CONVERGENCE OF EQUATION-FREE METHODS IN THE CASE OF FINITE TIME SCALE SEPARATION WITH APPLICATION TO DETERMINISTIC AND STOCHASTIC SYSTEMS

JAN SIEBER*, CHRISTIAN MARSCHLER†, AND JENS STARKE‡

Abstract. A common approach to studying high-dimensional systems with emergent low-dimensional behavior is based on lift-evolve-restrict maps (called equation-free methods): first, a user-defined lifting operator maps a set of low-dimensional coordinates into the high-dimensional phase space, then the high-dimensional (microscopic) evolution is applied for some time, and finally a user-defined restriction operator maps down into a low-dimensional space again. We prove convergence of equation-free methods for finite time-scale separation with respect to a method parameter, the so-called healing time. Our convergence result justifies equation-free methods as a tool for performing high-level tasks such as bifurcation analysis on high-dimensional systems.

More precisely, if the high-dimensional system has an attracting invariant manifold with smaller expansion and attraction rates in the tangential direction than in the transversal direction (normal hyperbolicity), and restriction and lifting satisfy some generic transversality conditions, then an implicit formulation of the lift-evolve-restrict procedure generates an approximate map that converges to the flow on the invariant manifold for healing time going to infinity. In contrast to all previous results, our result does not require the time scale separation to be large. A demonstration with Michaelis-Menten kinetics shows that the error estimates of our theorem are sharp.

The ability to achieve convergence even for finite time scale separation is especially important for applications involving stochastic systems, where the evolution occurs at the level of distributions, governed by the Fokker-Planck equation. In these applications the spectral gap is typically finite. We investigate a low-dimensional stochastic differential equation where the ratio between the decay rates of fast and slow variables is 2.

Key words. implicit equation-free methods, slow-fast systems, stochastic differential equations, Michaelis-Menten kinetics, dimension reduction

AMS subject classifications. 65Pxx, 37Mxx, 34E13

1. Introduction.

High-dimensional dynamical systems with time scale separation have under certain assumptions the potential to be studied and understood through a reduction to low-dimensional systems. In most cases these reduction methods are applied directly to the high-dimensional systems of equations [39]. The most common approaches are referred to as averaging and mean-field approximation [44], the slaving principle or adiabatic elimination [19] in the physics literature. The aim of these methods is to reduce the complexity of a high-dimensional (here also called *microscopic*) system to a relatively simple low-dimensional (here also called *macroscopic*) system. After reduction, the long-term dynamics of the system can be analyzed by studying the low-dimensional macroscopic system, using techniques that may only be available for low-dimensional deterministic systems (e.g., detailed bifurcation analysis). The underlying assumption is that a trajectory of the microscopic system will rapidly relax onto a low-dimensional manifold, which it will then track on a longer time scale, following the slower macroscopic equations. Thus, one speaks of *slow* variables, which are the coordinates on the slow low-dimensional manifold, and *fast* variables transversal to the slow manifold. The notion that the fast variables are "slaved" by

*College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter (Devon) EX4 4QF, United Kingdom (j.sieber@exeter.ac.uk).

†Department of Applied Mathematics and Computer Science, Technical University of Denmark, Matematiktorvet 303B, DK-2800 Kgs. Lyngby, Denmark.

‡Institute of Mathematics, University of Rostock, Ulmenstraße 69, 18057 Rostock, Germany (jens.starke@uni-rostock.de).

the slow variables describes that over long time the microscopic trajectories track the slow manifold.

The justification for this reduction is simplest and strongest if the underlying microscopic dynamical system possesses a low-dimensional attracting invariant manifold. In these cases mathematical theorems on persistence of invariant manifolds can be applied. Proofs were given by Fenichel [15] and Hirsch *et al.* [20] for finite-dimensional smooth dynamical systems such as ordinary differential equation (ODEs) and maps and by Bates *et al.* [6, 7] for general semiflows (covering certain classes of partial differential equations). Certain cases of averaging (such as periodic and quasi-periodic forcing) may also be reduced to invariant manifold persistence.

The case for model reduction is more subtle if the microscopic system is stochastic (or more generally, ergodic), for example, if the model is given by a multi-particle or agent-based simulation. The time-scale separation for these systems occurs if, for example, the number of particles is large. A model case for stochastic systems is the reduction of a high-dimensional system of stochastic differential equations (SDE) to a low-dimensional SDE acting on a slower time scale (smaller drift terms and smaller noise amplitude than the microscopic system). In this case, the arguments for model reduction look formally similar to the case of attracting manifolds in deterministic systems [36, 17]. However, the underlying mathematical convergence results are not as strong. Two aspects in which the deterministic results are stronger than the stochastic results will have implications on convergence results for computational methods:

Validity for finite time-scale separation: For invariant manifolds in a deterministic ODE the time scale separation (let us call it ε) is measured as the ratio between the rate of attraction along directions tangential to the manifold (d_{tan}) and transversal to the manifold (d_{tr}, so $\varepsilon = d_{\mathrm{tan}}/d_{\mathrm{tr}}$). As long as this ratio ε is less than unity, the manifold persists. Let us call this persistent low-dimensional manifold $\mathcal{C}$. Persistence implies that, even for a finite ε, a reduced model on this manifold $\mathcal{C}$ exists, describing some trajectories of the microscopic system with perfect accuracy (those that lie on $\mathcal{C}$). In practice the dynamics on the slow manifold $\mathcal{C}$ is often approximated by an expansion in ε.

Shadowing: Even more, every point u from an open neighborhood of $\mathcal{C}$ has a shadowing point $g(u) \in \mathcal{C}$. The difference between the trajectories starting from u and $g(u)$ goes to zero in time with a rate close to d_{tr}. This means that the reduced model describes *all* nearby trajectories even for positive ε with perfect accuracy except for rapidly decaying terms. The nonlinear projection $u \mapsto g(u) \in \mathcal{C}$ is called the *stable fiber projection*.

Compared to the above, the precise mathematical convergence statements in [36, 17] for stochastic systems with time scale separation are weaker. They are concerned with the limit $\varepsilon \to 0$ and prove that moments of the slow coordinates of the microscopic trajectory and of the trajectories of the slow model, derived by a formal expansion in ε, converge to each other for $\varepsilon \to 0$ [36].

On-demand computation of slow flow — Equation-free framework. The above mathematical theorems underpin the derivation of approximate low-dimensional models for large-dimensional systems. However, they also provide guidance for the convergence analysis of computational methods that avoid the explicit derivation of a low-dimensional model, but merely assume its existence. A general framework for analysing slow-time scale behaviour of systems with time scale separation was proposed by Kevrekidis *et al.* under the name "equation-free computations" [23, 16, 25]. The assumption behind equation-free computations is the existence of a slow low-dimensional description (in $\mathbb{R}^d$) for some macroscopic quantities of the high-

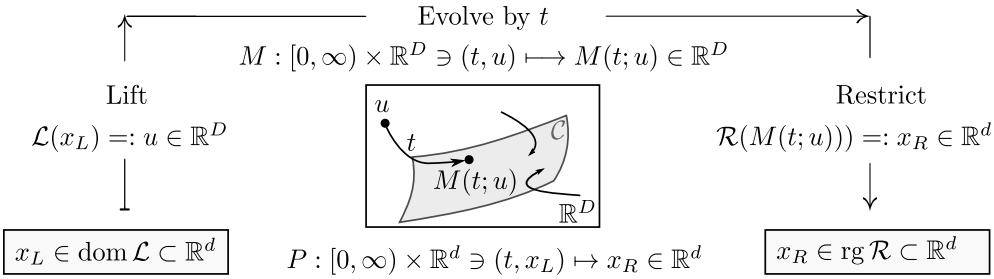


FIG. 1.1. *Lift-evolve-restrict map $P(t;\cdot)$ in low-dimensional space $\mathbb{R}^d$ used in equation-free computations.*

dimensional microscopic system (which is defined in $\mathbb{R}^D$) that contains a d-dimensional invariant manifold $\mathcal{C}$, which we will call the slow manifold. We do not append a subscript ε to $\mathcal{C}$, since our main result will only assume existence and smoothness of the invariant manifold $\mathcal{C}$ and its stable fiber projection, but not consider the limit of time scale separation $\varepsilon \to 0$. The framework, illustrated in Figure 1.1, only relies on the availability of a microscopic time stepper (a map $M(t;\cdot) : \mathbb{R}^D \mapsto \mathbb{R}^D$ for $t \geq 0$) that can be called at selected microscopic initial values $u \in \mathbb{R}^D$. The goal is to compose a macroscopic time stepper $\Phi_*(\delta;\cdot) : \mathbb{R}^d \mapsto \mathbb{R}^d$ for $\delta \in \mathbb{R}$ (possibly including $\delta < 0$) in some coordinates for the slow manifold $\mathcal{C}$, which is then amenable to higher-level tasks such as bifurcation analysis.

For equation-free computations the user also has to choose two operators, the *lifting* $\mathcal{L} : \mathbb{R}^d \mapsto \mathbb{R}^D$ and the *restriction* $\mathcal{R} : \mathbb{R}^D \mapsto \mathbb{R}^d$, which are maps between the original high-dimensional ($\mathbb{R}^D$) microscopic level and the low-dimensional ($\mathbb{R}^d$) macroscopic level. The user-defined lifting $\mathcal{L}$ and restriction $\mathcal{R}$, together with the time stepper $M(t;\cdot)$ define the central building block of equation-free methodology, the "lift-evolve-restrict" map,

$$P : [0,\infty) \times \operatorname{dom}\mathcal{L} \ni (t,x_L) \mapsto \mathcal{R}(M(t;\mathcal{L}(x_L))) \in \operatorname{rg}\mathcal{R}$$

(see Figure 1.1). For a given value $x_L \in \mathbb{R}^d$ of macroscopic quantities, one first applies the lifting $\mathcal{L}$ to x_L getting a microscopic state u, then one runs the microscopic simulation for time t starting from u (applying the microscopic evolution $M(t;u)$), and finally one applies the restriction $\mathcal{R}$ to the result $M(t;u)$.

The use of the lift-evolve-restrict map P assumes that the trajectory $t \mapsto M(t;u)$ of the time stepper will be close to the slow manifold $\mathcal{C}$ most of the time. Assuming this, equation-free methods aim to extract information about the slow flow along $\mathcal{C}$ by calling the lift-evolve-restrict map P judiciously. The simplest approach would be to use $P(t;\cdot)$ as an approximation for $M(t;\cdot)$ restricted to $\mathcal{C}$ (called *explicit* equation-free computation in [33]).

When using equation-free methods one faces several challenges, both analytical and in terms of implementation. First, as the slow manifold $\mathcal{C}$ cannot be assumed to be known to the user, the method cannot assume that the user provides a lifting operator $\mathcal{L}$ that maps onto $\mathcal{C}$. This leads to initial fast transients in the trajectory that will also change the supposedly slow variables, unless the stable fiber projection $g : \mathbb{R}^D \to \mathcal{C}$ keeps the restriction constant (the criterion would be $\mathcal{R} \circ g \circ \mathcal{L} \approx I$). Since the projection g cannot be assumed to be known either, this implies that an unknown nonlinear transformation is applied to the variables in $\operatorname{dom}\mathcal{L}$ before the slow

dynamics start. A detailed illustration of this problem is given in Figure 2.1 and its description in section 2.

Second, the justification for equation-free methods relies on the stronger results for classical attracting invariant manifolds of deterministic systems (including persistence of the slow manifold for finite time-scale separation and its shadowing properties via stable fiber projection). However, the methods are commonly applied to stochastic or deterministic chaotic systems with time scale separation, for which convergence results are weaker. In the stochastic case the microscopic time stepper $M(t;\cdot)$ applies to densities not single trajectories. Finally, for applications with stochastic microscopic systems the additional difficulty of low computational accuracy in the evaluation of $M(t;\cdot)$ and possibly $\mathcal{L}$ and $\mathcal{R}$ may impose practical limitations.

This paper addresses the first challenge, the unknown slow manifold and fiber projection. It proves convergence of the *implicit* approximation $y = \Phi_{t_{\text{skip}}}(\delta; x)$ for the slow flow $\Phi_*(\delta; x)$, given by the solution y of the d-dimensional nonlinear system

$$P(t_{\text{skip}}; y) = P(t_{\text{skip}} + \delta; x)$$

for sufficiently large *healing time* t_{skip} and a fixed finite time scale separation for the scenario of an attracting d-dimensional invariant manifold in $\mathbb{R}^D$ (strong reduction results are available in this scenario). We also give a demonstration how the implicit equation-free formulation behaves when it is applied to moments of distributions in a stochastic system. Starting from this demonstration, we outline in our subsequent discussion how convergence statements for stochastic systems may have to be formulated.

Applications and recent practical improvements. A motivation for using the equation-free framework is that it extends methods which are otherwise only applicable to low-dimensional dynamical systems directly to simulations of high-dimensional complex systems. Classical applications of equation-free methods were *macroscopic bifurcation analysis* for microscopic simulations in chemical engineering (see [24] for a review). Recently similar analysis was performed on stochastic network models of neurons [4, 30] or disease spread [18], or on agent-based models in ecology [45] and social sciences (for example, for consumer lock-in [3], for pedestrian flow [33, 32], or for trading [42]). Another example for a high-level task accessible via equation-free methods is control design [43, 42].

Recent modifications and improvements to equation-free methods in multi-particle or agent-based simulations are variance reduction [35, 3], restriction of computations to patches in space [40, 41, 28] (for which a-priori error estimates can be proven [40, 41]), and data-driven selection of the slow variables using diffusion maps [12, 33]. Debrabant *et al.* [13] construct an acceleration scheme for Monte-Carlo simulations of high-dimensional SDEs based on moments of densities (the macroscopic variables), and prove its convergence as the number of moments goes to infinity.

2. Current state of analysis.

Geometry of the idealized case of an attracting slow manifold. Analysis of the equation-free framework (based on lift-evolve-restrict) is still ongoing. Convergence analysis with general a-priori error estimates has been performed mostly for the idealized case where the D-dimensional microscopic problem has a d-dimensional attracting invariant slow manifold $\mathcal{C}$, which is rarely encountered in the practical applications listed above. Exceptions are, for example, a study of bursting neurons [8] and the application of implicit equation-free computations to generalize an algorithm for growing stable manifolds of fixed points of two-dimensional maps a delay-differential equation

with an unknown two-dimensional slow manifold [38]. Even for this idealized case one faces the geometric difficulty illustrated in Figure 2.1. The geometry shows an exam-

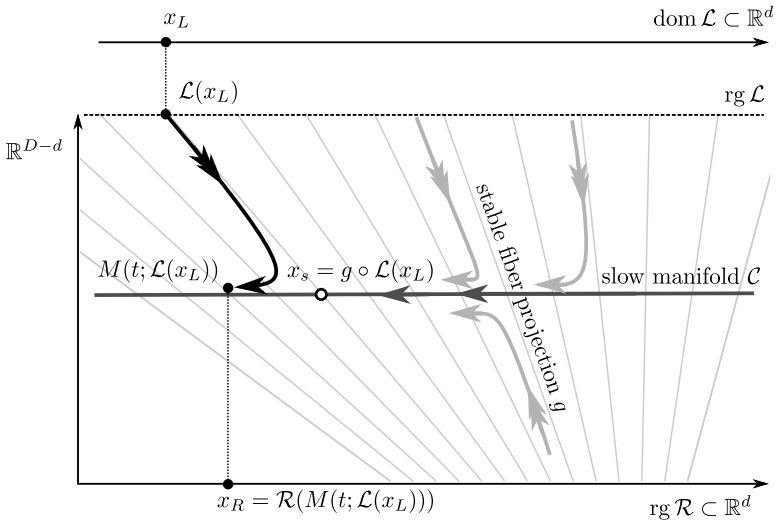


FIG. 2.1. *Geometry of lift-evolve-restrict map near slow manifold: macroscopic value $x_L \in \mathbb{R}^d$ gets lifted to $\mathcal{L}(x_L)$, then evolved to $M(t;\mathcal{L}(x_L))$, then restricted to $x_R = \mathcal{R}(M(t;\mathcal{L}(x_L)) \in \mathbb{R}^d$. The aim is to approximate the true slow flow $x_L \mapsto (g \circ \mathcal{L})^{-1} \circ M(t; g(\mathcal{L}(x_L)))$, which involves the unknown fiber projection g, using the map $\mathcal{R} \circ M(t;\cdot) \circ \mathcal{L}$, and assuming invertibility of $g \circ \mathcal{L} : \mathbb{R}^d \mapsto \mathcal{C}$. For this sketch, $d = 1$ and $D = 2$.*

ple scenario where the microscopic system is two-dimensional, and the slow manifold $\mathcal{C}$ is horizontal (and, thus, the slow motion is purely horizontal, drifting to the left). Here we choose a lifting $\mathcal{L}$ that maps also onto a horizontal line $\mathrm{rg}\,\mathcal{L}$. However, $\mathrm{rg}\,\mathcal{L}$ is at a distance to $\mathcal{C}$, because the precise location of $\mathcal{C}$ is in practice unknown. The restriction $\mathcal{R}$ is the horizontal component of any point $u \in \mathbb{R}^2$. The spaces $\mathrm{dom}\,\mathcal{L}$ and $\mathrm{rg}\,\mathcal{R}$ (both one-dimensional) are drawn separately for clarity in Figure 2.1, but they may be identical in examples. The fast motion of $M(t;\cdot)$ is not perfectly vertical, but has a significant horizontal component. Figure 2.1 also shows how the map P acts on a typical point x_L, showing its image $\mathcal{L}(x_L)$, the result of the evolution, $M(t;\mathcal{L}(x_L))$, and the result of the restriction $P(t;x_L) = x_R = \mathcal{R}(M(t;\mathcal{L}(x_L)))$.

The point $x_s \in \mathcal{C}$ in Figure 2.1 is defined as the unique point x_s on $\mathcal{C}$ such that $M(t;\mathcal{L}(x_L)) - M(t;x_s)$ converges at an exponential rate d_{tr} that is larger than the maximal rate of contraction d_{tan} tangential to $\mathcal{C}$ (which is horizontal). As mentioned in the introduction as shadowing, this mapping is defined for every point u in the neighborhood of $\mathcal{C}$: for every u near $\mathcal{C}$ there exists a point $g(u) \in \mathcal{C}$ such that $M(t;u) - M(t;g(u)) \sim \exp(-d_{\mathrm{tr}}t)$ (in the illustration $u = \mathcal{L}(x_L)$, $g(u) = x_s$). This point $g(u)$ is called the stable fiber projection of u. The map g is known to have the same regularity as $\mathcal{C}$ [15, 20]. For $d_{\mathrm{tan}} \ll d_{\mathrm{tr}}$ the map can be expanded in orders of $\varepsilon = d_{\mathrm{tan}}/d_{\mathrm{tr}}$. The thin grey lines (called stable fibers or isochrones) in Figure 2.1 indicate how points in the plane are projected onto $\mathcal{C}$ under the nonlinear projection g for the illustrative example.

Figure 2.1 makes clear that the dynamics of the map $P(t;x)$ is qualitatively different from the dynamics of $M(t;\cdot)$ restricted to $\mathcal{C}$, $M(t;\cdot)|_{\mathcal{C}}$. For the particular geometry shown in Figure 2.1 $P(t;\cdot)$ has a unique stable fixed point if the horizontal attraction/expansion rate d_{tan} of $M(t;\cdot)$ on $\mathcal{C}$ is sufficiently small compared to the

attraction rate d_{tr} transversal to $\mathcal{C}$. This fixed point is nearly independent of the dynamics of $M(t;\cdot)$ on $\mathcal{C}$.

More generally, if the lifting operator $\mathcal{L}$ does not map x_L into the low-dimensional slow manifold $\mathcal{C}$ then the initial part of the trajectory $t \mapsto M(t;\mathcal{L}(x_L))$, which is computed as part of the lift-evolve-restrict map P, is a rapidly changing transient toward the slow manifold $\mathcal{C}$, which will generically also change the resulting x_R.

In the limit of infinite time-scale separation (that is, the derivative of M with respect to time, $\partial_1 M(t;,u)$, goes to 0 for $u \in \mathcal{C}$) the dynamics of the lift-evolve-restrict map P is a small perturbation of the map $\mathcal{R}\circ g\circ\mathcal{L}$. Unless this limit map equals the identity, $P(t;\cdot)$ cannot be a good approximation of the slow flow along the manifold $\mathcal{C}$. Using x in the domain of the lifting $\mathcal{L}$ and the map $g\circ\mathcal{L} : \operatorname{dom}\mathcal{L} \mapsto \mathcal{C}$ onto the manifold $\mathcal{C}$ as the coordinate map, the slow flow Φ_* has the form

$$\begin{aligned} \Phi_*(\delta;\cdot) &= (g\circ\mathcal{L})^{-1}\circ M(\delta;\cdot)\circ(g\circ\mathcal{L}), && \text{or implicitly defined by} \\ \Phi_* &: \mathbb{R}\times\operatorname{dom}\mathcal{L} \ni (\delta,x) \mapsto y_* \in \operatorname{dom}\mathcal{L}\,, && \text{where } y_* \text{ solves} \\ \mathcal{R}(g(\mathcal{L}(y_*))) &= \mathcal{R}(M(\delta;g(\mathcal{L}(x)))) \end{aligned} \tag{2.1}$$

(using the notation $(\cdot)^{-1}$ for the inverse map). This definition is not directly computable since the nonlinear projection g is unknown in general.

Feasible approaches to construct an accurate approximation of $M(t;\cdot)$ restricted to $\mathcal{C}$ are *constrained runs*, as discussed by Gear, Zagaris *et al.* [16, 48, 49], or the introduction of a healing time t_{skip}. The latter approach is studied in this paper.

Constrained runs. The approach of [16, 48, 49] to ensuring that $\mathcal{R}\circ g\circ\mathcal{L}$ is close to the identity is to enforce that the lifting $\mathcal{L}$ maps onto the manifold $\mathcal{C}$ with sufficient accuracy for all x in its domain. Usually, this requires an additional scheme involving the iterative application of $\mathcal{L}$ and M; see [16, 48, 49]. The a-priori error estimates prove that the lift-evolve-restrict scheme with these additional iterations has an error of order $(d_{\tan}/d_{\mathrm{tr}})^m$ if the constrained runs scheme is of order m, where $d_{\tan}$ is the attraction/repulsion time scale tangential to the slow manifold $\mathcal{C}$ and d_{tr} is the transversal attraction rate. The ratio $d_{\tan}/d_{\mathrm{tr}}$ measures the time scale separation. It is assumed to be small when applying constrained runs (and called ε), and $O(\varepsilon^m)$ convergence is proven in [16, 48, 49] in the limit $\varepsilon\to 0$. This limit will not be required in our proof, later on.

Implicit formulation with healing time. A second, alternative, approach is to introduce a *healing time* t_{skip}, exploiting that M attracts along the fibers [25, 5]. Marschler *et al.* [31] show that the healing time t_{skip} can be motivated by introducing an additional shift $M(t_{\mathrm{skip}};\cdot)$ and its inverse into (2.1) (note that $M(t_{\mathrm{skip}};\cdot)$ is invertible on the slow manifold $\mathcal{C}$):

$$y_* = \Phi_*(\delta;x) = (g\circ\mathcal{L})^{-1}\circ M(t_{\mathrm{skip}};\cdot)^{-1}\circ M(\delta+t_{\mathrm{skip}};\cdot)\circ g\circ\mathcal{L}(x). \tag{2.2}$$

Removing the inverses in (2.2) leads to an implicit equation for $y_* = \Phi_*(\delta;x)$ with the healing time t_{skip} as an additional parameter:

$$\mathcal{R}\circ M(t_{\mathrm{skip}};\cdot)\circ g\circ\mathcal{L}(y_*) = \mathcal{R}\circ M(\delta+t_{\mathrm{skip}};\cdot)\circ g\circ\mathcal{L}(x) \tag{2.3}$$

In (2.3) the parameter t_{skip} has no effect since $M(t_{\mathrm{skip}};\cdot)$ is invertible on the slow manifold. However, the difference $M(t_{\mathrm{skip}};\cdot)\circ g - M(t_{\mathrm{skip}};\cdot)$ decreases with t_{skip} (at rate $\sim\exp(-d_{\mathrm{tr}}t_{\mathrm{skip}})$). In Figure 2.2 the distance between points along the trajectory starting from $\mathcal{L}(x)$ (in red) and their projections $g\circ\mathcal{L}(x)$ (white) illustrates this convergence. Thus, we may approximate $M(t_{\mathrm{skip}};\cdot)\circ g$ by $M(t_{\mathrm{skip}};\cdot)$ in (2.3). This

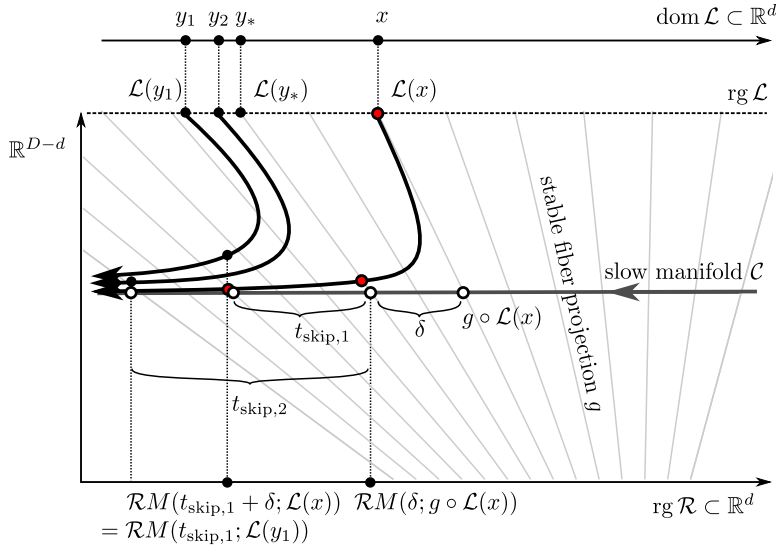


FIG. 2.2. *Trajectories involved in the implicit definition of y_* and y using the scenario of Figure 2.1: y_* is the image of x under the true slow flow, $y_* = \Phi_*(\delta; x)$; y_j are approximations for different t_{skip}, $y_j = \Phi_{t_{\text{skip}},j}(\delta; x)$.*

results in a computable approximation $y_{t_{\text{skip}}} = \Phi_{t_{\text{skip}}}(\delta; x)$ of y_*, given implicitly by the equation

$$\mathcal{R}(M(t_{\text{skip}}; \mathcal{L}(y_{t_{\text{skip}}}))) = \mathcal{R}(M(\delta + t_{\text{skip}}; \mathcal{L}(x))). \tag{2.4}$$

Figure 2.2 illustrates the effect of increasing healing time t_{skip} in the scenario introduced in Figure 2.1. The points y_1 and y_2 are the solutions of (2.4) for two different healing times $t_{\text{skip},1} < t_{\text{skip},2}$. Equation (2.4) means that the points y_j are defined as those elements of $\operatorname{dom}\mathcal{L}$ for which the trajectory starting from $\mathcal{L}(y_j)$ has the same horizontal component (restriction $\mathcal{R}$) as $M(t_{\text{skip},j} + \delta; \mathcal{L}(x))$ after time $t_{\text{skip},j}$.

The implicit approach was analyzed and illustrated in a traffic model in [31] and will also be studied in this paper. Vandekerckhove *et al.* [46] proposed and demonstrated a similar approach, but applied the healing time backward in time by fixing the image of the restriction: they solve $x = \mathcal{R}(M(t_{\text{skip}}; \mathcal{L}(x_b)))$ for x_b first and then set $y = \mathcal{R}(M(\delta + t_{\text{skip}}; \mathcal{L}(x_b)))$. This gives an (approximate) representation $\Phi_*^{\mathcal{R}}$ of the slow flow in the coordinates on the image of the restriction $\mathcal{R}$: $\Phi_*^{\mathcal{R}}(\delta; x) = \mathcal{R} \circ M(\delta; \cdot) \circ [\mathcal{R}\,|_{\mathcal{C}}]^{-1}(x)$.

The coordinates for the flow on the slow manifold $\mathcal{C}$ are somewhat arbitrary as the difference between the expressions used by Vandekerckhove *et al.* [46] and the implicit expression (2.4) for $\Phi_{t_{\text{skip}}}$ shows. For the coordinates in the space $\operatorname{dom}\mathcal{L}$ the diffeomorphism between $\operatorname{dom}\mathcal{L}$ and $\mathcal{C}$ is $g \circ \mathcal{L}$, where g is the stable fiber projection, as implied by (2.2). The diffeomorphism can be approximately computed by solving $\mathcal{R}(M(2t_{\text{skip}}; \mathcal{L}(x_g))) = \mathcal{R}(M(t_{\text{skip}}; \mathcal{L}(x)))$ for x_g and then using $M(t_{\text{skip}}; \mathcal{L}(x_g))$ as the approximation for $[g \circ \mathcal{L}](x)$. The approximate diffeomorphism for the expression of Vandekerckhove *et al.* [46] is $[\mathcal{R}\,|_{\mathcal{C}}]^{-1} : x \mapsto M(t_{\text{skip}}; \mathcal{L}(x_b))$.

Marschler *et al.* [31] proved that the approximation $y_{t_{\text{skip}}}$ is exponentially accurate if $d_{\tan}/d_{\text{tr}} \to 0$: $\|y_{t_{\text{skip}}} - y_*\| \sim \exp(-K d_{\text{tr}}/d_{\tan})$ (for some constant K depending on t_{skip}). The error estimates in [31] require that $t_{\text{skip}} d_{\tan}/d_{\text{tr}}$ and $(t_{\text{skip}} + \delta) d_{\tan}/d_{\text{tr}}$

stay bounded from above such that the convergence result is valid in the limit of infinite time scale separation $d_{\mathrm{tan}}/d_{\mathrm{tr}} \to 0$. This means that the assumptions of [31] are similar to those required by schemes involving constrained runs [16, 48, 49]. The analysis left open if the error goes to zero for $t_{\mathrm{skip}} \to \infty$ but the time scale separation stays finite: $d_{\mathrm{tan}}/d_{\mathrm{tr}} \in (0, 1)$.

Our paper will prove the general a-priori error estimate that $\|y_{t_{\mathrm{skip}}} - y_*\| \sim \exp((d_{\mathrm{tan}} - d_{\mathrm{tr}})t_{\mathrm{skip}})$ for $t_{\mathrm{skip}} \to \infty$ and fixed $d_{\mathrm{tan}} < d_{\mathrm{tr}}$ under some genericity conditions on $\mathcal{R}$ and $\mathcal{L}$. It will also give a convergence result for the derivatives of $y_{t_{\mathrm{skip}}}$ with respect to its argument x: $\|\partial^j y_{t_{\mathrm{skip}}} - \partial^j y_*\| \sim \exp(((2j+1)d_{\mathrm{tan}} - d_{\mathrm{tr}})t_{\mathrm{skip}})$ if $(2j+1)d_{\mathrm{tan}} < d_{\mathrm{tr}}$.

Analysis beyond attracting manifolds in slow-fast systems. As mentioned above, equation-free analysis based on lift-evolve-restrict maps is more commonly applied to problems that are assumed to have a fast subsystem, where the fast time scale converges only in a statistical sense to a stationary measure conditioned on the slow variables. In these cases the microscopic time stepper $M(\delta; \cdot)$ operates on measures (or densities). It may be approximated by Monte Carlo simulations on ensembles of initial conditions. Barkley *et al.* [5] investigated the behaviour of the lift-evolve-restrict map $P(\delta; \cdot) = \mathcal{R} \circ M(\delta; \cdot) \circ \mathcal{L}$ where the slow variables were leading moments (thus, P was called moment map in [5]) on prototype examples from the class of stochastic problems. The simplest example from [5] is a scalar stochastic differential equation (SDE), for which the evolution of the probability distribution is governed by a (linear) Fokker-Planck equation (FPE). Hence, the measure of time-scale separation is the size of the spectral gap in the right-hand side of the FPE. The analysis in [5] found that the dynamics of the map P was qualitatively different from the dynamics of the underlying linear FPE. For example, P was nonlinear and had several coexisting fixed points for certain choices of time δ.

Our paper will demonstrate for two different lifting operators $\mathcal{L}$ that the approximation $y_{t_{\mathrm{skip}}}$, defined by (2.4), behaves exactly as predicted by our convergence theorem. In particular, it preserves the metastability features and the linearity of the flow generated by the FPE, thus, addressing the problems highlighted in [5].

2.1. Outline of results. Section 3 states the precise assumptions (time scale separation for decay rates tangential and transversal to the invariant manifold $\mathcal{C}$ ($d_{\mathrm{tan}} < d_{\mathrm{tr}}$) and transversality of $\mathcal{R}$ and $\mathcal{L}$) for exponential convergence:

$$\partial^j y_{t_{\mathrm{skip}}} - \partial^j y_* \sim \exp(((2j+1)d_{\mathrm{tan}} - d_{\mathrm{tr}})t_{\mathrm{skip}}) \quad \text{for } t_{\mathrm{skip}} \to \infty \tag{2.5}$$

(using the convention that $\partial^0 y = y$ and assuming that the derivatives up to order $j+1$ exist). Estimate (2.5) predicts that convergence in t_{skip} is slower for derivatives of higher order. Section 4 demonstrates the convergence rates in t_{skip} for $y_{t_{\mathrm{skip}}}$ and its first two derivatives with respect to x for a singularly perturbed ODE modelling the Michaelis-Menten kinetics (which was also used by [16, 48, 49] for illustration). Section 5 studies the evolution of densities under a scalar SDE with a double-well potential drift term also considered by Barkley *et al.* [5]. We demonstrate global convergence of implicit equation-free methods for a linear lifting $\mathcal{L}_{\mathrm{lin}}$. We also demonstrate local convergence for the nonlinear lifting $\mathcal{L}_{\mathrm{Gauss}}$ used in [5].

Section 6 discusses differences between observations of the behaviour in the SDE and the predictions from the theoretical result. These are caused by the numerical errors in the evaluations of lifting, evolution and restriction and their growth along trajectories.

We conclude with an outlook on possible consequences of the results on application

of equation-free methods to Monte-Carlo simulations of multi-particle or agent-based systems. One important observation is that in some cases increasing the number of agents or particles does *not* increase the spectral gap (and, thus, the time scale separation). Didactic examples where the finiteness of the spectral gap is apparent are the dynamic networks as considered by Gross and Kevrekidis [18]. The slow system is an ODE derived from the pair-wise interaction approximation, cutting off an infinite series of ODEs of higher-order interaction terms. The spectral gap between pair-wise interaction terms and triplet interaction terms is finite even in the limit of infinitely large networks.

Thus, the results from Section 3 are potentially applicable to equation-free analysis of stochastic multi-particle systems, where distributions of microscopic initializations are studied. This is in contrast to previous convergence results on constrained runs [16, 48, 49] and implicit lifting [31], which only apply in the limit of infinite time scale separation.

3. Convergence in the case of finite time-scale separation. We consider a smooth dynamical system

$$\dot{u} = f(u), \quad u \in \mathbb{R}^D, \tag{3.1}$$

where D is large. We assume that the flow M generated by (3.1),

$$M : \mathbb{R} \times \mathbb{R}^D \to \mathbb{R}^D, \qquad (t; u) \mapsto M(t; u)$$

has a d-dimensional compact relatively invariant manifold $\mathcal{C}$ (possibly with boundary). That is, trajectories $M(t; u)$ starting in $u \in \mathcal{C}$ either stay in $\mathcal{C}$ for all times $t \in \mathbb{R}$, or they stay in $\mathcal{C}$ until they cross the boundary $\partial\mathcal{C}$ of $\mathcal{C}$. We assume that $\mathcal{C}$ is at least $k_{\max}$ times differentiable. For a point $u \in \mathcal{C}$, let us denote by $\mathcal{N}(u)$ the d-dimensional tangent space to $\mathcal{C}$. The following assumption states that attraction transversal to the manifold $\mathcal{C}$ is faster than attraction or expansion tangential to $\mathcal{C}$.

ASSUMPTION 3.1 (Hyperbolicity — Separation of time scales and transversal stability). *There exists an open neighborhood $\mathcal{U}$ of the manifold $\mathcal{C}$, a (possibly nonlinear) projection $g : \mathcal{U} \mapsto \mathcal{C}$ (the so-called* stable fiber projection*), a pair of constants (decay rates) $0 < d_{\tan} < d_{\mathrm{tr}}$, and a bound C such that the following two conditions hold.*

1. (tangential expansion/attraction rate) *For all points $u \in \mathcal{C}$ on the manifold, all tangent vectors $v_1, \ldots, v_{k_{\max}} \in \mathcal{N}(u)$ and all $t \in \mathbb{R}$ with $M([0, t]; u) \subset \mathcal{C}$*

   $$\|\partial_2^j M(t; u)[v_1, \ldots, v_{k_j}]\| \le C \exp(d_{\tan}|t|)\|v_1\| \cdot \ldots \cdot \|v_j\| \tag{3.2}$$

   for all $j \in \{1, \ldots, k_{\max}\}$.
2. (Stability along transversal fiber projections) *For all $u \in \mathcal{U}$ and all $t > 0$ with $M([0, t]; g(u)) \in \mathcal{C}$*

   $$\|\partial_2^j M(t; u) - \partial_2^j M(t; g(u))\| \le C \exp(-t d_{\mathrm{tr}})\|u - g(u)\| \tag{3.3}$$

   for all $j \in \{0, \ldots, k_{\max}\}$.

In (3.2) and (3.3) we use the convention that $\partial_k^j M$ is the jth-order partial derivative of M with respect to its kth argument, and that $\partial_2^0 M$ ($j = 0$) is the flow M itself. The norm on the left side of (3.3) is the usual operator norm for the multi-linear operators $\partial_2^j M(t, \cdot)$. The constants C, d_{tr} and $d_{\tan}$ are assumed to be independent of the point u and the time t. Assumption (3.2) is also made for negative times t (using

the convention that $M([0,t];u)$ means $M([t,0];u)$ for $t<0$) such that it is also an assumption about the inverse of M, when restricted to $\mathcal{C}$: $M(-t,\cdot)=M^{-1}(t,\cdot)$. The constant d_{tr} is the decay rate toward the manifold $\mathcal{C}$, the constant d_{tan} is the rate of attraction and expansion along the flow restricted to $\mathcal{C}$. The main requirement of Assumption 3.1 is that $d_{\mathrm{tr}} > d_{\mathrm{tan}}$.

Transversality of restriction and lifting. Second, we assume basic compatibility between

$$\begin{aligned} &\mathcal{R} : \mathcal{U} \subset \mathbb{R}^D \to \mathbb{R}^d && \text{the } \textit{restriction} \text{ operator,} \\ &\mathcal{L} : \operatorname{dom}\mathcal{L} \subset \mathbb{R}^d \to \mathbb{R}^D && \text{the } \textit{lifting} \text{ operator,} \end{aligned} \tag{3.4}$$

and the invariant manifold $\mathcal{C}$: the lifting $\mathcal{L}$ should map into the neighborhood $\mathcal{U}$ of $\mathcal{C}$ in which the stable fiber projection g is defined, and the restriction $\mathcal{R}$ should be defined on the projection g of the image of $\mathcal{L}$ along the stable fibers:

$$\mathcal{L}(\operatorname{dom}\mathcal{L}) \subset \mathcal{U}, \qquad\qquad g(\mathcal{L}(\operatorname{dom}\mathcal{L})) \subset \operatorname{dom}\mathcal{R} \cap \mathcal{C},$$

In addition to these compatibility conditions, we impose the following two transversality conditions on lifting $\mathcal{L}$ and restriction $\mathcal{R}$.

ASSUMPTION 3.2 (Transversality of $\mathcal{R}$ and $\mathcal{L}$).

1. *the projection g is a diffeomorphism between* $\operatorname{rg}\mathcal{L} = \mathcal{L}(\operatorname{dom}\mathcal{L})$ *and* $\mathcal{C}$. *In particular, for all* $x \in \operatorname{dom}\mathcal{L} \subset \mathbb{R}^d$

$$\operatorname{rank}\frac{\partial}{\partial x}\left[g(\mathcal{L}(x))\right] = \operatorname{rank}\left[\partial g(\mathcal{L}(x)) \circ \partial\mathcal{L}(x)\right] = d.$$

2. *The map $\mathcal{R}$, restricted to $\mathcal{C}$, is a diffeomorphism between $\mathcal{C}$ and $\mathbb{R}^d$. In particular, for all $u \in \mathcal{C}$ ($\mathcal{N}(u)$ is the tangent space to $\mathcal{C}$ in u)*

$$\dim \partial\mathcal{R}(u)\mathcal{N}(u) = d.$$

Coordinates on the slow manifold $\mathcal{C}$. The maps $\mathcal{R}$ and $\mathcal{L}$ create two natural ways to define local coordinate representations on the invariant manifold $\mathcal{C}$, one by a map from $\operatorname{dom}\mathcal{L}$ to $\mathcal{C}$, one by a map from $\mathcal{C}$ to $\operatorname{rg}\mathcal{R}$. For our presentation we choose the representation in $\operatorname{dom}\mathcal{L}$ coordinates:

$$g\circ\mathcal{L} : \operatorname{dom}\mathcal{L} \subset \mathbb{R}^d \mapsto \mathcal{C} \subset \mathbb{R}^D, \qquad\qquad x \mapsto g(\mathcal{L}(x)).$$

The inverse of $g\circ\mathcal{L}$ is defined implicitly. Assume that $u_0 = g(\mathcal{L}(x_0))$ for some $x_0 \in \operatorname{dom}\mathcal{L}$. Then for $u \in \mathcal{C}$ near u_0 the pre-image $x = (g\circ\mathcal{L})^{-1}(u)$ is found by solving $\mathcal{R}(u) = \mathcal{R}(g(\mathcal{L}(x)))$ for $x \approx x_0$, which has a locally unique solution by Assumption 3.2.

We can represent the flow M on $\mathcal{C}$ as a flow in $\operatorname{dom}\mathcal{L}$, denoting it by Φ_*:

$$\Phi_* : \mathbb{R}\times\operatorname{dom}\mathcal{L} \mapsto \operatorname{dom}\mathcal{L}, \quad \Phi_*(\delta;x) = [(g\circ\mathcal{L})^{-1}\circ M(\delta;\cdot)\circ g\circ\mathcal{L}](x) := y \tag{3.5}$$

(for $\delta\in\mathbb{R}$ and $x\in\operatorname{dom}\mathcal{L}$), where y is given implicitly as solution of a d-dimensional system of nonlinear equations

$$\mathcal{R}(g(\mathcal{L}(y))) = \mathcal{R}(M(\delta; g(\mathcal{L}(x)))). \tag{3.6}$$

Assumption 3.2 on transversality implies that Φ_* is well defined for small δ (since $y=x$ is a regular solution of (3.6) at $\delta=0$). For larger δ, one can break down

the solution into smaller steps by increasing δ gradually from 0 and tracking the curve $y(\delta)$ of solutions of (3.6), which is well parametrized by δ in every point by Assumption 3.2. If $\operatorname{dom}\mathcal{L}$ is simply connected then this continuation approach makes the implicit solution y used in the definition of Φ_* unique. Let us define the map

$$P_* : \mathbb{R} \times \operatorname{dom}\mathcal{L} \ni (t,x) \mapsto \mathcal{R}(M(t; g(\mathcal{L}(x)))) \in \mathbb{R}^d. \tag{3.7}$$

This map P_* is well defined and invertible for all $t \in \mathbb{R}$ and $x \in \operatorname{dom}\mathcal{L}$ for which the trajectory $s \mapsto M(s; g(\mathcal{L}(x)))$ stays in $\mathcal{C}$ for all s between 0 and t. The implicit definition (3.5) of the flow Φ_* on $\mathcal{C}$ has the following form when expressed with the help of this map P_* on $\operatorname{dom}\mathcal{L}$:

$$y = \Phi_*(\delta; x) \quad \text{if} \quad P_*(0; y) = P_*(\delta; x). \tag{3.8}$$

Since the flow $M(\delta;\cdot)$ is a diffeomorphism on $\mathcal{C}$, we can replace the times 0 and δ in the above implicit definition with t_{skip} and $t_{\text{skip}} + \delta$ for an arbitrary so-called *healing time* $t_{\text{skip}} \in \mathbb{R}$ (as long as $M([0, t_{\text{skip}}]; g(\mathcal{L}(x))) \subset \mathcal{C}$). So, equivalent to (3.8), we have for $t_{\text{skip}} > 0$ with $M([0, t_{\text{skip}}]; g(\mathcal{L}(x))) \subset \mathcal{C}$

$$y = \Phi_*(\delta; x) \quad \text{if} \quad P_*(t_{\text{skip}}; y) = P_*(t_{\text{skip}} + \delta; x). \tag{3.9}$$

Convergence Theorem for implicit equation-free computations with finite time-scale separation. The stable fiber projection g (which is part of the definition of P_*) is not known in most practical applications. Thus, implicit equation-free computations use the explicit macroscopic time-t map P instead of P_* in the equation defining y in (3.9):

$$P : [0,\infty) \times \operatorname{dom}\mathcal{L} \ni (t,x) \mapsto \mathcal{R}(M(t; \mathcal{L}(x))) \in \operatorname{rg}\mathcal{R} \tag{3.10}$$

such that we may define the approximate flow map

$$\Phi_{t_{\text{skip}}} : \mathbb{R} \times \operatorname{dom}\mathcal{L} \ni (\delta, x) \mapsto y \in \operatorname{dom}\mathcal{L}\,, \text{ where } y \text{ solves } P(t_{\text{skip}}; y) = P(t_{\text{skip}} + \delta; x) \tag{3.11}$$

implicitly in a similar way to (3.9). Our general convergence theorem, the following Theorem 3.3, states that $\Phi_{t_{\text{skip}}}$ is well defined for large t_{skip} (that is, the equation in (3.11), defining $\Phi_{t_{\text{skip}}}$ implicitly, has a locally unique solution), and that $\partial^j \Phi_{t_{\text{skip}}}$ is an approximation of $\partial^j \Phi_*$ of order $\exp(((2j+1)d_{\tan} - d_{\text{tr}})t_{\text{skip}})$ (including $j = 0$ for the map $\Phi_{t_{\text{skip}}}$).

THEOREM 3.3 (Convergence of approximate flow map at finite time-scale separation).
Let us assume that the microscopic flow M satisfies Assumption 3.1 on time-scale separation, and that the maps $\mathcal{R}$ and $\mathcal{L}$ satisfy Assumption 3.2 on transversality.

Let $\delta_{\max} > 0$ and $x \in \operatorname{dom}\mathcal{L}$ be arbitrary. Let us also assume that $x \in \operatorname{dom}\mathcal{L}$ maps to a point under $g \circ \mathcal{L}$ that keeps a positive distance from the boundary $\partial\mathcal{C}$ of $\mathcal{C}$ for all times $t \geq -\delta_{\max}$ under M. That is,

$$\operatorname{dist}(M(t; g(\mathcal{L}(x))), \partial\mathcal{C}) \geq c_\partial \quad \text{for all } t \geq -\delta_{\max} \text{ and some given } c_\partial > 0. \tag{3.12}$$

Then there exists a $t_0 \geq \delta_{\max}$ such that $y = \Phi_{t_{\text{skip}}}(\delta; x)$ is well defined by (3.11) for all $\delta \in [-\delta_{\max}, \delta_{\max}]$ and $t_{\text{skip}} > t_0$. The estimate

$$\|\partial_2^j \Phi_{t_{\text{skip}}}(\delta; x) - \partial_2^j \Phi_*(\delta; x)\| \leq C \exp(((2j+1)d_{\tan} - d_{\text{tr}})t_{\text{skip}}) \tag{3.13}$$

holds for all orders $j \in \{0, \dots, k_{\max} - 1\}$ satisfying $(2j+1)d_{\tan} < d_{\text{tr}}$. The constant C depends on $\delta_{\max}$ and x, but not on t_{skip}.

Assumption (3.12) in Theorem 3.3 is made to permit arbitrarily large t_{skip} while still having Assumption 3.1 and Assumption 3.2 uniformly satisfied. If one considers $x \in \operatorname{dom}\mathcal{L}$ for which the trajectory $t \mapsto M(t; g(\mathcal{L}(x)))$ leaves $\mathcal{C}$ (by crossing the boundary $\partial\mathcal{C}$) then one has to put restrictions on δ and t_{skip} to avoid crossing $\partial\mathcal{C}$. The theorem permits negative integration times δ shorter than $t_0 \le t_{\text{skip}}$ and positive integration times larger than t_{skip} as long as the factor $\exp(d_{\text{tan}}|\delta|)$ is of order 1. Since $1/d_{\text{tan}}$ and $1/d_{\text{tr}}$ are the time scales of the dynamics inside the invariant manifold $\mathcal{C}$ and transversal to it, the theorem covers time steps of length δ of order 1 in the slow $(1/d_{\text{tan}})$ time scale. The statement of Theorem 3.3 does not require that the time-scale separation $d_{\text{tan}}/d_{\text{tr}}$ goes to zero for convergence of the approximate map. It only requires that $d_{\text{tan}} < d_{\text{tr}}$, where d_{tr} is the attraction rate along fibers (see (3.3)) and d_{tan} is the attraction and expansion rate tangential to the invariant manifold $\mathcal{C}$. Since the constant C in (3.13) is independent of t_{skip} it can be chosen uniformly for compact domains $\operatorname{dom}\mathcal{L}$.

Outline of proof of Theorem 3.3. Existence and error of $\Phi_{t_{\text{skip}}}$: For the proof of Theorem 3.3 we have to analyze the difference between the two defining equations for the approximate solution y and the true solution $y_* = \Phi_*(\delta; x)$, both depending on x as a parameter:

$$\mathcal{R}(M(t_{\text{skip}}; \mathcal{L}(y))) = \mathcal{R}(M(t_{\text{skip}} + \delta; \mathcal{L}(x))), \tag{3.14}$$

$$\mathcal{R}(M(t_{\text{skip}}; g(\mathcal{L}(y_*)))) = \mathcal{R}(M(t_{\text{skip}} + \delta; g(\mathcal{L}(x)))). \tag{3.15}$$

Rearranging the difference between (3.14) and (3.15), we obtain an implicit fixed-point problem for y (recall that $P_*(t;\cdot) = \mathcal{R}\circ M(t;\cdot)\circ g\circ\mathcal{L}$ and $P(t;\cdot) = \mathcal{R}\circ M(t;\cdot)\circ\mathcal{L}$):

$$P_*(t_{\text{skip}}; y) = P_*(t_{\text{skip}}; y_*) + [P_*(t_{\text{skip}}; y) - P(t_{\text{skip}}; y)] + [P(t_{\text{skip}} + \delta; x) - P_*(t_{\text{skip}} + \delta; x)]. \tag{3.16}$$

The norms of both terms in square brackets, $P_*(t_{\text{skip}}; y) - P(t_{\text{skip}}; y)$ and $P(t_{\text{skip}} + \delta; x) - P_*(t_{\text{skip}} + \delta; x)$, are of order $\exp(-d_{\text{tr}} t_{\text{skip}})$ by Assumption 3.1, equation (3.3) (transversal stability with rate d_{tr} of $\mathcal{C}$). For the same reason, the Lipschitz constant of $P_*(t_{\text{skip}}; y) - P(t_{\text{skip}}; y)$ is of order $\exp(-d_{\text{tr}} t_{\text{skip}})$, too. By Assumption 3.1, equation (3.2) (tangential decay rate inside the manifold is less than d_{tan}), and Assumption 3.2 on transversality of $\mathcal{L}$ and $\mathcal{R}$, the inverse of $P_*(t_{\text{skip}};\cdot)$ has a local Lipschitz constant of order $\exp(d_{\text{tan}} t_{\text{skip}})$ near y_*. These two facts enable us to apply the Banach Contraction Mapping Principle to (3.16) to obtain a unique solution $y \approx y_*$ for large t_{skip}. More precisely, $y - y_*$ is of order $\exp((d_{\text{tan}} - d_{\text{tr}})t_{\text{skip}})$.

Inductive proof of error estimate for derivatives: We differentiate (3.16) with respect to x in its fixed point $y(t_{\text{skip}}; x)$ up to j times and then re-arrange the resulting equation for the jth-order derivatives of y and y_* into the form

$$\partial P_*(y)\left[\partial^j y - \partial^j y_*\right] = \left[\partial P_*(y) - \partial P_*(y_*)\right]\partial^j y_* + r. \tag{3.17}$$

In (3.17) we abbreviated $\partial_2 P_*(t_{\text{skip}};\cdot) = \partial P_*(\cdot)$, $\partial^j y = \partial_2^j y(t_{\text{skip}}; x)$ and dropped the argument x from y_* and the arguments t_{skip} and x from y. The remainder r is less than $C\exp(((2j-1)d_{\text{tan}} - d_{\text{tr}})t_{\text{skip}})$ for some constant C by induction hypothesis. The implicit expression (3.17) for $\partial^j y - \partial^j y_*$ shows why errors in derivatives of the solution can grow for increasing t_{skip} and insufficient time scale separation: the norms of $\partial P_*(y) - \partial P_*(y_*)$ and of $[\partial P_*(y)]^{-1}$ are of order $\exp(d_{\text{tan}} t_{\text{skip}})$ due to (3.2). The details of the proof are given in Appendix A.

4. Example: Michaelis-Menten kinetics. To illustrate the consequences of error estimate (3.13), we look at a model for Michaelis-Menten kinetics with explicit

time scale separation as studied in [34, 16, 48, 49]. The system is given in $\mathbb{R}^D$ with $D = 2$ as

$$\dot{x} = \varepsilon\left[-x + (x + \kappa - \lambda)y\right], \quad \dot{y} = x - (x + \kappa)y, \tag{4.1}$$

where $x \in \mathbb{R}$ is the slow variable, $y \in \mathbb{R}$ is the fast variable, and ε measures the time scale separation. The parameters $\kappa = 1, \lambda = 0.5$ and $\varepsilon = 0.01$ are kept fixed throughout this section.

In the singular case $\varepsilon = 0$, system (4.1) has a critical manifold $\mathcal{C}_0$ of equilibria, given by the graph $\mathcal{C}_0 = \{(x, y) : y = x/(x + \kappa)\}$. For positive ε, the system has a transversally stable invariant manifold, which can be represented as a graph $\mathcal{C}_\varepsilon = \{(x, y) : y = h_\varepsilon(x)\}$, such that $d = 1$. In this section we put a subscript ε on quantities to indicate their dependence on the parameter ε (so, writing, for example, $\mathcal{C}_\varepsilon$ instead of $\mathcal{C}$). The graph h_ε can be expanded in ε for small $\varepsilon > 0$:

$$y = h_\varepsilon(x) = \frac{x}{x + \kappa} + \frac{\kappa\lambda x}{(x + \kappa)^4}\varepsilon + \frac{\kappa\lambda x(2\kappa\lambda - 3\lambda x - \kappa x - \kappa^2)}{(x + \kappa)^7}\varepsilon^2 + \mathcal{O}(\varepsilon^3). \tag{4.2}$$

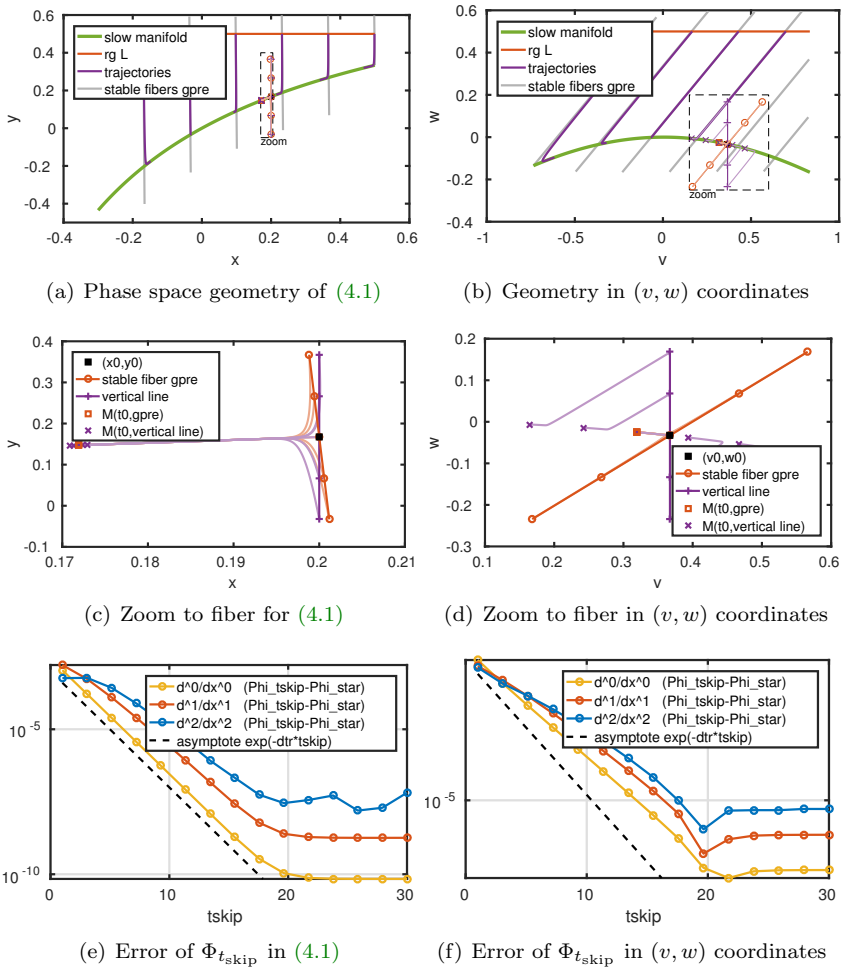
We plan to compare an equation-free approximate flow $\Phi_{t_{\text{skip}}}(\delta; \cdot)$, which is constructed below, to the true flow $\Phi_*(\delta; \cdot)$. For this simple example we may approximate the true flow $\Phi_*(\delta; \cdot)$ by obtaining an approximation of the stable fiber projection g_ε. For $\varepsilon \to 0$ in (4.1), g_ε has the limit $g_\varepsilon(x, y) \to (x, h_0(x))$. Thus, every point in phase space is approximately projected along vertical lines, as shown in Figure 4.1(a). For positive ε this stable fiber projection persists and is perturbed by terms of order ε. A general approximation algorithm for stable fibers in slow-fast systems was provided by Kristiansen *et al.* [26]. However, we need the stable fibers only to a degree of accuracy that permits us to compare $\Phi_*(\delta; \cdot)$ to $\Phi_{t_{\text{skip}}}(\delta; \cdot)$ for demonstration purposes. Thus, we expand $g_\varepsilon = (g_{x,\varepsilon}(x, y), g_{y,\varepsilon}(x, y))$ in ε. Since g_ε projects onto the manifold $\mathcal{C}_\varepsilon$, we know that $g_{y,\varepsilon}(x, y) = h_\varepsilon(g_{x,\varepsilon}(x, y))$. The first-order expansion of $g_{x,\varepsilon}$ is

$$g_{x,\varepsilon}(x, y) = x + \frac{(x + \kappa - \lambda)(y - 1)x + \kappa y}{x + \kappa}\varepsilon + O(\varepsilon^2).$$

Figure 4.1 was produced using a third-order expansion of h_ε and $g_{x,\varepsilon}$. The supplementary material provides Matlab code which reproduces the graphs in Figure 4.1 and computes the expansion coefficients for h_ε and $g_{x,\varepsilon}$ to third order (see also Appendix B). Figures 4.1(a) and 4.1(c) show the phase space geometry. The slow manifold $\mathcal{C}_\varepsilon = \{(x, y) : x = h_\varepsilon(x)\}$ is shown in green, the stable fibers (pre-images of selected points $(x_j, h_\varepsilon(x_j))$ on the slow manifold under the projection g_ε) are the almost straight grey lines, and some sample trajectories of (4.1) are shown in purple. After a rapid transient all trajectories approach the slow manifold $\mathcal{C}_\varepsilon$, given approximately by (4.2). Furthermore, Figure 4.1(c) shows in more detail how initial conditions on the same stable fiber, defined as the pre-image of x_0 under $g_{x,\varepsilon}$, $G_{\text{pre}} = \{(x, y) : g_{x,\varepsilon}(x, y) = x_0\}$, collapse onto the same slow limiting trajectory (trajectories shown in red in Figure 4.1(c)). In contrast, initial conditions with the same y-component do so only up to an error of order ε (trajectories shown in purple in Figure 4.1(c)).

To define the approximate flow $\Phi_{t_{\text{skip}}}$, we specify the restriction and lifting operators, $\mathcal{R}$ and $\mathcal{L}$, for the Michaelis-Menten system as

$$\mathcal{R} : \mathbb{R}^2 \mapsto \mathbb{R}, \qquad \mathcal{R}\begin{pmatrix} x \\ y \end{pmatrix} = x, \quad \text{and} \qquad \mathcal{L} : \mathbb{R} \mapsto \mathbb{R}^2, \qquad \mathcal{L}(x) = \begin{pmatrix} x \\ 0.5 \end{pmatrix}. \tag{4.3}$$



(a) Phase space geometry of (4.1)

(b) Geometry in (v, w) coordinates

(c) Zoom to fiber for (4.1)

(d) Zoom to fiber in (v, w) coordinates

(e) Error of $\Phi_{t_{\mathrm{skip}}}$ in (4.1)

(f) Error of $\Phi_{t_{\mathrm{skip}}}$ in (v, w) coordinates

FIG. 4.1. *Michaelis-Menten dynamics: panels* (a, c, e) *for* (4.1)*, panels* (b, d, f) *for rotated coordinates* (4.7). (a, b): *Location of slow manifold* $\mathcal{C}_\varepsilon = \{(x, y) : y = h_\varepsilon(x)\}$ *and stable fibers in phase space.* (c, d): *Trajectories* $M([0, \mathtt{t0}]; U_{\mathrm{ini}})$ $(\mathtt{t0} = 20)$ *starting from two sets* U_{ini} *of initial conditions: once from a subset of a vertical line, and once from a pre-image of a point under the stable fiber projection (called* gpre *in the legend).* (e, f): *Difference of* $\Phi_{t_{\mathrm{skip}}}(\delta; \cdot)$ *to third-order asymptotic expansion of* $\Phi_*(\delta; \cdot)$ *as a function of* t_{skip}*. Parameters:* $\lambda = 0.5, \kappa = 1, \varepsilon = 0.01$ *in* (4.1)*,* $x = -0.1 \in \mathrm{dom}\,\mathcal{L}$ *and* $\delta = 25$ *for* (e,f).

The approximate time-δ map $\Phi_{t_{\mathrm{skip}}}(\delta; \cdot)$ on the slow manifold is determined by the root $z_{t_{\mathrm{skip}}}$ of

$$F : \mathbb{R} \ni z \mapsto \mathcal{R}(M(t_{\mathrm{skip}}; \mathcal{L}(z))) - \mathcal{R}(M(t_{\mathrm{skip}} + \delta; \mathcal{L}(x))) \in \mathbb{R} \tag{4.4}$$

and setting $\Phi_{t_{\mathrm{skip}}}(\delta; x) := z_{t_{\mathrm{skip}}}$ (cf. (3.11)). We compare this to the true solution (or, rather, the alternative approximation by expansion) $\Phi_*(\delta; x)$, determined by the root z_* of

$$F_* : \mathbb{R} \ni z \mapsto \mathcal{R}(g_\varepsilon(\mathcal{L}(z))) - \mathcal{R}(M(\delta; g_\varepsilon(\mathcal{L}(x)))) \in \mathbb{R}, \tag{4.5}$$

setting $\Phi_*(\delta; x) := z_*$. Note, that F depends on t_{skip} and δ, and F_* depends on δ, which are not included in the list of arguments to simplify notation. We solve (4.4) and

(4.5) using a Newton iteration with tolerance 10^{-12}, where we approximate M with the fifth-order component of the `DOPRI45` Runge-Kutta scheme with fixed step size 0.1 for the ODE (4.1). We approximate the first two derivatives of Φ_* and $\Phi_{t_{\text{skip}}}$ (and the Jacobians needed inside the Newton iteration) by central finite differences with step size $\Delta_z = 10^{-4}$. The supplementary material contains didactic implementations of Φ_* and $\Phi_{t_{\text{skip}}}$ for this example in the form of matlab code and its published output. The error

$$E^j(t_{\text{skip}}) = |\partial_2^j \Phi_*(\delta; x_0) - \partial_2^j \Phi_{t_{\text{skip}}}(\delta; x_0)| \quad (\partial_2^0 \Phi \text{ refers to } \Phi) \tag{4.6}$$

for $x_0 = -0.1$ and $\delta = 25$ is shown in Figure 4.1(e) for a range of healing times $t_{\text{skip}} \in [0; 30]$. Since $\varepsilon = 10^{-2}$ and $d_{\tan} \sim \varepsilon$, the quantity $\exp(d_{\tan}\delta)$ is of order 1, as required by Theorem 3.3.

The error plot Figure 4.1(e) shows that $\Phi_{t_{\text{skip}}}$, $\partial_2 \Phi_{t_{\text{skip}}}$ and $\partial_2^2 \Phi_{t_{\text{skip}}}$ approach a limit at an exponential rate in t_{skip} up to an accuracy determined by the accuracy of the asymptotic expansion of Φ_* ($\sim \varepsilon^4$), round-off errors in the finite difference approximations of the derivatives ($\sim 10^{-4}$ for $\partial_2^2 \Phi_{t_{\text{skip}}}$), and the tolerance of the Newton iteration. Furthermore, the convergence rate is indeed lower for higher orders of the derivative of $\Phi_{t_{\text{skip}}}$ as the estimate (3.13) in Theorem 3.3 suggests. The slope of $\exp(-d_{\text{tr}} t_{\text{skip}})$ is included as a lower bound for the error for comparison.

We also observe that the error of the flow and its derivatives is acceptably small ($\approx 10^{-3}$) even for the minimal value $t_{\text{skip}} = 0$. Since $\mathcal{R} \circ \mathcal{L}$ equals the identity (see (4.3)), the implicit equation-free method turns into an explicit formulation if $t_{\text{skip}} = 0$. However, the geometry of system (4.1) is not generic. The system (4.1) is given in an explicit slow-fast form with one fast and one slow variable. This leads with our choice of lifting and restriction to the degenerate situation that the stable fiber projection g_ε is aligned with lifting and restriction to first order: $\mathcal{R} \circ g_0 \circ \mathcal{L} = I$ such that $\mathcal{R} \circ g_\varepsilon \circ \mathcal{L}$ is a small (order ε) perturbation of the identity. In this case the explicit equation-free method without healing time ($y = \mathcal{R}(M(0; \mathcal{L}(y))) = \mathcal{R}(M(\delta; \mathcal{L}(x)))$, such that $t_{\text{skip}} = 0$) is accurate up to order ε. To create a situation with a generic arrangement of the stable fiber projection g_ε, we study a rotated system of the Michaelis-Menten dynamics (which was also used by [16]).

We apply the rotation matrix R to the system in order to obtain the dynamics in the new coordinates $(v, w)^T \in \mathbb{R}^2$ by

$$\begin{pmatrix} v \\ w \end{pmatrix} = R \begin{pmatrix} x \\ y \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \text{ such that } \quad M_{\text{rot}}\left(t; (v, w)^T\right) = RM\left(t; R^{-1}(x, y)^T\right) \tag{4.7}$$

is the microscopic simulator in the new $(v, w)^T$ coordinates. In this rotated system the time scale separation is no longer visible between v and w, since the slow and fast variables are mixed. Figure 4.1(b) shows the phase space geometry with slow manifold (green), stable fibers (grey) and sample trajectories. The initial transients are no longer following a straight line parallel to a coordinate axis such that both, v and w, change rapidly during transients. This situation is expected in a generic situation when one applies equation-free methods without precise knowledge about the slow and fast variables. We use the same restriction and lifting operators as defined in (4.3) (but in the new coordinates (v, w): $\mathcal{L}(x) = (x, 0.5)^T$, $\mathcal{R}(v, w) = v$). All parameter values are as in the unrotated system (4.1), otherwise. The error E^j, defined in (4.6), is now much larger: it is of order 1 for $t_{\text{skip}} = 0$; see Figure 4.1(f). In the implicit framework the error decreases with increasing healing time t_{skip} down to

10^{-8} for $t_{\text{skip}} = 20$. Note again that the slope of the curves is smaller for higher-order derivatives as predicted by the estimate (3.13) in Theorem 3.3. The error for the flow is bounded from below again by the accuracy of the asymptotic expansion for Φ_*, the accuracy of the Newton iteration and round-off error caused by the finite difference approximations of $\partial_2^j \Phi_*$ and $\partial_2^j \Phi_{t_{\text{skip}}}$.

5. Application: stochastic dynamics. A common area where equation-free methods are applied are multi-particle systems where slow dynamics emerges for macroscopic (typically averaged) quantities, e. g. [29, 5]. More precisely, the macroscopic quantities are assumed to satisfy a low-dimensional stochastic differential equation (SDE). For example, the SDE could be assumed to be of the form $\mathrm{d}x = f(x)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$, where the noise term $\sigma\,\mathrm{d}W_t$ approximates the microscopic fluctuation as white noise and the deterministic part $f(x)$ is the systematic average drift of the macroscopic quantities. Givon *et al.* [17] review rigorous results concerning dimension reduction of SDEs.

Typically, a stochastic simulation is performed not just once, but for an ensemble of initial conditions and realizations (as part of a Monte Carlo simulation). At the level of an SDE, an ensemble of initial conditions corresponds to (a sampling of) an initial distribution density $\rho(x)$. In this section we restrict ourselves to the study of a scalar SDE of the form

$$\mathrm{d}Q = -V'(Q)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \tag{5.1}$$

where W_t is a Wiener process, an example for which explicit equation-free methods have been thoroughly analyzed by Barkley *et al.* [5]. As in [5], we set the noise strength σ equal to 1 in (5.1) without loss of generality. The potential

$$V(Q) = \frac{Q^4}{4} - \frac{\mu Q^2}{2} + \nu Q, \tag{5.2}$$

forms for $\mu > 0$ a double well with two local minima $Q_\pm$ and a local maximum Q_s (see lower panel of Figure 5.1 for a graph of V). The parameters μ and ν determine the depth and the asymmetry of the double-well potential, respectively. We use $\mu = 6$, $\nu = 0.3$ such that $Q_- < Q_s < Q_+$ and the well around Q_- is deeper than the well around Q_+. The microscopic simulation is a Monte Carlo simulation of (5.1) starting from initial (ensemble) density $\rho_0(Q)$ of initial conditions. Thus, the phase space is the space of possible initial distributions in Q, which has dimension D equal to infinity. Strictly, the infinite-dimensional case is outside of the scope of Theorem 3.3. However, the observations to follow agree with the convergence predicted by the theorem for reasons that will be discussed after defining lifting, evolution and restriction. We will make the connection to multi-particle systems or high-dimensional SDEs in section 6.

5.1. Lifting, evolution and restriction for distributions. The evolution of the probability density function (pdf) $\rho(Q, t)$ for the realization of (5.1) is determined by the Fokker-Planck equation with $\sigma = 1$,

$$\partial_t \rho = -\partial_Q(V'(Q)\rho) + \frac{1}{2}\partial_{QQ}\rho. \tag{5.3}$$

The right-hand side of (5.3) is linear, of the form

$$L\rho = -\partial_Q(V'\rho) + \frac{1}{2}\partial_Q^2\rho, \tag{5.4}$$

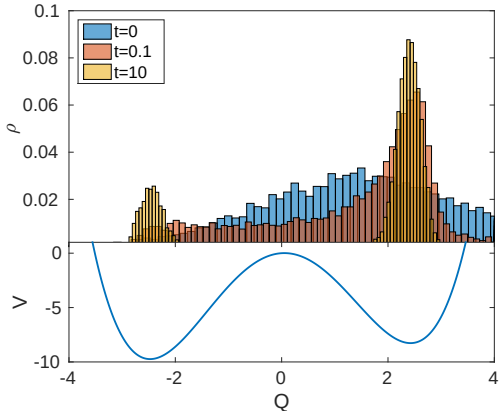


FIG. 5.1. (*Top*) *Dynamics of distributions for SDE model* (5.1) *for* $\mu = 6, \nu = 0.3, \sigma = 1$ (*sampled from* $N = 10000$ *realizations*). *On the microscopic level of distributions, the Gaussian distributed initial condition* ($t = 0$) *with mean* 1.5 *and variance* 3.5 *converges to a bimodal distribution by* $t = 10$. *Afterwards, the transition to the stationary distribution happens* (*see mode* 1 *in Figure* 5.2) *on a slow time scale corresponding to the second eigenvalue* $\lambda_2 \sim 10^{-8}$ *of the operator* L *given in* (5.4). (*Bottom*) *Shape of potential well* $V(Q)$, *as given in* (5.2).

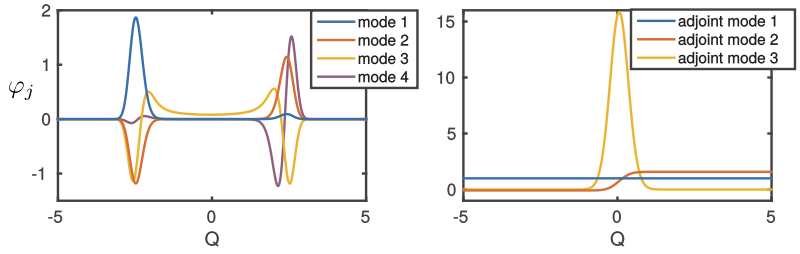


FIG. 5.2. *First four eigenfunctions (left panel) and first three* $\mathbb{L}^2$*-adjoint eigenfunctions (right panel) of the differential operator* L, *defined in* (5.4) *and computed with* `chebfun` [10, 14]. *The stationary solution is shown in blue for* $\lambda_1 = 0$ *as mode* 1 *in the left panel. The asymmetric eigenfunction corresponding to* λ_2 (*mode* 2) *is responsible for the transportation of mass from one well to the other. Eigenvalues* λ_k *are given in* (5.14): $(0, -8.63 \cdot 10^{-8}, -5.71, -10.3)$. *Parameters*: $\mu = 6$, $\nu = 0.3$, $\sigma = 1$. *The approximation of the density by Chebyshev polynomials is chosen automatically by* `chebfun`: *the degree is* 394 *for the left panel and* 941 *in the right panel.*

where the operator $L : \mathbb{H}^2_1(\mathbb{R};\mathbb{R}) \mapsto \mathbb{L}^2_1(\mathbb{R};\mathbb{R})$ is self-adjoint with respect to the scalar product

$$\langle \rho_1, \rho_2 \rangle_1 = \int_{\mathbb{R}} \frac{\rho_1(Q)\rho_2(Q)}{\varphi_1(Q)}\,\mathrm{d}Q \quad \text{where} \quad \varphi_1(Q) = \frac{\exp(-2V(Q))}{\int_{\mathbb{R}} \exp(-2V(q))\,\mathrm{d}q}. \tag{5.5}$$

The space $\mathbb{L}^2_1(\mathbb{R};\mathbb{R})$ is in our case the space of all measurable functions $u : \mathbb{R} \to \mathbb{R}$ with $\int_{\mathbb{R}} u^2(x)/\varphi_1(x)\,\mathrm{d}x < \infty$ (a subset of $\mathbb{L}^2(\mathbb{R};\mathbb{R})$, which has the scalar product $\langle \rho_1, \rho_2 \rangle = \int_{\mathbb{R}} \rho_1(Q)\rho_2(Q)\,\mathrm{d}Q$). The space $\mathbb{H}^\ell_1(\mathbb{R};\mathbb{R})$ is the space of all $u \in \mathbb{L}^2_1(\mathbb{R};\mathbb{R})$ with $u^{(j)} \in \mathbb{L}^2_1(\mathbb{R};\mathbb{R})$ for all $j \le \ell$. The spectrum of L is real and consists of point spectrum only. It has the form $0 = \lambda_1 > \lambda_2 > \ldots$ with eigenvectors $\varphi_j(Q) \in \mathbb{H}^2_1(\mathbb{R};\mathbb{R})$ that can be orthonormalized with respect to $\langle \cdot, \cdot \rangle_1$. The function φ_1 is the eigenvector for the trivial eigenvalue $\lambda_1 = 0$ (which is present due to the preservation of total probability $\int_{\mathbb{R}} \rho(Q,t)\,\mathrm{d}Q$ along trajectories). The spectrum and the corresponding eigenfunctions φ_j are shown in Figure 5.2 (left panel), together with the $\mathbb{L}^2$-adjoint

eigenfunctions φ_j/φ_1 (right panel). A solution of the Fokker-Planck equation (5.3) can be expanded in the eigenfunctions of L with time-dependent coefficients $a_j(t)$:

$$\rho(Q,t) = \sum_{j=1}^{\infty} a_j(t)\varphi_j(Q). \tag{5.6}$$

The coefficients satisfy $\dot{a}_j(t) = \lambda_j a_j(t)$ for all j, and the series $\sum_{j=1}^{\infty} a_j^2$ converges for all $t > 0$. The orthonormality of the basis $\{\varphi_j : j \geq 1\}$ with respect to $\langle\cdot,\cdot\rangle_1$, defined in (5.5), implies that

$$\int_{-\infty}^{\infty} \varphi_1(Q)\,\mathrm{d}Q = 1, \qquad \int_{-\infty}^{\infty} \varphi_j(Q)\,\mathrm{d}Q = 0 \quad \text{for } j \geq 2. \tag{5.7}$$

Since $\lambda_1 = 0$, $a_1(t)$ equals $a_1(0)$ for all times $t \geq 0$. One usually chooses $a_1(0) = 1$ such that $\rho(Q,t)$ converges to the *stationary density* $\varphi_1(Q)$ for $t \to \infty$. While Theorem 3.3 was only formulated for flows in $\mathbb{R}^D$, the linearity of L implies that statements identical to Theorem 3.3 can be made for the PDE (5.3). Instead of Fenichel's Theorem on invariant manifolds in ODEs [15] (persistence and regularity of invariant manifolds and fiber projections) we rely on the spectral mapping properties for the linear operator L. For any chosen dimension d of the slow variables, the slow manifold $\mathcal{C}$ is the subspace spanned by $\varphi_1, \ldots, \varphi_d$. Instead of the stable fiber projection in the finite-dimensional case, we have the linear spectral (for L) projection $g : \mathbb{L}^2_1(\mathbb{R};\mathbb{R}) \mapsto \mathcal{C}$ (M is also linear, such that we write $M(t)\rho$ and $g\rho$) which is explicitly known in terms of the eigenvectors of L:

$$g : \mathbb{L}^2_1(\mathbb{R};\mathbb{R}) \ni \rho \mapsto g\rho = \sum_{j=1}^{d} \langle\varphi_j, \rho\rangle_1 \varphi_j \in \mathcal{L}(\varphi_1, \ldots, \varphi_d) \subset \mathbb{L}^2_1(\mathbb{R};\mathbb{R}). \tag{5.8}$$

With this definition of $\mathcal{C}$ and g the decay and growth properties of the evolution map M replacing Assumption 3.1 are

$$\|M(t)|_{\mathcal{C}}\| \leq C\exp(\lambda_d t) \qquad \text{for all } t \leq 0, \quad \|M(t)|_{\mathcal{C}}\| \leq C \quad \text{for all } t \geq 0 \tag{5.9}$$

$$\|M(t) - M(t)\circ g\| \leq C\exp(\lambda_{d+1} t) \quad \text{for all } t \geq 0 \tag{5.10}$$

and some constant C, such that $d_{\tan} = -\lambda_d$, $d_{\mathrm{tr}} = -\lambda_{d+1}$. The approximation statement of Theorem 3.3 then follows immediately from error estimates for finite-dimensional matrices and will be derived after the definition of the restriction and lifting operators. The lifting and restriction operators are chosen to map from a macroscopic description of ρ, for example, by moments, to the full density ρ and vice versa. In particular, we will investigate the behaviour of implicit equation-free methods for $d = 3$ using the following restriction and two different lifting operators:

$$\mathcal{R} : \mathbb{L}^2_1(\mathbb{R};\mathbb{R}) \mapsto \mathbb{R}^d \qquad \mathcal{R}\,\rho = \left(\int_{\mathbb{R}} Q^{k-1}\rho(Q)\,\mathrm{d}Q\right)_{k=1,\ldots,d}, \tag{5.11}$$

$$\mathcal{L}_{\mathrm{lin}} : \mathbb{R}^d \mapsto \mathbb{L}^2_1(\mathbb{R};\mathbb{R}) \qquad \mathcal{L}_{\mathrm{lin}}(x)(Q) = \sum_{j=1}^{d} x_j\rho_j(Q), \tag{5.12}$$

$$\mathcal{L}_{\mathrm{Gauss}} : \mathbb{R}^3 \mapsto \mathbb{L}^2_1(\mathbb{R};\mathbb{R}) \qquad \mathcal{L}_{\mathrm{Gauss}}(x)(Q) = \frac{x_1}{\sqrt{2\pi x_3}}\exp\left(\frac{-(Q-x_2)^2}{2x_3}\right). \tag{5.13}$$

Thus, $\mathcal{R}$ projects a density onto its first d moments (counting from the zeroth moment, which is preserved by M since $\lambda_1 = 0$). In a Monte-Carlo simulation the zeroth moment would correspond to the (possibly scaled) number of realizations. The functions ρ_j in the definition (5.12) of $\mathcal{L}_{\text{lin}}$ are arbitrary in $\mathbb{L}^2_1(\mathbb{R};\mathbb{R})$ with $\int_{\mathbb{R}} \rho_j(Q)\,\mathrm{d}Q = 1$, which ensures that $\int_{\mathbb{R}} \mathcal{L}_{\text{lin}}(x)(Q)\,\mathrm{d}Q = \sum_{j=1}^{d} x_j$ is conserved under $M(t)$. For $\mathcal{L}_{\text{Gauss}}$, the x_1 component is preserved under $M(t)$ and becomes the first component of $\mathcal{R}$ such that always $[\mathcal{R}\, M(t)\, \mathcal{L}_{\text{Gauss}}\, x]_1 = x_1$.

For the combination of $\mathcal{L}_{\text{lin}}$ and $\mathcal{R}$ all components of the lift-evolve-restrict map $P(t;\cdot) = \mathcal{R} \circ M(t;\cdot) \circ \mathcal{L}$ and its exact counterpart $P_*(t;\cdot) = \mathcal{R} \circ M(t;\cdot) \circ g \circ \mathcal{L}$ from section 3 are linear such that we can reduce the study of convergence for arbitrary coordinates x to convergence estimates for matrices.

The combination of $\mathcal{L}_{\text{Gauss}}$ and $\mathcal{R}$ was studied in detail in [5] for explicit equation-free methods, where the authors observed that the nonlinearity of $\mathcal{L}_{\text{Gauss}}$ introduced a nonlinearity in the moment map and that the resulting flow depended qualitatively on the choice of the healing time t_{skip}. We will demonstrate that for $\mathcal{L}_{\text{Gauss}}$ the implicitly defined flow $\Phi_{\text{Gauss},t_{\text{skip}}}$ converges to a nonlinear transformation of the linear flow $M(t)|_{\mathcal{C}}$. Since the x_1 component does not change under $P(t;\cdot)$ and $P_*(t;\cdot)$ for $\mathcal{L} = \mathcal{L}_{\text{Gauss}}$, it can be ignored, making the choice of $\mathcal{L}_{\text{Gauss}}$ and $\mathcal{R}$ identical to the situation studied in [5].

We use the MATLAB [21] package `chebfun` [10, 14] to numerically compute the spectrum and eigenfunctions of L, the flow M, the projection g, restriction and lifting for the example potential V given in (5.2). The package `chebfun` uses Chebyshev polynomials of adaptive degree to approximate arbitrary functions on finite intervals to optimal precision. For a typical result, as shown in Figure 5.2 the degree is larger than 100 (394 for the left panel, 941 for the right panel). The numerically computed spectrum of L is

$$\text{spec}(L) = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \ldots) \approx (-2.37 \cdot 10^{-9}, -8.63 \cdot 10^{-8}, -5.71, -10.3, \ldots) \tag{5.14}$$

for V with the parameters

$$\mu = 6, \qquad \nu = 0.3. \tag{5.15}$$

Note that $\lambda_1 = 0$ is the correct value for the first eigenvalue on an infinite domain. In numerical computations we choose a bounded domain $[-10, 10]$ with Dirichlet boundary conditions, leading to a small probability of escape from the domain. The spectrum and the corresponding eigenfunctions φ_j are shown in Figure 5.2. The eigenvector φ_1 corresponds to the stationary solution of the Fokker-Planck equation and φ_2 is the mode representing escape from one well to another.

5.2. Convergence for the linear lifting operator $\mathcal{L}_{\text{lin}}$ with $d = 3$. We express the maps $P_*(t;\cdot)$ and $P(t;\cdot)$ in terms of M, the eigenvectors φ_j and the scalar product $\langle\cdot,\cdot\rangle_1$, initially for a general dimension d. The exact macroscopic flow Φ_* is defined using the map $P_*(t;\cdot)$ in (3.9), and the approximate macroscopic flow $\Phi_{t_{\text{skip}}}$ is defined using the map $P(t;\cdot)$ in (3.11). The definitions (5.11) for $\mathcal{R}$ and (5.12) for

$\mathcal{L}_{\mathrm{lin}}$ imply

$$[P_{\mathrm{lin},*}(t)x]_k = [\mathcal{R}\, M(t) g\, \mathcal{L}_{\mathrm{lin}}\, x]_k = \sum_{\ell,\, j=1}^{d} \exp(\lambda_\ell t) \int_{\mathbb{R}} Q^{k-1} \varphi_\ell(Q)\, \mathrm{d}Q \langle \varphi_\ell, \rho_j \rangle_1 x_j \tag{5.16}$$

$$[P_{\mathrm{lin}}(t)x]_k = [\mathcal{R}\, M(t)\, \mathcal{L}_{\mathrm{lin}}\, x]_k = \sum_{j=0}^{d} \left[\int_{\mathbb{R}} Q^{k-1} [M(t)\rho_j](Q)\, \mathrm{d}Q \right] x_j \tag{5.17}$$

where $k = 1, \ldots, d$. Using the $d \times d$ matrices $(k, \ell, j = 1, \ldots, d)$

(5.18)
$$(T_{\mathrm{lin}})_{\ell,j} = \langle \varphi_\ell, \rho_j \rangle_1, \quad M_d(t) = \mathrm{diag}\left[\exp(\lambda_\ell t)_{\ell=1,\ldots,d}\right], \quad (R_d)_{k,\ell} = \int_{\mathbb{R}} Q^{k-1} \varphi_\ell(Q)\, \mathrm{d}Q.$$

we can express the map $P_{\mathrm{lin},*}(t)$ and the exact slow flow $\Phi_{\mathrm{lin},*}(\delta)$ in the form

$$P_{\mathrm{lin},*}(t)x = R_d M_d(t) T_{\mathrm{lin}} x \tag{5.19}$$

$$\Phi_{\mathrm{lin},*}(\delta)x = P_{\mathrm{lin},*}(t_{\mathrm{skip}})^{-1} P_{\mathrm{lin},*}(t_{\mathrm{skip}} + \delta) = T_{\mathrm{lin}}^{-1} M_d(\delta) T_{\mathrm{lin}} x. \tag{5.20}$$

General Assumption 3.2 on transversality of $\mathcal{R}$ and $\mathcal{L}$ for Theorem 3.3, when applied to the SDE (5.1) and $\mathcal{R}$ and $\mathcal{L}_{\mathrm{lin}}$, demands the regularity of the matrices R_d and T_{lin}. If both matrices are indeed regular then $P_{\mathrm{lin},*}$ is invertible: $P_{\mathrm{lin},*}(t)^{-1} = T_{\mathrm{lin}}^{-1} M_d(-t) R_d^{-1}$. Thus, the claim of Theorem 3.3 can be simplified to a statement about perturbations of matrices using the quantities

(5.21)
$$n(t) := \|P_{\mathrm{lin},*}(t)^{-1}\| \le \|T_{\mathrm{lin}}^{-1}\| \|R_d^{-1}\| \exp(-\lambda_d t) \quad \text{for } t \ge 0, \text{ and}$$

(5.22)
$$r(t) := \|P_{\mathrm{lin},*}(t) - P_{\mathrm{lin}}(t)\| \le C \exp(\lambda_{d+1} t) \qquad \text{for all } t \ge 0 \text{ and a fixed } C > 0.$$

The estimate for $r(t)$ follows from (5.16) and (5.17):

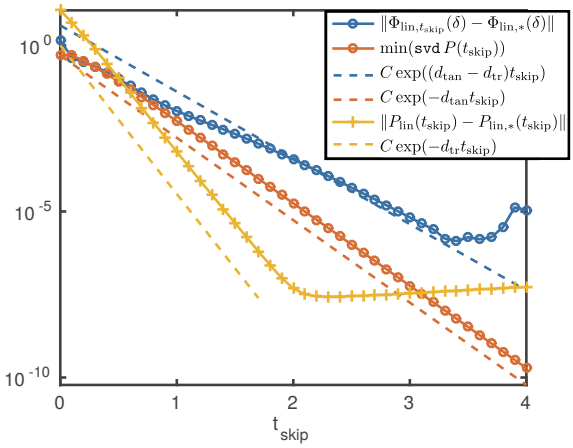
$$\begin{aligned} \left| [P_{\mathrm{lin}}(t)x]_k - [P_{\mathrm{lin},*}(t)x]_k \right| &= \left| \sum_{j=0}^{d} \left[\int_{\mathbb{R}} Q^{k-1} \left[M(t) \left[\rho_j - \sum_{\ell=1}^{d} \langle \varphi_\ell, \rho_j \rangle \varphi_\ell \right] \right] (Q)\, \mathrm{d}Q \right] x_j \right| \\ &\le C \exp(\lambda_{d+1} t) |x|. \end{aligned}$$

The integrand contains the spectral projection $\rho \mapsto \rho - \sum_{\ell=0}^{d} \langle \varphi_\ell, \rho_j \rangle \varphi_\ell$ onto the complement of the space spanned by $\varphi_1, \ldots, \varphi_d$. On the complement of $\mathcal{L}(\{\varphi_1, \ldots, \varphi_d\})$ the evolution operator $M(t)$ decays exponentially with rate λ_{d+1} in time. Together with the boundedness of $\mathcal{R}$ and the spectral projection, this decay of $M(t)$ implies estimate (5.22). Since $P_{\mathrm{lin}}(t; \cdot)$ is linear, the approximate flow map $\Phi_{\mathrm{lin},t_{\mathrm{skip}}}(\delta)$ is given by

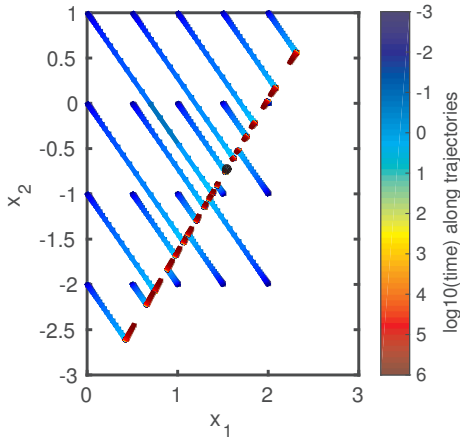
$$\Phi_{\mathrm{lin},t_{\mathrm{skip}}}(\delta) = P_{\mathrm{lin}}(t_{\mathrm{skip}})^{-1} P_{\mathrm{lin}}(t_{\mathrm{skip}} + \delta) \tag{5.23}$$

assuming the inverse of $P_{\mathrm{lin}}(t_{\mathrm{skip}})$ exists (all involved matrices have dimension $d \times d$). The linear expressions for the exact flow $\Phi_{\mathrm{lin},*}$, (5.20), and $\Phi_{\mathrm{lin},t_{\mathrm{skip}}}$, (5.23), imply

(5.24)
$$\|\Phi_{\mathrm{lin},*}(\delta) - \Phi_{\mathrm{lin},t_{\mathrm{skip}}}(\delta)\| \le \frac{n(t_{\mathrm{skip}})}{1 - n(t_{\mathrm{skip}}) r(t_{\mathrm{skip}})} \left[r\left(t_{\mathrm{skip}}\right) \|\Phi_{\mathrm{lin},*}(\delta)\| + r(t_{\mathrm{skip}} + \delta) \right].$$



(a) Convergence of $\Phi_{\mathrm{lin},t_{\mathrm{skip}}}$ to $\Phi_{\mathrm{lin},*}$



(b) Phase portrait of $\Phi_{\mathrm{lin},*}$

FIG. 5.3. *Analysis for* $\mathcal{L}_{\mathrm{lin}}$ *and* $\Phi_{\mathrm{lin},t_{\mathrm{skip}}}$ *with* $d = 3$. *(a) Matrix norm of* $\Phi_{\mathrm{lin},t_{\mathrm{skip}}}(\delta) - \Phi_{\mathrm{lin},*}(\delta)$ *and smallest singular value of* $P(t_{\mathrm{skip}})$*, compared to theoretical exponential estimates. (b): the plot shows trajectories of* $\Phi_{\mathrm{lin},*}$ *in the* (x_1, x_2)*-plane starting on a grid of initial points. The coloring indicates the time along the trajectory (in logarithmic scale). The black point is the fixed point (see text). Parameters:* $\delta = 0.1$*,* ρ_j $(j = 1 \dots 3)$ *are Gaussians with means* -1.5*,* -0.5 *and* 1 *and variance* 1*, shape parameters of potential* V *are* $\mu = 6$*,* $\nu = 0.3$*, see also* (5.14) *and* (5.15).

The exponential estimates $r(t) \leq C\exp(\lambda_{d+1}t)$ in (5.22) and $n(t) \leq \exp(-\lambda_d t)$ in (5.21) immediately imply the statement of Theorem 3.3 (given that $\Phi_{\mathrm{lin},*}$ is globally bounded).

The semilogarithmic plot in Figure 5.3(a) shows the difference between $\Phi_{\mathrm{lin},t_{\mathrm{skip}}}(\delta)$ and $\Phi_{\mathrm{lin},*}(\delta)$ (blue line with circles) for $d = 3$, for a linear basis of three Gaussians ρ_j (with variance 1 and means -1.5, -0.5 and 1), $\delta = 0.1$ and the double-well potential well V with parameters $\mu = 6$, $\nu = 0.3$. The decay rate inside the slow manifold $\mathcal{C} = \mathrm{span}(\varphi_1, \varphi_2, \varphi_3)$ is $d_{\mathrm{tan}} = -\lambda_3 \approx 5.71$ and the attraction rate toward $\mathcal{C}$ is $d_{\mathrm{tr}} = -\lambda_4 \approx 10.3$. Figure 5.3(a) also shows the two components of the error $\Phi_{\mathrm{lin},t_{\mathrm{skip}}}(\delta) - \Phi_{\mathrm{lin},*}(\delta)$ and their theoretical estimates:

1. (In yellow) The difference $r(t_{\mathrm{skip}})$ between $P_{\mathrm{lin}}(t_{\mathrm{skip}}) = \mathcal{R} \circ M(t_{\mathrm{skip}}) \circ \mathcal{L}_{\mathrm{lin}}$ and $P_{\mathrm{lin},*}(t_{\mathrm{skip}}) = \mathcal{R} \circ M(t_{\mathrm{skip}}) \circ g \circ \mathcal{L}_{\mathrm{lin}} = R_d M_d(t_{\mathrm{skip}}) T_{\mathrm{lin}}$, which decays according to the attraction toward $\mathcal{C}$ until it reaches the limits of numerical accuracy of `chebfun` ($\sim 10^{-8}$): $\|P_{\mathrm{lin}}(t_{\mathrm{skip}}) - P_{\mathrm{lin},*}(t_{\mathrm{skip}})\| \sim \exp(-d_{\mathrm{tr}} t_{\mathrm{skip}})$.
2. (In red) The norm $n(t_{\mathrm{skip}})$ of the inverse of $P_{\mathrm{lin}}(t_{\mathrm{skip}})$, which grows like $\exp(d_{\mathrm{tan}} t_{\mathrm{skip}})$. Figure 5.3(a) shows the inverse (the minimal singular value).

The overall error (5.24) is approximately the product of these two components, which is proportional to $\exp((d_{\mathrm{tan}} - d_{\mathrm{tr}}) t_{\mathrm{skip}})$ (shown as a blue dashed line in Figure 5.3(a)). In particular, the combination of $\|P_{\mathrm{lin},*}(t_{\mathrm{skip}})^{-1}\| \sim \exp(d_{\mathrm{tan}} t_{\mathrm{skip}})$, $\|P_{\mathrm{lin}}(t_{\mathrm{skip}}) - P_{\mathrm{lin},*}(t_{\mathrm{skip}})\| \sim \exp(-d_{\mathrm{tr}} t_{\mathrm{skip}})$ and $d_{\mathrm{tan}} < d_{\mathrm{tr}}$ implies that $P_{\mathrm{lin}}(t_{\mathrm{skip}})$ is invertible for sufficiently large t_{skip} and that $\|P_{\mathrm{lin}}(t_{\mathrm{skip}})^{-1}\| \sim \exp(d_{\mathrm{tan}} t_{\mathrm{skip}})$.

Figure 5.3(b) shows a phase portrait of the exact flow in the coordinates in $\mathrm{dom}\, \mathcal{L}_{\mathrm{lin}}$. Since $\Phi_{\mathrm{lin},*}$ and $\Phi_{\mathrm{lin},t_{\mathrm{skip}}}$ both preserve the quantity $\sum_{j=1}^{d} x_j$ (which corresponds to $\int_{\mathbb{R}} \mathcal{L}_{\mathrm{lin}}(x)(Q)\, \mathrm{d}Q$), we set $x_3 = 1 - x_1 - x_2$ in the initial values for the sample trajectories, keeping $\sum_{j=1}^{d} x_j = 1$ along trajectories without loss of generality. This leads to an affine flow in the (x_1, x_2)-plane with a non-trivial fixed point (shown in black in Figure 5.3(b)). The coloring along the sample trajectories illustrates the extreme difference in the time scale along the directions corresponding to λ_2 ($\approx -10^{-7}$; escape between wells), mostly evolving on time scales $\gg 10^4$ (dark red in Figure 5.3(b)), and λ_3 (≈ -5.71; relaxation into the nearest well), mostly decaying on time scale of order 1 and less (blue and light blue in Figure 5.3(b)).

Remark: Densities with sign changes in subsection 5.2. The phase portrait Figure 5.3(b) of the exact flow $\Phi_{\mathrm{lin},*}$ includes coordinates $x = (x_1, x_2, x_3)$ where the lifted initial density $\mathcal{L}_{\mathrm{lin}}(x)$ has sign changes. This is not unphysical. If one performs Monte Carlo simulations with ensembles on the example with the lifting operator $\mathcal{L}_{\mathrm{lin}}(x)$, one would run a Monte Carlo simulation on an ensemble for each of the three initial densities ρ_j. Then one would sum the densities at the end of the simulation with the weights x_j ($j = 1, \ldots, 3$). These weights can be negative to get a combined density.

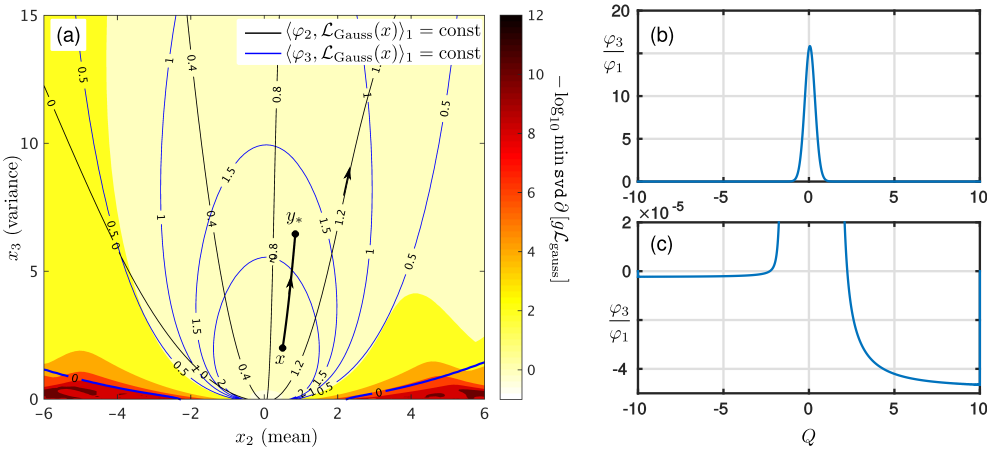


FIG. 5.4. *(a) Contour curves in coordinates in* $\mathrm{dom}\, \mathcal{L}_{\mathrm{Gauss}}$ *and condition of* $\partial T_{\mathrm{Gauss}} = \partial[g\, \mathcal{L}_{\mathrm{Gauss}}]$. *The piece of trajectory from* x *to* $y_* = \Phi_{\mathrm{Gauss},*}(\delta; x)$ *with* $\delta = 0.1$ *is used for convergence analysis in subsection 5.3. (b) Profile of* $\varphi_3(Q)/\varphi_1(Q)$ *in* Q. *Note that away from its peak around* 0, *the profile is slightly negative (see zoom in panel (c)). Parameters*: $\mu = 6$, $\nu = 0.3$.

5.3. Convergence for the nonlinear lifting operator $\mathcal{L}_{\text{Gauss}}$. The exact and approximate lift-evolve-restrict maps for lifting with a Gaussian distribution of mass x_1, mean x_2 and variance x_3, of the form $\mathcal{L}_{\text{Gauss}}(x) = q \mapsto x_1 \exp(-(q - x_2)^2/(2x_3))/\sqrt{2\pi x_3}$, are given by

$$\begin{aligned} P_{\text{Gauss},*}(t;x)_k &= [\mathcal{R}\, M(t) g\, \mathcal{L}_{\text{Gauss}}(x)]_k \\ &= \sum_{\ell=1}^{3} \exp(\lambda_\ell t) \int_{\mathbb{R}} Q^{k-1} \varphi_\ell(Q)\, \mathrm{d}Q \langle \varphi_\ell, \mathcal{L}_{\text{Gauss}}(x)\rangle_1 \end{aligned} \tag{5.25}$$

$$P_{\text{Gauss}}(t;x)_k = [\mathcal{R}\, M(t)\, \mathcal{L}_{\text{Gauss}}(x)]_k = \int_{\mathbb{R}} Q^{k-1} M(t)\, \mathcal{L}_{\text{Gauss}}(x)\, \mathrm{d}Q, \tag{5.26}$$

where $k = 1, \ldots, d$ ($d = 3$). The flow $M(t)$ preserves the integral of the initial distribution such that $P_{\text{Gauss}}(t;x)_1 = x_1$ and $P_{\text{Gauss},*}(t;x)_1 = x_1$. Thus, we can fix $x_1 = 1$ without loss of generality and focus on the dynamics in the (x_2, x_3)-plane in $\operatorname{dom} \mathcal{L}_{\text{Gauss}}$.

Phase portrait of the exact flow $\Phi_{\text{Gauss},*}$. The exact flow $\Phi_{\text{Gauss},*}$ on $\mathcal{C}$ in the coordinates of $\operatorname{dom} \mathcal{L}_{\text{Gauss}}$ is a nonlinear transformation of the linear map $M_d(t) = \operatorname{diag}\left[\exp(\lambda_\ell t)_{\ell=1}^{\ell=d}\right] : \mathbb{R}^d \mapsto \mathbb{R}^d$, defined in (5.18) (with $d = 3$). We call the nonlinear transformation

$$T_{\text{Gauss}} : \mathbb{R}^3 \mapsto \mathbb{R}^3 \qquad T_{\text{Gauss}}(x)_k = \langle \varphi_k, \mathcal{L}_{\text{Gauss}}(x)\rangle_1 \qquad (k = 1, \ldots, 3). \tag{5.27}$$

In particular $T_{\text{Gauss}}(x)_1 = x_1$ by construction. Using T_{Gauss}, M_d and the matrix R_d (defined in (5.18)), the map $P_{\text{Gauss},*}(t;x)$, and the exact flow $\Phi_{\text{Gauss},*}$ are given by (using the notation T_{Gauss}^{-1} for the inverse of the nonlinear map T_{Gauss})

$$\begin{aligned} P_{\text{Gauss},*}(t;x) &= R_d M_d(t) T_{\text{Gauss}}(x), \\ \Phi_{\text{Gauss},*}(\delta;x) &= T_{\text{Gauss}}^{-1}(M_d(\delta) T_{\text{Gauss}}(x)), \end{aligned} \tag{5.28}$$

where all involved quantities are maps from $\mathbb{R}^3$ to $\mathbb{R}^3$. Since the map T_{Gauss} is nonlinear, it is not clear if the inverse exists for all $x \in \mathbb{R}^3$, or if it is unique where it exists. Figure 5.4(a) shows the contours of $T_{\text{Gauss}}(x)_2$ (in black) and $T_{\text{Gauss}}(x)_3$ (in blue; remember that $T_{\text{Gauss}}(x)_1 = x_1$), and the norm of $[\partial T_{\text{Gauss}}(x)]^{-1}$ as color shading (in logarithmic scale). Since the difference in time scale between motion along φ_2 and motion along φ_3 is large ($0 > \lambda_2 \gg \lambda_3$), the flow $\Phi_{\text{Gauss},*}$ follows the black curves in the direction of the arrow until it reaches the zero-level of $T_{\text{Gauss}}(x)_3$ (slightly wider blue curve, only visible close to the bottom of Figure 5.4(a)).

Near-singularity of T_{Gauss}. The zero curve $\{x : T_{\text{Gauss}}(x)_3 = 0\}$ in the (x_2, x_3) plane (wide blue in Figure 5.4(a)) is given by $\int_{\mathbb{R}} \mathcal{L}_{\text{Gauss}}(x)(Q)\varphi_3(Q)/\varphi_1(Q)\, \mathrm{d}Q = 0$, where $\mathcal{L}_{\text{Gauss}}(x)$ is a Gaussian of mean x_2 and variance x_3 and $\varphi_3(Q)/\varphi_1(Q)$ is shown in Figure 5.4(b,c). From the profile of φ_3/φ_1 it is clear that the zero-level forms a single curve connecting the two pieces of the wide blue curve $\{x : T_{\text{Gauss}}(x)_3 = 0\}$ visible in Figure 5.4(a). However, this curve has a large radius (passing through the region $x_3 \gg 1$). For example, there exists a Gaussian $u = \mathcal{L}_{\text{Gauss}}(x)$ with mean $x_2 = 0$ and large variance x_3 such that $T_{\text{Gauss}}(x)_3 = 0$, because φ_3/φ_1 is negative everywhere outside its peak, but the negative values have small modulus (note the scaling of the vertical axis in the zoom of φ_3/φ_1 in Figure 5.4(c)). The fixed point of $x \mapsto \Phi_{\text{Gauss},*}(\delta;x)$ (assuming $x_1 = 1$) is the intersection of the two zero-level curves (not visible in Figure 5.4(a) as it has large x_3). The color shading in Figure 5.4(a) indicates that the nonlinear transformation T_{Gauss} is nearly singular close to the line

$x_3 = 0$, because the $\mathbb{L}^2$-adjoint modes φ_2/φ_1 and φ_3/φ_1 are both nearly constant away from the region around $Q \in [-2, 2]$ (see Figure 5.2, right panel) such that, when inverting T_{Gauss}, the mean x_2 is very sensitive for small changes in the coefficients for the $\mathbb{L}^2$-adjoint modes φ_2/φ_1 and φ_3/φ_1 .

Components of error $\Phi_{\mathrm{Gauss},t_{\mathrm{skip}}} - \Phi_{\mathrm{Gauss},*}$. We perform a detailed convergence analysis along the example trajectory of the exact flow $\Phi_{\mathrm{Gauss},*}$ shown in Figure 5.4(a): $y_* = \Phi_{\mathrm{Gauss},*}(\delta; x)$, where $x = (1, 0.5, 2)^T$ and $\delta = 0.1$ (thus, $y_* \approx (1, 0.8459, 6.4556)^T$). To understand the factors entering the practically achievable lower limit of the error

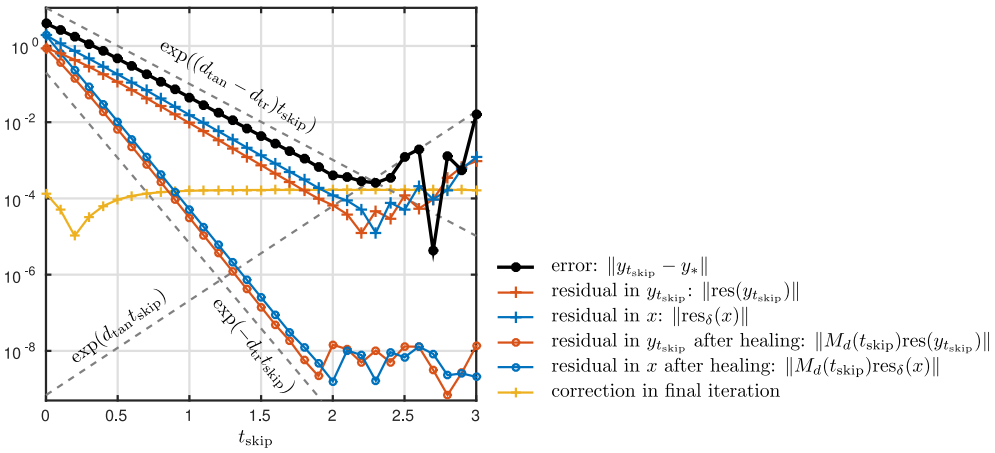


FIG. 5.5. *Convergence analysis along trajectory* $y_* = \Phi_{\mathrm{Gauss},*}(\delta; x)$, *shown in Figure* 5.4 (*a*). *Parameters*: $\mu = 6$, $\nu = 0.3$, $\delta = 0.1$, $x = (1, 0.5, 2)^T$.

$\|y_{t_{\mathrm{skip}}} - y_*\| = \|\Phi_{\mathrm{Gauss},t_{\mathrm{skip}}}(\delta; x) - \Phi_{\mathrm{Gauss},*}(\delta; x)\|$, we consider again the identity (3.16) used in the proof of Theorem 3.3:

$$P_*(t_{\mathrm{skip}}; y_{t_{\mathrm{skip}}}) = P_*(t_{\mathrm{skip}}; y_*) + [P_*(t_{\mathrm{skip}}; y_{t_{\mathrm{skip}}}) - P(t_{\mathrm{skip}}; y_{t_{\mathrm{skip}}})] + [P(t_{\mathrm{skip}} + \delta; x) - P_*(t_{\mathrm{skip}} + \delta; x)], \tag{5.29}$$

but re-arrange it using the concrete expressions for $P_{\mathrm{Gauss},*}$ and P_{Gauss}:

$$\begin{aligned} R_d M_d(t_{\mathrm{skip}}) T_{\mathrm{Gauss}}(y_{t_{\mathrm{skip}}}) = {} & R_d M_d(t_{\mathrm{skip}}) T_{\mathrm{Gauss}}(y_*) \dots \\ & + [R_d M_d(t_{\mathrm{skip}}) T_{\mathrm{Gauss}}(y_{t_{\mathrm{skip}}}) - \mathcal{R}\, M(t_{\mathrm{skip}})\, \mathcal{L}_{\mathrm{Gauss}}(y_{t_{\mathrm{skip}}})] \\ & + [\mathcal{R}\, M(t_{\mathrm{skip}} + \delta)\, \mathcal{L}_{\mathrm{Gauss}}(x) - R_d M_d(t_{\mathrm{skip}} + \delta) T_{\mathrm{Gauss}}(x)] \end{aligned} \tag{5.30}$$

Since the matrices R_d and $M_d(t_{\mathrm{skip}})$ are invertible, we can apply their inverses to both sides in (5.30). For a general distribution ρ, the composition of R_d^{-1} and $\mathcal{R}$

$$T_R\rho : \mathbb{L}^2_1(\mathbb{R}; \mathbb{R}) \mapsto \mathbb{R}^3 \qquad T_R\rho = R_d^{-1}\, \mathcal{R}\, \rho = R_d^{-1} \left[\int_{\mathbb{R}} Q^{k-1} \rho(Q)\, \mathrm{d}Q \right]_{k=1,2,3}$$

is a projection onto the slow manifold $\mathcal{C}$ in the coordinates $(\varphi_1, \varphi_2, \varphi_3)$. Furthermore, the nonlinear map T_{Gauss} is locally invertible in y_* (and, hence, also in $y_{t_{\mathrm{skip}}}$, if $y_{t_{\mathrm{skip}}}$ is near y_*). Its Jacobian is invertible in y_* with a moderate norm of its inverse $\|[\partial T_{\mathrm{Gauss}}(y_*)]^{-1}\| \approx 10$ for the chosen y_*. Hence, the identity (5.30) can be written

in the form

$$
\begin{aligned}
T_{\mathrm{Gauss}}(y_{t_{\mathrm{skip}}}) &= T_{\mathrm{Gauss}}(y_*) + \dots \\
&\underbrace{T_{\mathrm{Gauss}}(y_{t_{\mathrm{skip}}}) - M_d(-t_{\mathrm{skip}}) T_R M(t_{\mathrm{skip}})\, \mathcal{L}_{\mathrm{Gauss}}(y_{t_{\mathrm{skip}}})}_{\mathrm{res}(y_{t_{\mathrm{skip}}})} + \dots \\
&\underbrace{M_d(-t_{\mathrm{skip}}) T_R M(t_{\mathrm{skip}} + \delta)\, \mathcal{L}_{\mathrm{Gauss}}(x) - M_d(\delta) T_{\mathrm{Gauss}}(x)}_{\mathrm{res}_\delta(x)}.
\end{aligned} \tag{5.31}
$$

The two residual terms on the right-hand side, labelled $\mathrm{res}(y_{t_{\mathrm{skip}}})$ and $\mathrm{res}_\delta(x)$, are the two contributions to the error, before it gets amplified by a moderate factor ($\|[\partial T_{\mathrm{Gauss}}(y_*)]^{-1}\| \approx 10$) when inverting T_{Gauss}. The spectral properties of the flow M ensure that

$$
M_d(t) T_{\mathrm{Gauss}}(\eta) - T_R M(t)\, \mathcal{L}_{\mathrm{Gauss}}(\eta) \sim \exp(-d_{\mathrm{tr}} t),
$$

where $d_{\mathrm{tr}} = -\lambda_4 \approx 10.3$. Applying this estimate to $\eta = y_{t_{\mathrm{skip}}}$ and $t = t_{\mathrm{skip}}$, and to $\eta = x$ and $t = t_{\mathrm{skip}} + \delta$ gives the asymptotics $\sim \exp(-d_{\mathrm{tr}} t_{\mathrm{skip}})$ in t_{skip} for $M_d(t_{\mathrm{skip}})\,\mathrm{res}(y_{t_{\mathrm{skip}}})$ and $M_d(t_{\mathrm{skip}})\,\mathrm{res}_\delta(x)$, shown in Figure 5.5 (red and blue curves with circles). The healed residuals $M_d(t_{\mathrm{skip}})\,\mathrm{res}(y_{t_{\mathrm{skip}}})$ and $M_d(t_{\mathrm{skip}})\,\mathrm{res}_\delta(x)$ indeed decay with rate d_{tr} until computational errors for computing the distributions dominate (in this case 10^{-8}). The matrix $M_d(t_{\mathrm{skip}})^{-1} = M_d(-t_{\mathrm{skip}})$ has norm of order $\exp(d_{\tan} t_{\mathrm{skip}})$ (where $d_{\tan} = -\lambda_3 \approx 5.71$; see grey dashed line sloping upward in Figure 5.5) such that the residuals $\mathrm{res}(y_{t_{\mathrm{skip}}})$ and $\mathrm{res}_\delta(x)$ are of order $\sim \exp((d_{\tan} - d_{\mathrm{tr}}) t_{\mathrm{skip}})$ (blue and red curves with + marks in Figure 5.5). The residuals indeed decrease with rate $d_{\mathrm{tr}} - d_{\tan}$ for increasing t_{skip} until the amplification of the computational errors by $\exp(d_{\tan} t_{\mathrm{skip}})$ starts to dominate (at $t_{\mathrm{skip}} \approx 2$). The true error $y_{t_{\mathrm{skip}}} - y_*$ (shown in black in Figure 5.5) is then amplified approximately by the norm of $\|[\partial T_{\mathrm{Gauss}}(y_*)]^{-1}\| \approx 10$, because the residuals res and res_δ occur on the manifold $\mathcal{C}$ (in the coordinates $(\varphi_1, \varphi_2, \varphi_3)$), while the error $y_{t_{\mathrm{skip}}} - y_*$ is defined in $\mathrm{dom}\,\mathcal{L}$. The relation between the error $y_{t_{\mathrm{skip}}} - y_*$ and the residual errors is independent of t_{skip}. Overall, the error $y_{t_{\mathrm{skip}}} - y_*$ decays with rate $d_{\mathrm{tr}} - d_{\tan}$ asymptotically for increasing t_{skip}, but the computational error grows with rate $d_{\tan}$. The optimal healing time t_{skip} is when both errors are of the same order of magnitude.

The identity (5.31) becomes a nonlinear fixed-point problem after applying T_{Gauss}^{-1}, for which the right-hand side is a contraction for sufficiently large t_{skip} (see the proof of Theorem 3.3). For Figure 5.5 we applied this fixed-point iteration. The final fixed-point iteration correction (shown as a yellow curve in Figure 5.5 is always smaller than the error $y_{t_{\mathrm{skip}}} - y_*$.

5.4. The size of computational errors in ensemble computations. The results shown in Figures 5.3(a) and 5.5 show the qualitative behaviour of implicit lifting for increasing t_{skip}. Two sources contribute to the overall error. One source is the mismatch between the trajectory started from the lifted point and the projected (along the stable fiber) trajectory on the slow manifold. The size of this contribution is estimated in Theorem 3.3 as decaying with rate $d_{\mathrm{tr}} - d_{\tan}$ with increasing t_{skip} (also observed in Figures 5.3(a) and 5.5). The other source is the limited accuracy in the computations of the lifting $\mathcal{L}$, the microscopic flow M and the restriction $\mathcal{R}$. Errors introduced from this limited accuracy grow with rate $d_{\tan}$ for increasing t_{skip}. The analysis in Figures 5.3(a) and 5.5 illustrates the trade-off between these two sources of error when the computational error is small ($\approx 10^{-8}$, using `chebfun` [10, 14]).

If the microscopic flow M describes a multi-particle or high-dimensional stochastic system and is estimated using ensembles of realizations then the computational error of the flow estimate (and, possibly, the computation of $\mathcal{L}$ and $\mathcal{R}$) is determined by the ensemble size N. This error decreases asymptotically like $1/\sqrt{N}$ for increasing N, unless one is able to apply variance reduction techniques (see, for example, [3] for a technique to reduce noise in the computations of Jacobians needed to solve nonlinear systems). In this section we demonstrate that the error behavior can be expected to be qualitatively the same as in Figures 5.3(a) and 5.5, but with stricter limitations on t_{skip} due to larger computational errors in $\mathcal{L}$, $\mathcal{R}$, and the flow M. To keep the computations simple and comparable to the previous subsection, we perform a Monte-Carlo simulation directly for the SDE (5.1).

Figure 5.6(a) shows the overall behaviour of the error when performing computations based on random ensembles of finite size N, using the lifting operator $\mathcal{L}_{\text{Gauss}}$, based on Gaussians. For an ensemble size N, mean $\bar{Q}$ and we create a random set of initial conditions

$$[\mathcal{L}(N, \bar{Q}, \operatorname{var} Q)]_n = \bar{Q} + \sqrt{\operatorname{var} Q}\eta, \qquad n = 1, \dots, N \tag{5.32}$$

where $\eta \sim \mathcal{N}(0,1)$ is a random variable drawn from a standard normal distribution for each n. An ensemble of N realizations at positions Q_n is restricted according to

$$\mathcal{R}\left((Q_n)_{n=1}^N\right)_k = \sum_{n=1}^{N} Q_n^{k-1} \quad (k = 1, 2, 3). \tag{5.33}$$

Similar, to the definitions (5.11) and (5.13), the first component of the argument to $\mathcal{L}$ and of the output of $\mathcal{R}$ is the number of realizations, which is preserved. In order to solve (5.1) numerically, we use the Euler-Maruyama scheme

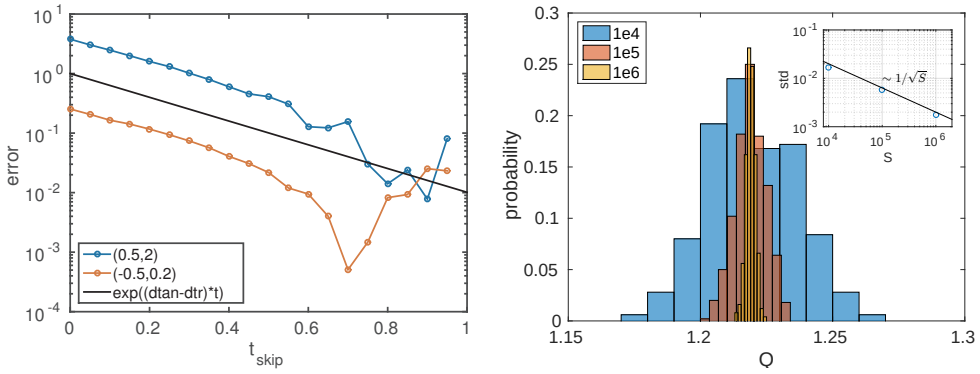
$$Q_n(t+h) = Q_n(t) + f(Q_n(t))h + \sqrt{h}\sigma\xi_n(t) \quad (n = 1, \dots, N), \tag{5.34}$$

where $h = 0.01$ is the step size, $f = -\partial_Q V$ and $\xi_n \sim \mathcal{N}(0,1)$ is standard normal random noise that is uncorrelated, i. e. $\langle \xi_n(t)\xi_n(t')\rangle = \delta(t-t')$.

The error for each t_{skip} in Figure 5.6(a) was estimated by comparing the value of $\Phi_{\text{Gauss},t_{\text{skip}}}(\delta; x)$ to the value $\Phi_{\text{Gauss},t_{\max}}(\delta; x)$ for the largest t_{skip} (called $t_{\max}$, equalling 1). Thus, the value of t_{skip} at which the error starts to grow and the growth rate may not have been captured accurately. However, we observe an exponential decay with increasing t_{skip} over approximately two orders of magnitude and the more stringent limitation on t_{skip}, as the error stops decreasing at $t_{\text{skip}} \approx 0.6$.

Two problems limit the computational accuracy of function evaluations.

1. In Monte Carlo simulations with ensemble size N the evaluation of the macroscopic lift-evolve-restrict map $P(t;\cdot)$ of the dynamics is noisy in $(\bar{Q}, \operatorname{var} Q)$. This is due to the inherent noise in (5.1) and due to the noise in the lifting procedure (5.32). Hence the evaluation of P with the same input parameters might yield different outputs. The result of P is a random variable with an ensemble-dependent distribution (see Figure 5.6(b), where the distribution of the second component (the mean) of $P(1; (N, -0.5, 0.2))$ is shown for a range of N). The standard deviation of P decreases with the ensemble size like $\sim 1/\sqrt{N}$.
2. Function evaluations for large $\operatorname{var} Q$ become computationally difficult since a large $\operatorname{var} Q$ implies sampling of trajectories far away from the minima of the potential. Since the potential is steep away from the minima, the drift forces



(a) (b)

FIG. 5.6. (a) *Error analysis for* $N = 10^7, h = 10^{-2}, \mu = 6, \nu = 0.3, \sigma = 1$, $\delta = 0.1$, *similar to Figure 4.1(f) for two different starting values* $(N, \bar{Q}, \operatorname{var} Q) = (10^7, 0.5, 2)$ *and* $(10^7, -0.5, 0.2)$. *The initial error is larger for the value with a larger variance. With increasing* t_{skip}, *the error shrinks with the same exponential rate for both initial conditions.* (b) *Distribution of function evaluations of the lift-evolve-restrict map* $\mathcal{R}(M(t; \mathcal{L}(\bar{Q}, \operatorname{var} Q)))_2/N$ *for* $t = 1$ *and various ensemble sizes* N *and* $(\bar{Q}, \operatorname{var} Q) = (-0.5, 0.2)$. *Inset: the uncertainty of the function evaluation scales as* $\sim 1/\sqrt{N}$.

V' become large, which results in stability problems of the numerical scheme (5.34) for a fixed step size h.

When solving $P(t_{\text{skip}}; y) = P(t_{\text{skip}} + \delta; x)$ for y in the analysis in Figure 5.6(a) we use a Newton iteration with damping $\gamma = 0.5$ on the macroscopic level with tolerance $\texttt{tol} = 5 \cdot 10^{-2}$ where Jacobians are computed by a central finite-difference scheme with $\Delta\bar{Q} = \Delta \operatorname{var} Q = 5 \cdot 10^{-2}$. The ensemble size is $N = 10^7$. The level of the minimal error is limited by the finite ensemble size N and the accuracy of function evaluations and approximations of the Jacobian in the Newton iterations (see [3] how the accuracy of the Jacobians can be improved).

6. Discussion.

6.1. General estimate for the influence of evaluation errors. While the theoretical convergence result in Theorem 3.3 appears to suggest that a larger t_{skip} always leads to a smaller error, the demonstrations for the Michaelis-Menten kinetics model in section 4 and the SDE in section 5 illustrate that there is a trade-off and, hence, an optimal value for t_{skip} in practice. One source for the difference between the estimates of Theorem 3.3 and numerical observations are numerical errors in the evaluation of lifting $\mathcal{L}$, evolution $M(t; \cdot)$ and restriction $\mathcal{R}$. The effect of these errors grow along trajectories inside the slow manifold $\mathcal{C}$ if the vector field tangent to $\mathcal{C}$ has non-zero expansion rates forward or backward in time. This becomes clear when looking at the arguments in the proof of Theorem 3.3. The approximate solution $y_{t_{\text{skip}}}$ is the fixed point of the map (see Equation (A.9))

$$y \mapsto P_*(t_{\text{skip}}; \cdot)^{-1}\Big(P_*(t_{\text{skip}}; y_*) + [P(t_{\text{skip}}; y) - P_*(t_{\text{skip}}; y)] + [P(t_{\text{skip}} + \delta; x) - P_*(t_{\text{skip}} + \delta; x)]\Big). \tag{6.1}$$

According to Theorem 3.3, $y_{t_{\text{skip}}} - y_* \sim \exp((d_{\tan} - d_{\text{tr}})t_{\text{skip}})$, where $d_{\tan}$ is defined as $\max\{d^+_{\tan}, d^-_{\tan}\}$, the maximum of the forward ($d^+_{\tan}$) and backward ($d^-_{\tan}$) expansion

rate of the flow $M|_{\mathcal{C}}$ tangential to $\mathcal{C}$, and d_{tr} is the rate of attraction transversal to $\mathcal{C}$. However, if we take into account evaluation errors, we have to distinguish between the exact and approximate operators. That is, $P_*(t;\cdot)$ equals $\mathcal{R}\circ M(t;\cdot)|_{\mathcal{C}}\circ g\circ\mathcal{L}$ (recall that g is the stable fiber projection) and $P(t;\cdot)$ equals $\mathcal{R}_\Delta\circ M_\Delta(t;\cdot)\circ\mathcal{L}_\Delta$ where we use the subscript Δ to indicate that the operator is affected by small errors. For $\mathcal{L}_\Delta$ and $\mathcal{R}_\Delta$ this means simply that they are perturbations of $\mathcal{L}$ and $\mathcal{R}$ of size Δ. The evaluation error in M along trajectories in $\mathcal{C}$ causes errors of size

$$\begin{aligned}\|M(t;\cdot)|_{\mathcal{C}} - M_\Delta(t;\cdot)|_{\mathcal{C}}\| &\sim \Delta\exp(d^+_{\tan}t) && \text{for } t\geq 0,\\ \|M(t;\cdot)|_{\mathcal{C}} - M_\Delta(t;\cdot)|_{\mathcal{C}}\| &\sim \Delta\exp(-d^-_{\tan}t) && \text{for } t<0.\end{aligned}$$

These errors in $\mathcal{L}_\Delta$, $\mathcal{R}_\Delta$ and $M_\Delta(t;\cdot)$ are all part of the term $P(t_{\mathrm{skip}}+\delta;x)$ in (6.1)) such that the error grows for increasing t_{skip} at the rate

$$\|P(t_{\mathrm{skip}}+\delta;x) - P_*(t_{\mathrm{skip}}+\delta;x)\| \sim \Delta\exp(d^+_{\tan}t_{\mathrm{skip}}),$$

which gets then amplified by the expansion rate of $M(-t_{\mathrm{skip}};\cdot)|_{\mathcal{C}}$ when applying $P_*(t_{\mathrm{skip}};\cdot)^{-1}$. Thus, there will be an error between the exact fixed point of the map (6.1) and the fixed point with evaluation errors. This error is of order $\Delta\exp((d^+_{\tan}+d^-_{\tan})t_{\mathrm{skip}})$, which is growing exponentially in t_{skip}. This is visible in all computational results:

- In the Michaelis-Menten kinetics model in section 4 the error Δ is of order 10^{-10} and $d_{\tan}$ is of order ε (which is 10^{-2}) such that the growth of the error with t_{skip} is not visible in the range of t_{skip} between 0 and 30 in Figures 4.1(e) and 4.1(f).
- For the stochastic differential equation in section 5, $d^+_{\tan}$ is zero and $d^-_{\tan} = -\lambda_3 \approx 5.71$. For Figures 5.3(a) and 5.5 we computed the evolution of densities directly using the Fokker-Plank equation and `chebfun` such that the evaluation error Δ is of the order 10^{-8} (visible as the lower bound on the residuals $\|P_{\mathrm{lin}}(t_{\mathrm{skip}}) - P_{\mathrm{lin},*}(t_{\mathrm{skip}})\|$ in Figure 5.3(a) and in the residuals after healing in Figure 5.5). Thus, the overall influence of the evaluation error is of order $\Delta\exp(t_{\mathrm{skip}}d^-_{\tan})$. The amplification factor reaches $\sim 10^5$ for $t_{\mathrm{skip}} = 2$. In Figure 5.3(a) evaluation errors dominate only from $t_{\mathrm{skip}} \approx 3$, while in Figure 5.5 they dominate from $t_{\mathrm{skip}} \approx 2.5$.
- In Figure 5.6(a) the growth rate of the evaluation error is the same as in Figure 5.5, but the basic evaluation error of a single time step of $M_\Delta(t,\cdot)$ and the lifting $\mathcal{L}_\Delta$ is larger (as they are generated from ensembles): $\Delta \sim 10^{-3.5}$ for ensemble size $N = 10^7$. Thus, the effects of evaluation error start to dominate already for $t_{\mathrm{skip}} \approx 0.7$. With smaller, more realistic, ensemble sizes the restriction on t_{skip} posed by evaluation errors will be even more severe. Since the necessary length of t_{skip} to reduce the projection error $y_{t_{\mathrm{skip}}} - y_*$ (from Theorem 3.3) is dictated by $d^-_{\tan} - d_{\mathrm{tr}}$, we have a general approximate optimal healing time for positive evaluation errors Δ of the order

$$t_{\mathrm{skip}} \sim \frac{-\log\Delta}{d^+_{\tan} + d_{\mathrm{tr}}},$$

resulting in an optimal error of the order

$$\max\left\{\Delta e^{(d^+_{\tan}+d^-_{\tan})t_{\mathrm{skip}}}, e^{(d^-_{\tan}-d_{\mathrm{tr}})t_{\mathrm{skip}}}\right\} \sim \Delta^p \quad \text{with} \quad p = \frac{d_{\mathrm{tr}} - d^-_{\tan}}{d_{\mathrm{tr}} + d^+_{\tan}}.$$

In the limit of large time scale separation ($d^{\pm}_{\text{tan}}/d_{\text{tr}} \to 0$) the power p of the error reaches 1 and the optimal t_{skip} is of order $-\log \Delta / d_{\text{tr}}$.

6.2. Consequences for equation-free analysis of stochastic systems. The lift-evolve-restrict map $P_{\text{Gauss}}(t;\cdot)$ in section 5 reduced the SDE $\mathrm{d}Q = -V'(Q) + \sigma\,\mathrm{d}W_t$ (or, more precisely, its Fokker-Planck equation) to the slow manifold (a linear subspace) spanned by its first 3 modes. Barkley *et al* [5] observed that the map $P_{\text{Gauss}}(t;\cdot)$ (called *moment map* in [5]) is nonlinear and, hence, suspected that the nonlinearity of P_{Gauss} may be the object of interest for nonlinear analysis (such as finding multiple equilibria, bifurcations under parameter changes, etc). However, as equation (5.28) shows, the exact flow map $\Phi_{\text{Gauss},*}(\delta;\cdot)$ of the low-order moments is still a nonlinear transformation (by T_{Gauss}) of a linear map such that there is no nonlinear dynamic behaviour present. More precisely, the phase portrait of the exact flow map $\Phi_{\text{Gauss},*}(\delta;\cdot)$ is topologically conjugate to the phase portrait of a linear system. Since the approximate flow $\Phi_{\text{Gauss},t_{\text{skip}}}(\delta;\cdot)$, computed with P_{Gauss}, converges to $\Phi_{\text{Gauss},*}(\delta;\cdot)$ for $t_{\text{skip}} \to \infty$ we do not expect nonlinear behavior for $\Phi_{\text{Gauss},t_{\text{skip}}}$ either.

This raises the question what the natural nonlinearity of the underlying system is in the case of equation-free methods applied to stochastic systems.

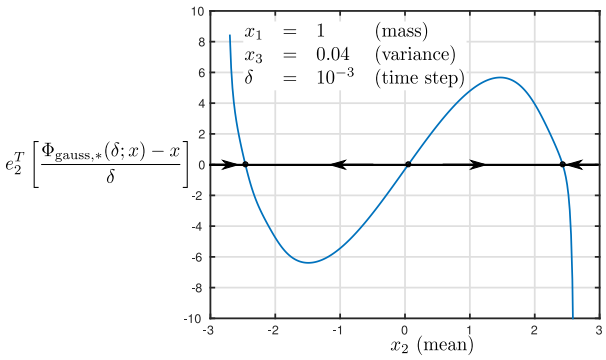


FIG. 6.1. *Apparent one-dimensional nonlinear phase portrait after projection of the nonlinearly transformed two-dimensional phase portrait for* $\Phi_{\text{Gauss},*}(\delta;\cdot)$ *in Figure 5.4(a) onto the horizontal line with variance* $x_3 = 0.04$ (*similar to Figure 14 (left) of Barkley* et al. *[5]*). *For healing time* t_{skip} *with* $1/\lambda_3 \ll t_{\text{skip}} \ll 1/\lambda_2 (\approx 10^6)$, *the approximate flow* $\Phi_{\text{Gauss},t_{\text{skip}}}$ *approximates the exact flow* $\Phi_{\text{Gauss},*}$ *accurately on the eigenspace of* L *corresponding to the* 3 *dominant eigenvalues, but not on the smaller space for the* 2 *dominant eigenvalues. Other parameters*: $\mu = 6$, $\nu = 0.3$, $\delta = 10^{-3}$, $x_1 = 1$, $x_3 = 4 \times 10^{-2}$.

6.2.1. Artificial nonlinearity. Since the Fokker-Planck equation is linear, the apparent nonlinear dynamics arises only due to artificial projections of nonlinearly transformed phase portraits of the linear Fokker-Planck equation when the healing time t_{skip} is not sufficiently large. For example, let us consider again the SDE with lifting to a Gaussian distribution from subsection 5.3. What happens if we choose a moment map for only the zeroth and first moment but an insufficiently large t_{skip} (which would have to be $\sim 1/\lambda_2 \approx 10^6$ to make Theorem 3.3 applicable)? For illustration we choose a lifting to near-delta Gaussian distributions, similar to [5]. In the notation from subsection 5.3 this means that we keep x_1 equal to 1 (mass), vary x_2 (mean) between -3 and 3, and keep $x_3 \ll 1$ (variance) fixed ($x_3 = 0.04$ for the illustration in Figure 6.1). The restriction is then the projection on the zeroth and first moment. If the healing time t_{skip} satisfies $1/\lambda_3 \ll t_{\text{skip}} \ll 1/\lambda_2$ (instead

of $t_{\text{skip}} \sim 1/\lambda_2$), then we obtain for the approximate flow $\Phi_{\text{Gauss},t_{\text{skip}}}$ a projection of the phase portrait Figure 5.4(a) onto the line with $x_3 = 0.04$. Figure 6.1 shows this projected phase portrait (arrows on the x-axis) and the associated right-hand side (in blue). It resembles a phase portrait of a scalar ODE with two coexisting stable fixed points, separated by an unstable fixed point. Of course, this nonlinearity is created artificially by projecting the accurate nonlinearly transformed two-dimensional phase portrait of a linear system onto an arbitrarily chosen line in $\mathbb{R}^2$.

6.2.2. Reduction of high-dimensional SDEs. While in high-dimensional SDEs there is at first sight no obvious nonlinearity present in the evolution of densities (see Fokker-Planck equation (5.3)), the reduction to low-order moments of a multi-particle system with randomness still gives a valid dimension reduction procedure. We give an informal outline of the argument for a particularly simple case in which dimension reduction is in theory possible according to Givon *et al.* [17] (see also textbook [36]).

Let us assume that the simulation (say, an agent-based simulation) can be modelled by a high-dimensional SDE (which is the microscopic model)

$$\mathrm{d}u = F(u)\,\mathrm{d}t + \sigma_u\,\mathrm{d}W_{u,t}, \tag{6.2}$$

where $u \in \mathbb{R}^{n_u}$ and (to keep the argument simple) σ_u is constant and regular, and $W_{u,t}$ are n_u independent instances of Brownian motion. Let us also assume that there exist coordinates $(x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$ $(n_x + n_y = n_u)$ for u such that in these coordinates we have a time scale separation:

$$\mathrm{d}x = \varepsilon f(x,y)\,\mathrm{d}t + \sqrt{\varepsilon}\sigma_x\,\mathrm{d}W_{x,t}, \qquad \mathrm{d}y = g(x,y)\,\mathrm{d}t + \sigma_y\,\mathrm{d}W_{y,t}, \tag{6.3}$$

and that for each x the random variable y converges to its stationary density with rate of order 1 (fast). Let $v_0(x,y)$ be the nullvector of the Fokker-Planck operator of the fast subsystem of (6.3), $p \mapsto L_0 p = \partial_y[\frac{1}{2}\sigma_y^T\sigma_y\partial_y p - gp]$, with $\int v_0(x,y)\,\mathrm{d}y = 1$. Any function of the form $v_0(x,y)p_x(x)$ is also a nullvector of L_0. If $(\varepsilon\lambda, p)$ (with $O(\lambda) = 1$) is an eigenpair of the Fokker-Planck operator $L_0 + \varepsilon L_1$ with $L_1 p = \partial_x[\frac{1}{2}\sigma_x^T\sigma_x\partial_x p - fp]$ for the combined system (6.3) in (x,y) coordinates, then $\lambda = \lambda_0 + O(\varepsilon)$, $p(x,y) = v_0(x,y)p_x(x) + O(\varepsilon)$, where $(\varepsilon\lambda_0, p_x)$ is an eigenpair of the of the right-hand side of the Fokker-Planck equation for the reduced SDE

$$\mathrm{d}x = \varepsilon\tilde{f}(x)\,\mathrm{d}t + \sqrt{\varepsilon}\tilde{\sigma}_x(x)\,\mathrm{d}W_{x,t}. \tag{6.4}$$

In (6.4) $\tilde{f}(x) = \int f(x,y)v_0(x,y)\,\mathrm{d}y$ is the conditional expectation with respect to x of the drift in x and $\tilde{\sigma}_x(x) = \sigma_x\left[\int v_0(x,y)\,\mathrm{d}y/2\right]^{1/2}$ is the standard deviation of x in the stationary distribution of y. Consequently, performing equation-free analysis on the high-dimensional SDE (6.3) using a small number d of variables gives the same results as equation-free analysis on the reduced system (6.4) (up to order ε^2).

Givon *et al.* [17] discuss dimension reduction more generally (independent of explicit spatial coordinates x and y) for Fokker-Planck operators of the form $L_0 + \varepsilon L_1$, assuming that the linear operator L_0 has a non-trivial kernel (dimension greater than 1, implying that ε is a singular perturbation parameter). Hence, equation-free-analysis based on implicit lifting and sufficiently large healing times can be used to perform closure-on-demand, as described in [25], rigorously. Convergence of the approximate system created by lift-evolve-restrict maps to the Fokker-Planck operator of the reduced system (6.4) occurs in the sense of classical singular perturbation

theory toward an attracting low-dimensional linear invariant subspace of densities in the domain of definition of $L_0 + \varepsilon L_1$, as ensured by Theorem 3.3 for sufficiently large healing times t_{skip}.

For the case that the high-dimensional SDE consists of a large number N of random variables (for example, describing agents) our analysis in section 5 and the above discussion raise an important point. Applying the equation-free procedure to initial densities and the Fokker-Planck operator $L_0 + \varepsilon L_1$ does not reduce the high-dimensional SDE to a low-dimensional SDE, but it reduces the high-dimensional SDE to a low-dimensional linear ODE for the coefficients of the leading modes of the Fokker-Planck equation. Hence, increasing the number of variables N (e.g., agents) does *not* increase the spectral gap or the time scale separation. This is obvious for the simple SDE example in section 5: decreasing the noise level will let λ_2/λ_3 converge to 0 (the time scale for escape from one well to the other), but λ_3/λ_4 will remain approximately 1/2. Hence, we need the convergence result for finite time scale separation to prove validity of the model reduction. Results for sufficiently large time scale separation such as those by Zagaris *et al.* [16, 48, 49] (using, for example, constrained runs) and Marschler *et al.* [31] are not applicable to equation-free methods operating on Fokker-Planck equations, if the aim is to extract the decay rate or shape of the dominant modes of the Fokker-Planck equation.

In summary, one possible work flow for analysing a high-dimensional SDE with generator splittable as $L_0 + \varepsilon L_1$ with equation-free methods is: (1) use the equation-free moment map to determine properties of the leading d eigenmodes φ_j and eigenvalues λ_j of $L_0 + \varepsilon L_1$; (2) if these φ_j and λ_j are also the leading eigenmodes and eigenvalues to an operator L_1 for a Fokker-Planck equation of a low-dimensional SDE, identify the properties of L_1 from the modes (for example, singular points of the potential).

7. Outlook. The arguments in section 5, studying the simple scalar SDE $\mathrm{d}Q = -V'(Q)\,\mathrm{d}t + \mathrm{d}W_t$, and the discussion in subsection 6.2 treat SDEs as linear evolution equations for densities. The sections below outline how one may have to modify the arguments of Theorem 3.3 for other tasks of equation-free analysis, which are beyond the scope of this paper.

7.1. Bifurcation analysis for the drift of the reduced system. Assume that we have access to a simulator of a system that can be modelled by a high-dimensional SDEs of type (6.2),

$$\mathrm{d}u = F(u) + \sigma_u\,\mathrm{d}W_{u,t}, \tag{7.1}$$

with time-scale separation as in (6.3). A sensible object for nonlinear equation-free analysis is a bifurcation analysis of the deterministic part $\dot{x} = \tilde{f}(x)$ of the reduced SDE (6.4),

$$\mathrm{d}x = \tilde{f}(x)\,\mathrm{d}t + \sigma_x\,\mathrm{d}W_{x,t}, \tag{7.2}$$

assuming a reduction as discussed in subsection 6.2.2 is possible. For example, one may want to determine its phase portraits and their parameter dependence. If one had a direct simulator of the low-dimensional reduced SDE (7.2), one could approximate $\tilde{f}$ in any given $x_0 \in \mathbb{R}^{n_x}$ via

$$\tilde{f}(x_0) = \lim_{\delta \to 0} \frac{1}{\delta}[EX_\delta - x_0], \tag{7.3}$$

where X_δ (a random variable in $\mathbb{R}^{n_x}$) is the solution of the SDE (7.2) at time δ starting from the deterministic x_0, and $EX_\delta \in \mathbb{R}^{n_x}$ is its expectation.

Equation-free analysis based on a lift-evolve-restrict map P with healing time provides an approximation for (7.3) if only a simulation of the high-dimensional SDE (7.1) is available. The healing time permits the fast variable y to settle to its stationary density $v_0(x,y)$ before one measures $\tilde{f}$. Since the slow-fast coordinate split of u into x and y is unknown, one has to define a lifting $\mathcal{L}$ and a restriction $\mathcal{R}$ between $\mathbb{R}^{n_x}$ and the space of random variables U in $\mathbb{R}^{n_u}$.

Let us assume that the lift $\mathcal{L}(x_L)$ of $x_L \in \mathbb{R}^{n_x}$ is a random variable U_0 in $\mathbb{R}^{n_u}$ with density p_0 on $\mathbb{R}^{n_u}$. The SDE (7.1) creates a Markov process $t \mapsto U_t$ for $t \geq 0$. Let us consider a restriction $\mathcal{R}$ of a random variable U_t that is the expectation $ER(U_t)$ of a map $R : \mathbb{R}^{n_u} \mapsto \mathbb{R}^{n_x}$. Thus, the lift-evolve-restrict map $P : \mathbb{R} \times \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_x}$ is $P(t; x_L) = E[R(U_t)|U_0 = \mathcal{L}(x_L)]$. A good approximation of the deterministic part of the slow flow (in x_L coordinates) would *not* be $(y - x_0)/\delta$ where y is the solution of $P(t_{\text{skip}} + \delta; x_0) = P(t_{\text{skip}}; y)$. Rather, a possible construction is to define $x_R = P(t_{\text{skip}}; x_L)$ and then compute

$$\tilde{f}_L(x_L) \approx \frac{1}{\delta}\left(E\left[R(U_{t_{\text{skip}}+\delta}) \middle| R(U_{t_{\text{skip}}}) = x_R \text{ and } U_0 = \mathcal{L}(x_L) \right] - x_R \right). \tag{7.4}$$

This means that one first solves the SDE for the healing time t_{skip}, then increases time to $t_{\text{skip}} + \delta$, and uses the conditional expectation of $R(U_{t_{\text{skip}}+\delta})$, with the condition that $R(U_{t_{\text{skip}}}) = x_R$. This conditional expectation enters the difference quotient for $\tilde{f}_L(x_L)$, which is otherwise similar to (7.3). Constructions of the form (7.4) do not fit into the framework of Theorem 3.3. Still, we conjecture that the function $\tilde{f}_L$ approximates $\tilde{f}$ (up to a coordinate change from x_L to x) for sufficiently small δ and large t_{skip}. The approximation will become accurate only in the limit of large time scale separation for a set of n_x slow variables (in contrast to Theorem 3.3), but we need only genericity conditions on $\mathcal{L}$ and $\mathcal{R}$.

7.2. Averaging deterministic high-dimensional systems. There is still another gap to applications for multi-particle systems, which are commonly deterministic at the microscopic level. For example, Barkley *et al.* [5] used the scalar SDE (5.1) as a simple model for a heat bath problem where the position Q of a heavy particle of mass M and generalized coordinates (Q, P) is coupled to a heat bath of N smaller particles of masses m_i and generalized coordinates (q_i, p_i) for $i = 1, \ldots, N$. The full system in [5] was described by the Hamiltonian

$$H(Q, P, q, p) = \frac{P^2}{2M} + V(Q) + \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \sum_{i=1}^{N} \frac{k_i}{2}(q_i - Q)^2, \tag{7.5}$$

where the number N of particles is large and the masses m_i and spring coupling constants k_i are small (with particular N-dependent distributions, see [5], eq. (2.2)). The necessary assumption to enable treatment of a fast deterministic subsystem as a stochastic system is some form of ergodicity: any distribution of initial conditions of the fast subsystem converges rapidly to a unique stationary distribution (conditioned on the slow variables). This condition is hard to verify (even empirically) for any particular system. In particular, it is not true for (7.5) if one treats the coordinates (Q, P) as the slow variables since the small masses are only coupled through the heavy particle. Convergence to an SDE is only guaranteed for the system with Hamiltonian (7.5) if the initial conditions for q_i and p_i are set according to the stationary measure

conditioned on P and Q (which was done in [5], see [5, 27, 37] for background results). Hence, the introduction of a healing time t_{skip} will not have an improving effect for equation-free analysis of the heat bath problem (7.5).

7.3. Approximation of stochastic slow manifolds. As mentioned already in the introduction, our convergence result for finite time scale separation relies on a result about model reduction that is valid for finite time scale separation, namely the persistence of normally hyperbolic invariant manifolds and their stable fibers. While the model reduction results for stochastic systems in [17, 36] provide only statements for the limit of infinite time scale separation, stronger results are available for stochastic systems, if one is able to fix the noise realization (for example, the Brownian path) [1, 2]. In this case, the microscopic map M has, for the example of an SDE of the form $\mathrm{d}u = F(u)\,\mathrm{d}t + \sigma_u\,\mathrm{d}W_{u,t}$, the form $M(t; u, \omega)$, where $\omega \in C([0,\infty); \mathbb{R}^D)$ is a realization of the Wiener process $W_{u,t}$ and M satisfies the invariance relation $M(t+s; x, \omega) = M(t; M(s; x, \omega), \omega(s+\cdot))$.

Invariant stochastic manifolds $\mathcal{C}$ are then invariant objects depending on the realization (one may write $\mathcal{C}(\omega)$). Their persistence and attraction properties have been proven for some cases such as finite-dimensional SDEs [9, 47] and SPDEs [11]. For these cases, an implicit equation-free scheme $y = \Phi_{t_{\text{skip}}}(\delta; x, \omega)$ defined implicitly via

$$\mathcal{R}\, M(t_{\text{skip}}; \mathcal{L}(y), \omega) = \mathcal{R}\, M(t_{\text{skip}} + \delta; \mathcal{L}(x), \omega) \tag{7.6}$$

may converge in a similar way as claimed in Theorem 3.3. However, the stochastic invariant manifold results and the implementation of (7.6) depend on the ability to use the same realization ω throughout the computation, as was done in [22] (for example, for different arguments y during a Newton iteration for (7.6)). While fixing the realization is possible for SDEs, for many of the applications for equation-free analysis [18, 24, 30, 32, 33, 42, 45] it is not clear how to do that.

8. Conclusion. This paper proves convergence of equation-free methods, based on lift-evolve-restrict maps $P(t;\cdot) = \mathcal{R} \cdot M(t;\cdot) \circ \mathcal{L}$. Our convergence proof does not assume that the time scale separation becomes large, in contrast to previous results [49, 31]. Rather, convergence is achieved for finite time scale separation, but in the limit of large healing time t_{skip} and an implicit approximation of the slow flow $\Phi_*(t; x)$: $P(t_{\text{skip}}; y) = P(t + t_{\text{skip}}; x)$ defines the approximation $\Phi_{t_{\text{skip}}}(t; x) := y$. The original explicit equation-free framework, as proposed by Kevrekidis *et al.*, corresponds to the case where $t_{\text{skip}} = 0$ and $\mathcal{R} \circ \mathcal{L} = I$. The analysis is performed for attracting slow manifolds in deterministic systems. However, we demonstrate on a simple SDE that our result may also be useful for stochastic systems, where the time scale separation is in the spectrum of the Fokker-Planck equation and is often only of order 1. In particular, for the prototype example investigated by [5] the implicit flow approximation $\Phi_{t_{\text{skip}}}$ converges to the true solution Φ_* of the linear Fokker-Planck equation for large healing times t_{skip}.

Acknowledgements. J. Sieber's research was supported by funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement number 643073, by the EPSRC Centre for Predictive Modelling in Healthcare (Grant Number EP/N014391/1) and by the EPSRC Fellowship EP/N023544/1.

C. Marschler and J. Starke would like to thank Civilingeniør Frederik Christiansens Almennyttige Fond for financial support. J. Starke would also like to thank the Villum Fonden (VKR-Centre of Excellence Ocean Life), the Technical University of Denmark and Queen Mary University of London for financial support.

REFERENCES

[1] L. Arnold, *Stochastic differential equations*, New York, (1974).

[2] L. Arnold, *Random dynamical systems*, Springer Science & Business Media, 2013.

[3] D. Avitabile, R. Hoyle, and G. Samaey, *Noise reduction in coarse bifurcation analysis of stochastic agent-based models: an example of consumer lock-in*, SIAM Journal on Applied Dynamical Systems, 13 (2014), pp. 1583–1619.

[4] D. Avitabile and K. Wedgwood, *Macroscopic coherent structures in a stochastic neural network: from interface dynamics to coarse-grained bifurcation analysis*, arXiv preprint arXiv:1603.04486, (2016).

[5] D. Barkley, I. G. Kevrikidis, and A. M. Stuart, *The moment map: nonlinear dynamics of density evolution via a few moments*, SIAM Journal on Applied Dynamical Systems, 5 (2006), pp. 403–434.

[6] P. W. Bates, K. Lu, and C. Zeng, *Persistence of overflowing manifolds for semiflow*, Comm. Pure Appl. Math., 52 (1999).

[7] P. W. Bates, K. Lu, and C. Zeng, *Invariant foliations near normally hyperbolic invariant manifolds for semiflows*, Trans. Amer. Math. Soc., 352 (2000), pp. 4641–4676.

[8] A. Ben-Tal and I. G. Kevrekidis, *Coarse-graining and simplification of the dynamics seen in bursting neurons*, SIAM Journal on Applied Dynamical Systems, 15 (2016), pp. 1193–1226.

[9] P. Boxler, *A stochastic version of center manifold theory*, Probability Theory and Related Fields, 83 (1989), pp. 509–545.

[10] Chebfun contributors, *Chebfun*, 2015, http://www.chebfun.org. Version 5.2.0 04-Jun-2015.

[11] X. Chen, A. J. Roberts, and J. Duan, *Centre manifolds for stochastic evolution equations*, Journal of Difference Equations and Applications, 21 (2015), pp. 606–632.

[12] R. R. Coifman, I. G. Kevrekidis, S. Lafon, M. Maggioni, and B. Nadler, *Diffusion maps, reduction coordinates, and low dimensional representation of stochastic systems*, Multiscale Modeling & Simulation, 7 (2008), pp. 842–864.

[13] K. Debrabant, G. Samaey, and P. Zielinski, *A micro-macro acceleration method for the monte carlo simulation of stochastic differential equations*, SIAM Journal on Numerical Analysis, 55 (2017), pp. 2745–2786.

[14] T. A. Driscoll, N. Hale, and L. N. Trefethen, *Chebfun guide*, Pafnuty Publications, Oxford, http://www.chebfun.org/, 2014.

[15] N. Fenichel, *Geometric singular perturbation theory for ordinary differential equations*, Journal of Differential Equations, 31 (1979), pp. 53–98.

[16] C. W. Gear, T. J. Kaper, I. G. Kevrekidis, and A. Zagaris, *Projecting to a slow manifold: Singularly perturbed systems and legacy codes*, SIAM Journal on Applied Dynamical Systems, 4 (2005), pp. 711–732, https://doi.org/10.1137/040608295.

[17] D. Givon, R. Kupferman, and A. Stuart, *Extracting macroscopic dynamics: model problems and algorithms*, Nonlinearity, 17 (2004), p. R55.

[18] T. Gross and I. G. Kevrekidis, *Robust oscillations in sis epidemics on adaptive networks: Coarse graining by automated moment closure*, EPL (Europhysics Letters), 82 (2008), p. 38004.

[19] H. Haken, *Advanced synergetics. Instability hierarchies of self-organizing systems and devices / Springer series in synergetics 20*, Springer, Berlin, 1983.

[20] M. Hirsch, C. Pugh, and M.Shub, *Invariant Manifolds*, vol. 583 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1977.

[21] http://www.mathworks.com, Matlab R2014b, *64-bit (glnxa64)*, http://www.mathworks.com.

[22] X. Kan, J. Duan, I. G. Kevrekidis, and A. J. Roberts, *Simulating stochastic inertial manifolds by a backward-forward approach*, SIAM Journal on Applied Dynamical Systems, 12 (2013), pp. 487–514.

[23] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos, *Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis*, Communications in Mathematical Sciences, 1 (2003), pp. 715 – 762, http://www.intlpress.com/CMS/journal/v1-4.html.

[24] Y. Kevrekidis and G. Samaey, *Equation-free multiscale computation: algorithms and applications*, Review of Physical Chemistry, 60 (2009), pp. 321–344.

[25] Y. Kevrekidis and G. Samaey, *Equation-free modeling*, Scholarpedia, 5 (2010), p. 4847.

[26] K. U. Kristiansen, M. Brøns, and J. Starke, *An iterative method for the approximation of fibers in slow-fast systems*, SIAM Journal on Applied Dynamical Systems, 13 (2014), pp. 861–900.

[27] R. Kupferman, A. M. Stuart, J. R. Terry, and P. F. Tupper, *Long-term behaviour of*

large mechanical systems with random initial data, Stochastics and Dynamics, 2 (2002), pp. 533–562, https://doi.org/10.1142/S0219493702000571.

[28] P. Liu, G. Samaey, C. W. Gear, and I. G. Kevrekidis, *On the acceleration of spatially distributed agent-based computations: A patch dynamics scheme*, Applied Numerical Mathematics, 92 (2015), pp. 54–69.

[29] A. G. Makeev, D. Maroudas, and I. G. Kevrekidis, *"Coarse" stability and bifurcation analysis using stochastic simulators: Kinetic Monte Carlo examples*, The Journal of Chemical Physics, 116 (2002), p. 10083, https://doi.org/10.1063/1.1476929.

[30] C. Marschler, C. Faust-Ellsässer, J. Starke, and J. van Hemmen, *Bifurcation of learning and structure formation in neuronal maps*, Europhysics Letters, 108 (2014).

[31] C. Marschler, J. Sieber, R. Berkemer, A. Kawamoto, and J. Starke, *Implicit methods for equation-free analysis: Convergence results and analysis of emergent waves in microscopic traffic models*, SIAM Journal on Applied Dynamical Systems, 13 (2014), pp. 1202–1238, https://doi.org/10.1137/130913961.

[32] C. Marschler, J. Sieber, P. Hjorth, and J. Starke, *Equation-free analysis of macroscopic behavior in traffic and pedestrian flow*, in Traffic and Granular Flow '13, M. Chraibi, M. Boltes, A. Schadschneider, and A. Seyfried, eds., Springer International Publishing, 2015, pp. 423–439, https://doi.org/10.1007/978-3-319-10629-8_48.

[33] C. Marschler, J. Starke, P. Liu, and I. G. Kevrekidis, *Coarse-grained particle model for pedestrian flow using diffusion maps*, Phys. Rev. E, 89 (2014), p. 013304, https://doi.org/10.1103/PhysRevE.89.013304.

[34] R. O'Malley, Jr., *Singular Perturbation Methods for Ordinary Differential Equations*, Springer-Verlag, 1991.

[35] A. Papavasiliou and I. G. Kevrekidis, *Variance reduction for the equation-free simulation of multiscale stochastic systems*, SIAM J. Multiscale Modeling & Simulation, 6 (2007), pp. 70–89.

[36] G. Pavliotis and A. M. Stuart, *Multiscale methods - Averaging and Homogenization*, Texts in Applied Mathematics, Springer, New York, 2008.

[37] G. A. Pavliotis and A. M. Stuart, *White noise limits for inertial particles in a random field*, Multiscale Modeling & Simulation, 1 (2003), pp. 527–553, https://doi.org/10.1137/S1540345903421076.

[38] C. Quinn, J. Sieber, and A. von der Heydt, *Effects of periodic forcing on a paleoclimate delay model*, arXiv preprint arXiv:1808.02310, (2018).

[39] A. Roberts, *Model emergent dynamics in complex systems*, vol. 20, SIAM, 2014.

[40] G. Samaey, I. G. Kevrekidis, and D. Roose, *Patch dynamics with buffers for homogenization problems*, Journal of Computational Physics, 213 (2006), pp. 264–287.

[41] G. Samaey, A. Roberts, and I. Kevrekidis, *Equation-free computation: an overview of patch dynamics*, Multiscale methods: bridging the scales in science and engineering, (2010), p. 216.

[42] C. Siettos, C. Gear, and I. Kevrekidis, *An equation-free approach to agent-based computation: bifurcation analysis and control of stationary states*, EPL (Europhysics Letters), 99 (2012), p. 48007.

[43] C. I. Siettos, D. Maroudas, and I. G. Kevrekidis, *Coarse Bifurcation Diagrams via Microscopic Simulators: a State-Feedback Control-Based Approach*, Int. J. of Bifurcation and Chaos, 14 (2004), pp. 207–220.

[44] H. E. Stanley, *Introduction to phase transitions and critical phenomena*, , 1 (1987), p. 336.

[45] S. Thomas, D. Lloyd, and A. Skeldon, *Equation-free analysis of agent-based models and systematic parameter determination: a Netlogo implementation*, preprint, (2016).

[46] C. Vandekerckhove, B. Sonday, A. Makeev, D. Roose, and I. G. Kevrekidis, *A common approach to the computation of coarse-scale steady states and to consistent initialization on a slow manifold*, Computers & Chemical Engineering, 35 (2011), pp. 1949 – 1958.

[47] W. Wang and A. Roberts, *Slow manifold and averaging for slow–fast stochastic differential system*, Journal of Mathematical Analysis and Applications, 398 (2013), pp. 822–839.

[48] A. Zagaris, C. W. Gear, T. J. Kaper, and Y. G. Kevrekidis, *Analysis of the accuracy and convergence of equation-free projection to a slow manifold*, ESAIM: Mathematical Modelling and Numerical Analysis, 43 (2009), pp. 757–784, https://doi.org/10.1051/m2an/2009026.

[49] A. Zagaris, C. Vandekerckhove, C. W. Gear, T. J. Kaper, and I. G. Kevrekidis, *Stability and stabilization of the constrained runs schemes for equation-free projection to a slow manifold*, Discrete and Continuous Dynamical Systems - Series A, 32 (2012), pp. 2759 – 2803.

Appendix A. Proof of Convergence Theorem 3.3.

For the proof of

Theorem 3.3 we have to analyze the two equations (for y and y_* respectively)

$$\mathcal{R}(M(t_{\text{skip}}; \mathcal{L}(y))) = \mathcal{R}(M(t_{\text{skip}} + \delta; \mathcal{L}(x))), \tag{A.1}$$

$$\mathcal{R}(M(t_{\text{skip}}; g(\mathcal{L}(y_*)))) = \mathcal{R}(M(t_{\text{skip}} + \delta; g(\mathcal{L}(x)))). \tag{A.2}$$

In both equations $x \in \mathbb{R}^d$ enters as a parameter. Assumption 3.2 ensures that the solution y_* of (A.2) is unique and independent of t_{skip}. For equation (A.1) we have to prove the existence of a solution y, and prove that it is close to y_* for sufficiently large t_{skip}. Throughout this appendix we will use the notations

$$u(t) = O(\exp(\alpha t)), \quad v(t) = o(\exp(\alpha t))$$

to describe that $\|u(t)\exp(-\alpha t)\|$ is bounded uniformly for all $t \geq 0$, and that the function $v(t)\exp(-\alpha t)$ tends to zero for $t \to \infty$. For the special case $\alpha = 0$ we write $O(1)$ and $o(1)$. If the quantity depends also on other parameters (say, $y \in \operatorname{dom}\mathcal{L}$) then the expression implies uniformity (for example, for y close to y_*) unless stated explicitly otherwise.

Using the definitions (3.7) of $P_*(t;x) = \mathcal{R}(M(t; g(\mathcal{L}(x))))$ and (3.10) for the map $P(t;x) = \mathcal{R}(M(t;\mathcal{L}(x)))$, equation (A.1) can be written in the form (using (A.2))

$$P_*(t_{\text{skip}}; y) = P_*(t_{\text{skip}}; y_*) + \exp(-d_{\text{tr}} t_{\text{skip}})\left[G(t_{\text{skip}}; y) + H(t_{\text{skip}}; x)\right], \quad \text{where} \tag{A.3}$$

$$\begin{aligned} G(t;y) &= -\exp(d_{\text{tr}} t)\left[P(t;y) - P_*(t;y)\right], \\ H(t;x) &= \exp(d_{\text{tr}} t)\left[P(t+\delta;x) - P_*(t+\delta;x)\right]. \end{aligned}$$

The operator P_* and the newly introduced G and H satisfy the following conditions on their derivatives by Assumption 3.1, (3.2) and (3.3) on separation of time scales for the flow M:

$$\partial_2^j P_*(t_{\text{skip}}; y) = O(\exp(d_{\text{tan}} t_{\text{skip}})), \tag{A.4}$$

$$\partial_2^j G(t_{\text{skip}}; y) = O(1), \tag{A.5}$$

$$\partial_2^j H(t_{\text{skip}}; x) = O(1) \tag{A.6}$$

for all $j \in \{0, \ldots, k_{\max}\}$ and all y in a neighborhood of y_*. In the case of H the bound is also uniform for $\delta \in [-\delta_{\max}, \delta_{\max}]$. Thus, the parameter δ has been dropped from the list of arguments in H. Combining the separation of time scales in Assumption 3.1, (3.2), with Assumption 3.2 on the uniform invertibility of $\mathcal{R}|_{\mathcal{C}}$ and $g \circ \mathcal{L} : \operatorname{dom}\mathcal{L} \mapsto \mathcal{C}$, we have a Lipschitz constant (C is independent of y_1, y_2 and t_{skip})

$$\|y_1 - y_2\| \leq C\exp(d_{\text{tan}} t_{\text{skip}}) \|P_*(t_{\text{skip}}; y_1) - P_*(t_{\text{skip}}; y_2)\|, \tag{A.7}$$

when inverting $P_*(t_{\text{skip}}; \cdot)$ for all y_1, y_2 in a neighborhood of y_* and all $t_{\text{skip}} \geq 0$. We also note that

$$\left\| \frac{\partial^j y_*(x)}{\partial x^j} \right\| = O(1) \tag{A.8}$$

Specifically, these derivatives depend only on $\delta \in [-\delta_{\min}, \delta_{\max}]$. Thus, $\partial^j y_*(x)$ are uniformly bounded due to (3.2), and because we required $\exp(d_{\text{tan}}\delta_{\max}) = O(1)$.

Abbreviating notation In the following all derivatives of the functions P_*, G and H are with respect to their second argument (y or x). The argument t_{skip} enters the

functions P_*, G and H as a parameter that we will drop in our notation such that we will write, for example, $\partial^3 P_*(y_*)[\partial y_*]^2[\partial^2 y_*]$ for $\partial_2^3 P_*(t_{\text{skip}}; y_*)[\partial y_*/\partial x]^2[\partial^2 y_*/(\partial x)^2]$. The parameter t_{skip} enters estimates via the bounds (A.4)–(A.8) for P_*, G and H.

The properties (A.4)–(A.8) make Banach's contraction mapping principle applicable to Equation (A.3) in a sufficiently small neighborhood of y_* and for sufficiently large t_{skip} (as shown in the paragraph that follows). We then estimate the error of the derivatives of y with respect to x.

Existence of solution y and its error. We apply the Banach contraction mapping principle to the map

$$N : y \mapsto P_*^{-1}\left(P_*(y_*) + \exp(-d_{\text{tr}} t_{\text{skip}})\,[G(y) + H(x)]\right) \tag{A.9}$$

($P_*^{-1}(\cdot)$ is the inverse of the diffeomorphism $P_* : U(y_*) \mapsto U(P_*(y_*))$). Let B be a closed ball around y_* of radius R in which all estimates (A.4)–(A.7) on P_*, G and H hold. Combining the estimate (A.7) for the Lipschitz constant of P_*^{-1} with $y_1 = y$ and $y_2 = y_*$, and the bound on the derivatives for G (w.r.t. y) gives an estimate for the difference of $N(y)$ from y_*:

$$\begin{aligned}\|N(y) - y_*\| &\le C\exp((d_{\tan} - d_{\text{tr}})t_{\text{skip}})\left[\max_B \|\partial G\|\|y\| + \|H(x)\|\right],\\ &\le C\exp((d_{\tan} - d_{\text{tr}})t_{\text{skip}})\left[\max_B \|\partial G\|(\|y_*\| + R) + \|H(x)\|\right].\end{aligned}$$

Thus, choosing t_{skip} sufficiently large, we can ensure that N maps B back into itself (since $d_{\tan} < d_{\text{tr}}$). Similarly, the Lipschitz constant of N in B can be estimated by

$$\|N(y_1) - N(y_2)\| \le C\exp((d_{\tan} - d_{\text{tr}})t_{\text{skip}}) \max_B \|\partial G\|\|y_1 - y_2\|,$$

where $C\exp((d_{\tan} - d_{\text{tr}})t_{\text{skip}}) \max_B \|\partial G\|$ is smaller than unity for sufficiently large t_{skip}. Consequently, N has a unique fixed point y in B, which solves the perturbed problem (A.1). Moreover, the difference $y - y_*$ satisfies

$$y - y_* = O(\exp((d_{\tan} - d_{\text{tr}})t_{\text{skip}})). \tag{A.10}$$

Error of derivatives. The smoothness of the coefficients in (A.3) ensures that y is also differentiable as a function of x up to order $k_{\max}$. We want to prove that for ℓ satisfying $\ell \le k_{\max} - 1$ (where $k_{\max}$ is the order of differentiability of the coefficients in (A.3)) and $(2\ell + 1)d_{\tan} < d_{\text{tr}}$ the bound on the error is

$$\partial^\ell y - \partial^\ell y_* = O(\exp(((2\ell + 1)d_{\tan} - d_{\text{tr}})t_{\text{skip}})). \tag{A.11}$$

We prove this by induction starting from $\ell = 1$, which we check first using the previous paragraph's results.

Assume that the bound (A.11) holds for all derivatives up to $\ell - 1$. This implies, in combination with (A.8), that y, ∂y, ..., $\partial^{\ell-1} y$ are bounded uniformly for all $t_{\text{skip}} \ge 0$ (just like $\partial^\ell y_*$ for $\ell = 1 \dots k_{\max}$ by (A.8)). In order to estimate the difference $\partial^\ell y - \partial^\ell y_*$, we return to (A.3) and differentiate each of the terms ℓ times with respect to x (noting that y_* and y are also functions of x):

$$\frac{\partial^\ell}{\partial x^\ell}[P_*(y(x))] - \frac{\partial^\ell}{\partial x^\ell}[P_*(y_*(x))] = \exp(-d_{\text{tr}} t_{\text{skip}})\left[\frac{\partial^\ell}{\partial x^\ell}[G(y(x))] + \partial^\ell H(x)\right]. \tag{A.12}$$

The term $\partial^\ell H(x)$ is $O(1)$ for all $t_{\text{skip}} \geq 0$ by (A.6). In the term $\partial^\ell/(\partial x^\ell)[G(y)]$ we extract the highest-order derivative of y by writing it in the form

$$\begin{aligned}\frac{\partial^\ell}{\partial x^\ell}[G(y)] &= O(1) + \partial G(y)\partial^\ell y = O(1) + \partial G(y)\partial^\ell y_* + \partial G(y)\left[\partial^\ell y - \partial^\ell y_*\right] \\ &= O(1) + \partial G(y)\left[\partial^\ell y - \partial^\ell y_*\right] \qquad \text{(A.13)}\end{aligned}$$

For (A.13) the boundedness of the $O(1)$ terms follows from the boundedness of all their parts: the derivatives of G are bounded by (A.5), $\partial^\ell y_*$ is bounded by (A.8), and y, ∂y,..., $\partial^{\ell-1} y$ are bounded by induction hypothesis. The pre-factor $\partial G(y)$ of $\partial^\ell y - \partial^\ell y_*$ is also bounded uniformly for all $t_{\text{skip}} \geq 0$.

Inserting the right-hand side of (A.13) into the right-hand side of (A.12), we obtain

$$\frac{\partial^\ell}{\partial x^\ell}[P_*(y(x))] - \frac{\partial^\ell}{\partial x^\ell}[P_*(y_*(x))] = \exp(-d_{\text{tr}} t_{\text{skip}})\partial G(y)\left[\partial^\ell y - \partial^\ell y_*\right] + O(\exp(-d_{\text{tr}} t_{\text{skip}})) \qquad \text{(A.14)}$$

Expanding the left-hand side of the above equation using the chain rule, we get a sequence of differences with equal powers of derivatives of P_*, y and y_*. From this sequence of differences we extract the difference between derivatives involving $\partial^\ell y$ and $\partial^\ell y_*$ and collect all other terms in a remainder r (which is present only for $\ell > 1$ and will later turn out to be of order $O(\exp((2\ell d_{\tan} - d_{\text{tr}})t_{\text{skip}}))$):

$$\frac{\partial^\ell}{\partial x^\ell}[P_*(y(x))] - \frac{\partial^\ell}{\partial x^\ell}[P_*(y_*(x))] = \partial P_*(y)\partial^\ell y - \partial P_*(y_*)\partial^\ell y_* + r. \qquad \text{(A.15)}$$

From the difference with the highest-order derivatives of y and y_* we extract the difference $\partial^\ell y - \partial^\ell y_*$ by adding zeroes. Using the notational convention

$$F\{x, y\} = \int_0^1 F(sx + (1-s)y)\,\mathrm{d}s$$

for the mean between two points of a single-argument function F in the following,

$$\begin{aligned}\partial P_*(y)\partial^\ell y - \partial P_*(y_*)\partial^\ell y_* &= \partial P_*(y)\left[\partial^\ell y - \partial^\ell y_*\right] + \left[\partial P_*(y) - \partial P_*(y_*)\right]\partial^\ell y_* \\ &= \partial P_*(y)\left[\partial^\ell y - \partial^\ell y_*\right] + \partial^2 P_*\{y, y_*\}[y - y_*]\,\partial^\ell y_* \qquad \text{(A.16)} \\ &= \partial P_*(y)\left[\partial^\ell y - \partial^\ell y_*\right] + O(\exp((2d_{\tan} - d_{\text{tr}})t_{\text{skip}})). \qquad \text{(A.17)}\end{aligned}$$

The order $O(\exp((2d_{\tan} - d_{\text{tr}})t_{\text{skip}}))$ of the second term follows from the bounds on $y - y_*$ (given in (A.10)), $\partial^2 P_*$ (given in (A.4)) and the boundedness of $\partial^\ell y_*$ (given in (A.8)). This immediately implies the estimate for the case $\ell = 1$: inserting (A.17) into (A.14), we have for $\ell = 1$

$$\partial P_*(y)\left[\partial y - \partial y_*\right] = \exp(-d_{\text{tr}} t_{\text{skip}})\partial G(y)\left[\partial y - \partial y_*\right] + O(\exp((2d_{\tan} - d_{\text{tr}})t_{\text{skip}})). \qquad \text{(A.18)}$$

In (A.18) we have collected the bounded terms with pre-factors $\exp(-d_{\text{tr}} t_{\text{skip}})$ and $\exp((2d_{\tan} - d_{\text{tr}})t_{\text{skip}})$ using the larger pre-factor $\exp((2d_{\tan} - d_{\text{tr}})t_{\text{skip}})$. Since (by (A.7)) the inverse of $\partial P_*(y)$ satisfies $\partial P_*(y)^{-1} = O(\exp(d_{\tan} t_{\text{skip}}))$ we can rearrange (A.18) to isolate $\partial y - \partial y_*$ for large t_{skip}, giving the estimate (note that $\partial G(y) = O(1)$)

$$\partial y - \partial y_* = O(\exp((3d_{\tan} - d_{\text{tr}})t_{\text{skip}})), \qquad \text{(A.19)}$$

which is what we had to prove for $\ell = 1$.

Error of higher-order derivatives. Let us assume that the assumptions of the theorem are satisfied for all $j < \ell$ with $\ell \geq 2$. By the conditions of the theorem we assume that $(2\ell+1)d_{\tan} < d_{\mathrm{tr}}$ and the conditions (A.4)–(A.8) are satisfied for $j \leq \ell$ (including existence of the corresponding derivatives).

For $\ell > 1$ we have to include the remainder r from (A.15) in our consideration. This remainder is a sum of expressions a_ν of the form

$$a_\nu = \partial^j P_*(y)\,[\partial^{\nu_1} y]\ldots[\partial^{\nu_j} y] - \partial^j P_*(y_*)\,[\partial^{\nu_1} y_*]\ldots[\partial^{\nu_j} y_*]\,, \tag{A.20}$$

where $2 \leq j \leq \ell$, and ν is a j-tuple of integers $\nu_i \in \{1,\ldots,\ell-1\}$ with $\sum_{i=1}^{j}\nu_i = \ell$. All factors $\partial^{\nu_i} y$ and $\partial^{\nu_i} y_*$ are of order $O(1)$ with respect to t_{skip} according to (A.8) and induction hypothesis. The terms $\partial^j P_*(y)$ and $\partial^j P_*(y_*)$ are of order $O(\exp(d_{\tan}t_{\mathrm{skip}}))$ according to (A.4). The difference in (A.20) can be expressed as a sum of $j+1$ differences involving $\partial^i y - \partial^i y_*$ for some $i \in \{0\ldots,\ell-1\}$ by adding $j+1$ zeros:

$$a_\nu = \partial^{j+1} P_*\{y, y_*\}[y-y_*]\,[\partial^{\nu_1} y]\ldots[\partial^{\nu_j} y] \tag{A.21}$$

$$+\sum_{i=1}^{j}\partial^j P_*(y_*)\left[\prod_{m<i}\partial^{\nu_m} y_*\right][\partial^{\nu_i} y - \partial^{\nu_i} y_*]\left[\prod_{m>i}\partial^{\nu_m} y\right] \tag{A.22}$$

The right-hand side in (A.21) is of order $O(\exp((2d_{\tan}-d_{\mathrm{tr}})t_{\mathrm{skip}}))$. The ith term in the sum in (A.22) is of order $O(\exp((d_{\tan}(1+(2\nu_i+1))-d_{\mathrm{tr}})t_{\mathrm{skip}}))$. So, since $\nu_i \leq \ell-1$ and $\ell > 1$, all terms in the sum for a_ν are at most of order $O(\exp((2\ell d_{\tan}-d_{\mathrm{tr}})t_{\mathrm{skip}}))$. Consequently,

$$r = O(\exp((2\ell d_{\tan} - d_{\mathrm{tr}})t_{\mathrm{skip}})). \tag{A.23}$$

Inserting this estimate in combination with (A.15) and (A.17) into (A.14), we obtain

$$\partial P_*(y)\left[\partial^\ell y - \partial^\ell y_*\right] = \exp(-d_{\mathrm{tr}}t_{\mathrm{skip}})\partial G(y)\left[\partial^\ell y - \partial^\ell y_*\right] + O(\exp((2\ell d_{\tan}-d_{\mathrm{tr}})t_{\mathrm{skip}})). \tag{A.24}$$

In (A.24) we have included the smaller error terms $O(\exp((2d_{\tan}-d_{\mathrm{tr}})t_{\mathrm{skip}}))$ and $O(\exp(-d_{\mathrm{tr}}t_{\mathrm{skip}}))$ into the (for $\ell > 1$) larger $O(\exp((2\ell d_{\tan}-d_{\mathrm{tr}})t_{\mathrm{skip}}))$. Since, $d_{\mathrm{tr}} < d_{\tan}$, $\partial G(y) = O(1)$ and $\partial P_*(y) = O(\exp(d_{\tan}t_{\mathrm{skip}}))$, we can isolate $\partial^\ell y - \partial^\ell y_*$ in (A.24). This results in the asymptotic estimate claimed in Theorem 3.3:

$$\partial^\ell y - \partial^\ell y_* = O(\exp(((2\ell+1)d_{\tan}-d_{\mathrm{tr}})t_{\mathrm{skip}})).$$

Appendix B. Brief description of supplementary material. The supplementary material contains `Matlab/octave` scripts and functions that reproduce Figure 4.1 from section 4. The provided `zip` file unpacks into folder `demo_Michaelis_Menten/`. The main script is `demo_Michaelis_Menten.m`, which will reproduce Figure 4.1, showing phase space geometry of the Michaelis-Menten kinetics (4.1) with explicit time scale separation as also studied by Gear *et al.* and others [34, 16, 48, 49].

- Folder `demo_Michaelis_Menten/rotated/` contains the published `html` output from the script for the rotated coordinate system (4.7) in file `demo_Michaelis_Menten.html`.
- Folder `demo_Michaelis_Menten/unrotated/` contains the published `html` output from the script for the coordinate system with explicit time scale separation (4.1) in file `demo_Michaelis_Menten.html`.
- Folder `tools/` contains some auxiliary functions called in the script (a simple Newton iteration `ScSolve.m`, an explicit initial-value-problem solver using the Dormand-Prince scheme and fixed step size `ScIVP.m`, and a function for approximating the Jacobian with finite differences `ScJacobian.m`.