# LIMIT THEORY FOR CONTROLLED MCKEAN–VLASOV DYNAMICS[*]

DANIEL LACKER[†]

**Abstract.** This paper rigorously connects the problem of optimal control of McKean–Vlasov dynamics with large systems of interacting controlled state processes. Precisely, the empirical distributions of near-optimal control-state pairs for the $n$-state systems, as $n$ tends to infinity, admit limit points in distribution (if the objective functions are suitably coercive), and every such limit is supported on the set of optimal control-state pairs for the McKean–Vlasov problem. Conversely, any distribution on the set of optimal control-state pairs for the McKean–Vlasov problem can be realized as a limit in this manner. Arguments are based on controlled martingale problems, which lend themselves naturally to existence proofs; along the way it is shown that a large class of McKean–Vlasov control problems admit optimal Markovian controls.

**Key words.** mean field control, McKean–Vlasov limit, controlled martingale problems, relaxed controls

**AMS subject classifications.** 93E20, 82C22, 60H30

**DOI.** 10.1137/16M1095895

**1. Introduction.** The past decade has seen a surge of interest in the optimal control of McKean–Vlasov dynamics, also known as mean field control. This problem can be described loosely as follows: The controller chooses a process $\alpha$, which in turn determines the state process $X$ via a McKean–Vlasov stochastic differential equation (SDE)

$$dX_t = b\left(t, X_t, \mathbb{P} \circ X_t^{-1}, \alpha_t\right) dt + \sigma\left(t, X_t, \mathbb{P} \circ X_t^{-1}, \alpha_t\right) dW_t,$$

where $W$ is a Brownian motion and $\mathbb{P} \circ X_t^{-1}$ denotes the law of $X_t$. The controller seeks to maximize a functional of the form

$$(1.1) \qquad \mathbb{E}\left[\int_0^T f\left(t, X_t, \mathbb{P} \circ X_t^{-1}, \alpha_t\right) dt + g\left(X_T, \mathbb{P} \circ X_T^{-1}\right)\right],$$

where $T > 0$ is a fixed time horizon. The unusual feature of this control problem is that the functions $(b, \sigma, f, g)$ depend on the law $\mathbb{P} \circ X_t^{-1}$ of the state process.

The study of McKean–Vlasov control problems is often justified by a heuristic connection to control problems involving large but finite numbers of interacting state processes. More precisely, imagine there are $n$ state processes interacting through their empirical measures via the following SDE system:

$$(1.2) \qquad dX_t^i = b\left(t, X_t^i, \widehat{\mu}_t^n, \alpha_t^i\right) dt + \sigma\left(t, X_t^i, \widehat{\mu}_t^n, \alpha_t^i\right) dW_t^i,$$

$$\widehat{\mu}_t^n = \frac{1}{n} \sum_{k=1}^n \delta_{X_t^k}.$$

†Division of Applied Mathematics, Brown University, Providence, RI 02906 (daniel_lacker@brown.edu).

Here $W^1, \ldots, W^n$ are independent Brownian motions, and $\alpha^1, \ldots, \alpha^n$ are controls chosen by a central planner. The objective of this central planner is to maximize the averaged objective

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ \int_0^T f\left(t, X_t^i, \widehat{\mu}_t^n, \alpha_t^i\right) dt + g\left(X_T^i, \widehat{\mu}_T^n\right) \right].$$

Of course, if $(X_0^i, W^i, \alpha^i)_{i=1}^n$ are suitably exchangeable (and the SDEs sufficiently well-posed), then each term in the average is equal, and the problem reduces to maximizing the objective corresponding to a single state. Throughout the paper, however, we avoid imposing any symmetry assumptions on the controls $(\alpha^1, \ldots, \alpha^n)$, as there may exist asymmetric optimizers.

When there is no control present (i.e., no $\alpha$ in $b$ or $\sigma$), it is by now well-established that the empirical measure flow $(\widehat{\mu}_t^n)_{t \in [0,T]}$ of (1.2) converges in a sense to the measure flow $(\mathbb{P} \circ X_t^{-1})_{t \in [0,T]}$ arising from (1.1); see, for instance [45, 22, 51]. This is true, at least, under reasonable continuity assumptions on $(b, \sigma)$, the most important of which is that the interactions are *weak* or *nonlocal* in the sense that the dependence on the measure argument is continuous with respect to weak convergence or a Wasserstein metric.

For the controlled model, however, is not obvious that this limit should commute with the optimization. The primary goal of this paper is to address this issue by providing general conditions under which a sequence of optimizers of the $n$-state system must converge (in the sense of empirical measure) to solutions of the McKean–Vlasov control problem. More precisely, Theorem 2.11 shows, under modest assumptions on $(b, \sigma, f, g)$, that the empirical measure flows $(\widehat{\mu}_t^n)_{t \in [0,T]}$ of optimally controlled $n$-state systems are tight, and every limit in distribution is supported on the set of measures flows $(\mathbb{P} \circ X_t^{-1})_{t \in [0,T]}$, where $X$ is an optimally controlled state in the McKean–Vlasov control problem. As an immediate corollary, whenever the McKean–Vlasov control problem admits a unique optimal control, we obtain a proper convergence result or *propagation of chaos* [51].

Our arguments are largely based on martingale problems, combining ideas from the McKean–Vlasov limit theory with a well-established compactification method for stochastic control. The state equations, both for the $n$-state problem and the McKean–Vlasov control problem, are formulated as controlled martingale problems and with *relaxed* (i.e., measure-valued) controls. For standard stochastic control problems, this formulation provides a certain compactness which has facilitated very general results on the existence of optimal controls. This idea originated with Fleming [19] and matured with the works of El Karoui, Nguyen, and Jeanblanc-Picqué [32] and Haussmann and Lepeltier [25], later seeing extensions to general state spaces [37]. Our Theorem 2.2 provides an analogous result on the existence of optimal relaxed controls for the McKean–Vlasov problem. Moreover, as in [32, 25], we show under an additional convexity hypothesis that there exists an optimal *Markovian* control. Remarkably, the mean field term does not complicate the arguments leading to Markovian controls, which are based on the mimicking theorem of Gyöngy [24], or rather the generalization due to Brunick and Shreve [6].

The proof of the main limit theorem follows the well-trodden path of formulating the limiting equation as a martingale problem in the sense of Stroock and Varadhan [50], with an additional nonlinearity stemming from the mean field term term. This particular approach to the study of McKean–Vlasov limits seems to have originated with Oelschläger [45], while the impressive paper of Gärtner [22] contains the most

broadly applicable results for models with continuous coefficients. Similar martingale arguments have been applied to a number of related models, including stronger interactions [46, 44, 29], rank-based models [49, 30], and Boltzmann-type models [43, 23]. The monograph of Sznitman [51] provides a general overview and a bird's eye view of some variants.

Our limit theorem appears to be the first its kind for controlled diffusions, and only three recent papers seem to touch on this: First, Fischer and Livieri [18] prove a limit theorem for a very special case of our model arising from mean-variance portfolio optimization. Second, Fornasier and Solombrino [20] treat a general class of related deterministic (i.e., $\sigma \equiv 0$) models; our results allow for degenerate volatility but do not subsume theirs. Last but not least, Budhiraja, Dupuis, and Fischer [8] study weak limits of empirical measures of controlled interacting diffusions with relaxed controls, en route to proving a large deviation principle for the McKean–Vlasov limit. Section 5 of their paper contains analysis similar to our section 5, but they encounter only particular types of coefficients, with linear-quadratic dependence on the control variable.

The literature on McKean–Vlasov optimal control problems is focused primarily on solution techniques. Only one paper [2] seems to adopt remotely similar techniques to ours, using relaxed controls (but not martingale problems) and much more restrictive assumptions on the form of the coefficients. The most popular techniques are based on extending Pontryagin's maximum principle [1, 7, 9, 4] or deriving a dynamic programming principle, and with it a form of a Hamilton–Jacobi–Bellman equation on a space of probability measures [41, 47, 3] (related to the so-called master equation studied in [10, 5]). Our solvability result does not provide any insight on how to construct an optimizer, and its strength lies rather in its generality, requiring not even Lipschitz assumptions.

While our assumptions on the model parameters are quite modest, several interesting extensions are left untouched. For instance, Pham and Wei in [47] study an extension of the basic model in which the coefficients depend on the law of the control, not just the state. More notably, we do not address models with *common noise* (cf. [48]), in which an additional independent Brownian motion $B$ appears in the dynamics, and the law $\mathbb{P} \circ X_t^{-1}$ in the coefficients is replaced by the conditional law $\mathbb{P}(X_t \in \cdot | B_s, \ s \le t)$. It may be possible to modify our methods to treat the limit theory for such models. The analogous mean field game limit theory of [38] suggests that we study the symmetrized measure $\frac{1}{n} \sum_{i=1}^{n} \mathrm{law}(X^i, \alpha^i, W^i, B, \widehat{\mu}^n)$ and identify a suitable notion of *weak solution* to account for the additional randomness which may appear when taking weak limits. The analysis would certainly be far more complex, and this is left for future work.

The optimal control of McKean–Vlasov dynamics is closely related to *mean field game theory*, which was developed by Lasry and Lions [40] and Huang, Malhamé, and Caines [28]. Mean field games are essentially concerned with the continuum limit of a *competitive* form of the $n$-state control problem, in which the controls $\alpha^1, \ldots, \alpha^n$ are chosen by different agents in Nash equilibrium. In several applications, in fact, controlled McKean–Vlasov dynamics are studied so that the competitive (decentralized) outcome can be compared with the Pareto optimal (centralized) one [26, 27]. The paper [12] and the forthcoming book [11] study and compare these two distinct regimes, highlighting the significant methodological overlap. It is worth mentioning in particular that martingale methods and relaxed controls have been applied in the study of mean field games, both for existence theory [39, 13] and limit theory [38, 17], and the present work borrows several technical points from these papers.

The paper is organized as follows. Section 2 carefully formulates both the McKean–Vlasov and $n$-state control problems, stating all of the main assumptions and results. The remaining sections are devoted to the proofs. Section 3 derives some preliminary estimates on the state processes, which are put to use in section 4 to prove the main existence theorems. The proofs of the main limit theorems comprise sections 5 and 6. Finally, section 7 contains the proof (of Theorem 2.4) that the optimal value of the control problem is the same for the usual strong formulation and for our preferred relaxed formulation, under suitable assumptions.

**2. Model setup and main results.** For a metric space $E$, let $\mathcal{P}(E)$ denote the set of Borel probability measures on $E$, and endow $\mathcal{P}(E)$ with the topology of weak convergence. Fix $p \geq 1$ throughout the paper. For a complete separable metric space $(E, d)$, let $\mathcal{P}^p(E)$ denote the set of $\mu \in \mathcal{P}(E)$ with $\int_E d(x, x_0)^p \mu(dx) < \infty$ for some $x_0 \in E$. Endow $\mathcal{P}^p(E)$ with the $p$-Wasserstein metric,

$$(2.1) \qquad \ell_{E,p}(\mu, \nu) = \inf \left\{ \int d^p \, d\pi : \pi \in \mathcal{P}(E \times E) \text{ has marginals } \mu \text{ and } \nu \right\}.$$

As is well known (cf. [52, Theorem 7.12]), $\ell_{E,p}(\mu_n, \mu) \to 0$ if and only if $\int \varphi \, d\mu_n \to \int \varphi \, d\mu$ for every continuous function $\varphi$ satisfying $|\varphi(x)| \leq c(1 + d(x, x_0)^p)$ for all $x \in E$, for some $c \geq 0$. The Borel $\sigma$-field of $\mathcal{P}^p(E)$ is the same as the one induced by the Borel $\sigma$-field of $\mathcal{P}(E)$, which is in turn equivalent to the $\sigma$-field induced by the evaluations $\mathcal{P}^p(E) \ni \mu \mapsto \mu(C)$ for Borel sets $C \subset E$. For our purposes, the most pertinent topological properties of $\mathcal{P}^p(E)$ are summarized in the appendix of [39], but see also [52, Chapter 7] for more details.

A time horizon $T > 0$ is fixed throughout, along with three exponents $(p', p, p_\sigma)$, an initial distribution $\lambda \in \mathcal{P}(\mathbb{R}^d)$, and functions

$$(b, \sigma, f) : [0, T] \times \mathbb{R}^d \times \mathcal{P}^p\left(\mathbb{R}^d\right) \times A \to \mathbb{R}^d \times \mathbb{R}^{d \times d_W} \times \mathbb{R},$$
$$g : \mathbb{R}^d \times \mathcal{P}^p\left(\mathbb{R}^d\right) \to \mathbb{R}.$$

Here $d$ and $d_W$ denote the respective dimensions of the state and noise processes. The following standing assumptions, heavily inspired by [39], are in force throughout the paper.

*Assumption* A.
(A.1) $A$ is a closed subset of a Euclidean space.
(A.2) The exponents satisfy $p' > p \geq 1 \vee p_\sigma$ and $p' \geq 2 \geq p_\sigma \geq 0$ and also $\lambda \in \mathcal{P}^{p'}(\mathbb{R}^d)$.
(A.3) The functions $b$ and $\sigma$ are jointly continuous, and $f$ and $g$ are upper semicontinuous.
(A.4) There exists $c_1 > 0$ such that, for all $(t, x, m, a)$,

$$|b(t, x, m, a)| \leq c_1 \left[ 1 + |x| + \left( \int_{\mathbb{R}^d} |z|^p m(dz) \right)^{1/p} + |a| \right],$$
$$|\sigma(t, x, m, a)|^2 \leq c_1 \left[ 1 + |x|^{p_\sigma} + \left( \int_{\mathbb{R}^d} |z|^p m(dz) \right)^{p_\sigma/p} + |a|^{p_\sigma} \right].$$

(A.5) There exist $c_2, c_3 > 0$ such that, for each $(t, x, m, a)$,

$$g(x, m) \leq c_2 \left( 1 + |x|^p + \int_{\mathbb{R}^d} |z|^p m(dz) \right),$$

$$g(x, m) \geq -c_2 \left( 1 + |x|^{p'} + \int_{\mathbb{R}^d} |z|^{p'} m(dz) \right),$$

$$f(t, x, m, a) \leq c_2 \left( 1 + |x|^p + \int_{\mathbb{R}^d} |z|^p m(dz) \right) - c_3 |a|^{p'},$$

$$f(t, x, m, a) \geq -c_2 \left( 1 + |x|^{p'} + \int_{\mathbb{R}^d} |z|^{p'} m(dz) + |a|^{p'} \right).$$

These minimal assumptions will suffice for an existence theorem. The least innocuous of these is the coercivity Assumption (A.5) on the running objective $f$, which is crucial for compactness purposes (see Lemmas 3.2 and 3.4). It should be noted that our methods also apply to a compact control space $A$ and bounded coefficients $(b, \sigma, f, g)$, continuous with respect to weak convergence, and in fact the proofs become significantly simpler in this case. The limit theorems require an additional assumption, mainly for providing uniqueness of the controlled McKean–Vlasov equations.

*Assumption* B. There exists $c_1' > 0$ such that, for all $(t, x, x', m, m', a)$,

$$|b(t, x, m, a) - b(t, x', m', a)| + |\sigma(t, x, m, a) - \sigma(t, x', m', a)|$$
$$\leq c_1' \left( |x - x'| + \ell_{\mathbb{R}^d, p}(m, m') \right).$$

Moreover, the functions $f$ and $g$ are continuous.

**2.1. Relaxed controls and canonical spaces.** The space $\mathcal{V}$ of *relaxed controls* is defined as the set of measures $q$ on $[0, T] \times A$ with first marginal equal to Lebesgue measure and with

$$\int_{[0,T] \times A} |a|^p q(dt, da) < \infty.$$

Noting that each $q \in \mathcal{V}$ has total mass $T$, we may endow $\mathcal{V}$ with a suitable scaling of the $p$-Wasserstein distance. More precisely, equip $\mathcal{V}$ with the metric

$$(2.2) \qquad d_{\mathcal{V}} \left( q^1, q^2 \right) = \ell_{[0,T] \times A, p} \left( q^1/T, q^2/T \right),$$

where $\ell$ is the Wasserstein distance defined in (2.1) relative to the metric on $[0, T] \times A$ given by $((t, a), (t', a')) \mapsto |t - t'| + |a - a'|$. Each $q \in \mathcal{V}$ is identified with a measurable function $[0, T] \ni t \mapsto q_t \in \mathcal{P}^p(A)$, defined uniquely up to almost sure equality by $q(dt, da) = dt q_t(da)$. Note that $\mathcal{V}$ is a Polish space because $A$ is. A relaxed control of the form $q(dt, da) = dt \delta_{\alpha_t}(da)$ for some measurable function $[0, T] \ni t \mapsto \alpha_t \in A$ is called a *strict control*. It is known that there exists a version of the map $[0, T] \times \mathcal{V} \ni (t, q) \mapsto q_t \in \mathcal{P}^p(A)$ which is predictable with respect to the filtration $\mathbb{F}^\Lambda = (\mathcal{F}_t^\Lambda)_{t \in [0,T]}$, where $\mathcal{F}_t^\Lambda$ is generated by the maps $q \mapsto q([0, s] \times C)$, where $s \leq t$ and $C \subset A$ is Borel (see, e.g., [39, Lemma 3.2]). In particular, this lets us freely identify any random element $\Lambda$ of $\mathcal{V}$ with a corresponding $\mathcal{P}^p(A)$-valued $\mathbb{F}^\Lambda$-predictable process $(\Lambda_t)_{t \in [0,T]}$.

We will work also with the path space $\mathcal{C}^d = C([0, T]; \mathbb{R}^d)$, equipped with the supremum norm $\|x\| = \sup_{t \in [0,T]} |x_t|$. For $m \in \mathcal{P}(\mathcal{C}^d \times \mathcal{V})$, let $m^x$ denote the $\mathcal{C}^d$-marginal. For $m^x \in \mathcal{P}(\mathcal{C}^d)$ and $t \in [0, T]$, let $m_t^x \in \mathcal{P}(\mathbb{R}^d)$ denote the time-$t$ marginal,

i.e., the image of $m^x$ under the map $x \mapsto x_t$. Equip $\mathcal{C}^d \times \mathcal{V}$ with the metric

$$(2.3) \qquad d_{\mathcal{C}^d \times \mathcal{V}}((x, q), (x', q')) = \|x - x'\| + d_{\mathcal{V}}(q, q'),$$

where $d_{\mathcal{V}}$ was defined in (2.2), and $\|\cdot\|$ is the supremum norm on $\mathcal{C}^d$. Then $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$ is endowed with the corresponding Wasserstein metric $\ell_{\mathcal{C}^d \times \mathcal{V}, p}$ induced by $d_{\mathcal{C}^d \times \mathcal{V}}$. We will state our main limit theorems in terms of $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$, which is equipped with the Wasserstein metric $\ell_{\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}), p}$ induced by equipping $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$ with the metric $\ell_{\mathcal{C}^d \times \mathcal{V}, p}$. See again the appendix of [39] for a more detailed discussion of this topology. For now, simply note that if $P_n \to P$ in $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$, then a fortiori $P_n \to P$ in $\mathcal{P}(\mathcal{P}(\mathcal{C}^d \times \mathcal{V}))$ (i.e., weakly).

**2.2. The mean field control problem.** We begin by describing the strong form of the McKean–Vlasov control problem. Suppose we are given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ supporting a $d_W$-dimensional $\mathbb{F}$-Wiener process $W$ and an $\mathcal{F}_0$-measurable $\mathbb{R}^d$-valued random variable $\xi$. Here $\mathbb{F}$ is the augmented filtration generated by the initial state and Wiener process, i.e., the (right-continuous) completion of $(\sigma(\xi, W_s : s \leq t))_{t \geq 0}$. An $\mathbb{F}$-progressively measurable $A$-valued process $\alpha$ is called an *admissible control* if it satisfies

$$\mathbb{E} \int_0^T |\alpha_t|^p dt < \infty$$

and if there exists a unique square-integrable strong solution on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ of the McKean–Vlasov SDE

$$dX_t = b\left(t, X_t, \mathbb{P} \circ X_t^{-1}, \alpha_t\right) dt + \sigma\left(t, X_t, \mathbb{P} \circ X_t^{-1}, \alpha_t\right) dW_t, \quad X_0 = \xi.$$

The strong form of the McKean–Vlasov control problem is to maximize

$$\mathbb{E}\left[\int_0^T f\left(t, X_t, \mathbb{P} \circ X_t^{-1}, \alpha_t\right) dt + g\left(X_T, \mathbb{P} \circ X_T^{-1}\right)\right]$$

over all admissible controls. Note that an admissible control induces a probability measure $\mathbb{P} \circ (X, dt\delta_{\alpha_t}(da))^{-1}$ on $\mathcal{C}^d \times \mathcal{V}$. Let $\mathcal{R}^s$ denote the set of such measures, and refer to an element of $\mathcal{R}^s$ as a *strong control*. The definition of $\mathcal{R}^s$ is insensitive to the choice of probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, provided that it satisfies the above requirements. Hence, we make no further reference to this particular $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

We next describe the relaxed form of the control problem, abandoning the probability space of the previous paragraph. Let $(X, \Lambda)$ denote the projection maps or canonical processes on $\mathcal{C}^d \times \mathcal{V}$. As in section 2.1, we may conflate the random measure $\Lambda(dt, da)$ and the $\mathcal{P}^p(A)$-valued process $(\Lambda_t)_{t \in [0,T]}$. The space $\mathcal{C}^d \times \mathcal{V}$ is equipped with the filtration generated by these canonical processes $(X_t, \Lambda_t)_{t \in [0,T]}$. Define the generator $\mathcal{L}$ to act on smooth compactly supported functions $\varphi$ by

$$(2.4) \qquad \mathcal{L}\varphi(t, x, m, a) = b(t, x, m, a) \cdot \nabla\varphi(x) + \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^\top(t, x, m, a)\nabla^2\varphi(x)\right],$$

where $\nabla$ and $\nabla^2$ denote the gradient and Hessian, respectively. That is, $\mathcal{L}$ maps a function of $\mathbb{R}^d$ to a function of $[0, T] \times \mathbb{R}^d \times \mathcal{P}^p(\mathbb{R}^d) \times A$. Define $\mathcal{R}$ to be the set of $m \in \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$ such that $m_0^x = \lambda$ (recalling that $m_t^x := m \circ X_t^{-1}$) and the process

$$\varphi(X_t) - \int_0^t \int_A \mathcal{L}\varphi\left(s, X_s, m_s^x, a\right) \Lambda_s(da)dt$$

is an $m$-martingale for every smooth compactly supported $\varphi$. It is a straightforward consequence of Itô's formula that $\mathcal{R}^s \subset \mathcal{R}$; that is, a strong control induces a relaxed control. An element of $\mathcal{R}$ is called a *control*, or a *relaxed control* for emphasis. We say a control $m \in \mathcal{R}$ is *Markovian* if there exists a measurable map $\hat{\alpha} : [0, T] \times \mathbb{R}^d \to A$ such that $m(\Lambda_t = \delta_{\hat{\alpha}(t, X_t)}, \ a.e. \ t \in [0, T]) = 1$. See Remark 2.6 below for a caveat regarding this use of the term *Markovian*.

*Remark* 2.1. Under Assumption A, the set $\mathcal{R}$ is nonempty. In particular, for any fixed constant control $a_0 \in A$, there exists $m \in \mathcal{R}$ such that $m(\Lambda_t = \delta_{a_0}, \ \text{a.e.} \ t \in [0, T]) = 1$. Indeed, this follows from the results of Gärtner [22, Theorem 2.9] or Funaki [21, Theorem 2.1] on McKean–Vlasov equations. When both Assumptions A and B hold, $\mathcal{R}^s$ is also nonempty, by standard Lipschitz arguments (see [51, section I] or the proof of [21, Lemma 3.1]).

The McKean–Vlasov control problem, in relaxed form, is to maximize

$$(2.5) \qquad \Gamma(m) := \mathbb{E}^m \left[ \int_0^T \int_A f\left(t, X_t, m_t^x, a\right) \Lambda_t(da) dt + g\left(X_T, m_T^x\right) \right]$$

over all choices of $m \in \mathcal{R}$. Note that $\Gamma : \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}) \to \mathbb{R} \cup \{-\infty\}$ is well-defined because of Assumption (A.5). Let $\mathcal{R}^*$ denote the set of optimal controls, i.e., the set of $m \in \mathcal{R}$ for which $\Gamma(m) \geq \Gamma(\tilde{m})$ for all $\tilde{m} \in \mathcal{R}$. We are now ready to state the main existence results, with proofs deferred to sections 6.

THEOREM 2.2. *Under Assumption* A, *the set* $\mathcal{R}^*$ *is nonempty. That is, there exists an optimal relaxed control.*

Next, we state an existence result for Markovian optimal controls, under an additional assumption, familiar in the control theory literature from the work of Filippov [16].

*Assumption* C. For each $(t, x, m) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}^p(\mathbb{R}^d)$, the following set is convex:

$$K(t, x, \mu) := \left\{ \left(b(t, x, m, a), \sigma\sigma^\top(t, x, m, a), z\right) : a \in A, \ z \leq f(t, x, m, a) \right\}$$
$$\subset \mathbb{R}^d \times \mathbb{R}^{d \times d} \times \mathbb{R}.$$

A notable class of examples satisfying Assumption C is when $A$ is a convex set and, for each fixed $(t, x, m)$, the functions $b(t, x, m, a)$ and $\sigma\sigma^\top(t, x, m, a)$ are affine in $a$ and $f(t, x, m, a)$ is concave in $a$. In this case, Assumption C follows from the fact that the hypograph of a concave function is a convex set.

THEOREM 2.3. *Suppose Assumptions* A *and* C *hold. Then, for each* $m \in \mathcal{R}$, *there exists a Markovian control* $\tilde{m} \in \mathcal{R}$ *satisfying* $\tilde{m}_t^x = m_t^x$ *for all* $t \in [0, T]$ *as well as* $\Gamma(\tilde{m}) \geq \Gamma(m)$. *In particular, there exists an optimal Markovian control.*

Theorem 2.3 illustrates that the relaxation of the McKean–Vlasov control problem is not too drastic. One might argue that it would be more natural to use closed-loop or Markovian controls as opposed to $\mathbb{F}$-progressively measurable controls in our formulation of the control problem. But Theorem 2.3 shows that, under suitable assumptions, this choice of information structure does not make a difference in terms of the optimal value. Moreover, a Markovian control exists which is optimal among all progressively measurable or even relaxed controls.

On a related note, the next theorem states that the optimal value of the strong and relaxed formulations are the same. The additional assumptions are minor and

can likely be dispensed with. The proof, deferred to section 7, requires some delicate approximations of martingale measures, for which we employ a result of Méléard [42]. An alternative proof is possible under less easily verifiable weak uniqueness assumptions, by adapting the methods of [32, 34].

THEOREM 2.4. *Suppose Assumptions* A *and* B *hold. Assume also that* $p' = 2$ *and that the initial condition* $\lambda$ *satisfies* $\int |x|^{p''} \lambda(dx) < \infty$ *for some* $p'' > p'$. *Then the optimal value of the McKean–Vlasov control problem is the same for both the relaxed and strong formulations. That is,*

$$\sup_{m \in \mathcal{R}^s} \Gamma(m) = \sup_{m \in \mathcal{R}} \Gamma(m).$$

Given the unusual nature of the martingale problems defining $\mathcal{R}$, it may be unclear to the uninitiated reader what exactly we have proved to exist in Theorems 2.2 and 2.3. This is clarified by the following proposition, which will be useful in the proofs as well. When the control is present in the volatility, the most useful SDE representation involves martingale measures. Only the very basics of the theory of martingale measures are needed, and these facts are reviewed as they come up. All of the relevant definitions and results are concisely summarized in [31], but refer to the original monograph of Walsh [53] for a more thorough treatment. When the control is absent from the volatility $\sigma$, the martingale measure $N(da, dt)$ in the following proposition can be replaced with a Wiener process $dW_t$.

PROPOSITION 2.5 (see [31, Theorem IV-2]). *The set* $\mathcal{R}$ *is precisely the set of laws* $\mathbb{P} \circ (X, \Lambda)^{-1}$, *where the following hold:*
   1. $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *is a filtered probability space supporting a* $d$-*dimensional adapted process* $X$, *a* $\mathcal{P}^p(A)$-*valued predictable process* $\Lambda$, *and a (column) vector* $N = (N^1, \ldots, N^{d_W})$ *of orthogonal* $\mathbb{F}$-*martingale measures on* $A \times [0, T]$, *each with intensity measure* $\Lambda_t(da)dt$.
   2. $\mathbb{P} \circ X_0^{-1} = \lambda$.
   3. $\mathbb{E}[\sup_{t \in [0,T]} |X_t|^p + \int_0^T \int_A |a|^p \Lambda_t(da)dt] < \infty$.
   4. *The state equation holds:*

$$dX_t = \int_A b\left(t, X_t, \mathbb{P} \circ X_t^{-1}, a\right) \Lambda_t(da)dt + \int_A \sigma\left(t, X_t, \mathbb{P} \circ X_t^{-1}, a\right) N(da, dt).$$

*The set of Markovian controls is precisely the set of laws* $\mathbb{P} \circ (X, \Lambda)^{-1}$, *where the following hold:*
   1. $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *is a filtered probability space supporting a* $d$-*dimensional adapted process* $X$ *and a* $d_W$-*dimensional* $\mathbb{F}$-*Wiener process* $W$.
   2. $\mathbb{P} \circ X_0^{-1} = \lambda$.
   3. *There exists a measurable function* $\hat{\alpha} : [0, T] \times \mathbb{R}^d \to A$ *such that the following hold:*
      (a) *The state equation holds:*

(2.6)
$$dX_t = b\left(t, X_t, \mathbb{P} \circ X_t^{-1}, \hat{\alpha}(t, X_t)\right) dt + \sigma\left(t, X_t, \mathbb{P} \circ X_t^{-1}, \hat{\alpha}(t, X_t)\right) dW_t.$$

      (b) $\mathbb{E}[\sup_{t \in [0,T]} |X_t|^p + \int_0^T |\hat{\alpha}(t, X_t)|^p dt] < \infty$.
      (c) $\Lambda_t = \delta_{\hat{\alpha}(t, X_t)}$ *for a.e.* $t \in [0, T]$, *almost surely.*

*Remark* 2.6. A word of caution about the terminology: Under a Markovian control $m \in \mathcal{R}$, it is not true in general that the state process $X$ is a Markov process.

Only when the state equation (2.6) is sufficiently well-posed is $X$ truly a Markov process. For instance, letting $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, X, \hat{\alpha})$ be as in the second part of Proposition 2.5, define new coefficients $\hat{b}(t,x) = b(t, x, \mathbb{P} \circ X_t^{-1}, \hat{\alpha}(t,x))$ and $\hat{\sigma}(t,x) = \sigma(t, x, \mathbb{P} \circ X_t^{-1}, \hat{\alpha}(t,x))$. If the martingale problem associated to $(\hat{b}, \hat{\sigma})$ is well-posed, then the unique in law weak solution $X$ is a Feller process [50, Chapter 12].

**2.3. $n$-state control problems.** This section states the main results on how the McKean–Vlasov control problem arises from $n$-state control problems as $n \to \infty$. Assume throughout this section that both Assumptions A and B are in force.

We first describe a standard *strong formulation* of the control problems. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space supporting independent $d_W$-dimensional $\mathbb{F}$-Wiener processes $W^1, \ldots, W^n$ as well as i.i.d. $\mathcal{F}_0$-measurable $\mathbb{R}^d$-valued random variables $\xi^1, \ldots, \xi^n$ with law $\lambda$. Assume the filtration $\mathbb{F}$ is generated by these initial states and Wiener processes, i.e., the (right-continuous) completion of $(\sigma(\xi^1, \ldots, \xi^n, W_s^1, \ldots, W_s^n : s \le t))_{t \ge 0}$. An $\mathbb{F}$-progressively measurable $A^n$-valued process $(\alpha^1, \ldots, \alpha^n)$ is called an *admissible control* if

$$\mathbb{E} \int_0^T |\alpha_t^k|^p dt < \infty, \ \ k = 1, \ldots, n,$$

and if there exists a unique square-integrable strong solution on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ of the SDE system

$$dX_t^i = b(t, X_t^i, \widehat{\mu}_t^n, \alpha_t^i)dt + \sigma(t, X_t^i, \widehat{\mu}_t^n, \alpha_t^i)dW_t^i,$$
$$\widehat{\mu}_t^n = \frac{1}{n} \sum_{k=1}^n \delta_{X_t^k}.$$

The strong form of the $n$-state control problem is to maximize

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ \int_0^T f(t, X_t^i, \widehat{\mu}_t^n, \alpha_t^i)dt + g(X_T^i, \widehat{\mu}_T^n) \right]$$

over all admissible controls. Note that an admissible control induces a probability measure $\mathbb{P} \circ ((X^i, dt\delta_{\alpha_t^i}(da))_{i=1}^n)^{-1}$ on $(\mathcal{C}^d \times \mathcal{V})^n$. Let $\mathcal{R}_n^s$ denote the set of such measures, and refer to an element of $\mathcal{R}_n^s$ as a *strong control*. As in the previous section, the definition of $\mathcal{R}_n^s$ is insensitive to the choice of probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, provided that it satisfies the above requirements. Hence, we make no further reference to this particular $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

The relaxed form of the $n$-state control problem is defined by working with relaxed controls and weak solutions of the SDEs. Let $(X^i, \Lambda^i)_{i=1}^n$ denote the canonical process on $(\mathcal{C}^d \times \mathcal{V})^n$. Define the empirical measures

$$(2.7) \qquad \widehat{\mu}^n = \frac{1}{n} \sum_{k=1}^n \delta_{(X^k, \Lambda^k)}, \qquad \qquad \widehat{\mu}_t^{n,x} = \frac{1}{n} \sum_{k=1}^n \delta_{X_t^k}, \ \ t \in [0, T].$$

Define $\mathcal{R}_n$ as the set of laws $P \in \mathcal{P}((\mathcal{C}^d \times \mathcal{V})^n)$ under which $(X_0^i)_{i=1}^n$ are i.i.d. with law $\lambda$, and

$$\varphi\left(X_t^1, \ldots, X_t^n\right) - \sum_{i=1}^n \int_0^t \int_A \mathcal{L}_i^n \varphi\left(s, X_s^1, \ldots, X_s^n, a\right) \Lambda_t^i(da)dt$$

is a martingale, where the generator $\mathcal{L}_i^n$ acts on smooth compactly supported functions $\varphi : (\mathbb{R}^d)^n \to \mathbb{R}$ by

$$\mathcal{L}_i^n \varphi(t, x_1, \dots, x_n, a) = b\left(t, x_i, \frac{1}{n}\sum_{k=1}^n \delta_{x_k}, a\right) \cdot \nabla_i \varphi(x_1, \dots, x_n)$$

(2.8)
$$+ \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^\top\left(t, x_i, \frac{1}{n}\sum_{k=1}^n \delta_{x_k}, a\right)\nabla_i^2\varphi(x_1, \dots, x_n)\right],$$

where $\nabla_i$ and $\nabla_i^2$ denote the gradient and Hessian with respect to the $i^{\text{th}}$ variable. An element of $\mathcal{R}_n$ is called a *control*, or an *(n-state) relaxed control* for emphasis. It is a straightforward consequence of Itô's formula that $\mathcal{R}_n^s \subset \mathcal{R}_n$; that is, a strong control induces a relaxed control. We have the following analog of Proposition 2.5 and again note that there is no need for martingale measures when $\sigma$ is uncontrolled.

PROPOSITION 2.7 (see [31, Theorem IV-2]).   *The set $\mathcal{R}_n$ equals the set of laws* $\mathbb{P} \circ ((X^i, \Lambda^i)_{i=1}^n)^{-1}$, *where the following hold:*
  1. *$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space supporting $n$ adapted $d$-dimensional processes $X^1, \dots, X^n$, $n$ predictable $\mathcal{P}^p(A)$-valued processes $\Lambda^1, \dots, \Lambda^n$, and $nd_W$ orthogonal $\mathbb{F}$-martingale measures $(N^{i,j})$ for $i = 1, \dots, n$ and $j = 1, \dots, d_W$, where $N^{i,j}$ has intensity $\Lambda_t^i(da)dt$ for each $j$.*
  2. *$X_0^1, \dots, X_0^n$ are i.i.d. with law $\lambda$.*
  3. *$\mathbb{E}[\sup_{t\in[0,T]}|X_t^i|^p + \int_0^T \int_A |a|^p \Lambda_t^i(da)dt] < \infty$ for each $i = 1, \dots, n$.*
  4. *The state equation holds, with $N^i = (N^{i,1}, \dots, N^{i,d_W})^\top$:*

$$dX_t^i = \int_A b(t, X_t^i, \widehat{\mu}_t^{n,x}, a)\Lambda_t^i(da)dt + \int_A \sigma(t, X_t^i, \widehat{\mu}_t^{n,x}, a)N^i(da, dt),$$

$$\widehat{\mu}_t^{n,x} = \frac{1}{n}\sum_{k=1}^n \delta_{X_t^k}.$$

*Remark* 2.8. Under Assumption A, the set $\mathcal{R}_n$ is nonempty for each $n$. In particular, for any fixed constant control $(a_0^1, \dots, a_0^n) \in A^n$, there exists $P \in \mathcal{R}_n$ such that, almost surely under $P_n$, $\Lambda_t^k = \delta_{a_0^k}$ for a.e. $t \in [0, T]$ and for each $k = 1, \dots, n$. Indeed, this follows essentially from the classical existence results for martingale problems of Stroock and Varadhan [50].

Recalling the definition of $\Gamma$ from (2.5), the reward of $P \in \mathcal{R}_n$ is given by

$$\mathbb{E}^P[\Gamma(\widehat{\mu}^n)] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[\int_0^T \int_A f(t, X_t^i, \widehat{\mu}_t^{n,x}, a)\Lambda_t^i(da)dt + g(X_T^i, \widehat{\mu}_T^{n,x})\right].$$

Note that $\mathbb{E}^P[\Gamma(\widehat{\mu}^n)]$ is well-defined in $[-\infty, \infty)$ for $P \in \mathcal{P}^p((\mathcal{C}^d \times \mathcal{V})^n)$, thanks to Assumption (A.5). Given $\epsilon \geq 0$, we say $P \in \mathcal{R}_n$ is a *(relaxed) n-state $\epsilon$-optimal control* if

$$\mathbb{E}^P[\Gamma(\widehat{\mu}^n)] \geq \sup_{Q \in \mathcal{R}_n} \mathbb{E}^Q[\Gamma(\widehat{\mu}^n)] - \epsilon.$$

If $\epsilon = 0$, we simply say $P$ is a *(relaxed) n-state optimal control*. The following result, at the present level of generality, is due to Haussmann and Lepeltier.

THEOREM 2.9 (see [25, Theorem 4.7]).   *Under Assumption* A, *for each $n$, there exists an optimal $n$-state relaxed control.*

In many cases, the optimal value of the relaxed control problem is the same as that of the strong formulation. Theorem 2.10 below is a step in this direction, analogous to Theorem 2.4. It is nearly a special case of the results of [32, section 4] when $(b, \sigma, f, g)$ are bounded.

THEOREM 2.10. *Suppose Assumptions* A *and* B *hold with* $p' = 2$. *Assume the initial condition* $\lambda$ *belongs to* $\mathcal{P}^{p''}(\mathbb{R}^d)$ *for some* $p'' > 2$. *Then, for every relaxed control* $P \in \mathcal{R}_n$, *there exists a sequence of strong controls* $P_n \in \mathcal{R}_n^s$ *with* $P_n \to P$ *in* $\mathcal{P}^p((\mathcal{C}^d \times \mathcal{V})^n)$ *and* $\mathbb{E}^{P_n}[\Gamma(\hat{\mu}^n)] \to \mathbb{E}^P[\Gamma(\hat{\mu}^n)]$. *In particular, the optimal value of the n-state control problem is the same for both the relaxed and strong formulations; that is,*

$$\sup_{P \in \mathcal{R}_n} \mathbb{E}^P[\Gamma(\widehat{\mu}^n)] = \sup_{P \in \mathcal{R}_n^s} \mathbb{E}^P[\Gamma(\widehat{\mu}^n)].$$

An additional hypothesis like Assumption C ensures that the $n$-state control problem admits an optimal relaxed control which is Markovian, analogous to Theorem 2.3. See [25, Theorem 5.15] or [32, Theorem 6.7 and Corollary 6.8]. We avoid making this claim any more precise, as we will have no use for it. But we merely wish to note that the decision to work with open-loop controls (i.e., $\mathbb{F}$-progressively measurable controls) as opposed to Markovian controls in the $n$-state control problem is often inconsequential.

**2.4. The main limit theorems.** Now that we understand the structure of the $n$-state and McKean–Vlasov control problems, we are ready to state the main results of the paper. Refer to section 2.1 for a discussion of convergence in the space $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$. Recall that $\mathcal{R}^* \subset \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$ denotes the set of (relaxed) optimal McKean–Vlasov controls. The first theorem states that $n$-state optimal controls converge to optimal McKean–Vlasov controls, while the second asserts that every optimal McKean–Vlasov control can be obtained as the limit of approximately optimal controls for the $n$-state system. It is worth noting that the $n$-state controls in Theorem 2.11 are not required to be symmetric in any sense.

THEOREM 2.11. *Suppose Assumptions* A *and* B *hold. For each n, let* $P_n \in \mathcal{R}_n$ *be a relaxed n-state* $\epsilon_n$*-optimal control for some sequence* $\epsilon_n \to 0$. *Then* $(P_n \circ (\widehat{\mu}^n)^{-1})$ *is precompact in* $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$, *and every limit is supported on* $\mathcal{R}^*$. *In particular,*

$$(2.9) \qquad \lim_{n \to \infty} P_n\left(\ell_{\mathcal{C}^d \times \mathcal{V}, p}(\widehat{\mu}^n, \mathcal{R}^*) \geq \epsilon\right) = 0 \quad \text{for all } \epsilon > 0,$$

*where* $\ell_{\mathcal{C}^d \times \mathcal{V}, p}(m, \mathcal{R}^*) := \inf_{m' \in \mathcal{R}^*} \ell_{\mathcal{C}^d \times \mathcal{V}, p}(m, m')$ *for* $m \in \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$.

THEOREM 2.12. *Suppose Assumptions* A *and* B *hold. Let* $P \in \mathcal{P}(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$ *be supported on* $\mathcal{R}^*$. *Then there exist* $\epsilon_n \to 0$ *and a sequence of relaxed n-state* $\epsilon_n$*-optimal controls* $P_n \in \mathcal{R}_n$ *such that* $P_n \circ (\widehat{\mu}^n)^{-1} \to P$. *Moreover, if* $p' = 2$ *and if* $\lambda \in \mathcal{P}^{p''}(\mathbb{R}^d)$ *for some* $p'' > 2$, *then the controls* $P_n$ *can be taken to be strong.*

If it happens that the optimal McKean–Vlasov control is unique, then an immediate corollary of Theorem 2.11 is a proper convergence theorem, stating that (the empirical measure sequence of) any $n$-state near-optimal $n$-state controls converge in probability to the unique optimal McKean–Vlasov control. An immediate corollary of Theorems 2.11 and 2.3 is the following statement, for which $C([0, T]; \mathcal{P}^p(\mathbb{R}^d))$ is endowed with the metric $(\mu, \nu) \mapsto \sup_{t \in [0, T]} \ell_{\mathbb{R}^d, p}(\mu_t, \nu_t)$.

COROLLARY 2.13. *Suppose Assumptions* A, B, *and* C *hold. For each n, let* $P_n$ *denote a weak n-state* $\epsilon_n$*-optimal control for some sequence* $\epsilon_n \to 0$. *Then* $(P_n \circ$

$(\widehat{\mu}_t^{n,x})_{t\in[0,T]}^{-1})$ is precompact in $\mathcal{P}^p(C([0,T];\mathcal{P}^p(\mathbb{R}^d)))$, and every weak limit is supported on the set $\{(m_t^x)_{t\in[0,T]} : m \in \mathcal{R}^*$ is Markovian$\}$.

The rest of the paper is devoted to the proofs. Section 3 is devoted to some useful preliminary results, including moment estimates on the state process and some continuity properties of the objective functional $\Gamma$. Section 4 proves the existence Theorems 2.2 and 2.3. Most of the work toward Theorems 2.11 and 2.12 is done in the preparatory section 5, with the main line of the proofs deferred to section 6. Finally, section 7 discusses Theorems 2.4 and 2.10.

**3. Some first estimates.** This section collects the essential estimates needed in the proofs of almost all of the theorems of the previous section. The first estimates of section 3.1 are in the mean field regime, whereas the estimates of section 3.2 pertain to the $n$-state games. The primary role of these estimates is in obtaining compactness. If the control space $A$ were assumed compact, and if the coefficients $b$, $\sigma$, $f$, and $g$ were assumed to be bounded, none of these estimates would be needed. In the following, let $\|x\|_t = \sup_{s\in[0,t]}|x_s|$ denote the truncated supremum norm for $x \in \mathcal{C}^d$, for $t \in [0,T]$, and recall that $\|\cdot\| = \|\cdot\|_T$.

**3.1. Mean field estimates.** The first lemma, stated without proof, is borrowed from [39]. The second shows how to use the coercivity Assumption (A.5) to translate optimality properties into moment bounds.

LEMMA 3.1 (see [39, Lemma 4.3]). *For each $\gamma \geq p$ such that $\int |x|^\gamma \lambda(dx) < \infty$, there exists a constant $C \geq 1$, depending only on $\gamma$, $p$, $p'$, $T$, the initial law $\lambda$, and the constant $c_1$ of Assumption (A.4) such that for all $m \in \mathcal{R}$ we have*

$$\int_{\mathcal{C}^d} \|x\|^p m^x(dx) = \mathbb{E}^m\left[\|X\|^\gamma\right] \leq C\mathbb{E}^m\left[1 + \int_0^T \int_A |a|^\gamma \Lambda_t(da)dt\right].$$

LEMMA 3.2. *Let $\epsilon > 0$. Suppose $\mathcal{R}^\epsilon$ is the set of $m \in \mathcal{R}$ satisfying*

$$\Gamma(m) \geq \sup_{m'\in\mathcal{R}} \Gamma(m') - \epsilon.$$

*Then*

(3.1) $$\sup_{m\in\mathcal{R}^\epsilon} \mathbb{E}^m \int_0^T \int_A |a|^{p'} \Lambda_t(da)dt < \infty.$$

*Moreover, $\sup_{m\in\mathcal{R}} \Gamma(m) < \infty$.*

*Proof.* Recall first that $\mathbb{E}^m \int_0^T \int_A |a|^p \Lambda_t(da)dt < \infty$ for all $m \in \mathcal{R}$ by assumption; this ensures that the following expressions are well-defined. Use the upper bounds on $f$ and $g$ from Assumption (A.5) along with Lemma 3.1 to find a constant $C > 0$ (which will change from line to line) such that, for all $m \in \mathcal{R}$,

$$\Gamma(m) \leq C\mathbb{E}^m\left[1 + \|X\|^p + \int_{\mathcal{C}^d} \|x\|^p m^x(dx)\right] - c_3\mathbb{E}^m \int_0^T \int_A |a|^{p'} \Lambda_t(da)dt$$

(3.2) $$\leq C\mathbb{E}^m\left[1 + \int_0^T \int_A |a|^p \Lambda_t(da)dt\right] - c_3\mathbb{E}^m \int_0^T \int_A |a|^{p'} \Lambda_t(da)dt.$$

This already proves $\sup_{m\in\mathcal{R}} \Gamma(m) < \infty$, as $a \mapsto C|a|^p - c_3|a|^{p'}$ is bounded from above. To prove the first claim, fix arbitrarily a constant control $a_0 \in A$, and let $m^0$ denote

an element of $\mathcal{R}$ satisfying $m^0(\Lambda_t = \delta_{a_0}, \ a.e. \ t) = 1$. (As in Remark 2.1, the existence of such an $m^0$ follows from a result of Gärtner [22, Theorem 2.9].) Lemma 3.1 implies

$$\int_{\mathcal{C}^d} \|x\|^{p'} \left(m^0\right)^x (dx) = \mathbb{E}^{m^0}\left[\|X\|^{p'}\right] \leq C\left(1 + T|a_0|^{p'}\right).$$

Then use the lower bounds of Assumption (A.5) to show $\Gamma(m^0) > -\infty$. For $m \in \mathcal{R}^\epsilon$ we have $\Gamma(m) \geq \Gamma(m^0) - \epsilon$, which combined with (3.2) yields

$$\sup_{m \in \mathcal{R}^\epsilon} \mathbb{E}^m \int_0^T \int_A \left(|a|^{p'} - C|a|^p\right) \Lambda_t(da)dt < \infty.$$

This is enough to complete the proof. $\quad\square$

**3.2. $n$-state estimates.** Here we derive an analogous pair of lemmas for the $n$-state control problem after first recalling some basic facts about martingale measures, all of which can be found in both [53] and [31]. Suppose $N$ is martingale measure with intensity measure $\Lambda_t(da)dt$, where $\Lambda$ is a predictable $\mathcal{P}(A)$-valued process, and $(h_t^i)_{t \in [0,T]}$ is a predictable random function on $A$.[1] If

$$\mathbb{E}\left[\int_0^T \int_A |h_t^i(a)|^2 \Lambda_t(da)dt\right] < \infty,$$

then

$$M_t^i = \int_0^t \int_A h_s^i(a) N(da, ds)$$

is a martingale for $i = 1, 2$. The covariation between $M^1$ and $M^2$ is

$$\left[M^1, M^2\right]_t = \int_0^t \int_A h_s^1(a) h_s^2(a) \Lambda_s(da)ds,$$

assuming this integral is well-defined. In particular, the quadratic variation of $M^1$ is

$$\left[M^1, M^1\right]_t = \int_0^t \int_A |h_s^1(a)|^2 \Lambda_s(da)ds,$$

assuming the latter is finite almost surely. If $\widetilde{N}$ and $N$ are *orthogonal* martingale measures, then the martingales

$$\int_0^t \int_A h_s(a) N(da, ds) \qquad \text{and} \qquad \int_0^t \int_A \widetilde{h}_s(a) \widetilde{N}(da, ds)$$

are themselves orthogonal (i.e., the covariation is identically zero) for square-integrable $h, \widetilde{h}$.

LEMMA 3.3. *For each $\gamma \geq p$ such that $\int |x|^\gamma \lambda(dx) < \infty$, there exists a constant $C \geq 1$, depending only on $\gamma$, $p$, $p'$, $T$, the initial law $\lambda$, and the constant $c_1$ of Assumption* (A.4) *such that, for all $n \geq 1$ and $P \in \mathcal{R}_n$, we have*

$$(3.3) \quad \mathbb{E}^P[\|X^k\|^\gamma] \leq C\mathbb{E}^P\left[1 + \int_0^T \int_A |a|^\gamma \Lambda_t^k(da)dt + \frac{1}{n}\sum_{i=1}^n \int_0^T \int_A |a|^\gamma \Lambda_t^i(da)dt\right]$$

---

[1]Assuming $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is the filtered probability space in the background, a *predictable random function on $A$* is a map $h : [0,T] \times \Omega \times A \to \mathbb{R}$ which is jointly measurable with respect to the $\mathbb{F}$-predictable $\sigma$-field on $[0,T] \times \Omega$ and the Borel $\sigma$-field on $A$. We suppress $\omega$ from the notation as usual, writing $h_t(a)$ in place of $h(t, \omega, a)$.

*for each $k = 1, \ldots, n$, and*

(3.4)

$$\mathbb{E}^P \left[ \int_{\mathcal{C}^d} \|x\|^\gamma \widehat{\mu}^{n,x}(dx) \right] = \frac{1}{n} \sum_{k=1}^n \mathbb{E}^P \left[ \|X^k\|^\gamma \right] \leq C \mathbb{E}^P \left[ 1 + \frac{1}{n} \sum_{k=1}^n \int_0^T \int_A |a|^\gamma \Lambda_t^k(da)dt \right].$$

*Proof.* Fix $n$ and $P \in \mathcal{R}_n$, and use Proposition 2.7 to express $(X^1, \ldots, X^n)$ (under $P$) as the solution of an SDE driven by $(N^1, \ldots, N^n)$, where each $N^k$ is a vector of $d_W$ orthogonal martingale measures with common intensity measure $\Lambda_t^k(da)dt$. That is,

$$dX_t^k = \int_A b\left(t, X_t^k, \widehat{\mu}_t^{n,x}, a\right) \Lambda_t^k(da)dt + \int_A \sigma\left(t, X_t^k, \widehat{\mu}_t^{n,x}, a\right) N^k(da, dt).$$

Recall the above remarks on quadratic variations of stochastic integrals with respect to martingale measures. Apply the Burkholder–Davis–Gundy inequality and Assumption (A.4) to find a universal constant $C$ (which will change from line to line) such that, for each $k = 1, \ldots, n$,

(3.5)

$$\mathbb{E} \left[ \|X^k\|_t^\gamma \right]$$

$$\leq C \mathbb{E} \left[ |X_0^k|^\gamma + \left( \int_0^t \int_A |b(t, X_s^k, \widehat{\mu}_s^{n,x}, a)| \Lambda_s^k(da)ds \right)^\gamma \right]$$

$$+ C \mathbb{E} \left[ \left( \int_0^t \int_A |\sigma(t, X_s^k, \widehat{\mu}_s^{n,x}, a)|^2 \Lambda_s^k(da)ds \right)^{\gamma/2} \right]$$

$$\leq C \mathbb{E} \left\{ 1 + |X_0^k|^\gamma + \int_0^t \left[ \|X^k\|_s^\gamma + \left( \int_{\mathcal{C}^d} \|x\|_s^p \widehat{\mu}^{n,x}(dx) \right)^{\gamma/p} + \int_A |a|^\gamma \Lambda_s^k(da) \right] ds \right\}$$

$$+ C \mathbb{E} \left\{ \left[ \int_0^t \int_A \left( \|X^k\|_s^{p_\sigma} + \left( \int_{\mathcal{C}^d} \|x\|_s^p \widehat{\mu}^{n,x}(dx) \right)^{p_\sigma/p} + |a|^{p_\sigma/p} \right) \Lambda_s^k(da)ds \right]^{\gamma/2} \right\}$$

$$\leq C \mathbb{E} \left\{ 1 + \int_0^t \left[ \|X^k\|_s^\gamma + \int_{\mathcal{C}^d} \|x\|_s^\gamma \widehat{\mu}^{n,x}(dx) + \int_A |a|^\gamma \Lambda_s^k(da) \right] ds \right\}.$$

The derivation of the last line used a number of facts. First, note that $\mathbb{E}[|X_0^1|^p] = \mathbb{E}[|X_0^k|^p]$ by symmetry, and this term was subsumed in the constant $C$ in the last line. Second, Jensen's inequality yielded $(\int_{\mathcal{C}^d} \|x\|_s^p \widehat{\mu}^{n,x}(dx))^{\gamma/p} \leq \int_{\mathcal{C}^d} \|x\|_s^\gamma \widehat{\mu}^{n,x}(dx)$, because $\gamma \geq p$. Finally, to deal with the exponent of $\gamma/2$ outside of the integral, there are two cases. First, if $\gamma \geq 2$, Jensen's inequality lets us bring the $\gamma/2$ inside of the time integral, and we then use the inequality $|x|^{p_\sigma \gamma/2} \leq 1 + |x|^\gamma$ which holds because $p_\sigma \leq 2$. The other alternative is $2 > \gamma \geq p \geq 1 \vee p_\sigma$, in which case we use the inequalities $|x|^{\gamma/2} \leq 1 + |x|$ and $|x|^{p_\sigma} \leq 1 + |x|^\gamma$. With (3.5) now justified, average over $k = 1, \ldots, n$ to get

$$\mathbb{E} \left[ \int_{\mathcal{C}^d} \|x\|_t^\gamma \widehat{\mu}^{n,x}(dx) \right] \leq C \mathbb{E} \left\{ 1 + \int_0^t \left[ \int_{\mathcal{C}^d} \|x\|_s^\gamma \widehat{\mu}^{n,x}(dx) + \frac{1}{n} \sum_{k=1}^n \int_A |a|^\gamma \Lambda_s^k(da) \right] ds \right\}.$$

The second claim (3.4) now follows from Gronwall's inequality. The first claim (3.3) follows from (3.5), (3.4), and Gronwall's inequality. $\qquad\square$

LEMMA 3.4. *Let $\epsilon \geq 0$. There exists a constant $C \geq 0$ such that, for all $n$ and all weak $\epsilon$-optimal controls $P \in \mathcal{R}_n$, we have*

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}^{P}\int_{0}^{T}\int_{A}|a|^{p'}\Lambda_{t}^{k}(da)dt \leq C.$$

*Proof.* Fix $P \in \mathcal{R}_n$, and recall that $\mathbb{E}^P \int_0^T \int_A |a|^p \Lambda_t^k(da)dt < \infty$ for all $k$, as this ensures that the following expressions are well-defined. Use the upper bounds on $f$ and $g$ from Assumption (A.5) along with Lemma 3.3 to get

$$\mathbb{E}^{P}[\Gamma(\widehat{\mu}^{n})] \leq C\mathbb{E}^{P}\left[1 + \int_{\mathcal{C}^{d}}\|x\|^{p}\widehat{\mu}^{n,x}(dx)\right] - c_{3}\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}^{P}\int_{0}^{T}\int_{A}|a|^{p'}\Lambda_{t}^{k}(da)dt$$

$$\leq C\left(1 + \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}^{P}\int_{0}^{T}\int_{A}|a|^{p}\Lambda_{t}^{k}(da)dt\right)$$

$$(3.6) \qquad - c_{3}\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}^{P}\int_{0}^{T}\int_{A}|a|^{p'}\Lambda_{t}^{k}(da)dt.$$

As usual, $C > 0$ is a constant, independent of $n$ and $P$, which can change from line to line. On the other hand, fix arbitrarily a constant control $a_0 \in A$, and let $P_n^0$ denote an element of $\mathcal{R}_n$ satisfying $P_n^0(\Lambda_t^k = \delta_{a_0}, \ a.e. \ t) = 1$ for all $k = 1, \ldots, n$ (as in Remark 2.8, the existence of such $P_n^0$ follows from the results of Stroock and Varadhan [50]). Lemma 3.3 implies

$$\mathbb{E}^{P_{n}^{0}}\left[\int_{\mathcal{C}^{d}}\|x\|^{p'}\widehat{\mu}^{n,x}(dx)\right] = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}^{P_{n}^{0}}\left[\|X^{k}\|^{p'}\right] \leq C\left(1 + T|a_{0}|^{p'}\right).$$

Then use the lower bounds of Assumption (A.5) to show $\inf_n \mathbb{E}^{P_n^0}[\Gamma(\widehat{\mu}^n)] > -\infty$. Now, if $P \in \mathcal{R}_n$ is $\epsilon$-optimal, we have $\mathbb{E}^P[\Gamma(\widehat{\mu}^n)] \geq \mathbb{E}^{P_n^0}[\Gamma(\widehat{\mu}^n)] - \epsilon$, which combined with (3.6) yields a constant $C \geq 0$ such that, for all $n$ and all $\epsilon$-optimal $P \in \mathcal{R}_n$,

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}^{P}\int_{0}^{T}\int_{A}\left(|a|^{p'} - C|a|^{p}\right)\Lambda_{t}^{k}(da)dt \leq C. \qquad \square$$

**3.3. A tightness criterion.** Finally, we state without proof a useful tailor-made compactness result for controlled Itô processes, which is essentially an application of Aldous's criterion. Its first claim is proved in [39, Proposition B.4] and its second in [38, Proposition 5.3], or rather an easy extension thereof (as it did not allow for control in the volatility).

PROPOSITION 3.5. *Fix $c > 0$. For $\kappa > 0$, let $\mathcal{Q}_\kappa$ denote the set of laws $\mathbb{P} \circ (X, \Lambda)^{-1}$, where the following hold:*
   1. *$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space supporting a $d$-dimensional $\mathbb{F}$-adapted process $X$, a $\mathcal{P}^p(A)$-valued $\mathbb{F}$-predictable process $\Lambda$, and a vector $N = (N^1, \ldots, N^{d_W})$ of orthogonal $\mathbb{F}$-martingale measures, each with intensity $\Lambda_t(da)dt$.*
   2. *The state equation holds:*

$$dX_{t} = \int_{A}B(t,a)\Lambda_{t}(da)dt + \int_{A}\Sigma(t,a)N(da,dt),$$

   *where $(B, \Sigma) : [0, T] \times \Omega \times A \to \mathbb{R}^d \times \mathbb{R}^{d \times d_W}$ are jointly measurable with respect to the predictable $\sigma$-field of $[0, T] \times \Omega$ and the Borel $\sigma$-field of $A$.*

3. *It holds for all $(t,a)$, a.s., that*

$$|B(t,a)| \le c\,(1+|X_t|+|a|), \qquad |\Sigma(t,a)|^2 \le c\,(1+|X_t|^{p_\sigma}+|a|^{p_\sigma}).$$

4. *Last, we have*

$$\mathbb{E}\left[|X_0|^{p'} + \int_0^T \int_A |a|^{p'}\Lambda_t(da)dt\right] \le \kappa.$$

*(That is, $\mathcal{Q}_\kappa$ is defined by varying the probability space as well as $B$ and $\Sigma$.) Then $\mathcal{Q}_\kappa$ is precompact in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$. Moreover, if a triangular array $\{\kappa_{n,i} : 1 \le i \le n\} \subset [0,\infty)$ satisfies $\sup_n \frac{1}{n}\sum_{i=1}^n \kappa_{n,i} < \infty$, then the set*

$$\left\{\frac{1}{n}\sum_{i=1}^n Q_i : n \ge 1,\ i=1,\dots,n,\ Q_i \in \mathcal{Q}_{\kappa_{n,i}}\right\}$$

*is precompact in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$.*

**4. Proofs of existence Theorems 2.2 and 2.3.** The existence Theorem 2.2 is an immediate consequence of the following Lemmas 4.1 and 4.2, which reduces the problem to maximizing an upper semicontinuous function on a compact set. Throughout the section, Assumption A is in force. The first lemma is essentially contained in [39, Lemma 4.5] and [38, Lemma 4.5], but we include the proof for the sake of transparency.

LEMMA 4.1. *Under Assumption* A, $\Gamma$ *is upper semicontinuous on $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$, and the map $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})) \ni P \mapsto \mathbb{E}^P[\Gamma(\mu)]$ is upper semicontinuous. When Assumption* B *holds as well, the latter function is continuous when restricted to any set $K \subset \mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$ satisfying*

$$\lim_{r\to\infty} \sup_{P\in K} \mathbb{E}^P\left[\int_{\mathcal{C}^d\times\mathcal{V}} Z 1_{\{Z\ge r\}}\,d\mu\right] = 0,$$

(4.1)
$$\text{where} \quad Z(x,q) = \|x\|^{p'} + \int_0^T \int_A |a|^{p'} q_t(da)dt.$$

*Proof.* Upper semicontinuity of the map $F$ defined on $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}) \times \mathcal{C}^d \times \mathcal{V}$ by

$$F(m,x,q) = \int_0^T \int_A f\,(t,x_t,m_t^x,a)\,q_t(da)dt + g\,(x_T,m_T^x)$$

follows from upper semicontinuity of $f$ and $g$ (Assumption (A.3)) and the growth assumption (Assumption (A.5)) (see [39, Corollary A.5] for details). This is enough to conclude (e.g., using Skorohod representation and Fatou's lemma) that $\Gamma(m) = \int F(m,\cdot)\,dm$ is upper semicontinuous. To prove the second claimed upper semciontinuity, note that Assumption (A.5) implies that there exists $C > 0$ (which can change from line to line) such that for all $m \in \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$ we have

$$\Gamma(m) = \int F(m,\cdot)\,dm \le C\int_{\mathcal{C}^d} m^x(dx)\left(1+\|x\|^p + \int_{\mathcal{C}^d}\|z\|^p m^x(dz)\right)$$
$$\le C\left(1+\int_{\mathcal{C}^d}\|z\|^p m^x(dz)\right)$$
$$= C\left(1+\ell_{\mathcal{C}^d,p}^p(m^x,\delta_0)\right)$$
$$\le C\left(1+\ell_{\mathcal{C}^d\times\mathcal{V},p}^p(m,\tilde m)\right),$$

where $\tilde{m} = \delta_0 \times \delta_{q^0}$ for an arbitrary choice of $q^0 \in \mathcal{V}$, and where $\ell_{\mathcal{C}^d \times \mathcal{V},p}$ was defined in (2.3). This is enough to prove the second claim; indeed, for a general complete separable metric space $(E, d)$, the map $\mathcal{P}^p(E) \ni \mu \mapsto \int \varphi \, d\mu$ is upper semicontinuous if $\varphi$ is upper semicontinuous and there exists $c > 0$ such that $\varphi(x) \leq c(1 + d(x, x_0))$ for all $x \in E$, for some $x_0 \in E$.

Finally, we prove the claimed restricted continuity, under the additional Assumption B, which says that $f$ and $g$ are jointly continuous. Let $P_n \to P$ in $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$ with $P_n, P \in K$. We show first that

$$(4.2) \qquad \left( \mathbb{E}^{P_n} - \mathbb{E}^P \right) \left[ \int_{\mathcal{C}^d \times \mathcal{V}} \int_0^T \int_A f(t, x_t, \mu_t^x, a) q_t(da) dt \mu(dx, dq) \right] \to 0.$$

To this end, define a probability measure $Q_n$ on $[0, T] \times \mathbb{R}^d \times \mathcal{P}^p(\mathbb{R}^d) \times A$ by

$$Q_n(B) = \frac{1}{T} \mathbb{E}^{P_n} \left[ \int_{\mathcal{C}^d \times \mathcal{V}} \int_0^T \int_A 1_{\{(t, x_t, \mu_t^x, a) \in B\}} q_t(da) dt \mu(dx, dq) \right],$$

and define $Q$ similarly in terms of $P$. It is clear that $Q_n \to Q$ weakly, because $P_n \to P$ weakly. Thus $Q_n \circ f^{-1} \to Q \circ f^{-1}$ weakly, as probability measures on $\mathbb{R}$. It follows from the assumption on the set $K$ and on the growth assumption (Assumption (A.5)) that

$$\lim_{r \to \infty} \sup_n \int_{\{|f| \geq r\}} |f| \, dQ_n = 0.$$

Thus $\int f \, dQ_n \to \int f \, dQ$, which is precisely (4.2). A similar argument shows

$$\left( \mathbb{E}^{P_n} - \mathbb{E}^P \right) \left[ \int_{\mathcal{C}^d} g(x, \mu) \mu^x(dx) \right] \to 0.$$

Combining this and (4.2) shows $\mathbb{E}^{P_n}[\Gamma(\mu)] \to \mathbb{E}^P[\Gamma(\mu)]$.          $\square$

The proof of the following compactness lemma makes some use of the following estimate, which follows immediately from Assumption (A.4) and the fact that $1 \vee p_\sigma \leq p$, by (A.2). Recall the definition of the generator $\mathcal{L}$ from (2.4). For every smooth compactly supported $\varphi$ on $\mathbb{R}^d$, there exists a constant $C > 0$ depending only on $\varphi$ and the constant $c_1$ of Assumption A such that

$$(4.3) \qquad |\mathcal{L}\varphi(t, x, m, a)| \leq C \left( 1 + |x|^p + \int_{\mathbb{R}^d} |z|^p m(dz) + |a|^p \right)$$

for all $(t, x, m, a) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}^p(\mathbb{R}^d) \times A$.

LEMMA 4.2. *Given $\epsilon \geq 0$, let $\mathcal{R}^\epsilon \subset \mathcal{R}$ be the of $\epsilon$-optimal controls, as in Lemma 3.2. Then $\mathcal{R}^\epsilon$ is compact in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$.*

*Proof.* Lemma 3.2 says that

$$\sup_{m \in \mathcal{R}^\epsilon} \mathbb{E}^m \int_0^T \int_A |a|^{p'} \Lambda_t(da) dt < \infty.$$

According to Lemma 3.1, this impies

$$\sup_{m \in \mathcal{R}^\epsilon} \int_{\mathcal{C}^d} \|x\|^{p'} m^x(dx) = \sup_{m \in \mathcal{R}^\epsilon} \mathbb{E}^m \left[ \|X\|^{p'} \right] < \infty.$$

It now follows easily from the first claim of Proposition 3.5 that $\mathcal{R}^\epsilon$ is precompact in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$. To show that $\mathcal{R}^\epsilon$ is closed, note that

$$\mathcal{R}^\epsilon = \{m \in \mathcal{R} : \Gamma(m) \geq V - \epsilon\}, \qquad \text{where} \qquad V = \sup_{m \in \mathcal{R}} \Gamma(m).$$

Now let $m^n \to m^\infty$ in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$ with $m^n \in \mathcal{R}^\epsilon$. Upper semicontinuity of $\Gamma$ (see Lemma 4.1) implies

$$\Gamma(m^\infty) \geq \limsup_{n \to \infty} \Gamma(m^n) \geq V - \epsilon,$$

and it remains only to show that $m^\infty$ belongs to $\mathcal{R}$. Since $X_0$ has law $\lambda$ under $m^n$, the same is true under $m^\infty$. For a smooth compactly supported function $\varphi$ and for $m \in \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$, define $M_t^{m,\varphi} : \mathcal{C}^d \times \mathcal{V} \to \mathbb{R}$ by

$$(4.4) \qquad M_t^{m,\varphi}(x,q) = \varphi(x_t) - \int_0^t \int_A \mathcal{L}\varphi(s, x_s, m_t^x, a) q_s(da) ds.$$

The estimate (4.3) yields

$$(4.5) \qquad |M_t^{m,\varphi}(x,q)| \leq C \left( 1 + \|x\|^p + \int_{\mathcal{C}^d} \|z\|^p m^x(dz) + \int_0^T \int_A |a|^p q_t(da) dt \right).$$

Using this and the continuity of $(b, \sigma)$, it is readily checked that $(m, x, q) \mapsto M_t^{m,\varphi}(x, q)$ is a continuous function for each $t$ and $\varphi$, e.g., using [39, Corollary A.5]. Since $m^n \to m^\infty$ in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$, it follows that

$$\mathbb{E}^{m^\infty} \left[ \left( M_t^{m^\infty, \varphi} - M_s^{m^\infty, \varphi} \right) h \right] = \lim_n \mathbb{E}^{m^n} \left[ \left( M_t^{m^n, \varphi} - M_s^{m^n, \varphi} \right) h \right]$$

for every smooth compactly supported $\varphi$ and every bounded continuous function $h$ on $\mathcal{C}^d \times \mathcal{V}$ which is measurable with respect to $\sigma(X_s, \Lambda_s : s \leq t)$. Because $m^n$ is in $\mathcal{R}$, the process $(M_t^{m^n, \varphi}(X, \Lambda))_{t \in [0,T]}$ is a martingale under $m^n$, and the above quantity is zero. This shows that $(M_t^{m^\infty, \varphi}(X, \Lambda))_{t \in [0,T]}$ is a martingale under $m^\infty$, and so $m^\infty \in \mathcal{R}$; see Appendix A for a short explanation of why it suffices here to consider only bounded continuous $h$.    $\square$

*Proof of Theorem* 2.2. Fix $\epsilon > 0$, and note that $\sup_{m \in \mathcal{R}} \Gamma(m) = \sup_{m \in \mathcal{R}^\epsilon} \Gamma(m)$. By Lemma 4.2, $\mathcal{R}^\epsilon$ is compact, and by Lemma 4.1, $\Gamma$ is upper semicontinuous. Therefore, the supremum is attained.    $\square$

*Proof of Theorem* 2.3. As in [32, Theorem 2.5(a)], there exists a measurable function $\tilde{\sigma} : [0, T] \times \mathbb{R}^d \times \mathcal{P}^p(\mathbb{R}^d) \times \mathcal{P}^p(A) \to \mathbb{R}^{d \times d}$ such that

$$\tilde{\sigma} \tilde{\sigma}^\top(t, x, m, q) = \sigma \sigma^\top(t, x, m, a) q(da) \quad \text{for all } (t, x, m, q)$$

and also $\tilde{\sigma}(t, x, m, \delta_a) = \sigma(t, x, m, a)$ for $a \in A$. Moreover, given $m \in \mathcal{R}$, we may find a filtered probability space $(\Omega^1, \mathcal{F}^1, \mathbb{F}^1, \mathbb{P}^1)$ supporting a $d$-dimensional $\mathbb{F}^1$-Wiener process $W^1$, a $d$-dimensional $\mathbb{F}^1$-adapted process $X^1$, and a $\mathcal{P}^p(A)$-valued $\mathbb{F}^1$-predictable process $\Lambda^1$ such that $\mathbb{P} \circ (X^1, \Lambda^1)^{-1} = m$ and

$$dX_t^1 = \int_A b\left(t, X_t^1, m_t^x, a\right) \Lambda_t^1(da) dt + \tilde{\sigma}\left(t, X_t^1, m_t^x, \Lambda_t^1\right) dW_t^1.$$

The convexity Assumption C entails that, almost surely,

$$(4.6) \qquad \int_A \left( b, \sigma\sigma^\top, f \right) \left( t, X_t^1, m_t^x, a \right) \Lambda_t(da) \in K \left( t, X_t^1, m_t^x \right).$$

By [25, Proposition 3.5], $K(t, x, m_t^x)$ is a closed set for each $(t, x)$. The measurable selection result of [25, Theorem A.9] (or rather an extension in [14, Lemma 3.1]) implies that there exist measurable functions $\hat\alpha : [0, T] \times \mathbb{R}^d \to A$ and $\hat{z} : [0, T] \times \mathbb{R}^d \to [0, \infty)$ such that

$$(4.7) \qquad \mathbb{E}\left[ \int_A b \left( t, X_t^1, m_t^x, a \right) \Lambda_t^1(da) \middle| X_t^1 \right] = b \left( t, X_t^1, m_t^x, \hat\alpha \left( t, X_t^1 \right) \right),$$

$$(4.8) \quad \mathbb{E}\left[ \int_A \sigma\sigma^\top \left( t, X_t^1, m_t^x, a \right) \Lambda_t^1(da) \middle| X_t^1 \right] = \sigma\sigma^\top \left( t, X_t^1, m_t^x, \hat\alpha \left( t, X_t^1 \right) \right),$$

$$(4.9) \qquad \mathbb{E}\left[ \int_A f \left( t, X_t^1, m_t^x, a \right) \Lambda_t^1(da) \middle| X_t^1 \right] = f \left( t, X_t^1, m_t^x, \hat\alpha \left( t, X_t^1 \right) \right) - \hat{z} \left( t, X_t^1 \right).$$

Note that (4.8) and (4.6) together imply

$$\mathbb{E}\left[ \int_A \tilde\sigma\tilde\sigma^\top \left( t, X_t^1, m_t^x, a \right) \Lambda_t^1(da) \middle| X_t \right] = \sigma\sigma^\top \left( t, X_t^1, m_t^x, \hat\alpha \left( t, X_t^1 \right) \right).$$

Thanks to (4.7), the mimicking theorem of Brunick and Shreve [6] (a generalization of a well-known result of Gyöngy [24]) then implies that there exists a filtered probability space $(\Omega^2, \mathcal{F}^2, \mathbb{F}^2, \mathbb{P}^2)$ supporting a $d_W$-dimensional $\mathbb{F}^2$-Wiener process $W^2$ and a $d$-dimensional $\mathbb{F}^2$-adapted process $X^2$ such that

$$dX_t^2 = b \left( t, X_t^2, m_t^x, \hat\alpha \left( t, X_t^2 \right) \right) dt + \sigma \left( t, X_t^2, m_t^x, \hat\alpha \left( t, X_T^2 \right) \right) dW_t^2,$$

and also $\mathbb{P}^2 \circ (X_t^2)^{-1} = \mathbb{P}^1 \circ (X_t^1)^{-1} = m_t^x$ for each $t \in [0, T]$. Define a $\mathcal{P}^p(A)$-valued process $\Lambda^2$ by $\Lambda_t^2 = \delta_{\hat\alpha(t, X_t^2)}$, and let $\widetilde{m} = \mathbb{P}^2 \circ (X^2, \Lambda^2)^{-1}$. Then $\widetilde{m}$ belongs to $\mathcal{R}$ and is Markovian, and also $\widetilde{m}_t^x = m_t^x$ for all $t \in [0, T]$. Finally, use Fubini's theorem and (4.9) to get, since $\hat{z} \geq 0$,

$$\begin{aligned}
\Gamma(m) &= \mathbb{E}^{\mathbb{P}^1}\left[ \int_0^T \int_A f \left( t, X_t^1, m_t^x, a \right) \Lambda_t^1(da)dt + g \left( X_T^1, m_T^x \right) \right] \\
&= \mathbb{E}^{\mathbb{P}^1}\left[ \int_0^T \left( f \left( t, X_t^1, m_t^x, \hat\alpha \left( t, X_t^1 \right) \right) - \hat{z} \left( t, X_t^1 \right) \right) dt + g \left( X_T^1, m_T^x \right) \right] \\
&= \mathbb{E}^{\mathbb{P}^2}\left[ \int_0^T \left( f \left( t, X_t^2, \widetilde{m}_t^x, \hat\alpha \left( t, X_t^2 \right) \right) - \hat{z} \left( t, X_t^2 \right) \right) dt + g \left( X_T^2, \widetilde{m}_T^x \right) \right] \\
&\leq \mathbb{E}^{\mathbb{P}^2}\left[ \int_0^T f \left( t, X_t^2, \widetilde{m}_t^x, \hat\alpha \left( t, X_t^2 \right) \right) dt + g \left( X_T^2, \widetilde{m}_T^x \right) \right] \\
&= \Gamma(\widetilde{m}). \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square
\end{aligned}$$

**5. Limits of $n$-state controls.** The proofs of both Theorems 2.11 and 2.12 involve similar constructions, detailed in the two propositions of this section. In fact, these two key results comprise the bulk of the proofs, by identifying limit points of various sequences of $n$-state controls. The first proposition proves all of Theorem 2.11

except for the claimed optimality of the limit points, and the second shows that every candidate control in the McKean–Vlasov control problem can be realized as a limit of $n$-state controls.

PROPOSITION 5.1. *Suppose $P_n \in \mathcal{R}_n$ satisfy*

$$(5.1) \qquad \sup_n \frac{1}{n} \sum_{k=1}^n \mathbb{E}^{P_n} \int_0^T \int_A |a|^{p'} \Lambda_t^k(da) dt < \infty.$$

*Then $(P_n \circ (\widehat{\mu}^n)^{-1})$ is precompact in $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$, and every limit point is supported on $\mathcal{R}$.*

PROPOSITION 5.2. *Let $m \in \mathcal{R}$. Then there exists $P_n \in \mathcal{R}_n$ such that $P_n \circ (\widehat{\mu}^n)^{-1} \to \delta_m$ in $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$ and $\mathbb{E}^{P_n}[\Gamma(\widehat{\mu}^n)] \to \Gamma(m)$.*

The proofs make some use of the metric $d_{\mathcal{V}}$ on $\mathcal{V}$ defined in (2.2). Fix arbitrarily some $a_0 \in A$, and let $q^0(dt, da) = dt\delta_{a_0}(da)$. Then any $q \in \mathcal{V}$ can be coupled with $q^0$ via the measure $\pi$ on $[0, T]^2 \times A^2$ given by $\pi(dt, dt', da, da') = dt\delta_t(dt')q_t(da)\delta_{a_0}(da)$, and this gives rise to the estimate

(5.2)

$$d_{\mathcal{V}}(q, q^0) \le \left( \frac{1}{T} \int_0^T \int_A |a - a_0|^p q_t(da) dt \right)^{1/p} \le |a_0| + \left( \frac{1}{T} \int_0^T \int_A |a|^p q_t(da) dt \right)^{1/p}.$$

Recall that $\mathcal{C}^d \times \mathcal{V}$ is equipped with the metric $d_{\mathcal{C}^d \times \mathcal{V}}$ defined in (2.3).

*Proof of Proposition* 5.1. We adapt to the controlled setting a martingale argument which is by now classical in McKean–Vlasov limit theory (cf. [45, 22] for uncontrolled and [8] for controlled diffusions). Throughout the proof, we will make use of the notation $\langle m, \varphi \rangle$ in place of $\int \varphi \, dm$. Let $q^0 \in \mathcal{V}$ be defined as above. According to [39, Corollary B.2], to prove precompactness it suffices to check that

$$(5.3) \qquad \sup_n \mathbb{E}^{P_n} \left[ \int_{\mathcal{C}^d \times \mathcal{V}} d_{\mathcal{C}^d \times \mathcal{V}} \left( (x, q), (0, q^0) \right)^{p'} \widehat{\mu}^n(dq, dx) \right] < \infty$$

and also that the mean measures $(\mathbb{E}^{P_n}[\widehat{\mu}^n])$ are tight. The mean measures are defined by, for bounded measurable functions $\varphi$ on $\mathcal{C}^d \times \mathcal{V}$,

$$\left\langle \mathbb{E}^{P_n}[\widehat{\mu}^n], \varphi \right\rangle = \mathbb{E}^{P_n} \left[ \langle \widehat{\mu}^n, \varphi \rangle \right] = \frac{1}{n} \sum_{k=1}^n \mathbb{E}^{P_n} \left[ \varphi \left( X^k, \Lambda^k \right) \right].$$

To prove (5.3), it suffices in light of (5.2) to show that

$$(5.4) \qquad \sup_n \frac{1}{n} \sum_{k=1}^n \mathbb{E}^{P_n} \left[ \|X^k\|^{p'} + \int_0^T \int_A |a|^{p'} \Lambda_t^k(da) dt \right] < \infty.$$

But this follows from the assumption (5.1) and Lemma 3.3. Finally, to show that the mean measures are tight, simply use the second assertion of Proposition 3.5.

The next task is to identify the limit points. Fix a limit point $P \in \mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$, and relabel the subsequence so that $P_n \to P$. First, note that $\widehat{\mu}_0^{n,x} = \frac{1}{n} \sum_{i=1}^n \delta_{X_0^i}$ converges weakly to $\delta_\lambda$, since $(X_0^i)$ are i.i.d. with law $\lambda$ by assumption. That is, $P(\mu_0^x = \lambda) = 1$, where $\mu$ denotes the identity map on $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$. To prove that

$P(\mu \in \mathcal{R}) = 1$, it remains to show that the martingale problem is satisfied at the limit. That is, defining $M_t^{m,\varphi}(x, q)$ as in (4.4), we must show that

(5.5) $$P\left((M_t^{\mu,\varphi})_{t \in [0,T]} \text{ is a martingale under } \mu, \ \forall \varphi\right) = 1,$$

where "$\forall \varphi$" means "for all smooth compactly supported functions $\varphi$." To this end, recall the useful estimate (4.5) as well as the discussion thereafter, namely, that $M_t^{m,\varphi}(x, q)$ is jointly continuous in $(m, x, q) \in \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}) \times \mathcal{C}^d \times \mathcal{V}$ for each fixed $t$ and $\varphi$.

Now, use the SDE representation of Proposition 2.7 along with Itô's formula to see that $M_t^{\widehat{\mu}^n,\varphi}(X^k, \Lambda^k)$ is a martingale under $P_n$ with quadratic variation

$$\int_0^t \int_A \left|\sigma\left(s, X_s^k, \widehat{\mu}_s^{n,x}, a\right) \nabla\varphi\left(X_s^k\right)\right|^2 \Lambda_s(da)ds.$$

In fact, for $k = 1, \ldots, n$, these martingales $M_t^{\widehat{\mu}^n,\varphi}(X^k, \Lambda^k)$ are orthogonal. Fix $s < t$, and let $h : \mathcal{C}^d \times \mathcal{V} \to \mathbb{R}$ be bounded, continuous, and $\sigma(X_s, \Lambda_s : s \le t)$-measurable. We will show that

(5.6) $$\mathbb{E}^P\left[\langle \mu, h(M_t^{\mu,\varphi} - M_s^{\mu,\varphi})\rangle^2\right] = 0.$$

First note that

$$\mathbb{E}^{P_n}\left[\left\langle \widehat{\mu}^n, h\left(M_t^{\widehat{\mu}^n,\varphi} - M_s^{\widehat{\mu}^n,\varphi}\right)\right\rangle^2\right]$$

$$= \mathbb{E}^{P_n}\left[\left(\frac{1}{n}\sum_{k=1}^n h\left(X^k, \Lambda^k\right)\left(M_t^{\widehat{\mu}^n,\varphi}\left(X^k, \Lambda^k\right) - M_s^{\widehat{\mu}^n,\varphi}\left(X^k, \Lambda^k\right)\right)\right)^2\right]$$

$$= \frac{1}{n^2}\sum_{k=1}^n \mathbb{E}^{P_n}\left[h\left(X^k, \Lambda^k\right)^2\left(M_t^{\widehat{\mu}^n,\varphi}\left(X^k, \Lambda^k\right) - M_s^{\widehat{\mu}^n,\varphi}\left(X^k, \Lambda^k\right)\right)^2\right]$$

$$\le \frac{1}{n^2}\sum_{k=1}^n \mathbb{E}^{P_n}\left[h\left(X^k, \Lambda^k\right)^2 \int_s^t \int_A \left|\sigma\left(u, X_u^k, \widehat{\mu}_u^{n,x}, a\right) \nabla\varphi\left(X_u^k\right)\right|^2 \Lambda_u(da)du\right].$$

Assumption (A.4) and $p_\sigma \le p'$ imply that there exists $C > 0$ (independent of $n$) such that

$$\int_s^t \int_A \left|\sigma\left(u, X_u^k, \widehat{\mu}_u^{n,x}, a\right) \nabla\varphi\left(X_u^k\right)\right|^2 \Lambda_u(da)du$$

$$\le C \int_s^t \int_A \left(1 + |X_u^k|^{p_\sigma} + \left(\int_{\mathbb{R}^d} |z|^p \widehat{\mu}_u^{n,x}(dz)\right)^{p_\sigma/p} + |a|^{p_\sigma}\right) \Lambda_u(da)du$$

$$\le C\left(1 + \|X^k\|^{p'} + \int_{\mathcal{C}^d} \|z\|^{p'} \widehat{\mu}^{n,x}(dz) + \int_0^T \int_A |a|^{p'} \Lambda_u^k(da)du\right).$$

Along with (5.4), this implies

$$\sup_n n\mathbb{E}^{P_n}\left[\left\langle \widehat{\mu}^n, h\left(M_t^{\widehat{\mu}^n,\varphi} - M_s^{\widehat{\mu}^n,\varphi}\right)\right\rangle^2\right] < \infty.$$

This in turn implies

$$\mathbb{E}^P\left[\langle \mu, h\left(M_t^{\mu,\varphi} - M_s^{\mu,\varphi}\right)\rangle^2\right] \le \liminf_{n \to \infty} \mathbb{E}^{P_n}\left[\left\langle \widehat{\mu}^n, h\left(M_t^{\widehat{\mu}^n,\varphi} - M_s^{\widehat{\mu}^n,\varphi}\right)\right\rangle^2\right] = 0,$$

because the map $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}) \ni m \mapsto \langle m, h(M_t^{m,\varphi} - M_s^{m,\varphi})\rangle^2$ is continuous and bounded from below; indeed, this follows from the aforementioned joint continuity of $M_t^{m,\varphi}(x,q)$ in $(m,x,q)$.

We have now proved (5.6), and it follows that $\langle \mu, h(M_t^{\mu,\varphi} - M_s^{\mu,\varphi})\rangle = 0$ holds $P$-almost surely, for each $s < t$, each smooth compactly supported $\varphi$, and each bounded continuous $\mathcal{F}_s$-measurable $h$. By applying this to a suitably dense countable set of $(s,t,\varphi,h)$ we can interchange the order of the quantifiers and conclude that $\langle \mu, h(M_t^{\mu,\varphi} - M_s^{\mu,\varphi})\rangle = 0$ for each $s < t$ and each $\varphi$, $P$-almost surely. This proves (5.5) and thus $P(\mu \in \mathcal{R}) = 1$. To elaborate on this last point, it is clear that we can restrict our attention to $s$ and $t$ belonging to a dense subset of $[0,T]$ and to $\varphi$ belonging to a dense set of smooth functions, whereas the separability of the class of functions $h$ is less immediate. See Appendix A for details. □

*Proof of Proposition* 5.2. The line of argument is often known as *trajectorial propagation of chaos* (see [51]), constructing an explicit coupling between the limit and prelimit state processes. To begin the proof, apply Proposition 2.5 to $m$, and then construct a sequence of independent copies. As a result we may find a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ supporting i.i.d. random variables $(X^i, \Lambda^i, N^i)$, satisfying the following:

1. Each $X^i$ is a $d$-dimensional $\mathbb{F}$-adapted process.
2. Each $\Lambda^i$ is an $\mathbb{F}$-predictable $\mathcal{P}^p(A)$-valued process.
3. Each $N^i = (N^{i,1}, \ldots, N^{i,d_W})$ is a vector of orthogonal $\mathbb{F}$-martingale measures, each with intensity measure $\Lambda_t^i(da)dt$.
4. The McKean–Vlasov equation holds for each $i$:

$$dX_t^i = \int_A b\left(t, X_t^i, m_t^x, a\right) \Lambda_t^i(da)dt + \int_A \sigma\left(t, X_t^i, m_t^x, a\right) N^i(da, dt).$$

5. The law of $(X^i, \Lambda^i)$ is precisely $m$ for each $i$.

In particular, Lemma 3.2 combined with Lemma 3.1 together imply that, for each $i$,

$$(5.7) \qquad \mathbb{E}\left[\|X^i\|^{p'} + \int_0^T \int_A |a|^{p'} \Lambda_t^i(da)dt\right] < \infty.$$

Thanks to the Lipschitz assumption (Assumption B), there exists a unique square-integrable (recall $p' \geq 2$) progressively measurable processes $(Z^{n,1}, \ldots, Z^{n,n})$ such that

$$dZ_t^{n,i} = \int_A b\left(t, Z_t^{n,i}, \widetilde{\mu}_t^{n,x}, a\right) \Lambda_t^i(da)dt + \int_A \sigma\left(t, Z_t^{n,i}, \widetilde{\mu}_t^{n,x}, a\right) N^i(da, dt), \ Z_0^{n,i} = X_0^i,$$

$$\widetilde{\mu}^n = \frac{1}{n} \sum_{k=1}^n \delta_{(Z^{n,k}, \Lambda^k)}.$$

(Indeed, existence and uniqueness here is an easy adaptation of now-standard arguments, which can be found in [21, 51].) It follows from Proposition 2.7 that the law $P_n = \mathbb{P} \circ ((Z^{n,i}, \Lambda^i)_{i=1}^n)^{-1}$ belongs to $\mathcal{R}_n$. Define $\widehat{\mu}^n = \frac{1}{n} \sum_{i=1}^n \delta_{(X^i, \Lambda^i)}$. We claim that

$$(5.8) \qquad \lim_{n \to \infty} \mathbb{E}\left[\|Z^{n,1} - X^1\|^{p'} + \ell_{\mathcal{C}^d \times \mathcal{V}, p'}^{p'}\left(\widehat{\mu}^n, \widetilde{\mu}^n\right)\right] \to 0.$$

Assuming for the moment that (5.8) holds, we complete the proof as follows: Because $(X^i, \Lambda^i)$ are i.i.d. with law $m$, the law of large numbers implies $\widehat{\mu}^n \to m$ almost surely in $\mathcal{P}(\mathcal{C}^d \times \mathcal{V})$, and the finite moments of (5.7) allow us upgrade this convergence to

$\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$. Hence, (5.8) implies $\mathbb{P} \circ (\widetilde{\mu}^n)^{-1} \to \delta_m$ in $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$. To conclude that $\mathbb{E}[\Gamma(\widetilde{\mu}^n)] \to \Gamma(m)$, simply note that (5.8) and (5.7), together with exchangeability of $(Z^{n,i}, \Lambda^i)_{i=1}^n$, verify the uniform integrability hypothesis of Lemma 4.1.

The rest of the proof is devoted to justifying (5.8). For $k = 1, \ldots, n$, use the Burkholder–Davis–Gundy inequality and the Lipschitz assumption (Assumption (A.4)) (noting that $\ell_{\mathbb{R}^d, p} \le \ell_{\mathbb{R}^d, p'}$) to find a constant $C$ (which will change from line to line) such that

$$
\mathbb{E}\left[ \left\| Z^{n,k} - X^k \right\|_t^{p'} \right]
$$

$$
\le C \mathbb{E}\left[ \left( \int_0^t \int_A \left| b\left(s, Z_s^{n,k}, \widetilde{\mu}_s^{n,x}, a\right) - b\left(s, X_s^k, m_s^x, a\right) \right| \Lambda_s^k(da)\,ds \right)^{p'} \right.
$$

$$
\left. + \left( \int_0^t \int_A \left| \sigma\left(s, Z_s^{n,k}, \widetilde{\mu}_s^{n,x}, a\right) - \sigma\left(s, X_s^k, m_s^x, a\right) \right|^2 \Lambda_s^k(da)\,ds \right)^{p'/2} \right]
$$

$$
\le C \mathbb{E} \int_0^t \left[ \left| Z_s^{n,k} - X_s^k \right|^{p'} + \ell_{\mathbb{R}^d, p'}^{p'}\left(\widetilde{\mu}_s^{n,x}, m_s^x\right) \right] ds.
$$

Use Gronwall's inequality to get

$$
(5.9) \qquad \mathbb{E}\left[ \left\| Z^{n,k} - X^k \right\|_t^{p'} \right] \le C \mathbb{E} \int_0^t \ell_{\mathbb{R}^d, p'}^{p'}\left(\widetilde{\mu}_s^{n,x}, m_s^x\right) ds.
$$

Define the truncated Wasserstein distance $\ell_{s, \mathcal{C}^d, p'}$ by

$$
(5.10) \quad \ell_{s, \mathcal{C}^d, p'}^{p'}(m, m')
$$

$$
= \inf \left\{ \int_{\mathcal{C}^d \times \mathcal{C}^d} \|x - y\|_s^{p'} \pi(dx, dy) : \pi \in \mathcal{P}(\mathcal{C}^d \times \mathcal{C}^d) \text{ has marginals } m,\ m' \right\}.
$$

It is straightforward to check that for every $m^1, m^2 \in \mathcal{P}(\mathcal{C}^d)$ and every $s \in [0, T]$ we have $\ell_{\mathbb{R}^d, p'}(m_s^1, m_s^2) \le \ell_{s, \mathcal{C}^d, p'}(m^1, m^2)$. Returning to (5.9), use the obvious coupling and the triangle inequality to get

$$
\mathbb{E}\left[ \ell_{t, \mathcal{C}^d, p'}^{p'}(\widetilde{\mu}^{n,x}, \widehat{\mu}^{n,x}) \right] \le \frac{1}{n} \sum_{k=1}^n \widetilde{\mathbb{E}}\left[ \left\| Z^{n,k} - X^k \right\|_t^{p'} \right]
$$

$$
\le C \mathbb{E} \int_0^t \left( \ell_{s, \mathcal{C}^d, p'}^{p'}(\widetilde{\mu}^{n,x}, \widehat{\mu}^{n,x}) + \ell_{s, \mathcal{C}^d, p'}^{p'}(\widehat{\mu}^{n,x}, m^x) \right) ds.
$$

Another application of Gronwall's inequality yields, for all $t \in [0, T]$,

$$
\mathbb{E}\left[ \ell_{t, \mathcal{C}^d, p'}^{p'}(\widetilde{\mu}^{n,x}, \widehat{\mu}^{n,x}) \right] \le C \mathbb{E} \int_0^t \ell_{s, \mathcal{C}^d, p'}^{p'}(\widehat{\mu}^{n,x}, m^x)\,ds \le C T \mathbb{E}\left[ \ell_{\mathcal{C}^d, p'}^{p'}(\widehat{\mu}^{n,x}, m^x) \right].
$$

The finite moment (5.7) and the almost sure weak convergence $\widehat{\mu}^{n,x} \to m^x$ together imply that the above expectation tends to zero. Recalling the definition (2.3) of the metric on $\mathcal{C}^d \times \mathcal{V}$, we have

$$
d_{\mathcal{C}^d \times \mathcal{V}}^{p'}((x, q), (x', q')) \le 2^{p'-1} \left( d_{\mathcal{C}^d}^{p'}(x, x') + d_{\mathcal{V}}^{p'}(q, q') \right).
$$

Since $\widetilde{\mu}^n$ and $\widehat{\mu}^n$ have the same $\mathcal{V}$-marginal, it follows that

$$
\mathbb{E}\left[ \ell_{\mathcal{C}^d \times \mathcal{V}, p'}^{p'}(\widetilde{\mu}^n, \widehat{\mu}^n) \right] \le 2^{p'-1} \mathbb{E}\left[ \ell_{\mathcal{C}^d, p'}^{p'}(\widetilde{\mu}^{n,x}, \widehat{\mu}^{n,x}) \right] \to 0.
$$

Recalling also (5.9), this completes the proof of (5.8).      $\square$

## 6. Proofs of the limit theorems.

*Proof of Theorem* 2.11. Given the preparations of the previous section, the proof of the main limit theorem is now straightforward. Note that the second claim (2.9) follows immediately from the first by an application of the Portmanteau theorem to the closed set $\{m \in \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}) : \ell_{\mathcal{C}^d \times \mathcal{V}}(m, \mathcal{R}^*) \geq \epsilon\}$. Let $P_n \in \mathcal{R}_n$ denote an $\epsilon_n$-optimal control for the $n$-state problem. Lemma 3.4 implies

$$\sup_n \frac{1}{n} \sum_{k=1}^n \mathbb{E}^{P_n} \int_0^T \int_A |a|^{p'} \Lambda_t^k(da) dt < \infty.$$

By Proposition 5.1, $(P_n)$ is precompact in $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$, and every limit point is concentrated on $\mathcal{R}$. Let $P$ denote a limit point, and relabel the subsequence so that $P_n \circ (\widehat{\mu}^n)^{-1} \to P$. Then $P(\mu \in \mathcal{R}) = 1$, and to prove $P(\mu \in \mathcal{R}^*) = 1$ it suffices to show that

(6.1)                    $\mathbb{E}^P[\Gamma(\mu)] \geq \Gamma(m)$  for all $m \in \mathcal{R}$.

First, use the upper semicontinuity of $\Gamma$ of Lemma 4.1 to get

$$\mathbb{E}^P[\Gamma(\mu)] \geq \limsup_{n \to \infty} \mathbb{E}^{P_n}[\Gamma(\widehat{\mu}^n)].$$

Fix $m \in \mathcal{R}$, and use Proposition 5.2 to find $Q_n \in \mathcal{R}_n$ such that $\mathbb{E}^{Q_n}[\Gamma(\widehat{\mu}^n)] \to \Gamma(m)$. The $\epsilon_n$-optimality of $P_n$ implies

$$\mathbb{E}^{P_n}[\Gamma(\widehat{\mu}^n)] \geq \mathbb{E}^{Q_n}[\Gamma(\widehat{\mu}^n)] - \epsilon_n.$$

Recalling that $\epsilon_n \to 0$, we have thus proved (6.1).                    □

*Proof of Theorem* 2.12. The second claim, that $P_n$ can be taken to be strong controls, follows from the first claim and from Theorem 2.10. Hence, we prove only the first claim. Define $L \subset \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$ to be the set of $\lim_{n \to} P_n \circ (\widehat{\mu}^n)^{-1}$ such that $P_n$ is an $\epsilon_n$-optimal control for each $n$ for some sequence $\epsilon_n \to 0$. Let $L_s$ denote the set of *subsequential* limits of such sequences. Naturally, write $\mathcal{P}(\mathcal{R}^*) = \{M \in \mathcal{P}(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})) : M(\mathcal{R}^*) = 1\}$. Because $\mathcal{R}^*$ is compact in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$ by Lemma 4.2, it follows that $\mathcal{P}(\mathcal{R}^*) \subset \mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$, that $\mathcal{P}^p(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$ and $\mathcal{P}(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$ induce the same topology on $\mathcal{P}(\mathcal{R}^*)$, and finally that $\mathcal{P}(\mathcal{R}^*)$ is compact with respect to either of these topologies. Henceforth, we work with the latter (weak convergence) topology on $\mathcal{P}(\mathcal{R}^*)$.

First note that

$$L \subset L_s \subset \mathcal{P}(\mathcal{R}^*),$$

where the first inclusion is obvious, and the second is the content of Theorem 2.11. The proof will be complete if we show that $\mathcal{P}(\mathcal{R}^*) \subset L$. Note that $\mathcal{P}(\mathcal{R}^*)$ is a compact convex set, and the set of extreme points is $\{\delta_m : m \in \mathcal{R}^*\}$. Hence, by the Krein–Milman theorem, to show that $\mathcal{P}(\mathcal{R}^*) \subset L$ it suffices to show that $L$ is closed and convex and that $\delta_m \in L$ for each $m \in \mathcal{R}^*$.

*Step* 1. We first show that $L$ is convex. Let $M^1, M^2 \in L$, so for $i = 1, 2$ we may find $\epsilon_n^i \to 0$ and $\epsilon_n^i$-optimal controls $P_n^i \in \mathcal{R}_n$ such that $M^i = \lim_{n \to \infty} P_n^i \circ (\widehat{\mu}^n)^{-1}$. Let $t \in (0, 1)$. Clearly $\epsilon_n = t\epsilon_n^1 + (1-t)\epsilon_n^2$ tends to zero. Moreover, $P_n := tP_n^1 + (1-t)P_n^2$ belongs to $\mathcal{R}_n$, because $\mathcal{R}_n$ is easily seen to be convex. Because the objective functional $\mathcal{R}_n \ni \mathcal{P} \mapsto \mathbb{E}^P[\Gamma(\widehat{\mu}^n)]$ is affine, $P_n$ is a $\epsilon_n$-optimal control. This shows $tM^1 + (1-t)M^2 = \lim_{n \to \infty} P_n$ belongs to $L$.

*Step* 2. We next show that $L$ is closed. For $n \geq 1$ and $\epsilon \geq 0$ define

$$S_n^\epsilon = \left\{ P_n \circ (\widehat{\mu}^n)^{-1} : P_n \in \mathcal{R}_n \text{ is an } \epsilon\text{-optimal control} \right\} \subset \mathcal{P}(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})).$$

Note that $S_n^\delta \subset S_n^\epsilon$ for all $\delta < \epsilon$. It is easy to see that $L$ is precisely the set of $M \in \mathcal{P}(\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V}))$ such that, for every $\epsilon > 0$ and every open set $U$ containing $M$, there exists $N$ such that $U \cap S_n^\epsilon \neq \emptyset$ for all $n \geq N$. Now fix $M \notin L$. Find $\epsilon > 0$, an open set $U$ containing $M$, and $n_k \to \infty$ such that $U \cap S_{n_k}^\epsilon = \emptyset$ for all $k$. Then $U \cap \overline{S}_{n_k}^\epsilon = \emptyset$. Define $V = U \backslash \bigcap_{k=1}^\infty \overline{S}_{n_k}^\epsilon$, and notice that $V$ is open. Clearly $V$ contains $M$. By construction, $V \cap S_{n_k}^\epsilon = \emptyset$ for each $k$, which shows that $V \subset L^c$. Thus $L^c$ is open.

*Step* 3. Finally, we show that $L$ contains the extreme points $\{\delta_m : m \in \mathcal{R}^*\}$ of $\mathcal{P}(\mathcal{R}^*)$. Fix $m^* \in \mathcal{R}^*$. By Proposition 5.2, there exist $P_n \in \mathcal{R}_n$ such that $P_n \circ (\widehat{\mu}^n)^{-1} \to \delta_{m^*}$ and $\mathbb{E}^{P_n}[\Gamma(\widehat{\mu}^n)] \to \Gamma(m^*)$. Let $P_n^* \in \mathcal{R}_n$ be an $n$-state optimal control for each $n$, the existence of which is guaranteed by Theorem 2.9. Let

$$\epsilon_n = \mathbb{E}^{P_n^*}[\Gamma(\widehat{\mu}^n)] - \mathbb{E}^{P_n}[\Gamma(\widehat{\mu}^n)].$$

Optimality of $P_n^*$ ensures that $\epsilon_n \geq 0$. By Theorem 2.11, every limit point of $P_n^* \circ (\widehat{\mu}^n)^{-1}$ is supported on $\mathcal{R}^*$. By the upper semicontinuity of $\Gamma$ (see Lemma 4.1), this implies

$$\limsup_{n\to\infty} \mathbb{E}^{P_n^*}[\Gamma(\widehat{\mu}^n)] \leq \sup_{m \in \mathcal{R}^*} \Gamma(m) = \Gamma(m^*) = \lim_{n\to\infty} \mathbb{E}^{P_n}[\Gamma(\widehat{\mu}^n)].$$

This shows $\epsilon_n \to 0$, completing the proof. $\qquad\square$

**7. Strong versus relaxed formulations.** This section is devoted to the proof of Theorem 2.4. The proof of Theorem 2.10 is nearly identical up to notational changes, so we omit it. Recall that we assume throughout that $p' = 2$. We start with a lemma due mostly to Méléard [42], modulo integrability issues.

LEMMA 7.1. *Suppose $M$ is a martingale measure with intensity $\Lambda_t(da)dt$, defined on some filtered probability space supporting a $\mathcal{P}^p(A)$-valued process $\Lambda$. Suppose $\Lambda_t^n(da)dt \to \Lambda_t(da)dt$ weakly, almost surely, for some other $\mathcal{P}^p(A)$-valued processes $\Lambda^n$ satisfying*

$$(7.1) \qquad \lim_{r\to\infty} \sup_n \mathbb{E}\left[ \int_0^T \int_{\{|a|>r\}} |a|^2 \Lambda_t^n(da)dt \right] < \infty.$$

*Then there exists a sequence of martingale measures $M^n$ (defined on an extension of the probability space) with intensities $\Lambda_t^n(da)dt$ such that*

$$(7.2) \qquad \lim_{n\to\infty} \mathbb{E}\left[ \left( \int_{A\times[0,T]} \varphi(t,a) M^n(da,dt) - \int_{A\times[0,T]} \varphi(t,a) M(da,dt) \right)^2 \right] = 0$$

*for every predictable function $\varphi(t,a)$, continuous in $a$, and satisfying $|\varphi(t,a)|^2 \leq c(Z + |a|^2)$ for all $a \in A$, for some integrable random variable $Z \geq 0$, and for some $c > 0$.*

*Proof.* The result of Méléard [42] provides a sequence of martingale measures $M^n$ with intensities $\Lambda_t^n(da)dt$ such that (7.2) holds for all *bounded* predictable $\varphi$ which

are continuous in $a$. For general $\varphi$, let $r > 0$ and compute

$$\mathbb{E}\left[\left(\int \varphi(t,a) M^n(da,dt) - \int \varphi(t,a) M(da,dt)\right)^2\right]$$

$$\leq 3\mathbb{E}\left[\left(\int \varphi(t,a) 1_{\{|\varphi|\leq r\}} M^n(da,dt) - \int \varphi(t,a) 1_{\{|\varphi|\leq r\}} M(da,dt)\right)^2\right]$$

$$(7.3) \quad + 3\mathbb{E}\left[\left(\int \varphi(t,a) 1_{\{|\varphi|>r\}} M^n(da,dt)\right)^2 + \left(\int \varphi(t,a) 1_{\{|\varphi|>r\}} M(da,dt)\right)^2\right].$$

The first term tends to zero, thanks to the aforementioned result of [42]. For the second term, notice that

$$\mathbb{E}\left[\left(\int \varphi(t,a) 1_{\{|\varphi|>r\}} M^n(da,dt)\right)^2\right]$$

$$= \mathbb{E}\left[\int |\varphi(t,a)|^2 1_{\{|\varphi|>r\}} \Lambda_t^n(da)dt\right]$$

$$\leq c\mathbb{E}\left[\int (Z + |a|^2) 1_{\{Z+|a|^2>r/c\}} \Lambda_t^n(da)dt\right]$$

$$\leq c\mathbb{E}\left[TZ1_{\{Z>r/2c\}} + \int |a|^2 1_{\{|a|^2>r/2c\}} \Lambda_t^n(da)dt\right].$$

Since $\mathbb{E}Z < \infty$ by assumption, this can be made arbitrarily small, uniformly in $n$, by sending $r \to \infty$. Similarly, by Fatou's lemma, the assumption (7.1) implies $\mathbb{E}\int_0^T \int_A |a|^2 \Lambda_t(da)dt < \infty$, and we see that choosing $r$ large can make the second term in line (7.3) arbitrarily small. $\quad\square$

*Proof of Theorem* 2.4. Let $m \in \mathcal{R}^*$, recalling that Theorem 2.2 ensures $\mathcal{R}^* \neq \emptyset$. By Proposition 2.5, there exists filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ supporting $(X, \Lambda, M)$ satisfying the following:

1. $X$ is a $d$-dimensional $\mathbb{F}$-adapted process.
2. $\Lambda$ is an $\mathbb{F}$-predictable $\mathcal{P}^p(A)$-valued process.
3. $M = (M^1, \ldots, M^{d_W})$ are orthogonal $\mathbb{F}$-martingale measures, each with intensity $\Lambda_t(da)dt$.
4. The McKean–Vlasov equation holds:

$$(7.4) \quad dX_t = \int_A b(t, X_t, m_t^x, a) \Lambda_t(da)dt + \int_A \sigma(t, X_t^i, m_t^x, a) M(da,dt).$$

5. The law of $(X, \Lambda)$ is precisely $m$.

By Lemma 3.2, we have (since $p' = 2$)

$$(7.5) \quad \mathbb{E}\left[\|X\|^2 + \int_0^T \int_A |a|^2 \Lambda_t(da)dt\right] < \infty.$$

We do a three-step approximation.

*Step* 1. We first approximate $\Lambda$ by bounded controls. For $n \geq 1$, let $\iota_n : A \to A$ be any measurable function such that $|\iota_n(a)| \leq n$ for all $a \in A$ and $\iota_n(a) = a$ when

$|a| \leq n$. Define $\Lambda_t^n = \Lambda_t \circ \iota_n^{-1}$, so that clearly $\Lambda_t^n(da)dt \to \Lambda_t(da)dt$ a.s. Note also that

(7.6)

$$\lim_{r \to \infty} \sup_n \mathbb{E}\left[\int_0^T \int_{\{|a|>r\}} |a|^2 \Lambda_t^n(da)dt\right] \leq \lim_{r \to \infty} \left[\int_0^T \int_{\{|a|>r\}} |a|^2 \Lambda_t(da)dt\right] = 0,$$

thanks to (7.5) and the simple observation that $|\iota_n(a)| \leq |a|$ for all $a \in A$. By Lemma 7.1, there exists (on an enlargement of the probability space) a sequence of orthogonal martingale measures $M^n = (M^{n,1}, \ldots, M^{n,d_W})$ such that $M^{n,i}$ has intensity measure $\Lambda_t^n(da)dt$ for each $i$ and

$$\lim_{n \to \infty} \mathbb{E}\left[\left(\int_{A \times [0,T]} \varphi(t,a)M^n(da,dt) - \int_{A \times [0,T]} \varphi(t,a)M(da,dt)\right)^2\right] = 0$$

for every predictable function $\varphi$ (from $[0,T] \times \Omega \times A$ to $\mathbb{R}^{d \times d_W}$) satisfying $|\varphi(t,a)|^2 \leq c(Z + |a|^2)$ for all $a \in A$, for some $c > 0$, and for some integrable random variable $Z$.

Let $X^n$ denote the unique solution of the McKean–Vlasov equation

$$dX_t^n = \int_A b\left(t, X_t^n, \mathbb{P} \circ (X_t^n)^{-1}, a\right) \Lambda_t^n(da)dt$$
$$+ \int_A \sigma\left(t, X_t^n, \mathbb{P} \circ (X_t^n)^{-1}, a\right) M^n(da,dt), \quad X_0^n = X_0.$$

Note that the Lipschitz assumption (Assumption B) ensures the well-posedness of this equation, by standard arguments. By Proposition 2.5 $\mathbb{P} \circ (X^n, \Lambda^n)^{-1}$ belongs to $\mathcal{R}$. The rest of this step is a long but straightforward proof, using the Lipschitz assumption, that $\mathbb{E}[\|X^n - X\|^2] \to 0$, from which the desired approximations will quickly follow. Recall the notation $\|x\|_t = \sup_{s \in [0,t]} |x_s|$ and also the truncated Wasserstein distance $\ell_{t,\mathcal{C}^d,p}$ from (5.10). Apply the Burkholder–Davis–Gundy inequality and Jensen's inequality to get (for a constant $C > 0$ which changes from line to line)

$$\mathbb{E}\left[\|X^n - X\|_t^2\right]$$
$$\leq C\mathbb{E}\left\{\int_0^t \int_A \left|b\left(s, X_s^n, \mathbb{P} \circ (X_s^n)^{-1}, a\right) - b(s, X_s, m_s^x, a)\right|^2 \Lambda_s^n(da)ds\right.$$
$$+ \left|\int_0^t \int_A b(s, X_s, m_s^x, a)\Lambda_s^n(da)ds - \int_0^t \int_A b(s, X_s, m_s^x, a)\Lambda_s(da)ds\right|^2$$
$$+ \left|\int_0^t \int_A \sigma\left(s, X_s^n, \mathbb{P} \circ (X_s^n)^{-1}, a\right) - \sigma(s, X_s, m_s^x, a)M^n(da,ds)\right|^2$$
$$+ \left.\left|\int_0^t \int_A \sigma(s, X_s, m_s^x, a)M^n(da,ds) - \int_0^t \int_A \sigma(s, X_s, m_s^x, a)M(da,ds)\right|^2\right\}$$
$$=: C\mathbb{E}[I_1 + I_2 + I_3 + I_4].$$

The Lipschitz assumption (Assumption B) yields

$$\mathbb{E}[I_1] \leq C \int_0^t \left(\|X^n - X\|_s^2 + \ell_{s,\mathcal{C}^d,2}^2\left(\mathbb{P} \circ (X^n)^{-1}, m^x\right)\right) dt$$

and also

$$\mathbb{E}[I_3] = \mathbb{E}\left[\int_0^t \int_A \left|\sigma\left(s, X_s^n, \mathbb{P}\circ(X_s^n)^{-1}, a\right) - \sigma(s, X_s, m_s^x, a)\right|^2 \Lambda_s^n(da)ds\right]$$

$$\leq C\int_0^t \left(\|X^n - X\|_s^2 + \ell_{s,\mathcal{C}^d,2}^2\left(\mathbb{P}\circ(X^n)^{-1}, m^x\right)\right)dt.$$

By Gronwall's inequality,

$$\mathbb{E}\left[\|X^n - X\|_t^2\right] \leq C\int_0^t \ell_{s,\mathcal{C}^d,2}^2\left(\mathbb{P}\circ(X^n)^{-1}, m^x\right)ds + C\mathbb{E}[I_2 + I_4].$$

Since $\mathbb{P}\circ(X^n, X)^{-1}$ is a coupling of $\mathbb{P}\circ(X^n)^{-1}$ and $m^x$, we have

$$\ell_{t,\mathcal{C}^d,2}^2\left(\mathbb{P}\circ(X^n)^{-1}, m^x\right) \leq \mathbb{E}\left[\|X^n - X\|_t^2\right],$$

and another application of Gronwall's inequality yields

$$\ell_{\mathcal{C}^d,2}^2\left(\mathbb{P}\circ(X^n)^{-1}, m^x\right) \leq C\mathbb{E}[I_2 + I_4].$$

Assumption (A.4) and $m^x = \mathbb{P}\circ X^{-1}$ imply (since $p < p' = 2$)

$$|b(s, X_s, m_s^x, a)|^2 \leq C\left(1 + \|X\|^2 + |a|^2\right),$$

and it follows from (7.5) and continuity of $b$ that $\mathbb{E}[I_2] \to 0$ as $n \to \infty$. Similarly, Assumption (A.4) implies (since $p_\sigma \leq 2$)

$$|\sigma(s, X_s, m_s^x, a)|^2 \leq C\left(1 + \|X\|^2 + |a|^2\right),$$

and it follows from (7.5), continuity of $\sigma$, and Lemma 7.1 that $\mathbb{E}[I_4] \to 0$ as $n \to \infty$. We finally conclude that

$$\lim_{n\to\infty}\mathbb{E}\left[\|X^n - X\|^2\right] = 0,$$

and so $\mathbb{P}\circ(X^n, \Lambda^n)^{-1} \to m$ in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$. It follows from Lemma 4.1 and (7.6) that $\Gamma(\mathbb{P}\circ(X^n, \Lambda^n)^{-1}) \to \Gamma(m)$.

*Step* 2. In light of Step 1, we may now assume without loss of generality that our original control $\Lambda$ is bounded, in the sense that there exists $r > 0$ such that $\Lambda([0,T] \times B_r) = 0$ a.s., where $B_r$ denotes the centered ball of radius $r$. We now use the chattering lemma [33, Theorem 2.2(b)] (originally due to Fleming [19]) to find a sequence of progressively measurable $B_r$-valued processes $(\alpha_t^n)_{t\in[0,T]}$ such that $\delta_{\alpha_t^n}(da)dt \to \Lambda_t(da)dt$ a.s., and by another result of Méléard [42] we can find a $d_W$-dimensional Wiener process $W$ (again by extending the probability space) such that

$$\lim_{n\to\infty}\mathbb{E}\left[\left(\int_0^T \varphi(t, \alpha_t^n)dW_t - \int_{B_r\times[0,T]}\varphi(t,a)M(da,dt)\right)^2\right] = 0$$

for every bounded predictable random $\mathbb{R}^{d\times d_W}$-valued function $\varphi$, where $M$ is again a vector $M = (M^1, \ldots, M^{d_W})$ of orthogonal martingale measures, each with intensity $\Lambda_t(da)dt$.[2] As in Lemma 7.1, we can relax the boundedness assumption, as long as

---

[2] This does not follow immediately from Lemma 7.1. The key point is that we have a *single* Wiener process $W$, whereas Lemma 7.1 would yield a *sequence* $W^n$.

there is an integrable random variable $Z \geq 0$ such that $|\varphi(t, a)| \leq Z$ a.s. Define $X^n$ as the unique solution of the McKean–Vlasov SDE

$$dX_t^n = b\left(t, X_t^n, \mathbb{P} \circ (X_t^n)^{-1}, \alpha_t^n\right) dt + \sigma\left(t, X_t^n, \mathbb{P} \circ (X_t^n)^{-1}, \alpha_t^n\right) dW_t, \ X_0^n = X_0.$$

We prove exactly as in Step 1 that $\mathbb{E}\|X^n - X\|^2 \to 0$, and so $\mathbb{P} \circ (X^n, \delta_{\alpha_t^n}(da)dt)^{-1} \to \mathbb{P} \circ (X, \Lambda)^{-1} = m$ in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$. Because the controls are uniformly bounded by $r$, the uniform integrability condition of Lemma 4.1 holds trivially, and we conclude that $\Gamma(\mathbb{P} \circ (X^n, \delta_{\alpha_t^n}(da)dt)^{-1}) \to \Gamma(m)$.

*Step* 3. In light of Step 2, we now assume without loss of generality that our original $m$ and our filtered probability space supports process $(X, \alpha, W)$ satisfying the following:

1. $X$ is a $d$-dimensional $\mathbb{F}$-adapted process.
2. $\alpha$ is an $\mathbb{F}$-predictable $A$-valued process, uniformly bounded in norm by a constant $r > 0$.
3. $W$ is a $d_W$-dimensional Wiener process.
4. The McKean–Vlasov equation holds:

$$(7.7) \qquad dX_t = b(t, X_t, m_t^x, \alpha_t)dt + \sigma(t, X_t, m_t^x, \alpha_t)dW_t.$$

5. The law of $(X, \delta_{\alpha_t}(da)dt)$ is precisely $m$.

The final step is to approximate $\alpha$ in a weak sense by controls which are *strong*, i.e., progressively measurable with respect to the filtration $\mathbb{F}^W = (\sigma(X_0, W_s : s \leq t))_{t \in [0,T]}$ generated by the Wiener process and initial state. For this we appeal to [13, Lemma 3.11] to find a sequence of $\mathbb{F}^W$-progressively measurable $A$ processes $(\alpha_t^n)_{t \in [0,T]}$ which share the same uniform bound as $\alpha$, such that

$$(7.8) \qquad \mathbb{P} \circ (dt\delta_{\alpha_t^n}(da), W)^{-1} \to \mathbb{P} \circ (\delta_{\alpha_t}(da)dt, W)^{-1}, \ \text{in } \mathcal{P}^p\left(\mathcal{V} \times \mathcal{C}^{d_W}\right).$$

In other words, $\alpha^n$ are *strong controls* which approximate $\alpha$ in joint law. Thanks to Assumption B, there exists a unique strong solution $X^k$ of the McKean–Vlasov SDE

$$dX_t^n = b\left(t, X_t^n, \mathbb{P} \circ (X_t^n)^{-1}, \alpha_t^n\right) dt + \sigma\left(t, X_t^n, \mathbb{P} \circ (X_t^n)^{-1}\right) dW_t, \qquad X_0^n = X_0.$$

Define $m^n = \mathbb{P} \circ (X^n, dt\delta_{\alpha_t^n}(da))^{-1}$, and note that $m^n$ belongs to the set $\mathcal{R}^s$ of strong controls. Now, suppose that we can show $m^n \to m$ in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$. Because $\alpha^n$ are uniformly bounded, it will then follow from Lemma 3.2 that

$$\sup_n \mathbb{E}\left[\|X^n\|^{p''}\right] < \infty,$$

where we recall that $p'' > p' = 2$ and $\int |x|^{p''} \lambda(dx) < \infty$ by assumption. This is enough to verify the uniform integrability condition of Lemma 4.1, which shows that $\Gamma(m^n) \to \Gamma(m)$. Hence, it remains to prove that $m^n \to m$ in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V})$.

It follows from Proposition 3.5 and the uniformly boundedness of $\alpha^n$ that the sequence $Q_n := \mathbb{P} \circ (X^n, dt\delta_{\alpha_t^n}(da), W)^{-1}$ is precompact in $\mathcal{P}^p(\mathcal{C}^d \times \mathcal{V} \times \mathcal{C}^{d_W})$. Now let $(\overline{X}, \overline{\Lambda}, \overline{W})$ denote the canonical process on $\mathcal{C}^d \times \mathcal{V} \times \mathcal{C}^{d_W}$. If $Q$ denotes any limit point of $Q_n$, it is clear from (7.8) that

(7.9)
$$Q \circ (\overline{\Lambda}, \overline{W}, \overline{X}_0)^{-1} = \lim_{n \to \infty} \mathbb{P} \circ (\delta_{\alpha_t^n}(da)dt, W, X_0)^{-1} = \mathbb{P} \circ (\delta_{\alpha_t}(da)dt, W, X_0)^{-1}.$$

Thus, under $Q$, there exists a $\sigma(\overline{X}_0, \overline{W}_s : s \leq t)$-progressively measurable $A$-valued process $\overline{\alpha}$ such that $\overline{\Lambda}_t = \delta_{\overline{\alpha}_t}$ for a.e. $t$, almost surely. We then argue using continuity

of the coefficients $(b, \sigma)$ (either using martingale problems or the results of Kurtz and Protter [36]) that under $Q$ the following SDE holds:

$$d\overline{X}_t = b\left(t, \overline{X}_t, \mathbb{P} \circ \overline{X}_t^{-1}, \overline{\alpha}_t\right) dt + \sigma\left(t, \overline{X}_t, \mathbb{P} \circ \overline{X}_t^{-1}, \overline{\alpha}_t\right) d\overline{W}_t.$$

Because of (7.9), the uniqueness in law of the SDE (7.7) implies $Q = \mathbb{P} \circ (X, \delta_{\alpha_t}(da) dt, W)^{-1}$. (See [35], particularly Example 2.14 therein, for a more careful discussion of the Yamada–Watanabe theorem in the context of McKean–Vlasov equations.) This holds for every limit point $Q$, and we conclude that

$$\mathbb{P} \circ (X^n, dt\delta_{\alpha_t^n}(da), W)^{-1} \to \mathbb{P} \circ (X, \delta_{\alpha_t}(da)dt, W)^{-1}, \ \text{ in } \mathcal{P}^p(\mathcal{C}^d \times \mathcal{V} \times \mathcal{C}^{d_W}).$$

Marginalizing yields

$$m^n = \mathbb{P} \circ (X^n, dt\delta_{\alpha_t^n}(da))^{-1} \to \mathbb{P} \circ (X, \delta_{\alpha_t}(da)dt)^{-1} = m \ \text{ in } \mathcal{P}^p\left(\mathcal{C}^d \times \mathcal{V}\right). \qquad \square$$

### Appendix A. A note on the filtration of $\mathcal{C}^d \times \mathcal{V}$.

The goal of this section is to clarify a technical point that arose in the proofs of Lemma 4.2 and Proposition 5.1. Recall that $(X, \Lambda)$ denotes the identity map on $\mathcal{C}^d \times \mathcal{V}$. Recall also that the natural filtration on $\mathcal{C}^d \times \mathcal{V}$ is defined as

$$\mathcal{F}_t = \sigma(X_s, \Lambda([0, s] \times B) : s \le t, B \subset A \text{ Borel}).$$

If $E$ is a Polish space and $\mathcal{G}$ a sub-$\sigma$-field of the Borel sets, let us say a family $\Phi$ of $\mathcal{G}$-measurable functions is a *separating class for $\mathcal{G}$* if $\int \varphi \, d\mu = \int \varphi \, d\nu$ for all $\varphi \in \Phi$ implies $\mu = \nu$ on $\mathcal{G}$, whenever $\mu, \nu \in \mathcal{P}(E)$. When $\mathcal{G}$ is the entire Borel $\sigma$-field, simply say that $\Phi$ is a *separating class*. It is known that every Polish space admits a countable separating class consisting of bounded continuous functions. Note that if $\Phi$ is separating, then $\int \varphi \, d\mu = \int \varphi \, d\nu$ for all $\varphi \in \Phi$ implies $\mu = \nu$ for all bounded signed measures $\mu, \nu$, which is spanned by the space of probability measures. In particular, if $\Phi$ is separating for $\mathcal{G}$, and if $X$ and $Y$ are random variables satisfying $\mathbb{E}[X\varphi] = \mathbb{E}[Y\varphi]$ for all $\varphi \in \Phi$, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[Y|\mathcal{G}]$ a.s.

LEMMA A.1. *For each $t \in [0, T]$, $\mathcal{F}_t$ admits a countable separating class $\Phi_t$ of bounded continuous functions.*

*Proof.* The statement is clearly true for $t = 0$, as $\mathcal{F}_0 = \sigma(X_0)$ can be identified with the Borel $\sigma$-field of $\mathbb{R}^d$. Fix $t > 0$. We deal with $X$ and $\Lambda$ separately. For $X$, notice that the restriction of $x \in \mathcal{C}^d$ to $x_{\cdot \wedge t} \in C([0, t]; \mathbb{R}^d)$ is a continuous operation and that the Polish space $C([0, t]; \mathbb{R}^d)$ admits a countable separating class $\Phi_t^X$ of bounded continuous functions. For $\Lambda$, define first the restriction of any $q \in \mathcal{V}$ by $q^t(\cdot) = t^{-1}q(([0, t] \times A) \cap \cdot) \in \mathcal{P}^p([0, t] \times A)$. The map $q \mapsto q^t$ is continuous from $\mathcal{V}$ to $\mathcal{P}^p([0, t] \times A)$ (see, e.g., [39, Corollary A.3]). As a Polish space, $\mathcal{P}^p([0, t] \times A)$ admits a countable separating class $\Phi_t^\Lambda$ of bounded continuous functions. Without loss of generality, assume that the constant function 1 belongs to both $\Phi_t^X$ and $\Phi_t^\Lambda$. Finally, define $\Phi_t$ to be the set of functions of the form $\mathcal{C}^d \times \mathcal{V} \ni (x, q) \mapsto \varphi(x_{\cdot \wedge t})\psi(q^t)$, where $\varphi \in \Phi_t^X$ and $\psi \in \Phi_t^\Lambda$. Then $\Phi_t$ fits the bill; see [15, Proposition 3.4.6] for a proof that $\Phi_t$ is separating. $\square$

## REFERENCES

[1] D. Andersson and B. Djehiche, *A maximum principle for SDEs of mean-field type*, Appl. Math. Optim., 63 (2011), pp. 341–356.

[2] K. Bahlali, M. Mezerdi, and B. Mezerdi, *Existence of optimal controls for systems governed by mean-field stochastic differential equations*, Afr. Stat., 9 (2014), pp. 627–645.

[3] E. Bayraktar, A. Cosso, and H. Pham, *Randomized Dynamic Programming Principle and Feynman–Kac Representation for Optimal Control of McKean–Vlasov Dynamics*, preprint, arXiv:1606.08204, 2016.

[4] A. Bensoussan, J. Frehse, and P. Yam, *Mean Field Games and Mean Field Type Control Theory*, Springer, Berlin, 2013.

[5] A. Bensoussan, J. Frehse, and P. Yam, *The master equation in mean field theory*, J. Math. Pures Appl., 103 (2015), pp. 1441–1474.

[6] G. Brunick and S. Shreve, *Mimicking an Itô process by a solution of a stochastic differential equation*, Ann. Appl. Probab., 23 (2013), pp. 1584–1628.

[7] R. Buckdahn, B. Djehiche, and J. Li, *A general stochastic maximum principle for SDEs of mean-field type*, Appl. Math. Optim., 64 (2011), pp. 197–216.

[8] A. Budhiraja, P. Dupuis, and M. Fischer, *Large deviation properties of weakly interacting processes via weak convergence methods*, Ann. Probab., 40 (2012), pp. 74–102.

[9] R. Carmona and F. Delarue, *Forward-backward stochastic differential equations and controlled McKean–Vlasov dynamics*, Ann. Prob., 43 (2015), pp. 2647–2700.

[10] R. Carmona and F. Delarue, *The master equation for large population equilibriums*, Stochastic Analysis and Applications 2014, Springer Proc. Math. Stat., Springer, Berlin, 2014, pp. 77–128.

[11] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications*, expected, Springer, Berlin, 2016.

[12] R. Carmona, F. Delarue, and A. Lachapelle, *Control of McKean–Vlasov dynamics versus mean field games*, Math. Financ. Econ., 7 (2013), pp. 131–166.

[13] R. Carmona, F. Delarue, and D. Lacker, *Mean field games with common noise*, Ann. Prob., 44 (2016), pp. 3740–3803.

[14] F. Dufour and R.H. Stockbridge, *On the existence of strict optimal controls for constrained, controlled Markov processes in continuous time*, Stochastics Internat. J. Probab. Stochastic Process., 84 (2012), pp. 55–78.

[15] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence*, 2nd ed., Vol. 282, Wiley-Interscience, New York, 2005.

[16] A.F. Filippov, *On certain questions in the theory of optimal control*, J. SIAM, Control, 1 (1962), pp. 76–84.

[17] M. Fischer, *On the connection between symmetric n-player games and mean field games*, Ann. Appl. Probab., to appear.

[18] M. Fischer and G. Livieri, *Continuous time mean-variance portfolio optimization through the mean field approach*, ESAIM Probab. Stat., 20 (2016), pp. 30–44.

[19] W.H. Fleming, *Generalized solutions in optimal stochastic control*, Differential Games and Control Theory, Kingston Conference 2, 30 (1977), pp. 147–165.

[20] M. Fornasier and F. Solombrino, *Mean-field optimal control*, ESAIM Control Optim. Calc. Var., 20 (2014), pp. 1123–1152.

[21] T. Funaki, *A certain class of diffusion processes associated with nonlinear parabolic equations*, Z. Wahr, Verwandte Gebiete, 67 (1984), pp. 331–348.

[22] J. Gärtner, *On the McKean–Vlasov limit for interacting diffusions*, Math. Nachr., 137 (1988), pp. 197–248.

[23] C. Graham and S. Méléard, *Stochastic particle approximations for generalized Boltzmann models and convergence estimates*, Ann. Probab., 25 (1997), pp. 115–132.

[24] I. Gyöngy, *Mimicking the one-dimensional marginal distributions of processes having an Itô differential*, Probab. Theory Related Fields, 71 (1986), pp. 501–516.

[25] U.G. Haussmann and J.P. Lepeltier, *On the existence of optimal controls*, SIAM J. Control Optim., 28 (1990), pp. 851–902.

[26] M. Huang, P. Caines, and R. Malhamé, *Individual and mass behaviour in large population stochastic wireless power control problems: Centralized and Nash equilibrium solutions*, in Proceedings of the 42nd IEEE Conference on Decision and Control 2003, Vol. 1, IEEE, Piscataway, NJ, 2003, pp. 98–103.

[27] M. Huang, P. Caines, and R. Malhamé, *Social optima in mean field LQG control: Centralized and decentralized strategies*, IEEE Trans. Automat. Control, 57 (2012), pp. 1736–1751.

[28] M. Huang, R. Malhamé, and P. Caines, *Large population stochastic dynamic games: Closed-loop McKean–Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst., 6 (2006), pp. 221–252.

[29] B. Jourdain and S. Méléard, *Propagation of chaos and fluctuations for a moderate model with smooth initial data*, Ann. Inst. Henri Poincaré Probab. Stat., 34 (1998), pp. 727–766.

[30] B. Jourdain and J. Reygner, *Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation*, Stoch. Partial Differ. Equ. Anal. Comput., 1 (2013), pp. 455–506.

[31] N. El Karoui and S. Méléard, *Martingale measures and stochastic calculus*, Probab. Theory Related Fields, 84 (1990), pp. 83–101.

[32] N. El Karoui, D.H. Nguyen, and M. Jeanblanc-Picqué, *Compactification methods in the control of degenerate diffusions: Existence of an optimal control*, Stochastics, 20 (1987), pp. 169–219.

[33] N. El Karoui, D.H. Nguyen, and M. Jeanblanc-Picqué, *Existence of an optimal Markovian filter for the control under partial observations*, SIAM J. Control Optim., 26 (1988), pp. 1025–1061.

[34] N. El Karoui and X. Tan, *Capacities, Measurable Selection and Dynamic Programming Part II: Application in Stochastic Control Problems*, preprint, arXiv:1310.3364, 2013.

[35] T.G. Kurtz, *Weak and strong solutions of general stochastic models*, Electron. Commun. Probab., 19 (2014), 58.

[36] T.G. Kurtz and P. Protter, *Weak limit theorems for stochastic integrals and stochastic differential equations*, Ann. Probab., 19 (1991), pp. 1035–1070.

[37] T.G. Kurtz and R.H. Stockbridge, *Existence of Markov controls and characterization of optimal Markov controls*, SIAM J. Control Optim., 36 (1998), pp. 609–653.

[38] D. Lacker, *A general characterization of the mean field limit for stochastic differential games*, Probab. Theory Related Fields, 165 (2014), pp. 581–648.

[39] D. Lacker, *Mean field games via controlled martingale problems: Existence of Markovian equilibria*, Stochastic Process. Appl., 125 (2015), pp. 2856–2894.

[40] J.M. Lasry and P.L. Lions, *Mean field games*, Japanese J. Math., 2 (2007), pp. 229–260.

[41] M. Lauriere and O. Pironneau, *Dynamic programming for mean-field type control*, C. R. Math., 352 (2014), pp. 707–713.

[42] S. Méléard, *Martingale Measure Approximation, Application to the Control of Diffusions*, University of Paris VI, 1992.

[43] S. Méléard, *Asymptotic behaviour of some interacting particle systems; McKean–Vlasov and Boltzmann models*, Probab. Models Nonlinear Partial Differential Equations, Springer, Berlin, 1996, pp. 42–95.

[44] S. Méléard and S. Roelly-Coppoletta, *A propagation of chaos result for a system of particles with moderate interaction*, Stochastic Process. Appl., 26 (1987), pp. 317–332.

[45] K. Oelschläger, *A martingale approach to the law of large numbers for weakly interacting stochastic processes*, Ann. Probab., 12 (1984), pp. 458–479.

[46] K. Oelschläger, *A law of large numbers for moderately interacting diffusion processes*, Z. Wahr. Verwandte Gebiete, 69 (1985), pp. 279–322.

[47] H. Pham and X. Wei, *Bellman equation and viscosity solutions for mean-field stochastic control problem*, preprint, arXiv:1512.07866, 2015.

[48] H. Pham and X. Wei, *Dynamic Programming for Optimal Control of Stochastic McKean–Vlasov Dynamics*, SIAM J. Control Optim., 55 (2017), pp. 1069–1101.

[49] M. Shkolnikov, *Large systems of diffusions interacting through their ranks*, Stochastic Process. Appl., 122 (2012), pp. 1730–1747.

[50] D.W. Stroock and S.R.S. Varadhan, *Multidimensional Diffusion Processes*, 2nd ed., Grundlehren Math. Wiss., Springer, Berlin, 1979.

[51] A.S. Sznitman, *Topics in propagation of chaos*, Ecole d'eté de Probabilités de Saint-Flour XIX—1989, Lecture Notes in Math. 1464, Springer, Berlin, 1991, pp. 165–251.

[52] C. Villani, *Topics in Optimal Transportation*, Grad. Stud. Math., AMS, Providence, RI, 2003.

[53] J.B. Walsh, *An Introduction to Stochastic Partial Differential Equations*, Springer, Berlin, 1986.