The Annals of Probability
2020, Vol. 48, No. 1, 211–263
https://doi.org/10.1214/19-AOP1359

FROM THE MASTER EQUATION TO MEAN FIELD GAME LIMIT THEORY: LARGE DEVIATIONS AND CONCENTRATION OF MEASURE

BY FRANÇOIS DELARUE[1], DANIEL LACKER[2] AND KAVITA RAMANAN[3]

[1]*Universite de Nice Sophia-Antipolis, francois.delarue@unice.fr*

[2]*Industrial Engineering and Operations Research, Columbia University, daniel.lacker@columbia.edu*

[3]*Division of Applied Mathematics, Brown University, kavita_ramanan@brown.edu*

We study a sequence of symmetric n-player stochastic differential games driven by both idiosyncratic and common sources of noise, in which players interact with each other through their empirical distribution. The unique Nash equilibrium empirical measure of the n-player game is known to converge, as n goes to infinity, to the unique equilibrium of an associated mean field game. Under suitable regularity conditions, in the absence of common noise, we complement this law of large numbers result with nonasymptotic concentration bounds for the Wasserstein distance between the n-player Nash equilibrium empirical measure and the mean field equilibrium. We also show that the sequence of Nash equilibrium empirical measures satisfies a weak large deviation principle, which can be strengthened to a full large deviation principle only in the absence of common noise. For both sets of results, we first use the master equation, an infinite-dimensional partial differential equation that characterizes the value function of the mean field game, to construct an associated McKean–Vlasov interacting n-particle system that is exponentially close to the Nash equilibrium dynamics of the n-player game for large n, by refining estimates obtained in our companion paper. Then we establish a weak large deviation principle for McKean–Vlasov systems in the presence of common noise. In the absence of common noise, we upgrade this to a full large deviation principle and obtain new concentration estimates for McKean–Vlasov systems. Finally, in two specific examples that do not satisfy the assumptions of our main theorems, we show how to adapt our methodology to establish large deviations and concentration results.

CONTENTS

Received April 2018; revised February 2019.

MSC2010 subject classifications. Primary 60F10, 60E15, 60H10, 91A13, 91A15; secondary 91G80, 60K35.

Key words and phrases. Mean field games, master equation, McKean–Vlasov limit, interacting particle systems, common noise, large deviation principle, concentration of measure, transport inequalities, linear-quadratic systems, systemic risk.

1. Introduction. *Description of the model*. In this article, we study Nash equilibria for a class of symmetric n-player stochastic differential games, for large n. To describe our main results, we first provide an informal description of the n-player game (see Section 2.3 for a complete description). Let the empirical measure of a vector $\boldsymbol{x} = (x_1, \dots, x_n)$ in $(\mathbb{R}^d)^n$ be denoted by

$$m^n_{\boldsymbol{x}} = \frac{1}{n}\sum_{k=1}^{n} \delta_{x_k},$$

where δ_x is the Dirac delta mass at $x \in \mathbb{R}^d$, which lies in $\mathcal{P}(\mathbb{R}^d)$, the space of probability measures on $\mathbb{R}^d$. Given independent $\mathbb{R}^{d_0}$-valued and $\mathbb{R}^d$-valued Wiener processes W and $B^1, \dots, B^n$ and $\mathbb{R}^d$-valued initial conditions $(X^1_0, \dots, X^n_0)$, a time horizon $T < \infty$, an action space A, a drift functional $b : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times A \to \mathbb{R}^d$ and two constant matrices $\sigma \in \mathbb{R}^{d\times d}$ and $\sigma_0 \in \mathbb{R}^{d\times d_0}$, with σ nondegenerate, the state of the n-player game at time t is given by $\boldsymbol{X}_t = (X^1_t, \dots, X^n_t)$, where the state X^i of the ith agent follows the dynamics

$$dX^i_t = b\big(X^i_t, m^n_{\boldsymbol{X}_t}, \alpha^i(t, \boldsymbol{X}_t)\big)\,dt + \sigma\,dB^i_t + \sigma_0\,dW_t. \tag{1.1}$$

Here, $\alpha^i : [0,T] \times (\mathbb{R}^d)^n \to A$ is a Markovian control that is chosen to minimize the ith objective function

$$J^n_i(\alpha^1, \dots, \alpha^n) = \mathbb{E}\bigg[\int_0^T f\big(X^i_t, m^n_{\boldsymbol{X}_t}, \alpha^i(t, \boldsymbol{X}_t)\big)\,dt + g\big(X^i_T, m^n_{\boldsymbol{X}_T}\big)\bigg], \tag{1.2}$$

for suitable cost functionals f and g. An n-tuple $(\alpha^1, \dots, \alpha^n)$ is said to be a Nash equilibrium of this game (in closed-loop strategies) if for every $i = 1, \dots, n$, and Markov control $\tilde{\alpha}$,

$$J^n_i(\alpha^1, \dots, \alpha^{i-1}, \alpha^i, \alpha^{i+1}, \dots, \alpha^n) \leq J^n_i(\alpha^1, \dots, \alpha^{i-1}, \tilde{\alpha}, \alpha^{i+1}, \dots, \alpha^n).$$

Under suitable conditions, it was shown in [10] this game has a unique Nash equilibria that can be characterized in terms of the classical solution of a certain partial differential equation (PDE) system called the Nash system, introduced in Section 2.3. If $\boldsymbol{X} = \{\boldsymbol{X}_t = (X_t^1, \dots, X_t^n), t \in [0,T]\}$, is the associated state process, then $(m^n_{\boldsymbol{X}_t})_{t\in[0,T]}$ is referred to as the associated Nash equilibrium empirical measure. Under additional regularity conditions, it was also shown in [10] that $(m^n_{\boldsymbol{X}_t})_{t\in[0,T]}$ converges, as n goes to infinity, to the unique equilibrium $(\mu_t)_{t\in[0,T]}$ of a certain associated mean field game (MFG), described in Section 2.4. The equilibrium $\mu = (\mu_t, t \in [0,T])$ is itself a stochastic flow of probability measures, and can be described in terms of the value function of the MFG, which is the unique solution to an infinite-dimensional PDE referred to as the so-called master equation (see Section 2.4 for full details). As we clarify below, the convergence of $(m^n_{\boldsymbol{X}_t})_{t\in[0,T]}$ to $(\mu_t)_{t\in[0,T]}$ must be regarded as a Law of Large Numbers (LLN) for games of type (1.1)–(1.2).

Main results and strategy of proof. This is the second article in a two-part series, with the first part [20] complementing the aforementioned LLN with a functional central limit theorem; see [20] for a more thorough introduction and bibliography. In this work, we refine the law of large numbers (LLN) convergence result of [10] mentioned above by establishing non-asymptotic concentration bounds and large deviation results.

We first construct a related interacting diffusion system $\overline{X} = (\overline{X}^1, \dots, \overline{X}^n)$ of McKean–Vlasov type:

$$d\overline{X}_t^i = \widetilde{b}(t, \overline{X}_t^i, m^n_{\overline{\boldsymbol{X}}_t})\,dt + \sigma\,dB_t^i + \sigma_0\,dW_t, \tag{1.3}$$

for a suitable drift $\widetilde{b}$ defined in terms of the drift b and the solution to the master equation. We then show that this McKean–Vlasov system is exponentially close to the Nash system. More precisely, under suitable assumptions (see Assumptions A, B and B′ below) we prove (see Theorem 4.3) that there exist constants $C < \infty$ and $\delta > 0$ such that for every $a > 0$ and $n \ge C/a$ we have

$$\mathbb{P}\big(\mathcal{W}_{2,\mathcal{C}^d}(m^n_{\boldsymbol{X}}, m^n_{\overline{\boldsymbol{X}}}) > a\big) \le 2ne^{-\delta a^2 n^2}, \tag{1.4}$$

where $\mathcal{W}_{p,\mathcal{C}^d}$ denotes the p-Wasserstein distance on the space of probability measures on the path space $\mathcal{C}^d := C([0,T];\mathbb{R}^d)$ with finite pth moment. This is a refinement of cruder estimates obtained in [10] and [20], relation (4.27), which are used to characterize LLN and (central limit) fluctuations of the Nash equilibrium empirical measure from the MFG equilibrium, respectively. The exponential equivalence estimate (1.4) reduces the problem of establishing concentration estimates or LDPs for the (sequence of) Nash systems to that of establishing analogous results for the (sequence of) McKean–Vlasov systems.

The following is the summary of our main results in the absence of common noise (i.e., when $\sigma_0 = 0$):

1. We obtain concentration results for McKean–Vlasov systems of the form (1.3) (see Section 5.2 and, in particular, Theorem 5.6), which are interesting in their own right. Prior works on concentration for McKean–Vlasov systems [7, 8], motivated mostly by questions of long-time convergence to equilibrium, restricted attention naturally to gradient drift coefficients. We thus adopt a new approach, for Lipschitz but nongradient drifts, that yields not only deviation probability bounds like those in [7, 8] but also full *concentration of measure*, in the sense that Lipschitz functions of $(\overline{X}^1, \dots, \overline{X}^n)$ concentrate around their means. The proofs rely on transport inequalities, crucially using a result of [22].

2. We use the exponential equivalence along with the result in (1) above to obtain concentration results for quantities like

$$\mathbb{P}\Big(\sup_{t\in[0,T]} \mathcal{W}_{p,\mathbb{R}^d}(m^n_{\boldsymbol{X}_t}, \mu_t) \ge \epsilon\Big),$$

for $\epsilon > 0$ and for exponents $p \in \{1, 2\}$ (see Corollaries 3.3 and 3.5); here, $\mathcal{W}_{p,\mathbb{R}^d}$ is the p-Wasserstein distance on the space of probability measures on $\mathbb{R}^d$ with finite pth moment. In fact, these bounds are consequences of more powerful results we obtain on concentration of Lipschitz functions of $\boldsymbol{X}$ (see Theorems 3.2 and 3.4). Notably, we show that as soon as the i.i.d. initial states $(X_0^i)_{i=1}^n$ obey a dimension-free concentration of measure property, then so do the Nash systems. In addition, under modest additional assumptions, we obtain comparable results on the rate of convergence of the equilibrium controls in the n-player game to the MFG equilibrium control (see Theorem 3.9).

3. We show (in Theorem 3.10) that the sequence $((m^n_{\boldsymbol{X}_t})_{t\in[0,T]})_{n\in\mathbb{N}}$ obeys a large deviation principle (LDP) in the space of continuous paths taking values in the space $\mathcal{P}(\mathbb{R}^d)$, equipped with the $\mathcal{W}_{1,\mathbb{R}^d}$ metric. We explicitly identify the rate function in a form similar to that of Dawson–Gärtner [19]. Our LDP can be obtained essentially by bootstrapping known large deviations results for McKean–Vlasov systems, such as those in [2, 9, 19]. Indeed, the result then nearly follows from the exponential equivalence (1.4) and [19], except that our drift coefficient $\widetilde{b}$ in (1.3) is (necessarily) time-dependent. In any case, we provide a complete proof because, in our setting with constant volatility coefficients, a relatively simple argument is available based on contraction mapping and, furthermore, because a similar argument is required for the LDP in the presence of common noise described below, for which there are no previous results.

In the presence of common noise (i.e., $\sigma_0 \neq 0$), the LDP we obtain for $((m^n_{\boldsymbol{X}_t})_{t\in[0,T]})_{n\in\mathbb{N}}$ is in fact a *weak LDP*, with a rate function that fails to be a good rate function; that is, the rate function does not have compact level sets (see Theorem 3.11).

Our results on concentration and large deviations appear to be the first of their kind for diffusion-based MFGs. Moreover, in the McKean–Vlasov setting, our concentration bounds and our weak LDP in the case with common noise appear to be new as well. The recent papers [1, 16, 17] develop similar techniques for MFGs with finite state space and without common noise, using the (finite-dimensional) master equation to connect the n-player equilibrium to a more classical interacting particle system, and then transferring limit theorems (specifically, a LLN, CLT and LDP) from the latter to the former. Notably, the second and third author recently developed in [31] a quite general LDP for static (i.e., one-shot) mean field games, but the methods used therein do not seem adaptable to dynamic settings. To the best of our knowledge, there are no prior results on LDPs in the presence of common noise or concentration bounds for MFGs, whether in finite or infinite state space, or for static or dynamic games.

Required assumptions and examples. As further elaborated in [20], the above results are all proven under admittedly very strong hypotheses, namely Assumptions A, and Assumption B or B′, which are spelled out in Section 2.5. That said, the same strategy of connecting the n-player equilibrium and a corresponding McKean–Vlasov system in order to transfer limit theorems seems to be more widely applicable than our rather restrictive assumptions might suggest. We illustrate this in Section 7 via two models, the linear-quadratic model of [14] and the Merton-type model of [32], which admit explicit solutions for both the n-player and mean field games. Taking advantage of the explicit solutions, we are able to derive similar concentration bounds and LDPs for these systems in spite of unbounded coefficients and other technical impediments.

Organization of the paper. In Section 2, we introduce common notation, describe the Nash system, the master equation, the MFG and the main sets of assumptions. In Section 3, we give precise statements of the main results, with the concentration bounds in Section 3.1, and the large deviations results in Section 3.2. The proofs of the concentration bounds and LDP are given in Sections 5 and 6, respectively. These rely on exponential estimates between the Nash system and the master equation, which are first developed in Section 4. Section 7 provides two examples that are not covered by the main theorem, but for which the general methodology can still be shown to apply. Finally, we discuss some open problems in Section 8.

2. Nash systems and master equations.

2.1. *Notation and model inputs.* For a topological space E, let $\mathcal{P}(E)$ denote the set of Borel probability measures on E. Throughout the paper, we make use of the standard notation $\langle \mu, \varphi\rangle := \int_E \varphi \, d\mu$ for integrable functions φ on E and measures μ on E. Given $n \in \mathbb{N}$, we often use boldface $\boldsymbol{X} = (x_1, \dots, x_n)$ for an element of E^n, and we write

$$m^n_{\boldsymbol{X}} := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$$

for the associated empirical measure, which lies in $\mathcal{P}(E)$. When $(E, \|\cdot\|)$ is a normed space, given $p \in [1,\infty)$, we write $\mathcal{P}^p(E, \|\cdot\|)$, or simply $\mathcal{P}^p(E)$ if the norm is understood, for the set of $\mu \in \mathcal{P}(E)$ satisfying $\langle \mu, \|\cdot\|^p\rangle < \infty$. For a separable Banach space $(E, \|\cdot\|)$, we always endow $\mathcal{P}^p(E, \|\cdot\|)$ with the p-Wasserstein metric $\mathcal{W}_{p,(E,\|\cdot\|)}$ defined by

$$\mathcal{W}_{p,(E,\|\cdot\|)}(\mu, \nu) := \inf_\pi \left(\int_{E\times E} \|x - y\|^p \pi(dx, dy)\right)^{1/p}, \tag{2.1}$$

where the infimum is over all probability measures π on $E \times E$ with marginals μ and ν. When the space E and/or the norm $\|\cdot\|$ is understood, we may omit it from the subscript in $\mathcal{W}_{p,(E,\|\cdot\|)}$, for example, by writing $\mathcal{W}_p$, or $\mathcal{W}_{p,E}$, or $\mathcal{W}_{p,\|\cdot\|}$.

For a positive integer k, we always equip $\mathbb{R}^k$ with the Euclidean norm, denoted $|\cdot|$, unless stated otherwise. For fixed $T \in (0,\infty)$, we will make use of the path spaces

$$\mathcal{C}^k := C([0,T]; \mathbb{R}^k), \quad k \in \mathbb{N},$$

which are always endowed with the supremum norm $\|x\|_\infty = \sup_{t\in[0,T]} |x_t|$. For $m \in \mathcal{P}(\mathcal{C}^k)$ and $t \in [0,T]$, we write m_t for the time-t marginal of m, that is, the image of m under the map $\mathcal{C}^k \ni x \mapsto x_t \in \mathbb{R}^k$.

2.2. *Derivatives on Wasserstein space.* The formulation of the master equation requires a suitable derivative for functions of probability measures. This section defines this notion of derivative, but it is worth noting that this paper will make no use of this notion of derivative except to state the master equation and the assumptions we impose on its solution. The main estimates derived in Section 4 of the companion paper [20] make use of properties of this derivative, but in this paper we simply apply these estimates.

For an exponent $q \in [1,\infty)$, we say that a function $V : \mathcal{P}^q(\mathbb{R}^d) \to \mathbb{R}$ is $\mathscr{C}^1$ if there exists a continuous map $\frac{\delta V}{\delta m} : \mathcal{P}^q(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ satisfying:

(i) For every $\mathcal{W}_{q,\mathbb{R}^d}$-compact set $K \subset \mathcal{P}^q(\mathbb{R}^d)$, there exists $c < \infty$ such that $\sup_{m\in K} |\frac{\delta V}{\delta m}(m, v)| \le c(1 + |v|^q)$ for all $v \in \mathbb{R}^d$.

(ii) For every $m, m' \in \mathcal{P}^q(\mathbb{R}^d)$,

$$V(m') - V(m) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta V}{\delta m}((1-t)m + tm', v)(m' - m)(dv)\, dt. \tag{2.2}$$

Note that the condition (i) is designed to make the integral in (ii) well-defined. Only one function $\frac{\delta V}{\delta m}$ can satisfy (2.2), up to a constant shift; that is, if $\frac{\delta V}{\delta m}$ satisfies (2.2) then so does $\frac{\delta V}{\delta m} + c$ for any $c \in \mathbb{R}$. For concreteness, we always choose the shift to ensure

$$\int_{\mathbb{R}^d} \frac{\delta V}{\delta m}(m, v) m(dv) = 0.$$

If $\frac{\delta V}{\delta m}(m, v)$ is continuously differentiable in v, we define its *intrinsic derivative* $D_m V : \mathcal{P}^q(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ by

$$D_m V(m, v) = D_v\left(\frac{\delta V}{\delta m}(m, v)\right),$$

where we use the notation D_v for the gradient in v. If, for each $v \in \mathbb{R}^d$, the map $m \mapsto \frac{\delta V}{\delta m}(m, v)$ is $\mathscr{C}^1$, then we say that V is $\mathscr{C}^2$ and let $\frac{\delta^2 V}{\delta m^2}$ denote its derivative, or more explicitly,

$$\frac{\delta^2 V}{\delta m^2}(m, v, v') = \frac{\delta}{\delta m}\left(\frac{\delta V}{\delta m}(\cdot, v)\right)(m, v').$$

We will also make some use of the derivative

$$D_v D_m V(m, v) = D_v\big[D_m V(m, v)\big],$$

when it exists, and we note that $D_v D_m V$ takes values in $\mathbb{R}^{d\times d}$; for some results, we will also consider higher order derivatives $D_v^k D_m V(m, v)$ with values in $\mathbb{R}^{d\times\cdots\times d} \cong \mathbb{R}^{d^{k+1}}$ for $k \in \mathbb{N}$. Finally, if V is $\mathscr{C}^2$ and if $\frac{\delta^2 V}{\delta m^2}(m, v, v')$ is twice continuously differentiable in (v, v'), we let

$$D_m^2 V(m, v, v') = D^2_{v,v'} \frac{\delta^2 V}{\delta m^2}(m, v, v')$$

denote the $d \times d$ matrix of partial derivatives $(\partial_{v_i} \partial_{v'_j} [\delta^2 V/\delta m^2](m, v, v'))_{i,j}$. Equivalently (see Lemma 2.4 in [10]),

$$D_m^2 V(m, v, v') = D_m\big(D_m V(\cdot, v)\big)(m, v').$$

2.3. *Nash systems and n-player games.* We fix throughout the paper a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, supporting independent $\mathbb{F}$-Wiener processes W of dimension d_0 (called *common noise*) and $(B^i)_{i=1}^\infty$ of dimension d (called *idiosyncratic noises*) (we choose the dimension of the idiosyncratic noises $(B^i)_{i=1}^\infty$ to be equal to the dimension of the state space for convenience only), as well as a sequence of i.i.d. $\mathcal{F}_0$-measurable $\mathbb{R}^d$-valued initial states $(X_0^i)_{i=1}^\infty$ with distribution μ_0.

We describe the n-player game and PDE systems first, deferring a precise statement of assumptions to Section 2.5. We are given an exponent $p^* \geq 1$, an action space A, assumed to be a Polish space and Borel measurable functions

$$(b, f) : \mathbb{R}^d \times \mathcal{P}^{p^*}(\mathbb{R}^d) \times A \to \mathbb{R}^d \times \mathbb{R},$$

$$g : \mathbb{R}^d \times \mathcal{P}^{p^*}(\mathbb{R}^d) \to \mathbb{R},$$

along with two matrices $\sigma \in \mathbb{R}^{d\times d}$ and $\sigma_0 \in \mathbb{R}^{d\times d_0}$, where σ is nondegenerate.

In the n-player game, players $i = 1, \dots, n$ control the state process $(\boldsymbol{X}_t = (X_t^1, \dots, X_t^n))_{t\in[0,T]}$, given by

$$dX_t^i = b\big(X_t^i, m^n_{\boldsymbol{X}_t}, \alpha^i(t, \boldsymbol{X}_t)\big)\,dt + \sigma\,dB_t^i + \sigma_0\,dW_t, \tag{2.3}$$

where we recall that $m^n_{\boldsymbol{X}_t}$ denotes the empirical measure associated with the vector $\boldsymbol{X}_t$. Here, α^i is the control chosen by player i in feedback form. The objective of player i is to try to choose α^i to minimize

$$J^{n,i}(\alpha^1, \dots, \alpha^n) = \mathbb{E}\left[\int_0^T f\big(X_t^i, m^n_{\boldsymbol{X}_t}, \alpha^i(t, \boldsymbol{X}_t)\big)\,dt + g\big(X_T^i, m^n_{\boldsymbol{X}_T}\big)\right].$$

A (closed-loop) Nash equilibrium is defined in the usual way as a vector of feedback functions $(\alpha^1, \dots, \alpha^n)$, where $\alpha^i : [0, T] \times (\mathbb{R}^d)^n \to A$ are such that the SDE (2.3) is unique in law, and

$$J^{n,i}(\alpha^1, \dots, \alpha^n) \le J^{n,i}(\alpha^1, \dots, \alpha^{i-1}, \widetilde{\alpha}, \alpha^{i+1}, \dots, \alpha^n),$$

for any alternative choice of feedback control $\widetilde{\alpha}$ such that the SDE (2.3), with α^i replaced by $\tilde{\alpha}$, is unique in law.

From the work of [3], we know that a Nash equilibrium can be built using a system of HJB equations. Define the Hamiltonian $H : \mathbb{R}^d \times \mathcal{P}^{p^*}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ by

$$H(x, m, y) = \inf_{a \in A} \big[b(x, m, a) \cdot y + f(x, m, a)\big].$$

Assume that this infimum is attained for each (x, m, y), and let $\widehat{\alpha}(x, m, y)$ denote a minimizer; we will place assumptions on the function $\widehat{\alpha}$ in the next section. Although we do not explicitly require $\widehat{\alpha}$ to be unique, the reader must be aware of the fact that Assumption A stated below is rather constraining. For instance, Assumption A requires the existence of a smooth solution to the master equation, which is described in detail in the next subsection. In all existing works on the subject, existence of a smooth solution to the master equation is in fact proven under the assumption that $\widehat{\alpha}(x, m, y)$ is unique (see, for instance, [13, 18, 26]). In [10], the Hamiltonian is smooth and $b(x, m, a) = a$: Following the proof of Theorem 2 in [35], uniqueness of the minimizer follows. It is convenient to define the functionals $\widehat{b}$ and $\widehat{f}$ on $\mathbb{R}^d \times \mathcal{P}^{p^*}(\mathbb{R}^d) \times \mathbb{R}^d$ by

$$\begin{aligned} &\widehat{b}(x, m, y) = b\big(x, m, \widehat{\alpha}(x, m, y)\big) \quad \text{and} \\ &\widehat{f}(x, m, y) = f\big(x, m, \widehat{\alpha}(x, m, y)\big), \end{aligned} \tag{2.4}$$

and note that then

$$H(x, m, y) = \widehat{b}(x, m, y) \cdot y + \widehat{f}(x, m, y). \tag{2.5}$$

The *n-player Nash system* is a PDE system for n functions, $(v^{n,i} : [0, T] \times (\mathbb{R}^d)^n \to \mathbb{R})_{i=1}^n$, given by

$$\begin{aligned} &\partial_t v^{n,i}(t, \boldsymbol{x}) + H\big(x_i, m^n_{\boldsymbol{x}}, D_{x_i} v^{n,i}(t, \boldsymbol{x})\big) \\ &\quad + \sum_{j=1, j\neq i}^{n} D_{x_j} v^{n,i}(t, \boldsymbol{x}) \cdot \widehat{b}\big(x_j, m^n_{\boldsymbol{x}}, D_{x_j} v^{n,j}(t, \boldsymbol{x})\big) \\ &\quad + \frac{1}{2} \sum_{j=1}^{n} \operatorname{Tr}\big[D^2_{x_j, x_j} v^{n,i}(t, \boldsymbol{x}) \sigma\sigma^\top\big] \\ &\quad + \frac{1}{2} \sum_{j,k=1}^{n} \operatorname{Tr}\big[D^2_{x_j, x_k} v^{n,i}(t, \boldsymbol{x}) \sigma_0\sigma_0^\top\big] = 0, \end{aligned} \tag{2.6}$$

with terminal condition $v^{n,i}(T, \boldsymbol{x}) = g(x_i, m^n_{\boldsymbol{x}})$.

Using (classical) solutions to the n-player Nash system, we may construct an equilibrium for the n-player game. The ith agent uses the feedback control

$$[0, T] \times \big(\mathbb{R}^d\big)^n \ni (t, \boldsymbol{x}) \mapsto \widehat{\alpha}\big(\boldsymbol{x}, m^n_{\boldsymbol{x}}, D_{x_i} v^{n,i}(t, \boldsymbol{x})\big).$$

As a result, the in-equilibrium state process $\boldsymbol{X} = (X^1, \dots, X^n)$ is governed by

$$dX^i_t = \widehat{b}\big(X^i_t, m^n_{\boldsymbol{X}_t}, D_{x_i} v^{n,i}(t, \boldsymbol{X}_t)\big)\, dt + \sigma\, dB^i_t + \sigma_0\, dW_t, \tag{2.7}$$

with $\widehat{b}$ defined in (2.4). Under Assumption A of Section 2.5 below, the SDE (2.7) is uniquely solvable. Indeed, due to Assumption A(4), $D_{x_i} v^{n,i}$ is at most of linear growth; moreover, the second derivatives of $v^{n,i}$ exist and are continuous, which ensures that $D_{x_i} v^{n,i}$ is locally Lipschitz. Also, Assumption A(1) and the fact that $\boldsymbol{x} \mapsto m^n_{\boldsymbol{x}}$ is a Lipschitz function from $(\mathbb{R}^d)^n$ to $(\mathcal{P}^{p^*}(\mathbb{R}^d), \mathcal{W}_{p^*,\mathbb{R}^d})$ ensure that the SDE system (2.7) has a unique strong solution.

2.4. *The mean field game and master equation.* The master equation is a PDE for a function $U : [0, T] \times \mathbb{R}^d \times \mathcal{P}^{p^*}(\mathbb{R}^d) \to \mathbb{R}$, given by

$$\begin{aligned} 0 = {} & \partial_t U(t,x,m) + H\big(x, m, D_x U(t,x,m)\big) \\ & + \frac{1}{2}\mathrm{Tr}\big[\big(\sigma\sigma^\top + \sigma_0\sigma_0^\top\big) D_x^2 U(t,x,m)\big] \\ & + \int_{\mathbb{R}^d} \widehat{b}\big(v, m, D_x U(t,v,m)\big) \cdot D_m U(t,x,m,v)\, dm(v) \\ & + \frac{1}{2}\int_{\mathbb{R}^d} \mathrm{Tr}\big[\big(\sigma\sigma^\top + \sigma_0\sigma_0^\top\big) D_v D_m U(t,x,m,v)\big]\, dm(v) \\ & + \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathrm{Tr}\big[\sigma_0\sigma_0^\top D_m^2 U\big(t,x,m,v,v'\big)\big]\, dm(v)\, dm\big(v'\big) \\ & + \int_{\mathbb{R}^d} \mathrm{Tr}\big[\sigma_0\sigma_0^\top D_x D_m U(t,x,m,v)\big]\, dm(v), \end{aligned} \tag{2.8}$$

for $(t,x,m) \in (0,T) \times \mathbb{R}^d \times \mathcal{P}^{p^*}(\mathbb{R}^d)$, with terminal condition $U(T,x,m) = g(x,m)$. The connection between the Nash system and the master equation is clarified in [10] and Proposition 4.1 of [20]; roughly speaking, $v^{n,i}(t,\boldsymbol{x})$ is expected to be close to $U(t, x_i, m^n_{\boldsymbol{x}})$ as n tends to infinity.

Just as the n-player Nash system was used to build an equilibrium for the n-player game, we will use the master equation to describe an equilibrium for the associated mean field game, described below. First, consider the McKean–Vlasov equation

$$\begin{aligned} d\mathcal{X}_t = {} & \widehat{b}\big(\mathcal{X}_t, \mu_t, D_x U(t, \mathcal{X}_t, \mu_t)\big)\, dt \\ & + \sigma\, dB^1_t + \sigma_0\, dW_t, \quad \mu = \mathcal{L}(\mathcal{X}|W), \end{aligned} \tag{2.9}$$

with initial condition $\mathcal{X}_0 = X^1_0$, where $\mathcal{L}(\mathcal{X}|W)$ denotes the conditional law of $\mathcal{X}$ given (the path) W, viewed as a random element of $\mathcal{P}^{p^*}(\mathcal{C}^d)$. Here, a solution $\mathcal{X} = (\mathcal{X}_t)_{t\in[0,T]}$ is required to be adapted to the filtration generated by the process $(X^1_0, W_t, B^1_t)_{t\in[0,T]}$. Notice that necessarily $\mu_t = \mathcal{L}(\mathcal{X}_t|W) = \mathcal{L}(\mathcal{X}_t|(W_s)_{s\in[0,t]})$ a.s., for each $t \in [0,T]$, because $(W_s - W_t)_{s\ge t}$ is independent of $(\mathcal{X}_s, W_s)_{s\le t}$. Assumptions A(1) and A(5), stated in Section 2.5 below, ensure that there is a unique strong solution to (2.9); this follows from a straightforward adaptation of the arguments in Chapter 1 of Sznitman [37] (cf. Section 7 of [15], Section 7, and Section 2.1 of [13]). For the reader who is more familiar with the PDE formulation of mean field games, we emphasize that the process $(\mu_t)_{t\in[0,T]}$ is a weak solution to the stochastic Fokker–Planck equation

$$\begin{aligned} d\mu_t = {} & -\mathrm{div}\big(\widehat{b}\big(\cdot, \mu_t, D_x U(t,\cdot,\mu_t)\big)\mu_t\big)\, dt + \frac{1}{2}\mathrm{Tr}\big[D_x^2 \mu_t\big(\sigma\sigma^\top + \sigma_0\sigma_0^\top\big)\big]\, dt \\ & - \big(\sigma_0^\top D_x \mu_t\big) \cdot dW_t, \end{aligned}$$

for $t \in [0,T]$, which follows from a straightforward application of Itô's formula to the process $(\phi(X_t))_{t\in[0,T]}$ for smooth test functions ϕ.

Since U is a classical solution to the master equation with bounded derivatives (see Assumptions A(1) and A(5) in Section 2.5 below), it is known that the measure flow μ constructed from the McKean–Vlasov equation (2.9) is the unique equilibrium of the mean field game; see, for instance, Proposition 5.106 in [12]. A mean field game equilibrium is usually defined as a fixed point of the map Φ that sends a W-measurable random measure μ on $\mathcal{C}^d$ (such that $(\mu_t)_{t\in[0,T]}$ is adapted to the filtration generated by W) to a new random measure $\Phi(\mu)$, defined as follows:

(i) Solve the stochastic optimal control problem, with μ fixed:

$$\begin{cases} \inf_{\alpha} \mathbb{E}\Big[\int_0^T f(X_t^\alpha, \mu_t, \alpha_t)\,dt + g(X_T^\alpha, \mu_T)\Big], \\ \quad \text{s.t. } dX_t = b(X_t^\alpha, \mu_t, \alpha_t)\,dt + \sigma\,dB_t^1 + \sigma_0\,dW_t, \end{cases}$$

where $(\alpha_t)_{t\in[0,T]}$ is an A-valued progressively measurable process (with respect to the filtration generated by X_0, B and W) such that SDE admits a unique strong solution and the cost functional makes sense (to simplify, we use strong solutions when dealing with stochastic optimal controls over open loop controls).

(ii) Letting X^* denote the optimally controlled state process, set $\Phi(\mu) = \mathcal{L}(X^*|W)$.

Note that if the optimization problem in step (i) has multiple solutions, the map Φ may be set-valued, and we seek μ such that $\mu \in \Phi(\mu)$. The original formulation of Lasry and Lions [33] is a forward–backward PDE system, which is essentially equivalent to this fixed-point procedure, when $\sigma_0 = 0$. When $\sigma_0 \neq 0$, the forward–backward PDE becomes stochastic, but the same connection remains. For more details on the connection between the master equation and more common PDE or probabilistic formulations of mean field games, see [4, 5, 11] or Section 1.2.4 in [10]. For our purposes, we simply take the McKean–Vlasov equation (2.9) as the definition of μ.

2.5. *Assumptions*. The following standing assumption holds throughout the paper, and this is notably the same standing assumption as in the companion paper [20] (specifically, Assumption A therein).

ASSUMPTION A.

1. A minimizer $\widehat{\alpha}(x,m,y) \in \arg\min_{a\in A}[b(x,m,a)\cdot y + f(x,m,a)]$ exists for every $(x,m,y) \in \mathbb{R}^d \times \mathcal{P}^{p^*}(\mathbb{R}^d) \times \mathbb{R}^d$, for some $p^* \in [1,2]$ such that the function $\widehat{b}(x,m,y)$ defined in (2.4) is Lipschitz in all variables. That is, there exists $C < \infty$ such that, for all $x, x', y, y' \in \mathbb{R}^d$ and $m, m' \in \mathcal{P}^{p^*}(\mathbb{R}^d)$,

$$\big|\widehat{b}(x,m,y) - \widehat{b}(x',m',y')\big| \le C\big(|x-x'| + \mathcal{W}_{p^*}(m,m') + |y-y'|\big),$$

where $\mathcal{W}_{p^*}$ is shorthand for $\mathcal{W}_{p^*,(\mathbb{R}^d,|\cdot|)}$.

2. The $d\times d$ matrix σ is nondegenerate.
3. The initial states $(X_0^i)_{i=1}^\infty$ are i.i.d. with law $\mu_0 \in \mathcal{P}^{p'}(\mathbb{R}^d)$ for some $p' > 4$.
4. For each n, the n-player Nash system (2.6) has a classical solution $(v^{n,i})_{i=1}^n$, in the sense that each function $v^{n,i}(t,\boldsymbol{x})$ is continuously differentiable in t and twice continuously differentiable in $\boldsymbol{x}$. Moreover, $D_{x_j} v^{n,i}$ has at most linear growth and $v^{n,i}$ has at most quadratic growth, for each fixed n, i, j. That is, there exist $L_{n,i} < \infty$ and $L_{n,i,j} < \infty$ such that, for all $t \in [0,T]$ and $\boldsymbol{x} \in (\mathbb{R}^d)^n$,

$$\big|D_{x_j} v^{n,i}(t,\boldsymbol{x})\big| \le L_{n,i,j}\big(1+|\boldsymbol{x}|\big),$$
$$\big|v^{n,i}(t,\boldsymbol{x})\big| \le L_{n,i}\big(1+|\boldsymbol{x}|^2\big).$$

5. The master equation admits a classical solution $U : [0, T] \times \mathbb{R}^d \times \mathcal{P}^2(\mathbb{R}^d) \ni (t, x, m) \mapsto U(t, x, m)$. The derivative $D_x U(t, x, m)$ exists and is Lipschitz in (x, m), uniformly in t (with respect to the metric $\mathcal{W}_{p^*}$ for the argument $m \in \mathcal{P}^{p^*}(\mathbb{R}^d)$), and U admits continuous derivatives $\partial_t U$, $D_x U$, $D_m U$, $D_x^2 U$, $D_v D_m U$, $D_x D_m U$ and $D_m^2 U$. Moreover, $D_x U$, $D_m U$, $D_x D_m U$ and $D_m^2 U$ are assumed to be bounded.

Recall that $|\boldsymbol{x}|$ in A(4) is the Euclidean norm of $\boldsymbol{x} \in (\mathbb{R}^d)^n$; in some places, we denote it by $\|\boldsymbol{x}\|_{n,2}$ in order to distinguish it explicitly from other norms, as in Section 3.1 below. We also need some assumptions on the growth of the function $\widehat{f}$, defined in (2.4), using of course the same function $\widehat{\alpha}$ from Assumption A(1). We provide two alternatives.

ASSUMPTION B. $\widehat{f}(x, m, y)$ is Lipschitz in y, uniformly in (x, m). That is, there exists $C < \infty$ such that, for all $x, y, y' \in \mathbb{R}^d$ and $m \in \mathcal{P}^{p^*}(\mathbb{R}^d)$,

$$|\widehat{f}(x, m, y) - \widehat{f}(x, m, y')| \leq C|y - y'|.$$

ASSUMPTION B′.

1. The solution U to the master equation is uniformly bounded.
2. The Nash system solutions $(v^{n,i})_{i=1}^n$ are bounded, uniformly in n and i.
3. $\widehat{f}(x, m, y)$ is locally Lipschitz in y with quadratic growth, uniformly in (x, m). That is, there exists $C < \infty$ such that, for all $x, y, y' \in \mathbb{R}^d$ and $m \in \mathcal{P}^{p^*}(\mathbb{R}^d)$,

$$|\widehat{f}(x, m, y) - \widehat{f}(x, m, y')| \leq C(1 + |y| + |y'|)|y - y'|.$$

These are admittedly very heavy assumptions, but they do cover a broad class of models. We refer the reader to the end of Section 1 and Section 2.4 in [20] for a detailed discussion and references. Notice that we do not place any assumptions directly on the terminal cost function g, but A(5) along with the boundary condition $U(T, x, m) = g(x, m)$ impose implicit requirements on g.

3. Statements of main results. This section summarizes the main results on the n-player Nash equilibrium empirical measures $(m^n_{\boldsymbol{X}})_{n\geq 1}$ and on their marginal flows $((m^n_{\boldsymbol{X}_t})_{t\in[0,T]})_{n\geq 1}$, defined by the SDE (2.7). Proofs are deferred to later sections. It is helpful to first recall the associated law of large numbers associated, regarding the convergence of $(m^n_{\boldsymbol{X}})_{n\geq 1}$ to μ, where μ is defined by the McKean–Vlasov equation (2.9). The first part is quoted from [20], and we elaborate here on the rate of convergence in various metrics. Define, for $p \in [1, 2]$, the constants

$$r_{n,p} = \begin{cases} n^{-1/2} & \text{if } d < 2p, \\ n^{-1/2}\log(1+n) & \text{if } d = 2p, \\ n^{-p/d} & \text{if } d > 2p. \end{cases} \tag{3.1}$$

The following law of large numbers is a slight elaboration on Theorem 3.1 of [20] and Theorem 2.13 of [10], with the short proof deferred to the end of Section 5.3.

THEOREM 3.1. *Suppose Assumption* A *holds, as well as either Assumption* B *or* B′. *Then, with* $p^* \in [1, 2]$ *as in Assumption* A,

$$\lim_{n\to\infty} \mathbb{E}[\mathcal{W}^2_{2,\mathcal{C}^d}(m^n_{\boldsymbol{X}}, \mu)] = 0,$$

and there exists $C < \infty$ such that, for each $n \geq 1$,

$$\sup_{t\in[0,T]} \mathbb{E}\big[\mathcal{W}^{p^*}_{p^*,\mathbb{R}^d}(m^n_{\boldsymbol{X}_t}, \mu_t)\big] \leq C r_{n,p^*},$$

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \mathcal{W}^2_{2,\mathbb{R}^d}(m^n_{\boldsymbol{X}_t}, \mu_t)\Big] \leq C n^{-2/(d+8)}.$$

The two different ways of estimating the rate of convergence in Theorem 3.1 (with the supremum over t inside or outside of the supremum) are somewhat standard in the theory of McKean–Vlasov equations and related particle systems. See, for instance, [10] and Chapter 6 in [13] for earlier applications in the framework of MFGs. A key point is that the distance between the initial sample in the n-player game and the initial theoretical distribution is kept stable under the Nash equilibrium dynamics. As a result, all known estimates for the rate of convergence in Theorem 3.1 do depend on the dimension d, which is a consequence of existing results on the fluctuations of the empirical distribution of a sample of i.i.d. random variables in $\mathbb{R}^d$ (see, for instance, [25]). In the central limit theorem of our companion paper [20] (see Theorem 3.2 therein) the dimension d also plays a notable role in the smoothness assumptions required of b and in the precise space in which the limit is formulated.

3.1. *Concentration inequalities in the absence of common noise.* We next look for a concentration bound for the empirical measure $m^n_{\boldsymbol{X}}$ of the Nash system, in the case of no common noise, that is, $\sigma_0 = 0$. Precisely, we work here with the empirical measure of the full paths, so that $m^n_{\boldsymbol{X}}$ is a random element of $\mathcal{P}(\mathcal{C}^d)$. We derive in this section an estimate on

$$\mathbb{P}(\mathcal{W}_{p^*,\mathcal{C}^d}(m^n_{\boldsymbol{X}}, \mu) > \epsilon), \quad \epsilon > 0.$$

The proofs of the main results, Theorems 3.2 and 3.4, of this section are given in Section 5.4.

In the following, we consider two different choices of norms on $(\mathcal{C}^d)^n$, namely the ℓ^1 and ℓ^2 norms. For $\boldsymbol{x} = (x^1, \dots, x^n) \in (\mathcal{C}^d)^n$, let

$$\|\boldsymbol{x}\|_{n,1} := \sum_{i=1}^n \|x^i\|_\infty, \qquad \|\boldsymbol{x}\|_{n,2} := \sqrt{\sum_{i=1}^n \|x^i\|_\infty^2}.$$

Note that we still always use the standard sup-norm $\|\cdot\|_\infty$ on $\mathcal{C}^d$, defined by $\|x\|_\infty = \sup_{t\in[0,T]} |x_t|$, where $|\cdot|$ is the usual Euclidean norm on $\mathbb{R}^d$. For a normed space $(E, \|\cdot\|)$, write $\mathrm{Lip}(E, \|\cdot\|)$ for the set of 1-Lipschitz functions, that is, the set of $f : E \to \mathbb{R}$ with $|f(x) - f(y)| \leq \|x - y\|$ for all $x, y \in E$. If the norm is understood, we write simply $\mathrm{Lip}(E)$.

Recall in the following that μ_0 is the law of the initial state (see Assumption A(3)). We now state our first concentration result.

THEOREM 3.2. *Assume $p^* = 1$ and $\sigma_0 = 0$, and suppose Assumption A holds, as well as either Assumption B or B′. Assume there exists $\kappa > 0$ such that*

$$\int_{\mathbb{R}^d} \exp(\kappa |x|^2) \mu_0(dx) < \infty. \tag{3.2}$$

Then there exist $C < \infty$, $\delta > 0$ such that, for every $a \geq C$, every $n \geq 1$ and every $\Phi \in \mathrm{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,1})$, we have

$$\mathbb{P}\big(\Phi(\boldsymbol{X}) - \mathbb{E}\Phi(\boldsymbol{X}) > a\big) \leq 3n \exp\big(-\delta a^2/n\big). \tag{3.3}$$

We quickly obtain a probabilistic rate of convergence, complementing Theorem 3.1.

COROLLARY 3.3. *Under the assumptions of Theorem 3.2, there exist $C < \infty$ and $\delta > 0$ such that, for every $a > 0$ and every $n \geq C/\min\{a, a^{d+8}\}$, we have*

$$\mathbb{P}\Big(\sup_{s\in[0,T]} \mathcal{W}_{1,\mathbb{R}^d}(m^n_{\boldsymbol{X}_s}, \mu_s) > a\Big) \leq 3n\exp(-\delta a^2 n). \tag{3.4}$$

PROOF. Note that $\boldsymbol{x} \mapsto \sup_{s\in[0,T]} \mathcal{W}_{1,\mathbb{R}^d}(m^n_{\boldsymbol{x}_s}, \mu_s)$ is $(1/n)$-Lipschitz from $((\mathcal{C}^d)^n, \|\cdot\|_{n,1})$ to $\mathbb{R}$. Observe also from Theorem 3.1 that

$$\mathbb{E}\Big[\sup_{s\in[0,T]} \mathcal{W}_{1,\mathbb{R}^d}(m^n_{\boldsymbol{X}_s}, \mu_s)\Big] \leq cn^{-1/(d+8)}$$

for some $c < \infty$. Then, for any $a > 0$,

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{s\in[0,T]} \mathcal{W}_{1,\mathbb{R}^d}(m^n_{\boldsymbol{X}_s}, \mu_s) > a\Big)\\
&\quad\leq \mathbb{P}\Big(\sup_{s\in[0,T]} \mathcal{W}_{1,\mathbb{R}^d}(m^n_{\boldsymbol{X}_s}, \mu_s) - \mathbb{E}\Big[\sup_{s\in[0,T]} \mathcal{W}_{1,\mathbb{R}^d}(m^n_{\boldsymbol{X}_s}, \mu_s)\Big] > a/2\Big)\\
&\qquad + \mathbb{P}\Big(\mathbb{E}\Big[\sup_{s\in[0,T]} \mathcal{W}_{1,\mathbb{R}^d}(m^n_{\boldsymbol{X}_s}, \mu_s)\Big] > a/2\Big).
\end{aligned}$$

The second term vanishes if $cn^{-1/(d+8)} \leq a/2$. The first term is bounded by the right-hand side of (3.4) when $an \geq 2\tilde{c}$, with $\tilde{c}$ being defined as the constant C in the statement of Theorem 3.2. The corollary then holds with $C = \max((2c)^{d+8}, 2\tilde{c})$. $\square$

The proof of Theorem 3.2 relies on the following well-known result of concentration of measure, borrowed from Theorem 2.3 of [22] and Theorem 3.1 of [6], which asserts that the following are equivalent:

(i) μ_0 satisfies (3.2) for some $\kappa > 0$.
(ii) There exists $\kappa > 0$ such that, for every $\varphi \in \mathrm{Lip}(\mathbb{R}^d)$, we have

$$\mu_0(\varphi - \langle\mu_0, \varphi\rangle > a) \leq \exp(-a^2/2\kappa).$$

(iii) There exists a finite constant $\kappa > 0$ such that

$$\mathcal{W}_{1,\mathbb{R}^d}(\mu_0, \nu) \leq \sqrt{2\kappa\mathcal{R}(\nu|\mu_0)}, \text{ for every } \nu \in \mathcal{P}^1(\mathbb{R}^d) \text{ with } \nu \ll \mu_0. \tag{3.5}$$

Here, $\mathcal{R}$ denotes relative entropy, defined by

$$\mathcal{R}(\nu|\mu_0) = \begin{cases} \displaystyle\int \frac{d\nu}{d\mu_0}\log\frac{d\nu}{d\mu_0}\,d\mu_0 & \text{if } \nu \ll \mu_0,\\ \infty & \text{otherwise,}\end{cases} \tag{3.6}$$

where $\nu \ll \mu_0$ denotes that ν is absolutely continuous with respect to μ_0. In fact, the change in the constant κ required between each of the conditions (i)–(iii) is universal, in particular independent of both μ_0 and the underlying metric space. We refer the reader to the book by Ledoux [34] for more discussion on concentration of measure and alternative formulations of (ii), some of which we collect in Section 5.1. The idea behind the proof of Theorem 3.2, given in Section 5.4, is to show that the law of the solution $\boldsymbol{X}$ on the path space $(\mathcal{C}^d)^n$ satisfies a transport inequality like (3.5) with a constant that depends optimally on the dimension n.

If we are willing to strengthen the condition (3.2), then we may sharpen Theorem 3.2 to make it *dimension-free*, in the sense that the bound will no longer depend on n. The proof of Theorem 3.4 below has a similar flavor to that of Theorem 3.2. The starting point for our strengthening of Theorem 3.2, in Theorem 3.4, is the remarkable result of Gozlan [27] that shows that *dimension-free* concentration is equivalent to the following quadratic transport inequality:

$$\mathcal{W}_{2,\mathbb{R}^d}(\mu_0,\nu) \le \sqrt{2\kappa \mathcal{R}(\nu|\mu_0)}, \quad \text{for every } \nu \in \mathcal{P}^2(\mathbb{R}^d) \text{ with } \nu \ll \mu_0. \tag{3.7}$$

More precisely, there exists a finite constant $\kappa > 0$ such that (3.7) holds if and only if there exists $\delta > 0$ such that for every $n \in \mathbb{N}$, every $f \in \mathrm{Lip}((\mathbb{R}^d)^n)$ (using the usual Euclidean metric on $(\mathbb{R}^d)^n$), and every $a > 0$ we have

$$\mu_0^n(f - \langle \mu_0^n, f\rangle > a) \le \exp(-\delta a^2).$$

By now, many probability measures are known to satisfy (3.7). The standard Gaussian measure on $\mathbb{R}^d$, for instance, satisfies (3.7) with $\kappa = 1$. More generally, if $\mu_0(dx) = e^{-V(x)}\,dx$ for some twice continuously differentiable function V on $\mathbb{R}^d$ with Hessian bounded below (in semidefinite order) by cI for some $c > 0$, then μ_0 satisfies (3.7) with $\kappa = 1/c$; see Corollary 7.2 in [28]. Of course, Dirac measures satisfy (3.7) trivially.

The following theorem is analogous to Theorem 3.2 but assumes (3.7) in place of (3.5), or equivalently (3.2).

THEOREM 3.4. *Assume $\sigma_0 = 0$, and suppose Assumption A holds, as well as either Assumptions B or B′. Assume there exists a finite constant $\kappa > 0$ such that (3.7) holds. Then there exist $C < \infty$ and $\delta_1, \delta_2 > 0$ such that, for every $a > 0$, every $n \ge C/a^2$, and every $\Phi \in \mathrm{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,2})$, we have*

$$\mathbb{P}\big(\Phi(\boldsymbol{X}) - \mathbb{E}\Phi(\boldsymbol{X}) > a\big) \le 2n \exp\big(-\delta_1 a^2 n\big) + 2\exp\big(-\delta_2 a^2\big). \tag{3.8}$$

We immediately obtain an improvement of Corollary 3.3.

COROLLARY 3.5. *Under the assumptions of Theorem 3.4, there exist $C < \infty$ and $\delta_1, \delta_2 > 0$ such that, for every $a > 0$ and every $n \ge C/\min(a, a^{d+8})$, we have*

$$\mathbb{P}\Big(\sup_{s\in[0,T]} \mathcal{W}_{2,\mathbb{R}^d}(m^n_{\boldsymbol{X}_s}, \mu_s) > a\Big) \le 2n \exp\big(-\delta_1 a^2 n^2\big) + 2\exp\big(-\delta_2 a^2 n\big). \tag{3.9}$$

PROOF. Similar to Corollary 3.3, this follows from Theorem 3.4: Note first that the mapping $\boldsymbol{x} \mapsto \sup_{s\in[0,T]} \mathcal{W}_{2,\mathbb{R}^d}(m^n_{\boldsymbol{x}_s}, \mu_s)$ is $n^{-1/2}$-Lipschitz from $((\mathcal{C}^d)^n, \|\cdot\|_{n,2})$ to $\mathbb{R}$. Then, by Theorem 3.1, we have

$$\mathbb{E}\Big[\sup_{s\in[0,T]} \mathcal{W}_{\mathbb{R}^d,2}(m^n_{\boldsymbol{X}_s}, \mu_s)\Big] \le c n^{-1/(d+8)}$$

for a constant $c < \infty$. □

A final notable corollary allows us to estimate the distance between the n-player and k-player games, for different population sizes n and k. This follows immediately from Corollaries 3.3 and 3.5, using the triangle inequality.

COROLLARY 3.6. *Under the assumptions of Theorem 3.2, there exist $C < \infty$ and $\delta > 0$ such that, for every $a > 0$ and every $n, k \ge C/\min\{a, a^{d+8}\}$, we have*

$$\mathbb{P}\Big(\sup_{s\in[0,T]} \mathcal{W}_{1,\mathbb{R}^d}(m^n_{\boldsymbol{X}_s}, m^k_{\boldsymbol{X}_s}) > a\Big) \le 3n \exp\big(-\delta a^2 n\big) + 3k \exp\big(-\delta a^2 k\big).$$

Alternatively, under the assumptions of Theorem 3.4, there exist $C < \infty$ and $\delta_1, \delta_2 > 0$ such that, for every $a > 0$ and every $n, k \ge C/\min(a, a^{d+8})$, we have

$$\begin{aligned}\mathbb{P}\Big(\sup_{s\in[0,T]} \mathcal{W}_{2,\mathbb{R}^d}(m^n_{\boldsymbol{X}_s}, m^k_{\boldsymbol{X}_s}) > a\Big) &\le 2n \exp\big(-\delta_1 a^2 n^2\big) + 2\exp\big(-\delta_2 a^2 n\big) \\ &\quad + 2k \exp\big(-\delta_1 a^2 k^2\big) + 2\exp\big(-\delta_2 a^2 k\big).\end{aligned}$$

REMARK 3.7. The exponent $d+8$ that appears in all of the corollaries of this section is suboptimal, stemming from our application of the second part of Theorem 3.1 (which hinges on results of [29]). But we obtained a better rate (coming from [25]) in Theorem 3.1 by taking the supremum *outside* of the expectation. With this in mind, one easily derives analogs of Corollaries 3.3, 3.5 and 3.6 in which the supremum is outside of the probability and expectation. For instance, in the setting of Corollary 3.3, there exist constants $C<\infty$ and $\delta>0$ such that for every $a>0$ and $n\in\mathbb{N}$ satisfying $a\ge C\max\{n^{-1},r_{n,1}\}$ we have

$$\sup_{s\in[0,T]}\mathbb{P}\big(\mathcal{W}_{1,\mathbb{R}^d}\big(m^n_{\boldsymbol{X}_s},\mu_s\big)>a\big)\le 3n\exp\big(-\delta na^2\big).$$

The key advantage is that the requirement $a\ge C\max\{n^{-1},r_{n,1}\}$ is much weaker; for a fixed a this inequality "kicks in" for much smaller n, as $r_{n,1}\le n^{-1/(d+8)}$.

REMARK 3.8. When there is common noise, it is natural to wonder what remains of these concentration bounds. One certainly cannot expect exactly the same results to hold, because concentration requires a degree of independence; for example, in the degenerate case where $X^i\equiv W$ for all i, and Theorems 3.2 and 3.4 clearly fail. See Remark 5.7 for a brief discussion of this possibility.

Lastly, we discuss similar concentration inequalities for the in-equilibrium controls themselves. We find that the results are most naturally stated in terms of the natural coupling of the in-equilibrium state processes with i.i.d. copies of the solution of the McKean–Vlasov equation, driven by the same Brownian motions and initial states

$$d\mathcal{X}^i_t=\widehat{b}\big(\mathcal{X}^i_t,\mu_t,D_xU\big(t,\mathcal{X}^i_t,\mu_t\big)\big)\,dt+\sigma\,dB^i_t+\sigma_0\,dW_t,\qquad \mu=\mathcal{L}\big(\mathcal{X}^i|W\big),$$

with initial condition $\mathcal{X}^i_0=X^i_0$, where $\mathcal{L}(\mathcal{X}^i|W)$ denotes the conditional law of $\mathcal{X}$ given (the path) W. The proof of the following is given in Section 5.4.

THEOREM 3.9. *Suppose Assumption A holds, as well as either Assumptions B or B′. Assume $\widehat{\alpha}(x,m,y)$ is Lipschitz on $\mathbb{R}^d\times\mathcal{P}^{p^*}(\mathbb{R}^d)\times\mathbb{R}^d$. Define the in-equilibrium controls*

$$\alpha^{n,i}_t=\widehat{\alpha}\big(X^i_t,m^n_{\boldsymbol{X}_t},D_{x_i}v^{n,i}(t,\boldsymbol{X}_t)\big),$$

and the limiting controls

$$\beta^i_t=\widehat{\alpha}\big(\mathcal{X}^i_t,\mu_t,D_xU\big(t,\mathcal{X}^i_t,\mu_t\big)\big),$$

for $t\in[0,T]$. Then, with $r_{n,p}$ defined as in (3.1), we have for each $n\ge 1$

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n\int_0^T\big|\alpha^{n,i}_t-\beta^i_t\big|^{p^*}\,dt\right]\le Cr_{n,p^*}. \tag{3.10}$$

If we assume also that $\sigma_0=0$, then:

(i) *If $p^*=1$ and (3.2) holds for some $\kappa>0$, then there exist $C<\infty$ and $\delta_1>0$ such that for every $a>0$ and every $n\ge C/\min\{a,a^{d+8}\}$,*

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n\int_0^T\big|\alpha^{n,i}_t-\beta^i_t\big|^2\,dt>a^2\right)\le 3ne^{-\delta_1a^2n}+2ne^{-\delta_2a^2n^2}.$$

(ii) *If (3.7) holds for some $\kappa>0$, then there exist $C<\infty$ and $\delta_1,\delta_2>0$ such that for every $a>0$ and every $n\ge C/\min\{a,a^{d+8}\}$,*

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n\int_0^T\big|\alpha^{n,i}_t-\beta^i_t\big|^2\,dt>a^2\right)\le 4ne^{-\delta_1a^2n^2}+2e^{-\delta_2a^2n}.$$

The assumption that $\widehat{\alpha}$ is Lipschitz in its arguments is not unreasonable; it is valid in the most common case where $b(x,m,a)=a$ and $f(x,m,a)=-\frac{1}{2}|a|^2$, for example. See also Section 3.1.4 of [12] for additional examples, at least when $\widehat{\alpha}$ does not depend on the measure, together with Lemma 6.18 in Chapter 6 of [12] for cases when $\widehat{\alpha}$ is allowed to depend on μ. Notably, the controls appear in Theorem 3.9 integrated in time; analogous results, with the integral replaced by a $\sup_{t\in[0,T]}$, would require a very different and likely more involved argument.

3.2. *Large deviations*. In this section, we state a large deviation principle (LDP) for the sequence $(m^n_{\boldsymbol{X}_t})_{t\in[0,T]}$ regarded as a sequence of random variables with values in the space $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$, where $\mathcal{P}^1(\mathbb{R}^d)$ is equipped with the 1-Wasserstein distance, and $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ is equipped with the resulting uniform topology. Below, let $C_c^\infty(\mathbb{R}^d)$ denote the space of smooth compactly supported functions on $\mathbb{R}^d$. It is convenient here to define

$$\begin{aligned}\widetilde{b}(t,x,m) &:= \widehat{b}\big(x,m,D_xU(t,x,m)\big)\\ &= b\big(x,m,\widehat{\alpha}\big(x,m,D_xU(t,x,m)\big)\big),\end{aligned} \tag{3.11}$$

with $\widehat{\alpha}$ being the minimizer in Assumption A(1).

Following [19], we now introduce the action functional, which requires the following definition: we say that a distribution-valued path $t\mapsto \nu_t$ defined on $[0,T]$ is absolutely continuous if, for each compact set $K\subset\mathbb{R}^d$, there exists a neighborhood U_K of 0 (for the inductive topology) in the space $C_K(\mathbb{R}^d)$ of functions in $C_c^\infty(\mathbb{R}^d)$ whose support is included in K and an absolutely continuous function $\delta_K:[0,T]\to\mathbb{R}$ such that

$$\big|\langle\mu_t,f\rangle-\langle\mu_s,f\rangle\big|\le\big|\delta_K(t)-\delta_K(s)\big|,\qquad s,t\in[0,T],\ f\in U_K.$$

We refer to [19] for more details. The action functional $I:C([0,T];\mathcal{P}^1(\mathbb{R}^d))\to[0,\infty]$ is then given by

$$I(\nu)=\begin{cases}\dfrac{1}{2}\displaystyle\int_0^T\big\|\dot{\nu}_t-\mathcal{L}^*_{t,\nu_t}\nu_t\big\|^2_{\nu_t}\,dt & \text{if } t\mapsto\nu_t \text{ is absolutely continuous,}\\ \infty & \text{otherwise,}\end{cases} \tag{3.12}$$

where, for $(t,m)\in[0,T]\times\mathcal{P}^1(\mathbb{R}^d)$, $\mathcal{L}^*_{t,m}$ is the formal adjoint of the operator

$$\mathcal{L}_{t,m}\varphi=\frac{1}{2}\operatorname{Tr}\big[\sigma\sigma^\top D_x^2\varphi\big]+D_x\varphi\cdot\widetilde{b}(t,\cdot,m),$$

for $\varphi\in C_c^\infty(\mathbb{R}^d)$, and the seminorm $\|\cdot\|_m$ acts on Schwartz distributions by

$$\|\gamma\|_m^2:=\sup_{\substack{\varphi\in C_c^\infty(\mathbb{R}^d)\\ \langle m,|D_x\varphi|^2\rangle\neq0}}\frac{\langle\gamma,\varphi\rangle^2}{\langle m,|D_x\varphi|^2\rangle},$$

the notation $\langle\cdot,\cdot\rangle$ here denoting the duality bracket.

We may now state the first main LDP, which covers the case without common noise ($\sigma_0=0$). Its proof is given in Section 6.1.3.

THEOREM 3.10. *Assume $p^*=1$ and $\sigma_0=0$, and suppose Assumption A and either Assumption B or B′ hold. Suppose also that*

$$\int_{\mathbb{R}^d}\exp(\lambda|x|)\mu_0(dx)<\infty\quad\textit{for all }\lambda>0.$$

Then the sequence $(m^n_{\boldsymbol{X}_t},t\in[0,T])_{n\in\mathbb{N}}$ satisfies a large deviation principle on $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$, with good rate function $\nu=(\nu_t)_{t\in[0,T]}\mapsto I(\nu)+\mathcal{R}(\nu_0|\mu_0)$, where I is given by (3.12) and $\mathcal{R}$ is as in (3.5).

This follows almost immediately from the results of [19] on large deviations for McKean–Vlasov particle systems, once the exponential equivalence of the Nash system and the McKean–Vlasov system is established. However, we revisit this classical question of large deviations from the McKean–Vlasov limit and provide a simpler self-contained proof based on the contraction principle, which is possible in our setting because the volatility coefficients are constant. Our main interest in providing our own proof is in addressing the case with common noise, for which there are no known results. This leads to the *weak LDP* of Theorem 3.11 below, for which we must first develop some notation.

We first introduce $(\tau_x : \mathbb{R}^d \ni z \mapsto z - x)_{x \in \mathbb{R}^d}$ the group of translations on $\mathbb{R}^d$, as well as the orthogonal projection $\Pi_{\sigma^{-1}\sigma_0}$ from $\mathbb{R}^d$ onto the image of $\sigma^{-1}\sigma_0$. Then, for any continuous path ϕ from $[0,T]$ into $\mathbb{R}^d$, we define $\widetilde{I}^\phi$ to be the rate function as given by (3.12), but modified by replacing the drift $\widetilde{b}$ with $(t,x,m) \mapsto \widetilde{b}(t, x+\phi_t, m \circ \tau_{-\phi_t}^{-1})$ where it appears in the operator $\mathcal{L}_{t,m}$. Also, for a path $\nu \in C([0,T]; \mathcal{P}^1(\mathbb{R}^d))$, we let

$$\mathbb{M}_t^{\widetilde{b},\nu} := \left(\sigma \Pi_{\sigma^{-1}\sigma_0} \sigma^{-1} \left(\int_{\mathbb{R}^d} x \, d(\nu_t - \nu_0)(x) - \int_0^t \langle \nu_s, \widetilde{b}(s, \cdot, \nu_s) \rangle \, ds \right) \right)_{t \in [0,T]}.$$

This allows us to define the following functional:

$$J^{\sigma_0}(\nu) = \widetilde{I}^{\mathbb{M}^{\widetilde{b},\nu}} \big((\nu_t \circ \tau^{-1}_{\mathbb{M}_t^{\widetilde{b},\nu}})_{t \in [0,T]} \big).$$

We may now state the weak LDP, valid even when there is common noise. Its proof is deferred to Section 6.2.2. Recall in the following that $\mathcal{R}$ denotes the relative entropy, defined in (3.6).

THEOREM 3.11. *Assume $p^* = 1$, and suppose Assumption A and either Assumptions B or B′ hold. Suppose also that*

$$\int_{\mathbb{R}^d} \exp(\lambda |x|) \mu_0(dx) < \infty, \quad \textit{for all } \lambda > 0.$$

Then the sequence $(m^n_{X_t}, t \in [0,T])_{n \in \mathbb{N}}$ satisfies the following weak large deviation principle in $C([0,T]; \mathcal{P}^1(\mathbb{R}^d))$:

(i) *For any open subset O of $C([0,T]; \mathcal{P}^1(\mathbb{R}^d))$,*

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}(m^n_{X.} \in O) \geq - \inf_{\nu \in O} \big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0 | \mu_0) \big).$$

(ii) *For any compact subset K of $C([0,T]; \mathcal{P}^1(\mathbb{R}^d))$,*

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(m^n_{X.} \in K) \leq - \inf_{\nu \in K} \big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0 | \mu_0) \big).$$

(iii) *For any closed subset F of $C([0,T]; \mathcal{P}^1(\mathbb{R}^d))$,*

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(m^n_{X.} \in F) \leq - \lim_{\delta \searrow 0} \inf_{\nu \in F_\delta} \big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0 | \mu_0) \big),$$

where $F_\delta = \{\nu \in C([0,T]; \mathcal{P}^1(\mathbb{R}^d)) : \inf_{\widetilde{\nu} \in F} \sup_{t \in [0,T]} \mathcal{W}_1(\widetilde{\nu}_t, \nu_t) \leq \delta\}$.

It must be stressed that J^{σ_0} coincides with I when $\sigma_0 = 0$ since the image of σ_0 reduces to $\{0\}$, the process $\mathbb{M}^{\widetilde{b},\nu}$ is null and $\widetilde{I}^0 = I$.

We also emphasize that other forms of the rate function J^{σ_0} are given in Section 6. For instance, the formulation provided in Proposition 6.5 is certainly more tractable than the one given just prior to Theorem 3.11, but it has the major drawback of holding only for a special class of paths ν. In fact, all these different expressions for J^{σ_0} convey the same idea: As soon as σ_0 differs from the null matrix, the rate function is not a good rate function, that is to say, its level sets are not compact. The reason is quite clear: the common noise permits to shift for free the mean of ν in the directions included in the image of σ_0. In words, $J^{\sigma_0}(\nu)$ may remain bounded even if the mean path of ν has higher and higher oscillations.

To illustrate the latter fact, let $\phi \in \mathcal{C}^d$ with $\phi_0 = 0$, call $\overline{X}^\phi$ the solution to the McKean–Vlasov equation:

$$d\overline{X}_t^\phi = \widetilde{b}\big(t, \overline{X}_t^\phi, \mathcal{L}(\overline{X}_t^\phi)\big)\,dt + \sigma\,dB_t^1 + \sigma_0\dot{\phi}_t\,dt, \quad t \in [0,T],$$

and let $\nu = (\mathcal{L}(\overline{X}_t^\phi))_{t\in[0,T]}$ denote its flow of marginal laws. In that case, $\mathbb{M}^{b,\nu}$ coincides with $\sigma_0\phi$, and thus $(\nu_t \circ \tau^{-1}_{\mathbb{M}_t^{b,\nu}})_{t\in[0,T]}$ is the flow of marginal laws of $(\overline{X}_t^\phi - \sigma_0\phi_t)_{t\in[0,T]}$, the latter solving the McKean–Vlasov equation (with no common noise) with drift $\widetilde{b}$ given by $(t,x,m) \mapsto b(t, x + \sigma_0\phi_t, m \circ \tau^{-1}_{-\sigma_0\phi_t})$. As a result, $\widetilde{I}^{\mathbb{M}^{\widetilde{b},\nu}}((\nu_t \circ \tau^{-1}_{\mathbb{M}_t^{\widetilde{b},\nu}})_{t\in[0,T]})$ is null, whatever ϕ is.

4. Main estimates. The results announced in Section 3 hinge on the estimates developed in this section. We begin by recalling two key estimates from [20], which we then use to derive the central exponential approximation of Theorem 4.3.

In the following results and proofs, U is the classical solution to the master equation (2.8). The letter C denotes a generic positive constant, which may change from line to line but is universal in the sense that it never depends on i or n, though it may of course depend on model parameters, including, for example, the bounds on the growth and the regularity of U and its derivatives, the Lipschitz constants of b and f, and the time horizon T.

To proceed, we define an n-particle SDE system of McKean–Vlasov type, which we will compare to the true Nash system. Precisely, let $\overline{\boldsymbol{X}} = (\overline{X}^1, \dots, \overline{X}^n)$ solve the approximating n-particle system

$$d\overline{X}_t^i = \widehat{b}\big(\overline{X}_t^i, m^n_{\overline{\boldsymbol{X}}_t}, D_xU\big(t, \overline{X}_t^i, m^n_{\overline{\boldsymbol{X}}_t}\big)\big)\,dt + \sigma\,dB_t^i + \sigma_0\,dW_t, \quad \overline{X}_0^i = X_0^i. \tag{4.1}$$

Because of Assumptions A(1) and A(5), this SDE system admits a unique strong solution.

We make the following abbreviations: For $(t,\boldsymbol{x}) \in [0,T] \times (\mathbb{R}^d)^n$, define

$$u^{n,i}(t,\boldsymbol{x}) = U\big(t, x_i, m^n_{\boldsymbol{x}}\big).$$

Also, in what follows, for $i = 1, \dots, n$, define

$$M_t^i = \int_0^t \sum_{j=1}^n \big(D_{x_j}v^{n,i}(s, \boldsymbol{X}_s) - D_{x_j}u^{n,i}(s, \boldsymbol{X}_s)\big) \cdot \sigma\,dB_s^j \tag{4.2}$$

$$+ \int_0^t \sum_{j=1}^n \big(D_{x_j}v^{n,i}(s, \boldsymbol{X}_s) - D_{x_j}u^{n,i}(s, \boldsymbol{X}_s)\big) \cdot \sigma_0\,dW_s, \tag{4.3}$$

$$N_t^i = \int_0^t \big(v^{n,i}(s, \boldsymbol{X}_s) - u^{n,i}(s, \boldsymbol{X}_s)\big)\,dM_s^i. \tag{4.4}$$

We may now state the main estimates from Theorems 4.2 and 4.6 of [20]. These two estimates are quite similar, but one holds under Assumption B and the other under Assumption B′.

THEOREM 4.1. *Suppose Assumptions A and B hold. Then there exists $C < \infty$ such that, for each n,*

$$\frac{1}{n}\sum_{i=1}^n \mathbb{E}\Big[\int_0^T \big|D_{x_i}v^{n,i}(t, \boldsymbol{X}_t) - D_xU\big(t, X_t^i, m^n_{\boldsymbol{X}_t}\big)\big|^2\,dt\Big] \le \frac{C}{n^2}, \tag{4.5}$$

$$\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^n \|X^i - \overline{X}^i\|_\infty^2\Big] \le \frac{C}{n^2}. \tag{4.6}$$

Moreover,

$$\frac{1}{n}\sum_{i=1}^{n}\|X^i-\overline{X}^i\|_\infty^2 \le \frac{C}{n}\sum_{i=1}^{n}[M^i]_T+\frac{C}{n^2}, \tag{4.7}$$

and for all $t\in[0,T]$,

$$\frac{1}{n}\sum_{i=1}^{n}[N^i]_t \le \frac{C}{n^3}\sum_{i=1}^{n}[M^i]_t, \quad \textit{and} \quad \frac{1}{n}\sum_{i=1}^{n}[M^i]_T \le \frac{C}{n^2}+\frac{C}{n}\sum_{i=1}^{n}|N_T^i|. \tag{4.8}$$

THEOREM 4.2. *Suppose Assumptions A and B′ hold. Then (4.7) holds, and, for sufficiently large n, the estimates (4.5) and (4.6) hold. For $i=1,\dots,n$ and a constant $\eta>0$, define M^i as in (4.2) and Q^i by*

$$\begin{aligned}Q_t^i=\int_0^t\big[&2\big(v^{n,i}(s,\boldsymbol{X}_s)-u^{n,i}(s,\boldsymbol{X}_s)\big)\\&+\eta\sinh\big(\eta\big(v^{n,i}(s,\boldsymbol{X}_s)-u^{n,i}(s,\boldsymbol{X}_s)\big)\big)\big]\,dM_s^i.\end{aligned}$$

Then, for sufficiently large n and η, we have for all $t\in[0,T]$,

$$\frac{1}{n}\sum_{i=1}^{n}[Q^i]_t \le \frac{C}{n^3}\sum_{i=1}^{n}[M^i]_t, \quad \textit{and} \quad \frac{1}{n}\sum_{i=1}^{n}[M^i]_T \le \frac{C}{n^2}+\frac{C}{n}\sum_{i=1}^{n}|Q_T^i|. \tag{4.9}$$

The main estimate for our purposes is the following theorem, which provides an exponential estimate of the distance between the solutions $\boldsymbol{X}$ and $\overline{\boldsymbol{X}}$ of the SDEs (2.7) and (4.1), respectively. These estimates will also serve us well in our study of large deviations in Section 6.

THEOREM 4.3. *Suppose Assumption A holds, as well as either Assumption B or B′. Then there exist constants $\kappa_1,\kappa_2\in(0,\infty)$ such that for every $\epsilon>0$ and $n\ge\kappa_1/\epsilon$ we have*

$$\begin{aligned}\mathbb{P}\big(\mathcal{W}_{2,\mathcal{C}^d}(m_{\boldsymbol{X}}^n,m_{\overline{\boldsymbol{X}}}^n)>\epsilon\big)&\le\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}\|X^i-\overline{X}^i\|_\infty^2>\epsilon^2\right)\\&\le 2n\exp\left(-\frac{\epsilon^2n^2}{\kappa_2}\right).\end{aligned}\tag{4.10}$$

The constants κ_1 and κ_2 depend (in an increasing manner) only on the Lipschitz constants and uniform bounds of the coefficients in Assumptions A and B or B′.

PROOF. The first inequality in (4.10) is an immediate consequence of the definition (2.1) of the 2-Wasserstein metric. Turning to the second inequality, we prove the case where Assumption B holds; the proof under Assumption B′ is obtained by simply replacing every occurrence of N^i, Theorem 4.1 and the estimate (4.8) with Q^i, Theorem 4.2, and (4.9), respectively. Recall the definitions of M^i and N^i from (4.2) and (4.4). Use (4.7) to get

$$\frac{1}{n}\sum_{i=1}^{n}\|X^i-\overline{X}^i\|_\infty^2 \le \frac{c_0}{n}\sum_{i=1}^{n}[M^i]_T+\frac{c_0}{n^2}, \tag{4.11}$$

where $c_0<\infty$ is a constant (independent of n), which we will now keep track of to clarify the following arguments. From Theorem 4.1, we have the estimates

$$\frac{1}{n}\sum_{i=1}^{n}[N^i]_t \le \frac{c_1}{n^3}\sum_{i=1}^{n}[M^i]_t, \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n}[M^i]_T \le \frac{c_2}{n^2}+\frac{c_3}{n}\sum_{i=1}^{n}|N_T^i|, \tag{4.12}$$

where the constants $c_1, c_2, c_3 < \infty$ do not depend on i or n. Fix i for the moment, as well as $\delta, \gamma > 0$, to be determined later. Note that for every continuous local martingale R, we have $\mathbb{E}[\exp(R_T - \frac{1}{2}[R]_T)] \le 1$. Combining this with Markov's inequality, we have for each $i = 1, \dots, n$,

$$\mathbb{P}\Big(\gamma N_T^i \ge \delta\gamma + \frac{\gamma^2}{2}[N^i]_T\Big) \le \exp(-\delta\gamma)$$

$$\text{and} \quad \mathbb{P}\Big(-\gamma N_T^i \ge \delta\gamma + \frac{\gamma^2}{2}[N^i]_T\Big) \le \exp(-\delta\gamma).$$

Thus, defining the event $A_n = \{\exists i \in \{1, \dots, n\} : \gamma|N_T^i| \ge \delta\gamma + \frac{\gamma^2}{2}[N^i]_T\}$, we have

$$\mathbb{P}(A_n) \le \sum_{i=1}^n \mathbb{P}\Big(\gamma|N_T^i| \ge \delta\gamma + \frac{\gamma^2}{2}[N^i]_T\Big) \le 2n\exp(-\delta\gamma).$$

On the other hand, on A_n^c,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n [M^i]_T &\le \frac{c_2}{n^2} + \frac{c_3}{n}\sum_{i=1}^n |N_T^i| \\ &\le \frac{c_2}{n^2} + c_3\delta + \frac{c_3\gamma}{2n}\sum_{i=1}^n [N^i]_T \\ &\le \frac{c_2}{n^2} + c_3\delta + c_1c_3\frac{\gamma}{2n^3}\sum_{i=1}^n [M^i]_T,\end{aligned}$$

and for $n^2 \ge (c_1c_3\gamma) \vee (c_2/c_3\delta)$ it holds that $\frac{1}{n}\sum_{i=1}^n [M^i]_T \le 4c_3\delta$. Thus, for any such n,

$$\mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^n [M^i]_T > 4c_3\delta\Big) \le \mathbb{P}(A_n) \le 2n\exp(-\delta\gamma). \tag{4.13}$$

Recalling (4.11), we may choose $\epsilon > 0$ and set $\delta = \epsilon^2/8c_3c_0$ to get

$$\begin{aligned}\mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^n \|X^i - \overline{X}^i\|_\infty^2 > \epsilon^2\Big) &\le \mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^n [M^i]_T > \frac{\epsilon^2}{c_0} - \frac{1}{n^2}\Big) \\ &\le \mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^n [M^i]_T > \frac{\epsilon^2}{2c_0}\Big) \\ &\le 2n\exp\Big(-\frac{\epsilon^2\gamma}{8c_3c_0}\Big),\end{aligned}$$

whenever $n^2 \ge (c_1c_3\gamma) \vee (8c_0c_2/\epsilon^2) \vee (2c_0/\epsilon^2)$. In particular, choose $\gamma = n^2/c_1c_3$ to deduce (4.10), with $\kappa_1 = \sqrt{(8c_2c_0) \vee (2c_0)}$ and $\kappa_2 = 16c_0c_1c_3^2$. $\square$

5. Proofs of concentration inequalities. In this section, we prove the claims of Section 3.1. Due to Theorem 4.3, it remains only to find concentration estimates for the McKean–Vlasov system $\overline{\boldsymbol{X}}$. We did not find directly applicable results for this, so we develop our own in Sections 5.1–5.3 below. Finally, in Section 5.4 we address the MFG system.

5.1. *Review of concentration inequalities*. We begin by reviewing known results characterizing concentration in terms of transport inequalities, combining well-known facts about sub-Gaussian random variables with Proposition 6.3 of [28] and Theorem 3.1 of [6]. Recall the definition of relative entropy $\mathcal{R}$ from (3.6).

THEOREM 5.1. *Let* $(E, \|\cdot\|)$ *be a separable Banach space and* $\theta \in \mathcal{P}^1(E)$. *Let* $\kappa > 0$. *Consider the following statements*:

(i) *For all* $\nu \ll \theta$,

$$\mathcal{W}_{1,E}(\theta, \nu) \le \sqrt{2\kappa \mathcal{R}(\nu|\theta)}.$$

(ii) *For every* $\lambda \in \mathbb{R}$ *and* $\varphi \in \mathrm{Lip}(E, \|\cdot\|)$,

$$\int_E \exp(\lambda(\varphi - \langle \theta, \varphi\rangle))\theta(dx) \le \exp(\kappa\lambda^2/2).$$

(iii) *For every* $a > 0$ *and* $\varphi \in \mathrm{Lip}(E, \|\cdot\|)$,

$$\theta\big(\varphi - \langle \theta, \varphi\rangle > a\big) \le \exp(-a^2/2\kappa).$$

(iv) *We have* $\int_E \exp(\|x\|^2/6\kappa)\theta(dx) < \infty$.

Then (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv). *Moreover, if* (iv) *holds for a given* κ, *then* (i) *holds with* κ *replaced by*

$$\kappa' = 6\left(1 + 4\log \int_E \exp(\|x\|^2/6\kappa)\mu(dx)\right).$$

In particular, (i)–(iv) *are equivalent up to a universal change in the constant* κ.

In addition, we will need two well-known tensorization results, both of which follow from Proposition 1.9 of [28]. In what follows, given a separable Banach space $(E, \|\cdot\|)$ and $p \ge 1$, by $(E^n, \|\cdot\|_{n,p})$ we will mean E^n equipped with the ℓ^p norm,

$$\|\boldsymbol{x}\|_{n,p} = \left(\sum_{i=1}^n \|x_i\|^p\right)^{1/p}, \tag{5.1}$$

for $\boldsymbol{x} = (x_1, \dots, x_n) \in E^n$. The subscript in $\|\cdot\|_{n,p}$ indicates that we are using the ℓ^p norm on the n-fold product space; while one might more descriptively include the space E^n itself in the subscript, the underlying space E should always be clear from context. Typically, p will be either 1 or 2.

THEOREM 5.2. *Let* $(E, \|\cdot\|)$ *be a separable Banach space,* $\kappa > 0$ *and* $\theta \in \mathcal{P}^1(E)$.

(i) *Suppose* $\mathcal{W}_{1,E}(\theta, \nu) \le \sqrt{2\kappa \mathcal{R}(\nu|\theta)}$, *for all* $\nu \ll \theta$. *Then, for all* $\nu \ll \theta^n$, *we have*

$$\mathcal{W}_{1,(E^n,\|\cdot\|_{n,1})}(\theta^n, \nu) \le \sqrt{2n\kappa \mathcal{R}(\nu|\theta^n)}.$$

(ii) *Suppose* $\mathcal{W}_{2,E}(\theta, \nu) \le \sqrt{2\kappa \mathcal{R}(\nu|\theta)}$, *for all* $\nu \ll \theta$. *Then, for all* $\nu \ll \theta^n$, *we have*

$$\mathcal{W}_{1,(E^n,\|\cdot\|_{n,2})}(\theta^n, \nu) \le \mathcal{W}_{2,(E^n,\|\cdot\|_{n,2})}(\theta^n, \nu) \le \sqrt{2\kappa \mathcal{R}(\nu|\theta^n)}.$$

The key difference between (i) and (ii) in Theorem 5.2 is of course that (ii) is *dimension-free*. Before we can apply these general principles to the study of concentration of interacting diffusions of McKean–Vlasov type, we first quote a slight modification of Corollary 4.1 of [22] (alternatively, see Theorem 1 of [38]).

THEOREM 5.3. *For $k \in \mathbb{N}$, suppose $b : [0,T] \times \mathbb{R}^k \to \mathbb{R}^k$ is jointly measurable and there exists $L < \infty$ such that*

$$|b(t,x) - b(t,y)| \le L|x-y|, \quad \text{for all } x, y \in \mathbb{R}^k.$$

Assume also that

$$\sup_{t\in[0,T]} |b(t,0)| < \infty. \tag{5.2}$$

For another $k' \in \mathbb{N}$, let $\sigma \in \mathbb{R}^{k\times k'}$, and let $\|\sigma\|_{\mathrm{op}} = \sup\{|\sigma x| : x \in \mathbb{R}^{k'}, |x| \le 1\}$ denote the operator norm. Fix a probability space supporting a k'-dimensional Wiener process W. Finally, let $X^x = (X_t^x)_{t\in[0,T]}$ denote the unique strong solution to the SDE

$$dX_t^x = b(t, X_t^x)\,dt + \sigma\,dW_t, \quad X_0 = x,$$

and let $P_x \in \mathcal{P}(C([0,T];\mathbb{R}^k))$ denote the law of X^x. Then there exists $\kappa < \infty$, depending only on T, L, and $\|\sigma\|_{\mathrm{op}}$ (and not on the values of k, k', (5.2)), such that, for all $x \in \mathbb{R}^k$ we have

$$\mathcal{W}_{1,(\mathcal{C}^k,\|\cdot\|_{k,2})}(Q, P_x) \le \sqrt{2\kappa\mathcal{R}(Q|P_x)}, \quad \text{for all } Q \in \mathcal{P}^1(\mathcal{C}^k) \text{ with } Q \ll P_x. \tag{5.3}$$

In particular, it holds for every $a > 0$ and $\Phi \in \mathrm{Lip}(\mathcal{C}^k, \|\cdot\|_{k,2})$ that

$$P_x(\Phi - \langle P_x, \Phi\rangle > a) \le \exp(-a^2/2\kappa).$$

PROOF. This would follow immediately from [22], Corollary 4.1 (or [38], Theorem 1), except that we are using the operator norm instead of the Hilbert–Schmidt (Frobenius) norm for σ. It is straightforward to check that their proof goes through with no change and that the constant κ does not depend on the values of k, k' or $\sup_{t\in[0,T]}|b(t,0)|$. The final claim ("in particular") follows from the implication (i) $\Rightarrow$ (iii) in Theorem 5.1. $\square$

5.2. *McKean–Vlasov concentration inequalities.* We now specialize this result to obtain concentration bounds for interacting diffusions. Let $B^1, \dots, B^n$ be i.i.d. standard Wiener processes of dimension d. We are given a parameter $p \in [1,2]$, to be specified later, and a drift $\widetilde{b} : [0,T] \times \mathbb{R}^d \times \mathcal{P}^p(\mathbb{R}^d) \to \mathbb{R}^d$ which is Lipschitz in the space and measure arguments; more precisely, there exists $\widetilde{L} < \infty$ such that

$$|\widetilde{b}(t,x,m) - \widetilde{b}(t',x',m')| \le \widetilde{L}(|x-x'| + \mathcal{W}_p(m,m')), \quad t \in [0,T]. \tag{5.4}$$

Assume also that

$$\sup_{t\in[0,T]} |\widetilde{b}(t,0,\delta_0)| < \infty. \tag{5.5}$$

Lastly, we are given $\sigma \in \mathbb{R}^{d\times d}$. Now, consider the n-particle system $\widetilde{\boldsymbol{X}} = (\widetilde{X}^1, \dots, \widetilde{X}^n)$ that is the unique strong solution to the SDE system

$$d\widetilde{X}_t^i = \widetilde{b}(t, \widetilde{X}_t^i, m^n_{\widetilde{\boldsymbol{X}}_t})\,dt + \sigma\,dB_t^i, \tag{5.6}$$

with initial conditions $\widetilde{X}_0^1, \dots, \widetilde{X}_0^n$ which are i.i.d. with law $\widetilde{\mu}_0$ satisfying $\mathbb{E}[|\widetilde{X}_0^1|^2] < \infty$.

For $\boldsymbol{x} \in (\mathbb{R}^d)^n$, let $P_{\boldsymbol{x}} \in \mathcal{P}((\mathcal{C}^d)^n)$ denote the law of the solution to the SDE system (5.6) started from initial states $(\widetilde{X}_0^1, \dots, \widetilde{X}_0^n) = \boldsymbol{x}$. Then $\boldsymbol{x} \mapsto P_{\boldsymbol{x}}$ is a version of the conditional law of $\widetilde{\boldsymbol{X}}$ given $\widetilde{\boldsymbol{X}}_0$. Moreover, for any $\boldsymbol{x}$ and $\boldsymbol{y}$ in $(\mathbb{R}^d)^n$ we can couple $P_{\boldsymbol{x}}$ and $P_{\boldsymbol{y}}$ in the usual way, by solving the SDE system from the two initial states with the same Brownian motion.

Let $\pi_{\boldsymbol{x},\boldsymbol{y}}$ denote this coupling. In what follows, we will make use of the following standard estimates: Under assumption (5.4), there exists a constant c that depends only on T, p, and $\widetilde{L}$ (and not on n or the value of (5.5)), such that

$$
(5.7)\qquad \big|\langle P_{\boldsymbol{x}}, \Phi\rangle - \langle P_{\boldsymbol{y}}, \Phi\rangle\big|^p \le \int \|x' - y'\|_{n,p}^p \pi_{\boldsymbol{x},\boldsymbol{y}}(dx', dy') \le c\|\boldsymbol{x} - \boldsymbol{y}\|_{n,p}^p,
$$

for all $\Phi \in \operatorname{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,p})$ and $\boldsymbol{x}, \boldsymbol{y} \in (\mathbb{R}^d)^n$. For our first concentration result, recall that $\|x\|_\infty = \sup_{s\in[0,T]} |x(s)|$, and that on $(\mathcal{C}^d)^n$ we make use of the corresponding ℓ^1 and ℓ^2 norms on the product space as in (5.1).

THEOREM 5.4. *Assume that the Lipschitz condition* (5.4) *holds with* $p = 2$. *Assume also that there exists* $\kappa_0 < \infty$ *such that*

$$
(5.8)\qquad \mathcal{W}_2(\widetilde{\mu}_0, \nu) \le \sqrt{2\kappa_0 \mathcal{R}(\nu|\widetilde{\mu}_0)}, \quad \text{for } \nu \ll \widetilde{\mu}_0.
$$

Then there exist a constant $\delta > 0$, *independent of* n, *such that for every* $a > 0$ *and every* $\Phi \in \operatorname{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,2})$ *we have*

$$
\mathbb{P}\big(\Phi(\widetilde{\boldsymbol{X}}) - \mathbb{E}\Phi(\widetilde{\boldsymbol{X}}) > a\big) \le 2e^{-\delta a^2}.
$$

PROOF. To apply Theorem 5.3, we first check that the constant κ in (5.3) does not grow with the dimension n. To this end, define $\beta_n : [0,T] \times (\mathbb{R}^d)^n \to (\mathbb{R}^d)^n$ by $\beta_n(t, \boldsymbol{x}) = (\widetilde{b}(t, x_1, m_{\boldsymbol{x}}^n), \dots, \widetilde{b}(t, x_n, m_{\boldsymbol{x}}^n))$. Define also the $nd \times nd$ volatility matrix Σ_n by

$$
\Sigma_n = \begin{pmatrix} \sigma & & & \\ & \sigma & & \\ & & \ddots & \\ & & & \sigma \end{pmatrix},
$$

with omitted entries understood to be zero. This way, we can write

$$
d\widetilde{\boldsymbol{X}}_t = \beta_n(t, \widetilde{\boldsymbol{X}}_t)\,dt + \Sigma_n\,d\boldsymbol{W}_t,
$$

where $\boldsymbol{W} = (B^1, \dots, B^n)$. We wish to show that $\beta_n(t, \cdot)$ is Lipschitz, uniformly in t and n, and that $\sup_n \|\Sigma_n\|_{\mathrm{op}} < \infty$. First, notice that for $\boldsymbol{x} = (x_1, \dots, x_n) \in (\mathbb{R}^d)^n$ and $\boldsymbol{y} = (y_1, \dots, y_n) \in (\mathbb{R}^d)^n$ we have for $t \in [0,T]$,

$$
\begin{aligned}
\big|\widetilde{b}(t, x_i, m_{\boldsymbol{x}}^n) - \widetilde{b}(t, y_i, m_{\boldsymbol{y}}^n)\big| &\le \widetilde{L}\big(|x_i - y_i| + \mathcal{W}_2(m_{\boldsymbol{x}}^n, m_{\boldsymbol{y}}^n)\big) \\
&\le \widetilde{L}\left(|x_i - y_i| + \sqrt{\frac{1}{n}\sum_{j=1}^n |x_j - y_j|^2}\right) \\
&= \widetilde{L}|x_i - y_i| + \widetilde{L}n^{-1/2}|\boldsymbol{x} - \boldsymbol{y}|,
\end{aligned}
$$

where $|\boldsymbol{x} - \boldsymbol{y}|$ as usual denotes the Euclidean distance. Hence,

$$
\begin{aligned}
\big|\beta_n(t, \boldsymbol{x}) - \beta_n(t, \boldsymbol{y})\big| &\le \sqrt{\sum_{i=1}^n \big(\widetilde{L}|x_i - y_i| + \widetilde{L}n^{-1/2}|\boldsymbol{x} - \boldsymbol{y}|\big)^2} \\
&\le 2\widetilde{L}|\boldsymbol{x} - \boldsymbol{y}|.
\end{aligned}
$$

This shows that the Lipschitz constant L of β_n is uniform in n. It is clear that $\|\Sigma_n\|_{\mathrm{op}} \le \|\sigma\|_{\mathrm{op}}$.

Now, for $\boldsymbol{x} \in (\mathbb{R}^d)^n$ recall that $\boldsymbol{x} \mapsto P_{\boldsymbol{x}}$ is a version of the conditional law of $\widetilde{\boldsymbol{X}}$ given $\widetilde{\boldsymbol{X}}_0 = \boldsymbol{x}$. By Theorem 5.3, there is a constant $\widetilde{c} > 0$, independent of n due to the above considerations, such that for any $\Phi \in \mathrm{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,2})$ we have

$$P_{\boldsymbol{x}}\big(\Phi - \langle P_{\boldsymbol{x}}, \Phi\rangle > a\big) \le \exp\big(-a^2/2\widetilde{c}\big), \quad \text{for all } a > 0.$$

Moreover, combining Theorem 5.2(ii) with Theorem 5.1, the assumption (5.8) ensures that for every $a > 0$ and $\varphi \in \mathrm{Lip}((\mathbb{R}^d)^n, \|\cdot\|_{n,2})$ we have

$$\widetilde{\mu}_0^n(\varphi - \langle\widetilde{\mu}_0^n, \varphi\rangle > a) \le \exp(-a^2/2\kappa_0).$$

Finally, fix any $\Phi \in \mathrm{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,2})$. Then by (5.7), the map $\boldsymbol{x} \mapsto \langle P_{\boldsymbol{x}}, \Phi\rangle$ is c-Lipschitz on $(\mathbb{R}^d)^n$ with respect to the Euclidean norm. Use this along with the previous two inequalities (together with the fact that $\widetilde{\mu}_0^n$ is the law of $\widetilde{\boldsymbol{X}}_0$) to conclude

$$\begin{aligned}\mathbb{P}\big(\Phi(\widetilde{\boldsymbol{X}}) - \mathbb{E}\Phi(\widetilde{\boldsymbol{X}}) > a\big) &\le \mathbb{E}\big[\mathbb{P}\big(\Phi(\widetilde{\boldsymbol{X}}) - \langle P_{\widetilde{\boldsymbol{X}}_0}, \Phi\rangle > a/2 | \widetilde{\boldsymbol{X}}_0\big)\big] \\ &\quad + \mathbb{P}\big(\langle P_{\widetilde{\boldsymbol{X}}_0}, \Phi\rangle - \mathbb{E}\langle P_{\widetilde{\boldsymbol{X}}_0}, \Phi\rangle > a/2\big) \\ &\le \exp\big(-a^2/8\widetilde{c}\big) + \exp\big(-a^2/8\kappa_0 c^2\big).\end{aligned}$$

The assertion of the theorem follows with $\delta = 1/(8\max\{\widetilde{c}, \kappa_0 c^2\})$. □

We now treat the case where $p = 1$ in (5.4) and $\widetilde{\mu}_0$ satisfies the much weaker assumption

$$\mathcal{W}_{1,\mathbb{R}^d}(\widetilde{\mu}_0, \nu) \le \sqrt{2\kappa_0 \mathcal{R}(\nu|\widetilde{\mu}_0)}, \quad \text{for } \nu \ll \widetilde{\mu}_0. \tag{5.9}$$

Adapting the proof of Theorem 5.4 yields the following.

THEOREM 5.5. *Assume that the Lipschitz condition* (5.4) *holds with* $p = 1$. *Assume also that* (5.9) *holds for some* $\kappa_0 < \infty$. *Then there exist constants* $c, \delta > 0$, *independent of* n, *such that for every* $a > 0$ *and every* $\Phi \in \mathrm{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,1})$, *we have*

$$\mathbb{P}\big(\Phi(\widetilde{\boldsymbol{X}}) - \mathbb{E}\Phi(\widetilde{\boldsymbol{X}}) > a\big) \le 2\exp\big(-\delta a^2/n\big).$$

PROOF. We proceed as in the proof of Theorem 5.4. It follows from (5.9) and Theorem 5.2(i) that

$$\mathcal{W}_{1,((\mathbb{R}^d)^n, \|\cdot\|_{n,1})}\big(\widetilde{\mu}_0^n, \nu\big) \le \sqrt{2n\kappa_0 \mathcal{R}\big(\nu|\widetilde{\mu}_0^n\big)}, \quad \text{for } \nu \ll \widetilde{\mu}_0^n. \tag{5.10}$$

Thus, for any function $\varphi \in \mathrm{Lip}((\mathbb{R}^d)^n, \|\cdot\|_{n,1})$, Theorem 5.1 yields

$$\widetilde{\mu}_0^n(\varphi - \langle\widetilde{\mu}_0^n, \varphi\rangle > a) \le \exp\big(-a^2/2n\kappa_0\big). \tag{5.11}$$

Fix $\Phi \in \mathrm{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,1})$, and note that Φ is $\sqrt{n}$-Lipschitz with respect to $\|\cdot\|_{n,2}$ because of the elementary inequality $\|\cdot\|_{n,1} \le \sqrt{n}\|\cdot\|_{n,2}$. Recall that $(\mathbb{R}^d)^n \ni \boldsymbol{x} \mapsto P_{\boldsymbol{x}}$ is a version of the conditional law of $\boldsymbol{X}$ given $\boldsymbol{X}_0$. By Theorem 5.3, there is a constant $\widetilde{c} > 0$, independent of n and Φ (as argued in the proof of Theorem 5.4), such that

$$P_{\boldsymbol{x}}\big(\Phi - \langle P_{\boldsymbol{x}}, \Phi\rangle > a\big) \le \exp\big(-a^2/2\widetilde{c}n\big), \quad \text{for all } a > 0. \tag{5.12}$$

Moreover, the map $\boldsymbol{x} \mapsto \langle P_{\boldsymbol{x}}, \Phi\rangle$ is c-Lipschitz on $((\mathbb{R}^d)^n, \|\cdot\|_{n,1})$ due to (5.7). Use (5.11) along with (5.12) to get

$$\begin{aligned}\mathbb{P}\big(\Phi(\widetilde{\boldsymbol{X}}) - \mathbb{E}\Phi(\widetilde{\boldsymbol{X}}) > a\big) &\le \mathbb{E}\big[\mathbb{P}\big(\Phi(\widetilde{\boldsymbol{X}}) - \langle P_{\widetilde{\boldsymbol{X}}_0}, \Phi\rangle > a/2 | \widetilde{\boldsymbol{X}}_0\big)\big] \\ &\quad + \mathbb{P}\big(\langle P_{\widetilde{\boldsymbol{X}}_0}, \Phi\rangle - \mathbb{E}\langle P_{\widetilde{\boldsymbol{X}}_0}, \Phi\rangle > a/2\big) \\ &\le \exp\big(-a^2/8n\widetilde{c}\big) + \exp\big(-a^2/8n\kappa_0 c^2\big).\end{aligned}$$

The assertion of the theorem follows with $\delta = 1/(8\max\{\widetilde{c}, \kappa_0 c^2\})$. □

5.3. *McKean–Vlasov expectation bounds*. The results of the previous subsection (the notation of which we keep here) pertain to the concentration of a function $\Phi(\widetilde{\boldsymbol{X}})$ around its mean but tell us nothing about the size of $\mathbb{E}\Phi(\widetilde{\boldsymbol{X}})$. In this section, we study the rate of convergence of $(m^n_{\widetilde{\boldsymbol{X}}_t})_{t\in[0,T]}$ to its limit $(\widetilde{\mu}_t)_{t\in[0,T]}$, defined through the McKean–Vlasov SDE

$$d\widetilde{Y}^1_t = \widetilde{b}(t, \widetilde{Y}^1_t, \widetilde{\mu}_t)\,dt + \sigma\,dB^1_t, \quad \widetilde{Y}^1_0 = \widetilde{X}^1_0, \widetilde{\mu}_t = \mathrm{Law}(\widetilde{Y}^1_t).$$

The assumptions on $\widetilde{b}$ in Section 5.2 ensure the existence of a unique strong solution $(\widetilde{Y}^1, \widetilde{\mu})$ to this equation (see, e.g., Section 7 of [15], or Section 2.1 in Chapter 2 of [13]). We next provide some quantitative bounds on $\mathbb{E}[\mathcal{W}^p_{p,\mathbb{R}^d}(m^n_{\widetilde{\boldsymbol{X}}_t}, \widetilde{\mu}_t)]$ for fixed t as well as a uniform bound, $\mathbb{E}[\sup_{t\in[0,T]} \mathcal{W}^p_{p,\mathbb{R}^d}(m^n_{\widetilde{\boldsymbol{X}}_t}, \widetilde{\mu}_t)]$. The results are essentially known but are provided for the sake of completeness.

THEOREM 5.6. *Fix $n \in \mathbb{N}$, and assume (5.4) holds for some $p \in [1,2]$. Recall the definition of $r_{n,p}$ from (3.1). If $\mathbb{E}[|X^1_0|^{2p+\delta}] < \infty$ for some $\delta > 0$, then there exists $C < \infty$ such that for each n and each $t \in [0,T]$ we have*

$$\mathbb{E}\big[\mathcal{W}^p_p(m^n_{\widetilde{\boldsymbol{X}}_t}, \widetilde{\mu}_t)\big] \le C r_{n,p}. \tag{5.13}$$

If $\mathbb{E}[|X^1_0|^{d+5}] < \infty$, then there exists $C < \infty$ such that for each n we have

$$\mathbb{E}\Big[\sup_{s\in[0,T]} \mathcal{W}^2_2(m^n_{\widetilde{\boldsymbol{X}}_s}, \widetilde{\mu}_s)\Big] \le C n^{-2/(d+8)}. \tag{5.14}$$

PROOF. The proof begins with a standard coupling argument. Construct i.i.d. copies of the unique solution $\widetilde{\boldsymbol{Y}}$ of the McKean–Vlasov equation, where $\widetilde{\boldsymbol{Y}} = (\widetilde{Y}^1, \dots, \widetilde{Y}^n)$, with

$$d\widetilde{Y}^i_t = \widetilde{b}(t, \widetilde{Y}^i_t, \widetilde{\mu}_t)\,dt + \sigma\,dB^i_t, \quad \widetilde{Y}^i_0 = \widetilde{X}^i_0, \widetilde{\mu}_t = \mathrm{Law}(\widetilde{Y}^i_t).$$

Together with (5.6), this implies

$$\begin{aligned}|\widetilde{X}^i_t - \widetilde{Y}^i_t| &\le \int_0^t \big|\widetilde{b}(s, \widetilde{X}^i_s, m^n_{\widetilde{\boldsymbol{X}}_s}) - \widetilde{b}(s, \widetilde{Y}^i_s, \widetilde{\mu}_s)\big|\,ds \\ &\le \widetilde{L}\int_0^t \big(|\widetilde{X}^i_s - \widetilde{Y}^i_s| + \mathcal{W}_p(m^n_{\widetilde{\boldsymbol{X}}_s}, \widetilde{\mu}_s)\big)\,ds.\end{aligned}$$

By Gronwall's inequality, we have

$$|\widetilde{X}^i_t - \widetilde{Y}^i_t| \le C\int_0^t \mathcal{W}_p(m^n_{\widetilde{\boldsymbol{X}}_s}, \widetilde{\mu}_s)\,ds.$$

Taking the power to the p and averaging the left-hand side of the last inequality over $i = 1, \dots, n$, we get

$$\mathcal{W}^p_p(m^n_{\widetilde{\boldsymbol{X}}_t}, m^n_{\widetilde{\boldsymbol{Y}}_t}) \le \frac{1}{n}\sum_{i=1}^n |\widetilde{X}^i_t - \widetilde{Y}^i_t|^p \le C\int_0^t \mathcal{W}^p_p(m^n_{\widetilde{\boldsymbol{X}}_s}, \widetilde{\mu}_s)\,ds.$$

Use the triangle inequality and Gronwall's inequality once more to obtain

$$\mathcal{W}^p_p(m^n_{\widetilde{\boldsymbol{X}}_t}, m^n_{\widetilde{\boldsymbol{Y}}_t}) \le C\int_0^t \mathcal{W}^p_p(m^n_{\widetilde{\boldsymbol{Y}}_s}, \widetilde{\mu}_s)\,ds.$$

Using again the triangle inequality, we have

$$\mathcal{W}^p_p(m^n_{\widetilde{\boldsymbol{X}}_t}, \widetilde{\mu}_t) \le C\mathcal{W}^p_p(m^n_{\widetilde{\boldsymbol{Y}}_t}, \widetilde{\mu}_t) + C\int_0^t \mathcal{W}^p_p(m^n_{\widetilde{\boldsymbol{Y}}_s}, \widetilde{\mu}_s)\,ds. \tag{5.15}$$

Now, (5.14) fits exactly Theorem 1.3 of [29]. To prove (5.13), it suffices to show that

$$\mathbb{E}[\mathcal{W}_p^p(m^n_{\widetilde{\boldsymbol{Y}}_t}, \widetilde{\mu}_t)] \le C r_{n,p}. \tag{5.16}$$

To this end, note that $\widetilde{Y}^i_t$ are i.i.d. with law $\widetilde{\mu}_t$. Hence, by Theorem 1 of [25],

$$\mathbb{E}[\mathcal{W}_p^p(m^n_{\widetilde{\boldsymbol{Y}}_t}, \widetilde{\mu}_t)] \le C r_{n,p} \mathbb{E}[|\widetilde{Y}^1_t|^{2p+\delta}]^{p/(2p+\delta)},$$

where C depends only on p, δ and d. Finally, it suffices to note that standard estimates yield

$$\sup_{t\in[0,T]} \mathbb{E}[|\widetilde{Y}^1_t|^{2p+\delta}] \le C(1 + \mathbb{E}[|\widetilde{Y}^1_0|^{2p+\delta}]) < \infty. \qquad \square$$

These estimates allow us to now provide a proof of the law of large numbers for the MFG system, stated in Theorem 3.1.

PROOF OF THEOREM 3.1. The first claim is proved in Theorem 3.1 of [20]. To prove the other two claims, note first that (4.6) implies

$$\mathbb{E}[W^2_{2,\mathcal{C}^d}(m^n_{\boldsymbol{X}}, m^n_{\overline{\boldsymbol{X}}})] \le \frac{C}{n^2}, \tag{5.17}$$

with $\boldsymbol{X}$ as in (2.7) and $\overline{\boldsymbol{X}}$ as in (4.1). We now simply simply use (5.17) along with the rates of convergence for the McKean–Vlasov empirical measures $m^n_{\overline{\boldsymbol{X}}}$, which were just identified in Theorem 5.6. $\square$

5.4. *Proofs of Theorems* 3.2, 3.4 *and* 3.9. Using the developments of Section 5.2, we are now ready to prove the main results on concentration for the MFG system.

PROOF OF THEOREM 3.2. Note that for $\Phi \in \mathrm{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,1})$ we have

$$\begin{aligned}\mathbb{P}(\Phi(\boldsymbol{X}) - \mathbb{E}\Phi(\boldsymbol{X}) > a) \le{}& \mathbb{P}\left(\Phi(\boldsymbol{X}) - \Phi(\overline{\boldsymbol{X}}) > \frac{a}{3}\right)\\ &+ \mathbb{P}\left(\Phi(\overline{\boldsymbol{X}}) - \mathbb{E}\Phi(\overline{\boldsymbol{X}}) > \frac{a}{3}\right)\\ &+ \mathbb{P}\left(\mathbb{E}\Phi(\overline{\boldsymbol{X}}) - \mathbb{E}\Phi(\boldsymbol{X}) > \frac{a}{3}\right),\end{aligned} \tag{5.18}$$

with $\boldsymbol{X}$ as in (2.7) and $\overline{\boldsymbol{X}}$ as in (4.1). Comparing (4.1) with $\sigma = 0$ to (5.6) with $\tilde{b}$ defined as in (3.11), and noting that when $p^* = 1$, Assumptions A.1–A.5 ensure that $\tilde{b}$ satisfies condition (5.4), the result of Theorem 5.5 can be applied to bound the second term by $2\exp(-\delta a^2/n)$. The third term vanishes for $a \ge 3\sqrt{C}$, with C as in Theorem 4.1, because by (4.6) therein and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}\mathbb{E}\Phi(\overline{\boldsymbol{X}}) - \mathbb{E}\Phi(\boldsymbol{X}) &\le \mathbb{E}\sum_{i=1}^n \|X^i - \overline{X}^i\|_\infty\\ &\le n^{1/2}\left(\mathbb{E}\sum_{i=1}^n \|X^i - \overline{X}^i\|^2_\infty\right)^{1/2} \le \sqrt{C}.\end{aligned}$$

Finally, using Theorem 4.3 with $\epsilon = a/3n$, we know there exist $\kappa_1 < \infty$, $\kappa_2 > 0$ such that for $a \ge \kappa_1$,

$$\mathbb{P}\left(\Phi(\boldsymbol{X}) - \Phi(\overline{\boldsymbol{X}}) > \frac{a}{3}\right) \le \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n \|X^i - \overline{X}^i\|_\infty > \frac{a}{3n}\right)$$

$$
\begin{aligned}
&\le \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}\|X^i - \overline{X}^i\|_\infty^2 > \frac{a^2}{9n^2}\right) \\
&\le 2n\exp\left(-\frac{a^2}{9\kappa_2}\right).
\end{aligned}
$$

Combining the above results, we find that for a suitable δ (smaller than the above, if necessary), and a sufficiently large, we have for $n \ge 2$

$$
\mathbb{P}\big(\Phi(\boldsymbol{X}) - \mathbb{E}\Phi(\boldsymbol{X}) > a\big) \le 3n\exp\left(-\frac{\delta a^2}{n}\right). \qquad \square
$$

PROOF OF THEOREM 3.4. Fix $\Phi \in \mathrm{Lip}((\mathcal{C}^d)^n, \|\cdot\|_{n,2})$. We start with the same inequality (5.18) as in the previous proof. Comparing (4.1) with $\sigma_0 = 0$ to (5.6) with $\tilde{b}$ defined as in (3.11), and noting that Assumptions A.1–A.5 ensure that $\tilde{b}$ satisfies condition (5.4), the result of Theorem 5.4 can be applied to bound the second term therein by $2\exp(-\delta a^2)$. The third term is zero for $n \ge 9C/a^2$, with C as in Theorem 4.1, because by (4.6) therein, and Jensen's inequality, we have

$$
\mathbb{E}\Phi(\overline{\boldsymbol{X}}) - \mathbb{E}\Phi(\boldsymbol{X}) \le \mathbb{E}\sqrt{\sum_{i=1}^{n}\|X^i - \overline{X}^i\|_\infty^2} \le \sqrt{\mathbb{E}\sum_{i=1}^{n}\|X^i - \overline{X}^i\|_\infty^2} \le \frac{\sqrt{C}}{\sqrt{n}}.
$$

Finally, use the Lipschitz continuity of Φ and Theorem 4.3 with $\epsilon = a/(3\sqrt{n})$ to get

$$
\begin{aligned}
\mathbb{P}\left(\Phi(\boldsymbol{X}) - \Phi(\overline{\boldsymbol{X}}) > \frac{a}{3}\right) &\le \mathbb{P}\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\|X^i - \overline{X}^i\|_\infty^2} > \frac{a}{3\sqrt{n}}\right) \\
&\le 2n\exp\left(-\frac{a^2 n}{9\kappa_2}\right).
\end{aligned}
$$

Letting $\delta_1 := 1/(9\kappa_2)$ and $\delta_2 := \delta$, we find for $n \ge 9C/a^2$:

$$
\mathbb{P}\big(\Phi(\boldsymbol{X}) - \mathbb{E}\Phi(\boldsymbol{X}) > a\big) \le 2n\exp\big(-\delta_1 a^2 n\big) + 2\exp\big(-\delta_2 a^2\big). \qquad \square
$$

REMARK 5.7. It is worth commenting on a natural idea for extending the arguments of this section to the case with common noise. For the McKean–Vlasov system $\overline{\boldsymbol{X}}$, one can bootstrap the arguments of Sections 5.2 and 5.3 by studying the shifted paths $\overline{X}_t^i - \sigma_0 W_t$. This line of reasoning leads to various *conditional* concentration estimates, for example, on expressions of the form

$$
\mathbb{P}\big(\Phi(\overline{\boldsymbol{X}}) - \mathbb{E}[\Phi(\overline{\boldsymbol{X}})|W] > \epsilon|W\big).
$$

However, we are unable to transfer such estimates to the Nash system $\boldsymbol{X}$, because our main estimate (Theorem 4.3) of the distance between the two systems $\boldsymbol{X}$ and $\overline{\boldsymbol{X}}$ does not appear to have a conditional analogue.

PROOF OF THEOREM 3.9. Define the controls

$$
\overline{\alpha}_t^{n,i} = \widehat{\alpha}\big(\overline{X}_t^i, m_{\overline{\boldsymbol{X}}_t}^n, D_x U\big(t, \overline{X}_t^i, m_{\overline{\boldsymbol{X}}_t}^n\big)\big).
$$

We will separately estimate $|\alpha^{n,i} - \overline{\alpha}^{n,i}|$ and then $|\overline{\alpha}^{n,i} - \beta^i|$. In the following, the constant $C < \infty$ can change from line but never depends on n or i. First, use the Lipschitz assumptions on $\widehat{\alpha}$ and $D_x U$ to get

$$\begin{aligned}
|\alpha_t^{n,i} - \overline{\alpha}_t^{n,i}| &\le C\big(|X_t^i - \overline{X}_t^i| + \mathcal{W}_{p^*,\mathbb{R}^d}(m_{\boldsymbol{X}_t}^n, m_{\overline{\boldsymbol{X}}_t}^n) \\
&\quad + |D_{x_i} v^{n,i}(t, \boldsymbol{X}_t) - D_x U(t, X_t^i, m_{\boldsymbol{X}_t}^n)| \\
&\quad + |D_x U(t, X_t^i, m_{\boldsymbol{X}_t}^n) - D_x U(t, \overline{X}_t^i, m_{\overline{\boldsymbol{X}}_t}^n)|\big) \\
&\le C\big(|X_t^i - \overline{X}_t^i| + \mathcal{W}_{p^*,\mathbb{R}^d}(m_{\boldsymbol{X}_t}^n, m_{\overline{\boldsymbol{X}}_t}^n) \\
&\quad + |D_{x_i} v^{n,i}(t, \boldsymbol{X}_t) - D_x U(t, X_t^i, m_{\boldsymbol{X}_t}^n)|\big).
\end{aligned}$$

Using (4.5) and (4.6) (recalling $1 \le p^* \le 2$), we conclude

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \int_0^T |\alpha_t^{n,i} - \overline{\alpha}_t^{n,i}|^2 \, dt\right] \le \frac{C}{n^2}. \tag{5.19}$$

Recall now the definitions of $u^{n,i}$ and M^i from the beginning of Section 4. Note next that

$$D_{x_i} u^{n,i}(t, \boldsymbol{x}) = D_x U(t, x_i, m_{\boldsymbol{x}}^n) + \frac{1}{n} D_m U(t, x_i, m_{\boldsymbol{x}}^n, x_i), \quad \boldsymbol{x} \in (\mathbb{R}^d)^n$$

(see equation (4.4) or Proposition 2.1 of [20]) and that $D_m U$ is bounded by assumption. Since σ is nondegenerate, we find

$$\begin{aligned}
&\int_0^T |D_{x_i} v^{n,i}(t, \boldsymbol{X}_t) - D_x U(t, X_t^i, m_{\boldsymbol{X}_t}^n)|^2 \, dt \\
&\quad \le \frac{C}{n^2} + 2\int_0^T |D_{x_i} v^{n,i}(t, \boldsymbol{X}_t) - D_{x_i} u^{n,i}(t, \boldsymbol{X}_t)|^2 \, dt \\
&\quad \le \frac{C}{n^2} + C[M^i]_T.
\end{aligned}$$

Thus, using inequality (4.13) from the proof of Theorem 4.3 (with $\gamma = n^2/c_1 c_3$), we find $\delta_1 > 0$ such that, for $a > 0$ and $n \ge C/a$,

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n \int_0^T |\alpha_t^{n,i} - \overline{\alpha}_t^{n,i}|^2 \, dt > a^2\right) \le 2n e^{-\delta_1 a^2 n^2}. \tag{5.20}$$

We next estimate $|\overline{\alpha}^{n,i} - \beta^i|$. Use again the Lipschitz assumptions on $\widehat{\alpha}$ and $D_x U$ to get

$$|\overline{\alpha}_t^{n,i} - \beta_t^i| \le C\big(|\overline{X}_t^i - \mathcal{X}_t^i| + \mathcal{W}_{p^*,\mathbb{R}^d}(m_{\overline{\boldsymbol{X}}_t}^n, \mu_t)\big).$$

A straightforward application of Gronwall's inequality, exactly as in the proof of Theorem 5.6, lets us bound this further by

$$C\int_0^t \mathcal{W}_{p^*,\mathbb{R}^d}(m_{\overline{\boldsymbol{X}}_s}^n, \mu_s) \, ds.$$

Integrate, average, take expectations, and apply Theorem 5.6 to get

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \int_0^T |\overline{\alpha}_t^{n,i} - \beta_t^i|^{p^*} \, dt\right] \le C r_{n,p^*}.$$

Combine this with (5.19), noting that $n^{-1} \le C r_{n,p^*}$, to complete the proof of (3.10). To prove the final claim, apply Corollary 3.3 or Corollary 3.5 (more precisely, the easier analogues for the uncontrolled dynamics $\overline{X}^i$) to find $\delta_2, \delta_3 > 0$ such that, for every $a > 0$ and every $n \geq C/\min\{a, a^{d+8}\}$,

$$\mathbb{P}\Big(\sup_{t\in[0,T]} \mathcal{W}_{p^*,\mathbb{R}^d}\big(m^n_{\overline{X}_t}, \mu_t\big) > a\Big) \leq \begin{cases} 3ne^{-\delta_2 a^2 n} & \text{in case (i)}, \\ 2ne^{-\delta_2 a^2 n^2} + 2e^{-\delta_3 a^2 n} & \text{in case (ii)}. \end{cases}$$

Combine this with (5.20) to complete the proof. □

6. Large deviations of the empirical measure. In this section, we prove a LDP for the sequence $(m^n_{\overline{X}})_{n\geq 1}$ regarded as a sequence of random variables with values in the space $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$, where $\mathcal{P}^1(\mathbb{R}^d)$ is equipped with the 1-Wasserstein distance and $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ is equipped with the resulting uniform topology. A key result is the following *exponential equivalence* of the sequences $(m^n_{\boldsymbol{X}_t})_{t\in[0,T]}$ and $(m^n_{\overline{\boldsymbol{X}}_t})_{t\in[0,T]}$, that is, the empirical measure flows associated with the n-player Nash equilibrium dynamics and the approximating n-particle system, respectively.

COROLLARY 6.1. *Suppose Assumptions* A *and either* B *or* B′ *hold, with* $p^* = 1$. *Then, for every* $\epsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\Big(\sup_{t\in[0,T]} \mathcal{W}_2\big(m^n_{\boldsymbol{X}_t}, m^n_{\overline{\boldsymbol{X}}_t}\big) > \epsilon\Big) = -\infty.$$

PROOF. This follows immediately from Theorem 4.3. □

6.1. *LDP for weakly interacting diffusions in the presence of common noise.* A simple and well-known result of large deviations theory is that if a sequence satisfies a LDP, then any exponentially equivalent sequence also satisfies a LDP with the same rate function (e.g., Theorem 4.2.13 of [21]). In particular, due to Corollary 6.1, to derive a LDP for the sequence $(m^n_{\boldsymbol{X}})_{n\geq 1}$ of empirical measure flows of the Nash equilibrium dynamics, it suffices to prove a LDP for the sequence $(m^n_{\overline{\boldsymbol{X}}})_{n\geq 1}$ of empirical measure flows of the approximating n-particle system of weakly interacting diffusions. While there exist several forms of LDPs for the empirical measures of McKean–Vlasov or weakly interacting diffusions [2, 9, 19], all of them are obtained in the absence of common noise (i.e., $\sigma_0 = 0$) and, strictly speaking, for time-independent coefficients and nonrandom initial states.

This prompts us to revisit the aforementioned results and to first establish an LDP for the sequence of empirical measures of a general n-particle system of weakly interacting diffusions that has the following form:

$$d\widetilde{X}^i_t = \widetilde{b}\big(t, \widetilde{X}^i_t, m^n_{\widetilde{\boldsymbol{X}}_t}\big)\,dt + \sigma\,dB^i_t + \sigma_0\,dW_t, \tag{6.1}$$

with some initial condition $\widetilde{X}^i_0$, where $\sigma \in \mathbb{R}^{d\times d}$, $\sigma_0 \in \mathbb{R}^{d\times d_0}$, B and W are independent Brownian motions as specified in Section 2.3, the families $(\widetilde{X}^i_0)_{i\geq 1}$ and $((B^i)_{i\geq 1}, W)$ are all independent, and the drift $\widetilde{b}$ maps $[0,T]\times\mathbb{R}^d\times\mathcal{P}^1(\mathbb{R}^d)$ to $\mathbb{R}^d$. As usual, we denote $\widetilde{\boldsymbol{X}}_t = (\widetilde{X}^1_t, \ldots, \widetilde{X}^n_t)$. Observe that, except for the fact that $\sigma_0 \neq 0$, (6.1) is similar to (5.6).

REMARK 6.2. Note that with the particular choice

$$\widetilde{b}(t,x,m) = \widehat{b}\big(x, m, D_x U(t,x,m)\big), \quad t\in[0,T], x\in\mathbb{R}^d, m\in\mathcal{P}^1\big(\mathbb{R}^d\big),$$

the general n-particle system $\widetilde{\boldsymbol{X}}$ coincides with $\overline{\boldsymbol{X}}$, the n-particle approximation to the Nash equilibrium dynamics proposed in (4.1), which is the primary object of interest.

We impose the following conditions on the general n-particle system dynamics.

CONDITION 6.3. *The following conditions are satisfied*:

1. *The initial conditions* $(\widetilde{X}_0^i)_{i\geq 1}$ *are i.i.d. random variables with common law* μ_0 *and finite exponential moments of any order, namely*

$$\forall \lambda > 0, \quad \mathbb{E}\big[\exp(\lambda|\widetilde{X}_0^1|)\big] = \int_{\mathbb{R}^d} \exp(\lambda|y|)\mu_0(dy) < \infty. \tag{6.2}$$

2. *The drift function* $\widetilde{b} : [0, T] \times \mathbb{R}^d \times \mathcal{P}^1(\mathbb{R}^d) \to \mathbb{R}^d$ *is bounded, continuous and Lipschitz continuous in the last two arguments, uniformly in time.*

6.1.1. *Form of the rate function.* In this section, we use informal arguments to conjecture the form of the rate function for $(m^n_{\widetilde{\boldsymbol{X}}})_{n\geq 1}$ (see Theorem 6.6 and Corollary 6.7 below), and then show in the subsequent section that $(m^n_{\widetilde{\boldsymbol{X}}})_{n\geq 1}$ does indeed satisfy a LDP with this rate function.

The general strategy to allow σ_0 to be nonzero entails first freezing the common noise. Indeed, by the standard support theorem for trajectories of Brownian motion (see, e.g., Lemma 3.1 of [36]), the path of W lives with positive probability in any open ball of the path space $\mathcal{C}_0^d := \{\phi \in \mathcal{C}^d : \phi_0 = 0\}$. Then, for any ϕ in the Cameron–Martin space $\mathcal{H}_0^1([0, T]; \mathbb{R}^d)$, let $(\widetilde{\boldsymbol{X}}_t^\phi = (\widetilde{X}_t^{1,\phi}, \dots, \widetilde{X}_t^{n,\phi}))_{t\in[0,T]}$ denote the unique strong solution to the SDE

$$d\widetilde{X}_t^{i,\phi} = \widetilde{b}\big(t, \widetilde{X}_t^{i,\phi}, m^n_{\widetilde{\boldsymbol{X}}_t^\phi}\big)\, dt + \sigma\, dB_t^i + \dot{\phi}_t\, dt, \tag{6.3}$$

with $\widetilde{X}_0^{i,\phi} = \widetilde{X}_0^i$ as initial condition. Here, recall that $\mathcal{H}_0^1([0, T]; \mathbb{R}^d) = \{\phi \in \mathcal{H}^1([0, T]; \mathbb{R}^d) : \phi_0 = 0\}$, where $\mathcal{H}^1([0, T]; \mathbb{R}^d)$ is the Hilbert space of $\mathbb{R}^d$-valued absolutely continuous functions ϕ on $[0, T]$ whose weak derivative $\dot{\phi}$ is also square integrable on $[0, T]$, equipped with the norm $\|\phi\|_{\mathcal{H}^1} = (\int_0^T |\phi(t)|^2\, dt)^{1/2} + (\int_0^T |\dot{\phi}(t)|^2\, dt)^{1/2}$.

The dynamics in (6.3) fail to fall under the scope of [9] because $\widetilde{b}$ is not continuous with respect to the weak topology on $\mathcal{P}(\mathbb{R}^d)$. Moreover, while the results of [19] permit more general continuity assumptions, they do not quite cover our dynamics (6.3) because of the time-dependence in $\widetilde{b}$ and $\dot{\phi}$ and the randomness of the initial states. Nevertheless, we borrow the associated rate function obtained in [19].

Recall from Section 3.2 the notation for the seminorm $\|\cdot\|_m$ acting on Schwartz distributions, for $m \in \mathcal{P}^1(\mathbb{R}^d)$, as well as the definition of absolutely continuous (abs. cont. in abbreviated form) distribution-valued functions. Following the notation in [19], for each $\phi \in \mathcal{H}^1([0, T]; \mathbb{R}^d)$, we define the corresponding action functional $I^\phi : C([0, T] : \mathcal{P}^1(\mathbb{R}^d)) \to [0, \infty]$ by

$$I^\phi(\nu) := \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \big\|\dot{\nu_t} - \mathcal{L}^*_{t,\nu_t}\nu_t + \operatorname{div}(\nu_t\dot{\phi}_t)\big\|^2_{\nu_t}\, dt & \text{if } t \mapsto \nu_t \text{ is abs. cont.,} \\ \infty & \text{otherwise,} \end{cases} \tag{6.4}$$

where, for $(t, m) \in [0, T] \times \mathcal{P}^1(\mathbb{R}^d)$, $\mathcal{L}^*_{t,m}$ is the formal adjoint of the operator

$$\mathcal{L}_{t,m}h(x) = \frac{1}{2}\operatorname{Tr}\big[\sigma\sigma^\top D^2h(x)\big] + Dh(x)\cdot \widetilde{b}(t, x, m), \quad h \in C_c^\infty(\mathbb{R}^d). \tag{6.5}$$

Observe that the operator $\mathcal{L}^*_{t,\nu_t}(\cdot) - \operatorname{div}(\dot{\phi}_t\cdot)$ in (6.4) is the adjoint of $\mathcal{L}_{t,\nu_t}(\cdot) + \dot{\phi}_t \cdot D(\cdot)$. Below, we will often use the action functional I^0, given by $I^0 = I^\phi$ for $\phi \equiv 0$.

The functional I^ϕ admits several alternative representations. Lemma 6.4 presents one that will be used to extend the definition of I^ϕ to continuous ϕ. To present this representation, we first need to introduce some more notation. Let $(\tau_x : \mathbb{R}^d \ni z \mapsto z - x)_{x\in\mathbb{R}}$ denote the group of translations on $\mathbb{R}^d$. For $(t, m) \in [0, T] \times \mathcal{P}^1(\mathbb{R}^d)$ and a path $\phi \in \mathcal{C}_0^d$, define $\widetilde{\mathcal{L}}^*_{t,m}[\phi]$ to be the formal adjoint of the operator

$$\widetilde{\mathcal{L}}_{t,m}[\phi]h(x) = \frac{1}{2}\operatorname{Tr}\big[\sigma\sigma^\top D^2h(x)\big] \\ + Dh(x) \cdot \widetilde{b}\big(t, x + \phi_t, m \circ \tau^{-1}_{-\phi_t}\big), \quad h \in C_c^\infty(\mathbb{R}^d).$$

Finally, define the modified action functional $\widetilde{I}^\phi : C([0, T]; \mathcal{P}^1(\mathbb{R}^d)) \to [0,\infty]$ by

$$\widetilde{I}^\phi(\nu) := \begin{cases} \frac{1}{2}\int_0^T \big\|\dot{\nu}_t - \widetilde{\mathcal{L}}^*_{t,\nu_t}[\phi]\nu_t\big\|^2_{\nu_t}\,dt & \text{if } t \mapsto \nu_t \text{ is abs. cont.,} \\ \infty & \text{otherwise.} \end{cases} \tag{6.6}$$

In other words, this is the action functional corresponding to the drift $(t, x, m) \mapsto \widetilde{b}(t, x + \phi_t, m \circ \tau^{-1}_{-\phi_t})$.

We then have the following relationship between I^ϕ and $\widetilde{I}^\phi$.

LEMMA 6.4. *For $\phi \in \mathcal{H}^1_0([0, T]; \mathbb{R}^d)$,*

$$I^\phi(\nu) = \widetilde{I}^\phi\big((\nu_t \circ \tau^{-1}_{\phi_t})_{t\in[0,T]}\big). \tag{6.7}$$

The proof of Lemma 6.4 is deferred to Section 6.4. Its importance arises from the fact that it allows one to extend the definition of the actional functional $I^\phi(\cdot)$ to functions ϕ that are merely continuous. Indeed, note that, whenever $\phi \in \mathcal{C}^d$ and $\nu \in C([0, T]; \mathcal{P}^1(\mathbb{R}^d))$, the path $(\nu_t \circ \tau^{-1}_{\phi_t})_{0\le t\le T}$ is continuous due to the fact that

$$\mathcal{W}_1\big(\nu_t \circ \tau^{-1}_{\phi_t}, \nu_s \circ \tau^{-1}_{\phi_s}\big) \le |\phi_t - \phi_s| + \mathcal{W}_1(\nu_t, \nu_s), \quad s, t \in [0, T]. \tag{6.8}$$

This ensures that the cost $\widetilde{I}^\phi(\nu)$ is well-defined. So, in the rest of the presentation of our main results, we take the identity in (6.7) as the definition of the cost functional I^ϕ for just continuous ϕ with $\phi_0 = 0$. Observe that this extension is especially meaningful since $I^\phi(\nu)$ may be finite even when ϕ does not lie in the Cameron–Martin space $\mathcal{H}^1_0([0, T]; \mathbb{R}^d)$. For instance, if $b \equiv 0$ and $(\nu_t = \delta_{\phi_t})_{0\le t\le T}$ for some $\phi \in \mathcal{C}^d_0$, then we have $\nu_t \circ \tau^{-1}_\phi = \delta_0$ for all $t \in [0, T]$ and then $I^\phi(\nu) = 0$.

Roughly speaking, [19] asserts that whenever the common law of $(\widetilde{X}^i_0)_{i\ge1}$ reduces to a Dirac mass, $(m^n_{\widetilde{X}^\phi})_{n\ge1}$ satisfies a LDP with I^ϕ as rate function. Returning to (6.1), and denoting $\sigma_0\phi$ by the path $t \mapsto \sigma_0\phi_t$, this leads naturally to the conjecture that the collection $(m^n_{\widetilde{X}})_{n\ge1}$ should then satisfy a LDP with rate function

$$J^{\sigma_0}(\nu) := \inf_{\phi\in\mathcal{C}^d_0} I^{\sigma_0\phi}(\nu), \tag{6.9}$$

provided that $\nu \in C([0, T]; \mathcal{P}^1(\mathbb{R}^d))$ is such that ν_0 is equal to the common law of $(\widetilde{X}^i_0)_{i\ge1}$. The intuitive argument behind this assertion is that, by the standard support theorem for Brownian motion, the common noise $(\sigma_0 W_t)_{t\in[0,T]}$ lives with a positive probability in the neighborhood of $\sigma_0\phi$, for any ϕ in $\mathcal{C}^d_0$. In other words, the cost for $(\sigma_0 W_t)_{t\in[0,T]}$ to be in the neighborhood of ϕ is null; as a result, the minimal cost for $m^n_{\widetilde{X}}$ to be in the neighborhood of some $\nu \in C([0, T]; \mathcal{P}^1(\mathbb{R}^d))$ is the infimum of $I^{\sigma_0\phi}(\nu)$ over all ϕ in $\mathcal{C}^d_0$. Of course, when $\sigma_0 = 0$, $I^{\sigma_0\phi}(\nu)$ is independent of ϕ and J^0 coincides with I^0. Observe that, whenever $\sigma_0 \ne 0$, $J^{\sigma_0}(\nu)$ depends on σ_0 only through its image space $\operatorname{Im}(\sigma_0)$ This latter fact becomes apparent with the following explicit expression for $J^{\sigma_0}(\nu)$ in Proposition 6.5, when ν is smooth. First, define the *mean path* of a measure flow $\nu \in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ by

$$\mathbb{M}^\nu = \left(\mathbb{M}^\nu_t := \int_{\mathbb{R}^d} x\, d\nu_t(x)\right)_{t\in[0,T]} \in \mathcal{C}^d. \tag{6.10}$$

In the following, let $\Pi_{\sigma^{-1}\sigma_0} \in \mathbb{R}^{d\times d}$ denote the orthogonal projection onto the image of $\sigma^{-1}\sigma_0$.

PROPOSITION 6.5. *Let $\nu \in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ be such that its mean path $\mathbb{M}^\nu$ from* (6.10) *lies in $\mathcal{H}^1([0,T];\mathbb{R}^d)$. Then the functionals I^0 defined in* (6.4), *with $\phi = 0$, and J^{σ_0} defined in* (6.9), *satisfy*

$$J^{\sigma_0}(\nu) = I^0(\nu) - \frac{1}{2}\int_0^T \bigl|\Pi_{\sigma^{-1}\sigma_0}\sigma^{-1}\bigl(\dot{\mathbb{M}}^\nu_t - \langle \nu_t, \widetilde{b}(t,\cdot,\nu_t)\rangle\bigr)\bigr|^2\, dt.$$

The proof of Proposition 6.5 is relegated to Section 6.6. In the general case, when the mean path is not necessarily absolutely continuous, we have another expression for J^{σ_0}, based on the same factorization as in Lemma 6.4. This may be regarded as our main statement on the form of the rate function. See the discussion following Theorem 3.11 for intuition regarding this form of the rate function.

THEOREM 6.6. *Take $\nu \in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ and with $\mathbb{M}^\nu$ as in* (6.10), *let*

$$\mathbb{M}^{\widetilde{b},\nu}_t := \sigma\Pi_{\sigma^{-1}\sigma_0}\sigma^{-1}\left(\mathbb{M}^\nu_t - \mathbb{M}^\nu_0 - \int_0^t \langle \nu_s, \widetilde{b}(s,\cdot,\nu_s)\rangle\, ds\right), \quad \textit{for } t\in[0,T].$$

Then J^{σ_0} in (6.9) *satisfies*

$$J^{\sigma_0}(\nu) = \begin{cases} \widetilde{I}^{\mathbb{M}^{\widetilde{b},\nu}}\bigl((\nu_t \circ \tau^{-1}_{\mathbb{M}^{\widetilde{b},\nu}_t})_{t\in[0,T]}\bigr) & \textit{if } \sigma_0 \neq 0, \\ I^0(\nu), & \textit{if } \sigma_0 = 0, \end{cases}$$

where I^0 and $\widetilde{I}^\phi$ are defined in (6.4) *and* (6.6), *respectively.*

The proof of Theorem 6.6 is given in Section 6.6. As this proof shows, the above expression may be restated in terms of the mean constant path $(\nu_t \circ \tau^{-1}_{\mathbb{M}^\nu_t - \mathbb{M}^\nu_0})_{t\in[0,T]}$. (Observe that, if $\widetilde{X}_t$ is a random variable with law ν_t, then $\widetilde{X}_t - \mathbb{E}[\widetilde{X}_t]$ has distribution $\nu_t \circ \tau^{-1}_{\mathbb{M}^\nu_t}$, which justifies the terminology, "mean constant path".)

As a corollary, we obtain the following result, whose proof is also deferred to Section 6.6.

COROLLARY 6.7. *Take $\nu \in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ and $\sigma_0 \neq 0$. Then*

$$\begin{aligned} J^{\sigma_0}(\nu) = {} & \widetilde{I}^{-\mathbb{M}^\nu + \mathbb{M}^\nu_0}\bigl((\nu_t \circ \tau^{-1}_{\mathbb{M}^\nu_t - \mathbb{M}^\nu_0})_{t\in[0,T]}\bigr) \\ & - \frac{1}{2}\int_0^T \bigl|\Pi_{\sigma^{-1}\sigma_0}\sigma^{-1}\langle \nu_t, \widetilde{b}(t,\cdot,\nu_t)\rangle\bigr|^2\, dt. \end{aligned}$$

Observe that the first term on the right-hand side does not depend upon σ_0. This is in contrast with the second term, which attains its minimum when σ_0 is null and its maximum when σ_0 has full rank.

6.1.2. *The form of the LDP.* We now provide the form of the LDP. The conjectured form of the rate function of the previous subsection did not take into account the random initial states $(\widetilde{X}_0^i)_{i\geq 1}$, which we recall are i.i.d. with law μ_0. Sanov's theorem suggests the true rate function should take the form

$$C([0,T];\mathcal{P}^1(\mathbb{R}^d)) \ni \nu \mapsto \widetilde{J}^{\sigma_0,\mu_0}(\nu) := J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0), \tag{6.11}$$

where $\mathcal{R}$ denotes relative entropy, defined in (3.6), and J^{σ_0} is as defined in (6.9).

The precise large deviation principle for the sequence $(m^n_{\widetilde{X}})_{n\geq 1}$ takes the following form; its proof is given in Section 6.3.

THEOREM 6.8. *Under the stated assumptions, the sequence $(m^n_{\widetilde{X}})_{n\geq 1}$, as defined by* (6.1), *satisfies a weak large deviation principle in $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ with rate function $\widetilde{J}^{\sigma_0,\mu_0}$ defined in* (6.11). *That is, the following hold*:

(i) *For any open subset O of $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$,*

$$\liminf_{n\to\infty} \frac{1}{n}\log \mathbb{P}(m^n_{\widetilde{X}} \in O) \geq \inf_{\nu\in O} \widetilde{J}^{\sigma_0,\mu_0}(\nu).$$

(ii) *For any closed subset F of $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$,*

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(m^n_{\widetilde{X}} \in F) \leq -\lim_{\delta\searrow 0}\inf_{\nu\in F_\delta} \widetilde{J}^{\sigma_0,\mu_0}(\nu),$$

where $F_\delta = \{\nu \in C([0,T];\mathcal{P}^1(\mathbb{R}^d)) : \inf_{\widetilde{\nu}\in F}\sup_{t\in[0,T]}\mathcal{W}_1(\widetilde{\nu}_t,\nu_t)\leq \delta\}$.

REMARK 6.9. It is worth mentioning that J^{σ_0} and, therefore, $\widetilde{J}^{\sigma_0,\mu_0}$, is not a good rate function (i.e., does not have compact level sets) except when $\sigma_0 = 0$; see Proposition 6.10 below. When $\sigma_0 \neq 0$, we can easily see that the level set $\{J^{\sigma_0}\leq 0\} = \{J^{\sigma_0} = 0\}$ is not compact. This can be seen either from Theorem 6.6 or via a direct computation (but very much in the spirit of the statement of the theorem). Indeed, for any $\phi \in \mathcal{H}_0^1([0,T];\mathbb{R}^d)$, as in Section 3.2, we may call $\bar{X}^\phi$ the unique solution to the McKean–Vlasov equation

$$d\bar{X}_t^\phi = \widetilde{b}(t,\bar{X}_t^\phi,\mathcal{L}(\bar{X}_t^\phi))\,dt + \sigma\,dB_t^1 + \sigma_0\dot{\phi}_t\,dt, \quad t\in[0,T],$$

with $\bar{X}_0^\phi = \widetilde{X}_0^1$ as initial condition. Then the path $(\nu_t^\phi = \mathcal{L}(\bar{X}_t^\phi))_{t\in[0,T]}$ solves the Fokker–Planck equation (see [37])

$$\dot{\nu}_t^\phi - \mathcal{L}^*_{t,\nu_t^\phi}\nu_t^\phi + \operatorname{div}(\nu_t^\phi\sigma_0\dot{\phi}_t) = 0, \quad t\in[0,T],$$

in the distributional sense, with the initial condition $\nu_0^\phi = \mu_0$. Also, $I^{\sigma_0\phi}(\nu^\phi) + \mathcal{R}(\nu_0^\phi|\mu_0) = 0$; hence, $J^{\sigma_0}(\nu^\phi) + \mathcal{R}(\nu_0^\phi|\mu_0) = 0$. However, taking the mean in the McKean–Vlasov dynamics, we see that

$$\dot{\mathbb{M}}_t^{\nu^\phi} = \langle \nu_t^\phi, \widetilde{b}(t,\cdot,\nu_t^\phi)\rangle + \sigma_0\dot{\phi}_t, \quad t\in[0,T].$$

Since $\widetilde{b}$ is bounded and ϕ may be arbitrarily chosen in $\mathcal{H}_0^1([0,T];\mathbb{R}^d)$, we deduce that $\{\mathbb{M}^{\nu^\phi} : \phi \in \mathcal{H}_0^1([0,T];\mathbb{R}^d)\}$ is unbounded, and in particular it is not precompact in $\mathcal{C}^d$. This clearly implies that the set $\{\nu^\phi : \phi\in\mathcal{H}_0^1([0,T];\mathbb{R}^d)\}$, which is contained in $\{J^{\sigma_0}\leq 0\}$ by construction, is not precompact in $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$.

As explained in Remark 6.9, the lack of compactness of the level sets of J^{σ_0} explains the need for the additional limit over δ in (ii) in the statement of Theorem 6.8. Fortunately, there is no longer need for such a relaxation when F is compact.

PROPOSITION 6.10. *Assume that $\sigma_0 \neq 0$ and that K is a compact subset of $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$. Then*

$$\lim_{\delta \searrow 0} \inf_{\nu \in K_\delta} \big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)\big) = \inf_{\nu \in K} \big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)\big).$$

If $\sigma_0 = 0$, the above holds true for any closed (instead of compact) set $F \subset C([0,T];\mathcal{P}^1(\mathbb{R}^d))$. In the latter case, $(m^n_{\boldsymbol{X}})_{n\geq 1}$ satisfies a standard LDP with a good rate function.

Although the level sets of J^{σ_0} are not compact when $\sigma_0 \neq 0$, we have the following weaker version. The proofs of both Propositions 6.10 and 6.11 are given in Section 6.6.

PROPOSITION 6.11. *For any $\sigma_0 \neq 0$ and $a \geq 0$, there exists a compact subset $K \subset C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ and a constant $\kappa < \infty$ such that, for any ν in the level set*

$$\{\gamma \in C\big([0,T];\mathcal{P}^1\big(\mathbb{R}^d\big)\big) : J^{\sigma_0}(\gamma) + \mathcal{R}(\gamma_0|\mu_0) \leq a\},$$

the following hold:

(i) $(\nu_t \circ \tau^{-1}_{\mathbb{M}^\nu_t})_{t\in[0,T]} \in K$.

(ii) *For any $\phi \in \mathcal{C}^d_0$ satisfying $I^{\sigma_0\phi}(\nu) \leq a$, the path $(\mathbb{M}^\nu_t - \sigma_0\phi_t)_{t\in[0,T]}$ lies in $\mathcal{H}^1([0,T];\mathbb{R}^d)$ and has $\mathcal{H}^1$-norm is less than κ.*

Proposition 6.11 shows that the counterexample that we constructed prior to the statement of the proposition to prove the lack of compactness of the level sets of J^{σ_0} is somehow typical, as boundedness of the rate function forces the "centered" path $(\nu_t \circ \tau^{-1}_{\mathbb{M}^\nu_t})_{t\in[0,T]}$ to live in a compact subset.

REMARK 6.12. Instead of a LDP for the marginal empirical measures of the system (6.1), we could also provide a LDP for the empirical measure of the paths, as done in [9] and [23] for the case $\sigma_0 = 0$.

In fact, our proof of Theorem 6.8 shows that the rate function for the latter would take the following variational form:

$$\mathcal{J}^{\sigma_0}(\mathcal{M}) = \inf\{\mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) : \phi \in \mathcal{C}^d_0, \mathcal{Q} \in \mathcal{P}^1\big(\mathbb{R}^d \times \mathcal{C}^d_0\big), \Psi(\mathcal{Q},\phi) = \mathcal{M}\},$$

for $\mathcal{M} \in \mathcal{P}^1(\mathcal{C}^d)$, where $\mathbb{W}$ stands for the Wiener measure, and Ψ maps a pair $(\mathcal{Q},\phi)$ onto the law under $\mathcal{Q}$ of the solution $x = (x_t)_{t\in[0,T]}$ of the following McKean–Vlasov equation:

$$x_t = e + \int_0^t \widetilde{b}\big(s, x_s, \mathcal{Q}\circ x_s^{-1}\big)\,ds + \sigma w_t + \sigma_0\phi_t, \quad t \in [0,T],$$

where $(e, w = (w_t)_{t\in[0,T]})$ denotes the canonical process on the space $\mathbb{R}^d \times \mathcal{C}^d_0$.

When $\sigma_0 = 0$ and $\mathcal{Q}$ has first marginal μ_0, this formulation essentially reduces to the one obtained in [9] and [23]. We prefer to focus on the LDP for the flow $m^n_{\boldsymbol{X}}$ of marginal empirical measures instead of empirical measures on the path space, for the following reasons. First, its rate function has a more pleasant form, though this is hardly more than a matter of taste. Second, it is precisely this quantity that governs the interactions between the players.

6.1.3. *Proof of the large deviations principle without common noise.* We now obtain Theorem 3.10 as a simple corollary of the results established above.

PROOF OF THEOREM 3.10. Observe that, due to Theorem 6.6, the rate function $I(\nu) + \mathcal{R}(\nu_0|\mu_0)$ in the statement of Theorem 3.10 coincides with the rate function $\tilde{J}^{\sigma_0,\mu_0}$ defined in (6.11). Thus, Theorem 3.10 is an immediate consequence of Theorem 6.8, Proposition 6.10 and the fact that $\sigma_0 = 0$. $\square$

6.2. *LDP for the sequence* $(m^n_{\boldsymbol{X}})_{n\ge 1}$. In Sections 6.2.1 and 6.2.2, we establish the weak LDP without and with common noise, respectively.

6.2.1. *A weak LDP.* By combining Corollary 6.1 and Theorem 6.8, we end up with the following statement.

THEOREM 6.13. *Suppose Assumptions* A *and either Assumption* B *or* B′ *hold, and that the common distribution* μ_0 *of the i.i.d. initial states* $(X^i_0)_{i\ge 1}$ *of the solutions* $(\boldsymbol{X}^n)_{n\ge 1}$ *to the Nash equilibrium dynamics satisfy the exponential integrability condition* (6.2). *Then the sequence* $(m^n_{\boldsymbol{X}})_{n\ge 1}$ *satisfies* (*as in the statement of Theorem* 6.8) *a weak LDP with rate function* $\widetilde{J}^{\sigma_0,\mu_0}$ *defined in* (6.11), *provided the drift* $\widetilde{b}$ *in* (6.5) *satisfies*

$$\widetilde{b}(t,x,m) = \widehat{b}\big(x,m,D_x U(t,x,m)\big), \quad t\in[0,T], x\in\mathbb{R}^d, m\in\mathcal{P}^1\big(\mathbb{R}^d\big).$$

REMARK 6.14. Note that the rate function $\widetilde{J}^{\sigma_0,\mu_0}$ is defined in terms of the quantities J^σ, I^ϕ and $\mathcal{L}_{t,m}$ specified in (6.9), (6.4) and (6.5), and that the dependence of $\widetilde{J}^{\sigma_0,\mu_0}$ on the drift $\widetilde{b}$ is reflected in the definition (6.5) of the operator $\mathcal{L}_{t,m}$.

PROOF. We first note that, as already observed in Remark 6.2, with the definition of $\widetilde{b}$ given as above, $\widetilde{X}$ of (6.1) coincides with $\overline{X}$ of (4.1). The basic idea behind the proof is to apply Theorem 6.8 to immediately obtain a weak LDP for $m^n_{\widetilde{\boldsymbol{X}}} = m^n_{\overline{\boldsymbol{X}}}$, and then apply Corollary 6.1 to transfer the weak LDP to $m^n_{\boldsymbol{X}}$. The proof is fairly standard, except that some care is needed because the rate function does not have compact level sets.

We first prove the lower bound, that is, the analogue of (i) in the statement of Theorem 6.8, but for $(m^n_{\boldsymbol{X}})_{n\ge 1}$. Without any loss of generality, we can assume that $\inf_{\nu\in O}\widetilde{J}^{\sigma_0,\mu_0}(\nu)<\infty$, as otherwise the lower bound is trivial. Then, for any $\eta>0$, using (6.11), we can find $\nu^{(\eta)}\in O$ such that

$$\inf_{\nu\in O}\widetilde{J}^{\sigma_0,\mu_0}(\nu) \ge J^{\sigma_0}\big(\nu^{(\eta)}\big) + \mathcal{R}\big(\nu^{(\eta)}_0|\mu_0\big) - \eta.$$

Since O is open, we can find $\varepsilon>0$ such that the ball $B(\nu^{(\eta)},\varepsilon) := \{\nu\in C([0,T];\mathcal{P}^1(\mathbb{R}^d)) : \sup_{t\in[0,T]}\mathcal{W}_1(\nu_t,\nu^{(\eta)}_t)<\varepsilon\}$ is contained in O. By (i) of Theorem 6.8, and the identity $m^n_{\widetilde{\boldsymbol{X}}} = m^n_{\overline{\boldsymbol{X}}}$, we have

$$\begin{aligned}
\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\big(m^n_{\overline{\boldsymbol{X}}}\in B\big(\nu^{(\eta)},\varepsilon/2\big)\big) &\ge -\inf_{\nu\in B(\nu^{(\eta)},\varepsilon/2)}\widetilde{J}^{\sigma_0,\mu_0}(\nu) \\
&\ge -\big[J^{\sigma_0}\big(\nu^{(\eta)}\big)+\mathcal{R}\big(\nu^{(\eta)}_0|\mu_0\big)\big] \\
&\ge -\inf_{\nu\in O}\big(J^{\sigma_0}(\nu)+\mathcal{R}(\nu_0|\mu_0)\big)-\eta.
\end{aligned}$$

Since the right-hand side of the last inequality is finite, using Corollary 6.1, we then obtain

$$\begin{aligned}
&\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\big(m^n_{\boldsymbol{X}}\in O\big) \\
&\quad\ge \liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\big(m^n_{\boldsymbol{X}}\in B\big(\nu^{(\eta)},\varepsilon\big)\big) \\
&\quad\ge \liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\Big(m^n_{\overline{\boldsymbol{X}}}\in B\big(\nu^{(\eta)},\varepsilon/2\big), \sup_{t\in[0,T]}\mathcal{W}_1\big(m^n_{\overline{\boldsymbol{X}}_t},m^n_{\boldsymbol{X}_t}\big)<\varepsilon/2\Big) \\
&\quad\ge -\inf_{\nu\in O}\big(J^{\sigma_0}(\nu)+\mathcal{R}(\nu_0|\mu_0)\big)-\eta.
\end{aligned}$$

Letting η tend to 0, this proves the lower bound.

We now turn to the proof of the upper bound, namely the analog of (ii) in Theorem 6.8. We know that, for any $\varepsilon > 0$ and for any closed subset $F \in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(m^n_{\boldsymbol{X}} \in F\big) \\
&\quad\le \limsup_{n\to\infty} \frac{1}{n} \log\Big(\mathbb{P}\Big(m^n_{\boldsymbol{X}} \in F,\ \sup_{t\in[0,T]} \mathcal{W}_1\big(m^n_{\boldsymbol{X}_t}, m^n_{\overline{\boldsymbol{X}}_t}\big) \le \varepsilon\Big) \\
&\qquad + \mathbb{P}\Big(\sup_{t\in[0,T]} \mathcal{W}_1\big(m^n_{\boldsymbol{X}_t}, m^n_{\overline{\boldsymbol{X}}_t}\big) > \varepsilon\Big)\Big) \\
&\quad\le \limsup_{n\to\infty} \frac{1}{n} \log\Big(\mathbb{P}\big(m^n_{\overline{\boldsymbol{X}}} \in F_\varepsilon\big) + \mathbb{P}\Big(\sup_{t\in[0,T]} \mathcal{W}_1\big(m^n_{\boldsymbol{X}_t}, m^n_{\overline{\boldsymbol{X}}_t}\big) > \varepsilon\Big)\Big) \\
&\quad\le \max\Big[\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(m^n_{\overline{\boldsymbol{X}}} \in F_\varepsilon\big), \\
&\qquad \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\Big(\sup_{t\in[0,T]} \mathcal{W}_1\big(m^n_{\boldsymbol{X}_t}, m^n_{\overline{\boldsymbol{X}}_t}\big) > \varepsilon\Big)\Big].
\end{aligned}$$

By Corollary 6.1, the second argument in the maximum is $-\infty$. Hence,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(m^n_{\boldsymbol{X}} \in F\big) \le \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(m^n_{\overline{\boldsymbol{X}}} \in F_\varepsilon\big).$$

Since F_ε is closed, Theorem 6.8(ii) and the identity $m^n_{\widetilde{\boldsymbol{X}}} = m^n_{\overline{\boldsymbol{X}}}$ yield

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(m^n_{\boldsymbol{X}} \in F\big) \le \lim_{\delta\searrow 0} \inf_{\mu\in(F_\delta)_\varepsilon} \big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)\big).$$

Obviously, $(F_\delta)_\varepsilon \subset F_{\delta+\varepsilon}$, from which we get

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(m^n_{\boldsymbol{X}} \in F\big) \le \lim_{\delta\searrow 0} \inf_{\mu\in F_{\delta+\varepsilon}} \big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)\big).$$

Letting ε tend to 0, we obtain, as required,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(m^n_{\boldsymbol{X}} \in F\big) \le \lim_{\delta\searrow 0} \inf_{\nu\in F_\delta} \big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)\big).$$

This completes the proof. □

6.2.2. *Proof of the weak LDP in the presence of common noise*. We are now in a position to complete the proof of the weak LDP in the presence of common noise.

PROOF OF THEOREM 3.11. The result is an immediate consequence of Theorem 6.13 after observing that the rate function $\tilde{J}^{\sigma_0,\mu_0}(\nu)$ therein coincides with the the rate function $J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)$ given above, due to Theorem 6.6. □

6.3. *Proof of Theorem* 6.8. Our proof relies on the so-called contraction principle, which is somewhat similar to the approach developed in [9] and [23]. In particular, the strategy used in this section may be adapted to obtain an LDP for the empirical distribution of the paths of (6.1) (instead of the marginal empirical distributions), with the rate function having a variational representation; see Remark 6.12.

6.3.1. *Case when* $\tilde{b}=0$. The first step of the proof is to focus on the case when the drift $\widetilde{b}$ is trivial. Then we can have a look at the pair

$$(\bar{\mathcal{Q}}^n, W) = \left(\frac{1}{n}\sum_{i=1}^n \delta_{(\widetilde{X}_0^i, B^i)}, W\right), \tag{6.12}$$

which we regard as a random element with values in the product space:

$$\mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \times \mathcal{C}_0^d.$$

As above, $\mathcal{C}_0^d$ is equipped throughout the paragraph with the uniform topology and $\mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d)$ is equipped with the corresponding 1-Wasserstein distance. Also, for a probability measure $\mathcal{Q}$ on $\mathbb{R}^d \times \mathcal{C}_0^d$, we denote by $\mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W})$ the relative entropy with respect to $\mu_0 \times \mathbb{W}$, where $\mathbb{W}$ is the Wiener measure on the space $\mathcal{C}_0^d$. Then we have the following statement.

PROPOSITION 6.15. *The pair* $(\bar{\mathcal{Q}}^n, W)_{n\geq 1}$ *satisfies the following weak LDP*:

(i) *For any open subset* O *of* $\mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \times \mathcal{C}_0^d$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big((\bar{\mathcal{Q}}^n, W) \in O\big) \geq - \inf_{(\mathcal{Q},\phi)\in O} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W});$$

(ii) *For any closed subset* F *of* $\mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \times \mathcal{C}_0^d$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big((\bar{\mathcal{Q}}^n, W) \in F\big) \leq - \lim_{\delta\searrow 0} \inf_{(\mathcal{Q},\phi)\in F_\delta} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}),$$

where

$$\begin{aligned} F_\delta = \Big\{ &(\mathcal{Q}, \phi) \in \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \times \mathcal{C}_0^d \\ &: \inf_{(\mathcal{Q}',\phi')\in F} \big[\max\big(\mathcal{W}_1(\mathcal{Q}, \mathcal{Q}'), \|\phi - \phi'\|_\infty\big)\big] \leq \delta \Big\}. \end{aligned}$$

PROOF. We start with the proof of (i). First, observe that for any $\varepsilon > 0$, $\mathcal{Q} \in \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d)$ and $\phi \in \mathcal{C}_0^d$, the independence of $\bar{\mathcal{Q}}^n$ and W implies

$$\begin{aligned} &\log \mathbb{P}\big(\mathcal{W}_1(\bar{\mathcal{Q}}^n, \mathcal{Q}) < \varepsilon, \|W - \phi\|_\infty < \varepsilon\big) \\ &\quad = \log \mathbb{P}\big(\mathcal{W}_1(\bar{\mathcal{Q}}^n, \mathcal{Q}) < \varepsilon\big) + \log \mathbb{P}\big(\|W - \phi\|_\infty < \varepsilon\big). \end{aligned} \tag{6.13}$$

By the support theorem for the trajectories of a Brownian motion (see Lemma 3.1 in [36]),

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\|W - \phi\|_\infty < \varepsilon\big) = 0.$$

Also, on dividing the first term in the second line of (6.13) by n and taking the limit inferior, Sanov's theorem in the 1-Wasserstein topology (see, for instance, [39]) implies that

$$\begin{aligned} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\mathcal{W}_1(\bar{\mathcal{Q}}^n, \mathcal{Q}) < \varepsilon\big) &\geq - \inf_{\mathcal{Q}'\in\mathcal{P}^1(\mathbb{R}^d\times\mathcal{C}_0^d):\mathcal{W}_1(\mathcal{Q},\mathcal{Q}')<\varepsilon} \mathcal{R}(\mathcal{Q}'|\mu_0 \times \mathbb{W}) \\ &\geq -\mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}). \end{aligned}$$

Now, given an open set $O \subset \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \times \mathcal{C}_0^d$, and $\eta > 0$, choose $(\mathcal{Q}, \phi) \in O$ such that

$$\inf_{(\mathcal{Q}',\phi')\in O} \mathcal{R}\big(\mathcal{Q}'|\mu_0 \times \mathbb{W}\big) \geq \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) - \eta.$$

By choosing $\varepsilon > 0$ such that the set

$$\{(\mathcal{Q}', \phi') \in \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \times \mathcal{C}_0^d : \max(\mathcal{W}_1(\mathcal{Q}', \mathcal{Q}), \|\phi' - \phi\|_\infty) < \varepsilon\}$$

is contained in O, we get

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}((\bar{\mathcal{Q}}^n, W) \in O) &\geq -\mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) \\ &\geq - \inf_{(\mathcal{Q}',\phi')\in O} \mathcal{R}(\mathcal{Q}'|\mu_0 \times \mathbb{W}) - \eta.\end{aligned}$$

The proof of (i) follows on sending η to 0.

We now prove the upper bound (ii). Consider a closed set F in the product space $\mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \times \mathcal{C}_0^d$, and let

$$F' = \{\mathcal{Q} : \exists \phi \in \mathcal{C}_0^d, (\mathcal{Q}, \phi) \in F\},$$

which may not be closed. Then the LDP for the sequence $(\bar{\mathcal{Q}}^n)_{n\geq 1}$ yields

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}((\bar{\mathcal{Q}}^n, W) \in F) \leq - \inf_{\mathcal{Q}\in \mathrm{cl}(F')} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}),$$

where $\mathrm{cl}(F')$ is the closure of F'. In order to complete the proof, it suffices to note that, if $\mathcal{Q} \in \mathrm{cl}(F')$, then there exists a sequence $(\mathcal{Q}^n, \phi^n) \in F$ such that $\mathcal{W}_1(\mathcal{Q}, \mathcal{Q}^n) \to 0$. Hence, for any $\delta > 0$, we can choose n large enough such that $(\mathcal{Q}, \phi^n) \in F_\delta$. Therefore,

$$\inf_{\mathcal{Q}\in \mathrm{cl}(F')} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) \geq \inf_{(\mathcal{Q},\phi)\in F_\delta} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}),$$

which completes the proof. $\square$

6.3.2. *Contraction principle for nonzero drift.* We now consider the general case with an arbitrary drift $\widetilde{b}$ that satisfies Condition 6.3. Let e and $w = (w_t)_{t\in[0,T]}$ denote the canonical variables on $\mathbb{R}^d \times \mathcal{C}_0^d$, and for $(\mathcal{Q}, \phi) \in \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \times \mathcal{C}_0^d$ as above, consider the McKean–Vlasov equation:

$$x_t = e + \int_0^t \widetilde{b}(s, x_s, \mathcal{Q} \circ x_s^{-1})\, ds + \sigma w_t + \sigma_0 \phi_t, \quad t \in [0, T],$$

on the space $\mathbb{R}^d \times \mathcal{C}^d$ equipped with the probability measure $\mathcal{Q}$ on the Borel σ-field. Here, $\mathcal{Q} \circ x_s^{-1}$ stands for the law of x_s under $\mathcal{Q}$. Under Condition 6.3, the above equation has a unique solution x. Let Ψ be the mapping that takes $(\mathcal{Q}, \phi)$ to the probability measure $\mathcal{Q} \circ x^{-1}$ on $\mathcal{C}^d$, and let Φ be the mapping that takes $(\mathcal{Q}, \phi)$ to the flow of marginal measures $(\mathcal{Q} \circ x_t^{-1})_{t\in[0,T]}$. Note that then $\Psi(\mathcal{Q}, \phi)$ is an element of $\mathcal{P}^1(\mathcal{C}^d)$ and $\Phi(\mathcal{Q}, \phi)$ is an element of $C([0, T]; \mathcal{P}^1(\mathbb{R}^d))$, and we have the following useful relation for each n:

$$m_X^n = \Phi(\bar{\mathcal{Q}}^n, W). \tag{6.14}$$

It is easily verified that the mapping Φ is continuous. Actually, we prove a slightly stronger property.

LEMMA 6.16. *The mapping Φ is uniformly continuous from the space $\mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \times \mathcal{C}_0^d$ into $C([0, T]; \mathcal{P}^1(\mathbb{R}^d))$.*

PROOF. Consider two probability measures $\mathcal{Q}$ and $\mathcal{Q}'$ on $\mathbb{R}^d \times \mathcal{C}_0^d$ and two paths ϕ and ϕ' in $\mathcal{C}_0^d$ such that $\mathcal{W}_1(\mathcal{Q}, \mathcal{Q}') < \varepsilon$ and $\|\phi - \phi'\|_\infty < \varepsilon$, for some $\varepsilon > 0$. By definition of the

1-Wasserstein distance, we know that there exists a probability measure $\mathcal{M}$ on $(\mathbb{R}^d \times \mathcal{C}_0^d)^2$, with $\mathcal{Q}$ and $\mathcal{Q}'$ as marginal distributions, such that

$$\int_{(\mathbb{R}^d \times \mathcal{C}_0^d)^2} \max(|e-e'|, \|w-w'\|_\infty)\, d\mathcal{M}((e,w),(e',w')) < \varepsilon.$$

Denoting by (e,w) and (e',w') the canonical processes on $(\mathbb{R}^d \times \mathcal{C}_0^d)^2$, we consider the system of two equations:

$$x_t = e + \int_0^t \widetilde{b}(s, x_s, \mathcal{M} \circ x_s^{-1})\, ds + \sigma w_t + \sigma_0 \phi_t,$$
$$x'_t = e' + \int_0^t \widetilde{b}(s, x'_s, \mathcal{M} \circ (x'_s)^{-1})\, ds + \sigma w'_t + \sigma_0 \phi'_t, \quad t \in [0,T].$$

By Gronwall's lemma, there exists $C < \infty$ (possibly depending on σ and σ_0) such that for every $t \in [0,T]$,

$$\begin{aligned}&|x_t - x'_t| \\ &\quad \le C\biggl(|e-e'| + \|w-w'\|_\infty + \|\phi-\phi'\|_\infty + \int_0^t ds \int_{(\mathbb{R}^d \times \mathcal{C}_0^d)^2} |x_s - x'_s|\, d\mathcal{M}\biggr).\end{aligned}$$

Integrating with respect to $\mathcal{M}$, applying Gronwall's lemma once again and allowing the constant C to increase from line to line, we obtain

$$\int_{(\mathbb{R}^d \times \mathcal{C}_0^d)^2} |x_t - x'_t|\, d\mathcal{M} \le 3C\varepsilon, \quad t \in [0,T],$$

which implies

$$\sup_{t\in[0,T]} \mathcal{W}_1(\mathcal{M} \circ x_t^{-1}, \mathcal{M} \circ (x'_t)^{-1}) \le 3C\varepsilon.$$

It is clear that, for all $t \in [0,T]$, $\mathcal{M} \circ x_t^{-1} = [\Phi(\mathcal{Q},\phi)]_t$ and $\mathcal{M} \circ (x'_t)^{-1} = [\Phi(\mathcal{Q}',\phi')]_t$, from which we conclude that

$$\sup_{t\in[0,T]} \mathcal{W}_1([\Phi(\mathcal{Q},\phi)]_t, [\Phi(\mathcal{Q}',\phi')]_t) \le 3C\varepsilon,$$

which completes the proof. $\square$

6.3.3. *Proof of Theorem* 6.8. We can now make use of the contraction principle to prove Theorem 6.8. We start with the proof of the lower bound (i) in the statement of Theorem 6.8. For any open set O of $C([0,T]; \mathcal{P}^1(\mathbb{R}^d))$, the relation (6.14), the continuity property of Φ established in Lemma 6.16 and Proposition 6.15 yield

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(m_X^n \in O) \ge -\inf_{\phi \in \mathcal{C}_0^d} \inf_{\mathcal{Q} \in \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d)) : \Phi(\mathcal{Q},\phi) \in O} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}).$$

By Lemma 6.17 below, the right-hand side is equal to

$$-\inf_{\nu \in O} \inf_{\phi \in \mathcal{C}_0^d} (I^{\sigma_0\phi}(\nu) + \mathcal{R}(\nu_0|\mu_0)),$$

where recall that $I^{\cdot}$ is the functional defined in (6.7). This completes the proof of the lower bound.

We turn to the proof of the upper bound (ii). Similarly, for any closed set $F \subset C([0,T]; \mathcal{P}^1(\mathbb{R}^d))$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(m_X^n \in F) \le -\lim_{\delta \searrow 0} \inf_{(\mathcal{Q},\phi) \in (\Phi^{-1}(F))_\delta} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}).$$

By the uniform continuity of Φ (Lemma 6.16), for any $\eta > 0$, we can choose $\delta > 0$ small enough such that for any $(\mathcal{Q}, \phi) \in (\Phi^{-1}(F))_\delta$, $\Phi(\mathcal{Q}, \phi)$ belongs to F_η. Therefore,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(m^n_X \in F) \le -\lim_{\eta\searrow 0} \inf_{\Phi(\mathcal{Q},\phi)\in F_\eta} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}).$$

To complete the proof, apply Lemma 6.17 once again to conclude that

$$\begin{aligned}\inf_{(\mathcal{Q},\phi):\Phi(\mathcal{Q},\phi)\in F_\eta} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) &= \inf_{\nu\in F_\eta} \inf_{\phi\in\mathcal{C}^d} \inf_{\mathcal{Q}:\Phi(\mathcal{Q},\phi)=\nu} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) \\ &= \inf_{\nu\in F_\eta} \widetilde{J}^{\sigma_0,\mu_0}(\nu),\end{aligned}$$

which completes the proof.

6.4. *Proof of auxiliary lemmas.* We now prove the auxiliary Lemma 6.17 below. This relies on Lemma 6.4, which we first prove.

PROOF OF LEMMA 6.4. Fix $\nu \in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ and $\phi \in \mathcal{H}^1_0([0,T];\mathbb{R}^d)$. It is straightforward to check that $\nu = (\nu_t)_{t\in[0,T]}$ is absolutely continuous if and only if $\widetilde{\nu} := (\nu_t \circ \tau^{-1}_{\phi_t})_{t\in[0,T]}$ is. Now, suppose that ν is absolutely continuous, and let us compute the time-derivative of $\widetilde{\nu}$. For any test function $h \in C^\infty_c(\mathbb{R}^d)$ and $0 \le s < t \le T$, we have

$$\begin{aligned}\langle \widetilde{\nu}_t - \widetilde{\nu}_s, h\rangle &= \langle \nu_t, h(\cdot - \phi_t)\rangle - \langle \nu_s, h(\cdot - \phi_s)\rangle \\ &= \langle \nu_t - \nu_s, h(\cdot - \phi_s)\rangle + \langle \nu_t, h(\cdot - \phi_t) - h(\cdot - \phi_s)\rangle.\end{aligned}$$

Assume first that ϕ is continuously differentiable. Then, by the absolute continuity of $t \mapsto \nu_t$, the continuity of h and ϕ and the fact that h has compact support, we may divide by $t - s$ and then send $s \to t$ (for a fixed value of t) in the above to obtain

$$\frac{d}{dt}\langle \widetilde{\nu}_t, h\rangle = \langle \dot{\nu}_t, h(\cdot - \phi_t)\rangle - \langle \nu_t, \dot{\phi}_t \cdot Dh(\cdot - \phi_t)\rangle, \tag{6.15}$$

where the derivative $\dot{\phi}_t$ is understood in a (time-)distributional sense. By approximation, noting that $\mathcal{H}^1$-convergence implies sup-norm convergence, we can lift the restriction that ϕ is continuously differentiable and merely require that $\phi \in \mathcal{H}^1_0([0,T];\mathbb{R}^d)$.

Next, we claim that, for any $h \in C^\infty_c(\mathbb{R}^d)$,

$$\langle \widetilde{\mathcal{L}}^*_{t,\widetilde{\nu}_t}[\phi]\widetilde{\nu}_t, h(\cdot)\rangle = \langle \mathcal{L}^*_{t,\nu_t}\nu_t, h(\cdot - \phi_t)\rangle. \tag{6.16}$$

The proof is simple:

$$\begin{aligned}\langle \widetilde{\nu}_t, \widetilde{\mathcal{L}}_{t,\widetilde{\nu}_t}[\phi]h\rangle &= \left\langle \nu_t \circ \tau^{-1}_{\phi_t}, \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top D^2 h(\cdot)] + Dh(\cdot)\cdot \widetilde{b}(t, \cdot + \phi_t, \widetilde{\nu}_t \circ \tau^{-1}_{-\phi_t})\right\rangle \\ &= \left\langle \nu_t, \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top D^2 h(\cdot - \phi_t)] + Dh(\cdot - \phi_t)\cdot \widetilde{b}(t, \cdot, \nu_t)\right\rangle \\ &= \langle \nu_t, \mathcal{L}_{t,\nu_t} h(\cdot - \phi_t)\rangle.\end{aligned}$$

Combining (6.15) and (6.16), we may calculate, for $h \in C^\infty_c(\mathbb{R}^d)$,

$$\begin{aligned}&\langle \dot{\widetilde{\nu}}_t - \widetilde{\mathcal{L}}^*_{t,\widetilde{\nu}_t}[\phi]\widetilde{\nu}_t, h\rangle \\ &\quad= \frac{d}{dt}\langle \widetilde{\nu}_t, h\rangle - \langle \widetilde{\mathcal{L}}^*_{t,\widetilde{\nu}_t}[\phi]\widetilde{\nu}_t, h\rangle \\ &\quad= \langle \dot{\nu}_t, h(\cdot - \phi_t)\rangle - \langle \nu_t, \dot{\phi}_t \cdot Dh(\cdot - \phi_t)\rangle - \langle \mathcal{L}^*_{t,\nu_t}\nu_t, h(\cdot - \phi_t)\rangle.\end{aligned}$$

Hence, changing h into $h(\cdot + \phi_t)$ to pass from the second to the third line below, we get

$$\begin{aligned}
&\|\dot{\widetilde{\nu}}_t - \widetilde{\mathcal{L}}^*_{t,\widetilde{\nu}_t}[\phi]\widetilde{\nu}_t\|^2_{\widetilde{\nu}_t} \\
&\quad = \sup_{\substack{h \in C_c^\infty(\mathbb{R}^d): \\ \langle \widetilde{\nu}_t, |Dh|^2\rangle \neq 0}} \frac{\langle \dot{\widetilde{\nu}}_t - \widetilde{\mathcal{L}}^*_{t,\widetilde{\nu}_t}[\phi]\widetilde{\nu}_t, h\rangle^2}{\langle \widetilde{\nu}_t, |Dh|^2\rangle} \\
&\quad = \sup_{\substack{h \in C_c^\infty(\mathbb{R}^d) \\ \langle \nu_t, |Dh|^2\rangle \neq 0}} \frac{(\langle \dot{\nu}_t, h\rangle - \langle \nu_t, \dot{\phi}_t \cdot Dh\rangle - \langle \mathcal{L}^*_{t,\nu_t}\nu_t, h\rangle)^2}{\langle \nu_t, |Dh|^2\rangle} \\
&\quad = \big\|\dot{\nu}_t - \mathcal{L}^*_{t,\nu_t}\nu_t + \operatorname{div}(\dot{\phi}_t\nu_t)\big\|^2_{\nu_t}.
\end{aligned}$$

Comparing the definitions of I^ϕ and $\widetilde{I}^\phi$, the proof is complete. □

LEMMA 6.17. *For $\nu \in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ and $\phi \in \mathcal{C}_0^d$,*

$$\inf_{\mathcal{Q} \in \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d):\Phi(\mathcal{Q},\phi)=\nu} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) = \widetilde{I}^{\sigma_0\phi}\big((\nu_t \circ \tau^{-1}_{\sigma_0\phi_t})_{t\in[0,T]}\big) + \mathcal{R}(\nu_0|\mu_0).$$

Observe that the first term on the right-hand side in Lemma 6.17 coincides with $I^{\sigma_0\phi}((\nu_t)_{t\in[0,T]})$ when $\phi \in \mathcal{H}_0^1([0,T];\mathbb{R}^d)$; when $\phi \notin \mathcal{H}_0^1([0,T];\mathbb{R}^d)$, we called it $I^{\sigma_0\phi}((\nu_t)_{t\in[0,T]})$).

PROOF OF LEMMA 6.17. First, let (e, w) be the coordinate maps on $\mathbb{R}^d \times \mathcal{C}^d$, as before, and let $\Phi^* : \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d) \to C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ be the mapping that takes, for a frozen $\phi \in \mathcal{C}_0^d$, $\mathcal{Q}$ to the flow of marginal laws of the solution $(y_t)_{t\in[0,T]}$ of the McKean–Vlasov equation:

$$y_t = e + \int_0^t \widetilde{b}\big(s, y_s + \sigma_0\phi_s, \mathcal{Q} \circ (\tau_{-\sigma_0\phi_s} y_s)^{-1}\big)\,ds + \sigma w_t, \quad t \in [0,T].$$

We now claim that $\Phi(\mathcal{Q},\phi) = \nu$ if and only $\Phi^*(\mathcal{Q})_t = \nu_t \circ \tau^{-1}_{\sigma_0\phi_t}$ for all $t \in [0,T]$, which can be seen by performing the change of variables $(x_t = y_t + \sigma_0\phi_t)_{t\in[0,T]}$ where $(x_t)_{t\in[0,T]}$ solves

$$x_t = e + \int_0^t \widetilde{b}\big(s, x_s, \mathcal{Q} \circ x_s^{-1}\big)\,ds + \sigma w_t + \sigma_0\phi_t, \quad t \in [0,T].$$

Hence, since $\phi_0 = 0$, it suffices now to show that

$$\inf_{\mathcal{Q} \in \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d):\Phi^*(\mathcal{Q})=\nu} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) = \widetilde{I}^{\sigma_0\phi}(\nu) + \mathcal{R}(\nu_0|\mu_0). \tag{6.17}$$

We start from the left-hand side of (6.17), for a fixed $\mathcal{Q} \in \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d)$. By Theorem D.13 in [21],

$$\mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) = \mathcal{R}(q|\mu_0) + \int_{\mathbb{R}^d} \mathcal{R}\big(\mathcal{Q}^{x_0}|\mathbb{W}\big)\,dq(x_0), \tag{6.18}$$

with $q \in \mathcal{P}(\mathbb{R}^d)$ denoting the first marginal of $\mathcal{Q} \in \mathcal{P}(\mathbb{R}^d \times \mathcal{C}_0^d)$, and with $(\mathcal{Q}^{x_0})_{x_0\in\mathbb{R}^d}$ denoting a regular conditional probability distribution of the $\mathcal{C}^d$ coordinate given the $\mathbb{R}^d$ coordinate, under $\mathcal{Q}$. In particular, replacing μ_0 by q in (6.18), we see that the second term in the right-hand side identifies with $\mathcal{R}(\mathcal{Q}|q \times \mathbb{W})$.

Now, for $(e, w) \in \mathbb{R}^d \times \mathcal{C}_0^d$, let $\Xi(e, w) \in \mathcal{C}^d$ denote the solution y of the equation

$$y_t = e + \int_0^t \widetilde{b}\big(s, y_s + \sigma_0\phi_s, \Phi^*(\mathcal{Q})_t \circ \tau^{-1}_{-\sigma_0\phi_s}\big)\,ds + \sigma w_t, \quad t \in [0,T], \tag{6.19}$$

noting of course that $\mathcal{Q} \circ y_t^{-1} = \Phi^*(\mathcal{Q})_t$ for each $t \in [0, T]$, by construction. The nondegeneracy of σ (see Assumption A(2)) ensures that the map $\Xi(x_0, \cdot)$ is one-to-one from $\mathcal{C}_0^d$ to $\mathcal{C}^d$, for a fixed $x_0 \in \mathbb{R}^d$. Hence, by the contraction property for relative entropy,

$$\mathcal{R}(\mathcal{Q}^{x_0}|\mathbb{W}) = \mathcal{R}(\mathcal{Q}^{x_0} \circ \Xi(x_0, \cdot)^{-1}|\mathbb{W} \circ \Xi(x_0, \cdot)^{-1}).$$

By the Donsker–Varadhan formula (see, for instance, Lemma 6.2.13 in [21]), we have

$$\begin{aligned}\mathcal{R}(\mathcal{Q}^{x_0}|\mathbb{W}) = \sup_{F \in C_b(\mathcal{C}^d)} \Big[& \int_{\mathcal{C}^d} F(\Xi(x_0, \cdot))\, d\mathcal{Q}^{x_0} \\ & - \log\Big(\int_{\mathcal{C}^d} e^{F(\Xi(x_0, \cdot))}\, d\mathbb{W} \Big) \Big],\end{aligned} \tag{6.20}$$

where $C_b(\mathcal{C}^d)$ is the set of bounded continuous functions on $\mathcal{C}^d$. The above right-hand side is denoted by $L^{(1)}_{\delta_{x_0}}(\mathcal{Q}^{x_0} \circ \Xi(x_0, \cdot)^{-1})$ in [19]; see Lemma 4.6 therein. Using that same notation here, by (6.18), we end up with

$$\begin{aligned}\mathcal{R}(\mathcal{Q}|q \times \mathbb{W}) &= \int_{\mathbb{R}^d} \mathcal{R}(\mathcal{Q}^{x_0}|\mathbb{W})\, dq(x_0) \\ &= \int_{\mathbb{R}^d} L^{(1)}_{\delta_{x_0}}(\mathcal{Q}^{x_0} \circ \Xi(x_0, \cdot)^{-1})\, dq(x_0).\end{aligned} \tag{6.21}$$

Now, passing the integral inside the supremum in (6.20), we obtain

$$\begin{aligned}&\mathcal{R}(\mathcal{Q}|q \times \mathbb{W}) \\ &\quad \geq \sup_{F \in C_b(\mathcal{C}^d)} \int_{\mathbb{R}^d} dq(x_0) \Big[\int_{\mathcal{C}^d} F(\Xi(x_0, \cdot))\, d\mathcal{Q}^{x_0} \\ &\qquad - \log\Big(\int_{\mathcal{C}^d} e^{F(\Xi(x_0, \cdot))}\, d\mathbb{W} \Big) \Big] \\ &\quad = \sup_{F \in C_b(\mathcal{C}^d)} \Big[\int_{\mathbb{R}^d \times \mathcal{C}^d} F(\Xi(\cdot, \cdot))\, d\mathcal{Q} \\ &\qquad - \int_{\mathbb{R}^d} dq(x_0) \log\Big(\int_{\mathcal{C}^d} e^{F(\Xi(x_0, \cdot))}\, d\mathbb{W} \Big) \Big] \\ &\quad =: L^{(1)}_q(\mathcal{Q} \circ \Xi^{-1}),\end{aligned} \tag{6.22}$$

where the definition in the last line agrees with the notation in Lemma 4.6 of [19]. In fact, the converse inequality holds as well: Because Ξ is a one-to-one map of $\mathbb{R}^d \times \mathcal{C}_0^d$ to $\mathcal{C}^d$, we again use the contraction property of relative entropy to get

$$\begin{aligned}\mathcal{R}(\mathcal{Q}|q \times \mathbb{W}) &= \mathcal{R}(\mathcal{Q} \circ \Xi^{-1}|(q \times \mathbb{W}) \circ \Xi^{-1})) \\ &= \sup_{F \in C_b(\mathcal{C}^d)} \Big[\int_{\mathbb{R}^d \times \mathcal{C}^d} F \circ \Xi\, d\mathcal{Q} - \log\Big(\int_{\mathbb{R}^d \times \mathcal{C}^d} e^{F \circ \Xi}\, d(q \times \mathbb{W}) \Big) \Big].\end{aligned}$$

By Jensen's inequality and concavity of log, this is bounded above by the right-hand side of (6.22), which shows that $\mathcal{R}(\mathcal{Q}|q \times \mathbb{W}) = L^{(1)}_q(\mathcal{Q} \circ \Xi^{-1})$. Using this along with (6.21) in (6.18), we end up with

$$\mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) = \mathcal{R}(q|\mu_0) + L^{(1)}_q(\mathcal{Q} \circ \Xi^{-1}).$$

Recalling that q denotes the first marginal of $\mathcal{Q}$ and that $\Phi^*(\mathcal{Q})_0 = q$, we have

$$\inf_{\mathcal{Q}: \Phi^*(\mathcal{Q}) = \nu} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) = \inf_{\mathcal{Q}: \Phi^*(\mathcal{Q}) = \nu} [\mathcal{R}(\nu_0|\mu_0) + L^{(1)}_q(\mathcal{Q} \circ \Xi^{-1})].$$

Finally, return to (6.19) and observe that $(\mathcal{Q} \circ \Xi_t^{-1})_{t\in[0,T]}$ coincides with $\Phi^*(\mathcal{Q})$. Also, for any two probability measures ν_0 and P on $\mathbb{R}^d$ and $\mathcal{C}^d$, with ν_0 being the image of P by the mapping $(x_t)_{t\in[0,T]} \mapsto x_0$, there exists a unique $\mathcal{Q} \in \mathcal{P}(\mathbb{R}^d \times \mathcal{C}_0^d)$ such that $P = \mathcal{Q} \circ \Xi^{-1}$; if P is integrable then $\mathcal{Q}$ is also integrable. Because, $t \mapsto \phi_t$ is continuous, the drift $(t,x) \mapsto b(t, x + \sigma_0\phi_t, \Phi^*(\mathcal{Q})_t \circ \tau^{-1}_{-\sigma_0\phi_s})$ is nice enough that we may apply Lemma 4.6 from [19], as well as Section 4.5 therein to conclude

$$\inf_{\mathcal{Q}:\Phi^*(\mathcal{Q})=\nu} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}) = \mathcal{R}(\nu_0|\mu_0) + \widetilde{I}^{\sigma_0\phi}(\nu).$$

Importantly, to check the above equality, we can assume that $\mathcal{R}(\nu_0|\mu_0) < \infty$, in which case $\nu_0 \in \mathcal{P}^1(\mathbb{R}^d)$; hence, by (4.11) in [19], with $\nu = \nu_0$, it is straightforward to verify that the minimum of the right-hand side of (4.10) in [19], may be restricted to the P's that are integrable. By the previous argument, those P can be written in the form $\mathcal{Q} \circ \Xi^{-1}$, with $\mathcal{Q} \in \mathcal{P}^1(\mathbb{R}^d \times \mathcal{C}_0^d)$, which yields the above identity. $\square$

6.5. *Proofs of Propositions* 6.10 *and* 6.11. We start with the proof of Proposition 6.11.

PROOF. Take a path ν such that $J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0) \le a$. Then, modifying without any loss of generality the value of a, we can find $\phi \in \mathcal{C}_0^d$ such that $I^{\sigma_0\phi}(\nu) + \mathcal{R}(\nu_0|\mu_0) \le a$. By Lemma 6.4, we deduce that the path $(\widetilde{\nu}_t = \nu_t \circ \tau^{-1}_{\sigma_0\phi_t})_{t\in[0,T]}$ is absolutely continuous. Also, for any test function $h \in C_c^\infty(\mathbb{R}^d)$ such that $|D_x h|$ and $|D_x^2 h|$ are bounded by 2, we have

$$\int_0^T |\langle \dot{\widetilde{\nu}}_t, h\rangle|^2 \, dt \le C(a),$$

where $C(a)$ is a constant only depending on a and the uniform bounds on b, σ, and σ_0. We can easily find a sequence of functions $(h_p)_{p\ge1}$ in $C_c^\infty(\mathbb{R}^d)$ converging to the identity function, uniformly on compact subsets, and satisfying at the same time the two constraints $\|D_x h_p\|_\infty \le 2$ and $\|D_x^2 h_p\|_\infty \le 2$. Using the fact that $\widetilde{\nu} \in C([0,T];\mathcal{P}^1(\mathbb{R}^d))$, we have

$$\lim_{p\to\infty} \sup_{t\in[0,T]} |\langle \widetilde{\nu}_t, h_p\rangle - \mathbb{M}_t^{\widetilde{\nu}}| = 0.$$

Since the set $\{\psi \in \mathcal{H}^1([0,T];\mathbb{R}^d) : \|\psi\|_{\mathcal{H}^1} \le \sqrt{C(a)}\}$ is closed for the uniform topology, we deduce that $\mathbb{M}^{\widetilde{\nu}} = \mathbb{M}^\nu - \sigma_0\phi$ is in $\mathcal{H}^1([0,T];\mathbb{R}^d)$ and has $\mathcal{H}^1$-norm bounded by $\sqrt{C(a)}$. This proves claim (ii).

Also, from Lemma 6.4 we know that

$$\widetilde{I}^{\sigma_0\phi}(\widetilde{\nu}) + \mathcal{R}(\widetilde{\nu}_0|\mu_0) = I^{\sigma_0\phi}(\nu) + \mathcal{R}(\nu_0|\mu_0) \le a.$$

Returning to the definition (6.4) of the action functional and using the fact that $\widetilde{b}$ is bounded, we can find a new constant, still denoted by $C(a)$ (and depending only on the same quantities as above), such that

$$I^0_{(0)}(\widetilde{\nu}) + \mathcal{R}(\widetilde{\nu}_0|\mu_0) \le C(a),$$

where $I^0_{(0)}$ is the action functional I^0 in the case when $\widetilde{b} \equiv 0$ (i.e., when $\mathcal{L}_{t,m} = \frac{1}{2}\operatorname{Tr}[\sigma\sigma^\top D_x^2]$). By Lemma 6.17,

$$I^0_{(0)}(\widetilde{\nu}) + \mathcal{R}(\widetilde{\nu}_0|\mu_0) = \inf_{\mathcal{Q}:\Phi_{(0)}(\mathcal{Q},0)=\widetilde{\nu}} \mathcal{R}(\mathcal{Q}|\mu_0 \times \mathbb{W}),$$

where $\Phi_{(0)}$ is the map Φ in the case when $\widetilde{b} \equiv 0$. By Sanov's theorem for the 1-Wasserstein topology (see [39]), $\mathcal{R}$ is a good rate function on $\mathcal{P}^1(\mathcal{C}^d)$. Hence, by the contraction principle, the left-hand side forms a good rate function on $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$. We deduce that there exists a compact set $K \subset C([0,T];\mathcal{P}^1(\mathbb{R}^d))$, depending only on $a > 0$, such that $\widetilde{\nu} \in K$. Now, $\nu_t \circ \tau^{-1}_{\mathbb{M}^\nu_t} = \widetilde{\nu}_t \circ \tau^{-1}_{\mathbb{M}^\nu_t - \sigma_0\phi_t}$ for all t. Using (6.8) and modifying the definition of K, we easily deduce that $(\nu_t \circ \tau^{-1}_{\mathbb{M}^\nu_t})_{t\in[0,T]}$ is in K, which completes the proof of (i). $\square$

We turn to the proof of Proposition 6.10.

PROOF OF PROPOSITION 6.11. We start with the first claim. We observe that the quantity $\inf_{\nu\in K_\delta}(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0))$ is nondecreasing as δ decreases. In particular,

$$\lim_{\delta\searrow 0}\inf_{\nu\in K_\delta}\big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)\big) \le \inf_{\nu\in K}\big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)\big).$$

In order to prove the converse bound, we proceed as follows. By the above inequality, we can assume that the left-hand side is finite, as otherwise there is nothing to prove. Recall from Lemma 6.17 that

$$\inf_{\Phi(\mathcal{Q},\phi)\in K_\delta}\mathcal{R}(\mathcal{Q}|\mu_0\times\mathbb{W}) = \inf_{\nu\in K_\delta}\big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)\big). \tag{6.23}$$

Since the right-hand side is less than some $C > 0$ independent of δ, the left-hand side can be rewritten as

$$\inf\big\{\mathcal{R}(\mathcal{Q}|\mu_0\times\mathbb{W}) : (\mathcal{Q},\phi) \text{ s.t. } \Phi(\mathcal{Q},\phi)\in K_\delta, \mathcal{R}(\mathcal{Q}|\mu_0\times\mathbb{W}) \le C\big\}.$$

Consider now a sequence $(\mathcal{Q}^n,\phi^n)_{n\ge 1}$ in $\mathcal{P}^1(\mathbb{R}^d\times\mathcal{C}^d_0)\times\mathcal{C}^d_0$, with $\phi^n\in\mathcal{C}^d_0$ and $\mathcal{R}(\mathcal{Q}^n|\mu_0\times\mathbb{W}) \le C$, yielding a $1/n$-approximation of the infimum when $\delta = 1/n$. Let $\nu^n = \Phi(\mathcal{Q}^n,\phi^n)\in K_{1/n}$, and notice that $(\nu^n)_{n\ge 1}$ is precompact in $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$ by compactness of K. Proposition 6.11 ensures that $(\sigma_0\phi^n)_{n\ge 1}$ must too be precompact in $\mathcal{C}^d_0$, and thus without loss of generality we may assume $(\phi^n)_{n\ge 1}$ is precompact as well. Finally, because $\mathcal{R}(\cdot|\mu_0\times\mathbb{W})$ is a good rate function on $\mathcal{P}^1(\mathbb{R}^d\times\mathcal{C}^d_0)$ by [39], we deduce that $(\mathcal{Q}^n)_{n\ge 1}$ is precompact. Relabel the subsequence and assume that $(\mu^n,\mathcal{Q}^n,\phi^n)_{n\ge 1}$ converges to some $(\mu,\mathcal{Q},\phi)$. By the continuity of Φ (see Lemma 6.16), $\nu = \Phi(\mathcal{Q},\phi)\in K$. Hence, by the lower semicontinuity of relative entropy, we get

$$\mathcal{R}(\mathcal{Q}|\mu_0\times\mathbb{W}) \le \liminf_{n\to\infty}\mathcal{R}\big(\mathcal{Q}^n|\mu_0\times\mathbb{W}\big) = \lim_{\delta\searrow 0}\inf_{\Phi(\mathcal{Q},\phi)\in K_\delta}\mathcal{R}(\mathcal{Q}|\mu_0\times\mathbb{W}).$$

Lemma 6.17 implies that (6.23) holds also without the δ, that is,

$$\inf_{\Phi(\mathcal{Q},\phi)\in K}\mathcal{R}(\mathcal{Q}|\mu_0\times\mathbb{W}) = \inf_{\nu\in K}\big(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)\big),$$

and the proof of the first claim is complete.

It remains to prove the second claim. In the case when $\sigma_0 = 0$, the fact that $J^0(\cdot) + \mathcal{R}(\cdot_0|\mu_0)$ is a good rate function is a consequence of the proof of Proposition 6.11. Equivalently, we can invoke Lemma 6.17, which asserts that

$$J^0(\nu) + \mathcal{R}(\nu_0|\mu_0) = \inf_{\mathcal{Q}:\Phi(\mathcal{Q},0)=\nu}\mathcal{R}(\mathcal{Q}|\mu_0\times\mathbb{W}).$$

Since $\mathcal{R}$ is a good rate function on $\mathcal{P}^1(\mathcal{C}^d)$ and Φ is continuous, the left-hand side forms a good rate function on $C([0,T];\mathcal{P}^1(\mathbb{R}^d))$. So, whenever $(\inf_{\nu\in F_\delta}(J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0)))_{\delta>0}$ is bounded, we may restrict ν in a compact set, and the passage to the limit works exactly as before. $\square$

6.6. *Proofs of Proposition* 6.5, *Theorem* 6.6 *and Corollary* 6.7. We start with the proof of Proposition 6.5.

PROOF OF PROPOSITION 6.5. The proof relies on another formulation of the rate function $I^{\sigma_0\phi}$. Let $C_c^{1,2}([0,T] \times \mathbb{R}^d)$ denote the set of compactly supported functions ϕ on $[0,T] \times \mathbb{R}^d$ possessing one time derivative and two space derivatives. By [19], Lemma 4.8, we claim that for $\phi \in C_0^2([0,T]; \mathbb{R}^d)$

$$\begin{aligned} I^{\sigma_0\phi}(\nu) = \sup_{\psi \in C_c^{1,2}([0,T]\times\mathbb{R}^d)} \Big[& \langle \nu_T, \psi_T \rangle - \langle \nu_0, \psi_0 \rangle \\ & - \int_0^T \Big\langle \nu_t, (\partial_t + \mathcal{L}_{t,\nu_t})\psi_t + \sigma_0 \dot{\phi}_t \cdot D_x \psi_t + \frac{1}{2} |\sigma^\top D_x \psi_t|^2 \Big\rangle dt \Big], \end{aligned}$$

where we write $\psi_t(x) = \psi(t,x)$. Since $\nu \in \mathcal{P}^1(C([0,T]; \mathcal{P}^1(\mathbb{R}^d)))$, we can allow ψ in the supremum to be at most of linear growth in x, uniformly in time, with bounded derivatives. Now consider the change of variables $\widetilde{\psi}_t(x) = \psi_t(x) - \sigma_0 \dot{\phi}_t \cdot (\sigma\sigma^\top)^{-1} x$. We then have

$$\begin{aligned} \partial_t \widetilde{\psi}_t(x) &= \partial_t \psi_t(x) - \sigma_0 \ddot{\phi}_t \cdot (\sigma\sigma^\top)^{-1} x, \\ D_x \widetilde{\psi}_t &= D_x \psi_t - (\sigma\sigma^\top)^{-1} \sigma_0 \dot{\phi}_t, \\ \mathcal{L}_{t,\nu_t} \widetilde{\psi}_t(x) &= \mathcal{L}_{t,\nu_t} \psi_t(x) - \widetilde{b}(t,x,\nu_t) \cdot [(\sigma\sigma^\top)^{-1} \sigma_0 \dot{\phi}_t]. \end{aligned}$$

We then find that

$$\begin{aligned} I^{\sigma_0\phi}(\nu) = \sup_{\psi \in C_c^{1,2}([0,T]\times\mathbb{R}^d)} \Big[& \langle \nu_T, \psi_T \rangle - \langle \nu_0, \psi_0 \rangle \\ & - \int_0^T \Big\langle \nu_t, (\partial_t + \mathcal{L}_{t,\nu_t})\psi_t + \frac{1}{2} |\sigma^\top D_x \psi_t|^2 \Big\rangle dt \Big] \\ & - (\mathbb{M}_T^\nu \cdot [(\sigma\sigma^\top)^{-1} \sigma_0 \dot{\phi}_T] - \mathbb{M}_0^\nu \cdot [(\sigma\sigma^\top)^{-1} \sigma_0 \dot{\phi}_0]) \\ & + \int_0^T \mathbb{M}_t^\nu \cdot [(\sigma\sigma^\top)^{-1} \sigma_0 \ddot{\phi}_t] \, dt \\ & + \int_0^T \Big(\langle \nu_t, \widetilde{b}(t,\cdot,\nu_t) \rangle \cdot [(\sigma\sigma^\top)^{-1} \sigma_0 \dot{\phi}_t] \\ & + \frac{1}{2} [\sigma_0 \dot{\phi}_t] \cdot [(\sigma\sigma^\top)^{-1} \sigma_0 \dot{\phi}_t] \Big) dt. \end{aligned}$$

The first term on the right-hand side is $I^0(\nu)$. By expanding the term on the second line by integration by parts, we get

$$\begin{aligned} I^{\sigma_0\phi}(\nu) = I^0(\nu) + \int_0^T \Big[& -\dot{\mathbb{M}}_t^\nu + \langle \nu_t, \widetilde{b}(t,\cdot,\nu_t) \rangle + \frac{1}{2} \sigma_0 \dot{\phi}_t \Big] \\ & \cdot [(\sigma\sigma^\top)^{-1} \sigma_0 \dot{\phi}_t] \, dt. \end{aligned} \tag{6.24}$$

Note that this shows that $I^{\sigma_0\phi}(\nu) < \infty$ if and only if $I^0(\nu) < \infty$. We wish to extend the identity (6.24) to $\phi \in \mathcal{H}_0^1([0,T]; \mathbb{R}^d)$. As the right-hand side above is clearly continuous in $\mathcal{H}_0^1([0,T]; \mathbb{R}^d)$, we must only show that the left-hand side is as well, at least when suitable terms are finite. Fix a sequence $\phi^n \in C_0^2([0,T]; \mathbb{R}^d)$, converging in $\mathcal{H}^1$-norm to some $\phi \in$

$\mathcal{H}_0^1([0,T];\mathbb{R}^d)$. First, use the definition to see that, for a finite constant C depending on σ_0,

$$\begin{aligned} I^{\sigma_0\phi^n}(\nu) &\le I^{\sigma_0\phi}(\nu) + C\int_0^T \|\dot{\nu}_t - \mathcal{L}^*_{t,\nu_t}\nu_t + \operatorname{div}(\nu_t\sigma_0\dot{\phi}_t)\|_{\nu_t}|\dot{\phi}_t - \dot{\phi}^n_t|\,dt \\ &\quad + \frac{C}{2}\int_0^T |\dot{\phi}_t - \dot{\phi}^n_t|^2\,dt \\ &\le I^{\sigma_0\phi}(\nu) + C\big[I^{\sigma_0\phi}(\nu)\big]^{1/2}\|\phi - \phi^n\|_{\mathcal{H}^1} + \frac{C}{2}\|\phi - \phi^n\|^2_{\mathcal{H}^1}. \end{aligned} \tag{6.25}$$

Similarly,

$$I^{\sigma_0\phi}(\nu) \le I^{\sigma_0\phi^n}(\nu) + \big[I^{\sigma_0\phi^n}(\nu)\big]^{1/2}\|\phi - \phi^n\|_{\mathcal{H}^1} + \frac{1}{2}\|\phi - \phi^n\|^2_{\mathcal{H}^1}. \tag{6.26}$$

If $I^{\sigma_0\phi}(\nu) = \infty$, then $I^{\sigma_0\phi^n}(\nu) = \infty$ for all n, and likewise $I^0(\nu) = \infty$. In this case, the identity (6.24) holds for ϕ. If $I^{\sigma_0\phi}(\nu) < \infty$, then (6.25) implies $\sup_n I^{\sigma_0\phi^n}(\nu) < \infty$. Then (6.25) and (6.26) together imply that $I^{\sigma_0\phi^n}(\nu) \to I^{\sigma_0\phi}(\nu)$, and again (6.24) holds for ϕ.

Now that we know (6.24) holds for all $\phi \in \mathcal{H}_0^1([0,T];\mathbb{R}^d)$, we take the infimum on both sides. To do this, note that if $S = R^\top R$ for some positive definite $d \times d$ matrix R, if V a subspace of $\mathbb{R}^d$, and if Π the orthogonal projection from $\mathbb{R}^d$ to the subspace RV, then for any $y \in \mathbb{R}^d$ we have $\inf_{x\in V} Sx \cdot (\frac{1}{2}x - y) = -\frac{1}{2}|\Pi R y|^2$. With $R = \sigma^{-1}$ and V equal to the image of σ_0, we find

$$\inf_{\phi\in\mathcal{H}_0^1([0,T];\mathbb{R}^d)} I^{\sigma_0\phi}(\nu) = I^0(\nu) - \frac{1}{2}\int_0^T \big|\Pi_{\sigma^{-1}\sigma_0}\sigma^{-1}\big(\dot{\mathbb{M}}^\nu_t - \langle \nu_t, \widetilde{b}(t,\cdot,\nu_t)\rangle\big)\big|^2\,dt.$$

In particular,

$$J^{\sigma_0}(\nu) \le I^0(\nu) - \frac{1}{2}\int_0^T \big|\Pi_{\sigma^{-1}\sigma_0}\sigma^{-1}\big(\dot{\mathbb{M}}^\nu_t - \langle \nu_t, \widetilde{b}(t,\cdot,\nu_t)\rangle\big)\big|^2\,dt.$$

If the left-hand side is infinite, the proof is over. If it is finite, we know from Proposition 6.11 that the infimum over $\mathcal{C}_0^d$ in the definition of J^{σ_0} can be reduced to an infimum over $\mathcal{H}^1([0,T];\mathbb{R}^d)$, since $\mathbb{M}^\nu$ is in $\mathcal{H}^1([0,T];\mathbb{R}^d)$. This completes the proof. □

We now turn to the proof of Theorem 6.6. The proof of Corollary 6.7 is similar, so we omit it.

PROOF OF THEOREM 6.6. Note that the operator $\sigma\Pi_{\sigma^{-1}\sigma_0}$ in the definition of $\mathbb{M}^{\widetilde{b},\nu}$ ensures that there exists $\widetilde{\phi} \in \mathcal{C}_0^d$ such that $\mathbb{M}^{\widetilde{b},\nu} = \sigma_0\widetilde{\phi}$. Thanks to Lemma 6.4, this permits the following change of variables:

$$\begin{aligned} J^{\sigma_0}(\nu) &= \inf_{\phi\in\mathcal{C}_0^d} \widetilde{I}^{\sigma_0\phi}\big((\nu_t \circ \tau^{-1}_{\sigma_0\phi_t})_{t\in[0,T]}\big) \\ &= \inf_{\phi\in\mathcal{C}_0^d} \widetilde{I}^{\sigma_0(\phi+\widetilde{\phi})}\big(((\nu_t \circ \tau^{-1}_{\sigma_0\widetilde{\phi}_t}) \circ \tau^{-1}_{\sigma_0\phi_t})_{t\in[0,T]}\big) \\ &=: \widetilde{J}^{\sigma_0,\widetilde{\phi}}\big((\nu_t \circ \tau^{-1}_{\sigma_0\widetilde{\phi}_t})_{t\in[0,T]}\big), \end{aligned}$$

where, for $\psi \in \mathcal{C}_0^d$, we define $\widetilde{J}^{\sigma_0,\psi}$ just like J^{σ_0} but with the drift modified to $(t,x,m) \mapsto \widetilde{b}(t, x + \sigma_0\psi_t, m \circ \tau^{-1}_{-\sigma_0\psi_t})$. More precisely,

$$\widetilde{J}^{\sigma_0,\psi}(\nu) := \inf_{\phi\in\mathcal{C}_0^d} \widetilde{I}^{\sigma_0(\phi+\psi)}\big((\nu_t \circ \tau^{-1}_{\sigma_0\phi_t})_{t\in[0,T]}\big).$$

The analog of Proposition 6.5 for this modified drift now implies that if ν has mean path in $\mathcal{H}_0^1([0,T];\mathcal{P}^1(\mathbb{R}^d))$ then

$$\begin{aligned}\widetilde{J}^{\sigma_0,\widetilde{\phi}}(\nu) &= \widetilde{I}^{\sigma_0\widetilde{\phi}}(\nu)\\ &\quad - \frac{1}{2}\int_0^T \big|\Pi_{\sigma^{-1}\sigma_0}\sigma^{-1}\big(\dot{\mathbb{M}}_t^\nu - \langle \nu_t, \widetilde{b}(t,\cdot+\sigma_0\widetilde{\phi}_t, \nu_t\circ\tau^{-1}_{-\sigma_0\widetilde{\phi}_t})\rangle\big)\big|^2\,dt.\end{aligned}$$

The mean path of $\tilde{\nu} = (\nu_t\circ\tau^{-1}_{\mathbb{M}_t^{\widetilde{b},\nu}} = \nu_t\circ\tau^{-1}_{\sigma_0\widetilde{\phi}_t})_{t\in[0,T]}$ is precisely

$$\mathbb{M}_t^{\tilde{\nu}} = \mathbb{M}_t^\nu - \sigma\Pi_{\sigma^{-1}\sigma_0}\sigma^{-1}\left(\mathbb{M}_t^\nu - \mathbb{M}_0^\nu - \int_0^t \langle \nu_s, \widetilde{b}(s,\cdot,\nu_s)\rangle\,ds\right),$$

so the above yields

$$\widetilde{J}^{\sigma_0,\widetilde{\phi}}\big((\nu_t\circ\tau^{-1}_{\sigma_0\widetilde{\phi}_t})_{t\in[0,T]}\big) = \widetilde{I}^{\sigma_0\widetilde{\phi}}(\tilde{\nu}) = \widetilde{I}^{\mathbb{M}^{\widetilde{b},\nu}}\big((\nu_t\circ\tau^{-1}_{\mathbb{M}_t^{\widetilde{b},\nu}})_{t\in[0,T]}\big). \qquad \square$$

7. Examples. This section discusses two explicitly solvable models that do not fit our assumptions A. Nonetheless, we show that our strategy for deriving limit theorems by comparison with a more classical McKean–Vlasov system is still successful in these cases.

7.1. *A linear-quadratic model.* In this section, we discuss how our ideas apply to the mean field game model of systemic risk proposed in [14]. Here, $d=1$, σ and σ_0 are positive constants, the action space $A=\mathbb{R}$, and for some $\bar{g},\epsilon,\bar{b}>0$ and $0\le q^2\le\epsilon$ we have

$$\begin{aligned} b(x,m,a) &= \bar{b}(\overline{m}-x)+a,\\ f(x,m,a) &= \frac{1}{2}a^2 - qa(\overline{m}-x) + \frac{\epsilon}{2}(\overline{m}-x)^2,\\ g(x,m) &= \frac{\bar{g}}{2}(\overline{m}-x)^2,\end{aligned}$$

where $\overline{m} = \int_{\mathbb{R}} y\,dm(y)$. Both the drift and cost functions induce a herding behavior toward the population average; see [14] for a thorough discussion.

It was shown in (3.24) of [14] that the unique closed loop Nash equilibrium dynamics is given by

$$\alpha_t^i = \left[q + \varphi_t^n\left(1-\frac{1}{n}\right)\right](\overline{X}_t - X_t^i), \quad t\in[0,T], \tag{7.1}$$

where $\overline{X}_t = \frac{1}{n}\sum_{i=1}^n X_t^i$, and where φ^n is the unique solution to the Riccati equation:

$$\dot{\varphi}_t^n = 2(\bar{b}+q)\varphi_t^n + \left(1-\frac{1}{n^2}\right)|\varphi_t^n|^2 - (\epsilon - q^2), \quad \varphi_T^n = \bar{g}.$$

The explicit solution takes the form

$$\varphi_t^n = \frac{-(\epsilon-q^2)(e^{(\delta_n^+-\delta_n^-)(T-t)}-1) - \bar{g}(\delta_n^+ e^{(\delta_n^+-\delta_n^-)(T-t)} - \delta_n^-)}{(\delta_n^- e^{(\delta_n^+-\delta_n^-)(T-t)} - \delta_n^+) - \bar{g}(1-\frac{1}{n^2})(e^{(\delta_n^+-\delta_n^-)(T-t)}-1)}, \tag{7.2}$$

where

$$\delta_n^\pm = -(\bar{b}+q) \pm \sqrt{(\bar{b}+q)^2 + \left(1-\frac{1}{n^2}\right)(\epsilon-q^2)}. \tag{7.3}$$

In particular, the Nash equilibrium state process is given by the solution $\boldsymbol{X} = (X^1, \dots, X^n)$ of the SDE system

$$dX_t^i = \left(\bar{b} + q + \varphi_t^n\left(1 - \frac{1}{n}\right)\right)(\overline{X}_t - X_t^i)\,dt + \sigma\,dB_t^i + \sigma_0\,dW_t, \qquad t \in [0, T]. \tag{7.4}$$

It is straightforward to show that $\varphi_t^n \to \varphi_t^\infty$ as $n \to \infty$, uniformly in $t \in [0, T]$, where φ^∞ is the unique solution to the Riccati equation

$$\dot{\varphi}_t^\infty = 2(\bar{b} + q)\varphi_t^\infty + |\varphi_t^\infty|^2 - (\epsilon - q^2), \qquad \varphi_T^\infty = \bar{g}.$$

The explicit solution is of the same form given by (7.2) and (7.3), with $n = \infty$. It follows that $\boldsymbol{X} = (X^1, \dots, X^n)$ should be "close" in some sense to the solution $\boldsymbol{Y} = (Y^1, \dots, Y^n)$ of the auxiliary SDE system

$$dY_t^i = (\bar{b} + q + \varphi_t^\infty)(\overline{Y}_t - Y_t^i)\,dt + \sigma\,dB_t^i + \sigma_0\,dW_t, \tag{7.5}$$

initialized at the same points $Y_0^i = X_0^i$. Of course, it should be noted that the process $\boldsymbol{Y}$ plays here the same role as the process $\overline{\boldsymbol{X}}$ in (4.1), the solution U to the master equation being given in the current framework by

$$U(t, x, m) = \frac{\varphi_t^\infty}{2}(\overline{m} - x)^2.$$

In this regard, the fact that $\boldsymbol{X}$ and $\boldsymbol{Y}$ should be "close" is completely analogous to the statements of Theorems 4.1 and 4.2. Here, we prefer to use $\boldsymbol{Y}$ instead of the notation $\overline{\boldsymbol{X}}$ used in previous sections, to avoid any confusion with the empirical mean process that appears in (7.1).

To compare (7.4) and (7.5), we use the fact that $\boldsymbol{X}_0 = \boldsymbol{Y}_0$, and we apply Gronwall's inequality to find a constant $C < \infty$ such that

$$\frac{1}{n}\sum_{i=1}^n \|X^i - Y^i\|_\infty \le C\left\|\left(1 - \frac{1}{n}\right)\varphi^n - \varphi^\infty\right\|_\infty \frac{1}{n}\sum_{i=1}^n \|X^i\|_\infty, \quad \text{a.s.}, \tag{7.6}$$

where, as usual, $\|\cdot\|_\infty$ denotes the supremum norm on $[0, T]$. On the other hand, the equation (7.4) and Gronwall's inequality yield

$$\frac{1}{n}\sum_{i=1}^n \|X^i\|_\infty \le C\left(1 + \frac{1}{n}\sum_{i=1}^n |X_0^i| + \frac{1}{n}\sum_{i=1}^n \|B^i\|_\infty + \|W\|_\infty\right) \quad \text{a.s.}$$

As soon as $(X_0^i)_{i\ge1}$ are i.i.d. and sub-Gaussian (e.g., $\mathbb{E}[\exp(\kappa|X_0^1|^2)] < \infty$ for some $\kappa > 0$), we find a uniform sub-Gaussian bound on these averages; that is, there exist constants $C < \infty$, $\delta > 0$, independent of n, such that

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n \|X^i\|_\infty > a\right) \le \exp(-\delta^2 a^2), \quad \text{for all } a \ge C, n \in \mathbb{N}.$$

Assuming without any loss of generality that the constant C in the last display coincides with the one in (7.6), and letting $r_n = C\|(1 - \frac{1}{n})\varphi^n - \varphi^\infty\|_\infty$, we find that, for $a \ge Cr_n$,

$$\begin{aligned}
\mathbb{P}(\mathcal{W}_{1,\mathcal{C}^d}(m_{\boldsymbol{X}}^n, m_{\boldsymbol{Y}}^n) > a) &\le \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n \|X^i - Y^i\|_\infty > a\right) \\
&\le \mathbb{P}\left(\frac{r_n}{n}\sum_{i=1}^n \|X^i\|_\infty > a\right) \\
&\le \exp(-\delta^2 a^2 / r_n^2).
\end{aligned} \tag{7.7}$$

It is straightforward to check that $r_n = O(1/n)$, which implies in particular the exponential equivalence of $(m^n_{\boldsymbol{X}})$ and $(m^n_{\boldsymbol{Y}})$, in the sense that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\mathcal{W}_{1,\mathcal{C}^d}(m^n_{\boldsymbol{X}}, m^n_{\boldsymbol{Y}}) > a\big) = -\infty \quad \text{for all } a > 0.$$

Moreover, the concentration estimates of Section 3.1 are all valid; all that was used in the proofs were the estimates in (7.7) and the concentration bounds for McKean–Vlasov systems of Sections 5.2 and 5.3.

Derivation of the LDP. As made clear in Section 6, the relation (7.7) is the cornerstone to get a LDP for $(m^n_{\boldsymbol{X}})_{n\ge1}$. Indeed, we can have the LDP for $(m^n_{\boldsymbol{Y}})_{n\ge1}$ by adapting the arguments of Section 6, but this requires some care as the drift here is no longer bounded.

Most of the derivation of Theorem 6.8 is based upon on the contraction principle: the fact that the drift is unbounded is not a problem for duplicating the proof. In fact, the assumption that b is bounded is used only a few times in Section 6, mainly for the derivation of Propositions 6.10 and 6.11. We explain below how to accommodate the unboundedness of b.

Notice in particular that, specialized to the present setting, the rate function of the weak LDP (see Theorem 6.8) has the form

$$J^{\sigma_0}(\nu) + \mathcal{R}(\nu_0|\mu_0),$$

where

$$J^{\sigma_0}(\nu) = \inf_{\phi\in\mathcal{C}^d_0} I^0\big((\nu_t \circ \tau^{-1}_{\sigma_0\phi_t})_{t\in[0,T]}\big),$$

I^0 standing for Dawson and Gartner's rate function as defined in the statement of Lemma 6.4 with the drift $\widetilde{b} : (t,x,\mu) \mapsto \bar{b} + q + \varphi^\infty_t(x - \overline{\mu})$ and with $\overline{\mu}$ denoting the mean of μ. Remarkably, since $\widetilde{b}(t, x+\phi_t, m\circ\tau^{-1}_{-\phi_t}) = \widetilde{b}(t,x,m)$, I^0 is completely independent of ϕ, which ultimately leads to nice formulas in our setting.

When $\sigma_0 = 0$, there is no need to push further the analysis. So, for the rest of this short discussion, we can assume $\sigma_0 > 0$. To proceed, we observe that, due to the special form of interaction in the dynamics, we can easily shift the path ϕ appearing in the definition of $J^{\sigma_0}(\nu)$. Indeed, we can rewrite $J^{\sigma_0}(\nu)$ (first changing $\sigma_0\phi$ into ϕ and then shifting ϕ) as

$$J^{\sigma_0}(\nu) = \inf_{\phi\in\mathcal{C}^d_0} I^0\big(\big((\nu_t \circ \tau^{-1}_{\mathbb{M}^\nu_t - \mathbb{M}^\nu_0}) \circ \tau^{-1}_{\phi_t}\big)_{t\in[0,T]}\big).$$

The key fact to observe here is that $\nu_t \circ \tau^{-1}_{\mathbb{M}^\nu_t - \mathbb{M}^\nu_0}$ has zero mean.

When ϕ is smooth enough, Lemma 6.4 provides another representation for $I^0((\nu_t \circ \tau^{-1}_{\phi_t})_{t\in[0,T]})$ and the relation (6.24) in the proof of Proposition 6.5 remains true as well. Thus, combining the special form of the drift together with (6.24), we see that, when $(\nu_t)_{t\in[0,T]}$ has a constant mean, we have

$$I^0\big((\nu_t \circ \tau^{-1}_{\phi_t})_{t\in[0,T]}\big) \ge I^0\big((\nu_t)_{t\in[0,T]}\big).$$

Arguing as in (6.25)–(6.26), the latter remains true when ϕ lies in $\mathcal{H}^1_0([0,T];\mathbb{R}^d)$. We now want to check that this remains true when $\phi \in \mathcal{C}^d_0$. To do so, we must revisit the first step in the proof of Proposition 6.10. If

$$I^0\big((\nu_t \circ \tau^{-1}_{\phi_t})_{t\in[0,T]}\big) \le a,$$

for some $a > 0$, we can find a constant $C(a,\nu)$ such that $\mathbb{M}^\nu - \phi$ lies in $\mathcal{H}^1([0,T];\mathbb{R}^d)$ with an $\mathcal{H}^1$ norm less than $C(a,\nu)$. The main difference with the proof of Proposition 6.10

is that the constant C here depends on ν, but it suffices to check that necessarily ϕ lies in $\mathcal{H}^1([0,T];\mathbb{R}^d)$. Therefore, (still in the case where $(\nu_t)_{t\in[0,T]}$ has a constant mean) we end up with

$$\inf_{\phi\in\mathcal{C}_0^d} I^0\big((\nu_t\circ\tau_{\phi_t}^{-1})_{t\in[0,T]}\big) = I^0\big((\nu_t)_{t\in[0,T]}\big).$$

In the general case when the mean is not constant, this yields

$$J^{\sigma_0}(\nu) = I^0\big((\nu_t\circ\tau^{-1}_{\mathbb{M}_t^\nu-\mathbb{M}_0^\nu})_{t\in[0,T]}\big).$$

Then, if needed, we can revisit the proof of Proposition 6.10 to specialize the upper bound in the case of compact sets. The only fact that is needed from Proposition 6.11 is that the aforementioned constant $C(a,\nu)$ is uniform in ν in compact subsets, which can be shown to be true. This suffices to obtain the complete form of the LDP, as stated in Theorem 3.11.

7.2. *A Merton-type model.* We now turn to one of the models of [32], which fails to fit our general assumptions for a number of reasons. As in Section 7.1, the coefficients are unbounded and the Hamiltonian is non-Lipschitz. But now both volatility terms are controlled, and agents are more heterogeneous in the sense that each is assigned a certain *type vector*, denoted by $\zeta_i = (X_0^i, \delta_i, \theta_i, \mu_i, \sigma_i, \nu_i)$ and belonging to the space

$$\mathcal{Z} := \big\{(x,\delta,\theta,\mu,\sigma,\nu)\in\mathbb{R}\times(0,\infty)\times[0,1]\times(0,\infty)\times[0,\infty)^2 : \sigma+\nu\geq c\big\},$$

where $c>0$ is fixed. Suppose henceforth that we are given an infinite sequence of deterministic type vectors $(\zeta_i)_{i\in\mathbb{N}}$. Assume also, for simplicity, that all of these parameters are uniformly bounded from above.

The n-player game is described by a state process $\boldsymbol{X} = (X^1,\dots,X^n)$ given by

$$dX_t^i = \alpha_t^i\big(\mu_i\,dt + \nu_i\,dB_t^i + \sigma_i\,dW_t\big),$$

where each X_t^i is one-dimensional. Agent i chooses $(\alpha_t^i)_{t\in[0,T]}$ to try to maximize the expected utility

$$-\mathbb{E}\Big[\exp\Big(-\frac{1}{\delta_i}(X_T^i - \theta_i\overline{X}_T)\Big)\Big],$$

where $\overline{X}_T = \frac{1}{n}\sum_{k=1}^n X_T^k$. This is essentially Merton's problem of portfolio optimization, under exponential utility, but with each agent concerned not only with absolute wealth but also with relative wealth, as measured by the average $\overline{X}_T$. The parameter $\theta_i\in[0,1]$ determines the tradeoff between absolute and relative performance concerns; see [32] for a complete discussion.

We express the equilibrium in terms of the constant

$$\eta_n := \frac{1}{n}\sum_{k=1}^n \frac{\delta_k\mu_k\sigma_k}{\sigma_k^2+\nu_k^2(1-\theta_k/n)} \Bigg/ \left(1-\frac{1}{n}\sum_{k=1}^n\frac{\theta_k\sigma_k^2}{\sigma_k^2+\nu_k^2(1-\theta_k/n)}\right),$$

assuming the denominator is nonzero (which certainly holds if $\theta_k<1$ for at least one k). It is shown in Theorem 3 of [32] that there exists a Nash equilibrium in which agent i chooses the constant (i.e., time- and state-independent) control

$$\alpha_i^n := \frac{\delta_i\mu_i + \eta_n\theta_i\sigma_i}{\sigma_i^2+\nu_i^2(1-\theta_i/n)}.$$

The corresponding state process is given by

$$X_t^i = X_0^i + \alpha_i^n\mu_i t + \alpha_i^n\nu_i B_t^i + \alpha_i^n\sigma_i W_t.$$

Now, as in the previous section, we can show that $\boldsymbol{X}$ is very close to a particle system $\boldsymbol{Y} = (Y^1, \dots, Y^n)$, where

$$Y_t^i = X_0^i + \widetilde{\alpha}_i^n \mu_i t + \widetilde{\alpha}_i^n \nu_i B_t^i + \widetilde{\alpha}_i^n \sigma_i W_t,$$

and where

$$\widetilde{\alpha}_i^n := \frac{\delta_i \mu_i + \eta_n \theta_i \sigma_i}{\sigma_i^2 + \nu_i^2},$$

$$\widetilde{\eta}_n := \frac{1}{n} \sum_{k=1}^n \frac{\delta_k \mu_k \sigma_k}{\sigma_k^2 + \nu_k^2} \Big/ \left(1 - \frac{1}{n} \sum_{k=1}^n \frac{\theta_k \sigma_k^2}{\sigma_k^2 + \nu_k^2}\right).$$

More precisely, note that the uniform bounds on the type parameters ensure that there exists $\widetilde{L} > 0$ such that $|\widetilde{\alpha}_i^n - \alpha_i^n| \le \widetilde{L}/n$ for all $n \ge 2$ and all i, and we conclude that

$$\|X^i - Y^i\|_\infty \le \frac{\widetilde{L}}{n}(\mu_i T + \nu_i \|B^i\|_\infty + \sigma_i \|W\|_\infty).$$

By assuming that $(X_0^i)_{i \ge 1}$ are i.i.d. and sub-Gaussian as in the previous subsection, it is straightforward to show that there exist constants $C, \delta > 0$, independent of n, such that

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n \|X^i - Y^i\|_\infty > a\right) \le \exp(-\delta^2 n^2 a^2), \quad \text{for all } a \ge C/n, n \ge 2. \tag{7.8}$$

Again, this estimate allows us to transfer limit theorems and concentration estimates for $\boldsymbol{Y}$ over to $\boldsymbol{X}$.

While $\boldsymbol{Y}$ is not exactly a standard McKean–Vlasov system because of the type parameters, it is close enough that we can do some similar analysis. Let us illustrate one simple way to study the limiting behavior of $m_{\boldsymbol{Y}}^n$. Define a map $\Psi : \mathcal{P}(\mathcal{Z} \times \mathcal{C}^1) \times \mathcal{C}^1 \to \mathcal{P}(\mathcal{C}^1)$ by setting $\Psi(Q, w)$ equal to the image of $Q \circ \widehat{Y}_w^{-1}$, where $\widehat{Y}_w : \mathcal{Z} \times \mathcal{C}^1 \to \mathcal{C}^1$ is defined for each $w \in \mathcal{C}^1$ by setting

$$\widehat{Y}_w(\zeta, \ell)(t) = x_0 + \frac{\delta\mu + \overline{Q}_1 \theta\sigma}{\sigma^2 + \nu^2}(\mu t + \nu \ell(t) + \sigma w(t)),$$

where $\zeta = (x_0, \delta, \theta, \mu, \sigma, \nu)$, and where

$$\overline{Q}_1 := \int_{\mathcal{Z} \times \mathcal{C}^1} \frac{\delta\mu\sigma}{\sigma^2 + \nu^2} \Big/ \left(1 - \frac{\theta\sigma^2}{\sigma^2 + \nu^2}\right) Q(d\zeta, d\ell).$$

We may then write

$$m_{\boldsymbol{Y}}^n = \Psi\left(\frac{1}{n}\sum_{i=1}^n \delta_{(\zeta_i, B^i)}, W\right).$$

For a fixed $M > 0$, it is easily checked that the map Ψ is continuous (with respect to weak convergence) when restricted to the subset of (Q, w) for which $\delta\mu\sigma \le M$ and $1 - \theta\sigma^2/(\sigma^2 + \nu^2) \ge 1/M$ holds for Q-a.e. (ζ, ℓ). Therefore, we may easily identify the limit of $m_{\boldsymbol{Y}}^n$ as $n \to \infty$, as long as $\frac{1}{n}\sum_{i=1}^n \delta_{(\zeta_i, B^i)}$ converges a.s. Moreover, if the type vectors ζ_i are i.i.d. then the sequence of empirical measures $\frac{1}{n}\sum_{i=1}^n \delta_{(\zeta_i, B^i)}$ satisfies a LDP, according to Sanov's theorem. If $\sigma_i = 0$ for all i, so there is no common noise, then $\Psi(Q, w)$ does not depend on w, and we may deduce a LDP for $m_{\boldsymbol{Y}}^n$ from the contraction principle. If the common noise is present, we can either deduce a LDP conditionally on W (i.e., quenched), or we can deduce an unconditional (i.e., annealed) *weak LDP*, as is done in Propositions 6.15 and Theorem 6.8 in a general setting.

8. Conclusions and open problems. In this paper and the companion [20], we have seen how a sufficiently well behaved solution to the master equation can be used to derive asymptotics for mean field games, in the form of a law of large numbers, central limit theorem and LDP, as well as nonasymptotic concentration bounds. This worked under a class of reasonable but restrictive assumptions, notably including boundedness of various derivatives of the master equation. Without this boundedess, it is not clear if we can always expect the Nash system $m^n_{\boldsymbol{X}}$ and the McKean–Vlasov system $m^n_{\overline{\boldsymbol{X}}}$ to share the same large deviations, or to be exponentially equivalent as in Theorem 4.3. In the two examples we presented in Section 7, there were no difficulties, but it is not clear how much regularity we really need for the master equation.

To comment more on this point, note that the proof of our main estimate Theorem 4.1 (given in Section 4 of [20]) was in many ways parallel to Lipschitz FBSDE estimates. To cover linear-quadratic models, we should allow the first derivatives of $U(t, x, m)$ to grow linearly in x and $\mathcal{W}_1(m, \delta_0)$ and the Hamiltonian to have quadratic growth in both x and α. This leads to a quadratic FBSDE system, as we encountered in the proof of Theorem 4.2 (given in Section 4 of [20]), but with unbounded coefficients controlled only in terms of the forward component. This would certainly require a much more delicate analysis.

Technical assumptions notwithstanding, there is an interesting gap in the current state of the limit theory for closed-loop versus open-loop equilibria. The papers [24, 30] provide laws of large numbers for open-loop equilibria, with the key advantage of addressing the nonunique regime, that is, when there are multiple mean field equilibria. A sequence of n-player equilibria may have multiple limit points as $n \to \infty$, but any such limit point is a mean field equilibrium in a suitable weak sense. In the closed-loop setting, there are no limit theorems addressing the nonunique regime, which is important in light of the fact that nonuniqueness is a key feature of many game theoretic models. On the other hand, we now have a central limit theorem and LDP for closed-loop equilibria, in the unique regime, and no such results are known for open-loop equilibria. However, it is worth mentioning that analogous LDPs have been established in the nonunique regime in the simpler setting of static games [31].

Acknowledgments. The first author was supported by ANR-16-CE40-0015-01 and Institut Universitaire de France.

The third author was supported in part by the National Science Foundation Grant DMS-1713032.

REFERENCES

[1] BAYRAKTAR, E. and COHEN, A. (2018). Analysis of a finite state many player game using its master equation. *SIAM J. Control Optim.* **56** 3538–3568. MR3860894 https://doi.org/10.1137/17M113887X

[2] BEN AROUS, G. and BRUNAUD, M. (1990). Méthode de Laplace: étude variationnelle des fluctuations de diffusions de type "champ moyen". *Stoch. Stoch. Rep.* **31** 79–144. MR1080535 https://doi.org/10.1080/03610919008833649

[3] BENSOUSSAN, A. and FREHSE, J. (1983). *Nonlinear Elliptic Systems in Stochastic Game Theory*, Universität Bonn. SFB 72. Approximation und Optimierung, 1983.

[4] BENSOUSSAN, A., FREHSE, J. and YAM, S. C. P. (2015). The master equation in mean field theory. *J. Math. Pures Appl.* (9) **103** 1441–1474. MR3343705 https://doi.org/10.1016/j.matpur.2014.11.005

[5] BENSOUSSAN, A., FREHSE, J. and YAM, S. C. P. (2017). On the interpretation of the Master Equation. *Stochastic Process. Appl.* **127** 2093–2137. MR3652408 https://doi.org/10.1016/j.spa.2016.10.004

[6] BOBKOV, S. G. and GÖTZE, F. (1999). Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** 1–28. MR1682772 https://doi.org/10.1006/jfan.1998.3326

[7] BOLLEY, F., GUILLIN, A. and MALRIEU, F. (2010). Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov–Fokker–Planck equation. *ESAIM Math. Model. Numer. Anal.* **44** 867–884. MR2731396 https://doi.org/10.1051/m2an/2010045

[8] BOLLEY, F., GUILLIN, A. and VILLANI, C. (2007). Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields* **137** 541–593. MR2280433 https://doi.org/10.1007/s00440-006-0004-7

[9] BUDHIRAJA, A., DUPUIS, P. and FISCHER, M. (2012). Large deviation properties of weakly interacting processes via weak convergence methods. *Ann. Probab.* **40** 74–102. MR2917767 https://doi.org/10.1214/10-AOP616

[10] CARDALIAGUET, P., DELARUE, F., LASRY, J.-M. and LIONS, P.-L. (2019) *The Master Equation and the Convergence Problem in Mean Field Games*. *Annals of Mathematics Studies* **12** Princeton Univ. Press.

[11] CARMONA, R. and DELARUE, F. (2014). The master equation for large population equilibriums. In *Stochastic Analysis and Applications* 2014. *Springer Proc. Math. Stat.* **100** 77–128. Springer, Cham. MR3332710 https://doi.org/10.1007/978-3-319-11292-3_4

[12] CARMONA, R. and DELARUE, F. (2017). *Probabilistic Theory of Mean Field Games*: *Vol. I*, *Mean Field FBSDEs*, *Control*, *and Games*. *Stochastic Analysis and Applications*, Springer, Berlin.

[13] CARMONA, R. and DELARUE, F. (2017). *Probabilistic Theory of Mean Field Games*: *Vol. II*, *Mean Field Games with Common Noise and Master Equations*. *Stochastic Analysis and Applications*, Springer, Berlin.

[14] CARMONA, R., FOUQUE, J.-P. and SUN, L.-H. (2015). Mean field games and systemic risk. *Commun. Math. Sci.* **13** 911–933. MR3325083 https://doi.org/10.4310/CMS.2015.v13.n4.a4

[15] CARMONA, R. and ZHU, X. (2016). A probabilistic approach to mean field games with major and minor players. *Ann. Appl. Probab.* **26** 1535–1580. MR3513598 https://doi.org/10.1214/15-AAP1125

[16] CECCHIN, A. and FISCHER, M. (2017). Probabilistic approach to finite state mean field games. Preprint. Available at arXiv:1704.00984.

[17] CECCHIN, A. and PELINO, G. (2019). Convergence, fluctuations and large deviations for finite state mean field games via the master equation. *Stochastic Process. Appl.* **129** 4510–4555.

[18] CHASSAGNEUX, J.-F., CRISAN, D. and DELARUE, F. (2014). A probabilistic approach to classical solutions of the master equation for large population equilibria. Preprint. Available at arXiv:1411.3009.

[19] DAWSON, D. A. and GÄRTNER, J. (1987). Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20** 247–308. MR0885876 https://doi.org/10.1080/17442508708833446

[20] DELARUE, F., LACKER, D. and RAMANAN, K. (2018). From the master equation to mean field game limit theory: A central limit theorem. *Electron. J. Probab.* **24** 1–54.

[21] DEMBO, A. and ZEITOUNI, O. (2010). *Large Deviations Techniques and Applications*. *Stochastic Modelling and Applied Probability* **38**. Springer, Berlin. MR2571413 https://doi.org/10.1007/978-3-642-03311-7

[22] DJELLOUT, H., GUILLIN, A. and WU, L. (2004). Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.* **32** 2702–2732. MR2078555 https://doi.org/10.1214/009117904000000531

[23] FISCHER, M. (2014). On the form of the large deviation rate function for the empirical measures of weakly interacting systems. *Bernoulli* **20** 1765–1801. MR3263089 https://doi.org/10.3150/13-BEJ540

[24] FISCHER, M. (2017). On the connection between symmetric N-player games and mean field games. *Ann. Appl. Probab.* **27** 757–810. MR3655853 https://doi.org/10.1214/16-AAP1215

[25] FOURNIER, N. and GUILLIN, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. MR3383341 https://doi.org/10.1007/s00440-014-0583-7

[26] GANGBO, W. and ŚWIĘCH, A. (2015). Existence of a solution to an equation arising from the theory of mean field games. *J. Differential Equations* **259** 6573–6643. MR3397332 https://doi.org/10.1016/j.jde.2015.08.001

[27] GOZLAN, N. (2009). A characterization of dimension free concentration in terms of transportation inequalities. *Ann. Probab.* **37** 2480–2498. MR2573565 https://doi.org/10.1214/09-AOP470

[28] GOZLAN, N. and LÉONARD, C. (2010). Transport inequalities. A survey. *Markov Process. Related Fields* **16** 635–736. MR2895086

[29] HOROWITZ, J. and KARANDIKAR, R. L. (1994). Mean rates of convergence of empirical measures in the Wasserstein metric. *J. Comput. Appl. Math.* **55** 261–273. MR1329874 https://doi.org/10.1016/0377-0427(94)90033-7

[30] LACKER, D. (2016). A general characterization of the mean field limit for stochastic differential games. *Probab. Theory Related Fields* **165** 581–648. MR3520014 https://doi.org/10.1007/s00440-015-0641-9

[31] LACKER, D. and RAMANAN, K. (2019). Rare Nash equilibria and the price of anarchy in large static games. *Math. Oper. Res.* **44** 400–422.

[32] LACKER, D. and ZARIPHOPOULOU, T. (2017). Mean field and n-agent games for optimal investment under relative performance criteria. Preprint. Available at arXiv:1703.07685.

[33] LASRY, J.-M. and LIONS, P.-L. (2007). Mean field games. *Jpn. J. Math.* **2** 229–260. MR2295621 https://doi.org/10.1007/s11537-007-0657-8

[34] LEDOUX, M. (2001). *The Concentration of Measure Phenomenon*. *Mathematical Surveys and Monographs* **89**. Amer. Math. Soc., Providence, RI. MR1849347

[35] STRÖMBERG, T. (2009). A note on the differentiability of conjugate functions. *Arch. Math.* (*Basel*) **93** 481–485. MR2563594 https://doi.org/10.1007/s00013-009-0050-3

[36] STROOCK, D. W. and VARADHAN, S. R. S. (1972). On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California*, *Berkeley*, *Calif.*, 1970/1971), *Vol. III*: *Probability theory* Univ. California Press, Berkeley, CA. 333–359. MR0400425

[37] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *École D'Été de Probabilités de Saint-Flour XIX*—1989. *Lecture Notes in Math.* **1464** 165–251. Springer, Berlin. MR1108185 https://doi.org/10.1007/BFb0085169

[38] ÜSTÜNEL, A. S. (2012). Transportation cost inequalities for diffusions under uniform distance. In *Stochastic Analysis and Related Topics*. *Springer Proc. Math. Stat.* **22** 203–214. Springer, Heidelberg. MR3236093 https://doi.org/10.1007/978-3-642-29982-7_9

[39] WANG, R., WANG, X. and WU, L. (2010). Sanov's theorem in the Wasserstein distance: A necessary and sufficient condition. *Statist. Probab. Lett.* **80** 505–512. MR2593592 https://doi.org/10.1016/j.spl.2009.12.003