

BEYOND SEPARABILITY: CONVERGENCE RATE OF VANISHING VISCOSITY APPROXIMATIONS TO MEAN FIELD GAMES VIA FBSDE STABILITY

WINSTON YU, QIANG DU, AND WENPIN TANG

ABSTRACT. We study 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 HJB 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 [45, 46, 47], and by Huang, Malhame, and Caines in [41], 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 infinitesimally small influence on all of the other players. Moreover, if we assume 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), then 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-Bellman (HJB) 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 $\mathbb{R}^d$:

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

where $\beta \geq 0$ is the idiosyncratic noise intensity¹, H is a possibly non-separable Hamiltonian, g is the terminal cost, and m_0 is the initial distribution. When $\beta > 0$, the system is of second

Date: July 27, 2026.

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

order. The system obtained by sending $\beta \rightarrow 0$ is called the *vanishing viscosity limit*, 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, 16, 53] and engineering [28, 40]. 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 pedestrian crowds and autonomous vehicles [35, 39], and Proof-of-Stake cryptocurrency mining [59, Section 5]. As prior works have noted, such as in [39], traditional methods (e.g., Newton's method) for solving first-order MFGs tend to be numerically unstable. Various approaches [14, 15, 50] 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 Hamilton-Jacobi equations (see e.g., [24, Section IV]) that as $\beta \rightarrow 0$, $u^\beta \rightarrow u$ in some topology. But the classical theory of viscosity solutions cannot immediately 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 to vanishing viscosity limit of the MFG (1.1) with a general, possibly non-separable Hamiltonian. Previous work [62] has 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 that under suitable conditions on the model parameters, 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 $C_{\mathcal{K}}$. Under certain regularity conditions, we can show that $C_{\mathcal{K}}$ grows at most quadratically in the diameter of $\mathcal{K}$. In other words, $u^\beta \rightarrow u$ at a rate of $\mathcal{O}(\beta)$ in the topology of uniform convergence on compact sets (Theorem 4.4), which matches the convergence rate from the classical viscosity theory of Hamilton-Jacobi equations. As intermediary steps, we prove: (1) when the initial condition is bounded, ∇u^β converges to ∇u at a rate of $\mathcal{O}(\beta)$ in the $L^\infty([0, T] \times U)$ metric, for any large enough bounded set $U \subseteq \mathbb{R}^d$ (Theorem 3.3), as well as in the $L^2(\rho_t)$ metric, uniformly in t (Corollary 3.5), and (2) ρ_t^β converges to ρ_t in the 2-Wasserstein distance (Corollary 3.4). Our analysis specifically requires neither the Lasry-Lions nor the displacement monotonicity condition; see Remark 2.8 for a discussion. In Section 5, we also show how to apply our result to various problems, such as particle system approximations and policy iteration for solving MFGs.

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 might 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

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., [22, 23, 58]).

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:

- (1) Using the FBSDE representation of the MFG, and temporarily assuming that the initial condition is bounded, we control the L^2 difference between ρ_t^β and ρ_t in terms of β and $\|\nabla u^\beta - \nabla u\|_{L^\infty(U)}$ (Lemma 3.2) for some large enough set $U \subseteq \mathbb{R}^d$.
- (2) Then, using the decoupling field of the FBSDE, we prove that $\|\nabla u^\beta - \nabla u\|_{L^\infty(U)} = \mathcal{O}(\beta)$ (Theorem 3.3). To remove the assumption that the initial condition is bounded, we apply a stability result for FBSDEs. This implies the convergence of ρ_t^β to ρ_t at a rate of $\mathcal{O}(\beta)$ in the L^2 , and hence W_2 , metric (Corollary 3.4).
- (3) We finally apply the previous two steps, in combination with a general PDE stability result (Theorem 4.1), to the PDE formulation of MFGs, in order to derive a convergence rate of $\mathcal{O}(\beta)$ for u^β to u in the topology of uniform convergence on compact sets (Theorem 4.4).

Finally, we comment that while we mostly use the L^2 and W_2 distances in the statements of our results, we expect that our results should easily extend to L^p and W_p distances when the initial distribution m_0 is only p -integrable, for $p \in [1, 2)$.

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. We make some concluding remarks 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 $\beta = 0$, [11, 12, 31] are some of the major works proving the well-posedness. In the separable, nonlocal case and when $\beta = 0$, [13] proved 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, 53], 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 [59], 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 [33], and later [52], 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. [18] 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 [43] 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 [17] for further developments in this direction.

While the classical setting of the convergence rate of vanishing viscosity approximations to pure Hamilton-Jacobi equations has been studied extensively, two recent papers [20, 19] were motivated by applications to mean-field control to provide an even sharper convergence rate of $\mathcal{O}(\beta^2 \log(\beta^2))$. Central to both papers is an estimate of the integral of the Laplacian of the value function, with respect to the solution of an adjoint equation [49], over $\mathbb{T}^d$ in [20] and $\mathbb{R}^d$ in [19]. Although they do not apply their results to mean field games, especially ones with a more general, non-separable Hamiltonian (however, in their setting, "non-separability" would mean that the Hamiltonian's momentum and time arguments cannot be separated like Equation (1.2)), it would be interesting to apply their technique to our setting as well.

Comparison to previous work: The only previous work addressing the convergence rate of vanishing viscosity approximations to MFGs is [62]. 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 $L^\infty([0, T] \times \mathcal{K})$ 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 [62] for MFGs with nonlocal and possibly non-separable Hamiltonians.

2. NOTATIONS, ASSUMPTIONS AND PROBLEM FORMULATION

2.1. Notations. For a metric space X , let $C^k(X)$ be the space of functions mapping X to $\mathbb{R}$, which are k -times differentiable and whose k -th order derivatives are continuous. $C_c^k(X)$ is the subset of $C^k(X)$ whose functions are compactly supported. 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 $p \in [1, \infty]$, a generic measure space $(X, \mathcal{B}, \mu)$, and a metric space $(Y, |\cdot|)$ (which will almost always be Euclidean space $\mathbb{R}^n$ in this paper), $L^p(X, \mathcal{B}, \mu; Y)$ is the space of Y -valued functions whose p -th power is integrable with respect to μ , i.e., all $f : X \rightarrow \mathbb{R}$ such that $\|f\|_{L^p}^p = \int_X |f|^p d\mu < \infty$. 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 $\|f\|_{L^\infty(X)}$. If we omit a σ -field $\mathcal{B}$, then it should be clear from context whether

it is the Borel σ -field or an element of some filtration generated by a stochastic process, for instance. We might also omit specifying the measure μ when it is clear whether it is, for example, Lebesgue measure on $\mathbb{R}^d$ or a probability measure $\mathbb{P}$ on some sample space Ω . If we do not specify a metric space Y , then it should be taken $\mathbb{R}$. 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 concept of the Wasserstein gradient of a function $U : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$. For a more extensive introduction, please refer to [17, Chapter 5], [13, Section 1.4], or [10, Section 6]. For $\mu \in \mathcal{P}_2(\mathbb{R}^n)$, the Wasserstein gradient of U at μ is denoted by $\nabla_\mu U(\mu, \cdot) : \mathbb{R}^n \mapsto \mathbb{R}$, and it is an element of the closure of gradients of $C^\infty(\mathbb{R}^n)$ functions, with respect to the $L^2(\mathbb{R}^n, \mu)$ metric. Moreover, it follows from [13, Definition 1.11]³ that $\nabla_\mu U$ satisfies: for all $\mathbb{R}^n$ -valued L^2 random variables ξ and η ,

$$U(\text{Law}(\xi + \eta)) - U(\text{Law}(\xi)) = \mathbb{E}[\langle \nabla_\mu U(\text{Law}(\xi), \xi), \eta \rangle] + o(\mathbb{E}[|\eta|^2]^{1/2}). \quad (2.1)$$

See also [34, 67] for related discussions.

Finally, for $a, b > 0$, the symbol $a = \mathcal{O}(b)$, or $a \lesssim b$ 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. Unless otherwise said, we work on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, $\mathbb{F} = \{\mathcal{F}_t\}_{t \in [0, T]}$, generated by a standard d -dimensional Brownian motion B . 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 , as well as a convexity assumption on H for the FBSDE representation (2.3) and a well-posedness assumption.

Assumption 2.1 (Regularity of H). *The derivatives $\nabla_\mu H$, $\nabla_{xx}^2 H$, $\nabla_{xp}^2 H$, $\nabla_{x\mu}^2 H$, $\nabla_{pp}^2 H$, 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:*

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

³The Wasserstein gradient is defined slightly differently in [13, Definition 1.11], but Equation 2.1 can be recovered by taking the transport plan π to be $(\xi, \xi + \eta)$ in their notation.

$\nabla_x H$ 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 $x, p \in \mathbb{R}^d$,

$$\begin{aligned} |\nabla_x H(x, p, \mu_1) - \nabla_x H(x, p, \mu_2)| &\leq \|\nabla_{x\mu}^2 H\|_\infty W_1(\mu_1, \mu_2), \\ |\nabla_p H(x, p, \mu_1) - \nabla_p H(x, p, \mu_2)| &\leq \|\nabla_{p\mu}^2 H\|_\infty W_1(\mu_1, \mu_2). \end{aligned}$$

Finally, we borrow these assumptions from [52, (2.5), (2.6)]: $H(\cdot, \cdot, \mu) \in C_{loc}^{1,1}(\mathbb{R}^d \times \mathbb{R}^d)$, uniformly in μ , and $H(x, p, \cdot)$ is continuous in μ with respect to W_1 , locally uniformly on $\mathbb{R}^d \times \mathbb{R}^d$.

Assumption 2.2 (Convexity of H). [52, (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 $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $|\nabla_p H(0, p, \mu)| \leq C(p)$.

Assumption 2.3 (Regularity of g). The derivatives $\nabla_\mu g$, $\nabla_{xx}^2 g$, 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:

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

Again, we borrow these assumptions from [52]: $\nabla_x g$ 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 $x \in \mathbb{R}^d$,

$$|\nabla_x g(x, \mu_1) - \nabla_x g(x, \mu_2)| \leq \|\nabla_{x\mu}^2 g\|_\infty W_1(\mu_1, \mu_2).$$

Finally, $g(\cdot, \mu) \in C_{loc}^1(\mathbb{R}^d)$ uniformly in μ , and $g(x, \cdot)$ is continuous with respect to W_1 , locally uniformly in $\mathbb{R}^d$.

Assumption 2.4. The initial condition m_0 is an element of $L^2(\Omega, \mathcal{F}_0, \mathbb{P})$.

Assumption 2.5. The McKean-Vlasov FBSDE (2.3) is stable with respect to the initial condition. To be precise, let $\xi, \tilde{\xi} \in L^2(\Omega, \mathcal{F}_0, \mathbb{P})$ have laws $m_0, \tilde{m}_0$ satisfying Assumption 2.4. Let the solution to Equation (2.3) with initial condition $\xi \sim m_0$ be denoted as (X, Y, Z) ; similarly, let $(\tilde{X}, \tilde{Y}, \tilde{Z})$ denote the solution to Equation (2.3) but with initial condition $\tilde{\xi} \sim \tilde{m}_0$. We assume that for some constant C depending only on T and the Lipschitz constants of $\nabla_p H$ and $\nabla_x H$,

$$\mathbb{E} \left[\sup_{t \in [0, T]} \{|X_t - \tilde{X}_t|^2\} \right] \leq C \mathbb{E}[|\xi - \tilde{\xi}|^2]. \quad (2.2)$$

This assumption is satisfied by FBSDEs covered by certain conditions sufficient for the well-posedness of the FBSDE associated with the MFG, such as displacement monotonicity [52, Theorem 4.5]. We would like to mention that the specific structure that displacement monotonicity is not the key to our results, as any other condition that implies the property of stability with respect to the initial condition could replace displacement monotonicity without changing our results.

Assumption 2.6. For all $\beta \geq 0$, the MFG (1.1) is well-posed with solution (u^β, ρ^β) 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.

Remark 2.7. We present two classes of Hamiltonians that satisfy Assumptions 2.1 and 2.2.

- (1) A quite general class of Hamiltonians is given by the following. Let $F, \gamma_1, \gamma_2 : \mathbb{R}^d \mapsto \mathbb{R}$ and $U_1, U_2 : \mathcal{P}_2(\mathbb{R}^d) \mapsto \mathbb{R}$. Then

$$H(x, p, \mu) = F(x) + \gamma_2(p) + \gamma_1(p)U_1(\mu) + U_2(\mu)$$

satisfies Assumptions 2.1 and 2.2 if $\nabla^2 F$, U_1 , $(\nabla^2 \gamma_1)U$, and $\nabla \gamma_1 \nabla_\mu U_1$ are bounded, and if there exists $C, c > 0$ such that $c \preccurlyeq \nabla^2 \gamma_2 \preccurlyeq C$ and $c \preccurlyeq \nabla^2 \gamma_1 U_1 \preccurlyeq C$.

- (2) The following Hamiltonian is an example that satisfies Assumptions 2.1 and 2.2 due to the addition of a large enough quadratic. Let $\Gamma > 0$, $F : \mathbb{R} \times \mathcal{P}_2(\mathbb{R}^d) \mapsto \mathbb{R}$, $\gamma : \mathbb{R}^d \mapsto \mathbb{R}$, and

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

Then Assumptions 2.1 and 2.2 are satisfied if $\nabla \gamma$, $\nabla^2 \gamma \cdot F$, $F(0, \cdot)$, $\nabla_x F$, $\nabla_{xx}^2 F$, $\nabla_\mu F$ are bounded and if Γ is large enough: $\Gamma > \frac{1}{2} \inf \{ \|\nabla^2 \gamma(p)\| F(x, \mu) \}$.

- (3) Consider the one-dimensional example

$$H(x, p, \mu) = f(x) + xp + (p - U(\mu))^2$$

for $f : \mathbb{R} \mapsto \mathbb{R}$, $U : \mathcal{P}_2(\mathbb{R}) \mapsto \mathbb{R}$. If $\nabla^2 f, \nabla_\mu U$ are bounded, then H satisfies Assumptions 2.1 and 2.2, and if there is some $C(p)$ such that $|p - U(\mu)| \leq C(p)$ for each μ . This example is inspired by [59], despite its Hamiltonian being time-dependent. While our assumptions do not allow for time-varying Hamiltonians, we expect our results to extend to them under suitable regularity conditions.

- (4) For a function $F : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ and a mollifier $\varphi \in C_{bc}^2(\mathbb{R}^d)$ (i.e., its derivatives up to the second order are bounded and compactly supported), $\eta > 0$, and $q > 2$, define

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

If $\gamma, \nabla \gamma, \nabla_x F, \nabla_\mu F$ are bounded, and if Γ is large enough, then Assumptions 2.1 and 2.2 are satisfied. 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. Ideally we would allow γ to be unbounded, such as setting $\gamma(p) = |p|^r/r$ as in [3], but this would violate the assumption that $\|\nabla_{p\mu}^2 H\|_\infty < \infty$ as well as the assumption of uniform convexity if $r \neq 2$.

Our 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.

As discussed in [52, Section 3], Assumption 2.6 can be satisfied by MFGs that do not necessarily possess displacement monotone H or g . Indeed, we allow any H and g that satisfy any conditions that are sufficient to guarantee well-posedness (and do not contradict Assumptions 2.1–2.4). See, for example, [37, 54] for additional conditions sufficient for well-posedness that are beyond the Lasry-Lions and displacement monotonicity conditions.

Remark 2.8 (Comparison with [62]). We compare our assumptions to those of [62] in greater detail. In terms of regularity, their condition (H1') on the 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 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 L -derivatives. However, observe that if the condition (in H_4' and H_4''):

$$\sup_{x \in \mathbb{T}^d} |f(x, \mu') - f(x, \mu)| \lesssim \left(\int_{\mathbb{T}^d} (f(x, \mu') - f(x, \mu)) d(\mu' - \mu)(x) \right)^{\frac{1}{2}},$$

were replaced with

$$\sup_{x \in \mathbb{T}^d} |f(x, \mu') - f(x, \mu)| \lesssim \int_{\mathbb{T}^d} (f(x, \mu') - f(x, \mu)) d(\mu' - \mu)(x),$$

then they would have achieved the same convergence rate as we did.

Moreover, we did not find necessary to assume the monotonicity conditions that [62] imposed on their Hamiltonian and coupling. For instance, we could impose any of the four conditions in [37] that guarantee well-posedness of the MFG system, while [62] did require the Lasry-Lions monotonicity condition for their proof⁴. Most importantly, [62] relied on the separable Hamiltonian structure, and their proof technique seems difficult to adapt to the non-separable case on account of the dual equation technique [49].

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

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

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 :

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

over all controls 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, 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 $\alpha : [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \mapsto \mathbb{R}^d$, defined by

$$\alpha_t(x, \rho) = \nabla_p H(x, -\nabla u^\beta(t, x), \rho),$$

and we can set the sequence of measures μ to be ρ^β , the solution to the Fokker-Planck equation in Equation (1.1). As both α and μ are fixed, and since our focus is on the dependence of u^β on β , we replace α and μ in $X^{\alpha, \mu}$ by the noise intensity β . We use X^β to refer to the stochastic

⁴The Lasry-Lions monotonicity condition is needed in their equations (6.10) and (6.11).

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

- (1) *for all $t \in [0, T]$, the Lipschitz constant of $u^\beta(t, \cdot)$ restricted to any compact set is finite, u^β is a viscosity solution to the HJB equation, and $\nabla_{xx}^2 u^\beta \in L^\infty([0, T] \times \mathbb{R}^d; \mathbb{R}^{d \times d})$.*
- (2) *$\rho^\beta : [0, T] \rightarrow (\mathcal{P}_1(\mathbb{R}^d), W_1)$ 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:*

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

Moreover, ρ_0^β satisfies the initial condition:

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

It is worth noting that from [52, Lemma 3.4], there is an estimate for $\nabla_{xx}^2 u^\beta$ which is uniform in β : $K := \sup_{\beta > 0} \|\nabla_{xx}^2 u^\beta\|_{L^\infty([0, T] \times \mathbb{R}^d; \mathbb{R}^{d \times d})} < \infty$. Moreover, according to [52, Theorem 4.1], under Assumption 2.2, the MFG (1.1) has an FBSDE representation:

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

and it 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 [62] 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^2 , and hence, in W_2 .

Before we state the results, we define the following FBSDE system for $t_0 < T$ and $\zeta \in L^2(\Omega, \mathcal{F}_{t_0}; \mathbb{P}; \mathbb{R}^d)$:

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

Its only differences compared to (2.3) is that the initial time t_0 is not necessarily 0 and that the initial condition does not need to have the law m_0 . To lighten notation, when $\beta = 0$, we denote by $X_t = X_t^0$ to be the solution to the FBSDE (2.3) or (3.1). For a bounded set $U \subseteq \mathbb{R}^d$, we write $A(t_0, T'; U) \in \mathcal{F}_{T'}$ to be the event that X does not exit U between $[t_0, T']$, i.e.

$$A(t_0, T'; U) = \{X_s \in U \text{ for all } s \in [t_0, T']\}.$$

Lemma 3.1. *Let $t_0 \in [0, T)$ and $\zeta \in L^\infty(\Omega, \mathcal{F}_{t_0}, \mathbb{P}; \mathbb{R}^d)$. If X is the solution to Equation (3.1) with $\beta = 0$ and initial condition ζ , then there exists a bounded set U such that $A(t_0, T; U)$ has probability 1. It is worth mentioning for future reference the obvious corollary that for any set $\tilde{U}$ containing U , $A(t_0, T; \tilde{U})$ also has probability 1.*

Proof. By a standard argument using Grönwall's inequality, $K < \infty$, the definition of X , and ∇u being a decoupling field,

$$\sup_{t \in [0, T]} \mathbb{E}[|X_t|^2] < \infty.$$

It follows that $(x, p, t) \mapsto \nabla_p H(x, p, \rho_t)$ has a Lipschitz constant uniform in t . By classical ODE theory, $X : [t_0, T] \mapsto \mathbb{R}^d$ is an element of $C^1([0, T]; \mathbb{R}^d)$ for all $\omega \in \Omega$. Since the law of ζ is compactly supported, there exists some bounded set U such that for all $t \in [t_0, T]$, $X_t \in U$ for all ω . $\square$

Lemma 3.2. *Let $\beta \geq 0$, $t_0 \in [0, T)$, and $\zeta \in L^\infty(\Omega, \mathcal{F}_{t_0}, \mathbb{P}; \mathbb{R}^d)$. Consider the uncoupled state variable dynamics:*

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

Under Assumptions 2.1 – 2.3, there exists a constant $C = C(H, g, T)$, where the dependence on H and g is only through the L^∞ norms of their second-order derivatives, such that for all $T' \in [t_0, T]$,

$$\sup_{t \in [t_0, T']} \mathbb{E}[|X_t^\beta - X_t|^2] \leq C \left\{ \beta^2 + \int_{t_0}^{T'} \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(U)}^2 ds \right\}. \quad (3.3)$$

Proof. Firstly, suppose U is defined as above the statement of the lemma: as a bounded set such that for all $t \in [t_0, T]$, $X_t \in U$. For any $\beta \geq 0$ and $s \in [t_0, T]$,

$$\|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(U)} < \infty$$

because $\nabla u^\beta(s, \cdot)$ is locally Lipschitz continuous in x , on account of $K < \infty$. Now, 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 [52, Theorem 4.1] and Assumption 2.2. Using Assumptions 2.1 and 2.3, as well as $K < \infty$, we have that for all $s \in [t_0, T']$:

$$\begin{aligned} & \mathbb{E}[\|\nabla_p H(X_s^\beta, -\nabla u^\beta(s, X_s^\beta), \rho_s^\beta) - \nabla_p H(X_s, -\nabla u(s, X_s), \rho_s)\|] \\ & \leq \mathbb{E}[\|\nabla_{xp}^2 H\|_\infty |X_s^\beta - X_s| + \|\nabla_{pp}^2 H\|_\infty (|\nabla u^\beta(s, X_s^\beta) - \nabla u^\beta(s, X_s)| \\ & \quad + \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(U)})] + \|\nabla_{p\mu}^2 H\|_\infty W_1(\rho_s^\beta, \rho_s) \\ & \leq C\{\mathbb{E}[|X_s^\beta - X_s|] + \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(U)}\} \end{aligned} \quad (3.4)$$

where the expectations are finite due to the assumption that $\zeta \in L^\infty$. As a result, there exists some constant C depending only on T and the Lipschitz constants of the gradients of H and g , such that for all $T' \leq T$:

$$\begin{aligned} & \mathbb{E}[|X_{T'}^\beta - X_{T'}|^2 \cdot \mathbf{1}_{A(t_0, T'; U)}] = \mathbb{E}[|X_{T'}^\beta - X_{T'}|^2] \\ & \leq C\mathbb{E}\left[\int_{t_0}^{T'} |\nabla_p H(X_s^\beta, -\nabla u^\beta(s, X_s^\beta), \rho_s^\beta) - \nabla_p H(X_s, -\nabla u(s, X_s), \rho_s)|^2 ds + \beta|B_{T'}|^2\right] \\ & \leq C\left\{\beta^2 + \int_{t_0}^{T'} \mathbb{E}[|X_s^\beta - X_s|^2] + \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(U)}^2 ds\right\} \end{aligned}$$

where we used the fact that $A(t_0, T'; U)$ has probability 1 in the first line, the definition of X^β and X in the second line, and $K < \infty$ and Equation (3.4) in the third line. By Grönwall's inequality, there exists another constant $C = C(H, g, T)$ such that

$$\mathbb{E}[|X_{T'}^\beta - X_{T'}|^2] \leq C\left\{\beta^2 + \int_{t_0}^{T'} \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(U)}^2 ds\right\}.$$

Using the fact that the right-hand side is non-decreasing in T' as well as the uniformity of C in $s \in [t_0, T]$ 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 $L^\infty(U)$. In the following theorem, we obtain a convergence rate of $\mathcal{O}(\beta)$ for ∇u^β to ∇u by using its role as the decoupling field.

Theorem 3.3. *Let u^β and u be solutions to the value function component of the MFG system (1.1) with $\beta > 0$ and $\beta = 0$, respectively. Suppose that Assumptions 2.1 – 2.3 hold and that $\zeta \in L^\infty(\Omega, \mathcal{F}_{t_0}, \mathbb{P}; \mathbb{R}^d)$. 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 , and some bounded set U that is large enough for $A(t_0, T; U)$ to have probability 1, such that*

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

Proof. Firstly, due to Lemma 3.1, we know that such a U exists and that the range of ζ is contained in U . Denote by $(X^\beta, Y^\beta, Z^\beta)$ and (X, Y) 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 , since we can take conditional expectation of the X and Y components of Equation (3.1), conditioned on the event that

$X_{t_0}^\beta = x$, as in [52, (4.13)]. Using the uniform convexity of H in p from Assumption 2.2, 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:

$$\begin{aligned} |\nabla u^\beta(t_0, x) - \nabla u(t_0, x)| &\leq \mathbb{E}[|\nabla_x g(X_T^\beta, \rho_T^\beta) - \nabla_x g(X_T, \rho_T)|] \\ &+ \int_{t_0}^T \mathbb{E}[|\nabla_x H(X_t^\beta, -\nabla u^\beta(t, X_t^\beta), \rho_t^\beta) - \nabla_x H(X_t, -\nabla u^\beta(t, X_t), \rho_t)|] dt. \end{aligned} \quad (3.5)$$

Using Lemma 3.2 and Assumption 2.3, we can bound the first term in (3.5) by

$$\begin{aligned} \mathbb{E}[|\nabla_x g(X_T^\beta, \rho_T^\beta) - \nabla_x g(X_T, \rho_T)|^2] &\leq \|\nabla_{xx}^2 g\|_\infty^2 \mathbb{E}[|X_T^\beta - X_T|^2] + \|\nabla_{x\mu}^2 g\|_\infty W_1(\rho_T^\beta, \rho_T)^2 \\ &\leq C \left\{ \beta^2 + \int_{t_0}^T \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(U)}^2 ds \right\}. \end{aligned} \quad (3.6)$$

By Assumption 2.1, there is some constant C depending on the Lipschitz constants of $\nabla_x H$ in x , p , and μ , as well as on K , such that we can bound the second term in (3.5) as:

$$\begin{aligned} &\int_{t_0}^T \mathbb{E}[|\nabla_x H(X_t^\beta, -\nabla u^\beta(t, X_t^\beta), \rho_t^\beta) - \nabla_x H(X_t, -\nabla u^\beta(t, X_t), \rho_t)|^2] dt \\ &\leq \int_{t_0}^T \|\nabla_{xx}^2 H\|_\infty^2 \mathbb{E}[|X_t^\beta - X_t|^2] + \|\nabla_{x\mu}^2 H\|_\infty^2 W_1(\rho_t^\beta, \rho_t)^2 \\ &\quad + \|\nabla_{xp}^2 H\|_\infty^2 \mathbb{E}[|\nabla u^\beta(t, X_t^\beta) - \nabla u^\beta(t, X_t)|^2 + |\nabla u^\beta(t, X_t) - \nabla u(t, X_t)|^2] dt \\ &\leq C \int_{t_0}^T \mathbb{E}[|X_t^\beta - X_t|^2] + \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_{L^\infty(U)}^2 dt. \end{aligned} \quad (3.7)$$

By applying Lemma 3.2 to $\mathbb{E}[|X_t^\beta - X_t|]$, using $K < \infty$, and collecting the time integral of $\mathbb{E}[|X_t^\beta - X_t|]$ into a supremum, we can continue from the last line in Equation (3.7) to get:

$$\begin{aligned} &\int_{t_0}^T \mathbb{E}[|\nabla_x H(X_t^\beta, -\nabla u^\beta(t, X_t^\beta), \rho_t^\beta) - \nabla_x H(X_t, -\nabla u^\beta(t, X_t), \rho_t)|^2] dt \\ &\leq C \left\{ \sup_{t \in [t_0, T]} \mathbb{E}[|X_t^\beta - X_t|^2] + \int_{t_0}^T \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_{L^\infty(U)}^2 dt \right\} \\ &\leq C \left\{ \beta^2 + \int_{t_0}^T \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_{L^\infty(U)}^2 dt \right\}. \end{aligned} \quad (3.8)$$

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 $x \in U$ and combining (3.6) and (3.8), Equation (3.5) is bounded by

$$\|\nabla u^\beta(t_0, \cdot) - \nabla u(t_0, \cdot)\|_{L^\infty(U)}^2 \leq C \left\{ \beta^2 + \int_{t_0}^T \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(U)}^2 ds \right\}.$$

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

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

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

Corollary 3.4. *Suppose that Assumptions 2.1 – 2.6 hold. Then, there exists a constant $C = C(H, g, T, m_0)$, where the dependence on H and g is only through their second derivatives, and whose dependence on m_0 is only through its second moment, such that:*

$$\sup_{t \in [0, T]} W_2(\rho_t^\beta, \rho_t) \leq C\beta. \quad (3.9)$$

Proof. Let $\varepsilon > 0$ and $\xi \sim m_0$. Since ξ is square-integrable, there exists some random variable $\tilde{\xi}$ whose range is contained in a bounded set U such that $\mathbb{E}[|\xi - \tilde{\xi}|^2] < \mathbb{E}[|\xi|^2]\varepsilon$.

To explain further why such a $\tilde{\xi}$ exists, consider the sequence of functions $\{\xi \cdot \mathbf{1}_{\{|\xi| \leq r\}}\}_{r>0}$, which converges pointwise to ξ as $r \rightarrow \infty$. By the dominated convergence theorem, we can find R large enough such that for $r \geq R$,

$$\varepsilon > \mathbb{E}[|\xi|^2 \cdot \mathbf{1}_{\{|\xi| \geq r\}}] \geq r^2 \mathbb{P}(|\xi| \geq r).$$

Take r large enough that U is contained in B_r . Define the random variable $\tilde{\xi}$ to be the product of ξ and the indicator function of B_r , so that

$$\mathbb{E}[|\zeta - \xi|^2] = \mathbb{E}[|\xi|^2 : |\xi| \geq r] \leq \mathbb{P}(|\xi| \geq r) \cdot \mathbb{E}[|\xi|^2] < r^{-2} \mathbb{E}[|\xi|^2] \varepsilon.$$

Thus, if we define $\tilde{U}$ to be the closure of the union of B_r and $\tilde{\xi} = \xi \cdot \mathbf{1}_{\tilde{U}}$, then ζ is the desired random variable. From now on, we can take $\tilde{U}$ to be U instead. For $\beta > 0$, denote $(X^\beta, Y^\beta, Z^\beta)$ and $(\tilde{X}^\beta, \tilde{Y}^\beta, \tilde{Z}^\beta)$ to be the solutions to Equation (2.3) with initial conditions ξ and $\tilde{\xi}$ respectively. (X, Y) and $(\tilde{X}, \tilde{Y})$ are the solutions to Equation (2.3) with initial conditions ξ and $\tilde{\xi}$, but for $\beta = 0$. With the above observations and using the abbreviation $\|\cdot\|$ for $\|\cdot\|_{L^2(\Omega, \mathcal{F}_t, \mathbb{P})}$, for all $t \in [0, T]$:

$$\begin{aligned} \|X_t^\beta - X_t\| &\leq \|X_t^\beta - \tilde{X}_t^\beta\| + \|\tilde{X}_t^\beta - \tilde{X}_t\| + \|\tilde{X}_t - X_t\| \\ &\leq C\mathbb{E}[|\xi - \tilde{\xi}|^2] + \left(\beta^2 + \int_t^T \|\nabla u^\beta(s, \cdot) - \nabla u(s, \cdot)\|_{L^\infty(U)}^2 ds \right)^{1/2} \\ &\leq C\{\mathbb{E}[|\xi|^2]\varepsilon + \beta\} = C(\beta + \varepsilon). \end{aligned} \quad (3.10)$$

In Equation (3.10), to transition from the first line to the second, we applied Assumption 2.5 to handle the first and third terms, and we applied Lemma 3.2 to handle the second term; the use of Lemma 3.2 was justified because their initial conditions $\tilde{X}_0^\beta$ and $\tilde{X}_0$ were assumed to be bounded. The constant in the second line depended on H , g , and T . To handle the term in the square root in the second line, we applied Theorem 3.3. As promised, in the third line, the dependence of C on ξ was only through absorbing $\mathbb{E}[|\xi|^2]$ into the constant of the second line. Since ε is arbitrary, the result follows. $\square$

The convergence result for ∇u^β may be slightly unsatisfying due to the metric in which it was stated. Thus, although we will not use this result for the rest of the paper, we briefly comment that Assumption 2.5 and Corollary 3.4 enable us to derive a rate of convergence of $O(\beta)$ for $\nabla u^\beta \rightarrow \nabla u$ in the $L^2(\rho_t)$ metric.

Corollary 3.5. *Under Assumptions 2.1–2.6, ∇u^β converges to ∇u at a rate of $\mathcal{O}(\beta)$ in $L^2(\rho_t)$ and $L^2(\rho_t^\beta)$, uniformly in t :*

$$\sup_{t \in [0, T]} \left\{ \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_{L^2(\rho_t)} + \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_{L^2(\rho_t^\beta)} \right\} \leq C\beta,$$

for some constant C depending only on H, g, T, m_0 .

Proof. The families of random variables $\{\nabla u^\nu(t, X_t) : \nu \in (0, \beta]\}$, indexed by t , are uniformly absolutely continuous in the sense that for all $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $A \in \mathcal{F}$ with $\mathbb{P}(A) < \delta$, we have

$$\sup_{\nu \in (0, \beta]} \int_A |\nabla u^\nu(t, X_t)|^2 d\mathbb{P} < \varepsilon,$$

since for a constant $C > 0$ independent of t and depending only on K and the supremum over t of $\mathbb{E}[|X_t|^2]$ (which is finite by a standard argument for an SDE whose initial condition is L^2 and whose coefficients are Lipschitz in space, uniformly in time, and satisfy a linear growth condition),

$$\sup_{t \in [0, T]} \int_A |\nabla u^\nu(t, X_t)|^2 d\mathbb{P} \leq 2K^2 \sup_{t \in [0, T]} \mathbb{E}[1 + |X_t|^2 : A] \leq C\delta < \varepsilon$$

for δ small enough. In combination with $\sup_{\nu, t} \mathbb{E}[|\nabla u^\nu(t, X_t)|^2] < \infty$, this is equivalent to the uniform integrability of $\{\nabla u^\nu(t, X_t)\}_{\nu, t}$ by [55, Theorem 6.5.1]. Moreover, due to Arzela-Ascoli, ∇u^ν converges to ∇u uniformly on compacts, so $\nabla u^\nu(t, X_t)$ converges to $\nabla u(t, X_t)$ in probability. By the Vitali convergence theorem (see, for example, [55, Theorem 6.6.1]),

$$\sup_{t \in [0, T]} \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_{L^2(\rho_t)}^2 = \lim_{\nu \rightarrow 0^+} \sup_{t \in [0, T]} \|\nabla u^\beta(t, \cdot) - \nabla u^\nu(t, \cdot)\|_{L^2(\rho_t)}^2 \quad (3.11)$$

For a constant C depending on K and on the constant from Assumption 2.5, we obtain

$$\sup_{t \in [0, T]} \|\nabla u^\beta(t, \cdot) - \nabla u^\nu(t, \cdot)\|_{L^2(\rho_t)}^2 \leq C\mathbb{E} \left[\sup_{t \in [0, T]} |X_t - X_t^\nu|^2 + |Y_t^\beta - Y_t^\nu|^2 \right] \leq C|\beta - \nu|^2, \quad (3.12)$$

where the first inequality is just by the triangle inequality, Fubini's theorem, and Holder's inequality. (In fact, we could have written ρ_t^β instead of ρ_t and X_t^β instead of X_t in everything above without any changes to the proof, which is how convergence in the $L^2(\rho_t^\beta)$ metric can be proven.) Taking the limit as $\nu \rightarrow 0^+$ in Equation (3.12) and then using Equation (3.11), we obtain $\sup_{t \in [0, T]} \|\nabla u^\beta(t, \cdot) - \nabla u(t, \cdot)\|_{L^2(\rho_t)}^2 \leq C\beta^2$. $\square$

4. CONVERGENCE OF u^β

We first present a general stability result concerning Hamilton-Jacobi equations with respect to the supremum norm on compact sets. Similar results exist in the literature but are not directly applicable in our scenario ⁵. To informally describe the result: if two HJ PDEs are well-posed with solutions u and v , Hamiltonians H_1 and H_2 , terminal cost functions g_1 and

⁵Closely related results are [57, 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 [62, Lemma 6.3]. But on account of their formulation not allowing for non-separable Hamiltonians, we also cannot directly apply this lemma.

g_2 , and viscosity parameters 0 and ν , 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) , on that compact set, and a $\mathcal{O}(\nu^{1/2})$ term.

Theorem 4.1. *Let $T > 0$. For $i = 1, 2$, 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 $H_i = H_i(x, p, t)$ is uniformly convex in the second variable $p \in \mathbb{R}^d$, in the sense that there exist constants $c, C > 0$ such that $cI_d \leq \nabla_{pp}^2 H_i \leq CI_d$. Assume that $\nabla_p H$ is Lipschitz in the first argument $x \in \mathbb{R}^d$, uniformly in the second and third arguments, i.e. $\|\nabla_{xp}^2 H\|_\infty$ is finite, and additionally assume that $\nabla_p H_i(0, 0, \cdot)$ is bounded. If the equations*

$$\begin{cases} -\partial_t u_i + H_i(x, -\nabla u_i(t, x), t) = 0, \\ u_i(T, x) = g_i(x), \end{cases} \quad (4.1)$$

are well-posed in the classical sense for $i = 1, 2$ with solution $u_i : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, then for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$ with diameter R and for any $x \in \mathcal{K}$, there exists a constant $\tilde{R} = C(H_1, H_2, g_1, g_2, T, \mathcal{K})$ growing at most linearly in R such that

$$|u_1(t, x) - u_2(t, x)| \leq (T - t)\|H_1 - H_2\|_{L^\infty(B_{\tilde{R}} \times B_{K(1+\tilde{R})+A/c} \times [0, T])} + \|g_1 - g_2\|_{L^\infty(B_{\tilde{R}})}, \quad (4.2)$$

where K satisfies $|\nabla u_i(t, x)| \leq K(1 + |x|)$ for all $(t, x) \in [0, T] \times \mathbb{R}^d$ and $A = A(H_i, \tilde{R}) > 0$ grows at most linearly in $\tilde{R}$.

Proof. Define the costate variable for $i = 1, 2$ as $p_i^*(s) = -\nabla u_i(s, \xi_i^*(s))$, where $\xi_i^* : [0, T] \mapsto \mathbb{R}^d$ is the solution to the following ODE

$$\begin{cases} \partial_s \xi_i^*(s) = \nabla_p H_i(\xi_i^*(s), p_i^*(s), s) & s \in [t, T] \\ \xi_i^*(t) = x. \end{cases} \quad (4.3)$$

In other words, ξ_i^* is the optimally controlled trajectory achieving u_i (c.f. Equation (3.1) with $\beta = 0$). From the assumption of $\nabla_{xp} H_i$ being bounded, Equation (4.3) is well-posed and defined for all $s \in [t, T]$. Let $|x| \leq R$. Due to $\nabla_p H_i(0, 0, \cdot)$ being bounded on $[0, T]$, the sublinear growth of p_i^* , and the uniform convexity of H_i , we can bound the right-hand side of Equation (4.3) as follows:

$$\begin{aligned} |\nabla_p H_i(\xi_i^*(s), p_i^*(s), s)| &\leq |\nabla_p H_i(0, 0, s)| + \|\nabla_{xp}^2 H_i\|_\infty |\xi_i^*(s)| + C_i(1 + |p_i^*(s)|) \\ &\leq \|\nabla_p H_i(0, 0, \cdot)\|_\infty + \|\nabla_{xp}^2 H_i\|_\infty |\xi_i^*(s)| + C_i K(1 + |\xi_i^*(s)|) \\ &\leq C'_i(1 + |\xi_i^*(s)|). \end{aligned} \quad (4.4)$$

for some C'_i depending on C_i , K , $\nabla_{xp}^2 H_i$ and $\|\nabla_p H_i(0, 0, \cdot)\|_\infty$. By a standard application of Grönwall's inequality, for all $s \in [t, T]$,

$$|\xi_i^*(s)| \leq (R + C'_i(T - t))(1 + C'_i e^{C'_i(T-t)}). \quad (4.5)$$

Defining the right hand side as $\tilde{R}$, the costate variable can be controlled as well: for all $s \in [t, T]$: $|p_i^*(s)| \leq K(1 + \tilde{R})$. Denote L_i as the Fenchel conjugate of H_i in the p -variable. Because ξ_2^* is a sub-optimal trajectory for u_1 but is optimal for u_2 , in conjunction with the Hopf-Lax representation of u_i , we get the inequality

$$(u_1 - u_2)(t, x) = (g_1 - g_2)(\xi_2^*(T)) + \int_t^T (L_1 - L_2)(\xi_2^*(s), \partial_s \xi_2^*(s), s) ds. \quad (4.6)$$

We want to control the integrand in Equation 4.6 by something depending on $H_1 - H_2$. Because L_i is the Fenchel conjugate of H_i in the p -variable, and because the supremum is attained uniquely due to the uniform convexity of H_i by some $p_i = p_i(x, v, s)$ solving the equation $v = \nabla_p H_i(x, p, s)$, it follows that for all $(x, v, s) \in \mathbb{R}^d \times \mathbb{R}^d \times [0, T]$,

$$\begin{aligned} (L_1 - L_2)(x, v, s) &\leq \langle p_1, v \rangle - H_1(x, p_1, s) + H_2(x, p_1, s) - \langle p_1, v \rangle \\ &= H_2(x, p_1, s) - H_1(x, p_1, s). \end{aligned} \quad (4.7)$$

Due to symmetry, the same equation can be proven with the subscripts $i = 1, 2$ swapped. Hence, we can bound $|L_1 - L_2|(x, v, s)$ with $x = \xi_2^*(s)$ and $v = \partial_s \xi_2^*(s)$:

$$\begin{aligned} |L_1 - L_2|(\xi_2^*(s), \partial_s \xi_2^*(s), s) &\leq \max \left\{ |H_2(\xi_2^*(s), p_1, s) - H_1(\xi_2^*(s), p_1, s)|, \right. \\ &\quad \left. |H_1(\xi_2^*(s), p_2, s) - H_2(\xi_2^*(s), p_2, s)| \right\} \\ &\leq \max \left\{ |H_2(\xi_2^*(s), p_1, s) - H_1(\xi_2^*(s), p_1, s)|, \|H_1 - H_2\|_{L^\infty(B_{\tilde{R}} \times B_{K(1+\tilde{R})} \times [0, T])} \right\}. \end{aligned} \quad (4.8)$$

The second inequality comes from the fact that $\partial_s \xi_2^*(s) = \nabla_p H_2(\xi_2^*(s), p_2, s)$ (see the definition of the optimizer in the Fenchel conjugate and the fact that it is unique, so that $p_2 = p_2^*$), using Equation (4.3), and then using $|\xi_2^*| \leq \tilde{R}$ and $p_2 \leq K(1 + \tilde{R})$.

Now we need to control $H_2(\xi_2^*, p_1, s) - H_1(\xi_2^*, p_1, s)$; since the only unknown quantity is p_1 , which does not coincide with p_1^* in general, we are done if we can contain p_1 in some ball whose radius depends on $\tilde{R}$ for all $s \in [t, T]$. Using the uniform convexity of H_1 , the fact that $p_1(s)$ is the solution to the equation $\partial_s \xi_2^*(s) = \nabla_p H_1(\xi_2^*(s), p_1(s), s)$, and Equation (4.3), we obtain:

$$\begin{aligned} |p_1(s) - p_2^*(s)| &\leq c^{-1} |\nabla_p H_1(\xi_2^*(s), p_1(s), s) - \nabla_p H_1(\xi_2^*(s), p_2^*(s), s)| \\ &\leq c^{-1} |\partial_s \xi_2^*(s) - \nabla_p H_1(\xi_2^*(s), p_2^*(s), s)| \\ &\leq c^{-1} |\nabla_p H_2(\xi_2^*(s), p_2^*(s), s) - \nabla_p H_1(\xi_2^*(s), p_2^*(s), s)|. \end{aligned} \quad (4.9)$$

Because $|p_2^*(s)| \leq K(1 + \tilde{R})$, $|\xi_2^*(s)| \leq \tilde{R}$, and $\nabla_p H_i$ is continuous, we see that the last line is bounded by some constant $A = A(\nabla_p H_i, \tilde{R}, K)$, and so

$$|p_1(s)| \leq |p_2^*(s)| + |p_1(s) - p_2^*(s)| \leq K(1 + \tilde{R}) + A/c, \quad (4.10)$$

and due to $|\nabla_{xp}^2 H_i|$ and $|\nabla_{pp}^2 H_i|$ being bounded, A grows at most linearly in $\tilde{R}$. This yields a bound on $H_2(\xi_2^*, p_1, s) - H_1(\xi_2^*, p_1, s)$, because

$$|H_2(\xi_2^*(s), p_1(s), s) - H_1(\xi_2^*(s), p_1(s), s)| \leq \|H_1 - H_2\|_{L^\infty(\bar{B}_{\tilde{R}} \times \bar{B}_{K(1+\tilde{R})+A/c} \times [0, T])}. \quad (4.11)$$

Therefore, substituting Equation (4.11) into Equation (4.8) and then into the Hopf-Lax representation in Equation 4.6, we obtain:

$$(u_1 - u_2)(t, x) \leq \|g_1 - g_2\|_{L^\infty(\bar{B}_{\tilde{R}})} + (T - t) \|H_1 - H_2\|_{L^\infty(\bar{B}_{\tilde{R}} \times \bar{B}_{K(1+\tilde{R})+A/c} \times [0, T])}. \quad (4.12)$$

The roles of u_1 and u_2 were symmetrical and can therefore be swapped, which proves the result. $\square$

The following result is well-known (see, for instance, e.g. [62, Lemma 7.1]), but we restate it since we want a precise dependence of its constants on the diameter of the set $\mathcal{K}$.

Proposition 4.2. *Let $T > 0$ and $\nu \geq 0$. Suppose $\bar{H} : \mathbb{R}^d \times \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}$ and $\bar{g} : \mathbb{R}^d \rightarrow \mathbb{R}$ satisfy the same assumptions as H_i and g_i do in Theorem 4.1. If the equations*

$$\begin{cases} -\partial_t u + \bar{H}(x, -\nabla u(t, x), t) = 0, \\ u(T, x) = \bar{g}(x), \end{cases} \quad \begin{cases} -\partial_t v + \bar{H}(x, -\nabla v(t, x), t) = \nu \Delta v, \\ v(T, x) = \bar{g}(x), \end{cases} \quad (4.13)$$

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

$$\|u - v\|_{L^\infty([0, T] \times \mathcal{K})} \leq C_{\mathcal{K}} \nu^{1/2}.$$

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

Proof. By modifying the proof of [62, Lemma 7.1], taking $\Omega = \mathcal{K}$, using the linear-in- x growth of ∇u and ∇v , and the quadratic-in- x growth of $\partial_t u$ and $\partial_t v$, we find that there exists a constant $C = C(\bar{H}, \bar{g}, T)$ such that for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$,

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

Technically, we should take the maximum of $\text{diam}(\mathcal{K})$ and its square, but for the sake of simplicity, we opt to omit this. $\square$

Corollary 4.3. *Let $H_0 : \mathbb{R}^d \times \mathbb{R}^d \times [0, T] \mapsto \mathbb{R}$ and $g_0 : \mathbb{R}^d \mapsto \mathbb{R}$ satisfy the assumptions in Theorem 4.1, and let $\bar{H}$ and $\bar{g}$ satisfy the assumptions in Theorem 4.2. Suppose v_0 satisfies Equation 4.1 with (H_0, g_0) , and suppose v satisfies the second PDE in Equation 4.13 with $(\bar{H}, \bar{g})$. Then for any compact set $\mathcal{K}$ with diameter R , there exists a constant $C_{\mathcal{K}} = C(R, H_0, \bar{H}, g_0, \bar{g}, T)$ growing at most quadratically in R , such that*

$$\|v_0 - v\|_{L^\infty([0, T] \times \mathcal{K})} \leq C_{\mathcal{K}} \{ \|H_0 - \bar{H}\|_{L^\infty(B_{\tilde{R}} \times B_{K(1+\tilde{R})+A/c} \times [0, T])} + \|g_0 - \bar{g}\|_{L^\infty(B_{\tilde{R}})} + \nu^{1/2} \}$$

where $\tilde{R}$, A , and K are as in Theorem 4.1.

Proof. This follows from applying the triangle inequality to the results of Theorem 4.1 and Proposition 4.2. $\square$

Now we combine the results in Section 3 with Corollary 4.3 and make some modifications to the proof of Theorem 4.1 that are particular to the case of measure-dependent H in order to derive the promised $\mathcal{O}(\beta)$ rate of convergence.

Theorem 4.4. *For $\beta \geq 0$, suppose the MFG system (1.1) satisfies Assumptions 2.1–2.6, whose solutions for the HJB equation are denoted u^β and u .*

- (1) *u^β converges uniformly to u on compacts at a rate of $\mathcal{O}(\beta)$: for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$, there exists a constant $C = C_{\mathcal{K}}$ depending on the data (H, g, T, m_0) such that*

$$\|u^\beta - u\|_{L^\infty([0, T] \times \mathcal{K})} \leq C_{\mathcal{K}} \beta, \quad (4.14)$$

when β is small enough.

- (2) *If additionally H is Lipschitz in the measure argument with respect to W_1 , then for some constant $C_{\mathcal{K}}$ depending on the data (H, g, T, m_0) that grows at most quadratically in $\text{diam}(\mathcal{K})$, for all $\beta \geq 0$,*

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

Proof. In Theorem 4.1, let us specialize to the case where $u_1 = u^\beta$ and $u_2 = u$, so that $H_1 = H(\cdot, \cdot, \rho^\beta)$ and $H_2 = H(\cdot, \cdot, \rho)$. Recall that (X, Y) is the solution to the FBSDE in Equation 2.3 with $\beta = 0$. Note that $H_1 - H_2$ is jointly continuous on $\bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T]$ for any $R_1, R_2 > 0$, since for any (x, p, s) , $\nabla_x H(\cdot, \cdot, \rho)$ is bounded on $\bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T]$, due to a Taylor expansion around $(0, 0, \rho_0)$, Assumption 2.1, and $\{\rho_s\}_{s=0}^T$ being bounded with respect to W_2 . For the same reasons, $\nabla_p H(\cdot, \cdot, \rho)$ is bounded on $\bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T]$. Noting that H is continuous in the measure variable with respect to W_1 and that the same argument holds when ρ is replaced with ρ^β closes the argument for joint continuity of H on $\bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T]$. It follows that the difference of H_1 and H_2 on $\bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T]$ is achieved at some $(x_*, p_*, s_*) \in \bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T]$, which we will abbreviate as (x, p, s) for convenience. Thus, for $R_1 = \tilde{R}$ and $R_2 = K(1 + \tilde{R}) + A/c$,

$$\begin{aligned} & \|H(\cdot, \cdot, \rho) - H(\cdot, \cdot, \rho^\beta)\|_{L^\infty(\bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T])} \\ & \leq \mathbb{E}[\langle \nabla_\mu H(x, p, \rho_s, X_s), X_s - X_s^\beta \rangle] + \mathcal{O}(\mathbb{E}[|X_s - X_s^\beta|^2]^{1/2}) \\ & \leq C\{\|\nabla_\mu H(0, 0, \rho_s, \cdot)\|_{L^2(\rho_s)} + \|\nabla_{x\mu}^2 H\|_\infty |x| + \|\nabla_{p\mu}^2 H\|_\infty |p|\}\beta + \mathcal{O}(\beta) \\ & \leq C(1 + R_1 + R_2)\beta + \mathcal{O}(\beta). \end{aligned} \quad (4.15)$$

The second line is via a Taylor expansion of H in the measure variable, which occurs only when ρ_t and ρ_t^β are close enough in W_2 - see Corollary 3.4. The third line comes from the same corollary, from Cauchy-Schwartz, and from the Taylor expansion of $\nabla_\mu H$ in x and p . The fourth line collects the relevant constants into C .

Now, we bound $K(1 + \tilde{R}) + A/c$ more precisely. Recall that A comes from Equation (4.9); since $H_2 = H(\cdot, \cdot, \rho)$, we can identify ξ_2^* in Theorem 4.1 with X solving Equation (3.1) with $\beta = 0$ and $t_0 = t$. Continuing from Equation (4.9), we are able to bound A as follows:

$$\begin{aligned} & |\nabla_p H_2(\xi_2^*(s), p_2^*(s), s) - \nabla_p H_1(\xi_2^*(s), p_2^*(s), s)| \\ & = |\nabla_p H(X_s, -\nabla u(s, X_s), \rho_s^\beta) - \nabla_p H(X_s, -\nabla u(s, X_s), \rho_s)| \\ & \leq \|\nabla_{p\mu}^2 H\|_\infty \mathbb{E}[|X_s^\beta - X_s|^2]^{1/2} + \mathcal{O}(\mathbb{E}[|X_s^\beta - X_s|^2]^{1/2}) \leq C\beta + \mathcal{O}(\beta). \end{aligned} \quad (4.16)$$

Therefore, $K(1 + \tilde{R}) + A/c$ is of at most linear growth in $\tilde{R}$, and

$$\|H(\cdot, \cdot, \rho^\beta) - H(\cdot, \cdot, \rho)\|_{L^\infty(\bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T])} \leq C(1 + \tilde{R})\beta + \mathcal{O}(\beta). \quad (4.17)$$

Now we want to bound $\|g(\cdot, \rho_T^\beta) - g(\cdot, \rho_T)\|_{L^\infty(\bar{B}_{\tilde{R}})}$, which is achieved for some $x_* \in \bar{B}_{\tilde{R}}$ due to the continuity of g in x . From Corollary 3.4, for β small enough,

$$\begin{aligned} & \|g(\cdot, \rho_T^\beta) - g(\cdot, \rho_T)\|_{L^\infty(\bar{B}_{\tilde{R}})} = \mathbb{E}[\langle \nabla_\mu g(x_*, \rho_T, X_T), X_T^\beta - X_T \rangle] + \mathcal{O}(\beta) \\ & \leq \{\|\nabla_\mu g(0, \rho_T, \cdot)\|_{L^1(\rho_T)} + \|\nabla_{x\mu}^2 g\|_\infty |x_*|\} \cdot \mathbb{E}[|X_T^\beta - X_T|^2]^{1/2} + \mathcal{O}(\beta) \\ & \leq C(1 + \tilde{R})\beta + \mathcal{O}(\beta). \end{aligned} \quad (4.18)$$

for some C depending on g and ρ_T (and so on H , T , and m_0 by extension). Substituting Equation 4.16 into Equation 4.15, and then substituting Equation 4.18 into the bound provided by Corollary 4.3, we obtain:

$$\begin{aligned} \|u^\beta - u\|_{L^\infty([0, T] \times \mathcal{K})} & \leq C_K\{\beta + \|g(\cdot, \rho_T^\beta) - g(\cdot, \rho_T)\|_{L^\infty(\bar{B}_{R_1})} \\ & \quad + T\|H(\cdot, \cdot, \rho^\beta) - H(\cdot, \cdot, \rho)\|_{L^\infty(\bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T])}\} \leq C_K\beta + \mathcal{O}(\beta), \end{aligned} \quad (4.19)$$

which concludes the proof of (1).

To prove (2), we first apply Corollary 4.3 with $H_0(x, p, t) = H(x, p, \rho_t)$, $\bar{H}(x, p, t) = H(x, p, \rho_t^\beta)$, $g_0(x) = g(x, \rho_T)$, $\bar{g}(x) = g(x, \rho_T^\beta)$, and $\nu = \beta^2/2$. Abbreviating the difference for H on $L^\infty(\bar{B}_{R_1} \times \bar{B}_{R_2} \times [0, T])$ and the difference for g on $L^\infty(B_{R_1})$, for a constant C_K having the same dependencies as the one from Corollary 4.3, we have:

$$\begin{aligned} \|u^\beta - u\|_{L^\infty([0, T] \times \mathcal{K})} &\leq C_K \left\{ \|H(\cdot, \cdot, \rho^\beta) - H(\cdot, \cdot, \rho)\|_\infty + \|g(\cdot, \rho_T^\beta) - g(\cdot, \rho_T)\|_\infty + \beta/\sqrt{2} \right\} \\ &\leq C_K \left\{ (\|\nabla_\mu H\|_\infty + \|\nabla_\mu g\|_\infty) \sup_{t \in [0, T]} W_2(\rho_t^\beta, \rho_t) + \beta \right\} \leq C_K \beta, \end{aligned}$$

where the third inequality used Corollary 3.4. $\square$

Remark 4.5. *If it is only assumed that H and g are continuous in W_1 , by the stability of viscosity solutions to the HJB equation, we can conclude that $u^\beta \rightarrow u$ uniformly on compacts, albeit without a rate.*

Remark 4.6. *If $\|\nabla_{\mu\mu}^2 H\|_\infty$ and $\|\nabla_{\mu\mu}^2 g\|_\infty$ are assumed to be finite, as [5] does, then we can derive a stronger result: in the Taylor expansions of Equations (4.15) and Equations (4.18), we can replace the term of $\mathcal{O}(\beta)$ by $C'\beta^2$, where C' depends on $\|\nabla_{\mu\mu}^2 H\|_\infty$ and $\|\nabla_{\mu\mu}^2 g\|_\infty$. Then,*

$$\|u^\beta - u\|_{L^\infty([0, T] \times \mathcal{K})} \leq C_K \beta + C' \beta^2.$$

5. APPLICATIONS

This section provides three applications of our result to 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 [29, 32, 44] that the limit is the MFG equilibrium, finding the convergence rate is a separate and difficult problem. The twin papers [26, 27] seem to 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-degenerate⁶. Here we apply Corollary 3.4 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 [27, (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 whose i -th component is:

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

⁶When 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.

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 are:

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

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

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

Corollary 5.1. *Let ρ satisfy the Fokker-Planck equation in the MFG (1.1) with $\beta = 0$. Let $\Sigma = \beta I$, and denote by $m_{X_t}^{N, \beta}$ the empirical measure in (5.2) corresponding to $\Sigma = \beta I$. Under the assumptions in Corollary 3.4, and under [27, Assumption A] and [27, Assumption B or B'], there exist $C_1 = C_1(H, g, T)$ and $C_2 = C_2(\beta, H, g, T)$ such that for all $t \in [0, T]$,*

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

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

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

By [27, Theorem 3.1], there is a constant $C_2 = C_2(\beta, H, g, T)$ such that

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

Combining Equations (5.4), (5.5) with Corollary 3.4 yields the desired bound. $\square$

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 [27] 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 at most $N(\varepsilon C_2^{-1}(\varepsilon))^{-(d+8)} \gg \varepsilon^{-(d+8)}$ 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 [25, Proposition 2.14], where a central planner seeks to control N particles by selecting a control process $\alpha = (\alpha^1, \dots, \alpha^N)$ from $\mathcal{A}^N$, which is the set of $(\mathbb{R}^d)^N$ -valued, progressively measurable processes, whose definition is immediately below [25, Equation 2.8]. Throughout this subsection, $\beta > 0$ is fixed, and N may vary. 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 satisfies the SDE

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

where $\bar{B}_t = N^{-1/2} \sum_{i=1}^N B_t^i$ is a d -dimensional Brownian motion. The objective of the central planner is to solve the optimization problem in [25, Equation 2.21]:

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

where $F, G : \mathbb{R}^d \rightarrow \mathbb{R}$ are assumed to be Lipschitz, $L \in C^2(\mathbb{R}^d)$ satisfies the second-derivative bounds $\frac{1}{C}I \leq \nabla^2 L \leq CI$ for some $C \geq 1$, and α^i is a function of both ω and x (though the optimizer is a deterministic function). An easy argument from [25] shows that the optimality in (5.8) is achieved by a deterministic control, and $V^N(t, x) = v^N(t, \bar{m}_x^N)$, where v^N solves the HJ equation:

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

and 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 3.4 hold. Let $\{X_0^i\}_{i=1}^N$ be independent and identically distributed according to m_0 with bounded support⁷, 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:*

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

Assume further that $\nabla_p H(-\nabla v(t, x))$ is Lipschitz in x bounded in t ⁸, i.e., there is $L > 0$ such that

$$|\nabla_p H(-\nabla v(t, x)) - \nabla_p H(-\nabla v(t, y))| \leq L|x - y| \quad \text{for all } t, x, y. \quad (5.12)$$

⁷For $d = 1$, 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 (5.17) is bounded by $C/\sqrt{N}$; see the discussion after [60, Theorem 3.4] or [56].

⁸The bound (5.13) is a conditional result on the Lipschitz assumption (5.12). The assumption implicitly requires v to be a classical solution, which is not true in general. A sufficient condition for this assumption to hold is that H is uniformly convex, F, G are convex, and H, F, G are smooth with bounded Hessians, i.e., $c_H I \leq \nabla_{pp}^2 H \leq C_H I$, $0 \leq \nabla_{xx}^2 F \leq C_F$ and $0 \leq \nabla_{xx}^2 G \leq C_G$. The uniform convexity of H , the convexity of F, G and their smoothness imply that v is a classical solution because the characteristics do not cross. The

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

$$W_2(\mu_t^N, \mu_t) \leq C \left(\frac{1}{\sqrt{N}} + \frac{\sqrt{\log N}}{N} + \sqrt{\frac{\text{Tr} \Sigma}{N}} \right). \quad (5.13)$$

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

$$\begin{cases} -\partial_t v^N(t, x) + H(-\nabla v^N(t, x)) - F(x) = \frac{\beta^2}{N} \Delta v^N(t, x), \\ \partial_t \mu_t^N + \text{div}_x \{ \mu_t^N \nabla_p H(-\nabla v^N(t, x)) \} = \frac{\beta^2}{N} \Delta \mu_t^N, \\ v^N(T, x) = G(x), \quad \mu_0^N = \text{Law}(\bar{X}_0^N), \end{cases} \quad (5.14)$$

Note that the HJ equation is not coupled with μ^N . Let $(\tilde{v}^N, \tilde{\mu}^N)$ be a solution to the MFG:

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

(So $\tilde{v}^N = v$ in the equation (5.10).) As a consequence of Corollary 3.4, we obtain:

$$W_2(\mu_t^N, \tilde{\mu}_t^N) \leq \frac{C}{\sqrt{N}} \quad \text{for some } C > 0 \text{ (independent of } N). \quad (5.15)$$

By the Lipschitz assumption (5.12) and the Cauchy-Lipschitz theory of the continuity equation (see e.g., [4, Section 2]):

$$W_2(\tilde{\mu}_t^N, \mu_t) \leq C W_2 \left(\text{Law}(\bar{X}_0^N), \delta_{\int x m_0(x) dx} \right). \quad (5.16)$$

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

$$\begin{aligned} W_2 \left(\text{Law}(\bar{X}_0^N), \delta_0 \right) &\leq W_2 \left(\text{Law}(\bar{X}_0^N), \mathcal{N} \left(0, \frac{\Sigma}{N} \right) \right) + W_2 \left(\mathcal{N} \left(0, \frac{\Sigma}{N} \right), \delta_0 \right) \\ &\leq \frac{C \sqrt{d \log N}}{N} + \sqrt{\frac{\text{Tr} \Sigma}{N}}, \end{aligned} \quad (5.17)$$

where the first term in the last inequality follows from [30, Theorem 1]⁹, and the W_2 distance of two Gaussian vectors (see e.g., [36, Proposition 7]). Combining the equations (5.15), (5.16) and (5.17) 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 [38, 42, 51, 65, 64]. 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 [48]. However, PI is not directly applicable to the first-order problems due to ill-posedness [61]. So a reasonable idea is to approximate first-order MFGs by second-order MFGs¹⁰, and a

boundedness of $\nabla_{xx}^2 F$ and $\nabla_{xx}^2 G$ further guarantees that $0 \leq \nabla_{xx}^2 v \leq \max \left(C_G, \sqrt{C_F / c_H} \right) I$. Combined with the boundedness of $\nabla_{pp}^2 H$ yields the assumption.

⁹A slightly looser bound $\mathcal{O}(\sqrt{d \log N} / N)$ (up to a $\log N$ factor) was proved in [68, Theorem 1.1].

¹⁰This idea was also proposed in [61] to solve deterministic control problems by PI.

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

$$\partial_t \rho_t^{n,\beta} - \operatorname{div}\{\rho_t^{n,\beta} q^n\} = \frac{\beta^2}{2} \Delta \rho_t^{n,\beta}, \quad \rho_0^{n,\beta} = m_0. \quad (5.18)$$

(ii) Solve

$$-\partial u^{n,\beta} + q^n \nabla u^{n,\beta} - \mathcal{L}(x, -\nabla u^{n,\beta}, q^n, \rho_t^{n,\beta}) = \frac{\beta^2}{2} \Delta u^{n,\beta}, \quad u^{n,\beta}(T, x) = g(x), \quad (5.19)$$

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

(iii) Update the policy

$$q^{n+1}(t, x) := \arg \max_{|q| \leq R} \left(q \cdot \nabla u^{n,\beta}(t, x) - \mathcal{L}(x, q, \rho_t^{n,\beta}) \right), \quad (5.20)$$

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

In all of the aforementioned works [6, 8, 9, 48], the convergence (rate) of PI (5.18)–(5.20) 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 [48], 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*

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

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 $\mathcal{K}, \beta$ 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 $\mathcal{K} \subset \mathbb{R}^d$ be a compact set. Under the assumptions in Theorem 4.4 and Claim 5.3, there exist $T = T(\beta) > 0$, $C_1 = C_1(\mathcal{K})$ and $C_2 = C_2(\beta)$ such that*

$$\frac{1}{T} \|u^{n,\beta} - u\|_{L^r([0,T] \times \mathcal{K})} \leq C_1 \beta + C_2(\beta) e^{-n} \quad \text{for } r > d + 2. \quad (5.22)$$

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:

$$\beta \asymp \varepsilon \quad \text{and} \quad C_2(\beta)e^{-n} \asymp \varepsilon, \quad (5.23)$$

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 vanishing viscosity approximations to MFGs matches the classically 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 of convergence.

Consider the following example from [7, 15]:

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

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 may also be extended to a class of displacement monotone MFGs by using the arguments in [21].

6.2. Numerical examples. We proved in Theorem 4.4 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 $\beta \in \{0.1, 0.2, \dots, 0.9, 1\}$). To solve the MFG, we used Picard iteration and added damping for stabilization purposes, with every iteration first solving for the Fokker-Planck equation and then the HJB equation. Since the Fokker-Planck equation is linear, we can use a generic linear solver for the system of equations derived from the equation's finite difference representation, but since the HJB equation is nonlinear, we used Newton's method to solve its system of equations¹¹.

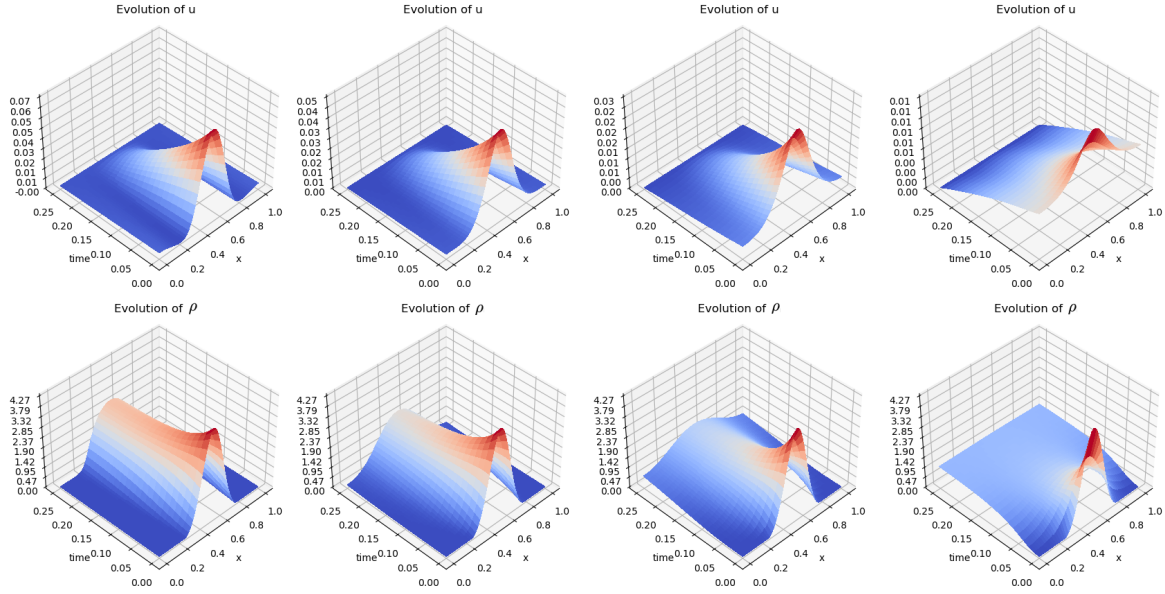


FIGURE 1. Plot of (u^β, ρ^β) for $\beta \in \{0.1, 0.3, 0.5, 1.0\}$ (left to right).

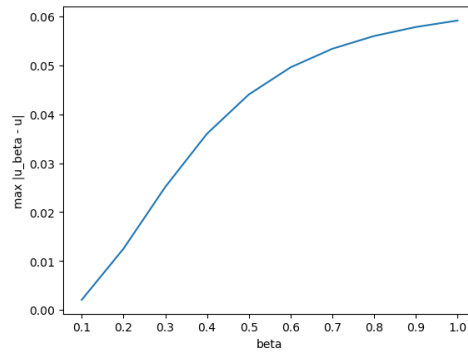


FIGURE 2. Plot of $\|u^\beta - u\|_\infty$ against β .

¹¹Our numerical results are based on the codes available at https://colab.research.google.com/drive/1shJWSD2MA5Fo7_rB625dAvNTdZS1a7bG?usp=sharing.

In [62], 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

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

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 [62].

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_2 metric, respectively. Our approach exploits both probabilistic and analytical arguments, where the FBSDE representation of the MFG is used to derive the convergence rate of ρ^β , and the rate of u^β follows from a stability property of the HJB 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 [60] 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 [60] 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 first order 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 [66] in the context of reinforcement learning, where the HJB equation is replaced with the exploratory equation under entropy regularization [63].

Acknowledgment: We thank the referee for their careful reading and valuable suggestions, which have greatly improved this work. We thank Alpár Mészáros for pointing out an error in a lemma in a previous version of this work concerning the stability of McKean-Vlasov FBSDEs, and we also thank Daniel Lacker for helpful discussions on how to resolve the issue. W. Yu is supported by the Columbia Innovation Hub grant and NSF grant DMS-2309245. Q. Du is supported in part by NSF grants DMS-2309245 and DMS-1937254. W. Tang is supported by

NSF CAREER Award DMS-2538791, the Tang Family Assistant Professorship and a Columbia-CityU/HK collaborative project that is supported by InnoHK Initiative, The Government of the HKSAR and the AIFT Lab.

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 and G. Crippa. Continuity equations and ODE flows with non-smooth velocity. *Proc. Roy. Soc. Edinburgh Sect. A*, 144(6):1191–1244, 2014.
- [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. Notes on mean field games. *Lectures by P.L. Lions at the College de France*, 2012.
- [11] 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.
- [12] P. Cardaliaguet and J. P. Graber. Mean field games systems of first order. *ESAIM Control Optim. Calc. Var.*, 21(3):690–722, 2015.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] 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.

- [18] R. Carmona and D. Lacker. A probabilistic weak formulation of mean field games and applications. *Ann. Appl. Probab.*, 25(3):1189–1231, 2015.
- [19] L.-P. Chaintron and S. Daudin. Optimal rate of convergence in the vanishing viscosity for uniformly convex Hamilton-Jacobi equations. *arxiv:2506.13255*, 2025.
- [20] M. Cirant and A. Goffi. Convergence rates for the vanishing viscosity approximation of Hamilton-Jacobi equations: the convex case. *arxiv:2502.15495*, 2025.
- [21] M. Cirant and A. R. Mészáros. Long time behavior and stabilization for displacement monotone mean field games. *arXiv:2412.14903*, 2024.
- [22] 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.
- [23] 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.
- [24] M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277(1):1–42, 1983.
- [25] 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.
- [26] 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.
- [27] 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.
- [28] B. Djehiche, A. Tchekam, and H. Tembine. Mean-field-type games in engineering. *AIMS Electronics and Electrical Engineering*, 1(1):18–73, 2017.
- [29] M. F. Djete. Extended mean field control problem: a propagation of chaos result. *Electron. J. Probab.*, 27:1–53, 2022.
- [30] R. Eldan, D. Mikulincer, and A. Zhai. The CLT in high dimensions: quantitative bounds via martingale embedding. *Ann. Probab.*, 48(5):2494–2524, 2020.
- [31] 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.
- [32] M. Fischer. On the connection between symmetric n-player games and mean field games. *Ann. Appl. Probab.*, 27(2):757–810, 2022.
- [33] 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.
- [34] W. Gangbo and A. Tudorascu. On differentiability in the Wasserstein space and well-posedness for Hamilton-Jacobi equations. *J. Math. Pures Appl.*, 9:119–174, 2019.
- [35] M. Ghattassi and N. Masmoudi. Non-separable mean field games for pedestrian flow: generalized Hughes model. *Math. Models Methods Appl. Sci.*, 35(7):1571–1607, 2025.
- [36] C. R. Givens and R. M. Shortt. A class of Wasserstein metrics for probability distributions. *Michigan Math. J.*, 31:231 – 240, 1984.
- [37] P. J. Graber and A. R. Mészáros. On monotonicity conditions for mean field games. *J. Funct. Anal.*, 285(9):Paper No. 110095, 45, 2023.
- [38] X. Guo, H. V. Tran, and Y. P. Zhang. Policy iteration for nonconvex viscous Hamilton-Jacobi equations. *arXiv:2503.02159*, 2025.
- [39] 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.
- [40] 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.
 - [41] 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.
 - [42] 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.
 - [43] D. Lacker. Mean field games via controlled martingale problems: existence of Markovian equilibria. *Stochastic Process. Appl.*, 125(7):2856–2894, 2015.
 - [44] D. Lacker. On the convergence of closed-loop Nash equilibria to the mean field game limit. *Ann. Appl. Probab.*, 30(4):1693–1761, 2020.
 - [45] 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.
 - [46] 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.
 - [47] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
 - [48] 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.
 - [49] C.-T. Lin and E. Tadmor. L^1 -stability and error estimates for approximate Hamilton-Jacobi solutions. *Numer. Math.*, 87(4):701–735, 2001.
 - [50] 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.
 - [51] J. Ma, G. Wang, and J. Zhang. Convergence analysis for entropy-regularized control problems: a probabilistic approach. *arXiv:2406.10959*, 2024.
 - [52] A. R. Mészáros and C. Mou. Mean field games systems under displacement monotonicity. *SIAM J. Math. Anal.*, 56(1):529–553, 2024.
 - [53] 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.
 - [54] 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.
 - [55] S. Resnick. *A Probability Path*. Modern Birkhäuser Classics. Birkhäuser Boston, MA, first edition, 2014.
 - [56] E. Rio. Upper bounds for minimal distances in the central limit theorem. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):802–817, 2009.
 - [57] P. E. Souganidis. Existence of viscosity solutions of Hamilton-Jacobi equations. *J. Differential Equations*, 56(5):345–390, 1985.
 - [58] E. Tadmor. Numerical viscosity and the entropy condition for conservative difference schemes. *Math. Comp.*, 43(168):369–381, 1984.
 - [59] 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.

- [60] W. Tang and F. Tang. The Poisson binomial distribution—old & new. *Statist. Sci.*, 38(1):108–119, 2023.
- [61] 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.
- [62] W. Tang and Y. P. Zhang. The convergence rate of vanishing viscosity approximations for mean field games. *SIAM J. Math. Anal.*, 57(3):3217–3254, 2025.
- [63] W. Tang, Y. P. Zhang, and X. Y. Zhou. Exploratory HJB equations and their convergence. *SIAM J. Control Optim.*, 60(6):3191–3216, 2022.
- [64] W. Tang and X. Y. Zhou. Regret of exploratory policy improvement and q -learning. *arXiv:2411.01302*, 2024.
- [65] 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.
- [66] 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.
- [67] C. Wu and J. Zhang. An elementary proof for the structure of Wasserstein derivatives. 2017. *arXiv:1705.08046*.
- [68] 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