

## Variational Koopman models: Slow collective variables and molecular kinetics from short off-equilibrium simulations

Hao Wu, Feliks Nüske, Fabian Paul, Stefan Klus, Péter Koltai, and Frank Noé

Citation: *The Journal of Chemical Physics* **146**, 154104 (2017); doi: 10.1063/1.4979344

View online: <http://dx.doi.org/10.1063/1.4979344>

View Table of Contents: <http://aip.scitation.org/toc/jcp/146/15>

Published by the [American Institute of Physics](#)

---

### Articles you may be interested in

[Introducing the mean field approximation to CDFT/MMpol method: Statistically converged equilibrium and nonequilibrium free energy calculation for electron transfer reactions in condensed phases](#)

*The Journal of Chemical Physics* **146**, 154101 (2017); 10.1063/1.4979895

[Perspective: Dissipative particle dynamics](#)

*The Journal of Chemical Physics* **146**, 150901 (2017); 10.1063/1.4979514

[Markov state models from short non-equilibrium simulations—Analysis and correction of estimation bias](#)

*The Journal of Chemical Physics* **146**, 094104 (2017); 10.1063/1.4976518

[Tests for, origins of, and corrections to non-Gaussian statistics. The dipole-flip model](#)

*The Journal of Chemical Physics* **146**, 154109 (2017); 10.1063/1.4981009

[Equation-of-motion coupled-cluster methods for atoms and molecules in strong magnetic fields](#)

*The Journal of Chemical Physics* **146**, 154105 (2017); 10.1063/1.4979624

[Incorporating nuclear vibrational energies into the “atom in molecules” analysis: An analytical study](#)

*The Journal of Chemical Physics* **146**, 154106 (2017); 10.1063/1.4979994

---



**COMPLETELY  
REDESIGNED!**

**PHYSICS  
TODAY**

*Physics Today* Buyer's Guide  
Search with a purpose.

# Variational Koopman models: Slow collective variables and molecular kinetics from short off-equilibrium simulations

Hao Wu,<sup>a),b)</sup> Feliks Nüske,<sup>a),c)</sup> Fabian Paul,<sup>d)</sup> Stefan Klus,<sup>e)</sup> Péter Koltai,<sup>f)</sup> and Frank Noé<sup>g)</sup>  
 Department of Mathematics and Computer Science, Freie Universität Berlin, Arnimallee 6,  
 14195 Berlin, Germany

(Received 13 September 2016; accepted 15 March 2017; published online 17 April 2017)

Markov state models (MSMs) and master equation models are popular approaches to approximate molecular kinetics, equilibria, metastable states, and reaction coordinates in terms of a state space discretization usually obtained by clustering. Recently, a powerful generalization of MSMs has been introduced, the variational approach conformation dynamics/molecular kinetics (VAC) and its special case the time-lagged independent component analysis (TICA), which allow us to approximate slow collective variables and molecular kinetics by linear combinations of smooth basis functions or order parameters. While it is known how to estimate MSMs from trajectories whose starting points are not sampled from an equilibrium ensemble, this has not yet been the case for TICA and the VAC. Previous estimates from short trajectories have been strongly biased and thus not variationally optimal. Here, we employ the Koopman operator theory and the ideas from dynamic mode decomposition to extend the VAC and TICA to non-equilibrium data. The main insight is that the VAC and TICA provide a coefficient matrix that we call Koopman model, as it approximates the underlying dynamical (Koopman) operator in conjunction with the basis set used. This Koopman model can be used to compute a stationary vector to reweight the data to equilibrium. From such a Koopman-reweighted sample, equilibrium expectation values and variationally optimal reversible Koopman models can be constructed even with short simulations. The Koopman model can be used to propagate densities, and its eigenvalue decomposition provides estimates of relaxation time scales and slow collective variables for dimension reduction. Koopman models are generalizations of Markov state models, TICA, and the linear VAC and allow molecular kinetics to be described without a cluster discretization. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4979344>]

## I. INTRODUCTION

With the ability to generate extensive and high-throughput molecular dynamics (MD) simulations,<sup>1–9</sup> the spontaneous sampling of rare-events such as protein folding, conformational changes, and protein-ligand association has become accessible.<sup>10–17</sup> Markov state models (MSMs),<sup>18–25</sup> master-equation models,<sup>26–28</sup> and closely related approaches<sup>29–33</sup> have emerged as powerful frameworks for the analysis of extensive MD simulation data. These methods do not require a very specific *a priori* definition of relevant reaction coordinates.<sup>23,34</sup> Furthermore, they allow a large variety of mechanistic information to be extracted,<sup>10,35,36</sup> and experimental observables to be computed and structurally interpreted.<sup>12,28,37–40</sup> Finally, they provide a direct approximation of the dynamic modes describing the slow conformational changes that are identical or closely related to the so-called reaction coordinates, depending on which notion of that term is employed.<sup>41–46</sup> An especially powerful feature

of MSMs and similar approaches is that the transition probabilities  $p_{ij}(\tau)$ , i.e., the probability that the trajectory is found in a set  $A_j$  a time lag  $\tau$  after which it has been found in a set  $A_i$ ,

$$p_{ij}(\tau) = \text{Prob} [\mathbf{x}_{t+\tau} \in A_j \mid \mathbf{x}_t \in A_i],$$

is a conditional transition probability.  $p_{ij}(\tau)$  can be estimated without bias even if the trajectory is not initiated from a global, but only a local equilibrium distribution.<sup>23</sup> Consequently, given  $c_{ij}(\tau)$  transition events between states  $i$  and  $j$  at lag time  $\tau$ , the maximum likelihood estimator of the transition probability can be easily shown to be

$$p_{ij}(\tau) = \frac{c_{ij}(\tau)}{\sum_k c_{ik}(\tau)}, \quad (1)$$

i.e., the fraction of the number of transitions to  $j$  conditioned on starting in  $i$ . This conditionality is a key reason why MSMs have become popular to analyze short distributed simulations that are started from arbitrary configurations whose relationship to the equilibrium distribution is initially unknown.

However, when estimating (1) from simulation data, one does not generally obtain a time-reversible estimate, i.e., the stationary probabilities of the transition matrix,  $\pi_i$ , will usually not fulfill the detailed balance equations  $\pi_i p_{ij} = \pi_j p_{ji}$ , even if the underlying dynamics are microscopically time-reversible.

<sup>a)</sup>H. Wu and F. Nüske contributed equally to this work.

<sup>b)</sup>hao.wu@fu-berlin.de

<sup>c)</sup>feliks.nueske@fu-berlin.de

<sup>d)</sup>fab@zedat.fu-berlin.de

<sup>e)</sup>stefan.klus@fu-berlin.de

<sup>f)</sup>peter.koltai@fu-berlin.de

<sup>g)</sup>frank.noel@fu-berlin.de

Compared to a reversible transition matrix, a transition matrix with independent estimates of  $p_{ij}$  and  $p_{ji}$  has more free parameters, resulting in larger statistical uncertainties, and may possess complex-valued eigenvalues and eigenvectors, which limits the application of some analysis tools designed for equilibrium molecular processes.<sup>47</sup> Since most molecular dynamics simulations are in thermal equilibrium and thus fulfill at least a generalized microscopic reversibility (Appendix B in Ref. 48), it is desirable to force  $p_{ij}$  to fulfill detailed balance, which both reduces statistical uncertainty and enforces a real-valued spectrum.<sup>47,49</sup> In old studies, the pragmatic solution to this problem was often to symmetrize the count matrix, i.e., to simply set  $c_{ij}^{\text{sym}} = c_{ij} + c_{ji}$ , which is equivalent to evaluating the simulation trajectory forward and backward, and which leads to a transition matrix with detailed balance when inserted into (1). However, it has been known since at least 2008 that this estimator is strongly biased, and therefore reversible maximum likelihood and Bayesian estimators have been developed.<sup>22,23,28,47,49,50</sup> These algorithms formulate the estimation problem as an optimization or sampling problem of the transition matrix constrained to fulfill detailed balance. The idea of these algorithms becomes clear when writing the reversible maximum likelihood estimator in two subsequent steps, as demonstrated in Ref. 47:

1. *Reweighting*: Estimate the stationary distribution  $\pi_i$  given all transition counts  $c_{ij}$  under the detailed balance condition.
2. *Estimation*: Insert  $\pi_i$  and  $c_{ij}$  into an equation for the reversible transition matrix to obtain a maximum likelihood estimate of  $p_{ij}$ .

Recently, a powerful extension to the Markov modeling framework has been introduced: the variational approach (VA) to approximate the slow components of reversible Markov processes.<sup>51</sup> Due to its relevance for molecular dynamics, it has also been referred to as VA for molecular kinetics<sup>52,53</sup> or VA for conformation dynamics (VAC).<sup>52,54</sup> It has been known for many years that Markov state models are good approximations to molecular kinetics if their largest eigenvalues and eigenvectors approximate the eigenvalues and eigenvectors of the Markov operator governing the full-phase space dynamics;<sup>18,34,55</sup> moreover, the first few eigenvalues and eigenvectors are sufficient to compute almost all stationary and kinetic quantities of interest.<sup>37,38,56–58</sup> The VAC has generalized this idea beyond discrete states and formulated the approximation problem of molecular kinetics in terms of an approach that is similar to the variational approach in quantum mechanics.<sup>51–53</sup> It is based on the following variational principle: If we are given a set of  $n$  orthogonal functions of state space and evaluate the autocorrelations of the molecular dynamics in these functions at lag time  $\tau$ , these will give us lower bounds to the true eigenvalues  $\lambda_1(\tau), \dots, \lambda_n(\tau)$  of the Markov operator, equivalent to an underestimate of relaxation time scales and an overestimate of relaxation rates. Only if the  $n$  functions used are the eigenfunctions themselves, then their autocorrelations will be maximal and identical to the true eigenvalues  $\lambda_1(\tau), \dots, \lambda_n(\tau)$ . Note that this statement is true in the correct

statistical limit—for finite data, the variational bound can be violated by problems in the estimation procedure. Sources of violation include systematic estimator bias, which is addressed in this work, and overfitting, which can be addressed by cross-validation.<sup>59</sup>

This principle allows us to formulate variational optimization algorithms to approximate the eigenvalues and eigenfunctions of the Markov operator. The linear VAC proceeds as follows:

1. Fix an arbitrary basis set  $\chi = [\chi_1(\mathbf{x}), \dots, \chi_n(\mathbf{x})]$  and evaluate the values of all basis functions for all sampled MD configurations  $\mathbf{x}$ .
2. Estimate two covariance matrices: the covariance matrix  $\mathbf{C}(0)$  and the time-lagged covariance matrix  $\mathbf{C}(\tau)$  from the basis-set-transformed data.
3. Solve a generalized eigenvalue problem involving both  $\mathbf{C}(0)$  and  $\mathbf{C}(\tau)$ , and obtain estimates for the eigenvalues and the optimal representation of eigenfunctions as a linear combination of basis functions.

Note that the functions  $\chi$  can be arbitrary nonlinear functions in the original coordinates  $\mathbf{x}$ , which allows complex nonlinear dynamics to be encoded even within this linear optimization framework. The variational approach has spawned a variety of follow-up works, for example, it has been shown that the algorithm called blind source separation, time-lagged, or time-structure based independent component analysis (TICA) established in signal processing and machine learning<sup>60–62</sup> is a special case of the VAC.<sup>52</sup> TICA is now widely used in order to reduce the dimensionality of MD data sets to a few slow collective coordinates, in which MSMs and other kinetic models can be built efficiently.<sup>52,63,64</sup> The VAC has been used to generate and improve guesses of collective reaction coordinates.<sup>46,65</sup> A VAC-based metric has been defined which transforms the simulation data into a space in which Euclidean distance corresponds to kinetic distance.<sup>66,67</sup> The importance of meaningful basis sets has been discussed, and a basis for peptide dynamics has been proposed in Ref. 54. Kernel versions of TICA have been proposed<sup>68,69</sup> and nonlinear deep versions have been proposed based on tensor approximations of products of simple basis functions.<sup>70</sup> Finally, the variational principle ranks kinetic models by the magnitude of their largest eigenvalues or derived quantities,<sup>51</sup> which can be used to select hyperparameters such as the basis functions  $\chi$ , or the number of states in a Markov state model.<sup>59,71</sup>

Despite the popularity of the VAC and TICA, their estimation from MD data is still in the stage that MSMs had been about a decade ago: A direct estimation of covariance matrices will generally provide a non-symmetric  $\mathbf{C}(\tau)$  matrix and complex eigenvalues/eigenfunction estimates that are not consistent with reversible molecular dynamics. In order to avoid this problem, the current state of the art is to enforce the symmetrization of covariance matrices directly.<sup>52,63,69</sup> In lack of a better estimator, this approach is currently used also with short MD simulations from distributed computing despite the fact that the resulting time scales and eigenfunctions may be biased and misleading. This problem is addressed in the present paper.

The algorithm of the linear VAC<sup>51,53</sup> is identical to more recently proposed extended dynamic mode decomposition (EDMD),<sup>72</sup> which is based on dynamic mode decomposition (DMD).<sup>73–76</sup> However, while the VAC has been derived for reversible and stationary dynamics, EDMD has been developed in the context of dynamical systems and fluid mechanics, where data are often nonreversible and non-stationary. Mathematically, these methods are based on the eigenvalue decomposition of the Koopman operator, which provides a theoretical description of non-stationary and non-equilibrium dynamics.<sup>77,78</sup> In the present paper, this theory is used in order to formulate robust equilibrium estimators for covariance matrices, even if the simulation data are generated in many short simulations that are not distributed according to equilibrium. Based on these estimates, a Koopman model is computed—a matrix model that approximates the dynamics of the Koopman operator on the basis functions used. Koopman models are proper generalizations of Markov state models—they do not rely on a state space clustering, but can still be used to propagate densities in time, and their eigenvalues and eigenvectors provide estimates of the equilibrium relaxation time scales and slow collective variables. We propose a reversible Koopman model estimator that proceeds analogously to reversible MSM estimation:

1. *Reweighting*: Estimate a reweighting vector  $u_i$  with an entry for each basis function given the covariance matrices  $\hat{\mathbf{C}}(0)$  and  $\hat{\mathbf{C}}(\tau)$  that have been empirically estimated without symmetrization.
2. *Estimation*: Insert  $u_i$  and  $\hat{\mathbf{C}}(0)$ ,  $\hat{\mathbf{C}}(\tau)$  into an equation for the symmetric equilibrium estimates of  $\mathbf{C}(0)$  and  $\mathbf{C}(\tau)$ . Then compute a Koopman model, and from its eigenvalue decomposition the relaxation time scales and slow collective variables.

In addition to this result, the reweighting vector  $u_i$  allows us to approximate *any* equilibrium estimate in terms of a linear combination of our basis functions from off-equilibrium data. The estimator is asymptotically unbiased in the limits of many short trajectories and an accurate basis set.

The new methods are illustrated on toy examples with stochastic dynamics and a benchmark protein-ligand binding problem. The methods described in this article are implemented in PyEMMA version 2.3 or later ([www.pyemma.org](http://www.pyemma.org)).<sup>71</sup>

## II. VARIATIONAL APPROACH OF MOLECULAR KINETICS

The VAC is an algorithmic framework to approximate the slow components of molecular dynamics—also called conformation dynamics or molecular kinetics—from data. It consists of two main ingredients: (1) a variational principle that provides a computable score of a model of the slow components and (2) an algorithm based on the variational principle that estimates slow components from simulation data.

### A. Variational principle of conformation dynamics

Simulations of molecular dynamics (MD) can be modeled as realizations of an ergodic and time-reversible Markov

process  $\{\mathbf{x}_t\}$  in a phase space  $\Omega$ .  $\mathbf{x}_t$  contains all variables that determine the conformational progression after time  $t$  (e.g., positions and velocities of all atoms). The time evolution of the probability distribution  $p_t(\mathbf{x})$  of the molecular ensemble can be decomposed into a set of relaxation processes as

$$p_{t+\tau}(\mathbf{x}) = \sum_{i=1}^{\infty} e^{-\frac{\tau}{t_i}} \mu(\mathbf{x}) \psi_i(\mathbf{x}) \langle \psi_i, p_t \rangle, \quad (2)$$

where  $\mu(\mathbf{x})$  is the stationary (Boltzmann) density of the system,  $t_i$  are relaxation time scales sorted in decreasing order,  $\psi_i$  are eigenfunctions of the backward operator or Koopman operator of  $\{\mathbf{x}_t\}$  with eigenvalues  $\lambda_i(\tau) = e^{-\frac{\tau}{t_i}}$  (see Section III A), and the inner product is defined as  $\langle \psi_i, p_t \rangle = \int d\mathbf{x} \psi_i(\mathbf{x}) p_t(\mathbf{x})$ . The first spectral component is given by the constant eigenfunction  $\psi_1(\mathbf{x}) = \mathbf{1}(\mathbf{x}) \equiv 1$  and infinite time scale  $t_1 = \infty > t_2$  corresponding to the stationary state of the system. According to this decomposition, the  $m$  dominant eigenfunctions  $\psi_1, \dots, \psi_m$  can be interpreted as  $m$  slow collective variables, which characterize the behavior of a molecular system on large time scales  $\tau \gg t_{m+1}$ .

The eigenvalues and eigenfunctions can also be formulated by the following variational principle:<sup>51,53</sup> *For any  $m \geq 1$ , the first  $m$  eigenfunctions  $\psi_1, \dots, \psi_m$  are the solution of the following optimization problem:*

$$R_m = \max_{f_1, \dots, f_m} \sum_{i=1}^m \mathbb{E}_{\mu} [f_i(\mathbf{x}_t) f_i(\mathbf{x}_{t+\tau})], \quad (3)$$

s.t.  $\mathbb{E}_{\mu} [f_i(\mathbf{x}_t)^2] = 1,$   
 $\mathbb{E}_{\mu} [f_i(\mathbf{x}_t) f_j(\mathbf{x}_t)] = 0, \text{ for } i \neq j,$

where  $\mathbb{E}_{\mu}[\cdot]$  denotes the expected value with  $\mathbf{x}_t$  sampled from the stationary density  $\mu$  and  $R_m$  is the generalized Rayleigh quotient, or Rayleigh trace, with maximum value  $\sum_{i=1}^m \lambda_i$ . Therefore, for every other set of functions that aims at approximating the true eigenfunctions, the eigenvalues will be underestimated, and we can use this variational principle in order to search for the best approximation of eigenfunctions and eigenvalues.

### B. Linear variational approach

In this paper, we focus on algorithms that approximate the eigenfunctions in the spectral decomposition (2) by a linear combination of real-valued basis functions, also called feature functions,  $\chi = (\chi_1, \dots, \chi_m)^T$ . Thus, we make the ansatz,

$$f_i(\mathbf{x}) = \sum_{j=1}^m b_{ij} \chi_j(\mathbf{x}) = \mathbf{b}_i^T \chi(\mathbf{x}) \quad (4)$$

with expansion coefficients  $\mathbf{b}_i = (b_{i1}, \dots, b_{im})^T$ . Note that the functions  $\chi$  are generally nonlinear in  $\mathbf{x}$ ; however, we will call the resulting algorithm a linear VAC because it is linear in the variables  $\mathbf{b}_i$ .

#### 1. Linear VAC algorithm

By solving (3) with the ansatz (4), we obtain the linear VAC to optimally approximate eigenvalues  $\lambda_i$  and eigenfunctions  $\psi_i$ .<sup>51,53</sup> We first estimate the equilibrium



covariance and time-lagged covariance matrices of the basis functions,

$$\mathbf{C}(0) = \mathbb{E}_\mu [\chi(\mathbf{x}_t) \chi(\mathbf{x}_t)^\top], \quad (5)$$

$$\mathbf{C}(\tau) = \mathbb{E}_\mu [\chi(\mathbf{x}_t) \chi(\mathbf{x}_{t+\tau})^\top], \quad (6)$$

then the solution of the generalized eigenvalue problem

$$\mathbf{C}(\tau) \mathbf{B} = \mathbf{C}(0) \mathbf{B} \hat{\Lambda} \quad (7)$$

provides the optimal approximation to eigenvalues  $\hat{\Lambda} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_m)$  and expansion coefficient  $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_m)$ . Inserting these coefficients into (4) results in the approximated eigenfunctions.<sup>51,53</sup> An important observation is that (7) is formally equivalent to the eigenvalue decomposition of  $\mathbf{K} = \mathbf{C}(0)^{-1} \mathbf{C}(\tau)$  if  $\mathbf{C}(0)$  has full rank.  $\mathbf{K}$  is the Koopman model that is the central object of the present paper and will provide the basis for constructing equilibrium estimates from short simulations (see Section III).

The linear VAC algorithm provides a general framework for the finite-dimensional approximation of spectral components of conformation dynamics, and two widely used analysis methods, time-lagged independent component analysis (TICA)<sup>52,60,63</sup> and Markov state models (MSMs),<sup>23</sup> are both special cases of the linear VAC, see also Fig. 1.

## 2. TICA

In TICA, basis functions are mean-free molecular coordinates,  $\chi(\mathbf{x}) = \mathbf{x} - \boldsymbol{\mu}$ , where  $\boldsymbol{\mu}$  are the means. In particular, the TICA basis set is linear in the input coordinates. Then the resulting estimates of eigenfunctions can be viewed as a set of linearly independent components (ICs) with autocorrelations  $\lambda_i(\tau)$ . The dominant ICs can be used to reduce the dimension of the molecular system. Notice that using mean free coordinates is equivalent to removing the stationary spectral component  $(\lambda_1, \psi_1) \equiv (1, \mathbb{1})$ , thus TICA will only contain the dynamical components, starting from  $(\lambda_2, \psi_2)$ .

In recent MD papers, the term TICA has also been used as the application of Eqs. (5)–(7) on trajectories of features, such as distances, contact maps, or angles, i.e., the transformation  $\chi(\mathbf{x}_t)$  has been applied.<sup>52,63</sup> In this paper, we will avoid using the term TICA when VAC is meant, because a main result here is that in order to obtain a good variational approximation of the spectral components in (2), it is necessary to employ specific estimation algorithms for (5) and (6) that require the stationary spectral component  $(\lambda_1, \psi_1)$  to be kept.

## 3. MSM

The MSM is a special case of the VAC while using the indicator functions as a basis set,

$$\chi_i(\mathbf{x}) = \begin{cases} 1, & \text{for } \mathbf{x} \in A_i, \\ 0, & \text{for } \mathbf{x} \notin A_i, \end{cases} \quad (8)$$

where  $A_1, \dots, A_m$  form a partition of the phase space  $\Omega$ . With such basis functions, the correlation matrix  $\mathbf{C}(0)$  is a diagonal matrix with  $[\mathbf{C}(0)]_{ii} = \Pr(\mathbf{x}_t \in A_i)$  being the equilibrium probability of  $A_i$ , and the  $(i, j)$ th element  $[\mathbf{C}(\tau)]_{ij} = \Pr(\mathbf{x}_t \in A_i, \mathbf{x}_{t+\tau} \in A_j)$  of the time-lagged correlation matrix

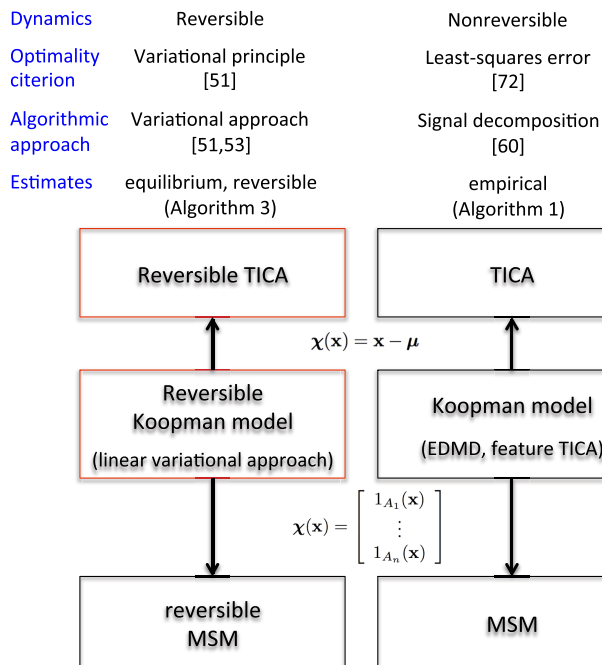


FIG. 1. Relationships between different methods for estimating the slow components of molecular kinetics: Methods for reversible dynamics are based on the variational principle, leading to the variational approach.<sup>51,53</sup> Methods for nonreversible dynamics can be derived by minimizing the least-squares error between the predicted and the observed dynamics and lead to a signal decomposition algorithm also called blind source separation.<sup>60</sup> Interestingly, the nonreversible estimates can also be obtained by implementing the using empirical estimates instead of reversible equilibrium estimates of covariance matrices. Amongst the nonreversible methods, the most general is the Koopman model estimation Algorithm (1 here), as they employ linear combinations of arbitrary basis functions. Their eigenvalue decompositions are known as EDMD and TICA in feature space. Regular TICA can be derived if the basis functions are linear in the original coordinates, and MSMs are obtained by using characteristic functions as basis. Amongst reversible methods, the variational approach leading to a reversible Koopman model Algorithm (3 here) is the most general, and reversible TICA/reversible MSM estimation methods can be derived by appropriate basis set choices. The methods in red boxes are derived in this paper, and the key to these algorithms is the ability to conduct a reversible equilibrium estimate of covariance matrices for general basis sets Algorithm (2 here).

$\mathbf{C}(\tau)$  is equal to the equilibrium frequency of the transition from  $A_i$  to  $A_j$ . Thus, a piecewise-constant approximation of eigenfunctions

$$\psi_j(\mathbf{x}) = b_{ij}, \text{ for } \mathbf{x} \in A_i, \quad (9)$$

and the corresponding eigenvalues are given by the generalized eigenvalue problem (7). When the equilibrium probability of each  $A_j$  is positive, this problem can be equivalently transformed into a simple eigenvalue problem by

$$\mathbf{C}(\tau) \mathbf{B} = \mathbf{C}(0) \mathbf{B} \mathbf{A} \Rightarrow \mathbf{P}(\tau) \mathbf{B} = \mathbf{B} \mathbf{A}. \quad (10)$$

Here,  $\mathbf{P}(\tau) = \mathbf{C}(0)^{-1} \mathbf{C}(\tau)$  is the transition matrix of the MSM with  $[\mathbf{P}(\tau)]_{ij} = \Pr(\mathbf{x}_{t+\tau} \in A_j | \mathbf{x}_t \in A_i)$  and is the Koopman model for the basis set (8). The viewpoint that MSMs can be viewed as an approximation to molecular kinetics via a projection of eigenfunctions to a basis of characteristic functions has been proposed earlier.<sup>34</sup>

The choice of more general basis functions for the VAC is beyond the scope of this paper, and some related work can be found in Refs. 53, 54, and 70.

### C. Estimation of covariance matrices

The remaining problem is how to obtain estimates of  $\mathbf{C}(0)$  and  $\mathbf{C}(\tau)$ . For convenience of notation, we take all sampled coordinates  $\mathbf{x}_t$  of a trajectory, evaluate their basis function values  $\chi(\mathbf{x}_t) = (\chi_1(\mathbf{x}_t), \dots, \chi_m(\mathbf{x}_t))^T$ , and define the following two matrices of size  $N \times m$ ,

$$\mathbf{X} = \begin{pmatrix} \chi_1(\mathbf{x}_1) & \cdots & \chi_m(\mathbf{x}_1) \\ \vdots & & \vdots \\ \chi_1(\mathbf{x}_{T-\tau}) & \cdots & \chi_m(\mathbf{x}_{T-\tau}) \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} \chi_1(\mathbf{x}_{\tau+1}) & \cdots & \chi_m(\mathbf{x}_{\tau+1}) \\ \vdots & & \vdots \\ \chi_1(\mathbf{x}_T) & \cdots & \chi_m(\mathbf{x}_T) \end{pmatrix}, \quad (11)$$

where each row corresponds to one stored time step. Thus,  $\mathbf{X}$  contains the first  $N = T - \tau$  time steps and  $\mathbf{Y}$  contains the last  $N = T - \tau$  time steps. Assuming that  $\{\mathbf{x}_t\}$  is ergodic,  $\mathbf{C}(0)$  and  $\mathbf{C}(\tau)$  can be directly estimated by time averages of  $\chi(\mathbf{x}_t)\chi(\mathbf{x}_t)^T$  and  $\chi(\mathbf{x}_t)\chi(\mathbf{x}_{t+\tau})^T$  over the trajectory,

$$\hat{\mathbf{C}}(0) = \frac{1}{N} \mathbf{X}^T \mathbf{X}, \quad (12)$$

$$\hat{\mathbf{C}}(\tau) = \frac{1}{N} \mathbf{X}^T \mathbf{Y}. \quad (13)$$

Furthermore, multiple trajectories  $k=1, \dots, K$  are trivially handled by adding up their contributions, e.g.,  $\hat{\mathbf{C}}(0) = \frac{1}{\sum_k N_k} \sum_k \mathbf{X}_k^T \mathbf{X}_k$ , etc. For all covariance estimates in this paper, we can employ the shrinkage approach<sup>79,80</sup> in order to reduce the sensitivity of estimated covariances to statistical noise<sup>81</sup> and improve the robustness of eigenvalues and eigenvectors computed from (12) and (13).

Due to statistical noise or non-equilibrium starting points, the time-lagged covariance matrix  $\hat{\mathbf{C}}(\tau)$  estimated by this method is generally not symmetric, even if the underlying dynamics are time-reversible. Thus, the eigenvalue problem (7) may yield complex eigenvalues and eigenvectors, which are undesirable in analysis of statistically reversible MD simulations. The relaxation time scales  $t_i$  can be computed from complex eigenvalues as  $t_i = -\tau / \ln |\lambda_i(\tau)|$  by using the norm of eigenvalues, but it is *a priori* unclear how to perform component analysis and dimension reduction as in TICA based on complex eigenfunctions.

In order to avoid the problem of complex estimates, a symmetric estimator is often used in applications, which approximates  $\mathbf{C}(0)$  and  $\mathbf{C}(\tau)$  by empirically averaging over all transition pairs  $(\mathbf{x}_t, \mathbf{x}_{t+\tau})$  and their reverses  $(\mathbf{x}_{t+\tau}, \mathbf{x}_t)$ , which is equivalent to averaging the time-forward and the time-inverted trajectory,

$$\hat{\mathbf{C}}_{\text{sym}}(0) \approx \frac{1}{2N} (\mathbf{X}^T \mathbf{X} + \mathbf{Y}^T \mathbf{Y}), \quad (14)$$

$$\hat{\mathbf{C}}_{\text{sym}}(\tau) \approx \frac{1}{2N} (\mathbf{X}^T \mathbf{Y} + \mathbf{Y}^T \mathbf{X}), \quad (15)$$

so that the estimate of  $\mathbf{C}(\tau)$  is always symmetric and the generalized eigenvalue problem (7) has real-valued solutions.

For equilibrium simulations, i.e., if the simulation starting points are sampled from the global equilibrium, or the simulations are much longer than the slowest relaxation times, Eqs. (14) and (15) are unbiased estimates of  $\mathbf{C}_\mu(0)$  and  $\mathbf{C}_\mu(\tau)$  and can also be derived from the maximum likelihood estimation by assuming a multivariate normal distribution of

$(\mathbf{x}_t, \mathbf{x}_{t+\tau})$ .<sup>69</sup> The major difficulty of this approach arises from non-equilibrium data, i.e., simulations whose starting points are not drawn from the equilibrium distribution and are not long enough to reach that equilibrium during the simulation. In this situation, (14) and (15) do not converge to the true covariance matrices in the limit of infinitely many trajectories and may thus provide biased estimates of the eigenvalues and eigenfunctions or independent components.

The difference between the direct estimation and symmetric estimation methods of covariance matrices becomes clear when considering the MSM special case. Since the transition matrix is  $\mathbf{P} = \mathbf{C}(0)^{-1} \mathbf{C}(\tau)$ , as shown in Section II B, transition matrices of MSMs given by the two estimators are

$$[\mathbf{P}]_{ij} = \frac{c_{ij}(\tau)}{\sum_{k=1}^m c_{ik}(\tau)} \quad (\text{direct estimation}), \quad (16)$$

$$[\mathbf{P}]_{ij} = \frac{c_{ij}(\tau) + c_{ji}(\tau)}{\sum_{k=1}^m c_{ik}(\tau) + c_{ki}(\tau)} \quad (\text{symmetric estimation}), \quad (17)$$

respectively. If the transition dynamics between discrete states  $A_1, \dots, A_m$  are exactly Markovian, the direct estimator converges to the true transition matrix in the large-data limit for non-equilibrium or even nonreversible simulations, whereas the symmetric estimator does not. However, the direct estimator may give a nonreversible transition matrix with complex eigenvalues, which is why the symmetric estimator has been frequently used before 2008 until it has been replaced by reversible maximum likelihood and Bayesian estimators.<sup>22,23,28,47,49,50</sup> How do we resolve this problem in the more general case of variational estimates with arbitrary basis functions  $\chi$ ? Below, we will introduce a solution based on the Koopman operator theory and dynamic mode decomposition (DMD).

### III. KOOPMAN MODELS OF EQUILIBRIUM KINETICS

A method equivalent to the linear VAC algorithm described in Refs. 51 and 53 and summarized in Sec. II B has more recently been introduced in the fluid mechanics field under the name extended Dynamic Mode Decomposition (EDMD).<sup>72</sup> EDMD also projects the data onto basis functions and approximates the same eigenvalue and eigenfunctions like the linear VAC. EDMD was developed independently of the VAC and is based on Dynamic Mode Decomposition (DMD).<sup>73,75,76</sup> EDMD and DMD approximate components of the Koopman operator which is a generalization of the backward operator usually used in molecular kinetics.<sup>74,78</sup>

The equivalence between the VAC and EDMD is striking, because the EDMD algorithm has been derived in a setting where dynamics are not reversible and may not even possess a unique stationary distribution. In practice, the EDMD algorithm effectively performs non-reversible empirical estimates of the covariances (12) and (13) and is used in non-equilibrium situations. EDMD is thus used in a regime for which the variational principle does not hold, and yet, it does make a useful approximation to eigenvalues and eigenfunctions of dynamical operators.<sup>72</sup> This has two important consequences:

1. The linear VAC is also usable for systems or data that are not reversible and not in equilibrium.
2. We can use ideas from the EDMD and Koopman operator theory to obtain equilibrium and reversible estimates from non-equilibrium, non-reversible data.

In this section, we will develop the second point and construct estimators for equilibrium expectations from non-equilibrium data. This will allow us to estimate relaxation time scales, slow collective variables, and equilibrium expectations using arbitrary basis sets and without requiring a cluster discretization as used in MSMs.

### A. Koopman operator description of conformation dynamics

According to the Koopman operator theory,<sup>77</sup> the dynamics of a Markov process  $\{\mathbf{x}_t\}$  can be fully described by an integral operator  $\mathcal{K}_\tau$ , called *Koopman operator*, which maps an observable quantity  $f(\mathbf{x}_t)$  at time  $t$ , to its expectation at time  $t + \tau$  as

$$\begin{aligned} \mathcal{K}_\tau f(\mathbf{x}) &= \mathbb{E}[f(\mathbf{x}_{t+\tau}) | \mathbf{x}_t = \mathbf{x}] \\ &= \int d\mathbf{y} p(\mathbf{x}, \mathbf{y}; \tau) f(\mathbf{y}). \end{aligned} \quad (18)$$

If the dynamics fulfill detailed balance, the spectral components  $\{(\lambda_i, \psi_i)\}$  discussed in Section II are in fact the eigenvalues and eigenfunctions of the Koopman operator,

$$\mathcal{K}_\tau \psi_i = \lambda_i \psi_i. \quad (19)$$

Notice that the operator description and decomposition of molecular kinetics can also be equivalently provided by the forward and backward operators, which propagate ensemble densities instead of observables.<sup>23</sup> We exploit the Koopman operator in this paper because it is the only one of these operators that can be reliably approximated from non-equilibrium data in general. See Section III B and Appendix A for a more detailed analysis.

Eq. (19) suggests the following way for spectral estimation: First, approximate the Koopman operator from data and then extract the spectral components from the estimated operator.

### B. Using linear VAC for non-equilibrium and non-reversible data: Extended dynamic mode decomposition

Like in the VAC, we can also approximate the Koopman operator  $\mathcal{K}_\tau$  by its projection  $\mathcal{K}_\tau^{\text{proj}}$  onto the subspace spanned

by basis functions  $\chi$  which satisfies

$$\mathcal{K}_\tau f \approx \mathcal{K}_\tau^{\text{proj}} f \in \text{span}\{\chi_1, \dots, \chi_m\} \quad (20)$$

for any function  $f$  in that subspace. As the Koopman operator is linear, even if the dynamics are nonlinear, it can be approximated by a matrix  $\mathbf{K} = (\mathbf{k}_1, \dots, \mathbf{k}_m) \in \mathbb{R}^{m \times m}$  as

$$\mathcal{K}_\tau^{\text{proj}} \left( \sum_{i=1}^m c_i \chi_i \right) = \sum_{i=1}^m c_i \mathbf{k}_i^\top \chi, \quad (21)$$

with

$$\mathbf{k}_i^\top \chi = \mathcal{K}_\tau^{\text{proj}} \chi_i \approx \mathcal{K}_\tau \chi_i \quad (22)$$

representing a finite-dimensional approximation of  $\mathcal{K}_\tau \chi_i$ . After a few algebraic steps,<sup>72</sup> it can be shown that eigenfunctions of  $\mathcal{K}_\tau^{\text{proj}}$  also have the form  $\psi_i = \mathbf{b}_i^\top \chi$ , and eigenvalues and eigenfunctions of  $\mathcal{K}_\tau^{\text{proj}}$  can be calculated by the eigenvalue problem,

$$\mathbf{K}\mathbf{B} = \mathbf{B}\mathbf{A}, \quad (23)$$

where definitions of  $\mathbf{A}$ ,  $\mathbf{B}$  are the same as in (7). Considering that

$$\mathbb{E}[\chi_i(\mathbf{x}_{t+\tau}) | \mathbf{x}_t] = \mathcal{K}_\tau \chi_i(\mathbf{x}_t) \approx \mathbf{k}_i^\top \chi(\mathbf{x}_t) \quad (24)$$

for each transition pair  $(\mathbf{x}_t, \mathbf{x}_{t+\tau})$  in simulations, the matrix  $\mathbf{K}$  can be determined via minimizing the mean square error between  $\mathbf{k}_i^\top \chi(\mathbf{x}_t)$  and  $\chi_i(\mathbf{x}_{t+\tau})$ ,

$$\begin{aligned} \mathbf{K} &= \arg \min_{\mathbf{K}} \frac{1}{N} \sum_{t=1}^{T-\tau} \sum_{i=1}^m \left\| \mathbf{k}_i^\top \chi(\mathbf{x}_t) - \chi_i(\mathbf{x}_{t+\tau}) \right\|^2 \\ &= \arg \min_{\mathbf{K}} \frac{1}{N} \|\mathbf{X}\mathbf{K} - \mathbf{Y}\|^2 \\ &= \hat{\mathbf{C}}(0)^{-1} \hat{\mathbf{C}}(\tau), \end{aligned} \quad (25)$$

where the covariance matrices are given by their direct estimates (12) and (13), and  $\|\cdot\|$  denotes the Frobenius norm of matrices.

Based on the above considerations, it makes sense to call the matrix  $\mathbf{K}$  together with the basis set  $\chi$  a *Koopman model* of the molecular kinetics. The Koopman model is a generalization of an MSM, as it can be constructed from any basis set  $\chi$ , not only from characteristic basis sets (Eq. (8)). Nonetheless, it shares the main features of an MSM as it can be used to propagate densities according to (20), and its eigenvalues can be used to compute relaxation time scales and its eigenvectors can be used to compute slow collective variables. Algorithm 1 summarizes the procedure of computing a nonreversible Koopman model from data. This algorithm is equivalent to the linear VAC and EDMD. If the feature trajectories are mean-free, it is also equivalent to TICA in feature space.

Please note that this pseudocode is given for illustrative purposes and should not be implemented literally. In particular, it assumes that the basis functions are linearly independent so that  $\hat{\mathbf{C}}(0)$  is invertible. In practice, linear independence can be achieved by de-correlation of basis functions—see Appendix F and specifically Algorithm 4 there for advice how to practically implement the Koopman estimator.

The above derivation shows that Koopman estimation as in Algorithm 1 has a key advantage: Suppose the points

Algorithm 1. Nonreversible Koopman estimation.

- 
1. Basis-transform input coordinates according to (11).
  2. Compute  $\hat{\mathbf{C}}(0) = \frac{1}{N} \mathbf{X}^\top \mathbf{X}$  and  $\hat{\mathbf{C}}(\tau) = \frac{1}{N} \mathbf{X}^\top \mathbf{Y}$ .
  3. Compute the Koopman model  $\mathbf{K} = \hat{\mathbf{C}}(0)^{-1} \hat{\mathbf{C}}(\tau)$ .
  4. Koopman decomposition: Solve eigenvalue problem  $\mathbf{K}\mathbf{B} = \mathbf{B}\mathbf{\Lambda}$ .
  5. Output the Koopman model  $\mathbf{K}$  and spectral components: Eigenvalues  $\lambda_i$  and eigenfunctions  $\psi_i = \mathbf{b}_i^\top \mathcal{X}$ . Both may have imaginary components that are either due to statistical noise or nonreversible dynamics.
- 

$\{\mathbf{x}_1, \dots, \mathbf{x}_{T-\tau}\}$  are sampled from a distribution  $\rho(\mathbf{x})$  which is not equal to the equilibrium distribution  $\mu$ . Although the empirical estimates of covariance matrices used in Algorithm 1 are biased with respect to the equilibrium expectations  $\mathbf{C}(0)$  and  $\mathbf{C}(\tau)$ , they are unbiased and consistent estimates of the non-equilibrium covariance matrices  $\mathbb{E}_\rho[\mathcal{X}(\mathbf{x}_t)\mathcal{X}(\mathbf{x}_t)^\top]$  and  $\mathbb{E}_\rho[\mathcal{X}(\mathbf{x}_t)\mathcal{X}(\mathbf{x}_{t+\tau})^\top]$ . Furthermore, the matrix  $\mathbf{K}$  given by (25) minimizes the error,

$$\sum_i \langle \mathbf{k}_i^\top \mathcal{X} - \mathcal{K}_\tau \mathcal{X}_i, \mathbf{k}_i^\top \mathcal{X} - \mathcal{K}_\tau \mathcal{X}_i \rangle_\rho, \quad (26)$$

with  $\langle f, g \rangle_\rho \triangleq \int d\mathbf{x} \rho(\mathbf{x}) f(\mathbf{x}) g(\mathbf{x})$ , as data size approaches infinity (see Appendices B and C). Therefore,  $\mathbf{K}$  is still a finite-dimensional approximation of  $\mathcal{K}_\tau$  with minimal mean square error with respect to  $\rho$ , which implies that Algorithm 1 is applicable to non-equilibrium data.

### C. Koopman reweighting: Estimation of the equilibrium distribution from non-equilibrium data

Not only is EDMD robust when using non-equilibrium data, we can also utilize the Koopman matrix  $\mathbf{K}$  to recover the equilibrium properties of the molecular system. According to the principle of importance sampling,<sup>82</sup> we can assign a weight

$$w(\mathbf{x}_t) \propto \frac{\mu(\mathbf{x}_t)}{\rho(\mathbf{x}_t)} \quad (27)$$

to each transition pair  $(\mathbf{x}_t, \mathbf{x}_{t+\tau})$  in the simulation data, such that the equilibrium ensemble average of a function  $h(\mathbf{x}_t, \mathbf{x}_{t+\tau})$  can be consistently estimated by the weighted mean as

$$\mathbb{E}_\mu[h(\mathbf{x}_t, \mathbf{x}_{t+\tau})] \approx \frac{\sum_{t=1}^{T-\tau} w(\mathbf{x}_t) h(\mathbf{x}_t, \mathbf{x}_{t+\tau})}{\sum_{t=1}^{T-\tau} w(\mathbf{x}_t)}. \quad (28)$$

Based on the finite-dimensional approximation (4) of spectral components, we can represent the weight function as a linear combination of basis functions  $\mathcal{X}$ ,

$$w(\mathbf{x}) = \mathbf{u}^\top \mathcal{X}(\mathbf{x}), \quad (29)$$

where  $\mathbf{u}$  satisfies

$$\hat{\mathbf{C}}(0)^{-1} \mathbf{K}^\top \hat{\mathbf{C}}(0) \mathbf{u} = \mathbf{u}, \quad (30)$$

i.e.,  $\mathbf{u}$  is the eigenvector of  $\hat{\mathbf{C}}(0)^{-1} \mathbf{K}^\top \hat{\mathbf{C}}(0)$  with eigenvalue 1, in the limit of large statistics. (See Appendix B for proofs of the above equations.)

In practice, the eigenvalue problem (30) cannot be solved for arbitrary choices of basis sets. If the basis set cannot represent the constant 1 eigenfunction, (30) does not have an

Algorithm 2. Koopman reweighting.

- 
1. Basis-transform input coordinates according to (11).
  2. Compute  $\hat{\mathbf{C}}(0) = \frac{1}{N} \mathbf{X}^\top \mathbf{X}$ ,  $\hat{\mathbf{C}}(\tau) = \frac{1}{N} \mathbf{X}^\top \mathbf{Y}$ , and  $\mathbf{K} = \hat{\mathbf{C}}(0)^{-1} \hat{\mathbf{C}}(\tau)$  as in Algorithm 1.
  3. Compute  $\mathbf{u}$  as eigenvector of  $\hat{\mathbf{C}}(0)^{-1} \mathbf{K}^\top \hat{\mathbf{C}}(0)$  with eigenvalue 1, and normalize it by  $\mathbf{1}^\top \mathbf{X} \mathbf{u}$ .
  4. Output weights:  $w(\mathbf{x}_t) = \mathbf{x}_t^\top \mathbf{u}$ .
- 

eigenvalue 1. In order to deal with general basis sets, we have two options. First, we can seek an approximate solution via the quadratic programming problem,

$$\min_{\mathbf{u}} \left\| \hat{\mathbf{C}}(0)^{-1} \mathbf{K}^\top \hat{\mathbf{C}}(0) \mathbf{u} - \mathbf{u} \right\|^2, \quad (31)$$

s.t.  $\mathbf{1}^\top \mathbf{X} \mathbf{u} = 1$

where the constraint  $\mathbf{1}^\top \mathbf{X} \mathbf{u}$  ensures that  $\sum_{t=1}^{T-\tau} w(\mathbf{x}_t) = 1$ , and  $\mathbf{1}$  denotes a column vector of ones.

We recommend a simpler way to solve this problem: Add the constant function  $\mathbb{1}$  to the basis set and change  $\mathbf{X}$  and  $\mathbf{Y}$  as  $\mathbf{X} := [\mathbf{X} \ \mathbb{1}]$  and  $\mathbf{Y} := [\mathbf{Y} \ \mathbb{1}]$  correspondingly so that the eigenvalue problem (31) can be exactly solved. The resulting method to compute equilibrium statistical weights of all samples,  $w(\mathbf{x}_t)$ , can be summarized by Algorithm 2.

Again, this algorithm is simplified for illustrative purposes. In our implementation, we ensure numerical robustness by adding the constant function to the decorrelated basis set—see Appendix E and Algorithm 5 there.

After the weights  $w(\mathbf{x}_t)$  have been estimated and normalized with  $\sum_{t=1}^{T-\tau} w(\mathbf{x}_t) = \mathbf{1}^\top \mathbf{X} \mathbf{u} = 1$ , we can compute equilibrium estimates for given observables  $f_1(\mathbf{x}_t)$  and  $f_2(\mathbf{x}_t)$  from non-equilibrium data. For example, ensemble average and time-lagged cross correlation can be approximated by

$$\mathbb{E}_\mu[f_1(\mathbf{x}_t)] \approx \sum_{t=1}^{T-\tau} w(\mathbf{x}_t) \cdot f_1(\mathbf{x}_t), \quad (32)$$

$$\mathbb{E}_\mu[f_1(\mathbf{x}_t) f_2(\mathbf{x}_{t+\tau})] \approx \sum_{t=1}^{T-\tau} w(\mathbf{x}_t) \cdot f_1(\mathbf{x}_t) f_2(\mathbf{x}_{t+\tau}). \quad (33)$$

### D. Reversible Koopman models and eigendecompositions

We now have the tools necessary to compute equilibrium covariance matrices while avoiding the bias of forced symmetrization described in Sec. II C and can conduct real-valued eigenvalue analysis for reversible dynamics using VAC or TICA. At the same time, our approach defines an equilibrium estimator of EDMD for time-reversible processes. We can obtain symmetrized equilibrium covariances from our off-equilibrium data by the following estimators:

$$\begin{aligned} \hat{\mathbf{C}}_{\text{rev}}(0) &= \frac{1}{2} \sum_{t=1}^{T-\tau} w(\mathbf{x}_t) \left( \mathcal{X}(\mathbf{x}_t) \mathcal{X}(\mathbf{x}_t)^\top + \mathcal{X}(\mathbf{x}_{t+\tau}) \mathcal{X}(\mathbf{x}_{t+\tau})^\top \right) \\ &= \frac{1}{2} \left( \mathbf{X}^\top \mathbf{W} \mathbf{X} + \mathbf{Y}^\top \mathbf{W} \mathbf{Y} \right), \end{aligned} \quad (34)$$



Algorithm 3. Reversible Koopman estimation.

- 
1. Basis-transform input coordinates according to (11).
  2. Use Koopman reweighting (Algorithm 2) to compute the equilibrium weights  $w(\mathbf{x}_t)$ .
  3. Compute  $\hat{\mathbf{C}}_{\text{rev}}(0)$  and  $\hat{\mathbf{C}}_{\text{rev}}(\tau)$  by (34) and (35).
  4. Compute the Koopman model  $\mathbf{K}_{\text{rev}} = \hat{\mathbf{C}}_{\text{rev}}(0)^{-1} \hat{\mathbf{C}}_{\text{rev}}(\tau)$ .
  5. Reversible Koopman decomposition: Solve eigenvalue problem  $\mathbf{K}_{\text{rev}} \mathbf{B} = \mathbf{B} \mathbf{\Lambda}$ .
  6. Output the Koopman model  $\mathbf{K}_{\text{rev}}$  and its spectral components: Eigenvalues  $\lambda_i$  and eigenfunctions  $\psi_i = \mathbf{b}_i^\top \chi$ . These eigenvalues and eigenfunctions are real-valued.
- 

$$\begin{aligned} \hat{\mathbf{C}}_{\text{rev}}(\tau) &= \frac{1}{2} \sum_{t=1}^{T-\tau} w(\mathbf{x}_t) \left( \chi(\mathbf{x}_t) \chi(\mathbf{x}_{t+\tau})^\top + \chi(\mathbf{x}_{t+\tau}) \chi(\mathbf{x}_t)^\top \right) \\ &= \frac{1}{2} (\mathbf{X}^\top \mathbf{W} \mathbf{Y} + \mathbf{Y}^\top \mathbf{W} \mathbf{X}). \end{aligned} \quad (35)$$

These estimators are based on time-reversibility and the reweighting approximation (28) for the equilibrium distribution (see Appendix D for proof). As a result, we obtain a time-reversible Koopman matrix,

$$\mathbf{K}_{\text{rev}} = \hat{\mathbf{C}}_{\text{rev}}(0)^{-1} \hat{\mathbf{C}}_{\text{rev}}(\tau). \quad (36)$$

By comparing (34) and (35) with (14) and (15), it is apparent that  $\hat{\mathbf{C}}_{\text{rev}}(0)$  and  $\hat{\mathbf{C}}_{\text{rev}}(\tau)$  are equal to the symmetrized direct estimates if weights of data are uniform with  $\mathbf{W} = \text{diag}(\frac{1}{N}, \dots, \frac{1}{N})$ . The weight function (29) used here can systematically reduce the bias of the symmetrized estimates for reversible dynamics. Under some weak assumptions, it can be shown that the spectral components calculated from  $\mathbf{K}_{\text{rev}}$  are real-valued and the largest eigenvalue is not larger than 1 even in the existence of statistical noise and modeling error. Furthermore, the procedure is self-consistent: If the estimation procedure is repeated while starting with weights  $w(\mathbf{x})$ , these weights remain fixed. (See Appendix E for more detailed analysis.)

The estimation algorithm for variationally optimal Koopman models of the reversible equilibrium dynamics can be summarized as Algorithm 3.

As before, this algorithm is presented in a pedagogical pseudocode. Taken literally, it will suffer from numerical instabilities if  $\hat{\mathbf{C}}_{\text{rev}}(0)$  is not positive-definite, which can also be overcome by reducing correlations between basis functions as mentioned in Section III B—see Appendix F and Algorithm 6 there.

## IV. APPLICATIONS

In this section, we compare three different estimators for molecular kinetics to the same data sets:

1. VAC or TICA in feature space symmetrization of covariance matrices (14) and (15), as proposed before.<sup>52,63</sup> Briefly we refer to this estimator as symmetrized VAC or symmetrized TICA.
2. Nonreversible Koopman estimation (Algorithm 1), which provides a nonreversible Koopman model whose eigendecomposition is equivalent to EDMD and (non-symmetrized) TICA in feature space.

3. Reversible Koopman estimation (Algorithm 3), which is consistent with the variational approach.<sup>51,53</sup>

In addition, we compare the estimated equilibrium distribution provided by Koopman reweighting (Algorithm 2) with the empirical distribution estimated from direct counting or histogramming the data in order to demonstrate the usefulness of the proposed reweighting method.

### A. One-dimensional diffusion process

As a first example, we consider a one-dimensional diffusion process  $\{x_t\}$  in a double-well potential restricted to the  $x$ -range  $[0, 2]$  as shown in Fig. 2(a). In order to validate the robustness of different estimators, we start all simulations far from equilibrium, in the region  $[0, 0.2]$  (shaded area in Fig. 2(a)). In order to apply the algorithms discussed here, we choose a basis set of 100 Gaussian functions with random parameters. For more details on the simulation model and experimental setup, see Appendix G 1.

Fig. 2(b) shows estimates of the slowest relaxation time scale  $\text{ITS}_2$  based on 500 independent short simulation trajectories with length 0.2 time units. The largest relaxation time scale  $t_2$  is computed from  $\lambda_2$  as  $t_2 = -\tau / \ln |\lambda_2(\tau)|$  and is a constant independent of lag time according to (2). For such non-equilibrium data, the symmetrized VAC significantly underestimates the relaxation time scale and gives even worse results with longer lag times. The Koopman models (both reversible and nonreversible), on the other hand, converge quickly to the true time scale before  $\tau = 0.01$  time units. The equilibrium distribution of  $\{x_t\}$  computed from Algorithm 2

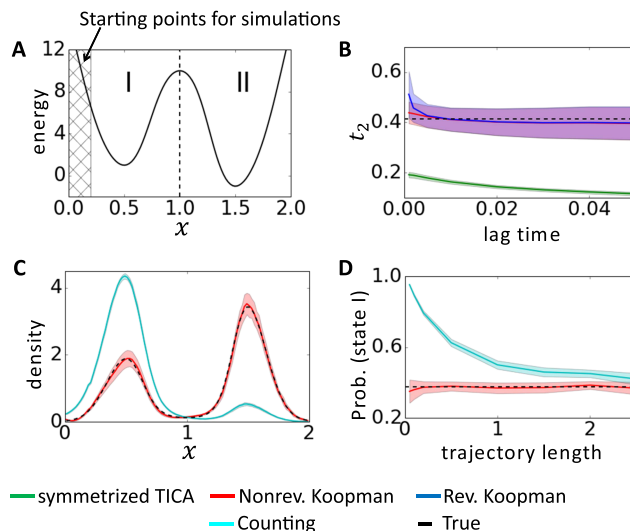


FIG. 2. Estimation results of a one-dimensional diffusion process. (a) Dimensionless energy  $U(x)$ , where the dashed line represents the border of the two potential wells I and II. The shaded area denotes the region where initial states are drawn for simulations. (b) The slowest relaxation time scale estimated by the previously used symmetrized TICA, nonreversible Koopman estimation (Algorithm 1), and reversible Koopman estimation (Algorithm 3) with different lag times. (c) Stationary density of states obtained from equilibrium probabilities of 100 uniform bins, where the probabilities are estimated using Koopman reweighting (Algorithm 2, red) and direct counting. (d) Estimates of the equilibrium probability of the potential well I given by direct counting and the Koopman reweighting (red) with different simulation trajectory lengths. In (b)–(d), solid lines and shaded regions indicate mean values and one standard deviation error intervals obtained from 30 independent experiments.

with lag time 0.01 is shown in Fig. 2(c). In contrast to the empirical histogram density given by direct counting, the direct estimator effectively recovers the equilibrium property of the process from non-equilibrium data.

Fig. 2(d) compares the empirical probability of the potential well I (i.e., by direct counting of the number of samples in well I) with the estimate from Koopman reweighting (Algorithm 2), for different simulation trajectory lengths, where the lag time is still 0.01 time units and the accumulated simulation time is kept fixed to be 100. Due to the ergodicity of the process, the empirical probability converges to the true value as the trajectory length increases. The convergence rate, however, is very slow as shown in Fig. 2(d), and empirical probability is close to the true value only for trajectories longer than 2 time units. When using the reweighting method proposed here, the estimated probability is robust with respect to changes in trajectory length, and accurate even for very short trajectories.

## B. Two-dimensional diffusion process

Next, we study a two-dimensional diffusion process  $\{(x_t, y_t)\}$  which has three potential wells as shown in Fig. 3(a), where all simulations are initialized with  $(x_0, y_0) \in [-2, -1.5] \times [-1.5, 2.5]$ , and the set of basis functions for spectral estimation consists of 100 Gaussian functions with random parameters (see Appendix G 2 for details).

We generate 8000 short simulation trajectories with length 1.25 and show the empirical free energy of the simulation data in Fig. 3(b). Comparing Figs. 3(b) and 3(a), it can be seen that most of the simulation data are distributed in the area  $x \leq 0$  and the empirical distribution of simulations is very different from the equilibrium distribution. Therefore,

eigenvalues/time scales and eigenfunctions estimated by the symmetrized VAC have large errors, whereas the nonreversible and reversible Koopman model provides accurate eigenvalues/time scales and eigenfunctions (Figs. 3(d) and 3(f)). Moreover the equilibrium density can be recovered with high accuracy using Koopman reweighting, although the data are far from equilibrium (Fig. 3(c)).

For such a two-dimensional process, it is also interesting to investigate the slow collective variables predicted by TICA, i.e., directly using the  $x$  and  $y$  coordinates as basis functions. Fig. 3(a) displays the TICA components from the exact equilibrium distribution with lag time  $\tau = 0.01$ . Notice that the slowest mode is parallel to  $x$ -axis, which is related to transitions between potential wells I and II, and the second IC is parallel to the  $y$ -axis, which is related to transitions between  $\{I, II\}$  and III. However, if we extract ICs from simulation data by using TICA with symmetrized covariance matrices, the result is significantly different as shown in Fig. 3(b), where the first IC characterizes transitions between I and III. The ICs given by nonreversible and reversible Koopman models (Algorithms 1 and 3 here) can be seen in Fig. 3(c). They are still different from those in Fig. 3(a) because the equilibrium distribution is difficult to approximate with only linear basis functions, but much more accurate than the estimates obtained by the previously used symmetric estimator in Fig. 3(b).

Fig. 3(e) shows the estimation errors of estimated equilibrium distribution obtained by using simulations with different trajectory lengths, where the accumulated simulation time is kept fixed to be  $10^4$ , the lag time for estimators is  $\tau = 0.005$ , and the error is evaluated as the total variation distance between the estimated probability distributions of the three potential wells and the true reference. Fig. 3(f) shows angles of linear

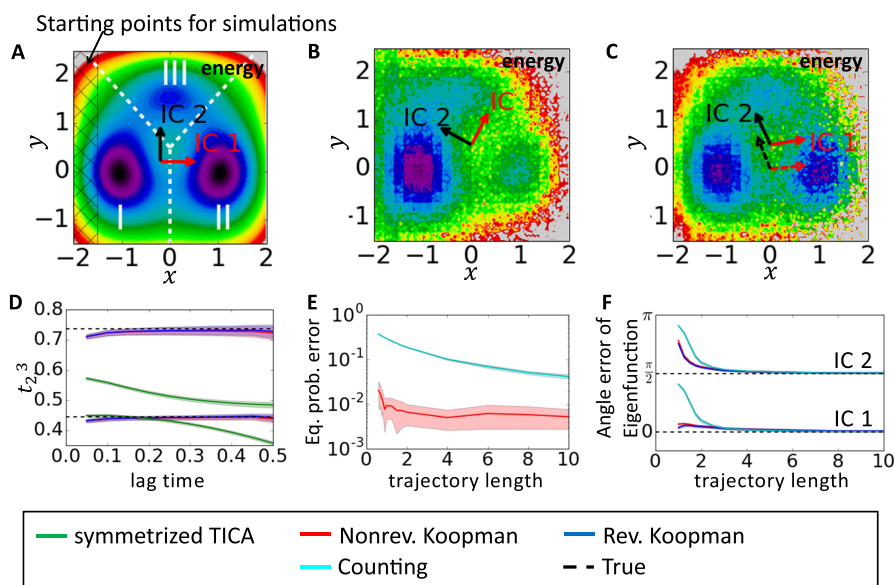


FIG. 3. Estimation results of a two-dimensional diffusion process. (a) Free energy of the process, where the dashed line represents the border of potential wells I, II, and III. The shaded area denotes the region where initial states are drawn for simulations, and the two linear ICs obtained from TICA with exact statistics. (b) Free energies computed from a histogram of the simulation data (direct counting). Arrows show the directions of TICA components computed from symmetrized TICA. (c) Free energies computed from Koopman reweighting (Algorithm 2). Arrows show the directions of the slowest modes computed from a reversible (solid arrows) and nonreversible (dashed arrows) Koopman estimation using  $\{x, y\}$  as a basis set. (d) Estimates of the two slowest relaxation time scales. (e) Estimation errors of equilibrium distributions using direct counting or the Koopman model (red). (f) Error in the angles of estimated eigenfunctions. Shaded area shows the standard deviation computed from 30 independent simulations.

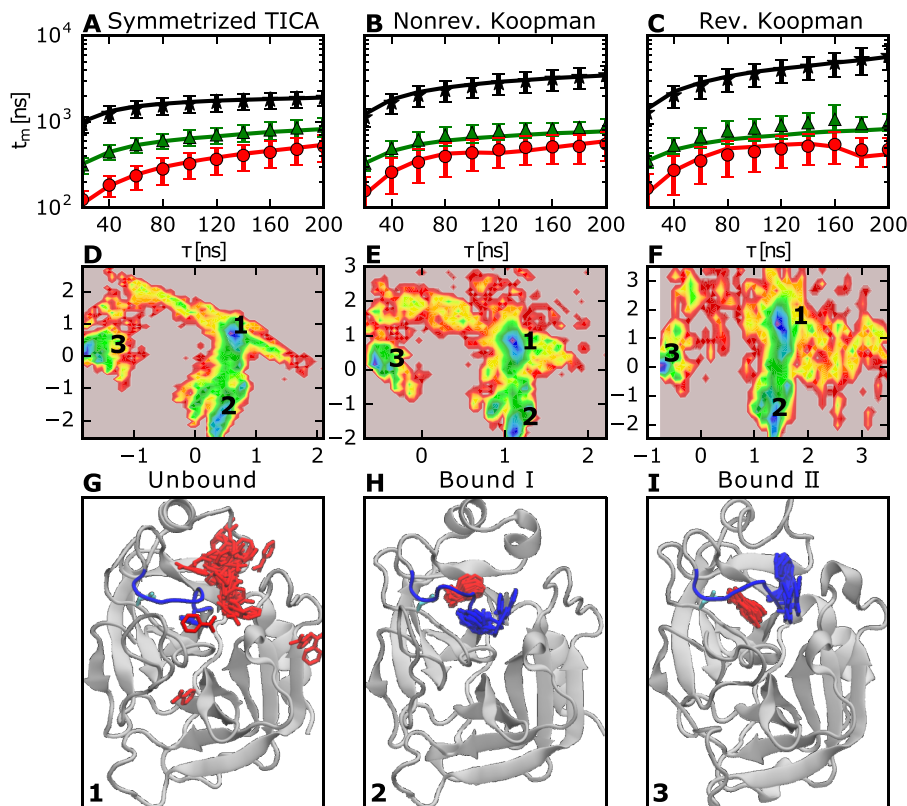


FIG. 4. Results for MD simulations of the trypsin-benzamidine binding process. ((a)–(c)) Relaxation time scales are estimated as a function of lag time. (a) TICA in feature space with the previously used symmetric estimator, (b) nonreversible Koopman model, equivalent to TICA in feature space without symmetrization (Algorithm 1 here), (c) variational reversible Koopman model suggested here (Algorithm 3). (d)–(f) Free energy landscapes (negative logarithm of the sampled densities) plotted on the two slowest process eigenfunctions. For all three methods, minima 1–3 correspond to the same macro-states of the system. Representative structures of these states are shown in (g)–(i). State 1 represents the ligand being unbound or loosely attached to the protein. States 2 and 3 are different conformational arrangements of the bound state, in particular of the binding loop including Trp 215.<sup>16</sup>

ICs approximated from the same simulation data with lag time  $\tau = 0.01$ . Both of the figures clearly demonstrate the superiority of the Koopman models suggested here.

### C. Protein-ligand binding

We revisit the binding process of benzamidine to trypsin which was studied previously in Refs. 11 and 71. The data set consists of 52 trajectories of  $2 \mu\text{s}$  and four trajectories of  $1 \mu\text{s}$  simulation time, resulting in a total simulation time of  $108 \mu\text{s}$ . From the simulations, we extract a feature set of 223 nearest neighbor heavy-atom contacts between all trypsin residues and the ligand. In this feature space, we then perform TICA using the symmetrized estimate (previous standard) and estimate a nonreversible Koopman model (Algorithm 1) and a reversible Koopman model (Algorithm 3). In order to obtain uncertainties, we compute 100 bootstrapping estimates in which outliers were rejected. In Figures 4(a)–4(c), we show the three slowest implied time scales as estimated by the three approaches discussed above. We observe that both the Koopman models provide a larger slowest implied time scale than symmetrized TICA. The slowest time scale estimated by the reversible estimator converges on relatively long lag times. This is likely due to the fact that the trypsin-benzamidine binding kinetics involves internal conformational changes of trypsin.<sup>16</sup> The first TICA components of the direct estimate are coincidentally purely real here. The eigenvectors used for the dimensionality reduction were estimated at lag time  $\tau = 100$  ns. The projections are qualitatively similar, revealing three minima of the landscape, labeled 1, 2, and 3. In all three cases, these centers correspond to the same macro-states of the system, shown underneath in Figures 4(g)–4(i).

Center 1 corresponds to the ligand being either unbound or loosely attached to the protein. The other two states are different conformational arrangements of the bound state of the ligand.

### V. CONCLUSION

Using dynamic mode decomposition theory, we have shown that the variational approach of conformation dynamics and the time-lagged independent component analysis can be made with small bias even if just empirical out-of-equilibrium estimates of the covariance matrices are available, i.e., they can be applied to ensembles of short MD simulations starting from an arbitrary starting point. A crucial point is that the forceful symmetrization of the empirical covariances practiced in previous studies must be avoided.

In order to facilitate a bias-corrected symmetric estimate of covariance matrices, we have proposed a Koopman reweighting technique in which the weights of sampled configurations can be estimated using a first pass over the data, during which empirical covariance matrices must be estimated. These weights can be applied in order to turn the empirical (out-of-equilibrium) estimates of covariance matrices into estimates of the equilibrium covariance matrices. These matrices can then be symmetrized without introducing a bias from the empirical distribution, resulting in real-valued eigenvalue and eigenfunction estimates.

With these algorithms, the variational approach and thus also the TICA algorithm inherit the same benefits that MSMs have enjoyed since nearly a decade: we can generate optimal and robust reversible and nonreversible estimates of the equilibrium kinetics from swarms of short trajectories

not started from equilibrium. Although this work focuses on the estimation of eigenvalues and eigenfunctions of the Koopman operator, the proposed Algorithms 1 and 3 provide Koopman models, which are discrete approximations of the Koopman operator, and that be used for other purposes, such as the propagation of densities. Koopman models are generalizations of Markov state models using arbitrary basis functions.

Besides the application to molecular kinetics highlighted in this paper, the Koopman reweighting principle described in Algorithm 2 can be used to compute variationally optimal estimates of any equilibrium property (expectation values, distributions) from out-of-equilibrium data using an approach that involves arbitrary sets of basis functions. While the viability of this approach critically depends on the suitability of the basis functions employed, it offers a very general way to computing equilibrium expectations that may lead to other applications and extensions in future studies.

## APPENDIX A: DYNAMICAL OPERATORS

Besides the Koopman operator  $\mathcal{K}_\tau$ , the conformation dynamics of a molecular system can also be described by the forward operator  $\mathcal{P}_\tau$  and backward operator, or called transfer operator,  $\mathcal{T}_\tau$ ,<sup>23</sup> which describe the evolution of ensemble densities as

$$\begin{aligned} p_{t+\tau}(\mathbf{x}) &= \mathcal{P}_\tau p_t(\mathbf{x}) \\ &= \int d\mathbf{y} p(\mathbf{y}, \mathbf{x}; \tau) p_t(\mathbf{y}) \end{aligned} \quad (\text{A1})$$

and

$$\begin{aligned} u_{t+\tau}(\mathbf{x}) &= \mathcal{T}_\tau u_t(\mathbf{x}) \\ &= \int d\mathbf{y} \frac{\mu(\mathbf{y})}{\mu(\mathbf{x})} p(\mathbf{y}, \mathbf{x}; \tau) u_t(\mathbf{y}), \end{aligned} \quad (\text{A2})$$

where  $p_t(\mathbf{x})$  denotes the probability density of  $\mathbf{x}_t$  and  $u_t(\mathbf{x}) = \mu(\mathbf{x})^{-1} p_t(\mathbf{x})$  denotes the density weighted by the inverse of the stationary density. The relationship between the three operators can be summarized as follows:

1.  $\mathcal{K}_\tau$  is adjoint to  $\mathcal{T}_\tau$  in the sense of

$$\langle \mathcal{K}_\tau f_1, f_2 \rangle_\mu = \langle f_1, \mathcal{T}_\tau f_2 \rangle_\mu \quad (\text{A3})$$

for any  $f_1, f_2 \in L^2_\mu$ . If  $\{\mathbf{x}_t\}$  is reversible,  $\mathcal{K}_\tau$  and  $\mathcal{T}_\tau$  are self-adjoint with respect to  $\langle \cdot, \cdot \rangle_\mu$ , i.e.,  $\mathcal{K}_\tau = \mathcal{T}_\tau$ .

2. Defining the multiplication operator  $\mathcal{M}_\mu : L^2_\mu \mapsto L^2_{\mu^{-1}}$  as  $\mathcal{M}_\mu f(\mathbf{x}) = \mu(\mathbf{x}) \cdot f(\mathbf{x})$ , the Markov propagator  $\mathcal{P}_\tau$  can be expressed as

$$\mathcal{P}_\tau = \mathcal{M}_\mu \mathcal{T}_\tau \mathcal{M}_\mu^{-1}. \quad (\text{A4})$$

Under the detailed balance condition,  $\mathcal{P}_\tau$  is self-adjoint with respect to  $\langle \cdot, \cdot \rangle_{\mu^{-1}}$ .

We can also find the finite-dimensional approximation  $\mathcal{P}_\tau \chi_i \approx \mathbf{p}_i^\top \chi$  and  $\mathcal{T}_\tau \chi_i \approx \mathbf{t}_i^\top \chi$  of  $\mathcal{P}_\tau$  and  $\mathcal{T}_\tau$  by minimizing errors  $\sum_i \langle \mathbf{p}_i^\top \chi - \mathcal{P}_\tau \chi_i, \mathbf{p}_i^\top \chi - \mathcal{P}_\tau \chi_i \rangle_\omega$  and  $\sum_i \langle \mathbf{t}_i^\top \chi - \mathcal{T}_\tau \chi_i, \mathbf{t}_i^\top \chi - \mathcal{T}_\tau \chi_i \rangle_\omega$  for some weight function  $\omega(\mathbf{x})$ . However, it is still unknown how to select the weight functions so that the approximation errors can be easily computed from simulation data as in the approximation of  $\mathcal{K}_\tau$ . For

example, if we select  $\omega(\mathbf{x}) = \rho(\mathbf{x})^{-1}$ , the approximation error of  $\mathcal{P}_\tau$  is

$$\begin{aligned} & \sum_i \langle \mathbf{p}_i^\top \chi - \mathcal{P}_\tau \chi_i, \mathbf{p}_i^\top \chi - \mathcal{P}_\tau \chi_i \rangle_{\rho^{-1}} \\ &= \sum_i \langle \mathbf{p}_i^\top \chi, \mathbf{p}_i^\top \chi \rangle_{\rho^{-1}} - 2 \sum_i \langle \mathbf{p}_i^\top \chi, \mathcal{P}_\tau \chi_i \rangle_{\rho^{-1}} \\ &+ \sum_i \langle \mathcal{P}_\tau \chi_i, \mathcal{P}_\tau \chi_i \rangle_{\rho^{-1}} = \sum_i \mathbb{E}_\rho \left[ \frac{\mathbf{p}_i^\top \chi(\mathbf{x}_t) \chi(\mathbf{x}_t)^\top \mathbf{p}_i}{\rho(\mathbf{x}_t)^2} \right] \\ &- 2 \sum_i \mathbb{E}_\rho \left[ \frac{\mathbf{p}_i^\top \chi(\mathbf{x}_{t+\tau}) \chi_i(\mathbf{x}_t)}{\rho(\mathbf{x}_{t+\tau}) \rho(\mathbf{x}_t)} \right] + \sum_i \langle \mathcal{P}_\tau \chi_i, \mathcal{P}_\tau \chi_i \rangle_{\rho^{-1}}, \end{aligned} \quad (\text{A5})$$

where the last term on the right hand side is a constant independent of  $\mathbf{p}_i$ , and the computation of the first two terms is infeasible for unknown  $\rho$ . For  $\mathcal{T}_\tau$ , the weight function is generally set to be  $\omega = \rho$ , and the corresponding approximation error is then

$$\begin{aligned} & \sum_i \langle \mathbf{t}_i^\top \chi - \mathcal{T}_\tau \chi_i, \mathbf{t}_i^\top \chi - \mathcal{T}_\tau \chi_i \rangle_\rho \\ &= \sum_i \langle \mathbf{t}_i^\top \chi, \mathbf{t}_i^\top \chi \rangle_\rho - 2 \sum_i \langle \mathbf{t}_i^\top \chi, \mathcal{T}_\tau \chi_i \rangle_\rho \\ &+ \sum_i \langle \mathcal{T}_\tau \chi_i, \mathcal{T}_\tau \chi_i \rangle_\rho = \sum_i \mathbb{E}_\rho \left[ \mathbf{t}_i^\top \chi(\mathbf{x}_t) \chi(\mathbf{x}_t)^\top \mathbf{t}_i \right] \\ &- 2 \sum_i \mathbb{E}_\rho \left[ \frac{\rho(\mathbf{x}_{t+\tau}) \mu(\mathbf{x}_t)}{\mu(\mathbf{x}_{t+\tau}) \rho(\mathbf{x}_t)} \cdot \mathbf{t}_i^\top \chi(\mathbf{x}_{t+\tau}) \chi_i(\mathbf{x}_t) \right] \\ &+ \sum_i \langle \mathcal{T}_\tau \chi_i, \mathcal{T}_\tau \chi_i \rangle_\rho \end{aligned} \quad (\text{A6})$$

which is difficult to estimate unless the empirical distribution  $\rho$  is consistent with  $\mu$  or the system is reversible. (For reversible systems,  $\mathcal{K}_\tau = \mathcal{T}_\tau$  and the finite-dimensional approximation of  $\mathcal{K}_\tau$  is therefore also that of  $\mathcal{T}_\tau$ .) In general cases, only the Koopman operator can be reliably estimated from the non-equilibrium data.

## APPENDIX B: PROPERTIES OF THE EMPIRICAL DISTRIBUTION

We first consider the case where the simulation data consist of  $M$  independent trajectories  $\{\mathbf{x}_t^1\}_{t=1}^T, \dots, \{\mathbf{x}_t^K\}_{t=1}^T$  of length  $T$  and the initial state  $\mathbf{x}_0^k \stackrel{\text{iid}}{\sim} p_0(\mathbf{x})$ . In this case,  $\rho$  can be given by

$$\rho = \frac{1}{T-\tau} \sum_{i=1}^{T-\tau} \mathcal{P}_i p_0, \quad (\text{B1})$$

where  $\mathcal{P}_i$  denotes the forward operator defined in Appendix A.

For an arbitrary function  $h$  of  $\mathbf{x}_t$  and  $\mathbf{x}_{t+\tau}$ , we have

$$\begin{aligned} & \mathbb{E} \left[ \frac{1}{K(T-\tau)} \sum_{k=1}^K \sum_{i=1}^{T-\tau} h(\mathbf{x}_t^k, \mathbf{x}_{t+\tau}^k) \right] \\ &= \frac{1}{T-\tau} \sum_{i=1}^{T-\tau} \mathbb{E}_{\mathcal{P}_i p_0} [h(\mathbf{x}_t, \mathbf{x}_{t+\tau})] = \mathbb{E}_\rho [h(\mathbf{x}_t, \mathbf{x}_{t+\tau})] \end{aligned} \quad (\text{B2})$$

and



$$\frac{1}{K(T-\tau)} \sum_{k=1}^K \sum_{t=1}^{T-\tau} h(\mathbf{x}_t^k, \mathbf{x}_{t+\tau}^k) = \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} \left( \frac{1}{K} \sum_{k=1}^K h(\mathbf{x}_t^k, \mathbf{x}_{t+\tau}^k) \right) \xrightarrow{p} \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} \mathbb{E}_{\mathcal{P}_t p_0} [h(\mathbf{x}_t, \mathbf{x}_{t+\tau})] = \mathbb{E}_\rho [h(\mathbf{x}_t, \mathbf{x}_{t+\tau})] \quad (\text{B3})$$

as  $M \rightarrow \infty$ , where “ $\xrightarrow{p}$ ” denotes the convergence in probability. Therefore  $\hat{\mathbf{C}}(0)$  and  $\hat{\mathbf{C}}(\tau)$  are unbiased and consistent estimates of

$$\mathbf{C}_\rho(0) = \mathbb{E}_\rho [\boldsymbol{\chi}(\mathbf{x}_t) \boldsymbol{\chi}(\mathbf{x}_t)^\top], \quad (\text{B4})$$

$$\mathbf{C}_\rho(\tau) = \mathbb{E}_\rho [\boldsymbol{\chi}(\mathbf{x}_t) \boldsymbol{\chi}(\mathbf{x}_{t+\tau})^\top]. \quad (\text{B5})$$

The importance sampling approximation provided by (28) is also consistent because

$$\begin{aligned} \frac{\sum_{k=1}^K \sum_{t=1}^{T-\tau} w(\mathbf{x}_t^k) h(\mathbf{x}_t^k, \mathbf{x}_{t+\tau}^k)}{\sum_{k=1}^K \sum_{t=1}^{T-\tau} w(\mathbf{x}_t^k)} &= \frac{\frac{1}{K(T-\tau)} \sum_{k=1}^K \sum_{t=1}^{T-\tau} w(\mathbf{x}_t^k) h(\mathbf{x}_t^k, \mathbf{x}_{t+\tau}^k)}{\frac{1}{K(T-\tau)} \sum_{k=1}^K \sum_{t=1}^{T-\tau} w(\mathbf{x}_t^k)} \xrightarrow{p} \frac{\mathbb{E}_\rho [w(\mathbf{x}_t) h(\mathbf{x}_t, \mathbf{x}_{t+\tau})]}{\mathbb{E}_\rho [w(\mathbf{x}_t)]} \\ &= \frac{\iint d\mathbf{x} d\mathbf{y} \frac{\mu(\mathbf{x})}{\rho(\mathbf{x})} \rho(\mathbf{x}) p(\mathbf{x}, \mathbf{y}) h(\mathbf{x}_t, \mathbf{x}_{t+\tau})}{\int d\mathbf{x} \frac{\mu(\mathbf{x})}{\rho(\mathbf{x})} \rho(\mathbf{x})} = \mathbb{E}_\mu [h(\mathbf{x}_t, \mathbf{x}_{t+\tau})]. \end{aligned} \quad (\text{B6})$$

If we further assume that the finite-dimensional approximation (4) of spectral components is exact, i.e.,  $\{\mathbf{x}_t\}$  has only  $m$  nonzero eigenvalues and  $\psi_i = \mathbf{b}_i^\top \boldsymbol{\chi}$  holds exactly for  $i = 1, \dots, m$ , we can get

$$\begin{aligned} \frac{\rho(\mathbf{x})}{\mu(\mathbf{x})} &= \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} \frac{\mathcal{P}_t p_0(\mathbf{x})}{\mu(\mathbf{x})} \\ &= \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} \frac{\sum_{i=1}^m \lambda_i(t) \mu(\mathbf{x}) \psi_i(\mathbf{x}) \langle \psi_i, p_0 \rangle}{\mu(\mathbf{x})} \\ &= \left[ \frac{\sum_{t=1}^{T-\tau} \lambda_1(t)}{T-\tau} \langle \psi_1, p_0 \rangle, \dots, \frac{\sum_{t=1}^{T-\tau} \lambda_m(t)}{T-\tau} \langle \psi_m, p_0 \rangle \right] \\ &\quad \times \mathbf{B}^\top \boldsymbol{\chi}(\mathbf{x}) \end{aligned} \quad (\text{B7})$$

which implies that (29) can be exactly satisfied with

$$\mathbf{u} = \mathbf{B} \left[ \frac{\sum_{t=1}^{T-\tau} \lambda_1(t)}{T-\tau} \langle \psi_1, p_0 \rangle, \dots, \frac{\sum_{t=1}^{T-\tau} \lambda_m(t)}{T-\tau} \langle \psi_m, p_0 \rangle \right]^\top. \quad (\text{B8})$$

Moreover, under the finite-dimensional assumption, we have

$$\begin{aligned} \mathcal{K}_\tau \boldsymbol{\chi}_i(\mathbf{x}) &= \int d\mathbf{y} p(\mathbf{x}, \mathbf{y}; \tau) \boldsymbol{\chi}_i(\mathbf{y}) \\ &= \int d\mathbf{y} \sum_{i=1}^m \lambda_i(t) \mu(\mathbf{y}) \psi_i(\mathbf{y}) \psi_i(\mathbf{x}) \boldsymbol{\chi}_i(\mathbf{y}) \\ &= \left( \int d\mathbf{y} \sum_{i=1}^m \lambda_i(t) \mu(\mathbf{y}) \psi_i(\mathbf{y}) \boldsymbol{\chi}_i(\mathbf{y}) \right) \mathbf{b}_i^\top \boldsymbol{\chi}(\mathbf{x}). \end{aligned} \quad (\text{B9})$$

Thus there exists a matrix  $\mathbf{K} = (\mathbf{k}_1, \dots, \mathbf{k}_m)$  so that  $\mathcal{K}_\tau \boldsymbol{\chi} = \mathbf{K}^\top \boldsymbol{\chi}$  holds exactly with  $\mathcal{K}_\tau \boldsymbol{\chi} = (\mathcal{K}_\tau \boldsymbol{\chi}_1, \dots, \mathcal{K}_\tau \boldsymbol{\chi}_m)^\top$ . Considering that

$$\begin{aligned} \mathbb{E}_\mu [\boldsymbol{\chi}(\mathbf{x}_{t+\tau})] &= \mathbb{E}_\mu [\mathcal{K}_\tau \boldsymbol{\chi}(\mathbf{x}_t)] \\ &= \mathbb{E}_\mu [\mathbf{K}^\top \boldsymbol{\chi}(\mathbf{x}_t)] \\ &= \int d\mathbf{x} \mathbf{u}^\top \boldsymbol{\chi}(\mathbf{x}) \cdot \rho(\mathbf{x}) \cdot \mathbf{K}^\top \boldsymbol{\chi}(\mathbf{x}) \\ &= \mathbf{K}^\top \left( \int d\mathbf{x} \rho(\mathbf{x}) \boldsymbol{\chi}(\mathbf{x}) \boldsymbol{\chi}(\mathbf{x})^\top \right) \mathbf{u} \\ &= \mathbf{K}^\top \mathbf{C}_\rho(0) \mathbf{u} \end{aligned} \quad (\text{B10})$$

and

$$\begin{aligned} \mathbb{E}_\mu [\boldsymbol{\chi}(\mathbf{x}_t)] &= \int d\mathbf{x} \mathbf{u}^\top \boldsymbol{\chi}(\mathbf{x}) \cdot \rho(\mathbf{x}) \cdot \boldsymbol{\chi}(\mathbf{x}) \\ &= \mathbf{C}_\rho(0) \mathbf{u}, \end{aligned} \quad (\text{B11})$$

we can obtain from  $\mathbb{E}_\mu [\boldsymbol{\chi}(\mathbf{x}_{t+\tau})] = \mathbb{E}_\mu [\boldsymbol{\chi}(\mathbf{x}_t)]$  that

$$\mathbf{C}_\rho(0)^{-1} \mathbf{K}^\top \mathbf{C}_\rho(0) \mathbf{u} = \mathbf{u}. \quad (\text{B12})$$

Since  $\mathbf{C}_\rho(0)^{-1} \mathbf{K}^\top \mathbf{C}_\rho(0)$  is similar to  $\mathbf{K}^\top$  and the largest eigenvalue of  $\mathbf{K}$  is 1, we can conclude that  $\mathbf{u}$  is the eigenvector of  $\mathbf{C}_\rho(0)^{-1} \mathbf{K}^\top \mathbf{C}_\rho(0)$  with the largest eigenvalue.

In more general cases, where, for example, trajectories are generated with different initial conditions and different lengths, the similar conclusions can be obtained by considering that

$$\begin{aligned} \frac{\mathcal{P}_t p_0(\mathbf{x})}{\mu(\mathbf{x})} &= [\lambda_1(t) \langle \psi_1, p_0 \rangle, \dots, \lambda_m(t) \langle \psi_m, p_0 \rangle] \\ &\quad \times \mathbf{B}^\top \boldsymbol{\chi}(\mathbf{x}) \in \text{span}\{\boldsymbol{\chi}_1, \dots, \boldsymbol{\chi}_m\} \end{aligned} \quad (\text{B13})$$

for all  $p_0$  and  $t$  if the finite-dimensional approximation (4) is assumed to be exact, i.e., the ratio between  $\rho$  and  $\mu$  can always be expressed as a linear combination of  $\boldsymbol{\chi}$  under this assumption.

### APPENDIX C: LIMIT OF THE KOOPMAN MODEL APPROXIMATION ERROR

The mean square error of the nonreversible Koopman model approximation is (see Ref. 72)

$$\text{MSE} = \frac{1}{N} \sum_{t=1}^{T-\tau} \sum_{i=1}^m \left\| \mathbf{k}_i^\top \chi_i(\mathbf{x}_t) - \chi_i(\mathbf{x}_{t+\tau}) \right\|^2. \quad (\text{C1})$$

Under the condition  $N \rightarrow \infty$ , we have

$$\begin{aligned} \text{MSE} &= \sum_{i=1}^m \int d\mathbf{x} \rho(\mathbf{x}) \left( \mathbf{k}_i^\top \chi - \mathcal{K}_\tau \chi_i \right)^\top \left( \mathbf{k}_i^\top \chi - \mathcal{K}_\tau \chi_i \right) \\ &= \sum_{i=1}^m \left\langle \mathbf{k}_i^\top \chi - \mathcal{K}_\tau \chi_i, \mathbf{k}_i^\top \chi - \mathcal{K}_\tau \chi_i \right\rangle_\rho. \end{aligned}$$

### APPENDIX D: PROOF OF (34) AND (35)

If  $\{\mathbf{x}_t\}$  is a time-reversible stochastic process, the time-lagged cross correlation between two arbitrary observable quantities  $f_1(\mathbf{x}_t)$  and  $f_2(\mathbf{x}_t)$  at equilibrium is symmetric in the sense of  $\mathbb{E}_\mu [f_1(\mathbf{x}_t) f_2(\mathbf{x}_{t+\tau})] = \mathbb{E}_\mu [f_2(\mathbf{x}_t) f_1(\mathbf{x}_{t+\tau})]$ . We can obtain that

$$\begin{aligned} \mathbf{C}(0) &= \mathbb{E}_\mu \left[ \chi(x_t) \chi(x_t)^\top \right] \\ &= \frac{1}{2} \mathbb{E}_\mu \left[ \chi(x_t) \chi(x_t)^\top + \chi(x_{t+\tau}) \chi(x_{t+\tau})^\top \right] \\ &\approx \frac{1}{2} \sum_{t=1}^{T-\tau} w(\mathbf{x}_t) \left( \chi(x_t) \chi(x_t)^\top + \chi(x_{t+\tau}) \chi(x_{t+\tau})^\top \right) \end{aligned}$$

and

$$\begin{aligned} \mathbf{C}(\tau) &= \mathbb{E}_\mu \left[ \chi(x_t) \chi(x_{t+\tau})^\top \right] \\ &= \frac{1}{2} \mathbb{E}_\mu \left[ \chi(x_t) \chi(x_{t+\tau})^\top + \chi(x_{t+\tau}) \chi(x_t)^\top \right] \\ &\approx \frac{1}{2} \sum_{t=1}^{T-\tau} w(\mathbf{x}_t) \left( \chi(x_t) \chi(x_{t+\tau})^\top + \chi(x_{t+\tau}) \chi(x_t)^\top \right), \end{aligned}$$

where the approximation steps in the above equations come from (28).

### APPENDIX E: ANALYSIS OF THE REVERSIBLE ESTIMATOR

Here we analyze properties of the reversible estimator under the following assumptions.

*Assumption 1.* The constant function  $\mathbb{1} \in \text{span}\{\chi_1, \dots, \chi_m\}$ , i.e., there is a vector  $\mathbf{v}$  so that  $\mathbf{v}^\top \chi = \mathbb{1}$ .

*Assumption 2.*  $\hat{\mathbf{C}}(0)$ ,  $\hat{\mathbf{C}}_{\text{rev}}(0)$  are positive-definite, and all weights  $w(\mathbf{x}_t)$  are positive.

Under Assumption 2,  $\mathbf{K}_{\text{rev}}$  is similar to

$$\hat{\mathbf{C}}_{\text{rev}}(0)^{\frac{1}{2}} \mathbf{K}_{\text{rev}} \hat{\mathbf{C}}_{\text{rev}}(0)^{-\frac{1}{2}} = \hat{\mathbf{C}}_{\text{rev}}(0)^{-\frac{1}{2}} \hat{\mathbf{C}}_{\text{rev}}(\tau) \hat{\mathbf{C}}_{\text{rev}}(0)^{-\frac{1}{2}}, \quad (\text{E1})$$

where  $\hat{\mathbf{C}}_{\text{rev}}(0)^{\frac{1}{2}}$  denotes the symmetric square root of  $\hat{\mathbf{C}}_{\text{rev}}(0)$ . Therefore the eigenvalue problem of  $\mathbf{K}_{\text{rev}}$  can be solved in the

real field. In addition, for any  $\lambda$  and nonzero  $\mathbf{b}$  which satisfy  $\mathbf{K}_{\text{rev}} \mathbf{b} = \lambda \mathbf{b}$ , we have

$$\begin{aligned} |\lambda| &= \left| \frac{\mathbf{b}^\top \hat{\mathbf{C}}_{\text{rev}}(\tau) \mathbf{b}}{\mathbf{b}^\top \hat{\mathbf{C}}_{\text{rev}}(0) \mathbf{b}} \right| \\ &= \frac{\left| \mathbf{b}^\top \mathbf{X}^\top \mathbf{W} \mathbf{Y} \mathbf{b} + \mathbf{b}^\top \mathbf{Y}^\top \mathbf{W} \mathbf{X} \mathbf{b} \right|}{\mathbf{b}^\top \mathbf{X}^\top \mathbf{W} \mathbf{X} \mathbf{b} + \mathbf{b}^\top \mathbf{Y}^\top \mathbf{W} \mathbf{Y} \mathbf{b}} \\ &\leq 1 \end{aligned}$$

which implies that the spectral radius of  $\mathbf{K}_{\text{rev}}$  is not larger than 1.

Under Assumption 1, the matrix  $\mathbf{K}$  given by (25) satisfies

$$\begin{aligned} \mathbf{K} \mathbf{v} &= \hat{\mathbf{C}}(0)^{-1} \hat{\mathbf{C}}(\tau) \mathbf{v} \\ &= \hat{\mathbf{C}}(0)^{-1} \left( \frac{1}{N} \mathbf{X}^\top \mathbf{Y} \mathbf{v} \right) \\ &= \hat{\mathbf{C}}(0)^{-1} \left( \frac{1}{N} \mathbf{X}^\top \mathbf{X} \mathbf{v} \right) \\ &= \hat{\mathbf{C}}(0)^{-1} \hat{\mathbf{C}}(0) \mathbf{v} \\ &= \mathbf{v}. \end{aligned} \quad (\text{E2})$$

So 1 is an eigenvalue of  $\mathbf{K}$  and the eigenvalue problem (30) can be exactly solved.

We now show that the weight function  $w(\mathbf{x})$  remains fixed after replacing  $\mathbf{K}$  by  $\mathbf{K}_{\text{rev}}$ , i.e.,

$$\hat{\mathbf{C}}(0)^{-1} \mathbf{K}_{\text{rev}}^\top \hat{\mathbf{C}}(0) \mathbf{u} = \mathbf{u} \quad (\text{E3})$$

for  $\mathbf{u}$  satisfying  $\mathbf{1}^\top \mathbf{X} \mathbf{u} = 1$  and

$$\hat{\mathbf{C}}(0)^{-1} \mathbf{K}^\top \hat{\mathbf{C}}(0) \mathbf{u} = \mathbf{u} \Leftrightarrow \mathbf{Y}^\top \mathbf{X} \mathbf{u} = \mathbf{X}^\top \mathbf{X} \mathbf{u}. \quad (\text{E4})$$

Considering that

$$\begin{aligned} \hat{\mathbf{C}}_{\text{rev}}(0) \mathbf{v} &= \frac{1}{2} \left( \mathbf{X}^\top \mathbf{W} \mathbf{X} + \mathbf{Y}^\top \mathbf{W} \mathbf{Y} \right) \mathbf{v} \\ &= \frac{1}{2} \left( \mathbf{X}^\top \mathbf{X} \mathbf{u} + \mathbf{Y}^\top \mathbf{X} \mathbf{u} \right) \\ &= N \hat{\mathbf{C}}(0) \mathbf{u}. \end{aligned} \quad (\text{E5})$$

Thus,

$$\begin{aligned} \hat{\mathbf{C}}(0)^{-1} \mathbf{K}_{\text{rev}}^\top \hat{\mathbf{C}}(0) \mathbf{u} &= \frac{1}{N} \hat{\mathbf{C}}(0)^{-1} \hat{\mathbf{C}}_{\text{rev}}(\tau) \hat{\mathbf{C}}_{\text{rev}}(0)^{-1} \hat{\mathbf{C}}_{\text{rev}}(0) \mathbf{v} \\ &= \frac{1}{2N} \hat{\mathbf{C}}(0)^{-1} \left( \mathbf{X}^\top \mathbf{W} \mathbf{Y} + \mathbf{Y}^\top \mathbf{W} \mathbf{X} \right) \mathbf{v} \\ &= \mathbf{u}. \end{aligned} \quad (\text{E6})$$

### APPENDIX F: DE-CORRELATION OF BASIS FUNCTIONS

In Section III B and Algorithm 1, the basis functions  $\chi$  are assumed to be linearly independent of the sampled data so that  $\hat{\mathbf{C}}(0)$  is invertible and the matrix  $\mathbf{K}$  given in (25) is well defined. In some publications, e.g., Ref. 72,  $\mathbf{K}$  is calculated as  $\mathbf{K} = \hat{\mathbf{C}}(0)^\dagger \hat{\mathbf{C}}(\tau)$  by using the pseudo-inverse  $\hat{\mathbf{C}}(0)^\dagger$  of  $\hat{\mathbf{C}}(0)$ ; however, this approach cannot completely avoid numerical instabilities. In this paper, we utilize principal component analysis (PCA)<sup>83</sup> to explicitly reduce correlations between basis functions as follows: First, we compute the empirical mean of basis functions and the covariance matrix of mean-centered basis functions,

Algorithm 4. Nonreversible Koopman estimation (with de-correlation of basis functions).

1. Basis-transform input coordinates according to (11).
2. Compute  $\boldsymbol{\pi}$  and COV by (F1) and (F2).
3. Let  $\boldsymbol{\chi} := \text{DC}[\boldsymbol{\chi}|\boldsymbol{\pi}, \text{COV}]$ , and recalculate  $\mathbf{X}$  and  $\mathbf{Y}$  according to the new basis functions.
4. Compute the matrix  $\mathbf{K} = \frac{1}{N}\mathbf{X}^\top\mathbf{Y}$  and solve the eigenvalue problem  $\mathbf{K}\mathbf{B} = \mathbf{B}\boldsymbol{\Lambda}$ .
5. Output spectral components: Eigenvalues  $\lambda_i$  and eigenfunctions  $\psi_i = \mathbf{b}_i^\top\boldsymbol{\chi}$ .

$$\boldsymbol{\pi} = \frac{1}{N}\mathbf{X}^\top\mathbf{1}, \quad (\text{F1})$$

$$\text{COV} = \frac{1}{N}\mathbf{X}^\top\mathbf{X} - \boldsymbol{\pi}\boldsymbol{\pi}^\top. \quad (\text{F2})$$

Next, perform the truncated eigendecomposition of the covariance matrix as

$$\text{COV} \approx \mathbf{Q}_d^\top \mathbf{S}_d \mathbf{Q}_d, \quad (\text{F3})$$

where the diagonal of matrix  $\mathbf{S}_d$  contains all positive eigenvalues that are larger than  $\epsilon_0$  and absolute values of all negative eigenvalues ( $\epsilon_0 = 10^{-10}$  in our applications). Last, the new basis functions are given by

$$\boldsymbol{\chi}^{\text{new}} = \begin{bmatrix} \mathbf{S}_d^{-0.5} \mathbf{Q}_d (\boldsymbol{\chi} - \boldsymbol{\pi}) \\ \mathbb{1} \end{bmatrix}. \quad (\text{F4})$$

Here  $\mathbf{S}_d^{-0.5} \mathbf{Q}_d (\boldsymbol{\chi} - \boldsymbol{\pi})$  is the PCA whitening transformation of the original basis functions, which transforms  $\boldsymbol{\chi}$  into all available principal components and scales each component to have a variance of 1, and the constant function  $\mathbb{1}$  is added to the basis function so that the eigenvalue problem (30) in the estimation of equilibrium distribution can be exactly solved (see Appendix E). It can be verified that the direct estimate  $\hat{\mathbf{C}}(0)$  of the covariance matrix obtained from  $\boldsymbol{\chi}^{\text{new}}(\mathbf{x}_t)$  is an identity matrix. The corresponding estimate of the Koopman operator is given by  $\mathbf{K} = \hat{\mathbf{C}}(0) = \frac{1}{N}\mathbf{X}^\top\mathbf{Y}$ .

For convenience of notation, we denote the transformation (F4) by

$$\boldsymbol{\chi}^{\text{new}} = \text{DC}[\boldsymbol{\chi}|\boldsymbol{\pi}, \text{COV}]. \quad (\text{F5})$$

Then the nonreversible Koopman estimation, which also perform EDMD, linear VA, and TICA in feature space, can be robustly implemented in Algorithm 4.

Furthermore, Koopman reweighting (Algorithm 2) can be robustly implemented in Algorithm 5.

Similarly, we can also guarantee the positive-definiteness of  $\hat{\mathbf{C}}_{\text{rev}}(0)$  by de-correlation of basis functions based on the

Algorithm 5. Koopman reweighting (with de-correlation of basis functions).

1. Basis-transform input coordinates according to (11).
2. Compute  $\mathbf{K}$  as in Algorithm 4.
3. Compute  $\mathbf{u}$  by solving  $\mathbf{K}^\top\mathbf{u} = \mathbf{u}$  and normalize it by  $\mathbf{1}^\top\mathbf{X}\mathbf{u}$ .
4. Output weights:  $w(\mathbf{x}_t) = \mathbf{x}_t^\top\mathbf{u}$ .

Algorithm 6. Variational Koopman model and eigendecomposition (with de-correlation of basis functions).

1. Basis-transform input coordinates according to (11).
2. Compute  $\mathbf{u}$  as in Algorithm 5 and let  $\mathbf{W} = \text{diag}(\mathbf{X}\mathbf{u})$ .
3. Compute  $\boldsymbol{\pi}_{\text{eq}}$  and  $\text{COV}_{\text{eq}}$  by (F7) and (F8).
4. Let  $\boldsymbol{\chi} := \text{DC}[\boldsymbol{\chi}|\boldsymbol{\pi}_{\text{eq}}, \text{COV}_{\text{eq}}]$ , and recalculate  $\mathbf{X}$  and  $\mathbf{Y}$  according to the new basis functions.
5. Compute  $\mathbf{K}_{\text{rev}} = \hat{\mathbf{C}}_{\text{rev}}(\tau) = \frac{1}{2}(\mathbf{X}^\top\mathbf{W}\mathbf{Y} + \mathbf{Y}^\top\mathbf{W}\mathbf{X})$  and solve the eigenvalue problem  $\mathbf{K}_{\text{rev}}\mathbf{B} = \mathbf{B}\boldsymbol{\Lambda}$ .
6. Output spectral components: Eigenvalues  $\lambda_i$  and eigenfunctions  $\psi_i = \mathbf{b}_i^\top\boldsymbol{\chi}$ .

transformation,

$$\boldsymbol{\chi}^{\text{new}} = \text{DC}[\boldsymbol{\chi}|\boldsymbol{\pi}_{\text{eq}}, \text{COV}_{\text{eq}}], \quad (\text{F6})$$

where

$$\boldsymbol{\pi}_{\text{eq}} = \frac{1}{2}(\mathbf{X} + \mathbf{Y})^\top\mathbf{W}\mathbf{1}, \quad (\text{F7})$$

$$\text{COV}_{\text{eq}} = \frac{1}{2}(\mathbf{X}^\top\mathbf{W}\mathbf{X} + \mathbf{Y}^\top\mathbf{W}\mathbf{Y}) - \boldsymbol{\pi}_{\text{eq}}\boldsymbol{\pi}_{\text{eq}}^\top \quad (\text{F8})$$

are the estimated equilibrium mean and covariance matrix of  $\boldsymbol{\chi}$ . The corresponding reversible Koopman estimator which is consistent with the variational approach can be robustly implemented in Algorithm 6.

## APPENDIX G: SIMULATION MODELS AND SETUPS

### 1. One-dimensional diffusion process

The diffusion processes in Section IV A are driven by the Brownian dynamics,

$$d\mathbf{x}_t = -\nabla U(\mathbf{x}_t)d\mathbf{t} + \sqrt{2\beta^{-1}}d\mathbf{W}_t, \quad (\text{G1})$$

where  $\beta = 0.3$ , the time step is 0.002,  $x_0$  is uniformly drawn in  $[0, 0.2]$ , and the potential function is given by

$$U(x) = \frac{\sum_{i=1}^5 (|x - c_i| + 0.001)^{-2} u_i}{\sum_{i=1}^5 (|x - c_i| + 0.001)^{-2}} \quad (\text{G2})$$

with  $c_{1:5} = (-0.3, 0.5, 1, 1.5, 2.3)$ . Simulations are implemented by a reversibility preserving numerical discretization scheme proposed in Ref. 84 with bin size 0.02. The basis functions for estimators are chosen to be

$$\chi_i(x) = \exp(-(w_i x + b_i)^2), \quad (\text{G3})$$

where  $w_i$  and  $b_i$  are randomly drawn in  $[-1, 1]$  and  $[0, 1]$ .

### 2. Two-dimensional diffusion process

The dynamics of the two-dimensional diffusion process in Section IV B has the same form as (G1), where  $\beta = 0.5$ , sample interval is 0.05,  $\mathbf{x}_0 = (x_0, y_0)$  is uniformly drawn in  $[-2, -1.5] \times [-1.5, 2.5]$ , and the potential function is chosen as in Ref. 85 by

$$\begin{aligned}
 U(x, y) = & 3 \exp\left(-x^2 - \left(y - \frac{1}{3}\right)^2\right) - 3 \exp\left(-x^2 - \left(y - \frac{5}{3}\right)^2\right) \\
 & - 5 \exp\left(-(x-1)^2 - y^2\right) - 5 \exp\left(-(x+1)^2 - y^2\right) \\
 & + \frac{1}{5}x^4 + \frac{1}{5}\left(y - \frac{1}{3}\right)^4. \quad (\text{G4})
 \end{aligned}$$

Simulations are implemented by the same algorithm as in [Appendix G 1](#) with bin size  $0.2 \times 0.2$ . The basis functions for estimators are also Gaussian functions,

$$\chi_i(\mathbf{x}) = \exp\left(-\left(\mathbf{w}_i^\top \mathbf{x} + b_i\right)^2\right), \quad (\text{G5})$$

with random weights  $\mathbf{w}_i \in [-1, 1] \times [-1, 1]$  and  $b_i \in [0, 1]$ .

- <sup>1</sup>M. Shirts and V. S. Pande, *Science* **290**, 1903 (2000).  
<sup>2</sup>J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, *J. Comput. Chem.* **26**, 1781 (2005).  
<sup>3</sup>M. Harvey, G. Giupponi, and G. D. Fabritiis, *J. Chem. Theory Comput.* **5**, 1632 (2009).  
<sup>4</sup>I. Buch, M. J. Harvey, T. Giorgino, D. P. Anderson, and G. De Fabritiis, *J. Chem. Inf. Model.* **50**, 397 (2010).  
<sup>5</sup>D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. Dror, M. Eastwood, J. Bank, J. Jumper, J. Salmon, Y. Shan, and W. Wriggers, *Science* **330**, 341 (2010).  
<sup>6</sup>P. Eastman, M. S. Friedrichs, J. D. Chodera, R. J. Radmer, C. M. Bruns, J. P. Ku, K. A. Beauchamp, T. J. Lane, L.-P. Wang, D. Shukla, T. Tye, M. Houston, T. Stich, C. Klein, M. R. Shirts, and V. S. Pande, *J. Chem. Theory Comput.* **9**, 461 (2013).  
<sup>7</sup>S. L. Grand, A. W. Goetz, and R. C. Walker, *Chem. Phys. Commun.* **184**, 374 (2013).  
<sup>8</sup>S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl, *Bioinformatics* **29**, 845 (2013).  
<sup>9</sup>S. Doerr, M. J. Harvey, F. Noé, and G. D. Fabritiis, *J. Chem. Theory Comput.* **12**, 1845 (2016).  
<sup>10</sup>F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl, *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19011 (2009).  
<sup>11</sup>I. Buch, T. Giorgino, and G. De Fabritiis, *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10184 (2011).  
<sup>12</sup>G. R. Bowman, V. A. Voelz, and V. S. Pande, *J. Am. Chem. Soc.* **133**, 664 (2011).  
<sup>13</sup>S. K. Sadiq, F. Noé, and G. De Fabritiis, *Proc. Natl. Acad. Sci. U. S. A.* **109**, 20449 (2012).  
<sup>14</sup>D.-A. Silva, D. R. Weiss, F. P. Avila, L.-T. Da, M. Levitt, D. Wang, and X. Huang, *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7665 (2014).  
<sup>15</sup>D. Shukla, Y. Meng, B. Roux, and V. S. Pande, *Nat. Commun.* **5**, 3397 (2014).  
<sup>16</sup>N. Plattner and F. Noé, *Nat. Commun.* **6**, 7653 (2015).  
<sup>17</sup>T. F. Reubold, K. Faelber, N. Plattner, Y. Posor, K. Branz, U. Curth, J. Schlegel, R. Anand, D. Manstein, F. Noé, V. Haucke, O. Daumke, and S. Eschenburg, *Nature* **525**, 404 (2015).  
<sup>18</sup>C. Schütte, A. Fischer, W. Huisinga, and P. Deuffhard, *J. Comput. Phys.* **151**, 146 (1999).  
<sup>19</sup>W. C. Swope, J. W. Pitera, and F. Suits, *J. Phys. Chem. B* **108**, 6571 (2004).  
<sup>20</sup>F. Noé, I. Horenko, C. Schütte, and J. C. Smith, *J. Chem. Phys.* **126**, 155102 (2007).  
<sup>21</sup>J. D. Chodera, K. A. Dill, N. Singhal, V. S. Pande, W. C. Swope, and J. W. Pitera, *J. Chem. Phys.* **126**, 155101 (2007).  
<sup>22</sup>G. R. Bowman, K. A. Beauchamp, G. Boxer, and V. S. Pande, *J. Chem. Phys.* **131**, 124101 (2009).  
<sup>23</sup>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).  
<sup>24</sup>C. Schütte, F. Noé, J. Lu, M. Sarich, and E. Vanden-Eijnden, *J. Chem. Phys.* **134**, 204105 (2011).  
<sup>25</sup>*An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, in *Advances in Experimental Medicine and Biology* Vol. 797, edited by G. R. Bowman, V. S. Pande, and F. Noé (Springer, Heidelberg, 2014).  
<sup>26</sup>D. S. Chekmarev, T. Ishida, and R. M. Levy, *J. Phys. Chem. B* **108**, 19487 (2004).  
<sup>27</sup>S. Sriraman, I. G. Kevrekidis, and G. Hummer, *J. Phys. Chem. B* **109**, 6479 (2005).  
<sup>28</sup>N. V. Buchete and G. Hummer, *J. Phys. Chem. B* **112**, 6057 (2008).  
<sup>29</sup>F. Noé, H. Wu, J.-H. Prinz, and N. Plattner, *J. Chem. Phys.* **139**, 184114 (2013).  
<sup>30</sup>H. Wu and F. Noé, *Multiscale Model. Simul.* **12**, 25 (2014).  
<sup>31</sup>E. Rosta and G. Hummer, *J. Chem. Theory Comput.* **11**, 276 (2015).  
<sup>32</sup>H. Wu and F. Noé, *J. Chem. Phys.* **142**, 084104 (2015).  
<sup>33</sup>H. Wu, F. Paul, C. Wehmeyer, and F. Noé, *Proc. Natl. Acad. Sci. U. S. A.* **113**, E3221 (2016).  
<sup>34</sup>M. Sarich, F. Noé, and C. Schütte, *Multiscale Model. Simul.* **8**, 1154 (2010).  
<sup>35</sup>P. Metzner, C. Schütte, and E. Vanden-Eijnden, *Multiscale Model. Simul.* **7**, 1192 (2009).  
<sup>36</sup>A. Berezhkovskii, G. Hummer, and A. Szabo, *J. Chem. Phys.* **130**, 205102 (2009).  
<sup>37</sup>F. Noé, S. Doose, I. Daidone, M. Löllmann, J. D. Chodera, M. Sauer, and J. C. Smith, *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4822 (2011).  
<sup>38</sup>B. G. Keller, J.-H. Prinz, and F. Noé, *Chem. Phys.* **396**, 92 (2012).  
<sup>39</sup>W. Zhuang, R. Z. Cui, D.-A. Silva, and X. Huang, *J. Phys. Chem. B* **115**, 5415 (2011).  
<sup>40</sup>B. Lindner, Z. Yi, J.-H. Prinz, J. C. Smith, and F. Noé, *J. Chem. Phys.* **139**, 175101 (2013).  
<sup>41</sup>B. Peters and B. L. Trout, *J. Chem. Phys.* **125**, 054108 (2006).  
<sup>42</sup>B. Peters, *J. Chem. Phys.* **125**, 241101 (2006).  
<sup>43</sup>J. D. Chodera and V. S. Pande, *Phys. Rev. Lett.* **107**, 098102 (2011).  
<sup>44</sup>M. A. Rohrdanz, W. Zheng, M. Maggioni, and C. Clementi, *J. Chem. Phys.* **134**, 124116 (2011).  
<sup>45</sup>M. A. Rohrdanz, W. Zheng, and C. Clementi, *Annu. Rev. Phys. Chem.* **64**, 295 (2013).  
<sup>46</sup>R. T. McGibbon, B. E. Husic, and V. S. Pande, *J. Chem. Phys.* **146**(4), 044109 (2017).  
<sup>47</sup>B. Trendelkamp-Schroer, H. Wu, F. Paul, and F. Noé, *J. Chem. Phys.* **143**, 174101 (2015).  
<sup>48</sup>A. Bittracher, P. Koltai, and O. Junge, *J. Appl. Dyn. Syst.* **14**, 1478 (2015).  
<sup>49</sup>F. Noé, *J. Chem. Phys.* **128**, 244103 (2008).  
<sup>50</sup>B. Trendelkamp-Schroer and F. Noé, *J. Phys. Chem.* **138**, 164113 (2013).  
<sup>51</sup>F. Noé and F. Nüske, *Multiscale Model. Simul.* **11**, 635 (2013).  
<sup>52</sup>G. Perez-Hernandez, F. Paul, T. Giorgino, G. D. Fabritiis, and F. Noé, *J. Chem. Phys.* **139**, 015102 (2013).  
<sup>53</sup>F. Nüske, B. G. Keller, G. Pérez-Hernández, A. S. J. S. Mey, and F. Noé, *J. Chem. Theory Comput.* **10**, 1739 (2014).  
<sup>54</sup>F. Vitalini, F. Noé, and B. G. Keller, *J. Chem. Theory Comput.* **11**, 3992 (2015).  
<sup>55</sup>S. M. W. Huisinga and C. Schuette, *Ann. Appl. Probab.* **14**, 419–458 (2004).  
<sup>56</sup>P. Deuffhard and M. Weber, ZIB Report 03–09 (2003).  
<sup>57</sup>S. Kube and M. Weber, *J. Chem. Phys.* **126**, 024103 (2007).  
<sup>58</sup>M. Cameron and E. Vanden-Eijnden, *J. Stat. Phys.* **156**, 427 (2014).  
<sup>59</sup>R. T. McGibbon and V. S. Pande, *J. Chem. Phys.* **142**, 124105 (2015).  
<sup>60</sup>L. Molgedey and H. G. Schuster, *Phys. Rev. Lett.* **72**, 3634 (1994).  
<sup>61</sup>A. Ziehe and K.-R. Müller, *ICANN 98* (Springer Science and Business Media, 1998), pp. 675–680.  
<sup>62</sup>E. O. Aapo Hyvärinen and J. Karhunen, *Independent Component Analysis* (John Wiley & Sons, 2001).  
<sup>63</sup>C. R. Schwantes and V. S. Pande, *J. Chem. Theory Comput.* **9**, 2000 (2013).  
<sup>64</sup>Y. Naritomi and S. Fuchigami, *J. Chem. Phys.* **134**, 065101 (2011).  
<sup>65</sup>L. Boninsegna, G. Gobbo, F. Noé, and C. Clementi, *J. Chem. Theory Comput.* **11**, 5947 (2015).  
<sup>66</sup>F. Noé and C. Clementi, *J. Chem. Theory Comput.* **11**, 5002 (2015).  
<sup>67</sup>F. Noé, R. Banisch, and C. Clementi, *J. Chem. Theory Comput.* **12**, 5620 (2016).  
<sup>68</sup>S. Harmeling, A. Ziehe, M. Kawanabe, and K.-R. Müller, *Neural Comput.* **15**, 1089 (2003).  
<sup>69</sup>C. R. Schwantes and V. S. Pande, *J. Chem. Theory Comput.* **11**, 600 (2015).  
<sup>70</sup>F. Nüske, R. Schneider, F. Vitalini, and F. Noé, *J. Chem. Phys.* **144**, 054105 (2016).  
<sup>71</sup>M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Perez-Hernandez, M. Hoffmann, N. Plattner, J.-H. Prinz, and F. Noé, *J. Chem. Theory Comput.* **11**, 5525 (2015).  
<sup>72</sup>M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, *J. Nonlinear Sci.* **25**, 1307 (2015).



- <sup>73</sup>P. J. Schmid and J. Sesterhenn, in *61st Annual Meeting of the APS Division of Fluid Dynamics* (American Physical Society, 2008).
- <sup>74</sup>C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, *J. Fluid Mech.* **641**, 115 (2009).
- <sup>75</sup>P. J. Schmid, *J. Fluid Mech.* **656**, 5 (2010).
- <sup>76</sup>J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, *J. Comput. Dyn.* **1**, 391 (2014).
- <sup>77</sup>I. Mezić, *Nonlinear Dyn.* **41**, 309 (2005).
- <sup>78</sup>S. Klus, P. Koltai, and C. Schütte, *J. Comput. Dyn.* **3**(1), 51–79 (2016).
- <sup>79</sup>O. Ledoit and M. Wolf, *J. Multivar. Anal.* **88**, 365 (2004).
- <sup>80</sup>J. Schäfer, K. Strimmer *et al.*, *Stat. Appl. Genet. Molec. Biol.* **4**, 32 (2005).
- <sup>81</sup>O. James and H.-N. Lee, e-print [arXiv:1402.6757](https://arxiv.org/abs/1402.6757) (2014).
- <sup>82</sup>P. W. Glynn and D. L. Iglehart, *Manag. Sci.* **35**, 1367 (1989).
- <sup>83</sup>K. Pearson, *Philos. Mag.* **2**, 559 (1901).
- <sup>84</sup>J. C. Latorre, P. Metzner, C. Hartmann, and C. Schütte, *Commun. Math. Sci.* **9**, 1051 (2011).
- <sup>85</sup>P. Metzner, C. Schütte, and E. Vanden-Eijnden, *J. Chem. Phys.* **125**, 084110 (2006).