Multiscale Method for Slow-Fast SDE Systems

Soon Hoe Lim*

Abstract

Many models of noisy systems are too complex to be solved analytically, or even numerically if a large range of time scales is involved. For a large class of high-dimensional systems it may be possible to derive lower-dimensional reduced models. The reduced model is often simpler to solve analytically and faster to integrate numerically, while still retaining the essential features of the full system. This short note demonstrates how one can apply a multiscale method to study model reduction for a class of slow-fast SDE systems as well as functionals along their trajectories. The main goal is to derive, at a formal level, a limiting SDE for the slow (resolved) variables *in the limit of infinite time scale separation*[1], thereby eliminating the fast (unresolved) variables from the description. The derived SDE can usually be justified rigorously by a homogenization theorem or verified by numerical experiments.

Contents

1 Notation and Background Materials

Consider a diffusion process $\boldsymbol{X}_t \in \mathbb{R}^d$, $t \geq 0$, satisfying the Itô SDE:

$$d\boldsymbol{X}_t = \boldsymbol{b}(t, \boldsymbol{X}_t)dt + \boldsymbol{\sigma}(t, \boldsymbol{X}_t)d\boldsymbol{W}_t, \tag{1}$$

*soon.hoe.lim@su.se

[1] It may be possible to derive effective SDEs in the case of finite time scale separation using the method in [7].

where $\boldsymbol{b} \in \mathbb{R}^d$, $\boldsymbol{\sigma} \in \mathbb{R}^{d\times m}$ is differentiable (in $\boldsymbol{X}$), and $\boldsymbol{W}_t \in \mathbb{R}^m$ is a Wiener process. Equivalently, it can be cast as the following Stratonovich SDE:

$$d\boldsymbol{X}_t = \boldsymbol{u}(t,\boldsymbol{X}_t)dt + \boldsymbol{\sigma}(t,\boldsymbol{X}_t) \circ d\boldsymbol{W}_t, \tag{2}$$

where $\boldsymbol{u}(t,\boldsymbol{X}_t) = \boldsymbol{b}(t,\boldsymbol{X}_t) - \boldsymbol{c}(t,\boldsymbol{X}_t)$, $\circ$ denotes Stratonovich convention, and, in index-free notation,

$$\boldsymbol{c} = \frac{1}{2}[\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\boldsymbol{\sigma}^T) - \boldsymbol{\sigma}\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}^T)], \tag{3}$$

or, in components,

$$c^i = \frac{1}{2}\frac{\partial \sigma^{ij}}{\partial X^k}\sigma^{kj}. \tag{4}$$

In the above, $\boldsymbol{\nabla}\cdot$ denotes divergence operator which contracts a matrix-valued function to a vector-valued function: for the matrix-valued function $\boldsymbol{A}(\boldsymbol{X})$, the ith component of its divergence is given by $(\boldsymbol{\nabla}\cdot\boldsymbol{A})^i = \sum_j \frac{\partial A^{ij}}{\partial X^j}$. The superscript T denotes transposition. We have used Einstein's summation convention for repeated indices.

2 Homogenization of Slow-Fast SDE Systems and Their Functionals

Consider the following Itô SDE system for $\boldsymbol{Z}_t = (\boldsymbol{X}_t, \boldsymbol{Y}_t, A_t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$:

$$d\boldsymbol{X}_t = \boldsymbol{u}_0(\boldsymbol{X}_t)dt + \frac{1}{\epsilon}\boldsymbol{U}_1(\boldsymbol{X}_t)\boldsymbol{Y}_t dt + \boldsymbol{\sigma}_0(\boldsymbol{X}_t)d\tilde{\boldsymbol{W}}_t, \tag{5}$$

$$d\boldsymbol{Y}_t = -\frac{1}{\epsilon^2}\boldsymbol{B}_2(\boldsymbol{X}_t)\boldsymbol{Y}_t dt + \frac{1}{\epsilon}\boldsymbol{b}_1(\boldsymbol{X}_t)dt + \frac{1}{\epsilon}\boldsymbol{\sigma}_1(\boldsymbol{X}_t)d\boldsymbol{W}_t, \tag{6}$$

$$dA_t = r(\boldsymbol{X}_t,\boldsymbol{Y}_t)dt + \frac{1}{\epsilon}q(\boldsymbol{X}_t,\boldsymbol{Y}_t)dt + \boldsymbol{p}(\boldsymbol{X}_t,\boldsymbol{Y}_t)\cdot d\boldsymbol{X}_t, \tag{7}$$

where $\tilde{\boldsymbol{W}}_t$ and $\boldsymbol{W}_t$ are independent Wiener processes. In the above, the variables $\boldsymbol{Z}_t$ are $O(1)$ as $\epsilon \to 0$. The coefficients in the SDEs above have sufficiently nice properties which allow us to justify our derivations later. The above systems are variants of the one considered in [1] (see also [3, 2]).

We assume that, for a given $\boldsymbol{X}$, the matrix-valued function $\boldsymbol{B}_2(\boldsymbol{X})$ is positive stable (i.e. all its eigenvalues have positive real part), and therefore it is invertible at any given $\boldsymbol{X}$. The diffusion matrices $\boldsymbol{\sigma}_i\boldsymbol{\sigma}_i^T$ $(i = 0, 1)$ are possibly degenerate, but it is still possible to work with the degenerate case under a more technical setting (see [6] for details), which we will bypass from now on. Here we have chosen to work with a specific fast process, which is the Vasicek model (viewing $\boldsymbol{X}$ as a parameter), to allow explicit expression for the effective coefficients to be derived later, rather than a general ergodic process such as the one considered in [6].

All the equations contain fast dynamics but the dynamics in $\boldsymbol{Y}$ is one order of magnitude faster than in $\boldsymbol{X}$ and A. Therefore, the $\boldsymbol{Z}$-dynamics can be simplified by eliminating the fast (unresolved) variable $\boldsymbol{Y}$ and described effectively by equations for the slow (resolved) variables $\boldsymbol{X}$ and A alone. *The main goal of this note is to derive an effective (homogenized) SDE for the slow process* $\boldsymbol{Q}_t = (\boldsymbol{X}_t, A_t)$ *in the limit* $\epsilon \to 0$.

We will study this problem using a formal perturbative expansion of infinitesimal generator (note that we can also study homogenization at the level of sample path) of the process $\boldsymbol{Z}_t$. The results obtained via this approach can be justified using the theorems in [6]. Although this gives very weak convergence result, the method is simple and fairly convenient to use. On the other hand, while the sample path approach leads to strong convergence result the analysis is often more technical, involving some intricate estimates – see, for instance, [4] or [5] (see also the relevant references therein for a list of previous works on homogenization).

The following statement is the main result of this note. We refer to Appendix A for a detailed derivation.

Result 2.1. For $\epsilon \ll 1$ and times up to $O(1)$, the process $\boldsymbol{Q}_t$, solving (5)-(7), is approximated by the solution to the following Itô SDE:

$$d\boldsymbol{Q}_t = \boldsymbol{F}(\boldsymbol{X}_t)dt + \boldsymbol{A}(\boldsymbol{X}_t)d\boldsymbol{W}_t, \tag{8}$$

where the drift vector $\boldsymbol{F} = (\boldsymbol{F}_1, F_2)$, with

$$\boldsymbol{F}_1(\boldsymbol{X}) = \boldsymbol{u}_0(\boldsymbol{X}) + \boldsymbol{U}_1(\boldsymbol{X})\boldsymbol{B}_2^{-1}(\boldsymbol{X})\boldsymbol{b}_1(\boldsymbol{X}) + \boldsymbol{S}(\boldsymbol{X}), \tag{9}$$

$$\begin{aligned} F_2(\boldsymbol{X}) &= \overline{r(\boldsymbol{X},\boldsymbol{Y})} + \overline{\boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y})} \cdot \boldsymbol{u}_0(\boldsymbol{X}) \\ &\quad + \boldsymbol{b}_1(\boldsymbol{X}) \cdot \overline{\boldsymbol{\nabla}_{\boldsymbol{Y}}(-\boldsymbol{L}_0^{-1}(\boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y}) \cdot (\boldsymbol{U}_1(\boldsymbol{X})\boldsymbol{Y}) + q(\boldsymbol{X},\boldsymbol{Y})))} \\ &\quad + \overline{(\boldsymbol{U}_1(\boldsymbol{X})\boldsymbol{Y}) \cdot \boldsymbol{\nabla}_{\boldsymbol{X}}(-\boldsymbol{L}_0^{-1}(\boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y}) \cdot (\boldsymbol{U}_1(\boldsymbol{X})\boldsymbol{Y}) + q(\boldsymbol{X},\boldsymbol{Y})))}. \end{aligned} \tag{10}$$

In the above,

- overbar denotes averaging with respect to the invariant density of a mean zero Gaussian process with the covariance matrix $\boldsymbol{J}$ that satisfies the Lyapunov equation[2] $\boldsymbol{B}_2\boldsymbol{J} + \boldsymbol{J}\boldsymbol{B}_2^T = \boldsymbol{\sigma}_1\boldsymbol{\sigma}_1^T$, and $\boldsymbol{L}_0 = -\boldsymbol{B}_2\boldsymbol{Y}\cdot\boldsymbol{\nabla}_{\boldsymbol{Y}} + \frac{1}{2}(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1^T) : \boldsymbol{\nabla}_{\boldsymbol{Y}}\boldsymbol{\nabla}_{\boldsymbol{Y}}$ is the infinitesimal generator associated with the fast dynamics, where $\boldsymbol{A} : \boldsymbol{\nabla}_{\boldsymbol{Y}}\boldsymbol{\nabla}_{\boldsymbol{Y}} := \sum_{i,j} A^{ij}\frac{\partial^2}{\partial Y^i \partial Y^j}$,
- $\boldsymbol{S} = \boldsymbol{\nabla}\cdot(\boldsymbol{U}_1\boldsymbol{B}_2^{-1}\boldsymbol{J}\boldsymbol{U}_1^T) - \boldsymbol{U}_1\boldsymbol{B}_2^{-1}\boldsymbol{\nabla}\cdot(\boldsymbol{J}\boldsymbol{U}_1^T)$ is the so-called *noise-induced drift*, whose ith component is
$$S^i = (U_1 J)^{lk}\frac{\partial}{\partial X^l}(U_1 B_2^{-1})^{ik}, \tag{11}$$
- the diffusion matrix $\boldsymbol{A}$ satisfies $\boldsymbol{A}\boldsymbol{A}^T = \frac{1}{2}(\boldsymbol{A}_0 + \boldsymbol{A}_0^T) + \boldsymbol{A}_1$, with
$$\boldsymbol{A}_0 = \begin{bmatrix} \boldsymbol{B}_0 & \boldsymbol{v}_0 \\ \boldsymbol{w}_0^T & u_0 \end{bmatrix}, \tag{12}$$
where
$$\boldsymbol{B}_0 = (\boldsymbol{U}_1\boldsymbol{\nu})(\boldsymbol{U}_1\boldsymbol{\nu})^T, \quad \text{with} \quad \boldsymbol{\nu} = \boldsymbol{B}_2^{-1}\boldsymbol{\sigma}_1, \tag{13}$$
$$\boldsymbol{v}_0 = 2\overline{\boldsymbol{U}_1\boldsymbol{y}(-\boldsymbol{L}_0^{-1}(\boldsymbol{p}\cdot(\boldsymbol{U}_1\boldsymbol{y}) + q))}, \tag{14}$$
$$\boldsymbol{w}_0^T = 2\overline{(\boldsymbol{p}\cdot(\boldsymbol{U}_1\boldsymbol{y}) + q)(-\boldsymbol{L}_0^{-1}(\boldsymbol{y}^T\boldsymbol{U}_1^T)}, \tag{15}$$
$$u_0 = 2\overline{(\boldsymbol{p}\cdot(\boldsymbol{U}_1\boldsymbol{y}) + q)(-\boldsymbol{L}_0^{-1}(\boldsymbol{p}\cdot(\boldsymbol{U}_1\boldsymbol{y}) + q))}, \tag{16}$$

[2]Positive stability of $\boldsymbol{B}_2$ ensures that there exists a unique solution to this equation.

and

$$\boldsymbol{A}_1 = \begin{bmatrix} \boldsymbol{\sigma}_0\boldsymbol{\sigma}_0^T & \frac{\boldsymbol{\sigma}_0\boldsymbol{\sigma}_0^T\overline{\boldsymbol{p}}}{(\boldsymbol{p}^T\boldsymbol{\sigma}_0)(\boldsymbol{p}^T\boldsymbol{\sigma}_0)^T} \\ (\boldsymbol{\sigma}_0\boldsymbol{\sigma}_0^T\overline{\boldsymbol{p}})^T & \end{bmatrix}. \tag{17}$$

It can be shown that $\boldsymbol{A}\boldsymbol{A}^T$ is positive semidefinite, for all $\boldsymbol{X}$. We remark that the convergence is only valid in the sense of weak convergence of probability measures and, therefore, the approximated bKe does not determine the limiting SDE uniquely (in particular, knowledge of $\boldsymbol{A}\boldsymbol{A}^T$ does not determine $\boldsymbol{A}$ uniquely).

Next we elaborate on the effective SDE for the functional A_t. Analogous results could be obtained in the case when the pre-limit SDEs have coefficients (for instance, r, $\boldsymbol{p}$ and q) that depend explicitly on time. In the case when the pre-limit SDE system is time-homogeneous, as was considered here, one can also represent the effective drift in the equation for the functional A_t in terms of time integral:

$$F_2(\boldsymbol{X}) = \overline{r(\boldsymbol{X},\boldsymbol{Y})} + \overline{\boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y})} \cdot \boldsymbol{u}_0(\boldsymbol{X}) + \int_0^\infty \mathbb{E}^{\mu_{\boldsymbol{X}}} G(\boldsymbol{X}, \phi^t_{\boldsymbol{X}}(\boldsymbol{Y}))dt, \tag{18}$$

whenever the time integral is well defined, where

$$\begin{aligned} G(\boldsymbol{X}, \phi^t_{\boldsymbol{X}}(\boldsymbol{Y})) = \boldsymbol{b}_1(\boldsymbol{X}) \cdot \boldsymbol{\nabla}_{\boldsymbol{Y}}(\boldsymbol{p}(\boldsymbol{X}, \phi^t_{\boldsymbol{X}}(\boldsymbol{Y})) \cdot (\boldsymbol{U}_1(\boldsymbol{X})\phi^t_{\boldsymbol{X}}(\boldsymbol{Y})) + q(\boldsymbol{X}, \phi^t_{\boldsymbol{X}}(\boldsymbol{Y}))) \\ + (\boldsymbol{U}_1(\boldsymbol{X})\phi^t_{\boldsymbol{X}}(\boldsymbol{Y})) \cdot \boldsymbol{\nabla}_{\boldsymbol{X}}(\boldsymbol{p}(\boldsymbol{X}, \phi^t_{\boldsymbol{X}}(\boldsymbol{Y})) \cdot (\boldsymbol{U}_1(\boldsymbol{X})\phi^t_{\boldsymbol{X}}(\boldsymbol{Y})) + q(\boldsymbol{X}, \phi^t_{\boldsymbol{X}}(\boldsymbol{Y}))), \end{aligned} \tag{19}$$

and $\mathbb{E}^{\mu_{\boldsymbol{X}}}$ denotes the product measure formed from distributing $\boldsymbol{Y}$ in its invariant measure, together with the Wiener process driving the equation for $\phi^t_{\boldsymbol{X}}(\boldsymbol{Y})$. An alternative representation, in terms of time averages, via the Birkhoff's ergodic theorem, can also be obtained – see [6].

We now discuss a particular case when $\boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y}) := \boldsymbol{p}_0(\boldsymbol{X})$ is independent of the fast variable. If, in addition, $q = 0$ and $\boldsymbol{\sigma}_0 = \boldsymbol{0}$, the effective SDE for A_t simplifies considerably. In particular, we have the following result.

Result 2.2. Let $\boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y}) := \boldsymbol{p}_0(\boldsymbol{X})$ be independent of the fast variable, $q = 0$ and $\boldsymbol{\sigma}_0 = \boldsymbol{0}$. Then, for $\epsilon \ll 1$ and times up to $O(1)$, the process A_t, solving (7), is approximated by the solution to the Itô SDE:

$$dA_t = \overline{r}dt + \boldsymbol{p}_0 \cdot d\boldsymbol{X}_t + dA'_t, \tag{20}$$

where $\boldsymbol{X}_t$ solves the Itô SDE:

$$d\boldsymbol{X}_t = (\boldsymbol{u}_0(\boldsymbol{X}_t) + \boldsymbol{U}_1(\boldsymbol{X}_t)\boldsymbol{B}_2^{-1}(\boldsymbol{X}_t)\boldsymbol{b}_1(\boldsymbol{X}_t) + \boldsymbol{S}(\boldsymbol{X}_t))dt + \boldsymbol{U}_1(\boldsymbol{X}_t)\boldsymbol{B}_2^{-1}(\boldsymbol{X}_t)\boldsymbol{\sigma}_1(\boldsymbol{X}_t)d\boldsymbol{U}_t, \tag{21}$$

with $\boldsymbol{S}$ the noise-induced drift as before, $\boldsymbol{U}_t$ a Wiener process, and

$$dA'_t = [\boldsymbol{\nabla} \cdot (\boldsymbol{p}_0^T \boldsymbol{U}_1 \boldsymbol{\mu} \boldsymbol{U}_1^T) - \boldsymbol{p}_0^T \boldsymbol{\nabla} \cdot (\boldsymbol{U}_1 \boldsymbol{\mu} \boldsymbol{U}_1^T)]dt = U_1^{ia} U_1^{jb} (B_2^{-1} J)^{ab} \frac{\partial p_0^i}{\partial X^j} dt, \tag{22}$$

or equivalently,

$$dA_t = \overline{r}dt + \boldsymbol{p}_0 \circ d\boldsymbol{X}_t + dA_t'', \tag{23}$$

where

$$dA_t'' = [\boldsymbol{\nabla} \cdot (\boldsymbol{p}_0^T \boldsymbol{U}_1 \boldsymbol{\mu}_A^T \boldsymbol{U}_1^T) - \boldsymbol{p}_0^T \boldsymbol{\nabla} \cdot (\boldsymbol{U}_1 \boldsymbol{\mu}_A^T \boldsymbol{U}_1^T)]dt = \frac{1}{2} U_1^{kb} U_1^{ja} \mu_A^{ab} \left(\frac{\partial p_0^j}{\partial X^k} - \frac{\partial p_0^k}{\partial X^j} \right) dt. \tag{24}$$

Therefore, in this special case we see that whenever $\boldsymbol{\mu}$ is symmetric, the effective SDE for the functional A_t can be expressed entirely in terms of the trajectory of the slow process with the Stratonovich definition[3].

3 References

[1] S. Bo and A. Celani, *White-noise limit of nonwhite nonequilibrium processes*, Physical Review E, 88 (2013), p. 062150.

[2] ——, *Entropy production in stochastic systems with fast and slow time-scales*, Journal of Statistical Physics, 154 (2014), pp. 1325–1351.

[3] ——, *Multiple-scale stochastic processes: decimation, averaging and beyond*, Physics Reports, 670 (2017), pp. 1–59.

[4] S. H. Lim and J. Wehr, *Homogenization for a class of generalized Langevin equations with an application to thermophoresis*, Journal of Statistical Physics, (2018).

[5] S. H. Lim, J. Wehr, and M. Lewenstein, *Homogenization for Generalized Langevin Equations with Applications to Anomalous Diffusion*, arXiv e-prints, (2019), p. arXiv:1902.06496.

[6] G. Pavliotis and A. Stuart, *Multiscale Methods: Averaging and Homogenization*, Springer Science & Business Media, 2008.

[7] J. Wouters and G. A. Gottwald, *Edgeworth expansions for slow–fast systems with finite time-scale separation*, Proceedings of the Royal Society A, 475 (2019), p. 20180358.

Appendix

A Derivation of Result 2.1

The infinitesimal generator of the process $\boldsymbol{Z}_t$ is

$$\boldsymbol{L} = \frac{1}{\epsilon^2} \boldsymbol{L}_0 + \frac{1}{\epsilon} \boldsymbol{L}_1 + \boldsymbol{L}_2, \tag{25}$$

[3] It turns out that the symmetry of $\boldsymbol{\mu}$ has intimate connection with reversibility of the fast process. This exploration kick-started our recent (and upcoming) work studying homogenization of functionals appearing in stochastic thermodynamics.

where

$$\boldsymbol{L}_0 = -\boldsymbol{B}_2(\boldsymbol{X})\boldsymbol{Y}\cdot\boldsymbol{\nabla}_{\boldsymbol{Y}} + \frac{1}{2}(\boldsymbol{\sigma}_1(\boldsymbol{X})\boldsymbol{\sigma}_1^T(\boldsymbol{X})) : \boldsymbol{\nabla}_{\boldsymbol{Y}}\boldsymbol{\nabla}_{\boldsymbol{Y}}, \tag{26}$$

$$\boldsymbol{L}_1 = \boldsymbol{U}_1(\boldsymbol{X})\boldsymbol{Y}\cdot\boldsymbol{\nabla}_{\boldsymbol{X}} + \boldsymbol{b}_1(\boldsymbol{X})\cdot\boldsymbol{\nabla}_{\boldsymbol{Y}} + \boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y})\cdot(\boldsymbol{U}_1(\boldsymbol{X})\boldsymbol{Y})\frac{\partial}{\partial A} + q(\boldsymbol{X},\boldsymbol{Y})\frac{\partial}{\partial A}, \tag{27}$$

$$\begin{aligned}\boldsymbol{L}_2 = {} & \boldsymbol{u}_0(\boldsymbol{X})\cdot\boldsymbol{\nabla}_{\boldsymbol{X}} + \frac{1}{2}(\boldsymbol{\sigma}_0(\boldsymbol{X})\boldsymbol{\sigma}_0^T(\boldsymbol{X})) : \boldsymbol{\nabla}_{\boldsymbol{Y}}\boldsymbol{\nabla}_{\boldsymbol{Y}} + r(\boldsymbol{X},\boldsymbol{Y})\frac{\partial}{\partial A} + \boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y})\cdot\boldsymbol{u}_0(\boldsymbol{X})\frac{\partial}{\partial A} \\ & + \frac{1}{2}(\boldsymbol{p}^T(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{\sigma}_0(\boldsymbol{X}))(\boldsymbol{p}^T(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{\sigma}_0(\boldsymbol{X}))^T\frac{\partial^2}{\partial A^2} + (\boldsymbol{\sigma}_0(\boldsymbol{X})\boldsymbol{\sigma}_0^T(\boldsymbol{X})\boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y}))\cdot\boldsymbol{\nabla}_{\boldsymbol{X}}\frac{\partial}{\partial A},\end{aligned} \tag{28}$$

where $\boldsymbol{A} : \boldsymbol{\nabla}_{\boldsymbol{Y}}\boldsymbol{\nabla}_{\boldsymbol{Y}} := \sum_{i,j} A^{ij}\frac{\partial^2}{\partial Y^i \partial Y^j}$.

The backward Kolmogorov equation (bKe) corresponding to the SDE system (5)-(7) is:

$$\frac{\partial \rho}{\partial t} = \boldsymbol{L}\rho, \tag{29}$$

where ρ is a function of $\boldsymbol{X}, \boldsymbol{Y}, A$ and t.

We seek a series expansion for the solution of (29) of the form $\rho = \rho_0 + \epsilon\rho_1 + \epsilon^2\rho_2 + \dots$. Substituting this expression into the bKe and equating terms of the same power in ϵ, we arrive at the following hierarchy of equations:

$$\boldsymbol{L}_0\rho_0 = 0, \tag{30}$$

$$\boldsymbol{L}_0\rho_1 + \boldsymbol{L}_1\rho_0 = 0, \tag{31}$$

$$\frac{\partial \rho_0}{\partial t} = \boldsymbol{L}_0\rho_2 + \boldsymbol{L}_1\rho_1 + \boldsymbol{L}_2\rho_0. \tag{32}$$

The generator $\boldsymbol{L}_0$, when viewed as a differential operator in $\boldsymbol{Y}$, in which $\boldsymbol{X}$ appears as a parameter, is the infinitesimal generator of an ergodic Markov process. It has one-dimensional null space characterized by $\boldsymbol{L}_0 1(\boldsymbol{Y}) = 0$ and $\boldsymbol{L}_0^*\rho^*(\boldsymbol{Y};\boldsymbol{X}) = 0$, where $1(\boldsymbol{Y})$ denotes constants in $\boldsymbol{Y}$, $\boldsymbol{L}_0^*$ is the adjoint of $\boldsymbol{L}_0$, and ρ^* is the density of an ergodic measure $\mu_{\boldsymbol{X}}(d\boldsymbol{Y}) = \rho^*(\boldsymbol{Y};\boldsymbol{X})d\boldsymbol{Y}$. Therefore, eqn. (30) and the ergodicity of the fast process imply that $\rho_0 = \rho(\boldsymbol{X}, A, t)$.

Eqn. (31) is the Poisson equation:

$$-\boldsymbol{L}_0\rho_1 = (\boldsymbol{U}_1(\boldsymbol{X})\boldsymbol{Y})\cdot\left(\boldsymbol{\nabla}_{\boldsymbol{X}}\rho + \boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y})\frac{\partial\rho}{\partial A}\right) + q(\boldsymbol{X},\boldsymbol{Y})\frac{\partial\rho}{\partial A}, \tag{33}$$

where $\boldsymbol{L}_0 = -(\boldsymbol{B}_2(\boldsymbol{X})\boldsymbol{Y})\cdot\boldsymbol{\nabla}_{\boldsymbol{Y}} + \frac{1}{2}(\boldsymbol{\sigma}_1(\boldsymbol{X})\boldsymbol{\sigma}_1(\boldsymbol{X})^T) : \boldsymbol{\nabla}_{\boldsymbol{Y}}\boldsymbol{\nabla}_{\boldsymbol{Y}}$. This is a PDE in $\boldsymbol{Y}$, with $\boldsymbol{X}$ a parameter. When we, in addition, require that the right hand side averages to zero with respect to ρ^* (centering condition), we refer to the Poisson equation (33) together with this condition as a cell problem [6].

Assuming that $\boldsymbol{p}$ and q are such that the right hand side in (33) satisfies the centering condition, the cell problem is then solvable (by the Fredholm alternative) and we solve it via separation of variables. The general solution of (33) has (up to a constant in the null space of $\boldsymbol{L}_0$, $Null(\boldsymbol{L}_0)$, and we are setting the constant to zero – this constant will not affect the limiting bKe) the form:

$$\rho_1 = \boldsymbol{\Phi}(\boldsymbol{X},\boldsymbol{Y})\cdot\boldsymbol{\nabla}_{\boldsymbol{X}}\rho + \phi(\boldsymbol{X},\boldsymbol{Y})\frac{\partial\rho}{\partial A} =: \tilde{\boldsymbol{\Phi}}(\tilde{\boldsymbol{X}},\boldsymbol{Y})\cdot\boldsymbol{\nabla}_{\tilde{\boldsymbol{X}}}\rho, \tag{34}$$

where $\tilde{\boldsymbol{X}} = (\boldsymbol{X}, A)$, for some vector-valued functions $\boldsymbol{\Phi}$, $\tilde{\boldsymbol{\Phi}}$, and scalar-valued function ϕ. We are going to use this solution and express the effective drift and diffusion coefficient of the limiting SDE in terms of $\boldsymbol{\Phi}$ and ϕ. We expect the term involving $\boldsymbol{U}_1$ in the $\boldsymbol{X}$ equation and the terms involving q and $\boldsymbol{p}$ in the A equation to contribute to (or, homogenize to) $O(1)$ effective drift and noise terms in the limiting equations in the limit $\epsilon \to 0$.

On the other hand, the formal solution of (33) can be written explicitly as:

$$\begin{aligned}\rho_1 &= -\boldsymbol{L}_0^{-1}\left(U_1^{ia}(\boldsymbol{X})Y^a\frac{\partial\rho}{\partial X^i} + p^i(\boldsymbol{X},\boldsymbol{Y})U_1^{ia}(\boldsymbol{X})Y^a\frac{\partial\rho}{\partial A} + q(\boldsymbol{X},\boldsymbol{Y})\frac{\partial\rho}{\partial A}\right) && (35)\\ &= -U_1^{ia}(\boldsymbol{X})\frac{\partial\rho}{\partial X^i}\boldsymbol{L}_0^{-1}(Y^a) - U_1^{ia}(\boldsymbol{X})\frac{\partial\rho}{\partial A}\boldsymbol{L}_0^{-1}(p^i(\boldsymbol{X},\boldsymbol{Y})Y^a) - \boldsymbol{L}_0^{-1}(q(\boldsymbol{X},\boldsymbol{Y}))\frac{\partial\rho}{\partial A}, && (36)\end{aligned}$$

where $-\boldsymbol{L}_0^{-1}$, being the inverse of a differential operator in the fast variable $\boldsymbol{Y}$ and containing $\boldsymbol{X}$ as a parameter, is an integral operator satisfying the time integral representation formula $(-\boldsymbol{L}_0)^{-1}(\boldsymbol{f}(\boldsymbol{X},\boldsymbol{Y})) = \int_0^\infty (e^{\boldsymbol{L}_0 t}\boldsymbol{f})(\boldsymbol{X},\boldsymbol{Y})dt = \int_0^\infty \mathbb{E}\boldsymbol{f}(\boldsymbol{X},\phi^t_{\boldsymbol{X}}(\boldsymbol{Y}))dt$ (see Result 11.8 in [6]), where $\mathbb{E}$ denotes expectation with respect to the Wiener measure and $\phi^t_{\boldsymbol{X}}(\boldsymbol{Y})$ is the solution operator of the fast dynamics (with $\boldsymbol{X}$ fixed) satisfying:

$$d\phi^t_{\boldsymbol{X}}(\boldsymbol{Y}) = -\boldsymbol{B}_2(\boldsymbol{X})\phi^t_{\boldsymbol{X}}(\boldsymbol{Y})dt + \boldsymbol{\sigma}_1(\boldsymbol{X})d\boldsymbol{W}_t,\ \phi^0_{\boldsymbol{X}}(\boldsymbol{Y}) = \boldsymbol{Y}. \tag{37}$$

Therefore, we can identify $\Phi^i = -U_1^{ia}(\boldsymbol{X})\boldsymbol{L}_0^{-1}(Y^a)$ and $\phi = -\boldsymbol{L}_0^{-1}(p^i(\boldsymbol{X},\boldsymbol{Y})U_1^{ia}(\boldsymbol{X})Y^a + q(\boldsymbol{X},\boldsymbol{Y}))$.

The solvability condition for (32), for each fixed $\boldsymbol{X}$, gives:

$$\frac{\partial\rho}{\partial t} = \int(\boldsymbol{L}_1\rho_1 + \boldsymbol{L}_2\rho_0)\rho^*(\boldsymbol{Y};\boldsymbol{X})d\boldsymbol{Y} = 0, \tag{38}$$

for all $\rho^* \in Null(L_0^*)$ and $\int \rho^*(\boldsymbol{Y};\boldsymbol{X})d\boldsymbol{Y} = 1$. In our case, ρ^* is the invariant density of a mean zero Gaussian process with the covariance matrix, $\boldsymbol{J} = \boldsymbol{J}(\boldsymbol{X})$, satisfying the Lyapunov equation

$$\boldsymbol{B}_2(\boldsymbol{X})\boldsymbol{J}(\boldsymbol{X}) + \boldsymbol{J}(\boldsymbol{X})\boldsymbol{B}_2^T(\boldsymbol{X}) = \boldsymbol{\sigma}_1(\boldsymbol{X})\boldsymbol{\sigma}_1^T(\boldsymbol{X}). \tag{39}$$

Writing $\boldsymbol{L}_1$ as $\boldsymbol{L}_1 = \boldsymbol{\theta}(\boldsymbol{X},\boldsymbol{Y})\cdot\boldsymbol{\nabla}_{\tilde{\boldsymbol{X}}} + \boldsymbol{b}_1(\boldsymbol{X})\cdot\boldsymbol{\nabla}_{\boldsymbol{Y}}$, where $\boldsymbol{\theta} = (\boldsymbol{U}_1(\boldsymbol{X})\boldsymbol{Y}, (\boldsymbol{U}_1(\boldsymbol{X})\boldsymbol{Y})\cdot\boldsymbol{p}(\boldsymbol{X},\boldsymbol{Y}) + q(\boldsymbol{X},\boldsymbol{Y}))$, and using the expression (34) for ρ_1, we find:

$$\begin{aligned}\boldsymbol{L}_1\rho_1 = \boldsymbol{\theta}(\boldsymbol{X},\boldsymbol{Y})\otimes\tilde{\boldsymbol{\Phi}}(\boldsymbol{X},\boldsymbol{Y}) : \boldsymbol{\nabla}_{\tilde{\boldsymbol{X}}}\boldsymbol{\nabla}_{\tilde{\boldsymbol{X}}}\rho + (\boldsymbol{\nabla}_{\tilde{\boldsymbol{X}}}\tilde{\boldsymbol{\Phi}}(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{\theta}(\boldsymbol{X},\boldsymbol{Y}))\cdot\boldsymbol{\nabla}_{\tilde{\boldsymbol{X}}}\rho\\ + (\boldsymbol{\nabla}_{\boldsymbol{Y}}\tilde{\boldsymbol{\Phi}}(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{b}_1(\boldsymbol{X}))\cdot\boldsymbol{\nabla}_{\tilde{\boldsymbol{X}}}\rho.\end{aligned} \tag{40}$$

Let overbar denote averaging with respect to ρ^* in the following. Working with (38), one computes the limiting bKe, from which one identifies the associated Itô SDE for the joint process $\boldsymbol{Q}_t = (\boldsymbol{X}_t, A_t)$ – this is exactly the SDE in Result 2.1.