Stoch PDE: Anal Comp (2024) 12:1265–1373
https://doi.org/10.1007/s40072-023-00301-0

Moderate deviations for fully coupled multiscale weakly interacting particle systems

Z. W. Bezemek[1] **· K. Spiliopoulos**[1]

Received: 25 March 2022 / Revised: 3 May 2023 / Accepted: 26 May 2023 /
Published online: 10 July 2023

Abstract

We consider a collection of fully coupled weakly interacting diffusion processes moving in a two-scale environment. We study the moderate deviations principle of the empirical distribution of the particles' positions in the combined limit as the number of particles grow to infinity and the time-scale separation parameter goes to zero simultaneously. We make use of weak convergence methods, which provide a convenient representation for the moderate deviations rate function in a variational form in terms of an effective mean field control problem. We rigorously obtain equivalent representation for the moderate deviations rate function in an appropriate "negative Sobolev" form, proving their equivalence, which is reminiscent of the large deviations rate function form for the empirical measure of weakly interacting diffusions obtained in the 1987 seminal paper by Dawson–Gärtner. In the course of the proof we obtain related ergodic theorems and we consider the regularity of Poisson type of equations associated to McKean–Vlasov problems, both of which are topics of independent interest. A novel "doubled corrector problem" is introduced in order to control derivatives in the measure arguments of the solutions to the related Poisson equations used to control behavior of fluctuation terms.

Keywords Interacting particle systems · Multiscale processes · Empirical measure · Moderate deviations

---

This work has been partially supported by the National Science Foundation (DMS 2107856) and Simons Foundation Award 672441. The authors of the paper would like to thank both reviewers for a very careful and constructive review of this article.

✉ K. Spiliopoulos
kspiliop@bu.edu

Z. W. Bezemek
bezemek@bu.edu

[1] Department of Mathematics and Statistics, Boston University, 111 Cummington Mall, Boston, MA 02215, USA

Mathematics Subject Classification 60F10 · 60F05

1 Introduction

The purpose of this paper is to study the moderate deviations principle (MDP) for slow–fast interacting particle systems. In particular, we consider the system

$$\begin{aligned} dX^{i,\epsilon,N}_t &= \left[\frac{1}{\epsilon}b(X^{i,\epsilon,N}_t, Y^{i,\epsilon,N}_t, \mu^{\epsilon,N}_t) + c(X^{i,\epsilon,N}_t, Y^{i,\epsilon,N}_t, \mu^{\epsilon,N}_t)\right]dt \\ &\quad + \sigma(X^{i,\epsilon,N}_t, Y^{i,\epsilon,N}_t, \mu^{\epsilon,N}_t)dW^i_t \\ dY^{i,\epsilon,N}_t &= \frac{1}{\epsilon}\left[\frac{1}{\epsilon}f(X^{i,\epsilon,N}_t, Y^{i,\epsilon,N}_t, \mu^{\epsilon,N}_t) + g(X^{i,\epsilon,N}_t, Y^{i,\epsilon,N}_t, \mu^{\epsilon,N}_t)\right]dt \\ &\quad + \frac{1}{\epsilon}\left[\tau_1(X^{i,\epsilon,N}_t, Y^{i,\epsilon,N}_t, \mu^{\epsilon,N}_t)dW^i_t + \tau_2(X^{i,\epsilon,N}_t, Y^{i,\epsilon,N}_t, \mu^{\epsilon,N}_t)dB^i_t\right] \\ &\quad (X^{i,\epsilon,N}_0, Y^{i,\epsilon,N}_0) = (\eta^x, \eta^y) \end{aligned} \tag{1}$$

on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ with $\{\mathcal{F}_t\}$ satisfying the usual conditions, where $b, c, \sigma, f, g, \tau_1, \tau_2 : \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$, B^i_t, W^i_t are independent standard 1-D $\mathcal{F}_t$-Brownian motions for $i = 1, \dots, N$, and $(\eta^x, \eta^y) \in \mathbb{R}^2$. Here and throughout $\mathcal{P}_2(\mathbb{R})$ denotes the space of probability measures on $\mathbb{R}$ with finite second moment, equipped with the 2-Wasserstein metric (see Appendix 4). $\mu^{\epsilon,N}$ is defined by

$$\mu^{\epsilon,N}_t = \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,\epsilon,N}_t}, t \in [0, T]. \tag{2}$$

In (1), $X^{i,\epsilon,N}$ and $Y^{i,\epsilon,N}$ represent the slow and fast motion respectively of the i^{th} component. Note that classical models of interacting particles in a two-scale potential, see [4, 16, 20, 38], can be thought of as special cases of (1) with $Y^{i,\epsilon,N} = X^{i,\epsilon,N}/\epsilon$.

Assume that $\epsilon(N) \to 0$ as $N \to \infty$. In our case, moderate deviations amounts to studying the behavior of the empirical measure of the particles, i.e., of $\mu^{\epsilon,N}$ in the regime between fluctuations and large deviations behavior. In particular, if we denote by $\mathcal{L}(X)$ the process at which $\mu^{\epsilon,N}$ converges to (the law of the averaged McKean–Vlasov Eq. 25) and consider the moderate deviation scaling sequence $\{a(N)\}_{N\in\mathbb{N}}$ such that $a(N) > 0, \forall N \in \mathbb{N}$ with $a(N) \to 0$ and $a(N)\sqrt{N} \to \infty$ as $N \to \infty$, the moderate deviations process is defined to be

$$Z^N_t := a(N)\sqrt{N}(\mu^{\epsilon,N}_t - \mathcal{L}(X_t)), t \in [0, T]. \tag{3}$$

The goal of this paper is to derive the large deviations principle with speed $a^{-2}(N)$ for the process Z^N_t, which is the moderate deviations principle for the measure-valued process $\mu^{\epsilon,N}_t$. Notice that if $a(N) = 1$ then we get the standard fluctuations process whose limiting behavior amounts to fluctuations around the law of large numbers, $\mathcal{L}(X_t)$, whereas if $a(N) = 1/\sqrt{N}$ then we would be in the large deviations regime.

We remark here that due to the effect of multiple scales, it turns out that a relation between ϵ and N is needed. So beyond requiring $a(N) \to 0$ and $a(N)\sqrt{N} \to \infty$, we also require that there exists $\rho \in (0, 1)$ and $\lambda \in (0, \infty]$ such that $a(N)\sqrt{N}\epsilon(N)^{\rho} \to \lambda$ as $N \to \infty$. Note that this should be viewed as a restriction on the scaling sequence $a(N)$, not on the relationship between ϵ and N, and in some regimes we expect this assumption can be weakened. See Remark 6.2.

The presence of multiple scales is a common feature in a range of models used in various disciplines ranging from climate modeling to chemical physics to finance, see for example [7, 26, 29, 42, 52, 74] for a representative, but by no means complete list of references. Interacting diffusions have also been the central topic of study in science and engineering, see for example [6, 34, 35, 43, 51, 53] to name a few. In the absence of multiple scales, i.e., when $\epsilon = 1$, law of large numbers, fluctuations and large deviations behavior as $N \to \infty$ has been studied in the literature, see [9, 16, 17]. Analogously, in the case of $N = 1$, the behavior as $\epsilon \downarrow 0$, have been extensively studied in the literature, see for example [1, 23, 30, 32, 44, 56–58, 60, 61, 66–68, 71, 72].

Homogenization of McKean–Vlasov equations (equations that are the limit of $N \to \infty$ with ϵ fixed) has also been recently studied in the literature, see e.g., [5, 40, 48, 65]. These results can be thought of as looking at the limit of the system (1) when first $N \to \infty$ and then $\epsilon \to 0$. Large deviations for a special case of (1) has been recently established in [4] and in the absence of multiple scales in [9]. In [59] the author studies large deviations for interacting particle systems in the absence of multiple scales but in the joint mean-field and small-noise limit. In the absence of multiple scales, i.e., when $\epsilon = 1$, moderate deviations for interacting particle systems have been studied in [10].

The contributions of this work are fourfold. Firstly, we investigate the combined limit $N \to \infty$ and $\epsilon \to 0$ for the fully coupled interacting particle system of McKean–Vlasov type (1) through the lens of moderate deviations. In order to do so, we use the weak convergence methodology developed in [22] which leads to the study of (appropriately linearized) optimal stochastic control problems of McKean–Vlasov type, see for example [13, 31, 50]. The first main result of this paper is Theorem 3.2 which provides a variational representation of the moderate deviations rate function.

Secondly, we rigorously re-express the obtained variational form of the rate function in the "negative Sobolev" form given in Theorem 5.1 of the seminal paper by Dawson–Gärtner [17] in the absence of multiple scales. Hence, we rigorously establish the equivalence of the two formulations in the moderate deviations setting, see Proposition 4.3. A connection of this form was recently established rigorously for the first time in the large deviations setting in [4].

Thirdly, in the process of establishing the MDP, we derive related ergodic theorems for multiscale interacting particle systems that are of independent interest. Due to the nature of moderate deviations, we need to consider certain solutions of Poisson equations whose properties are considered for the first time in this paper. In particular, we must control a term involving a derivative in the measure argument of the solution to the Poisson Eq. (22) (known as the Cell-Problem in the periodic setting). Such terms are unique to slow–fast interacting particle systems and slow–fast McKean–Vlasov SDEs, and thus do not appear whatsoever in proofs of averaging in the one-particle setting. Thus, the “doubled corrector problem” construction, (63), and the method of proof of Proposition 6.3 are novel ideas here, see also [5].

Fourthly, in contrast to [10], in this paper the coefficients of the model need not depend on the measure parameter in an affine way. We allow the coefficients of the interacting particle system (1) to have any dependence on the measure μ, so long that it is sufficiently smooth- see Corollary 3.4 and Remark 3.5. This is thanks to Lemma D.7, which is inspired by Lemma 5.10 in [19], and allows us to see that with sufficient regularity of a functional on $\mathcal{P}_2(\mathbb{R})$, the L^2 error of that functional evaluated at the empirical measure of N IID random variables and the Law of those random variables is $\mathcal{O}(1/N)$ as $N \to \infty$.

In Sect. 4.2, we make our general results concrete for a popular model of interacting particles in a two-scale potential, see also [38] for a motivating example in this direction. In addition, we present in Sect. 3 a number of concrete examples where the conditions of this paper hold.

The identification of the optimal change of measure in the moderate deviations lower bound through feedback controls together with the equivalence proof between the variational formulation and the “negative Sobolev” form of the rate function, open the door to a rigorous study of provably-efficient accelerated Monte–Carlo schemes for rare events computation, analogous to what has been accomplished in the one particle case, see e.g., [24, 57]. Exploring this is beyond the scope of this work and will be addressed elsewhere.

In addition, [16] remarks that phase transitions can occur at the level of fluctuations for interacting particle systems. Since the moderate deviations principle is essentially a large deviations statement around the fluctuations, the results obtained in this paper can potentially be related to phase transitions and allow to characterize them further. This dynamical systems direction is left for future work as it is also outside the scope of this paper.

In contrast to large deviations, the main difficulty with moderate deviations lies in the tightness proof, where we use an appropriate coupling argument, as well as in the fact that the space of signed measures is not completely metrizable in the topology of weak convergence (see [18] Remark 1.2, as well as [63] Remarks 2.2 and 2.3 for further discussion on related issues). Thus, as we will see, we will have to study Z^N as a distribution-valued process on a suitable weighted Sobolev space. In addition, the presence of the multiple scales complicates the required estimates because the ergodic behavior needs to be accounted for as well. The coupling argument used in the proof of tightness is non-standard in that the IID particle system used as an intermediary process between the empirical measure $\mu^{\epsilon,N}$ from Eq. (2) and its homogenized McKean–Vlasov limit $\mathcal{L}(X)$ from Eq. (25) is not equal in distribution to X. Instead, it is an IID system of slow–fast McKean Vlasov SDEs—see Eq. (57). Thus our proof of tightness is in a sense relying on the fact that the limits $N \to \infty$ and $\epsilon \downarrow 0$ for the empirical measure (2) commute at the level of the law of large numbers. For a further discussion of this, see Remark 5.1 and the discussion at the beginning of Sect. 7.

The rest of the paper is organized as follows. In Sect. 2, we introduce the appropriate topology for Z^N and lay out our main assumptions. We also introduce a quite useful multi-index notation that will allow us to circumvent notational difficulties with various combinations of mixed derivatives that appear throughout the paper. The derivation of the moderate deviations principle is based on the weak convergence approach of [22] which converts the large deviations problem to weak convergence of an appropriate stochastic control problem. The main result is presented in Sect. 3, Theorem 3.2. In Sect. 4 we prove an alternative form of the rate function. This form provides a rigorous connection in moderate deviations between the "variational form" of the rate function for the empirical measure of weakly interacting particle systems proved in Theorem 3.2 to the "negative Sobolev" form given in Theorem 5.1 of the seminal paper by Dawson–Gärtner [17]. Corollaries 3.4 and 4.6 specialize the discussion to the setting without multiscale structure and thus generalizes the results of [10]. Specific examples are presented in Sect. 4.2. Section 5 formulates the appropriate stochastic control problem.

Sections 6–10 are devoted to the proof of Theorem 3.2. Due to the presence of the multiple scales, ergodic theorems are needed to characterize the behavior as $\epsilon \downarrow 0$ of certain functionals of interest for the controlled multiscale interacting particle system; this is the content of Sect. 6. Tightness of the controlled system is proven in Sect. 7. In Sect. 8 we establish the limiting behavior of the controlled system. The Laplace principle lower bound is proven in Sect. 9. Section 10 contains the proof of the Laplace principle upper bound as well as compactness of level sets. Conclusions and directions for future work are in Sect. 11. Appendix 1 provides a list of technical notation used throughout the manuscript for convenience. A number of key technical estimates are presented in the remainder of the appendix. In particular, Appendix 2 contains moments bounds for the controlled system. Appendix 3 presents regularity results for the Poisson equation needed to study the fluctuations. Even though related results exist in the literature, the fully coupled McKean–Vlasov case is not covered by the existing results, and therefore Appendix 3 contains the appropriate discussion of the necessary extensions. Lastly, Appendix 4 contains necessary results on differentiation of functions on spaces of measures.

2 Notation, topologies, and assumptions

In order to construct an appropriate topology for the process Z^N from Eq. (3), we follow the method of [10, 39, 49]. Denote by $\mathcal{S}$ the space of functions $\phi : \mathbb{R} \to \mathbb{R}$ which are infinitely differentiable and satisfy $|x|^m \phi^{(k)}(x) \to 0$ as $|x| \to \infty$ for all $m, k \in \mathbb{N}$. On $\mathcal{S}$, consider the sequence of inner products $(\cdot, \cdot)_n$ and $\|\cdot\|_n$ defined by

$$(\phi, \psi)_n := \sum_{k=0}^{n} \int_{\mathbb{R}} (1 + x^2)^{2n} \phi^{(k)}(x) \psi^{(k)}(x) dx, \qquad \|\phi\|_n := \sqrt{(\phi, \phi)_n} \tag{4}$$

for each $n \in \mathbb{N}$. As per [36] p. 82 (this specific example on p. 84), this sequence of seminorms induces a nuclear Fréchet topology on $\mathcal{S}$. Let $\mathcal{S}_n$ be the completion of $\mathcal{S}$ with respect to $\|\cdot\|_n$ and $\mathcal{S}_{-n} = \mathcal{S}_n'$ the dual space of $\mathcal{S}_n$. We equip $\mathcal{S}_{-n}$ with dual norm $\|\cdot\|_{-n}$ and corresponding inner product $(\cdot, \cdot)_{-n}$. Then $\{\mathcal{S}_n\}_{n \in \mathbb{Z}}$ defines a sequence of nested Hilbert spaces with $\mathcal{S}_m \subset \mathcal{S}_n$ for $m \geq n$. In addition we have for each $n \in \mathbb{N}$, there exists $m > n$ such that the canonical embedding $\mathcal{S}_{-n} \to \mathcal{S}_{-m}$ is Hilbert–Schmidt. In particular, this holds for m sufficiently large that $\sum_{j=1}^{N} \left\|\phi_j^m\right\|_n < \infty$, where $\{\phi_j^m\}_{j\in\mathbb{N}}$ is a complete orthonormal system of $\mathcal{S}_m$. This allows us to use the results of [55] to see that $\{Z^N\}_{N\in\mathbb{N}}$ is tight as a sequence of $C([0,T]; S_{-m})$-valued random variables for sufficiently large m. In particular, we will require $m > 7$ to be sufficiently large so that the canonical embedding

$$\mathcal{S}_{-7} \to \mathcal{S}_{-m} \text{ is Hilbert-Schmidt.} \tag{5}$$

In the proof of the Laplace Principle, we will also make use of $w > 9$ such that

$$\mathcal{S}_{-m-2} \to \mathcal{S}_{-w} \text{ is Hilbert-Schmidt.} \tag{6}$$

When proving compactness of level sets of the rate function, we will in addition make use of $r > 11$ sufficiently large that the canonical embedding

$$\mathcal{S}_{-w-2} \to \mathcal{S}_{-r} \text{ is Hilbert-Schmidt.} \tag{7}$$

It will be useful to consider another system of seminorms on $\mathcal{S}$ given, for each $n \in \mathbb{N}$, by

$$|\phi|_n := \sum_{k=0}^{n} \sup_{x\in\mathbb{R}} |\phi^{(k)}(x)| \tag{8}$$

Via a standard Sobolev embedding argument, one can show that for each $n \in \mathbb{N}$, there exists $C(n)$ such that:

$$|\phi|_n \leq C(n) \, \|\phi\|_{n+1}, \quad \forall \phi \in \mathcal{S}. \tag{9}$$

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be Polish spaces, and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ be a measure space. We will denote by $\mathcal{P}(\boldsymbol{X})$ the space of probability measures on $\boldsymbol{X}$ with the topology of weak convergence, $\mathcal{P}_2(\boldsymbol{X}) \subset \mathcal{P}(\boldsymbol{X})$ the space of square integrable probability measures on $\boldsymbol{X}$ with the 2-Wasserstein metric (see Definition D.1), $\mathcal{B}(\boldsymbol{X})$ the Borel σ-field of $\boldsymbol{X}$, $C(\boldsymbol{X}; \boldsymbol{Y})$ the space of continuous functions from $\boldsymbol{X}$ to $\boldsymbol{Y}$, $C_b(\boldsymbol{X})$ the space of bounded, continuous functions from $\boldsymbol{X}$ to $\mathbb{R}$ with norm $\|\cdot\|_\infty$, and $L^p(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ the space of p-integrable functions on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ (where if $\tilde{\Omega} = \boldsymbol{X}$ and no σ-algebra is provided we assume it is $\mathcal{B}(\boldsymbol{X})$). For $\mu \in \mathcal{P}(\boldsymbol{X})$, $\nu \in \mathcal{P}(\boldsymbol{Y})$, we will denote the product measure induced by μ and ν on $\boldsymbol{X} \times \boldsymbol{Y}$ by $\mu \otimes \nu$. We will at times denote $L^2(\boldsymbol{X} \times \boldsymbol{X}, \mu \otimes \mu)$ by $L^2(\boldsymbol{X}, \mu) \otimes L^2(\boldsymbol{X}, \mu)$. We will denote by $L^1_{\text{loc}}(\boldsymbol{X}, \mu)$ the space of locally integrable functions on $\boldsymbol{X}$. For $U \subseteq \mathbb{R}^d$ open, we will denote by $C_c^\infty(U)$ the space of smooth, compactly supported functions on U. $C_b^k(\mathbb{R})$ for $k \in \mathbb{N}$ will note the space of functions with k continuous and bounded derivatives on $\mathbb{R}$, with norm $|\cdot|_k$ as in Eq. (8), and $C_b^{1,k}([0,T] \times \mathbb{R})$ will denote continuous functions ψ

on $[0, T] \times \mathbb{R}$ with a continuous, bounded time derivative on $(0, T)$, denoted $\dot{\psi}$, such that $\|\psi\|_{C_b^{1,k}([0,T]\times\mathbb{R})} := \sup_{t\in[0,T],x\in\mathbb{R}} |\dot{\psi}(t, x)| + \sup_{t\in[0,T]} \|\psi(t, \cdot)\|_{C_b^k(\mathbb{R})} < \infty$. $C_{b,L}^k(\mathbb{R}) \subset C_b^k(\mathbb{R})$ is the space of functions in $C_b^k(\mathbb{R})$ such that all k derivatives are Lipschitz continuous. For $\phi \in L^1(\boldsymbol{X}, \mu)$, $\mu \in \mathcal{P}(\mathbb{R})$ we define $\langle \mu, \phi\rangle := \int_{\boldsymbol{X}} \phi(x)\mu(dx)$. Similarly, for $Z \in \mathcal{S}_{-p}$, $\phi \in \mathcal{S}_p$, we will denote the action of Z on ϕ by $\langle Z, \phi\rangle$. For $a, b \in \mathbb{R}$, we will denote $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$. C will be used for a constant which may change from line to line throughout, and when there are parameters $a_1, \ldots, a_n$ which C depends on in an important manner, will denote this dependence by $C(a_1, \ldots, a_n)$. For all function spaces, the codomain is assumed to be $\mathbb{R}$ unless otherwise denoted.

In the construction of the controlled system, we will also make use of the space of measures on $\mathbb{R}^d \times [0, T]$ such that $Q(\mathbb{R}^d \times [0, t]) = t, \forall t \in [0, T]$. We will denote this space $M_T(\mathbb{R}^d)$. We equip $M_T(\mathbb{R}^d)$ with the topology of weak convergence of measures (thus making $M_T(\mathbb{R}^d)$ a Polish space by [22] Theorem A.3.3). See also the proof of Lemma 3.3.1 in [22] for the fact that $M_T(\mathbb{R}^d)$ is a closed subset of finite positive Borel measures on $\mathbb{R}^d \times [0, T]$). For when dealing with the occupation measures as defined in Eq. (56), we will in particular take $d = 4$ and will interpret $Q(dx, dy, dz, dt)$ as x denoting variable representing the first coordinate in $\mathbb{R}^4$, y the second, and z the third and fourth.

For a mapping $\vartheta : [0, T] \to \mathcal{P}(\mathbb{R}^d)$, it will be useful to define an element of $M_T(\mathbb{R}^d)$ induced by ϑ by

$$\nu_\vartheta(A \times [0, t]) := \int_0^t \vartheta(s)[A]ds, \forall t \in [0, T], A \in \mathcal{B}(\mathbb{R}^d). \tag{10}$$

Due to the nature of the space we consider the sequence $\{Z^N\}_{N\in\mathbb{N}}$ to live on, it is natural that we will have to restrict the growth of the coefficients which appear in Eqs. (1) and (25) in x. We will also need to ensure that the derivatives of the Poisson equation which appear in the definition the limiting coefficients in Eq. (23) exist and that the homogenized drift and diffusion coefficients in Eq. (24), which determine the limiting McKean–Vlasov Equation X_t from Eq. (25), are well-defined. In doing so, will be controlling many mixed derivatives of functions in the Lions sense [11] and in the standard sense, it will be useful for us to borrow the multi-index notation proposed in [15] and employed in [14]. For the reader's convenience, we have included in Appendix 4 a brief review on differentiation of functions on spaces of measures. For a more comprehensive exposition on this, we refer the interested reader to [13] Chapter 5.

Furthermore, since we prove the moderate deviations principle via use of the controlled particle system (55), we will only have up to second moments of the controlled fast system (see Appendix 2). It will be important to make sure that terms which the controlled fast process enters in the intermediate proofs of tightness, so naturally we will need some assumptions on the rate of polynomial growth in y of the coefficients which appear in Eqs. (1) and (25) (See Remark 2.7). We thus extend the multi-index notation from the aforementioned papers to track specific collections of mixed partial derivatives, and to give us a clean way of tracking the rate of polynomial growth in y for those mixed partials in the coming definitions.

Definition 2.1 Let n, l be non-negative integers and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$ be an $n-$dimensional vector of non-negative integers. We call any ordered tuple of the form $(n, \ell, \boldsymbol{\beta})$ a **multi-index**. For a function $G : \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$, we will denote for a multi-index $(n, \ell, \boldsymbol{\beta})$, if this derivative is well defined,

$$D^{(n,l,\boldsymbol{\beta})} G(x, \mu)[z_1, \ldots, z_n] = \partial_{z_1}^{\beta_1} \ldots \partial_{z_n}^{\beta_n} \partial_x^l \partial_\mu^n G(x, \mu)[z_1, \ldots, z_n],$$

As noted in the Remark D.3, for such a derivative to be well defined we require for it to be jointly continuous in $x, \mu, z_1, \ldots, z_n$ where the topology used in the measure component is that of $\mathcal{P}_2(\mathbb{R})$.

We also define $\boldsymbol{\delta}^{(n,l,\boldsymbol{\beta})} G(x, \mu)[z_1, \ldots, z_n]$ in the exact same way, with the Lions derivatives ∂_μ replaced by linear functional derivatives $\frac{\delta}{\delta m}$; see Appendix 4 for differentiation of functions on spaces of measures.

Definition 2.2 For $\boldsymbol{\zeta}$ a collection of multi-indices of the form $(n, l, \boldsymbol{\beta}) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}^n$, we will call $\boldsymbol{\zeta}$ a **complete** collection of multi-indices if for any $(n, l, \boldsymbol{\beta}) \in \boldsymbol{\zeta}$, $\{(k, j, \boldsymbol{\alpha}(k)) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}^k : j \le l, k \le n, \boldsymbol{\alpha}(k) = (\alpha_1, \ldots, \alpha_k), \exists \boldsymbol{\beta}(k) = (\beta(k)_1, \ldots, \beta(k)_k) \in \binom{\boldsymbol{\beta}}{k}$ such that $\alpha_p \le \beta(k)_p, \forall p = 1, \ldots, k\} \subset \boldsymbol{\zeta}$. Here for a vector of positive integers $\beta = (\beta_1, \ldots, \beta_n)$ and $k \in \mathbb{N}, k \le n$, we are using the notation $\binom{\boldsymbol{\beta}}{k}$ to represent the set of size $\binom{n}{k}$ containing all the k-dimensional vectors of positive integers which can be obtained from removing $n - k$ entries from $\boldsymbol{\beta}$.

Remark 2.3 Definition 2.2 is enforcing that if collection of multi-indices contains a multi-index representing some mixed derivative in (x, μ, z) as per Definition 2.1, then it also contains all lower-order mixed derivatives of the same type. For example, if $\boldsymbol{\zeta}$ is a collection of multi-indices containing $(2, 0, (1, 1))$ (corresponding to $\partial_{z_1} \partial_{z_2} \partial_\mu^2 G(x, \mu)[z_1, z_2]$) then, in order to be complete, it must also contain the terms $(2, 0, (1, 0))$, $(2, 0, (0, 1))$, $(2, 0, 0)$, $(1, 0, 1)$, $(1, 0, 0)$, and $(0, 0, 0)$ (corresponding to the terms $\partial_{z_1} \partial_\mu^2 G(x, \mu)[z_1, z_2]$, $\partial_{z_2} \partial_\mu^2 G(x, \mu)[z_1, z_2]$, $\partial_\mu^2 G(x, \mu)[z_1, z_2]$, $\partial_z \partial_\mu G(x, \mu)[z]$, $\partial_\mu G(x, \mu)[z]$, and $G(x, \mu)$ respectively). This is a technical requirement used in order to state the results in Appendix 1 in a way that allows the inductive arguments used therein to go through.

Using this multi-index notation, it will be useful to define some spaces regarding regularity of functions in regards to these mixed derivatives. We thus make the following modifications to Definition 2.13 in [14]:

Definition 2.4 For $\boldsymbol{\zeta}$ a collection of multi-indices of the form $(n, l, \boldsymbol{\beta}) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}^n$, we define $\mathcal{M}_b^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ to be the class of functions $G : \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ such that $D^{(n,l,\boldsymbol{\beta})} G(x, \mu)[z_1, \ldots, z_n]$ exists and satisfies

$$\|G\|_{\mathcal{M}_b^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))} := \sup_{(n,l,\boldsymbol{\beta}) \in \boldsymbol{\zeta}} \sup_{x, z_1, \ldots, z_n \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R})} |D^{(n,l,\boldsymbol{\beta})} G(x, \mu)[z_1, \ldots, z_n]| \le C. \quad (11)$$

We denote the class of functions $G \in \mathcal{M}_b^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ such that:

$$|D^{(n,l,\boldsymbol{\beta})} G(x,\mu)[z_1,\dots,z_n] - D^{(n,l,\boldsymbol{\beta})} G(x',\mu')[z_1',\dots,z_n']| \\ \leq C_L\left(|x-x'| + \sum_{i=1}^N |z_i - z_i'| + \mathbb{W}_2(\mu,\mu')\right) \tag{12}$$

for all $(n,l,\boldsymbol{\beta}) \in \boldsymbol{\zeta}$ and $x, x', z_1, \dots, z_n, z_1', \dots, z_n' \in \mathbb{R}, \mu, \mu' \in \mathcal{P}_2(\mathbb{R})$ by $\mathcal{M}_{b,L}^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$. We define $\mathcal{M}_b^{\boldsymbol{\zeta}}(\mathcal{P}_2(\mathbb{R}))$ and $\mathcal{M}_{b,L}^{\boldsymbol{\zeta}}(\mathcal{P}_2(\mathbb{R}))$ analogously, where instead here $\boldsymbol{\zeta}$ is a collection of multi-indices of the form $(n,\boldsymbol{\beta}) \in \mathbb{N} \times \mathbb{N}^n$, and we take the $l = 0$ in the above multi-index notation for the derivatives.

We will also make use of the class of functions $\mathcal{M}_p^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ which contains $G : \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ such that $G(\cdot, y, \cdot) \in \mathcal{M}_b^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ for all $y \in \mathbb{R}$, with all derivatives appearing in the definition of $\mathcal{M}_b^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ jointly continuous in $(x, y, \mathbb{W}_2)$, and for each multi-index $(n,l,\boldsymbol{\beta}) \in \boldsymbol{\zeta}$,

$$\sup_{x,z_1,\dots,z_n \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R})} |D^{(n,l,\boldsymbol{\beta})} G(x,y,\mu)[z_1,\dots,z_n]| \leq C(1+|y|)^{q_G(n,l,\boldsymbol{\beta})}, \tag{13}$$

where $q_G(n,l,\boldsymbol{\beta}) \in \mathbb{R}$. Similarly, $\mathcal{M}_{p,L}^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ is defined as $G \in \mathcal{M}_p^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ such that Eq. (12) holds for $G(\cdot, y, \cdot)$ for each $y \in \mathbb{R}$, where $C_L(y)$ grows at most polynomially in y.

We also define $\mathcal{M}_b^{\boldsymbol{\zeta}}([0,T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ to be the class of functions $G : [0,T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ such that $G(\cdot, x, \mu)$ is continuously differentiable on $(0,T)$ for all $x \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R})$ with time derivative denoted by $\dot{G}(t,x,\mu)$, $G(t,\cdot,\cdot) \in \mathcal{M}_b^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ for all $t \in [0,T]$, with (11) holding uniformly in t, and G, $\dot{G}$, and all derivatives involved in the definition of $\mathcal{M}_b^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ are jointly continuous in time, measure, and space. We define for $G \in \mathcal{M}_b^{\boldsymbol{\zeta}}([0,T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$

$$\|G\|_{\mathcal{M}_b^{\boldsymbol{\zeta}}([0,T]\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))} \\ := \sup_{t\in[0,T]} \|G(t,\cdot)\|_{\mathcal{M}_b^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))} + \sup_{t\in[0,T], x\in\mathbb{R}, \mu\in\mathcal{P}_2(\mathbb{R})} |\dot{G}(t,x,\mu)|.$$

We denote the class of functions $G \in \mathcal{M}_b^{\boldsymbol{\zeta}}([0,T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ such that (12) holds uniformly in t by $\mathcal{M}_{b,L}^{\boldsymbol{\zeta}}([0,T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$. Again, we define $\mathcal{M}_b^{\boldsymbol{\zeta}}([0,T] \times \mathcal{P}_2(\mathbb{R}))$, $\mathcal{M}_{b,L}^{\boldsymbol{\zeta}}([0,T] \times \mathcal{P}_2(\mathbb{R}))$, and $\mathcal{M}_p^{\boldsymbol{\zeta}}([0,T] \times \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ analogously.

At times we will want to consider Lions Derivatives bounded in $L^2(\mathbb{R},\mu)$ rather than uniformly in z. Thus we define $\tilde{\mathcal{M}}_b^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ to be the class of functions $G : \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ such that $D^{(n,l,\boldsymbol{\beta})} G(x,\mu)[z_1,\dots,z_n]$ exists and satisfies

$$\|G\|_{\tilde{\mathcal{M}}_b^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))} := \sup_{(n,l,\boldsymbol{\beta})\in\boldsymbol{\zeta}} \sup_{x\in\mathbb{R}, \mu\in\mathcal{P}_2(\mathbb{R})} \left\| D^{(n,l,\boldsymbol{\beta})} G(x,\mu)[\cdot] \right\|_{L^2(\mu,\mathbb{R})^{\otimes n}}$$

$$= \sup_{(n,l,\boldsymbol{\beta})\in\boldsymbol{\zeta}} \sup_{x\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} \left(\int_{\mathbb{R}} ... \int_{\mathbb{R}} |D^{(n,l,\boldsymbol{\beta})}G(x,\mu)[z_1,\dots,z_n]|^2 \mu(dz_1)...\mu(dz_n)\right)^{1/2}$$
$$\leq C. \tag{14}$$

We also define $\tilde{\mathcal{M}}_b^{\boldsymbol{\zeta}}([0,T]\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ analogously to $\mathcal{M}_b^{\boldsymbol{\zeta}}([0,T]\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$, $\tilde{\mathcal{M}}_p^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ analogously to $\mathcal{M}_p^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$, and $\tilde{\mathcal{M}}_p^{\boldsymbol{\zeta}}([0,T]\times\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ analogously to $\mathcal{M}_p^{\boldsymbol{\zeta}}([0,T]\times\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$. We will denote the polynomial growth rate for $G\in\tilde{\mathcal{M}}_p^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ and $(n,l,\boldsymbol{\beta})\in\boldsymbol{\zeta}$ as in Eq. (13) but with the $L^2(\mathbb{R},\mu)^{\otimes n}$-norm by $\tilde{q}(n,l,\boldsymbol{\beta})\in\mathbb{R}$ to avoid confusion with polynomial growth of the derivatives in the uniform norm. That is:

$$\sup_{x\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} \left(\int_{\mathbb{R}} ... \int_{\mathbb{R}} |D^{(n,l,\boldsymbol{\beta})}G(x,\mu)[z_1,\dots,z_n]|^2 \mu(dz_1)...\mu(dz_n)\right)^{1/2}$$
$$\leq C(1+|y|)^{\tilde{q}_G(n,l,\boldsymbol{\beta})}. \tag{15}$$

Lastly, we define $\mathcal{M}_{\boldsymbol{\delta},b}^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ and $\mathcal{M}_{\boldsymbol{\delta},p}^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ in the same way as $\mathcal{M}_b^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ and $\mathcal{M}_p^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ respectively, but with with $D^{(n,l,\boldsymbol{\beta})}$ replaced by $\boldsymbol{\delta}^{(n,l,\boldsymbol{\beta})}$. We also extend this in the natural way when the spatial components are in higher dimensions (i.e. taking gradients and using norms in $\mathbb{R}^d$).

Let us now introduce the main assumptions that are needed for the work of this paper to go through.

(A1) $0<\lambda_-\leq\tau_1^2(x,y,\mu)+\tau_2^2(x,y,\mu)\leq\lambda_+<\infty$, $\forall x,y\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})$, and τ_1,τ_2 have two uniformly bounded derivatives in y and which are jointly continuous in $(x,y,\mathbb{W}_2)$.

(A2) There exists $\beta>0$ and $\kappa>0$ such that:

$$f(x,y,\mu)=-\kappa y+\eta(x,y,\mu) \tag{16}$$

where η is uniformly bounded in x and μ, and Lipschitz in the sense of (A9) in x, μ, and y with

$$|\eta(x,y_1,\mu)-\eta(x,y_2,\mu)|\leq L_\eta|y_1-y_2|, \forall x\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})$$

for L_η such that $L_\eta-\kappa<0$, and

$$2(f(x,y_1,\mu)-f(x,y_2,\mu))(y_1-y_2)+3|\tau_1(x,y_1,\mu)-\tau_1(x,y_2,\mu)|^2$$
$$+3|\tau_2(x,y_1,\mu)-\tau_2(x,y_2,\mu)|^2\leq-\beta|y_1-y_2|^2,$$
$$\forall x,y_1,y_2\in\mathbb{R},\mu\in\mathcal{P}(\mathbb{R}). \tag{17}$$

Let $a(x,y,\mu)=\frac{1}{2}[\tau_1^2(x,y,\mu)+\tau_2^2(x,y,\mu)]$. For $x\in\mathbb{R}$, $\mu\in\mathcal{P}_2(\mathbb{R})$, we define the differential operator $L_{x,\mu}$ acting on $\phi\in C_b^2(\mathbb{R})$ by

$$L_{x,\mu}\phi(y)=f(x,y,\mu)\phi'(y)+a(x,y,\mu)\phi''(y). \tag{18}$$

Note that under assumptions (A1) and (A2), there is a constant C independent of x, y, μ such that:

$$2f(x, y, \mu)y + 3|\tau_1(x, y, \mu)|^2 + 3|\tau_2(x, y, \mu)|^2 \\ \leq -\frac{\beta}{2}|y|^2 + C, \forall x, y \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R}). \tag{19}$$

Thus by [60] Proposition 1 (see also [70]), there exists a $\pi(\cdot; x, \mu)$ which is the unique measure solving

$$L^*_{x,\mu}\pi = 0. \tag{20}$$

Moreover, for all $k > 0$, there is $C_k \geq 0$ such that $\sup_{x\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} \int_{\mathbb{R}} |y|^k \pi(dy; x, \mu) \leq C_k$.

(A3) For π as in Eq. (20),

$$\int_{\mathbb{R}} b(x, y, \mu)\pi(dy; x, \mu) = 0, \forall x \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R}), \tag{21}$$

b is jointly continuous in $(x, y, \mathbb{W}_2)$, grows at most polynomially in y uniformly in $x \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R})$.

Having introduced the notation above, we can now present the law of large numbers for the empirical measure $\mu^{\epsilon,N}$ from Eq. 2 in the joint limit as $\epsilon \downarrow 0, N \to \infty$. Under assumptions (A1)–(A3), by Lemma C.1 we consider Φ the unique classical solution to:

$$L_{x,\mu}\Phi(x, y, \mu) = -b(x, y, \mu), \qquad \int_{\mathbb{R}} \Phi(x, y, \mu)\pi(dy; x, \mu) = 0. \tag{22}$$

Let us define the functions

$$\begin{aligned}
\gamma(x, y, \mu) &:= \gamma_1(x, y, \mu) + c(x, y, \mu)\\
\gamma_1(x, y, \mu) &:= b(x, y, \mu)\Phi_x(x, y, \mu) + g(x, y, \mu)\Phi_y(x, y, \mu)\\
&\quad + \sigma(x, y, \mu)\tau_1(x, y, \mu)\Phi_{xy}(x, y, \mu)\\
D(x, y, \mu) &:= D_1(x, y, \mu) + \frac{1}{2}\sigma^2(x, y, \mu)\\
D_1(x, y, \mu) &= b(x, y, \mu)\Phi(x, y, \mu) + \sigma(x, y, \mu)\tau_1(x, y, \mu)\Phi_y(x, y, \mu).
\end{aligned} \tag{23}$$

and

$$\bar{\gamma}(x, \mu) := \left[\int_{\mathbb{R}} \gamma(x, y, \mu)\pi(dy; x, \mu)\right], \quad \bar{D}(x, \mu) := \left[\int_{\mathbb{R}} D(x, y, \mu)\pi(dy; x, \mu)\right]. \tag{24}$$

Then, by essentially the same arguments as in [4], under the conditions outlined below, $\mu^{\epsilon,N}$ converges in distribution to the deterministic limit $\mathcal{L}(X)$ where X satisfies the averaged McKean–Vlasov SDE

$$dX_t = \bar{\gamma}(X_t, \mathcal{L}(X_t))dt + \sqrt{2\bar{D}(X_t, \mathcal{L}(X_t))}dW_t^2 \quad X_0 = \eta^x. \tag{25}$$

Here W_t^2 is a Brownian motion on another filtered probability space satisfying the usual conditions. In fact, we see here in Lemma 8.2 that in fact this convergence occurs in $\mathcal{P}_2(\mathbb{R})$ for each $t \in [0, T]$.

Remark 2.5 Using an integration-by-parts argument, one can find that the diffusion coefficient $\bar{D}$ can be written in the alternative form

$$\bar{D}(x,\mu) = \frac{1}{2}\int_{\mathbb{R}} \left([\tau_2(x,y,\mu)\Phi_y(x,y,\mu)]^2 + [\sigma(x,y,\mu) + \tau_1(x,y,\mu)\Phi_y(x,y,\mu)]^2\right) \pi(dy;x,\mu), \tag{26}$$

and hence is non-negative. See [3] Chapter 3 Section 6.2 for a similar computation.

We now introduce the remaining assumptions. Since we are dealing with fluctuations, we will need to be able to obtain rates of averaging, and thus there are several auxiliary Poisson equations involved in the proof of tightness. When there is more specific structure to the system of Eq. (1), these assumptions may be able to be verified on a case-by-case basis. In Sect. 3 we provide concrete examples for which all of the conditions imposed in the paper hold. Remark 2.6 and mainly Remark 2.7 discuss the meaning of these assumptions more thoroughly. In doing so, it will be useful to define the following complete collections of multi-indices in the sense of Definitions 2.1 and 2.2:

$$\begin{aligned}
\hat{\zeta} &\ni \{(0,j_1,0),(1,j_2,j_3),(2,j_4,(j_5,j_6)),(3,0,(j_7,0,0)) \\
&\qquad : j_1 \in \{0,1,\dots 4\}, j_2+j_3 \le 4, j_4+j_5+j_6 \le 2, j_7 = 0,1\} \\
\tilde{\zeta} &\ni \{(0,j_1,0),(1,j_2,j_3),(2,0,0) : j_1 = 0,1,2, j_2+j_3 \le 1\} \\
\tilde{\zeta}_1 &\ni \{(j_1,j_2,0) : j_1+j_2 \le 1\} \\
\tilde{\zeta}_2 &\ni \{(j,0,0) : j = 0,1\} \\
\tilde{\zeta}_3 &\ni \{(0,j_1,0),(1,0,j_2) : j_1 = 0,1,2, j_2 = 0,1\} \\
\zeta_{x,l} &\ni \{(0,j,0) : j = 0,1,\dots,l\}, l \in \mathbb{N} \\
\bar{\zeta} &\ni \{(j,0,0) : j = 0,1,2\} \\
\bar{\zeta}_l &\ni \{(0,0,0),(1,0,j) : j = 0,1,\dots,l\}, l \in \mathbb{N}.
\end{aligned} \tag{27}$$

In the following set of assumptions, recall that for $G : \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ and a multi-index (n,l,β), $\tilde{q}_G(n,l,\beta)$ denotes the rate of polynomial growth in y of the mixed derivative of G corresponding to (n,l,β) as per Eq. (15) in Definition 2.4. Recall also the spaces of functions of measures from Definition 2.4.

(A4) Strong existence and uniqueness holds for the system of SDEs (1) for all $N \in \mathbb{N}$, for the Slow–Fast McKean–Vlasov SDEs (57), and for the limiting McKean–Vlasov SDE (25).

(A5) g and σ are uniformly bounded, and c, b grow at most linearly in y uniformly in $x \in \mathbb{R}$, $\mu \in \mathcal{P}_2(\mathbb{R})$. All coefficients are jointly continuous in $(x, y, \mathbb{W}_2)$.

(A6) There exists a unique strong solution $\Phi \in \tilde{\mathcal{M}}_p^{\tilde{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ to Eq. (22) with $\tilde{q}_\Phi(n, l, \boldsymbol{\beta}) \le 1, \forall (n, l, \boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}$, and $\Phi_y \in \tilde{\mathcal{M}}_p^{\tilde{\zeta}_2}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$, with $\tilde{q}_{\Phi_y}(n, l, \boldsymbol{\beta}) \le 1, \forall (n, l, \boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}_2$. In addition, this can be strengthened to $\tilde{q}_\Phi(0, k, 0) \le 0, k = 0, 1$ and $\tilde{q}_{\Phi_y}(0, 0, 0) \le 0$. (For Proposition 6.1 and Theorem 7.2).

(A7) There exists a unique strong solution $\chi \in \tilde{\mathcal{M}}_p^{\tilde{\zeta}}(\mathbb{R}^2 \times \mathbb{R}^2 \times \mathcal{P}_2(\mathbb{R}))$ to Eq. (63) with $\tilde{q}_\chi(n, l, \boldsymbol{\beta}) \le 1, \forall (n, l, \boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}$, and $\chi_y \in \tilde{\mathcal{M}}_p^{\tilde{\zeta}_1}(\mathbb{R}^2 \times \mathbb{R}^2 \times \mathcal{P}_2(\mathbb{R}))$, $\chi_{yy} \in \tilde{\mathcal{M}}_p^{(0,0,0)}(\mathbb{R}^2 \times \mathbb{R}^2 \times \mathcal{P}_2(\mathbb{R}))$ with $\tilde{q}_{\chi_y}(n, l, \boldsymbol{\beta}) \le 1, \forall (n, l, \boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}_1$, $\tilde{q}_{\chi_{yy}}(0, 0, 0) \le 1$. In addition, this can be strengthened to $\tilde{q}_\chi(0, k, 0) \le 0, k = 0, 1$ and $\tilde{q}_{\chi_y}(0, 0, 0) \le 0$. (For Proposition 6.3 and Theorem 7.2).

(A8) For $F = \gamma, D$, or $\sigma\psi_1 + [\tau_1\psi_1 + \tau_2\psi_2]\Phi_y$ for any $\psi_1, \psi_2 \in C_c^\infty([0, T] \times \mathbb{R} \times \mathbb{R})$, there exists a unique strong solution $\Xi \in \tilde{\mathcal{M}}_p^{\tilde{\zeta}}([0, T] \times \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ to Eq. (64) with each of these choices of F, $\tilde{q}_\Xi(n, l, \boldsymbol{\beta}) \le 2, \forall (n, l, \boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}$, and $\Xi_y \in \tilde{\mathcal{M}}_p^{\tilde{\zeta}_1}([0, T] \times \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ with $\tilde{q}_{\Xi_y}(n, l, \boldsymbol{\beta}) \le 2, \forall (n, l, \boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}_1$. Moreover, we assume for all choices of F, this can be strengthened to $\tilde{q}_\Xi(n, l, \boldsymbol{\beta}) \le 1, \forall (n, l, \boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}_1$ and $\tilde{q}_{\Xi_y}(0, 0, 0) \le 1$. (For Propositions 6.4/10.1 and Theorem 7.2).

(A9) For $F = \gamma, \sigma + \tau_1\Phi_y, \tau_2\Phi_y, \tau_1, \tau_2$:

$$
\begin{aligned}
&|F(x_1, y_1, \mu_1) - F(x_2, y_2, \mu_2)| \le C(|x_1 - x_2| + |y_1 - y_2| \\
&+ \mathbb{W}_2(\mu_1, \mu_2)), \forall x_1, x_2, y \in \mathbb{R}, \mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}).
\end{aligned}
$$

(Lemmas 7.4 and 7.5).

(A10) $\bar{\gamma}, \bar{D}^{1/2} \in \mathcal{M}_{b,L}^{\hat{\zeta}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$. (For Theorem 7.2).

(A11) Consider the Poisson equation

$$
\begin{aligned}
&L^2_{x,\bar{x},\mu}\tilde{\chi}(x, \bar{x}, y, \bar{y}, \mu) = -b(x, y, \mu)\Phi(\bar{x}, \bar{y}, \mu), \\
&\int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{\chi}(x, \bar{x}, y, \bar{y}, \mu)\pi(dy; x, \mu)\pi(d\bar{y}, \bar{x}, \mu) = 0. \qquad (28)
\end{aligned}
$$

where $L^2_{x,\bar{x},\mu}$ is as in Eq. (60). Assume there exists a unique strong solution $\tilde{\chi} \in \tilde{\mathcal{M}}_p^{\tilde{\zeta}_3}(\mathbb{R}^2 \times \mathbb{R}^2 \times \mathcal{P}_2(\mathbb{R}))$ and $\tilde{\chi_y} \in \tilde{\mathcal{M}}_p^{\zeta_{x,1}}(\mathbb{R}^2 \times \mathbb{R}^2 \times \mathcal{P}_2(\mathbb{R}))$ to Eq. (28). (For Theorem 7.2).

(A12) $\tau_1, \tau_2, f, \gamma, \sigma + \tau_1\Phi_y, \tau_2\Phi_y \in \mathcal{M}_{\delta,p}^{\bar{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$. (For Lemmas 7.4 and 7.5).

(A13) For w as in Eq. (6) and $\bar{\gamma}$,$\bar{D}$ as in Eq. (24), $\bar{\gamma}$, $\bar{D} \in \mathcal{M}_b^{\boldsymbol{\zeta}_{x,w+2}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R})) \cap \mathcal{M}_{\boldsymbol{\delta},b}^{\bar{\boldsymbol{\zeta}}_{w+2}}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$, and

$$\sup_{x\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} \left\| \frac{\delta}{\delta m}\bar{\gamma}(x,\mu)[\cdot]\right\|_{w+2} + \sup_{x\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} \left\| \frac{\delta}{\delta m}\bar{D}(x,\mu)[\cdot]\right\|_{w+2} < \infty.$$

(For Lemmas 7.7, 8.6, 8.7 and Proposition 8.3).

Remark 2.6 There is a current gap in the literature regarding rates of polynomial growth for derivatives of solutions to Poisson equations of the form (22), as outlined in [33] Remark A.1. Though in Proposition A.2 they state a result partially amending this issue, the bounds provided likely are not tight. In particular, under the assumption (A2) which we require for moment bounds of the fast process (and hence slow) process in Sect. 2, their result cannot provide boundedness of derivatives in y of Φ from (22), or any of the other auxiliary Poisson equations which we consider. This in turn also makes it difficult to gain good rates of polynomial growth for derivatives in the parameters x and μ. We need strict control of these rates of growth, for the reasons outlined in Remark 2.7. Stronger bounds are derived in the 1-D case in Proposition A.4 of [33], so this makes gaining the necessary control much easier in the current setting (see the results contained in Sect. 1 in the Appendix). Note also the much stricter assumptions imposed when handling the multi-dimensional cell problem in Lemma C.4 (which is required to establish sufficient conditions for (A7)).

Remark 2.7 Assumptions (A1) and (A2) are used in tandem for the existence and uniqueness of the invariant measure π from Eq. (20). Such an invariant measure exists under weaker recurrence conditions on f (see, e.g. [60] Proposition 1), but we use the near-Ornstein-Uhlenbeck structure assumed in (16) and the form of the retraction to the mean (17) in order to prove certain moment bounds on the controlled fast process in the Appendix 2, and (17) is also used in order to gain sufficient conditions for the required regularity of the Poisson Equations in Assumptions (A6)–(A13) in Appendix 3. In particular, (16) is inspired by Assumption 4.1 (iii) in [44] and is needed for Lemma B.2, and (17) is a standard assumption for control of moments of SDEs over infinite time horizons and for controlling solutions of related Cauchy problems (see e.g. [65] Assumption A.1 Eq. (2.3)).

The centering condition (A3) is standard in the theory of stochastic homogenization. Assumption (A4) is required in order to apply the weak-convergence approach to large deviations. In particular, it ensures that the prelimit control representation (51) holds. This is known to hold, for example, under global Lipschitz assumptions on all the coefficients (see, e.g. [73] Theorem 2.1 and Section 6.1 in [65]), though can also be proved under much weaker assumptions. These two assumptions, along with existence and uniqueness of the invariant measure π from Eq. (20) and the Poisson Equation Φ from Eq. (22), can be seen as the crucial hypothesis of this paper. The rest of the assumptions are technical and essentially used to have sufficient conditions for tightness of the controlled fluctuations processes $\tilde{Z}^N$ from Eq. (54) (and, in the case of Assumption (A13), to have uniqueness of solutions to its limit (32)).

The boundedness and linear growth of the coefficients from Assumption (A5) are used to restrict the growth of the coefficients so that second moments of the controlled fast process $\tilde{Y}^{i,\epsilon,N}$ from Eq. (55) can be proved in Appendix 2, and to ensure that only knowing these second moment bounds are sufficient for boundedness of the remainder terms in e.g. the ergodic-type theorems of Sect. 6. The joint continuity assumption is used to ensure that integrating the coefficients is a continuous function on the space of measures.

The Assumptions (A6)–(A13) are listed in terms of the Poisson Equations and averaged coefficients (and hence implicitly in terms of Φ from Assumption (A6)) because these assumptions can be verified on a case-by-case basis when the differential operator (18) or the inhomogeneities considered have some special structure. See the Examples provided in Appendix 3.

The growth required by the specific derivatives listed in Assumptions (A6)– (A8) are imposed in order to ensure that the remainder terms resulting form Itô's formula in the Ergodic-Type Theorems in Sect. 6 are bounded. In particular, in Sect. 6, we are dealing with the controlled slow–fast system (55), which due to the controls a priori being at best L^2 integrable (see the bound 53), we are only able to show that we have 2 bounded moments of the fast component (see Appendix 2). This is limiting, since the terms which show up in the Ergodic-Type Theorems are products of derivatives of the Poisson equation with the coefficients of the system (1), of which c and b may grow linearly as per assumption (A5), and with the L^2 controls.

Using Assumption (A6) as an example and unpacking the multi-index notation, we are requiring Φ, Φ_x, Φ_y are bounded, and Φ_{xx}, $\partial_\mu \Phi$, $\partial_\mu \Phi_x$, $\partial_\mu \Phi_y$, $\partial_z \partial_\mu \Phi$, $\partial^2_\mu \Phi$ grow at most linearly in y in their appropriate norms. Looking at the proof of Proposition 6.1, since we are taking the L^2 norm of the remainder terms $\tilde{B}_1 - \tilde{B}_8$, we are essentially ensuring all the products showing up in these terms are L^2 bounded. In particular, in $\tilde{B}_7$, the controls are multiplied by Φ and Φ_x, which is why we end up needing those derivatives to be bounded. Φ_y being bounded is needed elsewhere for essentially the same reason - see, e.g. the proof of Proposition 10.2, where we use that B^N_t is bounded in L^2. The reasoning behind the Assumptions (A7) and (A8) are the exact same, with additional regularity of χ_y and Ξ_y (replacing $\tilde{\zeta}_2$ by $\tilde{\zeta}_1$ means we are requiring χ_y and Ξ_y have an x derivative which grows at most linearly in addition to a μ derivative) and χ_{yy} required due to those additional terms showing up in $\bar{B}_2$ in Proposition 6.3, C_2 in Proposition 6.4, and $\bar{B}_{13}$ in Proposition 6.3 respectively.

The Lipschitz continuity imposed in Assumption (A9) and the existence of two linear functional derivatives which grow at most polynomially in y uniformly in x, μ imposed in Assumption (A12) are used to couple the controlled particles (55) to the auxiliary IID particles (57) in Sect. 7.2. In particular, the terms required to be Lipschitz are those which show up in the drift and diffusion of the processes which result from applying Proposition 6.1 to the controlled system and IID system respectively. The use of a Lipschitz property in such a coupling argument is standard - see, e.g. Lemma 1 in [39]. Since we don't assume that the coefficients have linear interaction with the measure, Assumption (A12) is being used to apply Lemma D.7 to the listed functions. The result of that Lemma is essentially the Assumption (S3) made in [49], which we are using in essentially the same manner that they are in their coupling argument in Theorem 2.4.

Assumption (A10) is tailored to ensure enough regularity of the coefficients of the Cauchy Problem on Wasserstein Space for Theorem 7.2 to hold- see [5] (in particular Lemma 5.1 therein). Assumption (A11) is used to apply the same result, and requires the introduction of the additional auxiliary Poisson equation (A11) which is defined similarly to χ from Assumption (A7) but with a different inhomogeneity due to an additional term which arises in [5] Proposition 4.4 due to the McKean–Vlasov dynamics. The use of this specific control over the derivatives of $\tilde{\chi}$ is discussed after the statement of Theorem 7.2.

Finally, Assumption (A13) is needed for well-definedness/uniqueness of the limiting Eq. (32). See the analogous Assumptions 2.2/2.3 in [10].

3 Main results

We are now ready to state our main result, which takes the form of Theorem 3.2 below. These results will be applied to a concrete class of examples of interacting particle systems of the form (1) in Sect. 4.2.

We prove the large deviations principle for fluctuations process $\{Z^N\}$ from Eq. (3) via means of the Laplace Principle. In other words, in Theorem 3.2, we identify the rate function $I : C([0,T];\mathcal{S}_{-r}) \to [0,+\infty]$ such that for w as in Eq. (6):

$$\lim_{N\to\infty} -a^2(N)\log\mathbb{E}\exp\left(-\frac{1}{a^2(N)}F(Z^N)\right) = \inf_{Z\in C([0,T];\mathcal{S}_{-w})}\{I(Z)+F(Z)\} \tag{29}$$

for all $F \in C_b(C([0,T];\mathcal{S}_{-\tau}))$, for any $\tau \geq w$. In particular, this holds for all $F \in C_b(C([0,T];\mathcal{S}_{-r}))$ for $r > w+2$ as in Eq. (7), and for such F the right hand side is equal to $\inf_{Z\in C([0,T];\mathcal{S}_{-r})}\{I(Z)+F(Z)\}$ by construction of I (see Theorem 3.2). The equality (29) along with compactness of level sets of I implies that $\{Z^N\}$ satisfies the large deviations principle with speed $a^{-2}(N)$ and rate function I via e.g. Theorem 1.2.3 in [22].

In order to show (29), we will show in Sect. 9 that the Laplace principle Lower Bound:

$$\begin{aligned}\liminf_{N\to\infty} -a^2(N)\log\mathbb{E}\exp\left(-\frac{1}{a^2(N)}F(Z^N)\right)\\ \geq \inf_{Z\in C([0,T];\mathcal{S}_{-w})}\{I(Z)+F(Z)\}, \forall F \in C_b(C([0,T];\mathcal{S}_{-\tau}))\end{aligned} \tag{30}$$

for any $\tau \geq w$, with w as in Eq. (6), holds.

Then, in Sect. 10 we will prove the Laplace principle Upper Bound:

$$\begin{aligned}\limsup_{N\to\infty} -a^2(N)\log\mathbb{E}\exp\left(-\frac{1}{a^2(N)}F(Z^N)\right)\\ \leq \inf_{Z\in C([0,T];\mathcal{S}_{-w})}\{I(Z)+F(Z)\}, \forall F \in C_b(C([0,T];\mathcal{S}_{-\tau}))\end{aligned} \tag{31}$$

for any $\tau \geq w$ holds and compactness of level sets of I in $C([0,T];\mathcal{S}_{-r})$, at which point the moderate deviations principle of Theorem 3.2 will be established.

We now formulate the rate function. Consider the controlled limiting equation:

$$\begin{aligned}\langle Z_t,\phi\rangle &= \int_0^t \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle ds\\ &\quad+\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \sigma(x,y,\mathcal{L}(X_s))z_1\phi'(x)Q(dx,dy,dz,ds)\\ &\quad+\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x,y,\mathcal{L}(X_s))z_1\\ &\quad+\tau_2(x,y,\mathcal{L}(X_s))z_2]\Phi_y(x,y,\mathcal{L}(X_s))\phi'(x)Q(dx,dy,dz,ds)\\ \bar{L}_\nu\phi(x) &:= \bar{\gamma}(x,\nu)\phi'(x)+\bar{D}(x,\nu)\phi''(x)\\ &\quad+\int_{\mathbb{R}}\left(\frac{\delta}{\delta m}\bar{\gamma}(z,\nu)[x]\phi'(z)+\frac{\delta}{\delta m}\bar{D}(z,\nu)[x]\phi''(z)\right)\nu(dz),\ \nu\in\mathcal{P}(\mathbb{R}).\end{aligned} \tag{32}$$

for all $\phi \in C_c^\infty(\mathbb{R})$. Here we recall the limiting coefficients $\bar{\gamma}, \bar{D}$ from Eq. (24), the limiting McKean–Vlasov Equation X_t from Eq. (25), and the linear functional derivative $\frac{\delta}{\delta m}$ from Definition D.4.

Theorem 3.1 *Let assumptions (A1)– (A13) hold. Then $\{Z^N\}_{N\in\mathbb{N}}$ satisfies the Laplace principle (29) with rate function I given by*

$$I(Z) = \inf_{Q\in P^*(Z)}\left\{\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(z_1^2+z_2^2\right)Q(dx,dy,dz,ds)\right\} \tag{33}$$

where $Q \in M_T(\mathbb{R}^4)$ (recall this space from above Eq. 10) is in $P^*(Z)$ if:

- P^*1 *(Z, Q) satisfies Eq. (32)*
- P^*2 $\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(z_1^2+z_2^2\right)Q(dx,dy,dz,ds) < \infty$
- P^*3 *Disintegrating $Q(dx,dy,dz,ds) = \kappa(dz;x,y,s)\lambda(dy;x,s)Q_{(1,4)}(dx,ds)$, $\lambda(dy;x,s) = \pi(dy;x,\mathcal{L}(X_s))$ $\nu_{\mathcal{L}(X.)}$-almost surely, where π is as in Eq. (20) and $\nu_{\mathcal{L}(X.)}$ is as in Eq. (10).*
- P^*4 $Q_{(1,4)} = \nu_{\mathcal{L}(X.)}$.

Here we use the convention that $\inf\{\emptyset\} = +\infty$.

Replacing assumption (A13) by the following:

(A'13) For r as in Eq. (7) and $\bar{\gamma}$,$\bar{D}$ as in Eq. (24), $\bar{\gamma}, \bar{D} \in \mathcal{M}_b^{\zeta_{x,r+2}}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R})) \cap \mathcal{M}_{\delta,b}^{\bar{\zeta}_{r+2}}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ (recalling these spaces from Definition 2.4 and these collections of multi-indices from Eq. 27), and

$$\sup_{x\in\mathbb{R},\mu\in\mathcal{P}(R)}\left\|\frac{\delta}{\delta m}\bar{\gamma}(x,\mu)[\cdot]\right\|_{r+2}+\sup_{x\in\mathbb{R},\mu\in\mathcal{P}(R)}\left\|\frac{\delta}{\delta m}\bar{D}(x,\mu)[\cdot]\right\|_{r+2}<\infty.$$

we can in addition prove compactness of level sets of the rate function given in (33) by extending it to a larger space. For a discussion of the necessity of this extension, see the comments below Eq. (2.10) and below Eq. (4.33) in [10]. This yields the main result:

Theorem 3.2 *Let assumptions (A1)–(A12) and (A'13) hold. Then $\{Z^N\}_{N\in\mathbb{N}}$ from Eq. (3) satisfies the large deviation principle on the space $C([0,T];\mathcal{S}_{-r})$, with r as in Eq. (7), speed $a^{-2}(N)$ and good rate function I given as in Eq. (33). Here we use the convention that $\inf\{\emptyset\}=+\infty$, and also impose that $I(Z)=+\infty$ for $Z\in C([0,T];\mathcal{S}_{-r})\backslash C([0,T];\mathcal{S}_{-w})$.*

As is typically the case when using the weak convergence approach of [22] to prove a large deviations principle, the rate function (33) can also be characterized by controls in feedback form:

Corollary 3.3 *In the setting of Theorem 3.1, we can alternatively characterize the rate function as:*

$$I^o(Z)=\inf_{h\in P^o(Z)}\left\{\frac{1}{2}\int_0^T\mathbb{E}\left[\int_{\mathbb{R}}|h(s,X_s,y)|^2\pi(dy;X_s,\mathcal{L}(X_s))\right]ds\right\} \tag{34}$$

where $h:[0,T]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}^2$ is in $P^o(Z)$ if:

(P^o1) (Z,h) satisfies Eq. (35) for all $t\in[0,T]$ and $\phi\in C_c^\infty(\mathbb{R})$

(P^o2) $\int_0^T\mathbb{E}\left[\int_{\mathbb{R}}|h(s,X_s,y)|^2\pi(dy;X_s,\mathcal{L}(X_s))\right]ds<\infty$.

Here we define:

$$\begin{aligned}\langle Z_t,\phi\rangle&=\int_0^t\langle Z_s,\bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle ds\\&\quad+\int_0^t\mathbb{E}\left[\int_{\mathbb{R}}\sigma(X_s,y,\mathcal{L}(X_s))h_1(s,X_s,y)\phi'(X_s)\pi(dy;X_s,\mathcal{L}(X_s))\right]ds\\&\quad+\int_0^t\mathbb{E}\left[\int_{\mathbb{R}}[\tau_1(X_s,y,\mathcal{L}(X_s))h_1(s,X_s,y)+\tau_2(X_s,y,\mathcal{L}(X_s))h_2(s,X_s,y)]\right.\\&\quad\left.\times\Phi_y(X_s,y,\mathcal{L}(X_s))\phi'(X_s)\pi(dy;X_s,\mathcal{L}(X_s))\right]ds\end{aligned} \tag{35}$$

Again, we use the convention that $\inf\{\emptyset\}=+\infty$. In the setting of Theorem 3.2, we also impose that $I^o(Z)=+\infty$ for $Z\in C([0,T];\mathcal{S}_{-r})\backslash C([0,T];\mathcal{S}_{-w})$.

Proof This follows from Jensen's inequality and the affine dependence of the coefficients on the controls. The details are omitted for brevity given that the argument is standard, e.g., see Section 5 in [23]. □

In addition, as a corollary to the proof of Theorem 3.2, we extend the results from [10] as follows:

Corollary 3.4 *(MDP without Multiscale Structure) Suppose that* $b = f = g = \tau_1 = \tau_2 \equiv 0$ *and* $c(x, y, \mu) = c(x, \mu), \sigma(x, y, \mu) = \sigma(x, \mu)$. *Let* $\upsilon > 4$ *be sufficiently large that the canonical embedding* $\mathcal{S}_{-4} \to \mathcal{S}_{-\upsilon}$ *is Hilbert–Schmidt,* $\rho > 6$ *be sufficiently large that the canonical embedding* $\mathcal{S}_{-\upsilon-2} \to \mathcal{S}_{-\rho}$ *is Hilbert–Schmidt, and* $\bar{\boldsymbol{\zeta}}$ *as in (27). Assume also that* $\sigma, c \in \mathcal{M}^{\boldsymbol{\zeta}}_{\delta,b}(\mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$ *and for* $F(x, \mu) = c(x, \mu)$ *or* $\sigma(x, \mu)$*:*

(1) $\sup_{\mu\in\mathcal{P}_2(\mathbb{R})} |F(\cdot, \mu)|_{\rho+2} < \infty$
(2) $\sup_{x\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} \left\| \frac{\delta}{\delta m} F(x, \mu)[\cdot]\right\|_{\rho+2} < \infty.$

Here we recall the space $\mathcal{M}_{\delta,b}$ *from Definition 2.4, the collection of multi-indices* $\boldsymbol{\zeta}$ *from Eq. (27), and the norms on* $\mathcal{S}$ *defined in Eqs. (4) and (8). Then* $\{Z^N\}_{N\in\mathbb{N}}$ *satisfies a large deviation principle on the space* $C([0, T]; \mathcal{S}_{-\rho})$ *with speed* $a^{-2}(N)$ *and good rate function* $\tilde{I}^o$ *given by*

$$\tilde{I}^o(Z) = \inf_{h\in\tilde{P}^o(Z)} \left\{ \frac{1}{2} \int_0^T \mathbb{E}\left[|h(s, X_s)|^2\right] ds \right\} \tag{36}$$

where $h : [0, T] \times \mathbb{R} \to \mathbb{R}$ *is in* $\tilde{P}^o(Z)$ *if:*

(P^o1) (Z, h) *satisfies Eq. (37) for all* $t \in [0, T]$ *and* $\phi \in C_c^\infty(\mathbb{R})$
(P^o2) $\int_0^T \mathbb{E}\left[|h(s, X_s)|^2\right] ds < \infty$

and $\inf\{\emptyset\} = +\infty$, $I(Z) = +\infty$ *for* $Z \in C([0, T]; \mathcal{S}_{-\rho})\backslash C([0, T]; \mathcal{S}_{-\upsilon})$. *Here we define:*

$$\begin{aligned}
\langle Z_t, \phi\rangle &= \int_0^t \langle Z_s, \tilde{L}_{\mathcal{L}(\tilde{X}_s)}\phi(\cdot)\rangle ds + \int_0^t \mathbb{E}\left[\sigma(\tilde{X}_s, \mathcal{L}(\tilde{X}_s))h(s, \tilde{X}_s)\phi'(\tilde{X}_s)\right] ds\\
\tilde{L}_\nu\phi(x) &= c(x, \nu)\phi'(x) + \frac{\sigma^2(x, \nu)}{2}\phi''(x)\\
&\quad + \int_{\mathbb{R}} \left(\frac{\delta}{\delta m} c(z, \nu)[x]\phi'(z) + \frac{1}{2}\frac{\delta}{\delta m}[\sigma^2(z, \nu)[x]]\phi''(z)\right) \nu(dz)\\
\tilde{X}_t &= \eta^x + \int_0^t c(\tilde{X}_s, \mathcal{L}(\tilde{X}_s))ds + \int_0^t \sigma(\tilde{X}_s, \mathcal{L}(\tilde{X}_s))dW_s.
\end{aligned} \tag{37}$$

Proof This follows from Theorem 3.2. The assumptions needed are vastly simplified due to the absence of multiscale structure. In particular, we have no need for the results from Sect. 6 and Sect. 7.1. The rate function can be posed on a smaller space $C([0, T]; \mathcal{S}_{-\rho})$ (agreeing with that of Theorem 2.1 in [10]) as opposed to the larger $C([0, T]; \mathcal{S}_{-r})$ of our Theorem 3.2 due to the IID system (57) not depending on ϵ in this regime. In particular, this means $\bar{X}^\epsilon_t \overset{d}{=} \tilde{X}_t$ in the proof of Lemma 7.6, and hence the result is improved $C(T)|\phi|_1^2$ instead of $C(T)|\phi|_4^2$. Similarly, in the result of Lemma 7.7, the bound on $R^N_t(\phi)$ can be improved from $\bar{R}(N, T)|\phi|_4$ to $\bar{R}(N, T)|\phi|_3$ using Lemma 7.5 and the proof method of Proposition 4.2 in [10]. At this point tightness of $\{\tilde{Z}^N\}_{N\in\mathbb{N}}$ from Eq. (54) can be proved in Proposition 7.8, but with the uniform 7-continuity of Eq. (67) improved to uniform 4-continuity, and hence the result holds with w replaced by v. The remainder of the proofs in the paper found in Sects. 7.4, 8 and 9 then go through verbatim with m and w replaced by v and r replaced with ρ, but with the simplifications assumed on the coefficients allowing us to set many terms equal to 0. In particular, in the controlled particle Eq. (55), we can set $\tilde{Y}^{i,\epsilon,N}\equiv 0$, and throughout the invariant measure π from Eq. (20) can be set to δ_0, which makes dealing with the second marginals of the occupation measures $\{Q^N\}_{N\in\mathbb{N}}$ from Eq. (56) trivial. Lastly, in Sect. 10, due to the lack of multiscale structure, there is no need for an approximation argument in the proof of Proposition 10.1, and hence existence of solutions to (37) can be established in $C([0,T];\mathcal{S}_{-v})$ and compactness of level sets established in $C([0,T];\mathcal{S}_{-\rho})$ exactly as in Sects. 4.4 and 4.5 of [10]. □

Remark 3.5 Note that, in contrast to [10], which assumes a linear-in-measure form of the coefficients of Eq. (1) (without multiscale structure), i.e. that there are $\beta,\alpha:\mathbb{R}^2\to\mathbb{R}$ such that $c(x,\mu)=\int_{\mathbb{R}}\beta(x,z)\mu(dz)$, $\sigma(x,\mu)=\int_{\mathbb{R}}\alpha(x,z)\mu(dz)$, we do not suppose any particular form of $c(x,\mu),\sigma(x,\mu)$ other than that they have sufficient regularity for the proof of tightness and existence/uniqueness of the limiting equation. We are able to do so via the use of Lemma D.7 (which holds also in the case without dependence of the function p on y) and the assumption that $\sigma,c\in\mathcal{M}^{\boldsymbol{\zeta}}_{\boldsymbol{\delta},b}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$. For the specific linear form of c and σ assumed by [10], $\frac{\delta}{\delta m}c(x,\mu)[z]=\beta(x,z)$ and $\frac{\delta}{\delta m}\sigma(x,\mu)[z]=\alpha(x,z)$, so the condition (2) from Corollary 3.4 in fact implies $\sigma,c\in\mathcal{M}^{\boldsymbol{\zeta}}_{\boldsymbol{\delta},b}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$. In addition, (1) and (2) are exactly the assumptions (a) and (b) from Condition 2.3 of [10] in this subcase, so indeed Corollary 3.4 provides a strict generalization of their result. See also Corollary 4.6 where we further extend this result to get an alternate form of the rate function analogous to that of Dawson-Gärtner [17].

It is also useful to characterize the way that the limiting Eqs. (32), (35), and (37) act on functions which depend both on time and space. Hence we make the following remark:

Remark 3.6 We can alternatively characterize the controlled limiting Eq. (32) (and analogously the ordinary controlled limiting Eqs. 35 and 37) in terms of how the Z acts on $\psi\in C_c^\infty(U\times\mathbb{R})$, where U is an open interval containing $[0,T]$. For Eq. (35), this characterization is:

$$
\begin{aligned}
\langle Z_T,\psi(T,\cdot)\rangle &= \int_0^T\langle Z_s,\dot{\psi}(s,\cdot)\rangle ds+\int_0^T\langle Z_s,\bar{L}_{\mathcal{L}(X_s)}\psi(s,\cdot)\rangle ds\\
&\quad+\int_0^T\mathbb{E}\left[\int_{\mathbb{R}}\sigma(X_s,y,\mathcal{L}(X_s))h_1(s,X_s,y)\psi_x(s,X_s)\pi(dy;X_s,\mathcal{L}(X_s))\right]ds\\
&\quad+\int_0^T\mathbb{E}\left[\int_{\mathbb{R}}[\tau_1(X_s,y,\mathcal{L}(X_s))h_1(s,X_s,y)+\tau_2(X_s,y,\mathcal{L}(X_s))h_2(s,X_s,y)]\right.\\
&\quad\left.\Phi_y(X_s,y,\mathcal{L}(X_s))\psi_x(s,X_s)\pi(dy;X_s,\mathcal{L}(X_s))\right]ds\\
Z_0&=0.
\end{aligned}
\tag{38}
$$

This is analogous to the form of the limiting equation seen in [59] (Remark 2.9) and [10] (Remark 2.2).

4 On the form of the rate function

4.1 Statement and proof of equivalent forms of the rate function

Here we prove an alternative form of the moderate deviations rate function (33), which is analogous to the "negative Sobolev" form of the large deviations rate function for the empirical measure of weakly interacting diffusions found in Theorem 5.1 of the classical work of Dawson-Gärtner [17]. This is the first time such a form of the rate function has been provided in the moderate deviations setting, both with and without multiscale structure. The result for the specialized case without multiscale structure can be found as Corollary 4.6 below.

A direct connection between the variational form of the large deviations rate function from [9] and the "negative Sobolev" form of [17] was recently made for the first time in [4] Section 5.2. In contrast to the large deviations setting, in the moderate deviations rate function (34), we already know the controls h are in feedback form, but rather than being feedback controls of the limiting controlled processes Z in Eq. (35), they are feedback controls of the law of large numbers $\mathcal{L}(X)$ from Eq. (25). Moreover, contrast to in the large deviations setting of [4], here we handle the dependence of the controls h on the parameter y do to the multiscale structure and obtaining the "negative Sobolev" form of the rate function uniformly.

In order to state the alternate form of the rate function we first need to recall the following definition:

Definition 4.1 (Definition 4.1 in [17]) For a compact set $K \subset \mathbb{R}$, we will denote the subspace of $C_c^\infty(\mathbb{R})$ which have compact support contained in K by $\mathcal{S}_K$. Let I be an interval on the real line. A map $Z : I \to \mathcal{S}'$ is called absolutely continuous if for each compact set $K \subset \mathbb{R}$, there exists a neighborhood of 0 in $\mathcal{S}_K$ and an absolutely continuous function $H_K : I \to \mathbb{R}$ such that

$$|\langle Z(u), \phi\rangle - \langle Z(v), \phi\rangle| \leq |H_K(u) - H_K(v)|$$

for all $u, v \in I$ and $\phi \in U_K$.

It is also useful to recall the following result:

Lemma 4.2 *(Lemma 4.2 in [17]) Assume the map $Z : I \to \mathcal{S}'$ is absolutely continuous. Then the real function $\langle Z, \phi\rangle$ is absolutely continuous for each $\phi \in C_c^\infty(\mathbb{R})$ and the derivative in the distribution sense*

$$\dot{Z}(t) := \lim_{h\downarrow 0} h^{-1}[Z(t+h) - Z(t)]$$

exists for Lebesgue almost-every $t \in I$.

Now we are ready to state the equivalent form of the rate function:

Proposition 4.3 *Let assumptions (A1)–(A12) and (A'13) hold. Assume also that $\bar{D}(x,\mu) > 0$ for all $x \in \mathbb{R}, \mu \in \mathcal{P}_2$, where $\bar{D}$ is as in Eq. (24). Let r be as in Eq. (7). Consider $I^{DG} : C([0,T];\mathcal{S}_{-r}) \to [0,+\infty]$ given by:*

$$I^{DG}(Z) = \frac{1}{4}\int_0^T \sup_{\phi \in C_c^\infty(\mathbb{R}):\mathbb{E}[\bar{D}(X_t,\mathcal{L}(X_t))|\phi'(X_t)|^2]\neq 0} \frac{\left|\langle \dot{Z}_t - \bar{L}^*_{\mathcal{L}(X_t)} Z_t, \phi\rangle\right|^2}{\mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\phi'(X_t)|^2\right]} dt \tag{39}$$

if $Z(0) = 0$, Z is absolutely continuous in the sense if Definition 4.1, and $Z \in C([0,T];\mathcal{S}_{-w})$, and $I^{DG}(Z) = +\infty$ otherwise. Here X_t is as in Eq. (25), $\dot{Z}$ is the time derivative of Z in the distribution sense from Lemma 4.2 and $\bar{L}^*_{\mathcal{L}(X_s)} : \mathcal{S}_{-w} \to \mathcal{S}_{-(w+2)}$ is the adjoint of $\bar{L}_{\mathcal{L}(X_s)} : \mathcal{S}_{w+2} \to \mathcal{S}_w$ given in Eq. (32) (using here Lemma 8.6).

Then $\{Z^N\}_{N\in\mathbb{N}}$ from Eq. (3) satisfies a large deviation principle on the space $C([0,T];\mathcal{S}_{-r})$ with speed $a^{-2}(N)$ and good rate function I^{DG}.

Remark 4.4 Note that the assumption that $\bar{D}(x,\mu) > 0$ for all $x \in \mathbb{R}$ and $\mu \in \mathcal{P}_2(\mathbb{R})$ is not very restrictive. In particular, via the representation for the density of the invariant measure π given in Eq. (72), we know it is strictly positive for all x, μ. Then via the representation for $\bar{D}(x,\mu)$ given in Eq. (26), we have that if there is x, μ such that $\bar{D}(x,\mu) = 0$, then for that x, μ, we must have

$$[\tau_2(x,y,\mu)\Phi_y(x,y,\mu)]^2 + [\sigma(x,y,\mu) + \tau_1(x,y,\mu)\Phi_y(x,y,\mu)]^2 = 0$$

for Lebesgue-almost every $y \in \mathbb{R}$. This will only happen if σ has a very specific relation to f, τ_1, τ_2, b and hence Φ_y.

In order to prove Proposition 4.3, we first prove the following Lemma, which gives us a form of the rate function analogous to Eq. (4.21) in [17]:

Lemma 4.5 *Assume the same setup as Proposition 4.3. For $\psi \in C_c^\infty(U \times \mathbb{R})$ and $Z \in S_{-w}$, define*

$$F_Z(\psi) = \langle Z_T, \psi(T,\cdot)\rangle - \int_0^T \langle Z_s, \dot{\psi}(s,\cdot)\rangle ds - \int_0^T \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\psi(s,\cdot)\rangle ds \tag{40}$$

and consider $J : C([0,T];\mathcal{S}_{-\rho}) \to [0,+\infty]$ given by:

$$J(Z) = \sup_{\psi \in C_c^\infty(U\times\mathbb{R})} \left\{ F_Z(\psi) - \int_0^T \mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x(t,X_t)|^2\right] dt \right\} \tag{41}$$

if $Z_0 = 0$ and $Z \in C([0,T];\mathcal{S}_{-w})$, and $J(Z) = +\infty$ otherwise. Here U is an open interval containing $[0,T]$. Then $\{Z^N\}_{N\in\mathbb{N}}$ satisfies a large deviation principle on the space $C([0,T];\mathcal{S}_{-r})$ with speed $a^{-2}(N)$ and good rate function J.

Proof Since by Theorem 1.3.1 in [22], the rate function for a sequence of random variables is unique, it suffices to show that $I^o = J$, where I^o is from Corollary 3.3. We note that by Remark 3.6, we can replace (P^o1) in definition of the multiscale ordinary rate function I^o by Z satisfying Eq. (38). We will also use the alternative form of $\bar{D}(x,\mu)$ provided by Eq. (26) in Remark 2.5.

First we show $J \leq I^o$. Take Z such $I^o(Z) < \infty$. Then $P^o(Z)$ is non-empty, and for any $h \in P^o(Z)$ and, by Eq. (38), for any $\psi \in C_c^\infty(U \times \mathbb{R})$:

$$
\begin{aligned}
|F_Z(\psi)| &= \Bigg|\int_0^T \mathbb{E}\Bigg[\int_{\mathbb{R}}\Big([\sigma(X_s,y,\mathcal{L}(X_s)) + \tau_1(X_s,y,\mathcal{L}(X_s))\Phi_y(X_s,y,\mathcal{L}(X_s))]h_1(s,X_s,y) \\
&\quad + \tau_2(X_s,y,\mathcal{L}(X_s))\Phi_y(X_s,y,\mathcal{L}(X_s))h_2(s,X_s,y)\Big)\psi_x(s,X_s)\pi(dy;X_s,\mathcal{L}(X_s))\Bigg]ds\Bigg| \\
&\leq \Bigg(\int_0^T \mathbb{E}\Bigg[\int_{\mathbb{R}}\Big([\sigma(X_s,y,\mathcal{L}(X_s)) + \tau_1(X_s,y,\mathcal{L}(X_s))\Phi_y(X_s,y,\mathcal{L}(X_s))]^2 \\
&\quad + [\tau_2(X_s,y,\mathcal{L}(X_s))\Phi_y(X_s,y,\mathcal{L}(X_s))]^2\Big)\pi(dy;X_s,\mathcal{L}(X_s))|\psi_x(s,X_s)|^2\Bigg]ds\Bigg)^{1/2} \\
&\quad \times \Bigg(\int_0^T \mathbb{E}\Bigg[\int_{\mathbb{R}} |h_1(s,X_s,y)|^2 + |h_2(s,X_s,y)|^2\pi(dy;X_s,\mathcal{L}(X_s))\Bigg]ds\Bigg)^{1/2} \\
&= \sqrt{2}\Bigg(\int_0^T \mathbb{E}\Bigg[\bar{D}(X_s,\mathcal{L}(X_s))|\psi_x(s,X_s)|^2\Bigg]ds\Bigg)^{1/2} \\
&\quad \Bigg(\int_0^T \mathbb{E}\Bigg[\int_{\mathbb{R}} |h(s,X_s,y)|^2\pi(dy;X_s,\mathcal{L}(X_s))\Bigg]ds\Bigg)^{1/2}
\end{aligned}
$$

so in particular, if $\int_0^T \mathbb{E}\Big[\bar{D}(X_s,\mathcal{L}(X_s))|\psi_x(s,X_s)|^2\Big]ds = 0$, then $F_Z(\psi) = 0$. Then, observing that $\psi \in C_c^\infty(U \times \mathbb{R})$ if and only if for any $c \in \mathbb{R}\backslash\{0\}$, $c\psi \in C_c^\infty(U \times \mathbb{R})$ and that F_Z is linear, we have:

$$
\begin{aligned}
J(Z) = &\sup_{\psi\in C_c^\infty(U\times\mathbb{R}):\int_0^T \mathbb{E}\Big[\bar{D}(X_s,\mathcal{L}(X_s))|\psi_x(s,X_s)|^2\Big]ds\neq 0} \\
&\left\{F_Z(\psi) - \int_0^T \mathbb{E}\Big[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x(t,X_t)|^2\Big]dt\right\} \vee 0 \\
= &\sup_{\psi\in C_c^\infty(U\times\mathbb{R}):\int_0^T \mathbb{E}\Big[\bar{D}(X_s,\mathcal{L}(X_s))|\psi_x(s,X_s)|^2\Big]ds\neq 0} \ \sup_{c\in\mathbb{R}} \\
&\left\{cF_Z(\psi) - c^2\int_0^T \mathbb{E}\Big[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x(t,X_t)|^2\Big]dt\right\} \vee 0 \\
= &\sup_{\psi\in C_c^\infty(U\times\mathbb{R}):\int_0^T \mathbb{E}\Big[\bar{D}(X_s,\mathcal{L}(X_s))|\psi_x(s,X_s)|^2\Big]ds\neq 0}
\end{aligned}
$$

$$\frac{|F_Z(\psi)|^2}{4\int_0^T \mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x(t,X_t)|^2\right]dt}.$$

Returning to the above inequality and squaring both sides, we have

$$\frac{|F_Z(\psi)|^2}{2\int_0^T \mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x(t,X_t)|^2\right]dt}$$
$$\leq \int_0^T \mathbb{E}\left[\int_{\mathbb{R}} |h(s,X_s,y)|^2\pi(dy;X_s,\mathcal{L}(X_s))\right]ds,$$

for all $\psi \in C_c^\infty(U\times\mathbb{R})$ such that $\int_0^T \mathbb{E}\left[\bar{D}(X_s,\mathcal{L}(X_s))|\psi_x(s,X_s)|^2\right]ds \neq 0$ and all $h \in P^o(Z)$. So $J(Z) \leq I^o(Z)$.

Now we prove $J \geq I^o$. Assume without loss of generality that $J(Z) \leq C < \infty$. Then, since

$$J(Z) = \sup_{\psi\in C_c^\infty(U\times\mathbb{R})}\sup_{c\in\mathbb{R}}\left\{cF_Z(\psi) - c^2\int_0^T \mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x(t,X_t)|^2\right]dt\right\}$$
$$= +\infty$$

if there exists $\psi \in C_c^\infty(U\times\mathbb{R})$ such that $F_Z(\psi)\neq 0$ and $\int_0^T \mathbb{E}[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x(t,X_t)|^2]dt = 0$, we have

$$J(Z) = \sup_{\psi\in C_c^\infty(U\times\mathbb{R}):\int_0^T \mathbb{E}\left[\bar{D}(X_s,\mathcal{L}(X_s))|\psi_x(s,X_s)|^2\right]ds\neq 0} \frac{|F_Z(\psi)|^2}{4\int_0^T \mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x(t,X_t)|^2\right]dt}.$$

This shows that for all $\psi \in C_c^\infty(U\times\mathbb{R})$

$$\left|F_Z(\psi)\right| \leq 2\sqrt{C}\left(\int_0^T \mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x(t,X_t)|^2\right]dt\right)^{1/2}. \tag{42}$$

Now we borrow some notation from [17] (see pp. 270–271). We let for $t\in[0,T]$ ∇_t, $(\cdot,\cdot)_t$, and $|\cdot|_t$ be (formally) the Riemannian gradient, inner product, and Riemannian norm in the tangent space of the Riemannian structure on $\mathbb{R}$ induced by the diffusion matrix $t\mapsto \bar{D}(\cdot,\mathcal{L}(X_t))$, i.e.:

$$\nabla_t f := \bar{D}(\cdot,\mathcal{L}(X_t))\frac{df}{dx}, \qquad (X,Y)_t := \bar{D}(\cdot,\mathcal{L}(X_t))^{-1}XY, \qquad |X|_t := (X,X)_t^{1/2}.$$

In particular, note that

$$|\nabla_t f|_t^2 = \bar{D}(\cdot, \mathcal{L}(X_t))|\frac{df}{dx}|^2.$$

Now, as on p. 279 in [17], we define $L^2[0,T]$ to be the Hilbert space of measurable maps $g : [0,T] \times \mathbb{R} \to \mathbb{R}$ with finite norm

$$\|g\| := \left(\int_0^T \langle \mathcal{L}(X_t), |g(t,\cdot)|_t^2\rangle dt\right)^{1/2} = \left(\int_0^T \mathbb{E}[\bar{D}(X_t, \mathcal{L}(X_t))^{-1}|g(t,X_t)|^2]dt\right)^{1/2}$$

and inner product

$$\begin{aligned}[g_1, g_2] &:= \int_0^T \langle \mathcal{L}(X_t), (g_1(t,\cdot), g_2(t,\cdot))_t\rangle dt \\ &= \int_0^T \mathbb{E}[\bar{D}(X_t, \mathcal{L}(X_t))^{-1} g_1(t,X_t) g_2(t,X_t)]dt.\end{aligned}$$

Denote by $L^2_\nabla[0,T]$ the closure in $L^2[0,T]$ of the linear subset L consisting of all maps $(s,x) \mapsto \nabla_s \psi(s,x)$, $\psi \in C_c^\infty(U \times \mathbb{R})$. Then F_Z can be viewed as a linear functional on L, and by the bound (42), is bounded. Then, by the Riesz Representation Theorem, there exists $\bar{h} \in L^2_\nabla[0,T]$ such that

$$\begin{aligned}F_Z(\psi) &= \int_0^T \langle \mathcal{L}(X_s), (\bar{h}(s,\cdot), \nabla_s\psi(s,\cdot))_s\rangle ds \\ &= \int_0^T \mathbb{E}[h(s,X_s)\psi_x(s,X_s)]ds,\ \psi \in C_c^\infty(U \times \mathbb{R}). \qquad (43)\end{aligned}$$

Note that actually, L^2_∇ must be considered not as a class of functions, but as a set of equivalence classes of functions agreeing $\nu_{\mathcal{L}(X_\cdot)}$-almost surely. This is of no consequence, however, since the bound (42) ensures that $F_Z(\psi) = F_Z(\tilde{\psi})$ if ψ_x and $\tilde{\psi}_x$ are in the same equivalence class (see p. 279 in [17] and Appendix D.5 in [28] for a more thorough treatment of the space $L^2_\nabla[0,T]$ and its dual).

Consider $\tilde{h}(s,x,y) : [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ given by

$$\begin{aligned}\tilde{h}_1(s,x,y) &= \frac{1}{2\bar{D}(x,\mathcal{L}(X_s))}[\sigma(x,y,\mathcal{L}(X_s)) + \tau_1(x,y,\mathcal{L}(X_s))\Phi_y(x,y,\mathcal{L}(X_s))]\bar{h}(s,x) \\ \tilde{h}_2(s,x,y) &= \frac{1}{2\bar{D}(x,\mathcal{L}(X_s))}\tau_2(x,y,\mathcal{L}(X_s))\Phi_y(x,y,\mathcal{L}(X_s))\bar{h}(s,x). \qquad (44)\end{aligned}$$

We have:

$$\begin{aligned}&\int_0^T \mathbb{E}\left[\int_{\mathbb{R}} |\tilde{h}(s,X_s,y)|^2 \pi(dy; X_s, \mathcal{L}(X_s))\right] ds \\ &= \int_0^T \mathbb{E}\left[\frac{|\bar{h}(s,X_s)|^2}{4|\bar{D}(x,\mathcal{L}(X_s))|^2}\int_{\mathbb{R}}\Big([\sigma(X_s,y,\mathcal{L}(X_s)) + \tau_1(X_s,y,\mathcal{L}(X_s))\Phi_y(X_s,y,\mathcal{L}(X_s))]^2\right.\end{aligned}$$

$$+ [\tau_2(X_s, y, \mathcal{L}(X_s))\Phi_y(X_s, y, \mathcal{L}(X_s))]^2\Big)\pi(dy; X_s, \mathcal{L}(X_s))\Big]ds$$
$$= \frac{1}{2}\int_0^T \mathbb{E}\Big[\bar{D}(x, \mathcal{L}(X_s))^{-1}|\bar{h}(s, X_s)|^2\Big]ds < \infty. \tag{45}$$

Moreover, for $\psi \in C_c^\infty(U \times \mathbb{R})$:

$$\int_0^T \mathbb{E}\Big[\int_{\mathbb{R}}\Big([\sigma(X_s, y, \mathcal{L}(X_s)) + \tau_1(X_s, y, \mathcal{L}(X_s))\Phi_y(X_s, y, \mathcal{L}(X_s))]\tilde{h}_1(s, X_s, y)$$
$$+ \tau_2(X_s, y, \mathcal{L}(X_s))\Phi_y(X_s, y, \mathcal{L}(X_s))\tilde{h}_2(s, X_s, y)\Big)\psi_x(s, X_s)\pi(dy; X_s, \mathcal{L}(X_s))\Big]ds$$
$$= \int_0^T \mathbb{E}\Big[\int_{\mathbb{R}}\Big([\sigma(X_s, y, \mathcal{L}(X_s)) + \tau_1(X_s, y, \mathcal{L}(X_s))\Phi_y(X_s, y, \mathcal{L}(X_s))]^2$$
$$+ [\tau_2(X_s, y, \mathcal{L}(X_s))\Phi_y(X_s, y, \mathcal{L}(X_s))]^2\Big)\pi(dy; X_s, \mathcal{L}(X_s))\frac{\bar{h}(s, X_s)}{2\bar{D}(X_s, \mathcal{L}(X_s)}\psi_x(s, X_s)\Big]ds$$
$$= \int_0^T \mathbb{E}\Big[\bar{h}(s, X_s)\psi_x(s, X_s)\Big]ds$$
$$= F_Z(\psi) \text{ by Eq. (43).}$$

Thus, $\tilde{h} \in P^o(Z)$ by definition. Take a sequence $\{\tilde{\psi}^n\} \subset L$ such that $\tilde{\psi}^n \to \bar{h}$ in $L^2[0, T]$. By virtue of $\tilde{\psi}^n \in L$, we have for each n, there is $\psi^n \in C_c^\infty(U \times \mathbb{R})$ such that $\tilde{\psi}^n(s, x) = \nabla_s\psi^n(s, x) = \bar{D}(x, \mathcal{L}(X_s))\psi_x^n(s, x)$.

In particular, we have $|\tilde{\psi}^n|_t^2 \to |\bar{h}|_t^2$, so

$$\int_0^T \mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))|\psi_x^n(t, X_t)|^2\Big]dt \to \int_0^T \mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))^{-1}|\bar{h}(t, X_t)|^2\Big]dt$$

and $(\tilde{\psi}^n, \bar{h})_t \to |\bar{h}|_t^2$, so

$$\int_0^T \mathbb{E}\Big[\psi_x^n(t, X_t)h(t, X_t)\Big]dt \to \int_0^T \mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))^{-1}|\bar{h}(t, X_t)|^2\Big]dt.$$

Note that if $\int_0^T \mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))^{-1}|\bar{h}(t, X_t)|^2\Big]dt = 0$, we have by Eq. (45), that the relation holds $\int_0^T \mathbb{E}\Big[\int_{\mathbb{R}}|\tilde{h}(s, X_s, y)|^2\pi(dy)\Big]ds = 0$, and hence $I^o(Z) = 0$, so the desired bound is trivial.

Assuming then that $\int_0^T \mathbb{E}\left[D(X_t,\mathcal{L}(X_t))^{-1}|\bar{h}(t,X_t)|^2\right]dt \neq 0$, we may choose a subsequence of $\{\psi_x^n\}$ such that $\int_0^T \mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x^n(t,X_t)|^2\right]dt \neq 0, \forall n$. Then:

$$
\begin{aligned}
J(Z) &\geq \frac{1}{4}\frac{\left(\int_0^T \mathbb{E}\left[\psi_x^n(s,X_s)\bar{h}(s,X_s)\right]ds\right)^2}{\int_0^T \mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\psi_x^n(t,X_t)|^2\right]dt} \text{ for all } n\in\mathbb{N}\\
&\to \frac{1}{4}\int_0^T \mathbb{E}\left[D(X_t,\mathcal{L}(X_t))^{-1}|\bar{h}(t,X_t)|^2\right]dt \text{ as } n\to\infty.
\end{aligned}
$$

By Eq. (45),

$$
\begin{aligned}
&\frac{1}{4}\int_0^T \mathbb{E}\left[D(X_t,\mathcal{L}(X_t))^{-1}|\bar{h}(t,X_t)|^2\right]dt\\
&\quad=\frac{1}{2}\int_0^T \mathbb{E}\left[\int_{\mathbb{R}}|\tilde{h}(s,X_s,y)|^2\pi(dy;X_s,\mathcal{L}(X_s))\right]ds,
\end{aligned}
$$

so since $\tilde{h}\in P^o(Z)$:

$$
J(Z)\geq \frac{1}{2}\int_0^T \mathbb{E}\left[\int_{\mathbb{R}}|\tilde{h}(s,X_s,y)|^2\pi(dy;X_s,\mathcal{L}(X_s))\right]ds \geq I^o(Z).
$$

□

Now we are ready to prove Proposition 4.3.

Proof of Proposition 4.3 As noted, the form of the rate function proved in Lemma 4.5 is analogous to that of Eq. (4.21) in [17]. We follow the proof of Lemma 4.8 in [17], making changes to account for the multiscale structure and the entry of $\mathcal{L}(X_s)$ rather than Z_s in the subtracted term in Eq. (4.24), which comes the fact that we are looking at moderate deviations rather than large deviations. We also use the specific information about the optimal control from the proof of Lemma 4.5.

Once again, it is sufficient to show $I^o = I^{DG}$, or equivalently, $J = I^{DG}$. First we show that $I^o = J \leq I^{DG}$. Let $Z\in C([0,T];\mathcal{S}_{-w})$ be such that $I^{DG}(Z)<\infty$. Note that

$$
\begin{aligned}
&\sup_{\phi\in C_c^\infty(\mathbb{R}):\mathbb{E}[\bar{D}(X_t,\mathcal{L}(X_t))|\phi'(X_t)|^2]\neq 0}\left\{\langle \dot{Z}_t-\bar{L}^*_{\mathcal{L}(X_t)}Z_t,\phi\rangle-\mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\phi'(X_t)|^2\right]\right\}\\
&=\sup_{\phi\in C_c^\infty(\mathbb{R}):\mathbb{E}[\bar{D}(X_t,\mathcal{L}(X_t))|\phi'(X_t)|^2]\neq 0}\ \sup_{c\in\mathbb{R}}\\
&\quad\left\{c\langle \dot{Z}_t-\bar{L}^*_{\mathcal{L}(X_t)}Z_t,\phi\rangle-c^2\mathbb{E}\left[\bar{D}(X_t,\mathcal{L}(X_t))|\phi'(X_t)|^2\right]\right\}
\end{aligned}
$$

$$= \frac{1}{4} \sup_{\phi \in C_c^\infty(\mathbb{R}):\mathbb{E}[\bar{D}(X_t,\mathcal{L}(X_t))|\phi'(X_t)|^2]\neq 0} \frac{\left|\langle \dot{Z}_t - \bar{L}^*_{\mathcal{L}(X_t)} Z_t, \phi\rangle\right|^2}{\mathbb{E}\left[\bar{D}(X_t, \mathcal{L}(X_t))|\phi'(X_t)|^2\right]}$$

for all $t \in [0, T]$. So for any $\psi \in C_c^\infty(U \times \mathbb{R})$:

$$\begin{aligned} I^{DG}(Z) &= \frac{1}{4}\int_0^T \sup_{\phi \in C_c^\infty(\mathbb{R}):\mathbb{E}[\bar{D}(X_t,\mathcal{L}(X_t))|\phi'(X_t)|^2]\neq 0} \frac{\left|\langle \dot{Z}_t - \bar{L}^*_{\mathcal{L}(X_t)} Z_t, \phi\rangle\right|^2}{\mathbb{E}\left[\bar{D}(X_t, \mathcal{L}(X_t))|\phi'(X_t)|^2\right]} dt \\ &= \int_0^T \sup_{\phi \in C_c^\infty(\mathbb{R}):\mathbb{E}[\bar{D}(X_t,\mathcal{L}(X_t))|\phi'(X_t)|^2]\neq 0} \\ &\quad \left\{\langle \dot{Z}_t - \bar{L}^*_{\mathcal{L}(X_t)} Z_t, \phi\rangle - \mathbb{E}\left[\bar{D}(X_t, \mathcal{L}(X_t))|\phi'(X_t)|^2\right]\right\} dt \\ &\geq \int_0^T \langle \dot{Z}_t - \bar{L}^*_{\mathcal{L}(X_t)} Z_t, \psi(t,\cdot)\rangle - \mathbb{E}\left[\bar{D}(X_t, \mathcal{L}(X_t))|\psi_x(t, X_t)|^2\right] dt \\ &= \langle Z_T, \psi(T,\cdot)\rangle - \int_0^T \langle Z_t, \dot{\psi}(t,\cdot)\rangle dt - \int_0^T \langle Z_t, \bar{L}_{\mathcal{L}(X_t)}\psi(t,\cdot)\rangle dt \\ &\quad - \mathbb{E}\left[\bar{D}(X_t, \mathcal{L}(X_t))|\psi_x(t, X_t)|^2\right] dt, \end{aligned}$$

where in the last step we used Lemma 4.3 in [17]. Then taking the supremum over all $\psi \in C_c^\infty(U \times \mathbb{R})$, we get $I^{DG}(Z) \geq J(Z)$, as desired.

Now we show that $I^{DG} \leq J = I^o$. Consider $Z \in C([0, T]; \mathcal{S}_{-w})$ such that $J(Z) < \infty$.

In Lemma 4.5, we proved for $\tilde{h}$ as in Eq. (44), $\tilde{h} \in P^o(Z)$. We also showed:

$$\begin{aligned} &\frac{1}{2}\int_0^T \mathbb{E}\left[\int_{\mathbb{R}} |\tilde{h}(s, X_s, y)|^2 \pi(dy)\right] ds \\ &\quad \leq J(Z) = I^o(Z) \leq \frac{1}{2}\int_0^T \mathbb{E}\left[\int_{\mathbb{R}} |h(s, X_s, y)|^2 \pi(dy)\right] ds, \forall h \in P^o(Z), \end{aligned}$$

so that in fact

$$\begin{aligned} J(Z) = I^o(Z) &= \frac{1}{2}\int_0^T \mathbb{E}\left[\int_{\mathbb{R}} |\tilde{h}(s, X_s, y)|^2 \pi(dy)\right] ds \\ &= \frac{1}{4}\int_0^T \mathbb{E}\left[\bar{D}(X_s, \mathcal{L}(X_s))^{-1} |\bar{h}(s, X_s)|^2\right] ds, \end{aligned} \tag{46}$$

where in the last inequality we used Eq. (45).

Now, by the fact that $\tilde{h} \in P^o(Z)$, we have by Eq. (35) that for all $0 \le s \le t \le T$ and $\phi \in C_c^\infty(\mathbb{R})$:

$$\begin{aligned}\langle Z_t, \phi\rangle - \langle Z_s, \phi\rangle &= \int_s^t \langle Z_u, \bar{L}_{\mathcal{L}(X_u)}\phi(\cdot)\rangle du \\ &+ \int_s^t \mathbb{E}\bigg[\int_{\mathbb{R}}\bigg([\sigma(X_u, y, \mathcal{L}(X_u)) + \tau_1(X_u, y, \mathcal{L}(X_u))\Phi_y(X_u, y, \mathcal{L}(X_u))]\tilde{h}_1(u, X_u, y) \\ &+ \tau_2(X_u, y, \mathcal{L}(X_u))\Phi_y(X_u, y, \mathcal{L}(X_u))\tilde{h}_2(u, X_u, y)\bigg)\pi(dy; X_u, \mathcal{L}(X_u))\phi'(X_u)\bigg]du \\ &= \int_s^t \langle Z_u, \bar{L}_{\mathcal{L}(X_u)}\phi(\cdot)\rangle du + \int_s^t \mathbb{E}\Big[\bar{h}(u, X_u)\phi'(X_u)\Big]du\end{aligned}$$

where $\bar{h}$ is as in Eq. (43), so by Definition 4.1 and Lemma 8.6, Z is an absolutely continuous map from $[0, T]$ to $\mathcal{S}'$. Then, using Lemma 4.2, we have for each $\phi \in C_c^\infty(\mathbb{R})$:

$$\langle \dot{Z}_t, \phi\rangle = \mathbb{E}\Big[\bar{h}(t, X_t)\phi'(X_t)\Big] + \langle Z_t, \bar{L}_{\mathcal{L}(X_t)}\phi(\cdot)\rangle. \tag{47}$$

Using a density argument, we can make sure this holds simultaneously for all $\phi \in C_c^\infty(\mathbb{R})$ and Lebesgue almost every $t \in [0, T]$ (see p. 280 of [17]). This gives:

$$I^{DG}(Z) = \frac{1}{4}\int_0^T \sup_{\phi \in C_c^\infty(\mathbb{R}) : \mathbb{E}[\bar{D}(X_t, \mathcal{L}(X_t))|\phi'(X_t)|^2] \neq 0} \frac{\left(\mathbb{E}\Big[\bar{h}(t, X_t)\phi'(X_t)\Big]\right)^2}{\mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))|\phi'(X_t)|^2\Big]} dt$$

For any $\phi \in C_c^\infty(\mathbb{R})$ and $t \in [0, T]$ such that $\mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))|\phi'(X_t)|^2\Big] \neq 0$, we have

$$\begin{aligned}\frac{\left(\mathbb{E}\Big[\bar{h}(t, X_t)\phi'(X_t)\Big]\right)^2}{\mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))|\phi'(X_t)|^2\Big]} &= \frac{\left(\mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))^{-1/2}\bar{h}(t, X_t)\bar{D}(X_t, \mathcal{L}(X_t))^{1/2}\phi'(X_t)\Big]\right)^2}{\mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))|\phi'(X_t)|^2\Big]} \\ &\le \mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))^{-1}|\bar{h}(t, X_t)|^2\Big]\end{aligned}$$

so

$$I^{DG}(Z) \le \frac{1}{4}\int_0^T \mathbb{E}\Big[\bar{D}(X_t, \mathcal{L}(X_t))^{-1}|\bar{h}(t, X_t)|^2\Big]dt,$$

and by Eq. (46) we are done. □

As a corollary to the above result, we also get an alternative form of the rate function in the setting without multiscale structure. This provides us with rate functions with which it is more feasible to compare the likelihood of rare events for the fluctuation process (3) as $N \to \infty$ in the multiscale and non-multiscale setting as opposed to the variational form given in Theorem 3.2 and Corollary 3.4. This analysis is outside the scope of this paper, but is an interesting avenue for future research.

Corollary 4.6 *In the setting of Corollary 3.4, assume in addition $\sigma^2(x,\mu) > 0$, for all $x \in \mathbb{R}$ and $\mu \in \mathcal{P}_2(\mathbb{R})$. Consider $\tilde{I}^{DG} : C([0,T];\mathcal{S}_{-\rho}) \to [0,+\infty]$ given by:*

$$\tilde{I}^{DG}(Z) := \frac{1}{2}\int_0^T \sup_{\phi\in C_c^\infty(\mathbb{R}):\mathbb{E}[\sigma^2(X_t,\mathcal{L}(X_t))|\phi'(\tilde{X}_t)|^2]\neq 0} \frac{|\langle \dot{Z}_t - \tilde{L}^*_{\mathcal{L}(\tilde{X}_t)} Z_t, \phi\rangle|^2}{\mathbb{E}[\sigma^2(X_t,\mathcal{L}(X_t))|\phi'(\tilde{X}_t)|^2]} dt, \tag{48}$$

if $Z(0) = 0$, Z is absolutely continuous in the sense if Definition 4.1, and $Z \in C([0,T];\mathcal{S}_{-v})$, and $I^{DG}(Z) = +\infty$ otherwise. Here $\tilde{X}_t$ is as in Corollary 3.4, $\dot{Z}$ is the time derivative of Z in the distribution sense from Lemma 4.2 and $\tilde{L}^*_{\mathcal{L}(\tilde{X}_s)} : \mathcal{S}_{-v} \to \mathcal{S}_{-(v+2)}$ is the adjoint of $\tilde{L}_{\mathcal{L}(\tilde{X}_s)} : \mathcal{S}_{v+2} \to \mathcal{S}_v$ given in Corollary 3.4 (using here Lemma 8.6).

Then $\{Z^N\}_{N\in\mathbb{N}}$ satisfies a large deviation principle on the space $C([0,T];\mathcal{S}_{-\rho})$ with speed $a^{-2}(N)$ and good rate function $\tilde{I}^{DG}$.

Proof This follows by the same proof as Proposition 4.3, removing the dependence of the control on y and setting $\Phi \equiv 0$.

□

4.2 Examples: a class of aggregation-diffusion equations

A common form for interacting particle systems which are widely used in many settings such as in biology, ecology, social sciences, economics, molecular dynamics, and in study of spatially homogeneous granular media (see e.g., [2, 34, 47, 54]) is:

$$dX_t^{i,N} = -V'(X_t^{i,N})dt - \frac{1}{N}\sum_{j=1}^N W'(X_t^{i,N} - X_t^{j,N})dt + \sigma dW_t^i, \quad X_0^{i,N} = \eta^x \tag{49}$$

where $V : \mathbb{R} \to \mathbb{R}$ is a sufficiently smooth confining potential and $W : \mathbb{R} \to \mathbb{R}$ is a sufficiently smooth interaction potential. The class of systems (49) contains the system in the seminal paper [16], where many mathematical aspects of a model for cooperative behavior in a bi-stable confining potential with attraction to the mean are explored. This leads us to our first example:

Example 4.7 Consider the system (49). Let v, ρ be as in Corollary 3.4. Suppose $V', W' \in C_b^{\rho+2}$, $W' \in \mathcal{S}_{\rho+2}$, and $\sigma > 0$. Then $\{Z^N\}_{N\in\mathbb{N}} = \{a(N)\sqrt{N}[\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}_\cdot} -$

$\mathcal{L}(\tilde{X}.)]\}_{N\in\mathbb{N}}$ satisfies a large deviation principle on the space $C([0,T];\mathcal{S}_{-\rho})$ with speed $a^{-2}(N)$ and good rate function $\tilde{I}^{DG}$ given by:

$$\tilde{I}^{DG}(Z) := \frac{1}{2\sigma^2}\int_0^T \sup_{\phi\in C_c^\infty(\mathbb{R}):\mathbb{E}[|\phi'(\tilde{X}_t)|^2]\neq 0} \frac{|\langle \dot{Z}_t - \tilde{L}^*_{\mathcal{L}(\tilde{X}_t)}Z_t,\phi\rangle|^2}{\mathbb{E}[|\phi'(\tilde{X}_t)|^2]}dt,$$

if $Z(0)=0$, Z is absolutely continuous in the sense if Definition 4.1, and $Z\in C([0,T];\mathcal{S}_{-v})$, and $I^{DG}(Z)=+\infty$ otherwise.

Here $\tilde{X}_t$ satisfies:

$$d\tilde{X}_t = -V'(\tilde{X}_t)dt - \bar{\mathbb{E}}[W'(x-\bar{X}_t)]|_{x=\tilde{X}_t}dt + \sigma dW_t, \quad \tilde{X}_0 = \eta^x$$

and $\tilde{L}_{\mathcal{L}(\tilde{X}_s)} : \mathcal{S}_{v+2}\to\mathcal{S}_v$ acts on $\phi\in C_c^\infty(\mathbb{R})$ by:

$$\tilde{L}_{\mathcal{L}(\tilde{X}_s)}\phi(x) = -[V'(x)+\mathbb{E}[W'(x-\tilde{X}_s)]]\phi'(x) + \frac{\sigma^2}{2}\phi''(x) - \mathbb{E}[W'(\tilde{X}_s - x)\phi'(\tilde{X}_s)].$$

We are denoting by $\bar{X}_t$ an independent copy of $\tilde{X}_t$ on another probability space $(\bar{\Omega},\bar{\mathcal{F}},\bar{\mathbb{P}})$, and by $\bar{\mathbb{E}}$ the expectation on that space.

Proof Noting that σ is constant and $\frac{\delta}{\delta m}c(x,\mu)[z] = -W'(x-z)$, the assumptions put forward in Corollary 3.4 can be directly verified. This example then immediately falls into the regime of Corollary 4.6. □

In [38], the authors make, among other modifications, a modification to V in Eq. (49) so that it is a so-called rough-potential (see also [74] and [4] Section 5), by letting $V^\epsilon(x) = V_1(x) + V_2(x/\epsilon)$, where V_2 is sufficiently smooth and periodic. The system becomes:

$$dX^{i,\epsilon,N}_t = -[V_1'(X^{i,\epsilon,N}_t) + \frac{1}{\epsilon}V_2'(X^{i,\epsilon,N}_t/\epsilon)]dt - \frac{1}{N}\sum_{j=1}^N W'(X^{i,\epsilon,N}_t - X^{j,\epsilon,N}_t)dt + \sigma dW^i_t.$$

Letting $Y^{i,\epsilon,N}_t = X^{i,\epsilon,N}_t/\epsilon$, we see this is a subclass of systems of the form (1) with

$$f(x,y,\mu) = b(x,y,\mu) = -V_2'(y), \quad g(x,y,\mu) = c(x,y,\mu) = -V_1'(x) - \langle \mu, W'(x-\cdot)\rangle$$
$$\sigma(x,y,\mu) = \tau_1(x,y,\mu) \equiv \sigma, \quad \tau_2 \equiv 0,$$

Keeping within our setting of a slow–fast system on $\mathbb{R}$, we consider a version of this system where the fast and slow dynamics are allowed to be different, and the fast system is not confined to the torus:

$$dX^{i,\epsilon,N}_t = -[V_1'(X^{i,\epsilon,N}_t) + \frac{1}{\epsilon}V_2'(Y^{i,\epsilon,N}_t)]dt - \frac{1}{N}\sum_{j=1}^N W_1'(X^{i,\epsilon,N}_t - X^{j,\epsilon,N}_t)dt + \sigma dW^i_t$$

$$
\begin{aligned}
dY_t^{i,\epsilon,N} &= -\frac{1}{\epsilon}[V_3'(X_t^{i,\epsilon,N}) + \frac{1}{\epsilon}V_4'(Y_t^{i,\epsilon,N})]dt - \frac{1}{\epsilon}\frac{1}{N}\sum_{j=1}^N W_2'(X_t^{i,\epsilon,N} - X_t^{j,\epsilon,N})dt \\
&\quad + \frac{1}{\epsilon}\tau_1 dW_t^i + \frac{1}{\epsilon}\tau_2 dB_t^i \\
(X_0^{i,\epsilon,N}, Y_0^{i,\epsilon,N}) &= (\eta^x, \eta^y).
\end{aligned} \tag{50}
$$

This falls into the class of systems (1) with

$$
\begin{aligned}
&b(x,y,\mu) = -V_2'(y), \quad c(x,y,\mu) = -V_1'(x) - \langle \mu, W_1'(x-\cdot)\rangle, \quad \sigma(x,y,\mu) \equiv \sigma \\
&f(x,y,\mu) = -V_4'(y), \quad g(x,y,\mu) = -V_3'(x) - \langle \mu, W_2'(x-\cdot)\rangle, \quad \tau_1(x,y,\mu) \equiv \tau_1, \\
&\qquad \tau_2(x,y,\mu) \equiv \tau_2.
\end{aligned}
$$

Example 4.8 Consider the system (50).

Suppose $V_4(y) = \frac{\kappa}{2}y^2 + \tilde{\eta}(y)$ where $\kappa > 0$ and $\tilde{\eta} \in C_b^2(\mathbb{R})$ is even with $\left\|\tilde{\eta}''\right\|_\infty < \kappa$, $V_1', V_3', W_1', W_2' \in C_b^{r+2}(\mathbb{R})$, $W_1', W_2' \in \mathcal{S}_{r+2}$ where r is as in Eq. (7), $\sigma, \tau_2 \neq 0$, V_2 is even, and V_2' is Lipschitz continuous and $O(|y|^{1/2})$ as $|y| \to \infty$.

Then $\{Z^N\}_{N\in\mathbb{N}} = \{a(N)\sqrt{N}[\frac{1}{N}\sum_{i=1}^N \delta_{X_\cdot^{i,\epsilon,N}} - \mathcal{L}(X_\cdot)]\}_{N\in\mathbb{N}}$ satisfies a large deviation principle on the space $C([0,T];\mathcal{S}_{-r})$ with speed $a^{-2}(N)$ and good rate function I^{DG} given by:

$$
I^{DG}(Z) = \frac{1}{2[\sigma^2 + 2\alpha a + 2\sigma\tau_1\tilde{\alpha}]}\int_0^T \sup_{\phi\in C_c^\infty(\mathbb{R}):\mathbb{E}[|\phi'(X_t)|^2]\neq 0} \frac{\left|\langle \dot{Z}_t - \bar{L}^*_{\mathcal{L}(X_t)}Z_t, \phi\rangle\right|^2}{\mathbb{E}\left[|\phi'(X_t)|^2\right]}dt
$$

if $Z(0) = 0$, Z is absolutely continuous in the sense if Definition 4.1, and $Z \in C([0,T];\mathcal{S}_{-w})$, and $I^{DG}(Z) = +\infty$ otherwise. Here X_t satisfies:

$$
\begin{aligned}
dX_t &= -[\tilde{\alpha}V_3'(X_t) + V_1'(X_t)]dt - \bar{\mathbb{E}}[\tilde{\alpha}W_2'(x - \bar{X}_t) \\
&\qquad + W_1'(x - \bar{X}_t)]|_{x=X_t}dt + [\sigma^2 + 2\alpha a + 2\sigma\tau_1\tilde{\alpha}]^{1/2}dW_t \\
X_0 &= \eta^x \\
\tilde{\alpha} &= \int_{\mathbb{R}}\Phi'(y)\pi(dy), \quad \alpha = \int_{\mathbb{R}}[\Phi'(y)]^2\pi(dy), \quad a = \frac{1}{2}[\tau_1^2 + \tau_2^2]
\end{aligned}
$$

and $\bar{L}_{\mathcal{L}(X_s)} : \mathcal{S}_{w+2} \to \mathcal{S}_w$ acts on $\phi \in C_c^\infty(\mathbb{R})$ by:

$$
\begin{aligned}
\bar{L}_{\mathcal{L}(X_s)}\phi(x) := &-[\tilde{\alpha}V_3'(x) + V_1'(x) + \mathbb{E}[\tilde{\alpha}W_2'(x - X_s) + W_1'(x - X_s)]]\phi'(x) \\
&+ \frac{1}{2}[\sigma^2 + 2\alpha a + 2\sigma\tau_1\tilde{\alpha}]\phi''(x) \\
&- \mathbb{E}[[\tilde{\alpha}W_2'(X_s - x) + W_1'(X_s - x)]\phi'(X_s)].
\end{aligned}
$$

Again, we are denoting by $\bar{X}_t$ an independent copy of X_t on another probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ and by $\bar{\mathbb{E}}$ the expectation on that space.

Proof Once we show $\int_{\mathbb{R}} V_2'(y)\pi(dy) = 0$ for π as in Eq. (20), it follows that assumptions (A1)–(A12) and (A'13) hold via Example C.6 in the appendix. Via Remark 4.4, we also have $\bar{D} > 0, \forall x \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R})$. Then this example is an immediate corollary of Proposition 4.3.

We know in this setting that π admits a density of the form $\pi(y) = C \exp\left(\frac{-V_4(y)}{a}\right)$, where C is a normalizing constant (see Eq. 72 in the appendix). Then since V_2, V_4 are assumed even and hence $V_2'\pi$ is odd, the result holds. □

5 Overview of the approach and formulation of the controlled system

We use the weak convergence approach of [22] in order to prove the large deviations principle for Z^N. As discussed in Sect. 3, we prove the large deviations principle via proving Z^N satisfies the Laplace principle with speed $a^{-2}(N)$ and good rate function I given by Eq. (33) (see, e.g. [22] Section 1.2).

The method for this is to use the variational representation from [8] to get that for each $N \in \mathbb{N}$ and $F \in C_b(C([0,T]; S_{-\tau}))$, $\tau \geq w$, where w is as in Eq. (6),

$$
\begin{aligned}
&- a^2(N) \log \mathbb{E} \exp\left(-\frac{1}{a^2(N)} F(Z^N)\right) \\
&= \inf_{\tilde{u}^N} \mathbb{E}\left[\frac{1}{2}\frac{1}{N}\sum_{i=1}^N \int_0^T \left(|\tilde{u}_i^{N,1}(s)|^2 + |\tilde{u}_i^{N,2}(s)|^2\right) ds + F(\tilde{Z}^N)\right] \qquad (51)
\end{aligned}
$$

where $\{\tilde{u}_i^{N,k}\}_{i\in\mathbb{N},k=1,2}$ are $\{\mathcal{F}_t\}$-progressively-measurable processes such that

$$
\sup_{N\in\mathbb{N}} \frac{1}{N}\mathbb{E}\left[\sum_{i=1}^N \int_0^T \left(|\tilde{u}_i^{N,1}(s)|^2 + |\tilde{u}_i^{N,2}(s)|^2\right) ds\right] < \infty. \qquad (52)
$$

One can see that in fact the results of [8] indeed imply the equality (51) by following an argument along the same lines as Proposition 3.3. in [4].

This bound on the controls can be improved when proving the Laplace principle Lower Bound (30) to:

$$
\sup_{N\in\mathbb{N}} \frac{1}{N}\sum_{i=1}^N \int_0^T \left(|\tilde{u}_i^{N,1}(s)|^2 + |\tilde{u}_i^{N,2}(s)|^2\right) ds < \infty, \mathbb{P}- \text{ almost surely.} \qquad (53)
$$

by the argument found in Theorem 4.4 of [8]. Here $\tilde{Z}^N$ is given by, for $\phi \in C_c^\infty(\mathbb{R})$:

$$
\langle \tilde{Z}_t^N, \phi\rangle = a(N)\sqrt{N}(\langle \tilde{\mu}_t^{\epsilon,N}, \phi\rangle - \langle \mathcal{L}(X_t), \phi\rangle), \quad \text{with}
$$

$$\tilde{\mu}_t^{\epsilon,N} = \frac{1}{N}\sum_{i=1}^{N}\delta_{\tilde{X}_t^{i,\epsilon,N}}, \quad t \in [0,T]. \tag{54}$$

$\tilde{X}_t^{i,\epsilon,N}$ are solutions to:

$$\begin{aligned}
d\tilde{X}_t^{i,\epsilon,N} &= \left[\frac{1}{\epsilon}b(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N}) + c(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})\right.\\
&\quad \left.+\sigma(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})\frac{\tilde{u}_i^{N,1}(t)}{a(N)\sqrt{N}}\right]dt\\
&\quad +\sigma(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})dW_t^i\\
d\tilde{Y}_t^{i,\epsilon,N} &= \frac{1}{\epsilon}\left[\frac{1}{\epsilon}f(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N}) + g(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})\right.\\
&\quad +\tau_1(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})\frac{\tilde{u}_i^{N,1}(t)}{a(N)\sqrt{N}}\\
&\quad \left.+\tau_2(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})\frac{\tilde{u}_i^{N,2}(t)}{a(N)\sqrt{N}}\right]dt\\
&\quad +\frac{1}{\epsilon}\left[\tau_1(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})dW_t^i + \tau_2(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})dB_t^i\right]\\
(\tilde{X}_0^{i,\epsilon,N},\tilde{Y}_0^{i,\epsilon,N}) &= (\eta^x,\eta^y).
\end{aligned} \tag{55}$$

We couple the controls to the joint empirical measures of the fast and slow process by defining occupation measures $\{Q^N\}_{N\in\mathbb{N}} \subset M_T(\mathbb{R}^4)$ in the following way: for $A, B \in \mathcal{B}(\mathbb{R})$ and $C \in \mathcal{B}(\mathbb{R}^2)$:

$$Q^N(A\times B\times C\times[0,t]) = \frac{1}{N}\sum_{i=1}^{N}\int_0^t \delta_{\tilde{X}_s^{i,\epsilon,N}}(A)\delta_{\tilde{Y}_s^{i,\epsilon,N}}(B)\delta_{(\tilde{u}_i^{N,1}(s),\tilde{u}_i^{N,2}(s))}(C)ds. \tag{56}$$

The proof of the Inequalities (30) and (31) are attained by identifying limit in distribution of $(\tilde{Z}^N, Q^N)$ as satisfying the limiting controlled Eq. (32). This identification of the limit is the subject of Sect. 8. In order to identify this limit, we first need to establish tightness of the sequence of random variables $\{(\tilde{Z}^N, Q^N)\}_{N\in\mathbb{N}}$, as done in Sect. 7. The proof of tightness relies on a combination of Ergodic-Type Theorems for the system of controlled interacting particles (55) as proved in Sect. 6 and on establishing rates of averaging for fully coupled McKean–Vlasov equations, as done in Sect. 7.1. These rates of averaging are needed do to a novel coupling argument made in the proof of tightness (see Lemma 7.6) to the following system of IID slow–fast McKean–Vlasov Equations:

$$d\bar{X}_t^{i,\epsilon} = \left[\frac{1}{\epsilon}b(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon})) + c(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\right]dt$$

$$
\begin{aligned}
&\quad + \sigma(\bar{X}_t^{i,\epsilon}, \bar{Y}_t^{i,\epsilon}, \mathcal{L}(\bar{X}_t^{\epsilon}))dW_t^i \\
d\bar{Y}_t^{i,\epsilon} &= \frac{1}{\epsilon}\left[\frac{1}{\epsilon} f(\bar{X}_t^{i,\epsilon}, \bar{Y}_t^{i,\epsilon}, \mathcal{L}(\bar{X}_t^{\epsilon})) + g(\bar{X}_t^{i,\epsilon}, \bar{Y}_t^{i,\epsilon}, \mathcal{L}(\bar{X}_t^{\epsilon}))\right]dt \\
&\quad + \frac{1}{\epsilon}\left[\tau_1(\bar{X}_t^{i,\epsilon}, \bar{Y}_t^{i,\epsilon}, \mathcal{L}(\bar{X}_t^{\epsilon}))dW_t^i + \tau_2(\bar{X}_t^{i,\epsilon}, \bar{Y}_t^{i,\epsilon}, \mathcal{L}(\bar{X}_t^{\epsilon}))dB_t^i\right] \\
(\bar{X}_0^{i,\epsilon}, \bar{X}_0^{i,\epsilon}) &= (\eta^x, \eta^y), \qquad (57)
\end{aligned}
$$

where $\bar{X}^\epsilon$ is any particle that has common law with the $\bar{X}^{i,\epsilon}$'s and W^i, B^i are the same driving Brownian motions as in Eqs. (1) and (55).

We will also make use of the empirical measure on N of the IID slow particles from Eq. (57):

$$
\bar{\mu}^{\epsilon,N} := \frac{1}{N}\sum_{i=1}^{N} \delta_{\bar{X}_t^{i,\epsilon}}. \qquad (58)
$$

Remark 5.1 Note that these IID particles are what we get from replacing $\mu^{\epsilon,N}$ by $\mathcal{L}(\bar{X}^\epsilon)$ in Eq. (1). Using such an auxiliary process is a traditional proof method for tightness of fluctuation processes related to empirical measures; See [39] Theorem 1/Lemma 1, [69] Section 8, [19] Section 5.1, [49] Theorem 2.4/3.1, [27] Lemma 3.2/Proposition 3.5/Section 4 for examples of this general approach. However, a key difference here form those proofs is that the IID particles are not copies of the limiting process (25), but instead are copies of the process we would obtain from keeping $\epsilon > 0$ fixed and sending $N \to \infty$. As seen in [4], the limit in distribution as $N \to \infty, \epsilon \downarrow 0$ of the empirical measure $\mu^{\epsilon,N}$ does not depend on the relative rates at which ϵ and N go to their respective limits. Hence, we are able to treat each of the problems separately, and obtain a rate of convergence of $\tilde{\mu}^{\epsilon,N}$ from Eq. (54) to $\bar{\mu}^{\epsilon,N}$ from Eq. (58) as $N \to \infty$ in L^2 (see Lemma 7.5), and a rate of convergence of $\mathcal{L}(\bar{X}_t^{1,\epsilon})$ from Eq. (57) to $\mathcal{L}(X_t)$ uniformly as an element of S_{-m}, where X_t is as in Eq. (25) and m is as in Eq. (5). The latter is a problem of independent interest in itself, and extends the current known results on averaging for SDEs and McKean–Vlasov SDEs, which can be found in, e.g. [64] and [65] respectively. The result is contained in Sect. 7.1 as Theorem 7.2, and its proof is the subject of the complimentary paper [5].

6 Ergodic-type theorems for the controlled system (55)

In this section, we use the method of auxiliary Poisson equations to derive rates of averaging in the form of Ergodic-Type Theorems for the controlled particles (55). These results are used in the proof of tightness of the controlled fluctuation process, as they allow us to couple the controlled particles (55) to the IID slow–fast McKean–Vlasov Eq. (57)—see Lemma 7.5. They also allow us to identify a prelimit representation for the controlled fluctuations processes $\tilde{Z^N}$ from Eq. (54) (see Lemma 7.7), which informs the controlled limit proved in Sect. 8. In particular, Proposition 6.1 is necessary to handle the terms $\frac{1}{\epsilon}b$ appearing in the drift of the slow particles $X^{i,N,\epsilon}$, $\tilde{X}^{i,N,\epsilon}$ in Eqs. (1), (55). This is where the terms involving the solution Φ to the Poisson Eq. (22) in the limiting coefficients (23) come from. The same analysis is performed in averaging fully-coupled standard diffusions—see e.g. [61] Theorem 4 and [64] Lemma 4.4—but here we must also account for the dependence of the coefficients on the empirical measure, and hence derivatives of Φ in its measure component appear in the remainder terms. One term involving the derivative in the measure component of Φ a priori seems to be $\mathcal{O}(1)$ in the limit, but is seen to vanish as $N \to \infty$ in Proposition 6.3. Naturally such a term does not appear in the setting without measure dependence of the coefficients, and is unique to slow–fast interacting particle systems and slow–fast McKean–Vlasov SDEs. Thus the "doubled Poisson equation" construction (see Eq. 63) and the proof of Proposition 6.3 are novel to this paper and the related paper [5]. Proposition 6.4 is used to see that drift and diffusion coefficients which depend on the fast particles $\tilde{Y}^{i,\epsilon,N}$ from Eq. (55) can be exchanged for those where dependence on $\tilde{Y}^{i,\epsilon,N}$ is replaced with integration against the invariant measure π from Eq. (20) at a cost of $\mathcal{O}(\epsilon)$. This method is employed when establishing rates of stochastic homogenization in the standard (one-particle) setting in e.g. [67] Lemma 4.1, [56] Lemma B.5, and [64] Lemma 4.2. There again, our setting is different than the standard case in that we must compensate for the dependence of the empirical measure of the coefficients, which yields terms involving the derivative in the measure component of the auxiliary Poisson Eq. (64).

Proposition 6.1 *Consider $\psi \in C_b^{1,2}([0,T]\times\mathbb{R})$. Under assumptions (A1)–(A6), we have for any $t \in [0,T]$:*

$$
\begin{aligned}
&\frac{a(N)}{\sqrt{N}}\sum_{i=1}^{N}\mathbb{E}\Bigg[\sup_{t\in[0,T]}\Bigg|\int_0^t \frac{1}{\epsilon}b(i)\psi(s,\tilde{X}_s^{i,\epsilon,N})ds\\
&-\int_0^t\Bigg(\gamma_1(i)\psi(s,\tilde{X}_s^{i,\epsilon,N})+D_1(i)\psi_x(s,\tilde{X}_s^{i,\epsilon,N})\\
&+[\frac{\tau_1(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s)+\frac{\tau_2(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)]\Phi_y(i)\psi(s,\tilde{X}_s^{i,\epsilon,N})\Bigg)ds\\
&-\int_0^t \tau_1(i)\Phi_y(i)\psi(s,\tilde{X}_s^{i,\epsilon,N})dW_s^i\\
&-\int_0^t \tau_2(i)\Phi_y(i)\psi(s,\tilde{X}_s^{i,\epsilon,N})dB_s^i-\int_0^t\frac{1}{N}\sum_{j=1}^{N}b(j)\partial_\mu\Phi(i)[j]\psi(s,\tilde{X}_s^{i,\epsilon,N})ds\Bigg|^2\Bigg]\\
&\le C[\epsilon^2 a(N)\sqrt{N}(1+T+T^2)+\frac{a(N)}{\sqrt{N}}T^2]\,\|\psi\|^2_{C_b^{1,2}}
\end{aligned}
$$

where here (i) denotes the argument $(\tilde{X}_s^{i,\epsilon,N},\tilde{Y}_s^{i,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})$ and similarly for (j), $[j]$ denotes the argument $\tilde{X}_s^{j,\epsilon,N}$, and Φ is as in (22). Here we recall the definitions of γ_1, D_1 from Eq. (23).

Proof Using Lemma C.1 to gain appropriate differentiablity of Φ, letting $\Phi^N : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ be the empirical projection of Φ and applying standard Itô's formula and Proposition D.6 to the composition $\Phi^N(\tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, (\tilde{X}_s^{1,\epsilon,N}, \ldots, \tilde{X}_s^{N,\epsilon,N}))$, we get:

$$\begin{aligned}
&\int_0^t \frac{1}{\epsilon} b(i)\psi(s, \tilde{X}_s^{i,\epsilon,N})ds - \int_0^t \Bigg(\gamma_1(i)\psi(s, \tilde{X}_s^{i,\epsilon,N}) + D_1(i)\psi_x(s, \tilde{X}_s^{i,\epsilon,N}) \\
&\quad + [\frac{\tau_1(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s) + \frac{\tau_2(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)]\Phi_y(i)\psi(s, \tilde{X}_s^{i,\epsilon,N})\Bigg)ds \\
&\quad - \int_0^t \tau_1(i)\Phi_y(i)\psi(s, \tilde{X}_s^{i,\epsilon,N})dW_s^i \\
&\quad - \int_0^t \tau_2(i)\Phi_y(i)\psi(s, \tilde{X}_s^{i,\epsilon,N})dB_s^i - \int_0^t \frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Phi(i)[j]\psi(s, \tilde{X}_s^{i,\epsilon,N})ds \\
&\quad = \sum_{k=1}^8 \tilde{B}_k^{i,\epsilon,N}
\end{aligned}$$

where:

$$\begin{aligned}
\tilde{B}_1^{i,\epsilon,N}(t) &= \epsilon[\Phi(\tilde{X}_0^{i,\epsilon,N}, \tilde{Y}_0^{i,\epsilon,N}, \tilde{\mu}_0^{\epsilon,N})\psi(0, \tilde{X}_0^{i,\epsilon,N}) \\
&\qquad - \Phi(\tilde{X}_t^{i,\epsilon,N}, \tilde{Y}_t^{i,\epsilon,N}, \tilde{\mu}_t^{\epsilon,N})\psi(t, \tilde{X}_t^{i,\epsilon,N})] \\
\tilde{B}_2^{i,\epsilon,N}(t) &= \frac{1}{N}\int_0^t \sigma(i)\tau_1(i)\partial_\mu\Phi_y(i)[i]\psi(s, \tilde{X}_s^{i,\epsilon,N})ds \\
\tilde{B}_3^{i,\epsilon,N}(t) &= \epsilon\int_0^t \Bigg(\Phi(i)\dot{\psi}(s, \tilde{X}_s^{i,\epsilon,N}) + c(i)[\Phi_x(i)\psi(s, \tilde{X}_s^{i,\epsilon,N}) + \Phi(i)\psi_x(s, \tilde{X}_s^{i,\epsilon,N})] \\
&\qquad + \frac{\sigma^2(i)}{2}[\Phi_{xx}(i)\psi(s, \tilde{X}_s^{i,\epsilon,N}) \\
&\qquad + 2\Phi_x(i)\psi_x(s, \tilde{X}_s^{i,\epsilon,N}) + \Phi(i)\psi_{xx}(s, \tilde{X}_s^{i,\epsilon,N})]\Bigg)ds \\
\tilde{B}_4^{i,\epsilon,N}(t) &= \epsilon\int_0^t \frac{\sigma^2(i)}{2}[\frac{2}{N}\partial_\mu\Phi(i)[i]\psi_x(s, \tilde{X}_s^{i,\epsilon,N}) + \frac{2}{N}\partial_\mu\Phi_x(i)[i]\psi(s, \tilde{X}_s^{i,\epsilon,N})]ds \\
\tilde{B}_5^{i,\epsilon,N}(t) &= \epsilon\int_0^t \frac{1}{N}\sum_{j=1}^N \Bigg\{c(j)\partial_\mu\Phi(i)[j]\psi(s, \tilde{X}_s^{i,\epsilon,N}) + \frac{1}{2}\sigma^2(j)[\frac{1}{N}\partial_\mu^2\Phi(i)[j,j] \\
&\qquad + \partial_z\partial_\mu\Phi(i)[j]]\psi(s, \tilde{X}_s^{i,\epsilon,N})\Bigg\}ds \\
\tilde{B}_6^{i,\epsilon,N}(t) &= \epsilon\Bigg[\int_0^t \sigma(i)[\Phi_x(i)\psi(s, \tilde{X}_s^{i,\epsilon,N}) + \Phi(i)\psi_x(s, \tilde{X}_s^{i,\epsilon,N})]dW_t^i \\
&\qquad + \frac{1}{N}\sum_{j=1}^N \Bigg\{\int_0^t \sigma(j)\partial_\mu\Phi(i)[j]\psi(s, \tilde{X}_s^{i,\epsilon,N})dW_s^j\Bigg\}\Bigg] \\
\tilde{B}_7^{i,\epsilon,N}(t) &= \epsilon\int_0^t \frac{\sigma(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s)[\Phi_x(i)\psi(s, \tilde{X}_s^{i,\epsilon,N}) + \Phi(i)\psi_x(s, \tilde{X}_s^{i,\epsilon,N})]ds
\end{aligned}$$

$$\tilde{B}_8^{i,\epsilon,N}(t) = \epsilon \int_0^t \frac{1}{N}\left\{\sum_{j=1}^N \frac{\sigma(j)}{a(N)\sqrt{N}} \tilde{u}_j^{N,1}(s)\partial_\mu \Phi(i)[j]\psi(s, \tilde{X}_s^{i,\epsilon,N})\right\}ds.$$

Via Lemma B.1, the assumed linear growth of b and c in y and boundedness of σ, and the assumed bound (53) on the controls, one can check that indeed $\tilde{\mu}_t^N \in \mathcal{P}_2(\mathbb{R})$ for each $t \in [0, T]$ and $N \in \mathbb{N}$, and so there is no issue with the domain of Φ and its derivatives being $\mathcal{P}_2(\mathbb{R})$.

Then, by multiple applications of Hölder's inequality, and using the assumed uniform in x, μ polynomial growth in y of Φ and its derivatives from Assumption (A6):

$$\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{B}_1^{i,\epsilon,N}(t)|^2\right] \le \epsilon^2 a(N)\sqrt{N}\,\|\psi\|_\infty^2$$

$$\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{B}_2^{i,\epsilon,N}(t)|^2\right] \le C\frac{a(N)}{\sqrt{N}}\frac{1}{N^2}\sum_{i=1}^N T\mathbb{E}\left[\int_0^T |\partial_\mu \Phi_y(i)[i]|^2 ds\right]\|\psi\|_\infty^2$$

$$\le C\frac{a(N)}{\sqrt{N}}\frac{1}{N}\sum_{i=1}^N T\mathbb{E}\left[\int_0^T \left\|\partial_\mu \Phi_y(i)[\cdot]\right\|^2_{L^2(\mathbb{R},\tilde{\mu}_s^{N,\epsilon})} ds\right]\|\psi\|_\infty^2$$

$$\le C\frac{a(N)}{\sqrt{N}}T^2\left(1+\frac{1}{N}\sum_{i=1}^N \sup_{t\in[0,T]}\mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N}|^{2\tilde{q}_{\Phi_y}(1,0,0)}\right]\right)\|\psi\|_\infty^2$$

$$\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{B}_3^{i,\epsilon,N}(t)|^2\right]$$

$$\le \epsilon^2 a(N)\sqrt{N}T^2\left(1+\frac{1}{N}\sum_{i=1}^N \sup_{t\in[0,T]}\mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N}|^2 + |\tilde{Y}_t^{i,\epsilon,N}|^{2q_\Phi(0,2,0)}\right]\right)\|\psi\|^2_{C_b^{1,2}}$$

$$\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{B}_4^{i,\epsilon,N}(t)|^2\right]$$

$$\le C\frac{a(N)}{\sqrt{N}}\frac{\epsilon^2}{N^2}\sum_{i=1}^N T\mathbb{E}\left[\int_0^T \left|\left(|\partial_\mu \Phi(i)[i]| + |\partial_\mu \Phi_x(i)[i]|\right)\right|^2 ds\right](\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2)$$

$$\le C\frac{a(N)}{\sqrt{N}}\frac{\epsilon^2}{N}\sum_{i=1}^N T\mathbb{E}\left[\int_0^T \left|\left(\left\|\partial_\mu \Phi(i)[\cdot]\right\|_{L^2(\mathbb{R},\tilde{\mu}_s^{N,\epsilon})}\right.\right.\right.$$

$$\left.\left.\left. + \left\|\partial_\mu \Phi_x(i)[\cdot]\right\|_{L^2(\mathbb{R},\tilde{\mu}_s^{N,\epsilon})}\right)\right|^2 ds\right](\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2)$$

$$\le C\frac{a(N)}{\sqrt{N}}\epsilon^2 T^2\left(1+\frac{1}{N}\sum_{i=1}^N \sup_{t\in[0,T]}\mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N}|^{2(\tilde{q}_\Phi(1,0,0)\vee\tilde{q}_\Phi(1,1,0))}\right]\right)(\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2)$$

Here for $\tilde{B}_1$, we used the assumed boundedness of Φ from (A6). For $\tilde{B}_2$ we used the assumed polynomial growth in y of $\partial_\mu \Phi$ from (A6) and the boundedness of σ and τ_1 from (A5) and (A1). For $\tilde{B}_3$ we used the assumed polynomial growth in y of Φ_{xx} and boundedness of Φ, Φ_x from (A6) and the boundedness of σ and the linear growth in y of c from (A5). In $\tilde{B}_4$ we used the assumed polynomial growth in y of $\partial_\mu \Phi$ and $\partial_\mu \Phi_x$ from (A6) and the assumed boundedness of σ from (A5).

For $\tilde{B}_5^{i,\epsilon,N}(t)$, we bound the two terms separately. For the first, we use the assumed linear growth in y of c and polynomial growth of $\partial_\mu \Phi$ in y to get:

$$
\begin{aligned}
&\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N \frac{\epsilon^2}{N^2}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sum_{j=1}^N c(j)\partial_\mu\Phi(i)[j]\psi(s,\tilde{X}_s^{i,\epsilon,N})ds\right|^2\right]\\
&\leq \epsilon^2 a(N)\sqrt{N}\frac{T}{N}\sum_{i=1}^N \mathbb{E}\left[\int_0^T \left\|\partial_\mu\Phi(i)[\cdot]\right\|^2_{L^2(\mathbb{R},\tilde{\mu}_s^{\epsilon,N})}\frac{1}{N}\sum_{j=1}^N |c(j)|^2 ds\right]\|\psi\|_\infty^2\\
&\leq C\epsilon^2 a(N)\sqrt{N}T^2\Bigg(1+\frac{1}{N}\sum_{i=1}^N \sup_{s\in[0,T]}\mathbb{E}\left[\frac{1}{N}\sum_{j=1}^N |\tilde{Y}_s^{i,\epsilon,N}|^2\right]\\
&\quad+\frac{1}{N}\sum_{i=1}^N \sup_{s\in[0,T]}\mathbb{E}\left[|\tilde{Y}_s^{i,\epsilon,N}|^{2\tilde{q}_\Phi(1,0,0)}\right]\\
&\quad+\sup_{s\in[0,T]}\mathbb{E}\left[\frac{1}{N^2}\sum_{j=1}^N\sum_{i=1}^N |\tilde{Y}_s^{i,\epsilon,N}|^{2\tilde{q}_\Phi(1,0,0)}|\tilde{Y}_s^{j,\epsilon,N}|^2\right]\Bigg)\|\psi\|_\infty^2\\
&\leq C\epsilon^2 a(N)\sqrt{N}T^2\Bigg(1+\frac{1}{N}\sum_{i=1}^N \sup_{s\in[0,T]}\mathbb{E}\left[\frac{1}{N}\sum_{j=1}^N |\tilde{Y}_s^{i,\epsilon,N}|^2\right]\\
&\quad+\frac{1}{N}\sum_{i=1}^N \sup_{s\in[0,T]}\mathbb{E}\left[|\tilde{Y}_s^{i,\epsilon,N}|^{2\tilde{q}_\Phi(1,0,0)}\right]\\
&\quad+\sup_{s\in[0,T]}\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^N |\tilde{Y}_s^{i,\epsilon,N}|^{2(\tilde{q}_\Phi(1,0,0)\vee 1)}\right)^2\right]\Bigg)\|\psi\|_\infty^2.
\end{aligned}
$$

For the second, we have by boundedness of σ and the assumed polynomial growth in y of $\partial_\mu^2\Phi$ and $\partial_z\partial_\mu\Phi$:

$$
\begin{aligned}
&\frac{a(N)}{\sqrt{N}}\frac{\epsilon^2}{N^2}\sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sum_{j=1}^N \frac{1}{2}\sigma^2(j)[\frac{1}{N}\partial_\mu^2\Phi(i)[j,j]+\partial_z\partial_\mu\Phi(i)[j]]\psi(s,\tilde{X}_s^{i,\epsilon,N})ds\right|^2\right]\\
&\leq \frac{a(N)}{\sqrt{N}}\epsilon^2 CT\sum_{i=1}^N \mathbb{E}\left[\int_0^T \frac{1}{N}\sum_{j=1}^N \frac{1}{N^2}|\partial_\mu^2\Phi(i)[j,j]|^2+|\partial_z\partial_\mu\Phi(i)[j]|^2 ds\right]\|\psi\|_\infty^2\\
&\leq \frac{a(N)}{\sqrt{N}}\epsilon^2 CT\sum_{i=1}^N \mathbb{E}\Bigg[\int_0^T \frac{1}{N}\left\|\partial_\mu^2\Phi(i)[\cdot,\cdot]\right\|^2_{L^2(\mathbb{R},\tilde{\mu}_s^{\epsilon,N})\otimes L^2(\mathbb{R},\tilde{\mu}_s^{\epsilon,N})}\\
&\quad+\left\|\partial_z\partial_\mu\Phi(i)[\cdot]\right\|^2_{L^2(\mathbb{R},\tilde{\mu}_s^{\epsilon,N})}ds\Bigg]\|\psi\|_\infty^2\\
&\leq C\epsilon^2 a(N)\sqrt{N}T^2\left[1+\frac{1}{N}\sum_{i=1}^N \sup_{s\in[0,T]}\mathbb{E}\left[|\tilde{Y}_s^{i,\epsilon,N}|^{2(\tilde{q}_\Phi(2,0,0)\vee\tilde{q}_\Phi(1,0,1))}\right]\right]\|\psi\|_\infty^2.
\end{aligned}
$$

For the martingale terms, by Burkholder-Davis-Gundy inequality, the assumed boundedness of σ, Φ, and Φ_x and assumed polynomial growth in y of $\partial_\mu \Phi$:

$$
\begin{aligned}
&\frac{a(N)}{\sqrt{N}} \sum_{i=1}^{N} \mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{B}_6^{i,\epsilon,N}(t)|^2\right] \\
&\quad \leq C\epsilon^2 a(N)\sqrt{N} T(\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2) \\
&\qquad + C\frac{a(N)}{\sqrt{N}} \sum_{i=1}^{N} \frac{\epsilon^2}{N^2} \sum_{j=1}^{N} \mathbb{E}\left[\int_0^T |\partial_\mu \Phi(i)[j]|^2 ds\right] \|\psi\|_\infty^2 \\
&\quad = C\epsilon^2 a(N)\sqrt{N} T(\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2) + C\frac{a(N)}{\sqrt{N}} \sum_{i=1}^{N} \frac{\epsilon^2}{N} \\
&\qquad \mathbb{E}\left[\int_0^T \left\|\partial_\mu \Phi(i)[\cdot]\right\|^2_{L^2(\mathbb{R},\tilde{\mu}_s^{\epsilon,N})} ds\right] \|\psi\|_\infty^2 \\
&\quad \leq C\epsilon^2 a(N)\sqrt{N} T(\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2) + C\frac{\epsilon^2 a(N)}{\sqrt{N}} T \\
&\qquad \times \left(1 + \frac{1}{N} \sum_{i=1}^{N} \sup_{t\in[0,T]} \mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N}|^{2\tilde{q}_\Phi(1,0,0)}\right]\right) \|\psi\|_\infty^2 .
\end{aligned}
$$

By the bound (52) and the assumed boundedness of Φ, Φ_x, we have also

$$
\begin{aligned}
&\frac{a(N)}{\sqrt{N}} \sum_{i=1}^{N} \mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{B}_7^{i,\epsilon,N}(t)|^2\right] \\
&\quad \leq \frac{a(N)}{\sqrt{N}} \sum_{i=1}^{N} \frac{\epsilon^2}{a^2(N) N} CT\mathbb{E}\left[\int_0^T |\tilde{u}_i^{N,1}(s)|^2 ds\right] (\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2) \\
&\quad \leq \frac{\epsilon^2}{a(N)\sqrt{N}} CT(\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2).
\end{aligned}
$$

Finally, by the assumed boundedness of σ and polynomial growth of $\partial_\mu \Phi$ in y:

$$
\begin{aligned}
&\frac{a(N)}{\sqrt{N}} \sum_{i=1}^{N} \mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{B}_8^{i,\epsilon,N}(t)|^2\right] \\
&\quad \leq \frac{a(N)}{\sqrt{N}} \sum_{i=1}^{N} \frac{\epsilon^2}{a^2(N) N^3} C\mathbb{E}\left[\left|\sum_{j=1}^{N} \int_0^T |\tilde{u}_j^{N,1}(s)||\partial_\mu \Phi(i)[j]| ds\right|^2\right] \|\psi\|_\infty^2 \\
&\quad \leq \frac{a(N)}{\sqrt{N}} \sum_{i=1}^{N} \frac{\epsilon^2}{a^2(N) N} C\mathbb{E}\left[\left(\frac{1}{N} \sum_{j=1}^{N} \int_0^T |\tilde{u}_j^{N,1}(s)|^2 ds\right)\right.
\end{aligned}
$$

$$\times\left(\int_0^T \left\|\partial_\mu\Phi(i)[\cdot]\right\|^2_{L^2(\mathbb{R},\tilde{\mu}_s^{\epsilon,N})}ds\right)\Bigg]\|\psi\|_\infty^2$$
$$\leq C\frac{\epsilon^2}{a(N)\sqrt{N}}T\left(1+\frac{1}{N}\sum_{i=1}^N\sup_{t\in[0,T]}\mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N}|^{2\tilde{q}_\Phi(1,0,0)}\right]\right)\|\psi\|_\infty^2$$

where we use the bound (53) in the last step. The result follows from Lemmas B.1 and B.3, using that the exponent of $|\tilde{Y}_t^{i,\epsilon,N}|$ in the expectation of all these bounds is less than or equal to 2 as imposed in Assumption (A6). Lemma B.3 is used to handle the last term appearing in the bound of the first part of $\tilde{B}_5$. □

Remark 6.2 Bounding the first term in $\tilde{B}_5$ in Proposition 6.1 is the only place where Lemma B.3 is required in this manuscript. The proof of Lemma B.3 is where it is required that there exists $\rho\in(0,1)$ such that $a(N)\sqrt{N}\epsilon^\rho\to\lambda\in(0,+\infty]$. Thus, if this term can be otherwise bounded (e.g. if c or $\partial_\mu\Phi$ is uniformly bounded), one can relax this technical assumption on the scaling sequence $a(N)$ to $a(N)\sqrt{N}\epsilon\to 0$. Moreover, $a(N)\sqrt{N}\epsilon\to 0$ is needed so that the term $\tilde{B}_1$ in Proposition 6.1 vanishes - without this, one cannot hope to prove tightness of $\{\tilde{Z}^N\}_{N\in\mathbb{N}}$, as in Proposition 7.8 there would be an $\mathcal{O}(1)$ term which is not uniformly continuous with respect to time. If $b\equiv 0$ and hence there is no need for Proposition 6.1, it is possible to prove tightness even when $a(N)\sqrt{N}\epsilon\to\lambda\in[0,\infty)$. Under this scaling, we expect to get a different formulation for the rate function in Theorem 3.2 when $\lambda>0$. This is an interesting avenue for future research which we do not pursue here for purposes of the presentation.

Proposition 6.3 *In the setup of Proposition 6.1, assume in addition (A7). Then*

$$\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t\frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Phi(i)[j]\psi(s,\tilde{X}_s^{i,\epsilon,N})ds\right|^2\right]$$
$$\leq C[\epsilon^2 a(N)\sqrt{N}(1+T+T^2)+\frac{a(N)}{N^{3/2}}T^2]\|\psi\|^2_{C_b^{1,2}}.$$

Proof Recall the operator $L_{x,\mu}$ from Eq. (18). For fixed $x\in\mathbb{R}$, $\mu\in\mathcal{P}(\mathbb{R})$, this is the generator of

$$dY_t^{x,\mu}=f(x,Y_t^{x,\mu},\mu)dt+\tau_1(x,Y_t^{x,\mu},\mu)dW_t+\tau_2(x,Y_t^{x,\mu},\mu)dB_t \tag{59}$$

for W_t, B_t independent 1-D Brownian motions.

We introduce a new generator $L^2_{x,\bar{x},\mu}$ parameterized by $x,\bar{x}\in\mathbb{R}$, $\mu\in\mathcal{P}_2$ which acts on $\psi\in C_b^2(\mathbb{R}^2)$ by

$$L^2_{x,\bar{x},\mu}\psi(y,\bar{y})=f(x,y,\mu)\psi_y(y,\bar{y})+f(\bar{x},\bar{y},\mu)\psi_{\bar{y}}(y,\bar{y})$$
$$+\frac{1}{2}[\tau_1^2(x,y,\mu)+\tau_2^2(x,y,\mu)]\psi_{yy}(y,\bar{y})$$

$$+\frac{1}{2}[\tau_1^2(\bar{x},\bar{y},\mu)+\tau_2^2(\bar{x},\bar{y},\mu)]\psi_{\bar{y}\bar{y}}(y,\bar{y}). \tag{60}$$

This is the generator associated to the 2-dimensional process solving 2 independent copies of Eq. (59) where the same parameter μ enters both equations, but different $x, \bar{x}$ enter each equation, i.e.

$$\begin{aligned} dY_t^{x,\mu} &= f(x,Y_t^{x,\mu},\mu)dt+\tau_1(x,Y_t^{x,\mu},\mu)dW_t+\tau_2(x,Y_t^{x,\mu},\mu)dB_t \\ d\bar{Y}_t^{\bar{x},\mu} &= f(\bar{x},\bar{Y}_t^{\bar{x},\mu},\mu)dt+\tau_1(\bar{x},\bar{Y}_t^{\bar{x},\mu},\mu)d\bar{W}_t+\tau_2(\bar{x},\bar{Y}_t^{\bar{x},\mu},\mu)d\bar{B}_t. \end{aligned} \tag{61}$$

for $W_t, B_t, \bar{W}_t, \bar{B}_t$ independent 1-D Brownian motions.

It is easy then to see that the unique distributional solution of the adjoint equation

$$L^2_{x,\bar{x},\mu}\bar{\pi}(\cdot;x,\bar{x},\mu)=0, \qquad \int_{\mathbb{R}^2}\bar{\pi}(dy,d\bar{y};x,\bar{x},\mu)=1, \forall x,\bar{x}\in\mathbb{R}, \mu\in\mathcal{P}(\mathbb{R})$$

is given by

$$\bar{\pi}(dy,d\bar{y};x,\bar{x},\mu)=\pi(dy;x,\mu)\otimes\pi(d\bar{y};\bar{x},\mu) \tag{62}$$

where π is as in Eq. (20). We now consider $\chi(x,\bar{x},y,\bar{y},\mu):\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathcal{P}(\mathbb{R})\to\mathbb{R}$ solving

$$\begin{aligned} &L^2_{x,\bar{x},\mu}\chi(x,\bar{x},y,\bar{y},\mu)=-b(x,y,\mu)\partial_\mu\Phi(\bar{x},\bar{y},\mu)[x] \\ &\int_{\mathbb{R}}\int_{\mathbb{R}}\chi(x,\bar{x},y,\bar{y},\mu)\pi(dy;x,\mu)\pi(d\bar{y},\bar{x},\mu)=0. \end{aligned} \tag{63}$$

Note that by the centering condition, Eq. (21), the right hand side of Eq. (63) integrates against $\bar{\pi}$ from Eq. (62) to 0 for all $x,\bar{x},\mu$. Also, the second order coefficient in $L^2_{x,\bar{x},\mu}$ is uniformly elliptic by virtue of Assumption (A1), and by virtue of Eq. (19), there is $R_{f_2}>0$ and $\Gamma_2>0$ such that

$$\sup_{x,\bar{x},\mu}(f(x,y,\mu)y+f(\bar{x},\bar{y},\mu)\bar{y})\le-\Gamma_2(|y|^2+|\bar{y}|^2), \forall y,\bar{y} \text{ such that } \sqrt{y^2+\bar{y}^2}>R_{f_2}.$$

Thus indeed we have a unique solution to (63) by Theorem 1 in [60] (which is a classical solution by assumption). Applying Itô's formula to $\chi^N(\tilde{X}_t^{j,\epsilon,N},\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{j,\epsilon,N},\tilde{X}_t^{i,\epsilon,N},(\tilde{X}_t^{1,\epsilon,N},\dots,\tilde{X}_t^{N,\epsilon,N}))\psi(t,\tilde{X}_t^{i,\epsilon,N})$, where $\chi^N:\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^N\to\mathbb{R}$ is the empirical projection of χ and using Proposition D.6, we get

$$\int_0^t\frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Phi(i)[j]\psi(s,\tilde{X}_s^{i,\epsilon,N})ds=\frac{1}{N}\sum_{j=1}^N\sum_{k=1}^{13}\bar{B}_k^{i,j,\epsilon,N}(t)$$

where

$$\bar{B}_1^{i,j,\epsilon,N}(t)=\epsilon^2[\chi(\tilde{X}_0^{j,\epsilon,N},\tilde{X}_0^{i,\epsilon,N},\tilde{Y}_0^{j,\epsilon,N},\tilde{Y}_0^{i,\epsilon,N},\tilde{\mu}_0^{\epsilon,N})\psi(0,\tilde{X}_0^{i,\epsilon,N})$$

$$- \chi(\tilde{X}_t^{j,\epsilon,N}, \tilde{X}_t^{i,\epsilon,N}, \tilde{Y}_t^{j,\epsilon,N}, \tilde{Y}_t^{i,\epsilon,N}, \tilde{\mu}_t^{\epsilon,N})\psi(t, \tilde{X}_t^{i,\epsilon,N})]$$

$$\bar{B}_2^{i,j,\epsilon,N}(t) = \epsilon \int_0^t \Bigg(b(j)\chi_x(i,j)\psi(s,i) + b(i)\Big[\chi_{\bar{x}}(i,j)\psi(s,i) + \chi(i,j)\psi_{\bar{x}}(s,i)\Big]$$
$$+ g(j)\chi_y(i,j)\psi(s,i) + g(i)\chi_{\bar{y}}(i,j)\psi(s,i)$$
$$+ \sigma(j)\tau_1(j)\chi_{xy}(i,j)\psi(s,i) + \sigma(i)\tau_1(i)\Big[\chi_{\bar{x}\bar{y}}(i,j)\psi(s,i) + \chi_{\bar{y}}(i,j)\psi_{\bar{x}}(s,i)\Big]\Bigg)ds$$

$$\bar{B}_3^{i,j,\epsilon,N}(t) = \epsilon \int_0^t \frac{1}{N}\sum_{k=1}^N b(k)\partial_\mu \chi(i,j)[k]\psi(s,i)ds$$

$$\bar{B}_4^{i,j,\epsilon,N}(t) = \frac{\epsilon}{N}\int_0^t \Big(\sigma(j)\tau_1(j)\partial_\mu \chi_y(i,j)[j]\psi(s,i) + \sigma(i)\tau_1(i)\partial_\mu \chi^2_{\bar{y}}(i,j)[i]\psi(s,i)\Big)ds$$

$$\bar{B}_5^{i,j,\epsilon,N}(t) = \epsilon^2 \int_0^t \Bigg(\chi(i,j)\dot{\psi}(s,i) + c(j)\chi_x(i,j)\psi(s,i)$$
$$+ c(i)\Big[\chi_{\bar{x}}(i,j)\psi(s,i) + \chi(i,j)\psi_{\bar{x}}(s,i)\Big]$$
$$+ \frac{1}{N}\sum_{k=1}^N \Big\{ c(k)\partial_\mu \chi(i,j)[k]\Big\}\psi(s,i) + \frac{1}{2}\sigma^2(j)\chi_{xx}(i,j)\psi(s,i)$$
$$+ \frac{1}{2}\sigma^2(i)\Big[\chi_{\bar{x}\bar{x}}(i,j)\psi(s,i) + 2\chi_{\bar{x}}(i,j)\psi_{\bar{x}}(s,i) + \chi(i,j)\psi_{\bar{x}\bar{x}}(s,i)\Big]$$
$$+ \frac{1}{2}\frac{1}{N}\sum_{k=1}^N \Big\{\sigma^2(k)\Big[\partial_z\partial_\mu \chi(i,j)[k] + \frac{1}{N}\partial_\mu^2 \chi^2(i,j)[k,k]\Big]\Big\}\psi(s,i)$$
$$+ \frac{1}{N}\sigma^2(j)\partial_\mu \chi_x(i,j)[j]\psi(s,i)$$
$$+ \frac{1}{N}\sigma^2(i)\Big[\partial_\mu \chi_{\bar{x}}(i,j)[i]\psi(s,i) + \partial_\mu \chi(i,j)[i]\psi_{\bar{x}}(s,i)\Big]\Bigg)ds$$

$$\bar{B}_6^{i,j,\epsilon,N}(t) = \epsilon \int_0^t \tau_1(j)\chi_y(i,j)\psi(s,i)dW_s^j$$
$$+ \epsilon \int_0^t \tau_2(j)\chi_y(i,j)\psi(s,i)dB_s^j + \epsilon \int_0^t \tau_1(i)\chi_{\bar{y}}(i,j)\psi(s,i)dW_s^i$$
$$+ \epsilon \int_0^t \tau_2(i)\chi_{\bar{y}}(i,j)\psi(s,i)dB_s^i$$

$$\bar{B}_7^{i,j,\epsilon,N}(t) = \epsilon^2 \int_0^t \sigma(j)\chi_x(i,j)\psi(s,i)dW_s^j$$
$$+ \epsilon^2 \int_0^t \sigma(i)\Big[\chi_{\bar{x}}(i,j)\psi(s,i) + \chi(i,j)\psi_{\bar{x}}(s,i)\Big]dW_s^i$$
$$+ \frac{\epsilon^2}{N}\sum_{k=1}^N \Big\{\int_0^t \sigma(k)\partial_\mu \chi(i,j)[k]\psi(s,i)dW_s^k\Big\}$$

$$\bar{B}_8^{i,j,\epsilon,N}(t) = \epsilon^2 \int_0^t \Bigg(\frac{\sigma(j)\tilde{u}_j^{N,1}(s)}{\sqrt{N}a(N)}\chi_x(i,j)\psi(s,i) + \frac{\sigma(i)\tilde{u}_i^{N,1}(s)}{\sqrt{N}a(N)}$$
$$\Big[\chi_{\bar{x}}(i,j)\psi(s,i) + \chi(i,j)\psi_{\bar{x}}(s,i)\Big]\Bigg)ds$$

$$\bar{B}_9^{i,j,\epsilon,N}(t) = \epsilon^2 \int_0^t \frac{1}{N}\sum_{k=1}^N \left\{ \frac{\sigma(k)\tilde{u}_k^{N,1}(s)}{\sqrt{N}a(N)} \partial_\mu \chi^2(i,j)[k] \right\} \psi(s,i)ds$$

$$\bar{B}_{10}^{i,j,\epsilon,N}(t) = \epsilon \int_0^t \Bigg(\Bigg[\frac{\tau_1(j)\tilde{u}_j^{N,1}(s)}{\sqrt{N}a(N)} + \frac{\tau_2(j)\tilde{u}_j^{N,2}(s)}{\sqrt{N}a(N)} \Bigg] \chi_y(i,j)\psi(s,i) + \left[\frac{\tau_1(i)\tilde{u}_i^{N,1}(s)}{\sqrt{N}a(N)} + \frac{\tau_2(i)\tilde{u}_i^{N,2}(s)}{\sqrt{N}a(N)} \right] \chi_{\bar{y}}(i,j)\psi(s,i) \Bigg) ds$$

$$\bar{B}_{11}^{i,j,\epsilon,N}(t) = \mathbb{1}_{i=j}\epsilon^2 \int_0^t \sigma(i)\sigma(j) \Big[\chi_{x\bar{x}}(i,j)\psi(s,i) + \chi_x(i,j)\psi_{\bar{x}}(s,i) \Big] ds$$

$$\bar{B}_{12}^{i,j,\epsilon,N}(t) = \mathbb{1}_{i=j}\epsilon \int_0^t \Big(\sigma(j)\tau_1(i)\chi_{x\bar{y}}(i,j)\psi(s,i) + \sigma(i)\tau_1(j) \Big[\chi_{\bar{x}y}(i,j)\psi(s,i) + \chi_y(i,j)\psi_{\bar{x}}(s,i) \Big] \Big) ds$$

$$\bar{B}_{13}^{i,j,\epsilon,N}(t) = \mathbb{1}_{i=j} \int_0^t \Big[\tau_1(i)\tau_1(j) + \tau_2(i)\tau_2(j) \Big] \chi_{y\bar{y}}(i,j)\psi(s,i)ds.$$

Here we have introduced the notation $\chi(i,j)$ to denote $\chi(\tilde{X}_s^{j,\epsilon,N}, \tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{j,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N})$, $\partial_\mu\chi(i,j)[k]$ to denote $\partial_\mu\chi(\tilde{X}_s^{j,\epsilon,N}, \tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{j,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N})[\tilde{X}_s^{k,\epsilon,N}]$, and similarly for $\partial_\mu\chi(i,j)[k,k]$. We also use $\psi(s,i)$ to denote $\psi(s,\tilde{X}_s^{i,\epsilon,N})$.

Using that σ, τ_1, τ_2, and g are bounded and (A7) on the growth of χ and its derivatives, the proof that

$$\frac{a(N)}{\sqrt{N}} \sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} \left| \frac{1}{N}\sum_{j=1}^N \sum_{k=1}^{12} \bar{B}_k^{i,j,\epsilon,N}(t) \right|^2 \right] \leq C\epsilon^2 a(N)\sqrt{N}(1+T+T^2)\,\|\psi\|^2_{C_b^{1,2}}$$

follows essentially in the same way as Proposition 6.1. For example, for $\bar{B}_2$, we can use the assumed linear growth in y of b and boundedness of g and σ from (A5), boundedness of τ_1 from (A1), and boundedness of $\chi, \chi_x, \chi_{\bar{x}}, \chi_y, \chi_{\bar{y}}$ and polynomial growth in y of χ_{xy} and $\chi_{\bar{x}\bar{y}}$ to get:

$$\begin{aligned}
&\frac{a(N)}{\sqrt{N}} \sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} \left| \frac{1}{N}\sum_{j=1}^N \bar{B}_2^{i,j,\epsilon,N}(t) \right|^2 \right] \\
&\leq C\epsilon^2 a(N)\sqrt{N}T^2 \left(1 + \frac{1}{N}\sum_{i=1}^N \sup_{t\in[0,T]} \mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N}|^2 + |\tilde{Y}_t^{i,\epsilon,N}|^{2q_{\chi_y}(0,1,0)} \right] \right) \\
&\quad \times (\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2) \\
&\leq C\epsilon^2 a(N)\sqrt{N}T^2 \left(1 + \frac{1}{N}\sum_{i=1}^N \sup_{t\in[0,T]} \mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N}|^2 \right] \right) (\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2) \\
&\leq C\epsilon^2 a(N)\sqrt{N}T^2 (\|\psi\|_\infty^2 + \|\psi_x\|_\infty^2),
\end{aligned}$$

where in the last step we used Lemma B.1.

The other bounds follow similarly. We omit the details for brevity. To handle the last term, we see by boundedness of τ_1, τ_2 from (A1) and linear growth of $\chi_{y\bar{y}}$ from (A7):

$$
\begin{aligned}
&\frac{a(N)}{\sqrt{N}}\sum_{i=1}^{N}\mathbb{E}\Bigg[\sup_{t\in[0,T]}\Bigg|\frac{1}{N}\sum_{j=1}^{N}\bar{B}_{13}^{i,j,\epsilon,N}\Bigg|^2\Bigg] \\
&= \frac{a(N)}{\sqrt{N}}\frac{1}{N^2}\sum_{i=1}^{N}\mathbb{E}\Bigg[\sup_{t\in[0,T]}\Bigg|\int_0^t\Big[\tau_1^2(i)+\tau_2^2(i)\Big]\chi_{y\bar{y}}(i,i)\psi(s,i)ds\Bigg|^2\Bigg] \\
&\le \frac{a(N)}{\sqrt{N}}\frac{1}{N^2}CT\sum_{i=1}^{N}\mathbb{E}\Bigg[\int_0^T|\chi_{y\bar{y}}(i,i)|^2ds\Bigg]\|\psi_\infty\|^2 \\
&\le \frac{a(N)}{N^{3/2}}CT^2(1+\frac{1}{N}\sum_{i=1}^{N}\sup_{t\in[0,T]}\mathbb{E}\Big[|\tilde{Y}_t^{i,\epsilon,N}|^2\Big])\|\psi_\infty\|^2 \\
&\le \frac{a(N)}{N^{3/2}}CT^2\|\psi_\infty\|^2 \text{ (by Lemma B.1).}
\end{aligned}
$$

□

Proposition 6.4 *Assume (A1)–(A5). Let F be any function such that Ξ satisfies assumption (A8). Then for $\bar{F}(x,\mu) := \int_{\mathbb{R}} F(x,y,\mu)\pi(dy;x,\mu)$, with π as in Eq. (20) and $\psi \in C_b^{1,2}([0,T]\times\mathbb{R})$*

$$
\begin{aligned}
&\frac{a(N)}{\sqrt{N}}\sum_{i=1}^{N}\mathbb{E}\Bigg[\sup_{t\in[0,T]}\Bigg|\int_0^t\Big(F(\tilde{X}_s^{i,\epsilon,N},\tilde{Y}_s^{i,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})-\bar{F}(\tilde{X}_s^{i,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})\Big)\psi(s,\tilde{X}_s^{i,\epsilon,N})dt\Bigg|\Bigg] \\
&\le C\epsilon a(N)\sqrt{N}(1+T+T^{1/2})\|\psi\|_{C_b^{1,2}}.
\end{aligned}
$$

Proof By Lemma C.1, we can consider $\Xi : \mathbb{R}\times\mathbb{R}\times\mathcal{P}(\mathbb{R})\to\mathbb{R}$ the unique classical solution to

$$
\begin{aligned}
L_{x,\mu}\Xi(x,y,\mu) &= -[F(x,y,\mu)-\int_{\mathbb{R}}F(x,y,\mu)\pi(dy;x,\mu)], \\
&\int_{\mathbb{R}}\Xi(x,y,\mu)\pi(dy;x,\mu)=0. \qquad (64)
\end{aligned}
$$

(Ξ and F may also depend on $t\in[0,T]$, but we suppress this in the notation here). Applying Itô's formula to $\Xi^N(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},(\tilde{X}_t^{1,\epsilon,N},\dots,\tilde{X}_t^{N,\epsilon,N}))\psi(t,\tilde{X}_t^{i,\epsilon,N})$, where again $\Xi^N : \mathbb{R}\times\mathbb{R}\times\mathbb{R}^N\to\mathbb{R}$ is the empirical projection of Ξ and using Proposition D.6, we get:

$$
\int_0^t\Big(F(\tilde{X}_s^{i,\epsilon,N},\tilde{Y}_s^{i,\epsilon,N},\tilde{\mu}^{\epsilon,N})-\bar{F}(\tilde{X}_s^{i,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})\Big)\psi(s,\tilde{X}_s^{i,\epsilon,N})dt=\sum_{k=1}^{10}C_k^{i,\epsilon,N}(t)
$$

$$C_1^{i,\epsilon,N}(t) = \epsilon^2[\Xi(\tilde{X}_0^{i,\epsilon,N}, \tilde{Y}_0^{i,\epsilon,N}, \tilde{\mu}_0^{\epsilon,N})\psi(0, \tilde{X}_0^{i,\epsilon,N}) - \Xi(\tilde{X}_t^{i,\epsilon,N}, \tilde{Y}_t^{i,\epsilon,N}, \tilde{\mu}_t^{\epsilon,N})\psi(t, \tilde{X}_t^{i,\epsilon,N})]$$

$$C_2^{i,\epsilon,N}(t) = \epsilon\int_0^t \Big(b(i)[\Xi_x(i)\psi(s,i) + \Xi(i)\psi_x(s,i)] + g(i)\Xi_y(i)\psi(s,i) + \sigma(i)\tau_1(i)[\Xi_{xy}(i)\psi(s,i) + \Xi_y(i)\psi_x(s,i)]\Big)ds$$

$$C_3^{i,\epsilon,N}(t) = \epsilon\int_0^t \frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Xi(i)[j]\psi(s,i)ds$$

$$C_4^{i,\epsilon,N}(t) = \frac{\epsilon}{N}\int_0^t \sigma(i)\tau_1(i)\partial_\mu\Xi_y(i)[i]\psi(s,i)ds$$

$$\begin{aligned}C_5^{i,\epsilon,N}(t) = \epsilon^2\int_0^t \Big(&\Xi(i)\dot{\psi}(s,i) + c(i)[\Xi_x(i)\psi(s,i) + \Xi(i)\psi_x(s,i)]\\ &+ \frac{\sigma^2(i)}{2}[\Xi_{xx}(i)\psi(s,i) + 2\Xi_x(i)\psi_x(s,i) + \Xi(i)\psi_{xx}(s,i)\\ &+ \frac{2}{N}\partial_\mu\Xi(i)[i]\psi_x(s,i) + \frac{2}{N}\partial_\mu\Xi_x(i)[i]\psi(s,i)]\\ &+ \frac{1}{N}\sum_{j=1}^N\Big\{c(j)\partial_\mu\Xi(i)[j]\psi(s,i) + \frac{1}{2}\sigma^2(j)[\frac{1}{N}\partial_\mu^2\Xi(i)[j,j] + \partial_z\partial_\mu\Xi(i)[j]]\psi(s,i)\Big\}\Big)ds\end{aligned}$$

$$C_6^{i,\epsilon,N}(t) = \epsilon\int_0^t \tau_1(i)\Xi_y(i)\psi(s,i)dW_s^i + \epsilon\int_0^t \tau_2(i)\Xi_y(i)\psi(s,i)dB_s^i$$

$$C_7^{i,\epsilon,N}(t) = \epsilon^2\Big[\int_0^t \sigma(i)[\Xi_x(i)\psi(s,i) + \Xi(i)\psi_x(s,i)]dW_t^i + \frac{1}{N}\sum_{j=1}^N\Big\{\int_0^t \sigma(j)\partial_\mu\Xi(i)[j]\psi(s,i)dW_s^j\Big\}\Big]$$

$$C_8^{i,\epsilon,N}(t) = \epsilon^2\int_0^t \frac{\sigma(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s)[\Xi_x(i)\psi(s,i) + \Xi(i)\psi_x(s,i)]ds$$

$$C_9^{i,\epsilon,N}(t) = \epsilon^2\int_0^t \frac{1}{N}\Big\{\sum_{j=1}^N \frac{\sigma(j)}{a(N)\sqrt{N}}\tilde{u}_j^{N,1}(s)\partial_\mu\Xi(i)[j]\psi(s,i)\Big\}ds$$

$$C_{10}^{i,\epsilon,N}(t) = \epsilon\int_0^t [\frac{\tau_1(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s) + \frac{\tau_2(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)]\Xi_y(i)\psi(s,i)ds.$$

Then using that σ, τ_1, τ_2, and g are bounded and that b, c grow at most linearly in y uniformly in x, μ, and the assumptions on the growth of Ξ and its derivatives from (A8), the proof follows in essentially the same way as Propositions 6.1 and 6.3.

Since Ξ is not necessarily bounded under Assumption (A8) ($\tilde{q}_\Xi(n,l,\boldsymbol{\beta}) \leq 1, \forall(n,l,\boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}_1$ allows Ξ to grow linearly in y), we need to handle the first term in the following way:

$$\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N \epsilon^2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\Xi(\tilde{X}_0^{i,\epsilon,N}, \tilde{Y}_0^{i,\epsilon,N}, \tilde{\mu}_0^{\epsilon,N})\psi(0, \tilde{X}_0^{i,\epsilon,N})$$

$$- \Xi(\tilde{X}^{i,\epsilon,N}_t, \tilde{Y}^{i,\epsilon,N}_t, \tilde{\mu}^{\epsilon,N}_t)\psi(t, \tilde{X}^{i,\epsilon,N}_t)\Big|\Big] \le$$

$$\le C\epsilon^2 a(N)\sqrt{N}\Bigg[1 + \frac{1}{N}\sum_{i=1}^N \mathbb{E}\Bigg[\sup_{t\in[0,T]} |\tilde{Y}^{i,\epsilon,N}_t|\Bigg]\Bigg] \|\psi\|_\infty$$

$$\le C(\rho)\epsilon^2 a(N)\sqrt{N}\Bigg[1 + \epsilon^{-\rho}\Bigg] \|\psi\|_\infty$$

for any $\rho \in (0, 2)$ by Lemma B.2. Taking any $\rho \in (0, 1]$, the desired bound holds.

The only other terms that are handled differently in a way that matters are $C^{i,\epsilon,N}_4(t)$, which corresponds to $\tilde{B}^{i,\epsilon,N}_2(t)$, where the difference of having a ϵ in front means that it is bounded by $\epsilon a(N) \le \epsilon a(N)\sqrt{N}$, hence there being no need to include $a(N)/\sqrt{N}$ in the definition of $C(N)$, and $C^{i,\epsilon,N}_2(t)$, $C^{i,\epsilon,N}_3(t)$, $C^{i,\epsilon,N}_6(t)$, and $C^{i,\epsilon,N}_{10}(t)$, which were $O(1)$ in Lemma 6.1 and hence were not shown to vanish. C_2 is handled as $\tilde{B}_3$ was, C_3 in the same way that $\tilde{B}_5$ was, C_6 in the same way that $\tilde{B}_6$ was, and C_{10} in the same way that $\tilde{B}_7$ was.

□

7 Tightness of the controlled pair

In this section we throughout fix any controls satisfying the bound (53) and prove tightness of the pair $(\tilde{Z}^N, Q^N)$ from Eqs. (54) and (56) under those controls. We will establish tightness in the appropriate spaces by proving tightness for each of the marginals.

As discussed in Sect. 2, in order to prove tightness of the controlled fluctuation process $\tilde{Z}^N$ in $C([0, T]; \mathcal{S}_{-m})$ for some $m \in \mathbb{N}$ sufficiently large (see Eq. 5), we will use the theory of Mitoma from [55]. In particular, we need to prove uniform m-continuity for sufficiently large m in the family of Hilbert norms (4), and tightness of $\langle Z^N, \phi\rangle$ as a $C([0, T]; \mathbb{R})$-valued random variables in order to apply Theorem 3.1 and Remark R1) in [55]. For the former, by definition we need some uniform in time control over the $\mathcal{S}_{-m}$-norm of $\tilde{Z}^N$. By Markov's inequality, it suffices to show that $\sup_{\phi\in\mathcal{S}:\|\phi\|_m=1} \mathbb{E}[\sup_{t\in[0,T]} |\langle \tilde{Z}^N_t, \phi\rangle|] \le C$ (see, e.g., the proof of [10] Theorem 4.7). As mentioned in Remark 5.1, we will do so in Lemma 7.6 via triangle inequality and establishing an L^2 rate of convergence of the controlled particle system (55) to the IID particle system (57), and a rate of convergence of the IID particle system (57) to the averaged McKean–Vlasov SDE (25). The convergence of the controlled particle system (55) to the IID particle system (57) is the subject of Sect. 7.2 and the convergence of the IID slow–fast McKean–Vlasov SDEs (57) to the averaged McKean–Vlasov SDE (25) is the subject of Sect. 7.1.

A major difference between the coupling arguments in the references listed in Remark 5.1 and ours is that the IID system in the listed references were all equal in distribution to the law of the system which they are considering fluctuations from. This is not the case for us, since, as is well-known, we do not expect in general to have L^2 convergence of fully-coupled slow–fast diffusions to their averaged limit (see

[3] Remark 3.4.4 for an illustrative example). In other words, Lemma 7.5 cannot hold with $\bar{X}^{i,\epsilon}$ replaced by IID copies of the averaged limiting McKean–Vlasov Equation X_t from Eq. (25). We are thus exploiting here the fact that the limits $\epsilon \downarrow 0$ and $N \to \infty$ commute, as shown in [4], and hence we can use an IID system of Slow–Fast McKean–Vlasov SDEs as our intermediate process for our proof of tightness. This commutativity of the limits will hold so long as sufficient conditions for the propagation of chaos and stochastic averaging respectively hold for the system of SDEs (55) and the invariant measure π from Eq. (20) is unique for all $x \in \mathbb{R}, \mu \in \mathcal{P}(\mathbb{R})$. Recall that the latter is a consequence of assumptions (A1) and (A2).

Tightness of Q^N is contained in Sect. 7.4, and is essentially a consequence of moment bounds on the controlled particles (55), which again follow from the results of Sect. (6).

7.1 On the rate of averaging for fully-coupled slow–fast McKean–Vlasov diffusions

Here we recall a result which allows us to establish closeness of the slow–fast McKean–Vlasov SDEs (57) to the averaged McKean–Vlasov SDE (25). This result will be used in the Lemma 7.6, which is a key ingredient in the proof of tightness of $\{\tilde{Z}^N\}_{N\in\mathbb{N}}$. Therein, the first term being bounded is essentially due to the propagation of chaos holding for the controlled particle system (55), as captured by Lemma 7.5. For the second term, the particles being IID means it is sufficient to gain control over convergence of $\sup_{\phi\in\mathcal{S}:\|\phi\|_m=1}|\mathbb{E}[\phi(\bar{X}^{1,\epsilon}_t)-\phi(X_t)]|$ as $\epsilon \downarrow 0$. There are very few results in the current literature in this direction. The existing averaging results for Slow–Fast McKean Vlasov SDEs can be found in [40, 48, 65] and in [5]. In [40, 48, 65] only systems where L^2 rates of averaging can be found are considered. Moreover, even for standard diffusion processes (which do not depend on their law), the only result for rates of convergence in distribution in the sense we desire for the fully-coupled setting is found in [64] Theorem 2.3. The fully coupled case for McKean–Vlasov diffusions is addressed in [5].

We mention here the main result from [5] that will be used in our case. In particular, we wish to establish a rate of convergence in distribution of

$$
\begin{aligned}
\bar{X}^\epsilon_t &= \eta^x + \int_0^t \left[\frac{1}{\epsilon}b(\bar{X}^\epsilon_s,\bar{Y}^\epsilon_s,\mathcal{L}(\bar{X}^\epsilon_s)) + c(\bar{X}^\epsilon_s,\bar{Y}^\epsilon_s,\mathcal{L}(\bar{X}^\epsilon_s))\right]ds \\
&\quad + \int_0^t \sigma(\bar{X}^\epsilon_s,\bar{Y}^\epsilon_s,\mathcal{L}(\bar{X}^\epsilon_s))dW_s \\
\bar{Y}^\epsilon_t &= \eta^y + \int_0^t \frac{1}{\epsilon}\left[\frac{1}{\epsilon}f(\bar{X}^\epsilon_s,\bar{Y}^\epsilon_s,\mathcal{L}(\bar{X}^\epsilon_s)) + g(\bar{X}^\epsilon_s,\bar{Y}^\epsilon_s,\mathcal{L}(\bar{X}^\epsilon_s))\right]dt \\
&\quad + \frac{1}{\epsilon}\left[\int_0^t \tau_1(\bar{X}^\epsilon_s,\bar{Y}^\epsilon_s,\mathcal{L}(\bar{X}^\epsilon_s))dW_s + \int_0^t \tau_2(\bar{X}^\epsilon_s,\bar{Y}^\epsilon_s,\mathcal{L}(\bar{X}^\epsilon_s))dB_s\right],
\end{aligned}
\tag{65}
$$

to the solution of Eq. (25). Note that a solution to Eq. (65) is equal in distribution to the IID particles from Eq. (57). The following moment bound holds.

Lemma 7.1 *Assume (A1)–(A2), (A4), and (A5). Then for any $p \in \mathbb{N}$:*

$$\sup_{\epsilon>0}\sup_{t\in[0,T]}\mathbb{E}\left[|\bar{Y}_t^{\epsilon}|^{2p}\right] \leq C(p,T) + |\eta^y|^{2p}.$$

Proof The proof of this lemma is omitted as it follows very closely the proof of Lemma 4.1 in [5]. □

Then, the main result of [5] that is relevant for our purposes is Theorem 7.2.

Theorem 7.2 *(Corollary 3.2 of [5]) Assume that assumptions (A1)–(A8) as well as (A10)–(A11) hold. Then for $\phi \in C^4_{b,L}(\mathbb{R})$, there is a constant $C = C(T)$ that is independent of ϕ such that*

$$\sup_{s\in[0,T]}\left|\mathbb{E}[\phi(\bar{X}_s^{\epsilon})] - \mathbb{E}[\phi(X_s)]\right| \leq \epsilon C(T)|\phi|_4,$$

where $\bar{X}^{\epsilon}$ is as in Eq. (65), X is as in Eq. (25), and $|\cdot|_4$ is as in Eq. (8).

Remark 7.3 Though in [5] the assumptions are stated in terms of sufficient conditions on the limiting coefficients for the needed regularity of Φ, χ, Ξ, $\bar{\gamma}$, $\bar{D}^{1/2}$, and $\tilde{\chi}$ in the proofs therein to hold (which is much easier to do in that situation since the lack of control eliminates the need for tracking specific rates of polynomial growth), it can be checked that the assumptions imposed on these functions by (A6), (A7), (A8), (A10), and (A11) respectively are sufficient. See also Remark 2.6 therein.

In particular, in [5], since specific rates of polynomial growth are not tracked, it is assumed the initial condition of $\bar{Y}^{\epsilon}$ has all moments bounded. This holds automatically here, since $\eta^y \in \mathbb{R}$ are deterministic. Then, due to Lemma 7.1, it is sufficient to show the derivatives of the Poisson equations which show up in the proof have polynomial growth in y uniformly in x, μ, z. In fact, the same Poisson equations are being used in Section 4 of [5] to gain ergodic-type theorems of the same nature as those of Sect. 6 here. The growth rates of Φ, χ, Ξ as imposed in (A6), (A7), and (A8) are already required here for the ergodic-type theorems for the controlled system (55) found in Sect. 6, and these conditions can be seen as more than sufficient for the results of [5] to go through. The solution $\tilde{\chi}$ to Eq. (28) does not, however, appear elsewhere in this paper, despite appearing in Proposition 4.4. of [5], which is fundamental to the result presented here as Theorem 7.2. This is why we can allow for the specified derivatives of $\tilde{\chi}$ (which are exactly those appearing in the proof of Proposition 4.4.) in Assumption (A11) to have polynomial growth of any order.

Lastly, we remark that the regularity of the limiting coefficients imposed by (A10) is used not for ergodic-type theorems, but instead to establish regularity a Cauchy-Problem on Wasserstein space in Lemma 5.1 of [5], which provides a refinement of Theorem 2.15 in [14]. As remarked therein, these assumptions can likely be relaxed via an alternative proof method, but as it stands these are the only results in this direction which provide sufficient regularity on the derivatives needed to prove Theorem 7.2.

7.2 Coupling of the controlled particles and the IID slow–fast McKean–Vlasov particles

Here we establish a coupling result, which we will use along with Theorem 7.2 in order to establish tightness for the controlled fluctuation process $\{\tilde{Z}^N\}$ from Eq. (54). Recall the processes $(\tilde{X}^{i,\epsilon,N}_t, \tilde{Y}^{i,\epsilon,N}_t)$ that satisfy (55) and $(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t)$ that satisfies (57).

Lemma 7.4 *Assume (A1)–(A7), (A9), and (A12). Then there exists $C > 0$ such that for all $t \in [0,T]$:*

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[|\tilde{Y}^{i,\epsilon,N}_t - \bar{Y}^{i,\epsilon}_t|^2\right] \le C\left\{\epsilon^2 + \frac{1}{N} + \frac{1}{Na^2(N)} + \sup_{s\in[0,t]}\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\left|\tilde{X}^{i,\epsilon,N}_s - \bar{X}^{i,\epsilon}_s\right|^2\right]\right\}$$

Proof We set $\tau_1 \equiv 0$, since terms involving τ_1 can be handled in the same way as those involving τ_2 in the proof. By Itô's formula, and given that the stochastic integrals are martingales (using Lemmas B.1 and 7.1),

$$\begin{aligned}
&\frac{d}{dt}\mathbb{E}\left[|\tilde{Y}^{i,\epsilon,N}_t - \bar{Y}^{i,\epsilon}_t|^2\right]\\
&\quad = 2\mathbb{E}\Bigg[\frac{1}{\epsilon^2}\left(f(\tilde{X}^{i,\epsilon,N}_t, \tilde{Y}^{i,\epsilon,N}_t, \tilde{\mu}^{\epsilon,N}_t) - f(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t, \mathcal{L}(\bar{X}^{\epsilon}_t))\right)(\tilde{Y}^{i,\epsilon,N}_t - \bar{Y}^{i,\epsilon}_t)\\
&\quad + \frac{1}{2\epsilon^2}|\tau_2(\tilde{X}^{i,\epsilon,N}_t, \tilde{Y}^{i,\epsilon,N}_t, \tilde{\mu}^{\epsilon,N}_t) - \tau_2(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t, \mathcal{L}(\bar{X}^{\epsilon}_t))|^2\\
&\quad + \frac{1}{\epsilon}\left(g(\tilde{X}^{i,\epsilon,N}_t, \tilde{Y}^{i,\epsilon,N}_t, \tilde{\mu}^{\epsilon,N}_t) - g(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t, \mathcal{L}(\bar{X}^{\epsilon}_t))\right)(\tilde{Y}^{i,\epsilon,N}_t - \bar{Y}^{i,\epsilon}_t)\\
&\quad + \frac{1}{\epsilon\sqrt{N}a(N)}\tau_2(\tilde{X}^{i,\epsilon,N}_t, \tilde{Y}^{i,\epsilon,N}_t, \tilde{\mu}^{\epsilon,N}_t)\tilde{u}^{N,2}_i(t)(\tilde{Y}^{i,\epsilon,N}_t - \bar{Y}^{i,\epsilon}_t)\Bigg]\\
&\quad \le 2\mathbb{E}\Bigg[\frac{1}{\epsilon^2}\left(f(\tilde{X}^{i,\epsilon,N}_t, \tilde{Y}^{i,\epsilon,N}_t, \tilde{\mu}^{\epsilon,N}_t) - f(\tilde{X}^{i,\epsilon,N}_t, \bar{Y}^{i,\epsilon}_t, \tilde{\mu}^{\epsilon,N}_t)\right)(\tilde{Y}^{i,\epsilon,N}_t - \bar{Y}^{i,\epsilon}_t)\\
&\quad + \frac{1}{\epsilon^2}\left(f(\tilde{X}^{i,\epsilon,N}_t, \bar{Y}^{i,\epsilon}_t, \tilde{\mu}^{\epsilon,N}_t) - f(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t, \bar{\mu}^{\epsilon,N}_t)\right)(\tilde{Y}^{i,\epsilon,N}_t - \bar{Y}^{i,\epsilon}_t)\\
&\quad + \frac{1}{\epsilon^2}\left(f(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t, \bar{\mu}^{\epsilon,N}_t) - f(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t, \mathcal{L}(\bar{X}^{\epsilon}_t))\right)(\tilde{Y}^{i,\epsilon,N}_t - \bar{Y}^{i,\epsilon}_t)\\
&\quad + \frac{1}{\epsilon^2}|\tau_2(\tilde{X}^{i,\epsilon,N}_t, \tilde{Y}^{i,\epsilon,N}_t, \tilde{\mu}^{\epsilon,N}_t) - \tau_2(\tilde{X}^{i,\epsilon,N}_t, \bar{Y}^{i,\epsilon}_t, \tilde{\mu}^{\epsilon,N}_t)|^2\\
&\quad + \frac{2}{\epsilon^2}|\tau_2(\tilde{X}^{i,\epsilon,N}_t, \bar{Y}^{i,\epsilon}_t, \tilde{\mu}^{\epsilon,N}_t) - \tau_2(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t, \bar{\mu}^{\epsilon,N}_t)|^2\\
&\quad + \frac{2}{\epsilon^2}|\tau_2(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t, \bar{\mu}^{\epsilon,N}_t) - \tau_2(\bar{X}^{i,\epsilon}_t, \bar{Y}^{i,\epsilon}_t, \mathcal{L}(\bar{X}^{\epsilon}_t))|^2
\end{aligned}$$

$$+\frac{1}{\epsilon}\bigg(g(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})-g(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\bigg)(\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon})$$
$$+\frac{1}{\epsilon\sqrt{N}a(N)}\tau_2(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})\tilde{u}_i^{N,2}(t)(\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon})\bigg]$$
$$\le 2\mathbb{E}\bigg[-\frac{\beta}{2\epsilon^2}|\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon}|^2$$
$$+\frac{1}{2\eta\epsilon^2}\bigg|f(\tilde{X}_t^{i,\epsilon,N},\bar{Y}_t^{i,\epsilon},\tilde{\mu}_t^{\epsilon,N})-f(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})\bigg|^2+\frac{\eta}{2\epsilon^2}|\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon}|^2$$
$$+\frac{1}{2\eta\epsilon^2}\bigg|f(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})-f(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\bigg|^2+\frac{\eta}{2\epsilon^2}|\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon}|^2$$
$$+\frac{2}{\epsilon^2}|\tau_2(\tilde{X}_t^{i,\epsilon,N},\bar{Y}_t^{i,\epsilon},\tilde{\mu}_t^{\epsilon,N})-\tau_2(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})|^2$$
$$+\frac{2}{\epsilon^2}|\tau_2(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})-\tau_2(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))|^2$$
$$+\frac{1}{2\eta}\bigg|g(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})-g(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\bigg|^2+\frac{\eta}{2\epsilon^2}|\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon}|^2$$
$$+\frac{2}{\eta Na^2(N)}|\tau_2(\tilde{X}_t^{i,\epsilon,N},\tilde{Y}_t^{i,\epsilon,N},\tilde{\mu}_t^{\epsilon,N})|^2|\tilde{u}_i^{N,2}(t)|^2$$
$$+\frac{\eta}{2\epsilon^2}|\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon})|^2\bigg]$$

for all $\eta>0$, where in the second inequality we used Eq. (17) of Assumption (A2). Taking $\eta=\beta/8$ and using the boundedness of g from Assumption (A5) and of τ_1,τ_2 from Assumption (A1), we get:

$$\frac{d}{dt}\mathbb{E}\bigg[|\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon}|^2\bigg]$$
$$\le-\frac{\beta}{2\epsilon^2}\mathbb{E}\bigg[|\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon}|^2\bigg]+\frac{C}{\epsilon^2}\mathbb{E}\bigg[\bigg|f(\tilde{X}_t^{i,\epsilon,N},\bar{Y}_t^{i,\epsilon},\tilde{\mu}_t^{\epsilon,N})-f(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})\bigg|^2$$
$$+\bigg|f(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})-f(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\bigg|^2$$
$$+|\tau_2(\tilde{X}_t^{i,\epsilon,N},\bar{Y}_t^{i,\epsilon},\tilde{\mu}_t^{\epsilon,N})-\tau_2(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})|^2$$
$$+|\tau_2(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})-\tau_2(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))|^2\bigg]+\frac{C}{Na^2(N)}\mathbb{E}\bigg[|\tilde{u}_i^{N,2}(s)|^2\bigg]+C$$

Now using the global Lipschitz property of f from Assumption (A2) and of τ_1 and τ_2 from Assumption (A9) to handle the terms of the form $|f(\tilde{X}_t^{i,\epsilon,N},\bar{Y}_t^{i,\epsilon},\tilde{\mu}_t^{\epsilon,N})-f(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})|^2$ and Assumption (A12) with Lemma D.7 for the terms of the form $|f(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})-f(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))|^2$, we have:

$$\frac{d}{dt}\mathbb{E}\bigg[|\tilde{Y}_t^{i,\epsilon,N}-\bar{Y}_t^{i,\epsilon}|^2\bigg]$$

$$\leq -\frac{\beta}{2\epsilon^2}\mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N} - \bar{Y}_t^{i,\epsilon}|^2\right] + \frac{C}{\epsilon^2}\mathbb{E}\left[\left|\tilde{X}_t^{i,\epsilon,N} - \bar{X}_t^{i,\epsilon}\right|^2 + \frac{1}{N}\sum_{i=1}^N\left|\tilde{X}_t^{i,\epsilon,N} - \bar{X}_t^{i,\epsilon}\right|^2\right] + \frac{C}{\epsilon^2 N} + \frac{C}{Na^2(N)}\mathbb{E}\left[|\tilde{u}_i^{N,2}(s)|^2\right] + C.$$

When using Lipschitz continuity of f, τ_1, and τ_2, we are also using that

$$\mathbb{W}_2(\tilde{\mu}_t^{\epsilon,N}, \bar{\mu}_t^{\epsilon,N}) \leq \frac{1}{N}\sum_{i=1}^N\left|\tilde{X}_t^{i,\epsilon,N} - \bar{X}_t^{i,\epsilon}\right|^2$$

by Eq. (76) in Appendix 4.

Now using a comparison theorem, dividing by $\frac{1}{N}$ and summing from $i = 1, \ldots, N$, we get

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^N\mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N} - \bar{Y}_t^{i,\epsilon}|^2\right] \\
&\leq C\left\{e^{-\frac{\beta}{2\epsilon^2}t}\int_0^t e^{\frac{\beta}{2\epsilon^2}s}ds + \frac{1}{\epsilon^2}e^{-\frac{\beta}{2\epsilon^2}t}\int_0^t\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N\left|\tilde{X}_s^{i,\epsilon,N} - \bar{X}_s^{i,\epsilon}\right|^2\right]e^{\frac{\beta}{2\epsilon^2}s}ds\right. \\
&\quad \left. + \frac{1}{N\epsilon^2}e^{-\frac{\beta}{2\epsilon^2}t}\int_0^t e^{\frac{\beta}{2\epsilon^2}s}ds + \frac{1}{Na^2(N)}e^{-\frac{\beta}{2\epsilon^2}t}\int_0^t\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N|\tilde{u}_i^{N,2}(s)|^2\right]e^{\frac{\beta}{2\epsilon^2}s}ds\right\} \\
&\leq C\left\{\epsilon^2 + \sup_{s\in[0,t]}\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N\left|\tilde{X}_s^{i,\epsilon,N} - \bar{X}_s^{i,\epsilon}\right|^2\right]\right. \\
&\quad \left. + \frac{1}{N} + \frac{1}{Na^2(N)}\int_0^T\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N|\tilde{u}_i^{N,2}(s)|^2\right]ds\right\}
\end{aligned}$$

and by the bound (52), we get:

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^N\mathbb{E}\left[|\tilde{Y}_t^{i,\epsilon,N} - \bar{Y}_t^{i,\epsilon}|^2\right] \\
&\quad \leq C\left\{\epsilon^2 + \frac{1}{N} + \frac{1}{Na^2(N)} + \sup_{s\in[0,t]}\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N\left|\tilde{X}_s^{i,\epsilon,N} - \bar{X}_s^{i,\epsilon}\right|^2\right]\right\}.
\end{aligned}$$

□

Lemma 7.5 *Under assumptions (A1)–(A9) and (A12) we have*

$$\sup_{s\in[0,T]}\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\left|\tilde{X}_s^{i,\epsilon,N}-\bar{X}_s^{i,\epsilon}\right|^2\right]\leq C(T)[\epsilon^2+\frac{1}{N}+\frac{1}{Na^2(N)}]$$

Proof Letting $(\bar{i})$ denote the argument $(\bar{X}_s^{i,\epsilon},\bar{Y}_s^{i,\epsilon},\mathcal{L}(\bar{X}_s^{\epsilon}))$, $(\tilde{i})$ denote the argument $(\tilde{X}_s^{i,\epsilon,N},\tilde{Y}_s^{i,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})$:

$$\begin{aligned}\tilde{X}_t^{i,\epsilon,N}-\bar{X}_t^{i,\epsilon}&=\frac{1}{\epsilon}\int_0^t\left(b(\tilde{i})-b(\bar{i})\right)ds+\int_0^t\left(c(\tilde{i})-c(\bar{i})\right)ds\\&\quad+\int_0^t\sigma(\tilde{i})\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}}ds+\int_0^t\left(\sigma(\tilde{i})-\sigma(\bar{i})\right)dW_s^i\\&=\int_0^t\left(\gamma(\tilde{i})-\gamma(\bar{i})\right)ds+\int_0^t\Big([\sigma(\tilde{i})\\&\quad+\tau_1(\tilde{i})\Phi_y(\tilde{i})]-[\sigma(\bar{i})+\tau_1(\bar{i})\Phi_y(\bar{i})]\Big)dW_s^i\\&\quad+\int_0^t\left(\tau_2(\tilde{i})\Phi_y(\tilde{i})-\tau_2(\bar{i})\Phi_y(\bar{i})\right)dB_s^i\\&\quad+R_1^{i,\epsilon,N}(t)-R_2^{i,\epsilon,N}(t)+R_3^{i,\epsilon,N}(t)-R_4^{i,\epsilon,N}(t)+R_5^{i,\epsilon,N}(t),\end{aligned}$$

where here we recall Φ from Eq. (22) and γ,γ_1 from Eq. (23), and:

$$\begin{aligned}R_1^{i,\epsilon,N}(t)&=\frac{1}{\epsilon}\int_0^t b(\tilde{i})ds-\int_0^t\left(\gamma_1(\tilde{i})+[\frac{\tau_1(\tilde{i})}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s)+\frac{\tau_2(\tilde{i})}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)]\Phi_y(\tilde{i})\right)ds\\&\quad-\int_0^t\tau_1(\tilde{i})\Phi_y(\tilde{i})dW_s^i-\int_0^t\tau_2(\tilde{i})\Phi_y(\tilde{i})dB_s^i\\&\quad-\int_0^t\frac{1}{N}\sum_{j=1}^N b(\tilde{X}_s^{j,\epsilon,N},\tilde{Y}_s^{j,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})\partial_\mu\Phi(\tilde{i})[\tilde{X}_s^{j,\epsilon,N}]ds\\R_2^{i,\epsilon,N}(t)&=\frac{1}{\epsilon}\int_0^t b(\bar{i})ds-\int_0^t\gamma_1(\bar{i})ds-\int_0^t\tau_1(\bar{i})\Phi_y(\bar{i})dW_s^i-\int_0^t\tau_2(\bar{i})\Phi_y(\bar{i})dB_s^i\\&\quad-\int_0^t\int_{\mathbb{R}^2}b(x,y,\mathcal{L}(\bar{X}_s^\epsilon))\partial_\mu\Phi(\bar{X}_s^{i,\epsilon},\bar{Y}_s^{i,\epsilon},\mathcal{L}(\bar{X}_s^\epsilon))[x]\mathcal{L}(\bar{X}_s^\epsilon,Y_s^\epsilon)(dx,dy)ds\\R_3^{i,\epsilon,N}(t)&=\int_0^t\frac{1}{N}\sum_{j=1}^N b(\tilde{X}_s^{j,\epsilon,N},\tilde{Y}_s^{j,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})\partial_\mu\Phi(\tilde{i})[\tilde{X}_s^{j,\epsilon,N}]ds\\R_4^{i,\epsilon,N}(t)&=\int_0^t\int_{\mathbb{R}^2}b(x,y,\mathcal{L}(\bar{X}_s^\epsilon))\partial_\mu\Phi(\bar{X}_s^{i,\epsilon},\bar{Y}_s^{i,\epsilon},\mathcal{L}(\bar{X}_s^\epsilon))[x]\mathcal{L}(\bar{X}_s^\epsilon,Y_s^\epsilon)(dx,dy)ds\\R_5^{i,\epsilon,N}(t)&=\int_0^t\left(\sigma(\tilde{i})\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}}+[\frac{\tau_1(\tilde{i})}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s)+\frac{\tau_2(\tilde{i})}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)]\Phi_y(\tilde{i})\right)ds\end{aligned}$$

By Proposition 6.1 with $\psi \equiv 1$, we have

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|R_1^{i,\epsilon,N}(t)|^2\Big] \leq C(T)[\epsilon^2 + \frac{1}{N}],$$

by Proposition 4.2 of [5] with $\psi \equiv 1$, we have

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|R_2^{i,\epsilon,N}(t)|^2\Big] \leq C(T)\epsilon^2,$$

by Proposition 6.3 with $\psi \equiv 1$, we have

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|R_3^{i,\epsilon,N}(t)|^2\Big] \leq C(T)[\epsilon^2 + \frac{1}{N^2}],$$

and by Proposition 4.3 of [5] with $\psi \equiv 1$, we have

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|R_4^{i,\epsilon,N}(t)|^2\Big] \leq C(T)\epsilon^2.$$

Here we are using that, under the assumed regularity of Φ and χ imposed by Assumptions (A6) and (A7) respectively along with the result of Lemma 7.1, the norm can brought inside the expectation in Propositions 4.2 and 4.3 of [5] with little modification to the proofs. Finally, since under Assumption (A6) Φ_y is bounded:

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|R_5^{i,\epsilon,N}(t)|^2\Big] &\leq C\frac{1}{a^2(N)N}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[\int_0^T |\tilde{u}_i^{N,1}(s)|^2 + |\tilde{u}_i^{N,2}(s)|^2 ds\Big]\\ &\leq C\frac{1}{a^2(N)N}\end{aligned}$$

by the assumed bound on the controls (52). Now we see that, by Itô Isometry:

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|\tilde{X}_t^{i,\epsilon,N} - \bar{X}_t^{i,\epsilon}|^2\Big]\\ &\leq \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[\Big|\int_0^t \Big(\gamma(\tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N}) - \gamma(\bar{X}_s^{i,\epsilon}, \bar{Y}_s^{i,\epsilon}, \mathcal{L}(\bar{X}_s^{\epsilon}))\Big) ds\Big|^2\Big]\\ &\quad + \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[\int_0^t \Big|[\sigma + \tau_1\Phi_y](\tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N}) - [\sigma + \tau_1\Phi_y](\bar{X}_s^{i,\epsilon}, \bar{Y}_s^{i,\epsilon}, \mathcal{L}(\bar{X}_s^{\epsilon}))\Big|^2 ds\Big]\\ &\quad + \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[\int_0^t \Big|[\tau_2\Phi_y](\tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N}) - [\tau_2\Phi_y](\bar{X}_s^{i,\epsilon}, \bar{Y}_s^{i,\epsilon}, \mathcal{L}(\bar{X}_s^{\epsilon}))\Big|^2 ds\Big] + R^N(t)\end{aligned}$$

where $R^N(t) \le C(T)[\epsilon^2 + \frac{1}{N}]$. We handle the terms from the martingales first.

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^N \mathbb{E}\Big[\int_0^t \Big|[\sigma+\tau_1\Phi_y](\tilde{X}^{i,\epsilon,N}_s,\tilde{Y}^{i,\epsilon,N}_s,\tilde{\mu}^{\epsilon,N}_s) - [\sigma+\tau_1\Phi_y](\bar{X}^{i,\epsilon}_s,\bar{Y}^{i,\epsilon}_s,\mathcal{L}(\bar{X}^{\epsilon}_s))\Big|^2 ds\Big]\\
&\le \frac{C}{N}\sum_{i=1}^N\Big\{\mathbb{E}\Big[\int_0^t \Big|[\sigma+\tau_1\Phi_y](\tilde{X}^{i,\epsilon,N}_s,\tilde{Y}^{i,\epsilon,N}_s,\tilde{\mu}^{\epsilon,N}_s) - [\sigma+\tau_1\Phi_y](\bar{X}^{i,\epsilon}_s,\bar{Y}^{i,\epsilon}_s,\bar{\mu}^{\epsilon,N})\Big|^2 ds\Big]\\
&+\mathbb{E}\Big[\int_0^t \Big|[\sigma+\tau_1\Phi_y](\bar{X}^{i,\epsilon}_s,\bar{Y}^{i,\epsilon}_s,\bar{\mu}^{\epsilon,N}) - [\sigma+\tau_1\Phi_y](\bar{X}^{i,\epsilon}_s,\bar{Y}^{i,\epsilon}_s,\mathcal{L}(\bar{X}^{\epsilon}_s))\Big|^2 ds\Big]\Big\}\\
&\le \frac{C}{N}\sum_{i=1}^N \mathbb{E}\Big[\int_0^t |\tilde{X}^{i,\epsilon,N}_s-\bar{X}^{i,\epsilon}_s|^2 + |\tilde{Y}^{i,\epsilon,N}_s-\bar{Y}^{i,\epsilon}_s|^2 ds\Big] + \frac{C}{N}
\end{aligned}$$

by Lipschitz continuity from Assumption (A9) and Assumption (A12) with Lemma D.7

$$\le C\Big\{\epsilon^2 + \frac{1}{N} + \frac{1}{Na^2(N)} + \int_0^t \sup_{\tau\in[0,s]}\mathbb{E}\Big[\frac{1}{N}\sum_{i=1}^N\Big|\tilde{X}^{i,\epsilon,N}_\tau - \bar{X}^{i,\epsilon}_\tau\Big|^2\Big]ds\Big\}$$

by Lemma 7.4 .

The exact same proof and bound applies to

$$\frac{1}{N}\sum_{i=1}^N \mathbb{E}\Big[\int_0^t \Big|[\tau_2\Phi_y](\tilde{X}^{i,\epsilon,N}_s,\tilde{Y}^{i,\epsilon,N}_s,\tilde{\mu}^{\epsilon,N}_s) - [\tau_2\Phi_y](\bar{X}^{i,\epsilon}_s,\bar{Y}^{i,\epsilon}_s,\mathcal{L}(\bar{X}^{\epsilon}_s))\Big|^2 ds\Big].$$

In a similar manner:

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^N \mathbb{E}\Big[\Big|\int_0^t \Big(\gamma(\tilde{X}^{i,\epsilon,N}_s,\tilde{Y}^{i,\epsilon,N}_s,\tilde{\mu}^{\epsilon,N}_s) - \gamma(\bar{X}^{i,\epsilon}_s,\bar{Y}^{i,\epsilon}_s,\mathcal{L}(\bar{X}^{\epsilon}_s))\Big) ds\Big|^2\Big]\\
&\le \frac{C(T)}{N}\sum_{i=1}^N\Big\{\mathbb{E}\Big[\int_0^t \Big|\gamma(\tilde{X}^{i,\epsilon,N}_s,\tilde{Y}^{i,\epsilon,N}_s,\tilde{\mu}^{\epsilon,N}_s) - \gamma(\bar{X}^{i,\epsilon}_s,\bar{Y}^{i,\epsilon}_s,\bar{\mu}^{\epsilon,N})\Big|^2 ds\Big]\\
&+\mathbb{E}\Big[\int_0^t \Big|\gamma(\bar{X}^{i,\epsilon}_s,\bar{Y}^{i,\epsilon}_s,\bar{\mu}^{\epsilon,N}) - \gamma(\bar{X}^{i,\epsilon}_s,\bar{Y}^{i,\epsilon}_s,\mathcal{L}(\bar{X}^{\epsilon}_s))\Big|^2 ds\Big]\Big\}\\
&\le \frac{C(T)}{N}\sum_{i=1}^N \mathbb{E}\Big[\int_0^t |\tilde{X}^{i,\epsilon,N}_s-\bar{X}^{i,\epsilon}_s|^2 + |\tilde{Y}^{i,\epsilon,N}_s-\bar{Y}^{i,\epsilon}_s|^2 ds\Big] + \frac{C(T)}{N}\\
&\le C(T)\Big\{\epsilon^2 + \frac{1}{N} + \frac{1}{Na^2(N)} + \int_0^t \sup_{\tau\in[0,s]}\mathbb{E}\Big[\frac{1}{N}\sum_{i=1}^N\Big|\tilde{X}^{i,\epsilon,N}_\tau - \bar{X}^{i,\epsilon}_\tau\Big|^2\Big]ds\Big\}.
\end{aligned}$$

Collecting these bounds, we have for all $p \in [0, T]$:

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|\tilde{X}_p^{i,\epsilon,N} - \bar{X}_p^{i,\epsilon}|^2\Big] \le C_1(T)\int_0^p \sup_{\tau\in[0,s]}\mathbb{E}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\tilde{X}_\tau^{i,\epsilon,N} - \bar{X}_\tau^{i,\epsilon}\Big|^2\Big]ds + C_2(T)[\epsilon^2 + \frac{1}{N} + \frac{1}{Na^2(N)}]$$

so

$$\sup_{p\in[0,t]}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|\tilde{X}_p^{i,\epsilon,N} - \bar{X}_p^{i,\epsilon}|^2\Big] \le C(T)\Big[\int_0^t \sup_{\tau\in[0,s]}\mathbb{E}\Big[\frac{1}{N}\sum_{i=1}^{N}\Big|\tilde{X}_\tau^{i,\epsilon,N} - \bar{X}_\tau^{i,\epsilon}\Big|^2\Big]ds + [\epsilon^2 + \frac{1}{N} + \frac{1}{Na^2(N)}]\Big]$$

and by Gronwall's inequality:

$$\sup_{p\in[0,T]}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|\tilde{X}_p^{i,\epsilon,N} - \bar{X}_p^{i,\epsilon}|^2\Big] \le C(T)[\epsilon^2 + \frac{1}{N} + \frac{1}{Na^2(N)}].$$

□

7.3 Tightness of $\tilde{Z}^N$

We now have the tools to prove tightness of $\{\tilde{Z}^N\}$ from Eq. (54). We first prove a uniform-in-time bound on $\langle \tilde{Z}_t^N, \phi\rangle$ in terms of $|\cdot|_4$ (recall Eq. 8) in Lemma 7.6. Then, using the results from Sect. 6, we provide a prelimit representation for $\tilde{Z}^N$ which is a priori $\mathcal{O}(1)$ in ϵ, N in Lemma 7.7. Combining these two lemmas, we are then able to establish tightness via the methods of [55] in Proposition 7.8.

Lemma 7.6 *Under Assumptions (A1)–(A12), there exists C independent of N such that for all* $\phi \in C_{b,L}^4(\mathbb{R})$

$$\sup_{N\in\mathbb{N}}\sup_{t\in[0,T]}\mathbb{E}\Big[|\langle \tilde{Z}_t^N, \phi\rangle|^2\Big] \le C(T)|\phi|_4^2.$$

Proof Let $\bar{\mu}_t^{\epsilon,N}$ be as in Eq. (58). Then, by triangle inequality

$$\mathbb{E}\Big[|\langle \tilde{Z}_t^N, \phi\rangle|^2\Big] \le 2a^2(N)N\mathbb{E}\Big[|\langle \tilde{\mu}_t^{\epsilon,N}, \phi\rangle - \langle \bar{\mu}_t^{\epsilon,N}, \phi\rangle|^2\Big] + 2a^2(N)N\mathbb{E}\Big[|\langle \bar{\mu}_t^{\epsilon,N}, \phi\rangle - \langle \mathcal{L}(X_t), \phi\rangle|^2\Big].$$

For the first term:

$$
\begin{aligned}
&a^2(N)N\mathbb{E}\left[|\langle\tilde{\mu}_t^{\epsilon,N},\phi\rangle-\langle\bar{\mu}_t^{\epsilon,N},\phi\rangle|^2\right]\\
&=a^2(N)N\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^N\phi(\tilde{X}_t^{i,\epsilon,N})-\phi(\bar{X}_t^{i,\epsilon})\right|^2\right]\\
&\le a^2(N)N\frac{1}{N}\sum_{i=1}^N\mathbb{E}\left[\left|\phi(\tilde{X}_t^{i,\epsilon,N})-\phi(\bar{X}_t^{i,\epsilon})\right|^2\right]\\
&\le a^2(N)N\frac{1}{N}\sum_{i=1}^N\mathbb{E}\left[\left|\tilde{X}_t^{i,\epsilon,N}-\bar{X}_t^{i,\epsilon}\right|^2\right]\left\|\phi'\right\|_\infty^2\\
&\le C(T)[\epsilon^2a^2(N)N+a^2(N)+1]\left\|\phi'\right\|_\infty^2\\
&\le C(T)\left\|\phi'\right\|_\infty^2,
\end{aligned}
$$

where in the second to last inequality we used Lemma 7.5.

For the second term, we use that $\{\bar{X}^{i,\epsilon}\}_{i\in\mathbb{N}}$ are IID to see:

$$
\begin{aligned}
a^2(N)N\mathbb{E}\left[|\langle\bar{\mu}_t^{\epsilon,N},\phi\rangle-\langle\mathcal{L}(X_t),\phi\rangle|^2\right]&=a^2(N)N\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^N\phi(\bar{X}_t^{i,\epsilon})-\mathbb{E}[\phi(X_t)]\right|^2\right]\\
&=a^2(N)\Big\{(N-1)(\mathbb{E}[\phi(\bar{X}_t^{\epsilon}]-\mathbb{E}[\phi(X_t)])^2\\
&+\mathbb{E}\left[\left|\phi(\bar{X}_t^{\epsilon})-\mathbb{E}[\phi(X_t)]\right|^2\right]\Big\}\\
&\le a^2(N)\Big\{N(\mathbb{E}[\phi(\bar{X}_t^{\epsilon})]-\mathbb{E}[\phi(X_t)])^2+4\left\|\phi\right\|_\infty^2\Big\}\\
&\le a^2(N)N\epsilon^2C(T)|\phi|_4^2+4a^2(N)\left\|\phi\right\|_\infty^2,
\end{aligned}
$$

where in the last inequality we used Theorem 7.2. This bound vanishes as $N\to\infty$. □

Lemma 7.7 *Assume (A1)–(A8), (A10), and (A13). Define $\bar{L}_{\nu_1,\nu_2}$ to be the operator parameterized by $\nu_1,\nu_2\in\mathcal{P}_2(\mathbb{R})$ which acts on $\phi\in C_b^2(\mathbb{R})$ by*

$$
\begin{aligned}
\bar{L}_{\nu_1,\nu_2}\phi(x)&=\bar{\gamma}(x,\nu_2)\phi'(x)+\bar{D}(x,\nu_2)\phi''(x)\\
&\quad+\int_{\mathbb{R}}\int_0^1\frac{\delta}{\delta m}\bar{\gamma}(z,(1-r)\nu_1+r\nu_2)[x]\phi'(z)\\
&\quad+\frac{\delta}{\delta m}\bar{D}(z,(1-r)\nu_1+r\nu_2)[x]\phi''(z)dr\nu_1(dz).
\end{aligned}\tag{66}
$$

Here we recall $\bar{\gamma}$, $\bar{D}$ from Eq. (24), the linear functional derivative from Definition D.4, Φ from Eq. (22), and the occupation measures Q^N from Eq. (56). Then we have *the representation: for $\phi \in C_c^\infty(\mathbb{R})$ and $t \in [0, T]$:*

$$\langle \tilde{Z}_t^N, \phi \rangle = \int_0^t \langle \tilde{Z}_s^N, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}_s^{\epsilon,N}} \phi(\cdot) \rangle ds + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \sigma(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 \phi'(x) Q^N(dx, dy, dz, ds) + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 + \tau_2(x, y, \tilde{\mu}_s^{\epsilon,N}) z_2] \Phi_y(x, y, \tilde{\mu}_s^{\epsilon,N}) \phi'(x) Q^N(dx, dy, dz, ds) + R_t^N(\phi)$$

where

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left| R_t^N(\phi) \right| \right] \le \bar{R}(N, T) |\phi|_4,$$

$\bar{R}(N, T) \to 0$ as $N \to \infty$, and $\bar{R}(N, T)$ is independent of ϕ.

Proof Recall $\tilde{\mu}^{N,\epsilon}$ from Eq. (54), X_t from Eq. (25), and that $\tilde{Z}^N = a(N)\sqrt{N}[\tilde{\mu}^{N,\epsilon} - \mathcal{L}(X)]$.

By Itô's formula,

$$\langle \tilde{\mu}_t^{\epsilon,N}, \phi \rangle = \phi(x) + \frac{1}{N} \sum_{i=1}^N \int_0^t \left(\frac{1}{\epsilon} b(i) \phi'(\tilde{X}_s^{i,\epsilon,N}) + c(i) \phi'(\tilde{X}_s^{i,\epsilon,N}) + \frac{1}{2} \sigma^2(i) \phi''(\tilde{X}_s^{i,\epsilon,N}) + \sigma(i) \frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}} \phi'(\tilde{X}_s^{i,\epsilon,N}) \right) ds + \int_0^t \sigma(i) \phi'(\tilde{X}_s^{i,\epsilon,N}) dW_s^i$$

where (i) denotes the argument $(\tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N})$ and

$$\langle \mathcal{L}(X_s), \phi \rangle = \phi(x) + \mathbb{E}\left[\int_0^t \left(\bar{\gamma}(X_s, \mathcal{L}(X_s)) \phi'(X_s) + \bar{D}(X_s, \mathcal{L}(X_s)) \phi''(X_s) \right) ds + \int_0^t \sqrt{2} \bar{D}^{1/2}(X_s, \mathcal{L}(X_s) \phi'(X_s) dW_s \right]$$
$$= \phi(x) + \mathbb{E}\left[\int_0^t \left(\bar{\gamma}(X_s, \mathcal{L}(X_s)) \phi'(X_s) + \bar{D}(X_s, \mathcal{L}(X_s)) \phi''(X_s) \right) ds \right]$$

since $\bar{D}^{1/2}$ is bounded as per Assumption (A10). Then

$$\langle \tilde{Z}_t^N, \phi \rangle = \frac{a(N)}{\sqrt{N}} \sum_{i=1}^N \left\{ \int_0^t \left(\frac{1}{\epsilon} b(i) \phi'(\tilde{X}_s^{i,\epsilon,N}) + c(i) \phi'(\tilde{X}_s^{i,\epsilon,N}) \right. \right.$$

$$
\begin{aligned}
&+\frac{1}{2}\sigma^2(i)\phi''(\tilde{X}_s^{i,\epsilon,N})+\sigma(i)\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}}\phi'(\tilde{X}_s^{i,\epsilon,N})\Big)ds\\
&+\int_0^t\sigma(i)\phi'(\tilde{X}_s^{i,\epsilon,N})dW_s^i\\
&-\mathbb{E}\Big[\int_0^t\left(\bar{\gamma}(X_s,\mathcal{L}(X_s))\phi'(X_s)+\bar{D}(X_s,\mathcal{L}(X_s))\phi''(X_s)\right)ds\Big]\Big\}\\
=&\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N\Big\{\int_0^t\Big(\gamma(i)\phi'(\tilde{X}_s^{i,\epsilon,N})-\mathbb{E}\Big[\bar{\gamma}(X_s,\mathcal{L}(X_s))\phi'(X_s)\Big]\Big)ds\\
&+\int_0^t\Big(D(i)\phi''(\tilde{X}_s^{i,\epsilon,N})-\mathbb{E}\Big[\bar{D}(X_s,\mathcal{L}(X_s))\phi''(X_s)\Big]\Big)ds\\
&+R_1^i(t)+R_2^i(t)+M^i(t)\Big\}\\
&+\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]}\sigma(x,y,\tilde{\mu}_s^{\epsilon,N})z_1\phi'(x)Q^N(dx,dy,dz,ds)\\
&+\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]}\Big([\tau_1(x,y,\tilde{\mu}_s^{\epsilon,N})z_1+\tau_2(x,y,\tilde{\mu}_s^{\epsilon,N})z_2]\\
&\times\Phi_y(x,y,\tilde{\mu}_s^{\epsilon,N})\phi'(x)\Big)Q^N(dx,dy,dz,ds)
\end{aligned}
$$

where here we recall γ_1,γ,D,D_1 from Eq. (23) and:

$$
\begin{aligned}
R_1^i(t):=&\int_0^t\Big(\frac{1}{\epsilon}b(i)\phi'(\tilde{X}_s^{i,\epsilon,N})-\int_0^t\gamma_1(i)\phi'(\tilde{X}_s^{i,\epsilon,N})+D_1(i)\phi''(\tilde{X}_s^{i,\epsilon,N})]\\
&+\Big[\frac{\tau_1(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s)+\frac{\tau_2(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)\Big]\Phi_y(i)\phi'(\tilde{X}_s^{i,\epsilon,N})\Big)ds\\
&-\int_0^t\tau_1(i)\Phi_y(i)\phi'(\tilde{X}_s^{i,\epsilon,N})dW_s^i-\int_0^t\tau_2(i)\Phi_y(i)\phi'(\tilde{X}_s^{i,\epsilon,N})dB_s^i\\
&-\int_0^t\frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Phi(i)[j]\phi'(\tilde{X}_s^{i,\epsilon,N})ds\\
R_2^i(t):=&\int_0^t\frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Phi(i)[j]\phi'(\tilde{X}_s^{i,\epsilon,N})ds\\
M^i(t):=&\int_0^t[\tau_1(i)\Phi_y(i)+\sigma(i)]\phi'(\tilde{X}_s^{i,\epsilon,N})dW_s^i+\int_0^t\tau_2(i)\Phi_y(i)\phi'(\tilde{X}_s^{i,\epsilon,N})dB_s^i.
\end{aligned}
$$

For $R_1^i(t)$, we have via Proposition 6.1 that

$$
\mathbb{E}\Big[\sup_{t\in[0,T]}\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N|R_1^i(t)|\Big]\le C[\epsilon a(N)\sqrt{N}(1+T+T^{1/2})+a(N)T]|\phi|_3.
$$

For $R_2^i(t)$, we have via Proposition 6.3 that

$$\mathbb{E}\Bigg[\sup_{t\in[0,T]}\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N|R_2^i(t)|\Bigg]\\ \leq C[\epsilon a(N)\sqrt{N}(1+T+T^{1/2})+\frac{a(N)}{\sqrt{N}}T]|\phi|_3.$$

For $M^i(t)$, we have by Burkholder-Davis-Gundy inequality that

$$\begin{aligned}&\mathbb{E}\Bigg[\sup_{t\in[0,T]}\Bigg|\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N M^i(t)\Bigg|\Bigg]\\ &\leq C\frac{a(N)}{\sqrt{N}}\Bigg\{\mathbb{E}\Bigg[\Bigg(\sum_{i=1}^N\int_0^T[\tau_1(i)\Phi_y(i)+\sigma(i)]^2|\phi'(i)|^2ds\Bigg)^{1/2}\Bigg]\\ &\quad+\mathbb{E}\Bigg[\Bigg(\sum_{i=1}^N\int_0^T|\tau_2(i)\Phi_y(i)|^2|\phi'(i)|^2ds\Bigg)^{1/2}\Bigg]\Bigg\}\\ &\leq Ca(N)T^{1/2}|\phi|_1\end{aligned}$$

since the integrand it bounded by Assumptions (A1),(A5), and (A6). Thus we have

$$\begin{aligned}\langle\tilde{Z}_t^N,\phi\rangle&=\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N\Bigg\{\int_0^t\Bigg(\gamma(i)\phi'(\tilde{X}_s^{i,\epsilon,N})-\mathbb{E}\Big[\bar{\gamma}(X_s,\mathcal{L}(X_s))\phi'(X_s)\Big]\Bigg)ds\\ &\quad+\int_0^t\Bigg(D(i)\phi''(\tilde{X}_s^{i,\epsilon,N})-\mathbb{E}\Big[\bar{D}(X_s,\mathcal{L}(X_s))\phi''(X_s)\Big]\Bigg)ds\Bigg\}\\ &\quad+\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]}\sigma(x,y,\tilde{\mu}_s^{\epsilon,N})z_1\phi'(x)Q^N(dx,dy,dz,ds)\\ &\quad+\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]}[\tau_1(x,y,\tilde{\mu}_s^{\epsilon,N})z_1+\tau_2(x,y,\tilde{\mu}_s^{\epsilon,N})z_2]\\ &\quad\times\Phi_y(x,y,\tilde{\mu}_s^{\epsilon,N})\phi'(x)Q^N(dx,dy,dz,ds)+R_t^{3,N}(\phi)\end{aligned}$$

where $\mathbb{E}\Big[\sup_{t\in[0,T]}|R_t^{3,N}(\phi)|\Big]\leq C(T)\epsilon a(N)\sqrt{N}\vee a(N)|\phi|_3$. We rewrite this as:

$$\begin{aligned}\langle\tilde{Z}_t^N,\phi\rangle=\frac{a(N)}{\sqrt{N}}\sum_{i=1}^N\Bigg\{\int_0^t\Bigg(&\bar{\gamma}(\tilde{X}_s^{i,\epsilon,N},\tilde{\mu}_s^{i,\epsilon,N})\phi'(\tilde{X}_s^{i,\epsilon,N})\\ &-\mathbb{E}\Big[\bar{\gamma}(X_s,\mathcal{L}(X_s))\phi'(X_s)\Big]\Bigg)ds\end{aligned}$$

$$+ \int_0^t \left(\bar{D}(\tilde{X}_s^{i,\epsilon,N}, \tilde{\mu}_s^{i,\epsilon,N})\phi''(\tilde{X}_s^{i,\epsilon,N}) - \mathbb{E}\left[\bar{D}(X_s, \mathcal{L}(X_s))\phi''(X_s) \right] \right) ds$$

$$+ R_4^i(t) + R_5^i(t) \Bigg\}$$

$$+ \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \sigma(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 \phi'(x) Q^N(dx, dy, dz, ds)$$

$$+ \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 + \tau_2(x, y, \tilde{\mu}_s^{\epsilon,N}) z_2]$$

$$\Phi_y(x, y, \tilde{\mu}_s^{\epsilon,N})\phi'(x) Q^N(dx, dy, dz, ds) + R_t^{3,N}(\phi)$$

where

$$R_4^i(t) = \int_0^t [\gamma(\tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N}) - \bar{\gamma}(\tilde{X}_s^{i,\epsilon,N}, \tilde{\mu}_s^{N,\epsilon})]\phi'(\tilde{X}_s^{i,\epsilon,N}) ds$$

$$R_5^i(t) = \int_0^t [D(\tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N}) - \bar{D}(\tilde{X}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N})]\phi''(\tilde{X}_s^{i,\epsilon,N}) ds.$$

By Proposition 6.4 (using here Assumption (A8)):

$$\mathbb{E}\left[\sup_{t\in[0,T]} \frac{a(N)}{\sqrt{N}} \sum_{i=1}^N |R_4^i(t)| \right] \le C\epsilon a(N)\sqrt{N}(1 + T^{1/2} + T)|\phi|_3$$

and

$$\mathbb{E}\left[\sup_{t\in[0,T]} \frac{a(N)}{\sqrt{N}} \sum_{i=1}^N |R_5^i(t)| \right] \le C\epsilon a(N)\sqrt{N}(1 + T^{1/2} + T)|\phi|_4.$$

Now, we arrive at

$$\langle \tilde{Z}_t^N, \phi \rangle = \frac{a(N)}{\sqrt{N}} \sum_{i=1}^N \Bigg\{ \int_0^t \left(\bar{\gamma}(\tilde{X}_s^{i,\epsilon,N}, \tilde{\mu}_s^{i,\epsilon,N})\phi'(\tilde{X}_s^{i,\epsilon,N}) - \mathbb{E}\left[\bar{\gamma}(X_s, \mathcal{L}(X_s))\phi'(X_s) \right] \right) ds$$

$$+ \int_0^t \left(\bar{D}(\tilde{X}_s^{i,\epsilon,N}, \tilde{\mu}_s^{i,\epsilon,N})\phi''(\tilde{X}_s^{i,\epsilon,N}) - \mathbb{E}\left[\bar{D}(X_s, \mathcal{L}(X_s))\phi''(X_s) \right] \right) ds \Bigg\}$$

$$+ \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \sigma(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 \phi'(x) Q^N(dx, dy, dz, ds)$$

$$+ \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 + \tau_2(x, y, \tilde{\mu}_s^{\epsilon,N}) z_2]$$

$$\Phi_y(x, y, \tilde{\mu}_s^{\epsilon,N})\phi'(x) Q^N(dx, dy, dz, ds) + R_t^N(\phi)$$

where $\mathbb{E}\left[\sup_{t\in[0,T]} |R_t^N(\phi)| \right] \le C(T)[\epsilon a(N)\sqrt{N} \vee a(N)]|\phi|_4$. By Assumption (A13), $\bar{\gamma}$ and $\bar{D}$ have well-defined linear functional derivatives (see Definition D.4).

Then we can rewrite $\langle \tilde{Z}_t^N, \phi \rangle$ as

$$
\begin{aligned}
\langle \tilde{Z}_t^N, \phi \rangle &= \int_0^t \langle \tilde{Z}_s^N, \bar{\gamma}(\cdot, \tilde{\mu}_s^{\epsilon,N})\phi'(\cdot) + \bar{D}(\cdot, \tilde{\mu}_s^{\epsilon,N})\phi''(\cdot)\rangle ds \\
&\quad + a(N)\sqrt{N}\Big[\int_0^t \langle \mathcal{L}(X_s), [\bar{\gamma}(\cdot, \tilde{\mu}_s^{\epsilon,N}) - \bar{\gamma}(\cdot, \mathcal{L}(X_s))]\phi'(\cdot) \\
&\quad + [\bar{D}(\cdot, \tilde{\mu}_s^{\epsilon,N}) - \bar{D}(\cdot, \mathcal{L}(X_s))]\phi''(\cdot)\rangle ds\Big] \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \sigma(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 \phi'(x) Q^N(dx, dy, dz, ds) \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 + \tau_2(x, y, \tilde{\mu}_s^{\epsilon,N}) z_2]\Phi_y(x, y, \tilde{\mu}_s^{\epsilon,N}) \\
&\quad \phi'(x) Q^N(dx, dy, dz, ds) + R_t^N(\phi) \\
&= \int_0^t \langle \tilde{Z}_s^N, \bar{\gamma}(\cdot, \tilde{\mu}_s^{\epsilon,N})\phi'(\cdot) + \bar{D}(\cdot, \tilde{\mu}_s^{\epsilon,N})\phi''(\cdot)\rangle ds \\
&\quad + a(N)\sqrt{N}\Big[\int_0^t \langle \mathcal{L}(X_s), \Big[\int_0^1 \int_{\mathbb{R}} \frac{\delta}{\delta m}\bar{\gamma}(\cdot, (1-r)\mathcal{L}(X_s) + r\tilde{\mu}_s^{\epsilon,N}) \\
&\quad [y](\tilde{\mu}_s^{N,\epsilon}(dy) - \mathcal{L}(X_s)(dy)) dr\Big]\phi'(\cdot) \\
&\quad + \Big[\int_0^1 \int_{\mathbb{R}} \frac{\delta}{\delta m}\bar{D}(\cdot, (1-r)\mathcal{L}(X_s) + r\tilde{\mu}_s^{\epsilon,N})[y](\tilde{\mu}_s^{N,\epsilon}(dy) \\
&\quad - \mathcal{L}(X_s)(dy)) dr\Big]\phi''(\cdot)\rangle ds\Big] \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \sigma(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 \phi'(x) Q^N(dx, dy, dz, ds) \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 + \tau_2(x, y, \tilde{\mu}_s^{\epsilon,N}) z_2]\Phi_y(x, y, \tilde{\mu}_s^{\epsilon,N}) \\
&\quad \phi'(x) Q^N(dx, dy, dz, ds) + R_t^N(\phi) \\
&= \int_0^t \langle \tilde{Z}_s^N, \bar{L}_{\mathcal{L}(X_s), \tilde{\mu}_s^{\epsilon,N}} \phi(\cdot)\rangle ds + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \\
&\quad \sigma(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 \phi'(x) Q^N(dx, dy, dz, ds) \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x, y, \tilde{\mu}_s^{\epsilon,N}) z_1 + \tau_2(x, y, \tilde{\mu}_s^{\epsilon,N}) z_2]\Phi_y(x, y, \tilde{\mu}_s^{\epsilon,N}) \\
&\quad \phi'(x) Q^N(dx, dy, dz, ds) + R_t^N(\phi).
\end{aligned}
$$

□

Proposition 7.8 *Under Assumptions (A1)–(A13),* $\{\tilde{Z}^N\}_{N\in\mathbb{N}}$ *is tight as a sequence of* $C([0, T]; \mathcal{S}_{-m})$*-valued random variables, where m is as in Eq.* (6).

Proof By Remark R.1 on p.997 of [55], it suffices to prove tightness of $\{\langle \tilde{Z}^N, \phi \rangle\}$ as a sequence of $C([0, T]; \mathbb{R})$-valued random variables for each $\phi \in \mathcal{S}$, along with uniform 7-continuity as defined in the same remark. By the argument found in the proof of [10] Theorem 4.7, to show the latter it suffices to prove:

$$\sup_{N\in\mathbb{N}} \mathbb{E}\left[\sup_{t\in[0,T]}\left|\langle \tilde{Z}_t^N, \phi\rangle\right|\right] \leq C(T)\, \|\phi\|_7\,, \forall \phi \in \mathcal{S}. \tag{67}$$

After these two results are shown, we will have established tightness of $\tilde{Z}^N$ as $C([0,T];\mathcal{S}_{-w})$-valued random variables for $m > 7$ such that the canonical embedding $\mathcal{S}_{-7} \to \mathcal{S}_{-m}$ is Hilbert–Schmidt. We start with showing tightness of $\{\langle \tilde{Z}^N, \phi\rangle\}$. By Lemma 7.7, we write for any $\phi \in \mathcal{S}$:

$$\begin{aligned}
\langle \tilde{Z}_t^N, \phi\rangle &= A_t^N(\phi) + R_t^N(\phi)\\
A_t^N(\phi) &:= \int_0^t \langle \tilde{Z}_s^N, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}_s^{\epsilon,N}}\phi(\cdot)\rangle ds\\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \sigma(x,y,\tilde{\mu}_s^{\epsilon,N})z_1\phi'(x)Q^N(dx,dy,dz,ds)\\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x,y,\tilde{\mu}_s^{\epsilon,N})z_1 + \tau_2(x,y,\tilde{\mu}_s^{\epsilon,N})z_2]\\
&\quad \Phi_y(x,y,\tilde{\mu}_s^{\epsilon,N})\phi'(x)Q^N(dx,dy,dz,ds)
\end{aligned}$$

where for each ϕ, $R^N(\phi) \to 0$ in $C([0,T];\mathbb{R})$ as $N \to \infty$. Thus, to prove tightness of $\{\langle \tilde{Z}^N, \phi\rangle\}$, it is sufficient to prove tightness of $\{A^N(\phi)\}$. We note that for any and $0 \leq \tau < t \leq T$:

$$\begin{aligned}
A_t^N(\phi) - A_\tau^N(\phi) &= B_{t,\tau}^N(\phi) + C_{t,\tau}^N(\phi) + D_{t,\tau}^N(\phi)\\
B_{t,\tau}^N(\phi) &:= \int_\tau^t \langle \tilde{Z}_s^N, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}_s^{\epsilon,N}}\phi(\cdot)\rangle ds\\
C_{t,\tau}^N(\phi) &= \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[\tau,t]} \sigma(x,y,\tilde{\mu}_s^{\epsilon,N})z_1\phi'(x)Q^N(dx,dy,dz,ds)\\
D_{t,\tau}^N(\phi) &= \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[\tau,t]} [\tau_1(x,y,\tilde{\mu}_s^{\epsilon,N})z_1 + \tau_2(x,y,\tilde{\mu}_s^{\epsilon,N})z_2]\\
&\quad \times \Phi_y(x,y,\tilde{\mu}_s^{\epsilon,N})\phi'(x)Q^N(dx,dy,dz,ds).
\end{aligned}$$

Then we have for $\delta > 0$,

$$\begin{aligned}
\mathbb{E}\left[\sup_{|t-\tau|\leq\delta}\left|B_{t,\tau}^N(\phi)\right|\right] &\leq \delta^{1/2}\mathbb{E}\left[\int_0^T \left|\langle \tilde{Z}_s^N, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}_s^{\epsilon,N}}\phi(\cdot)\rangle\right|^2 ds\right]^{1/2}\\
&\leq \delta^{1/2}T^{1/2}\sup_{s\in[0,T]}\mathbb{E}\left[\left|\langle \tilde{Z}_s^N, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}_s^{\epsilon,N}}\phi(\cdot)\rangle\right|^2\right]^{1/2}\\
&\leq \delta^{1/2}T^{1/2}\sup_{s\in[0,T]}\sup_{\nu_1,\nu_2\in\mathcal{P}_2(\mathbb{R})}\mathbb{E}\left[\left|\langle \tilde{Z}_s^N, \bar{L}_{\nu_1,\nu_2}\phi(\cdot)\rangle\right|^2\right]^{1/2}\\
&\leq C(T)\delta^{1/2}|\phi|_6
\end{aligned}$$

by boundedness of the first 5 derivatives in x of $\bar{\gamma}$, $\bar{D}$, and of the first 5 derivatives in z of $\frac{\delta}{\delta m}\bar{\gamma}$, $\frac{\delta}{\delta m}\bar{D}$ from Assumption (A13), the definition of $\bar{L}_{\nu_1,\nu_2}$ from Eq. (66), and Lemma 7.6. Also, we see:

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{|t-\tau|\le\delta}\left|D^N_{t,\tau}(\phi)\right|\right]\\
&\le C\frac{1}{N}\sum_{i=1}^N\mathbb{E}\left[\sup_{|t-\tau|\le\delta}\int_\tau^t\left([|\tilde{u}^{N,1}_i(s)|+|\tilde{u}^{N,2}_i(s)|]|\Phi_y(\tilde{X}^{i,\epsilon,N}_s,\tilde{Y}^{i,\epsilon,N}_s,\tilde{\mu}^{\epsilon,N}_s)|\right)ds\right]|\phi|_1\\
&\le C\frac{1}{N}\mathbb{E}\left[\sum_{i=1}^N\int_0^T|\tilde{u}^{N,1}_i(s)|^2+|\tilde{u}^{N,2}_i(s)|^2ds\right]^{1/2}\\
&\mathbb{E}\left[\sup_{|t-\tau|\le\delta}\sum_{i=1}^N\int_\tau^t|\Phi_y(\tilde{X}^{i,\epsilon,N}_s,\tilde{Y}^{i,\epsilon,N}_s,\tilde{\mu}^{\epsilon,N}_s)|^2ds\right]^{1/2}|\phi|_1\\
&\le C\mathbb{E}\left[\sup_{|t-\tau|\le\delta}\frac{1}{N}\sum_{i=1}^N\int_\tau^t|\Phi_y(\tilde{X}^{i,\epsilon,N}_s,\tilde{Y}^{i,\epsilon,N}_s,\tilde{\mu}^{\epsilon,N}_s)|^2ds\right]^{1/2}|\phi|_1 \text{ by the bound(52)}\\
&\le C\delta^{1/2}C(T)|\phi|_1 \text{ by the boundedness of}\Phi_y\text{from Assumption (A6).}
\end{aligned}
$$

The proof that $\mathbb{E}\left[\left|B^N_{t,\tau}(\phi)\right|\right]\le C\delta^{1/2}C(T)|\phi|_1$ holds in the same way.

So by the Arzelà-Ascoli tightness criterion on classical Wiener space (see, e.g. Theorem 4.10 in Chapter 2 of [46]), we have $\{A^N(\phi)\}$ and hence $\{\langle\tilde{Z}^N,\phi\rangle\}$ are tight as a sequence of $C([0,T];\mathbb{R})$-valued random variables for each ϕ.

Now we see by the same argument (fixing $\tau=0$) and the fact that, as shown in Lemma 7.7,
$\mathbb{E}\left[\sup_{t\in[0,T]}\left|R^N_t(\phi)\right|\right]\le\bar{R}(N,T)|\phi|_4$ with $\bar{R}(N,T)\to 0$ as $N\to\infty$:

$$
\sup_{N\in\mathbb{N}}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\langle\tilde{Z}^N_t,\phi\rangle\right|\right]\le C(T)|\phi|_6\le C(T)\,\|\phi\|_7
$$

for all $\phi\in\mathcal{S}$, where here we used the inequality (9). Thus the bound (67) holds, and tightness is established. □

7.4 Tightness of Q^N

The proof of tightness of $\{Q^N\}$ from Eq. (56) is standard, see [4, 9, 56]. We see that since the occupation measures Q^N involve $\{\tilde{X}^{i,\epsilon,N}\}_{N\in\mathbb{N}}$ from Eq. (55) as part of their definition, we will need some kind of uniform control on their expectation. Thus, we begin with a lemma:

Lemma 7.9 *Under assumptions (A1)–(A7) and (A9), we have* $\sup_{t\in[0,T]}\sup_{N\in\mathbb{N}}\frac{1}{N}\sum_{i=1}^N\mathbb{E}\left[|\tilde{X}^{i,\epsilon,N}_t|^2\right]<\infty$.

Proof Using that

$$
\begin{aligned}
\tilde{X}_t^{i,\epsilon,N} &= \eta^x + \int_0^t \left(\frac{1}{\epsilon}b(i) + c(i)\right) ds + \int_0^t \sigma(i)\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}}ds + \int_0^t \sigma(i)dW_s^i \\
&= \eta^x + \int_0^t \frac{1}{\epsilon}b(i)ds - \left\{\int_0^t \gamma_1(i)ds + \int_0^t \tau_1(i)\Phi_y(i)dW_s^i + \int_0^t \tau_2(i)\Phi_y(i)dB_s^i \right.\\
&\quad + \int_0^t \frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Phi(i)[j]ds \\
&\quad \left. + \int_0^t [\frac{\tau_1(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s) + \frac{\tau_2(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)]\Phi_y(i)ds\right\} \\
&\quad + \int_0^t \gamma(i)ds + \int_0^t \sigma(i)dW_s^i \\
&\quad + \int_0^t \tau_1(i)\Phi_y(i)dW_s^i + \int_0^t \tau_2(i)\Phi_y(i)dB_s^i + \int_0^t \frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Phi(i)[j]ds \\
&\quad + \int_0^t \left(\sigma(i)\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}} + [\frac{\tau_1(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s) + \frac{\tau_2(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)]\Phi_y(i)\right) ds,
\end{aligned}
$$

where here once again the argument (i) is denoting $(\tilde{X}_s^{i,\epsilon,N}, \tilde{Y}_s^{i,\epsilon,N}, \tilde{\mu}_s^{\epsilon,N})$ and similarly for j, and we recall Φ from Eq. (22) and γ_1, γ from Eq. (23).

So, by Itô Isometry and boundedness of σ from (A5), of τ_1 and τ_2 from (A1), and of Φ_y from (A6):

$$
\begin{aligned}
&\frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[|\tilde{X}_t^{i,\epsilon,N}|^2\right] \\
&\quad \le C(|\eta^x|^2 + T) + \frac{C}{N}\sum_{i=1}^N \left\{R_1^{i,N}(t) + R_2^{i,N}(t) + R_3^{i,N}(t) + R_4^{i,N}(t)\right\} \\
&R_1^{i,N}(t) := \mathbb{E}\left[\left|\int_0^t \frac{1}{\epsilon}b(i)ds - \left\{\int_0^t \gamma_1(i)ds + \int_0^t \tau_1(i)\Phi_y(i)dW_s^i + \int_0^t \tau_2(i)\Phi_y(i)dB_s^i \right.\right.\right.\\
&\qquad \left.\left.\left. + \int_0^t \frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Phi(i)[j]ds + \int_0^t [\frac{\tau_1(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s) + \frac{\tau_2(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)]\Phi_y(i)ds\right\}\right|^2\right] \\
&R_2^{i,N}(t) := \mathbb{E}\left[\left|\int_0^t \gamma(i)ds\right|^2\right] \\
&R_3^{i,N}(t) = \mathbb{E}\left[\left|\int_0^t \frac{1}{N}\sum_{j=1}^N b(j)\partial_\mu\Phi(i)[j]ds\right|^2\right] \\
&R_4^{i,N}(t) = \mathbb{E}\left[\left|\int_0^t \left(\sigma(i)\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}}\right.\right.\right.\\
&\qquad \left.\left.\left. + \left[\frac{\tau_1(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,1}(s) + \frac{\tau_2(i)}{a(N)\sqrt{N}}\tilde{u}_i^{N,2}(s)\right]\Phi_y(i)ds\right)\right|^2\right]
\end{aligned}
$$

Then applying Proposition 6.1 with $\psi = 1$, we have

$$\frac{1}{N}\sum_{i=1}^{N} R_1^{i,N}(t) \le C[\epsilon^2(1+T+T^2) + \frac{1}{N}T^2]$$

Using Assumption (A9):

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N} R_2^{i,N}(t) &\le \frac{1}{N}\sum_{i=1}^{N} T\mathbb{E}\left[\int_0^t |\gamma(i)|^2 ds\right]\\ &\le CT\int_0^t \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[|\tilde{X}_s^{i,\epsilon,N}|^2 + |\tilde{Y}_s^{i,\epsilon,N}|^2 + \frac{1}{N}\sum_{j=1}^{N}|\tilde{X}_s^{j,\epsilon,N}|^2\right]ds\\ &\le CT^2 + CT\int_0^t \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[|\tilde{X}_s^{i,\epsilon,N}|^2\right]ds\end{aligned}$$

by Lemma B.1. Applying Proposition 6.3 with $\psi = 1$:

$$\frac{1}{N}\sum_{i=1}^{N} R_3^{i,N}(t) \le C[\epsilon^2(1+T+T^2) + \frac{1}{N^2}T^2]$$

Using the boundedness of σ from (A5), of τ_1 and τ_2 from (A1), and of Φ_y from (A6) and the bound (53):

$$\frac{1}{N}\sum_{i=1}^{N} R_4^{i,N}(t) \le \frac{CT}{a^2(N)N}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\int_0^T \left(|\tilde{u}_i^{N,1}(s)|^2 + |\tilde{u}_i^{N,2}(s)|^2\right) ds\right] \le \frac{CT}{a^2(N)N}.$$

Then, by Gronwall's inequality:

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[|\tilde{X}_t^{i,\epsilon,N}|^2\right] \le C(T)[1+\epsilon^2 + \frac{1}{N} + \frac{1}{N^2} + \frac{1}{a^2(N)N}] \le C(T),$$

since all the above terms which depend on N, ϵ in the first bound vanish as $N \to \infty$. Since this holds uniformly in N and t, we are done. □

Now we can prove tightness of the occupation measures.

Proposition 7.10 *Under assumptions (A1)–(A7) and (A9),* $\{Q^N\}_{N\in\mathbb{N}}$ *is tight as a sequence of* $M_T(\mathbb{R}^4)$*-valued random variables (recall this space of measures introduced above Eq.* 10).

Proof Consider the function $G : \mathcal{P}(\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]) \to \mathbb{R}$ given by

$$G(\theta) = \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]} \left(|z|^2 + |y|^2 + |x|^2\right)\theta(dx, dy, dz, ds).$$

Then we have G is bounded below, and considering a given level set $A_L := \{\theta \in M_T(\mathbb{R}^4) : G(\theta) \leq L\}$, it follows by Chebyshev's inequality that $\sup_{\theta\in A_L} \theta((K_L^\epsilon)^c) \leq \epsilon$ where K_L^ϵ is the compact subset of $\mathbb{R}^4 \times [0, T]$

$$K_L^\epsilon := \{(x, y, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2 : |x|^2 + |y|^2 + |z|^2 \leq \frac{L}{\epsilon}\} \times [0, T].$$

We also see that any collection of measures on $\mathbb{R}^4 \times [0, T]$ which is in $M_T(\mathbb{R}^4)$ is uniformly bounded in the total variation norm, and that for $\{\theta^N\} \subset A_L$ such that $\theta^N \to \theta$ in $M_T(\mathbb{R}^4)$ (recalling here that we are using the topology of weak convergence), by a version of Fatou's lemma (see Theorem A.3.12 in [22])

$$G(\theta) \leq \liminf_{N\to\infty} G(\theta^N) \leq L,$$

so $\theta \in A_L$. Via Prokhorov's theorem, A_L is precompact, and we have shown that A_L is closed, and hence G has compact level sets. Thus G is a tightness function (see [22] p.309), and it suffices to prove

$$\begin{aligned}
&\sup_{N\in\mathbb{N}} \mathbb{E}\left[G(Q^N)\right] \\
&\quad= \sup_{N\in\mathbb{N}} \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\int_0^T \left(|\tilde{X}_s^{i,\epsilon,N}|^2 + |\tilde{Y}_s^{i,\epsilon,N}|^2 + |\tilde{u}_i^{N,1}(s)|^2 + |\tilde{u}_i^{N,2}(s)|^2\right) ds\right] \\
&\quad< \infty
\end{aligned}$$

to see that $\{Q^N\}$ is a tight sequence of $\mathcal{M}_R(\mathbb{R}^4)-$valued random variables. This follows immediately from the bound (52) and Lemmas B.1 and 7.9. □

8 Identification of the limit

Now having established tightness of $\{(\tilde{Z}^N, Q^N)\}_{\mathbb{N}}$, we take any sub-sequence that converges in distribution as $C([0, T]; \mathcal{S}_{-m}) \times M_T(\mathbb{R}^4)$-valued random variables, and call the random variable which is its limit (Z, Q). We will show that $Q \in P^*(Z)$, and that this uniquely characterizes the distribution of (Z, Q) for a given choice of controls in the construction of Q^N. We will at times apply Skorokhod's Representation Theorem to without loss of generality pose the problem on a probability space such that this subsequence converges to (Z, Q) almost surely. We also do not distinguish from the subsequence and the original sequence in the notation, nor the original probability space and that invoked by Skorokhod's Representation Theorem. We begin with two lemmas which allow us to identify convergence of the controlled empirical measure $\tilde{\mu}^N$ from (54) to the law of the averaged McKean–Vlasov Eq. (25):

Lemma 8.1 *In the setting of Proposition 7.9, we have for any $p \geq 1$:*

$$\sup_{\epsilon>0}\sup_{t\in[0,T]}\mathbb{E}\left[|\bar{X}^\epsilon_t|^p\right] \leq |\eta^x|^p + C(T,p)[1+|\eta^y|^p].$$

Here $\bar{X}^\epsilon$ is as in Eq. (65). That is, it is equal in distribution to the IID particles from Eq. (57).

Proof This follows in the same way as Lemmas 7.5 and 7.9, using Lemma 7.1 and the ergodic-type Theorems from Section 4 of [5]. We omit the proof for brevity. □

Lemma 8.2 *Assume (A1)–(A12). Let $\tilde{\mu}^{\epsilon,N}_t$ be as in Eq. (54), with controls satisfying (53). Then*

$$\mathbb{E}\left[\mathbb{W}_2(\tilde{\mu}^{\epsilon,N}_t,\mathcal{L}(X_t))\right] \to 0 \text{ as } N\to\infty, \forall t\in[0,T],$$

where X_t is as in Eq. (25). In particular, decomposing Q^N from Eq. (56) as $Q^N(dx,dy,dz,dt) = Q^N_t(dx,dy,dz)dt$, for any $t\in[0,T]$, the first marginal of Q^N_t converges to $\mathcal{L}(X_t)$ in probability as a sequence of $\mathcal{P}_2(\mathbb{R})$-valued random variables.

Proof Firstly, we note by Theorem 7.2, $\mathcal{L}(\bar{X}^\epsilon_t)\to\mathcal{L}(X_t)$ in $\mathcal{P}(\mathbb{R})$ (using here that $C^\infty_c(\mathbb{R})$ is convergence determining - see [25] Proposition 3,4.4). In addition, by Lemma 8.1, we have $\sup_{\epsilon>0}\int_{\mathbb{R}}|x|^p\mathcal{L}(\bar{X}^\epsilon_t)(dx) < \infty$, for some $p>2$. Thus, we have by uniform integrability, $\mathbb{E}\left[|\bar{X}^\epsilon_t|^2\right]\to\mathbb{E}\left[|X_t|^2\right]$ as $\epsilon\downarrow 0$, so $\mathbb{W}_2(\mathcal{L}(\bar{X}^\epsilon_t),\mathcal{L}(X_t))\to 0$ as $\epsilon\downarrow 0$ (Theorem 5.5 in [13]). By Lemma 7.5, we also have

$$\mathbb{E}\left[\mathbb{W}_2(\tilde{\mu}^{\epsilon,N}_t,\bar{\mu}^{\epsilon,N}_t)\right] \leq \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N\left|\tilde{X}^{i,\epsilon,N}_t - \bar{X}^{i,\epsilon}_t\right|^2\right] \to 0 \text{ as } N\to\infty,$$

where $\bar{\mu}^{\epsilon,N}$ is as in Eq. (58). Also, by Glivenko-Cantelli Convergence in the Wasserstein Distance (see, e.g. Section 5.1.2 in [13]):

$$\mathbb{E}\left[\mathbb{W}_2(\bar{\mu}^{\epsilon,N}_t,\mathcal{L}(\bar{X}^\epsilon_t))\right] \to 0 \text{ as } N\to\infty.$$

So by the triangle inequality (see, e.g. the proof of [13] Proposition 5.3), we have:

$$\mathbb{E}\left[\mathbb{W}_2(\tilde{\mu}^{\epsilon,N}_t,\mathcal{L}(X_t))\right] \leq \mathbb{E}\left[\mathbb{W}_2(\tilde{\mu}^{\epsilon,N}_t,\bar{\mu}^{\epsilon,N}_t)\right] + \mathbb{E}\left[\mathbb{W}_2(\bar{\mu}^{\epsilon,N}_t,\mathcal{L}(\bar{X}^\epsilon_t))\right] + \mathbb{W}_2(\mathcal{L}(\bar{X}^\epsilon_t),\mathcal{L}(X_t)) \to 0 \text{ as } N\to\infty.$$

The latter statement of the Lemma now follows from the construction of Q^N and Markov's inequality. □

Now we can use the prelimit representation for the controlled fluctuation process $\tilde{Z}^N$ from Lemma 7.7 in order to identify the limiting behavior of $(\tilde{Z}^N, Q^N)$.

Proposition 8.3 *Under assumptions (A1) - (A13),* (Z, Q) *satisfies Eq.* (32) *with probability 1.*

Proof We now invoke the Skorokhod's Representation Theorem as previously discussed. By a standard density argument, we can simply show that Eq. (32) holds with probability 1 for each $\phi \in C_c^\infty(\mathbb{R})$ and $t \in [0, T]$. This is using the fact that there exists a countable, dense collection of smooth, compactly supported functions in $\mathcal{S}_m$ (this follows from, e.g. Corollary 2.1.2 in [62]).

We note that by almost sure convergence of $\tilde{Z}^N$ to Z, we have for each $t \in [0, T]$ and $\phi \in C_c^\infty(\mathbb{R})$, $\langle \tilde{Z}_t^N, \phi\rangle \to \langle Z_t, \phi\rangle$ with probability 1. We also note that the prelimit representation given in Lemma 7.7 can be written solely in terms of Q^N and Z^N by replacing $\tilde{\mu}_s^{\epsilon,N}$ by the first marginal of Q_s^N. We can therefore take $\tilde{\mu}_s^{\epsilon,N}$ to also live on the new probability space from Skorokhod's Representation Theorem, and on that space we still have the convergence of $\tilde{\mu}_t^{\epsilon,N}$ to $\mathcal{L}(X_t)$ in probability proved in Lemma 8.2. Thus, by the representation provided by Lemma 7.7, we only need to show the limits in probability:

$$\int_0^t \langle \tilde{Z}_s^N, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}_s^{\epsilon,N}}\phi(\cdot)\rangle ds \to (N\to\infty) \int_0^t \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle ds \tag{68}$$

$$\begin{aligned}&\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \Big(\sigma(x,y,\tilde{\mu}_s^{\epsilon,N})z_1\phi'(x) + [\tau_1(x,y,\tilde{\mu}_s^{\epsilon,N})z_1 + \tau_2(x,y,\tilde{\mu}_s^{\epsilon,N})z_2]\\ &\Phi_y(x,y,\tilde{\mu}_s^{\epsilon,N})\phi'(x)\Big) Q^N(dx,dy,dz,ds)\\ &\to (N\to\infty)\\ &\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \Big(\sigma(x,y,\mathcal{L}(X_s))z_1\phi'(x) + [\tau_1(x,y,\mathcal{L}(X_s))z_1 + \tau_2(x,y,\mathcal{L}(X_s))z_2]\\ &\Phi_y(x,y,\mathcal{L}(X_s))\phi'(x)\Big) Q(dx,dy,dz,ds),\end{aligned} \tag{69}$$

where $\bar{L}_{\nu_1,\nu_2}$ is as in Eq. (66) and $\bar{L}_\nu$ is as in Eq. (32). By boundedness and continuity of $\bar{\gamma}, \bar{D}$ from assumption (A13) (see Definition D.4), along with Lemma 8.2, we have for each $s \in [0, T]$ and $\phi \in C_c^\infty(\mathbb{R})$, the limit in probability

$$\bar{L}_{\mathcal{L}(X_s),\tilde{\mu}_s^{\epsilon,N}}\phi(\cdot) \to \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot) \text{ in } \mathcal{S}_m$$

holds via the continuous mapping theorem. Thus, for each $s \in [0, T]$ and $\phi \in C_c^\infty(\mathbb{R})$, the limit in probability

$$\langle \tilde{Z}_s^N, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}_s^{\epsilon,N}}\phi(\cdot)\rangle \to \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle$$

holds. We have, then, for all $t \in [0, T]$:

$$\lim_{N\to\infty} \mathbb{E}\left[\left|\int_0^t \left(\langle \tilde{Z}^N_s, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}^{\epsilon,N}_s}\phi(\cdot)\rangle - \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle\right) ds\right|\right] \le$$
$$\le \lim_{N\to\infty} \mathbb{E}\left[\int_0^t \left|\langle \tilde{Z}^N_s, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}^{\epsilon,N}_s}\phi(\cdot)\rangle - \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle\right| ds\right],$$

and we have by Lemma 7.6 that

$$\sup_{N\in\mathbb{N}} \int_0^t \mathbb{E}\left[\left|\langle \tilde{Z}^N_s, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}^{\epsilon,N}_s}\phi(\cdot) - \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle\right|^2\right] ds < \infty,$$

so by uniform integrability we can pass to the limit to get

$$\lim_{N\to\infty} \mathbb{E}\left[\int_0^t \left|\langle \tilde{Z}^N_s, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}^{\epsilon,N}_s}\phi(\cdot)\rangle - \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle\right| ds\right]$$
$$= \mathbb{E}\left[\int_0^t \lim_{N\to\infty}\left|\langle \tilde{Z}^N_s, \bar{L}_{\mathcal{L}(X_s),\tilde{\mu}^{\epsilon,N}_s}\phi(\cdot)\rangle - \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle\right| ds\right] = 0.$$

Similarly, we can use that by the Lemma 7.6 and Fatou's lemma:

$$\sup_{N\in\mathbb{N}} \int_0^t \mathbb{E}\left[\left|\langle \tilde{Z}^N_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle - \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle\right|^2\right] ds < \infty$$

and to get

$$\lim_{N\to\infty} \mathbb{E}\left[\left|\int_0^t \left(\langle \tilde{Z}^N_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle - \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle\right) ds\right|\right]$$
$$\le \mathbb{E}\left[\int_0^t \lim_{N\to\infty}\left|\langle \tilde{Z}^N_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle - \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle\right| ds\right]$$
$$= 0.$$

Then, by Markov's inequality, we establish (68). The limit (69) follows immediately from the integrand being bounded by $C[|z_1| + |z_2|]$ and continuous in $\mathbb{W}_2$, along with the assumed bound on the controls (53) (see, e.g., [22] Theorem A.3.18). ◻

Proposition 8.4 *Under assumptions (A1) - (A13), $Q \in P^*(Z)$ with Probability 1.*

Proof By Proposition 8.3, P^*1 in the definition of $P^*(Z)$ holds. It remains to show P^*2-P^*4.

P^*4 is immediate from the fact that the last marginal of Q is Lebesgue measure by the definition of $M_T(\mathbb{R}^4)$ above Eq. (10), and the first marginal of $\tilde{Q}^N_s$ is $\tilde{\mu}^{\epsilon,N}_s$, which converges in $\mathcal{P}_2(\mathbb{R})$ and hence $\mathcal{P}(\mathbb{R})$ to $\mathcal{L}(X_s)$ by Lemma 8.2.

P^*2 follows from the version of Fatou's lemma from Theorem A.3.12 in [22], since $\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]} |z_1|^2+|z_2|^2 Q^N(dx,dy,dz,dt)$ is a non-negative random variable, and

$$\begin{aligned}&\mathbb{E}\Bigg[\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(|z_1|^2+|z_2|^2\right)Q(dx,dy,dz,dt)\Bigg]\\&\quad\le\liminf_{N\to\infty}\mathbb{E}\Bigg[\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(|z_1|^2+|z_2|^2\right)Q^N(dx,dy,dz,dt)\Bigg]\\&\quad\le\sup_{N\in\mathbb{N}}\int_0^T\mathbb{E}\Bigg[\frac{1}{N}\sum_{i=1}^N|\tilde{u}_i^{N,1}(s)|^2+|\tilde{u}_i^{N,2}(s)|^2\Bigg]ds<\infty\end{aligned}$$

by the assumed bound (52).

Lastly, to see P^*3, take $\psi\in C_c^\infty(U\times\mathbb{R})$ and $\phi\in C_c^\infty(\mathbb{R})$. Here U is an open interval in $\mathbb{R}$ containing $[0,T]$. Then applying Itô's formula (recalling here $\tilde{X}^{i,\epsilon,N}$, $\tilde{Y}^{i,\epsilon,N}$ from Eq. (55)):

$$\begin{aligned}&\phi(\tilde{Y}_T^{i,\epsilon,N})\psi(T,\tilde{X}_T^{i,\epsilon,N})=\phi(\tilde{Y}_0^{i,\epsilon,N})\psi(0,\tilde{X}_0^{i,\epsilon,N})+\int_0^T\Bigg(\dot{\psi}(s,\tilde{X}_s^{i,\epsilon,N})\phi(\tilde{Y}_s^{i,\epsilon,N})\\&+\frac{1}{\epsilon^2}\Bigg[f(i)\phi'(\tilde{Y}_s^{i,\epsilon,N})+\frac{1}{2}(\tau_1^2(i)+\tau_2^2(i))\phi''(\tilde{Y}_s^{i,\epsilon,N})\Bigg]\psi(s,\tilde{X}_s^{i,\epsilon,N})\\&+\frac{1}{\epsilon}\Bigg[g(i)+\tau_1(i)\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}}+\tau_2(i)\frac{\tilde{u}_2^{N,1}(s)}{a(N)\sqrt{N}}\Bigg]\phi'(\tilde{Y}_s^{i,\epsilon,N})\psi(s,\tilde{X}_s^{i,\epsilon,N})\\&+\Bigg[c(i)+\sigma(i)\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}}\Bigg]\phi(\tilde{Y}_s^{i,\epsilon,N})\psi_x(s,\tilde{X}_s^{i,\epsilon,N})+\frac{1}{2}\sigma^2(i)\phi(\tilde{Y}_s^{i,\epsilon,N})\psi_{xx}(s,\tilde{X}_s^{i,\epsilon,N})\\&+\frac{1}{\epsilon}b(i)\phi(\tilde{Y}_s^{i,\epsilon,N})\psi_x(s,\tilde{X}_s^{i,\epsilon,N})+\frac{1}{\epsilon}\sigma(i)\tau_1(i)\phi'(\tilde{Y}_s^{i,\epsilon,N})\psi_x(s,\tilde{X}_s^{i,\epsilon,N})\Bigg)ds\\&+\frac{1}{\epsilon}\int_0^T\tau_1(i)\phi'(\tilde{Y}_s^{i,\epsilon,N})\psi(s,\tilde{X}_s^{i,\epsilon,N})dW_s^i+\frac{1}{\epsilon}\int_0^T\tau_2(i)\phi'(\tilde{Y}_s^{i,\epsilon,N})\psi(s,\tilde{X}_s^{i,\epsilon,N})dB_s^i\\&+\int_0^T\sigma(i)\phi(\tilde{Y}_s^{i,\epsilon,N})\psi_x(s,\tilde{X}_s^{i,\epsilon,N})dW_s^i\end{aligned}$$

where (i) denotes the argument $(\tilde{X}_s^{i,\epsilon,N},\tilde{Y}_s^{i,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})$. So recalling the definition of $L_{x,\mu}$ from Eq. (18), multiplying both sides by $\frac{\epsilon^2}{N}$ and summing,

$$\begin{aligned}&\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}L_{x,\tilde{\mu}_s^{\epsilon,N}}\phi(y)\psi(s,x)Q^N(dx,dy,dz,ds)=\\&=\frac{1}{N}\sum_{i=1}^N\Bigg\{\epsilon^2[\phi(\tilde{Y}_0^{i,\epsilon,N})\psi(0,\tilde{X}_0^{i,\epsilon,N})-\phi(\tilde{Y}_T^{i,\epsilon,N})\psi(T,\tilde{X}_T^{i,\epsilon,N})]\\&\quad+\epsilon^2\int_0^T\Bigg(\dot{\psi}(s,\tilde{X}_s^{i,\epsilon,N})\phi(\tilde{Y}_s^{i,\epsilon,N})+\Bigg[c(i)+\sigma(i)\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}}\Bigg]\phi(\tilde{Y}_s^{i,\epsilon,N})\psi_x(s,\tilde{X}_s^{i,\epsilon,N})\\&\quad+\frac{1}{2}\sigma^2(i)\phi(\tilde{Y}_s^{i,\epsilon,N})\psi_{xx}(s,\tilde{X}_s^{i,\epsilon,N})\Bigg)ds\end{aligned}$$

$$+\epsilon\int_0^T\left(\left[g(i)+\tau_1(i)\frac{\tilde{u}_i^{N,1}(s)}{a(N)\sqrt{N}}+\tau_2(i)\frac{\tilde{u}_2^{N,1}(s)}{a(N)\sqrt{N}}\right]\phi'(\tilde{Y}_s^{i,\epsilon,N})\psi(s,\tilde{X}_s^{i,\epsilon,N})\right.$$
$$\left.+b(i)\phi(\tilde{Y}_s^{i,\epsilon,N})\psi_x(s,\tilde{X}_s^{i,\epsilon,N})+\sigma(i)\tau_1(i)\phi'(\tilde{Y}_s^{i,\epsilon,N})\psi_x(s,\tilde{X}_s^{i,\epsilon,N})\right)ds$$
$$+\epsilon\int_0^T\tau_1(i)\phi'(\tilde{Y}_s^{i,\epsilon,N})\psi(s,\tilde{X}_s^{i,\epsilon,N})dW_s^i+\epsilon\int_0^T\tau_2(i)\phi'(\tilde{Y}_s^{i,\epsilon,N})\psi(s,\tilde{X}_s^{i,\epsilon,N})dB_s^i$$
$$+\epsilon^2\int_0^T\sigma(i)\phi(\tilde{Y}_s^{i,\epsilon,N})\psi_x(s,\tilde{X}_s^{i,\epsilon,N})dW_s^i.$$

Since all terms in the right hand side are bounded other than b and c, which grow at most linearly in y as per Assumption (A5), we see after using the bound (52) that the right hand side is bounded in square expectation by

$$C(T)\epsilon^2(1+\sup_{N\in\mathbb{N}}\frac{1}{N}\sum_{i=1}^N\sup_{s\in[0,T]}\mathbb{E}\left[|\tilde{Y}_s^{i,\epsilon,N}|^2\right])\leq C(T)\epsilon^2$$

by Lemma B.1, and hence converges to 0 in probability.

We can see also by the fact that ϕ and ψ are compactly supported and the coefficients in $L_{x,\mu}$ are continuous in $(x,y,\mathbb{W}_2)$ by assumptions (A1) and (A2), we can use the definition of convergence in $M_T(\mathbb{R}^4)$ and Lemma 8.2 to see the left hand side converges in probability to

$$\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}L_{x,\mathcal{L}(X_s)}\phi(y)\psi(s,x)Q(dx,dy,dz,ds)$$

(see, e.g., [22] Theorem A.3.18). Thus, using that Q satisfies P^*4,

$$\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}L_{x,\mathcal{L}(X_s)}\phi(y)\psi(s,x)Q(dx,dy,dz,ds)$$
$$=\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}L_{x,\mathcal{L}(X_s)}\phi(y)\psi(s,x)\lambda(dy;x,s)\mathcal{L}(X_s)(dx)ds=0$$

for some stochastic kernel λ almost surely. Then noting that by boundedness of the coefficients and the derivatives of ϕ, we have $(s,x)\mapsto\int_{\mathbb{R}}L_{x,\mathcal{L}(X_s)}\phi(y)\lambda(dy;x,s)$ is in $L^1_{\text{loc}}([0,T]\times\mathbb{R},\nu_{\mathcal{L}(X.)})$ for all ϕ, and thus by Corollary 22.38 (2) in [21], for each ϕ, we have

$$\int_{\mathbb{R}}L_{x,\mathcal{L}(X_s)}\phi(y)\lambda(dy;x,s)=0$$

$\nu_{\mathcal{L}(X.)}$- almost surely. By a standard density argument (see [4] Section 6.2.1), we have by letting

$$A=\{(s,x):\int_{\mathbb{R}}L_{x,\mathcal{L}(X_s)}\phi(y)\lambda(dy;x,s)=0,\forall\phi\in C_c^\infty(\mathbb{R})\},$$

$\nu_{\mathcal{L}(X_\cdot)}(A\times[0,T]) = \int_0^T \int_{\mathbb{R}} \mathbb{1}_A \mathcal{L}(X_s)(dx)ds = 1$. This then characterizes $\lambda(dy; x, s)$ as $\nu_{\mathcal{L}(X_\cdot)}-$ almost surely satisfying $L^*_{x,\mathcal{L}(X_s)}\lambda(\cdot; x, s) = 0$ in the distributional sense, and by definition of stochastic kernels $\int_{\mathbb{R}} \lambda(dy; x, s) = 1, \forall x, s$, so $\lambda(dy; x, s)$ is an invariant measure associated to $L_{x,\mathcal{L}(X_s)}$. Since such an invariant measure is unique under assumptions (A1) and (A2) by [60] Proposition 1, we have in fact $\lambda(dy; x, s) = \pi(dy; x, \mathcal{L}(X_s))\ \nu_{\mathcal{L}(X_\cdot)}-$ almost surely.

□

8.1 Weak-sense uniqueness

In order to prove the Laplace Principle Lower bound (31) in Sect. 10 and compactness of level sets in Proposition 10.2, we will need to be able to identify a given $Z \in C([0, T]; \mathcal{S}_{-w/r})$ using only the information that Z solves the limiting controlled Eq. (32) for some fixed Q. Hence, in this subsection, we prove an appropriate notion of weak-sense uniqueness for Eq. (32). Recall the space spaces $\mathcal{S}_p, \mathcal{S}_{-p}$, and the related norms from the beginning of Sect. 2.

Lemma 8.5 *Let $p \in \mathbb{N}$ and consider $\phi \in \mathcal{S}_{p+2}$, $F \in C_b^p(\mathbb{R})$, and $G \in \mathcal{S}_p$. Then for any $\mu \in \mathcal{P}(\mathbb{R})$, we have:*

(1) $\langle \phi, F\phi' \rangle_p \le C \|\phi\|_p^2$

(2) $\langle \phi, F\phi'' \rangle_p \le C \|\phi\|_p^2 - \int_{\mathbb{R}} (1 + x^2)^p |\phi^{(p+1)}(x)|^2 F(x)dx$

(3) $\left\| \int_{\mathbb{R}} G(\cdot)\phi^{(k)}(z)\mu(dz) \right\|_p \le C \|\phi\|_{k+1}$, *for* $k \le p - 1$.

Proof The proof of (1) follows by the same integration by parts argument as A1) in the Appendix of [49]. Part 2 follows by the same integration by parts argument as A2) in the Appendix of [49]. It becomes evident upon reading those proofs that $w_p := (1 + x^2)^p$ can be replaced by any w_p such that $w_p^{-1} D^k w_p$ is bounded for all $k \le p$. The proof of 3 is similar to the proof of A4) in the Appendix of [49]. We recall it here:

$$\begin{aligned}
\left\| \int_{\mathbb{R}} G(\cdot)\phi^{(k)}(z)\mu(dz) \right\|_p &= \left(\sum_{j=0}^p \int_{\mathbb{R}} (1 + x^2)^{2p} \left| \int_{\mathbb{R}} G^{(j)}(x)\phi^{(k)}(z)\mu(dz) \right|^2 dx \right)^{1/2} \\
&\le \|G\|_p \left(\int_{\mathbb{R}} |\phi^{(k)}(z)|^2 \mu(dz) \right)^{1/2} \text{ by Hölder's inequality} \\
&\le \|G\|_p \, |\phi|_k \\
&\le \|G\|_p \, \|\phi\|_{k+1} \,.
\end{aligned}$$

□

Lemma 8.6 *Under assumption (A13), for any $p \in \{1, \ldots, w + 2\}$, where w is as in Eq. (6), and any $s \in [0, T]$, $\bar{L}_{\mathcal{L}(X_s)}$ as given in Eq. (32), where X_s is as in Eq. (25), is a bounded linear map from $\mathcal{S}_{p+2}$ to $\mathcal{S}_p$. In particular, there exists c_p such that for all $s \in [0, T]$ and $\phi \in \mathcal{S}_{p+2}$,*

$$\left\| \bar{L}_{\mathcal{L}(X_s)}\phi \right\|_p \le c_p \|\phi\|_{p+2} \,.$$

The same holds with w replaced by r from Eq. (7) if we in addition assume (A'13).

Proof Linearity is clear. For $\phi \in \mathcal{S}_{p+2}$ and $s \in [0, T]$,

$$\begin{aligned}\left\|\bar{\gamma}(\cdot, \mathcal{L}(X_s))\phi'(\cdot)\right\|_p^2 &= \sum_{k=0}^{p}\int_{\mathbb{R}}(1+x^2)^{2p}\Big([\bar{\gamma}(x, \mathcal{L}(X_s))\phi'(x)]^{(k)}\Big)^2 dx \\ &\leq c_p \sum_{k=0}^{p}\int_{\mathbb{R}}(1+x^2)^{2p}\Big(\phi^{(k+1)}(x)\Big)^2 dx \text{ by Assumption (A13)} \\ &\leq c_p \left\|\phi\right\|_{p+1}^2 .\end{aligned}$$

In the same way, we can see $\left\|\bar{D}(\cdot, \mathcal{L}(X_s))\phi''(\cdot)\right\|_p^2 \leq c_p \left\|\phi\right\|_{p+2}^2$. In addition, we have

$$\begin{aligned}&\left\|\int_{\mathbb{R}}\frac{\delta}{\delta m}\bar{\gamma}(z, \mathcal{L}(X_s))[\cdot]\phi'(z)\mathcal{L}(X_s)(dz)\right\|_p^2 \\ &= \sum_{k=0}^{p}\int_{\mathbb{R}}(1+x^2)^{2p}\left(\frac{\partial^k}{\partial x^k}\left[\int_{\mathbb{R}}\frac{\delta}{\delta m}\bar{\gamma}(z, \mathcal{L}(X_s))[x]\phi'(z)\mathcal{L}(X_s)(dz)\right]\right)^2 dx \\ &\leq \int_{\mathbb{R}}\left\|\frac{\delta}{\delta m}\bar{\gamma}(z, \mathcal{L}(X_s))[\cdot]\right\|_p^2 \mathcal{L}(X_s)(dz)|\phi|_1^2 \text{ by Jensen's inequality and Tonelli's Theorem} \\ &\leq c_p \left\|\phi\right\|_2^2 \text{ by Assumption (A13) and the inequality (9).}\end{aligned}$$

Again, in the same way, we can see

$$\left\|\int_{\mathbb{R}}\frac{\delta}{\delta m}\bar{D}(z, \mathcal{L}(X_s))[\cdot]\phi''(z)\mathcal{L}(X_s)(dz)\right\|_p^2 \leq c_p \left\|\phi\right\|_3^2 ,$$

so by definition of $\bar{L}_\nu$, the result follows. ◻

Lemma 8.7 *Under Assumption (A13), we have for any $p \in \{1, \ldots, w\}$ and $F \in \mathcal{S}_{-p}$, where w is as in Eq. (6),*

$$\sup_{s\in[0,T]} \langle F, \bar{L}^*_{\mathcal{L}(X_s)}F\rangle_{-(p+2)} \leq \left\|F\right\|_{-(p+2)}^2$$

where $\bar{L}^*_{\mathcal{L}(X_s)} : \mathcal{S}_{-p} \to \mathcal{S}_{-(p+2)}$ is the adjoint of $\bar{L}_{\mathcal{L}(X_s)} : \mathcal{S}_{p+2} \to \mathcal{S}_p$ given in Eq. (32) (using here Lemma 8.6). The same holds if instead we further assume (A'13) and replace w with r from Eq. (7).

Proof By the Riesz representation theorem we can take $\phi \in \mathcal{S}_p$ such that for all $\psi \in \mathcal{S}_p$, $\langle F, \psi\rangle = \langle \phi, \psi\rangle_p$ and $\left\|F\right\|_{-p} = \left\|\phi\right\|_p$. By a density argument, we may assume in fact that $\phi \in \mathcal{S}$, $\langle F, \psi\rangle = \langle \phi, \psi\rangle_{p+2}$, and $\left\|\phi\right\|_{p+2} = \left\|F\right\|_{-(p+2)}$. Then for any $s \in [0, T]$, $\langle F, \bar{L}^*_{\mathcal{L}(X_s)}F\rangle_{-(p+2)} = \langle F, \bar{L}_{\mathcal{L}(X_s)}\phi\rangle = \langle \phi, \bar{L}_{\mathcal{L}(X_s)}\phi\rangle_{p+2}$. Then,

$$\begin{aligned}
&\langle \phi, \bar{L}_{\mathcal{L}(X_s)}\phi\rangle_{p+2} \\
&\quad = \langle \phi, \bar{\gamma}(\cdot, \mathcal{L}(X_s))\phi'(\cdot)\rangle_{p+2} + \langle \phi, \bar{D}(\cdot, \mathcal{L}(X_s))\phi''(\cdot)\rangle_{p+2} \\
&\qquad + \langle \phi, \int_{\mathbb{R}} \frac{\delta}{\delta m}\bar{\gamma}(z, \mathcal{L}(X_s))[\cdot]\phi'(z)\mathcal{L}(X_s)(dz)\rangle_{p+2} \\
&\qquad + \langle \phi, \int_{\mathbb{R}} \frac{\delta}{\delta m}\bar{D}(z, \mathcal{L}(X_s))[\cdot]\phi''(z)\mathcal{L}(X_s)(dz)\rangle_{p+2} \\
&\quad \le \langle \phi, \bar{\gamma}(\cdot, \mathcal{L}(X_s))\phi'(\cdot)\rangle_{p+2} + \langle \phi, \bar{D}(\cdot, \mathcal{L}(X_s))\phi''(\cdot)\rangle_{p+2} \\
&\qquad + \|\phi\|_{p+2}\left\|\int_{\mathbb{R}} \frac{\delta}{\delta m}\bar{\gamma}(z, \mathcal{L}(X_s))[\cdot]\phi'(z)\mathcal{L}(X_s)(dz)\right\|_{p+2} \\
&\quad + \|\phi\|_{p+2}\left\|\int_{\mathbb{R}} \frac{\delta}{\delta m}\bar{D}(z, \mathcal{L}(X_s))[\cdot]\phi''(z)\mathcal{L}(X_s)(dz)\right\|_{p+2} \text{ by Cauchy Schwarz} \\
&\qquad \le C\left\{\|\phi\|_{p+2}^2 + \|\phi\|_{p+2}\|\phi\|_2 + \|\phi\|_{p+2}\|\phi\|_3\right\} \\
&\qquad \text{by Lemma 8.5 and Assumption (A13)} \\
&\quad \le C\|\phi\|_{p+2}^2 \\
&\quad = C\|F\|_{-(p+2)}^2 .
\end{aligned}$$

The proof follows in the same way if we replace w with r. □

Proposition 8.8 *Under Assumption (A13), for any (Z, Q) and $(\tilde{Z}, Q)$ such that $Q \in P^*(Z)$ and $Q \in P^*(\tilde{Z})$, $Z = \tilde{Z}$ as elements of $C([0, T]; \mathcal{S}_{-w})$. If we assume (A'13) instead of (A13), $Z = \tilde{Z}$ as elements of $C([0, T]; \mathcal{S}_{-r})$.*

Proof Consider $\eta = Z - \tilde{Z}$. Then by virtue of P^*1 in the definition of P^*, η almost surely satisfies

$$\langle \eta_t, \phi\rangle = \int_0^t \langle \eta_s, \bar{L}_{\mathcal{L}(X_s)}\phi(\cdot)\rangle ds$$

for all $t \in [0, T]$ and $\phi \in \mathcal{S}_w$. Let $\{\phi_j^{w+2}\}_{j\in\mathbb{N}}$ be an orthonormal basis for $\mathcal{S}_{-(w+2)}$. By chain rule, we have

$$\langle \eta_t, \phi_j^{w+2}\rangle^2 = 2\int_0^t \langle \eta_s, \phi_j^{w+2}\rangle\langle \eta_s, \bar{L}_{\mathcal{L}(X_s)}\phi_j^{w+2}(\cdot)\rangle ds.$$

Summing through j, we have using Parseval's identity, Riesz representation theorem, and linearity of η_s and $\bar{L}_{\mathcal{L}(X_s)}$ that

$$\begin{aligned}
\|\eta(t)\|_{-(w+2)} &= 2\int_0^t \langle \eta(s), \bar{L}^*_{\mathcal{L}(s)}\eta(s)ds\rangle_{-(w+2)}ds \\
&\le C\int_0^t \|\eta(s)\|_{-(w+2)}\, ds \text{ by Lemma 8.7}
\end{aligned}$$

so by Gronwall's inequality, $\|\eta(t)\|_{-(w+2)} = 0, \forall t \in [0, T]$, so $\|\eta(t)\|_{-w} = 0, \forall t \in [0, T]$, and hence $Z = \tilde{Z}$. The proof follows in the same way if we replace w with r. □

Remark 8.9 By P^*3 and P^*4 in the definition of P^*, we have that for any $Q \in P^*(Z)$ that disintegrating $Q(dx, dy, dz, ds) = \kappa(dz; x, y, s)\lambda(dy; x, s)Q_{(1,4)}(dx, ds)$, $\lambda(dy; x, s) = \pi(dy; x, \mathcal{L}(X_s))$ and $Q_{(1,4)}(dx, ds) = \mathcal{L}(X_s)(dx)ds = \nu_{\mathcal{L}(X_\cdot)}(dx, ds)$. Thus any $Q, \tilde{Q} \in P^*(Z)$ only differentiate in their control stochastic kernels, $\kappa(dz; x, y, s)$ and $\tilde{\kappa}(dz; x, y, s)$. These are, of course, entirely determined by the choice of controls in the construction of Q^N. In other words, keeping in mind the result of Proposition 8.4, the choice of controls in the prelimit system (55) determine uniquely the limit in distribution of $\tilde{Z}^N$.

9 Laplace principle lower bound

We now can prove the Laplace principle Lower Bound (30).

Proposition 9.1 *Under assumptions (A1)–(A13), Eq. (30) holds.*

Proof Take $\tau \geq w$, with w as in Eq. (6), $F \in C_b(C([0, T]; \mathcal{S}_{-\tau}))$ and $\eta > 0$. By Eq. (51) there exists $\{\tilde{u}^N\}_{N\in\mathbb{N}}$ such that for all N,

$$-a^2(N)\log\mathbb{E}\exp\left(-\frac{1}{a^2(N)}F(Z^N)\right) \geq \mathbb{E}\left[\frac{1}{2}\frac{1}{N}\sum_{i=1}^N\int_0^T|\tilde{u}_i^{N,1}(s)|^2 + |\tilde{u}_i^{N,2}(s)|^2ds + F(\tilde{Z}^N)\right] - \eta.$$

Where $\tilde{Z}^N$ is as in Eq. (54) and is controlled by $\{\tilde{u}^N\}_{N\in\mathbb{N}}$. Then letting Q^N be as in Eq. (56) with this choice of controls (recalling that we can assume the almost-sure bound (53) on the controls by the argument found in Theorem 4.4 of [8]), we have

$$\mathbb{E}\left[\frac{1}{2}\frac{1}{N}\sum_{i=1}^N\int_0^T\left(|\tilde{u}_i^{N,1}(s)|^2 + |\tilde{u}_i^{N,2}(s)|^2\right)ds + F(\tilde{Z}^N)\right] = \mathbb{E}\left[\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(|z_1|^2 + |z_2|^2\right)Q^N(dxdydzds) + F(\tilde{Z}^N)\right]$$

so by the version of Fatou's lemma from Theorem A.3.12 in [22], we have

$$\begin{aligned}&\liminf_{N\to\infty} -a^2(N)\log\mathbb{E}\exp\left(-\frac{1}{a^2(N)}F(Z^N)\right)\\ &\geq \liminf_{N\to\infty}\mathbb{E}\left[\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(|z_1|^2 + |z_2|^2\right)Q^N(dxdydzds) + F(\tilde{Z}^N)\right] - \eta\\ &\geq \mathbb{E}\left[\liminf_{N\to\infty}\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(|z_1|^2 + |z_2|^2\right)Q^N(dxdydzds) + F(\tilde{Z}^N)\right] - \eta\end{aligned}$$

$$
\begin{aligned}
&\geq \mathbb{E}\left[\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(|z_1|^2+|z_2|^2\right)Q(dxdydzds)+F(Z)\right]-\eta\\
&\geq \inf_{Z\in C([0,T];\mathcal{S}_{-m})}\left\{\inf_{Q\in P^*(Z)}\left\{\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(|z_1|^2+|z_2|^2\right)Q(dxdydzds)\right\}+F(Z)\right\}-\eta\\
&\geq \inf_{Z\in C([0,T];\mathcal{S}_{-w})}\left\{\inf_{Q\in P^*(Z)}\left\{\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(|z_1|^2+|z_2|^2\right)Q(dxdydzds)\right\}+F(Z)\right\}-\eta\\
&= \inf_{Z\in C([0,T];\mathcal{S}_{-w})}\{I(Z)+F(Z)\}-\eta,
\end{aligned}
$$

where in the second-to-last inequality we used Proposition 8.4. So Eq. (30) is established. □

10 Laplace principle upper bound and compactness of level sets

We now prove the Laplace principle Upper Bound and, under the additional assumption of (A'13), compactness of level sets.

Proposition 10.1 *Under assumptions (A1)–(A13), the Laplace principle Upper Bound* (31) *holds.*

Proof We use the ordinary formulation I^o from Eq. (34). We take $\eta > 0$, w as in Eq. (6), $\tau \geq w$, $F \in C_b(C([0,T];\mathcal{S}_{-\tau}))$, and Z^* such that

$$
I(Z^*)+F(Z^*)\leq \inf_{Z\in C([0,T];\mathcal{S}_{-w})}\mathbb{E}\left[I(Z)+F(Z)\right]+\frac{\eta}{2}.
$$

Then we can find $h \in P^o(Z^*)$ such that

$$
\frac{1}{2}\int_0^T\mathbb{E}\left[\int_{\mathbb{R}}|h(s,X_s,y)|^2\pi(dy;X_s,\mathcal{L}(X_s))\right]ds\leq I(Z^*)+\frac{\eta}{2}.
$$

Then since $\nu(\Gamma\times A\times B) := \int_\Gamma\int_{\mathbb{R}}\mathbb{1}_A(x)\int_{\mathbb{R}}\mathbb{1}_B(y)\pi(dy;x,\mathcal{L}(X_s))\mathcal{L}(X_s)(dx)ds$, $\Gamma \in \mathcal{B}(U)$, $A, B \in \mathcal{B}(\mathbb{R})$ is a finite Borel measure on $U\times\mathbb{R}\times\mathbb{R}$ for all $x \in \mathbb{R}$, by Corollary 22.38 (1) in [21], we can take $\{\psi^k_j\}_{k\in\mathbb{N}} \subset C^\infty_c(U\times\mathbb{R}\times\mathbb{R})$ such that $\psi^k_j \to h_j$ in $L^2(U\times\mathbb{R}\times\mathbb{R},\nu)$ for $j\in\{1,2\}$. Here we let U be any open interval containing $[0,T]$ and assume $\nu(\Gamma\times A\times B)$ is 0 when $\Gamma\cap[0,T]=\emptyset$.

Then letting $\tilde{u}^N_{i,k}(s,\omega) = \psi^k(s,\tilde{X}^{i,\epsilon,N,k}_s(\omega),\tilde{Y}^{i,\epsilon,N,k}_s(\omega))$, where $(\tilde{X}^{i,\epsilon,N,k}_s,\tilde{Y}^{i,\epsilon,N,k}_s)$ are as in Eq. (55) but controlled by $\frac{\tilde{u}^N_{i,k}(s)}{a(N)\sqrt{N}}$,

$$
\begin{aligned}
\sup_{N\in\mathbb{N}}\int_0^T\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N|\tilde{u}^N_{i,k}(s)|^2\right]ds &= \sup_{N\in\mathbb{N}}\int_0^T\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N|\psi_k(s,\tilde{X}^{i,\epsilon,N}_s,\tilde{Y}^{i,\epsilon,N}_s)|^2\right]ds\\
&\leq T\,\|\psi_k\|^2_\infty
\end{aligned}
$$

for each $k \in \mathbb{N}$, and in fact

$$\int_0^T \frac{1}{N} \sum_{i=1}^N |\tilde{u}_{i,k}^N(s)|^2 ds \le T \, \|\psi_k\|_\infty^2$$

for each $k \in \mathbb{N}$ (so the supposition (53) holds with this choice of controls).

Letting $(\tilde{Z}^{N,k}, Q^{N,k})$ be as in Eqs. (54) and (56) with this choice of controls, we want to establish that $(\tilde{Z}^{N,k}, Q^{N,k})$ converges in distribution as a sequence of $C([0,T]; \mathcal{S}_{-m}) \times M_T(\mathbb{R}^4)$-valued random variables to $(\tilde{Z}^k, Q^k)$ as $N \to \infty$, where $Q^k \in P^*(\tilde{Z}^k)$ (this is immediate since we prove this for all L^2 controls in Proposition 8.4) and such that

$$\begin{aligned} &Q^k(A \times B \times C \times \Gamma) \\ &\quad = \int_\Gamma \int_A \int_C \delta_{\psi^k(s,x,y)}(C) \pi(dy; x, \mathcal{L}(X_s)) \mathcal{L}(X_s)(dx) ds, \\ &\quad \forall A, B \in \mathcal{B}(\mathbb{R}), C \in \mathcal{B}(\mathbb{R}^2), \Gamma \in \mathcal{B}([0,T]). \end{aligned} \tag{70}$$

By the weak-sense uniqueness established in Proposition 8.8, this determines each $\tilde{Z}^k$ almost surely to be the unique element of $C([0,T]; \mathcal{S}_{-m})$ satisfying Eq. (32) with Q^k in the place of Q.

Then we will send $k \to \infty$ and show $(\tilde{Z}^k, Q^k)$ converges to $(\tilde{Z}, Q)$ in $C([0,T]; \mathcal{S}_{-w}) \times M_T(\mathbb{R}^4)$, where $Q \in P^*(\tilde{Z})$ and

$$\begin{aligned} &Q(A \times B \times C \times \Gamma) \\ &\quad = \int_\Gamma \int_A \int_C \delta_{h(s,x,y)}(C) \pi(dy; x, \mathcal{L}(X_s)) \mathcal{L}(X_s)(dx) ds, \\ &\quad \forall A, B \in \mathcal{B}(\mathbb{R}), C \in \mathcal{B}(\mathbb{R}^2), \Gamma \in \mathcal{B}([0,T]). \end{aligned} \tag{71}$$

Then by the weak-sense uniqueness established in Proposition 8.8, we have $\tilde{Z} \overset{d}{=} Z^*$. By reverse Fatou's lemma:

$$\begin{aligned} &\limsup_{N\to\infty} -a^2(N) \log \mathbb{E} \exp\left(-\frac{1}{a^2(N)} F(Z^N)\right) \\ &= \limsup_{N\to\infty} \inf_{\tilde{u}^N} \mathbb{E}\left[\frac{1}{2}\frac{1}{N}\sum_{i=1}^N \int_0^T \left(|\tilde{u}_i^{N,1}(s)|^2 + |\tilde{u}_i^{N,2}(s)|^2\right) ds + F(\tilde{Z}^N)\right] \text{ by Eq. (51)} \\ &\le \limsup_{N\to\infty} \mathbb{E}\left[\frac{1}{2}\frac{1}{N}\sum_{i=1}^N \int_0^T \left(|\tilde{u}_{i,k}^{N,1}(s)|^2 + |\tilde{u}_{i,k}^{N,2}(s)|^2\right) ds + F(\tilde{Z}^{N,k})\right], \forall k \in \mathbb{N} \\ &= \limsup_{N\to\infty} \mathbb{E}\left[\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]} \left(z_1^2 + z_2^2\right) Q^{N,k}(dx,dy,dz,ds) + F(\tilde{Z}^{N,k})\right], \forall k \in \mathbb{N} \\ &\le \mathbb{E}\left[\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]} \left(z_1^2 + z_2^2\right) Q^k(dx,dy,dz,ds) + F(\tilde{Z}^k)\right], \forall k \in \mathbb{N} \end{aligned}$$

$$= \frac{1}{2}\int_0^T \mathbb{E}\Big[\int_{\mathbb{R}} |\psi^k(s, X_s, y)|^2 \pi(dy; X_s, \mathcal{L}(X_s))\Big] ds + \mathbb{E}\Big[F(\tilde{Z}^k)\Big], \forall k \in \mathbb{N}$$

Then sending $k \to \infty$ and using the L^2 convergence of ψ^k to h and the boundedness F and convergence of $\tilde{Z}^k$ to Z^*, we get

$$\begin{aligned}
&\limsup_{N\to\infty} -a^2(N) \log \mathbb{E} \exp\left(-\frac{1}{a^2(N)} F(Z^N)\right) \\
&\quad \le \frac{1}{2}\int_0^T \mathbb{E}\Big[\int_{\mathbb{R}} |h(s, X_s, y)|^2 \pi(dy; X_s, \mathcal{L}(X_s))\Big] ds + \mathbb{E}\Big[F(Z^*)\Big] \\
&\quad \le I(Z^*) + F(Z^*) + \frac{\eta}{2} \\
&\quad \le \inf_{Z \in C([0,T];\mathcal{S}_{-w})} \mathbb{E}\Big[I(Z) + F(Z)\Big] + \eta
\end{aligned}$$

so Eq. (31) will be established.

Looking at the proof of Proposition 8.3, to see $(\tilde{Z}^{N,k}, Q^{N,k})$ converges to $(\tilde{Z}^k, Q^k)$ where $Q^k \in P^*(\tilde{Z}^k)$ satisfies Eq. (70). we just need to establish that

$$\begin{aligned}
&\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \sigma(x, y, \tilde{\mu}_s^{\epsilon,N,k}) z_1 \phi'(x) Q^{N,k}(dx, dy, dz, ds) \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x, y, \tilde{\mu}_s^{\epsilon,N,k}) z_1 + \tau_2(x, y, \tilde{\mu}_s^{\epsilon,N,k}) z_2] \\
&\quad \Phi_y(x, y, \tilde{\mu}_s^{\epsilon,N,k}) \phi'(x) Q^{N,k}(dx, dy, dz, ds)
\end{aligned}$$

converges in distribution to

$$\begin{aligned}
&\int_0^t \mathbb{E}\Big[\int_{\mathbb{R}} \sigma(X_s, y, \mathcal{L}(X_s)) \psi_1^k(s, X_s, y) \phi'(X_s) \pi(dy; X_s, \mathcal{L}(X_s))\Big] ds \\
&\quad + \int_0^t \mathbb{E}\Big[\int_{\mathbb{R}} \Big([\tau_1(X_s, y, \mathcal{L}(X_s)) \psi_1^k(s, X_s, y) + \tau_2(X_s, y, \mathcal{L}(X_s)) \psi_2^k(s, X_s, y)] \\
&\quad \Phi_y(X_s, y, \mathcal{L}(X_s)) \phi'(X_s) \Big) \pi(dy; X_s, \mathcal{L}(X_s))\Big] ds
\end{aligned}$$

for all $\phi \in C_c^\infty(\mathbb{R})$ and $t \in [0, T]$. Fix $k \in \mathbb{N}$ and ϕ and t. We have

$$\begin{aligned}
&\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} \sigma(x, y, \tilde{\mu}_s^{\epsilon,N,k}) z_1 \phi'(x) Q^{N,k}(dx, dy, dz, ds) \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,t]} [\tau_1(x, y, \tilde{\mu}_s^{\epsilon,N,k}) z_1 + \tau_2(x, y, \tilde{\mu}_s^{\epsilon,N,k}) z_2] \\
&\quad \Phi_y(x, y, \tilde{\mu}_s^{\epsilon,N,k}) \phi'(x) Q^{N,k}(dx, dy, dz, ds) \\
&= \int_0^t \frac{1}{N} \sum_{i=1}^N \sigma(\tilde{X}_s^{i,\epsilon,N,k}, \tilde{Y}_s^{i,\epsilon,N,k}, \tilde{\mu}_s^{\epsilon,N,k}) \psi_1^k(s, \tilde{X}_s^{i,\epsilon,N,k}, \tilde{Y}_s^{i,\epsilon,N,k}) \phi'(\tilde{X}_s^{i,\epsilon,N,k}) ds
\end{aligned}$$

$$+\int_0^t \frac{1}{N}\sum_{i=1}^N \Big[\tau_1(\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k},\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k})+$$
$$+\tau_2(\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k},\tilde{\mu}_s^{\epsilon,N,k})\psi_2^k(s,\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k})\Big]$$
$$\Phi_y(\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k},\tilde{\mu}_s^{\epsilon,N,k}\phi'(\tilde{X}_s^{i,\epsilon,N,k})ds$$

Then using Proposition 6.4 with

$$F(s,x,y,\mu)=\sigma(x,y,\mu)\psi_1^k(s,x,y)+[\tau_1(x,y,\mu)\psi_1^k(s,x,y)$$
$$+\tau_2(x,y,\mu)\psi_2^k(s,x,y)]\Phi_y(x,y,\mu)$$

using that s only appears as a parameter, in the same way as x, so that the same proof holds (using also the assumed bound on the time derivative of Ξ in (A8)), we get that

$$\mathbb{E}\Big[\Big|\int_0^t \frac{1}{N}\sum_{i=1}^N \sigma(\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k},\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k})\phi'(\tilde{X}_s^{i,\epsilon,N,k})ds$$
$$+\int_0^t \frac{1}{N}\sum_{i=1}^N \Big[\tau_1(\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k},\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k})+$$
$$+\tau_2(\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k},\tilde{\mu}_s^{\epsilon,N,k})\psi_2^k(s,\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k})\Big]$$
$$\Phi_y(\tilde{X}_s^{i,\epsilon,N,k},\tilde{Y}_s^{i,\epsilon,N,k},\tilde{\mu}_s^{\epsilon,N,k})\phi'(\tilde{X}_s^{i,\epsilon,N,k})ds$$
$$-\int_0^t \frac{1}{N}\sum_{i=1}^N \int_{\mathbb{R}} \sigma(\tilde{X}_s^{i,\epsilon,N,k},y,\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,\tilde{X}_s^{i,\epsilon,N,k},y)\phi'(\tilde{X}_s^{i,\epsilon,N,k})$$
$$+[\tau_1(\tilde{X}_s^{i,\epsilon,N,k},y,\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,\tilde{X}_s^{i,\epsilon,N,k},y)+\tau_2(\tilde{X}_s^{i,\epsilon,N,k},y,\tilde{\mu}_s^{\epsilon,N,k})$$
$$\times\,\psi_2^k(s,\tilde{X}_s^{i,\epsilon,N,k},y)]\Phi_y(\tilde{X}_s^{i,\epsilon,N,k},y,\tilde{\mu}_s^{\epsilon,N,k})\psi(s,\tilde{X}_s^{i,\epsilon,N,k})\pi(dy;\tilde{X}_s^{i,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})ds\Big|\Big]$$
$$\le C(T)\epsilon$$

Then noting that

$$\int_0^t \frac{1}{N}\sum_{i=1}^N \int_{\mathbb{R}} \sigma(\tilde{X}_s^{i,\epsilon,N,k},y,\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,\tilde{X}_s^{i,\epsilon,N,k},y)\phi'(\tilde{X}_s^{i,\epsilon,N,k})$$
$$+[\tau_1(\tilde{X}_s^{i,\epsilon,N,k},y,\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,\tilde{X}_s^{i,\epsilon,N,k},y)+\tau_2(\tilde{X}_s^{i,\epsilon,N,k},y,\tilde{\mu}_s^{\epsilon,N,k})\psi_2^k(s,\tilde{X}_s^{i,\epsilon,N,k},y)]\times$$
$$\times\,\Phi_y(\tilde{X}_s^{i,\epsilon,N,k},y,\tilde{\mu}_s^{\epsilon,N,k})\phi'(\tilde{X}_s^{i,\epsilon,N,k})\pi(dy;\tilde{X}_s^{i,\epsilon,N},\tilde{\mu}_s^{\epsilon,N})ds$$
$$=\int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}}\Big(\sigma(x,y,\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,x,y)\phi'(x)+[\tau_1(x,y,\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,x,y)$$
$$+\tau_2(x,y,\tilde{\mu}_s^{\epsilon,N,k})\psi_2^k(s,x,y)]\Phi_y(x,y,\tilde{\mu}_s^{\epsilon,N,k})\phi'(x)\Big)\pi(dy;x,\tilde{\mu}_s^{\epsilon,N})\tilde{\mu}_s^{\epsilon,N,k}(dx)ds$$

and using that the integrand of the first two integrals above is bounded by Assumptions (A1),(A5),and (A6) and continuous in $\mathbb{W}_2$ by Assumption (A12), along with the convergence of $\tilde{\mu}_s^{\epsilon,N}$ to $\mathcal{L}(X_s)$ from Lemma 8.2, we have by dominated convergence theorem (invoking here Skorokhod's representation theorem to assume $\tilde{\mu}_s^{\epsilon,N}$ to $\mathcal{L}(X_s)$ almost surely as in Proposition 8.3) and Theorem A.3.18 in [22]:

$$\lim_{N\to\infty}\mathbb{E}\Bigg[\Bigg|\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}\Big(\sigma(x,y,\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,x,y)\phi'(x)+[\tau_1(x,y,\tilde{\mu}_s^{\epsilon,N,k})\psi_1^k(s,x,y)$$
$$+\tau_2(x,y,\tilde{\mu}_s^{\epsilon,N,k})\psi_2^k(s,x,y)]\Phi_y(x,y,\tilde{\mu}_s^{\epsilon,N,k})\phi'(x)\Big)\pi(dy;x,\tilde{\mu}_s^{\epsilon,N})\tilde{\mu}_s^{\epsilon,N,k}(dx)ds$$
$$-\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}\Big(\sigma(x,y,\mathcal{L}(X_s))\psi_1^k(s,x,y)\phi'(x)+[\tau_1(x,y,\mathcal{L}(X_s))\psi_1^k(s,x,y)$$
$$+\tau_2(x,y,\mathcal{L}(X_s))\psi_2^k(s,x,y)]\Phi_y(x,y,\mathcal{L}(X_s))\phi'(x)\Big)\pi(dy;x,\mathcal{L}(X_s))\mathcal{L}(X_s)(dx)ds\Bigg|\Bigg]=0$$

so by triangle inequality, the desired convergence is shown.

Now we seek to establish that $(\tilde{Z}^k, Q^k)$ converges to $(\tilde{Z}, Q)$ in $C([0,T];\mathcal{S}_{-w})\times M_T(\mathbb{R}^4)$ where $Q\in P^*(\tilde{Z})$ and Q satisfies (71).

We first prove precompactness. We have since $\psi^k\to h$ in $L^2(U\times\mathbb{R}\times\mathbb{R},\nu)$,

$$\sup_{k\in\mathbb{N}}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(z_1^2+z_2^2\right)Q^k(dx,dy,dz,ds)=$$
$$=\sup_{k\in\mathbb{N}}\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}\left(|\psi_1^k(s,x,y)|^2+|\psi_2^k(s,x,y)|^2\right)\pi(dy;x,\mathcal{L}(X_s))\mathcal{L}(X_s)(dx)ds<\infty.$$

Moreover, by P^*3 and P^*4

$$\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(y^2+|x|^2\right)Q^k(dx,dy,dz,ds)$$
$$=\int_0^T\mathbb{E}\left[\int_{\mathbb{R}}y^2\pi(dy;X_s,\mathcal{L}(X_s))+|X_s|^2\right]ds<\infty,\forall k\in\mathbb{N},$$

where here we have used that $\pi(\cdot;x,\mu)$ from Eq. (20) has bounded moments of all orders uniformly in $x\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})$ and that for X_s from Eq. (25), $\sup_{s\in[0,T]}\mathbb{E}[|X_s|^2]<\infty$, which follows easily from the fact that $\bar{\gamma}$ and $\bar{D}$ are bounded as per Assumption (A13). Thus, via the same tightness function used for Proposition 7.10, $\{Q^k\}_{k\in\mathbb{N}}$ is tight in $M_T(\mathbb{R}^4)$.

To see that $\{\tilde{Z}^k\}_{k\in\mathbb{N}}$ is precompact, we use that for each k $(\tilde{Z}^k, Q^k)$ must satisfy Eq. (32). That is, for $\phi\in\mathcal{S}$ and $t\in[0,T]$:

$$\langle\tilde{Z}_t^k,\phi\rangle=\int_0^t\langle\tilde{Z}_s^k,\bar{L}_{\mathcal{L}(X_s)}\phi\rangle ds+\int_0^t\langle B_s^k,\phi\rangle ds$$
$$\langle B_t^k,\phi\rangle:=\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2}\Big(\sigma(x,y,\mathcal{L}(X_s))z_1\phi'(x)+[\tau_1(x,y,\mathcal{L}(X_s))z_1+\tau_2(x,y,\mathcal{L}(X_s))z_2]$$
$$\Phi_y(x,y,\mathcal{L}(X_s))\phi'(x)\Big)Q_t^k(dx,dy,dz).$$

Here $Q_t^k \in \mathcal{P}(\mathbb{R}^4)$ is such that $Q^k(dx, dy, dz, dt) = Q_t^k(dx, dy, dz)dt$. We can see that $B_t^k \in \mathcal{S}_{-(m+2)}$ for almost every $t \in [0, T]$, and in fact

$$\begin{aligned}
&\sup_{k\in\mathbb{N}} \int_0^T \left\| B_s^k \right\|_{-(m+2)}^2 ds \\
&\quad = \sup_{k\in\mathbb{N}} \int_0^T \sup_{\|\phi\|_{m+2}=1} |\langle B_s^k, \phi\rangle|^2 ds \\
&\quad \le \sup_{k\in\mathbb{N}} \int_0^T \sup_{\|\phi\|_{m+2}=1} \left\{ \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2} \left(|z_1|^2 + |z_2|^2\right) Q_s^k(dx, dy, dz) |\phi|_1^2 \right\} ds \\
&\quad \le \sup_{k\in\mathbb{N}} \int_0^T \sup_{\|\phi\|_{m+2}=1} \left\{ \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2} \left(|z_1|^2 + |z_2|^2\right) Q_s^k(dx, dy, dz) \|\phi\|_{m+2}^2 \right\} ds \\
&\quad = \sup_{k\in\mathbb{N}} \int_0^T \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2} \left(|z_1|^2 + |z_2|^2\right) Q_s^k(dx, dy, dz) ds < \infty
\end{aligned}$$

Thus, by the proof of Theorem 2.5.2 in [45], it suffices to show that for fixed $\phi \in \mathcal{S}$, $\langle \tilde{Z}_t^k, \phi\rangle$ is relatively compact in $C([0, T]; \mathbb{R})$, and $\tilde{Z}^k$ is uniformly $(m+2)$-continuous to get precompactness of $\tilde{Z}^k$ in $C([0, T]; \mathcal{S}_{-w})$ for $w > m + 2$ sufficiently large that the canonical embedding $\mathcal{S}_{-m-2} \to \mathcal{S}_{-w}$ is Hilbert–Schmidt (see Eq. 6).

We have that, in the same way as the proof of Proposition 8.8 (using here that $\tilde{Z}^k \in C([0, T]; \mathcal{S}_{-m})$),

$$\begin{aligned}
\left\| \tilde{Z}_t^k \right\|_{-(m+2)}^2 &= 2 \int_0^t \langle \tilde{Z}_s^k, L_{\mathcal{L}(X_s)}^* \tilde{Z}_s^k\rangle_{-(m+2)} ds + 2 \int_0^t \langle \tilde{Z}_s^k, B_s^k\rangle_{-(m+2)} ds \\
&\le C \int_0^t \left\| \tilde{Z}_s^k \right\|_{-(m+2)}^2 ds + 2 \int_0^t \left\| \tilde{Z}_s^k \right\|_{-(m+2)} \left\| B_s^k \right\|_{-(m+2)} ds \\
&\qquad \text{by Cauchy Schwarz and Lemma 8.7} \\
&\le C \left\{ \int_0^t \left\| \tilde{Z}_s^k \right\|_{-(m+2)}^2 ds + \int_0^t \left\| B_s^k \right\|_{-(m+2)}^2 ds \right\}
\end{aligned}$$

so by Gronwall's inequality,

$$\sup_{k\in\mathbb{N}} \sup_{t\in[0,T]} \left\| \tilde{Z}_t^k \right\|_{-(m+2)}^2 \le C(T).$$

This gives then that for $t_1, t_2 \in [0, T]$ and $\phi \in \mathcal{S}$:

$$\begin{aligned}
&|\langle \tilde{Z}_{t_2}^k, \phi\rangle - \langle \tilde{Z}_{t_1}^k, \phi\rangle| \\
&\quad \le 2|t_2 - t_1| \left\{ \int_0^T |\langle \tilde{Z}_s^k, \bar{L}_{\mathcal{L}(X_s)}\phi\rangle|^2 ds + \int_0^T |\langle B_s^k, \phi\rangle|^2 ds \right\} \\
&\quad \le 2|t_2 - t_1| \left\{ \int_0^T \left\| \tilde{Z}_s^k \right\|_{-(m+2)}^2 \left\| \bar{L}_{\mathcal{L}(X_s)}\phi \right\|_{m+2}^2 ds + \int_0^T \left\| B_s^N \right\|_{-(m+2)}^2 \|\phi\|_{m+2}^2 ds \right\}
\end{aligned}$$

$$\leq 2|t_2 - t_1| C(T) \|\phi\|_{m+4}^2 \text{ by Lemma 8.6,}$$

and precompactness of $\{\tilde{Z}^k\}_{k\in\mathbb{N}}$ is established.

Taking a convergent subsequence, which we do not relabel in the notation, we call its limit (Z, Q). The fact that P^*2-P^*4 in the definition of $P^*(Z)$ are satisfied follows in the exact same way as Proposition 10.2. It thus remains to show that (Z, Q) satisfies Eq. (32) with Q given in Eq. (71). At this point, by Proposition 8.8, we will have the limit is uniquely identified for every subsequence, and hence the lemma is proved. By a density argument, it suffices to show that for each $\phi \in C_c^\infty(\mathbb{R})$ and $t \in [0, T]$,

$$\begin{aligned}
&\lim_{k\to\infty} \langle \tilde{Z}_t^k, \phi\rangle \\
&\quad = \int_0^t \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi\rangle ds + \int_0^t \mathbb{E}\Bigg[\int_{\mathbb{R}} \Bigg(\sigma(x, y, \mathcal{L}(X_s))h_1(s, X_s, y)\phi'(X_s) \\
&\quad + [\tau_1(X_s, y, \mathcal{L}(X_s))h_1(s, X_s, y) + \tau_2(X_s, y, \mathcal{L}(X_s))h_2(s, X_s, y)] \\
&\quad \Phi_y(X_s, y, \mathcal{L}(X_s))\phi'(X_s)\Bigg)\pi(dy; X_s, \mathcal{L}(X_s))\Bigg]ds.
\end{aligned}$$

We have by dominated convergence theorem, L^2 convergence of ψ^k to h, and that under Assumption (A13) $\bar{L}_{\mathcal{L}(X_s)}\phi \in \mathcal{S}_w, \forall s \in [0, T]$:

$$\begin{aligned}
&\lim_{k\to\infty} \langle \tilde{Z}_t^k, \phi\rangle \\
&= \lim_{k\to\infty} \Bigg\{\int_0^t \langle \tilde{Z}_s^k, \bar{L}_{\mathcal{L}(X_s)}\phi\rangle ds + \int_0^t \langle B_s^k, \phi\rangle ds\Bigg\} \\
&= \int_0^t \lim_{k\to\infty} \langle \tilde{Z}_s^k, \bar{L}_{\mathcal{L}(X_s)}\phi\rangle ds + \lim_{k\to\infty} \int_0^t \mathbb{E}\Bigg[\int_{\mathbb{R}} \Bigg(\sigma(x, y, \mathcal{L}(X_s))\psi_1^k(s, X_s, y)\phi'(X_s) \\
&+ [\tau_1(X_s, y, \mathcal{L}(X_s))\psi_1^k(s, X_s, y) + \tau_2(X_s, y, \mathcal{L}(X_s))\psi_2^k(s, X_s, y)] \\
&\Phi_y(X_s, y, \mathcal{L}(X_s))\phi'(X_s)\Bigg)\pi(dy; X_s, \mathcal{L}(X_s))\Bigg]ds \\
&= \int_0^t \langle Z_s, \bar{L}_{\mathcal{L}(X_s)}\phi\rangle ds + \int_0^t \mathbb{E}\Bigg[\int_{\mathbb{R}} \Bigg(\sigma(x, y, \mathcal{L}(X_s))h_1(s, X_s, y)\phi'(X_s) \\
&+ [\tau_1(X_s, y, \mathcal{L}(X_s))h_1(s, X_s, y) + \tau_2(X_s, y, \mathcal{L}(X_s))h_2(s, X_s, y)] \\
&\Phi_y(X_s, y, \mathcal{L}(X_s))\phi'(X_s)\Bigg)\pi(dy; X_s, \mathcal{L}(X_s))\Bigg]ds
\end{aligned}$$

as desired. □

Proposition 10.2 *Under assumptions (A1)–(A12) and (A'13), I given in Theorems 3.1/3.2 is a good rate function on $C([0, T]; \mathcal{S}_{-r})$ for $r > w + 2$ as in Eq. (7).*

Proof We need to show that for any $L > 0$,

$$\Theta_L := \{Z \in C([0, T]; \mathcal{S}_{-r}) : I(Z) \leq L\}$$

is compact in $C([0,T];\mathcal{S}_{-r})$.

Let $\{Z^N\}_{N\in\mathbb{N}} \subset \Theta_L$. Then by the form of I, for each $N \in \mathbb{N}$, there exists $Q^N \in P^*(Z^N)$ such that

$$\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(z_1^2+z_2^2\right)Q^N(dx,dy,dz,ds) \le L+\frac{1}{N}$$

and by P^*3 and P^*4, we have as with the Q^k's in the proof of Proposition 9.1

$$\sup_{N\in\mathbb{N}}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(y^2+|x|^2\right)Q^N(dx,dy,dz,ds) < \infty.$$

Thus by the same tightness function used for Proposition 7.10, $\{Q^N\}_{N\in\mathbb{N}}$ is tight in $M_T(\mathbb{R}^4)$.

Taking a subsequence of $\{Q^N\}$ which converges to some $Q \in M_T(\mathbb{R}^4)$ (which we do not relabel in the notation), define $Z \in C([0,T];\mathcal{S}_{-w})$ to be the unique solution to Eq. (32) with this choice of Q. Here we are using that by the proof of Proposition 10.1 such a solution exists and that by Proposition 8.8 it is unique - see the discussion before Lemma 4.10 in [10]. We claim that (Z^N,Q^N) converges to (Z,Q) in $C([0,T];\mathcal{S}_{-r})\times M_T(\mathbb{R}^4)$ and $Q \in P^*(Z)$. At this point we will have that since Z^N has a limit, Θ_L is precompact, and by the version of Fatou's lemma from Theorem A.3.12 in [22]:

$$\begin{aligned}I(Z) &\le \frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(z_1^2+z_2^2\right)Q(dx,dy,dz,ds)\\ &\le \liminf_{N\to\infty}\frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2\times[0,T]}\left(z_1^2+z_2^2\right)Q^N(dx,dy,dz,ds) \le L\end{aligned}$$

so Θ_L is closed, and hence compact. Note that we have $I(Z^N) < \infty$ implies $Z^N \in C([0,T];\mathcal{S}_{-w})$, $\forall N \in \mathbb{N}$ and by definition $Z \in C([0,T];\mathcal{S}_{-w})$. Thus if we could show convergence of $Z^N \to Z$ in $C([0,T];\mathcal{S}_{-w})$, we would have compactness of level sets of I as a rate function on $C([0,T];\mathcal{S}_{-w})$. However, such convergence is not immediately obvious, hence the need for the additional assumption (A'13).

To see that $\{Z^N\}_{N\in\mathbb{N}}$ is precompact, we have that since $Q^N \in P^*(Z^N)$, for each N (Z^N,Q^N) must satisfy Eq. (32). That is, for $\phi \in \mathcal{S}$ and $t \in [0,T]$:

$$\begin{aligned}\langle Z_t^N,\phi\rangle &= \int_0^t \langle Z_s^N, \bar{L}_{\mathcal{L}(X_s)}\phi\rangle ds + \int_0^t \langle B_s^N,\phi\rangle ds\\ \langle B_t^N,\phi\rangle &:= \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2}\Big(\sigma(x,y,\mathcal{L}(X_s))z_1\phi'(x)\\ &\quad + [\tau_1(x,y,\mathcal{L}(X_s))z_1+\tau_2(x,y,\mathcal{L}(X_s))z_2]\Phi_y(x,y,\mathcal{L}(X_s))\phi'(x)\Big)\\ &\quad Q_t^N(dx,dy,dz).\end{aligned}$$

Thus precompactness of $\{Z^N\}_{N\in\mathbb{N}}$ in $C([0,T];\mathcal{S}_{-r})$ follows in the exact same way as precompactness of $\{\tilde{Z}^k\}_{k\in\mathbb{N}}$ in $C([0,T];\mathcal{S}_{-w})$ in the proof of Proposition 10.1, but replacing m by w. Note that there we knew that $\tilde{Z}^k$ was in $C([0,T];\mathcal{S}_{-m})$ for each k, where here we only know $Z^N \in C([0,T];\mathcal{S}_{-w})$ for each N. Along the way, we get:

$$\sup_{N\in\mathbb{N}}\sup_{t\in[0,T]}\left\|Z_t^N\right\|_{-(w+2)}^2 \leq C(T).$$

To see that $Q \in P^*(Z)$, we identify the point-wise limit of $\langle Z^N,\phi\rangle$ to satisfy the desired equation, i.e. (32) with our specific choice of Q. This uniquely characterizes the limit along the whole sequence by Lemma 8.8. This gives P^*1. P^*2 follows immediately from Fatou's lemma. P^*3 and P^*4 follow from convergence of the measure implying convergence of the marginals and uniqueness of the decomposition into stochastic kernels (see [22] Theorems A.4.2 and A.5.4).

To see (32) with our specific choice of Q holds, we may by a density argument consider fixed $\phi \in C_c^\infty(\mathbb{R})$ and $t \in [0,T]$. Then:

$$\begin{aligned}
\langle Z_t,\phi\rangle &= \lim_{N\to\infty}\langle Z_t^N,\phi\rangle = \lim_{N\to\infty}\left\{\int_0^t \langle Z_s^N,\bar{L}_{\mathcal{L}(X_s)}\phi\rangle ds + \int_0^t \langle B_s^N,\phi\rangle ds\right\}\\
&= \int_0^t \lim_{N\to\infty}\langle Z_s^N,\bar{L}_{\mathcal{L}(X_s)}\phi\rangle ds + \lim_{N\to\infty}\int_0^t \langle B_s^N,\phi\rangle ds\\
&\quad \text{(by boundedness of } \sup_{N\in\mathbb{N}}\sup_{t\in[0,T]}\left\|Z_t^N\right\|_{-(w+2)}^2\\
&\qquad \text{and Dominated Convergence Theorem)}\\
&= \int_0^t \langle Z_s,\bar{L}_{\mathcal{L}(X_s)}\phi\rangle ds + \int_0^t \langle B_s,\phi\rangle ds,\\
&\qquad \text{since under assumption} (A'13)\, \bar{L}_{\mathcal{L}(X_s)}\phi \in \mathcal{S}_r, \forall s \in [0,T].
\end{aligned}$$

Here

$$\begin{aligned}
\langle B_t,\phi\rangle := \int_{\mathbb{R}\times\mathbb{R}\times\mathbb{R}^2}\Big(&\sigma(x,y,\mathcal{L}(X_s))z_1\phi'(x)\\
&+[\tau_1(x,y,\mathcal{L}(X_s))z_1+\tau_2(x,y,\mathcal{L}(X_s))z_2]\Phi_y(x,y,\mathcal{L}(X_s))\phi'(x)\Big)\\
&Q_t(dx,dy,dz),
\end{aligned}$$

and to pass to the second limit, we use that the integrand appearing in $\int_0^t\langle B_s^N,\phi\rangle ds$ is bounded by $C[|z_1|+|z_2|]$, and hence is uniformly integrable with respect to Q^N. □

11 Conclusions and future work

In this paper we have derived a moderate deviations principle for the empirical measure of a fully coupled multiscale system of weakly interacting particles in the joint limit as number of particles increases and averaging due to the multiscale structure takes over. Using weak convergence methods we have derived a variational form of the rate function and have rigorously shown that the rate function can take equivalent forms analogous to the one derived in the seminal paper [17].

In this paper we have assumed that the particles are in dimension one. It is of great interest to extend this work in the multidimensional case. One source of difficulty here is that in higher dimensions we would probably have to consider a different space for the fluctuation process to live in (see, e.g. [27] and [69]). This is because in higher dimensions the result that for each v, there is $w \geq v$ such that $\mathcal{S}_{-v} \to \mathcal{S}_{-w}$ is Hilbert–Schmidt breaks down, and the bound (9) no longer holds true. See [19] Section 5.1 for a further discussion of this. The trade-off with using these alternative spaces is that they often require higher moments of the particles and limiting McKean–Vlasov Equation in order to establish tightness—see, e.g. Section 4.7 in [27], where the proofs depend crucially on Lemma 3.1 (even in one dimension this would require having bounded 8'th moments of the controlled particles $\tilde{X}^{i,N,\epsilon}$, with the required number of moments increasing with the dimension). This would seem to require strong assumptions on the coefficients Eq. (1) even in the absence of multiscale structure, since the controls are a priori only bounded in L^2.

Another potentially interesting direction is to derive the moderate deviations principle for the stochastic current. See [59] for some related results in the direction of large deviations for an interacting particle system in the joint mean field and small-noise limit. Also, we are hopeful that the results of this paper can also be used for the construction of provably-efficient importance sampling schemes for the computation of rare events for statistics of weakly interacting diffusions that are relevant to the moderate-deviations scaling. Lastly, as we also mentioned in the introduction, we believe that the results of this paper can be used to study dynamical questions related to phase transitions in the spirit of [16].

Appendix A. A list of technical notation

Here we provide a list of frequently used notation for the various processes, spaces, operators, ect. used throughout this manuscript for convenient reference. Other, more standard notation is introduced following Eq. (9) in Sect. 2.

- ϵ is the scale separation parameter which decreases to 0 as $N \to \infty$. N is the number of particles. $a(N)$ is moderate deviations the scaling sequence such that $a(N) \downarrow 0$ and $a(N)\sqrt{N} \to \infty$.
- $(X^{i,\epsilon,N}, Y^{i,\epsilon,N})$ is the slow–fast system of particles from Eq. (1).
- $\mu^{\epsilon,N}$ from Eq. (2) is the empirical measure on the slow particles $X^{i,\epsilon,N}$.
- X_t is the limiting averaged McKean–Vlasov Equation from Eq. (25). $\mathcal{L}(X_t)$ denotes its Law.
- Z^N is the fluctuations process from Eq. (3) for which we derive a large deviations principle.
- $(\tilde{X}^{i,\epsilon,N}, \tilde{Y}^{i,\epsilon,N})$ are the controlled slow–fast interacting particles from Eq. (55).
- $\tilde{\mu}^{\epsilon,N}$ is the empirical measure on the controlled slow particles $\tilde{X}^{i,\epsilon,N}$ from Eq. (54).
- $\tilde{Z}^N$ is the controlled fluctuations process from Eq. (54).

- Q^N are the occupation measures from Eq. (56).
- $(\bar{X}^{i,\epsilon}, \bar{Y}^{i,\epsilon})$ are the IID slow–fast McKean–Vlasov equations from Eq. (57). $\bar{X}^{\epsilon}$ is a random process with law Equal to that of the $\bar{X}^{i,\epsilon}$'s.
- $\bar{\mu}^{\epsilon,N}$ from Eq. (58) is the empirical measure on N of the IID slow particles $\bar{X}^{i,\epsilon}$.
- $\mathcal{P}_2(\mathbb{R})$ is the space of square integrable probability measures with the 2-Wasserstein metric $\mathbb{W}_2$ (Definition D.1).
- $M_T(\mathbb{R}^d)$ is the space of measures Q on $\mathbb{R}^d \times [0, T]$ such that $Q(\mathbb{R}^d \times [0, t]) = t, \forall t \in [0, T]$ equipped with the topology of weak convergence.
- For $p \in \mathbb{N}$, $\mathcal{S}_p$ is the completion of $\mathcal{S}$ with respect to $\|\cdot\|_p$ (see Eq. (4)) and $\mathcal{S}_{-p} = \mathcal{S}_p'$ the dual space of $\mathcal{S}_p$. We prove tightness of $\{\tilde{Z}^N\}_{N\in\mathbb{N}}$ in $C([0, T]; \mathcal{S}_{-m})$ for the choice of m found in Eq. (5), the Laplace Principle on $C([0, T]; \mathcal{S}_{-w})$ for the choice of w found in Eq. (6), and compactness of level sets of the rate function on $C([0, T]; \mathcal{S}_{-r})$ for the choice of r found in Eq. (7).
- For $n \in \mathbb{N}$, $|\cdot|_n$ is the sup norm defined in Eq. (8), which is related to $\|\cdot\|_{n+1}$ via Eq. (9).
- For $G : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ and $\nu \in \mathcal{P}_2(\mathbb{R})$, $\partial_\mu G(\nu)[\cdot] : \mathbb{R} \to \mathbb{R}$ denotes the Lions derivative of G at the point ν (Definition D.1) and $\frac{\delta}{\delta m}G(\nu)[\cdot] : \mathbb{R} \to \mathbb{R}$ denotes the Linear Functional Derivative of G at the point ν (Definition D.4).
- For $G : \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$, we use $D^{(n,l,\beta)}G$ to denote multiple derivatives of G in space and measure in the multi-index notation of Definition 2.1. Spaces (denoted by $\mathcal{M}$ with some sub or super-scripts) containing functions with different regularity of such mixed derivatives are found in Definition 2.4. When $G : \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$, polynomial growth of such derivatives in G's second coordinate, denoted by $q_G(n, l, \beta)$ or $\tilde{q}_G(n, l, \beta)$, are defined as in Eqs. (13) and (15).
- $L_{x,\mu}$ is the frozen generator associated to the fast particles from Eq. (18). π denotes its unique associated invariant measure from Eq. (20) and Φ denotes the solution to the associated Poisson Eq. (22).
- For $\nu \in \mathcal{P}_2(\mathbb{R})$ $\bar{L}_\nu$ is the linearized generator of the limiting averaged McKean–Vlasov Equation X_t at ν and is defined in Eq. (32).
- $\bar{\gamma}, \bar{D}$ from Eq. (24) are the drift and diffusion coefficients of the limiting averaged McKean–Vlasov Equation X_t, and are defined in terms of $\gamma_1, D_1, \gamma, D : \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ from Eq. (23).

Appendix B. A priori bounds on moments of the controlled process (55)

In this Appendix, we fix any controls satisfying the bound (53) and provide moment bounds on the fast component of the controlled particles (55). These are needed, among other places, to handle possible growth lack of boundedness in y of functions appearing in the remainders in the ergodic-type theorems of Sect. 6. The details of the proofs can be found in the extended arXiv version of the paper (arXiv:2202.08403).

Lemma B.1 *Under assumptions (A1)–(A2), (A4), and (A5), we have there is $C \geq 0$ such that:*

$$\sup_{N\in\mathbb{N}} \frac{1}{N}\sum_{i=1}^{N} \sup_{t\in[0,T]} \mathbb{E}\Big[|\tilde{Y}_t^{i,\epsilon,N}|^2\Big] \leq C + |\eta^y|^2.$$

Proof The proof of this lemma is omitted as it follows very closely the proof of Lemma 4.1 in [5]. The main additional step that is needed is to average over the N particles in order to obtain the final bound. □

Lemma B.2 *Under assumptions (A1)–(A2), (A4),and (A5), we have:*

$$\frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\Big[\sup_{0\leq t\leq T} |\tilde{Y}_t^{i,\epsilon,N}|^2\Big] \leq |\eta^y|^2 + C(\rho)\Big[1 + \epsilon^{-\rho} + \frac{1}{a^2(N)N}\Big]$$

for all $\rho \in (0, 2)$.

Proof The proof follows along the lines of that of Lemma B.4 in [44] and thus it is omitted here. Note that this is where the near-Ornstein-Uhlenbeck structure assumed in (16) plays an important role. □

Lemma B.3 *Under assumptions (A1)–(A2), (A4), and (A5), we have:*

$$\sup_{N\in\mathbb{N}} \sup_{0\leq t\leq T} \mathbb{E}\Bigg[\Bigg(\frac{1}{N}\sum_{i=1}^{N} |\tilde{Y}_t^{i,\epsilon,N}|^2\Bigg)^2\Bigg] < \infty.$$

Proof The proof is very similar to that of Lemma 4.1 in [5], but we need in addition to use Lemma B.2. In particular, one obtains:

$$\begin{aligned}
&\mathbb{E}\Bigg[\Bigg(\frac{1}{N}\sum_{i=1}^{N} |\tilde{Y}_t^{i,\epsilon,N}|^2\Bigg)^2\Bigg] \\
&\leq |\eta^y|^4 + C\Big[1 + \frac{1}{N}\Big] \sup_{s\in[0,T]} \mathbb{E}\Bigg[\frac{1}{N}\sum_{i=1}^{N} |\tilde{Y}_s^{i,\epsilon,N}|^2\Bigg] \\
&\quad + \frac{C}{a^2(N)N}\mathbb{E}\Bigg[\sup_{s\in[0,T]}\Bigg(\frac{1}{N}\sum_{i=1}^{N} |\tilde{Y}_s^{i,\epsilon,N}|^2\Bigg)\Bigg] \\
&\leq C(\rho)[1 + \frac{1}{N} + \frac{1}{a^2(N)N} + \frac{1}{a^4(N)N^2} + \frac{1}{a^2(N)N\epsilon^\rho}]
\end{aligned}$$

by Lemmas B.1 and B.2. This is where the assumption that we can take $\rho/2 \in (0, 1)$ such that $a(N)\sqrt{N}\epsilon^{\rho/2} \to \lambda \in (0, \infty]$ is needed. Details are omitted given that they are standard and for brevity of presentation. □

Appendix C. Regularity of the Poisson equations

As discussed in Remark 2.6, there is a current gap in the literature regarding rates of polynomial growth of derivatives of the Poisson equations used in Sect. 6. Nevertheless, it is important to verify that the assumptions imposed on these solutions in Sect. 2 are non-empty. For the reasons outlined in Remark 2.6, we handle the case of the 1D Poisson equations from Eqs. (22), (64) and the Multi-Dimensional Poisson Eqs. (28) and (63), separately in Sects. 1 and 2 below. In Sect. 3 we provide specific examples where the Assumptions in Sect. 2 hold. Proofs with details of the results presented in this section can be found in the extended arXiv version of this paper (arXiv:2202.08403); details are omitted here due to the standard form of the arguments and for reasons of brevity of presentation.

Results for the 1-dimensional Poisson equation

Throughout this subsection we assume (A1) and (A2). Recall the frozen generator $L_{x,\mu}$ from Eq. (18), the invariant measure π from Eq. (20), the multi-index derivative notation and associated spaces of functions from Definitions 2.1 and 2.4, and the definition of a from Eq. (18).

Lemma C.1 *Consider $B : \mathbb{R} \times \mathbb{R} \times \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}$ continuous such that*

$$\int_{\mathbb{R}} B(x, y, \mu)\pi(dy; x, \mu) = 0, \forall x \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R})$$

and $|B(x, y, \mu)| = O(|y|^{q_B})$ for $q_B \in \mathbb{R}$ uniformly in x, μ as $|y| \to \infty$. Then there exists a unique classical solution $u : \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ to

$$L_{x,\mu}u(x, y, \mu) = B(x, y, \mu)$$

such that u is continuous in $(x, y, \mathbb{W}_2)$, $\int_{\mathbb{R}} u(x, y, \mu)\pi(dy; x, \mu) = 0$, and u has at most polynomial growth as $|y| \to \infty$.

In addition,

$$|u(x, y, \mu)| = O(|y|^{q_B}) \text{ for } q_B \neq 0; \text{ if } q_B = 0, \text{ then } |u(x, y, \mu)| = O(\ln(|y|)$$
$$|u_y(x, y, \mu)| = O(|y|^{q_B - 1}), \quad |u_{yy}(x, y, \mu)| = O(|y|^{q_B})$$

as $|y| \to \infty$ uniformly in x, y, μ.

Furthermore if $B(x, y, \mu)$ is Lipschitz continuous in y uniformly in x, μ (so that necessarily $q_B \leq 1$), then so are u, u_y, u_{yy}.

Proof This follows by Proposition A.4 in [33], since in our setting Condition 2.1 (i) holds with $\alpha = 1$ and assumption A.3 holds with $\theta = 1$. The statement about Lipschitz continuity of u and its derivatives follows from the Lipschitz continuity of a, f under assumptions (A1) and (A2) and Theorem 9.19 in [37]. □

Lemma C.2 *Consider* $B : \mathbb{R} \times \mathbb{R} \times \mathbb{P}_2(\mathbb{R}) \to \mathbb{R}$ *continuous such that*

$$\int_{\mathbb{R}} B(x, y, \mu)\pi(dy; x, \mu) = 0, \forall x \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R}).$$

Suppose that for some complete collection of multi-indices $\boldsymbol{\zeta}$ *that* $B, a, f \in \mathcal{M}_p^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$. *Then for the unique classical solution* $u : \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ *to*

$$L_{x,\mu}u(x, y, \mu) = B(x, y, \mu)$$

such that u is continuous in $(x, y, \mathbb{W}_2)$, $\int_{\mathbb{R}} u(x, y, \mu)\pi(dy; x, \mu) = 0$, *and u has at most polynomial growth as* $|y| \to \infty$ *(which exists by Lemma C.1),*

(1) $u, u_y, u_{yy} \in \mathcal{M}_p^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$.
(2) If $B, a, f \in \mathcal{M}_{\boldsymbol{\delta},p}^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R})) \cap \mathcal{M}_p^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$, *then* $u, u_y, u_{yy} \in \mathcal{M}_{\boldsymbol{\delta},p}^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$.
(3) If $B, a, f \in \mathcal{M}_{p,L}^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$, *then* $u, u_y, u_{yy} \in \mathcal{M}_{p,L}^{\boldsymbol{\zeta}}(\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}))$.

Moreover, if we suppose that for all multi-indices $(n, l, \boldsymbol{\beta}) \in \boldsymbol{\zeta}$, $q_f(n, l, \boldsymbol{\beta}) \leq 1$ *and* $q_a(n, l, \boldsymbol{\beta}) \leq 0$ *(using here the notation of 13), we have control on the growth rate of the derivatives of u in terms of those of B. In particular, for any* $(n, l, \boldsymbol{\beta}) \in \boldsymbol{\zeta}$*:*

$$q_u(n, l, \boldsymbol{\beta}) \leq \max\{q_B(k, j, \boldsymbol{\alpha}(k)) : \alpha(k) \in \binom{\boldsymbol{\beta}}{k}, k \leq n, j \leq l\},$$

when the right hand side is nonzero, and the corresponding term grows at most like $\ln(|y|)$ *as* $|y| \to \infty$ *when the left hand side is zero. In addition,* $q_{u_y}(n, l, \boldsymbol{\beta}) \leq q_u(n, l, \boldsymbol{\beta}) - 1$, *and* $q_{u_{yy}}(n, l, \boldsymbol{\beta}) \leq q_u(n, l, \boldsymbol{\beta})$, *for all* $(n, l, \boldsymbol{\beta}) \in \boldsymbol{\zeta}$.

Proof For 1), the proof essentially uses the same tools and a similar method to Lemma A.2 in [5], so we will only check this in the case for $(n, l, \boldsymbol{\beta}) = (0, 1, 0)$ and then comment on how the rest of the terms follow. Importantly, Lemma A.2 in [5] only assumes existence and polynomial growth of derivatives of the solution u up to one order less than the derivative obtained there.

The result for $(n, l, \boldsymbol{\beta}) = (0, 0, 0)$ is just another way of writing Lemma C.1.

The differentiability and continuity of the derivatives is immediate via the explicit representation for u

$$v(x, y, \mu) = \int_{-\infty}^{y} \frac{1}{a(x, \bar{y}, \mu)\pi(\bar{y}; x, \mu)} \left[\int_{-\infty}^{\bar{y}} B(x, \tilde{y}, \mu)\pi(\tilde{y}; x, \mu)d\tilde{y}\right]d\bar{y}$$
$$\pi(y; x, \mu) = \frac{Z(x, \mu)}{a(x, y, \mu)} \exp\left(\int_0^y \frac{f(x, \bar{y}, \mu)}{a(x, \bar{y}, \mu)}d\bar{y}\right) \tag{72}$$

where $Z^{-1}(x, \mu) := \int_{\mathbb{R}} \frac{1}{a(x,y,\mu)} \exp\left(\int_0^y \frac{f(x,\bar{y},\mu)}{a(x,\bar{y},\mu)}d\bar{y}\right)dy$ is the normalizing constant.

To obtain the rate of polynomial growth of u_x, we differentiate the equation that u satisfies to get

$$
\begin{aligned}
L_{x,\mu}u_x(x,y,\mu) &= B_x(x,y,\mu) - f_x(x,y,\mu)u_y(x,y,\mu) - a_x(x,y,\mu)u_{yy}(x,y,\mu)\\
&= B_x(x,y,\mu) - L^{(0,1,0)}_{x,\mu}u(x,y,\mu)
\end{aligned}
$$

in the notation of Lemma A.2 in [5]. But by the centering condition on B, we have that letting $B = h$ in Lemma A.2 in [5], $u = v$ in the statement of that same lemma. Thus we have

$$
\begin{aligned}
&\int_{\mathbb{R}} \Big(B_x(x,y,\mu) - L^{(0,1,0)}_{x,\mu}u(x,y,\mu)\Big)\pi(dy;x,\mu)\\
&\quad = \frac{\partial}{\partial x}\int_{\mathbb{R}} B(x,y,\mu)\pi(dy;x,\mu) = 0,
\end{aligned}
$$

and the inhomogeneity of the elliptic PDE that u_x solves, in fact obeys the centering condition, and hence Lemma C.1 applies. From the same lemma we already know that $q_{u,y}(0,0,0) = q_B(0,0,0) - 1$ and $q_{yy} = q_B(0,0,0)$. This establishes that u_x grows at most polynomially in y uniformly in x, μ. Under the additional assumptions that $q_f(0,1,0) \leq 1$ and $q_a(0,1,0) \leq 0$, we have the inhomogeneity is $O(|y|^{q_B(0,0,0)\vee q_B(0,1,0)})$. So by Lemma C.1, $q_{u,x} = q_B(0,0,0)\vee q_B(0,1,0)$, $q_{u,x,y} = q_B(0,0,0)\vee q_B(0,1,0) - 1$, $q_{u,x,y,y} = q_B(0,0,0)\vee q_B(0,1,0)$.

All of the bounds work in the same way, with the inhomogeneity of the elliptic PDE of the desired derivative of u solves being the integrand of the expression for the corresponding derivative of $\bar{B}(x,y,\mu)$ from Lemma A.2 in [5]. Put explicitly:

$$
\begin{aligned}
&L_{x,\mu}D^{(n,l,\boldsymbol{\beta})}u(x,y,\mu)[z_1,\ldots,z_n] = D^{(n,l,\boldsymbol{\beta})}B(x,y,\mu)[z_1,\ldots,z_n]-\\
&-\sum_{k=0}^{n}\sum_{j=0}^{l}\sum_{\boldsymbol{p}_k} C_{(\boldsymbol{p}_k,j,n,l)}L^{(k,j,\boldsymbol{\alpha}(\boldsymbol{p}_k))}_{x,\mu}[z_{\boldsymbol{p}_k}]D^{(n-k,l-j,\boldsymbol{\alpha}(\boldsymbol{p}'_{n-k}))}u(x,y,\mu)[z_{\boldsymbol{p}'_{n-k}}],
\end{aligned}
\tag{73}
$$

where the constants $C_{(\boldsymbol{p}_k,j,n,l)}$ are defined inductively in Remark A.3 in [5] and $L^{(k,j,\boldsymbol{\alpha}(\boldsymbol{p}_k))}_{x,\mu}[z_{\boldsymbol{p}_k}]$ is the differential operator acting on $\phi \in C^2_b(\mathbb{R})$ by

$$
\begin{aligned}
&L^{(k,j,\boldsymbol{\alpha}(\boldsymbol{p}_k))}_{x,\mu}[z_{\boldsymbol{p}_k}]\phi(y)\\
&\quad = D^{(k,j,\boldsymbol{\alpha}(\boldsymbol{p}_k))}f(x,y,\mu)[z_{\boldsymbol{p}_k}]\phi'(y) + D^{(k,j,\boldsymbol{\alpha}(\boldsymbol{p}_k))}a(x,y,\mu)[z_{\boldsymbol{p}_k}]\phi''(y).
\end{aligned}
$$

The first y derivative of a lower order derivative in a parameter of u in the inhomogeneity is always multiplied by a derivative of f, and so if that derivative of f grows at most linearly in y, the growth of that term is at most that of that lower order derivative of u, and same for the second y derivative in a parameter of u in the inhomogeneity, which multiplied by a bounded lower order derivative of a. Thus it is clear the result follows by proceeding inductively on n, l.

The proof for 2) follows in the exact same way. We note here that Lemma A.2 in [5] holds for the linear functional derivatives $\boldsymbol{\delta}^{(n,l,\boldsymbol{\beta})}$ in place of the Lions derivatives $D^{(n,l,\boldsymbol{\beta})}$ if in addition we assume $h, a, f \in \mathcal{M}_{\boldsymbol{\delta},p}^{\boldsymbol{\zeta}}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$,$v_y, v_{yy} \in \mathcal{M}_{\boldsymbol{\delta},p}^{\boldsymbol{\zeta}'}(\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$.

The proof for 3) is similar to step 4 in the proof of Theorem 2.1 in [64]. For the case $(n, l, \boldsymbol{\beta}) = (0, 0, 0)$, we first note that

$$
\begin{aligned}
&L_{x,\mu_1}[u(x, y, \mu_1) - u(x, y, \mu_2)] = B(x, y, \mu_1) - L_{x,\mu_1}u(x, y, \mu_2)\\
&= B(x, y, \mu_1) - B(x, y, \mu_2) - [L_{x,\mu_1} - L_{x,\mu_2}]u(x, y, \mu_2).
\end{aligned}
$$

By the transfer formula in Lemma A.2 of [5] we have

$$
\begin{aligned}
&\int_{\mathbb{R}} \big(B(x, y, \mu_1) - B(x, y, \mu_2) - [L_{x,\mu_1} - L_{x,\mu_2}]u(x, y, \mu_2)\big)\,\pi(dy, x, \mu_1) =\\
&\quad = \int_{\mathbb{R}} B(x, y, \mu_1)\pi(dy; x, \mu_1) - \int_{\mathbb{R}} B(x, y, \mu_2)\pi(dy; x, \mu_2) = 0,
\end{aligned}
$$

so in fact the inhomogeneity in the above Poisson equation is centered. Now, rather than using Lemma C.1, we apply [60] Theorem 2 to get there is $k \in \mathbb{R}$ sufficiently large and $C > 0$ such that for all $x \in \mathbb{R}$, $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R})$

$$
\begin{aligned}
&\sup_{y\in\mathbb{R}} \frac{|u(x, y, \mu_1) - u(x, y, \mu_2)|}{(1 + |y|)^k}\\
&\quad \le C \sup_{y\in\mathbb{R}} \frac{\Big|B(x, y, \mu_1) - B(x, y, \mu_2) - [L_{x,\mu_1} - L_{x,\mu_2}]u(x, y, \mu_2)\Big|}{(1 + |y|)^k}\\
&\le C\mathbb{W}_2(\mu_1, \mu_2)
\end{aligned}
$$

by the Lipschitz assumptions on B, f, a. Thus for all $x, y \in \mathbb{R}$, $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R})$,

$$
|u(x, y, \mu_1) - u(x, y, \mu_2)| \le C\mathbb{W}_2(\mu_1, \mu_2)(1 + |y|)^k.
$$

To see then that there are $k', k'' \in \mathbb{R}$, $C', C'' > 0$ such that

$$
\begin{aligned}
|u_y(x, y, \mu_1) - u_y(x, y, \mu_2)| &\le C\mathbb{W}_2(\mu_1, \mu_2)(1 + |y|)^{k'}\\
|u_{yy}(x, y, \mu_1) - u_{yy}(x, y, \mu_2)| &\le C\mathbb{W}_2(\mu_1, \mu_2)(1 + |y|)^{k''},
\end{aligned}
$$

we can apply the result of [33] Lemma B.1 and Remark B.2, and the last line of Proposition A.4 in the same reference.

The proof with μ_1, μ_2 replaced by x_1, x_2 follows in the same way.

For the Lipschitz property in z, we first recall that for all $x, y, z \in \mathbb{R}$, $\mu \in \mathcal{P}_2(\mathbb{R})$

$$
L_{x,\mu}D^{(1,0,0)}u(x, y, \mu)[z] = D^{(1,0,0)}B(x, y, \mu)[z] - L_{x,\mu}^{(1,0,0)}[z]u(x, y, \mu)
$$

so

$$L_{x,\mu}\left[D^{(1,0,0)}u(x,y,\mu)[z_1]-D^{(1,0,0)}u(x,y,\mu)[z_2]\right]$$
$$=D^{(1,0,0)}B(x,y,\mu)[z_1]-L^{(1,0,0)}_{x,\mu}[z_1]u(x,y,\mu)$$
$$-\left[D^{(1,0,0)}B(x,y,\mu)[z_2]-L^{(1,0,0)}_{x,\mu}[z_2]u(x,y,\mu)\right].$$

By the transfer formula in Lemma A.2 of [5] we have for all $x,z\in\mathbb{R}$, $\mu\in\mathcal{P}_2(\mathbb{R})$:

$$\int_{\mathbb{R}}\left(D^{(1,0,0)}B(x,y,\mu)[z]-L^{(1,0,0)}_{x,\mu}[z]u(x,y,\mu)\right)\pi(dy;x,\mu)$$
$$=D^{(0,1,0)}\int_{\mathbb{R}}B(x,y,\mu)\pi(dy;x,\mu)[z]=0,$$

so the inhomogeneity in the Poisson equation above in centered. Thus, using the same argument as for the other Lipschitz continuity as well as the fact that $D^{(1,0,0)}B$, $D^{(1,0,0)}f$, $D^{(1,0,0)}a$ are Lipschitz in z and u_y, u_{yy} grow at most polynomially in y, we get there is $K\in\mathbb{R}$ and $C>0$ such that

$$\left|D^{(1,0,0)}u(x,y,\mu)[z_1]-D^{(1,0,0)}u(x,y,\mu)[z_2]\right|\le C|z_1-z_2|(1+|y|)^k,$$

and similarly for $D^{(1,0,0)}u_y$ and $D^{(1,0,0)}u_{yy}$.

Then using the Poisson equation the derivatives satisfy given in Eq. (73), we can iteratively use this same approach, along with the fact that products and sums of functions in $\mathcal{M}^{\boldsymbol{\zeta}}_{p,L}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ remain in $\mathcal{M}^{\boldsymbol{\zeta}}_{p,L}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$, to achieve the full result.

□

Lemma C.3 *Suppose that for some complete collection of multi-indices* $\boldsymbol{\zeta}$ *that* $h,a,f\in\mathcal{M}^{\boldsymbol{\zeta}}_p(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$. *Then* $\int_{\mathbb{R}}h(x,y,\mu)\pi(dy;x,\mu)\in\mathcal{M}^{\boldsymbol{\zeta}}_b(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$. *If in addition,* $h,a,f\in\mathcal{M}^{\boldsymbol{\zeta}}_{\delta,p}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$, *then* $\int_{\mathbb{R}}h(x,y,\mu)\pi(dy;x,\mu)\in\mathcal{M}^{\boldsymbol{\zeta}}_{\delta,b}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$. *Further, if we have that* $h,a,f\in\mathcal{M}^{\boldsymbol{\zeta}}_{p,L}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$, *then* $\int_{\mathbb{R}}h(x,y,\mu)\pi(dy;x,\mu)\in\mathcal{M}^{\boldsymbol{\zeta}}_{b,L}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$.

Proof This follows via Lemmas A.2 in [5] and C.1 in a similar way to Lemma C.2. The details are omitted. □

Result for the d-dimensional Poisson equation

Lemma C.4 *Suppose* $F:\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R})\to\mathbb{R}^d$, $G:\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R})\to\mathbb{R}$, $\tau:\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R})\to\mathbb{R}^{d\times m}$

$$|F(x_1,y_1,\mu_1)-F(x_2,y_2,\mu_2)|+|G(x_1,y_1,\mu_1)-G(x_2,y_2,\mu_2)|$$

$$
\begin{aligned}
&+ \|\tau(x_1, y_1, \mu_1) - \tau(x_2, y_2, \mu_2)\| \\
&\le C[|x_1 - x_2| + |y_1 - y_2| + \mathbb{W}_2(\mu_1, \mu_2)], \forall x_1, x_2, y_1, y_2 \in \mathbb{R}^d, \mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d),
\end{aligned}
$$

and there exists $\beta > 0$ *such that for all* $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R})$*:*

$$
\begin{aligned}
&2\langle F(x, y_1, \mu) - F(x, y_2, \mu), (y_1 - y_2)\rangle + 3\,\|\tau(x, y_1, \mu) - \tau(x, y_2, \mu)\|^2 \\
&\quad \le -\beta|y_1 - y_2|^2.
\end{aligned}
$$

Here $\langle \cdot, \cdot \rangle$ *is denoting the inner product on* $\mathbb{R}^d$ *and* $\|\cdot\|$ *the matrix norm. Also assume that* τ *is bounded, and*

$$
|G(x, y, \mu)|, |F(x, y, \mu)| \le C(1 + |y|), \forall x \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}).
$$

Define the differential operator $\tilde{\mathcal{L}}_{x,\mu}$ *which for each* $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R})$ *acts on* $\phi \in C_b^2(\mathbb{R})$ *by*

$$
\tilde{\mathcal{L}}_{x,\mu}\phi(y) := F(x, y, \mu) \cdot \nabla\phi(y) + \frac{1}{2}\tau\tau^\top(x, y, \mu) : \nabla^2\phi(y).
$$

Then there is a unique invariant measure $\nu(\cdot; x, \mu)$ *associated to* $\tilde{\mathcal{L}}_{x,\mu}$ *for each* x, μ*, and we assume the centering condition on* G*:*

$$
\int_{\mathbb{R}^d} G(x, y, \mu)\nu(dy; x, \mu) = 0, \forall x \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}).
$$

Finally, we assume the below derivatives all exist, are jointly continuous in $(x, y, \mathbb{W}_2)$ *and auxiliary variables where applicable, and satisfy:*

$$
\begin{aligned}
&\sup_{x\in\mathbb{R}^d, \mu\in\mathcal{P}_2(\mathbb{R})} \max\{|\partial_x G(x, \mu, y_1) - \partial_x G(x, \mu, y_2)|, |\partial_y G(x, \mu, y_1) - \partial_y G(x, \mu, y_2)|\} \\
&\le C|y_1 - y_2| \\
&\sup_{x\in\mathbb{R}^d, \mu\in\mathcal{P}_2(\mathbb{R})} \max\{\left\|\partial_x^2 G(x, \mu, y_1) - \partial_x^2 G(x, \mu, y_2)\right\|, \left\|\partial_y^2 G(x, \mu, y_1) - \partial_y^2 G(x, \mu, y_2)\right\|\} \\
&\le C|y_1 - y_2| \\
&\sup_{x\in\mathbb{R}^d, \mu\in\mathcal{P}_2(\mathbb{R})} \left\|\partial_x\partial_y G(x, \mu, y_1) - \partial_x\partial_y G(x, \mu, y_2)\right\| \\
&\le C|y_1 - y_2| \\
&\sup_{x\in\mathbb{R}^d, \mu\in\mathcal{P}_2(\mathbb{R})} \left\|\partial_\mu\partial_x G(x, \mu, y_1)[\cdot] - \partial_\mu\partial_x G(x, \mu, y_2)[\cdot]\right\|_{L^2(\mathbb{R},\mu)} \\
&\le C|y_1 - y_2| \\
&\sup_{x\in\mathbb{R}^d, \mu\in\mathcal{P}_2(\mathbb{R})} \left\|\partial_z\partial_\mu G(x, \mu, y_1)[\cdot] - \partial_z\partial_\mu G(x, \mu, y_2)[\cdot]\right\|_{L^2(\mathbb{R},\mu)} \le C|y_1 - y_2| \\
&\sup_{x\in\mathbb{R}^d, \mu\in\mathcal{P}_2(\mathbb{R})} \left\|\partial_\mu\partial_x G(x, \mu, y_1)[\cdot] - \partial_\mu\partial_x G(x, \mu, y_2)[\cdot]\right\|_{L^2(\mathbb{R},\mu)} \le C|y_1 - y_2|
\end{aligned}
$$

$$\sup_{x\in\mathbb{R}^d,\mu\in\mathcal{P}_2(\mathbb{R})} \left\| \partial_\mu \partial_y G(x,\mu,y_1)[\cdot] - \partial_\mu \partial_y G(x,\mu,y_2)[\cdot] \right\|_{L^2(\mathbb{R},\mu)} \le C|y_1-y_2|$$
$$\sup_{x\in\mathbb{R}^d,\mu\in\mathcal{P}_2(\mathbb{R})} \left\| \partial_\mu^2 G(x,\mu,y_1)[\cdot,\cdot] - \partial_\mu^2 G(x,\mu,y_2)[\cdot,\cdot] \right\|_{L^2(\mathbb{R},\mu)\otimes L^2(\mathbb{R},\mu)} \le C|y_1-y_2|$$
$$\sup_{x,y\in\mathbb{R}^d,\mu\in\mathcal{P}_2(\mathbb{R})} \max\left\{ \left\|\partial_y\partial_x G(x,\mu,y)\right\|, \left\|\partial_y^2 G(x,\mu,y)\right\|, \left\|\partial_\mu\partial_y G(x,\mu,y)[\cdot]\right\|_{L^2(\mathbb{R},\mu)} \right\} \le C$$

and same for G replaced by F and τ, and in addition

$$\sup_{x,y\in\mathbb{R}^d,\mu\in\mathcal{P}_2(\mathbb{R})} \max\Bigg\{ \left\|\partial_x^2 F(x,\mu,y)\right\|, \left\|\partial_x^2 \tau(x,\mu,y)\right\|,$$
$$\left\|\partial_z\partial_\mu F(x,\mu,y)[\cdot]\right\|_{L^2(\mu,\mathbb{R})}, \left\|\partial_z\partial_\mu \tau(x,\mu,y)[\cdot]\right\|_{L^2(\mu,\mathbb{R})},$$
$$\left\|\partial_\mu\partial_x F(x,\mu,y)[\cdot]\right\|_{L^2(\mu,\mathbb{R})}, \left\|\partial_\mu\partial_x \tau(x,\mu,y)[\cdot]\right\|_{L^2(\mu,\mathbb{R})},$$
$$\left\|\partial_\mu^2 F(x,\mu,y)[\cdot,\cdot]\right\|_{L^2(\mu,\mathbb{R})\otimes L^2(\mu,\mathbb{R})},$$
$$\left\|\partial_\mu^2 \tau(x,\mu,y)[\cdot,\cdot]\right\|_{L^2(\mu,\mathbb{R})\otimes L^2(\mu,\mathbb{R})} \Bigg\} \le C.$$

Then the partial differential equation

$$\tilde{\mathcal{L}}_{x,\mu}\chi(x,y,\mu) = -G(x,y,\mu)$$

admits a unique classical solution $\chi : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ which has all of the above derivatives, and

$$\sup_{x\in\mathbb{R}^d,\mu\in\mathcal{P}_2(\mathbb{R})} \max\Bigg\{ |\chi(x,y,\mu)|, \|\partial_x\chi(x,y,\mu)\|, \left\|\partial_\mu\chi(x,y,\mu)[\cdot]\right\|_{L^2(\mathbb{R},\mu)},$$
$$\left\|\partial_x^2\chi(x,y,\mu)\right\|, \left\|\partial_z\partial_\mu\chi(x,y,\mu)[\cdot]\right\|_{L^2(\mathbb{R},\mu)},$$
$$\left\|\partial_\mu\partial_x\chi(x,y,\mu)[\cdot]\right\|_{L^2(\mathbb{R},\mu)}, \left\|\partial_\mu^2\chi(x,y,\mu)[\cdot,\cdot]\right\|_{L^2(\mathbb{R},\mu)\otimes L^2(\mathbb{R},\mu)} \Bigg\}$$
$$\le C(1+|y|), \forall y \in \mathbb{R}^2,$$
$$\sup_{x,y\in\mathbb{R}^d,\mu\in\mathcal{P}_2(\mathbb{R})} \max\Bigg\{ \left\|\partial_y\chi(x,y,\mu)\right\|, \left\|\partial_y^2\chi(x,y,\mu)\right\|,$$
$$\left\|\partial_x\partial_y\chi(x,y,\mu)\right\|, \left\|\partial_\mu\partial_y\chi(x,y,\mu)\right\|_{L^2(\mathbb{R},\mu)} \le C.$$

Moreover, if all listed derivatives of F, G, τ are jointly continuous in $(x, y, \mathbb{W}_2)$, then so are listed derivatives of χ.

In the notation of Definition 2.4, this conclusion reads $\chi \in \tilde{\mathcal{M}}_p^{\tilde{\boldsymbol{\zeta}}}(\mathbb{R}^2\times\mathbb{R}^2\times\mathcal{P}_2(\mathbb{R}))$, $\chi_y \in \tilde{\mathcal{M}}_p^{\tilde{\boldsymbol{\zeta}}_1}(\mathbb{R}^2\times\mathbb{R}^2\times\mathcal{P}_2(\mathbb{R}))$, $\chi_{yy} \in \tilde{\mathcal{M}}_p^{(0,0,0)}(\mathbb{R}^2\times\mathbb{R}^2\times\mathcal{P}_2(\mathbb{R}))$ with $\tilde{q}_\chi(n,l,\boldsymbol{\beta}) \le 1, \forall(n,l,\boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}$, $\tilde{q}_{\chi_y}(n,l,\boldsymbol{\beta}) \le 0, \forall(n,l,\boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}_1$, and $\tilde{q}_{\chi_{yy}}(0,0,0) \le 0$ where $\tilde{\boldsymbol{\zeta}}, \tilde{\boldsymbol{\zeta}}_1$ are as in Eq. (27).

Proof The arguments here follow closely those in [65]. Existence and uniqueness for the invariant measure and strong solution from the Poisson equation are the subject of the beginning of Sects. 3.3 and 4.1 of [65]. The bound for $\chi, \partial_y\chi, \partial_x\chi, \partial_\mu\chi, \partial_x^2\chi$, and $\partial_z\partial_\mu\chi$ is also the subject of Proposition 4.1/Section 6.3 of [65], where we made the modification that τ, F (their g, f respectively) are bounded in x, μ, from which one can see that the bound on the solution is also uniform in x, μ.

Thus we just need to show the bounds for $\partial_y^2\chi, \partial_x\partial_y\chi, \partial_\mu\partial_y\chi, \partial_\mu\partial_x\chi$, and $\partial_\mu^2\chi$. The bounds for $\partial_y^2\chi$, $\partial_x\partial_y\chi$ and $\partial_\mu\partial_y\chi$ are established in the recent [41] Proposition 3.1.

For the mixed partial derivative in x and μ and the second partial derivative in μ, we can follow the proof of Proposition 4.1 of [65]. The details are omitted here due to the similarity of the argument. □

Some specific examples for which the assumptions of the paper hold

One can check, using the results of Lemmas C.2, C.3, and C.4, that examples C.5–C.7 below satisfy all of the assumptions made in the paper. See the extended arXiv version of this paper (arXiv:2202.08403) for more details.

Example C.5 (A case with full dependence of the coefficients on (x, y, μ)) Suppose $\tau_1, \tau_2, \sigma > 0$ are constant with σ large enough that $\bar{D}(x,\mu) > 0, \forall x \in \mathbb{R}, \mu \in \mathcal{P}_2(\mathbb{R})$, where $\bar{D}$ is as in Eq. (24), and the other coefficients take the form

$$b(x,y,\mu) = q\left(y - \frac{1}{\kappa}\langle \mu, \phi_f\rangle\right)p_b(x), \quad c(x,y,\mu) = r_c(y) + p_c(x) + \langle \mu, \phi_c\rangle$$
$$f(x,y,\mu) = -\kappa y + \langle \mu, \phi_f\rangle, \quad g(x,y,\mu) = r_g(y) + p_g(x) + \langle \mu, \phi_g\rangle,$$

where here $\kappa > 0$ and $\langle \phi, \mu\rangle := \int_{\mathbb{R}} \phi(z)\mu(dz)$. Suppose also that $q \in C^\infty(\mathbb{R})$ is odd, there is $\beta > 0$ such that $|q(z)|, |q'(z)|, |q''(z)|, |q'''(z)| \leq C(1+|z|)^{-\beta}$, $\left\|r_c'\right\|_\infty \leq C$, $|r_g|_{C_b^1(\mathbb{R})} \leq C$, $\phi_c, \phi_g, \phi_f \in \mathcal{S}_{w+2}$, and $p_c, p_g \in C_b^{w+2}(\mathbb{R})$, $p_b \in C_b^{w+3}(\mathbb{R})$. Then assumptions (A1)–(A13) hold, and (A'13) holds if w is replaced by r.

Example C.6 (A case where Φ and π are independent of x, μ) Consider the case:

$$b(x,y,\mu) = b(y), \quad c(x,y,\mu) = c_1(x) + \langle \mu, c_2(x - \cdot)\rangle, \quad \sigma(x,y,\mu) \equiv \sigma$$
$$f(x,y,\mu) = -\kappa y + \eta(y), \quad g(x,y,\mu) = g_1(x) + \langle \mu, g_2(x-\cdot)\rangle, \quad \tau_1(x,y,\mu) \equiv \tau_1,$$
$$\tau_2(x,y,\mu) \equiv \tau_2.$$

Suppose $\eta \in C_b^1(\mathbb{R})$ with $\left\|\eta'\right\|_\infty < \kappa$, $c_1, g_1, c_2, g_2 \in C_b^{w+2}(\mathbb{R})$, $c_2, g_2 \in \mathcal{S}_{w+2}$, $\tau_1^2 + \tau_2^2 > 0$, and b is Lipschitz continuous, $O(|y|^{1/2})$, and satisfies the centering condition (21). Then Assumptions (A1)–(A13) hold. Furthermore, if this holds with w replaced by r, then Assumption (A'13) holds.

Example C.7 (The case without full-coupling) Consider the case where

$$b(x,y,\mu) \equiv 0, \quad \sigma(x,y,\mu) = \sigma(x,\mu).$$

In this setting, it is known that when also $g \equiv 0$ and $\tau_1 \equiv 0$, under sufficient conditions on c, σ, f and τ_2, we can expect not only convergence in distribution of $\bar{X}^\epsilon \stackrel{d}{=} \bar{X}^{i,\epsilon}$ from Eq. (57) to X from Eq. (25), but also convergence in L^2. It is easily seen that this also holds when $g, \tau_1 \neq 0$ if they are sufficiently regular.

Note that in the limiting coefficients from Eq. (24), we have $\Phi \equiv 0$, so $\bar{\gamma}(x,\mu) = \bar{c}(x,\mu)$ and $\bar{D}(x,\mu) = \frac{1}{2}\sigma^2(x,\mu)$. In this setting, we can see immediately that there is no need for Assumptions (A6), (A7), and (A11). (A8) need only hold with $F = c$ and $F = \psi \in C_c^\infty([0,T]\times\mathbb{R}\times\mathbb{R})$. We will see that, since we can gain the aforementioned L^2 averaging, there is no need for Theorem 7.2, and hence for Assumption (A10).

Sufficient conditions for Theorem 3.2 to hold in this case are: (A1)–(A3), (A5), (A9), (A12), and

(1) $c, a, f \in \mathcal{M}_p^{\tilde{\boldsymbol{\zeta}}}(\mathbb{R}\times\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ with $q_f(n,l,\boldsymbol{\beta}) \leq 1$, $q_a(n,l,\boldsymbol{\beta}) \leq 0$, $q_c(n,l,\boldsymbol{\beta}) \leq 2, \forall (n,l,\boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}$ and $q_c(n,l,\boldsymbol{\beta}) \leq 1, \forall (n,l,\boldsymbol{\beta}) \in \tilde{\boldsymbol{\zeta}}_1$.
(2) $\sigma^2 \in \mathcal{M}_b^{\boldsymbol{\zeta}_{x,r+2}}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R})) \cap \mathcal{M}_{\boldsymbol{\delta},b}^{\bar{\boldsymbol{\zeta}}_{r+2}}(\mathbb{R}\times\mathcal{P}_2(\mathbb{R}))$ and $\sup_{x\in\mathbb{R},\mu\in\mathcal{P}(R)} \left\|\frac{\delta}{\delta m}\sigma^2(x,\mu)[\cdot]\right\|_{r+2} < \infty$.
(3) $f, a, c \in \mathcal{M}_p^{\boldsymbol{\zeta}_{x,r+2}}$ and

$$\left\|\frac{\delta}{\delta m} f(x,y,\mu)[\cdot]\right\|_{r+2}, \left\|\frac{\delta}{\delta m} a(x,y,\mu)[\cdot]\right\|_{r+2}, \left\|\frac{\delta}{\delta m} c(x,y,\mu)[\cdot]\right\|_{r+2} \leq C(1+|y|^k),$$

uniformly in $x \in \mathbb{R}$, $\mu \in \mathcal{P}_2(\mathbb{R})$ for some $k \in \mathbb{N}$.

In the above the referenced collections of multi-indices are from Eq. (27).

Appendix D. On differentiation of functions on spaces of measures

We will need the following two definitions from [13]:

Definition D.1 Given a function $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$, we may define a lifting of u to $\tilde{u} : L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}; \mathbb{R}^d) \to \mathbb{R}$ via $\tilde{u}(X) = u(\mathcal{L}(X))$ for $X \in L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}; \mathbb{R}^d)$. We assume $\tilde{\Omega}$ is a Polish space, $\tilde{\mathcal{F}}$ its Borel σ-field, and $\tilde{\mathbb{P}}$ is an atomless probability measure (since $\tilde{\Omega}$ is Polish, this is equivalent to every singleton having zero measure).

Here, denoting by $\mu(|\cdot|^r) := \int_{\mathbb{R}^d} |x|^r \mu(dx)$ for $r > 0$,

$$\mathcal{P}_2(\mathbb{R}^d) := \{\mu \in \mathcal{P}(\mathbb{R}^d) : \mu(|\cdot|^2) = \int_{\mathbb{R}^d} |x|^2 \mu(dx) < \infty\}.$$

$\mathcal{P}_2(\mathbb{R}^d)$ is a Polish space under the L^2-Wasserstein distance

$$\mathbb{W}_2(\mu_1,\mu_2) := \inf_{\pi\in\mathcal{C}_{\mu_1,\mu_2}} \left[\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \pi(dx,dy)\right]^{1/2},$$

where $\mathcal{C}_{\mu_1,\mu_2}$ denotes the set of all couplings of μ_1, μ_2.

We say u is **L-differentiable** or **Lions-differentiable** at $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ if there exists a random variable X_0 on some $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ satisfying the above assumptions, $\mathcal{L}(X_0) = \mu_0$ and $\tilde{u}$ is Fréchet differentiable at X_0.

The Fréchet derivative of $\tilde{u}$ can be viewed as an element of $L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}; \mathbb{R}^d)$ by identifying $L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}; \mathbb{R}^d)$ and its dual. From this, one can find that if u is L-differentiable at $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, there is a deterministic measurable function $\xi : \mathbb{R}^d \to \mathbb{R}^d$ such that $D\tilde{u}(X_0) = \xi(X_0)$, and that ξ is uniquely defined μ_0-almost everywhere on $\mathbb{R}^d$. We denote this equivalence class of $\xi \in L^2(\mathbb{R}^d, \mu_0; \mathbb{R}^d)$ by $\partial_\mu u(\mu_0)$ and call $\partial_\mu u(\mu_0)[\cdot] : \mathbb{R}^d \to \mathbb{R}^d$ the **Lions derivative** of u at μ_0. Note that this definition is independent of the choice of X_0 and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. See [13] Section 5.2.

To avoid confusion when u depends on more variables than just μ, if $\partial_\mu u(\mu_0)$ is differentiable at $z_0 \in \mathbb{R}^d$, we denote its derivative at z_0 by $\partial_z \partial_\mu u(\mu_0)[z_0]$.

Definition D.2 ([13] Definition 5.83) We say $u : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ is **Fully $\mathbf{C^2}$** if the following conditions are satisfied:

(1) u is C^1 in the sense of L-differentiation, and its first derivative has a jointly continuous version $\mathcal{P}_2(\mathbb{R}) \times \mathbb{R} \ni (\mu, z) \mapsto \partial_\mu u(\mu)[z] \in \mathbb{R}$.
(2) For each fixed $\mu \in \mathcal{P}_2(\mathbb{R})$, the version of $\mathbb{R} \ni z \mapsto \partial_\mu u(\mu)[z] \in \mathbb{R}$ from the first condition is differentiable on $\mathbb{R}$ in the classical sense and its derivative is given by a jointly continuous function $\mathcal{P}_2(\mathbb{R}) \times \mathbb{R} \ni (\mu, z) \mapsto \partial_z \partial_\mu u(\mu)[z] \in \mathbb{R}$.
(3) For each fixed $z \in \mathbb{R}$, the version of $\mathcal{P}_2(\mathbb{R}) \ni \mu \mapsto \partial_\mu u(\mu)[z] \in \mathbb{R}$ in the first condition is continuously L-differentiable component-by-component, with a derivative given by a function $\mathcal{P}_2(\mathbb{R}) \times \mathbb{R} \times \mathbb{R} \ni (\mu, z, \bar{z}) \mapsto \partial^2_\mu u(\mu)[z][\bar{z}] \in \mathbb{R}$ such that for any $\mu \in \mathcal{P}_2(\mathbb{R})$ and $X \in L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}; \mathbb{R})$ with $\mathcal{L}(X) = \mu$, $\partial^2 u(\mu)[z][X]$ gives the Fréchet derivative at X of $L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}; \mathbb{R}) \ni X' \mapsto \partial_\mu u(\mathcal{L}(X'))[z]$ for every $z \in \mathbb{R}$. Denoting $\partial^2_\mu u(\mu)[z][\bar{z}]$ by $\partial^2_\mu u(\mu)[z, \bar{z}]$, the map $\mathcal{P}_2(\mathbb{R}) \times \mathbb{R} \times \mathbb{R} \ni (\mu, z, \bar{z}) \mapsto \partial^2_\mu u(\mu)[z, \bar{z}]$ is also assumed to be continuous in the product topology.

Remark D.3 In this paper we will in fact also look at functions $u : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ which are required to have 3 Lions Derivatives. We will assume such functions are **Fully $\mathbf{C^2}$**, and satisfy:

(4) For each each fixed $\mu \in \mathcal{P}_2(\mathbb{R})$ the version of $\mathbb{R} \times \mathbb{R} \ni (z_1, z_2) \mapsto \partial^2_\mu u(\mu)[z_1, z_2] \in \mathbb{R}$ in Definition D.2 (3) is differentiable on $\mathbb{R}^2$ in the classical sense and its derivative is given by a jointly continuous function $\mathcal{P}_2(\mathbb{R}) \times \mathbb{R} \times \mathbb{R} \ni (\mu, z_1, z_2) \mapsto \nabla_z \partial^2_\mu u(\mu)[z_1, z_2] = (\partial_{z_1} \partial^2_\mu u(\mu)[z_1, z_2], \partial_{z_2} \partial^2_\mu u(\mu)[z_1, z_2]) \in \mathbb{R}^2$.
(5) For each fixed $(z_1, z_2) \in \mathbb{R}^2$, the version of $\mathcal{P}_2(\mathbb{R}) \ni \mu \mapsto \partial^2_\mu u(\mu)[z_1, z_2] \in \mathbb{R}$ in Definition D.2 (3) is continuously L-differentiable component-by-component, with a derivative given by a function $\mathcal{P}_2(\mathbb{R}) \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \ni (\mu, z_1, z_2, z_3) \mapsto \partial^3_\mu u(\mu)[z_1, z_2][z_3] \in \mathbb{R}$ such that for any $\mu \in \mathcal{P}_2(\mathbb{R})$ and $X \in L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}; \mathbb{R})$ with $\mathcal{L}(X) = \mu$, $\partial^3 u(\mu)[z_1, z_2][X]$ gives the Fréchet derivative at X of $L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}; \mathbb{R}) \ni X' \mapsto \partial^2_\mu u(\mathcal{L}(X'))[z_1, z_2]$ for every $(z_1, z_2) \in \mathbb{R}^2$. Denoting $\partial^3_\mu u(\mu)[z_1, z_2][z_3]$ by $\partial^2_\mu u(\mu)[z_1, z_2, z_3]$, the map $\mathcal{P}_2(\mathbb{R}) \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \ni (\mu, z_1, z_2, z_3) \mapsto \partial^3_\mu u(\mu)[z_1, z_2, z_3]$ is also assumed to be continuous in the product topology.

Though we don't require higher than 3 Lions derivatives in this paper, when we state general results for higher Lions derivatives in terms of the spaces from Definition 2.4, we assume the analogous higher continuity.

We will also make use of another notion of differentiation of functions of probability measures: the linear functional derivative.

Definition D.4 ([13] Definition 5.43) Let $p : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$. We say p has **Linear Functional Derivative** $\frac{\delta}{\delta m} p$ if there exists a function $(z, \mu) \ni \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \mapsto \frac{\delta}{\delta m} p(\mu)[z] \in \mathbb{R}$ continuous in the product topology on $\mathbb{R} \times \mathcal{P}_2(\mathbb{R})$ such that for any bounded subset $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R})$, the function $\mathbb{R} \ni z \mapsto \frac{\delta}{\delta m} p(\mu)[z]$ is of at most quadratic growth uniformly in μ for $\mu \in \mathcal{K}$, and for all $\nu_1, \nu_2 \in \mathcal{P}_2(\mathbb{R}^d)$:

$$p(\nu_2) - p(\nu_1) = \int_0^1 \int_{\mathbb{R}} \frac{\delta}{\delta m} p((1-r)\nu_1 + r\nu_2)[z](\nu_2(dz) - \nu_1(dz))dr.$$

Note in particular that this implies that p is continuous on $\mathcal{P}_2(\mathbb{R})$.

The second linear functional derivative is said to exist if the linear functional derivative of $\frac{\delta}{\delta m} p(\mu)[z_1]$ as defined above exists for each $z_1 \in \mathbb{R}$. For any bounded subset $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R})$, the function $(z_1, z_2) \ni \mathbb{R} \times \mathbb{R} \mapsto \frac{\delta}{\delta m}\left(\frac{\delta}{\delta m} p(\mu)[z_1]\right)[z_2] := \frac{\delta^2}{\delta m^2} p(\mu)[z_1, z_2] \in \mathbb{R}$, is of at most quadratic growth uniformly in μ for $\mu \in \mathcal{K}$, $(z_1, z_2, \mu) \ni \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \mapsto \frac{\delta^2}{\delta m^2} p(\mu)[z_1, z_2] \in \mathbb{R}$ and is assumed to be continuous in the product topology on $\mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R})$.

Remark D.5 See Section 5.4.1 of [13] for well-posedness of the above notion of differentiability and relation to Lions derivative. In particular, under sufficient regularity on $u : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$, $\partial_\mu u(\mu)[z] = \partial_z \frac{\delta}{\delta m} u(\mu)[z]$. For a formal understanding of the linear functional derivative as a Fréchet Derivative, see p. 21 of [12]. Lastly, it is important to note that the linear functional derivative is only defined up to a constant by definition. This is usually not of importance, at it normally arises when studying fluctuations of measures. In particular, applying $\tilde{Z}_t^N$ as defined in (3) to a constant function, we of course get 0 for any $N \in \mathbb{N}$ and $t \in [0, T]$, so shifting the linear functional derivative by a constant in Eq. (32) does not change the representation of the limiting process. A common means of fixing this constant for concreteness is to require that $\langle \mu, \frac{\delta}{\delta m} u(\mu)[\cdot] \rangle = 0, \forall \mu \in \mathcal{P}_2(\mathbb{R})$ (see p.31 of [12] or Section 2.2 of [19]. However, due to our choice of topology for the fluctuations process, correcting the constant for the linear functional derivative may break assumptions (A13) and (A'13). We thus interpret these assumptions to mean that there is a choice of constant when defining each of the linear functional derivatives of the functions in question which makes them satisfy the desired properties.

We recall a useful connection between the Lions derivative as defined in D.1 and the empirical measure.

Proposition D.6 *For $g : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ which is Fully C^2 in the sense of definition D.2, we can define the empirical projection of g, as $g^N : (\mathbb{R}^d)^N \to \mathbb{R}^d$ given by*

$$g^N(\beta_1, \ldots, \beta_N) := g(\frac{1}{N}\sum_{i=1}^N \delta_{\beta_i}).$$

Then g^N is twice differentiable on $(\mathbb{R}^d)^N$, and for each $\beta_1, \ldots, \beta_N \in \mathbb{R}^d$, $(i, j) \in \{1, \ldots, N\}^2$, $l \in \{1, \ldots, d\}$

$$\nabla_{\beta_i} g_l^N(\beta_1, \ldots, \beta_N) = \frac{1}{N}\partial_\mu g_l(\frac{1}{N}\sum_{i=1}^N \delta_{\beta_i})[\beta_i] \tag{74}$$

and

$$\begin{aligned}&\nabla_{\beta_i}\nabla_{\beta_j} g_l^N(\beta_1, ..., \beta_N)\\ &\quad = \frac{1}{N}\partial_z\partial_\mu g_l\left(\frac{1}{N}\sum_{i=1}^N \delta_{\beta_i}\right)[\beta_i]\mathbb{1}_{i=j} + \frac{1}{N^2}\partial_\mu^2 g_l\left(\frac{1}{N}\sum_{i=1}^N \delta_{\beta_i}\right)[\beta_i, \beta_j].\end{aligned} \tag{75}$$

Proof This follows from Propositions 5.35 and 5.91 of [13].

□

Finally, we provide a Lemma which allows us to couple the interacting particles (55) to the IID McKean–Vlasov Eqs. (57) knowing only information about the growth of the linear functional derivatives of the coefficients.

Lemma D.7 *Suppose $p : \mathbb{R} \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ satisfies*

$$\begin{aligned}&\sup_{x,z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta}{\delta m}p(x, y, \mu)[z]|\\ &\quad + \sup_{x,z,\bar{z}\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta^2}{\delta m^2}p(x, y, \mu)[z, \bar{z}]| \le C(1 + |y|^k)\end{aligned}$$

for some $C > 0, k \in \mathbb{N}$ independent of $y \in \mathbb{R}$, and that $p(x, y, \cdot) : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ is Lipschitz continuous in $\mathbb{W}_2$ for all $x, y \in \mathbb{R}$. Assume (A1)–(A7) and (A9). Then for $(\bar{X}^{i,\epsilon}, \bar{Y}^{i,\epsilon})$ as in Eq. (57) and $\bar{\mu}_t^{\epsilon,N}$ as in Eq. (58), we have there exists $C > 0$ independent of N such that for all $t \in [0, T]$:

$$\mathbb{E}\left[\left|p(\bar{X}_t^{i,\epsilon}, \bar{Y}_t^{i,\epsilon}, \bar{\mu}_t^{\epsilon,N}) - p(\bar{X}_t^{i,\epsilon}, \bar{Y}_t^{i,\epsilon}, \mathcal{L}(\bar{X}_t^{\epsilon}))\right|^2\right] \le \frac{C}{N-1}.$$

Here $\bar{X}^\epsilon \overset{d}{=} \bar{X}^{i,\epsilon}, \forall i, N \in \mathbb{N}$.

Proof This follows using the same conditional expectation argument as on p.26 in [19] and then following the proof of Lemma 5.10 in the same paper, but where we only require second order expansions rather than 4th. Since the argument and assumptions are slightly different, we present the proof here for completeness.

We first write

$$\begin{aligned}
&\mathbb{E}\left[\left|p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})-p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\right|^2\right]\\
&\quad\le 2\mathbb{E}\left[\left|p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N})-p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N,-i})\right|^2\right]\\
&\qquad+2\mathbb{E}\left[\left|p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N,-i})-p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\right|^2\right]\\
&\quad\le C\mathbb{E}\left[|\mathbb{W}_2(\bar{\mu}_t^{\epsilon,N},\bar{\mu}_t^{\epsilon,N,-i})|^2\right]\\
&\qquad+2\mathbb{E}\left[\left|p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N,-i})-p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\right|^2\right]
\end{aligned}$$

where here $\bar{\mu}_t^{\epsilon,N,-i}$ denotes $\bar{\mu}_t^{\epsilon,N}$ with the i'th particle removed, i.e.

$$\bar{\mu}_t^{\epsilon,N,-i}:=\frac{1}{N-1}\sum_{j=1,j\neq i}^{N}\delta_{\bar{X}_t^{j,\epsilon}}.$$

Recall the formula

$$\mathbb{W}_p^p(\mu_x^N,\mu_y^N)=\min_{\sigma}\frac{1}{N}\sum_{i=1}^{N}|x_i-\sigma(y)_i|^p \tag{76}$$

for $x,y\in\mathbb{R}^N$, $\mu_x^N=\frac{1}{N}\sum_{i=1}^N\delta_{x_i}$, $\mu_x^N=\frac{1}{N}\sum_{i=1}^N\delta_{x_i}$, and where $\sigma:\mathbb{R}^N\to\mathbb{R}^N$ denotes a permutation of the coordinates of a vector in $\mathbb{R}^N$ (see e.g. Eq. 2.8 in [59]). This suggests that the first term should be bounded due to the bound $C\mathbb{E}[|\bar{X}_t^{i,\epsilon}|^2]/N\le C/N$ from Lemma 8.1.

To see this is indeed true, we take $\mu_x^N,\mu_x^{N,-i}$ for any $x\in\mathbb{R}^N$, where here $\mu_x^{N,-i}$ is defined in the same way as $\bar{\mu}_t^{\epsilon,N,-i}$, and see

$$\begin{aligned}
\mathbb{W}_2^2(\mu_x^N,\mu_x^{N,-i})&\le\int_{\mathbb{R}^2}|x-y|^2\gamma^N(dx,dy)\\
\gamma^N(dx,dy)&:=\frac{1}{N}\sum_{j=1,j\neq i}^{N}\left[\delta_{x_j}(dx)+\frac{1}{N-1}\delta_{x_i}(dx)\right]\delta_{x_j}(dy)
\end{aligned}$$

We see that indeed γ^N is a coupling between μ_x^N, $\mu_x^{N,-i}$ since it is clearly non-negative,

$$\int_{\mathbb{R}^2} \gamma^N(dx, dy) = \frac{1}{N}[N-1][1 + \frac{1}{N-1}] = 1,$$

and for $f \in C_b(\mathbb{R})$,

$$\begin{aligned}
\int_{\mathbb{R}^2} f(x)\gamma^N(dx, dy) &= \frac{1}{N} \sum_{j=1, j\neq i}^{N} \left[f(x_j) + \frac{1}{N-1} f(x_i) \right] \\
&= \left\{ \frac{1}{N} \sum_{j=1, j\neq i}^{N} f(x_j) \right\} + \frac{1}{N} f(x_i) = \int_{\mathbb{R}} f(y)\mu_x^N(dy) \\
\int_{\mathbb{R}^2} f(y)\gamma^N(dx, dy) &= \frac{1}{N}\left[1 + \frac{1}{N-1}\right] \sum_{j=1, j\neq i}^{N} f(x_j) \\
&= \frac{1}{N-1} \sum_{j=1, j\neq i}^{N} f(x_j) = \int_{\mathbb{R}} f(y)\mu_x^{N,-i}(dy).
\end{aligned}$$

So indeed

$$\begin{aligned}
\mathbb{W}_2^2(\mu_x^N, \mu_x^{N,-i}) &\leq \int_{\mathbb{R}^2} |x-y|^2 \gamma^N(dx, dy) \\
&= \frac{1}{N} \sum_{j=1, j\neq i}^{N} \left\{ |x_j - x_j|^2 + \frac{1}{N-1}|x_j - x_i|^2 \right\} \\
&= \frac{1}{N(N-1)} \sum_{j=1, j\neq i}^{N} |x_j - x_i|^2.
\end{aligned}$$

Now, applying this to the first term we wish to bound,

$$\begin{aligned}
C\mathbb{E}\left[|\mathbb{W}_2(\bar{\mu}_t^{\epsilon,N}, \bar{\mu}_t^{\epsilon,N,-i})|^2 \right] &\leq \frac{C}{N(N-1)} \sum_{j=1, j\neq i}^{N} \mathbb{E}\left[\left| \bar{X}_t^{j,\epsilon} - \bar{X}_t^{i,\epsilon} \right|^2 \right] \\
&= \frac{C}{N}\left[\mathbb{E}[|\bar{X}_t^{\epsilon}|^2] - \mathbb{E}[\bar{X}_t^{\epsilon}]^2 \right] \\
&\leq \frac{C}{N}\mathbb{E}[|\bar{X}_t^{\epsilon}|^2] \leq \frac{C}{N}
\end{aligned}$$

where in the equality we use that the $\bar{X}_t^{i,\epsilon}$'s are IID, and in the last bound we used Lemma 8.1.

Now we turn to the second term. We have by independence,

$$\begin{aligned}&\mathbb{E}\left[\left|p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N,-i})-p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\right|^2\right]\\&\quad=\mathbb{E}\left[\mathbb{E}\left[\left|p(x,y,\bar{\mu}_t^{\epsilon,N,-i})-p(x,y,\mathcal{L}(\bar{X}_t^{\epsilon}))\right|^2\right]\Big|_{(x,y)=(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon})}\right].\end{aligned}$$

We will show that for $q:\mathcal{P}_2(\mathbb{R})\to\mathbb{R}$ with two bounded Linear Functional Derivatives, that for $\{\xi_i\}_{i\in\mathbb{N}}$ IID with $\xi_i\sim\mu\in\mathcal{P}_2(\mathbb{R})$, that letting $\xi=(\xi_1,\dots,\xi_N)$ and μ_ξ^N be as above with ξ in the place of x:

$$\begin{aligned}&\mathbb{E}\left[|q(\mu_\xi^N)-q(\mu)|^2\right]\\&\quad\le\frac{C}{N}\left[\sup_{z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})}|\frac{\delta}{\delta m}q(\mu)[z]|^2+\sup_{z,\bar{z}\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})}|\frac{\delta^2}{\delta m^2}q(\mu)[z,\bar{z}]|^2\right].\end{aligned}\tag{77}$$

Applying this to the above equality, we have there is $k\in\mathbb{N}$ such that

$$\begin{aligned}&\mathbb{E}\left[\left|p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\bar{\mu}_t^{\epsilon,N,-i})-p(\bar{X}_t^{i,\epsilon},\bar{Y}_t^{i,\epsilon},\mathcal{L}(\bar{X}_t^{\epsilon}))\right|^2\right]\\&\quad\le\frac{C}{N-1}\mathbb{E}\left[1+|\bar{Y}_t^{\epsilon}|^{2k}\right]\le\frac{C}{N-1}\end{aligned}$$

by Lemma 7.1, and the result will have been proved.

We now prove the bound (77). By definition of the linear functional derivative, we have:

$$q(\mu_\xi^N)-q(\mu)=\int_0^1\int_{\mathbb{R}}\frac{\delta}{\delta m}q(r\mu_\xi^N+(1-r)\mu)[z](\mu_\xi^N(dz)-\mu(dz))dr=S_1+S_2$$

where

$$\begin{aligned}S_1^N&=\int_{\mathbb{R}}\frac{\delta}{\delta m}q(\mu)[z](\mu_\xi^N(dz)-\mu(dz)),\\S_2^N&=\int_0^1\int_{\mathbb{R}}\left[\frac{\delta}{\delta m}q(r\mu_\xi^N+(1-r)\mu)[z]-\frac{\delta}{\delta m}q(\mu)[z]\right](\mu_\xi^N(dz)-\mu(dz))dr.\end{aligned}$$

For S_1, we have by independence:

$$\begin{aligned}\mathbb{E}\left[|S_1|^2\right]&=\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^N\frac{\delta}{\delta m}q(\mu)[\xi_i]-\mathbb{E}\left[\frac{\delta}{\delta m}q(\mu)[\xi_1]\right]\right|^2\right]\\&=\frac{1}{N}\left(\mathbb{E}\left[\left|\frac{\delta}{\delta m}q(\mu)[\xi_1]\right|^2\right]-\mathbb{E}\left[\frac{\delta}{\delta m}q(\mu)[\xi_1]\right]^2\right)\end{aligned}$$

$$\leq \frac{1}{N} \sup_{z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta}{\delta m} q(\mu)[z]|^2.$$

Now we set

$$\begin{aligned}\phi_r^i &:= \frac{\delta}{\delta m} q(r\mu_\xi^N + (1-r)\mu)[\xi_i] - \frac{\delta}{\delta m} q(\mu)[\xi_i] \\ &\quad - \tilde{\mathbb{E}}\left[\frac{\delta}{\delta m} q(r\mu_\xi^N + (1-r)\mu)[\tilde{\xi}] - \frac{\delta}{\delta m} q(\mu)[\tilde{\xi}]\right],\end{aligned}$$

where $\tilde{\xi}$ is an independent copy of the ξ_i's, $r \in [0, 1]$, and the expectation $\tilde{\mathbb{E}}$ is taken over the law of $\tilde{\xi}$.

Then we have $S_2^N = \frac{1}{N} \sum_{i=1}^N \int_0^1 \phi_r^i dr$, and

$$\mathbb{E}\left[\left|S_2^N\right|^2\right] \leq \frac{1}{N^2} \int_0^1 \mathbb{E}\left[\left|\sum_{i=1}^N \phi_r^i\right|^2\right] dr = \frac{1}{N^2} \int_0^1 \sum_{i=1}^N \sum_{j=1}^N \mathbb{E}\left[\phi_r^i \phi_r^j\right] dr = S_{2,1}^N + S_{2,2}^N$$

where

$$S_{2,1}^N = \frac{1}{N^2} \int_0^1 \sum_{i=1}^N \mathbb{E}\left[|\phi_r^i|^2\right] dr, \quad \text{and} \quad S_{2,2}^N = \frac{1}{N^2} \int_0^1 \sum_{i=1}^N \sum_{j=1, j\neq i}^N \mathbb{E}\left[\phi_r^i \phi_r^j\right] dr.$$

Observing that for all $i \in \mathbb{N}$, $r \in [0, 1]$ and $\omega \in \Omega$, $|\phi_r^i(\omega)|^2 \leq C \sup_{z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta}{\delta m} q(\mu)[z]|^2$, so we have

$$S_{2,1}^N \leq \frac{C}{N} \sup_{z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta}{\delta m} q(\mu)[z]|^2.$$

For $S_{2,2}^N$, we introduce the measures $\mu_\xi^{N,-(i_1,i_2)} := \frac{1}{N-2} \sum_{j=1, j\neq i_1,i_2}^N \delta_{\xi_j}$ for $i_1, i_2 \in \{1, \ldots, N\}$, and let

$$\begin{aligned}\phi_r^{i,-(i_1,i_2)} &:= \frac{\delta}{\delta m} q(r\mu_\xi^{N,-(i_1,i_2)} + (1-r)\mu)[\xi_i] - \frac{\delta}{\delta m} q(\mu)[\xi_i] \\ &\quad - \tilde{\mathbb{E}}\left[\frac{\delta}{\delta m} q(r\mu_\xi^{N,-(i_1,i_2)} + (1-r)\mu)[\tilde{\xi}] - \frac{\delta}{\delta m} q(\mu)[\tilde{\xi}]\right].\end{aligned}$$

Then

$$\begin{aligned}\phi_r^i \phi_r^j &= [\phi_r^i - \phi_r^{i,-(i,j)}][\phi_r^j - \phi_r^{j,-(i,j)}] + \phi_r^{j,-(i,j)}[\phi_r^i - \phi_r^{i,-(i,j)}] \\ &\quad + \phi_r^{i,-(i,j)}[\phi_r^j - \phi_r^{j,-(i,j)}] + \phi_r^{i,-(i,j)} \phi_r^{j,-(i,j)},\end{aligned}$$

so

$$S_{2,2}^N = S_{2,2,1}^N + S_{2,2,2}^N + S_{2,2,3}^N + S_{2,2,4}^N$$

$$S^N_{2,2,1} = \frac{1}{N^2}\int_0^1 \sum_{i=1}^N \sum_{j=1,j\neq i}^N \mathbb{E}\Big[\phi_r^{i,-(i,j)}\phi_r^{j,-(i,j)}\Big]dr$$

$$S^N_{2,2,2} = \frac{1}{N^2}\int_0^1 \sum_{i=1}^N \sum_{j=1,j\neq i}^N \mathbb{E}\Big[\phi_r^{j,-(i,j)}[\phi_r^i - \phi_r^{i,-(i,j)}]\Big]dr$$

$$S^N_{2,2,3} = \frac{1}{N^2}\int_0^1 \sum_{i=1}^N \sum_{j=1,j\neq i}^N \mathbb{E}\Big[\phi_r^{i,-(i,j)}[\phi_r^j - \phi_r^{j,-(i,j)}]\Big]dr$$

$$S^N_{2,2,4} = \frac{1}{N^2}\int_0^1 \sum_{i=1}^N \sum_{j=1,j\neq i}^N \mathbb{E}\Big[[\phi_r^i - \phi_r^{i,-(i,j)}][\phi_r^j - \phi_r^{j,-(i,j)}]\Big]dr.$$

For $S^N_{2,2,1}$, we have

$$\begin{aligned}\mathbb{E}\Big[\phi_r^{i,-(i,j)}\phi_r^{j,-(i,j)}\Big] &= \mathbb{E}\Big[\mathbb{E}\Big[\phi_r^{i,-(i,j)}\phi_r^{j,-(i,j)}|\xi_k, k\neq i,j\Big]\Big]\\ &= \mathbb{E}\Big[\mathbb{E}\Big[\phi_r^{i,-(i,j),x}\phi_r^{j,-(i,j),x}\Big]\Big|_{x=\xi}\Big]\\ &= \mathbb{E}\Big[\mathbb{E}\Big[\phi_r^{i,-(i,j),x}\Big]\Big|_{x=\xi}\mathbb{E}\Big[\phi_r^{j,-(i,j),x}\Big]\Big|_{x=\xi}\Big]\end{aligned}$$

where

$$\begin{aligned}\phi_r^{i,-(i_1,i_2),x} := &\frac{\delta}{\delta m}q(r\mu_x^{N,-(i_1,i_2)} + (1-r)\mu)[\xi_i] - \frac{\delta}{\delta m}q(\mu)[\xi_i]\\ &- \tilde{\mathbb{E}}\Big[\frac{\delta}{\delta m}q(r\mu_x^{N,-(i_1,i_2)} + (1-r)\mu)[\tilde\xi] - \frac{\delta}{\delta m}q(\mu)[\tilde\xi]\Big]\end{aligned}$$

and same for j. Then

$$\begin{aligned}\mathbb{E}\Big[\phi_r^{i,-(i,j),x}\Big]\Big|_{x=\xi} = &\Big\{\mathbb{E}\Big[\frac{\delta}{\delta m}q(r\mu_x^{N,-(i_1,i_2)} + (1-r)\mu)[\xi_i] - \frac{\delta}{\delta m}q(\mu)[\xi_i]\Big]\\ &- \tilde{\mathbb{E}}\Big[\frac{\delta}{\delta m}q(r\mu_x^{N,-(i_1,i_2)} + (1-r)\mu)[\tilde\xi] - \frac{\delta}{\delta m}q(\mu)[\tilde\xi]\Big]\Big\}\Big|_{(x=\xi)} = 0\end{aligned}$$

since $\xi_i \overset{d}{=} \tilde\xi$, and same for $\mathbb{E}\Big[\phi_r^{j,-(i,j),x}\Big]\Big|_{x=\xi}$. Thus in fact, $S^N_{2,2,1} = 0$.

To handle $S_{2,2,2} - S_{2,2,4}$, we need to see how to bound $|\phi_r^i - \phi_r^{i,-(i,j)}|$. We have that

$$\begin{aligned}\phi_r^i - \phi_r^{i,-(i,j)} = &\frac{\delta}{\delta m}q(r\mu_\xi^N + (1-r)\mu)[\xi_i] - \frac{\delta}{\delta m}q(r\mu_\xi^{N,-(i,j)} + (1-r)\mu)[\xi_i]\\ &+ \tilde{\mathbb{E}}\Big[\frac{\delta}{\delta m}q(r\mu_\xi^{N,-(i,j)} + (1-r)\mu)[\tilde\xi] - \frac{\delta}{\delta m}q(r\mu_\xi^N + (1-r)\mu)[\tilde\xi]\Big]\end{aligned}$$

$$
\begin{aligned}
&= r\int_0^1\int_{\mathbb{R}} \frac{\delta^2}{\delta m^2} q(rs\mu_\xi^N + r(1-s)\mu_\xi^{N,-(i,j)} + (1-r)\mu)[\xi_i, \bar{z}] \\
&[\mu_\xi^N(d\bar{z}) - \mu_\xi^{N,-(i,j)}(d\bar{z})]ds \\
&+ r\tilde{\mathbb{E}}\Big[\int_0^1\int_{\mathbb{R}} \frac{\delta^2}{\delta m^2} q(rs\mu_\xi^N + r(1-s)\mu_\xi^{N,-(i,j)} + (1-r)\mu) \\
&[\tilde{\xi}, \bar{z}][\mu_\xi^N(d\bar{z}) - \mu_\xi^{N,-(i,j)}(d\bar{z})]ds\Big].
\end{aligned}
$$

Then using

$$
\begin{aligned}
\mu_x^N - \mu_x^{N,-(i,j)} &= \frac{1}{N}\sum_{k=1}^N \delta_{x_k} - \frac{1}{N-2}\sum_{k=1,k\neq i,j}^N \delta_{x_k} \\
&= \frac{1}{N}\delta_{x_i} + \frac{1}{N}\delta_{x_j} - \frac{2}{N(N-2)}\sum_{k=1,k\neq i,j}^N \delta_{x_k}
\end{aligned}
$$

and that $r \in [0,1]$, we get

$$
|\phi_r^i(\omega) - \phi_r^{i,-(i,j)}(\omega)| \le \frac{4}{N}\sup_{z,\bar{z}\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta^2}{\delta m^2} q(\mu)[z,\bar{z}]|
$$

for all $\omega \in \Omega$, $r \in [0,1]$, $i, j \in \mathbb{N}$. This combined with the fact that $|\phi_r^{k,-(i,j)}(\omega)| \le C\sup_{z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta}{\delta m} q(\mu)[z]|$, $k = i, j$ for any $i, j \in \mathbb{N}$, $r \in [0,1]$, $\omega \in \Omega$ allows us to see:

$$
\begin{aligned}
S_{2,2,2}^N &\le C\frac{N(N-1)}{N^2}\sup_{z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta}{\delta m} q(\mu)[z]| \frac{1}{N}\sup_{z,\bar{z}\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta^2}{\delta m^2} q(\mu)[z,\bar{z}]| \\
&\le \frac{C}{N}[\sup_{z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta}{\delta m} q(\mu)[z]|^2 + \sup_{z,\bar{z}\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta^2}{\delta m^2} q(\mu)[z,\bar{z}]|^2] \\
S_{2,2,3}^N &\le C\frac{N(N-1)}{N^2}\sup_{z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta}{\delta m} q(\mu)[z]| \frac{1}{N}\sup_{z,\bar{z}\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta^2}{\delta m^2} q(\mu)[z,\bar{z}]| \\
&\le \frac{C}{N}[\sup_{z\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta}{\delta m} q(\mu)[z]|^2 + \sup_{z,\bar{z}\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta^2}{\delta m^2} q(\mu)[z,\bar{z}]|^2] \\
S_{2,2,4}^N &\le C\frac{N(N-1)}{N^2}\frac{1}{N^2}\sup_{z,\bar{z}\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta^2}{\delta m^2} q(\mu)[z,\bar{z}]|^2 \\
&\le \frac{C}{N^2}\sup_{z,\bar{z}\in\mathbb{R},\mu\in\mathcal{P}_2(\mathbb{R})} |\frac{\delta^2}{\delta m^2} q(\mu)[z,\bar{z}]|^2.
\end{aligned}
$$

So the bound (77) is proved. □

Remark D.8 Note we could have polynomial growth in x for the Linear Functional Derivatives as well and the result above would still hold, so long as we have sufficient bounded moments for $\bar{X}^{\epsilon}_{t}$. Also, the result is independent of the fact that the particles depend on ϵ, and of the fact that the particles are one-dimensional. See Lemma 5.10 in [19] and Theorem 2.11 in [14] for similar results in the higher-dimensional setting.

References

1. Baldi, P.: Large deviations for diffusions processes with homogenization and applications. Ann. Probab. **19**(2), 509–524 (1991)
2. Benedetto, D., Caglioti, E., Carrillo, J.A., Pulvirenti, M.: A non-Maxwellian steady distribution for one-dimensional granular media. J. Stat. Phys. **91**, 979–990 (1998)
3. Bensoussan, A., Lions, J.L., Papanicolau, G.: Asymptotic Analysis for Periodic Structures. North Holland, Amsterdam (1978)
4. Bezemek, Z., Spiliopoulos, K.: Large deviations for interacting multiscale particle systems. Stoch. Process. Appl. **155**, 27–108 (2023)
5. Bezemek, Z., Spiliopoulos, K.: Rate of homogenization for fully-coupled McKean–Vlasov SDEs. Stoch. Dyn. **23**(2). (2023)
6. Binney, J., Tremaine, S.: Galactic Dynamics. Princeton University Press, Princeton (2008)
7. Bryngelson, J.D., Onuchic, J.N., Socci, N.D., Wolynes, P.G.: Funnels pathways and the energy landscape of protein folding: a synthesis. Proteins Struct. Funct. Bioinform. **21**(3), 167–195 (1995)
8. Budhiraja, A., Dupuis, P.: A variational representation for positive functionals of infinite dimensional Brownian motion. Probab. Math. Stat. **20**, 39–61 (2001)
9. Budhiraja, A., Dupuis, P., Fischer, M.: Large devation properties of weakly interacting particles via weak convergence methods. T.A. Prob. **40**, 74–100 (2012)
10. Budhiraja, A., Wu, R.: Moderate deviation principles for weakly interacting particle systems. Probab. Theory Relat. Fields **168**, 721–771 (2016)
11. Cardaliaguet, P.: Notes on mean field games (from P. L. Lions' lectures at Coll'ege de France), 2013. https://www.ceremade.dauphine.fr/cardaliaguet/MFG20130420.pdf, unpublished
12. Cardaliaguet, P., Delarue, F., Lasry, J.M., Lions, P.L.: The Master Equation and the Convergence Problem in Mean Field Games. Princeton University Press, Princeton (2019)
13. Carmona, R., Delarue, F.: Probabilistic Theory of Mean Field Games with Applications I. Springer, New York (2018)
14. Chassagneux, J.F., Szpruch, L., Tse, A.: Weak quantitative propagation of chaos via differential calculus on the space of measures. Ann. Appl. Probab. **32**(3), 1929–1969 (2022)
15. Crisan, D., McMurray, E.: Smoothing properties of McKean–Vlasov SDEs. Probab. Theory Relat. Fields **171**(2), 97–148 (2018)
16. Dawson, D.A.: Critical dynamics and fluctuations for a mean-field model of cooperative behavior. J. Stat. Phys. **31**, 29–85 (1983)
17. Dawson, D.A., Gäartner, J.: Large deviations from the Mckean–Vlasov limit for weakly interacting diffusions. Stochastics **20**(4), 247–308 (1987)
18. Del Barrio, E., Deheuvels, P., Van De Geer, S.: Lectures on Empirical Processes: Theory and Statistical Applications. European Mathematical Society Publishing House, Zurich (2007)
19. Delarue, F., Lacker, D., Ramanan, K.: From the master equation to mean field game limit theory: a central limit theorem. Electron. J. Probab. **24**, 1–54 (2019)
20. Delgadino, M.G., Gvalani, R.S., Pavliotis, G.A.: On the diffusive-mean field limit for weakly interacting diffusions exhibiting phase transitions. Arch. Ration. Mech. Anal. **241**, 91–148 (2021)
21. Driver, B.K.: Analysis Tools with Examples (2004). http://www.math.ucsd.edu/bdriver/DRIVER/Book/anal.pdf, unpublished
22. Dupuis, P., Ellis, R.S.: A Weak Convergence Approach to the Theory of Large Deviations. Wiley, New York (1997)
23. Dupuis, P., Spiliopoulos, K.: Large deviations for multiscale diffusion via weak convergence methods. Stoch. Process. Appl. **122**(4), 1947–1987 (2012)

24. Dupuis, P., Spiliopoulos, K., Wang, H.: Importance sampling for multiscale diffusions. SIAM Multiscale Model. Simul. **12**(1), 1–27 (2012)
25. Ethier, S., Kurtz, T.: Markov Processes: Characterization and Convergence. Wiley, New York (1986)
26. Feng, J., Fouque, J.P., Kumar, R.: Small-time asymptotics for fast mean-reverting stochastic volatility models. Ann. Appl. Probab. **22**(4), 1541–1575 (2012)
27. Fernandez, B., Méléard, S.: A Hilbertian approach for fluctuations on the McKean–Vlasov model. Stoch. Process. Appl. **71**, 33–53 (1997)
28. Feng, J., Kurtz, T.: Large Deviations for Stochastic Processes. AMS, Providence (2006)
29. Fouque, J.P., Papanicolaou, G., Sircar, K.R.: Derivatives in Financial Markets with Stochastic Volatility. Cambridge University Press, Cambridge (2000)
30. Freidlin, M.I., Sowers, R.: A comparison of homogenization and large deviations, with applications to wavefront propagation. Stoch. Process Appl. **82**(1), 23–52 (1999)
31. Fischer, M.: On the connection between symmetric N-player games and mean field games. Ann. Appl. Probab. **27**(2), 757–810 (2017)
32. Gaitsgor, V., Nguyen, M.T.: Multiscale singularly perturbed control systems: limit occupational measures sets and averaging. SIAM J. Control. Optim. **41**(3), 954–974 (2002)
33. Ganguly, A., Sundar, P.: Inhomogeneous functionals and approximations of invariant distributions of ergodic diffusions: central limit theorem and moderate deviation asymptotics. Stoch. Process. Appl. **133**, 74–110 (2021)
34. Garnier, J., Papanicolaou, G., Yang, T.W.: Large deviations for a mean field model of systemic risk. SIAM J. Financ. Math. **4**(1), 151–184 (2013)
35. Garnier, J., Papanicolaou, G., Yang, T.W.: Consensus convergence with stochastic effects. Vietnam J. Math. **45**(1–2), 51–75 (2017)
36. Gel'fand, I.M., Vilenkin, N.Y.: Generalized Functions, vol. 4. AMS, Providence (1964)
37. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Springer, New York (2001)
38. Gomes, S.N., Pavliotis, G.A.: Mean field limits for interacting diffusions in a two-scale potential. J. Nonlinear Sci. **28**, 905–941 (2018)
39. Hitsuda, M., Mitoma, I.: Tightness problem and stochastic evolution equation arising from fluctuation phenomena for interacting diffusions. J. Multivar. Anal. **19**(2), 311–328 (1986)
40. Hong, W., Li, S., Liu, W.: Strong convergence rates in averaging principle for slow–fast McKean–Vlasov SPDEs. J. Differ. Equ. **316**, 94–135 (2022)
41. Hong, W., Li, S., Liu, W., Sun, X.: Central limit type theorem and large deviations for multi-scale McKean–Vlasov SDEs (2021). arXiv:2112.08203 [math.PR]
42. Hyeon, C., Thirumalai, D.: Can energy landscapes roughness of proteins and RNA be measured by using mechanical unfolding experiments? Proc. Natl. Acad. Sci. **100**(18), 10249–10253 (2003)
43. Isaacson, S.A., Ma, J., Spiliopoulos, K.: Mean field limits of particle-based stochastic reaction-diffusion models. SIAM J. Math. Anal. **54**(1), 453–511 (2022)
44. Jacquier, A., Spiliopoulos, K.: Pathwise moderate deviations for option pricing. Math. Financ. **30**(2), 426–463 (2020)
45. Kallianpur, G., Xiong, J.: Stochastic differential equations in infinite dimensional spaces. Lecture Notes-Monograph Series, vol. 26, pp. iii–342 (1995)
46. Karatzas, I., Shreve, S.: Brownian Motion and Stochastic Calculus. Springer, New York (1998)
47. Kolokolnikov, T., Bertozzi, A., Fetecau, R., Lewis, M.: Emergent behaviour in multi-particle systems with non-local interactions. Physica D **260**, 1–4 (2013)
48. Kong, M., Shi, Y., Sun, X.: Well-posedness and averaging principle of McKean–Vlasov SPDEs driven by cylindrical α-stable process (2021). arXiv:2106.05561 [math.PR]
49. Kurtz, T.G., Xiong, J.: A stochastic evolution equation arising from the fluctuations of a class of interacting particle systems. Commun. Math. Sci. **2**(3), 325–358 (2004)
50. Lacker, D.: Limit theory for controlled McKean–Vlasov dynamics. SIAM J. Control. Optim. **55**(3), 1641–1672 (2017)
51. Lućcon, E., Stannat, W.: Transition from Gaussian to non-Gaussian fluctuations for mean-field diffusions in spatial interaction. Ann. Probab. **26**(6), 3840–3909 (2016)
52. Majda, A.J., Franzke, C., Khouider, B.: An applied mathematics perspective on stochastic modelling for climate. Philos. Trans. R. Soc. A **336**(1875), 2429–2455 (2008)
53. Motsch, S., Tadmor, E.: Heterophilious dynamics enhances consensus. SIAM Rev. **56**(4), 577–621 (2014)

54. Muntean, A., Toschi (eds.), F.: Collective dynamics from bacteria to crowds: an excursion through modeling, analysis and simulation. CISM International Centre for Mechanical Sciences. Courses and Lectures, vol. 533. Springer, Vienna (2014)
55. Mitoma, I.: Tightness of probabilities on C([0, 1]; $\mathcal{S}'$) and D([0, 1]; $\mathcal{S}'$). Ann. Probab. **11**(4), 989–999 (1983)
56. Morse, M.R., Spiliopoulos, K.: Moderate deviations principle for systems of slow–fast diffusions. Asymptot. Anal. **105**(3–4), 97–135 (2017)
57. Morse, M.R., Spiliopoulos, K.: Importance sampling for slow–fast diffusions based on moderate deviations. SIAM J. Multiscale Model. Simul. **18**(1), 315–350 (2020)
58. Lipster, R.: Large deviations for two scaled diffusions. Probab. Theory Relat. Fields **106**(1), 71–104 (1996)
59. Orrieri, C.: Large deviations for interacting particle systems: joint mean-field and small-noise limit. Electron. J. Probab. **25**(11), 1–44 (2020)
60. Pardoux, E., Veretennikov, A.Y.: On Poisson equation and diffusion approximation I. Ann. Probab. **29**(3), 1061–1085 (2001)
61. Pardoux, E., Veretennikov, A.Y.: On Poisson equation and diffusion approximation II. Ann. Probab. **31**(3), 1166–1192 (2003)
62. Rauch, J.: Partial Differential Equations. Springer, New York (1991)
63. Réemillard, B., Vaillancourt, J.: On signed measure valued solutions of stochastic evolution equations. Stoch. Process. Appl. **124**(1), 101–122 (2014)
64. Röockner, M., Xie, L.: Diffusion approximation for fully coupled stochastic differential equations. Ann. Probab. **49**(3), 101–122 (2021)
65. Röockner, M., Sun, X., Xie, Y.: Strong convergence order for slow–fast McKean–Vlasov stochastic differential equations. Ann. Inst. Henri Poincaré Probab. Stat. **57**(1), 547–576 (2021)
66. Spiliopoulos, K.: Large deviations and importance sampling for systems of slow–fast motion. Appl. Math. Optim. **67**, 123–161 (2013)
67. Spiliopoulos, K.: Fluctuation analysis and short time asymptotics for multiple scales diffusion processes. Stoch. Dyn. **14**(3), 1350026 (2014)
68. Spiliopoulos, K.: Quenched large deviations for multiscale diffusion processes in random environments. Electron. J. Probab. **20**(15), 1–29 (2015)
69. Spiliopoulos, K., Sirignano, J.A., Giesecke, K.: Fluctuation analysis for the loss from default. Stoch. Process. Appl. **124**(7), 2322–2362 (2014)
70. Veretennikov, AYu.: Bounds for the mixing rates in the theory of stochastic equations. Theory Probab. Appl. **32**, 273–281 (1987)
71. Veretennikov, A.Y.: On large deviations in the averaging principle for SDEs with a “full dependence”. Correction (2005) arXiv:math/0502098 [math.PR]. Initial article in Annals of Probability, Vol. 27, No. 1, pp. 284-296. (1999)
72. Veretennikov, A.Y.: On large deviations for SDEs with small diffusion and averaging. Stoch. Process. Appl. **89**(1), 69–79 (2000)
73. Wang, F.Y.: Distribution dependent SDEs for Landau type equations. Stoch. Process. Appl. **128**(2), 595–621 (2017)
74. Zwanzig, R.: Diffusion in a rough potential. Proc. Natl. Acad. Sci. **85**, 2029–2030 (1988)