega = \mathcal{K}$ and using the linear-in- x growth of ∇u , ∇u^β , and $\partial_t u$ (see, for example, [47, Lemma 3.4]), we find that there exists a constant $C = C(H, g, T)$ such that for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$,

$$\|v^1 - v^2\|_{L^\infty([0, T] \times \mathcal{K})} \leq C(1 + \text{diam}(\mathcal{K}))\sqrt{|\nu^1 - \nu^2|}.$$

Step 6: For $i = 1, 2$, suppose that \tilde{u}^1 and \tilde{u}^2 satisfy (4.1) with data (H^i, g^1, ν^1) , \tilde{w}^1 and \tilde{w}^2 satisfy (4.1) with data (H^2, g^i, ν^1) , and \tilde{v}^1 and \tilde{v}^2 satisfy (4.1) with data (H^2, g^2, ν^i) . Because $u^1 = \tilde{u}^1$, $u^2 = \tilde{v}^2$, $\tilde{u}^2 = \tilde{w}^1$, and $\tilde{w}^2 = \tilde{v}^1$, by the triangle inequality we get:

$$\begin{aligned} \|u^1 - u^2\|_{L^\infty([0, T] \times \mathcal{K})} &\leq \|\tilde{u}^1 - \tilde{u}^2\|_{L^\infty([0, T] \times \mathcal{K})} + \|\tilde{w}^1 - \tilde{w}^2\|_{L^\infty([0, T] \times \mathcal{K})} + \|\tilde{v}^1 - \tilde{v}^2\|_{L^\infty([0, T] \times \mathcal{K})} \\ &\leq C(1 + \text{diam}(\mathcal{K}))\{\|H^1 - H^2\|_{L^\infty(\mathcal{K} \times \mathbb{R}^d \times [0, T])} + \|g^1 - g^2\|_{L^\infty(\mathcal{K})} + \sqrt{|\nu^1 - \nu^2|}\}, \end{aligned}$$

which proves the desired result. \square

Now we combine the results in Section 3 with Theorem 4.1. The first-derivative bounds in Assumptions 2.1 and 2.5 were crucial, since they allowed the use of $W_1(\rho_t^\beta, \rho_t)$ being $O(\beta)$.

Theorem 4.2. *Under Assumptions 2.1–2.7, there exists a constant $C = C(H, g, T)$ such that for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$:*

$$\|u^\beta - u\|_{L^\infty([0, T] \times \mathcal{K})} \leq C(1 + \text{diam}(\mathcal{K}))\beta. \quad (4.5)$$

Proof. First, we apply Theorem 4.1 with a constant $C_{\mathcal{K}}$ growing at most linearly in $\text{diam}(\mathcal{K})$, $H^1(x, p, t) = H(x, p, \rho_t^\beta)$, $H^2(x, p, t) = H(x, p, \rho_t)$, $g^1(x) = g(x, \rho_T^\beta)$, $g^2(x) = g(x, \rho_T)$, $\nu^1 = \beta^2/2$, and $\nu^2 = 0$. Abbreviating the difference for H on $L^\infty(\mathcal{K} \times \mathbb{R}^d \times [0, T])$ and the difference for g on $L^\infty(\mathcal{K})$, we have:

$$\begin{aligned} \|u^\beta - u\|_{L^\infty([0, T] \times \mathcal{K})} &\leq C_{\mathcal{K}} \left\{ \|H(\cdot, \cdot, \rho^\beta) - H(\cdot, \cdot, \rho)\|_\infty + \|g(\cdot, \rho_t^\beta) - g(\cdot, \rho_T)\|_\infty + \frac{\beta}{\sqrt{2}} \right\} \\ &\leq C_{\mathcal{K}} \left\{ (\|\nabla_\mu H\|_\infty + \|\nabla_\mu g\|_\infty) \sup_{t \in [0, T]} W_1(\rho_t^\beta, \rho_t) + \frac{\beta}{\sqrt{2}} \right\} \\ &\leq (C_{\mathcal{K}} + (\|\nabla_\mu H\|_\infty + \|\nabla_\mu g\|_\infty)C + \frac{1}{\sqrt{2}})\beta, \end{aligned}$$

where we use the assumption that $\|\nabla_\mu H\|_\infty$ and $\|\nabla_\mu g\|_\infty$ are finite (from Assumptions 2.1 and 2.5), and Corollary 3.3. \square

5. APPLICATIONS

This section provides three applications of our result concerning N -player games, mean field control and policy iteration.

5.1. N -player games. MFGs arise as the limit of N -player games, as the number of players N increases to infinity. Although it is known in various circumstances [27, 30, 39] that the limit is the MFG equilibrium, finding the convergence rate is a separate and difficult problem. The twin papers [24, 25] comprise the most recent progress on determining the convergence rate. However, their results cannot be directly applied to the N -player convergence rate problem if the agents follow deterministic dynamics, because one of their assumptions, namely A.2 in both papers, is that the volatility coefficient Σ is non-degenerate⁵. Here we apply Corollary 3.3 to approximate the probability flow ρ_t of the first-order MFG by the empirical measures of an N -player system with non-degenerate volatility.

To simplify our discussion, we only consider the linear drift $b(t, x, a) = a$. So by [25, (2.6)], the value functions of all N players, $\{v^{N,i} : [0, T] \times (\mathbb{R}^d)^N \rightarrow \mathbb{R}\}_{i=1}^N$, satisfy the N -player system of PDEs:

$$\begin{cases} \partial_t v^{N,i}(t, x_i) + H(x_i, \nabla_x v^{N,i}(t, x_i), m_x^N) + \frac{1}{2} \sum_{j=1}^N \text{Tr}(\Sigma \Sigma^T \nabla_{x_j x_j}^2 v^{N,j}(t, x_j)) \\ \quad - \sum_{j \neq i} \langle \nabla_p H(x_j, \nabla_{x_j} v^{N,i}(t, x_j), m_x^N), \nabla_{x_i} v^{N,i}(t, x_i) \rangle = 0, \\ v^{N,i}(T, x) = g(x_i, m_x^N), \end{cases}$$

where $m_x^N := \frac{1}{N} \sum_{i=1}^N \delta_{x_i}$ is the empirical measure of $x = (x^1, \dots, x^N) \in (\mathbb{R}^d)^N$. Specializing to the case of $b(t, x, a) = a$, the i -th player's dynamics is:

$$dX_t^i = \alpha_t^i dt + \sigma dB_t^i = \nabla_p H(X_t^i, -\nabla u^\sigma(t, X_t^i), m_{X_t^i}^{N,\Sigma}) dt + \Sigma dB_t^i, \quad (5.1)$$

where $\{B^i\}_{i=1}^N$ are independent d -dimensional Brownian motions, and $m_{X_t^i}^{N,\Sigma}$ is the (random) empirical measure of the N -player system (5.1) at time $t \in [0, T]$:

$$m_{X_t^i}^{N,\Sigma} = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}. \quad (5.2)$$

Corollary 5.1. *Let ρ_t satisfy the Fokker-Planck equation in the MFG (1.1) with $\beta = 0$. Let $\Sigma = \beta I$, and denote by $m_{X_t^i}^{N,\beta}$ the empirical measure in (5.2) corresponding to $\Sigma = \beta I$. Under the assumptions in Corollary 3.3, there exist $C_1 = C_1(H, g, T)$ and $C_2 = C_2(\beta, H, g, T)$ such that for all $t \in [0, T]$,*

$$W_1(\rho_t, m_{X_t^i}^{N,\beta}) \leq C_1 \beta + C_2 N^{-\frac{1}{d+8}}. \quad (5.3)$$

Proof. Let ρ_t^β satisfy the Fokker-Planck equation in the MFG (1.1) with $\beta > 0$. We have:

$$W_1(\rho_t, m_{X_t^i}^{N,\beta}) \leq W_1(\rho_t, \rho_t^\beta) + W_1(\rho_t^\beta, m_{X_t^i}^{N,\beta}). \quad (5.4)$$

By [25, Theorem 3.1], there is a constant $C_2 = C_2(\beta, H, g, T)$ such that

$$\sup_{t \in [0, T]} W_1(\rho_t^\beta, m_{X_t^i}^{N,\beta}) \leq C_2 N^{-\frac{1}{d+8}}. \quad (5.5)$$

Combining the equations (5.4), (5.5) with Corollary 3.3 yields the desired bound. \square

⁵When we say that the volatility Σ is non-degenerate, we mean that its minimum eigenvalue is positive. Moreover, if the minimum eigenvalue of Σ is allowed to vanish, then their upper bounds for the distance between the probability distribution of the finite player system and that of the MFG limit become infinite.

As a result of Corollary 5.1, we obtain the population level to approximate the probability flow ρ_t of the first-order MFG via large player system. Assume that an accuracy of $\varepsilon > 0$ is needed, i.e., $W_1(\rho_t, m_{X_t}^{N,\beta}) \leq \varepsilon$. Then we set:

$$C_1\sigma \asymp \varepsilon \quad \text{and} \quad C_2(\sigma)N^{-\frac{1}{d+8}} \asymp \varepsilon. \quad (5.6)$$

Here we assume that (H, g, T) are given, so C_2 only depends on σ . A close scrutiny of the proofs (in particular, Equations 4.16 and 4.17) in [25] indicates that $C_2(\beta)$ blows up (in a rather complicated way), as $\beta \rightarrow 0$. So we first take $\sigma \asymp \varepsilon$, and then take $N \asymp (\varepsilon C_2^{-1}(\varepsilon))^{-(d+8)}$. That is, it requires $N (\varepsilon C_2^{-1}(\varepsilon))^{-(d+8)} \gg \varepsilon^{-(d+8)}$ players to approximate the probability flow of the first-order MFG with accuracy ε .

5.2. Mean field control. Next we consider a mean field control problem [23, Example 2], where a central planner seeks to control N particles by selecting a progressively measurable process $\alpha = (\alpha^1, \dots, \alpha^N) \in (\mathbb{R}^d)^N$. The dynamics of the i -th particle evolve as:

$$\begin{cases} dX_t^i = \alpha_t^i(X_t^i)dt + \beta dB_t^i & \text{for } t \in [t_0, T], \\ X_{t_0}^i = x_0^i, \end{cases}$$

where $\{B_t^i\}_{i=1}^N$ are independent d -dimensional Brownian motions. Denote the average state of the particles by $\bar{X}_t^N = \frac{1}{N} \sum_{i=1}^N X_t^i$, which is governed by

$$\begin{cases} d\bar{X}_t^N = \frac{1}{N} \sum_{i=1}^N \alpha_t^i dt + \frac{\beta}{\sqrt{N}} d\bar{B}_t & \text{for } t \in [t_0, T], \\ \bar{X}_{t_0}^N = \frac{1}{N} \sum_{i=1}^N x_0^i, \end{cases} \quad (5.7)$$

where $\bar{B}_t = \frac{\sum_{i=1}^N B_t^i}{\sqrt{N}}$ is d -dimensional Brownian motion. The objective of the central planner is to solve the optimization problem:

$$V^N(t_0, x_0) = \inf_{\alpha} \mathbb{E} \left[\int_{t_0}^T \frac{1}{N} \sum_{i=1}^N L(\alpha_t^i(X_t^i)) + F(\bar{X}_t) dt + G(\bar{X}_T) \Big| X_{t_0} = x_0 \right], \quad (5.8)$$

where $F, G : \mathbb{R}^d \rightarrow \mathbb{R}$ are assumed to be Lipschitz, and $L \in \mathcal{C}^2(\mathbb{R}^d)$ with $\frac{1}{C}I \leq \nabla^2 L \leq CI$ for some $C \geq 1$. An easy argument from [23] shows that the optimality in (5.8) is achieved by deterministic control, and $V^N(t, x) = v^N(t, \bar{m}_x^N)$, where v^N solves the HJ equation:

$$\begin{cases} -\partial_t v^N(t, x) + H(-\nabla v^N(t, x)) - F(x) = \frac{\beta^2}{2N} \Delta v^N(t, x), \\ v^N(T, x) = G(x), \end{cases} \quad (5.9)$$

where $H(-p)$ is the Legendre transform of L . By classical viscosity theory, v^N converges to v , which is the solution to the first-order equation:

$$\begin{cases} -\partial_t v(t, x) + H(-\nabla v(t, x)) - F(x) = 0, \\ v(T, x) = G(x). \end{cases} \quad (5.10)$$

Furthermore, $\sup_{[0, T] \times \mathbb{R}^d} |v^N - v| = \mathcal{O}(N^{-\frac{1}{2}})$.

Let $\mu_t^N := \text{Law}(\bar{X}_t^N)$ be the probability density of the average state \bar{X}_t^N . The following result specifies the limit of μ_t^N , as $N \rightarrow \infty$.

Corollary 5.2. *Let the aforementioned assumptions and those in Corollary 3.3 hold. Let $\{X_0^i\}_{i=1}^N$ be independent and identically distributed according to m_0 with bounded support⁶, and covariance matrix Σ . Then for all $t \in [0, T]$, μ_t^N converges to μ_t in W_1 , where μ_t is the solution to the equation:*

$$\begin{cases} \partial_t \mu_t + \operatorname{div}_x \{ \mu_t \nabla_p H(-\nabla v(t, x)) \} = 0, \\ \mu_0 \sim \delta_{\int x m_0(x) dx}. \end{cases} \quad (5.11)$$

Moreover, there exists a constant $C > 0$ (independent of N) such that

$$W_1(\mu_t^N, \mu_t) \leq C \left(\frac{1}{\sqrt{N}} + \frac{\sqrt{d \log N}}{N} + \sqrt{\frac{\operatorname{Tr} \Sigma}{N}} \right). \quad (5.12)$$

In particular, if $\Sigma = I$ then the bound (5.12) specializes to $\mathcal{O}(\sqrt{d/N})$, as $N, d \rightarrow \infty$.

Proof. First observe that (v^N, μ_t^N) solves the (degenerate) MFG:

$$\begin{cases} -\partial_t v^N(t, x) + H(-\nabla v^N(t, x)) - F(x) = \frac{\sigma^2}{N} \Delta v^N(t, x), \\ \partial_t \mu_t^N + \operatorname{div}_x \{ \mu_t^N \nabla H(-\nabla v^N(t, x)) \} = \frac{\sigma^2}{N} \Delta \mu_t^N, \\ v^N(T, x) = G(x), \quad \mu_0^N = \operatorname{Law}(\bar{X}_0^N), \end{cases} \quad (5.13)$$

where the HJ component does not involve the probability flow μ_t^N . Let $(\tilde{v}^N, \tilde{\mu}_t^N)$ be a solution to the MFG:

$$\begin{cases} -\partial_t \tilde{v}^N(t, x) + H(-\nabla \tilde{v}^N(t, x)) - F(x) = 0, \\ \partial_t \tilde{\mu}_t^N + \operatorname{div}_x \{ \tilde{\mu}_t^N \nabla H(-\nabla \tilde{v}^N(t, x)) \} = 0, \\ \tilde{v}^N(T, x) = G(x), \quad \tilde{\mu}_0^N = \operatorname{Law}(\bar{X}_0^N). \end{cases}$$

As a consequence of Corollary 3.3, we obtain:

$$W_1(\mu_t^N, \tilde{\mu}_t^N) \leq \frac{C}{\sqrt{N}} \quad \text{for some } C > 0 \text{ (independent of } N). \quad (5.14)$$

By classical viscosity theory, we know that $|\nabla v|$ is bounded (see e.g., [59, Theorem 1.9]). Combining with the fact that $\nabla_p H$ is continuous implies $(t, x) \rightarrow \nabla_p H(-\nabla v(t, x))$ is bounded. Further applying Grönwall's inequality, we get the stability estimate:

$$W_1(\tilde{\mu}_t^N, \mu_t) \leq C W_1 \left(\operatorname{Law}(\bar{X}_0^N), \delta_{\int x m_0(x) dx} \right), \quad (5.15)$$

for some $C > 0$ (independent of N). Without loss of generality, assume that X_0^i has mean 0, i.e., $\int x m_0(x) dx = 0$. We have:

$$\begin{aligned} W_1 \left(\operatorname{Law}(\bar{X}_0^N), \delta_0 \right) &\leq W_1 \left(\operatorname{Law}(\bar{X}_0^N), \mathcal{N} \left(0, \frac{\Sigma}{N} \right) \right) + W_1 \left(\mathcal{N} \left(0, \frac{\Sigma}{N} \right), \delta_0 \right) \\ &\leq \frac{C \sqrt{d \log N}}{N} + \sqrt{\frac{\operatorname{Tr} \Sigma}{N}}, \end{aligned} \quad (5.16)$$

⁶For $d = 1$, the assumption of bounded support can be removed, and $W_1(\mu_t^N, \mu_t) \leq C/\sqrt{N}$ for some $C > 0$ (independent of N). This is because the first term in the last inequality of (5.16) is bounded by C/\sqrt{N} , see the discussion after [54, Theorem 3.4], or [50].

where the first term in the last inequality follows from [28, Theorem 1] ⁷, and the second term is by the W_2 distance of two Gaussian vectors. Combining the equations (5.14), (5.15) and (5.16) yields the desired bound. \square

5.3. Policy iteration. As mentioned in the Introduction, there has been growing interest in first-order MFG models, but solving first-order MFGs numerically poses challenges.

Policy iteration (PI) is a class of approximate dynamic programming algorithms that have been used to solve stochastic control problems with provable guarantees [33, 37, 46, 60, 58]. In a series of papers [6, 8, 9], PI was proposed to solve second-order MFGs with separable Hamiltonians. An extension to second-order MFGs with non-separable Hamiltonians was considered in [43]. However, PI is not directly applicable to the first-order problems due to ill-posedness [55]. So a reasonable idea is to approximate first-order MFGs by second-order MFGs ⁸, and a convergence rate of second-order MFGs to the vanishing viscosity limit gives the approximation error.

Now, let us specify the PI for solving the MFG (1.1) with $\beta > 0$. For simplicity, we assume that the terminal data $g(x, \rho) = g(x)$ depend only on x . There are three steps: given $R > 0$ and a measurable function $q^0 : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ with $\|q^0\|_\infty \leq R$, we iterate for $n \geq 0$,

(i) Solve

$$\partial_t \rho_t^{n,\beta} - \operatorname{div}\{\rho_t^{n,\beta} q^n\} = \frac{\beta^2}{2} \Delta \rho_t^{n,\beta}, \quad \rho_0^{n,\beta} = m_0. \quad (5.17)$$

(ii) Solve

$$-\partial u^{n,\beta} + q^n \nabla u^{n,\beta} - \mathcal{L}(x, -\nabla u^{n,\beta}, q^n, \rho_t^{n,\beta}) = \frac{\beta^2}{2} \Delta u^{n,\beta}, \quad u^{n,\beta}(T, x) = g(x), \quad (5.18)$$

where $\mathcal{L}(x, p, q, \rho) := p \cdot q - H(x, p, \rho)$.

(iii) Update the policy

$$q^{n+1}(t, x) := \arg \max_{|q| \leq R} \left(q \cdot \nabla u^{n,\beta}(t, x) - \mathcal{L}(x, q, \rho_t^{n,\beta}) \right), \quad (5.19)$$

where $\mathcal{L}(x, q, \rho) := \max_p \mathcal{L}(x, p, q, \rho)$.

In all of the aforementioned works [6, 8, 9, 43], the convergence (rate) of PI (5.17)–(5.19) for MFGs was proved on the torus $\mathbb{R}^d/\mathbb{Z}^d$, rather than the whole space \mathbb{R}^d to avoid boundary effects. Nevertheless, a review of the methods in these papers allow to prove the convergence of PI for solving MFGs on \mathbb{R}^d . The extension is technical, and goes beyond the scope of this paper. The claim below, extending [43], summarizes the “expected” convergence results of PI for solving second-order MFGs on \mathbb{R}^d . We plan to prove it rigorously in the future.

Claim 5.3. *Under suitable conditions on $H(x, p, \rho)$, $m_0(x)$ and $g(x)$ (e.g., H and its derivatives are Lipschitz and H is strictly convex in p , and m_0, g have some Sobolev regularity), for any compact set $\mathcal{K} \subset \mathbb{R}^d$, there exists $T = T(\mathcal{K}, \beta) > 0$ and $C = C(\mathcal{K}, \beta)$ such that*

$$\|u^{n,\beta} - u^\beta\|_{W_r^{1,2}([0,T] \times \mathcal{K})} + \|\rho^{n,\beta} - \rho^\beta\|_{W_r^{1,2}([0,T] \times \mathcal{K})} \leq C e^{-n}, \quad \text{for } r > d + 2, \quad (5.20)$$

⁷A slightly looser bound $\mathcal{O}(\sqrt{d} \log N/N)$ (up to a $\log N$ factor) was proved in [62, Theorem 1.1].

⁸This idea was also proposed in [55] to solve deterministic control problems by PI.

where $W_r^{1,2}(Q)$ denotes the space of functions f such that $\partial_t^\delta \partial_x^\sigma f \in L^r(Q)$ for all multi-indices (δ, σ') with $2\delta + \sigma' \leq 2$, and

$$\|f\|_{W_r^{1,2}(Q)} := \left(\int_Q \sum_{2\delta + \sigma' \leq 2} |\partial_t^\delta \partial_x^{\sigma'} f|^r dt dx \right)^{\frac{1}{r}}.$$

The constants $T(\mathcal{K}, \beta), C(\mathcal{K}, \beta) > 0$ depend on \mathcal{K}, β in a complicated way. Given \mathcal{K} and as $\beta \rightarrow 0$, $C(\mathcal{K}, \beta)$ is typically of order $e^{\frac{C}{\beta^2}}$ for some $C > 0$, and $T(\mathcal{K}, \beta)$ is typically of order $\beta^{-\kappa}$ for some $\kappa > 0$.

With Claim 5.3 in place, we derive the (time-weighted) convergence rate of $u^{\beta,n}$ to u by simply applying the triangle inequality.

Corollary 5.4. *Let $\mathcal{K} \in \mathbb{R}^d$ be a compact set. Under the assumptions in Theorem 4.2 and Claim 5.3, there exist $T = T(\beta) > 0$, $C_1 = C_1(\mathcal{K})$ and $C_2 = C_2(\beta)$ such that*

$$\frac{1}{T} \|u^{n,\beta} - u\|_{L^r([0,T] \times \mathcal{K})} \leq C_1 \beta + C_2(\beta) e^{-n} \quad \text{for } r > d + 2. \quad (5.21)$$

As a consequence of Corollary 5.4, we get the complexity of PI for solving the first-order MFGs. Assume that an accuracy of $\varepsilon > 0$ is required, i.e., $\frac{1}{T} \|u^{n,\beta} - u\|_{L^r([0,T] \times \mathcal{K})} \leq \varepsilon$. Then we set:

$$\beta \asymp \varepsilon \quad \text{and} \quad C_2(\beta) e^{-n} \asymp \varepsilon, \quad (5.22)$$

so $n \asymp \log(C(\varepsilon)/\varepsilon)$. The discussion at the end of Claim 5.3 suggests that $C_2(\varepsilon)$ be of order $e^{\frac{C}{\varepsilon^2}}$ for some $C > 0$, as $\varepsilon \rightarrow 0$. Therefore, we have $n \asymp \varepsilon^{-2}$, i.e., it takes the order of ε^{-2} steps for PI to approximate u^0 with accuracy ε .

6. EXAMPLES AND NUMERICAL RESULTS

6.1. A closed-form example. As mentioned in the Introduction, the convergence rate of the vanishing viscosity approximation to MFGs matches the optimal rate of that to HJ equations, so it is hard to expect a better rate in the general setting. Nevertheless, this does not rule out some MFGs with special structures, which may have sharper rates for vanishing viscosity approximations.

Consider the following example from [7, 14]:

$$\begin{cases} -\partial_t u^\beta + \frac{1}{2} |\nabla u^\beta|^2 - \frac{1}{2} \left(x - \int y \rho_t^\beta(y) dy \right)^2 = \frac{\beta^2}{2} \Delta u & \text{on } [0, T] \times \mathbb{R}^d, \\ \partial_t \rho_t^\beta - \operatorname{div}_x \{ \rho_t^\beta \nabla u^\beta \} = \frac{\beta^2}{2} \Delta \rho_t^\beta & \text{on } [0, T] \times \mathbb{R}^d, \\ u^\beta(T, x) = 0, \quad \rho_0^\beta(x) \sim \mathcal{N}(m, \sigma^2 I) & \text{on } \mathbb{R}^d. \end{cases} \quad (6.1)$$

That is, the MFG (6.1) has a nonlocal and separable Hamiltonian

$$H(x, p, \mu) = \frac{1}{2} |p|^2 - \frac{1}{2} \left(x - \int_y y \mu(y) dy \right)^2 \quad (6.2)$$

with $g(x, \mu) = 0$ and $m_0(x)$ being Gaussian with mean m and covariance matrix $\sigma^2 I$. Interestingly, this MFG has a closed-form solution:

$$u^\beta(t, x) = \frac{e^{2T-t} - e^t}{2(e^{2T-t} + e^t)} |x - m|^2 - \frac{\beta^2 d}{2} \ln \left(\frac{2e^T}{e^{2T-t} + e^t} \right), \quad (6.3)$$

and

$$\rho_t^\beta(x) \sim \mathcal{N} \left(m, \left(\sigma^2 \left(\frac{e^{2T-t} + e^t}{e^{2T} + 1} \right)^2 + \beta^2 \frac{(e^{2T-t} + e^t)^2 (e^{2t} - 1)}{2(e^{2T} + 1)(e^{2T} + e^{2t})} \right) I \right). \quad (6.4)$$

As a consequence,

$$\|u^\beta - u\|_\infty \leq C\beta^2 \quad \text{and} \quad W_1(\rho_t^\beta, \rho_t) \leq C\beta^2, \quad (6.5)$$

for some $C > 0$ (independent of β). The same rate $\mathcal{O}(\beta^2)$ for vanishing viscosity can also be extended to a class of displacement monotone MFGs by using the arguments in [18].

6.2. Numerical examples. We have proved in Theorem 4.2 that u^β of MFGs with a nonlocal Hamiltonian converges at a rate of $\mathcal{O}(\beta)$. Here we compare the rate to that of MFGs with a local coupling.

We consider the following example on $[0, 0.25] \times \mathbb{T}^1$ (i.e., $T = 0.25$) with:

$$H(x, p, \mu(x)) = 0.01 \left\{ |p|^2 - \mu(x)^2 - \cos(4\pi x) - 0.1 \cos(2\pi x) - 0.1 \sin \left(2\pi \left(x - \frac{\pi}{8} \right)^2 \right) \right\}, \quad (6.6)$$

and $g(x) = 0$, and m_0 being Gaussian center at 0 with variance 0.01 truncated to have Dirichlet boundary conditions. Figure 1 plots the solutions to this local and separable MFG, with $\beta \in \{0.1, 0.3, 0.5, 1.0\}$, and Figure 2 illustrates how $\|u^\beta - u\|_\infty$ varies against β (for $\beta \in \{0.1, 0.2, \dots, 0.9, 1\}$). Observe that the dependence $\|u^\beta - u\|_\infty$ on β is much less, as β gets larger.

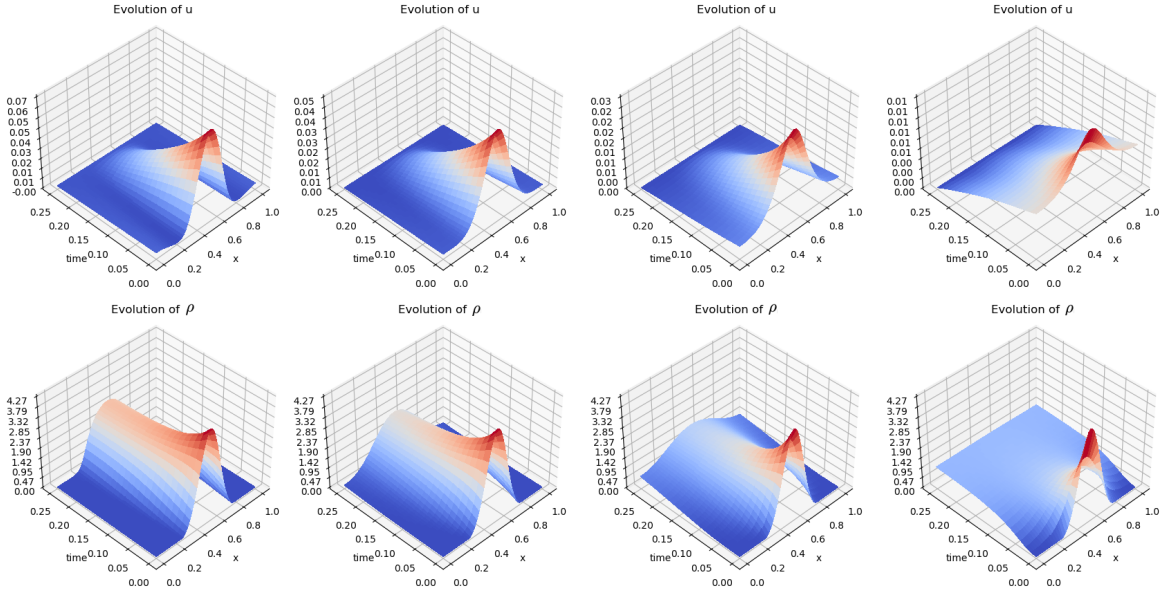
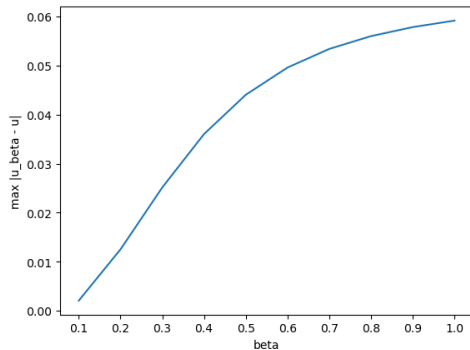


FIGURE 1. Plot of (u^β, ρ^β) for $\beta \in \{0.1, 0.3, 0.5, 1.0\}$ (left to right).

In [56], it was proved that u^β converges at a rate of $\mathcal{O}(\beta^{\frac{1}{4}})$ in some weighted L^2 norm. Now by regressing $\log \|u^\beta - u\|_\infty$ over $\log \beta$, we find that the slope is 1.050 using all $\beta \in \{0.1, \dots, 1.0\}$, while the slope is 1.162 using the first half $\beta \in \{0.1, \dots, 0.5\}$. It is natural to expect that

$$\|u^\beta - u\|_\infty \asymp \beta^{1+\delta} \quad \text{as } \beta \rightarrow 0, \quad (6.7)$$

FIGURE 2. Plot of $\|u^\beta - u\|_\infty$ against β .

for some $0 < \delta < 1$. The rate (6.7) is better than the proved $\mathcal{O}(\beta^{\frac{1}{4}})$ -rate for MFGs with a local Hamiltonian, and is between the $\mathcal{O}(\beta)$ -rate for MFGs with a general nonlocal Hamiltonian and the $\mathcal{O}(\beta^2)$ -rate for the example in Section 6.1. An interesting question is to find suitable conditions on model data to achieve the rate in (6.7) (with an explicit δ), hence improving the bounds in [56].

7. CONCLUSION

This paper studies the convergence rate of the vanishing viscosity approximation to MFGs with a nonlocal, and possibly non-separable Hamiltonian. With β^2 as the diffusivity constant, we prove that u^β and ρ^β converge a rate of $\mathcal{O}(\beta)$ in the topology of uniform convergence on compact sets and the W_1 metric, respectively. Our approach exploits both probabilistic and analytical arguments, where an FBSDE representation of the MFG is used to derive the convergence rate of ρ^β , and the rate of u^β follows by a stability property of the HJ equation. We also apply our result to N -player games, mean field control, and policy iteration for MFGs.

There are several directions to extend this work:

- (1) First, our result is proved for MFGs with a nonlocal and possibly non-separable Hamiltonian. It would be interesting to establish the convergence result for MFGs with a local Hamiltonian, underpinning the numerical results in Section 6.2.
- (2) Second, we prove in this work the convergence rate of vanishing viscosity for MFGs in \mathbb{R}^d ; while [54] considered the case in \mathbb{T}^d . The main difference between these two papers is that our work uses an FBSDE representation of the MFG together with a PDE stability result, while [54] relies exclusively on PDE arguments. A natural question is whether the FBSDE approach can be extended to other domains, so that the convergence can be established for MFGs on domains other than \mathbb{T}^d and \mathbb{R}^d .
- (3) Finally, the vanishing viscosity approximation to MFGs can be regarded as a “perturbation” of MFGs, where the perturbation is to add the operator $\frac{\beta^2}{2}\Delta$. We expect that the tools in this paper can also be used to analyze other types of perturbation, e.g., perturbation on the Hamiltonian. A notable example is the entropy-regularized relaxed control [61] in the context of reinforcement learning, where the HJ equation is replaced with the exploratory equation under entropy regularization [57].

Acknowledgment: W. Yu is supported by the Columbia Innovation Hub grant and NSF grant DMS-2309245. Q. Du is supported by NSF grant DMS-2309245. W. Tang is supported by NSF grant DMS-2206038, the Columbia Innovation Hub grant, and the Tang Family Assistant Professorship.

REFERENCES

- [1] Y. Achdou, F. J. Buera, J.-M. Lasry, P.-L. Lions, and B. Moll. Partial differential equation models in macroeconomics. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 372(2028):20130397, 19, 2014.
- [2] Y. Achdou and M. Laurière. Mean field games and applications: numerical aspects. In *Mean field games*, volume 2281 of *Lecture Notes in Math.*, pages 249–307. Springer, Cham, 2020.
- [3] Y. Achdou and A. Porretta. Mean field games with congestion. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 35(2):443–480, 2018.
- [4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [5] M. Bansil, A. R. Mészáros, and C. Mou. Global well-posedness of displacement monotone degenerate mean field games master equations. *SIAM J. Control Optim.*, 63(2):993–1021, 2025.
- [6] S. Cacace, F. Camilli, and A. Goffi. A policy iteration method for mean field games. *ESAIM Control Optim. Calc. Var.*, 27:Paper No. 85, 19, 2021.
- [7] E. Calzola, E. Carlini, and F. J. Silva. A high-order scheme for mean field games. *J. Comput. Appl. Math.*, 445:Paper No. 115769, 19, 2024.
- [8] F. Camilli. A policy iteration method for mean field games. *IFAC-PapersOnLine*, 55(30):406–411, 2022.
- [9] F. Camilli and Q. Tang. Rates of convergence for the policy iteration method for mean field games systems. *J. Math. Anal. Appl.*, 512(1):Paper No. 126138, 18, 2022.
- [10] P. Cardaliaguet. Weak solutions for first order mean field games with local coupling. In *Analysis and geometry in control theory and its applications*, volume 11 of *Springer INdAM Ser.*, pages 111–158. Springer, Cham, 2015.
- [11] P. Cardaliaguet and J. P. Graber. Mean field games systems of first order. *ESAIM Control Optim. Calc. Var.*, 21(3):690–722, 2015.
- [12] P. Cardaliaguet and A. Porretta. An introduction to mean field game theory. In *Mean field games*, volume 2281 of *Lecture Notes in Math.*, pages 1–158. Springer, Cham, 2020.
- [13] E. Carlini and F. J. Silva. A fully discrete semi-Lagrangian scheme for a first order mean field game problem. *SIAM J. Numer. Anal.*, 52(1):45–67, 2014.
- [14] E. Carlini, F. J. Silva, and A. Zorkot. A Lagrange-Galerkin scheme for first order mean field game systems. *SIAM J. Numer. Anal.*, 62(1):167–198, 2024.
- [15] R. Carmona. Applications of mean field games in financial engineering and economic theory. In *Mean field games*, volume 78 of *Proc. Sympos. Appl. Math.*, pages 165–219. Amer. Math. Soc., Providence, RI, 2021.
- [16] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. I*, volume 83 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. Mean field FBSDEs, control, and games.

- [17] R. Carmona and D. Lacker. A probabilistic weak formulation of mean field games and applications. *Ann. Appl. Probab.*, 25(3):1189–1231, 2015.
- [18] M. Cirant and A. R. Mészáros. Long time behavior and stabilization for displacement monotone mean field games. *arXiv:2412.14903*, 2024.
- [19] M. Colombo, G. Crippa, M. Graff, and L. V. Spinolo. On the role of numerical viscosity in the study of the local limit of nonlocal conservation laws. *ESAIM Math. Model. Numer. Anal.*, 55(6):2705–2723, 2021.
- [20] M. Colombo, G. Crippa, and L. V. Spinolo. On the singular local limit for conservation laws with nonlocal fluxes. *Arch. Ration. Mech. Anal.*, 233(3):1131–1167, 2019.
- [21] M. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1), 1992.
- [22] M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277(1):1–42, 1983.
- [23] S. Daudin, F. Delarue, and J. Jackson. On the optimal rate for the convergence problem in mean field control. *J. Funct. Anal.*, 287(12):Paper No. 110660, 94, 2024.
- [24] F. Delarue, D. Lacker, and K. Ramanan. From the master equation to mean field game limit theory: a central limit theorem. *Electron. J. Probab.*, 24:1–54, 2019.
- [25] F. Delarue, D. Lacker, and K. Ramanan. From the master equation to mean field game limit: large deviations and concentration of measure. *Ann. Probab.*, 48(1):211–263, 2020.
- [26] B. Djehiche, A. Tchekam, and H. Tembine. Mean-field-type games in engineering. *AIMS Electronics and Electrical Engineering*, 1(1):18–73, 2017.
- [27] M. F. Djete. Extended mean field control problem: a propagation of chaos result. *Electron. J. Probab.*, 27:1–53, 2022.
- [28] R. Eldan, D. Mikulincer, and A. Zhai. The CLT in high dimensions: quantitative bounds via martingale embedding. *Ann. Probab.*, 48(5):2494–2524, 2020.
- [29] R. Ferreira, D. Gomes, and T. Tada. Existence of weak solutions to time-dependent mean-field games. *Nonlinear Anal.*, 212:Paper No. 112470, 31, 2021.
- [30] M. Fischer. On the connection between symmetric n-player games and mean field games. *Ann. Appl. Probab.*, 27(2):757–810, 2022.
- [31] W. Gangbo, A. R. Mészáros, C. Mou, and J. Zhang. Mean field games master equations with nonseparable Hamiltonians and displacement monotonicity. *Ann. Probab.*, 50(6):2178–2217, 2022.
- [32] P. J. Graber and A. R. Mészáros. On monotonicity conditions for mean field games. *J. Funct. Anal.*, 285(9):Paper No. 110095, 45, 2023.
- [33] X. Guo, H. V. Tran, and Y. P. Zhang. Policy iteration for nonconvex viscous Hamilton–Jacobi equations. *arXiv:2503.02159*, 2025.
- [34] K. Huang, X. Di, Q. Du, and X. Chen. A game-theoretic framework for autonomous vehicles velocity control: Bridging microscopic differential games and macroscopic mean field games. *Discrete Contin. Dyn. Syst. Ser. B*, 25(12):4869–4903, 2020.
- [35] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ -Nash equilibria. *IEEE Trans. Automat. Control*, 52(9):1560–1571, 2007.
- [36] M. Huang, R. Malhame, and P. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–252, 2006.

- [37] B. Kerimkulov, D. Siska, and L. Szpruch. Exponential convergence and stability of Howard’s policy improvement algorithm for controlled diffusions. *SIAM J. Control Optim.*, 58(3):1314–1340, 2020.
- [38] D. Lacker. Mean field games via controlled martingale problems: existence of Markovian equilibria. *Stochastic Process. Appl.*, 125(7):2856–2894, 2015.
- [39] D. Lacker. On the convergence of closed-loop nash equilibria to the mean field game limit. *Ann. Appl. Probab.*, 30(4):1693–1761, 2020.
- [40] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.
- [41] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II – horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.
- [42] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
- [43] M. Laurière, J. Song, and Q. Tang. Policy iteration method for time-dependent mean field games systems with non-separable Hamiltonians. *Appl. Math. Optim.*, 87(17), 2023.
- [44] C.-T. Lin and E. Tadmor. L^1 -stability and error estimates for approximate Hamilton-Jacobi solutions. *Numer. Math.*, 87(4):701–735, 2001.
- [45] S. Liu, M. Jacobs, W. Li, L. Nurbekyan, and S. J. Osher. Computational methods for first-order nonlocal mean field games with applications. *SIAM J. Numer. Anal.*, 59(5):2639–2668, 2021.
- [46] J. Ma, G. Wang, and J. Zhang. Convergence analysis for entropy-regularized control problems: a probabilistic approach. *arXiv:2406.10959*, 2024.
- [47] A. R. Mészáros and C. Mou. Mean field games systems under displacement monotonicity. *SIAM J. Math. Anal.*, 56(1):529–553, 2024.
- [48] B. Moll. Mean field games in macroeconomics. *Econ Meets Math Workshop*, 2019. Available at https://benjaminmoll.com/wp-content/uploads/2020/10/MFG_macro.pdf.
- [49] C. Mou and J. Zhang. Mean field game master equations with anti-monotonicity conditions. *arXiv:2201.10762. To appear in J. Eur. Math. Soc.*, 2024.
- [50] E. Rio. Upper bounds for minimal distances in the central limit theorem. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):802–817, 2009.
- [51] P. E. Souganidis. Existence of viscosity solutions of Hamilton-Jacobi equations. *J. Differential Equations*, 56(5):345–390, 1985.
- [52] E. Tadmor. Numerical viscosity and the entropy condition for conservative difference schemes. *Math. Comp.*, 43(168):369–381, 1984.
- [53] W. Tang. Trading and wealth evolution in the proof of stake protocol. In *Proof-of-Stake for Blockchain Networks: Fundamentals, Challenges and Approaches*, pages 135–161. IET, 2024.
- [54] W. Tang and F. Tang. The Poisson binomial distribution—old & new. *Statist. Sci.*, 38(1):108–119, 2023.
- [55] W. Tang, H. V. Tran, and Y. P. Zhang. Policy iteration for the deterministic control problems—a viscosity approach. *SIAM J. Control Optim.*, 63(1):375–401, 2025.
- [56] W. Tang and Y. P. Zhang. The convergence rate of vanishing viscosity approximations for mean field games. *arXiv:2303.14560. To appear in SIAM J. Math. Anal.*, 2025.
- [57] W. Tang, Y. P. Zhang, and X. Y. Zhou. Exploratory HJB equations and their convergence. *SIAM J. Control Optim.*, 60(6):3191–3216, 2022.
- [58] W. Tang and X. Y. Zhou. Regret of exploratory policy improvement and q -learning. *arXiv:2411.01302*, 2024.

- [59] H. V. Tran. *Hamilton-Jacobi equations—theory and applications*, volume 213 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2021.
- [60] H. V. Tran, Z. Wang, and Y. P. Zhang. Policy iteration for exploratory Hamilton–Jacobi–Bellman equations. *Appl. Math. Optim.*, 91(2):Paper No. 50, 2025.
- [61] H. Wang, T. Zariphopoulou, and X. Y. Zhou. Reinforcement learning in continuous time and space: A stochastic control approach. *J. Mach. Learn. Res.*, 21:1–34, 2020.
- [62] A. Zhai. A high-dimensional CLT in \mathcal{W}_2 distance with near optimal convergence rate. *Probab. Theory Related Fields*, 170(3-4):821–845, 2018.

DEPARTMENT OF APPLIED PHYSICS AND APPLIED MATHEMATICS, COLUMBIA UNIVERSITY.

Email address: wy2436@columbia.edu

DEPARTMENT OF APPLIED PHYSICS AND APPLIED MATHEMATICS, COLUMBIA UNIVERSITY.

Email address: qd2125@columbia.edu

DEPARTMENT OF INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH, COLUMBIA UNIVERSITY.

Email address: wt2319@columbia.edu