6. Deviations from reduced models: correcting Brownian motion

Recall the idea of the Averaging Principle as introduced in Section 3.2. Remember moreover that we have obtained averaged diffusion equations for the reaction coordinates by artificially accelerating the dynamics of the unresolved modes. The averaged equations can then be considered the asymptotic result of the (singular) limit of infinite time scale separation between the reaction coordinate and the remaining degrees of freedom.

Correspondingly, we shall briefly sketch possible scenarios where this asymptotic strategy fails to capture the effective dynamics, e.g., due to a lack time scale separation or due to metastability in the unresolved modes. Anyway in realistic examples there is no control over the scale separation; in fact there is no small parameter at all.

6.1. Moderate deviations from the Averaging Principle

For the sake of clarity we restrict our attention to the case of a linear reaction coordinate. (The generalization to curvilinear reaction coordinates is straightforward using the results of Section 2.3.) Accordingly, we consider a Smoluchowski equation with separated slow and fast modes (x, y) and a potential $V : \mathbf{R}^s \times \mathbf{R}^k \to \mathbf{R}$, viz.,

$$\begin{aligned} \dot{x}_\epsilon(t) &= -\mathbf{D}_1 V(x_\epsilon(t), y_\epsilon(t)) + \sigma \, \dot{W}_1(t) \\ \dot{y}_\epsilon(t) &= -\frac{1}{\epsilon} \mathbf{D}_2 V(x_\epsilon(t), y_\epsilon(t)) + \frac{\sigma}{\sqrt{\epsilon}} \, \dot{W}_2(t) \,. \end{aligned} \tag{6.1}$$

Here $\sigma^2 = 2/\beta$. As we know from Proposition 3.9, for $\epsilon \to 0$ the slow process $x_\epsilon(t)$ converges pathwise to a Markov process $x(t)$ for all $t \in [0, T]$ which is governed by

$$\dot{x}(t) = -\nabla G(x(t)) + \sigma \, \dot{W}_1(t) \,, \tag{6.2}$$

where the averaged potential G is the (geometric) free energy,

$$G(x) = -\beta^{-1} \ln \int \exp(-\beta V(x, y)) \, dy \,.$$

6.1.1. Central Limit Theorem: fluctuations from equilibrium We want to study the error of the averaged motion, $x_\epsilon(t) - x(t)$, on a fixed time interval $[0, T]$. It was discovered by Khas'minskii [85] that the normalized error

$$\xi_\epsilon(t) = \frac{x_\epsilon(t) - x(t)}{\sqrt{\epsilon}}$$

has a limiting distribution for $\epsilon \to 0$ that is Gaussian.[28] What is interesting to note is the difference to deterministic systems, for which it is possible to asymptotically expand the error $x_\epsilon(t) - x(t) = \epsilon \xi_1(t) + \epsilon^2 \xi_2(t) + \dots$ in powers of the small parameter. For diffusive systems it can be shown however, that the error is of order $\sqrt{\epsilon}$, where all higher-order error terms are exponentially small [24]. Hence no further terms of the asymptotic expansion can be written down.

Suppose that $\mathbf{D}_1 V(x, y)$ has bounded first and second derivatives in x. We continue the analysis of the error by setting $x_\epsilon(t) = x(t) + \sqrt{\epsilon} \xi_\epsilon(t)$. Then

$$\dot{\xi}_\epsilon(t) = \frac{1}{\sqrt{\epsilon}} \left(\mathbf{D}_1 V(x(t) + \sqrt{\epsilon} \xi_\epsilon(t), y_\epsilon(t)) - \nabla G(x(t)) \right) . \tag{6.3}$$

[28] We write $\xi_\epsilon(t)$ rather than $\xi_\epsilon(t, \omega)$, where $\omega \in \Omega$ is an element of some set of elementary events. However the reader should keep in mind that $\xi_\epsilon(t)$ is a stochastic process that depends on the realizations of the white noise, i.,e., $\xi_\epsilon(t, \cdot)$ is a random variable for each $t \in [0, T]$.

Expanding the right hand side formally in terms of ϵ, we obtain [185]

$$\begin{aligned}\dot{\xi}_\epsilon(t) = \frac{1}{\sqrt{\epsilon}} &\left(\mathbf{D}_1 V(x(t), y_\epsilon(t)) - \nabla G(x(t))\right) \\ &+ \mathbf{D}_1^2 V(x(t), y_\epsilon(t)) \cdot \xi_\epsilon(t) + \mathcal{O}(\epsilon^\infty)\,.\end{aligned}$$

The remainder $\mathcal{O}(\epsilon^\infty)$ is far from obvious, but we refer the reader to the article [86] and the references given there. Now recall that $y_\epsilon(t) = y(t/\epsilon)$. Integrating yields

$$\begin{aligned}\xi_\epsilon(t) = \frac{1}{\sqrt{\epsilon}} &\int_0^t \left(\mathbf{D}_1 V(x(s), y(s/\epsilon)) - \nabla G(x(s))\right) ds \\ &+ \int_0^t \mathbf{D}_1^2 V(x(s), y(s/\epsilon)) \cdot \xi_\epsilon(s)\, ds + \mathcal{O}(\epsilon^\infty)\,.\end{aligned}$$

Obviously the second term on the right hand side converges in the way that

$$\int_0^t \mathbf{D}_1^2 V(x(s), y(s/\epsilon)) \cdot \xi_\epsilon(s)\, ds \to \int_0^t \nabla^2 G(x(s)) \cdot \xi(s)\, ds \tag{6.4}$$

as $\epsilon \to 0$. Upon rescaling time according to $s \mapsto \epsilon s$, the first term becomes

$$\frac{1}{\sqrt{\epsilon}} \int_0^{t/\epsilon} \left(\mathbf{D}_1 V(x(\epsilon s), y(s)) - \nabla G(x(\epsilon s))\right) ds\,.$$

Letting ϵ going to zero, the last expression would be simply an instance of the ordinary Central Limit Theorem, if there were no time dependence in x. But it has been proved that the integral (6.4) converges weakly to a Gaussian Markov process [85]. Abbreviating $f(x, y) = \mathbf{D}_1 V(x, y)$, this Gaussian process has the covariance matrix

$$a(x)a(x)^T = \lim_{T\to\infty} \int_0^T \int_0^T \operatorname{cov}\left(f(x, y_x(s)), f(x, y_x(t))\right) ds dt\,, \tag{6.5}$$

where $y_x(t)$ is the solution of the fast dynamics for a fixed value of x, and $\operatorname{cov}(f, g)$ denotes the covariance of two random vector fields f, g. (Note that $y_x(t)$ for fixed t is a random variable by virtue of the different realizations of the Brownian motion.) The limit of the normalized deviation satisfies the following family of equations

$$\dot{\xi}(t) = \nabla^2 G(x) \cdot \xi(t) + a(x) \cdot \dot{W}_\xi(t)$$

with $\dot{W}_\xi(t)$ denoting standard Brownian motion in $\mathbf{R}^k$. As a result the averaged equation (6.2) is replaced by the following pair of equations

$$\begin{aligned}\dot{x}(t) &= -\nabla G(x(t)) + \sigma\, \dot{W}_1(t) \\ \dot{\xi}(t) &= \nabla^2 G(x(t)) \cdot \xi(t) + a(x(t)) \cdot \dot{W}_\xi(t)\,.\end{aligned} \tag{6.6}$$

Equation (6.6) has an interesting structure, a so-called *skew product structure*. That is, the slow equation for x is decoupled from the equation for ξ, while the deviations depend on the solution of the averaged equation. Therefore it is by no means obvious how to recover the full solution from the averaged one, including the deviations. One possibility to do so is by employing *van Kampen's approximation* [296]

$$x_\epsilon(t) \approx x(t) + \sqrt{\epsilon}\xi(t)\,, \tag{6.7}$$

where the "$\approx$" symbol should not be taken literally, for convergence $\xi_\epsilon(t) \to \xi(t)$ was only in the sense of probability distributions, whereas we had convergence in probability with regard to $x_\epsilon(t) \to x(t)$. Another frequently used approach, e.g., [297], is to add some extra noise to the averaged solution, while omitting the drift:

$$\dot{x}_\epsilon(t) = -\nabla G(x_\epsilon(t)) + \left(\sigma + \sqrt{\epsilon} a(x_\epsilon(t))\right) \dot{W}_1(t) \tag{6.8}$$

All these approximations are in some sense obvious, but *ad-hoc*. Therefore we refer to the steps (6.7) and (6.8) as *remodelling* of the full dynamics. Nevertheless the method offers interesting opportunities for the practical implementation, since the covariance matrix (6.5) can be computed numerically running constrained simulations of the fast dynamics at fixed values of the slow coordinate.

Example 6.1. For the sake of illustration we reconsider our guiding example 3.20:

$$\begin{aligned}\dot{x}_\epsilon(t) &= -\partial_x V(x_\epsilon(t), y_\epsilon(t)) + \sigma\, \dot{W}_1(t)\\ \dot{y}_\epsilon(t) &= -\frac{1}{\epsilon}\partial_y V(x_\epsilon(t), y_\epsilon(t)) + \frac{\sigma}{\sqrt{\epsilon}}\, \dot{W}_2(t)\,.\end{aligned}$$

Here $(x, y) \in \mathbf{R} \times \mathbf{R}$, and the potential is given by

$$V(x, y) = \frac{1}{4}\left(x^2 - 1\right)^2 + \frac{1}{2}\,\lambda(x)^2 y^2$$

with the function $\lambda(x) \geq c > 0$ defined by

$$\lambda(x) = 1 + C \exp\left(-\alpha(x - x_b)^2\right)\,.$$

Note that the ϵ-scaling is slightly different from the previous occurrences (see Example 3.20, for instance). The skew-structured limit equation then reads

$$\begin{aligned}\dot{x}(t) &= -\partial_x G(x(t)) + \sigma\, \dot{W}_1(t)\\ \dot{\xi}(t) &= \partial_x^2 G(x(t))\xi(t) + a(x(t))\, \dot{W}_\xi(t)\,,\end{aligned} \tag{6.9}$$

where G is the free energy exhibiting the entropic (dynamical) barrier,

$$G(x) = \frac{1}{4}\left(x^2 - 1\right)^2 + \beta^{-1} \ln \lambda(x)\,,$$

and

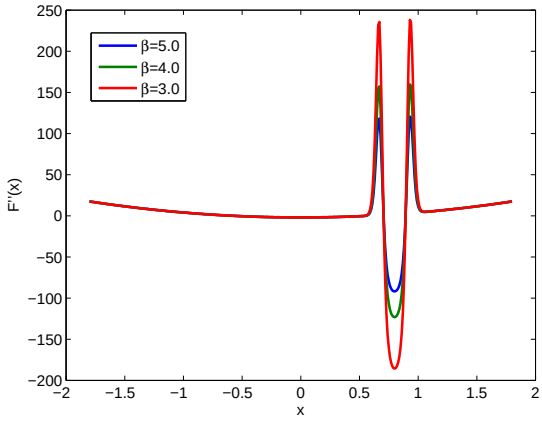
$$a(x) = \frac{1}{\beta}\left|\frac{\lambda'(x)}{\lambda(x)^2}\right|\,.$$

is the standard deviation of the Gaussian error which can be computed analytically, if we assume that the fast process is ergodic (see Appendix F). Note that the second derivative of $F(x)$, which appears in the limit equation of the error is not negative definite (see Figure 39). In fact the error is completely governed by the dynamical barrier at $x = x_b$, where both $|\lambda'(x)|$ and $|\lambda''(x)|$ attain their maximum values. This is in good agreement with the numerical results in Example 3.20.

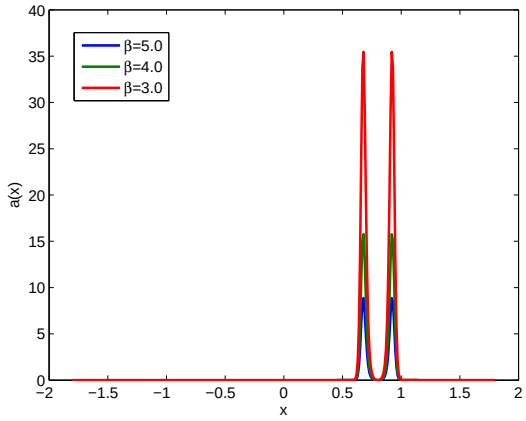
We emphasize that the averaged dynamics $x(t)$ already provides the *best-approximation* to the full dynamics $x_\epsilon(t)$ for reasonably small ϵ. Therefore, including the error does by no means improve the approximation quality of the limit dynamics. It rather serves as a vehicle to increase the variability of the averaged model, in order to incorporate certain fluctuations, that would not be present otherwise. Moreover it turns out that incorporating moderate deviations can account for situations where the scale separation is not good, i.e., ϵ is not small.

6.2. Large deviations from the Averaging Principle

The method of averaging and the Central Limit Theorem make assertions for the dynamics on the finite time intervals. Typically they both may become invalid on diverging time intervals of order $1/\epsilon$ or even $\exp(1/\epsilon)$. However there are many interesting phenomena that occur on such time scales (rare events), and which are not always captured by the Averaging Principle. One such case is the *hopping* of the fast dynamics between local attractors (metastable conformations), which may happen on time scales that are beyond the time scale of the slowest modes in the system.



(a)



(b)

Figure 39. These plots illustrate the coefficients in the equations of motion (6.9) for the error for various inverse temperatures $\beta \in \{5.0, 4.0, 3.0\}$. Again $\beta = 3.0$ labels the most noticeable peak at $x = x_b$, whereas the little one corresponds to $\beta = 5.0$. Note first of all the error achieves its maximum in the vicinity of the dynamical barrier, and furthermore that the second derivative of $F(x)$ that drives the averaging error is not negative definite.

6.2.1. Diffusive limits Reconsider the generic slow-fast system (3.1). Here we seek to derive a reduced equation on the longer, diffusive, time scale of order $1/\epsilon$. This may be of interest, if the averaged system is deterministic on the $\mathcal{O}(1)$ time scale with vanishing drift, whereas the fast dynamics is stochastic. (One such case is the high-friction limit of the Langevin equation on the observation time scale of order 1.)

Changing the free variable in (6.1) according to $t \mapsto t/\epsilon$ while omitting the noise in the slow equation, we obtain the following system on the diffusive time scale

$$\begin{aligned} \dot{x}_\epsilon(t) &= -\frac{1}{\epsilon}\mathbf{D}_1 V(x_\epsilon(t), y_\epsilon(t)) \\ \dot{y}_\epsilon(t) &= -\frac{1}{\epsilon^2}\mathbf{D}_2 V(x_\epsilon(t), y_\epsilon(t)) + \frac{\sigma}{\epsilon}\,\dot{W}_2(t)\,. \end{aligned} \tag{6.10}$$

We assume that the slow dynamics averages to zero under the fast process, i.e.,

$$\int_{\mathbf{R}^k} \mathbf{D}_1 V(x,y)\mu_x(dy) = 0\,, \tag{6.11}$$

where $\mu_x(dy) \propto \exp(-\beta V(x,y))dy$ is the invariant Gibbs measure of the fast process $y_x(t)$. It can then be shown that the slow drift gives significant contributions on the diffusive time scale by coupling to the noise of the fast equation as $\epsilon \to 0$, hence the term *diffusive limit*; see the rich literature, e.g., [33, 298, 23]. Associated with the system (6.10) is the Kolmogorov backward generator

$$\mathcal{A}^\epsilon = \frac{1}{\epsilon^2}\mathcal{A}_1 + \frac{1}{\epsilon}\mathcal{A}_2$$

with

$$\begin{aligned}\mathcal{A}_1 &= \frac{\sigma^2}{2}\,\mathrm{tr}\,\mathbf{D}_2^2 - \mathbf{D}_2 V(x,y)\cdot\mathbf{D}_2\\ \mathcal{A}_2 &= -\mathbf{D}_1 V(x,y)\cdot\mathbf{D}_1\,.\end{aligned}$$

By adopting arguments from semigroup perturbation theory, e.g., [299], we can expand the solution of the backward equation into a power series in ϵ. Equating coefficients of equal powers gives rise to reduced equations in terms of the slow variable x, namely

$$\dot{x}(t) = b(x(t)) + a(x(t))\,\dot{W}(t)\,. \tag{6.12}$$

Here $\dot{W}(t)$ denotes standard Brownian motion in $\mathbf{R}^s$. As is shown in Appendix E, the drift and the noise coefficients are given by the expressions

$$\begin{aligned}a(x)a(x)^T &= 2\int \mathbf{D}_1 V_1(x,y)^T \int_0^\infty \mathbf{E}_y \mathbf{D}_1 V(x, y_x(t))\,dt\,\mu_x(dy)\\ b(x) \qquad &= \int \mathbf{D}_1 V(x,y) \int_0^\infty \mathbf{E}_y \mathbf{D}_1^2 V(x, y_x(t))\,dt\,\mu_x(dy)\,,\end{aligned}$$

where $y_x(t)$ denotes the fast process at time t starting at $y_x(0) = y$, and $\mathbf{E}_y$ labels the average over all realizations up to time t conditional on the initial value y.

A few important bibliographical remarks are in order: Firstly, giving a rigorous proof of the diffusive limit equation is far beyond the scope of the present thesis, and refer to the appendix for a rough sketch of derivation using a perturbative expansion of the backward equation. Yet another issue is convergence: the vast majority of the results in the literature deals with weak convergence $x_\epsilon(t) \to x(t)$ on condition that the centering condition (6.11) is satisfied. Stronger results are available though; see, e.g., the original paper by Khas'minskii [33] or the textbook [125]. Other papers, like [34], relax the centering condition demanding for a two-scale expansion of the backward equation (which is much more challenging). See the paper [300] for a numerical scheme, which is based upon a multiple time stepping discretization of the original system (6.10) that can be used to simulate the limit system (6.12) quite effectively.

Remark 6.2. *The reader may wonder what happens if the slow equation already contains some noise on the $\mathcal{O}(1)$ time scale, i.e., if we face a situation like*

$$\begin{aligned}\dot{x}_\epsilon(t) &= -\frac{1}{\epsilon}\mathbf{D}_1 V(x_\epsilon(t), y_\epsilon(t)) + \frac{\sigma}{\sqrt{\epsilon}}\dot{W}_1(t)\\ \dot{y}_\epsilon(t) &= -\frac{1}{\epsilon^2}\mathbf{D}_2 V(x_\epsilon(t), y_\epsilon(t)) + \frac{\sigma}{\epsilon}\,\dot{W}_2(t)\,.\end{aligned}$$

Apparently the system admits an invariant measure, the Gibbs measure

$$\mu(dx,dy) = \frac{1}{Z}\exp(-\beta V(x,y))\,dxdy\,,$$

where the normalization constant Z is defined by

$$Z = \int_{\mathbf{R}^s \times \mathbf{R}^k} \exp(-\beta V(x,y))\, dxdy\,.$$

Existence of the Gibbs measure, however, contradicts the centering condition (6.11), as can be easily seen by interchanging the order of integration in the last equation,

$$Z = \int_{\mathbf{R}^s} \underbrace{\left(\int_{\mathbf{R}^k} \exp(-\beta V(x,y))\, dy \right)}_{\text{constant due to (6.11)}} dx\,.$$

The centering condition is crucial for the derivation of the diffusive limit equation in the present form (see Remark F.1 below). In fact it seems that diffusive limits are more targeted on systems with deterministic right hand side, subject to random perturbations stemming from the fast variables. This is slightly different from problems usually considered in molecular dynamics. However the method is useful for studying hypo-elliptic diffusion processes, such as the high-friction limit of the Langevin equation.

6.2.2. Metastability and conditional averaging Roughly speaking the Averaging Principle relies on the idea that the fast dynamics explores its state space, sampling its invariant measure, while the slow dynamics is at rest. Therefore averaging fails, if there is some subset of state space, in which the fast dynamics is likely to get trapped. Consider the slow-fast system (6.1), and assume that the fast subsystem

$$\dot{y}_x(t) = -\mathbf{D}_2 V(x, y_x(t)) + \sigma\, \dot{W}_2(t)$$

is ergodic with respect to the conditional Gibbs measure $\mu_x(dy)$. Assume further that for some values of $x \in \mathbf{R}^s$ the fast dynamics shows metastability. For example, we can assume that for a particular value x_{crit} the potential $V(x,y)$, considered along the fast direction, has a significant potential barrier $\Delta V(x_{\text{crit}})$ separating two wells. If $2\Delta V(x_{\text{crit}}) \gg -\sigma^2 \ln \epsilon$ then Large Deviation Theory explains that the exit time from the potential wells induces an additional time scale that is of the order

$$\tau_{\epsilon,\sigma} \sim \epsilon \exp\left(\beta \Delta V(x_{\text{crit}})\right) \quad (\beta = 2/\sigma^2 \gg 1)\,. \tag{6.13}$$

The last equation can be considered as some sort of Arrhenius law; for a mathematical justification we refer to the standard textbook on Large Deviation Theory by Freidlin and Wentzell [24]. It is easy to see that, at low temperature, $\tau_{\epsilon,\sigma}$ exceeds any other time scale in the system. In turn, rapid mixing of the fast variable then requires ϵ to be exponentially small as compared to the noise level σ . We would like to study the system (6.1) in the case that $\tau_{\epsilon,\sigma} = \mathcal{O}(1)$. Fixing the order of exit times in that way amounts to a scaling relation between ϵ and σ by virtue of (6.13).

We shall briefly explain the basic idea of the *conditional averaging* approach that has been put forward in [35], and which has been refined recently in [165, 301]. To this end assume that for each x, we can identify two more or less metastable sets $B_1(x), B_2(x) \subset \mathbf{R}^k$. (Here $\mathbf{R}^k$ denotes the state space of the fast variables.) As one observes that the fast process $y_x(t)$ is rapidly mixing inside each metastable set $B_1(x)$ or $B_2(x)$, respectively, it makes sense to average the slow dynamics with respect to the (almost invariant) probability measures on each metastable set. This results in two locally averaged models, one for $B_1(x)$ and another one for $B_2(x)$, which are coupled by means of a rate matrix that governs the transitions between the metastable sets. The reduced model is then a time inhomogeneous model of the form

$$\dot{x}(t) = -\nabla G_{i(t)}(x(t)) + \sigma \dot{W}_1(t)\,, \tag{6.14}$$

where G_i is the local free energy

$$G_i(x) = -\beta^{-1} \ln \int_{\mathbf{R}^k} \mu_{x,i}(dy) \tag{6.15}$$

with $\mu_{x,i}(dy) = (\mu_x | B_i(x))(dy)$, appropriately normalized. The switching $i(t) = i(t,x)$ is a two-state Markov jump process, that mimics the transition between $B_1(x)$ and $B_2(x)$. Under certain conditions the rates of the jump process are determined by the second dominant eigenvalue $\lambda_2(x)$ of the infinitesimal generator of the fast process $y_x(t)$, as has been demonstrated in [35] (however in a non-rigorous fashion).

A more detailed multiscale analysis is carried out in the PhD thesis [165]. There the author also exposes how the metastable fast dynamics can be approximated in some L^1-sense by a family of Ornstein-Uhlenbeck processes that are coupled by appropriately designed Markov jump processes; cf. also the results in [302, 303].

Realization as a stochastic particle method The method of conditional averaging allows for an elegant numerical realization as a stochastic particle method which makes it accessible for practical applications. The discretization is based on a Trotter splitting of the generator associated with the conditionally averaged system (6.14), an idea borrowed from so-called *surface hopping* algorithms in quantum-classical dynamics [304, 305]. To derive the numerical scheme, consider the Fokker-Planck equation associated with the original system (6.1)

$$\partial_t \rho(x,y,t) = \mathcal{L}\rho(x,y,t)\,, \quad u(x,y,0) = g(x)\,,$$

where $\mathcal{L}$ is the backward generator in the semi-weighted Hilbert space $L^2(\mu_x)$. Given two families of metastable sets $B_1(x)$ and $B_2(x)$, we seek for a Galerkin decomposition of the full solution of the Fokker-Planck equation in the form

$$\rho(x,y,t) = c_1(x,t)\chi_1(x,y) + c_2(x,t)\chi_2(x,y)\,.$$

Here χ_1, χ_2 span the two-dimensional dominant subspace of the fast dynamics' generator. Provided certain conditions are met (e.g., regularity of the boundary between the metastable sets), then it has been shown in [35] that the Fokker-Planck equation associated with the reduced system (6.14)–(6.15) has the following intuitive representation in terms of the coefficient vector $c = (c_1, c_2)^T$, namely

$$\partial c(x,t) = \left(\bar{\mathcal{L}} + \bar{\mathcal{R}}\right) c(x,t)$$

with $\bar{\mathcal{L}}$ containing the generators $\mathcal{L}_i$ of the locally averaged systems (6.14)

$$\bar{\mathcal{L}} = \begin{pmatrix} \mathcal{L}_1 & 0 \\ 0 & \mathcal{L}_2 \end{pmatrix},$$

and a rate matrix $\bar{\mathcal{R}} \in \mathbf{R}^{2\times 2}$ that provides the switching between the states $i \in \{1,2\}$. Here the rate depends on the second eigenvalue $\lambda_2(x)$ of the generator of the fast dynamics $y_x(t)$. A time discretization at time step $h = \mathcal{O}(\epsilon^2)$ is obtained by a splitting,

$$\exp\left(h(\bar{\mathcal{L}} + \bar{\mathcal{R}})\right) = \exp\left(h\bar{\mathcal{L}}\right) \exp\left(h\bar{\mathcal{R}}\right) + \mathcal{O}(h^2)\,,$$

of the propagator. The thus defined propagator has a nice pathwise interpretation: apparently the first exponential simply gives the propagation according to the locally averaged equations (6.14) up to time h, where the second one represents the exchange between the two states $i = 1$ and $i = 2$. In point of fact, $\exp(h\bar{\mathcal{R}})$ is a stochastic matrix for all $h > 0$, i.e., it is the transition matrix of the Markov chain $\{i(0), i(h), i(2h), \ldots\}$.

Considering an ensemble of N particles $x_k(t)$, $k = 1, \dots N$ in the respective states $i_k \in \{1, 2\}$, the system (6.14)–(6.15) has the following realization as a stochastic particle method: Propagate each particle, $x_k(t) \mapsto x_k(t+h)$, by solving

$$\dot{x}_k(t) = -\nabla G_{i_k(t)}(x_k(t)) + \sigma \dot{W}_1(t)\,.$$

Then let each particle make a transition $i_k \mapsto j_k$ according to the transition probabilities contained in the stochastic matrix $\exp(h\bar{\mathcal{R}})$. If the $x_k(t)$ represent the ensemble $c(x,t)$ at time t, then the resulting ensemble at time $t+h$ yields a representation of $c(x, t+h)$. For the details we refer to [301].

Finally, we claim that conditional averaging provides a useful extension of the proposed reduction scheme for diffusion at low temperature. This presupposes that it is possible to estimate the transition rates between possible metastable sets along the fast variables which is typically difficult, whenever the state space of the unresolved variables is high-dimensional. One possible way then may be to resort to iterative schemes like [47], and confine the attention to very few fast variables, e.g., certain torsion angles, to get a rough estimate of metastabilities or exit times, respectively.

Unlike the extensions based on Central Limit Theorem or diffusive limits, the generalization of conditional averaging to Riemannian manifolds and curvilinear reaction coordinates is not straightforward; it requires a careful study of the boundary between the metastable sets and the restriction of the fast generator to these sets. This problem is under current investigation by the author.