

PÉTER KOLTAI¹ AND CHRISTOF SCHÜTTE^{1,2}

¹*Department of Mathematics and Computer Science, Freie Universität Berlin, Germany*

²*Zuse Institute Berlin, Germany*

A MULTI SCALE PERTURBATION
EXPANSION APPROACH FOR MARKOV
STATE MODELING OF
NON-STATIONARY MOLECULAR
DYNAMICS

Zuse Institute Berlin
Takustr. 7
14195 Berlin
Germany

Telephone: +49 30-84185-0
Telefax: +49 30-84185-125

E-mail: bibliothek@zib.de
URL: <http://www.zib.de>

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

A multi scale perturbation expansion approach for Markov state modeling of non-stationary molecular dynamics

Péter Koltai¹ and Christof Schütte^{1,2}

¹Department of Mathematics and Computer Science, Freie Universität Berlin, Germany

²Zuse Institute Berlin, Germany

September 6, 2017

Abstract

We investigate metastable dynamical systems subject to non-stationary forcing as they appear in molecular dynamics for systems driven by external fields. We show, that if the strength of the forcing is inversely proportional to the length of the slow metastable time scales of the unforced system, then the effective behavior of the forced system on slow time scales can be described by a low-dimensional reduced master equation. Our construction is explicit and uses the multiscale perturbation expansion method called two-timing, or method of multiple scales. The reduced master equation—a Markov state model—can be assembled by constructing two equilibrium Markov state models; one for the unforced system, and one for a slightly perturbed one.

Keywords: Markov state model, non-equilibrium molecular dynamics, two timescale master equation, non-stationary forcing, metastability

AMS-Classification: 34E13, 60J20, 60J60

1. Introduction

This article aims at studying metastable dynamical systems with non-stationary external forcing. The goal is to understand the effect of a weak forcing on the transition probabilities between the metastable sets of the dynamical system. This question is of high interest for understanding the dynamical behavior of biomolecular systems under non-equilibrium conditions caused by external fields. Such systems have attracted a considerable attention recently, for example, the potential effects of electromagnetic radiation on the human body tissue, in particular on the molecular and cellular level

(e.g. on DNA, proteins), has been extensively investigated in a vast number of articles, e.g. [dPSD⁺03, BD14, AGT12].

Molecular dynamics (MD) simulations have proved particularly useful for understanding the response of biomolecular conformations to external fields because of their ability to resolve molecular details that cannot be resolved in experiments and allows for studying non-equilibrium processes in a statistically reliable and thermodynamically consistent way [WSCDS14]. Despite their many advantages, MD simulations have severe limitations. One of the main problems is that the maximal possible simulation length often is shorter than the timescale of interest.

This article is mainly concerned with circumventing the latter obstacle by introducing non-equilibrium Markov State Models for MD with weak external forcing. Markov State Models (MSMs) have been well developed over the past decade in theory and applications, see [SS14, BPN13], but for systems under equilibrium conditions or in non-equilibrium steady states only! The principal idea of equilibrium MSMs is to approximate the original high-dimensional MD system by a reduced Markovian dynamics whose (discrete) states represent the dominant metastable sets of the MD system which are identical to the main conformations of the molecular system under consideration [SS14]. It has been shown that for molecular systems exhibiting such metastable sets the dynamics given by a MSM allows very close approximation of the longest relaxation processes of the underlying molecular system, at least under equilibrium conditions [SNS10]. Moreover, it has been demonstrated that in such cases MSM building requires short equilibrium MD trajectories only, much shorter than the timescales of interest [NSVE⁺09, BPN13]. Thus, standard MSM building exhibits two main advantages: It allows to study the dynamical behavior on long timescales (i) *without* requiring MD trajectories of comparable length, and (ii) the MSM has to be constructed *only once to study all times*, at least under equilibrium/steady state conditions.

MSM building for MD systems with external forcing have already been discussed, in particular for strictly periodic forcing [WS15] and for the case of strong forcing in which the metastable sets of the system are moved in state space and become coherent sets [KCS16]. The first case can be mapped into the equilibrium/steady case by introducing Floquet states, but is of limited interest for applications since most external fields show slowly varying amplitudes of the oscillatory field strength. In the latter case the resulting MSM has to be repeatedly re-calculated during the system’s evolution, thus destroying the advantage (ii) of MSMs. Furthermore, in many applications, a forcing so strong that it allows to move the metastable sets in state space would be unreasonable since the associated influx of energy would destroy the molecular structure itself. In contrast, most interesting applications belong to cases of highly oscillatory external fields with slowly varying small amplitudes [AGT12]. This is the case we consider in this article. We will demonstrate it allows for a version of MSM building that inherits the two main advantages (i) and (ii) from MSM building in equilibrium.

More precisely, we will show that a certain class of forced systems—where the forcing influences the system’s conformation dynamics in a non-trivial way—can be modeled by a MSM that can be constructed *once to study the forced system for all times*. The construction is based on two equilibrium Markov state models; one for the un-

forced equilibrium system, and one for a slightly perturbed one. To this end we use the multiscale perturbation expansion method called two-timing, or method of multiple scales [Lic69, O'M70], [Kue15, Section 9.8]. In order to set this up we choose a specific scaling of the external forcing that explicitly depends on the dominant metastable timescales of the underlying unforced system. This seems reasonable since this article aims at showing that there exist cases in which a weak and slow forcing induces non-trivial and significant changes in the conformation dynamics of the molecular system under consideration, instead of aiming at the most general analysis. In order to specify this setting we first discuss details of the unforced dynamical system and the scaling of the forcing in Section 2, before we review equilibrium MSM building in Section 3. This prepares the ground for approaching MSM building for the forced system in Section 4. Algorithms for constructing the resulting non-equilibrium MSMs are discussed in Section 5; here the focus is on showing that such algorithms really require the construction of two equilibrium Markov state models only, and *not* on algorithmic efficiency. In Section 6 we give a detailed analysis of a simple test case: a diffusion in a one-dimensional potential landscape. The article ends with concluding remarks about limitations of our approach and future work.

2. Dynamics

We consider diffusive dynamics in a smooth, bounded from below energy landscape V ,

$$d\mathbf{X}_t = (b(\mathbf{X}_t, t) - \nabla_x V(\mathbf{X}_t))dt + \sqrt{2\beta^{-1}}d\mathbf{B}_t, \quad (1)$$

where $\mathbf{X}_t \in \mathbb{R}^d$ is a time-dependent random variable¹ describing the state, \mathbf{B}_t denotes a standard Brownian motion in d dimensions, and $\beta > 0$ the parameter of the inverse temperature. The time-dependent drift term b represents external forces on the system under consideration.

The associated Fokker–Planck equation for probability density functions $u = u(x, t)$ has the form

$$\frac{\partial}{\partial t}u = \mathcal{A}u, \quad (2)$$

where \mathcal{A} denotes the generator given by

$$\mathcal{A}(b)u(x, t) = \beta^{-1}\Delta_x u(x, t) + \nabla_x \cdot ((\nabla_x V(x, t) - b(x, t))u(x, t)), \quad (3)$$

where the argument b indicates that \mathcal{A} depends on the external forces and, via them, explicitly on time t .

As pointed out above, we are interested in perturbations that influence the system's conformation dynamics in equilibrium but does not reshape the energy landscape completely. The literature on MSMs indicates that the conformation dynamics in equilibrium is determined by the dominant spectrum of the generator $\mathcal{A}(0)$ for $b = 0$, or, respectively, by the associated transfer operator $T_t = \exp(t\mathcal{A}(0))$: The longest time scales of the conformation dynamics are encoded by the low-lying spectrum of $\mathcal{A}(0)$ (which corresponds

¹Bold face capital letters denote random variables.

to the dominant eigenvalues of T_t); since $\mathcal{A}(0)$ is the generator of a reversible process, its eigenvalues λ all are real-valued and non-positive, $\lambda \leq 0$, where the largest eigenvalue $\lambda = 0$ is simple if the process is ergodic [SS14]. In the following we assume that the *spectral gap condition* is satisfied: there are m low-lying eigenvalues of $\mathcal{A}(0)$, that are of the order of some small parameter ε , while the remaining eigenvalues are significantly larger in modulus (order 1). A precise statement is given in Section 3 below. In order to stress this assumption notation-wise, we use the small parameter ε as an index to the generator, that is, we denote the generator by $\mathcal{A}_\varepsilon(b)$ in the following. The m low-lying eigenvalues induce m slow time scales of order ε^{-1} in the solution of (2) that are the dominant relaxation time scales of the conformation dynamics in equilibrium.

What kind of relation between the external force b and the internal conformation dynamics in equilibrium do we consider? On one hand, we do not allow for the driving force to shift the main wells of the energy landscape around. That is, we restrict ourselves to the consideration of weak external forces. On the other hand, as we will see later, the interaction of the external forces b with this internal conformation dynamics will only be “non-trivial” if they also depend on the slow time scale ε^{-1} and not only on faster time scales (otherwise the external driving force and the internal conformation dynamics would be adiabatically decoupled). These two aspects lead to the following form of the external forcing that we will consider in the following:

$$b(x, t) = \varepsilon f(t, \varepsilon t) \tilde{b}(x),$$

where f denotes a time-dependent pre-factor that may depend on fast and slow time scales, and \tilde{b} depends a force that depends on the state of the system but no longer explicitly on time. Choosing the strength of the forcing to be $\mathcal{O}(\varepsilon)$ ensures that the total energy injected into the system on the slow time scales of length $\mathcal{O}(\varepsilon^{-1})$ is at most $\mathcal{O}(1)$, thus the forcing does not overwhelm the original potential, as desired. Of course, other magnitudes of the forcing strength might be possible here too, but already this choice will lead to interesting and non-trivial behavior.

Solving the Fokker–Planck equation exactly is infeasible in general. Therefore, we want to approximate its solution by spatial discretization, e.g., based on a finite element or finite volume discretization of the generator \mathcal{A}_ε . By choosing an arbitrarily fine discretization one can in principle reduce the discretization error as much as wanted, i.e., a very fine discretization will yield an arbitrarily accurate solution of (2). This might not be practically realizable in some example cases (e.g. due to the dimension d of the state space being too large), but theoretically possible. Also, this allows us to work with matrices instead of operators, and hence the presentation will not be obscured by functional-analytic technicalities, as generators of the form (3) are unbounded operators.

When discretizing the generator \mathcal{A}_ε appropriately, the transpose $\mathcal{L}_\varepsilon \in \mathbb{R}^{n \times n}$ of the resulting discretization matrix has the form of a *generator or rate matrix* whose dimension n is given by the number of discretization elements used. A matrix $L \in \mathbb{R}^{n \times n}$ is

called rate matrix if it satisfies the following properties:

$$L_{ij} \geq 0 \quad \text{for } i \neq j, \quad (4a)$$

$$L_{ii} \leq 0 \quad \text{for all } i, \quad (4b)$$

$$\sum_{j=1}^n L_{ij} = 0 \quad \text{for all } i. \quad (4c)$$

That is, by such discretization we replace the Fokker–Planck equation (2) by the master equation,

$$\frac{d}{dt}u(t) = \mathcal{L}_\varepsilon^T u(t), \quad (5)$$

on a finite dimensional state space, i.e., $u(t) \in \mathbb{R}^n$, where n typically is very large. With our above assumptions, the rate matrix $\mathcal{L}_\varepsilon \in \mathbb{R}^{n \times n}$ will have the form²

$$\mathcal{L}_\varepsilon(t) = L_\varepsilon + \varepsilon f(t, \varepsilon t)L,$$

where L_ε denotes the rate matrix that results from discretization of $\mathcal{A}_\varepsilon(0)$ alone, and L the remaining part of $\mathcal{A}_\varepsilon(b)$ originating from the external force \tilde{b} . We introduce time t as an argument of \mathcal{L}_ε in order to mark that fact that \mathcal{L}_ε is explicitly depending on t in contrast to L_ε .

3. Stationary equilibrium MSM

3.1. Key assumptions and properties

If external forces can be ignored, i.e., if $b \equiv 0$, then, under weak growth condition³ on V , the process is ergodic and has the unique invariant measure $\mu(x) \propto \exp(-\beta V(x))$. In addition it is reversible, i.e., it satisfies the detailed balance condition. This is the case of molecular dynamics in equilibrium. In this case $\mathcal{L}_\varepsilon = L_\varepsilon$ is a self-adjoint matrix due to the reversibility of the underlying process.

We will consider this situation first and show how to construct a Markov State Model for $\mathcal{L}_\varepsilon = L_\varepsilon$ that allows to replace (5) by a low-dimensional coarse grained master equation. Subsequently, we will return to the general case with non-zero b , where \mathcal{L}_ε depends explicitly on time t and where we no longer can use properties like ergodicity, existence of an invariant measure or reversibility.

In the equilibrium situation, $\mathcal{L}_\varepsilon = L_\varepsilon$ inherits the form of the spectrum of \mathcal{A}_ε . That is, we assume that L_ε exhibits the following properties:

²Note that if L is a rate matrix, $-L$ is not, but we would like fL to be a rate matrix independent of the sign of f . This can be achieved, as done also earlier in [KV14], by having a discretization L_+ and L_- for \tilde{b} and $-\tilde{b}$, respectively, and using $|f|L_{\text{sign}(f)}$ as discretization of the corresponding term in $\mathcal{A}(b)$. Here, $\text{sign}(f) \in \{+, -\}$. We will refrain from pointing this out again, and use the simpler notation, fL , keeping this comment in mind. If ε is sufficiently small, \mathcal{L}_ε will be a rate matrix, even if fL is not.

³E.g., $V(x) \geq c\|x\|^2$ for some $c > 0$ and sufficiently large $\|x\|$ implies the desired properties [MS02, MSH02]. Here, $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^d .

(A1) The generator L_ε is irreducible and aperiodic with unique invariant measure μ_ε , furthermore, due to reversibility the detailed balance condition holds, i.e., $\mu_{\varepsilon,i} L_{\varepsilon,ij} = \mu_{\varepsilon,j} L_{\varepsilon,ji}$ for all $i, j = 1, \dots, n$.

(A2) For small enough ε the generator L_ε exhibits m eigenvalues

$$0 = \lambda_{\varepsilon,1} > \lambda_{\varepsilon,2} = \varepsilon \hat{\lambda}_2 \geq \dots \geq \lambda_{\varepsilon,m} = \varepsilon \hat{\lambda}_m, \quad (6)$$

where the identities are up to order ε^2 and the $\hat{\lambda}_j < 0$, $j = 2, \dots, m$, denote ε -independent constants. In contrast, the other eigenvalues satisfy

$$\lambda_{\varepsilon,i} < -r, \quad i = m+1, \dots, n, \quad (7)$$

with some ε -independent $r > 0$.

(A3) The dependence on ε is smooth and for $\varepsilon \rightarrow 0$ the family L_ε converges to a generator L_0 (entry-wise).⁴

While (A1) directly results from the assumptions on V and a structure-preserving discretization, e.g. [LMHS11], and (A2) from the spectral gap condition, (A3) allows us to study the asymptotic behavior of L_ε for $\varepsilon \rightarrow 0$ which will show instrumental for deriving the coarse grained master equation for the equilibrium/stationary case as well as for the explicitly time-dependent case.

Note that ε does not need to be a parameter with physical meaning. It is tempting but misleading to think of ε being proportional to the temperature β^{-1} . In that case, every local minimum of the potential V would become arbitrarily metastable as the temperature decreases to zero, and potentially many more eigenvalues than just the first m would accumulate at zero. Instead, L_0 is an artificial rate matrix not necessarily belonging to a diffusion of form (1), which decouples exactly those metastable sets which we want to model for a *given* ε .

Example 1. *Let us consider the double well potential*

$$V(x) = (x^2 - 1)^2,$$

for which (1) exhibits two metastable sets, the neighborhoods of $x = \pm 1$, if $b \equiv 0$. Although, as mentioned above, we do not view the inverse temperature as the small parameter ε , in general, here it allows for a simple construction by doing so. We consider the dominant eigenvalues $\lambda_{\beta,i}$, $i = 1, 2, \dots$, of $\mathcal{A}(0)$ for $\beta = 10^\xi$, $\xi = -1, -0.9, \dots, 1.8, 1.9, 2$, by discretizing $\mathcal{A}(0)$ on a very fine grid (2000 grid points on $[-5, 5]$ with Dirichlet zero boundary conditions). While $\lambda_{\beta,1} = 0$ (numerically) and $\lambda_{\beta,3} < -3$ for every β , the second eigenvalue decreases in magnitude exponentially in β , cf. Figure 1 (right). As long as the magnitude of $\lambda_{\beta,2}$ is larger than machine precision, we observe exponential decay, i.e., $|\lambda_{\beta,2}| \propto \omega^\beta$ for some $0 < \omega < 1$. Thus, here we can set $\varepsilon = \omega^\beta$.

⁴In the undiscretized, operator-based case such a strong assumption would hardly hold, since \mathcal{A}_ε is already an unbounded operator. In that case we would have to reformulate (A3) and turn to operator theoretic methods [PS08].

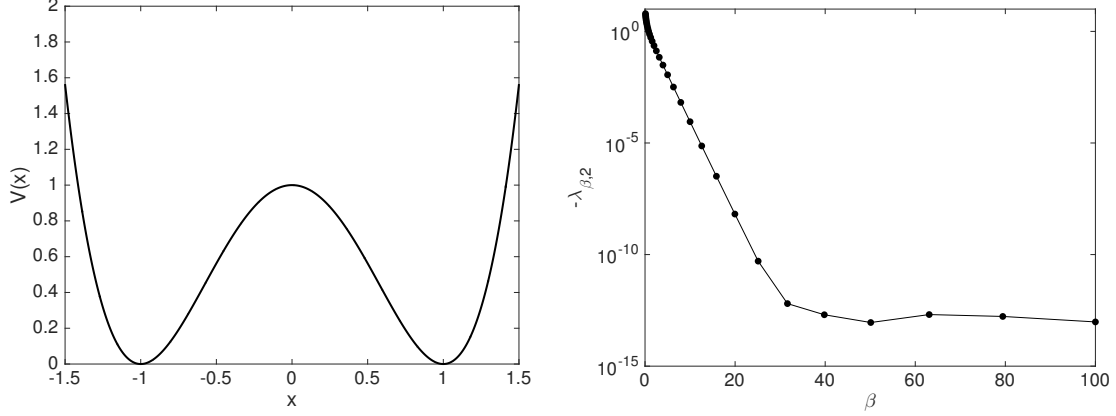


Figure 1: Left: the double well potential. Right: the dependence of the second eigenvalue, $\lambda_{\beta,2}$ on the inverse temperature β . Before reaching the order of magnitude of machine precision, $|\lambda_{\beta,2}| \propto \omega^\beta$ for some $0 < \omega < 1$.

Under above assumptions on L_ε , the matrix L_0 has an m -fold eigenvalue $\lambda = 0$. Since L_0 is a rate matrix this means that the state space $\{1, \dots, n\}$ decomposes into m (disjoint) invariant sets $\mathbb{A}_1, \dots, \mathbb{A}_m$ such that by (4c) we have

$$L_0 \mathbf{1}_j = 0, \quad j = 1, \dots, m.$$

Here, $\mathbf{1}_j \in \mathbb{R}^n$ denotes the indicator function (vector) of \mathbb{A}_j , i.e., $\mathbf{1}_j(i) = 1$ for $i \in \mathbb{A}_j$ and $= 0$ otherwise. L_0 is self-adjoint, irreducible and aperiodic if restricted to any of the \mathbb{A}_j . Thus, the m -dimensional (right-hand) eigenspace \mathbb{E}_r of L_0 for the eigenvalue $\lambda = 0$ is spanned by the indicator functions $\mathbf{1}_1, \dots, \mathbf{1}_m$. Moreover, for every $j = 1, \dots, m$, L_0 has a unique invariant measure μ_j with $\mu_j^T \mathbf{1}_j = 1$ that is supported on \mathbb{A}_j , that is, μ_j is 0 outside of \mathbb{A}_j . Then, every measure of the form $\mu_0 = \sum_{j=1}^m \alpha_j \mu_j$ with non-negative coefficients α_j is an invariant measure of L_0 . Thus, the (left-hand) eigenspace $\mathbb{E}_l \subset \mathbb{R}^n$ of L_0 is spanned by the μ_j .

The projections Π_l and Π_r on the eigenspaces \mathbb{E}_l and \mathbb{E}_r , respectively, are thus given by

$$\Pi_l = \sum_{j=1}^m \mu_j \mathbf{1}_j^T, \quad \Pi_r = \sum_{j=1}^m \mathbf{1}_j \mu_j^T = \Pi_l^T.$$

Since L_0 is reversible with respect to μ_0 (for every feasible coefficient vector $\alpha = (\alpha_1, \dots, \alpha_m)^T$), the (right) eigenvectors of L_0 span the whole space \mathbb{R}^n , and they are mutually orthogonal with respect to the μ_0 -weighted scalar product $\langle \cdot, \cdot \rangle_{\mu_0}$, where $\langle u, v \rangle_{\mu_0} = \sum_{i=1}^n u(i)v(i)\mu_0(i)$. For later use, we also introduce the orthogonal complement \mathbb{E}_l^\perp of \mathbb{E}_l that consists of all $v \in \mathbb{R}^n$ that satisfy $\Pi_l u = 0$. Also, let us collect following properties of the projections, which we will need later. They follow from straightforward calculations.

Lemma 1. *We have*

a) $\Pi_l L_0^T = L_0^T \Pi_l = 0$; and

b) $\exp(tL_0^T)\Pi_l = \Pi_l \exp(tL_0^T) = \Pi_l$ for all $t \geq 0$.

The same expressions hold if we replace both Π_l by Π_r and L_0^T by L_0 .

With these preparations we return to the master equation in equilibrium,

$$\frac{d}{dt}u_\varepsilon(t) = L_\varepsilon^T u_\varepsilon(t), \quad u_\varepsilon(0) = \rho. \quad (8)$$

3.2. Perturbation expansion and secular terms

By assumption (A3) we have that $L_\varepsilon = L_0 + \varepsilon G + \mathcal{O}(\varepsilon^2)$ for some G , and let us assume that the solution of (8) admits the *perturbation expansion*

$$u_\varepsilon(t) = u_0(t) + \varepsilon u_1(t) + \mathcal{O}(\varepsilon^2)$$

in ε , with coefficient functions u_0, u_1 bounded uniformly in t . Substituting these expansions into (8) and collecting the terms with the same order of ε gives

$$\begin{aligned} \varepsilon^0 : \quad & \frac{d}{dt}u_0(t) = L_0^T u_0(t), \quad u_0(0) = \rho, \\ \varepsilon^1 : \quad & \frac{d}{dt}u_1(t) = L_0^T u_1(t) + G^T u_0(t), \quad u_1(0) = 0. \end{aligned}$$

Solving the first, then the second equation yields⁵

$$u_\varepsilon(t) = \exp(tL_0^T)\rho + \varepsilon \int_0^t \exp((t-s)L_0^T) G^T \exp(sL_0^T)\rho ds + \mathcal{O}(\varepsilon^2). \quad (9)$$

This result has two major shortcomings. First, the best a priori bound on u_1 , the integral term in (9), is $\mathcal{O}(t)$, since L_0 being a rate matrix gives that $\exp(tL_0^T)$ is bounded uniformly in t . That means, the expansion is only valid for times $t = \mathcal{O}(1)$ as $\varepsilon \rightarrow 0$, since from this time on we cannot guarantee the validity of our expansion, because the term εu_1 , assumed to be of size $\mathcal{O}(\varepsilon)$ during the derivation, may grow beyond this bound. Such unbounded terms in perturbation expansions are called *secular* [Kue15, Section 9.6], and we shall try to avoid them.

Second, expression (9) does not tell anything about the behavior of the system on *slow time scales*. The eigenvalues of L_ε of magnitude $\mathcal{O}(\varepsilon)$ tell us that some components of the solution of the master equation (8) decay on time scales $\mathcal{O}(\varepsilon^{-1})$. However, the term $\exp(tL_0^T)\rho$ of (9) lacks exactly these components by the construction of L_0 ; it describes only evolution on the *fast time scale*. The integral term, hence containing the slow processes, stays uninformative in this respect.

We could try to obtain a perturbation expansion directly on the slow time scale $\tau := \varepsilon t$. For this, let $v_\varepsilon(\tau) := u_\varepsilon(\tau/\varepsilon)$. Carrying out the same procedure as above, we obtain, to our dismay, that

$$\frac{d}{d\tau}(v_0(\tau) + \varepsilon v_1(\tau)) = \varepsilon^{-1} L_0^T v_0(\tau) + \dots,$$

⁵The solution of $\frac{d}{dt}u(t) = Au(t) + g(t)$ is known to be $u(t) = \exp(tA)u(0) + \int_0^t \exp((t-s)A)g(s)ds$.

i.e., we are unable to match the first term on the right-hand side, because there is no other term with the same order of ε . This makes perfect sense, since u_ε has fast components in it, which cannot be modeled on the slow time scale otherwise. Thus, there is no *regular perturbation expansion* [Kue15, Def. 5.1.4] of our solution on the slow time scale. Such problems are called *singular perturbation problems*. The literature on singularly perturbed dynamical systems gives us a partial answer. According to [CR79], e.g., on the slow time scale ε^{-1} we have

$$\lim_{\varepsilon \rightarrow 0} u_\varepsilon(t/\varepsilon) = \exp(t\Pi_l G^T \Pi_l) \Pi_l \rho, \quad (10)$$

see Appendix A. However, such kind of limit results do not allow to understand the dynamics on faster scales outside of \mathbb{E}_l .

3.3. Coarse grained master equation

The following statement is a simple consequence of Theorem 2, which we will derive later in Section 4. Note that it splits the evolution of the full master equation into two parts: one on the slow and on the fast time scales.

Theorem 1. *Under the assumptions (A1)-(A3) the n -dimensional master equation (8) has a solution u_ε of the form*

$$u_\varepsilon(t) = \hat{u}(\varepsilon t) + \exp(tL_0^T) \delta(\varepsilon t) + \mathcal{O}(\varepsilon), \quad (11)$$

valid for times $t = \mathcal{O}(\varepsilon^{-1})$, where $\hat{u} \in \mathbb{E}_l$ is given by the coarse grained master equation of essential dimension $m \ll n$,

$$\frac{d}{dt} \hat{u}(t) = \Pi_l G^T \Pi_l \hat{u}(t), \quad \hat{u}(0) = \Pi_l \rho, \quad (12)$$

with $\delta(\varepsilon t) \in \mathbb{E}_l^\perp$ and an $n \times n$ matrix

$$G = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (L_\varepsilon - L_0). \quad (13)$$

Remark 1. *Our result incorporates (10), since $\delta(\varepsilon t) \in \mathbb{E}_l^\perp$, thus is contained in the span of L_0 -eigenspaces corresponding to purely negative eigenvalues, and so $\lim_{\varepsilon \rightarrow 0} \exp((t/\varepsilon)L_0^T) \delta(t) = 0$ for $t > 0$.*

The generator $\Pi_l G^T \Pi_l = (\Pi_r G \Pi_r)^T$ of the coarse grained master equation is an $n \times n$ -matrix with rank $m \ll n$. The matrix representation \hat{G} of $\Pi_r G \Pi_r$ with respect to the basis $\{\mathbf{1}_1, \dots, \mathbf{1}_m\}$ of \mathbb{E}_r is an $m \times m$ matrix given by

$$\hat{G}_{ij} = \mu_i^T G \mathbf{1}_j, \quad i, j = 1, \dots, m. \quad (14)$$

To see (14), we compute

$$\Pi_r G \Pi_r v = \sum_{k,j=1}^m \mathbf{1}_k \mu_k^T G \mathbf{1}_j \mu_j^T v,$$

so that by inserting the basis expression $v = \sum_i c_i \mathbf{1}_i$ of an arbitrary vector $v \in \mathbb{E}_r$ we get

$$\Pi_r G \Pi_r v = \sum_{k,j,i=1}^m \mathbf{1}_k \mu_k^T G \mathbf{1}_j \mu_j^T \mathbf{1}_i c_i = \sum_{k=1}^m \left(\sum_{j=1}^m \mu_k^T G \mathbf{1}_j c_j \right) \mathbf{1}_k,$$

that is, $\Pi_r G \Pi_r v = \sum_k a_k \mathbf{1}_k$ with coefficients $a_k = \sum_j \hat{G}_{kj} c_j$ which shows that \hat{G} is the desired matrix representation.

As a consequence of (14), \hat{G}^T is the matrix representation of $\Pi_l G^T \Pi_l$ in the dual basis $\{\mu_1, \dots, \mu_m\}$. Therefore, the coarse grained master equation (12) has the following m -dimensional representation in the basis $\{\mu_1, \dots, \mu_m\}$:

$$\frac{d}{dt} \hat{p}(t) = \hat{G}^T \hat{p}(t), \quad \hat{p}(0) = (\mathbf{1}_j^T \rho)_{j=1, \dots, m}, \quad (15)$$

such that $\hat{u}(t) = \sum_{j=1}^m \hat{p}_j(t) \mu_j$. Equation (15) is thus equivalent to (12), therefore we call (12) the coarse grained master equation of essential dimension m .

Remark 2. G need not be a rate matrix, because its off-diagonal entries need not be non-negative and it may even have positive diagonal entries. In contrast, \hat{G} is always a rate matrix. Moreover, it is self-adjoint, and has eigenvalues $0 = \hat{\lambda}_1 > \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_m$ where the $\hat{\lambda}_j$ are the constants introduced in (6). See Appendix B for details.

3.4. Coarse grained transfer operator

By far the most articles on Markov State Models consider transfer operators or stochastic matrices instead of generators or rate matrices. Thus, next we show how the coarse grained master equation (12) is related to the coarse grained transfer operator for the same system.

The forward transfer operator T_t^ε associated with (8) is its solution operator, i.e.,

$$(T_t^\varepsilon)^T = \exp(t L_\varepsilon^T),$$

while its transpose T_t^ε is called the associated backward transfer operator. With (11), (12), and Lemma 1b) we immediately get

$$\begin{aligned} \Pi_l (T_{t/\varepsilon}^\varepsilon)^T \Pi_l \rho &= \Pi_l u_\varepsilon(t/\varepsilon) = \Pi_l \exp(t \Pi_l G^T \Pi_l) \Pi_l \rho + \mathcal{O}(\varepsilon) \\ &= \exp(t \Pi_l G^T \Pi_l) \Pi_l \rho + \mathcal{O}(\varepsilon). \end{aligned}$$

Therefore,

$$\Pi_r T_{t/\varepsilon}^\varepsilon \Pi_r = \exp(t \Pi_r G \Pi_r) + \mathcal{O}(\varepsilon). \quad (16)$$

The entries of the matrix representation \hat{T}_t of the coarse grained transfer operator $\Pi_r T_{t/\varepsilon}^\varepsilon \Pi_r$ in the basis $\{\mathbf{1}_1, \dots, \mathbf{1}_m\}$ are given by

$$\hat{T}_{t,kj} = \mu_k^T T_t^\varepsilon \mathbf{1}_j. \quad (17)$$

Now, let μ_0 denote the limit for $\varepsilon \rightarrow 0$ of the invariant measures μ_ε . Its existence and form are given in Appendix B, in particular (42). Then

$$\mu_0 = \sum_{j=1}^m \alpha_j \mu_j, \quad \alpha_j = \mathbf{1}_j^T \mu_0.$$

Next, recalling the scalar product $\langle \cdot, \cdot \rangle_{\mu_0}$ from Section 3.1, the matrix representation of the coarse grained transfer operator is given by

$$\hat{T}_{t,kj} = \frac{\langle \mathbf{1}_k, T_t^\varepsilon \mathbf{1}_j \rangle_{\mu_0}}{\langle \mathbf{1}_k, \mathbf{1}_k \rangle_{\mu_0}},$$

which is the classical formula appearing in articles on Markov State Modelling for the coarse grained transfer operator. In addition, due to (16), the coarse grained transfer operator is identical to the solution operator of the coarse grained master equation up to order $\mathcal{O}(\varepsilon)$, i.e.,

$$\hat{T}_{t/\varepsilon} = \exp(t\hat{G}) + \mathcal{O}(\varepsilon). \quad (18)$$

Remark 3. *In the literature on Markov State Models one repeatedly finds remarks stating that the dynamics given by the coarse grained transfer operator \hat{T}_t is only correct, if the so-called lag time t is large enough [BPN13]. Our result (18) exhibits this fact in a very clear way: The coarse grained system approximates the full solution of the original master equation only on the slow time scale of order ε^{-1} .*

4. Non-stationary MSM

We will discuss now the non-stationary case, when L_ε is complemented by an explicitly time-dependent part,

$$\mathcal{L}_\varepsilon(t) = L_\varepsilon + \varepsilon f(t, \varepsilon t)L,$$

where the stationary part L_ε has the properties outlined before, f is a function depending on time and possibly on ε too, and L is an additional generator/rate matrix—we think of L being the part of the generator coming from the perturbing force field \tilde{b} , cf. Section 2. We again consider the associated master equation,

$$\frac{d}{dt} u_\varepsilon(t) = \mathcal{L}_\varepsilon(t)^T u_\varepsilon(t), \quad u_\varepsilon(0) = \rho, \quad (19)$$

and try to understand its long-term behavior. First, we realize that the generator L_0 is still the limit of $\mathcal{L}_\varepsilon(t)$ for $\varepsilon \rightarrow 0$, since

$$\mathcal{L}_\varepsilon(t) = L_0 + \varepsilon G + \varepsilon f(t, \varepsilon t)L + \mathcal{O}(\varepsilon^2),$$

with G introduced in (13).

We will approach (19) again by perturbation expansion, but now taking care of secular terms. We have different tools at our disposal, such as the *composite expansion*, where

expansions on slow and fast time scales are combined to be valid on both scales, to a large extent motivated by boundary layers [O'M69, Fra69], [Kue15, Section 9.1]; the *Lindstedt–Poincaré method* [Lin82, Poi93], [Kue15, Section 9.8], where a “strained time coordinate” is used to incorporate both scales and avoid secular terms; the *renormalization group method* [CGO94, CGO96], [Kue15, Section 9.9]; or *two-timing*, also called the *method of multiple scales* [Lic69, O'M70], [Kue15, Section 9.8], where multiple time scales are used in the very same expansion as independent variables. We will use two-timing, because it is fairly straight-forward, and serves our goal well, namely to have access to both fast and slow scales directly. For further introductory texts on two-timing the reader is referred to [Joh05, KC12, Jak16].

4.1. Driving force depends on slow time scale only

First, we restrict our attention to functions f that depend only on the slow time εt , meaning,

$$f(t, \varepsilon t) = f(\varepsilon t), \quad (20)$$

with a bounded function f , i.e., there is a constant $C > 0$ such that $|f(t)| \leq C$ for all times t . Later we will also consider the more general form $f(t) = f(t, \varepsilon t)$ where f also depends on the fast time scale.

However, even with f depending only on the slow time scale, the dynamics will evolve on both time scales: the fast one, t , and the slow one, $\tau = \varepsilon t$. Thus, for the solution of the master equation (19) we make the ansatz

$$u_\varepsilon = u_\varepsilon(t, \tau), \quad (21)$$

where we consider t and τ as independent variables. We are overloading the notation, since we use u_ε for the solution of (19), depending on *one* variable, and also for this new, auxiliary function, depending on *two* variables, t and τ . By counting the number of variable, no confusion should arise. The connection between them is

$$u_\varepsilon(t) = u_\varepsilon(t, \tau) \Big|_{\tau=\varepsilon t} = u_\varepsilon(t, \varepsilon t).$$

This connection shows how to substitute the ansatz into the equations:

$$\frac{d}{dt} u_\varepsilon(t) = \frac{\partial}{\partial t} u_\varepsilon(t, \tau) + \varepsilon \frac{\partial}{\partial \tau} u_\varepsilon(t, \tau),$$

and other appearances of $u_\varepsilon(t)$ are simply replaced by $u_\varepsilon(t, \tau)$. Naturally, $u_\varepsilon(0, 0) = \rho$. Now, (19) reads

$$\partial_t u_\varepsilon(t, \tau) + \varepsilon \partial_\tau u_\varepsilon(t, \tau) = (L_0 + \varepsilon L(\tau))^T u_\varepsilon(t, \tau) + \mathcal{O}(\varepsilon^2), \quad (22)$$

with the time-dependent generator

$$L(\tau) := G + f(\tau)L, \quad (23)$$

and shorthand notation

$$\partial_t u := \frac{\partial}{\partial t} u, \quad \text{and} \quad \partial_\tau u := \frac{\partial}{\partial \tau} u.$$

Using an expansion of the form

$$u_\varepsilon(t, \tau) = u_0(t, \tau) + \varepsilon u_1(t, \tau) + \mathcal{O}(\varepsilon^2)$$

we obtain the following result. We defer the proof to Appendix C.

Theorem 2. *Under the assumption (A1)-(A3) on the properties of L_ε , the driven n -dimensional master equation (22) has a solution $u_\varepsilon(t, \tau)$ of the form*

$$u_\varepsilon(t, \tau) = u_0(t, \tau) + \mathcal{O}(\varepsilon) = \hat{u}(\tau) + \exp(tL_0^T)\delta(\tau) + \mathcal{O}(\varepsilon),$$

valid for all t and $\tau = \mathcal{O}(1)$ as $\varepsilon \rightarrow 0$. Here, $\hat{u}(\tau) \in \text{Kern}(L_0^T) = \mathbb{E}_l$ is given by the m -dimensional (coarse grained) master equation

$$\frac{d}{d\tau} \hat{u}(\tau) = \Pi_l(G^T + f(\tau)L^T)\Pi_l \hat{u}(\tau), \quad \hat{u}(0) = \Pi_l \rho, \quad (24)$$

and $\delta(\tau) \in \mathbb{E}_l^\perp$ with $\delta(0) = \Pi_l^\perp \rho$. With (7) we get

$$u_\varepsilon(t, \tau) = \hat{u}(\tau) + \mathcal{O}(e^{-rt}) + \mathcal{O}(\varepsilon), \quad (25)$$

that is, the solution converges exponentially fast in t to \mathbb{E}_l and then follows the coarse grained master equation up to errors in $\mathcal{O}(\varepsilon)$.

Remark 4. First, note that Theorem 1 is a corollary of Theorem 2 by setting $f \equiv 0$, because $u_\varepsilon(t) = u_\varepsilon(t, \varepsilon t)$ solves the master equation (8), which is the same as (19) with $f \equiv 0$.

Second, the restriction $\tau = \mathcal{O}(1)$ is not a problem in practice, since one wants to consider the system on slow time scales $\mathcal{O}(\varepsilon^{-1})$, and that is possible. There is no need to go beyond. If this would be desired, one could invoke higher order terms in the expansion in Appendix C, at the cost of additional non-trivial calculations.

To gain more intuition for our result, let us briefly consider the limit $\varepsilon \rightarrow 0$ on the fast and slow time scales. Fix t , and note that $\tau = \varepsilon t \rightarrow 0$ as $\varepsilon \rightarrow 0$, thus we obtain

$$\lim_{\varepsilon \rightarrow 0} u_\varepsilon(t, \varepsilon t) = u_0(t, 0) = \underbrace{\hat{u}(0)}_{=\Pi_l \rho} + \exp(tL_0^T) \underbrace{\delta(0)}_{=\Pi_l^\perp \rho} \stackrel{\text{Lem. 1b)}}{=} \exp(tL_0^T)\rho.$$

So, on the fast time scale we see in the limit the evolution driven by the rate matrix that has the m invariant sets \mathbb{A}_j . Transitions between these sets are not observed on this time scale. Meanwhile, on the slow time scale τ , we have $t = \tau/\varepsilon \rightarrow \infty$ as $\varepsilon \rightarrow 0$, and we obtain

$$\lim_{\varepsilon \rightarrow 0} u_\varepsilon(\tau/\varepsilon, \tau) = u_0(\infty, \tau) = \hat{u}(\tau) + \underbrace{\exp(\infty L_0^T)\delta(\tau)}_{=0, \text{ since } \delta \in \mathbb{E}_l^\perp} = \hat{u}(\tau).$$

Thus, on the slow time scale, the fluctuations inside the sets \mathbb{A}_j caused by L_0 are equilibrated, and we only see the slow transition processes between the \mathbb{A}_j .

These latter slow processes have a low-dimensional coarse-grained representation, if projected on the sets \mathbb{A}_j . The matrix representation of $\Pi_l L(\tau)^T \Pi_l$ with respect to the basis $\{\mu_1, \dots, \mu_m\}$ of \mathbb{E}_l is given by $\hat{R}(\tau)^T$ with $\hat{R}(\tau)$ being the matrix representation of $\Pi_r L(\tau) \Pi_r$ in the basis $\{\mathbf{1}_1, \dots, \mathbf{1}_m\}$, and satisfying

$$\hat{R}(\tau) = \hat{G} + f(\tau) \hat{L}, \quad \hat{G}_{jk} = \mu_j^T G \mathbf{1}_k, \quad \hat{L}_{jk} = \mu_j^T L \mathbf{1}_k, \quad (26)$$

which can be seen based on the same derivation as for (14) above. That is, the m -dimensional coarse grained master equation reads as

$$\frac{d}{dt} \hat{p}(\tau) = \hat{R}(\tau)^T \hat{p}(\tau), \quad \hat{p}(0) = (\mathbf{1}_j^T \rho)_{j=1, \dots, m}, \quad (27)$$

with $\hat{u}(\tau) = \sum_{j=1}^m \hat{p}_j(\tau) \mu_j$.

4.2. Driving force depends both on the fast and slow time scale

What if f depends on the fast time also, i.e., if $f = f(t, \varepsilon t)$? Just a slight modification of the derivation of Theorem 2 is needed. With $L(s, \tau) := f(s, \tau) L$, we make the following assumption.

Assumption 1. *Let for all $t \geq 0$ hold that*

$$\int_0^t L(s, \tau)^T ds = t \bar{L}(\tau)^T + \Delta L(t, \tau)^T,$$

where \bar{L} is a possibly τ -dependent generator / rate matrix, and $\Delta L(t, \tau)$ is bounded in t and τ .

We obviously do not need to assume that the time-dependent perturbation of \mathcal{L}_ε is of the form $\varepsilon f(t, \tau) L$. In fact, we can allow for a more general form

$$\mathcal{L}_\varepsilon = L_0 + \varepsilon G + \varepsilon L(t, \tau), \quad (28)$$

where $L(t, \tau)$ is a rate matrix for all times t, τ . We have the following result, shown in Appendix D.

Theorem 3. *Assume that $L_\varepsilon = L_0 + \varepsilon G + \mathcal{O}(\varepsilon^2)$ has properties (A1)-(A3), listed initially. Moreover, let $L(t, \tau)$ satisfy Assumption 1. Then, the solution of the driven n -dimensional master equation,*

$$\frac{d}{dt} u_\varepsilon(t) = (L_0 + \varepsilon(G + L(t, \varepsilon t)) + \mathcal{O}(\varepsilon^2)) u_\varepsilon(t), \quad u_\varepsilon(0) = \rho,$$

has the form

$$u_\varepsilon(t, \tau) = \hat{u}(\tau) + \exp(t L_0^T) \delta(\tau) + \mathcal{O}(\varepsilon),$$

for all t and $\tau = \mathcal{O}(1)$ as $\varepsilon \rightarrow 0$, where $\delta(\tau) \in \mathbb{E}_l^\perp$, and $\delta(0) = \Pi_l^\perp \rho$. There, $\hat{u}(\tau) \in \mathbb{E}_l$ is given by the essentially m -dimensional (coarse grained) master equation

$$\frac{d}{d\tau} \hat{u}(\tau) = \Pi_l (G^T + \bar{L}(\tau)^T) \Pi_l \hat{u}(\tau), \quad \hat{u}(0) = \Pi_l \rho. \quad (29)$$

Let us first consider two exemplary cases, to see what can happen.

Case 1: External field is not changing the effective equilibrium dynamics. Let

$$f(t, \tau) = F(\tau) \sin(\omega t),$$

with a bounded F , so that

$$\int_0^t L(s, \tau)^T ds = t \underbrace{G^T}_{=\bar{L}^T} + \underbrace{F(\tau) \omega^{-1} (1 - \cos(\omega t)) L^T}_{=\Delta L(t, \tau)^T}.$$

Then, (56) gives us $\partial_\tau \hat{u}(\tau) = \Pi_l G^T \Pi_l \hat{u}(\tau)$, i.e., the resulting coarse grained master equation is identical to the coarse grained master equation of the equilibrium case. In general, if the fast-scale average $\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(s, \tau) ds$ of the forcing term is equal to zero, the coarse grained master equation is unaffected from the forcing. Things change though, if this is not the case.

Case 2: External field is changing the effective equilibrium dynamics. Consider now

$$f(t, \tau) = F(\tau) \sin^2(\omega t),$$

again with a bounded F . We have

$$\int_0^t L(s, \tau)^T ds = t G^T + F(\tau) \left(\frac{1}{2} t - \frac{1}{2\omega} \sin \omega t \cos \omega t \right) L^T$$

such that Assumption 1 dictates the choice

$$= t \left(\underbrace{G^T + \frac{1}{2} F(\tau) L^T}_{=\bar{L}^T} \right) - \underbrace{\frac{F(\tau)}{2\omega} \sin \omega t \cos \omega t L^T}_{=\Delta L(t, \tau)^T}.$$

This yields the coarse grained master equation

$$\partial_\tau \hat{u}(\tau) = \Pi_l \left(G^T + \frac{1}{2} F(\tau) L^T \right) \Pi_l \hat{u}(\tau). \quad (30)$$

Note that in this case $\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(s, \tau) ds = \frac{1}{2} F(\tau) \neq 0$.

Remark 5. Assumption 1 for the generators can be equivalently stated in terms of the forcing function as follows. Let

$$\frac{1}{t} \int_0^t f(s, \tau) ds = \hat{f}(\tau) + h(t, \tau),$$

with the function h that decays sufficiently fast in its first argument, i.e.,

$$th(t, \tau) \leq Z(\tau),$$

for all times t and for some Z that may depend on τ but is independent of t . The assertions of Theorem 3 would follow with $\bar{L}(\tau) = \hat{f}(\tau) L$.

4.3. Illustrative Example

For illustration of this result let us consider the following block diagonal 9×9 matrix with $m = 3$:

$$L_0 = \begin{pmatrix} -1 & 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3/2 & 3/4 & 3/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.9 & -1.8 & 0.9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & 1.5 & 1.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & -2 \end{pmatrix},$$

which is self-adjoint, irreducible and aperiodic in each of the three blocks $\mathbb{A}_1 = \{1, 2, 3\}$, $\mathbb{A}_2 = \{4, 5, 6\}$, and $\mathbb{A}_3 = \{7, 8, 9\}$. For simplicity, we set $L_\varepsilon = L_0$, i.e., the only transitions between the invariant sets \mathbb{A}_i will only appear due to the external perturbation. This has the form

$$L = \begin{pmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1.25 & 0 & 0 & 0 & 1.25 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We consider the solution u_ε of (22) with initial probability distribution

$$u_\varepsilon(0) = \rho = (1, 0, 0, 0, 0, 0, 0, 0, 0)^T,$$

and

$$f(t, \varepsilon t) = 2 \sin^2(2\pi \varepsilon t).$$

We compare the evolution of the probability of being in the set \mathbb{A}_1 ,

$$p_\varepsilon(t, \mathbb{A}_1) = \sum_{j \in \mathbb{A}_1} u_{\varepsilon,j}(t),$$

where $u_{\varepsilon,j}(t)$ denote the entries of the full solution $u_\varepsilon(t) \in \mathbb{R}^n$ of the master equation (22), with its coarse grained approximation

$$\hat{p}(t, \mathbb{A}_1) = \hat{p}_1(t),$$

where \hat{p} denotes the solution of (27), the representation of \hat{u} in the basis $\{\mu_1, \mu_2, \mu_3\}$. Figure 2 shows that $p(t, \mathbb{A}_1)$ and $\hat{p}(t, \mathbb{A}_1)$ are indeed ε -close to each other.

According to (25), the difference $u_\varepsilon(t) - \hat{u}(t)$ decays exponentially fast in t and then stays $\mathcal{O}(\varepsilon)$ -small. This behavior is illustrated in Figure 3, where the 1-norm of the error is shown.

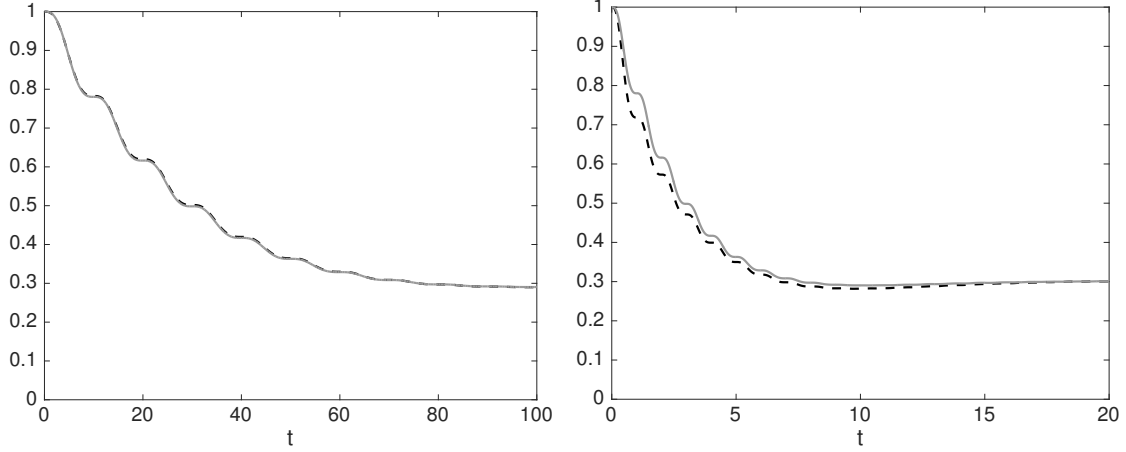


Figure 2: Comparison of the time-dependent probabilities $p(t, \mathbb{A}_1)$ (black, dashed) and $\hat{p}_t(\mathbb{A}_1)$ (gray, solid) for $\varepsilon = 0.05$ (left) and $\varepsilon = 0.5$ (right).

Next we consider the same case with f depending on the fast time scale t explicitly;

$$f(t, \varepsilon t) = 2 \sin^2(2\pi \varepsilon t) \sin^2(2\pi t).$$

The coarse grained master equation can be obtained from (30), since the forcing is of the form $F(\tau) \sin^2(\omega t)$. Figure 4 shows that also here the agreement is excellent.

5. Algorithms for coarse graining

In this section we very briefly outline how one could proceed to apply our approach for non-artificial, realistic systems, where an arbitrarily fine discretization of the generators used in the above exposition might be computationally out of reach. Note that the single steps below might contain non-trivial tasks to perform, where one would need to consult further literature. Thus, to give a detailed “recipe” would lead way beyond the scope of this work. We also expect a thorough treatment of this issue to differ from case to case, and thus will be topic of future work.

Now, we want to determine the coarse grained master equation (24) or (29) of the non-equilibrium process

$$d\mathbf{X}_t = \left[-\nabla_x V(\mathbf{X}_t) + \varepsilon f(t, \varepsilon t) \tilde{b}(\mathbf{X}_t) \right] dt + \sqrt{2\beta^{-1}} d\mathbf{B}_t, \quad (31)$$

where (for the sake of simplicity) subsequently we restrict ourselves to potential forces, i.e., $\tilde{b} = -\nabla_x U$ with some smooth potential U . To this end, we need to determine ε , m , \hat{G} and \hat{L} , in order to get the coarse grained generator $\hat{R} = \hat{G} + \varepsilon f(t, \tau) \hat{L}$ associated with (31), cf. (26). We do this in four steps:

(S1) First we consider the equilibrium process \mathbf{X}_t given by

$$d\mathbf{X}_t = -\nabla V(\mathbf{X}_t) dt + \sqrt{2\beta^{-1}} d\mathbf{B}_t. \quad (32)$$

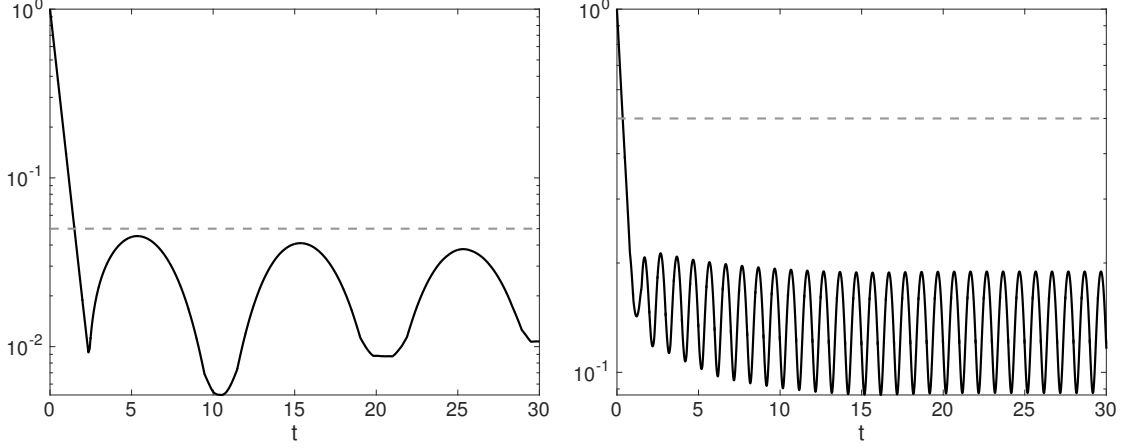


Figure 3: Evolution of $\|u_\varepsilon(t) - \hat{u}(t)\|_1$. The gray dashed line indicates the value of ε ; on the left 0.05, and on the right 0.5. Note that the error decays exponentially in $\mathcal{O}(1)$ time, then stays bounded by $\mathcal{O}(\varepsilon)$.

Using standard MSM building techniques [BPN13, PBB10, SS14] we construct an appropriate n -dimensional discretization T_θ of the transfer operator associated with (32) for some appropriately chosen short time $t = \theta$. Next, we determine m and ε by determining the dominant spectrum $\Lambda_1, \dots, \Lambda_m$ of T_θ , and setting, e.g., $\varepsilon = -\frac{1}{\theta} \log(\Lambda_2)$. Let us choose some $0 < \gamma < 1$ (by default, say, $\gamma = \frac{1}{2}$), and recompute a MSM T_{θ_ε} with $\theta_\varepsilon = \varepsilon^{-\gamma}$. Note, that the restriction $\gamma < 1$ is only due to practical reasons, since for $\gamma \geq 1$ we would need to compute a MSM on the slow time scales, or even longer ones, and that is in general computationally infeasible.

- (S2) Based on the dominant eigenvectors of T_{θ_ε} we identify its metastable sets $\mathbb{A}_1, \dots, \mathbb{A}_m$ and the local measures μ_j by standard MSM algorithm as, e.g., PCCA+ [DW04]. This allows us to determine the coarse grained transfer operator $\hat{T}_{\theta_\varepsilon}$ according to (17). We identify the generator \hat{G} associated with $\hat{T}_{\theta_\varepsilon}$ using (18), i.e., via

$$\exp(\varepsilon \theta_\varepsilon \hat{G}) = \hat{T}_{\theta_\varepsilon} + \mathcal{O}(\varepsilon), \quad \text{that is}^6, \quad \hat{G} = \frac{1}{\varepsilon \theta_\varepsilon} \log(\hat{T}_{\theta_\varepsilon}) + \underbrace{\frac{1}{\varepsilon \theta_\varepsilon} \mathcal{O}(\varepsilon)}_{=\mathcal{O}(\varepsilon^\gamma)}.$$

Note that the transfer operator is a semigroup generated by a self-adjoint infinitesimal generator, thus it has spectrum lying in the interval $[0, 1]$, whence $\log \hat{T}_{\theta_\varepsilon}$ is well-defined as a real-valued matrix. Alternatively, one could use the approximation $\hat{G} = \frac{1}{\varepsilon \theta_\varepsilon} (\hat{T}_{\theta_\varepsilon} - \text{Id})$, which is expected to be less accurate. With this choice of θ_ε we can make sure that T_{θ_ε} is still computationally tractable, and the error

⁶For $\gamma \leq 1$ we gave that all eigenvalues of $\hat{T}_{\theta_\varepsilon}$ are of order one, and since $\log(x + \mathcal{O}(\varepsilon)) = \log(x) + \mathcal{O}(\varepsilon)$ for x of order 1, the $\mathcal{O}(\varepsilon)$ error in (18) remains under taking logarithms of the two sides of the equation.

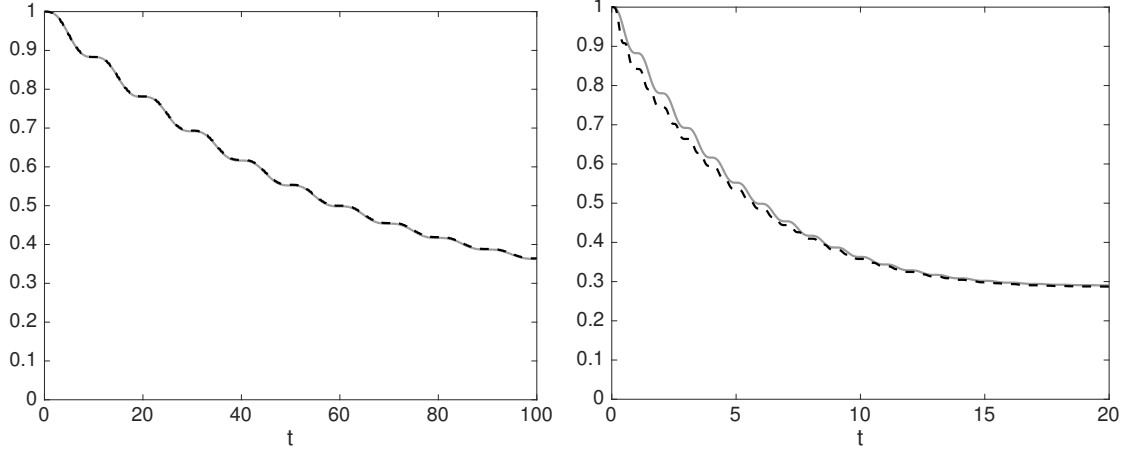


Figure 4: Comparison of the time-dependent probabilities $p(t, \mathbb{A}_1)$ (black, dashed) and $\hat{p}(t, \mathbb{A}_1)$ (gray, solid) for $\varepsilon = 0.05$ (left) and $\varepsilon = 0.5$ (right) for f depending on both time scales t and τ .

of \hat{G} vanishes for $\varepsilon \rightarrow 0$. Note that for $\gamma = 0$, we get error $\mathcal{O}(1)$, and we could not trust our results.

(S3) Next, we consider another equilibrium process \mathbf{Y}_t , given by

$$d\mathbf{Y}_t = -\left[\nabla V(\mathbf{Y}_t) + \varepsilon \nabla U(\mathbf{Y}_t)\right]dt + \sqrt{2\beta^{-1}}d\mathbf{B}_t. \quad (33)$$

In order to approximate the transfer operator $P_{\theta_\varepsilon}^\varepsilon$ associated with (33), we use the approach introduced in [SNW15] where it is shown that the coarse grained transfer operator $\hat{P}_{\theta_\varepsilon}^\varepsilon$ can be computed based on $\hat{T}_{\theta_\varepsilon}$ by reweighting the trajectories of (32) that have already been computed in step (S1) for T_{θ_ε} . More precisely, a procedure is given for the computation of the entries of $P_{\theta_\varepsilon}^\varepsilon$, only using information already having been computed in (S1) along with repeated evaluations of U and ∇U .

(S4) Knowing the $\mathbf{1}_j, \mu_j$, $j = 1, \dots, m$, we compute $\hat{P}_{\theta_\varepsilon}^\varepsilon$ from $P_{\theta_\varepsilon}^\varepsilon$ as we compute \hat{T}_θ from T_θ in (17). Note that if we take $f \equiv 1$ in (31), then by (26) and (18) we have $\hat{P}_{\theta_\varepsilon}^\varepsilon = \exp(\varepsilon \theta_\varepsilon \hat{Q}) + \mathcal{O}(\varepsilon)$ with $\hat{Q} = \hat{G} + \hat{L}$. Again, $\hat{P}_{\theta_\varepsilon}^\varepsilon$ is self-adjoint, thus its matrix logarithm is well defined and numerically stable to compute by taking the logarithms of the eigenvalues in its eigenvalue decomposition. Thus, we compute

$$\hat{L} = \frac{1}{\varepsilon \theta_\varepsilon} \log(\hat{P}_{\theta_\varepsilon}^\varepsilon) - \hat{G}, \quad (34)$$

and again, this identity holds up to an error $\mathcal{O}(\varepsilon^\gamma)$.

In light of Remark 3 it is interesting to see in (S1) and (S4) that we require a MSM on an intermediate time scale between slow and fast time scales in order to be able to assemble a master equation that is asymptotically correct as $\varepsilon \rightarrow 0$.

6. Example: diffusion in a potential landscape

We now return to the diffusive dynamics (1). Let us consider the double well potential $V(x) = (x^2 - 1)^2$, as we have done for $b \equiv 0$ in Example 1. Set $\beta = 5$, which makes the two wells of the potential metastable sets. As discussed above, L_ε denotes the (fine-grid) discretization of the generator $\mathcal{A}(0)$. More specifically, we use a stable finite-volume based discretization [FJK13, LeV02] of this generator on the interval $[-5, 5]$ with $n = 2000$. The two dominant eigenvalues of L_ε are

$$\lambda_1 = 0, \quad \lambda_2 = -0.011,$$

and we therefore set $\varepsilon = 0.011$ for the next steps. The other eigenvalues λ all satisfy

$$\lambda < -3,$$

thus $m = 2$.

The equilibrium dynamics is turned into a non-equilibrium process by considering

$$d\mathbf{X}_t = -\nabla_x \left[V(\mathbf{X}_t) + \varepsilon f(\varepsilon t) U(\mathbf{X}_t) \right] dt + \sqrt{2\beta^{-1}} d\mathbf{B}_t \quad (35)$$

with

$$U(x) = x, \quad f(t) = 15 \sin(t),$$

such that we have a tilting potential with periodic switches between deepening the left or right well, see Figure 5 for illustration. Despite the external time-dependent excitation the sets $\mathbb{A}_1 = (-\infty, 0]$ and $\mathbb{A}_2 = [0, \infty)$ are metastable sets of the dynamics.

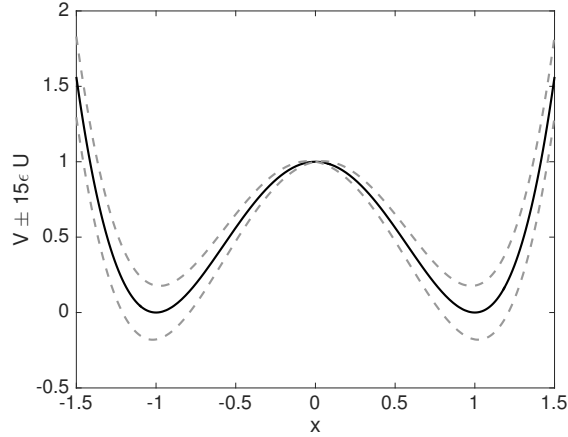


Figure 5: Unperturbed potential V (black solid) and tilted double well potentials $V \pm 15\varepsilon U$ (gray, dashed).

We are now going to compute the coarse grained master equation associated with (35). Thereby, we will use the steps from Section 5. However, since we are dealing with a one-dimensional example, we have a choice between possibly multiple computationally

tractable ways of doing one single step. We will make use of this freedom to keep the exposition simple, and ask the reader to keep in mind that this might be infeasible in higher dimensions.

As discussed before, we need to compute the coarse grained rate matrices \hat{G} and \hat{L} . To obtain the former, we proceed as in (S1)-(S2) in Section 5, and compute first a MSM for the unforced system, i.e. (8). As described there, we quickly obtain sufficient approximations for $\mathbf{1}_1, \mathbf{1}_2, \mu_1, \mu_2$ from the appropriate eigenvectors of L_ε . Since the fast fluctuations die out on time scales $\mathcal{O}(|\lambda_3^{-1}|)$, we take $\theta_\varepsilon = 10 \approx \varepsilon^{-1/2}$ to compute the auxiliary MSM T_{θ_ε} ; i.e., we choose $\gamma = 1/2$ in step (S2). Note that θ_ε is still an order of magnitude smaller than the slow time scale $\varepsilon^{-1} \approx 100$. We compute the coarse grained MSM $\hat{T}_{\theta_\varepsilon}$ via (17), such that we solve the full master equation (8) for initial conditions μ_1 and μ_2 to obtain $T_{\theta_\varepsilon}^T \mu_1$ and $T_{\theta_\varepsilon}^T \mu_2$, respectively. For this we used an implicit Euler method with step size $\theta_\varepsilon/20 = 0.5$. Finally, as in (S2), we set $\hat{G} = \frac{1}{\varepsilon\theta_\varepsilon} \log(\hat{T}_{\theta_\varepsilon})$. Figure 6 (left) compares the convergence of the probabilities $p(t, \mathbb{A}_1)$ and $\hat{p}_1(t)$ towards their limit, $\frac{1}{2}$, in the unforced case for initial densities $u(0) = \mu_1$ and $\hat{u}(0) = (1, 0)^T$, respectively. The slopes of the graphs therein show the convergence rates $\lambda_2 \approx \varepsilon \hat{\lambda}_2$, where $\hat{\lambda}_2$ is the second eigenvalue of \hat{G} .

Now, to obtain \hat{L} , we repeat the same procedure, just for the equilibrium system having the potential $V + \varepsilon U$. Its coarse grained transfer operator $\hat{P}_{\theta_\varepsilon}$ has the generator $\hat{Q} = \varepsilon(\hat{G} + \hat{L})$, thus we obtain $\hat{L} = \varepsilon^{-1}\hat{Q} - \hat{G}$. With this, all is set up for the coarse grained master equation of (35). To validate our approach, we compute the temporal evolution of probabilities in the metastable set \mathbb{A}_1 by three different means:

- A) Highly accurate approximation by determining the solution $u_\varepsilon(t)$ of the full master equation (19) generated by $\mathcal{L}_\varepsilon(t) = L_\varepsilon + \varepsilon f(\varepsilon t)L$, where L is the generator associated with the tilting potential U , computed by the same method and on the same grid as L_ε . The master equation is computed by the implicit Euler method with step size 0.1.
- B) Trajectory-based Monte Carlo approximation by computing $N = 1000$ random realizations $\mathbf{X}_t^{(j)}$, $j = 1, \dots, N$ of (35) by the Euler–Maruyama method with step size 0.005, and subsequent histogram approximation u_{MC} of the time-dependent probability density functions.
- C) Solution of the two-dimensional coarse grained master equation (27).

Based on these approximations we can provide several ways to compute the evolution of the probability to be in the left well:

$$\begin{aligned} p(t, \mathbb{A}_1) &= \sum_{j \in \mathbb{A}_1} u_{\varepsilon,j}(t), \\ p^{\text{MC}}(t, \mathbb{A}_1) &= \#\{j = 1, \dots, N : \mathbf{X}_t^{(j)} \in \mathbb{A}_1\}, \\ \hat{p}(t, \mathbb{A}_1) &= \hat{p}_1(t). \end{aligned}$$

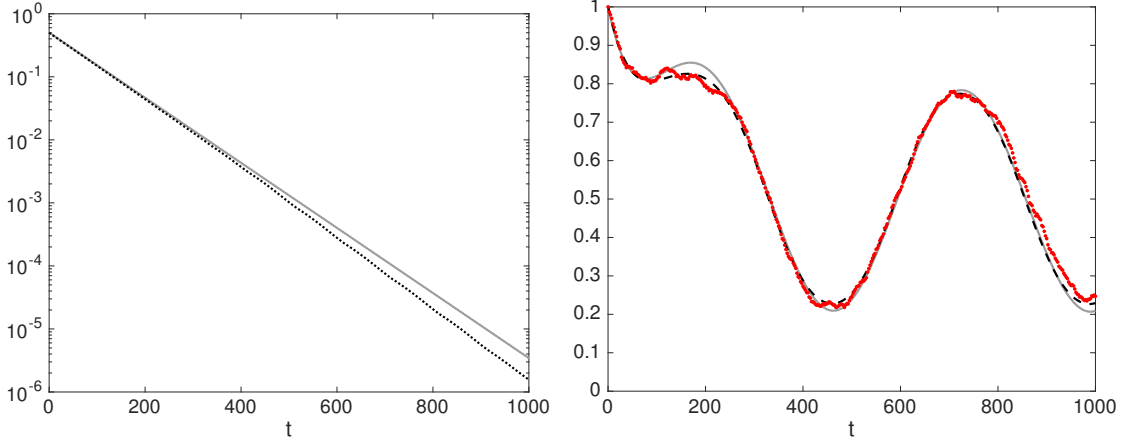


Figure 6: Left: semilogarithmic plot of the errors $|p(t, \mathbb{A}_1) - \frac{1}{2}|$ (gray, solid) and $|\hat{p}_1(t) - \frac{1}{2}|$ (black, dotted) for the unforced system, showing that the rate of convergence to equilibrium is essentially the same for L_ε and \hat{G} . Right: evolution of probabilities $p(t, \mathbb{A}_1)$, $p^{\text{MC}}(t, \mathbb{A}_1)$, and $\hat{p}(t, \mathbb{A}_1)$ for the highly accurate approximation A) of the non-equilibrium dynamics (black, dashed), its trajectory-based Monte Carlo approximation B) (red dots), and its coarse grained approximation C) (gray, solid), respectively.

Figure 6 (right) shows these different approximations. We observe that the computationally expensive approximations A) and B) are accurately approximated by the coarse grained, just two-dimensional version C).

Note that even though the amplitude of the forcing is $\mathcal{O}(\varepsilon)$, its effect on the occupation probabilities is of order one: the occupation probabilities vary by 0.6. We would like to stress, that the Markov state modeling technique we developed here relies on *time scale separation* by a small factor ε , but is describing variations of the coarse grained probabilities on order one.

To see that it is instrumental for the approximation of the coarse grained generators \hat{G} and \hat{L} to obtain them from a MSM that is computed on a sufficiently large time scale, we repeat our calculations, only now with $\gamma = 0$, i.e. $\theta_\varepsilon = 1$. Note also, that the μ_j and $\mathbf{1}_j$ are unaffected by this, because here we compute those from the full generator L_ε . The results are shown in Figure 7. Indeed, larger errors arise. For $\theta = 0.1$, even the qualitative agreement between coarse grained and actual solution vanishes (not shown).

As a last test, we consider a case where the forcing acts both on the slow and fast scales:

$$f(t, \tau) = 15 \sin(4\tau) \sin(t)^2.$$

We can reuse the already computed coarse-grained generators \hat{G}, \hat{L} , but we need to assemble the coarse grained master equation according to (30). The results are shown in Figure 8. The inclusion of forcing on the fast time scale decreases the accuracy of the Monte Carlo estimation (still using 1000 realizations), while the solution of our coarse grained master equation is indistinguishable from the exact solution.

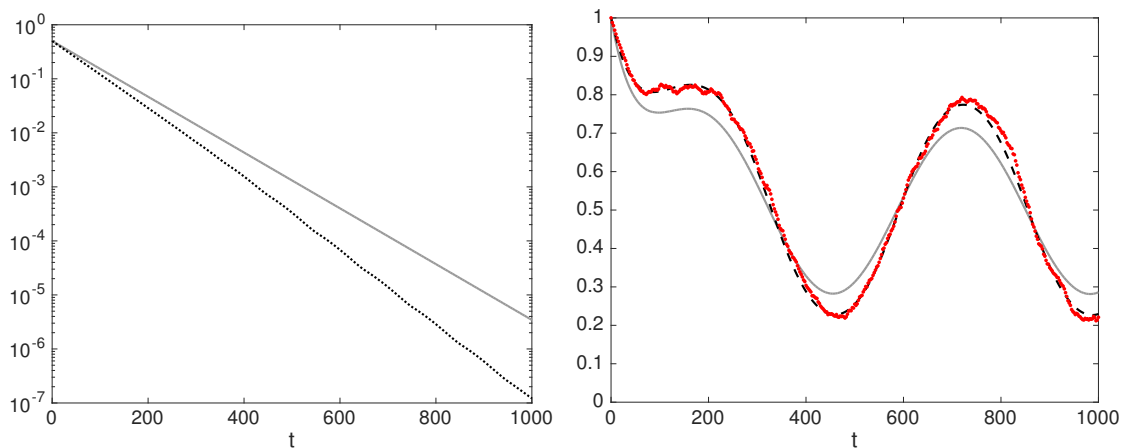


Figure 7: The same as Figure 6, only now for $\theta = 1$.

7. Conclusions

We show, that if the strength of the forcing of a metastable system is inversely proportional to the length of the slow metastable time scales of the unforced system, then the effective behavior of the forced system on slow time scales can be described by a low-dimensional reduced master equation that can be algorithmically constructed based on relatively short trajectories of the unforced system only.

The main limitations of our approach lie in the explicit scaling of the forcing and the multi scale perturbation expansion method utilized for analyzing its long-term effect, and in the fact that we did not consider the perturbation approach for the (unbounded) operators that appear in the Fokker–Planck equation but for their finite-dimensional matrix discretizations only. The first limitation is a feature since we aimed at showing that there exist cases in which a weak and slow forcing induces non-trivial and significant changes in the conformation dynamics of the molecular system under consideration. The second one, if overcome, would allow to understand better where the operator L_0 , that “defines” the metastable sets of the system, originates from, and how it can be defined explicitly.

Forthcoming work will also have to discuss the consequences of the insight that even weak fields can significantly alter the conformational dynamics of biomolecular systems, and the discussion of potential effects of electromagnetic radiation on the human body tissue. Furthermore, our results—at least in principle—would allow for designing an external forcing of a biomolecular system with specific effect on certain conformations, for example, one may design a forcing in order to control the occupation probability of a desired conformation of long timescales, in order to stabilize certain molecular function that is associated with this conformation. Because of the enormous complexity of real-world biomolecular systems, however, applications regarding optimal control of conformation dynamics seem feasible only if based on a low-dimensional reduced master equation.

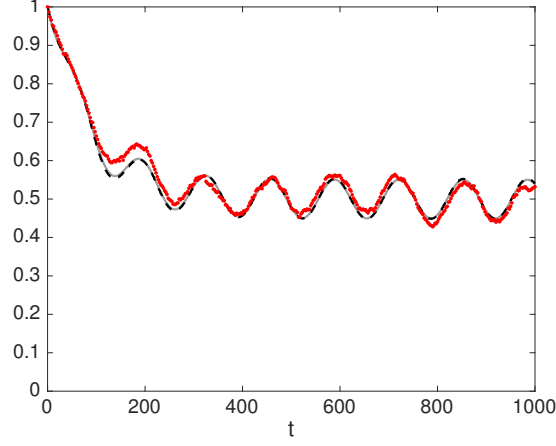


Figure 8: Evolution of probabilities $p(t, \mathbb{A}_1)$, $p^{\text{MC}}(t, \mathbb{A}_1)$, and $\hat{p}(t, \mathbb{A}_1)$ for the system with forcing on both fast and slow times scales, $f(t, \tau) = 15 \sin(4\tau) \sin(t)^2$. Colors and line styles are as in Figure 6.

Acknowledgments

This work is supported by the Deutsche Forschungsgemeinschaft (DFG) through the CRC 1114 “Scaling Cascades in Complex Systems”, and by the Einstein Foundation Berlin (Einstein Center ECMath).

A. Singularly perturbed linear systems

Let us consider to the master equation (8) in its state before we started the multiscale asymptotics:

$$\frac{d}{dt}u(t) = (L_0 + \varepsilon L)^T u(t),$$

let us scale time $\tau = \varepsilon t$ and consider $v = u^T$ to get

$$\frac{d}{d\tau}v_\varepsilon(\tau) = (\varepsilon^{-1}L_0 + L)v_\varepsilon(\tau),$$

where the index of u, v denotes the dependence of the solution on ε . This latter is a singularly perturbed linear system and so the literature on such systems should be helpful. Obviously

$$v_\varepsilon(\tau) = \exp(\tau(\varepsilon^{-1}L_0 + L))v_\varepsilon(0).$$

Since L_0 is semistable (all nonzero eigenvalues are negative), from [CR79] we get that if $v_\varepsilon(0) \rightarrow v_{\text{init}}$ for $\varepsilon \rightarrow 0$, then in the limit of $\varepsilon \rightarrow 0$,

$$v_\varepsilon(\tau) \rightarrow \exp(\tau A)(\text{Id} - L_0^D L_0)v_{\text{init}}, \quad A = (\text{Id} - L_0^D L_0)L(\text{Id} - L_0^D L_0).$$

Here, L_0^D denotes the *Drazin pseudoinverse* of L_0 , which in our case has the form

$$L_0^D = \left(L_0|_{\mathbb{E}_r^\perp} \right)^{-1} \Pi_r^\perp$$

where $\Pi_r^\perp = \text{Id} - \Pi_r$, and $\left(L_0|_{\mathbb{E}_r^\perp} \right)^{-1}$ is the inverse of the restriction $L_0|_{\mathbb{E}_r^\perp}$ of L_0 onto \mathbb{E}_r^\perp . Recall, that \mathbb{E}_r is the eigenspace of L_0 for the eigenvalue zero, thus $\left(L_0|_{\mathbb{E}_r^\perp} \right)^{-1}$ exists. As a consequence, we obtain

$$\begin{aligned} \text{Id} - L_0^D L_0 &= \text{Id} - \left(L_0|_{\mathbb{E}_r^\perp} \right)^{-1} \Pi_r^\perp L_0 (\Pi_r + \Pi_r^\perp) \\ &\stackrel{\text{Lem. 1a)}}{=} \text{Id} - \left(L_0|_{\mathbb{E}_r^\perp} \right)^{-1} \underbrace{\Pi_r^\perp L_0 \Pi_r^\perp}_{= L_0|_{\mathbb{E}_r^\perp} \Pi_r^\perp} \\ &= \text{Id} - \Pi_r^\perp = \Pi_r. \end{aligned}$$

Thus,

$$A = (\text{Id} - L_0^D L_0) L (\text{Id} - L_0^D L_0) = \Pi_r L \Pi_r,$$

such that we obtain (10) and (11) & (12), namely

$$v_\varepsilon(\tau) \rightarrow \exp(\tau \Pi_r L \Pi_r) \Pi_r v_{\text{init}}, \quad \varepsilon \rightarrow 0.$$

B. Properties of \hat{G}

\hat{G} is a rate matrix. Let L_ε, L_0 satisfy the assumptions (A1)-(A3) from the main text, and let $G = \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (L_\varepsilon - L_0)$. Since L_ε, L_0 are both rate matrices, we have that $L_\varepsilon \mathbf{1} = 0$ and that $L_0 \mathbf{1} = 0$, thus

$$G \mathbf{1} = \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (L_\varepsilon - L_0) \mathbf{1} = 0.$$

Each row sum of G is thus zero. Note, however, that G need not be a rate matrix itself, we only require $L_\varepsilon = L_0 + \varepsilon G + \mathcal{O}(\varepsilon^2)$ to be one for sufficiently small ε , thus also $L_0 + \varepsilon G$.

Recalling the definition (4) of a rate matrix, and noting that multiplication by $\mathbf{1}_j$ means summing up columns whose index belonging to \mathbb{A}_j , we have

$$L_0 \mathbf{1}_j = 0, \quad (L_\varepsilon \mathbf{1}_j)|_{\mathbb{A}_j} \leq 0, \quad (L_\varepsilon \mathbf{1}_j)|_{\mathbb{A}_j^c} \geq 0,$$

where $v|_{\mathbb{A}} \leq (\geq) 0$ means that the vector v restricted to its components given by the index set \mathbb{A} is componentwise non-positive (non-negative), and \mathbb{A}_j^c denotes the complement of \mathbb{A}_j , i.e., $\{1, \dots, n\} \setminus \mathbb{A}_j$. This implies

$$\varepsilon^{-1} (L_\varepsilon - L_0) \mathbf{1}_j \begin{cases} \leq 0 & \text{on } \mathbb{A}_j, \\ \geq 0 & \text{on } \mathbb{A}_j^c. \end{cases} \quad (36)$$

Recall that we defined μ_j as $\mu_j^T L_0 = 0$ and $\mu_j|_{\mathbb{A}_j^c} = 0$. Additionally, they have the property $\mu_j \geq 0$ componentwise. This readily gives with (36) that

$$\hat{G}_{ij} = \mu_i^T G \mathbf{1}_j = \mu_i^T \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (L_\varepsilon - L_0) \mathbf{1}_j \begin{cases} \leq 0 & i = j, \\ \geq 0 & i \neq j. \end{cases}$$

According to (14), these are the entries of the $m \times m$ discretization matrix \hat{G} of $\Pi_r G \Pi_r$. Moreover, $\Pi_r G \Pi_r \mathbf{1} = \Pi_r G \mathbf{1} = 0$ implies $\hat{G} \mathbf{1} = 0$. Thus, \hat{G} is a rate matrix.

Eigenmodes. By assumption (A1) the eigenvalue $\lambda = 0$ of L_ε is simple with unique right eigenvector $\mathbf{1}$. Let the (right) eigenvectors of L_ε be denoted by v_ε^j with eigenvalues λ_ε^j ,

$$L_\varepsilon v_\varepsilon^i = \lambda_\varepsilon^i v_\varepsilon^i. \quad (37)$$

Next, let us concentrate on the first m eigenvalues of L_ε that according to our assumptions will have to be perturbations of the m -fold eigenvalue $\lambda = 0$ of L_0 : For small enough ε our assumptions yield that there are coefficients c_j^i assembled into coefficient vectors $c^i \in \mathbb{R}^m$, and $\alpha_j > 0$ in

$$v_0^i = \sum_{j=1}^m c_j^i \mathbf{1}_j, \quad \mu_0 = \sum_{j=1}^m \alpha_j \mu_j,$$

such that the v_0^i , $i = 1, \dots, m$, span the right eigenspace \mathbb{E}_r of L_0 associated with eigenvalue $\lambda = 0$, and we have an asymptotic expansion of the form

$$\begin{aligned} \lambda_\varepsilon^i &= 0 + \varepsilon \lambda_1^i + \mathcal{O}(\varepsilon^2), \\ v_\varepsilon^i &= v_0^i + \varepsilon v_1^i + \mathcal{O}(\varepsilon^2), \\ \mu_\varepsilon &= \mu_0 + \varepsilon \nu + \mathcal{O}(\varepsilon^2). \end{aligned}$$

For later use, we emphasize that $v_0^1 = \mathbf{1}$ and therefore $c_j^1 = 1$ for all $j = 1, \dots, m$, moreover, $v_1^1 = 0$, and $\lambda_1^1 = 0$. Putting the asymptotic expansion into (37), using $\mu_\varepsilon^T L_\varepsilon = \mu_\varepsilon^T$, and comparing orders 0 and 1 in ε yields

$$\begin{aligned} \varepsilon^0 : \quad & \mu_0^T L_0 = 0, \\ & L_0 v_0^i = 0 \end{aligned} \quad (38)$$

$$\varepsilon^1 : \quad \mu_0^T G + \nu^T L_0 = 0 \quad (39)$$

$$G v_0^i + L_0 v_1^i = \lambda_1^i v_0^i \quad (40)$$

While the two equations for order 0 are satisfied automatically (see Section 3.1 for a discussion), equations (39) and (40) result in conditions on the unknown coefficients c_j^i and α_j .

In order to see this, let $\eta := \sum_{j=1}^m \gamma_j \mu_j \in \mathbb{R}^n$ with $\gamma \in \mathbb{R}^m$ arbitrary, and multiply (40) by η^T to obtain

$$\sum_{j,k=1}^m \gamma_j \mu_j^T G \mathbf{1}_k c_k^i = \lambda_1^i \sum_{j,k=1}^m \gamma_j \mu_j^T \mathbf{1}_k c_k^i \stackrel{\mu_j^T \mathbf{1}_k = \delta_{jk}}{=} \lambda_1^i \sum_{j=1}^m \gamma_j c_j^i,$$

or, equivalently,

$$\gamma^T \hat{G} c^i = \lambda_1^i \gamma^T c^i \quad \text{for all } \gamma \in \mathbb{R}^m.$$

Since this has to hold for arbitrary γ , we obtain

$$\hat{G} c^i = \lambda_1^i c^i. \quad (41)$$

Similarly simple is to obtain a condition on the coefficient vector α of μ_0 . Multiplying (39) from the right by $\sum_{j=1}^m z_j \mathbf{1}_j$, where $z \in \mathbb{R}^m$ is arbitrary, we obtain after an analogous calculation as above that $\alpha^T \hat{G} z = 0$ for any z , implying

$$\alpha^T \hat{G} = 0. \quad (42)$$

Thus, the eigenvalue problems (42) and (41) determine the coefficients of the zeroth order expansion terms μ_0 and v_0 uniquely up to a multiplicative scalar. Uniqueness is achieved by adding normalizing conditions; such as $\sum_{i=1}^m \alpha_i = 1$ for μ_0 to be a probability measure.

Reversibility. Reversibility in assumption (A1) means the self-adjointness of L_ε with respect to the scalar product $\langle \cdot, \cdot \rangle_{\mu_\varepsilon}$, cf. Section 3.1 for this notation. That can be expressed in two other equivalent ways, namely

$$\langle L_\varepsilon v, w \rangle_{\mu_\varepsilon} = \langle v, L_\varepsilon w \rangle_{\mu_\varepsilon} \quad \forall v, w \in \mathbb{R}^n, \quad \text{or} \quad D_{\mu_\varepsilon} L_\varepsilon = L_\varepsilon^T D_{\mu_\varepsilon},$$

where D_{μ_ε} is the diagonal matrix with the entries of μ_ε on its diagonal. Due to the assumed smoothness of μ_ε and L_ε in ε , the latter expression gives $D_{\mu_0} L_0 = L_0^T D_{\mu_0}$ in the $\varepsilon \rightarrow 0$ limit. Thus L_0 is μ_0 -reversible. Self-adjointness of a matrix means orthogonal diagonalizability in the given scalar product, so we have that eigenvectors to different eigenvalues are orthogonal, and we obtain for $i \neq j$:

$$\begin{aligned} 0 &= \langle L_\varepsilon v_\varepsilon^i, v_\varepsilon^j \rangle_{\mu_\varepsilon} \\ &= \langle L_0 v_0^i + \varepsilon G v_0^i + \varepsilon L_0 v_1^i, v_0^j + \varepsilon v_1^j \rangle_{\mu_0 + \varepsilon \nu} + \mathcal{O}(\varepsilon^2) \\ &= \langle L_0 v_0^i, v_0^j \rangle_{\mu_0 + \varepsilon \nu} + \varepsilon \langle G v_0^i, v_0^j \rangle_{\mu_0} + \varepsilon \langle L_0 v_1^i, v_0^j \rangle_{\mu_0} + \mathcal{O}(\varepsilon^2). \end{aligned}$$

In the last line, the first term on the right hand side vanishes due to $L_0 v_0^i = 0$, cf (38). The third term vanishes, because L_0 is self-adjoint with respect to $\langle \cdot, \cdot \rangle_{\mu_0}$, hence $\langle L_0 v_1^i, v_0^j \rangle_{\mu_0} = \langle v_1^i, L_0 v_0^j \rangle_{\mu_0} = 0$ due to (38) again. Since $\langle L_\varepsilon v_\varepsilon^i, v_\varepsilon^j \rangle_{\mu_\varepsilon}$ vanishes uniformly in ε , it follows that

$$\langle G v_0^i, v_0^j \rangle_{\mu_0} = 0 \quad (43)$$

has to hold for $i \neq j$. Using the expression of the v_0^i in the basis $\{\mathbf{1}_1, \dots, \mathbf{1}_m\}$, this gives with (14) that

$$\langle c^i, \hat{G}c^j \rangle_\alpha = 0, \quad \forall i \neq j, \quad i, j = 1, \dots, m. \quad (44)$$

Since the v_0^i are mutually $\langle \cdot, \cdot \rangle_{\mu_0}$ -orthogonal (as eigenvectors of the self-adjoint matrix L_0), they form a linearly independent collection of vectors, hence the c^i form a basis of \mathbb{R}^m . Thus, (44) shows that \hat{G} is self-adjoint with respect to $\langle \cdot, \cdot \rangle_\alpha$, and so it generates as a rate matrix an α -reversible process.

Summarizing, \hat{G} is a rate matrix and has a unique, positive left eigenvector α associated with the eigenvalue 0, and \hat{G} is symmetric with respect to $\langle \cdot, \cdot \rangle_\alpha$ and thus has real-valued spectrum. Let $0 = \hat{\lambda}_1 > \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_m$ be the eigenvalues of \hat{G} and $c^1 = \mathbf{1}, c^2, \dots, c^m$ be the associated eigenvectors. Then the m leading eigenvalues of L_ε satisfy

$$\lambda_\varepsilon^i = \varepsilon \hat{\lambda}_i + \mathcal{O}(\varepsilon^2),$$

and the corresponding right eigenvectors fulfill

$$v_\varepsilon^i = v_0^i + \mathcal{O}(\varepsilon) = \sum_{k=1}^m c_k^i \mathbf{1}_k + \mathcal{O}(\varepsilon).$$

Remark 6. In this context (only for transfer operators instead of generators), Deufhard and Weber [DW04, Lemma 1.1] show that the $\mathcal{O}(\varepsilon)$ term in the expansion of v_ε^i is also a linear combination of the vectors $\mathbf{1}_j$, thus one has

$$v_\varepsilon^i = \sum_{k=1}^m \tilde{c}_k^i \mathbf{1}_k + \mathcal{O}(\varepsilon^2),$$

for some coefficients \tilde{c}_k^i .

C. Proof of Theorem 2

Inserting the asymptotic expansion

$$u_\varepsilon(t, \tau) = u_0(t, \tau) + \varepsilon u_1(t, \tau) + \mathcal{O}(\varepsilon^2).$$

into (22), and matching the ε -terms of equal order yields the following equations for the expansion terms:

$$\varepsilon^0: \quad \partial_t u_0 = L_0^T u_0, \quad u_0(0, 0) = \rho \quad (45)$$

$$\varepsilon^1: \quad \partial_t u_1 + \partial_\tau u_0 = L_0^T u_1 + L(\tau)^T u_0, \quad u_1(0, 0) = 0 \quad (46)$$

From (45) we immediately get

$$u_0(t, \tau) = \exp(tL_0^T)u_0(0, \tau). \quad (47)$$

As a consequence we find that $\Pi_l u_0$ does not depend on t , but only on τ ,

$$\Pi_l u_0(t, \tau) \stackrel{\text{Lem. 1b)}}{=} \Pi_l u_0(0, \tau) =: \hat{u}(\tau). \quad (48)$$

Next, multiplication of (46) from the left with Π_l yields

$$\begin{aligned} \Pi_l \partial_t u_1(t, \tau) + \Pi_l \partial_\tau u_0(t, \tau) &\stackrel{\text{Lem. 1a)}}{=} f(\tau) \Pi_l L^T u_0(t, \tau) \\ &= \Pi_l L(\tau)^T \exp(t L_0^T) u_0(0, \tau), \end{aligned} \quad (49)$$

whose integration with respect to t gives

$$\Pi_l u_1(t, \tau) = \Pi_l u_1(0, \tau) - t \partial_\tau \hat{u}(\tau) + \int_0^t \Pi_l L(\tau)^T \exp(s L_0^T) u_0(0, \tau) ds. \quad (50)$$

Next, we define

$$\delta(\tau) := u_0(0, \tau) - \hat{u}(\tau) \stackrel{(48)}{=} \Pi_l^\perp u_0(0, \tau), \quad (51)$$

which, since $\hat{u} \in \mathbb{E}_l$ and thus $\Pi_l \delta(\tau) = 0$, is an element of the orthogonal complement \mathbb{E}_l^\perp of \mathbb{E}_l . Thus,

$$\begin{aligned} \exp(s L_0^T) u_0(0, \tau) &= \exp(s L_0^T) (\hat{u}(\tau) + \delta(\tau)) \\ &= \hat{u}(\tau) + \exp(s L_0^T) \delta(\tau). \end{aligned}$$

The advantage of this splitting is that the first term in the right-hand side, $\hat{u}(\tau)$, does not depend on the fast time, here s , hence can be pulled out of the integral in (50), giving

$$\begin{aligned} \Pi_l u_1(t, \tau) &= \Pi_l u_1(0, \tau) + t (\Pi_l L(\tau)^T \hat{u}(\tau) - \partial_\tau \hat{u}(\tau)) \\ &\quad + \Pi_l L(\tau)^T \int_0^t \exp(s L_0^T) \delta(\tau) ds. \end{aligned} \quad (52)$$

At this point, we need to make sure that no secular terms appear. The last term in (52) is bounded in t . To see this, note that $\delta(\tau) \in \mathbb{E}_l^\perp$, thus it is spanned by eigenvectors associated with purely negative eigenvalues of L_0 , so the integrand $\exp(s L_0^T) \delta(\tau)$ decays exponentially in s , and hence the integral is uniformly bounded in t . To assure that the second term on the right-hand side of (52) is bounded, the expression in the brackets has to vanish, that is,

$$\partial_\tau \hat{u}(\tau) = \Pi_l L(\tau)^T \hat{u}(\tau), \quad (53)$$

must hold. With this, we made sure that $\Pi_l u_1$ is not secular in t . Fortunately, u_1 itself also cannot be secular in t , because by (46) u_1 is governed by a forced linear system of equations, cf. Footnote 5, where the linear part L_0^T is asymptotically stable on \mathbb{E}_l^\perp (i.e., decaying exponentially), and the forcing term has the same properties as u_0 , which is bounded in t by (47) and that L_0 is a rate matrix. By considering higher order expansion terms one could obtain conditions that allow bounding u_1 in τ . Without doing this, the validity of our expansion can only be secured for times $\tau = \tau(\varepsilon)$, such that $u_1(t, \tau) = \mathcal{O}(1)$ as $\varepsilon \rightarrow 0$ for fixed t . As we argued above, since the order equations

are system linear equations, u_1 can grow at most exponentially in τ as well, and this guarantees that the expansion is valid for $\tau = \mathcal{O}(1)$.

To put everything together, we recall that $\hat{u}(\tau) = \Pi_l u_0(0, \tau)$, thus $\Pi_l \hat{u} = \hat{u}$, and we can write (53) as

$$\partial_\tau \hat{u}(\tau) = \Pi_l L(\tau)^T \Pi_l \hat{u}(\tau). \quad (54)$$

In more details, the generator of the coarse grained master equation (54) satisfies

$$\Pi_l L(\tau)^T \Pi_l = \Pi_l G^T \Pi_l + f(\tau) \Pi_l L^T \Pi_l.$$

D. Proof of Theorem 3

The derivation follows that from Appendix C until (50). With $L(s, \tau) := f(s, \tau)L$, (50) reads as

$$\Pi_l u_1(t, \tau) = \Pi_l u_1(0, \tau) - t \partial_\tau \hat{u}(\tau) + \int_0^t \Pi_l L(s, \tau)^T \exp(s L_0^T) u_0(0, \tau) ds,$$

and we get by using (51) that

$$\begin{aligned} \Pi_l u_1(t, \tau) &= \Pi_l u_1(0, \tau) + \Pi_l \left(\int_0^t L(s, \tau)^T ds \hat{u}(\tau) - t \partial_\tau \hat{u}(\tau) \right) \\ &\quad + \Pi_l \int_0^t L(s, \tau)^T \exp(s L_0^T) \delta(\tau) ds. \end{aligned} \quad (55)$$

Thus, since $\delta \in \mathbb{E}_l^\perp$, $\exp(s L_0^T) \delta$ decays exponentially fast in s , so that the last integral is bounded in time t . The second term on the right-hand side of (55) can be guaranteed to be non-secular if Assumption 1 holds, by setting

$$\partial_\tau \hat{u}(\tau) = \bar{L}(\tau)^T \hat{u}(\tau) \quad \hat{u}(0) = \Pi_l \rho, \quad (56)$$

such that \hat{u} is bounded for all τ because $\bar{L}(\tau)$ is assumed to be a generator. Thus, the remainder of the term is bounded, since $\Delta L(t, \tau)^T \hat{u}(\tau)$ is, also by this assumption. Hence, no secular terms appear, and the rest of the derivation would follow that of Theorem 2.

References

- [AGT12] Loukas G Astrakas, Christos Gousias, and Margaret Tzaphlidou. Structural destabilization of chignolin under the influence of oscillating electric fields. *J. Appl. Phys.*, 111(7):074702–074702, 2012.
- [BD14] I. Bekard and D.E. Dunstan. Electric field induced changes in protein conformation. *Soft Matter*, pages –, 2014.
- [BPN13] Gregory R Bowman, Vijay S Pande, and Frank Noé. *An introduction to Markov state models and their application to long timescale molecular simulation*, volume 797. Springer Science & Business Media, 2013.

- [CGO94] Lin Yuan Chen, Nigel Goldenfeld, and Yoshitsugu Oono. Renormalization group theory for global asymptotic analysis. *Physical review letters*, 73(10):1311, 1994.
- [CGO96] Lin-Yuan Chen, Nigel Goldenfeld, and Yoshitsugu Oono. Renormalization group and singular perturbations: Multiple scales, boundary layers, and reductive perturbation theory. *Physical Review E*, 54(1):376, 1996.
- [CR79] Stephen L Campbell and Nicholas J Rose. Singular perturbation of autonomous linear systems. *SIAM Journal on Mathematical Analysis*, 10(3):542–551, 1979.
- [dPSD⁺03] David I de Pomerai, Brette Smith, Adam Dawe, Kate North, Tim Smith, David B Archer, Ian R Duce, Donald Jones, and E Peter M Candido. Microwave radiation can alter protein conformation without bulk heating. *FEBS letters*, 543(1):93–97, 2003.
- [DW04] Peter Deuffhard and Marcus Weber. Robust Perron cluster analysis in conformation dynamics. *Linear Algebra Appl.*, 398:161–184, 2004. Special Issue on Matrices and Mathematical Biology.
- [FJK13] Gary Froyland, Oliver Junge, and Péter Koltai. Estimating long term behavior of flows without trajectory integration: the infinitesimal generator approach. *SIAM J. Numer. Anal.*, 51(1):223–247, 2013.
- [Fra69] LE Fraenkel. On the method of matched asymptotic expansions. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 65, pages 233–261. Cambridge University Press, 1969.
- [Jak16] Per Jakobsen. Introduction to the method of multiple scales. *arXiv preprint arXiv:1312.3651*, 2016.
- [Joh05] Robin Stanley Johnson. *Singular perturbation theory: mathematical and analytical techniques with applications to engineering*. Springer Science & Business Media, 2005.
- [KC12] Jirair K Kevorkian and Julian D Cole. *Multiple scale and singular perturbation methods*, volume 114. Springer Science & Business Media, 2012.
- [KCS16] P. Koltai, G. Ciccotti, and C. Schütte. On metastability and Markov state models for non-stationary molecular dynamics. *J. Chem. Phys.*, 145(17):174103, 2016. 2016 Editors’ Choice article.
- [Kue15] Christian Kuehn. *Multiple time scale dynamics*, volume 191. Springer, 1 edition, 2015.
- [KV14] Péter Koltai and Alexander Volf. Optimizing the stable behavior of parameter-dependent dynamical systems — maximal domains of attraction, minimal absorption times. *Journal of Computational Dynamics*, 1(2):339–356, 2014.
- [LeV02] Randall J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.
- [Lic69] Wilbert Lick. Two-variable expansions and singular perturbation problems. *SIAM Journal on Applied Mathematics*, 17(4):815–825, 1969.
- [Lin82] Anders Lindstedt. Abh. K. Akad. Wiss. St. Petersburg, 31(4), 1882.

- [LMHS11] Juan C Latorre, Philipp Metzner, Carsten Hartmann, and Christof Schütte. A structure-preserving numerical discretization of reversible diffusions. *Commun. Math. Sci.*, 9(4):1051–1072, 2011.
- [MS02] Jonathan C Mattingly and Andrew M Stuart. Geometric ergodicity of some hypo-elliptic diffusions for particle motions. *Markov Process. Related Fields*, 8(2):199–214, 2002.
- [MSH02] Jonathan C Mattingly, Andrew M Stuart, and Desmond J Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic processes and their applications*, 101(2):185–232, 2002.
- [NSVE⁺09] F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl. Constructing the full ensemble of folding pathways from short off-equilibrium simulations. *Proc. Natl. Acad. Sci. USA*, 106:19011–19016, 2009.
- [O’M69] Robert E O’Malley. On a boundary value problem for a nonlinear differential equation with a small parameter. *SIAM J. Appl. Math.*, 17(3):569–581, 1969.
- [O’M70] Robert E O’Malley. A non-linear singular perturbation problem arising in the study of chemical flow reactors. *IMA J. Appl. Math.*, 6(31):12–20, 1970.
- [PBB10] V. S. Pande, K. Beauchamp, and G. R. Bowman. Everything you wanted to know about Markov state models but were afraid to ask. *Methods*, 52(1):99–105, 2010.
- [Poi93] Henri Poincaré. *Les méthodes nouvelles de la mécanique céleste: Méthodes de MM. Newcomb, Glydén, Lindstedt et Bohlin. 1893*, volume 2. Gauthier-Villars it fils, 1893.
- [PS08] Grigorios A Pavliotis and Andrew Stuart. *Multiscale methods: averaging and homogenization*. Springer Science & Business Media, 2008.
- [SNS10] Marco Sarich, Frank Noé, and Christof Schütte. On the approximation quality of Markov state models. *Multiscale Modeling & Simulation*, 8(4):1154–1177, 2010.
- [SNW15] Christof Schütte, Adam Nielsen, and Marcus Weber. Markov state models and molecular alchemy. *Molecular Physics*, 113(1):69–78, 2015.
- [SS14] Ch. Schütte and M. Sarich. *Metastability and Markov State Models in Molecular Dynamics: Modeling, Analysis, Algorithmic Approaches*. Courant Lecture Notes. American Mathematical Society/Courant Institute of Mathematical Science, 2014.
- [WS15] Han Wang and Christof Schuette. Building Markov state models for periodically driven non-equilibrium systems. *J. Chem. Theory Comput.*, 11 (4):18191831, 2015.
- [WSCDS14] Han Wang, Christof Schütte, Giovanni Ciccotti, and Luigi Delle Site. Exploring the conformational dynamics of alanine dipeptide in solution subjected to an external electric field: A nonequilibrium molecular dynamics simulation. *Journal of Chemical Theory and Computation*, 10(4):1376–1386, 2014.