

# A homogenization theorem for Langevin systems with an application to Hamiltonian dynamics

Jeremiah Birrell and Jan Wehr

**Abstract.** This paper studies homogenization of stochastic differential systems. The standard example of this phenomenon is the small mass limit of Hamiltonian systems. We consider this case first from the heuristic point of view, stressing the role of detailed balance and presenting the heuristics based on a multiscale expansion. This is used to propose a physical interpretation of recent results by the authors, as well as to motivate a new theorem proven here. Its main content is a sufficient condition, expressed in terms of solvability of an associated partial differential equation (“the cell problem”), under which the homogenization limit of an SDE is calculated explicitly. The general theorem is applied to a class of systems, satisfying a generalized detailed balance condition with a position-dependent temperature.

**Mathematics Subject Classification (2010).** 60H10, 82C31.

**Keywords.** homogenization, stochastic differential equation, Hamiltonian system, small mass limit, noise-induced drift.

## 1. Introduction and background

This paper studies the small mass limit of a general class of Langevin equations. Langevin dynamics is defined in terms of canonical variables—positions and momenta—by adding damping and (Itô) noise terms to Hamiltonian equations. In the limit when the mass, or masses, of the system’s particles, go to zero, the momenta homogenize, and one obtains a limiting equation for the position variables only. This is a great simplification which often allows one to see the nature of the dynamics more clearly. If the damping matrix of the original system depends on its state, a *noise-induced drift* arises in the limit. We analyze and interpret this term from several points of view. The paper consists of four parts. The first part contains general background on stochastic differential equations. In the second part, the small-mass limit

of Langevin equations is studied using a multiscale expansion. This method requires making additional assumptions, but it leads to correct results in all cases in which rigorous proofs are known. The third part presents a new rigorous result about homogenization of a general class of singularly perturbed SDEs. The final part applies this result to prove a theorem about the homogenization of a large class of Langevin systems.

### 1.1. Stochastic differential equations

Let us start from a general background on Langevin equations. The material presented here is not new, and its various versions can be found in many textbooks, see for example [1]. We do not strive for complete precision or a listing of all necessary assumptions in our discussions here. The aim of the first two sections is to motivate and facilitate reading the remainder of the paper. Detailed technical considerations will be reserved for Sections 3 and 4, where we present our new results.

Consider the stochastic differential equation

$$dy_t = b(y_t) dt + \sigma(y_t) dW_t. \quad (1.1)$$

The process  $y_t$  takes values in  $\mathbb{R}^m$ ,  $b$  is a vector field in  $\mathbb{R}^m$ ,  $W$  is an  $n$ -dimensional Wiener process and  $\sigma$  is an  $m \times n$ -matrix-valued function. Define an  $m \times m$  matrix  $\Sigma$  by  $\Sigma = \sigma\sigma^T$ . The equation Eq. (1.1) defines a Markov process with the infinitesimal operator

$$(Lf)(y) = \frac{1}{2} \sum_{i,j} \partial_i \partial_j f + b_i \partial_i f \quad (1.2)$$

where we are writing  $\partial_i$  for  $\partial_{y_i}$  and suppressing the dependence of  $\Sigma$ ,  $b_i$  and  $f$  on  $y$  from the notation. Summation over repeating indices is implied.

We assume that this process has a unique stationary probability measure with a  $C^2$ -density  $h(y)$ . Under this assumption  $h$  satisfies the equation

$$L^* h = 0 \quad (1.3)$$

where  $L^*$  denotes the formal adjoint of  $L$ ,

$$L^* f = \frac{1}{2} \partial_i \partial_j (\Sigma_{ij} f) - \partial_i (b_i f). \quad (1.4)$$

That is, we have

$$\partial_i \left( \frac{1}{2} \partial_j (\Sigma_{ij} h) - b_i h \right) = 0. \quad (1.5)$$

Consider the special case when  $h$  solves the equation

$$\frac{1}{2} \partial_j (\Sigma_{ij} h) - b_i h = 0. \quad (1.6)$$

In this case the operator  $L$  is symmetric on the space  $L^2(\mathbb{R}^m, h) \equiv L_h^2$  of square-integrable functions with the weight  $h$ , as the following calculation shows. Using product formula, we have

$$\int (Lf) gh = \int f L^* (gh) = \int f \partial_i \left( \frac{1}{2} \partial_j (\Sigma_{ij} gh) - b_i gh \right). \quad (1.7)$$

The expression in parentheses equals

$$\frac{1}{2}\partial_j g (\Sigma_{ij} h) + g \frac{1}{2}\partial_j (\Sigma_{ij} h) - g b_i h = \frac{1}{2}\partial_j g (\Sigma_{ij} h) \quad (1.8)$$

by Eq. (1.6). Applying product formula again, we obtain

$$\int (Lf) gh = \int f \left( \frac{1}{2}\Sigma_{ij} (\partial_i \partial_j g) h + \frac{1}{2}\partial_i (\Sigma_{ij} h) \partial_j g \right) \quad (1.9)$$

which, by another application of Eq. (1.6), equals

$$\int f \left( \frac{1}{2}\Sigma_{ij} \partial_i \partial_j g - b_j \partial_j g \right) h = \int f (Lg) h. \quad (1.10)$$

Here is a more complete discussion:

## 1.2. Detailed balance condition and symmetry of the infinitesimal operator

We have

$$\begin{aligned} \int (Lf) gh &= \int \left( \frac{1}{2}\Sigma_{ij} \partial_i \partial_j f + b_i \partial_i f \right) gh \\ &= -\frac{1}{2} \int \partial_i f \partial_j (\Sigma_{ij} gh) + \int (\partial_i f) b_i gh \\ &= -\frac{1}{2} \int \partial_i f [\partial_j (\Sigma_{ij} h) g + \Sigma_{ij} h \partial_j g] + \partial (\partial_i f) b_i gh \\ &= \int \partial_i f \left[ -\frac{1}{2}\partial_j (\Sigma_{ij} h) + b_i h \right] g - \frac{1}{2} \int \Sigma_{ij} \partial_i f \partial_j g h. \end{aligned} \quad (1.11)$$

Interchanging the roles of  $f$  and  $g$  and canceling the term symmetric in  $f$  and  $g$ , we obtain

$$\int (Lf) gh - \int f (Lg) h = \int [(\partial_i f) g - (\partial_i g) f] \left( -\frac{1}{2}\partial_j (\Sigma_{ij} h) + b_i h \right). \quad (1.12)$$

If  $h$  is a solution to the equation

$$-\frac{1}{2}\partial_j (\Sigma_{ij} h) + b_i h = 0 \quad (1.13)$$

then the above expression is zero, showing that the operator  $L$  is symmetric on the space  $L_h^2$ . Conversely, for this symmetry to hold, the  $\mathbb{R}^m$ -valued function  $-\frac{1}{2}\partial_j (\Sigma_{ij} h) + b_i h$  has to be orthogonal to all elements of the space  $L^2$  (of functions with values in  $\mathbb{R}^m$ ) of the form  $(\partial_i f) g - (\partial_i g) f$ . It is not hard to prove that every  $C^1$  function with this property must vanish, and thus, that  $\frac{1}{2}\partial_j (\Sigma_{ij} h) - b_i h = 0$ . Here is a sketch of a proof: suppose  $\phi$  is  $C^1$  and orthogonal to all such functions. That is, for every  $f$  and  $g$ ,

$$\int [\phi_i (\partial_i f) g - \phi_i (\partial_i g) f] = 0. \quad (1.14)$$

Integrating the first term by parts we obtain

$$\int [-(\partial_i \phi_i) g - 2\phi_i \partial_i g] f = 0. \quad (1.15)$$

Since this holds for all  $f$ , it follows that

$$-(\partial_i \phi_i) g - 2\phi_i \partial_i g = 0 \quad (1.16)$$

and thus also

$$\int [-(\partial_i \phi_i) g - 2\phi_i \partial_i g] = 0. \quad (1.17)$$

Integrating the second term by parts, we get

$$\int (\partial_i \phi_i) g = 0 \quad (1.18)$$

and, since this is true for every  $g$ , it follows that  $\partial_i \phi_i$  vanishes. We thus have, for every  $g$

$$\phi_i \partial_i g = 0 \quad (1.19)$$

and this implies that  $\phi$  vanishes. In summary:

**Proposition:** If the density  $h$  of the stationary probability measure is  $C^2$ , then  $h$  satisfies the stationary Fokker-Planck equation

$$\partial_i \left[ -\frac{1}{2} \partial_j (\Sigma_{ij} h) + b_i h \right] = 0. \quad (1.20)$$

The stronger statement

$$-\frac{1}{2} \partial_j (\Sigma_{ij} h) + b_i h = 0 \quad (1.21)$$

is equivalent to symmetry of the operator  $L$  on the space  $L_h^2$ .

We are now going to relate the above symmetry statement to the detailed balance property of the stationary dynamics. First, it is clearly equivalent to the analogous property for the backward Kolmogorov semigroup:

$$\int (P_t f) g h = \int f (P_t g) h \quad (1.22)$$

since  $P_t = \exp(tL)$ . Now,  $(P_t f)(x)$  is the expected value of  $f(x_t)$  for the process, starting at  $x$  at time 0. In particular, for  $f = \delta_y$ , we obtain  $P_t f(x) = p_t(x, y)$ —the density of the transition probability from  $x$  to  $y$  in time  $t$ . Using the above symmetry of  $P_t$  with  $f = \delta_y$  and  $g = \delta_x$ , we obtain the detailed balance condition:

$$h(x) p_t(x, y) = h(y) p_t(y, x) \quad (1.23)$$

which, conversely, implies the symmetry statement for arbitrary  $f$  and  $g$ .

### 1.3. The case of a linear drift and constant noise

When both  $b(y)$  and  $\sigma(y)$  are constant or depend linearly on  $y$ , Eq. (1.1) can be solved explicitly [2] and an explicit formula for its stationary distribution can be found, when it exists. We consider the special case  $b(y) = -\gamma y$  and  $\sigma(y) \equiv \sigma$ , where  $\gamma$  and  $\sigma$  are constant matrices and the eigenvalues of  $\gamma$  have positive real parts. The stationary Fokker-Planck equation, Eq. (1.5), reads

$$\nabla \cdot \left( \frac{1}{2} \Sigma \nabla h + (\gamma y) h \right) = 0 \quad (1.24)$$

where  $\Sigma = \sigma \sigma^T$ . It has a Gaussian solution

$$h(y) = (2\pi)^{-\frac{m}{2}} (\det M)^{-\frac{1}{2}} \exp \left( -\frac{1}{2} (M^{-1} y, y) \right) \quad (1.25)$$

with the covariance matrix  $M$  which is the unique solution of the Lyapunov equation

$$\gamma M + M\gamma^T = \Sigma \quad (1.26)$$

and can be written as (see, for example, [3])

$$M = \int_0^\infty \exp(-t\gamma) \Sigma \exp(-t\gamma^T) dt. \quad (1.27)$$

This result can be verified by a direct calculation. We emphasize that it holds without assuming the detailed balance condition. The latter condition is satisfied if and only if the above  $h$  solves the equation

$$\frac{1}{2} \Sigma \nabla h + (\gamma y) h = 0 \quad (1.28)$$

which is equivalent to  $M = \frac{1}{2}\gamma^{-1}\Sigma$  or, in terms of the coefficients of the system, to

$$\Sigma \gamma^T = \gamma \Sigma \quad (1.29)$$

To see the physical significance of this condition, let us go back to the general case and write (adapting the discussion in [4] to our notation)

$$\gamma = \frac{1}{2} \Sigma M^{-1} - i\Omega. \quad (1.30)$$

$\Omega$  represents the “oscillatory degrees of freedom” of the diffusive system. The above calculations show that the detailed balance condition is equivalent to  $\Omega = 0$ , in agreement with the physical intuition that there are no macroscopic currents in the stationary state.

## 2. Small mass limit—a perturbative approach

We are now going to apply the general facts about Langevin equations to a model of a mechanical system, interacting with a noisy environment. The dynamical variables of this system are positions and momenta, and, in general, the Langevin equations which describe its time evolution, are not linear. However, when investigating the small mass limit of the system by a perturbative method, we will encounter equations closely related to those studied above. This will be explained later, when we interpret the limiting equations.

Consider a mechanical system with the Hamiltonian  $\mathcal{H}(q, p)$  where  $q, p \in \mathbb{R}^n$ . We want to study a small mass limit of this system, coupled to a damping force and the noise. Therefore, we introduce the variable  $z = \frac{p}{\sqrt{m}}$  and assume the Hamiltonian can be written  $\mathcal{H}(q, p) = H(q, z)$  where  $H$  is independent of  $m$ . We thus have

$$dq_t = \frac{1}{\sqrt{m}} \nabla_z H(q_t, z_t) dt \quad (2.1)$$

$$dz_t = -\frac{1}{\sqrt{m}} \nabla_q H(q_t, z_t) dt - \frac{1}{m} \gamma(q_t) \nabla_z H(q_t, z_t) dt + \frac{1}{\sqrt{m}} \sigma(q_t) dW_t.$$

$\gamma$  is  $n \times n$ -matrix-valued,  $\sigma$  is  $n \times k$ -matrix-valued and  $W$  is a  $k$ -dimensional Wiener process. We emphasize that  $\sigma$  does not play here the same role that

it played in our discussion of the general Langevin equation, since the noise term enters only the equation for  $dz_t$ . The number  $k$  of the components of the driving noise does not have to be related to the dimension of the system in any particular way. The corresponding backward Kolmogorov equation for a function  $\rho(q, z, t)$  is

$$\partial_t \rho = L \rho \quad (2.2)$$

where the differential operator  $L$  equals

$$L = \frac{1}{m} L_1 + \frac{1}{\sqrt{m}} L_2 \quad (2.3)$$

with

$$\begin{aligned} L_1 &= \frac{1}{2} \Sigma \nabla_z \cdot \nabla_z - \gamma \nabla_z H \nabla_z \\ L_2 &= \nabla_z H \cdot \nabla_q - \nabla_q H \cdot \nabla_z \end{aligned} \quad (2.4)$$

where  $\Sigma(q) = \sigma(q)\sigma(q)^T$ . We represent the solution of the Kolmogorov equation as a formal series

$$\rho = \rho_0 + \sqrt{m}\rho_1 + m\rho_2 + \dots \quad (2.5)$$

Equating the expressions, proportional to  $m^{-1}$ ,  $m^{-\frac{1}{2}}$  and  $m^0$ , we obtain the equations:

$$\begin{aligned} L_1 \rho_0 &= 0, \\ L_1 \rho_1 &= -L_2 \rho_0, \\ \partial_t \rho_0 &= L_1 \rho_2 + L_2 \rho_1. \end{aligned} \quad (2.6)$$

To satisfy the first equation it is sufficient to choose  $\rho_0$  which does not depend on  $z$ :

$$\rho_0 = \rho_0(q, t). \quad (2.7)$$

If we now search for  $\rho_1$  which is linear in  $z$ , the second equation simplifies to

$$\gamma \nabla_z H \cdot \nabla_z \rho_1 = \nabla_z H \cdot \nabla_q \rho_0 \quad (2.8)$$

which has a solution

$$\rho_1(q, z) = (\gamma^{-1})^T \nabla_q \rho_0 \cdot z = \nabla_q \rho_0 \cdot \gamma^{-1} z. \quad (2.9)$$

Writing the third equation as

$$\partial_t \rho_0 - L_2 \rho_1 = L_1 \rho_2 \quad (2.10)$$

and applying the identity

$$Ran L_1 = (Ker L_1^*)^\perp \quad (2.11)$$

to the space  $L^2$  with respect to the  $z$  variable, we see that  $\partial_t \rho_0 - L_2 \rho_1 = L_1 \rho_2$  must be orthogonal in this space to any function  $h$  in the null space of  $L_1^*$ . We have

$$L_1^* h = \nabla_z \cdot \left( \frac{1}{2} \Sigma \nabla_z h + (\gamma \nabla_z H) h \right) \quad (2.12)$$

where  $\Sigma = \sigma \sigma^T$ .

It is impossible to continue the analysis without further, simplifying assumptions. We are first going to study the case of a general  $H$ , assuming a form of the detailed balance condition in the variable  $z$ , at fixed  $q$ .

**Assumption 1:** for every  $q$  there exists a nonnegative solution of the equation

$$\frac{1}{2} \Sigma \nabla_z h + (\gamma \nabla_z H) h = 0 \quad (2.13)$$

of finite  $L^1(dz)$ -norm. We can thus choose

$$\int h(q, z) dz = 1. \quad (2.14)$$

We will say in this case that the system satisfies the *conditional detailed balance property* in the variable  $z$ . Since  $\rho_0$  does not depend on  $z$ , the orthogonality condition can be written as

$$\partial_t \rho_0 = \int L_2 \rho_1(q, z) h(q, z) dz. \quad (2.15)$$

We have the following explicit formula for  $L_2 \rho_1$  (summation over repeated indices is implied):

$$\begin{aligned} L_2 \rho_1 = & \partial_{z_i} H (\partial_{q_i} \partial_{q_j} \rho_0) (\gamma^{-1})_{jk} z_k + \partial_{z_i} H (\partial_{q_j} \rho_0) \partial_{q_i} ((\gamma^{-1})_{jk}) z_k \\ & - \partial_{q_i} H (\partial_{q_j} \rho_0) (\gamma^{-1})_{ji}. \end{aligned} \quad (2.16)$$

To integrate it against  $h(q, z)$ , we will use the following consequence of Eq. (2.13)

$$\begin{aligned} \int (\partial_{z_i} H) z_k h(q, z) dz &= -\frac{1}{2} \int (\gamma^{-1} \Sigma \nabla_z h)_i z_k dz \\ &= -\int (\gamma^{-1} \Sigma)_{ij} (\partial_{z_j} h) z_k dz = -\frac{1}{2} (\gamma^{-1} \Sigma)_{ij} \int (\partial_{z_j} h) z_k dz \\ &= \frac{1}{2} (\gamma^{-1} \Sigma)_{ij} \int h \delta_{jk} dz = \frac{1}{2} (\gamma^{-1} \Sigma)_{ik}. \end{aligned} \quad (2.17)$$

The orthogonality condition is thus

$$\begin{aligned} \partial_t \rho_0 = & - (\gamma^{-1})_{ji} \langle \partial_{q_i} H \rangle \partial_{q_j} \rho_0 + \frac{1}{2} (\gamma^{-1} \Sigma)_{ik} \partial_{q_i} ((\gamma^{-1})_{jk}) \partial_{q_j} \rho_0 \\ & + \frac{1}{2} (\gamma^{-1} \Sigma)_{ik} (\gamma^{-1})_{jk} (\partial_{q_i} \partial_{q_j} \rho_0). \end{aligned} \quad (2.18)$$

In this formula, which is more general than the detailed-balance case of the rigorous result of [5],  $\langle - \rangle$  denotes the average (i.e. the integral over  $z$  with the density  $h(q, z)$ ). This notation is used only in the term in which the average has not been calculated explicitly. Passing from the Kolmogorov equation to the corresponding SDE, we obtain the effective Langevin equation in the  $m \rightarrow 0$  limit:

$$dq_t = -\gamma(q_t)^{-1} (\langle \nabla_q H \rangle(q_t) + S(q_t)) dt + \gamma^{-1}(q_t) \sigma(q_t) dW_t \quad (2.19)$$

where the components of the noise-induced drift,  $S(q)$ , are given by

$$S_i(q) = \frac{1}{2} (\gamma^{-1} \Sigma)_{jk} \partial_{q_j} ((\gamma^{-1})_{ik}) \quad (2.20)$$

and we have used

$$\gamma^{-1}\sigma(\gamma^{-1}\sigma)^T = \gamma^{-1}\Sigma(\gamma^{-1})^T. \quad (2.21)$$

We are now going to interpret the limiting equation Eq. (2.19), using the stationary probability measure  $h(q, z) dz$ , as follows: from the original equations for  $q_t$  and  $z_t$  we obtain

$$dq_t = -\gamma(q_t)^{-1}\nabla_q H dt + \gamma(q_t)^{-1}\sigma(q_t) dW_t - \sqrt{m}\gamma(q_t)^{-1} dz_t. \quad (2.22)$$

Integrating the last term by parts, we obtain

$$\sqrt{m}(\gamma^{-1}(q_t))_{ij} dz_t^j = d(\sqrt{m}(\gamma_{ij}^{-1}(q_t)) z_t^j) - \sqrt{m} d((\gamma^{-1})_{ij}) z_t^j. \quad (2.23)$$

We leave the first term out, since, under fairly general natural assumptions, it is of order  $m^{\frac{1}{2}}$  [5]. The second term equals

$$-\partial_{q_k}((\gamma^{-1})_{ij})(\partial_{z_k}H) z_j dt. \quad (2.24)$$

We substitute this into the equation for  $dq_t$  and average, multiplying by  $h(q, z)$  and integrating over  $z$ . The calculation is as in Eq. (2.17) and the result is thus the same as the equation obtained by the multiscale expansion Eq. (2.19). This provides the following heuristic physical interpretation of the perturbative result: the smaller  $m$  is, the faster the variation of  $z$  becomes, and in the limit  $m \rightarrow 0$ ,  $z$  homogenizes instantaneously, with  $q$  changing only infinitesimally.

Let us now discuss conditions, under which one may expect our conditional detailed balance assumption to hold. As seen above, at fixed  $q$  this assumption is equivalent to existence of a non-negative, integrable solution of the equation

$$\frac{1}{2}\Sigma\nabla_z h + \gamma(\nabla_z H) h = 0. \quad (2.25)$$

This equation can be rewritten as

$$\frac{\nabla_z h}{h} = -2\Sigma^{-1}\gamma\nabla_z H. \quad (2.26)$$

The left-hand side equals  $\nabla_z \log h$ . Letting  $B = -2\Sigma^{-1}\gamma$  to simplify notation, we see that a necessary condition for existence of a solution is that  $B\nabla_z H$  be a gradient. This requires

$$\partial_{z_k}(b_{ij}\partial_{z_j}H) = \partial_{z_i}(b_{kj}\partial_{z_j}H) \quad (2.27)$$

for all  $i, k$ , where  $b_{ij}$  are matrix elements of  $B$ . Introducing the matrix  $R = (r_{ij})$  of second derivatives of  $H$ ,

$$r_{ij} = \partial_{z_i}\partial_{z_j}H \quad (2.28)$$

we see that solvability of Eq. (2.25) is equivalent to symmetry of the product  $BR$ :

$$BR = RB^T. \quad (2.29)$$

For the system to satisfy the conditional detailed balance property, this relation has to be satisfied for all  $q$  and  $z$ . When  $H$  is a quadratic function of  $z$ , the matrix  $R$  is constant. Even though in this case we will derive the limiting

equation without assuming conditional detailed balance, let us remark that the above approach provides a method of determining when that condition holds, different from that used earlier. Namely, let

$$H(q, z) = V(q) + \frac{1}{2} Q(q) z \cdot z \quad (2.30)$$

where  $Q(q)$  is a symmetric matrix. We then have  $R = Q$  and the solvability condition becomes

$$BQ = QB^T. \quad (2.31)$$

In a still more special—but the most frequently considered—case when  $Q$  is a multiple of identity, this reduces to

$$B = B^T \quad (2.32)$$

which is easily seen to be equivalent to the relation

$$\gamma \Sigma = \Sigma \gamma^T. \quad (2.33)$$

We have derived this condition earlier by a different argument Eq. (1.29).

If  $\gamma$  is symmetric, this becomes the commutation relation

$$\gamma \Sigma = \Sigma \gamma. \quad (2.34)$$

Note that if  $\gamma \Sigma = \Sigma \gamma^T$ , the solution of the Lyapunov equation

$$J \gamma^T + \gamma J = \Sigma \quad (2.35)$$

is given by  $J = \frac{1}{2} \gamma^{-1} \Sigma$ . In this the case the linear Langevin equation in the  $z$  variable, whose conditional equilibrium at fixed value of  $q$  we are studying, has no “oscillatory degrees of freedom”, as discussed earlier (see also [4]).

In the case when  $H$  is not a quadratic function of  $z$ , the matrix  $BR(q, z)$  has to be symmetric for all  $q$  and  $z$ , which means satisfying a continuum of conditions for every fixed  $q$ . It is interesting to ask whether there exist physically natural examples in which this happens, without each  $B(q)$  being a multiple of identity. We are not going to pursue this question here.

In the case when  $B(q)$  is a multiple of identity, we can write

$$A = 2\beta(q)^{-1} \gamma \quad (2.36)$$

with  $\beta(q)^{-1} = k_B T(q)$  and call the scalar function  $T(q)$  *generalized temperature*. The limiting Kolmogorov equation reads then

$$\partial_t \rho_0 = -(\gamma^{-1})_{ji} \langle \partial_{q_i} H \rangle \partial_{q_j} \rho_0 + k_B T \partial_{q_k} (\gamma^{-1})_{jk} \partial_{q_j} \rho_0 + k_B T (\gamma^{-1})_{ij} (\partial_{q_i} \partial_{q_j} \rho_0) \quad (2.37)$$

and the components of the noise-induced drift are thus

$$S_j(q) = k_B T \partial_{q_k} (\gamma^{-1})_{jk}. \quad (2.38)$$

The above applies in particular in the one-dimensional case, in which  $\sigma(q)^2$  and  $\gamma(q)$  are scalars and hence one is always an ( $q$ -dependent) multiple of the other:

$$k_B T(q) = \frac{\sigma(q)^2}{2\gamma(q)}. \quad (2.39)$$

The limiting Langevin equation is in this case

$$dq_t = -\frac{\langle \nabla_q H \rangle}{\gamma} dt - \frac{1}{2} \frac{\nabla_q \gamma}{\gamma^3} \sigma^2 dt + \frac{\sigma}{\gamma} dW_t. \quad (2.40)$$

For a Hamiltonian equal to a sum of potential and quadratic kinetic energy,  $H = V(q) + \frac{z^2}{2}$ , the first term equals  $\frac{F}{\gamma}$ , where  $F = -\nabla_q V dt$ , in agreement with earlier results.

The second situation, in which the perturbative treatment of the original system can be carried out explicitly is the general quadratic kinetic energy case.

**Assumption 2:**  $H = V(q) + \frac{z^2}{2}$

If we follow the singular perturbation method used above, we again need to find the integral Eq. (2.17), where  $\partial_{z_i} H = z_i$ . In this case we know the solution of  $L_1^* h = 0$  explicitly:

$$h(q, z) = (2\pi)^{-\frac{n}{2}} (\det M)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} M^{-1} z \cdot z\right) \quad (2.41)$$

so the integral in Eq. (2.17) is the mean of  $z_i z_k$  in the Gaussian distribution with the covariance  $M = (m_{ik})$ , that is,  $m_{ik}$ . The second-order term in the Kolmogorov equation is thus  $m_{ik} (\gamma^{-1})_{jk} \partial_{q_i} \partial_{q_j} \rho_0$ . The corresponding Langevin equation, which has been derived rigorously in [5] is in this case

$$dq_t = -\gamma^{-1}(q_t) \nabla_q V(q_t) dt + S(q_t) dt + \gamma^{-1}(q_t) \sigma(q_t) dW_t. \quad (2.42)$$

The homogenization heuristics proposed under Assumption 1 applies here as well: the limiting Langevin equation can be interpreted as a result of averaging over the conditional stationary distribution of the  $z$  variable. A rigorous result, corroborating this picture has recently been proven in [6].

### 3. A rigorous homogenization theorem

We now develop a framework for the homogenization of Langevin equations that is able to make many of the heuristic results from the previous two sections rigorous. Our results will concern Hamiltonians of the form

$$H(t, x) = K(t, q, p - \psi(t, q)) + V(t, q) \quad (3.1)$$

where  $x = (q, p) \in \mathbb{R}^n \times \mathbb{R}^n$ ,  $K = K(t, q, z)$  and  $V = V(t, q)$  are  $C^2$ ,  $\mathbb{R}$ -valued functions,  $K$  is non-negative, and  $\psi$  is a  $C^2$ ,  $\mathbb{R}^n$ -valued function. The splitting of  $H$  into  $K$  and  $V$  does not have to correspond physically to any notion of kinetic and potential energy, although we will use those terms for convenience. The splitting is not unique; it will be constrained further as we continue. We now define the family of scaled Hamiltonians, parameterized by  $\epsilon > 0$  (generalizing the above mass parameter):

$$H^\epsilon(t, q, p) \equiv K^\epsilon(t, q, p) + V(t, q) \equiv K(t, q, (p - \psi(t, q))/\sqrt{\epsilon}) + V(t, q). \quad (3.2)$$

Consider the following family of SDEs:

$$dq_t^\epsilon = \nabla_p H^\epsilon(t, x_t^\epsilon) dt, \quad (3.3)$$

$$dp_t^\epsilon = (-\gamma(t, x_t^\epsilon) \nabla_p H^\epsilon(t, x_t^\epsilon) - \nabla_q H^\epsilon(t, x_t^\epsilon) + F(t, x_t^\epsilon)) dt + \sigma(t, x_t^\epsilon) dW_t, \quad (3.4)$$

where  $\gamma : [0, \infty) \times \mathbb{R}^{2n} \rightarrow \mathbb{R}^{n \times n}$  and  $\sigma : [0, \infty) \times \mathbb{R}^{2n} \rightarrow \mathbb{R}^{n \times k}$  are continuous,  $\gamma$  is positive definite, 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 [7].

Our objective in this section is to develop a method for investigating the behavior of  $x_t^\epsilon$  in the limit  $\epsilon \rightarrow 0^+$ ; more precisely, we wish to prove the existence of a limiting “position” process  $q_t$  and derive a homogenized SDE that it satisfies. In fact, the method we develop is applicable to a more general class SDEs that share certain properties with Eq. (3.3)-Eq. (3.4). In the following subsection, we discuss some prior results concerning Eq. (3.3)-Eq. (3.4). This will help motivate the assumptions made in the development of our general homogenization method, starting in subsection .

### 3.1. Summary of prior results

Let  $x_t^\epsilon$  be a family of solutions to the SDE Eq. (3.3)-Eq. (3.4) with initial condition  $x_0^\epsilon = (q_0^\epsilon, p_0^\epsilon)$ . We assume that a solution exists for all  $t \geq 0$  (i.e. there are no explosions). See Appendix B in [8] for assumptions that guarantee this. Under Assumptions 1-3 in [8] (repeated as Assumptions A1-A3 in A, we showed that for any  $T > 0$ ,  $p > 0$ ,  $0 < \beta < p/2$  we have

$$\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^{p/2}) \text{ and } E\left[\sup_{t \in [0, T]} \|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p\right] = O(\epsilon^\beta) \quad (3.5)$$

as  $\epsilon \rightarrow 0^+$  i.e. the point  $(p, q)$  is attracted to the surface defined by  $p = \psi(t, q)$ .

Adding Assumption 4 (Assumption Eq. (A4) in the appendix) we also showed that

$$\begin{aligned} d(q_t^\epsilon)^i &= (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon)(-\partial_t \psi_j(t, q_t^\epsilon) - \partial_{q^j} V(t, q_t^\epsilon) + F_j(t, x_t^\epsilon)) dt \\ &\quad + (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \sigma_{j\rho}(t, x_t^\epsilon) dW_t^\rho - (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{q^j} K(t, q_t^\epsilon, z_t^\epsilon) dt \\ &\quad + (z_t^\epsilon)_j \partial_{q^l} (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{z_l} K(t, q_t^\epsilon, z_t^\epsilon) dt - d((\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon)(u_t^\epsilon)_j) \\ &\quad + (u_t^\epsilon)_j \partial_t (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) dt, \end{aligned} \quad (3.6)$$

where  $u_t^\epsilon \equiv p_t^\epsilon - \psi(t, q_t^\epsilon)$ ,  $z_t^\epsilon \equiv u_t^\epsilon / \sqrt{\epsilon}$ , and

$$\tilde{\gamma}_{ik}(t, q) \equiv \gamma_{ik}(t, q) + \partial_{q^k} \psi_i(t, q) - \partial_{q^i} \psi_k(t, q). \quad (3.7)$$

We define the components of  $\tilde{\gamma}^{-1}$  such that

$$(\tilde{\gamma}^{-1})^{ij} \tilde{\gamma}_{jk} = \delta_k^i, \quad (3.8)$$

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

Under the additional Assumptions 5-7 in [8], which include further restrictions on the form of the Hamiltonian, we were then able to show that

$q_t^\epsilon$  converges in an  $L^p$ -norm as  $\epsilon \rightarrow 0^+$  to the solution of a lower dimensional SDE,

$$\begin{aligned} dq_t &= \tilde{\gamma}^{-1}(t, q_t)(-\partial_t \psi(t, q_t) - \nabla_q V(t, q_t) + F(t, q_t, \psi(t, q_t)))dt + S(t, q_t)dt \\ &\quad + \tilde{\gamma}^{-1}(t, q_t)\sigma(t, q_t, \psi(t, q_t))dW_t. \end{aligned} \quad (3.9)$$

The *noise-induced drift* term,  $S(t, q)$ , that arises in the limit is the term of greatest interest here. Its form is given in Eq. (3.26) in [8].

The homogenization technique used in [8] to arrive at Eq. (4.21) relies heavily on the specific structural assumptions on the form of the Hamiltonian. Those assumptions cover a wide variety of important systems, such as a particle in an electromagnetic field, and motion on a Riemannian manifold, but it is desirable to search for a more generally applicable homogenization method. In this paper, we develop a significantly more general technique, adapted from the methods presented in [9], that is capable of homogenizing terms of the form  $G(t, q_t^\epsilon, (p_t^\epsilon - \psi(t, q_t^\epsilon))/\sqrt{\epsilon})dt$  for a general class of SDEs that satisfy the property Eq. (3.5), as well as prove convergence of  $q_t^\epsilon$  to the solution of a limiting, homogenized SDE. In particular, it will be capable of homogenizing  $q_t^\epsilon$  from the Hamiltonian system Eq. (3.3)-Eq. (3.4) under less restrictive assumptions on the form of the Hamiltonian, than those made in [8]. We emphasize that the convergence statements are proven in the strong sense, see Section 3.2.

### 3.2. General homogenization framework

Here we describe our homogenization technique in a more general context than the Hamiltonian setting from the previous section. This method is related to the cell problem method from [9], our proof applies to a larger class of SDEs and demonstrates  $L^p$ -convergence rather than weak convergence.

We will denote an element of  $\mathbb{R}^n \times \mathbb{R}^m$  by  $x = (q, p)$ , where we no longer require the  $q$  and  $p$  degrees of freedom to have the same dimensionality, though we still employ the convention of writing  $q$  indices with superscripts and  $p$  indices with subscripts. We let  $W_t$  be an  $\mathbb{R}^k$ -valued Wiener process,  $\psi : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^m$  be  $C^2$  and  $G_1, F_1 : [0, \infty) \times \mathbb{R}^{n+m} \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $G_2, F_2 : [0, \infty) \times \mathbb{R}^{n+m} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ ,  $\sigma_1 : [0, \infty) \times \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n \times k}$ , and  $\sigma_2 : [0, \infty) \times \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{m \times k}$  be continuous. With these definitions, we consider the following family of SDEs, depending on a parameter  $\epsilon > 0$ :

$$dq_t^\epsilon = \left( \frac{1}{\sqrt{\epsilon}} G_1(t, x_t^\epsilon, z_t^\epsilon) + F_1(t, x_t^\epsilon, z_t^\epsilon) \right) dt + \sigma_1(t, x_t^\epsilon) dW_t, \quad (3.10)$$

$$dp_t^\epsilon = \left( \frac{1}{\sqrt{\epsilon}} G_2(t, x_t^\epsilon, z_t^\epsilon) + F_2(t, x_t^\epsilon, z_t^\epsilon) \right) dt + \sigma_2(t, x_t^\epsilon) dW_t, \quad (3.11)$$

where we define  $z_t^\epsilon = (p_t^\epsilon - \psi(t, q_t^\epsilon))/\sqrt{\epsilon}$ . We will assume, in analogy with Eq. (3.5), that:

**Assumption 3.1.** For any  $T > 0$ ,  $p > 0$ ,  $0 < \beta < p/2$  we have

$$\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^{p/2}) \text{ and } E\left[\sup_{t \in [0, T]} \|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p\right] = O(\epsilon^\beta) \quad (3.12)$$

as  $\epsilon \rightarrow 0^+$ .

In words, we assume that the  $p$  degrees of freedom are attracted to the values defined by  $p = \psi(t, q)$ . This is an appropriate setting to expect some form of homogenization, as it suggests that the dynamics in the limit  $\epsilon \rightarrow 0^+$  can be characterized by fewer degrees of freedom—the  $q$ -variables.

**3.2.1. Homogenization of integral processes.** In this section we derive a method capable of homogenizing processes of the form

$$M_t^\epsilon \equiv \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds \quad (3.13)$$

in the limit  $\epsilon \rightarrow 0^+$ . More specifically, our aim is to find conditions under which there exists some function,  $S(t, q)$ , such that

$$\int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t S(s, q_s^\epsilon) ds \rightarrow 0 \quad (3.14)$$

in some norm, as  $\epsilon \rightarrow 0^+$ , i.e. only the  $q$ -degrees of freedom are needed to characterize  $M_t^\epsilon$  in the limit. We will call a family of processes,  $S(t, q_t^\epsilon) dt$ , that satisfies such a limit, a homogenization of  $G(t, x_t^\epsilon, z_t^\epsilon) dt$ . The technique we develop will also be useful for proving existence of a limiting process  $q_s$  (i.e.  $q_s^\epsilon \rightarrow q_s$ ), and showing that

$$\int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds \rightarrow \int_0^t S(s, q_s) ds. \quad (3.15)$$

as  $\epsilon \rightarrow 0^+$ . We will consider this second question in Section 3.2.3. Here, our focus is on Eq. (3.14).

As a starting point, let  $\chi(t, x, z) : [0, \infty) \times \mathbb{R}^{n+m} \times \mathbb{R}^m \rightarrow \mathbb{R}$  be  $C^{1,2}$ , where  $C^{1,2}$  is defined as follows:

- If  $\sigma_1 \neq 0$  then we take this to mean  $\chi$  is  $C^1$  and, for each  $t$ ,  $\chi(t, x, z)$  is  $C^2$  in  $(x, z)$  with second derivatives continuous jointly in all variables.
- If  $\sigma_1 = 0$  then we take this to mean  $\chi$  is  $C^1$  and, for each  $t, q$ ,  $\chi(t, q, p, z)$  is  $C^2$  in  $(p, z)$  with second derivatives continuous jointly in all variables.

Eventually, we will need to carefully choose  $\chi$  so that we achieve our aim, but for now we simply use Itô's formula to compute  $\chi(t, x_t^\epsilon, z_t^\epsilon)$ . We defined  $C^{1,2}$  precisely so that Itô's formula is justified. For this computation, we will define  $\chi^\epsilon(t, x) = \chi(t, x, (p - \psi(t, q)) / \sqrt{\epsilon})$ , and

$$\Sigma_{11}^{ij} = \sum_{\rho} (\sigma_1)_{\rho}^i (\sigma_1)_{\rho}^j, \quad (\Sigma_{12})_j^i = \sum_{\rho} (\sigma_1)_{\rho}^i (\sigma_2)_{j\rho}, \quad (\Sigma_{22})_{ij} = \sum_{\rho} (\sigma_2)_{i\rho} (\sigma_2)_{j\rho}. \quad (3.16)$$

Itô's formula gives

$$\begin{aligned}
 \chi(t, x_t^\epsilon, z_t^\epsilon) &= \chi(0, x_0^\epsilon, z_0^\epsilon) + \int_0^t \partial_s \chi^\epsilon(s, x_s^\epsilon) ds + \int_0^t \nabla_q \chi^\epsilon(s, x_s^\epsilon) \cdot dq_s^\epsilon \quad (3.17) \\
 &+ \int_0^t \nabla_p \chi^\epsilon(s, x_s^\epsilon) \cdot dp_s^\epsilon + \frac{1}{2} \int_0^t \partial_{q^i} \partial_{q^j} \chi^\epsilon(s, x_s^\epsilon) \Sigma_{11}^{ij}(s, x_s^\epsilon) ds \\
 &+ \frac{1}{2} \int_0^t \partial_{q^i} \partial_{p_j} \chi^\epsilon(s, x_s^\epsilon) (\Sigma_{12})_j^i(s, x_s^\epsilon) ds + \frac{1}{2} \int_0^t \partial_{p_i} \partial_{q^j} \chi^\epsilon(s, x_s^\epsilon) (\Sigma_{12})_i^j(s, x_s^\epsilon) ds \\
 &+ \frac{1}{2} \int_0^t \partial_{p_i} \partial_{p_j} \chi^\epsilon(s, x_s^\epsilon) (\Sigma_{22})_{ij}(s, x_s^\epsilon) ds.
 \end{aligned}$$

Note that if  $\sigma_1 = 0$  then only the second derivatives that we have assumed exist are involved in this computation.

We can compute these terms as follows:

$$\partial_t \chi^\epsilon(t, x) = \partial_t \chi(t, x, z) - \partial_{z_i} \chi(t, x, z) \partial_t \psi_i(t, q) / \sqrt{\epsilon}, \quad (3.18)$$

$$\partial_{q^i} \chi^\epsilon(t, x) = (\partial_{q^i} \chi)(t, x, z) - \epsilon^{-1/2} \partial_{q^i} \psi_k(t, q) (\partial_{z_k} \chi)(t, x, z), \quad (3.19)$$

$$\partial_{p_i} \chi^\epsilon(t, x) = (\partial_{p_i} \chi)(t, x, z) + \epsilon^{-1/2} (\partial_{z_i} \chi)(t, x, z), \quad (3.20)$$

$$\begin{aligned}
 \partial_{q^i} \partial_{q^j} \chi^\epsilon(t, x) &= (\partial_{q^i} \partial_{q^j} \chi)(t, x, z) + \epsilon^{-1/2} (-\partial_{q^j} \psi_k(t, q) (\partial_{q^i} \partial_{z_k} \chi)(t, x, z) \\
 &\quad - \partial_{q^i} \partial_{q^j} \psi_k(t, q) (\partial_{z_k} \chi)(t, x, z) - \partial_{q^i} \psi_k(t, q) (\partial_{q^j} \partial_{z_k} \chi)(t, x, z)) \quad (3.21)
 \end{aligned}$$

$$\begin{aligned}
 \partial_{p_i} \partial_{p_j} \chi^\epsilon(t, x) &= (\partial_{p_i} \partial_{p_j} \chi)(t, x, z) + \epsilon^{-1/2} ((\partial_{z_j} \partial_{p_i} \chi)(t, x, z) \\
 &\quad + (\partial_{p_j} \partial_{z_i} \chi)(t, x, z)) + \epsilon^{-1} (\partial_{z_i} \partial_{z_j} \chi)(t, x, z), \quad (3.22)
 \end{aligned}$$

$$\begin{aligned}
 \partial_{q^i} \partial_{p_j} \chi^\epsilon(t, x) &= (\partial_{q^i} \partial_{p_j} \chi)(t, x, z) + \epsilon^{-1/2} ((\partial_{q^i} \partial_{z_j} \chi)(t, x, z) \\
 &\quad - \partial_{q^i} \psi_k(t, q) (\partial_{p_j} \partial_{z_k} \chi)(t, x, z)) \\
 &\quad - \epsilon^{-1} \partial_{q^i} \psi_k(t, q) (\partial_{z_j} \partial_{z_k} \chi)(t, x, z), \quad (3.23)
 \end{aligned}$$

where  $z$  is evaluated at  $z(t, x, \epsilon) = (p - \psi(t, q)) / \sqrt{\epsilon}$  in each of the above formulae.

Using these expressions, together with the SDE Eq. (3.10)-Eq. (3.11) we find

$$\begin{aligned}
\chi(t, x_t^\epsilon, z_t^\epsilon) &= \chi(0, x_0^\epsilon, z_0^\epsilon) + \int_0^t \partial_s \chi(s, x_s^\epsilon, z_s^\epsilon) - \epsilon^{-1/2} (\partial_{z_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) \partial_s \psi_i(s, q_s^\epsilon) ds \\
&+ \int_0^t \left( (\partial_{q^i} \chi)(s, x_s^\epsilon, z_s^\epsilon) - \epsilon^{-1/2} \partial_{q^i} \psi_k(s, q_s^\epsilon) (\partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right) \\
&\quad \times \left[ \left( \frac{1}{\sqrt{\epsilon}} G_1(s, x_s^\epsilon, z_s^\epsilon) + F_1(s, x_s^\epsilon, z_s^\epsilon) \right) ds + \sigma_1(s, x_s^\epsilon) dW_s \right]^i \\
&+ \int_0^t \left( (\partial_{p_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) + \epsilon^{-1/2} (\partial_{z_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right) \\
&\quad \times \left[ \left( \frac{1}{\sqrt{\epsilon}} G_2(s, x_s^\epsilon, z_s^\epsilon) + F_2(s, x_s^\epsilon, z_s^\epsilon) \right) ds + \sigma_2(s, x_s^\epsilon) dW_s \right]_i \\
&+ \frac{1}{2} \int_0^t \Sigma_{11}^{ij}(s, x_s^\epsilon) \left[ (\partial_{q^i} \partial_{q^j} \chi)(s, x_s^\epsilon, z_s^\epsilon) + \epsilon^{-1/2} \left( -\partial_{q^j} \psi_k(s, q_s^\epsilon) (\partial_{q^i} \partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right. \right. \\
&\quad \left. \left. - \partial_{q^i} \partial_{q^j} \psi_k(s, q_s^\epsilon) (\partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon) - \partial_{q^i} \psi_k(s, q_s^\epsilon) (\partial_{q^j} \partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right) \right. \\
&\quad \left. + \epsilon^{-1} \partial_{q^i} \psi_k(s, q_s^\epsilon) \partial_{q^j} \psi_l(s, q_s^\epsilon) (\partial_{z_k} \partial_{z_l} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right] ds \\
&+ \int_0^t (\Sigma_{12})_j^i(s, x_s^\epsilon) \left[ (\partial_{q^i} \partial_{p_j} \chi)(s, x_s^\epsilon, z_s^\epsilon) + \epsilon^{-1/2} \left( (\partial_{q^i} \partial_{z_j} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right. \right. \\
&\quad \left. \left. - \partial_{q^i} \psi_k(t, q) (\partial_{p_j} \partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right) - \epsilon^{-1} \partial_{q^i} \psi_k(s, q_s^\epsilon) (\partial_{z_j} \partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right] ds \\
&+ \frac{1}{2} \int_0^t (\Sigma_{22})_{ij}(s, x_s^\epsilon) \left[ (\partial_{p_i} \partial_{p_j} \chi)(s, x_s^\epsilon, z_s^\epsilon) + \epsilon^{-1/2} \left( (\partial_{z_j} \partial_{p_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right. \right. \\
&\quad \left. \left. + (\partial_{p_j} \partial_{z_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right) + \epsilon^{-1} (\partial_{z_i} \partial_{z_j} \chi)(s, x_s^\epsilon, z_s^\epsilon) \right] ds.
\end{aligned} \tag{3.24}$$

Multiplying by  $\epsilon$  and collecting powers, we arrive at

$$\int_0^t (L\chi)(s, x_s^\epsilon, z_s^\epsilon) ds = \epsilon^{1/2} (R_1^\epsilon)_t + \epsilon (\chi(t, x_t^\epsilon, z_t^\epsilon) - \chi(0, x_0^\epsilon, z_0^\epsilon) + (R_2^\epsilon)_t), \tag{3.25}$$

where we define

$$\begin{aligned}
(L\chi)(t, x, z) &= \left( \frac{1}{2} \Sigma_{11}^{ij}(t, x) \partial_{q^i} \psi_k(t, q) \partial_{q^j} \psi_l(t, q) \right. \\
&\quad \left. - (\Sigma_{12})_l^i(t, x) \partial_{q^i} \psi_k(t, q) + \frac{1}{2} (\Sigma_{22})_{kl}(t, x) \right) (\partial_{z_k} \partial_{z_l} \chi)(t, x, z) \\
&+ ((G_2)_k(t, x, z) - \partial_{q^i} \psi_k(t, q) G_1^i(t, x, z)) (\partial_{z_k} \chi)(t, x, z),
\end{aligned} \tag{3.26}$$

$$\begin{aligned}
& (R_1^\epsilon)_t & (3.27) \\
& = \int_0^t (\partial_{z_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) \partial_s \psi_i(s, q_s^\epsilon) ds - \int_0^t (\partial_{q^i} \chi)(s, x_s^\epsilon, z_s^\epsilon) G_1^i(s, x_s^\epsilon, z_s^\epsilon) ds \\
& + \int_0^t \partial_{q^i} \psi_k(s, q_s^\epsilon) (\partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon) [F_1(s, x_s^\epsilon, z_s^\epsilon) ds + \sigma_1(s, x_s^\epsilon) dW_s]^i \\
& - \int_0^t (\partial_{z_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) [F_2(s, x_s^\epsilon, z_s^\epsilon) ds + \sigma_2(s, x_s^\epsilon) dW_s]_i \\
& - \int_0^t (\partial_{p_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) (G_2)_i(s, x_s^\epsilon, z_s^\epsilon) ds \\
& - \frac{1}{2} \int_0^t \Sigma_{11}^{ij}(s, x_s^\epsilon) (-\partial_{q^j} \psi_k(s, q_s^\epsilon) (\partial_{q^i} \partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon) \\
& - \partial_{q^i} \partial_{q^j} \psi_k(s, q_s^\epsilon) (\partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon) - \partial_{q^i} \psi_k(s, q_s^\epsilon) (\partial_{q^j} \partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon)) ds \\
& - \int_0^t (\Sigma_{12})_j^i(s, x_s^\epsilon) ((\partial_{q^i} \partial_{z_j} \chi)(s, x_s^\epsilon, z_s^\epsilon) - \partial_{q^i} \psi_k(t, q) (\partial_{p_j} \partial_{z_k} \chi)(s, x_s^\epsilon, z_s^\epsilon)) ds \\
& - \int_0^t (\Sigma_{22})_{ij}(s, x_s^\epsilon) (\partial_{z_j} \partial_{p_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) ds,
\end{aligned}$$

and

$$\begin{aligned}
& (R_2^\epsilon)_t & (3.28) \\
& = - \int_0^t \partial_s \chi(s, x_s^\epsilon, z_s^\epsilon) ds \\
& - \int_0^t (\partial_{q^i} \chi)(s, x_s^\epsilon, z_s^\epsilon) [F_1(s, x_s^\epsilon, z_s^\epsilon) ds + \sigma_1(s, x_s^\epsilon) dW_s]^i \\
& - \int_0^t (\partial_{p_i} \chi)(s, x_s^\epsilon, z_s^\epsilon) [F_2(s, x_s^\epsilon, z_s^\epsilon) ds + \sigma_2(s, x_s^\epsilon) dW_s]_i \\
& - \frac{1}{2} \int_0^t \Sigma_{11}^{ij}(s, x_s^\epsilon) (\partial_{q^i} \partial_{q^j} \chi)(s, x_s^\epsilon, z_s^\epsilon) ds \\
& - \int_0^t (\Sigma_{12})_j^i(s, x_s^\epsilon) (\partial_{q^i} \partial_{p_j} \chi)(s, x_s^\epsilon, z_s^\epsilon) ds \\
& - \frac{1}{2} \int_0^t (\Sigma_{22})_{ij}(s, x_s^\epsilon) (\partial_{p_i} \partial_{p_j} \chi)(s, x_s^\epsilon, z_s^\epsilon) ds.
\end{aligned}$$

First, think of simply homogenizing Eq. (3.13) to a quantity of the form  $\int_0^t \tilde{G}(s, x_s^\epsilon) ds$ . Suppose we have a candidate for  $\tilde{G}$ . If we can find a  $C^{1,2}$  solution,  $\chi$ , to the PDE

$$(L\chi)(t, x, z) = G(t, x, z) - \tilde{G}(t, x) \quad (3.29)$$

then substituting this into Eq. (3.25) gives

$$\begin{aligned}
& \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, x_s^\epsilon) ds & (3.30) \\
& = \epsilon^{1/2} (R_1^\epsilon)_t + \epsilon (\chi(t, x_t^\epsilon, z_t^\epsilon) - \chi(0, x_0^\epsilon, z_0^\epsilon) + (R_2^\epsilon)_t).
\end{aligned}$$

Given sufficient growth bounds for  $\chi$  and its derivatives, one anticipates that the right hand side of Eq. (3.30) vanishes in the limit. If in addition,  $\tilde{G}$  is

Lipschitz in  $p$ , uniformly in  $(t, q)$ , then, based on Assumption 3.1, one expects

$$\int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, q_s^\epsilon, \psi(s, q_s^\epsilon)) ds \rightarrow 0 \quad (3.31)$$

as  $\epsilon \rightarrow 0^+$ .

We make this informal discussion precise in Theorem 3.1, below. For this, we will need the following assumptions:

**Assumption 3.2.** *For all  $T > 0$ , the following quantities are polynomially bounded in  $z$ , with the bounds uniform on  $[0, T] \times \mathbb{R}^{n+m}$ :*

*$G_1, F_1, G_2, F_2, \sigma_1, \sigma_2, \partial_t \psi, \partial_{q^i} \psi, \partial_{q^i} \partial_{q^j} \psi$ . If  $\sigma_1 = 0$  then we can remove the requirement on  $\partial_{q^i} \partial_{q^j} \psi$ .*

Recall that an  $\mathbb{R}^l$ -valued function,  $\phi(t, x, z)$ , is called *polynomially bounded* in  $z$ , uniformly on  $[0, T] \times \mathbb{R}^{n+m}$  if there exists  $q, C > 0$  such that

$$\|\phi(t, x, z)\| \leq C(1 + \|z\|^q) \quad (3.32)$$

for all  $(t, x, z) \in [0, T] \times \mathbb{R}^{n+m} \times \mathbb{R}^m$ . In particular, if  $\phi$  is independent of  $z$ , this just means it is bounded on  $[0, T] \times \mathbb{R}^{n+m}$ . Applying this to  $\psi$ , we note that Assumption 3.2 implies  $\psi$  is Lipschitz in  $q$ , uniformly in  $t \in [0, T]$ .

**Assumption 3.3.** *Given a continuous  $G : [0, \infty) \times \mathbb{R}^{n+m} \times \mathbb{R}^m \rightarrow \mathbb{R}$ , assume that there exists a  $C^{1,2}$  function  $\chi : [0, \infty) \times \mathbb{R}^{n+m} \times \mathbb{R}^m \rightarrow \mathbb{R}$  and a continuous function  $\tilde{G}(t, x) : [0, \infty) \times \mathbb{R}^{n+m} \rightarrow \mathbb{R}$  that together satisfy the PDE*

$$(L\chi)(t, x, z) = G(t, x, z) - \tilde{G}(t, x), \quad (3.33)$$

where the differential operator,  $L$ , is defined in Eq. (3.26).

Assume that, for a given  $T > 0$ ,  $\tilde{G}$  is Lipschitz in  $p$ , uniformly for  $(t, q) \in [0, T] \times \mathbb{R}^n$ . Also suppose that  $\chi$ , its first derivatives, and the second derivatives  $\partial_{q^i} \partial_{q^j} \chi, \partial_{q^i} \partial_{p_j} \chi, \partial_{q^i} \partial_{z_j} \chi, \partial_{p_i} \partial_{p_j} \chi$ , and  $\partial_{p_i} \partial_{z_j} \chi$  are polynomially bounded in  $z$ , uniformly for  $(t, x) \in [0, T] \times \mathbb{R}^{n+m}$ . If  $\sigma_1 = 0$  then the only second derivatives that we require to be polynomially bounded are  $\partial_{p_i} \partial_{p_j} \chi$  and  $\partial_{p_i} \partial_{z_j} \chi$ .

**Theorem 3.1.** *Fix  $T > 0$ . Let Assumptions 3.1-3.3 hold and  $x_t^\epsilon = (q_t^\epsilon, p_t^\epsilon)$  satisfy the SDE Eq. (3.10)-Eq. (3.11). Then for any  $p > 0$  we have*

$$E \left[ \sup_{t \in [0, T]} \left| \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, q_s^\epsilon, \psi(s, q_s^\epsilon)) ds \right|^p \right] = O(\epsilon^{p/2}). \quad (3.34)$$

and

$$E \left[ \sup_{t \in [0, T]} \left| \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, q_s^\epsilon, \psi(s, q_s^\epsilon)) ds \right|^p \right] = O(\epsilon^{p/2}) \quad (3.35)$$

as  $\epsilon \rightarrow 0^+$ .

*Proof.* Fix  $T > 0$ . First let  $p \geq 2$ . Eq. (3.30) gives

$$\begin{aligned} & E \left[ \sup_{t \in [0, T]} \left| \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, x_s^\epsilon) ds \right|^p \right] \\ & \leq 3^{p-1} \left( \epsilon^{p/2} E \left[ \sup_{t \in [0, T]} |(R_1^\epsilon)_t|^p \right] + 2^p \epsilon^p E \left[ \sup_{t \in [0, T]} |\chi(t, x_t^\epsilon, z_t^\epsilon)|^p \right] \right. \\ & \quad \left. + \epsilon^p E \left[ \sup_{t \in [0, T]} |(R_2^\epsilon)_t|^p \right] \right). \end{aligned} \quad (3.36)$$

From Eq. (3.27) and Eq. (3.28) we see that  $R_1^\epsilon$  and  $R_2^\epsilon$  have the forms

$$(R_i^\epsilon)_t = \int_0^t V_i(s, x_s^\epsilon, z_s^\epsilon) ds + \int_0^t Q_{ij}(s, x_s^\epsilon, z_s^\epsilon) dW_s^j, \quad (3.37)$$

where  $V_i$  and  $Q_{ij}$  are linear combinations of products of (components of) one or more terms from the following list:

$G_1, F_1, G_2, F_2, \sigma_1, \sigma_2, \partial_t \psi, \partial_{q^i} \psi, \partial_{q^i} \partial_{q^j} \psi, \partial_t \chi, \partial_{q^i} \chi, \partial_{z_i} \chi, \partial_{p_i} \chi, \partial_{q^i} \partial_{q^j} \chi, \partial_{q^i} \partial_{p_j} \chi, \partial_{q^i} \partial_{z_j} \chi, \partial_{p_i} \partial_{p_j} \chi, \partial_{p_i} \partial_{z_j} \chi$ . Also note that if  $\sigma_1 = 0$  then the only second derivatives terms that are involved are  $\partial_{p_i} \partial_{p_j} \chi$  and  $\partial_{p_i} \partial_{z_j} \chi$ .

By assumption, these are all polynomially bounded in  $z$ , uniformly on  $[0, T] \times \mathbb{R}^{n+m}$ , as is  $\chi$ . Therefore, letting  $\tilde{C}$  denote a constant that potentially varies line to line, there exists  $r > 0$  such that

$$\begin{aligned} & E \left[ \sup_{t \in [0, T]} \left| \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, x_s^\epsilon) ds \right|^p \right] \\ & \leq \tilde{C} \epsilon^{p/2} \left( E \left[ \left( \int_0^T |V_1(s, x_s^\epsilon, z_s^\epsilon)| ds \right)^p \right] + E \left[ \sup_{t \in [0, T]} \left| \int_0^t Q_{1j}(s, x_s^\epsilon, z_s^\epsilon) dW_s^j \right|^p \right] \right) \\ & \quad + \tilde{C} \epsilon^p \left( E \left[ \left( \int_0^T |V_2(s, x_s^\epsilon, z_s^\epsilon)| ds \right)^p \right] + E \left[ \sup_{t \in [0, T]} \left| \int_0^t Q_{2j}(s, x_s^\epsilon, z_s^\epsilon) dW_s^j \right|^p \right] \right. \\ & \quad \left. + 1 + E \left[ \sup_{t \in [0, T]} \|z_t^\epsilon\|^{rp} \right] \right). \end{aligned} \quad (3.38)$$

Hölder's inequality and polynomial boundedness yields

$$\begin{aligned} & E \left[ \left( \int_0^T |V_i(s, x_s^\epsilon, z_s^\epsilon)| ds \right)^p \right] \leq T^{p-1} E \left[ \int_0^T |V_i(s, x_s^\epsilon, z_s^\epsilon)|^p ds \right] \\ & \leq \tilde{C} T^p \left( 1 + \sup_{t \in [0, T]} E [\|z_t^\epsilon\|^{rp}] \right). \end{aligned} \quad (3.39)$$

Applying the Burkholder-Davis-Gundy inequality to the terms involving  $Q_{ij}$ , (as found in, for example, Theorem 3.28 in [7]), and then Hölder's inequality,

we obtain

$$\begin{aligned}
& E \left[ \sup_{t \in [0, T]} \left| \int_0^t Q_{ij}(s, x_s^\epsilon, z_s^\epsilon) dW_s^j \right|^p \right] \\
& \leq \tilde{C} E \left[ \left( \int_0^T \|Q_i(s, x_s^\epsilon, z_s^\epsilon)\|^2 ds \right)^{p/2} \right] \\
& \leq \tilde{C} T^{p/2-1} E \left[ \int_0^T \|Q_i(s, x_s^\epsilon, z_s^\epsilon)\|^p ds \right] \\
& \leq \tilde{C} T^{p/2} \left( 1 + \sup_{t \in [0, T]} E[\|z_t^\epsilon\|^{rp}] \right).
\end{aligned} \tag{3.40}$$

Combining these bounds, and using Assumption 3.1, we find

$$\begin{aligned}
& E \left[ \sup_{t \in [0, T]} \left| \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, x_s^\epsilon) ds \right|^p \right] \\
& \leq \tilde{C} \epsilon^{p/2} \left( 1 + \sup_{t \in [0, T]} E[\|z_t^\epsilon\|^{rp}] \right) \\
& \quad + \tilde{C} \epsilon^p \left( 1 + \sup_{t \in [0, T]} E[\|z_t^\epsilon\|^{rp}] + E \left[ \sup_{t \in [0, T]} \|z_t^\epsilon\|^{rp} \right] \right) \\
& \leq \tilde{C} \epsilon^{p/2} (1 + O(1)) + \tilde{C} \epsilon^p (1 + O(1) + O(\epsilon^{-\delta}))
\end{aligned} \tag{3.41}$$

for any  $\delta > 0$ . Letting  $\delta = p/2$  we find

$$E \left[ \sup_{t \in [0, T]} \left| \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, x_s^\epsilon) ds \right|^p \right] = O(\epsilon^{p/2}). \tag{3.42}$$

Now use Hölder's inequality, the uniform Lipschitz property of  $\tilde{G}$ , and Assumption 3.1 again to compute

$$\begin{aligned}
& E \left[ \sup_{t \in [0, T]} \left| \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, q_s^\epsilon, \psi(s, q_s^\epsilon)) ds \right|^p \right] \\
& \leq O(\epsilon^{p/2}) + \tilde{C} E \left[ \sup_{t \in [0, T]} \left| \int_0^t \tilde{G}(s, x_s^\epsilon) - \tilde{G}(s, q_s^\epsilon, \psi(s, q_s^\epsilon)) ds \right|^p \right] \\
& \leq O(\epsilon^{p/2}) + \tilde{C} T^{p-1} E \left[ \int_0^T |\tilde{G}(s, x_s^\epsilon) - \tilde{G}(s, q_s^\epsilon, \psi(s, q_s^\epsilon))|^p ds \right] \\
& \leq O(\epsilon^{p/2}) + \tilde{C} T^p \sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] \\
& = O(\epsilon^{p/2}).
\end{aligned} \tag{3.43}$$

This proves the claim for  $p \geq 2$ . The result for arbitrary  $p > 0$  then follows from an application of Hölder's inequality.  $\square$

**3.2.2. Formal derivation of  $\tilde{G}$ .** Formally applying the Fredholm alternative to Eq. (3.33) motivates the form that  $\tilde{G}$  must have in order for  $\chi$  and its derivatives to possess the growth bounds required by Theorem 3.1. The formal calculation is simple enough that we repeat it here:

Let  $L^*$  be the formal adjoint to  $L$  and suppose we have a solution,  $h(t, x, z)$ , to

$$L^*h = 0, \quad \int h(t, x, z)dz = 1. \quad (3.44)$$

If  $\chi$  and its derivatives grow slowly enough and  $h$  and its derivatives decay quickly enough, then  $\int h L \chi dz$  will exist, the boundary terms from integration by parts will vanish at infinity, and we find

$$0 = \int (L^*h)\chi dz = \int h L(\chi) dz = \int h(G - \tilde{G}) dz = \int h G dz - \tilde{G}. \quad (3.45)$$

Therefore we must have

$$\tilde{G}(t, x) = \int h(t, x, z)G(t, x, z)dz. \quad (3.46)$$

In essence, the homogenized quantity is obtained by averaging over  $h$ , the instantaneous equilibrium distribution for the fast variables,  $z$ . This corroborates the heuristic discussion in Section 2.

**3.2.3. Limiting equation.** We now apply the above framework to prove existence of a limiting process  $q_t^\epsilon \rightarrow q_s$  and deriving an SDE satisfied by  $q_s$ . Specifically, we have:

**Theorem 3.2.** *Let  $T > 0$ ,  $p \geq 2$ ,  $0 < \beta \leq p/2$ ,  $x_t^\epsilon = (q_t^\epsilon, p_t^\epsilon)$  satisfy the SDE Eq. (3.10)-Eq. (3.11), suppose Assumptions 3.1-3.3 hold, and that the SDE for  $q_t^\epsilon$ , Eq. (3.10), can be rewritten in the form*

$$q_t^\epsilon = q_0^\epsilon + \int_0^t \tilde{F}(s, x_s^\epsilon) ds + \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds + \int_0^t \tilde{\sigma}(s, x_s^\epsilon) dW_s + R_t^\epsilon \quad (3.47)$$

where the components of  $G$  have the properties described in Assumption 3.3,  $\tilde{F}(t, x) : [0, \infty) \times \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ ,  $\tilde{\sigma}(t, x) : [0, \infty) \times \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n \times k}$  are continuous, Lipschitz in  $x$ , uniformly in  $t \in [0, T]$ , and  $R_t^\epsilon$  are continuous semimartingales that satisfy

$$E \left[ \sup_{t \in [0, T]} \|R_t^\epsilon\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (3.48)$$

Suppose  $\tilde{G}$  (from Assumption 3.3) is Lipschitz in  $x$ , uniformly in  $t \in [0, T]$ , and we have initial conditions  $E[\|q_0^\epsilon\|^p] < \infty$ ,  $E[\|q_0\|^p] < \infty$ , and  $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2})$ . Then

$$E \left[ \sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+ \quad (3.49)$$

where  $q_t$  satisfies the SDE

$$\begin{aligned} q_t &= q_0 + \int_0^t \tilde{F}(s, q_s, \psi(s, q_s)) ds + \int_0^t \tilde{G}(s, q_s, \psi(s, q_s)) ds \\ &\quad + \int_0^t \tilde{\sigma}(s, q_s, \psi(s, q_s)) dW_s. \end{aligned} \quad (3.50)$$

*Proof.* We will prove this theorem by verifying all the hypotheses of Lemma A3. Define

$$\tilde{R}_t^\epsilon = R_t^\epsilon + \int_0^t G(s, x_s^\epsilon, z_s^\epsilon) ds - \int_0^t \tilde{G}(s, x_s^\epsilon) ds. \quad (3.51)$$

Then

$$q_t^\epsilon = q_0^\epsilon + \int_0^t \tilde{F}(s, x_s^\epsilon) ds + \int_0^t \tilde{G}(s, x_s^\epsilon) ds + \int_0^t \tilde{\sigma}(s, x_s^\epsilon) dW_s + \tilde{R}_t^\epsilon \quad (3.52)$$

where  $\tilde{F} + \tilde{G}$  and  $\tilde{\sigma}$  are Lipschitz in  $x$ , uniformly for  $t \in [0, T]$  and

$$E \left[ \sup_{t \in [0, T]} \|\tilde{R}_t^\epsilon\|^p \right] = O(\epsilon^\beta) \quad (3.53)$$

by Theorem 3.1.  $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^\beta)$  as  $\epsilon \rightarrow 0^+$  by assumption and

$$\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^{p/2}) \text{ as } \epsilon \rightarrow 0^+ \quad (3.54)$$

by Assumption 3.1. Note that the assumptions also imply that a solution  $q_t$  to Eq. (3.50) exists for all  $t \geq 0$  [1].

For any  $\epsilon > 0$ , using the Burkholder-Davis-Gundy inequality and Hölder's inequality we obtain the bound

$$\begin{aligned} &E \left[ \sup_{t \in [0, T]} \|q_t^\epsilon\|^p \right] \\ &\leq 4^{p-1} \left( E[\|q_0^\epsilon\|^p] + \epsilon^{-p/2} E \left[ \left( \int_0^T \|G_1(s, x_s^\epsilon, z_s^\epsilon)\| ds \right)^p \right] \right. \\ &\quad \left. + E \left[ \left( \int_0^T \|F_1(s, x_s^\epsilon, z_s^\epsilon)\| ds \right)^p \right] + E \left[ \sup_{t \in [0, T]} \left\| \int_0^t \sigma_1(s, x_s^\epsilon) dW_s \right\|^p \right] \right) \\ &\leq 4^{p-1} \left( E[\|q_0^\epsilon\|^p] + \epsilon^{-p/2} T^{p-1} \int_0^T E[\|G_1(s, x_s^\epsilon, z_s^\epsilon)\|^p] ds \right. \\ &\quad \left. + T^{p-1} \int_0^T E[\|F_1(s, x_s^\epsilon, z_s^\epsilon)\|^p] ds + \tilde{C} T^{p/2-1} \int_0^T E[\|\sigma_1(s, x_s^\epsilon)\|_F^p] ds \right). \end{aligned} \quad (3.55)$$

Polynomial boundedness (see Assumption 3.2) gives

$$E \left[ \sup_{t \in [0, T]} \|q_t^\epsilon\|^p \right] \leq 4^{p-1} \left( E[\|q_0^\epsilon\|^p] + \tilde{C} \int_0^T E[(1 + \|z_s^\epsilon\|^q)^p] ds \right), \quad (3.56)$$

where we absorbed all factors of  $T$  and  $\epsilon$  into the constant  $\tilde{C}$ . Using Assumption 3.1 then gives

$$E \left[ \sup_{t \in [0, T]} \|q_t^\epsilon\|^p \right] < \infty \quad (3.57)$$

for all  $\epsilon$  sufficiently small.

Finally, for  $n > 0$  define the stopping time  $\tau_n = \inf\{t \geq 0 : \|q_t\| \geq n\}$ . Then for  $0 \leq t \leq T$  the Lipschitz properties together with the Burkholder-Davis-Gundy and Hölder's inequalities imply

$$E \left[ \sup_{s \in [0, t]} \|q_s^{\tau_n}\|^p \right] \quad (3.58)$$

$$\leq 3^{p-1} \left( E[\|q_0\|^p] + E \left[ \left( \int_0^{t \wedge \tau_n} \|(\tilde{F} + \tilde{G})(s, q_s^{\tau_n}, \psi(s, q_s^{\tau_n}))\| ds \right)^p \right] \right)$$

$$+ E \left[ \sup_{t \in [0, t]} \left\| \int_0^{t \wedge \tau_n} \tilde{\sigma}(s, q_s^{\tau_n}, \psi(s, q_s^{\tau_n})) dW_s \right\|^p \right] \right)$$

$$\leq 3^{p-1} E[\|q_0\|^p] + \tilde{C} \int_0^t E[\|q_s^{\tau_n}\|^p] ds \quad (3.59)$$

$$+ \tilde{C} \int_0^t \|(\tilde{F} + \tilde{G})(s, 0, \psi(s, 0))\|^p + \|\tilde{\sigma}(s, 0, \psi(s, 0))\|_F^p ds$$

$$\leq \tilde{C} \left( 1 + \int_0^t E \left[ \sup_{r \in [0, s]} \|q_r^{\tau_n}\|^p \right] ds \right), \quad (3.60)$$

where  $\tilde{C}$  changes line to line, and is independent of  $t$ .

The definition of  $\tau_n$ , together with  $E[\|q_0\|^p] < \infty$ , implies that

$$\sup_{s \geq 0} E \left[ \sup_{r \in [0, s]} \|q_r^{\tau_n}\|^p \right] < \infty. \quad (3.61)$$

Therefore we can apply Gronwall's inequality to get

$$E \left[ \sup_{t \in [0, T]} \|q_t^{\tau_n}\|^p \right] \leq \tilde{C} e^{\tilde{C} T}, \quad (3.62)$$

where the constant  $\tilde{C}$  is independent of  $n$ . Hence, the monotone convergence theorem yields

$$E \left[ \sup_{t \in [0, T]} \|q_t\|^p \right] \leq \tilde{C} e^{\tilde{C} T} < \infty. \quad (3.63)$$

This completes the verification that the hypotheses of Lemma A3 hold, allowing us to conclude that

$$E \left[ \sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (3.64)$$

□

## 4. Homogenization of Hamiltonian systems

In this final section, we apply the above framework to our original Hamiltonian system, Eq. (3.3)-Eq. (3.4) (in particular,  $m = n$  in this section) in order to prove the existence of a limiting process  $q_t^\epsilon \rightarrow q_t$  and derive a homogenized SDE for  $q_t$ . Specifically, in Sections 4.1 and 4.2 we will study a class of Hamiltonian systems for which the PDEs Eq. (3.44) for  $h$  and Eq. (3.33) for  $\chi$  that are needed to derive the limiting equation are explicitly solvable and the required bounds can be verified by elementary means.

The SDE Eq. (3.3)-Eq. (3.4) can be rewritten in the general form Eq. (3.10)-Eq. (3.11):

$$dq_t^\epsilon = \frac{1}{\sqrt{\epsilon}} \nabla_z K(t, q_t^\epsilon, z_t^\epsilon) dt, \quad (4.1)$$

$$dp_t^\epsilon = \left( -\frac{1}{\sqrt{\epsilon}} (\gamma_l(t, x_t^\epsilon) - \nabla_q \psi_l(t, q_t^\epsilon)) \partial_{z_l} K(t, q_t^\epsilon, z_t^\epsilon) - \nabla_q K(t, q_t^\epsilon, z_t^\epsilon) \right. \\ \left. - \nabla_q V(t, q_t^\epsilon) + F(t, x_t^\epsilon) \right) dt + \sigma(t, x_t^\epsilon) dW_t, \quad (4.2)$$

where  $\gamma_l$  denotes the vector obtained by taking the  $l$ th column of  $\gamma$ . Specifically,

$$F_1 = 0, \quad \sigma_1 = 0, \quad \sigma_2 = \sigma, \quad G_1(t, x, z) = \nabla_z K(t, q, z), \quad (4.3)$$

$$F_2(t, x, z) = -\nabla_q K(t, q, z) - \nabla_q V(t, q) + F(t, x), \quad (4.4)$$

$$G_2(t, x, z) = -(\gamma_l(t, x) - \nabla_q \psi_l(t, q)) \partial_{z_l} K(t, q, z). \quad (4.5)$$

In particular,  $\sigma_1 = 0$ , so below we use the definition of  $C^{1,2}$  applicable to this case.

The operator  $L$ , Eq. (3.26), and its formal adjoint have the following form:

$$(L\chi)(t, x, z) = \frac{1}{2} \Sigma_{kl}(t, x) (\partial_{z_k} \partial_{z_l} \chi)(t, x, z) \quad (4.6)$$

$$- \tilde{\gamma}_{kl}(t, x) \partial_{z_l} K(t, q, z) (\partial_{z_k} \chi)(t, x, z),$$

$$(L^* h)(t, x, z) = \partial_{z_k} \left( \frac{1}{2} \Sigma_{kl}(t, x) \partial_{z_l} h(t, x, z) \right. \quad (4.7)$$

$$\left. + \tilde{\gamma}_{kl}(t, x) \partial_{z_l} K(t, q, z) h(t, x, z) \right),$$

where

$$\Sigma_{ij} = \sum_{\rho} \sigma_{i\rho} \sigma_{j\rho} \quad (4.8)$$

and  $\tilde{\gamma}$  was defined in Eq. (3.7). Here  $\sigma$  and  $\Sigma$  denote what were  $\sigma_2$  and  $\Sigma_{22}$  respectively in Eq. (3.16) and  $\sigma_1 = 0$ . In particular, the indices on  $\Sigma$  have the meaning  $\Sigma_{ij} \equiv (\Sigma_{22})_{ij}$ .

#### 4.1. Computing the noise induced drift

In general, an explicit solution to  $L^*h = 0$  is not available, and so the homogenized equation can only be defined implicitly, as in Theorem 3.2. However, there are certain classes of systems where we can explicitly derive the form of the additional vector field,  $\tilde{G}$ , appearing in the homogenized equation. In [8], one such class was studied by a different method. Here, we explore the case where the noise and dissipation satisfy the fluctuation dissipation relation pointwise for a time and state dependent generalized temperature  $T(t, q)$ ,

$$\Sigma_{ij}(t, q) = 2k_B T(t, q) \gamma_{ij}(t, q). \quad (4.9)$$

where  $\Sigma$  was defined in Eq. (4.8). We will make Assumptions A1-A4, but make no further constraints on the form of the Hamiltonian here.

As can be verified by a direct calculation, under the assumption Eq. (4.9), the adjoint equation Eq. (4.7) is solved by

$$h(t, q, z) = \frac{1}{Z(t, q)} \exp[-\beta(t, q)K(t, q, z)], \quad (4.10)$$

where we define  $\beta(t, q) = 1/(k_B T(t, q))$  and  $Z$ , the ‘partition function’, is chosen so that  $\int h dz = 1$ . Note that Assumption A3 ensures such a normalization exists. We also point out that in this case, the antisymmetric part of  $\tilde{\gamma}$  does not contribute to the right hand side of Eq. (4.7).

An interesting point to note is that when the antisymmetric part of  $\tilde{\gamma}$  vanishes (physically, for  $K$  quadratic in  $z$  this means a vanishing magnetic field), the vector field that we are taking the divergence of in Eq. (4.7) vanishes identically. When  $\tilde{\gamma}$  has a non-vanishing antisymmetric part, only once we take the divergence does the expression in Eq. (4.7) vanish.

From Eq. (3.6), we see that the terms that require homogenization are

$$\begin{aligned} G(t, q_t^\epsilon, z_t^\epsilon) = & -(\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{q^j} K(t, q_t^\epsilon, z_t^\epsilon) dt \\ & + (z_t^\epsilon)_j \partial_{q^j} (\tilde{\gamma}^{-1})^{ij}(t, q_t^\epsilon) \partial_{z^i} K(t, q_t^\epsilon, z_t^\epsilon) dt. \end{aligned} \quad (4.11)$$

Using Eq. (4.10), the formal calculation of Section 3.2.2 gives

$$\tilde{G}(t, q) = -(\tilde{\gamma}^{-1})^{ij}(t, q) \langle \partial_{q^j} K(t, q, z) \rangle + \frac{\partial_{q^l} (\tilde{\gamma}^{-1})^{il}(t, q)}{\beta(t, q)}, \quad (4.12)$$

where we define

$$\langle \partial_{q^j} K(t, q, z) \rangle = \frac{1}{Z(t, q)} \int \partial_{q^j} K(t, q, z) \exp[-\beta(t, q)K(t, q, z)] dz. \quad (4.13)$$

Of course, this calculation is only formal. In the next section, we study a particular case where everything can be made rigorous.

#### 4.2. Rigorous Homogenization of a class of Hamiltonian systems

In this section we explore a class of Hamiltonian systems for which Assumption 3.3 can be rigorously verified via an explicit solution to the PDE for

$\chi$ . We will work with Hamiltonian systems that satisfy Assumptions A1-A5, A7. In particular, we are restricting to the class of Hamiltonians with

$$K(t, q, z) = \tilde{K}(t, q, A^{ij}(t, q)z_i z_j), \quad (4.14)$$

where  $A(t, q)$  is valued in the space of positive definite  $n \times n$ -matrices. We will write  $\tilde{K} \equiv \tilde{K}(t, q, \zeta)$  and  $\tilde{K}' \equiv \partial_\zeta \tilde{K}$ .

We will also need the following relations between  $\Sigma$ ,  $\gamma$ , and  $A$  to hold:

**Assumption 4.1.**  $\sigma$  is independent of  $p$  and

$$\Sigma(t, q) = b_1(t, q)A^{-1}(t, q), \quad \gamma(t, q) = b_2(t, q)A^{-1}(t, q) \quad (4.15)$$

where, for every  $T > 0$ , the  $b_i$  are bounded,  $C^2$  functions that have positive lower bounds and bounded first derivatives, both on  $[0, T] \times \mathbb{R}^n$ .

Note that these relations imply a fluctuation-dissipation relation with a time and state dependent generalized temperature  $T = \frac{b_1}{2k_B b_2}$ .

In [8], we showed that Assumptions A1-A5 imply:

$$d(q_t^\epsilon)^i = \tilde{F}^i(t, x)dt + \tilde{\sigma}_\rho^i(t, x_t^\epsilon)dW_t^\rho + G^i(t, x_t^\epsilon, z_t^\epsilon)dt + d(R_t^\epsilon)^i,$$

where

$$\begin{aligned} \tilde{F}^i(t, x) &= (\tilde{\gamma}^{-1})^{ij}(t, q)(-\partial_t \psi_j(t, q) - \partial_{q^j} V(t, q) + F_j(t, x)) + S^i(t, q), \\ (4.16) \end{aligned}$$

$$G^i(t, q, z) = -(\tilde{\gamma}^{-1})^{ij}(t, q)(\partial_{q^j} \tilde{K})(t, q, A^{ij}(t, q)z_i z_j), \quad (4.17)$$

$$\tilde{\sigma}_\rho^i(t, x) = (\tilde{\gamma}^{-1})^{ij}(t, q)\sigma_{jp}(t, x), \quad (4.18)$$

$$S^i(t, q) = k_B T(t, q) \left( \partial_{q^j} (\tilde{\gamma}^{-1})^{ij}(t, q) - \frac{1}{2} (\tilde{\gamma}^{-1})^{ik}(t, q) A_{jl}^{-1}(t, q) \partial_{q^k} A^{jl}(t, q) \right), \quad (4.19)$$

and  $R_t^\epsilon$  is a family of continuous semimartingales.  $S(t, q)$  is called the *noise induced drift* (see Eq. 3.26 in [8]).

Note, that with  $K(t, q, z)$  defined by Eq. (4.14), the first term in Eq. (4.11) consists of two contributions—one coming from the  $q$ -dependence of  $\tilde{K}$  and one coming from the  $q$ -dependence of  $A$ . The  $G$  defined here comprises only the first contribution. The method of [8] is able to homogenize the second term in Eq. (4.11), as well the first contribution of the first term, leading to the noise induced drift,  $S$ , but fails when  $\tilde{K}$  depends explicitly on  $q$ . However, under certain circumstances, the method developed in Section 3.2 is succeeds in homogenizing the system when  $\tilde{K}$  has  $q$  dependence, as we now show.

We will need one final assumption:

**Assumption 4.2.** For every  $T > 0$ :

1. There exists  $\zeta_0 > 0$  and  $C > 0$  such that  $\tilde{K}'(t, q, \zeta) \geq C$  for all  $(t, q, \zeta) \in [0, T] \times \mathbb{R}^n \times [\zeta_0, \infty)$ .
2.  $\tilde{K}(t, q, \zeta)$ ,  $\partial_t \partial_{q^i} \tilde{K}(t, q, \zeta)$ ,  $\partial_{q^i} \partial_\zeta \tilde{K}(t, q, \zeta)$ , and  $\partial_{q^i} \partial_{q^j} \tilde{K}(t, q, \zeta)$  are polynomially bounded in  $\zeta$ , uniformly in  $(t, q) \in [0, T] \times \mathbb{R}^n$ .

We are now prepared to prove the following homogenization result:

**Theorem 4.1.** Let  $x_t^\epsilon = (q_t^\epsilon, p_t^\epsilon)$  satisfy the Hamiltonian SDE Eq. (3.3)-Eq. (3.4) and suppose Assumptions A1-A5, A7, 4.1, and 4.2 hold. Let  $p \geq 2$  and suppose we have initial conditions that satisfy  $E[\|q_0^\epsilon\|^p] < \infty$ ,  $E[\|q_0\|^p] < \infty$ , and  $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2})$ . Then for any  $T > 0$ ,  $0 < \beta < p/2$  we have

$$E \left[ \sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+ \quad (4.20)$$

where  $q_t$  is the solution to the SDE

$$\begin{aligned} dq_t^i &= (\tilde{\gamma}^{-1})^{ij}(t, q_t)(-\partial_t \psi_j(t, q_t) - \partial_{q^j} V(t, q_t) + F_j(t, q_t, \psi(t, q_t)))dt \\ &\quad + S^i(t, q_t)dt + \tilde{G}^i(t, q_t)dt + (\tilde{\gamma}^{-1})^{ij}(t, q_t)\sigma_{j\rho}(t, q_t)dW_t^\rho \end{aligned} \quad (4.21)$$

with initial condition  $q_0$ . See Eq. (3.7), Eq. (4.19), and Eq. (4.23) for the definitions of  $\tilde{\gamma}$ ,  $S$ , and  $\tilde{G}$ , respectively.

*Proof.* From [8], Assumptions A1-A5, A7 imply:

1.  $q_t^\epsilon$  satisfies an equation of the form Eq. (3.47), where, for every  $T > 0$ ,  $\tilde{F}$ ,  $\tilde{\sigma}$  are bounded, continuous, and Lipschitz in  $x$ , on  $[0, T] \times \mathbb{R}^{2n}$ , with Lipschitz constant uniform in  $t$ .
2. Assumptions 3.1 holds.
3. For any  $p > 0$ ,  $T > 0$ ,  $0 < \beta < p/2$  we have

$$E \left[ \sup_{t \in [0, T]} \|R_t^\epsilon\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (4.22)$$

Combined with polynomial boundedness of  $\tilde{K}$  (Assumption 4.2) we see that Assumption 3.2 also holds. Therefore, to apply Theorem 3.2, we have to verify Assumption 3.3 and that the  $\tilde{G}$  referenced therein is Lipschitz in  $x$ , uniformly in  $t \in [0, T]$ .

From Section 3.2.2, we expect that

$$\tilde{G}^i(t, q) = -(\tilde{\gamma}^{-1})^{ij}(t, q)\langle \partial_{q^j} \tilde{K}(t, q, A^{ij}(t, q)z_i z_j) \rangle \quad (4.23)$$

where, similarly to Eq. (4.13),

$$\langle \partial_{q^j} \tilde{K}(t, q, \|z\|_A^2) \rangle = \frac{1}{Z(t, q)} \int \partial_{q^j} \tilde{K}(t, q, \|z\|_A^2) \exp[-\beta(t, q)\tilde{K}(t, q, \|z\|_A^2)] dz. \quad (4.24)$$

Here we use the shorthand  $\|z\|_A^2 \equiv A^{ij}(t, q)z_i z_j$  when the implied values of  $t, q$  are apparent from the context.

Using our assumptions, along with several applications of the DCT, one can see that  $\tilde{G}$  is  $C^1$  and, for every  $T > 0$ , is bounded with bounded first derivatives on  $[0, T] \times \mathbb{R}^n$ . In particular, it is Lipschitz in  $q$ , uniformly in  $t \in [0, T]$ .

We now turn to solving the equation

$$L\chi = G - \tilde{G}. \quad (4.25)$$

Since  $G - \tilde{G}$  is independent of  $p$  and depends on  $z$  only through  $\|z\|_A^2$ , we look for  $\chi$  with the same behavior. Using the ansatz  $\chi(t, q, z) = \tilde{\chi}(t, q, \|z\|_A^2)$ ,

and defining  $G^i(t, q, \zeta) = -(\tilde{\gamma}^{-1})^{ij}(t, q)(\partial_{q^j}\tilde{K})(t, q, \zeta)$ , leads (on account of the antisymmetry of the matrix  $\tilde{\gamma} - \gamma$ ) to the ODE in the variable  $\zeta$ :

$$\begin{aligned} \zeta \tilde{\chi}''(t, q, \zeta) + \left( \frac{n}{2} - \beta(t, q)\zeta \tilde{K}'(t, q, \zeta) \right) \tilde{\chi}'(t, q, \zeta) \\ = \frac{1}{2b_1(t, q)}(G(t, q, \zeta) - \tilde{G}(t, q)). \end{aligned} \quad (4.26)$$

This has the solution

$$\begin{aligned} \tilde{\chi}(t, q, \zeta) = \frac{1}{2b_1(t, q)} \int_0^\zeta \zeta_1^{-n/2} \exp[\beta(t, q)\tilde{K}(t, q, \zeta_1)] \\ \times \int_0^{\zeta_1} \zeta_2^{(n-2)/2} \exp[-\beta(t, q)\tilde{K}(t, q, \zeta_2)] (G(t, q, \zeta_2) - \tilde{G}(t, q)) d\zeta_2 d\zeta_1 \end{aligned} \quad (4.27)$$

Therefore  $\chi(t, q, z) \equiv \tilde{\chi}(t, q, \|z\|_A^2)$  solves the PDE Eq. (4.25). One can show that it is  $C^{1,2}$  and that  $\chi$  and its first derivatives are polynomially bounded in  $z$ , uniformly for  $(t, q) \in [0, T] \times \mathbb{R}^n$ . As a representative example, in B we outline the proof that  $\tilde{\chi}(t, q, \zeta)$  is polynomially bounded in  $\zeta$ , uniformly in  $(t, q) \in [0, T] \times \mathbb{R}^n$ . The remainder of the computations are similar and we leave them to the reader.

$\chi$  is independent of  $p$ , so  $\partial_{p_i}\partial_{p_j}\chi = 0$  and  $\partial_{p_i}\partial_{z_j}\chi = 0$ . Therefore, this completes the verification of Assumption 3.3 and we are justified in using Theorem 3.2 to conclude

$$E \left[ \sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+, \quad (4.28)$$

where  $q_t$  satisfies the SDE

$$q_t = q_0 + \int_0^t \tilde{F}(s, q_s, \psi(s, q_s)) ds + \int_0^t \tilde{G}(s, q_s) ds + \int_0^t \tilde{\sigma}(s, q_s) dW_s$$

as claimed.  $\square$

Lastly, we give an example of a general class of Hamiltonians that satisfy the hypotheses of Theorem 4.1. The proof of this corollary is straightforward, so we leave it to the reader.

**Corollary 4.1.** *Consider the class of Hamiltonians of the form*

$$H(t, q, p) = \sum_{l=k_1}^{k_2} d_l(t, q) [A^{ij}(t, q)(p - \psi(t, q))_i(p - \psi(t, q))_j]^l + V(t, q) \quad (4.29)$$

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  $t \in [0, T]$ .
2.  $\psi$  is  $C^3$  and  $\partial_t \psi$ ,  $\partial_{q^i} \psi$ ,  $\partial_t \partial_{q^i} \psi$ ,  $\partial_{q^i} \partial_{q^j} \psi$ ,  $\partial_t \partial_{q^j} \partial_{q^i} \psi$ , and  $\partial_{q^i} \partial_{q^j} \partial_{q^i} \psi$  are bounded.
3.  $d_l$  are  $C^2$ , non-negative, bounded, and have bounded first and second derivatives.
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.

Also suppose that

1.  $\sigma$  is independent of  $p$  and

$$\Sigma(t, q) = b_1(t, q)A^{-1}(t, q), \quad \gamma(t, q) = b_2(t, q)A^{-1}(t, q) \quad (4.30)$$

where, for every  $T > 0$ , the  $b_i$  are bounded,  $C^2$  functions with positive lower bounds and bounded first derivatives.

2.  $\gamma$  is  $C^2$ , is independent of  $p$ , and  $\partial_t \gamma$ ,  $\partial_{q^i} \gamma$ ,  $\partial_t \partial_{q^j} \gamma$ ,  $\partial_{q^i} \partial_{q^j} \gamma$  are bounded on  $[0, T] \times \mathbb{R}^n$ .

3. The eigenvalues of  $\gamma$  are bounded below by some  $\lambda > 0$ .

4.  $\gamma$ ,  $F$ , and  $\sigma$  are bounded.

5.  $F$  and  $\sigma$  are Lipschitz in  $x$  uniformly in  $t \in [0, T]$ .

6. 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  $\omega \in \Omega$ .

7. There is a  $p \geq 2$  such that

$$E[\|q_0^\epsilon\|^p] < \infty, \quad E[\|q_0\|^p] < \infty, \quad \text{and} \quad E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^{p/2}). \quad (4.31)$$

Then all the hypotheses of Theorem 4.1 hold, in particular Assumptions A1-A5, A7, 4.1, and 4.2 hold, and hence, for any  $\beta \in (0, \frac{p}{2})$ ,

$$E\left[\sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p\right] = O(\epsilon^\beta) \quad \text{as } \epsilon \rightarrow 0^+, \quad (4.32)$$

where  $x_t^\epsilon = (q_t^\epsilon, p_t^\epsilon)$  satisfy the Hamiltonian SDE Eq. (3.3)-Eq. (3.4) and  $q_t$  satisfies the homogenized SDE, Eq. (4.21).

## Appendix A. Material from [8]

In this appendix, we collect several useful assumptions and results from [8]. The assumptions listed here are *not* used in the entirety of this current work. When they are needed for a particular result we explicitly reference them.

**Assumption A1.** We assume that the Hamiltonian has the form given in Eq. (3.1) where  $K$  and  $\psi$  are  $C^2$  and  $K$  is non-negative. For every  $T > 0$ , we assume the following bounds hold on  $[0, T] \times \mathbb{R}^{2n}$ :

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

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

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

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

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

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

**Assumption A2.** 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.  $\gamma$  is symmetric with eigenvalues bounded below by some  $\lambda > 0$ .
3.  $\gamma$ ,  $F$ ,  $\partial_t \psi$ , and  $\sigma$  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 A3.** We assume that for every  $T > 0$  there exists  $c > 0$ ,  $\eta > 0$  such that

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

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

**Assumption A4.** We assume that  $\gamma$  is  $C^1$  and is independent of  $p$ .

**Assumption A5.** We assume that  $K$  has the form

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

where  $\tilde{K}(t, q, \zeta)$  is  $C^2$  and non-negative on  $[0, \infty) \times \mathbb{R}^n \times [0, \infty)$  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  $[0, T] \times \mathbb{R}^n$ .

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.

**Assumption A7.** We assume that, for every  $T > 0$ ,  $\nabla_q V$ ,  $F$ , and  $\sigma$  are Lipschitz in  $x$  uniformly in  $t \in [0, T]$ . We also assume that  $A$  and  $\gamma$  are  $C^2$ ,  $\psi$  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  $[0, T] \times \mathbb{R}^{2n}$  for every  $T > 0$ .

Note that, combined with Assumptions A1-A4, 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.

**Lemma A1.** Under Assumptions A1 and A2, for any  $T > 0$ ,  $p > 0$  we have

$$E \left[ \sup_{t \in [0, T]} \|q_t^\epsilon\|^p \right] < \infty. \quad (\text{A.6})$$

**Lemma A2.** Under Assumptions A1-A3, for any  $T > 0$ ,  $p > 0$  we have

$$\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^{p/2}) \text{ as } \epsilon \rightarrow 0^+ \quad (\text{A.7})$$

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

$$E \left[ \sup_{t \in [0, T]} \|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (\text{A.8})$$

The following is a slight variant of the result from [8], but the proof is identical.

**Lemma A3.** Let  $T > 0$  and suppose we have continuous functions  $\tilde{F}(t, x) : [0, \infty) \times \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ ,  $\tilde{\sigma}(t, x) : [0, \infty) \times \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n \times k}$ , and  $\psi : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^m$  that are Lipschitz in  $x$ , uniformly in  $t \in [0, T]$ .

Let  $W_t$  be a  $k$ -dimensional Wiener process,  $p \geq 2$  and  $\beta > 0$  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.  $E[\|q_0^\epsilon - q_0\|^p] = O(\epsilon^\beta)$  as  $\epsilon \rightarrow 0^+$ .
4.  $E[\sup_{t \in [0, T]} \|\tilde{R}_t^\epsilon\|^p] = O(\epsilon^\beta)$  as  $\epsilon \rightarrow 0^+$ .
5.  $\sup_{t \in [0, T]} E[\|p_t^\epsilon - \psi(t, q_t^\epsilon)\|^p] = O(\epsilon^\beta)$  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

$$E \left[ \sup_{t \in [0, T]} \|q_t^\epsilon - q_t\|^p \right] = O(\epsilon^\beta) \text{ as } \epsilon \rightarrow 0^+. \quad (\text{A.9})$$

## Appendix B. Polynomial boundedness of $\tilde{\chi}$

Changing variables,  $\tilde{\chi}$  can be expressed as

$$\begin{aligned} \tilde{\chi}(t, q, \zeta) &= \frac{1}{2b_1(t, q)} \zeta \int_0^1 \exp[\beta(t, q) \tilde{K}(t, q, s\zeta)] \\ &\quad \times \int_0^1 r^{(m-2)/2} \exp[-\beta(t, q) \tilde{K}(t, q, rs\zeta)] (G(t, q, rs\zeta) - \tilde{G}(t, q)) dr ds. \end{aligned} \quad (\text{B.1})$$

Applying the DCT to this expression several times, one can prove that  $\tilde{\chi}$  is  $C^{1,2}$ . Using the fact that  $\tilde{K}$  and  $\partial_{q^i} \tilde{K}$  are polynomially bounded in  $\zeta$ , uniformly in  $(t, q) \in [0, T] \times \mathbb{R}^n$ , one can see that  $\tilde{\chi}(t, q, \zeta)$  is bounded on  $[0, T] \times \mathbb{R}^n \times [0, \zeta_0]$  for any  $\zeta_0 > 0$ . From Assumption 4.2, there exists  $\zeta_0$  and  $C > 0$  such that  $\tilde{K}'(t, q, \zeta) \geq C$  for all  $(t, q, \zeta) \in [0, T] \times \mathbb{R}^n \times [\zeta_0, \infty)$ .

By combining Eq. (4.27) with Eq. (4.23), one finds that for  $\zeta \geq \zeta_0$ ,  $\tilde{\chi}$  can alternatively be written as

$$\begin{aligned} \tilde{\chi}(t, q, \zeta) &= \tilde{\chi}(t, q, \zeta_0) + \frac{1}{2b_1(t, q)} \int_{\zeta_0}^{\zeta} \zeta_1^{-m/2} \exp[\beta(t, q) \tilde{K}(t, q, \zeta_1)] \\ &\quad \times \int_{\zeta_1}^{\infty} \exp[-\beta(t, q) \tilde{K}(t, q, \zeta_2)] \zeta_2^{(m-2)/2} (\tilde{G}(t, q) - G(t, q, \zeta_2)) d\zeta_2 d\zeta_1. \end{aligned} \quad (\text{B.2})$$

Therefore, if we can show that the second term has the polynomial boundedness property then so does  $\tilde{\chi}$ , and hence  $\chi$ .

Letting  $\tilde{C}$  denote a constant that potentially changes in each line and choosing  $\zeta_0$  as in Assumption 4.2, we have

$$\begin{aligned}
& \left\| \frac{1}{2b_1(t, q)} \int_{\zeta_0}^{\zeta} \zeta_1^{-m/2} \exp[\beta(t, q)\tilde{K}(t, q, \zeta_1)] \right. \\
& \quad \times \left. \int_{\zeta_1}^{\infty} \exp[-\beta(t, q)\tilde{K}(t, q, \zeta_2)] \zeta_2^{(m-2)/2} (\tilde{G}(t, q) - G(t, q, \zeta_2)) d\zeta_2 d\zeta_1 \right\| \\
& \leq \tilde{C} \int_{\zeta_0}^{\zeta} \zeta_1^{-m/2} \exp[\beta(t, q)\tilde{K}(t, q, \zeta_1)] \\
& \quad \times \int_{\zeta_1}^{\infty} \exp[-\beta(t, q)\tilde{K}(t, q, \zeta_2)] \zeta_2^{(m-2)/2+q} d\zeta_2 d\zeta_1 \\
& \leq \tilde{C} \zeta_0^{-m/2} \int_{\zeta_0}^{\zeta} \exp[\beta(t, q)\tilde{K}(t, q, \zeta_1)] \\
& \quad \times \int_{\zeta_1}^{\infty} \exp[-\beta(t, q)\tilde{K}(t, q, \zeta_2)] \tilde{K}(t, q, \zeta_2)^{((m-2)/2+q)/\eta} \tilde{K}'(t, q, \zeta_2) d\zeta_2 d\zeta_1 \\
& = \tilde{C} \zeta_0^{-m/2} \int_{\zeta_0}^{\zeta} \exp[\beta(t, q)\tilde{K}(t, q, \zeta_1)] \\
& \quad \times \int_{\tilde{K}(t, q, \zeta_1)}^{\infty} \exp[-\beta(t, q)u] u^{((m-2)/2+q)/\eta} du d\zeta_1
\end{aligned} \tag{B.3}$$

for some  $q > 0$ . To obtain the first inequality, we use polynomial boundedness of  $\partial_{q^i} \tilde{K}$ . For the second, we used Assumption Eq. (A3) together with the fact that  $\tilde{K}' \geq C > 0$  on  $[0, T] \times \mathbb{R}^n \times [\zeta_0, \infty)$ .

Therefore we obtain

$$\|\tilde{\chi}(t, q, \zeta)\| \leq \tilde{C} \left( 1 + \int_{\zeta_0}^{\zeta} P(\tilde{K}(t, q, \zeta_1)) d\zeta_1 \right) \tag{B.4}$$

for some polynomial  $P(x)$  with positive coefficients that are independant of  $t$  and  $q$ . Polynomial boundedness of  $\tilde{K}$  then implies

$$\|\tilde{\chi}(t, q, \zeta)\| \leq \tilde{C} \left( 1 + \int_{\zeta_0}^{\zeta} Q(\zeta_1) d\zeta_1 \right) \tag{B.5}$$

for some polynomial  $Q(\zeta)$ . This proves the desired polynomial boundedness property for  $\tilde{\chi}$ .

### Acknowledgments

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

### References

- [1] R. Khasminskii, *Stochastic Stability of Differential Equations*, vol. 66, Springer Science & Business Media, 2011.
- [2] L. Arnold, *Stochastic Differential Equations: Theory and Applications*, Krieger Pub Co, 1992.
- [3] J.M. Ortega, *Matrix Theory: A Second Course*, Springer, 2013.

- [4] R. Zwanzig, *Nonequilibrium statistical mechanics*, Oxford University Press, 2001.
- [5] S. Hottovy, A. McDowell, G. Volpe and J. Wehr, *The Smoluchowski-Kramers limit of stochastic differential equations with arbitrary state-dependent friction*, *Commun. Math. Phys.* (2015) 336: 1259.
- [6] J. Birrell, J. Wehr, Phase space homogenization of noisy Hamiltonian systems. *Annales Henri Poincaré*, 19(4):1081–1114, Apr 2018.
- [7] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, Springer New York, 2014.
- [8] J. Birrell and J. Wehr, Homogenization of dissipative, noisy, Hamiltonian dynamics. *Stochastic Processes and their Applications*, 2017.
- [9] G. Pavliotis and A. Stuart, *Multiscale Methods: Averaging and Homogenization*, Texts in Applied Mathematics, Springer New York, 2008.

Jeremiah Birrell  
Department of Mathematics  
University of Arizona  
Tucson, AZ, 85721, USA  
e-mail: [jbirrell@math.arizona.edu](mailto:jbirrell@math.arizona.edu)

Jan Wehr  
Department of Mathematics and Program in Applied Mathematics  
University of Arizona  
Tucson, AZ, 85721, USA  
e-mail: [wehr@math.arizona.edu](mailto:wehr@math.arizona.edu)