Homogenization of Dissipative, Noisy, Hamiltonian Dynamics

Jeremiah Birrell^a, Jan Wehr^{a,b}

^a*Department of Mathematics,*

^b*Program in Applied Mathematics*

University of Arizona

Tucson, AZ, 85721, USA

Abstract

We study the dynamics of a class of Hamiltonian systems with dissipation, coupled to noise, in a singular (small mass) limit. We derive the homogenized equation for the position degrees of freedom in the limit, including the presence of a *noise-induced drift* term. We prove convergence to the solution of the homogenized equation in probability and, under stronger assumptions, in an L^p -norm. Applications cover the overdamped limit of particle motion in a time-dependent electromagnetic field, on a manifold with time-dependent metric, and the dynamics of nuclear matter.

Keywords: Hamiltonian system, homogenization, small mass limit, noise-induced drift

2010 MSC: 60H10, 82C31

1. Introduction

In the simplest case, the motion of a diffusing particle of non-zero mass, m , is governed by a stochastic differential equation (SDE) of the form

$$dq_t = v_t dt, \quad m dv_t = -\gamma v_t dt + \sigma dW_t, \quad (1.1)$$

where γ and σ are the dissipation (or drag) and diffusion coefficients respectively and W_t is a Wiener process. The study of diffusive systems in the limit $m \rightarrow 0$ was initiated by Smoluchowski in [1] and continued by Kramers in [2]. The field has grown to explore a large array of models and phenomena, including coupled fluid-particle systems [3], relativistic diffusion [4, 5], and a variety

of processes and convergence modes on manifolds [6, 7, 8, 9, 10, 11, 12, 13]. History of the subject and a review of the early literature can be found in [14]. Such problems can be classified under the broad umbrella of homogenization, for which [15] is an excellent reference.

Recently, there has been increased interest in the phenomenon of *noise-induced drift*, which arises when the drag and noise coefficients are state dependent. In such cases, the equation governing the process in the limit $m \rightarrow 0$ possesses an additional drift term that was not present in the original system. First derived in [16], this has been observed experimentally in [17] and derived rigorously for one dimensional systems [18], systems satisfying the fluctuation-dissipation relation [16], in Euclidean space of arbitrary dimension [19, 20], and on compact Riemannian manifolds of arbitrary dimension [21]. Further references to work on the phenomenon of noise-induced drift are found in [19].

Statistical mechanics of fluctuating systems, as reviewed in [22, 23], covers systems more general than those governed by the Hamiltonians with quadratic kinetic energy,

$$H(q, p) = \frac{\|p\|^2}{2m} + V(q), \quad (1.2)$$

but to this point, the study of noise-induced drift has been restricted to Hamiltonians quadratic in p . In this paper, we extend the theory to a large class of Hamiltonian systems generalizing Eq. (1.2). See Section 2 for examples of the type of systems that are covered. We prove that solutions to these more general Hamiltonian systems converge in probability and, under stronger assumptions, in an L^p -norm to solutions of a homogenized limiting equation with a noise-induced drift term, for which we derive an explicit formula. This is a far-reaching generalization of the previous results about the $m \rightarrow 0$ limit of the equations Eq. (1.1).

1.1. Dissipative Hamiltonian System with Noise

Here, we review the basic equations and properties of dissipative, noisy Hamiltonian systems. See also [22]. Given a time-dependent Hamiltonian $H(t, x)$ which is C^1 jointly in $t \in \mathbb{R}$ and $x = (q, p) \in \mathbb{R}^n \times \mathbb{R}^n$, a positive-semi-definite continuous matrix-valued function $\Gamma(t, x)$, the matrix

$$\Pi = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad (1.3)$$

and a continuous vector field $G(t, x)$, we first consider the following deterministic equation

$$\dot{x}_t = -\Gamma(t, x_t)\nabla H(t, x_t) + \Pi\nabla H(t, x_t) + G(t, x_t). \quad (1.4)$$

This equation describes the dynamics of a dissipative Hamiltonian system with drag matrix Γ and external forcing G . We will refer to q as the position degrees of freedom and to p as the momentum degrees of freedom.

The rate of change of the Hamiltonian along a solution is given by

$$\begin{aligned} \frac{d}{dt}H(t, x_t) & \\ = \partial_t H(t, x_t) - \nabla H(t, x_t) \cdot \Gamma(t, x_t) \nabla H(t, x_t) & \\ + \nabla H(t, x_t) \cdot \Pi \nabla H(t, x_t) + \nabla H(t, x_t) \cdot G(t, x_t) & \\ \leq \partial_t H(t, x_t) + \nabla H(t, x_t) \cdot G(t, x_t), & \end{aligned} \quad (1.5)$$

where we used the anti-symmetry of Π and the positive semi-definiteness of Γ . In particular, if H is time independent and G vanishes then the energy is non-increasing and if Γ also vanishes then energy is conserved. This justifies the interpretation of Eq. (1.4) as a dissipative Hamiltonian system with external forcing G and drag matrix Γ .

We specialize to the case where the dissipation and external force enter only the momentum equation:

$$\Gamma(t, x) = \begin{pmatrix} 0 & 0 \\ 0 & \gamma(t, x) \end{pmatrix}, \quad (1.6)$$

and $G(t, x) = (0, F(t, x))$. With this, the dissipation couples linearly to the generalized velocity $v = \nabla_p H$, since the equations are now

$$\dot{q}_t = \nabla_p H(t, x_t), \quad \dot{p}_t = -\gamma(t, x_t) \nabla_p H(t, x_t) - \nabla_q H(t, x_t) + F(t, x_t). \quad (1.7)$$

We will be interested in families of Hamiltonians depending on some parameter $\epsilon \in (0, \epsilon_0]$ of the form

$$H^\epsilon(t, q, p) \equiv K^\epsilon(t, q, p) + V(t, q) \equiv K(\epsilon, t, q, (p - \psi(t, q))/\sqrt{\epsilon}) + V(t, q), \quad (1.8)$$

where $V = V(t, q)$ is C^2 , $K = K(\epsilon, t, q, z)$ is non-negative and C^2 in (t, q, z) for each ϵ , and ψ is a C^2 , $\mathbb{R}^n$ -valued function.

Remark 1. *The momentum-dependent term, K , and the momentum-independent term, V , into which we split the Hamiltonian, do not have to carry with it the physical interpretation of kinetic and potential energy respectively, though we will use that terminology. The splitting will become constrained (though not quite unique) by further assumptions we will make below, but at this point it is largely arbitrary.*

Families of Hamiltonians of the form Eq. (1.8) generalize the Hamiltonian of a classical particle coupled to an electromagnetic field (or in the case of vanishing vector potential, simply a Newtonian particle in a potential):

$$H(t, q, p) = \frac{\|p - e\phi(t, q)\|^2}{2m} + eV(t, q), \quad (1.9)$$

where e is the charge, ϕ is the vector potential, and V is the electrostatic potential. Scaling $p - e\phi(t, q)$ with $\sqrt{\epsilon}$ as in Eq. (1.8) is equivalent to replacing m with ϵm , hence taking $\epsilon \rightarrow 0^+$ is equivalent to the small mass limit, $m \rightarrow 0$. In this case $K(\epsilon, t, q, z)$ does not depend on ϵ , but in general it is useful to allow an additional ϵ dependence. The form of this dependence will be somewhat constrained as we proceed.

Adding a noise term to the momentum components of Hamilton's equations, we arrive at the following family of SDEs:

$$dq_t^\epsilon = \nabla_p H^\epsilon(t, x_t^\epsilon) dt, \quad (1.10)$$

$$dp_t^\epsilon = (-\gamma(t, x_t^\epsilon) \nabla_p H^\epsilon(t, x_t^\epsilon) - \nabla_q H^\epsilon(t, x_t^\epsilon) + F(t, x_t^\epsilon)) dt + \sigma(t, x_t^\epsilon) dW_t, \quad (1.11)$$

where $\sigma : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times k}$ is continuous and W_t is a $\mathbb{R}^k$ -valued Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ satisfying the usual conditions [24]. In this paper we investigate the behavior of x_t^ϵ in the limit $\epsilon \rightarrow 0^+$ and derive a homogenized SDE satisfied by the limiting position process, q_t .

1.2. Summary of the Main Results

To prove our first main theorem, Theorem 5.1, we will require several assumptions — namely Assumptions 1-7 in Appendix A. Each one is restated in the body of the paper when it is first used, as not all are required for each result. These assumptions will constrain the initial conditions, the analytical properties and form of the Hamiltonian, the drag matrix, γ , and the noise coefficients, σ . In particular, we will eventually require γ to be independent of

p and its eigenvalues to satisfy a positive lower bound. The latter coercivity requirement will be crucial in proving the kinetic energy and momentum bounds in Section 3. Under these assumptions we will prove the following:

Let x_t^ϵ be a family of solutions to the SDE 1.10-1.11 with initial condition $x_0^\epsilon = (q_0^\epsilon, p_0^\epsilon)$. In this paper, we work under the assumption that a unique solution (pathwise uniqueness) exists for all $t \geq 0$ (i.e. there are no explosions). See Appendix C for assumptions that guarantee this.

Then, as $\epsilon \rightarrow 0$, q_t^ϵ approaches the solution, q_t , to the SDE

$$dq_t = \tilde{\gamma}^{-1}(t, q_t)(-\partial_t \psi(t, q_t) - \nabla_q V(t, q_t) + F(t, q_t, \psi(t, q_t)))dt + S(t, q_t)dt + \tilde{\gamma}^{-1}(t, q_t)\sigma(t, q_t, \psi(t, q_t))dW_t. \quad (1.12)$$

The objects appearing in the SDE are defined as follows (here, and in the rest of the paper, we employ the summation convention on repeated indices):

1. $\tilde{\gamma}_{ik}(t, q) \equiv \gamma_{ik}(t, q) + \partial_{q^k} \psi_i(t, q) - \partial_{q^i} \psi_k(t, q)$
2. $S^i(t, q) \equiv Q^{ijl}(t, q)J_{jl}(t, q, \psi(t, q))$
3. $Q^{ijl}(t, q) \equiv \partial_{q^k}(\tilde{\gamma}^{-1})^{ij}(t, q)A^{kl}(t, q) - \frac{1}{2}(\tilde{\gamma}^{-1})^{ik}(t, q)\partial_{q^k}A^{jl}(t, q)$, where A^{ij} is the matrix-valued function from Assumption 5 and the index placement on $\tilde{\gamma}^{-1}$ is defined by $(\tilde{\gamma}^{-1})^{ij}\tilde{\gamma}_{jk} = \delta_k^i$.
4. $J_{ij}(t, x) \equiv G_{ij}^{kl}(t, q)\Sigma_{kl}(t, x)$
5. $G_{ij}^{kl}(t, q) \equiv \int_0^\infty (e^{-y(A\tilde{\gamma})(t, q)})_i^k (e^{-y(A\tilde{\gamma})(t, q)})_j^l dy$, where $A\tilde{\gamma}$ is the matrix $(A\tilde{\gamma})_j^i = \tilde{\gamma}_{jk}A^{ki}$.
6. $\Sigma_{ij} \equiv \sum_\rho \sigma_{i\rho}\sigma_{j\rho}$

It is interesting to note that the only feature of the kinetic energy function, K , that plays a part in the limiting equation is the “metric tensor” A^{ij} from Assumption 5.

The convergence is in the following sense:

Suppose that for all $\epsilon > 0$ and all $p > 0$ we have $E[\|q_0^\epsilon\|^p] < \infty$, $E[\|q_0\|^p] < \infty$, and $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2})$. Then for any $T > 0$, $p > 0$, $0 < \beta < p/2$ we have

$$E\left[\sup_{t \in [0, T]} \|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p\right] = O(\epsilon^\beta) \text{ and } E\left[\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p\right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (1.13)$$

We also prove a convergence in probability result,

$$\lim_{\epsilon \rightarrow 0^+} P\left(\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\| > \delta\right) = 0 \text{ for all } T > 0, \delta > 0, \quad (1.14)$$

under less restrictive assumptions than the above L^p result. See Theorem 6.1 for details.

The drift term, $S(t, q)$, that appears in the limiting equation is called the *noise-induced drift* and is nonzero when σ is nonzero and a particular combination of A , γ , and ψ have non-trivial state dependence. Other works studying the small mass limit of inertial systems, both in Euclidean space [18, 16, 19, 20], and on manifolds [21], have found analogous phenomena.

In addition to applying to a much larger class of Hamiltonians, our derivation here gives a unified treatment of two previously studied systems: a particle in an electromagnetic field, [19], and a particle on a Riemannian manifold, [21]. These previous works used different, and somewhat specialized, methods. Our results here also expand on these by allowing the metric tensor, forces, drag, and diffusion to be time-dependent. For Riemannian manifolds, we differ here by considering the non-compact case. See Section 2 for details and further examples.

1.3. Outline of the Proof

The full details of the proof begin in Section 3. Here we outline our strategy and main ideas.

- In Section 3 we prove several results bounding the expectation of the kinetic energy. For example, we show that for $T > 0$, $q > 0$ we have

$$\sup_{t \in [0, T]} E[K^\epsilon(t, x_t^\epsilon)^q] = O(1) \text{ as } \epsilon \rightarrow 0^+. \quad (1.15)$$

We use this to conclude several convergence results involving the momentum, for example

$$\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^q] = O(\epsilon^{q/2}) \text{ as } \epsilon \rightarrow 0^+. \quad (1.16)$$

The main tools for computing the estimates are Itô's formula, several well known (stochastic) integral inequalities, and a lesser known P -a.s. stochastic integral inequality, Lemma 3.3.

- In Section 4 we derive the proposed form of the limiting equation for the position variables by solving Eq. (1.11) for $\nabla_p H^\epsilon(t, x_t^\epsilon) dt$, substituting into Eq. (1.10), and integrating by parts to separate out the components that depend on $p_t^\epsilon - \psi(t, q_t^\epsilon)$. This is done with the aim of later using the results of Section 3 to prove that these terms vanish in the limit.

The main complication here is that, in general, the required separation is possible only through formulating and solving an appropriate Lyapunov equation, Eq. (4.15).

- In Section 5 we prove L^p -convergence of q_t^ϵ to the solution, q_t , of the proposed limiting equation from Section 4. This is accomplished by a Gronwall's inequality argument. The estimates of Section 3 are the critical ingredient here, allowing us to prove that the error terms converge to zero as $\epsilon \rightarrow 0^+$. This result relies on the assumption that the gradient of the potential is bounded (among others).
- Finally, in Section 6 we use a technique adapted from [20] to prove convergence in probability, Eq. (1.14), for a much wider range of systems, including many whose potentials have unbounded gradient.

The core idea is that, by modifying the objects in the SDE to be compactly supported (or at least have compactly supported derivatives) we can use the L^p convergence result from Section 5 together with a limiting argument to prove convergence in probability. A crucial ingredient is that none of the solutions (of the original or modified SDEs) explode in finite time. This can be proven using the results of Appendix C.

2. Examples

Before we begin the proof, we first discuss several examples that fit within the above framework.

2.1. Particle in a Electromagnetic Field

The Hamiltonian of a particle of mass ϵm and charge e in an electromagnetic field with vector potential $\phi(t, q)$ and electrostatic potential $V(t, q)$ is

$$H^\epsilon(t, q, p) = \frac{1}{2\epsilon m} \|p - e\phi(t, q)\|^2 + eV(t, q). \quad (2.1)$$

Allowing for an additional forcing, F , Hamilton's equations for this system are

$$dq_t^\epsilon = \frac{1}{\epsilon m} (p_t^\epsilon - e\phi(t, q_t^\epsilon)) dt, \quad (2.2)$$

$$d(p_t^\epsilon)_i = \left(-\frac{1}{\epsilon m} \gamma_{ij}(t, q_t^\epsilon) \delta^{jk} ((p_t^\epsilon)_k - e\phi_k(t, q_t^\epsilon)) + F_i(t, x_t^\epsilon) - e\partial_{q^i} V(t, q_t^\epsilon) \right. \\ \left. + \frac{e}{\epsilon m} \partial_{q^i} \phi_k(t, q_t^\epsilon) \delta^{jk} ((p_t^\epsilon)_j - e\phi_j(t, q_t^\epsilon)) \right) dt + \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho. \quad (2.3)$$

The homogenized equation in the small mass limit is difficult to simplify further than Eq. (1.12) in general. However, in the case where γ and σ are independent of p and the fluctuation dissipation relation holds pointwise for a time and position dependent “temperature” $T(t, q)$,

$$\Sigma_{ij}(t, q) = 2k_B T(t, q) \gamma_{ij}(t, q), \quad (2.4)$$

one can show that

$$G_{kl}^{ab}(t, q) \Sigma_{ab}(t, q) = k_B T(t, q) \delta_{kl}, \quad (2.5)$$

where G was defined in Eq. (4.19), $\Sigma_{ij} = \sum_\rho \sigma_{i\rho} \sigma_{j\rho}$, and k_B is Boltzmann's constant.

The noise induced drift, Eq. (4.27), can therefore be simplified to

$$S^i(t, q) = k_B T(t, q) \partial_{q^j} (\tilde{\gamma}^{-1})^{ij}(t, q_t). \quad (2.6)$$

Recall that we defined

$$\tilde{\gamma}_{ik}(t, q) \equiv \gamma_{ik}(t, q) + \partial_{q^k} \psi_i(t, q) - \partial_{q^i} \psi_k(t, q), \quad (2.7)$$

where here, $\psi = e\phi$.

The homogenized equation in the small mass limit is then

$$dq_t^i = (\tilde{\gamma}^{-1})^{ij}(t, q_t) (-\partial_t \psi_j(t, q_t) - e\partial_{q^j} V(t, q_t) + F_j(t, q_t, \psi(t, q_t))) dt \\ + k_B T(t, q) \partial_{q^j} (\tilde{\gamma}^{-1})^{ij}(t, q_t) dt + (\tilde{\gamma}^{-1})^{ij}(t, q_t) \sigma_{j\rho}(t, q_t) dW_t^\rho. \quad (2.8)$$

The time independent case was studied in [19] by a different method and coincides with the results in this paper.

2.2. Particle on a Riemannian Manifold

Another case that is covered by the framework developed here is the inertial motion of a particle in $\mathbb{R}^n$, but with geometry specified by a time-dependent Riemannian metric tensor, $g_{ij}(t, q)$. The family of Hamiltonians describing this system is

$$H^\epsilon(t, q, p) = \frac{1}{2\epsilon m} g^{ij}(t, q) p_i p_j. \quad (2.9)$$

Note that the inverse metric tensor, $g^{ij}(t, q)$, is playing the role of $A^{ij}(t, q)$ in our formalism, and so all the assumptions that are required of A^{ij} there must be satisfied by g^{ij} here.

Allowing for external forcing, F , Hamilton's equations are

$$d(q_t^\epsilon)^i = \frac{1}{\epsilon m} g^{ij}(t, q_t^\epsilon) (p_t^\epsilon)_j dt, \quad (2.10)$$

$$d(p_t^\epsilon)_i = \left(-\frac{1}{\epsilon m} \gamma_{ij}(t, x_t^\epsilon) g^{jk}(t, q_t^\epsilon) (p_t^\epsilon)_k - \frac{1}{2\epsilon m} \partial_{q^i} g^{kl}(t, q_t^\epsilon) (p_t^\epsilon)_k (p_t^\epsilon)_l + F_i(t, x_t^\epsilon) \right) dt + \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho. \quad (2.11)$$

Again, the homogenized equation in the small mass limit, Eq. (4.26), can be simplified if γ and σ are independent of p and the fluctuation dissipation relation holds pointwise for a time and position dependent “temperature” $T(t, q)$,

$$\Sigma_{ij}(t, q) = 2k_B T(t, q) \gamma_{ij}(t, q). \quad (2.12)$$

In this case one finds that

$$G_{kl}^{ab}(t, q) \Sigma_{ab}(t, q) = k_B T(t, q) g_{kl}(t, q) \quad (2.13)$$

and hence Eq. (4.26) becomes

$$dq_t^i = (\gamma^{-1})^{ij}(t, q_t) F_j(t, q_t, 0) dt + S^i(t, q_t) dt + (\gamma^{-1})^{ij}(t, q_t) \sigma_{j\rho}(t, q_t) dW_t^\rho, \quad (2.14)$$

where the noise induced drift is

$$S^i(t, q) = k_B T(t, q) \left(\partial_{q^j} (\gamma^{-1})^{ij}(t, q) - \frac{1}{2} (\gamma^{-1})^{ij}(t, q) g_{kl}(t, q) \partial_{q^j} g^{kl}(t, q) \right). \quad (2.15)$$

See also [21], which treats the case of a smooth, compact, connected, manifold without boundary (but otherwise arbitrary topology) and with time independent metric via a more geometrically motivated approach.

One can argue that the present approach is simpler, as we only rely on tools from (stochastic) analysis; we avoid the geometrical machinery used in [21]. In addition, here we don't require compactness; on the other hand, we do lose the ability to handle non-trivial topology of the manifold.

2.3. Hamiltonian that are Polynomials in the Momentum

Generalizing the above two quadratic cases, our convergence result applies to Hamiltonians that are polynomials in $p - \psi(t, q)$. Specifically, Theorem 5.1 implies L^p convergence if the family of Hamiltonians has the form

$$H^\epsilon(t, q, p) = \sum_{l=k_1}^{k_2} d_l(t) \left[A^{ij}(t, q) (p - \psi(t, q))_i (p - \psi(t, q))_j / \epsilon \right]^l + V(t, q) \quad (2.16)$$

where $1 \leq k_1 \leq k_2$ are integers and the following properties hold on $[0, T] \times \mathbb{R}^n$ for every $T > 0$:

1. V is C^2 and $\nabla_q V$ is bounded and Lipschitz in q , uniformly in t .
2. ψ is C^3 and $\partial_t \psi$, $\partial_{q^i} \psi$, $\partial_{q^i} \partial_{q^j} \psi$, $\partial_t \partial_{q^i} \psi$, $\partial_t \partial_{q^j} \partial_{q^i} \psi$, and $\partial_{q^i} \partial_{q^j} \partial_{q^i} \psi$ are bounded.
3. d_l are C^2 and non-negative.
4. d_{k_1} and d_{k_2} are uniformly bounded below by a positive constant.
5. A is C^2 , positive-definite, and A , $\partial_t A$, $\partial_{q^i} A$, $\partial_t \partial_{q^i} A$, and $\partial_{q^i} \partial_{q^j} A$ are bounded.
6. The eigenvalues of A are uniformly bounded below by a positive constant.

As mentioned above, convergence in probability holds under much weaker assumptions on the potential. See Theorem 6.1.

2.4. Effective Nuclear Interactions

Mean field models of nuclear interactions can lead to non-quadratic momentum dependence in the Hamiltonian. For example, in [25] a contribution to the effective potential of the form

$$U(p) = c_1 \ln^2 [1 + c_2 \|p\|^2] \quad (2.17)$$

was calculated. For non-relativistic particle motion, a term of this form can be accommodated in our framework in several ways, depending on which parameters one wishes to scale. For example, one can let

$$K^\epsilon(p) = \frac{\|p\|^2}{2\epsilon m} + c_1 \ln^2 [1 + c_2 \|p\|^2] \quad (2.18)$$

or

$$K^\epsilon(p) = \frac{\|p\|^2}{2\epsilon m} + c_1 \ln^2 [1 + c_2 \|p\|^2 / \epsilon]. \quad (2.19)$$

In either case, the assumptions of Theorem 5.1 relating to K are satisfied.

3. Kinetic Energy and Momentum Bounds

We now begin working towards the proof of our main results. In this section, we derive bounds on the behavior of the kinetic energy in the limit $\epsilon \rightarrow 0^+$. As a consequence we will obtain a convergence result for the canonical momentum, the first formula in Eq. (1.13). Some assumptions on the structure of the Hamiltonian are required. As usual, here and in the sequel, generic symbols, denoting constants, such as C , M etc., do not have to have the same value in all equations.

Assumption 1. *We assume that σ , F , and γ are continuous, the Hamiltonian has the form given in Eq. (1.8) where $K(\epsilon, t, q, z)$ is non-negative and C^2 in (t, q, z) for each ϵ , ψ is C^2 , and the solutions, x_t^ϵ , to the SDE 1.10-1.11 exist for all $t \geq 0$.*

For every $T > 0$, we assume the following bounds hold on $(0, \epsilon_0] \times [0, T] \times \mathbb{R}^{2n}$:

1. *There exist $C > 0$ and $M > 0$ such that*

$$\max\{|\partial_t K(\epsilon, t, q, z)|, \|\nabla_q K(\epsilon, t, q, z)\|\} \leq M + CK(\epsilon, t, q, z). \quad (3.1)$$

2. There exist $c > 0$ and $M \geq 0$ such that

$$\|\nabla_z K(\epsilon, t, q, z)\|^2 + M \geq cK(\epsilon, t, q, z). \quad (3.2)$$

3. For every $\delta > 0$ there exists an $M > 0$ such that

$$\max \left\{ \|\nabla_z K(\epsilon, t, q, z)\|, \left(\sum_{i,j} |\partial_{z_i} \partial_{z_j} K(\epsilon, t, q, z)|^2 \right)^{1/2} \right\} \leq M + \delta K(\epsilon, t, q, z). \quad (3.3)$$

We will also need the following assumptions concerning the potential, dissipation, noise, external forcing, and initial conditions. Some of these will be relaxed in Section 6.

Assumption 2. For every $T > 0$, we assume that the following hold uniformly on $[0, T] \times \mathbb{R}^n$:

1. V is C^2 and $\nabla_q V$ is bounded.
2. γ is symmetric with eigenvalues bounded below by some $\lambda > 0$.
3. γ , F , $\partial_t \psi$, and σ are bounded.
4. There exists $C > 0$ such that the (random) initial conditions satisfy $K^\epsilon(0, x_0^\epsilon) \leq C$ for all $\epsilon > 0$ and all $\omega \in \Omega$.

We now state and prove the kinetic energy bound which underlies our main results. As with our notation K^ϵ , for any function $f(\epsilon, t, q, z)$ we define

$$f^\epsilon(t, x) \equiv f(\epsilon, t, q, (p - \psi(t, q))/\sqrt{\epsilon}). \quad (3.4)$$

For example,

$$(\partial_{z_i} K)^\epsilon(t, x) \equiv \partial_{z_i} K(\epsilon, t, q, (p - \psi(t, q))/\sqrt{\epsilon}) \quad (3.5)$$

and similarly for $(\nabla_z K)^\epsilon$, $(\partial_{z_i} \partial_{z_j} K)^\epsilon$, etc.

Lemma 3.1. Under Assumptions 1 and 2, for any $q \in \mathbb{N}$, $q \geq 1$ and any $T > 0$ there exist $\alpha_0 > 0$, $\epsilon_0 > 0$ and $\kappa > 0$ such that for all $0 < \alpha \leq \alpha_0$, $0 < \epsilon \leq \epsilon_0$, $0 \leq t \leq T$ we have the P -a.s. inequality

$$K^\epsilon(t, x_t^\epsilon)^q \leq \frac{\kappa}{\alpha} + \frac{q}{\sqrt{\epsilon}} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s. \quad (3.6)$$

Proof. Take $T > 0$, $q \in \mathbb{N}$, $q \geq 1$, $\alpha > 0$, and apply Itô's formula to $e^{\alpha t/\epsilon} K^\epsilon(t, x_t^\epsilon)^q$:

$$\begin{aligned}
e^{\alpha t/\epsilon} K^\epsilon(t, x_t^\epsilon)^q &= K^\epsilon(0, x_0^\epsilon)^q + \frac{\alpha}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^q ds \\
&+ q \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\partial_s K)^\epsilon(s, x_s^\epsilon) ds \\
&- \frac{q}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \gamma(s, x_s^\epsilon) (\nabla_z K)^\epsilon(s, x_s^\epsilon) ds \\
&+ \frac{q}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot (-\partial_s \psi(s, q_s^\epsilon) - \nabla_q V(s, q_s^\epsilon) + F(s, x_s^\epsilon)) ds \\
&+ \frac{q(q-1)}{2\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-2} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \Sigma(s, x_s^\epsilon) (\nabla_z K)^\epsilon(s, x_s^\epsilon) ds \\
&+ \frac{q}{2\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\partial_{z_i} \partial_{z_j} K)^\epsilon(s, x_s^\epsilon) \Sigma_{ij}(s, x_s^\epsilon) ds \\
&+ \frac{q}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s.
\end{aligned} \tag{3.7}$$

Here we used the fact that $\nabla_p H^\epsilon = \nabla_p K^\epsilon$ to cancel the terms involving $\nabla_q K^\epsilon$.

Using Assumption 2, for any $t \in [0, T]$ we have

$$\begin{aligned}
&e^{\alpha t/\epsilon} K^\epsilon(t, x_t^\epsilon)^q \\
&\leq K^\epsilon(0, x_0^\epsilon)^q + \frac{\alpha}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^q ds \\
&+ q \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\partial_s K)^\epsilon(s, x_s^\epsilon) ds \\
&- \frac{q\lambda}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^2 ds \\
&+ \frac{q}{\sqrt{\epsilon}} \|-\partial_t \psi - \nabla_q V + F\|_\infty \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\| ds \\
&+ \frac{q(q-1)}{2\epsilon} \|\Sigma\|_\infty \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-2} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^2 ds \\
&+ \frac{q}{2\epsilon} \|\Sigma\|_{F,\infty} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} \left(\sum_{i,j} (\partial_{z_i} \partial_{z_j} K)^\epsilon(s, x_s^\epsilon)^2 \right)^{1/2} ds \\
&+ \frac{q}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s.
\end{aligned} \tag{3.8}$$

Here and in the following, $\|Y\|_F$ will denote the Frobenius (or Hilbert-Schmidt) norm of a matrix Y , i.e. $\|Y\|_F = \left(\sum_{i,j} Y_{ij}^2 \right)^{\frac{1}{2}}$. For any matrix or vector-valued

quantity Y we write $\|Y\|_\infty \equiv \sup_{(t,x) \in [0,T] \times \mathbb{R}^{2n}} \|Y(t,x)\|$ and similarly for $\|\cdot\|_{F,\infty}$. The implied value of T will be clear from the context.

For any $\delta > 0$, Assumption 1 implies the existence of $C > 0$, $c > 0$, and $M > 0$ such that

$$\begin{aligned}
& e^{\alpha t/\epsilon} K^\epsilon(t, x_t^\epsilon)^q \\
& \leq K^\epsilon(0, x_0^\epsilon)^q + \frac{\alpha}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^q ds \\
& \quad + q \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (M + CK^\epsilon(s, x_s^\epsilon)) ds \\
& \quad - \frac{q\lambda}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (cK^\epsilon(s, x_s^\epsilon) - M) ds \\
& \quad + \frac{q}{\sqrt{\epsilon}} \|\partial_t \psi - \nabla_q V + F\|_\infty \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (M + \delta K^\epsilon(s, x_s^\epsilon)) ds \\
& \quad + \frac{q(q-1)}{2\epsilon} \|\Sigma\|_\infty \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-2} (M + \delta K^\epsilon(s, x_s^\epsilon))^2 ds \\
& \quad + \frac{q}{2\epsilon} \|\Sigma\|_{F,\infty} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (M + \delta K^\epsilon(s, x_s^\epsilon)) ds \\
& \quad + \frac{q}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s.
\end{aligned} \tag{3.9}$$

In the estimate that follows, the first two terms and the last term of the above expression will be left unchanged. To estimate the remaining terms, we will use the elementary inequalities $(a+b)^2 \leq 2(a^2 + b^2)$ and

$$K^{q-1} \leq \left(\frac{M}{\delta}\right)^{q-1} + \frac{\delta}{M} K^q, \tag{3.10}$$

$$K^{q-2} \leq \left(\frac{M}{\delta}\right)^{q-2} + \left(\frac{\delta}{M}\right)^2 K^q. \tag{3.11}$$

The inequalities involving K are obtained by looking at the cases $K \leq M/\delta$ and $K > M/\delta$. The first holds for every $q \geq 1$ and the second for every $q \geq 2$. Note that for $q = 1$, the term containing $K^\epsilon(s, x_s^\epsilon)^{q-2}$ vanishes. Applying these inequalities yields

$$\begin{aligned}
& K^\epsilon(t, x_t^\epsilon)^q \\
& \leq e^{-\alpha t/\epsilon} K^\epsilon(0, x_0^\epsilon)^q + \frac{D}{\alpha} - \frac{d}{\epsilon} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^q ds \\
& \quad + \frac{q}{\sqrt{\epsilon}} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s,
\end{aligned} \tag{3.12}$$

where

$$\begin{aligned}
D &= qM \left(\frac{M}{\delta} \right)^{q-1} \left[\lambda + \epsilon + \sqrt{\epsilon} \|\partial_t \psi - \nabla_q V + F\|_\infty + \frac{1}{2} \|\Sigma\|_{F,\infty} \right] \\
&\quad + q(q-1)M^2(M/\delta)^{q-2} \|\Sigma\|_\infty, \\
d &= qc\lambda - \alpha - qC\epsilon - q\delta\sqrt{\epsilon} \|\partial_t \psi - \nabla_q V + F\|_\infty - \frac{q\delta}{2} \|\Sigma\|_{F,\infty} - q(q-1)\delta^2 \|\Sigma\|_\infty \\
&\quad - q(q-1)\delta^2 \|\Sigma\|_\infty - \left(q\lambda + q\epsilon + q\sqrt{\epsilon} \|\partial_t \psi - \nabla_q V + F\|_\infty + \frac{q}{2} \|\Sigma\|_{F,\infty} \right) \delta.
\end{aligned} \tag{3.13}$$

For all ϵ, δ, α sufficiently small, d is non-negative, and hence

$$\begin{aligned}
K^\epsilon(t, x_t^\epsilon)^q &\leq K^\epsilon(0, x_0^\epsilon)^q + \frac{D}{\alpha} \\
&\quad + \frac{q}{\sqrt{\epsilon}} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s.
\end{aligned} \tag{3.14}$$

By Assumption 2, $K^\epsilon(0, x_0^\epsilon)$ is bounded, so we are done. $\square$

We will use this bound to prove several results about the behavior of the kinetic energy and momentum as $\epsilon \rightarrow 0^+$.

3.1. Integrability of the Kinetic Energy

Lemma 3.2. *Under Assumptions 1 and 2, $E[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p]$ is finite for any $T > 0$, $\epsilon > 0$, and $p > 0$.*

Proof. Fix $T > 0$, $\epsilon > 0$. First, let $p > 2$. Given $M > 0$, define the stopping time $\tau_M = \inf\{t : K^\epsilon(t, x_t^\epsilon) = M\}$. By Assumption 2 we can take M large enough so that $K^\epsilon(0, x_0^\epsilon) < M$.

Let $t \leq T$ and raise Eq. (3.12) (with $q = 1$) to the p th power to obtain

$$\begin{aligned}
&K^\epsilon(t \wedge \tau_M, (x^\epsilon)_t^{\tau_M})^p \\
&\leq C_1 + C_2 \left(\int_0^t 1_{s \leq \tau_M} K^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M}) ds \right)^p \\
&\quad + C_3 \left| \int_0^t 1_{s \leq \tau_M} e^{\alpha(s \wedge \tau_M)/\epsilon} (\nabla_z K)^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M}) \cdot \sigma(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M}) dW_s \right|^p
\end{aligned} \tag{3.15}$$

where C_i are constants (that depend on ϵ and T).

Therefore, applying Hölder's inequality to the second term and the Burkholder-Davis-Gundy inequality to the third term, (see, for example, Theorem 3.28 in [24]), we obtain

$$\begin{aligned} & E \left[\sup_{s \in [0, t]} K^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M})^p \right] \\ & \leq C_1 + C_2 T^{p-1} \int_0^t E[K^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M})^p] ds \end{aligned} \quad (3.16)$$

$$\begin{aligned} & + C_4 E \left[\left(\int_0^t 1_{r \leq \tau_M} e^{2\alpha(r \wedge \tau_M)/\epsilon} \|(\nabla_z K)^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})\|^2 \right. \right. \\ & \quad \left. \left. \times \|\sigma(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})\|^2 dr \right)^{p/2} \right] \\ & \leq C_1 + C_2 T^{p-1} \int_0^t E \left[\sup_{r \in [0, s]} K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})^p \right] ds \\ & + C_5 E \left[\left(\int_0^t (1 + K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M}))^2 dr \right)^{p/2} \right]. \end{aligned} \quad (3.17)$$

By assumption, $p > 2$, so we can use Hölder's inequality again to obtain

$$\begin{aligned} & E \left[\sup_{s \in [0, t]} K(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M})^p \right] \\ & \leq C_1 + C_2 T^{p-1} \int_0^t E \left[\sup_{r \in [0, s]} K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})^p \right] ds \end{aligned} \quad (3.18)$$

$$\begin{aligned} & + C_5 T^{p/2-1} E \left[\int_0^t (1 + K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M}))^p dr \right] \\ & \leq C_6 + C_7 \int_0^t E \left[\sup_{r \in [0, s]} K^\epsilon(r \wedge \tau_M, (x^\epsilon)_r^{\tau_M})^p \right] ds, \end{aligned} \quad (3.19)$$

where C_i are independent of t and M .

By the definition of τ_M ,

$$\sup_{s \in [0, t]} K^\epsilon(s \wedge \tau_M, (x^\epsilon)_s^{\tau_M})^p \leq M \quad (3.20)$$

for all t . Therefore the integral in Eq. (3.19) is finite for all $t \leq T$ and Gronwall's inequality gives

$$E \left[\sup_{t \in [0, T]} K^\epsilon(t \wedge \tau_M, (x^\epsilon)_t^{\tau_M})^p \right] \leq C_6 e^{C_7 T}. \quad (3.21)$$

The C_i are independent of M , so taking $M \rightarrow \infty$ and using the Monotone Convergence Theorem implies

$$E \left[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p \right] \leq C_6 e^{C_7 T} < \infty. \quad (3.22)$$

This gives the result for $p > 2$. It follows for all $p > 0$ by an application of Hölder's inequality. $\square$

3.2. Supremum of the Expectation of the Kinetic Energy

Combining Lemmas 3.1 and 3.2 we can prove the following bound for the supremum of the expected value of the kinetic energy.

Proposition 3.1. *Under the Assumptions 1 and 2, for any $T > 0$, $q > 0$ we have*

$$\sup_{t \in [0, T]} E[K^\epsilon(t, x_t^\epsilon)^q] = O(1) \text{ as } \epsilon \rightarrow 0^+. \quad (3.23)$$

Proof. First take $T > 0$, $q \in \mathbb{N}$, $q \geq 1$. The following computation shows that

$$M_t \equiv \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \quad (3.24)$$

is a martingale (see [24]):

$$\begin{aligned} & E \left[\int_0^t \|e^{\alpha s/\epsilon} K^\epsilon(s, x_s^\epsilon)^{q-1} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon)\|^2 ds \right] \\ & \leq e^{2\alpha t/\epsilon} \|\sigma\|_\infty^2 t E \left[\sup_{s \in [0, t]} K^\epsilon(s, x_s^\epsilon)^{2(q-1)} (M + K^\epsilon(s, x_s^\epsilon))^2 \right] \\ & \leq 2e^{2\alpha t/\epsilon} \|\sigma\|_\infty^2 t \left(M^2 E \left[\sup_{s \in [0, t]} K^\epsilon(s, x_s^\epsilon)^{2(q-1)} \right] + E \left[\sup_{s \in [0, t]} K^\epsilon(s, x_s^\epsilon)^{2q} \right] \right) < \infty, \end{aligned} \quad (3.25)$$

where we used Assumption 1 and Lemma 3.2.

Therefore, taking the expectation of Eq. (3.6), we see that there exists $\kappa > 0$ such that for all $t \leq T$ and all α and ϵ sufficiently small, we have

$$E[K^\epsilon(t, x_t^\epsilon)^q] \leq \frac{\kappa}{\alpha}. \quad (3.26)$$

This proves the result for q a positive integer. The result then follows for arbitrary $q > 0$ by an application of Hölder's inequality. $\square$

Corollary 3.1. *We note that if the constants involved in the bounds from Assumptions 1-2 are valid uniformly for $t \in [0, \infty)$ (and not just $t \in [0, T]$) then we obtain the stronger bound*

$$\sup_{t \in [0, \infty)} E[K^\epsilon(t, x_t^\epsilon)^q] = O(1) \text{ as } \epsilon \rightarrow 0^+ \quad (3.27)$$

for any $q > 0$.

3.3. Expectation of the Supremum of the Kinetic Energy

We now have the ingredients to derive a bound on the expectation of the supremum of the kinetic energy. For this, we need to recall a special case of Lemma 5.1 from [21]:

Lemma 3.3. *Let $V \in L_{loc}^2(dt)$ be an $\mathbb{R}^k$ -valued process. For any $\alpha > 0$, $T \geq \delta > 0$ we have the P -a.s. bound*

$$\begin{aligned} & \sup_{t \in [0, T]} \left| \int_0^t e^{-\alpha(t-s)} V_s dW_s \right| \\ & \leq 5 \left(e^{-\alpha\delta} \sup_{t \in [0, T]} \left| \int_0^t V_r dW_r \right| + \max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t V_r dW_r \right| \right) \end{aligned} \quad (3.28)$$

where $N = \max\{k \in \mathbb{Z} : k\delta < T\}$.

Proposition 3.2. *Under Assumptions 1 and 2, for any $T > 0$, $p > 0$, $\beta > 0$ we have*

$$E \left[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p \right] = O(\epsilon^{-\beta}) \text{ as } \epsilon \rightarrow 0^+. \quad (3.29)$$

Proof. By Lemma 3.1 with $q = 1$, there exist $\alpha > 0$ and $\kappa > 0$ such that for all ϵ sufficiently small and all $t \in [0, T]$, the following bound holds a.s.:

$$K^\epsilon(t, x_t^\epsilon) \leq \frac{\kappa}{\alpha} + \frac{1}{\sqrt{\epsilon}} e^{-\alpha t/\epsilon} \int_0^t e^{\alpha s/\epsilon} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s. \quad (3.30)$$

We will first prove the proposition under the additional assumption $p > 2$. The general case $p > 0$ will follow by an application of Hölder's inequality.

$$\begin{aligned} & E \left[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p \right] \\ & \leq 2^{p-1} (\kappa/\alpha)^p + \frac{2^{p-1}}{\epsilon^{p/2}} E \left[\sup_{t \in [0, T]} \left| \int_0^t e^{-\alpha(t-s)/\epsilon} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right]. \end{aligned} \quad (3.31)$$

For any $T \geq \delta > 0$, Lemma 3.3 (with α/ϵ in place of α) implies

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t e^{-\alpha(t-s)/\epsilon} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right] \\
& \leq 5^p 2^{p-1} \left(e^{-p\alpha\delta/\epsilon} E \left[\sup_{t \in [0, T]} \left| \int_0^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right] \right. \\
& \quad \left. + E \left[\max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right] \right)
\end{aligned} \tag{3.32}$$

where $N = \max\{k \in \mathbb{Z} : k\delta < T\}$.

The Burkholder-Davis-Gundy inequality, applied to the first term on the right side of the inequality, implies existence of a constant $\tilde{C} > 0$ such that

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \\
& \leq \tilde{C} E \left[\left(\int_0^T \|(\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon)\|^2 dr \right)^{p/2} \right] \\
& \leq \tilde{C} \|\sigma\|_\infty^p E \left[\left(\int_0^T \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^2 dr \right)^{p/2} \right] \\
& \leq \tilde{C} \|\sigma\|_\infty^p E \left[\left(\int_0^T (M + K^\epsilon(s, x_s^\epsilon))^2 dr \right)^{p/2} \right].
\end{aligned} \tag{3.33}$$

In the last line, we used Assumption 1.

We have assumed $p > 2$, so we can use Hölder's inequality with exponents $p/(p-2)$ and $p/2$, to get

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \\
& \leq \tilde{C} \|\sigma\|_\infty^p T^{p/2-1} E \left[\int_0^T (M + K^\epsilon(s, x_s^\epsilon))^p dr \right] \\
& \leq 2^{p-1} \tilde{C} \|\sigma\|_\infty^p T^{p/2} \left(M^p + \sup_{s \in [0, T]} E[K^\epsilon(s, x_s^\epsilon)^p] \right) \\
& = O(1)
\end{aligned} \tag{3.34}$$

as $\epsilon \rightarrow 0^+$ by Proposition 3.1.

We now work on the second term in Eq. (3.32). Using the fact that the ℓ^∞ -norm on $\mathbb{R}^N$ is bounded by the $\ell^{\tilde{p}}$ norm for any $\tilde{p} \geq 1$, and then applying Hölder's inequality and the Burkholder-Davis-Gundy inequality, we derive the bound

$$\begin{aligned}
& E \left[\max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \quad (3.35) \\
& \leq E \left[\left(\sum_{k=0}^{N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^{p\tilde{p}} \right)^{1/\tilde{p}} \right] \\
& \leq \left(\sum_{k=0}^{N-1} E \left[\sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^{p\tilde{p}} \right] \right)^{1/\tilde{p}} \\
& \leq \left(\sum_{k=0}^{N-1} \tilde{C} E \left[\left(\int_{k\delta}^{(k+2)\delta} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon)\|^2 dr \right)^{p\tilde{p}/2} \right] \right)^{1/\tilde{p}} \\
& \leq \tilde{C}^{1/\tilde{p}} \|\sigma\|_\infty^p \left(\sum_{k=0}^{N-1} E \left[\left(\int_{k\delta}^{(k+2)\delta} \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^2 dr \right)^{p\tilde{p}/2} \right] \right)^{1/\tilde{p}}.
\end{aligned}$$

Note that for $0 \leq k < N$ we have $0 \leq (k+2)\delta \leq (N+1)\delta \leq 2T$. So here, the time interval corresponding to $\|\cdot\|_\infty$ can be taken to be $[0, 2T]$.

By assumption, $p\tilde{p} > 2$, so using Hölder's inequality again with exponents $p\tilde{p}/(p\tilde{p}-2)$ and $p\tilde{p}/2$, along with Assumption 1, we get

$$\begin{aligned}
& E \left[\max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \quad (3.36) \\
& \leq \tilde{C}^{1/\tilde{p}} \|\sigma\|_\infty^p \left(\sum_{k=0}^{N-1} (2\delta)^{p\tilde{p}/2-1} \int_{k\delta}^{(k+2)\delta} E[\|(\nabla_z K)^\epsilon(s, x_s^\epsilon)\|^{p\tilde{p}}] dr \right)^{1/\tilde{p}} \\
& \leq \tilde{C}^{1/\tilde{p}} \|\sigma\|_\infty^p \left((2\delta)^{p\tilde{p}/2} N \right)^{1/\tilde{p}} \sup_{s \in [0, (N+1)\delta]} E[(M + K^\epsilon(s, x_s^\epsilon))^{p\tilde{p}}]^{1/\tilde{p}}.
\end{aligned}$$

Using $N < T/\delta$ we obtain

$$\begin{aligned}
& E \left[\max_{k=0, \dots, N-1} \sup_{t \in [k\delta, (k+2)\delta]} \left| \int_{k\delta}^t (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_r \right|^p \right] \quad (3.37) \\
& \leq 2^{p/2} \tilde{C}^{1/\tilde{p}} T^{1/\tilde{p}} \|\sigma\|_\infty^p \delta^{p/2-1/\tilde{p}} \sup_{s \in [0, 2T]} E[(M + K^\epsilon(s, x_s^\epsilon))^{p\tilde{p}}]^{1/\tilde{p}} \\
& = \delta^{p/2-1/\tilde{p}} O(1),
\end{aligned}$$

where we used Proposition 3.1.

Combining these results we see that for $\epsilon > 0$ sufficiently small and any $T \geq \delta > 0$, $\tilde{p} \geq 1$ we have

$$E \left[\sup_{t \in [0, T]} \left| \int_0^t e^{-\alpha(t-s)/\epsilon} (\nabla_z K)^\epsilon(s, x_s^\epsilon) \cdot \sigma(s, x_s^\epsilon) dW_s \right|^p \right] \leq e^{-p\alpha\delta/\epsilon} O(1) + \delta^{p/2-1/\tilde{p}} O(1), \quad (3.38)$$

where the big-O terms do not depend on δ .

Now let $0 < \xi < 1$ and choose $\delta = \epsilon^{1-\xi}$. Then

$$\begin{aligned} E \left[\sup_{t \in [0, T]} K^\epsilon(t, x_t^\epsilon)^p \right] &\leq 2^{p-1} (\kappa/\alpha)^p + \frac{2^{p-1}}{\epsilon^{p/2}} \left(e^{-p\alpha/\epsilon^\xi} O(1) + \epsilon^{(1-\xi)(p/2-1/\tilde{p})} O(1) \right) \\ &= 2^{p-1} \left((\kappa/\alpha)^p + \epsilon^{-p/2} e^{-p\alpha/\epsilon^\xi} O(1) + \epsilon^{(1-\xi)(p/2-1/\tilde{p})-p/2} O(1) \right). \end{aligned} \quad (3.39)$$

For any $\beta > 0$ there exists $\tilde{p} \geq 1$ and $0 < \xi < 1$ such that

$$(1 - \xi)(p/2 - 1/\tilde{p}) - p/2 = -\xi \frac{p}{2} - \frac{1 - \xi}{\tilde{p}} > -\beta. \quad (3.40)$$

Hence the term $\epsilon^{(1-\xi)(p/2-1/\tilde{p})-p/2} O(1)$ diverges more slowly than $\epsilon^{-\beta}$. Also, $\epsilon^{-p/2} e^{-p\alpha/\epsilon^\xi} = o(1)$ for all $\xi > 0$. This proves the result for $p > 2$. The result for all $p > 0$ again follows by an application of Hölder's inequality. $\square$

3.4. Decay of Momentum

Starting in this section, we will assume that the difference between the canonical momentum and ψ is bounded by the kinetic energy in the following sense:

Assumption 3. *We assume that for every $T > 0$ there exists $c > 0$, $\eta > 0$ such that*

$$K(\epsilon, t, q, z) \geq c \|z\|^{2\eta} \quad (3.41)$$

on $(0, \epsilon_0] \times [0, T] \times \mathbb{R}^{2n}$.

With the addition of Assumption 3, the bounds on the kinetic energy from Propositions 3.1 and 3.2 yield the following decay rates for the momentum to the submanifold defined by $p = \psi(t, q)$:

Lemma 3.4. *Under Assumptions 1-3, for any $T > 0$, $p > 0$ we have*

$$\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^{p/2}) \text{ as } \epsilon \rightarrow 0^+ \quad (3.42)$$

and for any $p > 0$, $T > 0$, $0 < \beta < p/2$ we have

$$E\left[\sup_{t \in [0, T]} \|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p\right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (3.43)$$

Our final momentum decay rate result concerns a class of integrals with respect to products of the components of $u_t^\epsilon \equiv p_t^\epsilon - \psi(t, q_t^\epsilon)$.

Proposition 3.3. *Let $f : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}$ be a C^1 function, such that for every $T > 0$, f , $\partial_t f$, and $\nabla_q f$ are bounded on $[0, T] \times \mathbb{R}^n$. Define $u_t^\epsilon = p_t^\epsilon - \psi(t, q_t^\epsilon)$. Under Assumptions 1-3, for any $p > 0$, $T > 0$, $i, j = 1, \dots, n$ we have*

$$E\left[\sup_{t \in [0, T]} \left|\int_0^t f(s, q_s^\epsilon) d((u_s^\epsilon)_i (u_s^\epsilon)_j)\right|^p\right] = O(\epsilon^{p/2}) \text{ as } \epsilon \rightarrow 0^+. \quad (3.44)$$

Proof. $f(s, q_s^\epsilon)$ is a C^1 -semimartingale. Therefore integration by parts gives

$$\begin{aligned} \int_0^t f(s, q_s^\epsilon) d((u_s^\epsilon)_i (u_s^\epsilon)_j) &= f(t, q_t^\epsilon) (u_t^\epsilon)_i (u_t^\epsilon)_j - f(0, q_0^\epsilon) (u_0^\epsilon)_i (u_0^\epsilon)_j \\ &\quad - \int_0^t (u_s^\epsilon)_i (u_s^\epsilon)_j (\partial_s f(s, q_s^\epsilon) + \nabla_q f(s, q_s^\epsilon) \cdot \nabla_p H^\epsilon(s, x_s^\epsilon)) ds. \end{aligned} \quad (3.45)$$

Hence, for $p \geq 1$, using Assumption 3 we obtain:

$$\begin{aligned} &E\left[\sup_{t \in [0, T]} \left|\int_0^t f(s, q_s^\epsilon) d((u_s^\epsilon)_i (u_s^\epsilon)_j)\right|^p\right] \\ &\leq 3^{p-1} \left(E\left[\sup_{t \in [0, T]} |f(t, q_t^\epsilon) (u_t^\epsilon)_i (u_t^\epsilon)_j|^p\right] + E[|f(0, q_0^\epsilon) (u_0^\epsilon)_i (u_0^\epsilon)_j|^p] \right. \\ &\quad \left. + E\left[\sup_{t \in [0, T]} \left|\int_0^t (u_s^\epsilon)_i (u_s^\epsilon)_j (\partial_s f(s, q_s^\epsilon) + \nabla_q f(s, q_s^\epsilon) \cdot \nabla_p K^\epsilon(s, x_s^\epsilon)) ds\right|^p\right] \right) \\ &\leq 3^{p-1} \left(2\|f\|_\infty^p E\left[\sup_{t \in [0, T]} \epsilon^p (K^\epsilon(t, x_t^\epsilon)/c)^{p/\eta}\right] \right. \\ &\quad \left. + E\left[\left(\int_0^T \|u_s^\epsilon\|^2 (\|\partial_s f\|_\infty + \|\nabla_q f\|_\infty \|(\nabla_z K)^\epsilon(s, x_s^\epsilon)/\sqrt{\epsilon}\|) ds\right)^p\right] \right). \end{aligned} \quad (3.46)$$

Now, using Assumption 1 and Proposition 3.2, for any $\beta > 0$ we find

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t f(s, q_s^\epsilon) d((u_s^\epsilon)^i (u_s^\epsilon)^j) \right|^p \right] \\
& \leq O(\epsilon^{p-\beta}) + 3^{p-1} E \left[\left(\int_0^T \epsilon (K^\epsilon(s, x_s^\epsilon)/c)^{1/\eta} \right. \right. \\
& \quad \left. \left. \times (\|\partial_s f\|_\infty + \epsilon^{-1/2} \|\nabla_q f\|_\infty (M + K^\epsilon(s, x_s^\epsilon))) ds \right)^p \right].
\end{aligned} \tag{3.47}$$

Hölder's inequality and Proposition 3.1 allow us to bound the second term:

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left| \int_0^t f(s, q_s^\epsilon) d((u_s^\epsilon)_i (u_s^\epsilon)_j) \right|^p \right] \\
& \leq O(\epsilon^{p-\beta}) + 3^{p-1} c^{-p/\eta} T^{p-1} \epsilon^p E \left[\int_0^T K^\epsilon(s, x_s^\epsilon)^{p/\eta} (\|\partial_s f\|_\infty \right. \\
& \quad \left. + \epsilon^{-1/2} \|\nabla_q f\|_\infty (M + K^\epsilon(s, x_s^\epsilon)))^p ds \right] \\
& \leq O(\epsilon^{p-\beta}) + 3^{2(p-1)} c^{-p/\eta} T^p \epsilon^p \sup_{t \in [0, T]} E [K^\epsilon(s, x_s^\epsilon)^{p/\eta} (\|\partial_s f\|_\infty^p \\
& \quad + \epsilon^{-p/2} \|\nabla_q f\|_\infty^p (M^p + K^\epsilon(s, x_s^\epsilon)^p))] \\
& = O(\epsilon^{p-\beta}) + 3^{2(p-1)} c^{-p/\eta} T^p \epsilon^p O(\epsilon^{-p/2}).
\end{aligned} \tag{3.48}$$

Taking $\beta = p/2$ gives the result when $p \geq 1$. The result for any $p > 0$ follows from Hölder's inequality. $\square$

4. Derivation of the Limiting Equation

In this section, we derive the equation satisfied by q_t^ϵ in the limit $\epsilon \rightarrow 0^+$. The actual convergence proof will be given in the following section. The derivation is an adaptation of the methods used in [19, 21]. We will need the following:

Assumption 4. *We assume that γ is C^1 and is independent of p .*

The starting point for the derivation is a rewriting of Hamilton's equation of motion in terms of the variables $u_t^\epsilon \equiv p_t^\epsilon - \psi(t, q_t^\epsilon)$:

$$d(u_t^\epsilon)_i = -\gamma_{ij}(t, q_t^\epsilon) \partial_{p_j} H^\epsilon(t, x_t^\epsilon) dt + (-\partial_{q^i} H^\epsilon(t, x_t^\epsilon) + F_i(t, x_t^\epsilon)) dt \quad (4.1)$$

$$\begin{aligned} & -\partial_t \psi_i(t, q_t^\epsilon) dt - \partial_{q^k} \psi_i(t, q_t^\epsilon) \partial_{p_k} H^\epsilon(t, x_t^\epsilon) dt + \sigma_{ij}(t, x_t^\epsilon) dW_t^j \\ & = -\tilde{\gamma}_{ik}(t, q_t^\epsilon) \partial_{p_k} K^\epsilon(t, x_t^\epsilon) dt - (\partial_{q^i} K)^\epsilon(t, x_t^\epsilon) dt \\ & + (-\partial_t \psi_i(t, q_t^\epsilon) - \partial_{q^i} V(t, q_t^\epsilon) + F_i(t, x_t^\epsilon)) dt + \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho \end{aligned} \quad (4.2)$$

where

$$\tilde{\gamma}_{ik}(t, q) \equiv \gamma_{ik}(t, q) + \partial_{q^k} \psi_i(t, q) - \partial_{q^i} \psi_k(t, q). \quad (4.3)$$

The second and third terms in $\tilde{\gamma}$ together form an antisymmetric matrix, hence the eigenvalue bound for γ from Assumption 2 implies invertibility of $\tilde{\gamma}$. See Lemma B1. We define the components of $\tilde{\gamma}^{-1}$ such that

$$(\tilde{\gamma}^{-1})^{ij} \tilde{\gamma}_{jk} = \delta_k^i, \quad (4.4)$$

and for any $v \in \mathbb{R}^n$ we define $(\tilde{\gamma}^{-1}v)^i = (\tilde{\gamma}^{-1})^{ij} v_j$.

This lets us solve for $\nabla_p H^\epsilon(t, x_t^\epsilon) dt$ to get

$$\begin{aligned} d(q_t^\epsilon)^i &= \partial_{p_i} H^\epsilon(t, x_t^\epsilon) dt \\ &= (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ &\quad - (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (\partial_{q^j} K)^\epsilon(t, x_t^\epsilon) dt + (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho \\ &\quad - (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) d(u_t^\epsilon)_j. \end{aligned} \quad (4.5)$$

$\tilde{\gamma}^{-1}(t, q_t^\epsilon)$ is pathwise C^1 , so integrating the last term by parts results in

$$\begin{aligned} -(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) d(u_t^\epsilon)_j &= -d((\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (u_t^\epsilon)_j) + (u_t^\epsilon)_j \partial_t (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) dt \\ &+ (u_t^\epsilon)_j \partial_{q^i} (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{p_i} H^\epsilon(t, x_t^\epsilon) dt. \end{aligned} \quad (4.6)$$

Therefore

$$\begin{aligned} d(q_t^\epsilon)^i &= (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ &\quad - (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (\partial_{q^j} K)^\epsilon(t, x_t^\epsilon) dt + (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho \\ &\quad - d((\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) (u_t^\epsilon)_j) + (u_t^\epsilon)_j \partial_t (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) dt \\ &\quad + (u_t^\epsilon)_j \partial_{q^i} (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{p_i} H^\epsilon(t, x_t^\epsilon) dt. \end{aligned} \quad (4.7)$$

In order to homogenize $(u_t^\epsilon)_j \partial_{p_i} H^\epsilon(t, x_t^\epsilon) dt$, we make the additional assumption:

Assumption 5. We assume that K has the form

$$K(\epsilon, t, q, z) = \tilde{K}(\epsilon, t, q, A^{ij}(t, q)z_i z_j) \quad (4.8)$$

where $\tilde{K}(\epsilon, t, q, \zeta)$ is C^2 in (t, q, ζ) for every ϵ , non-negative on $(0, \epsilon_0] \times [0, \infty) \times \mathbb{R}^n \times [0, \infty)$, $A(t, q)$ is a C^2 function whose values are symmetric $n \times n$ -matrices. We also assume that for every $T > 0$, the eigenvalues of A are bounded above and below by some constants $C > 0$ and $c > 0$ respectively, uniformly on $[0, T] \times \mathbb{R}^n$.

We will write $\tilde{K}'$ for $\partial_\zeta \tilde{K}$ and will use the abbreviation $\|z\|_A^2$ for $A^{ij}(t, q)z_i z_j$ when the implied values of t and q are apparent from the context.

With this assumption,

$$\begin{aligned} (\partial_{q^i} K)^\epsilon(t, x_t^\epsilon) &= \partial_{q^i} \tilde{K}(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) \\ &\quad + \tilde{K}'(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) \partial_{q^i} A^{kl}(t, q_t^\epsilon) (u_t^\epsilon)_k (u_t^\epsilon)_l / \epsilon \end{aligned} \quad (4.9)$$

and

$$\partial_{p_i} H^\epsilon(t, x_t^\epsilon) = \frac{2}{\epsilon} A^{lk}(t, q_t^\epsilon) \tilde{K}'(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) (u_t^\epsilon)_k. \quad (4.10)$$

To simplify $(u_t^\epsilon)_j \partial_{p_i} H^\epsilon(t, x_t^\epsilon) dt$, we compute

$$\begin{aligned} d((u_t^\epsilon)_i (u_t^\epsilon)_j) &= (u_t^\epsilon)_i d(u_t^\epsilon)_j + (u_t^\epsilon)_j d(u_t^\epsilon)_i + d[u_i^\epsilon, u_j^\epsilon]_t \\ &= (-(u_t^\epsilon)_i \tilde{\gamma}_{jk}(t, q_t^\epsilon) - (u_t^\epsilon)_j \tilde{\gamma}_{ik}(t, q_t^\epsilon)) \frac{2}{\epsilon} \tilde{K}'(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) A^{kl}(t, q_t^\epsilon) (u_t^\epsilon)_l dt \\ &\quad - (u_t^\epsilon)_i (\partial_{q^j} K)^\epsilon(t, x_t^\epsilon) dt - (u_t^\epsilon)_j (\partial_{q^i} K)^\epsilon(t, x_t^\epsilon) dt \\ &\quad + (u_t^\epsilon)_i (-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ &\quad + (u_t^\epsilon)_j (-\partial_t \psi_i(t, q_t^\epsilon) - \partial_{q^i} V(t, q_t^\epsilon) + F_i(t, x_t^\epsilon)) dt \\ &\quad + (u_t^\epsilon)_i \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho + (u_t^\epsilon)_j \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho + \Sigma_{ij}(t, x_t^\epsilon) dt, \end{aligned} \quad (4.11)$$

where we employed the equation for u_t^ϵ , Eq. (4.2), and Eq. (4.10). We isolate the u -dependent terms that appear in $(u_t^\epsilon)_j \partial_{p_i} H^\epsilon(t, x_t^\epsilon) dt$ to find

$$\begin{aligned} &\frac{2}{\epsilon} \tilde{K}'(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) (\tilde{\gamma}_{jk}(t, q_t^\epsilon) A^{kl}(t, q_t^\epsilon) (u_t^\epsilon)_l (u_t^\epsilon)_i \\ &\quad + \tilde{\gamma}_{ik}(t, q_t^\epsilon) A^{kl}(t, q_t^\epsilon) (u_t^\epsilon)_l (u_t^\epsilon)_j) dt \\ &= -d((u_t^\epsilon)_i (u_t^\epsilon)_j) - (u_t^\epsilon)_i (\partial_{q^j} K)^\epsilon(t, x_t^\epsilon) dt - (u_t^\epsilon)_j (\partial_{q^i} K)^\epsilon(t, x_t^\epsilon) dt \\ &\quad + (u_t^\epsilon)_i (-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ &\quad + (u_t^\epsilon)_j (-\partial_t \psi_i(t, q_t^\epsilon) - \partial_{q^i} V(t, q_t^\epsilon) + F_i(t, x_t^\epsilon)) dt \\ &\quad + (u_t^\epsilon)_i \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho + (u_t^\epsilon)_j \sigma_{i\rho}(t, x_t^\epsilon) dW_t^\rho + \Sigma_{ij}(t, x_t^\epsilon) dt. \end{aligned} \quad (4.12)$$

We will solve this equation for $\tilde{K}'(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon)(u_t^\epsilon)_j(u_t^\epsilon)_k dt$ using a Lyapunov equation technique, as in [19, 21].

The formula for the left-hand side of Eq. (4.12) clearly represents a differential of a C^1 -function. Therefore the integral from 0 to t of the right-hand side, which we denote by $(C_t)_{ij}$, is a C^1 -function P -a.s. Differentiating both sides with respect to t , we obtain

$$\begin{aligned} & \frac{2}{\epsilon} \tilde{K}'(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon)(A\tilde{\gamma})_j^l(t, q_t^\epsilon)(u_t^\epsilon)_l(u_t^\epsilon)_i \\ & + \frac{2}{\epsilon} \tilde{K}'(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon)(A\tilde{\gamma})_i^l(t, q_t^\epsilon)(u_t^\epsilon)_l(u_t^\epsilon)_j = (\dot{C}_t)_{ij}, \end{aligned} \quad (4.13)$$

where we define $(A\tilde{\gamma})_j^i = \tilde{\gamma}_{jk}A^{ki}$.

Defining the matrix

$$(V_t)_{ij} = \frac{2}{\epsilon} \tilde{K}'(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon)(u_t^\epsilon)_i(u_t^\epsilon)_j \quad (4.14)$$

we rewrite Eq. (4.13) as

$$(A\tilde{\gamma})_i^l V_{lj} + V_{il} (A\tilde{\gamma})_j^l = \dot{C}_{ij}. \quad (4.15)$$

This is a Lyapunov equation for V .

For every $T > 0$, there exists $c > 0$ and $\lambda > 0$ such that $-A\tilde{\gamma}$ has eigenvalues with real parts bounded above by $-c\lambda$, uniformly on $[0, T] \times \mathbb{R}^n$. See Lemma B2. Hence, we can solve uniquely for V ,

$$V_{ij} = \int_0^\infty (e^{-yA\tilde{\gamma}})_i^k \dot{C}_{kl} (e^{-yA\tilde{\gamma}})_j^l dy. \quad (4.16)$$

See, for example, Theorem 6.4.2 in [26].

Remark 2. *Assumption 5, or something else that accomplishes a similar purpose, is necessary in the above computation. If one tries to solve for $(\tilde{V}_t)_j^i \equiv (u_t^\epsilon)_j \partial_{p_i} H^\epsilon(t, x_t^\epsilon)$ directly, then one is led to the linear equation*

$$\tilde{\gamma}_{jk} \tilde{V}_i^k + \tilde{\gamma}_{ik} \tilde{V}_j^k = \dot{C}_{ij}. \quad (4.17)$$

The left-hand side of this equation has a non-trivial kernel, consisting of all $\tilde{V}$ for which $\tilde{\gamma}_{ik} \tilde{V}_j^k$ is antisymmetric. Therefore, just knowing that $\tilde{V}$ satisfies Eq. (4.17) does not allow us to uniquely solve for $\tilde{V}$. Some additional constraint must be combined with Eq. (4.17) in order to solve for $\tilde{V}$.

Integrating Eq. (4.16) with respect to time, we obtain

$$\begin{aligned} & \frac{2}{\epsilon} \int_0^t \tilde{K}'(\epsilon, s, q_s^\epsilon, \|u_s^\epsilon\|_A^2/\epsilon)(u_s^\epsilon)_i(u_s^\epsilon)_j ds \\ &= \int_0^t \int_0^\infty \left(e^{-y(A\tilde{\gamma})(s, q_s^\epsilon)}\right)_i^k \left(e^{-y(A\tilde{\gamma})(s, q_s^\epsilon)}\right)_j^l dy (\dot{C}_s)_{kl} ds. \end{aligned} \quad (4.18)$$

The functions

$$G_{ij}^{kl}(t, q) = \int_0^\infty \left(e^{-y(A\tilde{\gamma})(t, q)}\right)_i^k \left(e^{-y(A\tilde{\gamma})(t, q)}\right)_j^l dy \quad (4.19)$$

are C^1 , hence $G_{ij}^{kl}(t, q_t^\epsilon)$ are semimartingales and

$$\begin{aligned} & \frac{2}{\epsilon} \tilde{K}'(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon)(u_t^\epsilon)_i(u_t^\epsilon)_j dt = G_{ij}^{ab}(t, q_t^\epsilon) d(C_t)_{ab} \\ &= G_{ij}^{ab}(t, q_t^\epsilon) \Sigma_{ab}(t, x_t^\epsilon) dt - G_{ij}^{ab}(t, q_t^\epsilon) d((u_t^\epsilon)_a(u_t^\epsilon)_b) \\ & \quad - G_{ij}^{ab}(t, q_t^\epsilon)(u_t^\epsilon)_a(\partial_{q^b} K)^\epsilon(t, x_t^\epsilon) dt - G_{ij}^{ab}(t, q_t^\epsilon)(u_t^\epsilon)_b(\partial_{q^a} K)^\epsilon(t, x_t^\epsilon) dt \\ & \quad + G_{ij}^{ab}(t, q_t^\epsilon)(u_t^\epsilon)_a(-\partial_t \psi_b(t, q_t^\epsilon) - \partial_{q^b} V(t, q_t^\epsilon) + F_b(t, x_t^\epsilon)) dt \\ & \quad + G_{ij}^{ab}(t, q_t^\epsilon)(u_t^\epsilon)_b(-\partial_t \psi_a(t, q_t^\epsilon) - \partial_{q^a} V(t, q_t^\epsilon) + F_a(t, x_t^\epsilon)) dt \\ & \quad + G_{ij}^{ab}(t, q_t^\epsilon)(u_t^\epsilon)_a \sigma_{b\rho}(t, x_t^\epsilon) dW_t^\rho + G_{ij}^{ab}(t, q_t^\epsilon)(u_t^\epsilon)_b \sigma_{a\rho}(t, x_t^\epsilon) dW_t^\rho. \end{aligned} \quad (4.20)$$

Combining Eq. (4.7) with Eq. (4.9), Eq. (4.10), and Eq. (4.20) we see that q_t^ϵ satisfies the equation

$$\begin{aligned} d(q_t^\epsilon)^i &= (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon)(-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ & \quad + (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho - (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{q^j} \tilde{K}(\epsilon, t, q_t^\epsilon, \|u_t^\epsilon\|_A^2/\epsilon) dt \\ & \quad + Q^{ikl}(t, q_t^\epsilon) J_{kl}(t, x_t^\epsilon) dt + d(R_t^\epsilon)^i, \end{aligned} \quad (4.21)$$

where

$$J_{ij}(t, x) \equiv G_{ij}^{kl}(t, q) \Sigma_{kl}(t, x), \quad (4.22)$$

$$Q^{ijl}(t, q) \equiv \partial_{q^k} (\tilde{\gamma}^{-1})^{ij}(t, q) A^{kl}(t, q) - \frac{1}{2} (\tilde{\gamma}^{-1})^{ik}(t, q) \partial_{q^k} A^{jl}(t, q), \quad (4.23)$$

and

$$\begin{aligned}
d(R_t^\epsilon)^i &\equiv -d((\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon)(u_t^\epsilon)_j) + (u_t^\epsilon)_j \partial_t (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) dt \\
&\quad - Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) d((u_t^\epsilon)_a (u_t^\epsilon)_b) \\
&\quad + Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_a (-\partial_t \psi_b(t, q_t^\epsilon) - \partial_{q^b} V(t, q_t^\epsilon) \\
&\quad \quad - (\partial_{q^b} K)^\epsilon(t, x_t^\epsilon) + F_b(t, x_t^\epsilon)) dt \\
&\quad + Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_b (-\partial_t \psi_a(t, q_t^\epsilon) - \partial_{q^a} V(t, q_t^\epsilon) \\
&\quad \quad - (\partial_{q^a} K)^\epsilon(t, x_t^\epsilon) + F_a(t, x_t^\epsilon)) dt \\
&\quad + Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_a \sigma_{b\rho}(t, x_t^\epsilon) dW_t^\rho \\
&\quad + Q^{ikl}(t, q_t^\epsilon) G_{kl}^{ab}(t, q_t^\epsilon) (u_t^\epsilon)_b \sigma_{a\rho}(t, x_t^\epsilon) dW_t^\rho.
\end{aligned} \tag{4.24}$$

Based on our knowledge of the decay rate of u_t^ϵ , we expect R_t^ϵ to go to zero in the limit $\epsilon \rightarrow 0^+$. In general, one would still need to extract the portion of $(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{q^j} \tilde{K}(\epsilon, t, q_t^\epsilon, \|z\|_A^2/\epsilon) dt$ that survives in the limit. We will address this question in a future work [27], but in this paper we will assume:

Assumption 6. $\tilde{K} = \tilde{K}(\epsilon, t, z)$ i.e. $\tilde{K}$ is independent of q , and hence

$$(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{q^j} \tilde{K}(\epsilon, t, q_t^\epsilon, \|z\|_A^2/\epsilon) dt = 0. \tag{4.25}$$

Along with Lemma 3.4, the above calculations motivate the proposed limiting equation

$$\begin{aligned}
dq_t^i &= (\tilde{\gamma}^{-1})^{ij}(t, q_t) (-\partial_t \psi_j(t, q_t) - \partial_{q^j} V(t, q_t) + F_j(t, q_t, \psi(t, q_t))) dt \\
&\quad + Q^{ikl}(t, q_t) J_{kl}(t, q_t, \psi(t, q_t)) dt + (\tilde{\gamma}^{-1})^{ij}(t, q_t) \sigma_{j\rho}(t, q_t, \psi(t, q_t)) dW_t^\rho.
\end{aligned} \tag{4.26}$$

Note that an additional *noise induced drift* term,

$$S^i(t, q) \equiv Q^{ijl}(t, q) J_{jl}(t, q, \psi(t, q)), \tag{4.27}$$

arises in the limit when Σ is nonzero and (generally) when $\tilde{\gamma}$ and/or A have nontrivial q -dependence. This is in addition to the forcing term, $-\partial_t \psi - \nabla_q V + F$, and is another manifestation of the phenomenon derived in [19, 21].

Remark 3. *Assumption 6 determines the splitting of the Hamiltonian into $K^\epsilon(t, x)$ and $V(t, q)$, up to a function of time i.e. if $H^\epsilon = K_1^\epsilon + V_1 = K_2^\epsilon + V_2$ are two splittings then $V_1(t, q) = V_2(t, q) + c(t)$. This ambiguity does not impact the limiting equation Eq. (4.26), and so the limiting equation is uniquely defined by the original SDE, Eq. (1.10)-Eq. (1.11), as it has to be, of course.*

5. Convergence Proof

In this final section, we prove convergence of q_t^ϵ to the solution of the proposed limiting equation, Eq. (4.26). This will be accomplished by using the following lemma.

Lemma 5.1. *Let $T > 0$ and suppose we have continuous functions $\tilde{F}(t, x) : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\tilde{\sigma}(t, x) : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times k}$, and $\psi : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ that are Lipschitz in x , uniformly in $t \in [0, T]$.*

Let W_t be a k -dimensional Wiener process, $p \geq 2$ and $\delta > 0$ and suppose that we have continuous semimartingales q_t and, for each $0 < \epsilon \leq \epsilon_0$, $\tilde{R}_t^\epsilon$, $x_t^\epsilon = (q_t^\epsilon, p_t^\epsilon)$ that satisfy the following properties:

1. $q_t^\epsilon = q_0^\epsilon + \int_0^t \tilde{F}(s, x_s^\epsilon) ds + \int_0^t \tilde{\sigma}(s, x_s^\epsilon) dW_s + \tilde{R}_t^\epsilon$
2. $q_t = q_0 + \int_0^t \tilde{F}(s, q_s, \psi(s, q_s)) ds + \int_0^t \tilde{\sigma}(s, q_s^\epsilon, \psi(s, q_s)) dW_s$
3. $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^\delta)$ as $\epsilon \rightarrow 0^+$.
4. $E[\sup_{t \in [0, T]} \|\tilde{R}_t^\epsilon\|^p] = O(\epsilon^\delta)$ as $\epsilon \rightarrow 0^+$.
5. $\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^\delta)$ as $\epsilon \rightarrow 0^+$.
6. $E[\sup_{t \in [0, T]} \|q_t^\epsilon\|^p] < \infty$ for all $\epsilon > 0$ sufficiently small.
7. $E[\sup_{t \in [0, T]} \|q_t\|^p] < \infty$

Then

$$E \left[\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p \right] = O(\epsilon^\delta) \text{ as } \epsilon \rightarrow 0^+. \quad (5.1)$$

If we replace properties 4, 6, 7 with

- 4* $\sup_{t \in [0, T]} E[\|\tilde{R}_t^\epsilon\|^p] = O(\epsilon^\delta)$ as $\epsilon \rightarrow 0^+$,
- 6* $\sup_{t \in [0, T]} E[\|q_t^\epsilon\|^p] < \infty$ for all $\epsilon > 0$ sufficiently small,
- 7* $\sup_{t \in [0, T]} E[\|q_t\|^p] < \infty$,

then we instead arrive at

$$\sup_{t \in [0, T]} E[\|q_t^\epsilon - q_t\|^p] = O(\epsilon^\delta) \text{ as } \epsilon \rightarrow 0^+. \quad (5.2)$$

Proof. The equations for q_t^ϵ and q_t imply

$$\begin{aligned} (q_t^\epsilon)^i - (q_t)^i &= (q_0^\epsilon)^i - (q_0)^i + \int_0^t \tilde{F}^i(s, x_s^\epsilon) - \tilde{F}^i(s, q_s, \psi(s, q_s)) ds \\ &+ \int_0^t [\tilde{\sigma}_\rho^i(s, x_s^\epsilon) - \tilde{\sigma}_\rho^i(s, q_s, \psi(s, q_s))] dW_s^\rho + (\tilde{R}_t^\epsilon)^i. \end{aligned} \quad (5.3)$$

Using properties 3 and 4, along with the Hölder and Burkholder-Davis-Gundy inequalities, for $0 \leq t \leq T$ we have

$$\begin{aligned} &E \left[\sup_{s \in [0, t]} \|q_s^\epsilon - q_s\|^p \right] \\ &\leq 4^{p-1} \left(E[\|q_0^\epsilon - q_0\|^p] + E \left[\left(\int_0^t \|\tilde{F}(s, x_s^\epsilon) - \tilde{F}(s, q_s, \psi(s, q_s))\| ds \right)^p \right] \right. \\ &\quad \left. + E \left[\sup_{s \in [0, t]} \left\| \int_0^s \tilde{\sigma}_\rho^i(r, x_r^\epsilon) - \tilde{\sigma}_\rho^i(r, q_r, \psi(r, q_r)) dW_r^\rho \right\|^p \right] + E \left[\sup_{s \in [0, t]} \|\tilde{R}_s^\epsilon\|^p \right] \right) \\ &\leq 4^{p-1} \left(T^{p-1} E \left[\int_0^t \|\tilde{F}(s, x_s^\epsilon) - \tilde{F}(s, q_s, \psi(s, q_s))\|^p ds \right] \right. \\ &\quad \left. + \tilde{C} E \left[\left(\int_0^t \|\tilde{\sigma}(r, x_r^\epsilon) - \tilde{\sigma}(r, q_r, \psi(r, q_r))\|_F^2 dr \right)^{p/2} \right] \right) + O(\epsilon^\delta) \\ &\leq 4^{p-1} \left(T^{p-1} \int_0^t E[\|\tilde{F}(s, x_s^\epsilon) - \tilde{F}(s, q_s, \psi(s, q_s))\|^p] ds \right. \\ &\quad \left. + \tilde{C} T^{p/2-1} \int_0^t E[\|\tilde{\sigma}(r, x_r^\epsilon) - \tilde{\sigma}(r, q_r, \psi(r, q_r))\|_F^p] dr \right) + O(\epsilon^\delta). \end{aligned} \quad (5.4)$$

By assumption, $\tilde{\sigma}$, $\tilde{F}$, and ψ are Lipschitz in x , uniformly on $[0, T]$. Hence, using property 5,

$$\begin{aligned} &E \left[\sup_{s \in [0, t]} \|q_s^\epsilon - q_s\|^p \right] \leq \tilde{C} \int_0^t E[\|q_s^\epsilon - q_s\|^p + \|p_s^\epsilon - \psi(s, q_s)\|^p] ds + O(\epsilon^\delta) \\ &\leq \tilde{C} \int_0^t E[\|q_s^\epsilon - q_s\|^p + \|p_s^\epsilon - \psi(s, q_s^\epsilon)\|^p + \|\psi(s, q_s^\epsilon) - \psi(s, q_s)\|^p] ds + O(\epsilon^\delta) \\ &\leq \tilde{C} \int_0^t E[\|q_s^\epsilon - q_s\|^p] ds + \tilde{C} \sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] + O(\epsilon^\delta) \\ &= \tilde{C} \int_0^t E \left[\sup_{r \in [0, s]} \|q_r^\epsilon - q_r\|^p \right] ds + O(\epsilon^\delta) \end{aligned} \quad (5.5)$$

for all $0 \leq t \leq T$, where the constants change from line to line and are all independent of t .

Properties 6 and 7 imply that $E[\sup_{s \in [0, t]} \|q_s^\epsilon - q_s\|^p] \in L^1([0, T])$ for ϵ sufficiently small, and hence Gronwall's inequality applied to Eq. (5.5) gives

$$E \left[\sup_{s \in [0, t]} \|q_s^\epsilon - q_s\|^p \right] \leq O(\epsilon^\delta) e^{\tilde{C}t} \quad (5.6)$$

for $0 \leq t \leq T$.

The proof of Eq. (5.2) under assumptions 1-3, 4*, 5, 6*, and 7* is almost identical. $\square$

To prove that the hypotheses of Lemma 5.1 hold, we will need the following assumption:

Assumption 7. *We assume that, for every $T > 0$, $\nabla_q V$, F , and σ are Lipschitz in x uniformly in $t \in [0, T]$. We also assume that A and γ are C^2 , ψ is C^3 , and $\partial_t \psi$, $\partial_{q^i} \psi$, $\partial_{q^i} \partial_{q^j} \psi$, $\partial_t \partial_{q^i} \psi$, $\partial_t \partial_{q^j} \partial_{q^i} \psi$, $\partial_{q^i} \partial_{q^j} \partial_{q^i} \psi$, $\partial_t \gamma$, $\partial_{q^i} \gamma$, $\partial_t \partial_{q^j} \gamma$, $\partial_{q^i} \partial_{q^j} \gamma$, $\partial_t A$, $\partial_{q^i} A$, $\partial_t \partial_{q^i} A$, and $\partial_{q^i} \partial_{q^j} A$ are bounded on $[0, T] \times \mathbb{R}^n$ for every $T > 0$.*

Note that, combined with our prior assumptions, this implies $\tilde{\gamma}$, $\tilde{\gamma}^{-1}$, $\partial_t \tilde{\gamma}^{-1}$, $\partial_{q^i} \tilde{\gamma}^{-1}$, $\partial_t \partial_{q^j} \tilde{\gamma}^{-1}$, and $\partial_{q^i} \partial_{q^j} \tilde{\gamma}^{-1}$ are bounded on compact t intervals. Additionally, using the formula for the derivative of the matrix exponential found in [28], one can prove that our assumptions also imply that the G_{kl}^{ij} 's are bounded and Lipschitz in q , uniformly on compact t intervals.

As a step towards using Lemma 5.1 to prove our convergence result, we now show that R_t^ϵ from Eq. (4.24) converges to zero in the appropriate sense.

Lemma 5.2. *Under Assumptions 1-5, 7, for any $p > 0$, $T > 0$, $0 < \beta < p/2$ we have*

$$E \left[\sup_{t \in [0, T]} \|R_t^\epsilon\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+ \quad (5.7)$$

and

$$\sup_{t \in [0, T]} E [\|R_t^\epsilon\|^p] = O(\epsilon^{p/2}) \text{ as } \epsilon \rightarrow 0^+, \quad (5.8)$$

where R_t^ϵ was defined in Eq. (4.24).

Proof. Let us first assume that $p > 2$. Let $0 < \beta < p/2$. Define

$$Y(t, x) = -\partial_t \psi(t, q) - \nabla_q V(t, q) + F(t, x). \quad (5.9)$$

Our assumptions imply that Y is bounded on $[0, T] \times \mathbb{R}^{2n}$.

From Eq. (4.24),

$$\begin{aligned} E \left[\sup_{t \in [0, T]} \|R_t^\epsilon\|^p \right] &\leq 8^{p-1} \left(E \left[\sup_{t \in [0, T]} \|(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon)(u_t^\epsilon)_j\|^p \right] \right. \\ &+ E[\|(\tilde{\gamma}^{-1})^{ij}(0, q_0^\epsilon)(u_0^\epsilon)_j\|^p] + E \left[\left(\int_0^T \|(u_s^\epsilon)_j \partial_s (\tilde{\gamma}^{-1})^{ij}(s, q_s^\epsilon)\| ds \right)^p \right] \\ &+ E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) d((u_s^\epsilon)_a (u_s^\epsilon)_b) \right\|^p \right] \\ &+ E \left[\left(\int_0^T \|Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a (-\partial_{q^b} K)^\epsilon(s, x_s^\epsilon) + Y_b(s, x_s^\epsilon)\| ds \right)^p \right] \\ &+ E \left[\left(\int_0^T \|Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_b (-\partial_{q^a} K)^\epsilon(s, x_s^\epsilon) + Y_a(s, x_s^\epsilon)\| ds \right)^p \right] \\ &+ E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a \sigma_{b\rho}(s, x_s^\epsilon) dW_s^\rho \right\|^p \right] \\ &+ E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_b \sigma_{a\rho}(s, x_s^\epsilon) dW_s^\rho \right\|^p \right] \Bigg), \end{aligned} \quad (5.10)$$

where the norm is the 2-norm of vectors, with components indexed by i , resulting from summation over other, repeated indices. We now show that all of these terms are $O(\epsilon^\beta)$.

Boundedness of $\tilde{\gamma}^{-1}$ together with Lemma 3.4 implies that the first two terms satisfy

$$\begin{aligned} &E \left[\sup_{t \in [0, T]} \|(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon)(u_t^\epsilon)_j\|^p \right] + E[\|(\tilde{\gamma}^{-1})^{ij}(0, q_0^\epsilon)(u_0^\epsilon)_j\|^p] \\ &\leq 2\|\tilde{\gamma}^{-1}\|_\infty^p E \left[\sup_{t \in [0, T]} \|u_t^\epsilon\|^p \right] = O(\epsilon^\beta). \end{aligned} \quad (5.11)$$

By Hölder's inequality, boundedness of $\partial_t \tilde{\gamma}^{-1}$, and Lemma 3.4, the third

term satisfies

$$\begin{aligned}
& E \left[\left(\int_0^T \|(u_s^\epsilon)_j \partial_t(\tilde{\gamma}^{-1})^{ij}(s, q_s^\epsilon)\| ds \right)^p \right] \\
& \leq T^{p-1} \|\partial_t \tilde{\gamma}^{-1}\|_\infty^p \int_0^T E[\|u_s^\epsilon\|^p] ds \\
& \leq T^p \|\partial_t \tilde{\gamma}^{-1}\|_\infty^p \sup_{s \in [0, T]} E[\|u_s^\epsilon\|^p] = O(\epsilon^{p/2}).
\end{aligned} \tag{5.12}$$

The functions $Q^{ijl}(t, q)G_{jl}^{ab}(t, q)$ are C^1 , bounded, with bounded first derivatives on $[0, T] \times \mathbb{R}^n$. Therefore, by Proposition 3.3 we have

$$E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) d((u_s^\epsilon)_a (u_s^\epsilon)_b) \right\|^p \right] = O(\epsilon^{p/2}). \tag{5.13}$$

Using Hölder's inequality, Eq. (3.1), and our various boundedness assumptions, the fifth term can be bounded as follows:

$$\begin{aligned}
& E \left[\left(\int_0^T \|Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a (-\partial_{q^b} K)^\epsilon(s, x_s^\epsilon) + Y_b(s, x_s^\epsilon)\| ds \right)^p \right] \\
& \leq \tilde{C} T^{p-1} E \left[\int_0^T \|u_s^\epsilon\|^p (\|(\partial_{q^b} K)^\epsilon(s, x_s^\epsilon)\| + \|Y\|_\infty)^p ds \right] \\
& \leq \tilde{C} T^p \sup_{s \in [0, T]} E[\|u_s^\epsilon\|^p (M + C K^\epsilon(s, x_s^\epsilon) + \|Y\|_\infty)^p].
\end{aligned} \tag{5.14}$$

Again, here and in the following, we let $\tilde{C}$ denote a constant that may vary from line to line. We now use the Cauchy-Schwarz inequality, Lemma 3.4, and Proposition 3.1 to obtain

$$\begin{aligned}
& E \left[\left(\int_0^T \|Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a (-\partial_{q^b} K)^\epsilon(s, x_s^\epsilon) + Y_b(s, x_s^\epsilon)\| ds \right)^p \right] \\
& \leq \tilde{C} T^p \sup_{s \in [0, T]} E[\|u_s^\epsilon\|^{2p}]^{1/2} \sup_{s \in [0, T]} E[(M + C K^\epsilon(s, x_s^\epsilon) + \|Y\|_\infty)^{2p}]^{1/2} \\
& = O(\epsilon^p)^{1/2} O(1) = O(\epsilon^{p/2}).
\end{aligned} \tag{5.15}$$

A similar argument shows that the sixth term is also $O(\epsilon^{p/2})$.

Using the Burkholder-Davis-Gundy and Hölder inequalities together with

the boundedness assumptions and Lemma 3.4, the seventh term satisfies

$$\begin{aligned}
& E \left[\sup_{t \in [0, T]} \left\| \int_0^t Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a \sigma_{b\rho}(s, x_s^\epsilon) dW_s^\rho \right\|^p \right] \quad (5.16) \\
& \leq \tilde{C} E \left[\left(\int_0^T \sum_{i, \rho} (Q^{ijl}(s, q_s^\epsilon) G_{jl}^{ab}(s, q_s^\epsilon) (u_s^\epsilon)_a \sigma_{b\rho}(s, x_s^\epsilon))^2 ds \right)^{p/2} \right] \\
& \leq \tilde{C} E \left[\left(\int_0^T \|u_s^\epsilon\|^2 ds \right)^{p/2} \right] \leq \tilde{C} T^{p/2-1} E \left[\int_0^T \|u_s^\epsilon\|^p ds \right] \\
& \leq \tilde{C} \sup_{s \in [0, T]} E[\|u_s^\epsilon\|^p] = O(\epsilon^{p/2}),
\end{aligned}$$

where the power of T can be absorbed into the constant, since we are working on a fixed time interval. A similar estimate applies to the final, eighth term. Therefore, we have proven the claim for $p > 2$. An application of Hölder's inequality proves it for all $p > 0$. The proof of Eq. (5.8) is nearly identical. $\square$

We now have all the ingredients to prove convergence of q_t^ϵ to q_t .

Theorem 5.1. *Let x_t^ϵ be a family of solutions to the SDE 1.10-1.11 with initial condition $(q_0^\epsilon, p_0^\epsilon)$ and q_t be a solution to the proposed limiting SDE, Eq. (4.26), with initial condition q_0 . Suppose that for all $\epsilon > 0$ and all $p > 0$ we have $E[\|q_0^\epsilon\|^p] < \infty$, $E[\|q_0\|^p] < \infty$, and $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2})$. Also suppose that Assumptions 1-7 hold. Then for any $T > 0$, $p > 0$, $0 < \beta < p/2$ we have*

$$E \left[\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+ \quad (5.17)$$

and

$$\sup_{t \in [0, T]} E[\|q_t^\epsilon - q_t\|^p] = O(\epsilon^{p/2}) \text{ as } \epsilon \rightarrow 0^+. \quad (5.18)$$

Proof. First let $p > 2$. Define the generalized force vector field

$$\tilde{F}(t, x)^i = (\tilde{\gamma}^{-1})^{ij}(t, q) (-\partial_t \psi_i(t, q) - \partial_{q^j} V(t, q) + F_j(t, x)) + Q^{ijl}(t, q) J_{jl}(t, x) \quad (5.19)$$

and noise coefficients

$$\tilde{\sigma}_\eta^i(t, x) = (\tilde{\gamma}^{-1})^{ij}(t, q) \sigma_{j\eta}(t, x). \quad (5.20)$$

Our assumptions imply that these are bounded on $[0, T] \times \mathbb{R}^{2n}$ and, along with ψ , they are Lipschitz in x , uniformly in $t \in [0, T]$.

We will now check all of the properties that are required to use Lemma 5.1. By Eq. (4.21) and Eq. (4.26), q_t^ϵ and q_t satisfy the equations

$$(q_t^\epsilon)^i = (q_0^\epsilon)^i + \int_0^t \tilde{F}^i(s, x_s^\epsilon) ds + \int_0^t \tilde{\sigma}_\eta^i(s, x_s^\epsilon) dW_s^\eta + (R_t^\epsilon)^i \quad (5.21)$$

and

$$(q_t)^i = (q_0)^i + \int_0^t \tilde{F}^i(s, q_s, \psi(s, q_s)) ds + \int_0^t \tilde{\sigma}_\eta^i(s, q_s, \psi(s, q_s)) dW_s^\eta. \quad (5.22)$$

Note, that the term involving $\partial_{q^j} \tilde{K}$, present in Eq. (4.21), vanishes under Assumption 6. In addition, for each $T > 0$, $\tilde{F}(t, q, \psi(t, q))$ and $\tilde{\sigma}(t, q, \psi(t, q))$ are bounded and Lipschitz in q , uniformly in t on $[0, T] \times \mathbb{R}^n$, so a unique solution to Eq. (5.22) exists and is defined for all $t \geq 0$ [29].

By assumption, the initial conditions satisfy $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2}) = O(\epsilon^\beta)$. Lemma 5.2 implies that

$$E \left[\sup_{t \in [0, T]} \|R_t^\epsilon\|^p \right] = O(\epsilon^\beta), \quad \sup_{t \in [0, T]} E[\|R_t^\epsilon\|^p] = O(\epsilon^{p/2}) \quad (5.23)$$

as $\epsilon \rightarrow 0^+$ and, by Lemma 3.4,

$$\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^{p/2}) = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (5.24)$$

For any $\epsilon > 0$, $p > 1$ we have

$$\begin{aligned} & \sup_{t \in [0, T]} E[\|q_t^\epsilon\|^p] \leq E \left[\sup_{t \in [0, T]} \|q_t^\epsilon\|^p \right] \\ &= E \left[\sup_{t \in [0, T]} \left\| q_0^\epsilon + \int_0^t \nabla_p H^\epsilon(r, x_r^\epsilon) dr \right\|^p \right] \\ &\leq 2^{p-1} E[\|q_0^\epsilon\|^p] + 2^{p-1} E \left[\left(\int_0^T \|\nabla_p K^\epsilon(r, x_r^\epsilon)\| dr \right)^p \right] \\ &= 2^{p-1} E[\|q_0^\epsilon\|^p] + 2^{p-1} E \left[\left(\int_0^T \epsilon^{-1/2} \|(\nabla_z K)^\epsilon(r, x_r^\epsilon)\| dr \right)^p \right] \\ &\leq 2^{p-1} E[\|q_0^\epsilon\|^p] + 2^{p-1} T^{p-1} \epsilon^{-p/2} E \left[\int_0^T (M + K^\epsilon(r, x_r^\epsilon))^p dr \right] \\ &\leq 2^{p-1} E[\|q_0^\epsilon\|^p] + 4^{p-1} T^p \epsilon^{-p/2} \left(M^p + \sup_{r \in [0, T]} E[K^\epsilon(r, x_r^\epsilon)^p] \right) < \infty, \end{aligned} \quad (5.25)$$

where we used Assumption 1 to bound $\nabla_z K$ and Proposition 3.1 in the last line.

We also have

$$\begin{aligned} \sup_{t \in [0, T]} E[\|q_t\|^p] &\leq E\left[\sup_{t \in [0, T]} \|q_t\|^p\right] \\ &\leq 3^{p-1} \left(E[\|q_0\|^p] + E\left[\left(\int_0^T \|\tilde{F}(s, q_s, \psi(s, q_s))\| ds\right)^p\right] \right. \\ &\quad \left. + E\left[\sup_{t \in [0, T]} \left\|\int_0^t \tilde{\sigma}(s, q_s, \psi(s, q_s)) dW_s\right\|^p\right] \right). \end{aligned} \quad (5.26)$$

$\tilde{F}$ and $\tilde{\sigma}$ are bounded uniformly up to time T , so applying the Burkholder-Davis-Gundy and Hölder inequalities to the last term, we get

$$\sup_{t \in [0, T]} E[\|q_t\|^p] \leq E\left[\sup_{t \in [0, T]} \|q_t\|^p\right] \leq 3^{p-1} (E[\|q_0\|^p] + T^p \|\tilde{F}\|_\infty^p + \tilde{C} T^{p/2} \|\tilde{\sigma}\|_{F, \infty}^p) < \infty. \quad (5.27)$$

Hence we have verified properties 1-7 from Lemma 5.1 for every $T > 0$, $p > 2$, and $\delta = \beta \in (0, p/2)$, and properties 1-3, 4*, 5, 6*, 7* for every $T > 0$, $p > 2$, with $\delta = p/2$. We can therefore conclude the convergence results Eq. (5.17) and Eq. (5.18). The results for any $p > 0$ follows from an application of Hölder's inequality. $\square$

6. Extension to Unbounded Forces

In this section, we focus on relaxing some of the boundedness assumptions in Theorem 5.1 by adapting the method developed in [20]. We make no claim that the assumptions here are as weak or general as possible and there are many variations on this idea that one could pursue, weakening the boundedness assumptions on the various objects appearing in the SDEs; here we focus on accommodating unbounded forces, F and $\nabla_q V$. Specifically, we will be able to prove convergence for potentials that are confining, or at least not too unstable, meaning that there exists $a \geq 0, b \geq 0$ such that $a + b\|q\|^2 + V(t, q)$ is non-negative.

Theorem 6.1. *Suppose that the following hold:*

1. The family of Hamiltonians has the form

$$H^\epsilon(t, x) = K(\epsilon, t, q, (p - \psi(t, q))/\sqrt{\epsilon}) + V(t, q) \quad (6.1)$$

where V is a C^2 , $\mathbb{R}$ -valued function, ψ is a C^3 , $\mathbb{R}^n$ -valued function, and

$$K(\epsilon, t, q, z) = \tilde{K}(\epsilon, t, A^{ij}(t, q)z_iz_j) \quad (6.2)$$

for a non-negative function $\tilde{K}(\epsilon, t, \zeta)$ that is C^2 in $(t, \zeta) \in [0, \infty) \times [0, \infty)$ and a C^2 , positive definite $n \times n$ matrix-valued function, A .

2. For every $T > 0$, the following bounds hold on $(0, \epsilon_0] \times [0, \infty) \times \mathbb{R}^{2n}$:

(a) There exist $C > 0$ and $M > 0$ such that

$$\max\{|\partial_t K(\epsilon, t, q, z)|, \|\nabla_q K(\epsilon, t, q, z)\|\} \leq M + CK(\epsilon, t, q, z). \quad (6.3)$$

(b) There exist $c > 0$ and $M \geq 0$ such that

$$\|\nabla_z K(\epsilon, t, q, z)\|^2 + M \geq cK(\epsilon, t, q, z). \quad (6.4)$$

(c) For every $\delta > 0$ there exists an $M > 0$ such that

$$\begin{aligned} & \max\left\{\|\nabla_z K(\epsilon, t, q, z)\|, \left(\sum_{i,j} |\partial_{z_i} \partial_{z_j} K(\epsilon, t, q, z)|^2\right)^{1/2}\right\} \\ & \leq M + \delta K(\epsilon, t, q, z). \end{aligned} \quad (6.5)$$

(d) There exists $c > 0$, $\eta > 0$ such that

$$K(\epsilon, t, q, z) \geq c\|z\|^{2\eta}. \quad (6.6)$$

3. There exists $a \geq 0, b \geq 0$ such that $\tilde{V}(t, q) \equiv a + b\|q\|^2 + V(t, q)$ is non-negative.

4. γ is C^2 , independent of p , and symmetric with eigenvalues bounded below by some $\lambda > 0$.

5. The eigenvalues of A are bounded below by some $c > 0$.

6. $\sigma(t, x)$ and $F(t, x)$ are continuous and Lipschitz in x with the Lipschitz constant uniform on compact time intervals.

7. $\sigma, \gamma, \partial_{q^i}\psi, A, \partial_{q^i}A, \partial_t A, \partial_{q^i}A, \partial_t \partial_{q^i}A$, and $\partial_{q^i}\partial_{q^j}A$ are bounded.

8. There exists $C > 0, M > 0$ such that

$$|\partial_t V(t, q)| \leq M + C(\|q\|^2 + \tilde{V}(t, q)), \quad (6.7)$$

$$\|\partial_t \psi(t, q)\|^2 \leq M + C(\|q\|^2 + \tilde{V}(t, q)), \quad (6.8)$$

$$\|F(t, x)\|^2 \leq M + C(\|q\|^2 + \tilde{V}(t, q)), \quad (6.9)$$

$$\|\partial_{q^i} \tilde{\gamma}(t, q)\| \leq M + C(\|q\|^2 + \tilde{V}(t, q)), \quad i = 1, \dots, n, \quad (6.10)$$

and

$$\left(\sum_{i,j} |\partial_{q^i} \partial_{q^j} V(t, q)|^2 \right)^{1/2} \leq M + C(\|q\|^2 + \tilde{V}(t, q)). \quad (6.11)$$

9. We have $\mathbb{R}^n$ -valued initial conditions $x_0^\epsilon = (q_0^\epsilon, p_0^\epsilon)$ and q_0 that satisfy the following:

For some $R > 0, C > 0$ we have $\|q_0\| \leq R, \|q_0^\epsilon\| \leq R$ and $K^\epsilon(0, x_0^\epsilon) \leq C$ for all $\epsilon > 0$ and all $\omega \in \Omega$.

For all $p > 0$ we have and $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2})$.

Let x_t^ϵ be the family of solutions to the SDE 1.10-1.11 with initial condition x_0^ϵ and q_t be a solution to the proposed limiting SDE, Eq. (4.26), with initial condition q_0 . For any $T > 0, \delta > 0$, we have

$$\lim_{\epsilon \rightarrow 0^+} P \left(\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\| > \delta \right) = 0. \quad (6.12)$$

Proof. By Lemmas C1 and C2, the maximal solutions, x_t^ϵ , to the SDE Eq. (1.10)-Eq. (1.11) and q_t to the SDE Eq. (1.12) are unique a.s. and a.s. exist for all $t \geq 0$.

Let $\chi : \mathbb{R}^n \rightarrow [0, 1]$ be a C^∞ bump function, equal to 1 on $B_1(0) \equiv \{\|q\| \leq 1\}$ and zero outside $B_2(0)$. Given $r > 0$ let $\chi_r(q) = \chi(q/r)$. Define

$$\begin{aligned} V_r(t, q) &= \chi_r(q)V(t, q), \quad F_r(t, x) = \chi_r(q)F(t, x), \quad \psi_r(t, q) = \chi_r(q)\psi(t, q), \\ \gamma_r(t, q) &= \chi_r(q)\gamma(t, q) + (1 - \chi_r(q))\lambda I. \end{aligned} \quad (6.13)$$

Replacing V with V_r , F with F_r etc., we arrive at an SDE satisfying the hypotheses of Theorem 5.1. Let $x_t^{r,\epsilon}$ be the corresponding solution to Eq. (1.10)-Eq. (1.11) and q_t^r the limit of $q_t^{r,\epsilon}$, both using the same initial conditions as the original systems. Theorem 5.1 then implies that for all $T > 0$, $p > 0$, $0 < \beta < p/2$ and all $r > R$ we have

$$E \left[\sup_{t \in [0, T]} \|q_t^{r,\epsilon} - q_t^r\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (6.14)$$

We will now use this result to prove that q_t^ϵ converges to q_t in probability.

For each $r > R$ define the stopping times $\tau_r^\epsilon = \inf\{t : \|q_t^\epsilon\| \geq r\}$ and $\tau_r = \inf\{t : \|q_t\| \geq r\}$. The drifts and diffusions of the modified and unmodified SDEs agree on the ball $\{\|q\| \leq r\}$, so uniqueness of solutions implies

$$q_{\tau_r^\epsilon \wedge t}^\epsilon = q_{\tau_r^\epsilon \wedge t}^{r,\epsilon} \text{ for all } t \geq 0 \text{ a.s.} \quad (6.15)$$

and

$$q_{\tau_r \wedge t} = q_{\tau_r \wedge t}^r \text{ for all } t \geq 0 \text{ a.s.} \quad (6.16)$$

Using this, given $T > 0$, $\delta > 0$ we can calculate

$$\begin{aligned} & P \left(\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\| > \delta \right) \\ &= P \left(\tau_r \wedge \tau_r^\epsilon > T, \sup_{t \in [0, T]} \|q_{\tau_r^\epsilon \wedge t}^\epsilon - q_{\tau_r \wedge t}\| > \delta \right) + P \left(\tau_r \wedge \tau_r^\epsilon \leq T, \sup_{t \in [0, T]} \|q_t^\epsilon - q_t\| > \delta \right) \\ &= P \left(\tau_r \wedge \tau_r^\epsilon > T, \sup_{t \in [0, T]} \|q_t^{r,\epsilon} - q_t^r\| > \delta \right) + P \left(\tau_r \wedge \tau_r^\epsilon \leq T, \sup_{t \in [0, T]} \|q_t^\epsilon - q_t\| > \delta \right) \\ &\leq P \left(\sup_{t \in [0, T]} \|q_t^{r,\epsilon} - q_t^r\| > \delta \right) + P(\tau_r \wedge \tau_r^\epsilon \leq T). \end{aligned} \quad (6.17)$$

The first term converges to zero as $\epsilon \rightarrow 0^+$ by Eq. (6.14), so we focus on the

second.

$$\begin{aligned}
P(\tau_r \wedge \tau_r^\epsilon \leq T) &\leq P\left(\sup_{t \in [0, T]} \|q_t^{r, \epsilon} - q_t^r\| > 1\right) + P(\tau_r \leq T) \\
&\quad + P\left(\tau_r > T, \tau_r^\epsilon \leq T, \sup_{t \in [0, T]} \|q_t^{r, \epsilon} - q_t^r\| \leq 1\right) \\
&\leq P\left(\sup_{t \in [0, T]} \|q_t^{r, \epsilon} - q_t^r\| > 1\right) + P\left(\sup_{t \in [0, T]} \|q_t\| \geq r\right) \\
&\quad + P\left(\tau_r^\epsilon \leq T, \|q_{\tau_r^\epsilon \wedge T}^{r, \epsilon} - q_{\tau_r^\epsilon \wedge T}^r\| \leq 1\right) \\
&= P\left(\sup_{t \in [0, T]} \|q_t^{r, \epsilon} - q_t^r\| > 1\right) + P\left(\sup_{t \in [0, T]} \|q_t\| \geq r\right) \\
&\quad + P\left(\tau_r^\epsilon \leq T, \|q_{\tau_r^\epsilon \wedge T}^\epsilon - q_{\tau_r^\epsilon \wedge T}^r\| \leq 1\right),
\end{aligned} \tag{6.18}$$

where we again used the uniqueness result, Eq. (6.15).

On the event where $\tau_r^\epsilon \leq T$ and $\|q_{\tau_r^\epsilon \wedge T}^\epsilon - q_{\tau_r^\epsilon \wedge T}^r\| \leq 1$ we have $\|q_{\tau_r^\epsilon \wedge T}^\epsilon\| \geq r$ and hence

$$\|q_{\tau_r^\epsilon \wedge T}^\epsilon\| \geq \|q_{\tau_r^\epsilon \wedge T}^\epsilon\| - \|q_{\tau_r^\epsilon \wedge T}^\epsilon - q_{\tau_r^\epsilon \wedge T}^r\| \geq r - 1. \tag{6.19}$$

Therefore $\sup_{t \in [0, T]} \|q_t\| \geq r - 1$ on this event.

Combining the above calculation with Eq. (6.14), for any $r > 0$ we obtain

$$\begin{aligned}
&\limsup_{\epsilon \rightarrow 0^+} P\left(\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\| > \delta\right) \\
&\leq \limsup_{\epsilon \rightarrow 0^+} P\left(\sup_{t \in [0, T]} \|q_t^{r, \epsilon} - q_t^r\| > \delta\right) + \limsup_{\epsilon \rightarrow 0^+} P\left(\sup_{t \in [0, T]} \|q_t^{r, \epsilon} - q_t^r\| > 1\right) \\
&\quad + P\left(\sup_{t \in [0, T]} \|q_t\| \geq r\right) + P\left(\sup_{t \in [0, T]} \|q_t\| \geq r - 1\right) \\
&\leq 2P\left(\sup_{t \in [0, T]} \|q_t\| \geq r - 1\right).
\end{aligned} \tag{6.20}$$

Non-explosion of q_t implies that $P\left(\sup_{t \in [0, T]} \|q_t\| \geq r - 1\right) \rightarrow 0$ as $r \rightarrow \infty$ and so we have proven the claimed result. $\square$

Appendix A. Assumptions

In this appendix we collect all the assumptions that are needed for one of the main results, Theorem 5.1. In the body of the paper, we will restate each assumption when it is first used.

Assumption 1. *We assume that σ , F , and γ are continuous, the Hamiltonian has the form given in Eq. (1.8) where $K(\epsilon, t, q, z)$ is non-negative and C^2 in (t, q, z) for each ϵ , ψ is C^2 , and the solutions, x_t^ϵ , to the SDE 1.10-1.11 exist for all $t \geq 0$.*

For every $T > 0$, we assume the following bounds hold on $(0, \epsilon_0] \times [0, T] \times \mathbb{R}^{2n}$:

1. *There exist $C > 0$ and $M > 0$ such that*

$$\max\{|\partial_t K(\epsilon, t, q, z)|, \|\nabla_q K(\epsilon, t, q, z)\|\} \leq M + CK(\epsilon, t, q, z). \quad (\text{A.1})$$

2. *There exist $c > 0$ and $M \geq 0$ such that*

$$\|\nabla_z K(\epsilon, t, q, z)\|^2 + M \geq cK(\epsilon, t, q, z). \quad (\text{A.2})$$

3. *For every $\delta > 0$ there exists an $M > 0$ such that*

$$\max\left\{\|\nabla_z K(\epsilon, t, q, z)\|, \left(\sum_{i,j} |\partial_{z_i} \partial_{z_j} K(\epsilon, t, q, z)|^2\right)^{1/2}\right\} \leq M + \delta K(\epsilon, t, q, z). \quad (\text{A.3})$$

Assumption 2. *For every $T > 0$, the following hold uniformly on $[0, T] \times \mathbb{R}^n$:*

1. *V is C^2 and $\nabla_q V$ is bounded.*
2. *γ is symmetric with eigenvalues bounded below by some $\lambda > 0$.*
3. *γ , F , $\partial_t \psi$, and σ are bounded.*
4. *There exists $C > 0$ such that the (random) initial conditions satisfy $K^\epsilon(0, x_0^\epsilon) \leq C$ for all $\epsilon > 0$ and all $\omega \in \Omega$.*

Assumption 3. *For every $T > 0$ there exists $c > 0$, $\eta > 0$ such that*

$$K(\epsilon, t, q, z) \geq c\|z\|^{2\eta} \quad (\text{A.4})$$

on $(0, \epsilon_0] \times [0, T] \times \mathbb{R}^{2n}$.

Assumption 4. γ is C^1 and is independent of p .

Assumption 5. K has the form

$$K(\epsilon, t, q, z) = \tilde{K}(\epsilon, t, q, A^{ij}(t, q)z_i z_j) \quad (\text{A.5})$$

where $\tilde{K}(\epsilon, t, q, \zeta)$ is C^2 in (t, q, ζ) for every ϵ , non-negative on $(0, \epsilon_0] \times [0, \infty) \times \mathbb{R}^n \times [0, \infty)$ and $A(t, q)$ is a C^2 function whose values are symmetric $n \times n$ -matrices. We also assume that for every $T > 0$, the eigenvalues of A are bounded above and below by some constants $C > 0$ and $c > 0$ respectively, uniformly on $[0, T] \times \mathbb{R}^n$.

Assumption 6. $\tilde{K}$ is independent of q

Assumption 7. For every $T > 0$, $\nabla_q V$, F , and σ are Lipschitz in x uniformly in $t \in [0, T]$. We also assume that A and γ are C^2 , ψ is C^3 , and $\partial_t \psi$, $\partial_{q^i} \psi$, $\partial_{q^i} \partial_{q^j} \psi$, $\partial_t \partial_{q^i} \psi$, $\partial_t \partial_{q^i} \partial_{q^j} \psi$, $\partial_{q^i} \partial_{q^j} \partial_{q^k} \psi$, $\partial_t \gamma$, $\partial_{q^i} \gamma$, $\partial_t \partial_{q^i} \gamma$, $\partial_{q^i} \partial_{q^j} \gamma$, $\partial_t A$, $\partial_{q^i} A$, $\partial_t \partial_{q^i} A$, and $\partial_{q^i} \partial_{q^j} A$ are bounded on $[0, T] \times \mathbb{R}^n$ for every $T > 0$.

Appendix B. Linear Algebra Lemmas

For the benefit of the reader we collect some more or less well linear algebra lemmas in this appendix.

Lemma B1. Let A be an $n \times n$ -real or complex matrix with symmetric part $A^s = \frac{1}{2}(A + A^*)$. If the eigenvalues of A^s are bounded above (resp. below) by α then the real parts of the eigenvalues of A are bounded above (resp. below) by α .

Proof. Suppose y is an eigenvector of A with norm 1 corresponding to the eigenvalue λ . Let A^a be the antisymmetric part of A .

$$\Re(y^* A^a y) = \Re(\overline{y^* A^a y}) = \Re(y^* (A^a)^* y) = -\Re(y^* A^a y). \quad (\text{B.1})$$

Therefore $\Re(y^* A^a y) = 0$ and

$$\Re(\lambda) = \Re(y^* A y) = \Re(y^* A^s y). \quad (\text{B.2})$$

If the eigenvalues of A^s are bounded above by α then

$$\Re(\lambda) = \Re(y^* A^s y) \leq \alpha \|y\|^2 = \alpha \quad (\text{B.3})$$

and if they are bounded below by α then

$$\Re(\lambda) = \Re(y^* A^s y) \geq \alpha \|y\|^2 = \alpha. \quad (\text{B.4})$$

□

Lemma B2. *Let A be a positive definite $n \times n$ -real matrix with eigenvalues bounded below by $\lambda > 0$ and B be an $n \times n$ -real matrix whose symmetric part has eigenvalues bounded below by $\gamma > 0$. Then the eigenvalues of AB have real parts bounded below by $\gamma\lambda$.*

Proof. A is positive definite, so we can factor it as $A = DD^T$ where D is a real valued, invertible, $n \times n$ -matrix. AB and the conjugation $D^{-1}ABD = D^TBD$ have the same eigenvalues. The symmetric part of D^TBD is D^TB^sD and for any $y \in \mathbb{R}^n$,

$$y^T D^T B^s D y \geq \gamma y^T D^T D y. \quad (\text{B.5})$$

$D^T D$ is a positive definite matrix that has the same eigenvalues as $A = DD^T$ (both are equal to the squared singular values of D). The eigenvalues of A are bounded below by λ , so

$$y^T D^T D y \geq \lambda \|y\|^2. \quad (\text{B.6})$$

Therefore the eigenvalues of the symmetric part of D^TBD are bounded below by $\gamma\lambda$. Hence, by Lemma B1, the real parts of the eigenvalues of D^TBD , and hence of AB , are bounded below by $\gamma\lambda$. $\square$

Lemma B3. *Let A be an $n \times n$ -real or complex matrix whose symmetric part has eigenvalues bounded below by $\lambda > 0$. Then A is invertible and the symmetric part of A^{-1} has eigenvalues bounded below by $\lambda/\|A\|^2$.*

Proof. Lemma B1 implies that A is invertible. Let $v \in V$ be non-zero.

$$\begin{aligned} \langle v, (A^{-1})^s v \rangle &= \frac{1}{2} \Re(\langle v, A^{-1} v \rangle + \langle v, (A^*)^{-1} v \rangle) \\ &= \frac{1}{2} \Re(\langle AA^{-1} v, A^{-1} v \rangle + \langle A^* (A^*)^{-1} v, (A^*)^{-1} v \rangle) \\ &= \frac{1}{2} (\Re(\langle A^{-1} v, (A^*)^s (A^{-1} v) \rangle) + \Re(\langle A^{-1} v, (A^*)^a (A^{-1} v) \rangle)) \\ &\quad + \frac{1}{2} (\Re(\langle (A^*)^{-1} v, A^s ((A^*)^{-1} v) \rangle) + \Re(\langle (A^*)^{-1} v, A^a ((A^*)^{-1} v) \rangle)) \\ &= \frac{1}{2} (\langle A^{-1} v, A^s (A^{-1} v) \rangle + 0) + \frac{1}{2} (\langle (A^*)^{-1} v, A^s ((A^*)^{-1} v) \rangle + 0) \\ &\geq \frac{\lambda}{2} (\langle v, (AA^*)^{-1} v \rangle + \langle v, (A^* A)^{-1} v \rangle). \end{aligned} \quad (\text{B.7})$$

Using the singular value decomposition of A we see that both $(AA^*)^{-1}$ and $(A^*A)^{-1}$ are positive definite with eigenvalues bounded below by $1/\|A\|^2$. The result follows. $\square$

Appendix C. Non-Explosion of Solutions

In the course of proving our main result, Theorem 5.1, we showed that the limiting process, q_t , exist for all $t \geq 0$ with probability one, at least under Assumptions 1-7. Though we have assumed it to be the case throughout this paper, the same is not obvious for the family of solutions, x_t^ϵ , to the SDE 1.10-1.11. However, existence for all $t \geq 0$ can be proven under a collection of assumptions that are very similar to our Assumptions 1-7 from the main text, as shown in the following lemma. We emphasize that in this appendix, we do not employ any of the Assumptions 1-7 per se. The assumptions we do use are all listed below, in the statement of the lemma.

Lemma C1. *Suppose:*

1. *The family of Hamiltonians have the form*

$$H^\epsilon(t, x) = K(\epsilon, t, q, (p - \psi(t, q))/\sqrt{\epsilon}) + V(t, q) \quad (\text{C.1})$$

where $K(\epsilon, t, q, z)$ is non-negative, C^2 in (t, q, z) for every ϵ , V is C^2 , and ψ is a C^2 , $\mathbb{R}^n$ -valued function.

2. *There exists $a \geq 0, b \geq 0$ such that $\tilde{V}(t, q) \equiv a + b\|q\|^2 + V(t, q)$ is non-negative.*
3. *$\gamma(t, x)$, $\sigma(t, x)$, and $F(t, x)$ are continuous and locally Lipschitz in x with the Lipschitz constant uniform on compact time intervals.*
4. *σ is bounded.*
5. *The eigenvalues of γ (which are real, since γ is symmetric) are bounded below by some $\lambda > 0$.*
6. *For every $\epsilon \in (0, \epsilon_0]$, $t \geq 0$ there exists $c > 0$, $N > 0$, $\eta > 0$ such that*

$$K(\epsilon, t, q, z) \geq c\|z\|^{2\eta} \quad (\text{C.2})$$

for all $\|z\| \geq N$.

7. There exist $C > 0$ and $M > 0$ such that

$$\left(\sum_{i,j} |\partial_{z_i} \partial_{z_j} K(\epsilon, t, q, z)|^2 \right)^{1/2} \leq M + CK(\epsilon, t, q, z), \quad (\text{C.3})$$

$$|\partial_t K(\epsilon, t, q, z) + \partial_t V(t, q)| \leq M + C(\|q\|^2 + \tilde{V}(t, q) + K(\epsilon, t, q, z)), \quad (\text{C.4})$$

and

$$\| -\partial_t \psi(t, q) + F(t, x) \|^2 \leq M + C(\|q\|^2 + \tilde{V}(t, q)). \quad (\text{C.5})$$

Then the maximal solution, x_t^ϵ , to the SDE 1.10-1.11 is unique a.s. and a.s. exist for all $t \geq 0$.

Proof. Fix $\epsilon > 0$. The assumptions imply that an a.s. unique, maximal solution x_t^ϵ exists up to explosion time e^ϵ (see Section 3.4 in [29]). Non-explosion of x_t^ϵ (i.e. $e^\epsilon = \infty$ a.s.) will follow from the existence of a Lyapunov function (see Theorem 3.5 in [29]), a non-negative C^2 function, $U(t, x)$, that satisfies:

1. For any $t \geq 0$, $\lim_{x \rightarrow \infty} U(t, x) = \infty$,
- 2.

$$L[U](t, x) \leq \tilde{M} + \tilde{C}U(t, x) \quad (\text{C.6})$$

for some $\tilde{M} \geq 0$, $\tilde{C} > 0$, where L is the time-dependent generator

$$\begin{aligned} L[U](t, x) = & \partial_t U(t, x) + \nabla_p H^\epsilon(t, x) \cdot \nabla_q U(t, x) \\ & + (-\gamma(t, x) \nabla_p H^\epsilon(t, x) - \nabla_q H^\epsilon(t, x) + F(t, x)) \cdot \nabla_p U(t, x) \\ & + \frac{1}{2} \Sigma_{ij}(t, x) \partial_{p_i} \partial_{p_j} U(t, x) \end{aligned} \quad (\text{C.7})$$

and $\Sigma_{ij} = \sum_\rho \sigma_{i\rho} \sigma_{j\rho}$.

We note that to connect with the result as stated in [29], one needs to work with $\tilde{M}/\tilde{C} + U$ in place of U , but we find the above formulation more convenient here.

Fix $\epsilon > 0$. As our candidate Lyapunov function, we define

$$U(t, x) \equiv \|q\|^2 + \tilde{V}(t, q) + K^\epsilon(t, x) = a + (1 + b)\|q\|^2 + H^\epsilon(t, x). \quad (\text{C.8})$$

By assumption, $\tilde{V}(t, q)$ and $K^\epsilon(t, x)$ are non-negative and C^2 , therefore U is as well.

Fix $t \geq 0$. By assumption there exists $c > 0$, $N > 0$, $\eta > 0$ such that

$$K(\epsilon, t, q, z) \geq c\|z\|^{2\eta} \quad (\text{C.9})$$

for all $\|z\| \geq N$.

Given $R > 0$, let

$$\max\{\|q\|, \|p\|\} \geq \max\left\{R^{1/2}, \sup_{\|q\| \leq R^{1/2}} \|\psi(t, q)\| + \epsilon^{1/2} \left(N + (R/c)^{1/2\eta}\right)\right\}. \quad (\text{C.10})$$

If $\|q\| \geq R^{1/2}$ then $U(t, x) \geq \|q\|^2 \geq R$. If $\|q\| < R^{1/2}$ then

$$\|p - \psi(t, q)\|/\sqrt{\epsilon} \geq \left(\|p\| - \sup_{\|q\| \leq R^{1/2}} \|\psi(t, q)\|\right)/\sqrt{\epsilon} \geq N. \quad (\text{C.11})$$

Hence

$$\begin{aligned} U(t, x) &\geq K^\epsilon(t, x) \geq c\|p - \psi(t, q)\|^{2\eta}/\epsilon^\eta \geq \frac{c}{\epsilon^\eta} \left(\|p\| - \sup_{\|q\| \leq R^{1/2}} \|\psi(t, q)\|\right)^{2\eta} \\ &\geq \frac{c}{\epsilon^\eta} \left(\epsilon^{\frac{1}{2}} \left(\frac{R}{c}\right)^{\frac{1}{2\eta}}\right)^{2\eta} = R. \end{aligned} \quad (\text{C.12})$$

Therefore $U(t, x) \rightarrow \infty$ as $x \rightarrow \infty$.

Using the inequality $\alpha\beta \leq \frac{1}{2}(\delta\alpha^2 + \frac{1}{\delta}\beta^2)$ for any $\alpha > 0$, $\beta > 0$, $\delta > 0$, we can compute

$$\begin{aligned} L[U](t, x) &= \partial_t H^\epsilon(t, x) + \nabla_p H^\epsilon(t, x) \cdot \nabla_q H^\epsilon(t, x) \\ &\quad + (-\gamma(t, x) \nabla_p H^\epsilon(t, x) - \nabla_q H^\epsilon(t, x) + F(t, x)) \cdot \nabla_p H^\epsilon(t, x) \\ &\quad + \frac{1}{2} \Sigma_{ij}(t, x) \partial_{p_i} \partial_{p_j} H^\epsilon(t, x) + 2(1 + b) \nabla_p H^\epsilon(t, x) \cdot q \end{aligned} \quad (\text{C.13})$$

$$\begin{aligned}
&\leq (\partial_t K)^\epsilon(t, x) + \partial_t V(t, q) - \frac{\lambda}{\epsilon} \|(\nabla_z K)^\epsilon(t, x)\|^2 \\
&\quad + \frac{1}{\sqrt{\epsilon}} (2(1+b)q + F(t, x) - \partial_t \psi(t, q)) \cdot (\nabla_z K)^\epsilon(t, x) \\
&\quad + \frac{1}{2\epsilon} \Sigma_{ij}(t, x) (\partial_{z_i} \partial_{z_j} K)^\epsilon(t, x) \\
&\leq M + CU(t, x) - \frac{\lambda}{\epsilon} \|(\nabla_z K)^\epsilon(t, x)\|^2 + \frac{1}{2\epsilon} \|\Sigma\|_{F, \infty} (M + CK^\epsilon(t, x)) \\
&\quad + \frac{1}{4\lambda} \|2(1+b)q + F(t, x) - \partial_t \psi(t, q)\|^2 + \frac{\lambda}{\epsilon} \|(\nabla_z K)^\epsilon(t, x)\|^2 \\
&\leq M + CU(t, x) + \frac{\lambda}{\epsilon} \|(\nabla_z K)^\epsilon(t, x)\|^2 \\
&\quad + \frac{1}{2\lambda} (4(1+b)^2 \|q\|^2 + M + C(\|q\|^2 + \tilde{V}(t, q))).
\end{aligned}$$

The right hand side is bounded by $\tilde{M} + \tilde{C}U$ for some $\tilde{M} \geq 0$, $\tilde{C} > 0$. This completes the proof that U is a Lyapunov function and allows us to conclude that $e^\epsilon = \infty$ a.s. i.e. the solution x_t^ϵ exists for all $t \geq 0$ a.s. $\square$

A similar result holds for proving non-explosion of the limiting equation, Eq. (4.26), under assumptions weaker than Assumptions 1-7. The following lemma can be proven by using the Lyapunov function $U(t, q) \equiv a + (1+b) \|q\|^2 + V(t, q)$ and Lemma B3. The proof closely follows that of the previous lemma and so we leave the details to the reader.

Lemma C2. *Consider an SDE on $\mathbb{R}^n$ of the form*

$$dq_t = \tau^{-1}(t, q_t) (-\nabla_q V(t, q_t) + \tilde{F}(t, q_t)) dt + \tilde{\sigma}(t, q_t) dW_t \quad (\text{C.14})$$

where

1. V is C^2 and there exists $a > 0$, $b > 0$ such that $\tilde{V}(t, q) \equiv a + b\|q\|^2 + V(t, q)$ is non-negative,
2. $\tilde{\sigma}(t, q)$, and $\tilde{F}(t, q)$ are continuous and locally Lipschitz in q with the Lipschitz constant uniform on compact time intervals,
3. $\tilde{\sigma}$ is bounded,
4. τ is C^1 , bounded, and its symmetric part has eigenvalues bounded below by $\lambda > 0$,

5. there exists $M > 0$ and $C > 0$ such that

$$\left(\sum_{i,j} |\partial_{q^i} \partial_{q^j} V(t, q)|^2 \right)^{1/2} \leq M + C (\|q\|^2 + \tilde{V}(t, q)), \quad (\text{C.15})$$

$$|\partial_t V(t, q)| \leq M + C (\|q\|^2 + \tilde{V}(t, q)), \quad (\text{C.16})$$

and

$$\|\tilde{F}(t, q)\|^2 \leq M + C (\|q\|^2 + \tilde{V}(t, q)). \quad (\text{C.17})$$

Then the maximal solution to the SDE C.14 is unique a.s. and a.s. exist for all $t \geq 0$.

Proof. By Lemma B3, τ is invertible and the symmetric part of τ^{-1} has eigenvalues bounded below by $\lambda/\|\tau\|^2$. τ^{-1} is also C^1 , therefore the drift and diffusion are continuous and locally Lipschitz. Therefore we have existence and uniqueness up to explosion time.

Define $U(t, q) = \|q\|^2 + \tilde{V}(t, q) = a + (1+b)\|q\|^2 + V(t, q)$. U is C^2 and non-negative since $\tilde{V}$ is and, for any t , we have $U(t, q) \geq \|q\|^2 \rightarrow \infty$ as $q \rightarrow \infty$. We also have

$$L[U](t, q) \quad (\text{C.18})$$

$$= \partial_t U(t, q) + \nabla_q U(t, q) \cdot \tau^{-1}(t, q_t) (-\nabla_q V(t, q_t) + \tilde{F}(t, q_t)) + \frac{1}{2} \tilde{\Sigma}_{i,j} \partial_{q^i} \partial_{q^j} U(t, q) \quad (\text{C.19})$$

$$= \partial_t V(t, q) + (\nabla_q V(t, q) + 2(1+b)q) \cdot \tau^{-1}(t, q_t) (-\nabla_q V(t, q_t) + \tilde{F}(t, q_t)) \quad (\text{C.20})$$

$$+ \frac{1}{2} \tilde{\Sigma}_{i,j} (2(1+b)\delta^{i,j} + \partial_{q^i} \partial_{q^j} V(t, q)) \quad (\text{C.21})$$

where $\tilde{\Sigma}_{i,j} = \sum_\rho \tilde{\sigma}_{i\rho} \tilde{\sigma}_{j\rho}$. We can bound this as follows

$$L[U](t, q) \quad (\text{C.22})$$

$$\leq M + CU + (\nabla_q V(t, q) + 2(1+b)q) \cdot \tau^{-1}(t, q) (-\nabla_q V(t, q) + \tilde{F}(t, q)) \quad (\text{C.23})$$

$$+ (1+b)\|\tilde{\sigma}(t, q)\|_F^2 + \frac{1}{2}(M + CU(t, q)) \quad (\text{C.24})$$

$$\leq M + CU(t, q) + (\nabla_q V(t, q) + 2(1+b)q) \cdot \tau^{-1}(t, q) (-\nabla_q V(t, q) + \tilde{F}(t, q)) \quad (\text{C.25})$$

τ^{-1} is bounded since τ is and it is valued in set of matrices whose symmetric part has eig. $\geq \lambda > 0$ -see dynamical systems notes (a topological argument based on image of a compact set under a continuous map). Using $ab \leq \frac{1}{4\delta}a^2 + \delta b^2$ for any $\delta > 0$ we can compute

$$(\nabla_q V(t, q) + 2(1+b)q) \cdot \tau^{-1}(t, q)(-\nabla_q V(t, q) + \tilde{F}(t, q)) \quad (\text{C.26})$$

$$= -\nabla_q V(t, q) \cdot \tau^{-1}(t, q) \nabla_q V(t, q) - 2(1+b)q \cdot \tau^{-1}(t, q) \nabla_q V(t, q) \quad (\text{C.27})$$

$$+ \nabla_q V(t, q) \cdot \tau^{-1}(t, q) \tilde{F}(t, q) + 2(1+b)q \cdot \tau^{-1}(t, q) \tilde{F}(t, q) \quad (\text{C.28})$$

$$\leq -\frac{\lambda}{\|\tau(t, q)\|^2} \|\nabla_q V(t, q)\|^2 + \frac{\|\tau(t, q)\|^2}{4\lambda} (\|\tau^{-1}(t, q) \tilde{F}(t, q)\| + 2(1+b)\|(\tau^{-1})^T(t, q)q\|)^2 \quad (\text{C.29})$$

$$+ \frac{\lambda}{\|\tau(t, q)\|^2} \|\nabla_q V(t, q)\|^2 + 2(1+b)\|\tau^{-1}\|_\infty \|q\| \|\tilde{F}(t, q)\| \quad (\text{C.30})$$

$$\leq \frac{\|\tau(t, q)\|^2}{4\lambda} (\|\tau^{-1}(t, q)\| \|\tilde{F}(t, q)\| + 2(1+b)\|\tau^{-1}(t, q)\| \|q\|)^2 \quad (\text{C.31})$$

$$+ (1+b)\|\tau^{-1}\|_\infty (\|q\|^2 + \|\tilde{F}(t, q)\|^2) \leq C\|\tilde{F}(t, q)\|^2 + \|q\|^2 \leq \tilde{M} + \tilde{C}U(t, q). \quad (\text{C.32})$$

□

Corollary C1. *The above result holds for the limiting SDE if we have:*

1. V is C^2 and there exists $a > 0$, $b > 0$ such that $\tilde{V}(t, q) \equiv a + b\|q\|^2 + V(t, q)$ is non-negative.
2. $\sigma(t, x)$, and $F(t, x)$ are continuous and locally Lipschitz in x with the Lipschitz constant uniform on compact time intervals.
3. σ is bounded
4. γ is bounded, C^2 , symmetric, and its eigenvalues are bounded below by $\lambda > 0$.
5. ψ is C^3 and $\partial_{q^i} \psi(t, q)$ are bounded for $i = 1, \dots, n$.
6. $A(t, q)$ is C^2 , symmetric, and its eigenvalues are bounded below by $c > 0$.
7. $A(t, q)$ and $\partial_{q^i} A(t, q)$ are bounded, $i = 1, \dots, n$.

8. There exists $M > 0$ and $C > 0$ such that

$$\left(\sum_{i,j} |\partial_{q^i} \partial_{q^j} V(t, q)|^2 \right)^{1/2} \leq M + C (\|q\|^2 + \tilde{V}(t, q)), \quad (\text{C.33})$$

$$|\partial_t V(t, q)| \leq M + C (\|q\|^2 + \tilde{V}(t, q)), \quad (\text{C.34})$$

$$\|\partial_{q^i} \tilde{\gamma}(t, q)\|^2 \leq M + C (\|q\|^2 + \tilde{V}(t, q)), \quad (\text{C.35})$$

and

$$\| -\partial_t \psi(t, q) + F(t, q, \psi(t, q)) \|^2 \leq M + C (\|q\|^2 + \tilde{V}(t, q)). \quad (\text{C.36})$$

Proof. Recall the SDE

$$\begin{aligned} dq_t = & \tilde{\gamma}^{-1}(t, q_t) (-\partial_t \psi(t, q_t) - \nabla_q V(t, q_t) + F(t, q_t, \psi(t, q_t))) dt + S(t, q_t) dt \\ & + \tilde{\gamma}^{-1}(t, q_t) \sigma(t, q_t, \psi(t, q_t)) dW_t. \end{aligned} \quad (\text{C.37})$$

1. $\tilde{\gamma}_{ik}(t, q) \equiv \gamma_{ik}(t, q) + \partial_{q^k} \psi_i(t, q) - \partial_{q^i} \psi_k(t, q)$
2. The index placement on $\tilde{\gamma}^{-1}$ is defined by $(\tilde{\gamma}^{-1})^{ij} \tilde{\gamma}_{jk} = \delta_k^i$.
3. $S^i(t, q) \equiv Q^{ijl}(t, q) J_{jl}(t, q, \psi(t, q))$.
4. $Q^{ijl}(t, q) \equiv \partial_{q^k} (\tilde{\gamma}^{-1})^{ij}(t, q) A^{kl}(t, q) - \frac{1}{2} (\tilde{\gamma}^{-1})^{ik}(t, q) \partial_{q^k} A^{jl}(t, q)$, where A^{ij} is the matrix-valued function from Assumption 5.
5. $J_{ij}(t, x) \equiv G_{ij}^{kl}(t, q) \Sigma_{kl}(t, x)$
6. $G_{ij}^{kl}(t, q) = \int_0^\infty (e^{-y(A\tilde{\gamma})(t, q)})_i^k (e^{-y(A\tilde{\gamma})(t, q)})_j^l dy$, where $A\tilde{\gamma}$ is the matrix $(A\tilde{\gamma})_j^i = \tilde{\gamma}_{jk} A^{ki}$.
7. $\Sigma = \sigma \sigma^T$

Therefore

$$\tilde{F}(t, q) = -\partial_t \psi(t, q) + F(t, q, \psi(t, q)) + S(t, q), \quad (\text{C.38})$$

$$\tilde{\sigma}(t, q) = \tilde{\gamma}^{-1}(t, q) \sigma(t, q, \psi(t, q)). \quad (\text{C.39})$$

γ and $\nabla_q \psi$ are bounded and C^1 , so $\tilde{\gamma}$ is. Hence $\tilde{\gamma}^{-1}$ is. So $\tilde{\sigma}$ is. So now we just need to prove:

1. $S(t, q)$ is continuous and locally Lipschitz in q with the Lipschitz constant uniform on compact time intervals.
2. There exists $M > 0$ and $C > 0$ such that

$$\|S(t, q)\|^2 \leq M + C (\|q\|^2 + \tilde{V}(t, q)). \quad (\text{C.40})$$

$Q^{ijl}(t, q) \equiv \partial_{q^k}(\tilde{\gamma}^{-1})^{ij}(t, q)A^{kl}(t, q) - \frac{1}{2}(\tilde{\gamma}^{-1})^{ik}(t, q)\partial_{q^k}A^{jl}(t, q)$ is continuous and locally Lipschitz. $G_{i,j}^{k,l}$ are C^1 and Σ is cont and locally Lip. So we need to bound

$$\|S(t, q)\| \leq C \int_0^\infty \|e^{-y(A\tilde{\gamma})(t,q)}\|^2 dy (\|\nabla_q \tilde{\gamma}(t, q)\| \|A(t, q)\| + \|\nabla_q A(t, q)\|) \quad (\text{C.41})$$

$$\|e^{-y(A\tilde{\gamma})(t,q)}\| \leq (\|A(t, q)\|/c)^{1/2} e^{-yc\lambda} \quad (\text{C.42})$$

so

$$\|S(t, q)\| \leq C (\|A(t, q)\|/c) \int_0^\infty e^{-2yc\lambda} dy (\|\nabla_q \tilde{\gamma}(t, q)\| \|A(t, q)\| + \|\nabla_q A(t, q)\|) \quad (\text{C.43})$$

$$\leq M + C \|\nabla_q \tilde{\gamma}(t, q)\| \quad (\text{C.44})$$

So

$$\|S(t, q)\|^2 \leq M + C \|\nabla_q \tilde{\gamma}(t, q)\|^2 \leq M + C (\|q\|^2 + \tilde{V}(t, q)), \quad (\text{C.45})$$

□

Acknowledgments

J.W. was partially supported by NSF grants DMS 131271 and DMS 1615045.

- [1] M. Smoluchowski, “Drei vortrage uber diffusion, brownsche bewegung und koagulation von kolloidteilchen,” *Zeitschrift fur Physik*, vol. 17, pp. 557–585, 1916.
- [2] H. Kramers, “Brownian motion in a field of force and the diffusion model of chemical reactions,” *Physica*, vol. 7, no. 4, pp. 284 – 304, 1940.

- [3] G. A. Pavliotis and A. M. Stuart, “White noise limits for inertial particles in a random field,” *Multiscale Modeling & Simulation*, vol. 1, no. 4, pp. 527–553, 2003.
- [4] C. Chevalier and F. Debbasch, “Relativistic diffusions: A unifying approach,” *Journal of Mathematical Physics*, vol. 49, no. 4, 2008.
- [5] I. Bailleul, “A stochastic approach to relativistic diffusions,” in *Annales de l’institut Henri Poincaré (B)*, vol. 46, pp. 760–795, 2010.
- [6] M. A. Pinsky, “Isotropic transport process on a riemannian manifold,” *Transactions of the American Mathematical Society*, vol. 218, pp. 353–360, 1976.
- [7] M. A. Pinsky, “Homogenization in stochastic differential geometry,” *Publications of the Research Institute for Mathematical Sciences*, vol. 17, no. 1, pp. 235–244, 1981.
- [8] E. Jørgensen, “Construction of the brownian motion and the ornstein-uhlenbeck process in a riemannian manifold on basis of the gangolli-mc.kean injection scheme,” *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 44, no. 1, pp. 71–87, 1978.
- [9] R. M. Dowell, *Differentiable approximations to Brownian motion on manifolds*. PhD thesis, University of Warwick, 1980.
- [10] X.-M. Li, “Random Perturbation to the Geodesic Equation,” *Ann. Prob.*, vol. 44, no. 1, pp. 544–566, 2016.
- [11] J. Angst, I. Bailleul, and C. Tardif, “Kinetic brownian motion on riemannian manifolds,” *arXiv preprint arXiv:1501.03679*, 2015.
- [12] J.-M. Bismut, “The hypoelliptic laplacian on the cotangent bundle,” *Journal of the American Mathematical Society*, vol. 18, no. 2, pp. 379–476, 2005.
- [13] J.-M. Bismut, “Hypoelliptic laplacian and probability,” *J. Math. Soc. Japan*, vol. 67, pp. 1317–1357, 10 2015.
- [14] E. Nelson, *Dynamical Theories of Brownian Motion*. Mathematical Notes - Princeton University Press, Princeton University Press, 1967.

- [15] G. Pavliotis and A. Stuart, *Multiscale Methods: Averaging and Homogenization*. Texts in Applied Mathematics, Springer New York, 2008.
- [16] P. Hänggi, “Nonlinear fluctuations: The problem of deterministic limit and reconstruction of stochastic dynamics,” *Phys. Rev. A*, vol. 25, pp. 1130–1136, Feb 1982.
- [17] G. Volpe, L. Helden, T. Brettschneider, J. Wehr, and C. Bechinger, “Influence of noise on force measurements,” *Physical review letters*, vol. 104, no. 17, p. 170602, 2010.
- [18] J. M. Sancho, M. S. Miguel, and D. Dürr, “Adiabatic elimination for systems of brownian particles with nonconstant damping coefficients,” *Journal of Statistical Physics*, vol. 28, no. 2, pp. 291–305, 1982.
- [19] S. Hottovy, A. McDaniel, G. Volpe, and J. Wehr, “The Smoluchowski-Kramers Limit of Stochastic Differential Equations with Arbitrary State-Dependent Friction,” *Communications in Mathematical Physics*, vol. 336, no. 3, pp. 1259–1283, 2014.
- [20] D. P. Herzog, S. Hottovy, and G. Volpe, “The small-mass limit for Langevin dynamics with unbounded coefficients and positive friction,” *Journal of Statistical Physics*, vol. 163, no. 3, pp. 659–673, 2016.
- [21] J. Birrell, S. Hottovy, G. Volpe, and J. Wehr, “Small Mass Limit of a Langevin Equation on a Manifold,” *ArXiv e-prints*, Apr. 2016.
- [22] R. Chetrite and K. Gawędzki, “Fluctuation relations for diffusion processes,” *Communications in Mathematical Physics*, vol. 282, no. 2, pp. 469–518, 2008.
- [23] K. Gawędzki, “Fluctuation relations in stochastic thermodynamics,” *arXiv preprint arXiv:1308.1518*, 2013.
- [24] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics, Springer New York, 2014.
- [25] D. T. Khoa, N. Ohtsuka, M. Matin, A. Faessler, S. Huang, E. Lehmann, and R. K. Puri, “In-medium effects in the description of heavy-ion collisions with realistic nn interactions,” *Nuclear Physics A*, vol. 548, no. 1, pp. 102 – 130, 1992.

- [26] J. Ortega, *Matrix Theory: A Second Course*. University Series in Mathematics, Springer US, 2013.
- [27] J. Birrell and J. Wehr, “A homogenization theorem for langevin systems with an application to hamiltonian dynamics,” *arXiv preprint arXiv:1707.02884*, 2017.
- [28] R. M. Wilcox, “Exponential operators and parameter differentiation in quantum physics,” *Journal of Mathematical Physics*, vol. 8, no. 4, 1967.
- [29] R. Khasminskii, *Stochastic stability of differential equations*, vol. 66. Springer Science & Business Media, 2011.