



## UvA-DARE (Digital Academic Repository)

### A maximum caliber approach for continuum path ensembles

Bolhuis, P.G.; Brotzakis, Z.F.; Vendruscolo, M.

**DOI**

[10.1140/epjb/s10051-021-00154-3](https://doi.org/10.1140/epjb/s10051-021-00154-3)

**Publication date**

2021

**Document Version**

Final published version

**Published in**

The European Physical Journal B. Condensed Matter Physics

**License**

CC BY

[Link to publication](#)

**Citation for published version (APA):**

Bolhuis, P. G., Brotzakis, Z. F., & Vendruscolo, M. (2021). A maximum caliber approach for continuum path ensembles. *The European Physical Journal B. Condensed Matter Physics*, 94(9), Article 188. <https://doi.org/10.1140/epjb/s10051-021-00154-3>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



# A maximum caliber approach for continuum path ensembles

Peter G. Bolhuis<sup>1,a</sup> , Z. Faidon Brotzakis<sup>2</sup> , and Michele Vendruscolo<sup>2</sup> 

<sup>1</sup> van 't Hoff Institute for Molecular Sciences, University of Amsterdam, PO Box 94157, 1090 GD Amsterdam, The Netherlands

<sup>2</sup> Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK

Received 23 April 2021 / Accepted 5 July 2021 / Published online 23 September 2021  
© The Author(s) 2021

**Abstract.** The maximum caliber approach implements the maximum entropy principle for trajectories by maximizing a path entropy under external constraints. The maximum caliber approach can be applied to a diverse set of equilibrium and non-equilibrium problems concerning the properties of trajectories connecting different states of a system. In this review, we recapitulate the basic concepts of the maximum entropy principle and of its maximum caliber implementation for path ensembles, and review recent applications of this approach. In particular, we describe how we recently used this approach to introduce a framework, called here the continuum path ensemble maximum caliber (CoPE-MaxCal) method, to impose kinetic constraints in molecular simulations, for instance to include experimental information about transition rates. Such incorporation of dynamical information can ameliorate inaccuracies of empirical force fields, and lead to improved mechanistic insights. We conclude by offering an outlook for future research.

## 1 Introduction

The maximum entropy principle is a general variational principle to maximize the entropy of a system under given constraints [1, 2]. This principle provides an effective route to a statistical mechanical description of complex molecular systems. The maximum caliber approach represents an implementation of the maximum entropy principle for path ensembles [3, 4]. This approach is becoming increasingly used to deal with the equilibrium and non-equilibrium statistical mechanics of trajectories [5–11].

The trajectories of a molecular system can be produced using computer simulation methods, as for instance molecular dynamics. Such simulations have become a standard tool for predicting equilibrium properties of molecular systems in structural biology, physics, chemistry, and material science, yielding stable and metastable structures and their populations, as given by the free energy differences between states. While the application of molecular dynamics provides such information by sampling the Boltzmann distribution, the timescales that such simulations can access are often insufficient to observe the reactive processes of interest [12–15]. This is known as the sampling problem, and is a well-known source of statistical and systematic errors [16]. Therefore, enhanced sampling is often needed to overcome the high free energy barriers that are underlying the long timescales [13, 17, 18]. A second source of error arises from the approxima-

tions used in the definition of the force field used in the simulations. In the last decade, it has become therefore common to correct for force field inaccuracies in a system-dependent manner by incorporating experimental data as restraints in the simulations. To implement this strategy, powerful tools have been developed based on the maximum entropy principle [19–30] leading to numerous applications for example in protein folding and assembly [31–36] as well as in cases where force fields are less accurate [37], such as for intrinsically disordered proteins (IDPs) and nucleic acids [28, 30, 38–44].

The maximum entropy (MaxEnt) principle can also be applied to accurately determine dynamical processes in complex systems undergoing rare transitions, such as biomolecular isomerization, association, self-assembly or nucleation phenomena. In these approaches, experimental rate constants or time-dependent data are incorporated in discrete time, biased molecular dynamics simulations using the maximum caliber (MaxCal) approach [9, 45] (for an explanation of the terminology see Table 1). We recently explored a new avenue by developing a method based on the MaxCal approach, which utilizes unbiased atomistic classical molecular dynamics in combination with trajectory-based rare event methods. This framework can be used to ameliorate the effects of force field errors on the kinetic properties of complex molecular systems [46].

Dill and coworkers [5, 9, 11] have reviewed the broad scope of applications of the MaxCal approach in the area of non-equilibrium dynamics, including Green–Kubo relations, and Onsager’s reciprocal relations, as

<sup>a</sup> e-mail: [p.g.bolhuis@uva.nl](mailto:p.g.bolhuis@uva.nl) (corresponding author)

**Table 1** Terminology used in this work

MaxEnt principle	Maximum entropy principle
<i>Configuration space</i>	
MaxEnt approach	MaxEnt principle applied to configuration space
MaxEnt method	A method based on the MaxEnt approach
<i>Trajectory space</i>	
MaxCal approach	Application of the MaxEnt principle to discrete or continuous trajectory space
CoPE-MaxCal	Application of the MaxCal approach to continuous path ensembles
CoPE-MaxCal for transition rates	Application of the CoPE-MaxCal method for imposing transition rates

well as applications in equilibrium situations, including the optimization of Markov state models [47]. The latter approach aims to infer the interconversion rate matrix between metastable states from limited information on the population of the states themselves. Using MaxCal optimization, a classical-mechanics connection between trajectories and Langevin/Fokker–Planck parameters was recently derived [8]. Furthermore, a way to construct such Markov state models for non-equilibrium steady state processes was also reported [48, 49].

These approaches concern discrete systems [9], in the sense that continuous molecular dynamics trajectories are discretized into matrices of transition probabilities between metastable states, which can then be analyzed with MaxCal approaches. In many cases, however, one would like to keep the microscopic atomistic information of the pathways, and use the concept of altering the path probabilities, also known as trajectory reweighting. This concept is well known in the literature, e.g., in large deviation theory [50, 51] and path reweighting [48, 52, 53]. However, the application of the MaxCal approach to continuous space trajectories has proven difficult.

We recently introduced the continuum path ensemble maximum caliber (CoPE-MaxCal) method to reweight existing ensembles of trajectories to match experimental rate constants [46]. Given an ensemble of pathways harvested by unbiased molecular dynamics or enhanced trajectory sampling methods that do not bias the dynamics, we introduced a consistent reweighting scheme in which not only the reactive paths are assigned altered weights, but also all other paths in the ensemble, so that they match the corrected free energy landscape obtained from a simultaneous MaxEnt-based reweighting. Transition Path Sampling (TPS) [54–56] and the related Transition Interface Sampling (TIS) [57–59] techniques are rigorous enhanced trajectory sampling methods that can provide the required path ensembles. These methods elucidated the kinetics of a variety of complex molecular systems, including biomolecular conformational changes [60–75], ion dissociation [76–78] and dissolution [79], water isomerization [80] and evaporation [81], polymer collapse [82, 83], nucleation processes [84–92], polymorph transitions [93–95], chemical reactions [96, 97], enzyme reaction kinetics [98–101], colloidal cluster assembly [85, 102–104], and glassy systems [51, 105–110], leading in many cases to novel insights. Other trajectory-based methods such as Forward Flux

Sampling [111, 112] or Weighted Ensemble [113–115], stochastic process rare event sampling [116], milestoneing [117, 118], adaptive multilevel splitting [119], and non-equilibrium umbrella sampling [120], can in principle also be used to collect a prior path ensembles.

An important feature of the CoPE-MaxCal method is that it can be applied a posteriori, i.e., by post-processing an existing prior path ensemble. Importantly, such an approach addresses simultaneously the sampling and the force field problems. Thus, the CoPE-MaxCal method enables a system-dependent characterization when experimental data are available, without solving the extremely challenging problem of optimizing the force field itself in a system-independent, transferable manner, even though this might be preferred from a fundamental point of view.

In this review, we provide a perspective of the CoPE-MaxCal method by explaining the basic MaxEnt and MaxCal concepts, and how they lead to the CoPE-MaxCal reweighting scheme. Our aim is to recapitulate, and partly recast, this method, put it in perspective, and to provide an outlook on future research directions. The remainder of the paper is organized as follows. In Sects. 2 and 3, we introduce the MaxEnt principle and the MaxCal approach, and describe how they apply to continuous time and coordinate spaces. Section 4 reviews the specific application to the incorporation of rate constants. Section 5 discusses several extensions, and connects some loose ends. We end with a conclusion and an outlook.

## 2 Maximum entropy

### 2.1 The maximum entropy principle

According to the MaxEnt principle [1, 2], the condition of equilibrium for a system is achieved by a maximization of the Shannon entropy

$$S = - \sum_i p_i \ln p_i, \quad (1)$$

where  $p_i$  is the probability distribution for state  $i$ . Note that this entropy is related to the Gibbs entropy by a factor  $k_B$ . For a closed and isolated system with only the constraint of an overall normalization, obtaining the

probability distribution that optimizes the entropy can efficiently be done by the method of Lagrange multipliers. The Lagrange function is

$$L = - \sum_i p_i \ln p_i - \nu \left( \sum_i p_i - 1 \right), \tag{2}$$

where  $\nu$  is the Lagrange multiplier, which is used to impose the normalization condition. The extremum of this function follows from setting to zero the derivative with respect to  $p_i$

$$\frac{\delta S}{\delta p_i} = - \ln p_i - 1 - \nu = 0. \tag{3}$$

Solving this for  $p_i$  gives

$$p_i = e^{-\nu-1}, \tag{4}$$

which is a constant factor. Applying the normalization condition leads to

$$\sum_i p_i = \sum_i e^{-\nu-1} = W e^{-\nu-1} = 1, \tag{5}$$

where  $W$  is the number of accessible states, so that the Lagrange multiplier is given by  $e^{\nu+1} = W$ , leading to the probability  $p_i = 1/W$ . This condition corresponds to the standard assumption in statistical mechanics that in a closed and isolated system each microstate has the same probability, which corresponds to the Boltzmann identity  $S = \ln W$ , in units of the Boltzmann constant  $k_B$ .

Often the MaxEnt principle is concerned with the relative entropy, which is

$$S = - \sum_i p_i \ln \frac{p_i}{p_i^0}, \tag{6}$$

where  $p^0$  denotes the prior distribution, e.g., the constant probability of the microcanonical ensemble. The relative entropy is the negative of the Kullback–Leibler (KL) divergence, and reaches its maximum of zero only when the prior and posterior distribution are identical. The MaxEnt principle can be applied subject to additional constraints, for instance that an average of a fluctuating observable  $s_i$  is fixed to a certain value

$$\sum_i p_i s_i = s. \tag{7}$$

The MaxEnt principle states that in equilibrium, the entropy should be maximized subject to this constraint. Using the method of Lagrange multipliers, the optimization function becomes

$$L = - \sum_i p_i \ln \frac{p_i}{p_i^0} - \eta \left( \sum_i p_i s_i - s \right) - \nu \left( \sum_i p_i - 1 \right),$$

which again can be optimized by setting the derivative to zero

$$\frac{\delta S}{\delta p_i} = - \ln \frac{p_i}{p_i^0} - 1 - \eta s_i - \nu = 0. \tag{8}$$

Rearranging leads to

$$p_i \propto p_0 e^{-\eta s_i}, \tag{9}$$

where we ignored the normalization constant set by  $\nu$ . Assuming a constant prior probability  $p_0$ , and normalizing, gives the distribution

$$p_i = \frac{e^{-\eta s_i}}{Z}, \tag{10}$$

where the partition function is

$$Z = \sum_i e^{-\eta s_i}. \tag{11}$$

As an example, we can take the energy  $E_i$  of the system to be the observable  $s$ , and identify  $\eta$  as  $\beta$ . This leads to

$$p_i = \frac{e^{-\beta E_i}}{\sum_i e^{-\beta E_i}}, \tag{12}$$

which is the Boltzmann distribution, with  $\beta = 1/k_B T$  being the inverse temperature of the surrounding reservoir that the system can exchange heat with. In addition, the derivative of  $Z$  with respect to the variable  $\eta$

$$\frac{\partial \ln Z}{\partial \eta} = - \langle s_i \rangle, \tag{13}$$

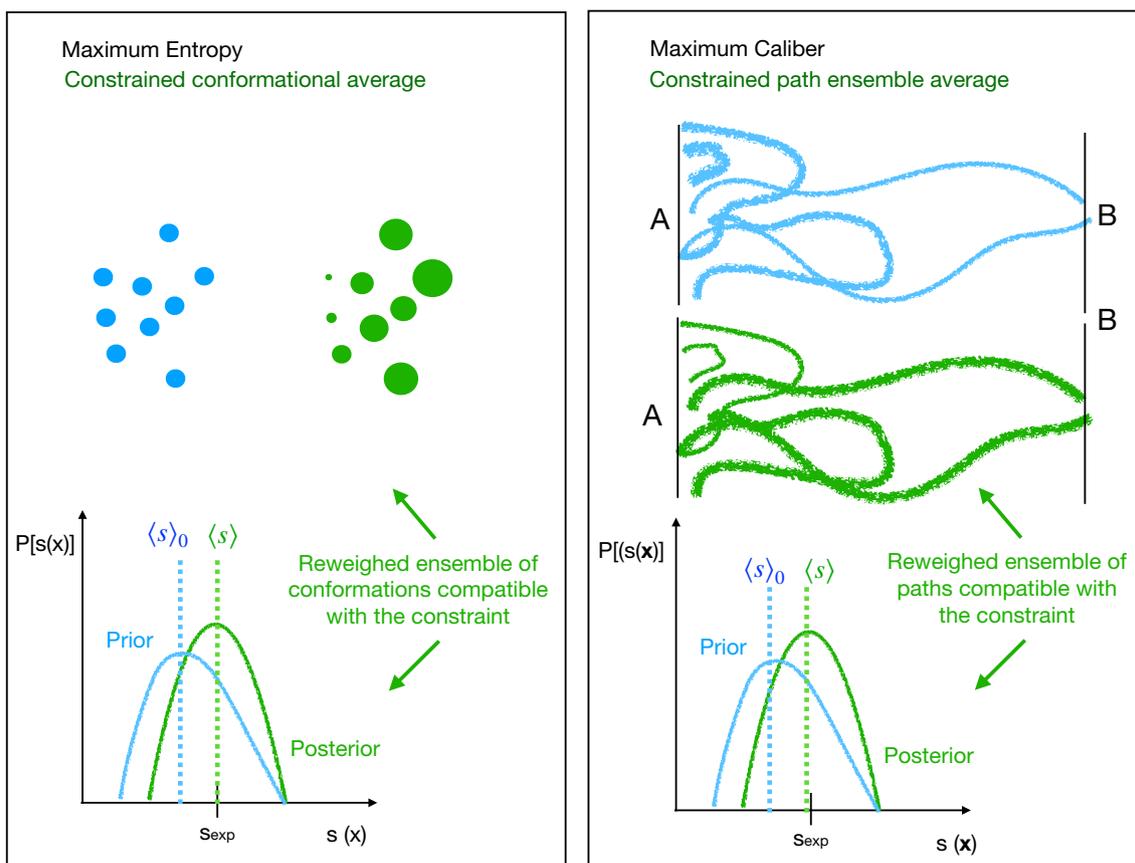
gives the average of the observable.

In summary, the above outline illustrates how applying the MaxEnt principle to systems with a large number of degrees of freedom leads to standard statistical mechanics. For further details, we refer to Refs. [1, 2, 9].

### 2.2 MaxEnt in configuration space

In this section, we discuss the application of the MaxEnt principle in continuous space [28, 29]. Consider a continuous state space  $x$ , where  $x$  denotes the positions and velocities of all the particles in a system. The maximum relative entropy principle states that, given a prior probability distribution (density)  $\rho^0(x)$ , the optimal distribution  $\rho(x)$  maximally compatible with a set of given experimental constraints is the one that which maximizes the entropy, or the negative of the KL divergence

$$S[\rho||\rho^0] = - \int dx \rho(x) \ln \frac{\rho(x)}{\rho^0(x)}. \tag{14}$$



**Fig. 1** Schematic comparison of the MaxEnt and MaxCal approaches for reweighting continuous space conformational and path ensembles, respectively. Left: The prior conformational ensemble (blue circles) changes its weights of conformations, illustrated by the different size of the circles in the posterior ensemble (green circles). This change of weights also alters the distribution along a conformational observable  $s(x)$  and shifts the prior conformational ensemble average  $\langle s(x) \rangle_0$  to the posterior average  $\langle s(x) \rangle = s_{exp}$  (in green). Right: The same principle holds for path reweighting, where now instead of a conformational ensemble, a prior path ensemble (blue paths) is reweighted (green paths) according to a constraint in an experimental (dynamical) observable, such as a kinetic rate constant

The entropy should, thus, be maximized subject to  $M$  experimental constraints

$$\rho^{ME}(x) = \underset{P(x)}{\operatorname{argmax}} S[\rho || \rho^0]$$

$$\text{subject to: } \begin{cases} \int dx \rho(x) s_i(x) = \langle s_i(x) \rangle = s_i^{exp} \\ \int dx \rho(x) = 1, \end{cases} \quad (16)$$

where the  $s_i$  denote the experimental observables ( $1 \leq i \leq M$ ), and  $s_i^{exp}$  is the value of the experimental constraint. The last line is the normalization condition. Again, the relative entropy is always non-positive, and zero only if the two distributions are identical. In Fig. 1 we illustrate the application of the MaxEnt principle to the reweighting of the prior conformational distribution, thereby identifying a posterior distribution which meets a given experimental constraints  $s_i^{exp}$ . The maximization can be carried out using the method of the

Lagrangian multipliers, by looking for the stationary point of the Lagrange function

$$\mathcal{L} = S[\rho || \rho^0] - \sum_{i=1}^M \mu_i \left( \int dx s_i(x) \rho(x) - s_i^{exp} \right) - \nu \left( \int dx \rho(x) - 1 \right), \quad (17)$$

where the  $\mu_i$  and  $\nu$  are the multipliers, which are related to each of the  $M$  experimental constraints, and the normalization condition, respectively. A functional derivative with respect to  $\rho(x)$  gives

$$\frac{\delta \mathcal{L}}{\delta \rho(x)} = -\ln \frac{\rho(x)}{\rho^0(x)} - 1 - \sum_{i=1}^M \mu_i s_i(x) - \nu. \quad (18)$$

Setting the derivative to zero and solving for  $\rho(x)$  leads to

$$\rho^{ME}(x) = \frac{1}{Z} e^{-\sum_{i=1}^M \mu_i s_i(x)} \rho^0(x), \quad (19)$$

where  $Z$  is a normalization factor akin to the partition function,

$$Z = \int dx \rho^0(x) e^{-\sum_{i=1}^M \mu_i s_i(x)}. \tag{20}$$

The posterior MaxEnt density is, thus, a reweighing of the prior distribution with an exponential factor involving the experimental observable.

The unknown multipliers can be determined by noting that the solution of the Lagrange function is equivalent to minimizing the following function with respect to  $\mu$ , as shown by Cesari and coworkers [28]

$$\Gamma(\boldsymbol{\mu}) = -\ln Z + \boldsymbol{\mu} \cdot \mathbf{s}^{exp}, \tag{21}$$

where bold symbols denote a vector of the  $M$  multipliers and observables. This can be seen as a Legendre transformation. Minimization of  $\Gamma(\boldsymbol{\mu})$  with respect to each multiplier yields a set of  $M$  equations for the ensemble averages in the posterior ensemble

$$\frac{\int dx \rho^0(x) e^{-\sum_{i=1}^M \mu_i s_i(x)} s_i(x)}{\int dx \rho^0(x) e^{-\sum_{i=1}^M \mu_i s_i(x)}} = s_i^{exp}, \tag{22}$$

which can be solved to give the sought multipliers  $\mu_i$ . For further details on the MaxEnt approach, we refer the interested reader to Ref. [28].

### 2.3 MaxEnt approach for equilibrium constants

In this review, we are concerned with the application of the MaxEnt principle to incorporate kinetic information, for example rate constants, in the characterization of a system. While the phase space distributions in the MaxEnt approach preclude the description of time dependent properties, one can use MaxEnt to impose the *ratio* of rate constants, which is related to the equilibrium constant.

In what follows, we focus on systems that exhibit two-state kinetics between two well-defined stable states, A and B, for which there is a separation between the molecular timescale and the reaction time [12, 13]. This guarantees that well-defined rate constants,  $k_{AB}$  and  $k_{BA}$ , exist for the interconversion from A to B, and vice versa. Using the detailed balance condition  $k_{AB}\pi_A = k_{BA}\pi_B$ , with  $\pi$  being the equilibrium population of the two states, it follows that

$$K_{eq} = \frac{\pi_B}{\pi_A} = \frac{k_{AB}}{k_{BA}}.$$

This equilibrium constant is related to the equilibrium fraction  $K = \pi_B/(\pi_A + \pi_B) = 1/(1 + K_{eq})$ , which reduces to  $K = \pi_B$ , for normalized populations. This is an equilibrium quantity that can be treated with the MaxEnt procedure described above by setting  $M = 1$  and setting the  $s_1^{exp} = \pi_B$ . Next, we should find a microscopic function  $s_1(x)$  that can measure the equilibrium

fraction  $K$ . In the following, we denote this microscopic function  $g(x)$ . Substituting these variables in Eq. 22 gives

$$\frac{\int dx g(x) \rho(x)}{\int dx \rho(x)} = \pi_B, \tag{23}$$

which states that the ensemble average of  $\langle g(x) \rangle$  should be equal to  $\pi_B$ , the fraction of configurations in state B. Therefore,  $g(x)$  should be a function that determines whether or not a configuration is part of B or not. A possible, quite natural, definition is the (configurational) committor function  $p_B(x)$ , which gives the probability for ending in state B, when a trajectory is initiated in  $x$  with random velocities [56, 76, 121] (see Fig. 2). While this is not an easy function to compute, it is well defined and extensively studied [56, 122, 123]. Note that the committor is also sometimes denoted Onsager’s splitting probability [124], or p-fold [125] in the context of protein folding. Here, following the transition path theory convention [126], we denote  $p_B(x)$  as the committor function. Using the committor function  $p_B(x)$ , it is possible to compute the density in the basin of attraction of state A, and that of state B, respectively, as

$$\rho_A(x) = \rho(x)p_A(x) \tag{24}$$

$$\rho_B(x) = \rho(x)p_B(x), \tag{25}$$

where  $p_A(x) = 1 - p_B(x)$ . Setting  $g(x) = p_B(x)$ , we now can apply the MaxEnt principle to obtain the reweighted posterior distributions

$$\rho_A(x) = \rho_A^0(x) e^{-\mu_A p_A(x)} \tag{26}$$

$$\rho_B(x) = \rho_B^0(x) e^{-\mu_B p_B(x)}. \tag{27}$$

The posterior densities will alter the committor function as [127]

$$p_B(x) = \rho_B(x)/(\rho_A(x) + \rho_B(x)), \tag{28}$$

with respect to the original prior committor

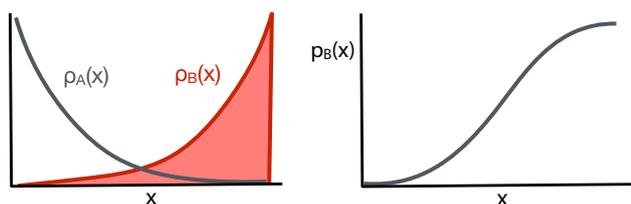
$$p_B^0(x) = \rho_B^0(x)/(\rho_A^0(x) + \rho_B^0(x)), \tag{29}$$

or, by substituting the densities

$$p_B(x) = \frac{\rho_B^0(x)}{\rho_A^0(x) e^{-\mu_A} e^{(\mu_A + \mu_B) p_B(x)} + \rho_B^0(x)}. \tag{30}$$

As we need the  $g(x)$  to be correct in the posterior ensemble average, it is clear that  $g(x) \equiv p_B(x)$  is not equal to  $p_B^0$ , but instead follows from numerically solving Eq. 30.

The multipliers  $\mu_A$  and  $\mu_B$  derive from the ensemble average conditions, and are related to the ratio of the



**Fig. 2** Schematic illustration of how the configurational density in state  $B$  (red area) is related to the committor through Eq. 28

equilibrium constants as

$$e^{\mu_A - \mu_B} = K_{eq}^{exp} / K_{eq}^0 = \frac{\pi_B^{exp} \pi_A^0}{\pi_A^{exp} \pi_B^0}. \quad (31)$$

Using the definitions in terms of rates, this is equal to

$$e^{\mu_A - \mu_B} = \frac{k_{AB}^{exp} / k_{BA}^0}{k_{BA}^{exp} / k_{AB}^0}. \quad (32)$$

We can then identify the multipliers  $\mu_A$  and  $\mu_B$  as

$$e^{\mu_A} = \frac{k_{AB}^{exp}}{k_{AB}^0} \quad e^{\mu_B} = \frac{k_{BA}^{exp}}{k_{BA}^0}. \quad (33)$$

We will later see that this is correct by applying the MaxCal approach.

### 3 Maximum caliber

#### 3.1 The maximum caliber approach

In 1980, Jaynes proposed an extension of the maximum entropy principle for trajectories [3]. Following the notation of Dill and collaborators [5, 9, 11], the path entropy is defined as

$$S = - \sum_{\Gamma} \ln \frac{p_{\Gamma}}{q_{\Gamma}}, \quad (34)$$

where  $\Gamma = \{X_0, X_1, \dots, X_L\}$  denotes a sequence of discrete state variables  $X$ , at different time indices,  $p_{\Gamma}$  is the probability of such sequence, and  $q_{\Gamma}$  denotes a prior or reference distribution. The summation is over all possible sequences of states, or trajectories.

Consider a dynamical feature  $s(\Gamma)$  in the system that can be computed over the trajectory ensemble

$$\langle s \rangle = \sum_{\Gamma} p_{\Gamma} s(\Gamma). \quad (35)$$

According to the MaxCal approach, imposing this average as a constraint on a system can be carried out according to the maximum entropy principle, by optimizing the distribution  $p_{\Gamma}$  such that it maximizes the

path entropy in Eq. 34. To do so, we again use the method of Lagrange multipliers. The optimization function, or caliber is

$$\mathcal{L} = - \sum_{\Gamma} \ln \frac{p_{\Gamma}}{q_{\Gamma}} - \eta \left( \sum_{\Gamma} p_{\Gamma} s(\Gamma) - s_i \right) - \nu \left( \sum_{\Gamma} p_{\Gamma} - 1 \right). \quad (36)$$

This constrained caliber can, in analogy with the MaxEnt approach, be optimized by setting the derivative with respect to  $p_{\Gamma}$  to zero, and solving for  $p_{\Gamma}$ ,

$$p_{\Gamma} = \frac{e^{-\eta s(\Gamma)}}{\mathcal{Z}}, \quad (37)$$

with the partition function analog

$$\mathcal{Z} = \sum_{\Gamma} e^{-\eta s(\Gamma)}. \quad (38)$$

The path ensemble average of  $s$  is recovered by taking the derivative of  $\mathcal{Z}$  with respect to  $\eta$

$$\frac{\partial \ln \mathcal{Z}}{\partial \eta} = - \langle s_i \rangle. \quad (39)$$

Solving for  $\eta$  gives the value of the Lagrange multiplier.

Dill and collaborators reviewed the MaxCal approach extensively, and illustrated how it can be invoked to obtain general principles for non-equilibrium systems, including the Onsager's and Green-Kubo relations, Prigogine's entropy production, and the diffusion and Fokker-Planck equations. In addition, the MaxCal approach can be used for optimization of transition rate matrices in Markov state models, for equilibrium and non-equilibrium steady states [48, 49]. The MaxCal approach also has been used for reaction coordinate optimization [128, 129].

These studies define the state space as a discrete space, which makes the approach suitable for Markov state models, but less so for molecular dynamics trajectories in continuous space. However, as we are interested in atomistic microscopic trajectories, our purpose is to recast the MaxCal approach in continuous space.

#### 3.2 Continuum path ensemble (CoPE) MaxCal

In this section, we discuss how the MaxCal approach can be extended to the space of continuous trajectories. Such space can be mapped using molecular dynamics simulations, Langevin dynamics or Brownian dynamics. A trajectory can be denoted as an ordered sequence of frames  $\mathbf{x} = \{x_0, x_1, \dots, x_L\}$ , where  $x$  is a continuous space configuration as in Sect. 2.2, and the subscripts denote the time index. Each frame is separated from the next by a short time interval  $\Delta t$ , so that the total duration of a trajectory is  $\mathcal{T} = L\Delta t$ . The probability distribution

**Table 2** Definitions of variables and functions used in this work

CV	Description
$\rho(x)$	A configurational probability distribution, with $x$ denoting the configuration
$\rho_0^0(x)$	The prior configurational probability distribution
$\rho^{ME}(x)$	The maximum entropy posterior configurational probability distribution, after applying the constraints
$s_i(x)$	Collective variable/observable $i$ as a function of the microscopic configurations.
$s_i^{exp}$	The experimental observation/data $i$
$k_{AB}^{exp}$	Experimental observation/data of transition $k_{AB}$
$\sigma_i$	Level of confidence in the data $i$
$\sigma_{k_{AB}}$	Level of confidence in the rate constant data $k_{AB}$
$\mu_i$	Lagrange multiplier associated with experimental data $i$
$\mu_{A,B}$	Lagrange multiplier associated with experimental rate constant data for A ( $k_{AB}$ ), or B ( $k_{BA}$ )
$p(x_i \rightarrow x_{i+1})$	Short time Markovian probability representing the evolution of the probability density
$\mathcal{P}[\mathbf{x}]$	A full path distribution, with $\mathbf{x}$ denoting the trajectory
$\mathcal{P}^0[\mathbf{x}]$	The prior full path distribution
$\mathcal{P}^{MC}[\mathbf{x}]$	The maximum caliber posterior path distribution, after applying the constraints
$h_{A,B}(x)$	Indicator function identifying whether a configuration is in state (A,B)
$\mathcal{P}_{A,B}[\mathbf{x}]$	A partial path distribution of trajectories starting in A, or B
$\mathcal{P}_{A,B}^0[\mathbf{x}]$	The prior partial path distribution of trajectories starting in A, B
$\mathcal{P}_A^{MC}[\mathbf{x}]$	The MaxCal posterior partial path distribution of trajectories starting in A, after applying the constraints related to state A, i.e $k_{AB}$
$\mathcal{P}_B^{MC}[\mathbf{x}]$	The MaxCal posterior partial path distribution of trajectories starting in B, after applying the constraints related to state B i.e $k_{BA}$
$\lambda$	Collective variable used to parameterize ordered set of interfaces between state A and B
$\lambda(x)$	Value of collective variable $\lambda$ evaluated at configuration $x$
$\lambda_{max}[\mathbf{x}]$	Maximum value of $\lambda$ along an individual path $\mathbf{x}$
$\lambda_{min}[\mathbf{x}]$	Minimum value of $\lambda$ along an individual path $\mathbf{x}$
$P_A(\lambda_n \lambda_i)$	Probability that a trajectory originating from A and crossing interface $i$ , reaches interface $n$ (i.e., B)
$P_A(\lambda_i \lambda_0)$	Probability that a trajectory originating from A and crossing the first interface ( $\lambda_0$ ), reaches interface $i$
$P_A(\lambda \lambda_0)$	Probability that a trajectory originating from A and crossing the first interface, reaches a value $\lambda$ .
$\phi_0$	Flux through innermost interface $\lambda_0$ for trajectories starting in A
$f_A(\lambda \mathbf{x})$	MaxCal biasing/reweighting function for pathways starting in A
$f_B(\lambda \mathbf{x})$	MaxCal biasing/reweighting function for pathways starting in B
$P_A^0(\lambda \lambda_0)$	Probability a trajectory of the prior path ensemble originating from A and crossing the first interface, reaches a value $\lambda$
$P_A^{MC}(\lambda \lambda_0)$	Probability a trajectory of the MaxCal posterior path ensemble originating from A and crossing the first interface, reaches a value $\lambda$
$g(\lambda)$	MaxEnt biasing function of the configurational densities that reassures the experimentally determined equilibrium constant
$g^{exp}$	Experimentally determined value of the configurational density observable, i.e. the equilibrium constant
$\rho(\lambda)$	Configurational density projected onto $\lambda$
$\rho^0(\lambda)$	The prior configurational density projected onto $\lambda$
$\rho^{MC}(\lambda)$	The MaxCal configurational density projected onto $\lambda$
$\rho_{A,B}(\lambda)$	A configurational density projected onto $\lambda$ for trajectories originating in A, B
$\rho_{A,B}^0(\lambda)$	The prior configurational density projected onto $\lambda$ for trajectories originating in A, B
$\rho_{A,B}^{ME}(\lambda)$	The MaxEnt configurational density projected onto $\lambda$ for trajectories originating in A, B
$\rho_{A,B}^{MC}(\lambda)$	The MaxCal configurational density projected onto $\lambda$ for pathways originating in A, B
$\hat{\rho}(\lambda \mathbf{x})$	Instantaneous density operator projecting the path $\mathbf{x}$ on $\lambda$
$R_A^0(\lambda \lambda_0)$	The reaching histogram for pathways originating from state A, crossing the innermost interface $\lambda_0$ and just reach the value of $\lambda$
$R_B^0(\lambda \lambda_n)$	The reaching histogram for pathways originating from state B, crossing the innermost interface $\lambda_n$ and just reach the value of $\lambda$
$K_{eq}$	The equilibrium constant
$K$	The equilibrium fraction
$K_{eq}^{exp}$	The experimental equilibrium constant
$K_{eq}^0$	The prior equilibrium constant
$p_B(x)$	The probability that a trajectory initiated at configuration $x$ (with random momenta) ends up in B, a.k.a. the committor
$p_B(\lambda)$	The probability to end up in B, a.k.a. the committor, projected onto the collective variable $\lambda$
$S(\mathcal{P}  \mathcal{P}^0)$	Relative path entropy, or Caliber, equivalent to negative KL divergence between path distributions $\mathcal{P}$ and $\mathcal{P}^0$

for a trajectory  $x$  is [17, 55, 56]

$$\mathcal{P}^0[\mathbf{x}] = \rho(x_0) \prod_{i=0}^{L-1} p(x_i \rightarrow x_{i+1}), \quad (40)$$

where  $\rho(x_0)$  is the initial distribution, and  $p(x_i \rightarrow x_{i+1})$  is the short time Markovian propagator that describes the transition from state  $x_i$  to the next  $x_{i+1}$  in the time interval.

In molecular dynamics, the initial condition is often the Boltzmann distribution  $\rho(x) = \exp(-\beta H(x))$ , with  $H(x)$  the Hamiltonian of the system. The short-time Markovian probability  $p(x_i \rightarrow x_{i+1})$  depends on the dynamics used, for example for molecular dynamics in the microcanonical ensemble this probability is a  $\delta$  function given by the molecular dynamics integrator or propagator  $\phi$ ,  $p(x_i \rightarrow x_{i+1}) = \delta(x_{i+1} - \phi(x_i))$  [56]. For stochastic dynamics  $p(x_i \rightarrow x_{i+1}) \propto e^{-\delta x_R^2 / (2\sigma^2)}$ , where  $\delta_R$  denotes the random displacement in the stochastic (Wiener) process, and  $\sigma^2$  is the variance of that displacement [54]. One can also define a Metropolis–Hasting Monte Carlo-based dynamics  $p(x_i \rightarrow x_{i+1}) = \min[1, e^{-\beta(H(x_{i+1}) - H(x_i))}]$ , where the min function returns the lower of its arguments [54]. Besides enabling the estimate of equilibrium properties, molecular dynamics trajectories can also offer information on dynamical and non-equilibrium properties. Since this type of study is affected by the approximations involved in the definition of the force field, here we are concerned with improving such description by incorporating in the simulations dynamical information in the form of constraints on dynamical properties such as the rate constants.

The path entropy is now defined as

$$S[\mathcal{P}||\mathcal{P}^0] = - \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] \ln \frac{\mathcal{P}[\mathbf{x}]}{\mathcal{P}^0[\mathbf{x}]}, \quad (41)$$

where  $\mathcal{D}\mathbf{x}$  indicates an integral over all trajectories or paths  $\mathbf{x}$ . Following MaxEnt for continuous space, the MaxCal approach states that the optimal distribution  $\mathcal{P}[\mathbf{x}]$  is given by the maximization

$$\begin{aligned} \mathcal{P}^{MC}[\mathbf{x}] &= \operatorname{argmax}_{\mathcal{P}[\mathbf{x}]} S[\mathcal{P}||\mathcal{P}^0] \\ \text{subject to: } &\begin{cases} \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] s_i[\mathbf{x}] = \langle s_i[\mathbf{x}] \rangle = s_i^{exp} \\ \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] = 1. \end{cases} \end{aligned} \quad (42)$$

Again, the  $s_i^{exp}$  are experimental constraints encoded in the microscopic observable  $s_i[\mathbf{x}]$ . The ensemble average  $\langle s_i[\mathbf{x}] \rangle$  refers to either static/thermodynamic or dynamic/kinetic observables. Note that  $s[\mathbf{x}]$  is now a function of the entire path, which includes autocorrelation functions.

Following the same procedure as for the MaxEnt approach, we can perform this optimization by applying the method of Lagrange multipliers, and finding the

stationary point of the Lagrange function, or caliber

$$\begin{aligned} \mathcal{L} &= - \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] \ln \frac{\mathcal{P}[\mathbf{x}]}{\mathcal{P}^0[\mathbf{x}]} - \nu \left( \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] - 1 \right) \\ &\quad - \sum_i \mu_i \left( \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] s_i[\mathbf{x}] - s_i^{exp} \right), \end{aligned} \quad (43)$$

by taking the functional derivative with respect to the path distribution. This yields

$$\frac{\delta \mathcal{L}}{\delta \mathcal{P}[\mathbf{x}]} = - \ln \frac{\mathcal{P}[\mathbf{x}]}{\mathcal{P}^0[\mathbf{x}]} - 1 - \sum_i \mu_i s_i[\mathbf{x}] - \nu, \quad (44)$$

which gives in turn the posterior distribution

$$\mathcal{P}^{MC}[\mathbf{x}] \propto e^{-\sum_i \mu_i s_i[\mathbf{x}]} \mathcal{P}^0[\mathbf{x}], \quad (45)$$

as a reweighting of a given prior distribution  $\mathcal{P}^0[\mathbf{x}]$ . In Fig. 1 we illustrate the effect of the CoPE-MaxCal procedure in reweighting the prior path distribution, thereby identifying a posterior distribution which meets a given experimental constraints  $s_i^{exp}$ , which can be for instance a rate constant. It is important to realize that the trajectories are not changing, but only the weights with which they occur in the ensemble. This is analogous to the fact that the configurations are not changing in the MaxEnt procedure.

### 3.3 CoPE-MaxCal for dynamical constraints

In this section, we discuss the application of the CoPE-MaxCal approach to incorporate dynamical information as constraints in the description of a system. This can be achieved for instance using a correlation function  $c(t)$  defined as

$$c(t) = \langle s_i(0) s_j(t) \rangle = \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] s_i(x_0) s_j(x_\tau), \quad (46)$$

where the observable  $s_i$  at time  $t = 0$  is correlated with observable  $s_j$  a later time  $t$ . As  $i$  and  $j$  can be identical, this definition includes autocorrelation functions. In the second equation, we have used the explicit dependence  $s_{i,j}(x_\tau)$  as a function of the configuration  $x_\tau$ , with  $\tau = t/\Delta t$  corresponding to the frame index for time  $t$ .

Experimental information about the dynamics of a system in the form of correlation data  $c^{exp}(t)$  can be imposed as a constraint on the path ensemble distribution, giving the Lagrange function

$$\begin{aligned} \mathcal{L} &= - \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] \ln \frac{\mathcal{P}[\mathbf{x}]}{\mathcal{P}^0[\mathbf{x}]} - \nu \left( \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] - 1 \right) \\ &\quad - \sum_\tau \mu_\tau \left( \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] s_i(x_0) s_j(x_\tau) - c^{exp}(t) \right). \end{aligned} \quad (47)$$

Following the optimization procedure gives

$$\frac{\delta \mathcal{L}}{\delta \mathcal{P}[\mathbf{x}]} = -\ln \frac{\mathcal{P}[\mathbf{x}]}{\mathcal{P}^0[\mathbf{x}]} - 1 - \sum_{\tau} \mu_{\tau} s_i(x_0) s_j(x_{\tau}) - \nu, \tag{48}$$

yielding the posterior distribution

$$\mathcal{P}^{MC}[\mathbf{x}] \propto e^{-\sum_{\tau} \mu_{\tau} s_i(x_0) s_j(x_{\tau})} \mathcal{P}^0[\mathbf{x}]. \tag{49}$$

As an example, consider the mobility  $\mathcal{K}_{\tau}[\mathbf{x}]$ , measuring the mean square displacement at a particular time  $\tau$  with respect to time  $\tau = 0$ . As this correlation only has to be constrained at  $\tau$ , the posterior is

$$\mathcal{P}^{MC}[\mathbf{x}] \propto e^{-\mu_{\tau} \mathcal{K}_{\tau}[\mathbf{x}]} \mathcal{P}^0[\mathbf{x}]. \tag{50}$$

Note that this is also the expression for the s-ensemble (with  $\mu = s$ , where  $s$  should not be confused with the observable function  $s[\mathbf{x}]$ ). In the s-ensemble path ensembles are biased according to a time correlation function [51, 105–110]. This s-ensemble is usually presented in the context of large deviation theory, but clearly follows also from the MaxCal approach. The s-ensemble biases all paths with a field  $s$  conjugate to the function  $\mathcal{K}$ . In the MaxCal approach, the Lagrange multiplier  $\mu$  follows from the constraint imposed. Thus, the s-ensemble might also be interpreted as the field that imposes a certain constraint.

### 3.4 MaxCal for thermodynamic constraints

Since equilibrium properties are not time dependent, they can be computed as time averages over path ensembles distributions

$$\langle s \rangle = \frac{1}{\langle L \rangle} \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] \sum_t s(x_t), \tag{51}$$

with  $\langle L \rangle$  being the average path length, and  $x_t$  the coordinates at each time step of the path. Constraining an equilibrium property  $s^{exp}$  then leads to a posterior distribution

$$\mathcal{P}^{MC}[\mathbf{x}] \propto e^{-\mu \sum_t s(x_t)} \mathcal{P}^0[\mathbf{x}]. \tag{52}$$

An alternative way of constraining equilibrium properties is to first reduce the path space back to a configurational density  $\rho(x) \equiv P(x)$  by

$$\rho(x) \propto \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] \sum_t \delta(x_t - x). \tag{53}$$

The average then becomes

$$\langle s \rangle = \frac{\int dx \rho(x) s(x)}{\int dx \rho(x)}. \tag{54}$$

Indeed, substitution of Eqs. 52 and 53 in Eq. 54 yields the same result as Eq. 51.

## 4 Continuum path ensemble maximum caliber for rate constants

### 4.1 Independence of partial path distributions

In the above sections, we did not specify the path ensembles precisely. When considering systems that show two-state kinetics between two stable states, A and B, with forward and backward rate constants,  $k_{AB}$  and  $k_{BA}$  respectively, we can think about partial path ensembles  $\mathcal{P}_A[\mathbf{x}]$  and  $\mathcal{P}_B[\mathbf{x}]$ , consisting, respectively, of all paths that start in A and paths that start in B. We can define these partial ensembles as  $\mathcal{P}_A[\mathbf{x}] \equiv \mathcal{P}[\mathbf{x}] h_A(x_0)$  and  $\mathcal{P}_B[\mathbf{x}] \equiv \mathcal{P}[\mathbf{x}] h_B(x_0)$ , where  $h_{A,B}(x)$  are the indicator functions, which are unity when the configuration  $x$  is in state A(B), and zero otherwise. Restricting all paths to start and end in one of the stable states, the total path distribution is simply the sum of the non-normalized partial path distributions  $\mathcal{P}[\mathbf{x}] = \mathcal{P}_A[\mathbf{x}] + \mathcal{P}_B[\mathbf{x}]$ .

We apply different constraints  $s_A$  and  $s_B$  to both partial ensembles we arrive at the Lagrange function

$$\begin{aligned} \mathcal{L} = & - \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] \ln \frac{\mathcal{P}[\mathbf{x}]}{\mathcal{P}^0[\mathbf{x}]} - \nu \left( \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] - 1 \right) \\ & - \mu_A \left( \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] h_A(x_0) s_A[\mathbf{x}] - s_A^{exp} \right) \\ & - \mu_B \left( \int \mathcal{D}\mathbf{x} \mathcal{P}[\mathbf{x}] h_B(x_0) s_B[\mathbf{x}] - s_B^{exp} \right), \end{aligned} \tag{55}$$

where we used the definition of the partial ensembles. Maximization of the caliber yields the posterior

$$\mathcal{P}^{MC}[\mathbf{x}] \propto e^{-\mu_A h_A(x_0) s_A[\mathbf{x}] - \mu_B h_B(x_0) s_B[\mathbf{x}]} \mathcal{P}^0[\mathbf{x}], \tag{56}$$

or, expressing it in partial ensembles

$$\begin{aligned} & \mathcal{P}_A^{MC}[\mathbf{x}] + \mathcal{P}_B^{MC}[\mathbf{x}] \propto \\ & \propto e^{-\mu_A h_A(x_0) s_A[\mathbf{x}] - \mu_B h_B(x_0) s_B[\mathbf{x}]} (\mathcal{P}_A^0[\mathbf{x}] + \mathcal{P}_B^0[\mathbf{x}]). \end{aligned} \tag{57}$$

For paths belonging to partial ensemble A,  $h_A(x_0) = 1$  and thus  $h_B(x_0) = 0$ , leading to

$$\mathcal{P}_A^{MC}[\mathbf{x}] \propto e^{-\mu_A s_A[\mathbf{x}]} \mathcal{P}_A^0[\mathbf{x}], \tag{58}$$

while for paths from partial ensemble B,  $h_A(x_0) = 0$  and  $h_B(x_0) = 1$  it holds

$$\mathcal{P}_B^{MC}[\mathbf{x}] \propto e^{-\mu_B s_B[\mathbf{x}]} \mathcal{P}_B^0[\mathbf{x}]. \tag{59}$$

Thus, both partial ensembles can be optimized and normalized independently.

### 4.2 CoPE-MaxCal for rate constants

One can apply a rate constant as a kinetic constraint by setting  $s_A^{exp} \equiv k_{AB}^{exp}$ , and  $s_B^{exp} \equiv k_{BA}^{exp}$ . However, we also need a microscopic dynamical correlation function for this constraint

$$C(t) = \langle h_A(x_0)h_B(x_\tau) \rangle / \langle h_A(x_0) \rangle. \tag{60}$$

The observed rate constant is the time derivative of this function, measured at the plateau value [12]

$$k_{AB} = \frac{dC(t)}{dt} = \frac{\langle h_A(x_0)\dot{h}_B(x_\tau) \rangle}{\langle h_A(x_0) \rangle}. \tag{61}$$

The rate is, thus, the flux through the boundary of state B, given that the trajectory started in state A.

At this point, we are switching to the language of transition path sampling [54–56], and in particular transition interface sampling [57], as this allows us to write the rate constant in terms of the path probability  $\mathcal{P}[\mathbf{x}]$ . In the TIS method, the configuration space is foliated by a set of non-intersecting interfaces parametrized with a collective variable  $\lambda(x)$ . This set of  $n + 1$  interfaces is denoted  $\{\lambda_0, \lambda_1 \dots \lambda_n\}$ . Taking  $\lambda_0$  and  $\lambda_n$  as the boundary of A and B, respectively, the rate is expressed as the effective positive flux  $\phi$  through these interfaces

$$k_{AB} = \phi_n = \phi_0 P_A(\lambda_n | \lambda_0), \tag{62}$$

where  $\phi_i$  denotes the effective positive flux through interface  $i$ . In the second equation, we have defined the crossing probability  $P_A(\lambda_n | \lambda_0)$  for a trajectory to reach interface  $\lambda_n$ , under the condition that it has already crossed  $\lambda_0 = \lambda_A$ . Note that  $\phi_0$  is usually large, and easy to obtain from a molecular dynamics simulation in state A, while  $P_A(\lambda_n | \lambda_0)$  is usually very small, and difficult to compute. While it is in principle (and sometimes even in practice) possible to obtain this also from a brute force molecular dynamics simulation, it is much more efficient to compute via TPS [17, 54, 56, 130, 131], TIS [57–59], Forward Flux Sampling [111, 112], or other trajectory-based methods. In these sampling approaches, we can construct a path ensemble by reweighting the paths, yielding the so-called reweighted path ensemble [132], which constitutes a collection of paths with an associated weight, representing the probability to observe that path in an unbiased simulation, i.e.,  $\mathcal{P}_A[\mathbf{x}]$ . This reweighted path ensemble is related to the crossing probability by

$$P_A(\lambda | \lambda_0) = \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] \theta(\lambda_{max}[\mathbf{x}] - \lambda), \tag{63}$$

where  $\lambda_{max}[\mathbf{x}]$  returns the maximum value of  $\lambda(x)$  along the trajectory, and the Heaviside step function  $\theta$  returns unity for the paths that cross the  $\lambda$  interface, assuming that  $\lambda$  monotonically increases when going from A to B. This function can take the place of the microscopic

correlation function in the Lagrange function, giving

$$\begin{aligned} \mathcal{L} = & - \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] \ln \frac{\mathcal{P}_A[\mathbf{x}]}{\mathcal{P}_A^0[\mathbf{x}]} - \nu \left( \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] - 1 \right) \\ & - \mu'_A \left( \phi_0 \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] \theta(\lambda_{max}[\mathbf{x}] - \lambda_B) - k_{AB}^{exp} \right). \end{aligned} \tag{64}$$

Optimizing this function as before leads to the posterior

$$\mathcal{P}_A^{MC}[\mathbf{x}] \propto e^{-\mu'_A \phi_0 \theta(\lambda_{max}[\mathbf{x}] - \lambda_B)} \mathcal{P}_A^0[\mathbf{x}], \tag{65}$$

We can now solve for the Lagrange multiplier by constructing a log partition function as in Eq. 21

$$\Gamma^{MC}(\mu'_A) = \ln \left[ \int \mathcal{D}\mathbf{x} \mathcal{P}_A^{MC}[\mathbf{x}] \right] + \mu'_A k_{AB}^{exp}, \tag{66}$$

taking the derivative with respect to the  $\mu'_A$  and setting it to zero

$$\frac{- \int \mathcal{D}\mathbf{x} \mathcal{P}_A^{MC}[\mathbf{x}] \phi_0 \theta(\lambda_{max}[\mathbf{x}] - \lambda_B)}{\int \mathcal{D}\mathbf{x} \mathcal{P}_A^{MC}[\mathbf{x}]} + k_{AB}^{exp} = 0. \tag{67}$$

Since only the paths that reach the final interface are contributing to the  $\theta$  function, this changes into

$$\frac{\int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] e^{-\mu'_A \phi_0} \phi_0 \theta(\lambda_{max}[\mathbf{x}] - \lambda_B)}{\int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}]} = k_{AB}^{exp}, \tag{68}$$

where we replaced the  $\mathcal{P}$  with  $\mathcal{P}^0$  in the normalization, which hardly affects it. The exponent on the left hand side can be taken out of the integral, yielding

$$\begin{aligned} \frac{\int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] \phi_0 \theta(\lambda_{max}[\mathbf{x}] - \lambda_B)}{\int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}]} e^{-\mu'_A \phi_0} &= k_{AB}^0 \cdot e^{-\mu'_A \phi_0} \\ &= k_{AB}^{exp}. \end{aligned} \tag{69}$$

The exponential factor needs to be larger than unity for an increase in the rate, which is then associated with a negative Lagrange multiplier  $\mu'_A$ . As we prefer to associate a positive variable with an increase in rate, we define  $\mu_A \equiv -\mu'_A \phi_0$  yielding

$$k_{AB}^0 e^{\mu_A} = k_{AB}^{exp}, \tag{70}$$

Indeed, a positive  $\mu_A$  now increases the rate.

### 4.3 Imposing a rate constraint for all $\lambda$

In the above procedure, only the reactive AB paths in the prior ensemble are reweighted to yield a given rate constant. However, while this imposes the experimental rate constant, it does lead to an undesirable discontinuity at the boundary of  $\lambda_B$ . Moreover, the rate cannot

be strongly dependent on where precisely this boundary is located. More precisely, the rate can be expressed as the product of two crossing probabilities [57]

$$k_{AB} = \phi_0 P_A(\lambda_B|\lambda_i) P_A(\lambda_i|\lambda_0). \tag{71}$$

We can now interpret the  $\lambda_i$  as the interface that we put the kinetic constraint on. Since this location is arbitrary for  $0 < i < n$  we impose constraints on all these  $\lambda$  values simultaneously. This leads to the following Lagrange, or caliber, function

$$\begin{aligned} \mathcal{L} = & - \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] \ln \frac{\mathcal{P}_A[\mathbf{x}]}{\mathcal{P}_A^0[\mathbf{x}]} - \nu \left( \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] - 1 \right) - \sum_{i=0}^n \mu_i \\ & \times \left( \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] \theta(\lambda_{max}[\mathbf{x}] - \lambda_i) P_A(\lambda_n|\lambda_i) - k_{AB}^{exp} \right), \end{aligned} \tag{72}$$

where the sum runs over the  $n + 1$  interfaces. In each of the constraints in the summation the paths that cross  $\lambda_i$  are counted, and  $P_A(\lambda_n|\lambda_i)$  is the crossing probability to reach B from the  $\lambda_i$  interface. Optimization leads to the MaxCal posterior path distribution

$$\mathcal{P}_A^{MC}[\mathbf{x}] \propto \exp \left[ - \sum_{i=0}^n \mu_i \theta(\lambda_{max}[\mathbf{x}] - \lambda_i) P_A(\lambda_n|\lambda_i) \right] \mathcal{P}_A^0[\mathbf{x}]. \tag{73}$$

Note that the  $P_A(\lambda_n|\lambda_i)$  in this equation is dependent on the posterior distribution itself, and hence this is an implicit definition. The sum in the exponent is only dependent on  $\lambda_{max}[\mathbf{x}]$  (and of course on  $P_A(\lambda_n|\lambda)$ ), but for a given system  $P_A(\lambda_n|\lambda)$  is a function of  $\lambda$ , so the sum can be written as

$$- \sum_{i=0}^n \mu_i \theta(\lambda_{max}[\mathbf{x}] - \lambda_i) P_A(\lambda_n|\lambda_i) \equiv f_A(\lambda_{max}[\mathbf{x}]), \tag{74}$$

where the  $P_A(\lambda_n|\lambda)$  dependence is implicit in the function  $f_A$ . The interpretation is that the weight of each trajectory in the posterior path ensemble is entirely dependent on  $\lambda_{max}[\mathbf{x}]$ .

The  $n + 1$  constraints imposed are given for  $k = 0 \dots n$ ,

$$\int \mathcal{D}\mathbf{x} \mathcal{P}_A^{MC}[\mathbf{x}] \theta(\lambda_{max}[\mathbf{x}] - \lambda_k) P_A(\lambda_n|\lambda_k) = k_{AB}^{exp}, \tag{75}$$

with

$$\mathcal{P}_A^{MC}[\mathbf{x}] \propto e^{f_A(\lambda_{max}[\mathbf{x}])} \mathcal{P}_A^0[\mathbf{x}], \tag{76}$$

In Ref. [46] it is shown that these constraints can be fulfilled for any reasonably continuous function  $f_A(\lambda)$ , as long as  $f_A(\lambda_B) = \mu_A$ , and  $f_A(\lambda_0) = 0$ .

We can simplify the function by considering all paths that reach a maximum at  $\lambda_j$ . Due to the  $\theta$  function in

Eq. 74 only interfaces  $0 < i < j$  contribute to the sum, leading to

$$f_A(\lambda_j) = - \sum_{i=1}^j \mu_i P_A(\lambda_n|\lambda_i). \tag{77}$$

The function  $f_A(\lambda)$  will be determined in the next section. Figure 3 shows schematically how the TIS, reweighted path ensemble and CoPE-MaxCal approaches are related in terms of interface ensembles based on  $\lambda$ .

The prior path distributions can be projected on the collective variable (CV) or order parameter  $\lambda$ . In particular, the crossing probability  $P_A^0(\lambda|\lambda_0)$  and configurational density  $\rho_A^0(\lambda)$  are expressed as

$$P_A^0(\lambda|\lambda_0) = \int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] \theta(\lambda_{max}[\mathbf{x}] - \lambda), \tag{78}$$

$$\rho_A^0(\lambda) \propto \int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] \hat{\rho}(\lambda|\mathbf{x}), \tag{79}$$

where we defined the instantaneous density projection operator

$$\hat{\rho}(\lambda|\mathbf{x}) = \sum_{k=0}^{L[\mathbf{x}]} \delta(\lambda(x_k) - \lambda), \tag{80}$$

to project a single path  $\mathbf{x}$  on the CV  $\lambda$ . These projections can also be done for the posterior distribution, leading to the configurational density  $\rho_A^{MC}(\lambda)$

$$\rho_A^{MC}(\lambda) \propto \int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] e^{f_A(\lambda_{max}[\mathbf{x}])} \hat{\rho}(\lambda|\mathbf{x}), \tag{81}$$

and for the crossing probability

$$P_A^{MC}(\lambda|\lambda_0) = \int_{\lambda_n}^{\lambda} R_A^0(\lambda|\lambda_0) e^{f_A(\lambda)} d\lambda, \tag{82}$$

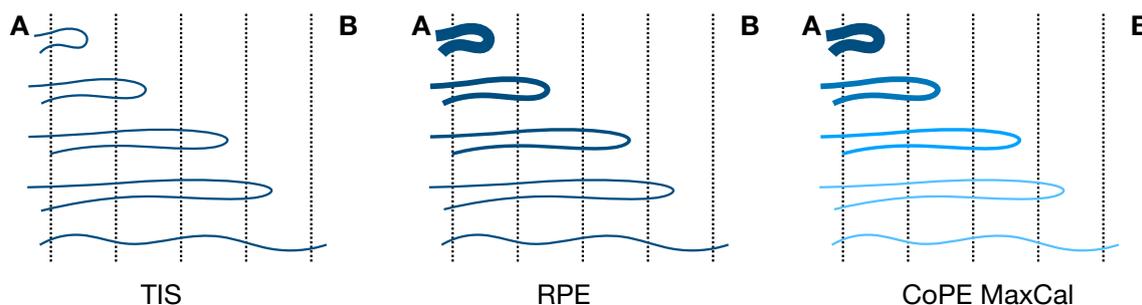
where we defined the prior probability  $R_A^0(\lambda|\lambda_0)$  for a path to just reach the interface  $\lambda$

$$R_A^0(\lambda|\lambda_0) = \int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] \delta(\lambda_{max}[\mathbf{x}] - \lambda). \tag{83}$$

#### 4.4 The MaxCal bias function $f_A(\lambda)$ follows from MaxEnt for the density

The function  $f_A(\lambda)$  above is not completely specified. However, we can make use of the density obtained by a MaxEnt reweighting of the same system. In particular, we can set

$$\rho_A^{MC}(\lambda) = \rho_A^{ME}(\lambda). \tag{84}$$



**Fig. 3** Schematic illustration of the relationship between the TIS, reweighted path ensemble (RPE) and CoPE-MaxCal approaches. (Left) TIS simulations samples paths that cross a certain interfaces. The single path shown for each interface represents a whole ensemble. (Center) The RPE follows by reweighting the path ensembles sampled in TIS. Weights are schematically indicated by the line thickness. (Right) After application of the CoPE-MaxCal to impose a rate constant constraint the paths in the RPE are reweighted with the  $f_A$  function, as indicated by color. Lighter colors indicate more strongly affected weights

In the MaxEnt Sect. 2.3, we imposed the equilibrium fraction by

$$\rho_A^{ME}(\lambda) = e^{-\mu_A g(\lambda)} \rho_A^0(\lambda), \quad (85)$$

where we projected the configurational density on the  $\lambda$  parameter. Like in Sect. 2.3, we can use the committor to impose the constraint,  $g(\lambda) = p_B(\lambda)$ , where  $p_B(\lambda)$  is now the projected committor. This committor follows from the implicit equation Eq. 30. This procedure leads to the following condition for the densities

$$\int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] e^{f_A(\lambda_{max}[\mathbf{x}])} \hat{\rho}(\lambda|\mathbf{x}) = e^{-\mu g(\lambda)} \int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] \hat{\rho}(\lambda|\mathbf{x}), \quad (86)$$

which should be solved numerically.

We can also use the full (configurational) committor  $p_B(x)$ , which gives

$$\int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] e^{f_A(p_B^{max}[\mathbf{x}])} \hat{\rho}(x|\mathbf{x}) = e^{-\mu p_B(x)} \int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] \hat{\rho}(x|\mathbf{x}), \quad (87)$$

where  $p_B^{max}[\mathbf{x}]$  returns the maximum committor value along the trajectory  $\mathbf{x}$ . Clearly, this might be very difficult to obtain in practice.

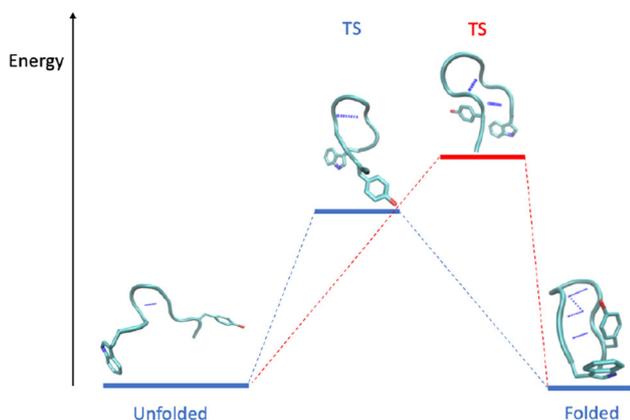
#### 4.5 Application of CoPE-MaxCal to protein folding

The CoPE-MaxCal approach enables the investigation of mechanistic properties for instance the location of the transition state, by imposing experimental transition rates. In this section we briefly highlight an application of the MaxCal approach on the folding of a small protein, which we presented in Ref. [46]. Taking information from a long molecular dynamics simulation [133], we constructed the prior path distributions and predicted rates [46]. Since the force field did not reproduce the folding rate correctly at the expected

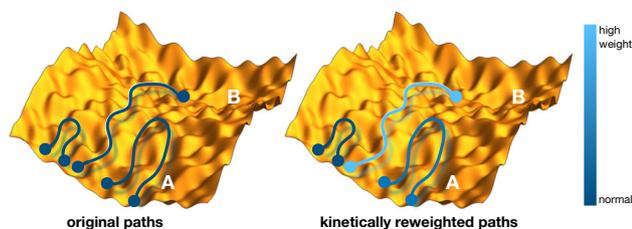
melting temperature, we imposed the correct experimental folding rate using the CoPE-MaxCal procedure [46]. The reweighted path ensembles led to a reweighted free energy and committor landscape. Importantly, the imposed rate caused a shift in predicted transition state (see Fig. 4), that provides a qualitative change in mechanistic insight. We refer to Ref. [46] for more details and examples.

#### 4.6 Interpretation of the CoPE-MaxCal method

The CoPE-MaxCal method takes as input an unbiased ensemble of paths from molecular dynamics simulations, or a reweighted path ensemble [132] from TIS [57] or Virtual Interface Exchange TPS [134], and reweights each trajectory in this ensemble according to how far it progresses along a predefined collective variable (see a rendering in Fig. 5). This includes the paths that cross the barrier and reach the final state, so the rate constants are automatically constrained to the correct value via the functions  $f_{A,B}(\lambda)$ . The more involved part of the framework is to ensure that the thermodynamic properties are correct, in particular the equilibrium constant. This requires a specific bias function  $g(\lambda)$  based on the committor function, which produces the least perturbed path ensemble, while still obeying the constraints (see Sect. 2.3). So, imposing  $g(\lambda)$  can be viewed as responsible for constraining equilibrium conditions, whereas  $f_{A,B}(\lambda)$  takes care of the dynamical corrections. The interpretation of the reweighting procedure is that trajectories are artificially made more, or less, probable in the path ensembles. This is analogous to changing the weight of each conformation in the Boltzmann distribution, using the MaxEnt approach. Note that the trajectories themselves do not change in the reweighting procedure. Rather, the distribution of initial conditions is adjusted. This is analogous, but not identical, to how microcanonical trajectories can be reweighted to give a canonical distributed path ensemble (see, e.g., Ref. [135]).



**Fig. 4** Illustration of the movement of the transition state for protein folding when the folding rate constant is imposed using the CoPE-MaxCal method



**Fig. 5** Cartoon of how the method reweights paths in a complex landscape

## 5 Extensions

### 5.1 Multiple state descriptions

The MaxCal method can be extended to multiple states. For multiple states, the kinetics is described by a rate matrix  $\mathbf{K}$  with entries  $k_{ij}$  for each pair of states  $i$  and  $j$ . Incorporating multiple rate constants in the ensembles adds additional constraints to the caliber function, which ensures that the fluxes through the final interfaces are correct, and yields the following path reweighting

$$\mathcal{P}_i^{MC}[\mathbf{x}] \propto e^{f_i(\lambda_i^{max}[\mathbf{x}])} \mathcal{P}_i^0[\mathbf{x}], \tag{88}$$

where the function  $f_i$  is now defined for each state, as function of the maximum value along an order parameter, that is possibly different for each state. Note that by changing the weight of state  $i$  all rates out of  $i$  are changed in the same way. Thus  $k_{i,j} = c_i k_{i,j}^0$ , i.e., all the rates from state  $i$  change proportionally by the same factor. This leads to a consistent description.

The other ingredient, the MaxEnt part, is based on the committor. Applying a correction to the imposed equilibrium constant  $K_{ij} = \pi_j/\pi_i = k_{ij}/k_{ji}$  means that the densities for each basin of attraction  $\rho_i(x)$  are reweighted by

$$\rho_i(x) = \rho_i^0(x) e^{-\mu_i p_i(x)}, \tag{89}$$

where  $p_i$  is the committor to state  $i$  with  $\sum_i p_i = 1$  [136]. Using the definition

$$p_i(x) = \frac{\rho_i(x)}{\sum_i \rho_i(x)}. \tag{90}$$

This implicit equation for  $p_i(x)$  can be solved numerically for a given  $\rho_i^0(x)$ . This gives rise to the MaxEnt reweighting by setting  $g_i(x) = p_i(x)$

$$\rho_i^{ME} = \rho_i^0(x) e^{-\mu_i p_i(x)}, \tag{91}$$

which in turn can be related to the MaxCal condition

$$\rho_i^{MC}(x) = e^{-\mu_i p_i(x)} \rho_i^0(x), \tag{92}$$

or

$$\begin{aligned} & \int \mathcal{D}\mathbf{x} \mathcal{P}_i^0[\mathbf{x}] e^{f_i(p_{i,min}[\mathbf{x}])} \hat{\rho}(x|\mathbf{x}) \\ &= e^{-\mu_i p_i(x)} \int \mathcal{D}\mathbf{x} \mathcal{P}_i^0[\mathbf{x}] \hat{\rho}(x|\mathbf{x}), \end{aligned} \tag{93}$$

where the minimum of the committor  $p_i$  to state  $i$  is searched for along the path. Note that we have to use the minimum along the pathway now, since the committor starts high, being close to the initial state, and then slowly decreases to zero as the reaction progresses.

It is also possible to project the committor to an order parameter  $\lambda_i$ , connected to each state. A numerical solution should then lead to a set of committor-based bias functions  $p_i(\lambda_i)$ . We leave the specific details for future work.

### 5.2 An alternative formulation of the MaxCal approach

In this section, we discuss whether one could establish a MaxCal approach by including the equilibrium density constraint directly in the MaxCal procedure. This would give an additional constraint to fulfill

$$\begin{aligned} & \zeta_A \left( \int dx \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] \hat{\rho}(x|\mathbf{x}) - \pi_A^{exp} \right) \\ &= \zeta_A \left( \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] L[\mathbf{x}] - \pi_A^{exp} \right). \end{aligned} \tag{94}$$

One could include this constraint in the MaxCal approach, which would lead to an additional weighting factor

$$\mathcal{P}[\mathbf{x}] = \mathcal{P}^0[\mathbf{x}] e^{f_A(\lambda_{max}[\mathbf{x}])} e^{-\zeta_A L[\mathbf{x}]}, \tag{95}$$

where the  $\zeta$  multiplier would be determined by solving

$$\int \mathcal{P}_i[\mathbf{x}] L[\mathbf{x}] = \pi_i, \tag{96}$$

and where one should use the non-normalized distributions to obtain the correct relative path ensemble weights. This could in principle be done but has a serious downside, namely that the paths are weighted with the instantaneous path length  $L[\mathbf{x}]$ , which is sensitive to the A and B definition, and hence seems arbitrary.

Moreover, this is not what we do in the CoPE-MaxCal approach. Instead, we state that the MaxCal method should reproduce the MaxEnt density  $\rho_A^{ME}(x)$  at each configuration  $x$ . This means that, in fact, we are introducing many constraints, one for each configuration point  $x$ , yielding an additional set of constraints

$$\begin{aligned} & \int dx \zeta_A(x) \left( \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] \hat{\rho}(x|\mathbf{x}) - \rho_A^{ME}(x) \right) \\ &= \int dx \zeta_A(x) \left( \int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] \hat{\rho}(x|\mathbf{x}) \right. \\ & \quad \left. - e^{-\mu_A g(x)} \int \mathcal{D}\mathbf{x} \mathcal{P}_A^0[\mathbf{x}] \hat{\rho}(x|\mathbf{x}) \right), \end{aligned} \tag{97}$$

where  $g(x)$  is again the MaxEnt biasing function defined in Sect. 2.3. This set of constraints ensures that the densities match at all configurations  $x$ . The path weight is then

$$\mathcal{P}[\mathbf{x}] \propto \mathcal{P}^0[\mathbf{x}] e^{f_A(\lambda_{max}[\mathbf{x}])} e^{-\int dx \zeta_A(x) \hat{\rho}(x|\mathbf{x})}, \tag{98}$$

where the  $\zeta_A(x)$  multiplier would be determined by solving the root of the derivative of the log partition function with respect to the  $\zeta_A(x)$

$$\int \mathcal{D}\mathbf{x} \mathcal{P}_A[\mathbf{x}] \hat{\rho}(x|\mathbf{x}) = \rho_A^{ME}(x). \tag{99}$$

Indeed, imposing this equation is at the heart of the CoPE-MaxCal method. Note that this implies that the equilibrium constant is also correctly reproduced. Note also that the constraint is stronger than what is needed from MaxCal alone. However, we stress for an equilibrium system the MaxEnt approach should give a distribution that is consistent with the MaxCal one.

### 5.3 Effects of the errors in the measurements

The MaxEnt principle can be used to model uncertainties in the experimental data [24, 28–30]. This is done by adding to the constraint average  $\langle s_i \rangle$  the expected error  $\langle e_i \rangle$  due to the perturbed distribution, leading to

$$\langle s_i \rangle = s_i^{exp} + \langle e_i \rangle. \tag{100}$$

When the error is Gaussian distributed with a standard deviation  $\sigma_i$ , this average expected error is given by

$$\langle e_i \rangle = -\mu_i \sigma_i^2, \tag{101}$$

where  $\sigma_i$  can be interpreted as the level of confidence in the experimental measurements or data. One can

include this into the Langrange function and obtain

$$\Gamma(\boldsymbol{\mu}) = \ln \left[ \int dx P^{ME}(x) \right] + \boldsymbol{\mu} \cdot \mathbf{s}^{exp} + \frac{1}{2} \sum_{i=1}^M \mu_i^2 \sigma_i^2. \tag{102}$$

Optimizing this function, and solving for the Lagrange multipliers  $\mu_i$  thus takes the level of confidence into account. Setting  $\sigma = 0$ , the level of confidence is so high that it is actually a constraint, and leads to the original Eq. 21. When  $\sigma$  is set large, the Lagrange multiplier  $\mu_i$  will become smaller, and the optimized distributor will hardly differ from the original distribution. Thus, this procedure changes the constraint into a restraint, tuned by the level of confidence in the data.

The MaxCal equation can also be adjusted to incorporate the errors in the data, leading to the analog of Eq. 102

$$k_{AB}^0 e^{\mu_A} = k_{AB}^{exp} + \mu_A \sigma_{k_{AB}}^2, \tag{103}$$

where  $\sigma_{k_{AB}}$  represents the level of confidence in the rate constant data. Thus, one can turn the constraint condition into a restraint condition.

### 5.4 CoPE-MaxCal and machine learning

The MaxCal method aims at incorporating experimental data in molecular simulations by a path reweighting function  $f_A(\lambda)$ . Since the computation of  $f_A(\lambda)$  from Eq. 86 might be not always trivial, one might use advanced regression procedures, such as those provided by machine learning methods. In addition, when considering the more general form of the function,  $f_A(p_{B,max}[\mathbf{x}])$ , machine learning methods might be invoked to optimize both the  $f_A$  function as well as the committor description in terms of CVs. This direction of research is left for future exploration.

### 5.5 Optimization of a collective variable with CoPE-MaxCal

By inspecting the CoPE-MaxCal approach, one realizes that the path reweighting depends on the choice of the collective variable  $\lambda$ . For a different choice of  $\lambda$ , one would get a different weight function, and hence caliber. So, it should be possible to optimize the caliber function over all possible CVs.

In principle, it is possible to vary the collective variable and maximize the entropy and caliber as function of this CV. The most optimal collective variable is then the one that leads to the least perturbed path distribution. To do so, we can substitute the optimized MaxCal distributions  $\mathcal{P}_A^{MC}[\mathbf{x}] = C_A^{-1} \mathcal{P}_A^0[\mathbf{x}] \exp[f_A(\lambda_{max}[\mathbf{x}])]$  and  $\mathcal{P}_B^{MC}[\mathbf{x}] = C_B^{-1} \mathcal{P}_B^0[\mathbf{x}] \exp[f_B(\lambda_{min}[\mathbf{x}])]$ , with  $C_A, C_B$  normalization constants for these distributions, into the path entropy expression. Here, we can make use

of the reaching histograms approach  $R_A^0(\lambda|\lambda_0)$ , and  $R_B^0(\lambda|\lambda_n)$ . It follows that

$$S[\mathcal{P}_A||\mathcal{P}_A^0] = -\frac{1}{C_A} \int d\lambda R_A^0(\lambda|\lambda_0) e^{f_A(\lambda)} (f_A(\lambda) - \ln C_A)$$

$$S[\mathcal{P}_B||\mathcal{P}_B^0] = -\frac{1}{C_B} \int d\lambda R_B^0(\lambda|\lambda_n) e^{f_B(\lambda)} (f_B(\lambda) - \ln C_B),$$

(104)

where the normalization  $C_A = \int d\lambda R_A^0(\lambda|\lambda_0) e^{f_A(\lambda)}$ , is now in terms of the reaching histograms. While in the above all sub-distributions  $\mathcal{P}_A^0, \mathcal{P}_A, \mathcal{P}_B, \mathcal{P}_B^0$  are supposed to be normalized, we require the normalized total path distributions  $\mathcal{P}$  and  $\mathcal{P}^0$ , when computing the total entropy In terms of  $S[\mathcal{P}_A||\mathcal{P}_A^0]$  and  $S[\mathcal{P}_B||\mathcal{P}_B^0]$ , defined above, the total path entropy becomes

$$S[\mathcal{P}||\mathcal{P}^0] = \alpha S[\mathcal{P}_A||\mathcal{P}_A^0] + (1 - \alpha) S[\mathcal{P}_B||\mathcal{P}_B^0]$$

$$+ \alpha \ln \frac{\alpha}{\alpha_0} + (1 - \alpha) \ln \frac{1 - \alpha}{1 - \alpha_0},$$

(105)

with  $\alpha = C_A/(C_A + C_B)$ , and  $\alpha_0 = C_A^0/(C_A^0 + C_B^0)$ .

It is now possible to compare combinations of different CVs with each other and select the best combination that optimizes the caliber, i.e disturbs the original distribution as little as possible. This gives thus an additional variational principle. Going back to the original caliber maximization over the distribution, we can add a second level of optimization

$$\mathcal{P}^{MC}[\mathbf{x}] = \operatorname{argmax}_{\mathcal{P}[\mathbf{x}], \lambda} S[\mathcal{P}||\mathcal{P}^0]$$

subject to:  $\begin{cases} \int D\mathbf{x} \mathcal{P}_B[\mathbf{x}] f_A^{(\lambda_i)}[\mathbf{x}] = k_{AB}^{exp} \\ \int D\mathbf{x} \mathcal{P}_B[\mathbf{x}] f_B^{(\lambda_i)}[\mathbf{x}] = k_{BA}^{exp} \\ \int D\mathbf{x} \mathcal{P}[\mathbf{x}] = 1. \end{cases}$

(106)

with  $f_A^{(\lambda_i)}[\mathbf{x}] = \theta((\lambda_{max}[\mathbf{x}] - \lambda_i)P(\lambda_B|\lambda_i))$ , and  $f_B^{(\lambda_i)}[\mathbf{x}] = \theta((\lambda_{min}[\mathbf{x}] - \lambda_i)P(\lambda_A|\lambda_i))$ , the instantaneous measures for the rates. The additional maximization over the CV space gives, thus, the best set of CVs. Such a procedure is very similar to reaction coordinate optimization [128, 137–141].

### 5.6 CoPE-MaxCal as a variational principle?

In the previous section, we discussed how varying the collective variables until the path entropy is maximized yields the best set of CVs, i.e., the reaction coordinate. It seems, therefore, natural to identify the committor  $p_B$  as the best collective variable, or reaction coordinate, since it predicts for each configurations the probability to end in B. But why should the MaxCal procedure give the same reaction coordinate? In other words, why would the path entropy be optimal when the committor is chosen as a CV?

This question can be addressed by showing that when deviating from the optimal CV, i.e., the committor function, one always increases the KL divergence, i.e., lowers the path entropy. This would be the case even for very small perturbations, i.e., in the limit  $\mu \rightarrow 0$ . Such a finding would amount to establishing a variational principle.

To justify such a principle, we can take a closer look at the path entropy and its behavior for a given path distribution. We consider first an illustrative example in Fig. 6.

For a path ensemble in 2D space of a  $3 \times 3$  grid, we consider two cases, a diagonal and a horizontal CV. From the symmetry of the problem, the left panel should represent the best CV, with the maximum path entropy  $S$  or the lowest KL divergence, the negative of  $S$ , which is always positive (or zero). To show that we reconsider the expressions for the negative entropy, and change from a integral expression to a discrete version for the sake of this illustration. The KL divergence is then written as

$$D_{KL} = -S = \frac{1}{C_A} \sum_i R_i e^{f_i} (f_i - \ln \frac{C_A}{C_A^0}),$$

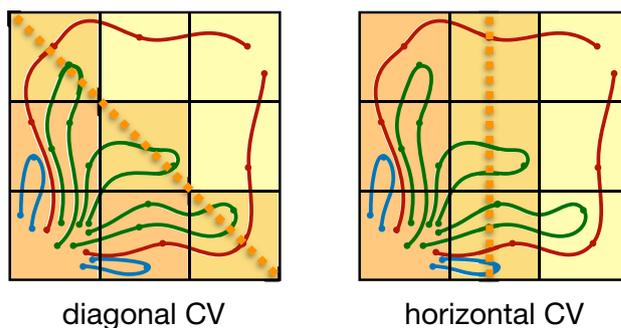
(107)

where we used  $R_i$  as the discrete version of the reaching histogram  $R(\lambda|\lambda_0)$ , and  $f_i$  as the discrete version of  $f(\lambda)$ . The normalization constant  $C_A^0 = \sum_i^n R_i$ ,  $C_A = \sum_i^n R_i e^{f_i}$ . Substitution of these expressions into the above equation gives

$$D_{KL} = \frac{\sum_{i=0} R_i e^{f_i} f_i}{\sum_{i=0} R_i e^{f_i}} - \ln \frac{\sum_{i=0} R_i e^{f_i}}{\sum_{i=0} R_i}.$$

(108)

We can simplify this expression by taking a factor  $f_0$  out of the first and a factor  $e^{f_0}$  out of the second fraction,



**Fig. 6** CoPE-MaxCal as a variational principle. We show a path ensemble in 2D space of a  $3 \times 3$  grid emanating from the bottom left cell, which is part of state A. In both panels, the same prior path ensemble is considered. The coloring of the paths shows the prior weight of the trajectories. Blue indicates a higher weight, while green and red lower weights. The CV foliation is shown as shades of yellow cells (from dark to light is bin 0, 1 and 2). Also shown is a indication of the location of the barrier in the projection on the CV. This is the best estimate of the location of the barrier in each CV

and canceling out several factors: where the  $f_0$  terms cancel

$$D_{KL} = \frac{\sum_{i=1} R_i e^{f_i} (f_i - f_0)}{\sum_{i=0} R_i e^{f_i}} - \ln \left( 1 + \frac{\sum_{i=1} R_i (e^{f_i - f_0} - 1)}{\sum_{i=0} R_i} \right) \quad (109)$$

Up to now, we made no assumptions. Next, we consider the small perturbation limit  $\mu \ll 1$ . This means that  $|f| \ll 1$ . The logarithm can then be replaced, yielding

$$D_{KL} = \frac{\sum_{i=1} R_i e^{f_i} (f_i - f_0)}{\sum_{i=0} R_i e^{f_i}} - \frac{\sum_{i=1} R_i (e^{f_i - f_0} - 1)}{\sum_{i=0} R_i} \quad (110)$$

Combining the two fractions gives

$$\begin{aligned} D_{KL} &= \frac{\sum_{j=0} \sum_{i=1} R_j R_i e^{f_i} (f_i - f_0) - R_j e^{f_j} R_i (e^{f_i - f_0} - 1)}{(\sum_{i=0} R_i e^{f_i})(\sum_{i=0} R_i)} \\ &= \frac{\sum_{j=0} \sum_{i=1} R_j R_i (e^{f_i} (f_i - f_0) - e^{f_j} (e^{f_i - f_0} - 1))}{(\sum_{i=0} R_i e^{f_i})(\sum_{i=0} R_i)} \end{aligned} \quad (111)$$

The second assumption that we can make is that the  $R_0$  is the largest contribution to the path ensemble. This represents the paths that make small excursions from the stable state. In fact,  $R_0$  can be made exponentially larger than the next terms in the summation, if the barrier is high enough. This means that we can approximate  $D_{KL}$  by only keeping the  $j = 0$  term.

$$D_{KL} \approx \frac{\sum_{i=1} R_0 R_i (e^{f_i} (f_i - f_0) - (e^{f_i} - e^{f_0}))}{R_0 R_0} \quad (112)$$

The next step is to realize that the bias  $f_0 \rightarrow 0$ , which means that we obtain

$$D_{KL} = \frac{1}{R_0} \sum_{i=1} R_i (e^{f_i} (f_i) - (e^{f_i} - 1)) \quad (113)$$

Finally, since all biases are small, due the fact that we choose  $\mu \ll 1$ , we can expand the exponential giving

$$D_{KL} = \frac{1}{R_0} \sum_{i=1} \frac{1}{2} R_i f_i^2 \quad (114)$$

The above reasoning shows that in the limit of small  $\mu$ , the path entropy, i.e., the KL divergence, scales with the square of the bias  $f$ -function (thus, being always positive as expected), and is linear in the value of the reaching histogram, the number of paths reaching a certain bin. This derivation shows that this KL divergence should be minimal for the optimal CV.

The  $f$ -function is changing with the choice of the CV, but not strongly, as we assume  $|f| \ll 1$ . This means

that the most important factor is the  $R_i$ . This then establishes a variational principle. Either  $f$  goes to zero (i.e., no perturbation), which is a trivial solution, or the CV is chosen such that the path reaching histogram  $R_i$  becomes minimal.

This path density is minimal for a CV that is aligned with the committor, for the following reason. Reactive paths have to go through a bottleneck where the path density  $R_i$  is low. When the CV is the committor, this  $R_i$  is the density of reaching paths that is measured. Any deviation from this optimal CV will include paths that are not part of the bottleneck, with a higher path density. This will then increase the KL divergence, or lower the path entropy. This is illustrated in the figure 6, where only 3 bins are considered:  $R_0, R_1$  and  $R_2$  (indicated by the yellow shades in the cells). The diagonal CV has only 3 paths in  $R_1$  (the green paths), but the horizontal CV also picks up the blue paths in the lower middle cells, which have a much higher weight (as they do not reach the barrier). This increase in  $R_i$  is much more influential than the marginal change in  $f_i$  due to the change in variable. In fact, for this example the function  $f$  only changes by 1% between the CVs.

While this is just an example, it is possible that this variational principle holds more generally. In fact, the concept of lowest path density  $R_i$  for the best CV is reminiscent of the variational transition state theory [142], which states that the best CV is the one that has the highest barrier, or lowest rate. In our treatment, the rate constant does not change since the reactive pathways do not change. Still, the configurational density for both the A and B ensembles projected on the CV, directly gives the free energy as we have seen above. Thus, optimizing the CV using MaxCal is consistent with optimizing the free energy barrier in the sense of the variational transition state theory.

## 6 Conclusions and outlook

### 6.1 Summary

In this review, we have described the MaxEnt and MaxCal approaches and their extensions to continuum path ensembles through the CoPE-MaxCal method. This strategy enables the definition of the least perturbed path ensemble consistent with given external information, i.e., the trajectory ensemble of maximum relative path entropy under given constraints. We have focused in particular on imposing kinetic rate constants on the continuous path ensembles obtained by molecular dynamics simulations of rare events. We have shown how this can be done by a reweighting function depending on the maximum value of a collective variable along the trajectory. To obtain this function, we made use of the MaxEnt approach to impose the equilibrium constant.

The CoPE-MaxCal optimized trajectory ensemble can yield new information for other observables. One example is a shift in transition state, which we dis-

cussed in the previous section. Other kinetic properties might include residual dipolar couplings in NMR spectroscopy.

Besides reviewing the CoPE-MaxCal method, we have discussed several further additional directions of investigations of this method, including an extension to multiple states, and the relationship of the MaxCal collective variable optimization with reaction coordinate optimization methods.

To provide an outlook, we present here some open questions and future directions.

## 6.2 Incorporation of time-resolved experimental measurements

An potential extension and application of the CoPE-MaxCal framework involves the addition of constraints on the evolution of time-dependent variables. In this way, one goes beyond adding a constraint on the final reactive flux into a certain state (the rate constant) for trajectories starting in a different initial state. This extension can be obtained for example by starting from the correlation function in Eq. 46. This would result in a reweighting of the existing trajectories based on constraints added at each time, as derived in Eq. 49. Such time-dependent experimental data could be provided for example by time-dependent small angle X-ray scattering (SAXS) [33, 41, 42, 143] and time-dependent solution cryo-electron microscopy [144].

## 6.3 Rational design of mutations and lead compounds in drug discovery

We envision that a combination of our CoPE-MaxCal framework (providing more accurate path ensembles) with perturbative interaction models can result in kinetics-mutation landscapes and/or kinetics-lead-compound-optimization landscapes. This will, thus, constitute a relatively inexpensive way to non-random design of biomolecular mutation and lead compound derivatives that modify the barrier in the desired direction, as opposed to random design of mutations that have to be tested experimentally, or in expensive additional simulations.

## 6.4 Biased molecular dynamics simulations

In its initial implementation, the CoPE-MaxCal method utilizes unbiased molecular dynamics simulations, possibly with the help of enhanced path sampling such as TPS, thereby preserving the correct kinetics up to the accuracy of the force field. Yet, advances in computational structural biology methods enable incorporating experimental data in the form of restraints or constraints [28–32, 34–36, 38, 39, 43] by biasing the Hamiltonian, while providing free energies that maximally meet the experimental restraints or constraints. These observations raise the question whether the CoPE-MaxCal approach can be combined with such biasing methods, even if they cannot reproduce the natural kinetics.

Obtaining a free energy landscape using enhanced sampling, and then performing unbiased molecular dynamics, might be a viable way to reconstruct unbiased kinetics. For instance, trajectories from infrequent metadynamics or frequency adaptive metadynamics [145–148] can subsequently be reweighted to meet given kinetic constraints.

Finally, the Filizola and Keller groups recently explored the use of biased molecular dynamics simulations and subsequent application of constraints on configurational ensemble averages that result in the alteration of an transition-probability matrix to reconstruct the unbiased kinetics [149, 150].

## 6.5 Force field optimization for molecular kinetics

Molecular dynamics is in principle capable to provide quantitative predictions about the thermodynamics and kinetics of a system at the atomistic level. However, the accuracy of the calculations is limited by the quality of the force fields, which are typically parameterized using quantum mechanical calculations and experimental measurements of equilibrium properties [14]. There are several methods using MaxEnt, Bayesian inference and force-matching techniques that pursue the development of such force fields [28, 151–155]. However, there is still a need of further methodological advances in this area for the accurate reproduction of time-dependent properties, such as the rate constants. These advances would have a large impact in the growing field of integrative structural biology, since for example they would potentially better predict populations of conformations that lie in the transition states [46]. Such a force field optimization for molecular kinetics is currently not possible within CoPE-MaxCal approach, as the prior dynamical trajectories, and hence the force field, are not affected in the reweighting procedure. Yet, it might be made possible by introducing novel forward models for the rate as a function of the force field parameters. By computing the derivative of the rate constant with respect to these parameters, one could minimize an error function to meet target kinetics. This is an avenue that are we currently actively pursuing.

## 6.6 Open questions in statistical mechanics

The extension of the MaxEnt principle to path ensembles that we have discussed in this review prompts a series of questions of general nature, some of which are listed in the following.

- (1) We have presented the CoPE-MaxCal method as a reweighting scheme. This method affects only the initial conditions (see Sect. 4.6), but not the trajectories themselves. It would be interesting to investigate the analogy between this type of reweighting and the coupling to a heat bath, such as done in the derivation of the canonical ensemble from the microcanonical ensemble. One can also ask whether

the MaxCal approach can be used to modify the trajectories themselves. This might be used in an on-the-fly optimization such as proposed in [156].

- (2) In the CoPE-MaxCal application to rate constants we have used the Transition Path Sampling and Transition Interface Sampling framework. While these methods are efficient tools to obtain and manipulate path ensembles, one should be able to formulate the methodology independently.
- (3) The investigation of the relationship between the CoPE-MaxCal approach and other path reweighting methodologies, such as the s-ensemble or other path reweighting methods, could provide further insight for establishing more general methods.
- (4) The CoPE-MaxCal framework may be extended to non-equilibrium steady states, e.g., for the driven dynamics observed in active systems. For that situation the MaxEnt approach cannot describe the correct steady state density, and one has to use the MaxCal approach in order to construct the reweighting function.

### 6.7 Final thoughts

We anticipate that the application of the MaxCal approach to ensembles of continuum trajectories, as for example provided by molecular dynamics simulations, will offer new opportunities to address equilibrium and non-equilibrium problems that involve the trajectory space. We undoubtedly will see more research in this area in the coming years.

Z.F.B. would like to acknowledge the Federation of European Biochemical Societies (FEBS) for financial support (LTF).

### Author contributions

ZFB and PGB designed research; ZFB and PGB performed research; and ZFB, MV, and PGB wrote the paper. All authors have read and approved the final manuscript.

**Data Availability Statement** This manuscript has no associated data or the data will not be deposited [Authors' comment: This review has no associated data as all data discussed has been published elsewhere.]

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statu-

tory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### References

1. E.T. Jaynes, Phys. Rev. **108**, 171 (1957)
2. E. T. Jaynes, Phys. Rev. **106**, 620 (1957)
3. E. T. Jaynes (1980) Ann. Rev. Phys. Chem. 31 579
4. E.T. Jaynes, H. Haken, *Complex Systems- Operational Approaches* (Springer-Verlag, Berlin, 1985), pp. 254–259
5. S. Pressé, K.Ghosh, J. Lee, K. A. Dill. Rev. Mod. Phys. 85 1115 (2013)
6. A. A. Filyukov, V. Y. Karpov (1967) J. Eng. Phys. 13 416
7. C. Monthus, J. Stat. Mech. 2011 P03008 (2011)
8. S. Davis, D. Gonzalez, J. Phys. A: Math. Theor. **48**, 425003 (2015)
9. P.D. Dixit, J. Wagoner, C. Weistuch, S. Pressé, K. Ghosh, K.A. Dill, J. Chem. Phys. **148**, 010901 (2018)
10. D.M. Rogers, J. Stat. Mech.: Theory Exp. **2019**, 084010 (2019)
11. K. Ghosh, P.D. Dixit, L. Agozzino, K.A. Dill, Annu. Rev. Phys. Chem. **71**, 213 (2020)
12. D. Chandler, *Introduction to modern statistical mechanics*. Oxford University Press (1987)
13. D. Chandler. Classical and quantum dynamics in condensed phase simulations. (World Scientific, Singapore, 1998), Chap. Barrier crossings: classical theory of rare but important events, pgs. 3–23
14. D. Frenkel, B. Smit, *Understanding Molecular Simulation*, 2nd ed (Academic Press, Inc., Orlando, FL, USA, 2001)
15. D.E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R.O. Dror, M.P. Eastwood, J.A. Bank, J.M. Jumper, J.K. Salmon, Y. Shan, W. Wriggers Sci. **330**, 341 (2010)
16. M. Bonomi, M. Vendruscolo, Current Opinion in Structural Biology. **56**, 37 (2019)
17. C. Dellago, P. G. Bolhuis. Adv. Polym. Sci. 221 167 (2009)
18. M. Bonomi, G. Bussi, C. Camilloni, G. Tribello, P. Banáš, A. Barducci, M. Bernetti et al., Nat. Methods **16**, 670 (2019)
19. J. W. Pitera, J. D. Chodera, J. Chem. Theory Comput. 8 3445 (2012)
20. A. Cavalli, C. Camilloni, M. Vendruscolo, J. Chem. Phys. **138**, 094112 (2013)
21. B. Roux, J. Weare, J. Chem. Phys. **138**, 084107 (2013)
22. W. Boomsma, J. Ferkinghoff-Borg, K. Lindorff-Larsen, PLoS Computational Biology **10**, 1 (2014)
23. G. Hummer, J. Köfinger, J. Chem. Phys. **143**, 243150 (2015)
24. M. Bonomi, C. Camilloni, A. Cavalli, M. Vendruscolo. Sci. Adv. 2 1 (2015)
25. F. Marinelli, J. D. Faraldo-Gómez. Biophys. J. 108 2779 (2015)
26. S. Olsson, H. Wu, F. Paul, C. Clementi, F. Noé, Proceedings of the National Academy of Sciences. **114**, 8265 (2017)

27. M. Bonomi, R. Pellarin, M. Vendruscolo, *Biophys. J.* **114**, 0605018 (2018)
28. A. Cesari, S. Reißer, G. Bussi, *Computation* **6**, 15 (2018)
29. S. Bottaro, K. Lindorff-Larsen, *Science* **361**, 355 (2018)
30. S. Bottaro, T. Bengtsen, K. Lindorff-Larsen Integrating Molecular Simulation and Experimental Data: A Bayesian/Maximum Entropy Reweighting Approach, in *Structural Bioinformatics: Methods and Protocols* edited by Z.Gáspári (Springer US New York, NY, 2020) pp. 219–240
31. M. Bonomi, C. Camilloni, A. Cavalli, M. Vendruscolo, *Sci. Adv.* **21** (2016)
32. S. Vahidi, Z.A. Ripstein, M. Bonomi, T. Yuwen, M.F. Mabanglo, J.B. Juravsky, K. Rizzolo, A. Velyvis, W.A. Houry, M. Vendruscolo, J.L. Rubinstein, L.E. Kay, *Proc. Nat. Acad. Sci. USA* **115**, E6447 (2018)
33. M. R. Hermann, J. S. Hub. *J. Chem. Theory Comput.* **15** 5103 (2019)
34. L. Eshun-Wilson, R.Zhang, D.Portran, M. V. Nachury, D. B. Toso, T.Löhr, M.Vendruscolo, M.Bonomi, J. S. Fraser, E.Nogales, *Proc. Nat. Acad. Sci. USA.* **116** 10366 (2019)
35. Z. F. Brotzakis, T. Lohr, M. Vendruscolo, *Chem. Sci.* **12**, 9168–9175 (2021)
36. Z.F. Brotzakis, P.R. Lindstedt, R. Taylor, G.J.L. Bernardes, M. Vendruscolo, *bioRxiv* <https://doi.org/10.1101/2020.11.10.376285> (2020)
37. P. Robustelli, S.Piana, D. E. Shaw, *Proc. Natl. Acad. Sci. U. S. A.* **115** E4758 (2018)
38. G.T. Heller, F.A. Aprile, M. Bonomi, C. Camilloni, A. De Simone, M. Vendruscolo, *Journal of Molecular Biology.* **429**, 2772 (2017)
39. A. N. Borkar, M. F. Bardaro, C.Camilloni, F. A. Aprile, G.Varani, M.Vendruscolo, *Proc. Nat. Acad. Sci. USA.* **113** 7171 (2016)
40. H. Lou, R.I. Cukier, *J. Chem. Phys.* **149**, 234106 (2018)
41. T. Bengtsen, V.L. Holm, L.R. Kjølbye, S.R. Midtgaard, N.T. Johansen, S. Bottaro, B. Schjøtt, L. Arleth, K. Lindorff-Larsen, *Elife.* **9**, e56518 (2020)
42. C. Paissoni, A. Jussupow, C. Camilloni, *J. Chem. Theory Comput.* **16** 2825 (2020)
43. G. T. Heller, F. A. Aprile, T. C. Michaels, R.Limbocker, M.Perni, F. S. Ruggeri, B.Mannini, T.Löhr, M.Bonomi, C.Camilloni, A.de Simone, I. C. Felli, R.Pierattelli, T. P. Knowles, C. M. Dobson, M.Vendruscolo, *Science Adv.* **6**, 1 (2020)
44. S. Orioli, A. H. Larsen, S.Bottaro, K. Lindorff-Larsen. *Prog Mol Biol Transl Sci.* **170**, 123 (2020)
45. R. Capelli, G. Tiana, C. Camilloni, *J. Chem. Phys.* **148**, 184114 (2018)
46. Z.F. Brotzakis, M. Vendruscolo, P.G. Bolhuis, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2012423118 (2021)
47. J.-H. Prinz, H. Wu, M.Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C.Schütte, F. Noé,*J. Chem. Phys.* **134**, 174105 (2011)
48. M. Bause, T. Wittenstein, K. Kremer, T. Bereau, *Phys. Rev. E.***100**, 60103 (2019)
49. M. Bause T. Bereau, *J. Chem. Phys.* **154**, 134105 (2021)
50. H. Touchette, *Physics Reports.* **478** 1 (2009)
51. L. O. Hedges, R. L. Jack, J. P. Garrahan, D.Chandler, *Science.* **323** 1309 (2009)
52. A.B. Adib, *J. Phys. Chem. B.* **112**, 5910 (2008)
53. S. Kieninger B. G. Keller, *J. Chem. Phys.* **154**, 094102 (2021)
54. C. Dellago, P. G. Bolhuis, F. S. Csajka, D. Chandler, *J. Chem. Phys.* **108**, 1964 (1998)
55. P. G. Bolhuis, D .Chandler, C. Dellago, P. Geissler. *Ann. Rev. Phys. Chem.* **53**, 291 (2002)
56. C. Dellago, P.G. Bolhuis, P. L. Geissler. *Adv. Chem. Phys.* **123** 1 (2002)
57. T. S. van Erp, D.Moroni, P. G. Bolhuis. *J. Chem. Phys.* **118** 7762 (2003)
58. T. S. van Erp. *Adv. Chem. Phys.* **151** 27 (2012)
59. R. Cabriolu, K.M.S. Refsnes, P.G. Bolhuis, T.S. van Erp, *J. Chem. Phys.* **147**, 152722 (2017)
60. P. G. Bolhuis, *Proc. Nat. Acad. Sci. USA.* **100** 12129 (2003)
61. J. Juraszek, P.G.. Bolhuis, *Proc. Nat. Acad. Sci. USA* **103**, 15859 (2006)
62. J. Juraszek, P. G. Bolhuis, *Biophys. J.* **95** 4246 (2008)
63. W. Du, P.G. Bolhuis, *J. Chem. Phys.* **140**, 195102 (2014)
64. W. Du, P.G. Bolhuis. *Biophys. J.* **108** 368 (2015)
65. J. Vreede, J. Juraszek, P. G. Bolhuis, *Proc. Natl. Acad. Sci. U.S.A.* **107** 2397 (2010)
66. R. B. Best, G.Hummer, *Proc. Natl. Acad. Sci. U.S.A.* **113** 3263 (2016)
67. Z.F. Brotzakis, P.G. Bolhuis, *J. Phys. Chem. B.* **123**, 1883 (2019)
68. J. Domański, M.S. Sansom, P.J. Stansfeld, R.B. Best, *PLOS Comput. Bio.* **16**, e1007919 (2020)
69. M. Hagan, A.Dinner, D.Chandler, A.Chakraborty, *Proc. Natl. Acad. Sci. U.S.A.* **100** 13922 (2003)
70. R. Radhakrishnan, T.Schlick, *Proc. Natl. Acad. Sci. U.S.A.* **101** 5970 (2004)
71. R.J. Dimelow, N. A. Burton, I.H. Hillier, *Phys. Chem. Chem. Phys.* **9** 1318 (2007)
72. Y. Li, B.D. Freudenthal, W. A. Beard, S. H. Wilson, T.Schlick, *J. Am. Chem. Soc.* **136** 3630 (2014)
73. J. Vreede, A. P. de Alba Ortíz, P. G. Bolhuis, D.W.H. Swenson, *Nucleic Acids Res.* **47**, 11069 (2019)
74. E. Riccardi, E.C. van Mastbergen, W.W. Navarre, J. Vreede, *PLOS Comput. Bio.* **15**, e1006845 (2019)
75. K.-i. Okazaki, D.Wöhlert, J.Warnau, H.Jung, Özkan Yildiz, W.Kühlbrandt, G.Hummer, *Nat. Commun.* **10**, 4322 (2019)
76. P. L. Geissler, C.Dellago, D.Chandler, *J. Phys. Chem. B.* **103** 3706 (1999)
77. A. J. Ballard and C.Dellago, *J. Phys. Chem. B.* **116** 13490 (2012)
78. R. G. Mullen, J.-E. Shea, B.Peters, *J. Chem. Theory. Comput.* **10** 659 (2014)
79. M. N. Joswiak, M. F. Doherty, B.Peters, *Proc. Natl. Acad. Sci. U.S.A.* **115** 656 (2018)
80. D. Laria, J. Rodriguez, C.Dellago, D.Chandler, *J. Phys. Chem. A.* **105** 2646 (2001)
81. P. Varilyly and D.Chandler, *J. Phys. Chem. B.* **117** 1419 (2013)
82. P.R. ten Wolde, D. Chandler, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6539 (2002)
83. C. Leitold, C. Dellago, *J. Chem. Phys.* **141**, 134901 (2014)
84. D. Moroni, P.R. Ten Wolde, P.G. Bolhuis, *Phys. Rev. Lett.* **94** 1 (2005)

85. W. Lechner, C. Dellago, P.G. Bolhuis, *Phys. Rev. Lett.* **106**, 085701 (2011)
86. G. T. Beckham, B.Peters, *J. Phys. Chem. Lett.* **2** 1133 (2011)
87. B.C. Barnes, B.C. Knott, G.T. Beckham, D.T. Wu, A.K. Sum, *J. Phys. Chem. B.* **118** 13236 (2014)
88. G.D. Leines, R. Drautz, J. Rogal, *J. Chem. Phys.* **146**, 154702 (2017)
89. G.D. Leines, J. Rogal, *J. Phys. Chem. B.* **122** 10934 (2018)
90. Arjun, T.A. Berendsen, P. G. Bolhuis, *Proc. Natl. Acad. Sci. U.S.A.* **116** 19305 (2019)
91. Y. Liang, G.D. Leines, R. Drautz, J. Rogal, *J. Chem. Phys.* **152**, 224504 (2020)
92. A. Arjun, P. G. Bolhuis, *J. Phys. Chem. B.* **124** 8099 (2020)
93. D. Zahn, Y. Grin, S. Leoni, *Phys. Rev. B.* **72** 064110 (2005)
94. G.T. Beckham, B.Peters, C.Starback, N.Variankaval, B. L. Trout, *J. Am. Chem. Soc.* **129** 4714 (2007)
95. G.T. Beckham, B.Peters, B.L. Trout, *J. Phys. Chem. B.* **112** 7460 (2008)
96. P.L. Geissler, C. Dellago, D. Chandler, J. Hutter, M. Parrinello, *Science* **291**, 2121 (2001)
97. A. Tiwari, B.Ensing, *Faraday Discuss.* **195** 291 (2016)
98. J.E. Basner, S. D. Schwartz, *J. Am. Chem. Soc.* **127** 13822 (2005)
99. B. C. Knott, M. H. Momeni, M. F. Crowley, L. F. Mackenzie, A.W. Götz, M.Sandgren, S. G. Withers, J. Ståhlberg, G. T. Beckham, *J. Am. Chem. Soc.* **136** 321 (2013)
100. M. Dzierlenga, M. Varga, S. Schwartz in *Methods in Enzymology* (Elsevier 2016) pp. 21–43
101. T.K. Paul, S. Taraphder, *ChemPhysChem* **21**, 1455 (2020)
102. A. C. Newton, J. Groenewold, W. K. Kegel, P. G. Bolhuis, *Proc. Nat. Acad. Sci. USA.* **112** 15308 (2015)
103. A.C. Newton, J. Groenewold, W.K. Kegel, P.G. Bolhuis, *J. Chem. Phys.* **146**, 234901 (2017)
104. A.C. Newton, R. Kools, D.W. Swenson, P.G. Bolhuis, *J. Chem. Phys.* **147**, 155101 (2017)
105. A. S. Keys, L. O. Hedges, J. P. Garrahan, S. C. Glotzer, D.Chandler, *Phys. Rev. X*, **1** 021013 (2011)
106. R. L. Jack, L. O. Hedges, J. P. Garrahan, D.Chandler, *Phys. Rev. Lett.* **107**, 275702 (2011)
107. A.S. Mey, P.L. Geissler, J.P. Garrahan, *Phys. Rev. E.* **89**, 032109 (2014)
108. D. Coslovich, R.L. Jack, *J. Stat. Mech: Theory Exp.* **2016**, 074012 (2016)
109. F.Turci, C. P. Royall, T.Speck, *Phys. Rev. X.* **7**, 031028 (2017)
110. M. Campo, T. Speck, *J. Chem. Phys.* **152**, 014501 (2020)
111. R.J. Allen, P.B. Warren, P.R. ten Wolde, *Phys. Rev. Lett.* **94**, 018104 (2005)
112. R.Allen, D.Frenkel, P.R.ten Wolde, *J. Chem. Phys.* **124**, 024102 (2006)
113. G. Huber, S.Kim, *Biophys. J.* **70** 97 (1996)
114. B.W. Zhang, D. Jasnow, D.M. Zuckerman, *J. Chem. Phys.* **132**, 054107 (2010)
115. L. T. Chong, A. S. Saglam, D. M. Zuckerman, *Curr. Opin. Struct. Biol.* **43** 88 (2017)
116. J. T. Berryman, T. Schilling, *J. Chem. Phys.* **133**, 244101 (2010)
117. A. K. Faradjian, R. Elber, *J. Chem. Phys.* **120**, 10880 (2004)
118. R. Elber, *Annu. Rev. Biophys.* **49** 69 (2020)
119. F.Cerou, A.Guyader, T.Lelievre, D.Pommier, *J. Chem. Phys.* **134**, 054108 (2011)
120. A. Dickson, A. Warmflash, A.R. Dinner, *J. Chem. Phys.* **131**, 154104 (2009)
121. P. G. Bolhuis, C.Dellago, D.Chandler, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5877 (2000)
122. W.E, W.Ren, E.Vanden-Eijnden, *J. Phys. Chem. B.* **109**, 6688 (2005)
123. W.E., E.Vanden-Eijnden, *J. Stat. Phys.* **123**, 503 (2006)
124. L. Onsager, *Physical Review.* **54** 554 (1938)
125. R. Du, V. S. Pande, A. Y. Grosberg, T.Tanaka, E. S. Shakhnovich, *J. Chem. Phys.* **108** 334 (1998)
126. W.E, E.Vanden-Eijnden, *Annu. Rev. Phys. Chem.* **61**, 391 (2010)
127. P.G. Bolhuis, W.Lechner, *J. Stat. Phys.* **145** 841 (2011)
128. P. Tiwary, B. J. Berne, *Proc. Nat. Acad. Sci. USA.* **113** 2839 (2016)
129. Y. Mori, K.I. Okazaki, T. Mori, K. Kim, N. Matubayasi, *J. Chem. Phys.* **153**, 054115 (2020)
130. P.G. Bolhuis, C.Dellago, *Eur. Phys. J. Spec. Top.* **224** 2409 (2015)
131. P.G. Bolhuis, D.W.H. Swenson, *Adv Theory Simul.* **4**, 2000237 (2021)
132. J. Rogal, W. Lechner, J. Juraszek, B. Ensing, P.G. Bolhuis, *J. Chem. Phys.* **133**, 174109 (2010)
133. K. Lindorff-Larsen, S. Piana, R.O. Dror, D.E. Shaw, *Science* **334**, 517 (2011)
134. Z.F. Brotzakis, P.G. Bolhuis, *J. Chem. Phys.* **151**, 174111 (2019)
135. P.G. Bolhuis, G. Csányi, *Phy. Rev. Lett.* **120**, 1 (2018)
136. P. Király, D.J. Kiss, G. Tóth, *J. Chem. Phys.* **148**, 134107 (2018)
137. A. Ma, A. R. Dinner, *J. Phys. Chem. B.* **109** 6769 (2005)
138. B. Peters, B.L. Trout, *J. Chem. Phys.* **125**, 054108 (2006)
139. B. Peters, *Annu. Rev. Phys. Chem.* **67** 669 (2016)
140. F. Noé, F. Nüske, *Multiscale Model. and Simul.* **11**, 635 (2013)
141. J. Marcelo, L.Ribeiro, P. B. Collado, Y.Wang, P.Tiwary, *J. Chem. Phys.* **149** 1 (2018)
142. D. G. Truhlar, B. C. Garrett, *Ann. Rev. Phys. Chem.* **35** 159 (1984)
143. L. Pollack, *Biopolymers.* **95** 543 (2011)
144. G.Marchello, C.De Pace, S.Acosta-Gutierrez, C.Lopez-Vazquez, N.Wilkinson, F. L. Gervasio, L.Ruiz-Perez, G.Battaglia, C. I. Building, *bioRxiv*, 2021.01.21.427613 (2021)
145. P. Tiwary, M. Parrinello, *Phy. Rev. Lett.* **111**, 230602 (2013)
146. P. Tiwary, V. Limongelli, M. Salvalaglio, M. Parrinello, *Proc. Nat. Acad. Sci. USA* **112**, E386 (2015)

147. Y.Wang, O.Valsson, P.Tiwary, M.Parrinello, K.Lindorff-Larsen, *J. Chem. Phys.* **149**, 072309 (2018)
148. Z. F. Brotzakis, V.Limongelli, M.Parrinello, *J. Chem. Theory Comput.* **15**, 743 (2019)
149. D.Meral, D.Provasi, M.Filizola. *J. Chem. Phys.* **149**, 224101 (2018)
150. L. Donati, B.G. Keller, *J. Chem. Phys.* **149**, 072335 (2018)
151. A. Cesari, A.Gil-Ley, G.Bussi, *J. Chem. Theory Comput.* **12**, 6192 (2016)
152. R.Dutta, Z. F. Brotzakis, A.Mira, *J. Chem. Phys.* **149**, 154110 (2018)
153. J. Chen, J. Chen, G. Pinamonti, C. Clementi, *J. Chem. Theory Comput.* **14**, 3849 (2018)
154. J.Wang, S.Olsson, C.Wehmeyer, A.Pérez, N. E. Charon, G.De Fabritiis, F.Noé, C.Clementi, *ACS Central Science* **5**, 755 (2019)
155. S. Doerr, M.Majewski, A.Pérez, A.Krämer, C.Clementi, F.Noé, T.Giorgino, G.De Fabritiis, *J. Chem. Theory Comput.* **17**, 2355 (2021)
156. D.C. Rose, J.F. Mair, J.P. Garrahan, *New J. Phys.* **23**, 013013 (2021)