A generalized model for the projected dynamics of metastable Markov state models

Jan-Hendrik Prinz,[1] Hao Wu,[1,*] John D. Chodera,[2,†] and Frank Noé[1,‡]

[1]*DFG Research Center Matheon, Free University Berlin, Arnimallee 6, 14195 Berlin, Germany*[§]

[2]*Memorial Sloan-Kettering Cancer Center, Zuckerman Research Institute, Zuckerman Research Center, 417 East 68th Street, New York, NY 10065, US*

(Dated: February 20, 2015)

The application of small-sized Markov state models (MSMs) has become a standard tool for the analysis of the dynamics in molecular dynamics (MD) simulations and computation of key properties. As has been proven before, the predictive quality of the MSMs depends strongly on the selected discretization, which can results in unfeasible parametrizations especially when low-dimensional experimental results are used. For the case of meta-stable dynamics we present a generalized theory to correctly describe the projection of markovian dynamics. This allows a much improved estimation of timescales and eigenvectors projected on the observable space. We demonstrate the presented approach on a numerical example.

I. INTRODUCTION

In recent years Markov State Models (MSMs) have matured into a useful tool for the descriptions of dynamics that shows stochastic, but metastable behavior on the timescales of interest [1]. This includes protein dynamics, chemical [cite], physical systems [cite], etc. Its application spans dynamics that can be written using a differential equation that is first order in time and has also been applied for various types of dynamics, e.g. Brownian dynamics, Langevin dynamics, etc... Once a MSM has been successfully parametrized from data it allows to quickly compute a large variety of target properties like relaxation timescales, mean first passage times, metastable subsets [2, 3] and path related properties like reactive pathways and committor probabilities [4, 5]. The usage of conditional jump probabilities also allows to effectively split the generation of necessary statistics into independent subproblems that require only local equilibrium and can be thus be parallelized as used in adaptive sampling techniques.

A MSM requires a crisp or fuzzy decomposition of the available state space into finite set of states between which the system can jump. If done carefully this can reduce the dominant dynamics to a very small system of transition probabilities. However, the prediction quality crucially depends on the choice of this subsets [6, 7] and the task of finding an optimal decomposition requires full knowledge of the dominant eigenvectors and eigenvalues of the associated propagator. Often this propagator is approximated by a fine discretization that can then be clustered using a metastability anaylsis [2, 3] from the data in a finite timeseries. Even if we have perfect statistics, an optimal decomposition requires access to the dynamics in the full state space which is only available in atomistic simulations and even then usually a dimension reduction is unavoidable to keep the system size tractable. In most feasible cases – especially dynamics observed from experiments – the available timeseries might be deteriorated by stochastic noise and more important a projection onto the observable subspace. In these cases the MSM approach fails since dominant eigenvectors cannot be approximated with sufficient accuracy.

Single-molecule experiments, e.g FRET, AFM and Optical Tweezer [cite examples], provide access to exactly these kind of projected dynamics that can already span much larger timescales than accessible with simulations. To make use of these data a generalization of MSMs for these projected Markovian dynamics is necessary. In this paper we present a novel approach to treat these projected dynamics in the case of metastability correct and overcome some of the previously described deficiencies of MSMs, leading to a deeper understanding of the projection error.

II. THEORY

Let $\{x_t \mid x_t \in \Omega, t \in \mathbb{R}^+\}$ be the (stochastic) process of interest in full state space Ω which we require to be Markovian, time-homogeneous, ergodic and reversible wrt. its unique stationary distribution μ. For a time-homogeneous Markov process the time evolution of the probability density $p_t(x) \equiv \mathbb{P}(x_t = x)$ can be expressed using a time-step τ (also called lag time) dependent propagator $\mathcal{P}_\tau$ by [8, 9]

$$\begin{aligned} p_t(x) &= \mathcal{P}_\tau \left[p_t\right](x) \\ &\equiv \int_\Omega \mathrm{d}x' \, \mathbb{P}\left[x_{t+\tau} = x \mid x_t = x'\right] p_t(x'). \end{aligned}$$

Ergodicity and reversibility of $\{x_t\}$ then implies that $\mathcal{P}_\tau$ is compact and self-adjoint wrt. the stationary density μ on a Hilbert space defined by the inner product

$$\langle u, v \rangle \equiv \int_\Omega \mathrm{d}x \, u(x) v(x) \mu^{-1}(x)$$

and we can expand the propagator into a countable sum

$$\mathcal{P}_\tau p_t = \sum_{k=1}^{\infty} \lambda_i(\tau) \phi_i \langle \phi_i, p_t \rangle$$

* hao.wu@fu-berlin.de
† choderaj@mskcc.org
‡ frank.noe@fu-berlin.de
§ jan-hendrik.prinz@fu-berlin.de

of orthogonal eigenfunction/eigenvalue pairs $\{(\lambda_i(\tau), \phi_i) \mid i \in \mathbb{N}\}$ with $\langle \phi_i, \phi_i \rangle = \delta_{ij}$ and $\|\lambda_i\| \leq 1$. For convenience we order all pairs with descending magnitude of the eigenvalue where the uniqueness of the stationary distribution ensures that the eigenvalue of one λ_1 is single and so $\phi_1 = \mu$. The semi-group property of a markov chain also implies a relation $\lambda_i(k\tau) = \lambda_i^k(\tau)$ s.t. we can define $\lambda_i^\tau \equiv \lambda_i^\tau(1)$ as the eigenvalues at the native lag time $\tau = 1$.

To describe metastability which we required we use that metastability leads to gaps in the spectrum of the propagator [10] where the finite set of m eigenvalue/eigenfunction pairs above the gap (being slower) describe the metastable dynamics while the remaining faster pairs describe the fast dynamics within metastable sets. That these separation is possible and the dynamics do not mix follows from the orthogonality of the eigenfunctions. Thus we can split the propagator

$$\mathcal{P}_\tau = \mathcal{P}_\tau^{\text{meta}} + \mathcal{P}_\tau^{\text{fast}} \\ \sum_{k=1}^{m} \lambda_i(\tau)\phi_i\langle\phi_i, \cdot\rangle + \mathcal{P}_\tau^{\text{fast}} \tag{1}$$

with $\|\mathcal{P}_\tau^{\text{fast}}\| \in \mathcal{O}\left(\lambda_{m+1}^\tau\right)$ and conclusively, if $\tau \gg -1/\log\lambda_{m+1}$ the fast part of $\mathcal{P}_\tau$ has decayed and the dynamics can be well approximated by the metastable part $\mathcal{P}_\tau \approx \mathcal{P}_\tau^{\text{meta}}$. In this paper we will be dealing with this type of dynamics and call this an "m-metastable" system.

Usually, the full dynamics is not directly observable, either by an artificially induced discretization, the experimental setup or unavoidable sideeffects like noise. Instead we have to deal with a projected process $\{y_t \mid y_t \in \Upsilon, t \in \mathbb{R}^+\}$ on an observable space $\Upsilon = \{1, \ldots, K\}$ with K states which we choose to be finite. To describe the probabilistic relation between x_t and y_t we use a output probability functions χ_i so that

$$\chi_k(x) \equiv \mathbb{P}\left(y_t = k \mid x_t = x\right)$$

which explicitely assumes that the output probabilities are time independent and also contain no memory s.t. each state $x \in \Omega$ in the full state space always maps into Υ in the same way. The combination of the reversible Markov state model and output probabilities will be called a *projected Markov model* (PMM), if in addition the full state dynamics is m-metastable we will refer to m-metastable PMM or short m-PMM. Note, that $\{y_t\}$ is *not* a Markov chain anymore unless the output probability functions exactly span the space of the m dominant eigenfunctions, which is either impossible due to experimental constraints or requires explicit knowledge of the propagator. Therefore the popular modelling using a Markov state model (MSM) is often only approximate [7].

To capture the dynamics of a PMM we compute the state-to-state correlation matrix $C_{ij}(\tau)$ in the observable space

$$C_{ij}(\tau) \equiv \mathbb{P}\left[y_{t+\tau} = j \wedge y_t = i\right] \\ = \iint_{\Omega^2} \mathrm{d}x\, \mathrm{d}x'\, \chi_i(x)\mathbb{P}\left[x_{t+\tau} = x \wedge x_t = x'\right]\chi_j(x')$$

for the stochstic process y_t which implicitely assumes infinite statistics. In this case the expression can be rewritten as

$$C_{ij}(\tau) = \sum_{k=1}^{\infty} \lambda_i(\tau) \int_\Omega \mathrm{d}x\, \chi_i(x)\phi_k(x) \int_\Omega \mathrm{d}x'\, \phi_k(x')\chi_j(x') \\ = \sum_{k=1}^{\infty} \lambda_i(\tau) Q_{ki}\, Q_{kj}$$

with the matrix Q of *projected eigenfunctions*

$$Q_{ki} \equiv \int_\Omega \mathrm{d}x\, \chi_i(x)\phi_k(x).$$

Using the diagonal matrix of eigenvalues $\Lambda \equiv \mathrm{diag}\,(\lambda_1, \ldots)$ we can finally write

$$C(\tau) = Q^\mathsf{T}\Lambda(\tau)Q$$

which is exact for any PMM but might require countable, but infinite sized Q and Λ. For a m-PMM the metastability reduces this size to $Q \in \mathbb{R}^{m\times K}$ and $\Lambda \in \mathbb{R}^{m\times m}$ which is much more tractable. We conclude that all correlation functions on the observable space of a m-PMM can exactly be represented by the parametrization into Q and Λ.

III. SPECTRAL ESTIMATION

The goal will be to construct a learning algorithm to estimate Q and Λ from a given observation $\{y_t\}$. This will allow us to correctly estimate all dominant eigenvalues λ_i of the full propagator. As input for the learning algorithm we use the state-to-state correlation matrices $C(\tau)$, which can easily be estimated from a finite timeseries. Since $C(\tau)$ is by construction symmetric [1], the number of independent variables in a single $C(\tau)$ is smaller than the variables to be parametrized in Q and $\Lambda(\tau)$. Conclusively, a single $C(\tau)$ does not contain enough information to allow for a unique decomposition of a given $C(\tau)$ into Q and Λ. In MSMs this lack of information is filled by an orthogonality constraint

$$Q^T Q = \mathrm{diag}(\pi)$$

wrt. the observed stationary distribution π. Other decompositions of this type include the singular value decomposition for symmetric matrices (SVD) or, equivalently, the eigenvalue decomposition (EVD) which both assume orthogonality $Q^T Q = \mathrm{Id}$ or the Cholesky decomposition that instead uses a triangular shaped Q. It can be shown (see SI), that two correlation matrices at two different lagtimes τ_1 and τ_2 provide enough information to solve this problem uniquely in the case of an m-PMM and positive eigenvalues.

[1] The real $C(\tau)$ is symmetric for reversible dynamics, while an estimated $C(\tau)$ might not, depending on the estimation procedure. Also if the the used data is finite and is not sampled from global equilibrium the estimated $C(\tau)$ is not symmetric!

We will first address the special case where the number of dominant eigenvalues m equals the number of observable states K so that Q is square. This case can always be achieved by an additional projection onto m states provided that initially the number of observed states is larger than the number of processes $K > m$. The problem then reduces to solving the generalized hermitian eigenvalue problem (GHEP)

$$C(\tau_1)u_i = \nu_i C(\tau_2)u_i$$

for the generalized eigenvalues ν and eigenvectors u_i that are related to the dominant eigenvalues by

$$\lambda_i = \nu_i^{1/(\tau_1-\tau_2)}$$

to compute the matrix Q of projected eigenvectors by

$$Q = U^{-1}$$

where $U = [u_1, \ldots, u_m]$ is the matrix of the generalized eigenvectors u_i that are normalized so that

$$\|u_i\|^2 = \lambda_i^{\tau_1}\left(u_i^T C(\tau_1)u_i\right)^{-1}$$

holds. Note, that a GHEP can also be solved using an inversion of one matrix, but this is not necessary. Also this transforms the problem into a non-symmetric problem and the inversion might be badly conditioned. Nonetheless, we can compute the solutions using an ordinary EVD of $C^{-1}(\tau_2)C(\tau_2)$.

IV. LOW-RANK APPROXIMATION

In most cases we are only interested in the dominant processes that govern the dynamics as stated in Eq. 1. Since the spectral estimation method is (almost) independent of the projection we can project onto a set of N states, then apply the spectral estimation method and retrieve the dominant timescales. If we also want to reconstruct the dominant processes on the complete observable space we need to use a generating set of vectors for the space spanned by the projected left eigenvectors. Since we have no direct access to the eigenvectors we can use the singular value decomposition

$$C(\tau) \to V(\tau)\Sigma(\tau)V^T(\tau) \tag{2}$$

which for a symmetric matrix provides for a decomposition into a matrix V and a diagonal matrix Σ. If for a particular $\tau > \tau_{\text{min}}$ only m singular values exist then V must span the same subspace of Υ as Q which follows from rank considerations. As stated before, the decomposition in Eq. 2 is not unique and the SVD requires $V^TV = \text{Id}$ which is not the correct solution. The most general form the correct solution can take is $Q = AV$ with A being an $m \times m$ matrix that captures the linear recombination and a second correlation matrix $C(\tau_2)$ makes the decomposition unique. Finally, we propose the following procedure to compute the dominant Q and Λ

1. Estimate a reasonable range of lagtimes using the singular value decomposition (SVD) which hints to the number N of processes needed to approximate the dynamics,
2. Select two lagtimes τ_1 and τ_2 which lie in the range that at least m singular values are non-zero.
3. Compute the two symmetric correlation matrices $C(\tau_1)$ and $C(\tau_2)$
4. Compute the low-rank approximation of rank m using the SVD of $C(\tau_1) = V\Sigma V^T$ with $V \in \mathbb{R}^{K\times m}$ and $\Lambda \in \mathbb{R}^{m\times m}$
5. Compute the projections
$$B(\tau) \equiv V^T C(\tau_1)V$$
onto the sub-space spanned by the dominant m singular vectors for both lagtimes $B_1 \equiv B(\tau_1) = \Sigma$, $B_2 \equiv B(\tau_2)$
6. Compute the spectral decomposition (SD) by solving the GHEP $B_1u_i = \nu B_2u_i$ as given above to get A and Λ so that we can write $B(\tau) = A^T\Lambda(\tau)A$.
7. The final approximation is given by $Q = V\,A^T$ and Λ.

V. EXAMPLE

We demonstrate the improved estimation qualities of the new method at a simple 2D-model that represents a case where MSMs fail. Fig. V shows the stationary distribution of the overdamped langevin process in a potential U at inverse temperature $\beta = 1$. The system is designed to consist of 4 distinct metastable states labeled A to D in Fig. V. The process $x_t \in \Omega$ in the full state space is also projected onto the second (y) coordinate which is then clustered onto 28 equidistant bins ($\Delta x = 0.05$) as y_t

$$x_t \in \Omega \mapsto x_t^{(2)} \mapsto y_t \in \Upsilon = \{1, \ldots, 28\}$$

where the bins comprise the observable space $\Upsilon = \{1, \ldots, 28\}$. For the computation of reference values we also discretized the system into a set of 28×28 equidistant bins while the SE and MSM is applied to the observed space only. Two cases of data are considered (I) infinite statistics $L \to \infty$ and (II) a timeseries of length $L = 10^6$ which is long compared to the slowest relaxation timescale of $t_2 = 382.4$ and can thus be considered a case of converged and good, but finite statistics.

Fig. V shows the singular values of $C(t)$ computed from the timeseries and show the presence of the 4 distinct timescales as anticipated from the shape of the stationary distribution. The plot allows to estimate that for lagtimes $\tau > 125$ the low-rank approximation of the 4 dominant processes (eigenvectors) is best.

Fig. V presents the estimation results (see caption for details) indicating the the SE estimator allows for a much faster

convergence of the estimated timescales to the true values in comparison to the MSM approach. It is also able to much more accurately predict the non-orthogonal projected eigenvectors Q.

VI. MULTI-LAGTIME ESTIMATION

To address the issue of statistical uncertainties we rewrite the problem into an optimization problem. We use the list of time scales used in the estimation $\mathcal{T} = \{\tau_1, \tau_2\}$ and write

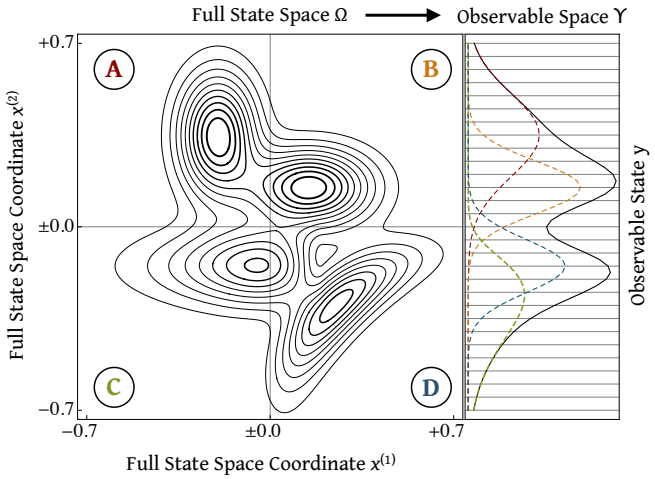
$$\{Q, \Lambda\} = \underset{Q,\Lambda}{\operatorname{argmin}} \sum_{\tau \in \mathcal{T}} \|C(\tau) - Q^{\mathsf{T}} \Lambda(\tau) Q\|_{\mathsf{F}}$$

which is equivalent of estimation the parameters of a PMM as described above. This can be generalized to more than 2 lagtimes by expanding $\mathcal{T}$. Fig. VI shows estimation results from the multi-τ estimation for the 2D example from Fig. V. For this we simulated 100 trajectories starting in the center of state B for 6 different trajectory lengths $L \in \{50k, 80k, 100k, 150k, 200k, 500k\}$. With increasing length L the distribution becomes more narrow providing a more reliable estimate. The MSM estimate (lighter colors) show a systematically too small estimate compared to the spectral estimations (darker colors).

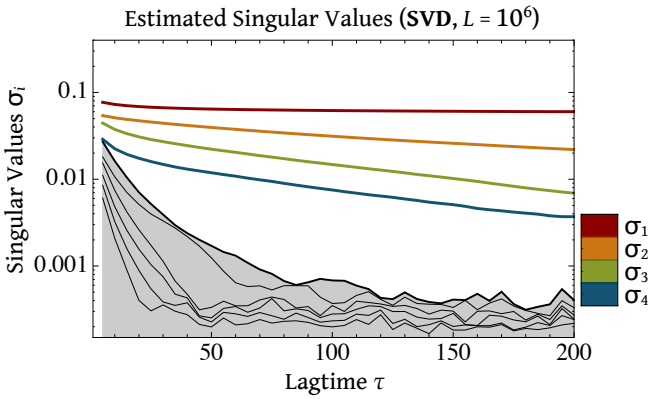
VII. APO-MYOGLOBIN EXAMPLE

We demonstrate the method on the Apo-Myoglobin. [Copy experimental details from PRX publication or give a reference].

In order to illustrate the performance of spectral estimation on real data, it is applied to optical tweezer measurements of the extension fluctuations of a biomolecule examined in a recent optical force spectroscopy study: the H36Q mutant of sperm whale apo-myoglobin at low pH [11]. The apo-myoglobin [crystal structure shown in Fig. 3(4)] hops between unfolded and molten globule states at the experimental pH. Experimental force trajectory data were generously provided by the authors of Refs. [11]. Experimental details are given therein, but we briefly summarize aspects of the apparatus and experimental data collection procedure relevant to our analysis. The instrument used to collect the data sets was a dual-beam counterpropagating optical trap [12]. The molecule of interest was tethered to polystyrene beads by means of dsDNA handles, with one bead suctioned onto a pipette and the other held in the optical trap. A piezoactuator controlled the position of the trap and allowed position resolution to within 0.5 nm, with the instrument operated in passive (equilibrium) mode such that the trap was stationary relative to the pipette during data collection. The force on the bead held in the optical trap was recorded at 50 kHz, with each recorded force trajectory 60 s in duration resulting in $3 \cdot 10^6$ data points each. From the set of presentes fibers we chose to analyze fiber no. 6. [check how this is mentioned in Ref. [11]]



Stationary probability. Energy Landscape of illustrative 2D potential with 4 distinct basins (metastable states) A, B, C and D. The dynamics in full space $x_t \in \Omega$ is mapped onto coordinate $x^{(2)}$ and discretized into 28 bins constituting the observable space $y_t \in \Upsilon$

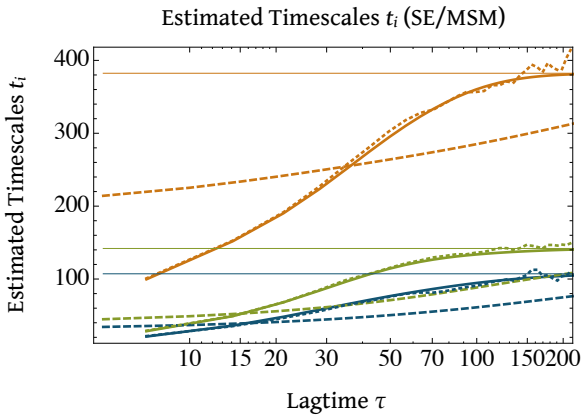


Estimated Singular values. Singular values of the correlation matrices $C(\tau)$ symmetrically estimated from a trajectory $(L = 10^6)$ in the 2D illustrative potential. Even for small lag times τ four disctint singular values are present. For lagtimes $\tau \gtrapprox 150$ all but the 4 dominant singular values have reached noise level and remain approximately constant.

Fig. VII shows estimation results for $\mathcal{T} = \{150, \ldots, 500\}$ and computing the projective SVD also at $\tau = 150$. The estimations show a consistent estimation of two processes [maybe we can find a fourth one for 15-17? Actually I checked and it looks like it, but haven't finished this yet.] For #IDs 15-17 the blue process seems to switch to a different one at lower bead extensions.

VIII. CONCLUSION

We presented a novel approach to estimate the parameters Q and Λ of a projected Markov state model (PMM). The method has been demonstated at a simple illustrative model and shows much improved convergence in the estimation of the dominant timescales as well as related pro-



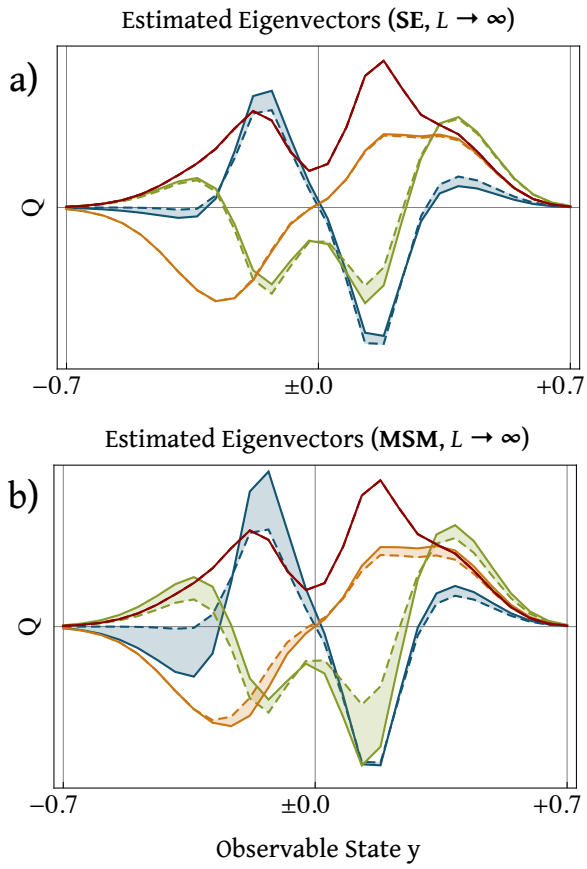
Estimated timescales Various estimation results for a simple diffusion process in a potential defined by the stationary distribution in Fig. V. Comparison of timescales estimation for t_2 (yellow), t_3 (green) and t_4 (blue) in 3 cases: (thin solid and constant) exact solution from 28x28 bin discretization of full space, (thick, solid) *spectral estimate* (SE) for infinite statistics and using $\tau_2 = 2\tau_1$, (dashed) SE estimate for finite statistics ($L = 10^6$) using $\tau_2 = 2\tau_1$ and (dashed) MSM estimate using single τ. The corresponding arrows in the middle plot indicate locations of the the used lagtimes τ_1, τ_2. The SE estimates converge much faster ($\tau \approx 200$) compared to the MSM method.

jected eigenvectors. Although the sensitivity to statistical noise has increased the method still provides better results compared to the MSM approach. The sensitivity can be reduced if the estimation is reformulated into an optimization problem that fits the result at more than two lagtimes. We have applied the optimizational approach to data from Apo-Myoglobin and showed a consistent set of two processes (maybe four). On first sight the estimation for m-PMMs looks identical to and m-state HMM, but it can be shown that HMMs are not able to capture all possible dynamics of m-PMMs and HMMs can model a subset of the dynamics of PMMs. Thus the HMM solution might be close, but rarely exact[13]. Finally, the cases of low-statistics and non-reversible dynamics is interesting and subject of active research.

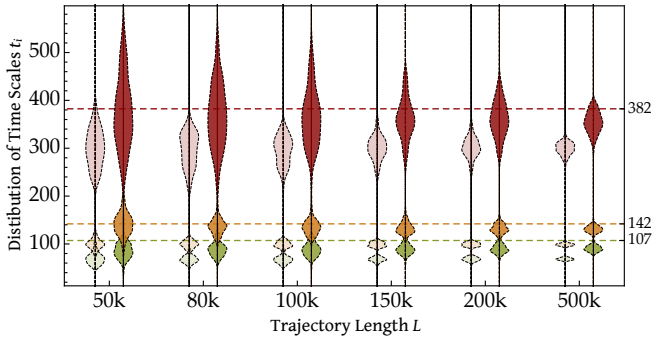
IX. ACKNOWLEDGMENTS

The authors would like to thank Ralf Banisch for stimulating discussions on the topic. JHP acknowledges support from the DFG research center MATHEON. HW [missing] . FN acknowledges funding from DFG grant NO 825/2-2 and ERC starting grant 'pcCell' ". JDC acknowledges funding from [@JDC: missing]

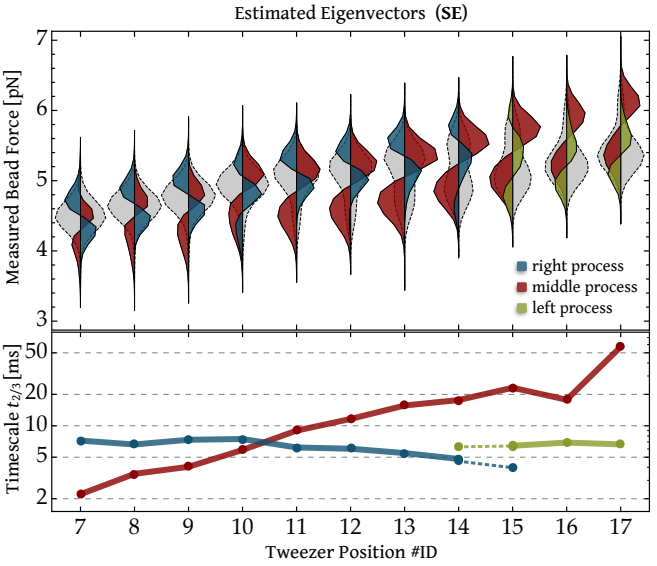
[1] J.-H. Prinz, H. Wu, M. Sarich, B. G. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé, J. Chem. Phys. **134**, 174105 (2011).
[2] S. Röblitz and M. Weber, Adv Data Anal Classif **7**, 147 (2013).
[3] P. Deuflhard and M. Weber, ZIB Report **09** (2003).
[4] E. Vanden-Eijnden, *Transition Path Theory*, Vol. 703 (Springer Verlag, 2006).
[5] P. Metzner, C. Schütte, and E. Vanden-Eijnden, Multiscale Model. Sim. **7**, 1192 (2009).
[6] N. Djurdjevac, M. Sarich, and C. Schütte, Multiscale Model. Sim. **10**, 61 (2012).
[7] M. Sarich, F. Noé, and C. Schütte, Multiscale Model. Sim. **8**, 1154 (2010).
[8] C. Schütte, A. Fischer, W. Huisinga, and P. Deuflhard, J. Comp. Phys. **151**, 146 (1999).
[9] C. Schütte and W. Huisinga, *Biomolecular Conformations can be Identified as Metastable Sets of Molecular Dynamics* (2003).
[10] A. Bovier, V. Gayrard, and M. Klein, J. Eur. Math. Soc. **76**, 69 (2005).
[11] P. J. Elms, J. D. Chodera, C. Bustamante, and S. Marqusee, Proc. Nat. Acad. Sci. USA **109**, 3796 (2012).
[12] C. Bustamante and S. B. Smith, United States Patent No. 7133132 (2006).
[13] F. Noé, H. Wu, J.-H. Prinz, and N. Plattner, submitted to J. Chem. Phys. (2013).



Estimated projected eigenvectors Q. (***a***) SE at $\tau_1 = 125$, $\tau_2 = 175$ using infinite statistics ($L = \infty$), (**b**) MSM at $\tau = 175$ estimate using infinite statistics. The (**solid**) lines correspond to the estimation, while the (**dashed**) lines are the reference solution. The smaller the colored areas, the better the approximation. Although both approaches correctly predict the stationary distribution, the MSM estimate performs worse than the SE estimate. SE can almost predict the slowest process (λ_2, yellow) correctly while process 3 (green) and 4 (blue) are less accurate.



Statistical Uncertainty Variance in the estimation of timescales due to statistical variation for a timeseries in the potential of Fig. V. All timeseries start in the center of state B. Shown are the distributions of the estimated time scales t_i of 100 trajectories for 6 different trajectory lengths $L \in \{50k, 80k, 100k, 150k, 200k, 500k\}$. Colors indicate the processes (red: slowest, yellow: 2nd slowest, green:3rd slowest). Darker colors refer to the result of Spectral Estimation with $\mathcal{T} = \{50, \ldots, 150\}$ and lighter color is the MSM estimation of a model computed for $\tau = 150$. With incresing trajectory length the estimation becomes more accurate. While the MSM estimation is sharper the estimation is significantly less accurate. Both methods seem to converge to underestimated timescales, as the theory predicts.



Apo-Myoglobin estimation results Results for a series of 11 timeseries recorded in a optical tweezer experiment of Apo-Myoglobin [cite ElmsEtAl, fiber 6]. **(upper)** 3 dominant eigenvectors, which have been assigned according to their overall shape (**gray**: stationary distribution, **red**: intermediate process, **blue**: higher force process, **green**: lower force process). The shift in measured bead position is due to an increase in the pulling force, moving the whole protein toward the force exerting bead. **(lower)** Time scales for the slowest processes. Colors are the same as in the upper part. The smaller extended process shows an almost constant time scale while the other process gets slower with increasing force [check if this is consistent with prior publication]. For positions 14 and 15 also the estimation with 4 states produced reasonable results which are shown using dashed lines.