Defining, calculating and converging observables of kinetic transition networks

T D Swinburne, David J Wales

Defining, Calculating and Converging Observables of a Kinetic Transition Network

Thomas D Swinburne
Aix-Marseille Université, CNRS, CINaM UMR 7325, Campus de Luminy, 13288 Marseille, France

David J Wales
Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

(Dated: March 2, 2020)

Kinetic transition networks of local minima and transition states are able to capture the dynamics of numerous systems in chemistry, biology and materials science. However, extracting observables is numerically challenging for large networks and will in general be sensitive to additional computational discovery. To have any measure of convergence for observables, these sensitivities must be regularly calculated. We present a matrix formulation of the discrete path sampling framework for kinetic transition networks, deriving expressions for branching probabilities, transition rates, and waiting times. Using the concept of the quasistationary distribution a clear hierarchy of expressions for network observables is established, from exact results to steady state approximations. We use these results in combination with the graph transformation method to derive the sensitivity with respect to perturbations of the known kinetic transition network, giving explicit terms for the pairwise sensitivity, and discussing the pathwise sensitivity. These results provide guidelines for converging the network with respect to additional sampling, focusing on the estimates obtained for the overall rate coefficients between product and reactant states. We demonstrate this procedure for two atomic clusters representative of single and double-funnel landscapes.

I. INTRODUCTION

The dynamics of a condensed phase system can often be represented as transitions between local minima in the potential energy landscape. When local minima are sufficiently deep that the system thermalizes before escape, interminima transitions are Markovian,[?] permitting a linear master equation representation of the dynamics. Transition rates $k = \kappa\omega_0 \exp(-\beta\Delta F)$ are decomposed into a dynamical prefactor $\kappa\omega_0$ and a free energy barrier $\Delta F = F^\dagger_{ji} - F_i$ between a given local minimum i and the local free energy maximum on the $j \leftarrow i$ minimum free energy path, known as the transition state.[? ? ?] The simplest transition state theory assumes that trajectories that pass the transition state do not recross and free energies can be calculated in the harmonic approximation. The set of all minima and transition states, characterized by their free energies, form a kinetic transition network (KTN).[? ? ? ?]

As the escape time from a minimum i scales as $\exp(\beta(\min_j F^\dagger_{ji} - F_i))$, construction and analysis of a KTN can, in principle, be much more efficient than direct molecular dynamics simulation. In practice, the number of local minima is exponential in system size,[? ?] meaning that it may be computationally challenging to sample the thermodynamically important minima and numerically challenging to reliably extract observables from the landscape[?] .

The variety of strategies to build KTNs can be considered a subclass of rare event simulation techniques that focus on identifying transition states. Starting with some set of minima produced from experimental insight or global optimisation[? ? ?] we can search for transition pathways from each minimum, using single-ended methods, or seek connections between two minima, using double-ended methods.

For the purposes of KTN construction, double-ended searches are typically static methods that find pathways by minimising a continuous or discrete chain of configurations connecting two specified states. Popular approaches include the string,[?] nudged elastic band (`NEB`),[? ?] and doubly-nudged elastic band (`DNEB`)[? ?] methods. The `DNEB` approach has been extended to treat distant initial and final states where many intermediate minima may be involved, resulting in pathways including multiple transition states. Here we employ a missing connection algorithm[?] along with hybrid eigenvector-following to refine the transition states accurately.[? ? ? ?] There are also dynamical double-ended methods, such as milestoning,[?] transition path sampling,[? ?] and forward flux sampling,[?] amongst others.[? ? ?]

Single-ended search methods can be roughly classified into unbiased dynamical approaches, such as temperature accelerated dynamics[? ? ?] and parallel replica dynamics with state recognition,[? ?] potential biasing dynamical approaches, such as hyperdynamics[?] and metadynamics,[?] and geometry optimisation methods such as hybrid eigenvector-following (the dimer method)[? ? ? ?] and the closely related activation-relaxation technique,[? ? ?] which use curvature information to find transition states directly. In dynamical single-ended methods transition rates can be estimated directly if the process of interest can actually be simulated with meaningful statistics.[? ? ? ?] Some further discussion of these methods can be found in a recent review.[?]

For any chosen construction method, the resulting KTN will always be finite, meaning that predictions will change when additional computational effort is expended to grow the network. This problem is well

recognized;[? ? ? ?] previous work has focused on single-ended searches using high temperature molecular dynamics, as the Poisson distribution of rare event escape times allows the benefit of additional computational effort to be quantified. Recently, one of us has used the Poisson law to derive a monotonically increasing Bayesian estimator of statewise sampling completeness.[?] This approach was combined with a KTN to determine where additional sampling should be performed to maximise the predictive time horizon as efficiently as possible, yielding a autonomous global sampling strategy.

Single-ended search strategies are particularly useful for exploring complex but locally single funnel free energy landscapes,[?] such as those in defect transport,[?] where the modeling goal is to build a KTN to study the long time evolution.

However, many problems in materials science, protein folding, and numerous other fields have competing global structures, giving rise to multi-funnel potential and free energy landscapes.[? ?] In this setting, the modeling goal is to build a KTN to capture the kinetics, via multiple complex pathways, between these potentially distant funnels. Whilst single ended search methods may eventually find these connections, double-ended strategies such as `DNEB` are in practice much more efficient, as putative pathways between basins can be rapidly found, then additional sampling performed to refine and explore around the key paths. The discrete path sampling (DPS) framework[? ? ?] can exploit both single- and double-ended methods to expand a KTN. Observables, such as average traversal times or effective transition rates, are then expressed as averages over all possible pathways. To simplify the resulting expressions local equilibrium is usually assumed for the reactants;[? ? ?] in the present contribution we will compare this approach with a direct treatment of the large, sparse matrices that encode the underlying master equation.[? ?] As we detail below, these matrices regularly suffer from severe conditioning issues due to the exponential sensitivity of transition rates, motivating the development of the graph transformation technique[? ?] to iteratively remove states, producing lower rank and better conditioned KTN matrices with identical observables. Typically, the renormalisation is continued until the desired observables can be simply extracted. Whilst numerically stable, this complete renormalisation is undesirable for sensitivity analysis, as we wish to specify which (pairs of) minima should be subject to further sampling in order to converge the observable of interest.

In this contribution, we unify the various approaches to calculating KTN observables and look in detail at the their sensitivity to additional sampling effort, in the form of double-ended transition pathway searches. In section **??** we define the terminology used to describe a KTN and introduce the master equation, before introducing the graph transformation method in §**??**. We then derive exact results for KTN observables as a function of transition matrices before and after an arbitrarily complete graph transformation in section **??**. Invoking the concept of the quasistationary distribution,[? ? ?] we find a clear hierarchy of approximations and provide a precise equivalence to the various results of DPS. Much of the working is assigned to appendices for clarity of exposition.

In §**??** we show how the graph transformation method is used to produce a partially renormalised, well conditioned KTN, which retains the ensemble of found reaction pathways, and we identify the total branching probability (or the sum of all committor probabilities) as the observable we use for the sensitivity analysis. This analysis is presented in section **??**, with explicit expressions for changes in the total branching probability of the renormalised KTN following additional connections found during a double-ended saddle search between an arbitrary state pair. We show that the matrix of all possible pairwise sensitivities can be rapidly calculated following two linear solves involving the renormalised KTN. The optimal deployment of sampling tasks using these pairwise sampling tasks is discussed briefly in section **??**, and approximate confidence bounds on the resulting KTN observables are analyzed in section **??**. Further development of confidence bounds will be the subject of future work. Finally, in section **??** we apply our framework to simulate the sampling the KTN of Lennard-Jones clusters that exhibit single and double-funnel landscapes, showing how our observable converges.

II. KINETIC TRANSITION NETWORKS

A kinetic transition network (KTN) is constructed from a set of metastable states and the transition states or rates that connect them. The states are considered sufficiently metastable that local equilibrium is achieved before escape and thus the state-to-state dynamics are Markovian,[?] providing the master equation representation that is used as the starting point for DPS.[? ? ?] In this section, we define our notation and introduce the master equation in terms of large, typically sparse matrix products.

Consider a tripartition of the set of all minima $\mathcal{S} = \mathcal{A} \cup \mathcal{B} \cup \mathcal{I}$ into two (possibly directly connected) regions $\mathcal{A}, \mathcal{B}$ containing $N_\mathcal{A}, N_\mathcal{B}$ minima respectively, and an intervening region $\mathcal{I}$ containing $N_\mathcal{I}$ minima. We assume that $\mathcal{A}$ and $\mathcal{B}$ are metastable, a statement which will be made precise when considering KTN observables. Without loss of generality, we focus on transitions from $\mathcal{B}$ to $\mathcal{A}$, which we denote as $\mathcal{A} \leftarrow \mathcal{B}$. Let $\mathbb{I}_\mathcal{X}$ be the identity matrix in $\mathbb{R}^{N_\mathcal{X} \times N_\mathcal{X}}$, i.e. of dimension equal to the number of minima in region $\mathcal{X}$, and $\mathbf{1}_\mathcal{X} \in \mathbb{R}^{N_\mathcal{X}}$ a row vector of ones of the same dimension. Before reaching equilibrium, the probability density vector $\mathbf{P}_\mathcal{X} \in \mathbb{R}^{\mathbb{N}_\mathcal{X}}$ in a region $\mathcal{X} \in \{\mathcal{A}, \mathcal{B}, \mathcal{I}\}$ evolves according to the master

equation[?]

$$\dot{\mathbf{P}}_{\mathcal{X}} = -\mathbf{D}_{\mathcal{X}}\mathbf{P}_{\mathcal{X}} + \sum_{\mathcal{Y}\in\{\mathcal{A},\mathcal{B},\mathcal{I}\}} \mathbf{K}_{\mathcal{X}\mathcal{Y}}\mathbf{P}_{\mathcal{Y}}, \tag{1}$$

where $\mathbf{K}_{\mathcal{X}\mathcal{Y}} \in \mathbb{R}^{\mathbb{N}_{\mathcal{X}}\times\mathbb{N}_{\mathcal{Y}}}$ is a (rectangular) matrix of all the minimum-to-minimum rates $\mathcal{Y} \to \mathcal{X}$, and $\mathbf{D}_{\mathcal{X}} \in \mathbb{R}^{\mathbb{N}_{\mathcal{X}}\times\mathbb{N}_{\mathcal{X}}}$ is a diagonal matrix whose entries are the total escape rate from each state in $\mathcal{X}$. Conservation of probability requires $\sum_{\mathcal{X}\in\{\mathcal{A},\mathcal{B},\mathcal{I}\}} \mathbf{1}_{\mathcal{X}}\dot{\mathbf{P}}_{\mathcal{X}} = 0$, for any probability density, which in turn implies that

$$\mathbf{1}_{\mathcal{X}}\mathbf{D}_{\mathcal{X}} = \sum_{\mathcal{Y}\in\{\mathcal{A},\mathcal{B},\mathcal{I}\}} \mathbf{1}_{\mathcal{Y}}\mathbf{K}_{\mathcal{Y}\mathcal{X}}. \tag{2}$$

Assuming detailed balance holds, the probability distribution $\mathbf{P}_{\mathcal{X}}$ will relax exponentially to the equilibrium occupation probabilities $\boldsymbol{\pi}_{\mathcal{X}}$ for all the component minima. It is useful to define the normalised restricted equilibrium distributions $\widehat{\boldsymbol{\pi}}_{\mathcal{X}} = \boldsymbol{\pi}_{\mathcal{X}}/\mathbf{1}_{\mathcal{X}}\boldsymbol{\pi}_{\mathcal{X}}$, such that $\mathbf{1}_{\mathcal{X}}\widehat{\boldsymbol{\pi}}_{\mathcal{X}} = 1$. Hence a component, such as $\widehat{\pi}_b$, is the equilibrium conditional probability that the system is associated with local minimum b given that we are in region $\mathcal{B}$. These conditional probabilities appear in all the rate constant formulations previously derived in the DPS framework,[? ? ? ? ?] as discussed below.

We now introduce the branching probability $\mathrm{B}_{ij} = \mathrm{P}(j \to i|j)$ that the next step of a Markov jump process will be $j \to i$. The branching probability plays a key role in DPS[? ? ?] and kinetic Monte Carlo simulations.[?] In terms of the transition rates defined above, branching probability matrices are given by $\mathbf{B}_{\mathcal{X}\mathcal{Y}} \equiv \mathbf{K}_{\mathcal{X}\mathcal{Y}}\mathbf{D}_{\mathcal{Y}}^{-1}$. Due to the (assumed) Markovinity of the state-to-state dynamics, just as $[\mathbf{B}_{\mathcal{X}\mathcal{X}}]_{ij}$ gives the probability of reaching $i \in \mathcal{X}$ from $j \in \mathcal{X}$ in one jump, the sum

$$\sum_l [\mathbf{B}_{\mathcal{X}\mathcal{X}}]_{il}[\mathbf{B}_{\mathcal{X}\mathcal{X}}]_{lj} = [\mathbf{B}^2_{\mathcal{X}\mathcal{X}}]_{ij}, \tag{3}$$

gives the probability of reaching i from j in exactly two jumps within $\mathcal{X}$. More generally, we consider the sum of all possible path probabilities, conditional on not leaving $\mathcal{X}$. Making the eigendecomposition $\mathbf{B}_{\mathcal{X}\mathcal{X}} = \sum_l \lambda_l \mathbf{v}_l^R \otimes \mathbf{v}_l^L$ it is straightforward to show the total probability of all $\mathcal{X} \to \mathcal{X}$ paths (with recrossings) reads

$$\mathbf{G}_{\mathcal{X}} \equiv \sum_{n=0}^{\infty} \mathbf{B}^n_{\mathcal{X}\mathcal{X}} = \sum_l \frac{\mathbf{v}_l^R \otimes \mathbf{v}_l^L}{1-\lambda_l} = [\mathbb{I}_{\mathcal{X}} - \mathbf{B}_{\mathcal{X}\mathcal{X}}]^{-1}, \tag{4}$$

where we define the matrix $\mathbf{G}_{\mathcal{X}}$ to simplify later expressions and $\otimes$ is the diadic (outer) product; for two vectors $\mathbf{a}$, $\mathbf{b}$ the outer product is the matrix $\mathbf{a}\otimes\mathbf{b}$ with elements $[\mathbf{a}\otimes\mathbf{b}]_{ij} = \mathrm{a}_i\mathrm{b}_j$. This inversion is nonsingular ($\lambda_l < 1$) when escape from $\mathcal{X}$ is possible.

The Green's matrices (**??**) are very useful when deriving KTN observables for transitions between $\mathcal{A}$ and $\mathcal{B}$; in particular, the Green's matrix $\mathbf{G}_{\mathcal{I}}$ provides a compact manner to write the branching probability $\mathbf{B}^{\mathcal{I}}_{\mathcal{X}\mathcal{Y}}$ of leaving $\mathcal{Y} \in \{\mathcal{A},\mathcal{B}\}$ in one jump, executing a path of arbitrary (possibly zero) length within $\mathcal{I}$ and ending in $\mathcal{X} \in \{\mathcal{A},\mathcal{B}\}$. This result reads

$$\mathbf{B}^{\mathcal{I}}_{\mathcal{X}\mathcal{Y}} \equiv \mathbf{B}_{\mathcal{X}\mathcal{Y}} + \mathbf{B}_{\mathcal{X}\mathcal{I}}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Y}}, \tag{5}$$

where $\mathbf{B}_{\mathcal{X}\mathcal{Y}}$ allows for the possibility of a direct $\mathcal{X} \leftarrow \mathcal{Y}$ jump, bypassing $\mathcal{I}$. As in previous work,[?] we distinguish quantities that implicitly account for intervening minima by the superscript $\mathcal{I}$, which immediately defines another key quantity in DPS, the committor vector:

$$\mathbf{C}^{\mathcal{A}}_{\mathcal{B}} = \mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}} \in \mathbb{R}^{N_{\mathcal{B}}}, \tag{6}$$

with component $[\mathbf{C}^{\mathcal{A}}_{\mathcal{B}}]_b \equiv C_b^{\mathcal{A}}$ the probability of leaving state $b \in \mathcal{B}$ in the first jump, then reaching $\mathcal{A}$ before $\mathcal{B}$.

The compound branching probabilities $\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}$ and $\mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}$ will play a central role when deriving the main KTN observables in this work. Indeed, the total committor probability $\mathbf{1}_{\mathcal{B}}\mathbf{C}^{\mathcal{A}}_{\mathcal{B}}$ will be the objective function for our convergence analysis. However, we will see in later sections that evaluation of the Green's matrices $\mathbf{G}_{\mathcal{X}}$ routinely suffer from severe conditioning issues due to the exponential sensitivity of transition rates,[?] meaning that direct solution is numerically unstable. To overcome these issues we use the graph transformation (GT) method,[? ? ? ?] detailed in the next section.

III. THE GRAPH TRANSFORMATION METHOD

The graph transformation (GT) method[? ? ? ?] is a deterministic technique to remove a state $l \in \mathcal{S}$, giving a new state space $\mathcal{S}\setminus l$ with renormalised branching probabilities and escape times. If the current branching probability matrix is $\mathbf{B}_{\mathcal{S}}$, where by definition $\sum_i [\mathbf{B}_{\mathcal{S}}]_{il} = 1$ for all $l \in \mathcal{S}$, the GT procedure to remove a state l to give $\mathcal{S}\setminus l$ is[?]

$$\begin{aligned}
[\mathbf{B}_{\mathcal{S}}]_{ij} \to [\mathbf{B}_{\mathcal{S}\setminus l}]_{ij} &= [\mathbf{B}_{\mathcal{S}}]_{ij} + \frac{[\mathbf{B}_{\mathcal{S}}]_{il}[\mathbf{B}_{\mathcal{S}}]_{lj}}{1-[\mathbf{B}_{\mathcal{S}}]_{ll}}, \\
[\mathbf{D}^{-1}_{\mathcal{S}}]_{jj} \to [\mathbf{D}^{-1}_{\mathcal{S}\setminus l}]_{jj} &= [\mathbf{D}^{-1}_{\mathcal{S}}]_{jj} + \frac{[\mathbf{D}^{-1}_{\mathcal{S}}]_{ll}[\mathbf{B}_{\mathcal{S}}]_{lj}}{1-[\mathbf{B}_{\mathcal{S}}]_{ll}}.
\end{aligned} \tag{7}$$

Whilst self-transitions $l \to l$ are initially zero (i.e. we start with $[\mathbf{B}_{\mathcal{S}}]_{ll} = 0$), such 'self transition' terms emerge after repeated application of the GT renormalisation, representing the branching probability to paths solely on removed states that start and end on a state l. Including degenerate rearrangements[? ?] rescales the initial branching probabilities and waiting times, leaving the ratios unchanged.[?] The original GT procedure excluded self-transitions from the possible events.[? ?] The branching probabilities and waiting times then scale by a common factor, and the expected waiting time for escape is unaffected.[?] However, when self-transitions are included in the events we obtain a direct connection to committor probabilities.[?]

The GT method was specifically designed[? ? ?] as an alternative to matrix representations of Markov chain observables to overcome the numerical conditioning issues described above. In particular, when $[\mathbf{B}_{\mathcal{S}}]_{ll}$ becomes very close to unity, evaluation of the denominator $1-[\mathbf{B}_{\mathcal{S}}]_{ll}$ induces floating point error. The GT method overcomes this problem by exploiting a state-by-state removal scheme, replacing $1-[\mathbf{B}_{\mathcal{S}}]_{ll}$ with the equivalent term $\sum_{i\neq l}[\mathbf{B}_{\mathcal{S}}]_{il}$, which does not suffer from these issues. The GT approach has indeed been shown to have far superior numerical stability to direct linear algebra solutions across a wide range of systems,[?] and has been exploited to overcome trapping in kinetic Monte Carlo simulations.[?]

A matrix generalization of the GT method has recently been analyzed,[?] where blocks of states are removed simultaneously instead of individually. In fact, the renormalisation procedure can preserve branching probabilities and waiting times for removal of any subsets of states in any sequence. Specifically, suppose that we start with states $\mathcal{I}\cup\mathcal{Z}$ and wish to remove all states belonging to $\mathcal{I}$. Equation (**??**) gives the sum of the products of branching probabilities for all paths starting at z_1, ending at z_2, with $z_2, z_1 \in \mathcal{Z}$ and any number of steps in the $\mathcal{I}$ region in between the end points:

$$\left[\mathbf{B}^{\mathcal{I}}_{\mathcal{Z}\mathcal{Z}}\right]_{z_2 z_1} \equiv \left[\mathbf{B}_{\mathcal{Z}\mathcal{Z}} + \mathbf{B}_{\mathcal{Z}\mathcal{I}}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}}\right]_{z_2 z_1}. \tag{8}$$

We can therefore conserve the probability associated with these paths using the renormalised branching probability matrix $\mathbf{B}^{\mathcal{I}}_{\mathcal{Z}\mathcal{Z}}$ with the elements defined above in $\mathbb{R}^{N_{\mathcal{Z}}\times N_{\mathcal{Z}}}$.

The expected waiting time for a single step transition between z_1 and z_2 can be obtained by summing the waiting time $[\mathbf{D}^{-1}_{\mathcal{Z}}]_{zz}$ or $[\mathbf{D}^{-1}_{\mathcal{I}}]_{ii}$ for each state $z\in\mathcal{Z}$ or $i\in\mathcal{I}$ along a given path, then performing a weighted average over all paths using the product of branching probabilities. Here we assuming Markovian transitions within the KTN, so that the expected time to traverse each path is simply the sum of the expected escape times for each state. To achieve this averaging, we define[? ? ?]

$$\widetilde{\mathbf{B}}_{\mathcal{X}\mathcal{Y}} \equiv \mathbf{B}_{\mathcal{X}\mathcal{Y}}\exp(\zeta\mathbf{D}^{-1}_{\mathcal{Y}}), \quad \widetilde{\mathbf{G}}_{\mathcal{I}} \equiv \left[\mathbb{I}_{\mathcal{I}} - \widetilde{\mathbf{B}}_{\mathcal{I}\mathcal{I}}\right]^{-1}, \tag{9}$$

where $\mathcal{X},\mathcal{Y}\in\{\mathcal{I},\mathcal{Z}\}$, which weights each step by the correct waiting time when taking a derivative with respect to ζ and then setting $\zeta=0$. Since branching probabilities must sum to one over all possible connections we have $\mathbf{1}_{\mathcal{I}} = \mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}} + \mathbf{1}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{I}}$, so that $\mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}} = \mathbf{1}_{\mathcal{I}}(\mathbb{I}_{\mathcal{I}} - \mathbf{B}_{\mathcal{I}\mathcal{I}})$ and hence $\mathbf{1}_{\mathcal{I}} = \mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}}\mathbf{G}_{\mathcal{I}}$. In Appendix **??** we provide a summary of other useful relations between these quantities.

We can use these identities to simplify the expression for the expected waiting time for a transition from any state $z\in\mathcal{Z}$ to any another state in $\mathcal{Z}$ via an arbitrary number of steps between states in $\mathcal{I}$. This waiting time becomes the expected escape time from z when the $\mathcal{I}$ states are renormalised away and is given by the z component of (see Appendix **??** for a full derivation). Here we note that $\mathbf{1}_{\mathcal{Z}}\mathbf{B}^{\mathcal{I}}_{\mathcal{Z}\mathcal{Z}} = \mathbf{1}_{\mathcal{Z}}$, so the sum of path weights in question for any component of $\mathcal{Z}$ is unity, so the expected escape times are:

$$\begin{aligned}\mathbf{1}_{\mathcal{Z}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{Z}}]^{-1} \equiv& \mathbf{1}_{\mathcal{Z}}\frac{\partial}{\partial\zeta}\left(\widetilde{\mathbf{B}}^{\mathcal{I}}_{\mathcal{Z}\mathcal{Z}}\right)\Big|_{\zeta=0}\\ =& \mathbf{1}_{\mathcal{Z}}\frac{\partial}{\partial\zeta}\left(\widetilde{\mathbf{B}}_{\mathcal{Z}\mathcal{Z}} + \widetilde{\mathbf{B}}_{\mathcal{Z}\mathcal{I}}\widetilde{\mathbf{G}}_{\mathcal{I}}\widetilde{\mathbf{B}}_{\mathcal{I}\mathcal{Z}}\right)\Big|_{\zeta=0}\\ =& \mathbf{1}_{\mathcal{Z}}\mathbf{D}^{-1}_{\mathcal{Z}} + \mathbf{1}_{\mathcal{I}}\mathbf{D}^{-1}_{\mathcal{I}}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}}\end{aligned} \tag{10}$$

which defines the renormalised waiting times $\mathbf{1}_{\mathcal{Z}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{Z}}]^{-1}$. Hence the expected waiting times associated with direct transitions between $\mathcal{Z}$ states when the $\mathcal{I}$ states are removed are conserved if we replace the original values by the diagonal elements of $[\mathbf{D}^{\mathcal{I}}_{\mathcal{Z}}]^{-1}$. The renormalisation conserves the path probabilities as branching ratios between all $\mathcal{Z}$ states exactly, but not individual first passage times. As for the previous derivation of state-by-state renormalisation, the sum over end points in $\mathcal{Z}$ conserves the escape time.

The renormalisation only changes values for $\mathcal{Z}$ states that are first or second neighbours of $\mathcal{I}$ states, but the formulae can be be applied for all $\mathcal{Z}$. We will also obtain the same results if we remove sets of states in any order, so long as the final state space is the same. In particular, we recover the results in equation (**??**) if we remove a single state. Furthermore, we can partition $\mathcal{Z}$ arbitrarily, for example, into product and reactant portions, which we indicate by $\mathcal{Z}=\mathcal{A}\cup\mathcal{B}$. We then obtain renormalised branching probabilities and escape times:

$$\begin{bmatrix}\mathbf{B}_{\mathcal{A}\mathcal{A}} & \mathbf{B}_{\mathcal{A}\mathcal{I}} & \mathbf{B}_{\mathcal{A}\mathcal{B}}\\ \mathbf{B}_{\mathcal{I}\mathcal{A}} & \mathbf{B}_{\mathcal{I}\mathcal{I}} & \mathbf{B}_{\mathcal{I}\mathcal{B}}\\ \mathbf{B}_{\mathcal{B}\mathcal{A}} & \mathbf{B}_{\mathcal{B}\mathcal{I}} & \mathbf{B}_{\mathcal{B}\mathcal{B}}\end{bmatrix} \rightarrow \mathbf{B}^{\mathcal{I}}_{\mathcal{Z}\mathcal{Z}} = \begin{bmatrix}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{A}} & \mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}\\ \mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{A}} & \mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}\end{bmatrix}, \tag{11}$$

$$\begin{bmatrix}\mathbf{1}_{\mathcal{A}}[\mathbf{D}_{\mathcal{A}}]^{-1}\\ \mathbf{1}_{\mathcal{I}}[\mathbf{D}_{\mathcal{I}}]^{-1}\\ \mathbf{1}_{\mathcal{B}}[\mathbf{D}_{\mathcal{B}}]^{-1}\end{bmatrix} \rightarrow \mathbf{1}_{\mathcal{Z}}\left[\mathbf{D}^{\mathcal{I}}_{\mathcal{Z}}\right]^{-1} = \begin{bmatrix}\mathbf{1}_{\mathcal{A}}\left[\mathbf{D}^{\mathcal{I}}_{\mathcal{A}}\right]^{-1}\\ \mathbf{1}_{\mathcal{B}}\left[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}\end{bmatrix}.$$

The components of $\mathbf{1}_{\mathcal{B}}\left[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}$ correspond to the renormalised escape times $\tau^{\mathcal{I}}_b$ in previous work;[?] we will use them again in §**??**.

As we shall see, the flexibility to remove blocks of states and conserve the branching probabilities and waiting times of interest is useful for analysis, but suffers from the same numerical issues as those that the state-by-state GT method overcomes.

In the next section we derive some formally exact results from the full Markov chain (**??**), then show how these may be generalized to produce expressions for the waiting time and branching probabilities found in previous work. The extended GT results described above will then be used to study how reaction rates emerge in a KTN. We note here that the renormalisation procedure conserves the escape time for a transition within the remaining state space, while approximations involving steady state assumptions for $\mathcal{I}$ states do not. Nevertheless, the steady state approximation can be useful for

analysis: it has been exploited in DPS,[? ?] and the same formulae result from a coarse-graining approach.[?]

IV. EXACT OBSERVABLES FOR A KTN

When $\mathcal{A}$ and $\mathcal{B}$ are metastable, $\mathcal{A} \leftrightarrow \mathcal{B}$ transitions will be rare and typically followed by long periods in the product region. It is therefore meaningful to ask for the expected *waiting time* to reach, say $\mathcal{A}$, given that we prepare the initial distribution in $\mathcal{B}$, i.e. $\mathbf{P}_{\mathcal{I}}(0) = \mathbf{0}_{\mathcal{I}}$ and $\mathbf{P}_{\mathcal{A}}(0) = \mathbf{0}_{\mathcal{A}}$.

The $\mathcal{A} \leftarrow \mathcal{B}$ waiting time can be evaluated exactly by studying an artificial system where all transitions out of $\mathcal{A}$ are set to zero, i.e. $\mathbf{K}_{\mathcal{X}\mathcal{A}} = \mathbf{0}$. Whilst this system then clearly violates detailed balance (as all trajectories will eventually end in $\mathcal{A}$ for a connected network), the dynamics before reaching $\mathcal{A}$ are unperturbed. In this limit the dynamics in $\mathcal{I} \cup \mathcal{B}$ follow a master equation analogous to (**??**)

$$\begin{bmatrix}\dot{\mathbf{P}}_{\mathcal{I}}\\ \dot{\mathbf{P}}_{\mathcal{B}}\end{bmatrix} = \begin{bmatrix}\mathbf{K}_{\mathcal{I}\mathcal{I}} - \mathbf{D}_{\mathcal{I}} & \mathbf{K}_{\mathcal{I}\mathcal{B}}\\ \mathbf{K}_{\mathcal{B}\mathcal{I}} & \mathbf{K}_{\mathcal{B}\mathcal{B}} - \mathbf{D}_{\mathcal{B}}\end{bmatrix}\begin{bmatrix}\mathbf{P}_{\mathcal{I}}\\ \mathbf{P}_{\mathcal{B}}\end{bmatrix}, \tag{12}$$

where all quantities are exactly as defined in (**??**), meaning transitions to $\mathcal{A}$ are encoded only in the diagonal matrices $\mathbf{D}_{\mathcal{I}}$ and $\mathbf{D}_{\mathcal{B}}$. With no further approximations it is possible to evaluate the expected waiting time to reach $\mathcal{A}$ conditional on starting in $\mathcal{B}$ analytically, i.e. with initial conditions $\mathbf{P}_{\mathcal{X}}(0)$ such that $\mathbf{1}_{\mathcal{B}}\mathbf{P}_{\mathcal{I}}(0) = 0$, $\mathbf{1}_{\mathcal{A}}\mathbf{P}_{\mathcal{A}}(0) = 0$ and $\mathbf{1}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}}(0) = 1$. The probability density for the waiting time τ is simply the probability flux out of $\mathcal{I} \cup \mathcal{B}$:

$$\mathrm{P}(\tau \in [t, t + \mathrm{d}t]) = -(\mathbf{1}_{\mathcal{I}}\dot{\mathbf{P}}_{\mathcal{I}} + \mathbf{1}_{\mathcal{B}}\dot{\mathbf{P}}_{\mathcal{B}})\mathrm{d}t, \tag{13}$$

The integral of this density is the total change of probability in $\mathcal{I} \cup \mathcal{B}$ in the limit $t \to \infty$, which is clearly unity. We can therefore express the expected waiting time as[?]

$$\begin{aligned}\langle \tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle &\equiv \frac{\int_0^\infty tP(\tau \in [t, t + \mathrm{d}t])}{\int_0^\infty P(\tau \in [t, t + \mathrm{d}t])}, \\ &= \begin{bmatrix}\mathbf{1}_{\mathcal{I}}\\ \mathbf{1}_{\mathcal{B}}\end{bmatrix}^\top \begin{bmatrix}\mathbf{D}_{\mathcal{I}} - \mathbf{K}_{\mathcal{I}\mathcal{I}} & -\mathbf{K}_{\mathcal{I}\mathcal{B}}\\ -\mathbf{K}_{\mathcal{B}\mathcal{I}} & \mathbf{D}_{\mathcal{B}} - \mathbf{K}_{\mathcal{B}\mathcal{B}}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{0}_{\mathcal{I}}\\ \mathbf{P}_{\mathcal{B}}(0)\end{bmatrix},\end{aligned} \tag{14}$$

where the second line arrives by considering the evolution equation (**??**). We derive this result in appendix **??**, verifying that the denominator of the first line above is -1, and then solve the matrix inversion analytically to give

$$\begin{aligned}\langle \tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle &= \left[\mathbf{1}_{\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{B}} + \mathbf{1}_{\mathcal{B}}\mathbf{D}_{\mathcal{B}}^{-1}\right]\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}}(0)\\ &= \mathbf{1}_{\mathcal{B}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]^{-1}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}}(0),\end{aligned} \tag{15}$$

where the Green's matrix $\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}$ is defined as

$$\mathbf{G}^{\mathcal{I}}_{\mathcal{B}} = \left[\mathbb{I}_{\mathcal{B}} - \mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}\right]^{-1} \tag{16}$$

i.e. the Green's matrix corresponding to the compound branching probability of all possible non-reactive paths starting and ending in $\mathcal{B}$ without reaching $\mathcal{A}$. The matrix $\mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}$ corresponds to the definition in equation (**??**) and motivates the notation employed for $\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}$.

A. Exact waiting time from pathwise averages

To make a connection with previous work on KTN observables, we wish to interpret the matrix expression (**??**) in terms of weighted averages over all possible paths connecting $\mathcal{B}$ to $\mathcal{A}$. These path weights form the starting point for all the mean first passage time and rate constant derivations in DPS theory based upon the graph transformation renormalisation approach.[? ? ?] In particular, we previously wrote the product of branching probabilities for any discrete path ξ as $\mathcal{W}_\xi$.[?] Element b of the vector $\mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}$ is then identified with the sum of path weights $\mathcal{W}_\xi$ over $a \in \mathcal{A}$ and paths $\xi \in a \leftarrow b$. Since $[\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}]_{b'b}$ gives the sum of all probabilities for paths starting at $b \in \mathcal{B}$ and ending in $b' \in \mathcal{B}$ without reaching $\mathcal{A}$, and $[\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}]_{ab'}$ is the probability of reaching $a \in \mathcal{A}$ from b' summed over all paths that do not return to $\mathcal{B}$, the product is the sum of probabilities over all possible $a \leftarrow b$ paths. Every element of this vector is unity: by conservation of probability $\mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}} = \mathbf{1}_{\mathcal{B}} - \mathbf{1}_{\mathcal{B}}\mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}} = \mathbf{1}_{\mathcal{B}}\left[\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}$, giving

$$\mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}} = \mathbf{1}_{\mathcal{B}}\left[\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}} = \mathbf{1}_{\mathcal{B}}, \tag{17}$$

which in turn implies that $\mathbf{C}^{\mathcal{A}}_{\mathcal{B}}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}}(0) = 1$ for any initial condition. To obtain the expected waiting time we follow the procedure of §**??**, and in appendix **??** we show that the exact waiting time can be written

$$\begin{aligned}\langle \tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle &= \frac{\partial}{\partial\zeta}\left(\mathbf{1}_{\mathcal{A}}\widetilde{\mathbf{B}}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}\widetilde{\mathbf{G}}^{\mathcal{I}}_{\mathcal{B}}\right)\Big|_{\zeta=0}\mathbf{P}_{\mathcal{B}}(0)\\ &= \mathbf{1}_{\mathcal{B}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]^{-1}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}}(0),\end{aligned} \tag{18}$$

which connects the expected waiting time to the sums of path weights. Here we identify the components $\left[\mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]_b$ with the waiting times $\mathcal{T}_{\mathcal{A}b}$ in previous work.[?]

B. Exact waiting time using the GT method

The exact waiting time (**??**) can also be obtained using the graph transformed branching probabilities (**??**) and escape times **??**, with $\mathbf{K}_{\mathcal{X}\mathcal{A}} = \mathbf{0}$, to form a reduced Markov chain executing dynamics in a state space $\mathcal{B}^{\mathcal{I}}$ the same size as $\mathcal{B}$, which subsumes excursions into $\mathcal{I}$ before eventual absorption in $\mathcal{A}$. The $\mathcal{I}$ superscript indicates that all the $\mathcal{I}$ minima have been renormalised away, as above.

The transition rate matrix between renormalised states in $i, j \in \mathcal{B}^{\mathcal{I}}$ is given by $\mathbf{K}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}} = \mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}$, i.e. the branching probability multiplied by the total escape rate. The absorbing master equation (**??**) in $\mathcal{I} \cup \mathcal{B}$ is then transformed to a renormalised master equation in $\mathcal{B}^{\mathcal{I}}$ evolving via $\mathbf{K}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}} - \mathbf{D}^{\mathcal{I}}_{\mathcal{B}}$; using the definition (**??**) of $\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}$ the evolution equation for $\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(t)$ becomes

$$\dot{\mathbf{P}}_{\mathcal{B}^{\mathcal{I}}}(t) = -\left[\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(t). \tag{19}$$

Employing the same procedure as employed to derive the exact waiting time (**??**) in Appendix **??**, we can write the exact waiting time of (**??**) as

$$\langle \tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle = \mathbf{1}_{\mathcal{B}}\left[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(0), \tag{20}$$

which is identical to the exact result of (**??**) for the full system when $\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(0) = \mathbf{P}_{\mathcal{B}}(0)$, consistent with the definitions in (**??**) and (§**??**). The required initial condition in $\mathcal{B}^{\mathcal{I}}$ to reproduce the exact waiting time when $\mathbf{P}_{\mathcal{I}}(0) \neq \mathbf{0}_{\mathcal{I}}$ is given in appendix **??**. The notation $\mathcal{B}^{\mathcal{I}}$ for the probabilities is used to emphasise that although (**??**) has exact escape statistics to $\mathcal{A}$, the GT procedure changes the nature of the remaining state space. The exact waiting time of the full (or reduced) model is the expected time spent in $\mathcal{I}\cup\mathcal{B}$ (or $\mathcal{B}^{\mathcal{I}}$) before reaching $\mathcal{A}$, which is clearly not equal to the expected time spent in $\mathcal{B}$. The state space $\mathcal{B}^{\mathcal{I}}$ therefore has the same rank as $\mathcal{B}$, but evolves with branching probabilities and escape times that account for all possible sojourns into $\mathcal{I}$. This distinction is important when defining the metastability of $\mathcal{B}$ and transition rates $k_{\mathcal{A}\leftarrow\mathcal{B}}$, as we see in the next section.

C. Defining an exact transition rate

The matrix formalism allows us to define exact expressions for the expected waiting time and branching probabilities for any initial condition and energy landscape; the metastability of $\mathcal{A}$ and $\mathcal{B}$ determine the utility, not the accuracy, of (**??**) and (**??**). However, the existence of a constant reaction rate $k_{\mathcal{A}\leftarrow\mathcal{B}}$ on a suitable observation time scale is more subtle, as it is only well defined (i.e. has a constant value) when the decay into $\mathcal{A}$ follows single exponential kinetics. A weak condition for such kinetics is metastability in $\mathcal{I}\cup\mathcal{B}$; a stronger condition is metastability in $\mathcal{B}$ alone.

In this section, we first derive the conditions for an exact transition rate, dependent on metastability in $\mathcal{I}\cup\mathcal{B}$, before seeing how this relates to previous work.

When $\mathcal{I}\cup\mathcal{B}$ is metastable, the rate matrix in (**??**) will have a spectral gap. More precisely, if we write the eigendecomposition of the rate matrix as

$$\mathbf{M} = \begin{bmatrix}\mathbf{K}_{\mathcal{I}\mathcal{I}} - \mathbf{D}_{\mathcal{I}} & \mathbf{K}_{\mathcal{I}\mathcal{B}}\\ \mathbf{K}_{\mathcal{B}\mathcal{I}} & \mathbf{K}_{\mathcal{B}\mathcal{B}} - \mathbf{D}_{\mathcal{B}}\end{bmatrix} = -\sum_l \nu_l \mathbf{w}^R_l \otimes \mathbf{w}^L_l, \tag{21}$$

single exponential decay from $\mathcal{I}\cup\mathcal{B}$ will emerge if the ordered eigenvalues, numbering from zero, satisfy $0 < \nu_0 \ll \nu_1 < \nu_2$...., i.e. have a simple spectral gap $\nu_0 \ll \nu_1$. To verify this limit, we solve for the probability distribution in the eigenbasis. Writing $\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(t) \equiv [\mathbf{P}_{\mathcal{I}}(t), \mathbf{P}_{\mathcal{B}}(t)]$, we find the multiexponential solution

$$\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(t) = \sum_l \exp(-t\nu_l)\left(\mathbf{w}^L_l\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0)\right)\mathbf{w}^R_l, \tag{22}$$

which will decay to a projection along the slowest eigenmode $\mathbf{w}^R_0$ on a timescale $(\nu_1-\nu_0)^{-1}$. The limiting distribution in $\mathcal{I}\cup\mathcal{B}$, conditional on not being absorbed, is therefore

$$\lim_{t\to\infty}\frac{\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(t)}{\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(t)} = \frac{\mathbf{w}^R_0}{\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{w}^R_0} \equiv \widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{I}\cup\mathcal{B}}, \tag{23}$$

where the superscript QSD signifies that we have defined the *quasistationary* distribution $\boldsymbol{\pi}^{QSD}_{\mathcal{I}\cup\mathcal{B}} = \mathbf{w}^R_0$ for $\mathcal{I}\cup\mathcal{B}$. The QSD is defined as the limiting distribution in a region with absorbing boundaries conditional on not being absorbed, a natural and useful generalization of the local equilibrium distribution for metastable states.[? ?] Importantly, if prepared in the QSD, basin escape statistics are single exponential, as we demonstrate below. For any eigenspectrum, i.e. any degree of metastability, any initial conditions will decay to the QSD on a timescale $(\nu_1-\nu_0)^{-1}$. The spectral gap condition means that the escape time scale for the slowest mode, ν_0^{-1}, is much longer than the time required to establish the relative probabilities corresponding to the QSD, namely $(\nu_1-\nu_0)^{-1}$ When we have simple metastability, $\nu_0 \ll \nu_1$, KTN observables approach single exponential kinetics on this timescale, giving an *exact* escape rate of

$$-\lim_{t\to\infty}\frac{\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\dot{\mathbf{P}}_{\mathcal{I}\cup\mathcal{B}}(t)}{\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(t)} = \nu_0 \equiv k^*_{\mathcal{A}\leftarrow\mathcal{B}}. \tag{24}$$

The exact waiting time (**??**) retains dependence on the initial conditions $\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0)$. Using the eigendecomposition (**??**) we can rewrite the formula for $\langle\tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle$ involving the inverse matrix obtained in Appendix **??** as

$$\langle \tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle = \mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\sum_l \nu_l^{-1}\mathbf{w}^R_l\left(\mathbf{w}^L_l\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0)\right). \tag{25}$$

Note that the eigenvalues of the rate matrix $\mathbf{M}$ in (**??**) are ≤ 0, and the minus sign is chosen so that $\nu_l \geq 0$. This sign cancels the minus sign in the first line of Appendix C equation (**??**) to give (**??**). When $\nu_0 \ll \nu_1$ the first term dominates and we can use (**??**) to replace $\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{w}^R_0$

$$\begin{aligned}\langle \tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle &= \nu_0^{-1}\left(\mathbf{w}^L_0\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0)\right)\left(\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{w}^R_0\right) + \mathcal{O}(\nu_0/\nu_1)\\ &= \frac{1}{k^*_{\mathcal{A}\leftarrow\mathcal{B}}}\frac{\mathbf{w}^L_0\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0)}{\mathbf{w}^L_0\widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{I}\cup\mathcal{B}}} + \mathcal{O}(\nu_0/\nu_1),\end{aligned} \tag{26}$$

which is the inverse rate multiplied by the ratio of projections onto the slowest mode for the initial distribution and the QSD. Hence $\langle\tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle = 1/k^*_{\mathcal{A}\leftarrow\mathcal{B}} \equiv \langle\tau^{*,QSD}_{\mathcal{A}\leftarrow\mathcal{B}}\rangle$ if we simply prepare the system in the QSD, with $\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0) = \widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{I}\cup\mathcal{B}}$.

Furthermore, when $\mathcal{I}\cup\mathcal{B}$ is very metastable, i.e. as the decay rate to $\mathcal{A}$ vanishes ($\mathbf{K}_{\mathcal{A}\mathcal{I}} \to 0, \mathbf{K}_{\mathcal{A}\mathcal{B}} \to 0$), the slowest eigenvalue ν_0 will also tend to zero. In this limit $\mathbf{w}^R_0$, and by definition $\boldsymbol{\pi}^{QSD}$, will be proportional to the invariant local equilibrium distribution of (**??**), which before any GT renormalisation is simply $\boldsymbol{\pi}_{\mathcal{I}\cup\mathcal{B}}$. Since $\dot{\mathbf{P}}_{\mathcal{I}\cup\mathcal{B}}(t) = \mathbf{M}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(t) = \mathbf{0}_{\mathcal{I}\cup\mathcal{B}}$ for any $\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(t)$ we also know that the corresponding left eigenvector $\mathbf{w}^L_0$ is proportional to $\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}$. The orthonormality relation $\mathbf{w}^L_p\mathbf{w}^R_q = \delta_{pq}$ provides the constants of proportionality

as $\mathbf{w}_0^R = \boldsymbol{\pi}_{\mathcal{I}\cup\mathcal{B}}/\mathbf{w}_0^L\boldsymbol{\pi}_{\mathcal{I}\cup\mathcal{B}}$ and $\mathbf{w}_0^L = \mathbf{1}_{\mathcal{I}\cup\mathcal{B}}/\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{w}_0^R = \mathbf{1}_{\mathcal{I}\cup\mathcal{B}}/\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\boldsymbol{\pi}^{QSD}$. The limiting form for $\mathbf{w}_0^L$ holds even when the limiting form for $\boldsymbol{\pi}^{QSD}$ is no longer the local equilibrium distribution due a GT renormalisation, a point we return to below. The proportionality of $\mathbf{w}_0^L$ to $\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}$ is sufficient to give $\langle\tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle \to 1/k^*_{\mathcal{A}\leftarrow\mathcal{B}}$ by substitution in (**??**), as expected for single exponential decay.

The same result, again consistent with (**??**), is obtained by calculating the expected waiting time from $\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(t)$ given by (**??**), given the system has not decayed to $\mathcal{A}$ after a time t. This waiting time, which corresponds to $\langle\tau^{*,QSD}_{\mathcal{A}\leftarrow\mathcal{B}}\rangle$ defined above, would typically be measured in experiment. Using the notation of (**??**) and multiexponential expansion (**??**) we find

$$\langle\tau^{*,QSD}_{\mathcal{A}\leftarrow\mathcal{B}}\rangle \equiv \lim_{t\to\infty}\frac{\int_t^\infty (t'-t)P(\tau\in[t',t'+\mathrm{d}t'])}{\int_t^\infty P(\tau\in[t',t+\mathrm{d}t'])} = \frac{1}{k^*_{\mathcal{A}\leftarrow\mathcal{B}}}, \tag{27}$$

meaning we can identify $\langle\tau^{*,QSD}_{\mathcal{A}\leftarrow\mathcal{B}}\rangle$ as the correct inverse rate $k^*_{\mathcal{A}\leftarrow\mathcal{B}}$ when decay to $\mathcal{A}$ is simple exponential. This result requires t sufficiently large to retain only the first term in (**??**), and $\nu_0 t$ sufficiently small to be neglected.

Whilst the expressions (**??**) and (**??**) are formally exact, as discussed above the rate matrix (**??**) suffers from significant numerical conditioning issues meaning that in practice evaluation of $k^*_{\mathcal{A}\leftarrow\mathcal{B}}$ is rarely possible.

To proceed, we apply an identical analysis to the graph transformed evolution equation (**??**), which reduces $\mathcal{I}\cup\mathcal{B}$ to $\mathcal{B}^{\mathcal{I}}$, the set of renormalised $\mathcal{B}$ minima, whilst retaining the exact waiting time (**??**). As above, when $\mathcal{B}^{\mathcal{I}}$ is metastable the rate matrix in (**??**) will have a spectral gap. The eigendecomposition of the transformed rate matrix reads

$$-\left[\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}\mathbf{D}^{\mathcal{I}}_{\mathcal{B}} = -\sum_l \nu_l^{\mathcal{I}}\mathbf{z}_l^R\otimes\mathbf{z}_l^L, \tag{28}$$

and single exponential decay from $\mathcal{B}^{\mathcal{I}}$ will again emerge if the ordered eigenvalues satisfy $0<\nu_0^{\mathcal{I}}\ll\nu_1^{\mathcal{I}}<\nu_2^{\mathcal{I}}$..... In this limit we can identity the renormalised QSD, the limiting distribution in $\mathcal{B}$, as

$$\lim_{t\to\infty}\frac{\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(t)}{\mathbf{1}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(t)} = \frac{\mathbf{z}_0^R}{\mathbf{1}_{\mathcal{B}}\mathbf{z}_0^R} \equiv \frac{\boldsymbol{\pi}^{QSD}_{\mathcal{B}^{\mathcal{I}}}}{\mathbf{1}_{\mathcal{B}}\boldsymbol{\pi}^{QSD}_{\mathcal{B}^{\mathcal{I}}}} = \widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{B}^{\mathcal{I}}}, \tag{29}$$

which gives a renormalised escape rate

$$-\lim_{t\to\infty}\frac{\mathbf{1}_{\mathcal{B}}\dot{\mathbf{P}}_{\mathcal{B}^{\mathcal{I}}}(t)}{\mathbf{1}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(t)} = \nu_0^{\mathcal{I}} \equiv k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}. \tag{30}$$

To compare $k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}$ to $k^*_{\mathcal{A}\leftarrow\mathcal{B}}$, the exact waiting time (**??**) can be written in a form closely resembling (**??**)

$$\langle\tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle = \mathbf{1}_{\mathcal{B}}\sum_l(\nu_l^{\mathcal{I}})^{-1}\mathbf{z}_l^R\left(\mathbf{z}_l^L\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(0)\right), \tag{31}$$

which will again be dominated by the first term when $\mathcal{B}^{\mathcal{I}}$ is metastable, giving

$$\langle\tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle = \frac{1}{k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}}\frac{\mathbf{z}_0^L\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(0)}{\mathbf{z}_0^L\widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{B}^{\mathcal{I}}}} + \mathcal{O}(\nu_0^{\mathcal{I}}/\nu_1^{\mathcal{I}}) \tag{32}$$

In the highly metastable limit, when the probability flux out of $\mathcal{B}^{\mathcal{I}}$ vanishes, we have again $\mathbf{z}_0^L\to\mathbf{1}_{\mathcal{B}}/(\mathbf{1}_{\mathcal{B}}\boldsymbol{\pi}^{QSD}_{\mathcal{B}^{\mathcal{I}}})$ and thus $\langle\tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle\to 1/k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}$, meaning that $k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}} = k^*_{\mathcal{A}\leftarrow\mathcal{B}}$.

In this limit $\boldsymbol{\pi}^{QSD}_{\mathcal{B}^{\mathcal{I}}}\to\boldsymbol{\pi}_{\mathcal{B}^{\mathcal{I}}}$, the invariant distribution of the transformed rate equation (**??**) when decay to $\mathcal{A}$ is vanishing. Importantly, $\boldsymbol{\pi}_{\mathcal{B}^{\mathcal{I}}}$ is *not* equal to $\boldsymbol{\pi}_{\mathcal{B}}$, the Boltzmann distribution in $\mathcal{B}$, which will become relevant when comparing the exact rate to previous approximations.

The correspondence between the QSD rate $k^*_{\mathcal{A}\leftarrow\mathcal{B}}$, equation (**??**), and the transformed QSD rate $k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}$, equation (**??**), can also be obtained as for (**??**) by calculating the expected waiting time conditional on not decaying to $\mathcal{A}$,

$$\langle\tau^{*,QSD}_{\mathcal{A}\leftarrow\mathcal{B}}\rangle \equiv \lim_{t\to\infty}\frac{\mathbf{1}_{\mathcal{B}}\left[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(t)}{\mathbf{1}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(t)} = \frac{1}{k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}}. \tag{33}$$

Comparing the two equivalent exact results for the expected waiting time, we see that the transformed QSD rate $k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}$ (which we can calculate) and the exact QSD rate $k^*_{\mathcal{A}\leftarrow\mathcal{B}}$ (which is typically impossible to calculate) are equal in the metastable limit. Hence $k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}$ can be considered an exact rate in the vast majority of cases, as the GT transformation is typically required only to calculate KTN observables in the highly metastable (rare event) limit, where we have demonstrated exponentially accurate correspondence between the GT QSD rate $k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}$ and the formally exact, but typically incalculable, QSD transition rate $k^*_{\mathcal{A}\leftarrow\mathcal{B}}$. We also show in the next section that it is only in this limit, when these rates agree with the transition rate defined as the exact reference in previous work, that they also agree with each other.

To facilitate the comparison to previous results for the transition rate, we end this section by deriving an equivalent expression for the transformed QSD rate (**??**). As discussed above, the QSD projects out $k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}$ from the transformed rate matrix, equation (**??**). Substituting for $\dot{\mathbf{P}}_{\mathcal{B}^{\mathcal{I}}}(t)$ from this equation and using $\boldsymbol{\pi}^{QSD}_{\mathcal{B}^{\mathcal{I}}}$ for $\mathbf{P}_{\mathcal{B}^{\mathcal{I}}}(t)$ in the numerator and denominator of (**??**) gives

$$k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}} = \frac{\mathbf{1}_{\mathcal{B}}\left[\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}\boldsymbol{\pi}^{QSD}_{\mathcal{B}^{\mathcal{I}}}}{\mathbf{1}_{\mathcal{B}}\boldsymbol{\pi}^{QSD}_{\mathcal{B}^{\mathcal{I}}}}. \tag{34}$$

By the conservation of branching probabilities and the committor definition (**??**) we have $\mathbf{C}^{\mathcal{A}}_{\mathcal{B}} = \mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}} = \mathbf{1}_{\mathcal{B}}\left[\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}$ (see also (**??**) above). We can therefore express the transformed QSD transition rate in the suggestive form

$$k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}} = \mathbf{C}^{\mathcal{A}}_{\mathcal{B}}\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}\widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{B}^{\mathcal{I}}}. \tag{35}$$

Equation (**??**) is a very useful result, an expression for the KTN transition rate that is exact in the metastable limit where a transition rate is well defined, in a form that we can relate to existing approximations identified

in previous work as $k^{SS}_{\mathcal{A}\leftarrow\mathcal{B}}$ for steady state (SS)[?] and $k^{NSS}_{\mathcal{A}\leftarrow\mathcal{B}}$ non-steady state (NSS)[?] , as detailed in the next section.

We note that the exact rate in (**??**) can be considered a sum over the branching probabilities $\mathbf{C}^{\mathcal{A}}_{\mathcal{B}}$ for each reactive trajectory in $\mathcal{B}$ (i.e. those that reach $\mathcal{A}$ before returning to $\mathcal{B}$) weighted by the total escape rate and QSD weight for the corresponding renormalised state in $\mathcal{B}^{\mathcal{I}}$. This formulation is consistent with the probability distribution for reactive trajectories.[? ?] We also note the similarity to formulae based on the flux over a dividing surface.[? ? ? ?]

D. Comparison of the exact rate to previous work

The steady state approximation used in the original DPS derivation[? ?] assumes that the intervening region $\mathcal{I}$ is in a steady state, $\dot{\mathbf{P}}_{\mathcal{I}} = \mathbf{0}$, and that the $\mathcal{A},\mathcal{B}$ distributions are in local equilibrium, $\mathbf{P}_{\mathcal{X}}(t) = \widehat{\boldsymbol{\pi}}_{\mathcal{X}}\mathrm{P}_{\mathcal{X}}(t)$ for $\mathcal{X} = \mathcal{A},\mathcal{B}$. In appendix **??** we show that the transition rate in this approximation is given by

$$k^{SS}_{\mathcal{A}\leftarrow\mathcal{B}} = \mathbf{C}^{\mathcal{A}}_{\mathcal{B}}\mathbf{D}_{\mathcal{B}}\widehat{\boldsymbol{\pi}}_{\mathcal{B}} = \sum_{b\in\mathcal{B}}\frac{[\mathbf{C}^{\mathcal{A}}_{\mathcal{B}}]_b[\widehat{\boldsymbol{\pi}}_{\mathcal{B}}]_b}{\tau_b}, \tag{36}$$

where the final expression uses $\tau_b = 1/[\mathbf{D}_{\mathcal{B}}]_{bb} = [\mathbf{D}^{-1}_{\mathcal{B}}]_{bb}$, the expected waiting time for a transition out of minimum b, demonstrating that $k^{SS}_{\mathcal{A}\leftarrow\mathcal{B}}$ is precisely that derived in previous work.[?] Through comparison with (**??**) we see that the effect of the steady state approximation is to replace the GT renormalised escape times $1/[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]_{bb}$ with $1/[\mathbf{D}_{\mathcal{B}}]_{bb}$ and the normalised QSD $\widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{B}^{\mathcal{I}}}$ is replaced by the local equilibrium occupation probability $\widehat{\boldsymbol{\pi}}_{\mathcal{B}}$. This result amounts to assuming that traversal of the $\mathcal{I}$ region is instantaneous, and that the presence of nonzero escape rates from $\mathcal{B}$ affects the limiting distribution (which is the content of the QSD). We expect that the latter assumption is acceptable for sufficiently metastable basins, but this first assumption is typically accurate only for simple landscapes.

The non-steady state (NSS) formula for the transition rate was derived in previous work by considering transitions within the state space of $\mathcal{A}\cup\mathcal{B}$ after renormalising away the $\mathcal{I}$ minima one-by-one.[?] If we denote the corresponding rate constants by an $\mathcal{I}$ superscript, and treat all the transitions as competing Poisson processes, then the expected waiting time for a transition from b to $\mathcal{A}\cup\mathcal{B}$ is the renormalised value $\tau^{\mathcal{I}}_b = 1/(K^{\mathcal{I}}_{\mathcal{A}b} + K^{\mathcal{I}}_{\mathcal{B}b})$ and we identify the committor probability $C^{\mathcal{A}}_b = K^{\mathcal{I}}_{\mathcal{A}b}/(K^{\mathcal{I}}_{\mathcal{A}b} + K^{\mathcal{I}}_{\mathcal{B}b})$, which is obtained from the renormalised branching probability.[? ?] The required rate is then $K^{\mathcal{I}}_{\mathcal{A}b} = C^{\mathcal{A}}_b/\tau^{\mathcal{I}}_b$ and we obtained a mean rate constant by averaging over the local equilibrium distribution in $\mathcal{B}$:[? ?]

$$k^{NSS}_{\mathcal{A}\leftarrow\mathcal{B}} = \sum_{b\in\mathcal{B}}\frac{[\mathbf{C}^{\mathcal{A}}_{\mathcal{B}}]_b[\widehat{\boldsymbol{\pi}}_{\mathcal{B}}]_b}{[\boldsymbol{\tau}^{\mathcal{I}}]_b}. \tag{37}$$

$[\boldsymbol{\tau}^{\mathcal{I}}]_b$ is component b of the vector of expected escape times produced by GT. In §**??** we showed that these expected times can be written as $\boldsymbol{\tau}^{\mathcal{I}} = \mathbf{1}_{\mathcal{B}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]^{-1}$. The NSS escape rate then reads

$$k^{NSS}_{\mathcal{A}\leftarrow\mathcal{B}} = \mathbf{C}^{\mathcal{A}}_{\mathcal{B}}\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}\widehat{\boldsymbol{\pi}}_{\mathcal{B}}. \tag{38}$$

which is the exact rate (**??**) derived above with the normalised QSD $\widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{B}^{\mathcal{I}}}$ replaced by the local equilibrium conditional probabilities $\widehat{\boldsymbol{\pi}}_{\mathcal{B}}$. The NSS is thus a significant improvement on the SS rate, as we only assume that the local equilibrium $\widehat{\boldsymbol{\pi}}_{\mathcal{B}}$ is a good approximation to the transformed QSD $\widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{B}^{\mathcal{I}}}$. Here we note that the matrix $\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}$ is always considered to be diagonal, with non-zero elements defined in terms of the reciprocal waiting times from the elements of the diagonal matrix that appear in $\mathbf{1}_{\mathcal{B}}\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}$.

We have seen in the previous section that in the highly metastable limit, the QSD $\widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{B}^{\mathcal{I}}} \to \widehat{\boldsymbol{\pi}}_{\mathcal{B}^{\mathcal{I}}}$, the invariant distribution for the transformed rate matrix (**??**). When there is no metastability in $\mathcal{I}$ we expect the approximation $\widehat{\boldsymbol{\pi}}_{\mathcal{B}^{\mathcal{I}}} \simeq \widehat{\boldsymbol{\pi}}_{\mathcal{B}}$ to be accurate, and thus $k^{NSS}_{\mathcal{A}\leftarrow\mathcal{B}} \simeq k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}$, which in turn will be equal to $k^{*}_{\mathcal{A}\leftarrow\mathcal{B}}$. However, if there is significant metastability in $\mathcal{I}$ it is possible that $\widehat{\boldsymbol{\pi}}_{\mathcal{B}^{\mathcal{I}}}$ could be different from $\widehat{\boldsymbol{\pi}}_{\mathcal{B}}$, inducing error in $k^{NSS}_{\mathcal{A}\leftarrow\mathcal{B}}$.

Finally, we consider the rate obtained from the waiting time $\mathcal{T}_{\mathcal{A}b}$, which is calculated by graph transformation removal of all $\mathcal{I}$ minima and all other sources $b' \neq b$ in $\mathcal{B}$, using local equilibrium conditional occupation probabilities:[?]

$$k^{F}_{\mathcal{A}\leftarrow\mathcal{B}} \equiv \sum_{b\in\mathcal{B}}\frac{[\widehat{\boldsymbol{\pi}}_{\mathcal{B}}]_b}{\mathcal{T}_{\mathcal{A}b}}. \tag{39}$$

Here, $\mathcal{T}_{\mathcal{A}b}$ is the waiting time for a transition from b to $\mathcal{A}$, including revisits to b, obtained from the renormalised branching probabilities and escape time (see Appendix **??**). $k^{NSS}_{\mathcal{A}\leftarrow\mathcal{B}}$ agrees with $k^{F}_{\mathcal{A}\leftarrow\mathcal{B}}$ in (**??**) if the $\mathcal{B}$ minima are in rapid local equilibrium compared to the time scale for transitions to $\mathcal{A}$.[?] Alternatively, considering the first-step relation[?] we can prove that $\mathcal{T}_{\mathcal{A}b} = \tau^{\mathcal{I}}_b/C^{\mathcal{A}}_b$ if $\mathcal{T}_{\mathcal{A}b}$ is the same for all b, and the two rate formulations in (**??**) and (**??**) are then equivalent (Appendix **??**).

In the present QSD framework we can also show that (**??**) will agree with $k^{*}_{\mathcal{A}\leftarrow\mathcal{B}}$ if the system relaxes to the highly metastable limit of the QSD, $\widehat{\boldsymbol{\pi}}^{QSD}_{\mathcal{B}^{\mathcal{I}}} \to \widehat{\boldsymbol{\pi}}_{\mathcal{B}^{\mathcal{I}}}$. Furthermore, if this limiting distribution $\widehat{\boldsymbol{\pi}}_{\mathcal{B}^{\mathcal{I}}} \simeq \widehat{\boldsymbol{\pi}}_{\mathcal{B}}$, then $k^{F}_{\mathcal{A}\leftarrow\mathcal{B}} \simeq k^{NSS}_{\mathcal{A}\leftarrow\mathcal{B}}$.

To see this correspondence we use our exact waiting time expression (**??**) for an initial condition $[\mathbf{P}_{\mathcal{B}}]_{b'}(0) = \delta_{bb'}$ to write

$$\mathcal{T}_{\mathcal{A}b} = \left[\mathbf{1}_{\mathcal{B}}\left[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]_b. \tag{40}$$

However, we know that the system will decay to the QSD for any initial condition if the basin is metastable. In this limit the first term in the series for equation (**??**) gives

$$\mathcal{T}_{\mathcal{A}b} \to (k^{*}_{\mathcal{A}\leftarrow\mathcal{B}})^{-1}\left[\mathbf{z}^{L}_0\right]_b\left(\mathbf{1}_{\mathcal{B}}\boldsymbol{\pi}^{QSD}_{\mathcal{B}^{\mathcal{I}}}\right). \tag{41}$$

In the highly metastable limit $\boldsymbol{\pi}_{\mathcal{B}^{\mathcal{I}}}^{QSD} \to \boldsymbol{\pi}_{\mathcal{B}^{\mathcal{I}}}$, we know $\mathbf{z}_0^L \to \mathbf{1}_{\mathcal{B}}/(\mathbf{1}_{\mathcal{B}}\boldsymbol{\pi}_{\mathcal{B}^{\mathcal{I}}})$, meaning that

$$\mathcal{T}_{\mathcal{A}b} \to (k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}})^{-1} \Rightarrow k^{F}_{\mathcal{A}\leftarrow\mathcal{B}} \to k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}. \tag{42}$$

In addition, using the exact waiting time expression (**??**) for the full $\mathcal{I}\cup\mathcal{B}$ rate matrix (**??**), analogous manipulations yield $k^{F}_{\mathcal{A}\leftarrow\mathcal{B}} \to k^{*}_{\mathcal{A}\leftarrow\mathcal{B}}$, showing that in the metastable limit, $k^{F}_{\mathcal{A}\leftarrow\mathcal{B}}$, $k^{*}_{\mathcal{A}\leftarrow\mathcal{B}}$ and $k^{*,\mathcal{I}}_{\mathcal{A}\leftarrow\mathcal{B}}$ are all in agreement.

Hence, a well defined transition rate emerges when we allow the initial conditions to relax to the QSD, and all previous results are recovered by first assuming that the QSD becomes the local equilibrium distribution before any additional assumptions. The difference between the QSD and the local equilibrium view arises from the treatment of the region as isolated, or with an absorbing boundary. In both cases the limiting distribution corresponds to the eigenvalue of the rate matrix with the smallest magnitude (zero for local equilibrium). The calculated rates will agree when the region defined as reactants in the experiment is sufficiently metastable, so that a local equilibrium setup can be achieved.

In previous work we considered rates obtained from averaging over kinetic Monte Carlo runs as the exact reference for any given initial distribution, and referred to this rate as $k^{kMC}_{\mathcal{A}\leftarrow\mathcal{B}}$. However, since the exact rate can be calculated in other ways we prefer the notation $k^{*}_{\mathcal{A}\leftarrow\mathcal{B}}$, which is not associated with any particular numerical approach. We finally note that an absorbing boundary condition has previously been used in combination with master equation dynamics to guide the construction of a kinetic transition network by defining boundary states.[?]

V. PATH-BASED GT REGULARIZATION FOR SENSITIVITY

To recover a numerically tractable system upon which sensitivity analysis can be performed we use the GT method to remove states from $\mathcal{I}$ which do not participate or significantly influence a given reaction pathway, leaving a reduced set of states $\mathcal{P}\subset\mathcal{I}$ with renormalised branching probabilities. We thus remove $\mathcal{I}\setminus\mathcal{P}$, i.e. those in $\mathcal{I}$ but not in $\mathcal{A}\cup\mathcal{B}\cup\mathcal{P}$ through the GT method. All branching probabilities and waiting times remain unchanged; the branching probability matrix will reduce in rank, thus typically becoming more amenable to linear algebra manipulations, though it will in general become less sparse. The precise definition of which states (local minima) to retain in $\mathcal{P}$ is flexible; indeed, the definition is free to change arbitrarily throughout the computation. In previous applications that remove states one-by-one, it proved much more efficient to remove $\mathcal{I}$ states with the fewest connections first.[? ?] In the present work, a natural choice is to select all states in $\mathcal{I}$ that either lie on or are directly connected to at least one known reaction pathway. Following the notation convention for $\mathbf{B}^{\mathcal{I}}_{\mathcal{X}\mathcal{Y}}$ the GT procedure detailed in (**??**) yields renormalised branching probability matrices $\mathbf{B}_{\mathcal{A}\mathcal{I}} \to \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{A}\mathcal{P}}$, $\mathbf{B}_{\mathcal{I}\mathcal{B}} \to \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{P}\mathcal{B}}$, $\mathbf{B}_{\mathcal{I}\mathcal{I}} \to \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{P}\mathcal{P}}$, and $\mathbf{B}_{\mathcal{A}\mathcal{B}} \to \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{A}\mathcal{B}}$, where the latter matrix does not change dimension, but the entries are renormalised.

With corresponding renormalisations for the waiting times as in §**??**, defined as the inverse of the total escape rate $\mathbf{D}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{Y}}$, we also obtain renormalised rate matrices through $\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{X}\mathcal{Y}} \equiv \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{X}\mathcal{Y}}\mathbf{D}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{Y}}$. The branching probability matrix from $\mathcal{B}$ to $\mathcal{A}$ can be exactly expressed in an identical form to equation (**??**), namely

$$\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}} = \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{A}\mathcal{B}} + \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{A}\mathcal{P}}\mathbf{G}_{\mathcal{P}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{P}\mathcal{B}}, \tag{43}$$

where $\mathbf{G}_{\mathcal{P}} \equiv \left[\mathbb{I}_{\mathcal{P}} - \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{P}\mathcal{P}}\right]^{-1}$. This formulation naturally defines the solution matrices $\mathbf{X}\in\mathbb{R}^{N_{\mathcal{P}}\times N_{\mathcal{B}}}$ and $\mathbf{Y}\in\mathbb{R}^{N_{\mathcal{A}}\times N_{\mathcal{P}}}$ that solve the renormalised linear equations

$$\left[\mathbb{I}_{\mathcal{P}} - \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{P}\mathcal{P}}\right]\mathbf{X} \equiv \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{P}\mathcal{B}}, \tag{44}$$

to give $\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}} = \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{A}\mathcal{B}} + \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{A}\mathcal{P}}\mathbf{X}$, or equivalently

$$\mathbf{Y}\left[\mathbb{I}_{\mathcal{P}} - \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{P}\mathcal{P}}\right] \equiv \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{A}\mathcal{P}}, \tag{45}$$

to give $\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}} = \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{A}\mathcal{B}} + \mathbf{Y}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{P}\mathcal{B}}$. We present both $\mathbf{X}$ and $\mathbf{Y}$ in anticipation of results below.

Equation (**??**) is the central object of analysis in this contribution, as all other network observables can be calculated in the manner detailed in section **??**. In the next section we derive the general convergence and sensitivity criteria with respect to the introduction of additional transition rates in the network. In particular, we focus on a scalar contraction of (**??**), the total branching probability, using the committor vector defined in equation (**??**):

$$\mathrm{C}^{\mathcal{A}}_{\mathcal{B}} \equiv \mathbf{C}^{\mathcal{A}}_{\mathcal{B}}\mathbf{1}^{\top}_{\mathcal{B}} \equiv \mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}\mathbf{1}^{\top}_{\mathcal{B}}. \tag{46}$$

Crucially, although only states in $\mathcal{A}\cup\mathcal{B}\cup\mathcal{P}$ are explicitly enumerated in (**??**), the sensitivity of $\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}$ to the introduction of additional transitions involving states in $\mathcal{I}\setminus\mathcal{P}$ will still be present in the effective transition rates of the renormalised network. However, kinetically important states should be retained in $\mathcal{P}$ to focus the sensitivity analysis. A detailed analysis of the optimal strategy to determine $\mathcal{P}$ on-the-fly will be the subject of a future contribution. In the following analysis, we assume a suitable $\mathcal{P}$ has been chosen, and derive explicit terms for the sensitivity to the introduction of new transitions between previously unconnected states in $\mathcal{A}\cup\mathcal{B}\cup\mathcal{P}$.

VI. SENSITIVITY AND CONVERGENCE

Additional sampling will usually produce new minima and transition states, changing network properties such

as $\mathbf{B}^{\mathcal{I}}_{\mathcal{AB}}$ and overall rates. A key goal of this paper is to derive expressions for the sensitivity of such quantities to additional sampling, thereby allowing the construction of some measure of convergence. This is a difficult problem as the kinetic transition network is a complex object; rate convergence is in principle a global optimisation problem. Furthermore, any sensitivity measure will necessarily depend on the chosen sampling strategy,[?] which determines the nature of the additional (possibly redundant) information returned by additional sampling tasks.

In this work, we consider use of the `OPTIM`[?] program to perform double-ended transition state searches between candidate state pairs (l, m) via the doubly-nudged elastic band (`DNEB`) method.[?] In this procedure, initial pathways are found by launching `DNEB` searches for 'direct' pairs (l, m), where $l \in \mathcal{A}$ and $m \in \mathcal{B}$, which will in general return indirect pathways with many intervening minima in $\mathcal{I}$. This will then affect the graph transformed KTN for $\mathcal{A} \cup \mathcal{P} \cup \mathcal{B}$.

Our goal is to derive a pathwise local sampling sensitivity measure, once some initial $\mathcal{B} \leftrightarrow \mathcal{A}$ pathways are found, which can both propose target pairs $(l, m) \in \mathcal{A} \cup \mathcal{P} \cup \mathcal{B}$ for additional `DNEB` tasks and estimate the expected change in network observables upon the incorporation of new sampling data. If the expected changes can be bounded, a convergence measure then becomes possible.

In general, there often be additional criteria that can reduce the number of candidate (l, m) pairs for sensitivity and convergence measures, based on e.g. the distance in configuration space, or some other metric such as the change in bonding topology. This does not preclude the possibility of $l \leftrightarrow m$ pathways, only the existence of direct $l \leftrightarrow m$ transitions. Furthermore, in the steady state approximation for intervening minima $\mathcal{I}$, and thus necessarily the renormalised region $\mathcal{P} \subset \mathcal{I}$, $k_{\mathcal{A}\leftarrow\mathcal{B}}$ is unchanged to first order by additional transitions entirely within $\mathcal{A}$ or $\mathcal{B}$, and $\mathbf{B}_{\mathcal{AB}}$ is unchanged by transitions within $\mathcal{A}$. Whilst the discovery of new $\mathcal{A}, \mathcal{B}$ states and their connections to $\mathcal{P}$ could also be considered, we will focus here on transitions corresponding to $\mathcal{B} \to \mathcal{P}$, $\mathcal{P} \to \mathcal{P}$ and $\mathcal{P} \to \mathcal{A}$.

If a `DNEB` search returns an indirect pathway where all the intervening minima are new, the sensitivity expression is equivalent to that for the discovery of a direct transition (with some effective rate). However, in the general case, especially as sampling reaches local convergence, it is more likely that searches will produce indirect pathways involving already discovered minima, modifying a range of branching probabilities. Locality criteria for pair selection will likely reduce the possible number of intervening minima, but in general the central complication to deriving a sensitivity measure remains: a sampling task starting from a given pair (l, m) will in general yield an indirect pathway with multiple intervening minima, producing matrix modifications $\delta\mathbf{K}_{\mathcal{XY}}$ that affect a larger number of states.

Before analysing the exact form of the $\delta\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XY}}$ matrices following a `DNEB` search, we look at the most general expression for the change in $\mathbf{B}^{\mathcal{I}}_{\mathcal{AB}}$ as defined in (**??**), and thus the total branching probability $\mathrm{C}^{\mathcal{A}}_{\mathcal{B}}$ as defined in (**??**), under additional sampling. We employ component form below where ambiguity could arise. Using (**??**) and the identity $[\mathbf{D}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{X}}]_{ij} = \delta_{ij}[\mathbf{1}_{\mathcal{X}}\mathbf{D}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{X}}]_i$, the change in $\mathbf{D}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{X}}$ under a general perturbation is

$$[\delta\mathbf{D}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{X}}]_{ij} = \delta_{ij}\sum_{\mathcal{Y}}\left[\mathbf{1}_{\mathcal{Y}}\delta\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{YX}}\right]_i, \tag{47}$$

giving a change in branching probabilities of

$$\begin{aligned}\left[\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XY}}\right]_{ij} = &\left[\delta\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XY}}\right]_{ij}\left[\mathbf{D}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{Y}}\right]^{-1}_{jj} \\ &- \left[\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XY}}\right]_{ij}\left[\delta\mathbf{D}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{Y}}\right]_{jj}\left[\mathbf{D}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{Y}}\right]^{-2}_{jj}.\end{aligned} \tag{48}$$

To propagate these changes to $\mathbf{B}^{\mathcal{I}}_{\mathcal{AB}}$ defined in (**??**) we need to calculate $\delta\mathbf{G}_{\mathcal{P}}$. For any matrix $\mathbf{M}$, applying the chain rule to $\delta(\mathbf{M}^{-1}\mathbf{M}) = \mathbf{0}$ yields $\delta(\mathbf{M}^{-1}) = -\mathbf{M}^{-1}\,\delta\mathbf{M}\,\mathbf{M}^{-1}$. For the renormalised Green's function $\mathbf{G}_{\mathcal{P}}$, for a given $\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}$ we have

$$\mathbf{G}_{\mathcal{P}} = \left[\mathbb{I}_{\mathcal{P}} - \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}\right]^{-1} \Rightarrow \delta\mathbf{G}_{\mathcal{P}} = \mathbf{G}_{\mathcal{P}}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}\mathbf{G}_{\mathcal{P}}. \tag{49}$$

This gives a total change in $\mathbf{B}^{\mathcal{I}}_{\mathcal{AB}}$ of

$$\begin{aligned}\delta\mathbf{B}^{\mathcal{I}}_{\mathcal{AB}} =&\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AB}} + \delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}\mathbf{G}_{\mathcal{P}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}} + \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}\mathbf{G}_{\mathcal{P}}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}} \\ &+ \mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}\mathbf{G}_{\mathcal{P}}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}\mathbf{G}_{\mathcal{P}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}},\end{aligned} \tag{50}$$

where we revert to matrix products for clarity of presentation. In terms of the solution matrices $\mathbf{X}$ and $\mathbf{Y}$ defined in (**??**) and (**??**), we can write $\delta\mathbf{B}^{\mathcal{I}}_{\mathcal{AB}}$ in the compact form

$$\delta\mathbf{B}^{\mathcal{I}}_{\mathcal{AB}} = \delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AB}} + \delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}\mathbf{X} + \mathbf{Y}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}} + \mathbf{Y}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}\mathbf{X}, \tag{51}$$

which motivates the linear algebra formulation. The total branching probability $\mathrm{C}^{\mathcal{A}}_{\mathcal{B}}$ therefore undergoes a total change

$$\begin{aligned}\delta\mathrm{C}^{\mathcal{A}}_{\mathcal{B}} = &\mathbf{1}_{\mathcal{A}}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AB}}\mathbf{1}^{\top}_{\mathcal{B}} + \mathbf{1}_{\mathcal{A}}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}\mathbf{x} + \mathbf{y}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}}\mathbf{1}^{\top}_{\mathcal{B}} \\ &+ \mathbf{y}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}\mathbf{x}\end{aligned} \tag{52}$$

where $\mathbf{x} = \mathbf{X}\mathbf{1}^{\top}_{\mathcal{B}}$ and $\mathbf{y} = \mathbf{1}_{\mathcal{A}}\mathbf{Y}$ are each obtained with a single linear solve. In the following we focus on the sensitivity to new transitions which affect $\mathcal{A} \leftarrow \mathcal{P} \leftarrow \mathcal{B}$ paths, i.e. those that pass through $\mathcal{P}$, meaning that we assume $\mathbf{1}_{\mathcal{A}}\delta\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AB}}\mathbf{1}^{\top}_{\mathcal{B}} = 0$ in the following expressions for $\delta\mathrm{C}^{\mathcal{A}}_{\mathcal{B}}$. It is straightforward to include this contribution to the general sensitivities (**??**).

A. Sensitivity to any direct or pseudo-direct transition

When a sampling task returns a direct transition between (l, m), the rate matrix modifications $\delta\mathbf{K}_{\mathcal{XY}}$ have

a simple closed form. Furthermore, this closed form is also valid when the sampling task returns an indirect transition where all intermediate states were previously unknown, a case we label 'pseudo-direct'. With a forward ($l \to m$) rate of k^U, the reverse ($m \to l$) rate is defined to be $\phi^{ml}\mathrm{k}^U$. Before any GT renormalisation, detailed balance implies $\phi^{ml} = [\boldsymbol{\pi}_\mathcal{Y}]_l/[\boldsymbol{\pi}_\mathcal{X}]_m$, i.e. the ratio of components of the steady state Boltzmann distribution. The GT renormalisation $\mathcal{I} \to \mathcal{P}$ will in general modify the stationary distribution as the effective rates between states changes. Using the renormalised rate matrices $\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{X}\mathcal{Y}}$, we estimated the change in the steady state distribution by using the Boltzmann distributions $\boldsymbol{\pi}_\mathcal{A}, \boldsymbol{\pi}_\mathcal{B}$ and $\boldsymbol{\pi}_\mathcal{P}$ of the retained states as a preconditioner for an iterative minimization. Whilst this effect could give large changes if only sparsely connected, high energy states are retained during the GT procedure, as we typically retain highly connected, low energy states (those that participate in reaction pathways), the proportional changes in the ϕ^{lm} were extremely small, meaning that we use the Boltzmann distributions to estimate the ϕ^{ml}. The modification to the existing rate matrix is therefore given by

$$\begin{aligned}
[\delta^{ml}_{\mathcal{Y}\mathcal{X}}\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{Y}\mathcal{X}}]_{ij} &= \mathrm{k}^U\delta_{im}\delta_{jl}, \quad l \in \mathcal{X}, m \in \mathcal{Y} \\
[\delta^{ml}_{\mathcal{Y}\mathcal{X}}\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{X}\mathcal{Y}}]_{ij} &= \phi^{ml}\mathrm{k}^U\delta_{il}\delta_{jm}, \quad l \in \mathcal{X}, m \in \mathcal{Y} \\
[\delta^{ml}_{\mathcal{X}\mathcal{X}}\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{X}\mathcal{X}}]_{ij} &= \mathrm{k}^U\delta_{im}\delta_{jl} + \phi^{ml}\mathrm{k}^U\delta_{il}\delta_{jm}, \quad l, m \in \mathcal{X},
\end{aligned} \tag{53}$$

where $\delta^{ml}_{\mathcal{Y}\mathcal{X}}$ is the finite difference operator for $l \in \mathcal{X}, m \in \mathcal{Y}$. In (**??**), k^U is a 'test rate' that represents the expected value of the as yet undiscovered transition rate. When using local saddle point search routines driven by high temperature molecular dynamics[?] it is possible to derive monotonically increasing Bayesian estimators for sampling completeness,[?] which can be used to estimate k^U. However, when minima and saddle points are found using geometry optimisation procedures, such as `DNEB` calculations, sampling completeness estimators are not available. In the numerical examples below we discuss various approaches to determining an appropriate k^U, the simplest being an expected upper bound based on prior knowledge of the system under study, such as $\mathrm{k}^U = 10\,\mathrm{THz}$ for thermally activated processes in metals.

The rate matrix modifications given by (**??**) cause changes $\delta^{ml}\mathbf{B}_{\mathcal{Y}\mathcal{X}}$ in the branching probabilities as detailed in (**??**). Inserting these into (**??**) gives the total propagated change $\delta^{ml}_{\mathcal{Y}\mathcal{X}}\mathrm{C}^\mathcal{A}_\mathcal{B}$ for a single direct transition pathway $(l, m) \in (\mathcal{X}, \mathcal{Y})$. We will derive explicit expressions these direct transitions in the three cases of interest: $(l, m) \in (\mathcal{P}, \mathcal{P})$, $(\mathcal{B}, \mathcal{P})$ and $(\mathcal{P}, \mathcal{A})$.

Whilst the first term in (**??**) accounts for possible changes due to direct connections bypassing $\mathcal{P}$, our focus is on the remaining terms that gauge the effect of changes in the structure of the intervening region of the network. For all valid candidate pairs we can obtain a predicted change in the branching probability through (**??**) and thus rank all pairs as candidates for a double-ended saddle search. However, this approach applies for a direct or pseudo-direct pathway. As a result, before giving explicit expressions for direct transitions we will first consider indirect paths.

B. Sensitivity to an indirect transition

In the general case a sampling task targeting $(l, m) \in \mathcal{X}, \mathcal{Y}$ will produce a pathway through M intervening minima $p_1,\ p_2, \ldots,\ p_M$. Considering a general summation $\delta\mathbf{K} = \sum_n \delta\mathbf{K}_n$, with all superscripts suppressed for brevity, we note that the change in branching probabilities is a linear sum to first order:

$$\delta\mathbf{B} = \sum_n \delta\mathbf{K}_n\mathbf{D}^{-1} - \mathbf{K}\,\mathrm{diag}\left(\mathbf{1}\delta\mathbf{K}_n\right)\mathbf{D}^{-2} = \sum_n \delta\mathbf{B}_n, \tag{54}$$

where we have used $\delta(\mathbf{D}^{-1}) = \mathrm{diag}(\mathbf{1}\delta\mathbf{K})\mathbf{D}^{-2}$ to first order, with $\mathrm{diag}(\mathbf{1}\delta\mathbf{K})$ the diagonal matrix with elements given by the components of the vector $[\mathbf{1}\delta\mathbf{K}]_j = \sum_i \delta k_{ij}$, the total change in escape rate from minimum j. If new intervening minima p_i are found, the dimensions of $\mathbf{B}$, $\mathbf{K}$ and $\mathbf{D}$ must reflect the final dimension of the space after the new path is added to the database. We can therefore decompose the change in total escape rate into contributions from direct transitions. As the change in total branching probability is linear in the $\delta\mathbf{B}_{\mathcal{X}\mathcal{Y}}$, the first order propagated change to network observables for indirect transitions will then simply be the sum of the propagated changes due to the composite direct transitions. Hence, to investigate the effect of indirect transitions it is sufficient to evaluate and rank all of the direct transitions in the manner described above, which is the task of the next section.

C. Sensitivity to direct transitions

Details are collected in Appendix **??**. The final expressions for changes in the scalar committor probability are

$$\begin{aligned}
(l, m) \in (\mathcal{P}, \mathcal{P}) \qquad \delta^{m\leftarrow l}_{\mathcal{P}\mathcal{P}}\mathrm{C}^\mathcal{A}_\mathcal{B} &= \mathrm{k}^U\left([\mathbf{y}]_m - [\mathbf{y}]_l\right)[\mathbf{D}^{-1}_\mathcal{P}\mathbf{x}]_l \\
(l, m) \in (\mathcal{B}, \mathcal{P}) \qquad \delta^{ml}_{\mathcal{A}\mathcal{P}}\mathrm{C}^\mathcal{A}_\mathcal{B} &= \mathrm{k}^U\left([\mathbf{y}]_m - [\mathbf{y}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{P}\mathcal{B}}]_l\right)[\mathbf{D}^{-1}_\mathcal{B}]_{ll} \\
&\quad - \mathrm{k}^U\phi^{ml}[\mathbf{y}]_m[\mathbf{D}^{-1}_\mathcal{P}\mathbf{x}]_m \\
(l, m) \in (\mathcal{P}, \mathcal{A}) \qquad \delta^{ml}_{\mathcal{A}\mathcal{P}}\mathrm{C}^\mathcal{A}_\mathcal{B} &= \mathrm{k}^U\left(1 - [\mathbf{y}]_l\right)[\mathbf{D}^{-1}_\mathcal{P}\mathbf{x}]_l,
\end{aligned} \tag{55}$$

where $\delta^{m\leftarrow l}_{\mathcal{P}\mathcal{P}}$ is a one-sided difference operator, which includes only the changes due to the $l \to m$ path. As expected, self-transitions $l = m$ in $\mathcal{P}$, which are permitted after renormalisation, do not affect the committor probability. The result for $\delta^{ml}_{\mathcal{A}\mathcal{P}}\mathrm{C}^\mathcal{A}_\mathcal{B}$ is independent of $m \in \mathcal{A}$ because of the sum over $\mathcal{A}$ states in the definition of the committor vector in (**??**).

The convergence of these changes in the committor is investigated below for two benchmark systems involving

atomic clusters of N atoms bound by the Lennard-Jones potential, LJ_N, using the regularization techniques developed above. We have tested the sensitivity expressions by removing known rates and looking at the predicted and actual change in branching probabilities. In all these tests excellent agreement was found, demonstrating the self-consistency of our approach.

VII. SELECTION, DISTRIBUTION AND ALLOCATION OF SAMPLING TASKS

When an accurate sensitivity can be evaluated for all candidate direct transitions, their influence on network properties can be ranked according to the change in the scalar committor value $\delta^{ml}\mathrm{C}_{\mathcal{B}}^{\mathcal{A}}$. With a given computational budget, `DNEB` calculations can then be assigned according to this ranking. However, this approach may not ensure optimal computational efficiency; it is likely that the number of `DNEB` images needed to represent a pathway will scale approximately linearly with the distance between candidate minima, increasing the computational cost. A possible strategy is to weight the ranking according to the number N_{im} of `DNEB` images that would be assigned for the task, and consider $\delta^{ml}\mathrm{C}_{\mathcal{B}}^{\mathcal{A}}/N_{im}$, to reflect the fact that multiple less sensitive transitions may yield better computational return than a single sensitive transition between distant minima.

As a `DNEB` calculation is a formally deterministic process, once a given pair (l, m) has been targeted for sampling it should be removed from all future selections, i.e. set $\delta^{ml}\mathrm{C}_{\mathcal{B}}^{\mathcal{A}} = 0$. However, we note that many strategies could yield multiple pathways on repeated `DNEB` searches between the same two minima. For example, multiple initial pathways could be attempted or stochastic forces or energetic penalty functions could be applied during the `DNEB` minimization. In this case, multiple `DNEB` requests could return different results and therefore the $\delta^{ml}\mathrm{C}_{\mathcal{B}}^{\mathcal{A}}$ should be reevaluated each sampling cycle, or multiple `DNEB` requests could be assigned to a given pair in one cycle.

VIII. CONSTRUCTION OF CONFIDENCE INTERVALS FOR $\mathrm{C}_{\mathcal{B}}^{\mathcal{A}}$

The local sensitivity analysis detailed here clearly cannot solve the global problem of whether there is some other distant set of unexplored pathways that will drastically change the $\mathcal{A} \leftrightarrow \mathcal{B}$ kinetics. Rather, we aim to make some statement on the convergence of the transition rate associated with the current database of sampled pathways.

The sensitivity metric works well if the to-be-discovered transition is either a direct connection between two known states, or an indirect connection involving only newly discovered states; in either case, by postulating an upper bound k^U on the to-be-discovered transition rate between (l, m) an upper bound on the absolute change in the branching probabilities is obtained, to first order. We use the absolute change to emphasise that this does *not* mean additional connections necessarily increase $\mathrm{C}_{\mathcal{B}}^{\mathcal{A}}$.

Our convergence measure for the path finding search procedure involves upper and lower confidence intervals $\sigma_{\pm}$ for the branching probability $\mathrm{C}_{\mathcal{B}}^{\mathcal{A}}$. This sensitivity machinery was combined with two estimators of the network structure, namely the expected value of newly discovered transition rates, k^U above, and the connection sparsity ξ. Rigorous, monotonic estimators of unseen rates have been developed in dynamic sampling strategies,[?] where the dynamical trajectories provide a well defined probability law for discovering transitions. In the present context, where a database of stationary points is harvested using geometry optimisation, no such law exists, meaning that the rate estimator $\widehat{k^U}$ (defined below) will be nonmonotonic, with uncontrolled fluctuations upon the incorporation of new sampling data. A sensible strategy in this scenario is to consider multiple estimators in parallel, using the collective information to guide decisions on convergence. In this final section we test some preliminary estimators, demonstrate the sensitivity framework that is a main object of this paper; future work will concentrate on the optimal form of estimation and thus how to deduce more rigorous convergence bounds. The present contribution gives the computational framework upon which such convergence concepts can be tested.

Our estimator $\widehat{k^U}$ has one hyperparameter, a postulated maximum unknown rate k^U_{max}, which we set to $k^U_{max} = \omega_0 \exp(-3)$, where $\omega_0 = 5$ in reduced units for the LJ potential. This value corresponds to an energy barrier of $3/\beta$, around the limit of the rare event regime. Future work will investigate the dependence on the final sensitivities to estimates of k^U_{max}, and when it is beneficial to spend more effort in the estimation. In addition, we measure the logarithmic mean and variance of the observed rates, correcting for differences in the free energy of initial and final states through

$$\widetilde{k}_{ij} \equiv \frac{1}{\beta h} \exp\left[-\beta({F^{\dagger}}_{ij} - \max(F_i, F_j))\right], \tag{56}$$

$$\langle (\ln \widetilde{k})^n \rangle = \sum_{\text{rates } ij} (\ln \widetilde{k}_{ij})^n / N_{rates} \tag{57}$$

where ${F^{\dagger}}_{ij} = {F^{\dagger}}_{ji}$ is the saddle point free energy, $\beta = 1/\mathrm{k_B T}$, with $\mathrm{k_B}$ the Boltzmann constant and T the temperature, meaning that ${F^{\dagger}}_{ij} - \max(F_i, F_j)$ is the lower free energy barrier for the ij transition. We consider the first and second moments, corresponding to $n = 1$ and 2 in constructing the rate estimator below.

The logarithmic mean $\langle \ln \widetilde{k} \rangle$ was chosen to account for the wide range in observed rates; it can be considered as estimating the average free energy barrier. Nevertheless, only if the distribution is suitably well peaked should this mean value be taken as informative. As a result, our

preliminary estimator $\widehat{k^U}$ for the newly discovered rates reads

$$\widehat{k^U} = k^U_{max} + \left(e^{\langle \ln \widetilde{k}\rangle} - k^U_{max}\right) \exp\left(1 - \langle (\ln \widetilde{k})^2\rangle / \langle \ln \widetilde{k}\rangle^2\right), \tag{58}$$

meaning that a large geometric variance in the observed rates suppresses the influence of the geometric mean observed rate $e^{\langle \ln \widetilde{k}\rangle}$.

The sensitivity analysis assigns an expected change $\delta^{ml}_{\mathcal{XY}} \mathrm{C}^{\mathcal{A}}_{\mathcal{B}}$ to the branching probabilities upon the discovery of a new transition for every possible transition $m \leftrightarrow l$ in the network, where $m \in \mathcal{X}, l \in \mathcal{Y}$ apart from already sampled pairs, where we set the sensitivity to zero.

Our first estimator bounds $\sigma^{\mathrm{tot}}_{\pm}$ are therefore simply the sum of all possible positive or negative changes:

$$\sigma^{\mathrm{tot}}_{+} \equiv \sum_{\delta \mathrm{C}^{\mathcal{A}}_{\mathcal{B}} > 0} \delta^{ml}_{\mathcal{XY}} \mathrm{C}^{\mathcal{A}}_{\mathcal{B}}, \quad \sigma^{\mathrm{tot}}_{-} \equiv \sum_{\delta \mathrm{C}^{\mathcal{A}}_{\mathcal{B}} < 0} |\delta^{ml}_{\mathcal{XY}} \mathrm{C}^{\mathcal{A}}_{\mathcal{B}}|, \tag{59}$$

giving the projected change $\sigma^{\mathrm{tot}} \equiv \sigma^{\mathrm{tot}}_{+} - \sigma^{\mathrm{tot}}_{-}$. We can also define the maximal bounds $\sigma^{1}_{\pm}$ due to the discovery of a single transition as

$$\sigma^{1}_{+} \equiv \max \delta^{ml}_{\mathcal{XY}} \mathrm{C}^{\mathcal{A}}_{\mathcal{B}}, \quad \sigma^{1}_{-} \equiv |\min \delta^{ml}_{\mathcal{XY}} \mathrm{C}^{\mathcal{A}}_{\mathcal{B}}|, \tag{60}$$

giving the total single change $\sigma^{1} \equiv \sigma^{1}_{+} - \sigma^{1}_{-}$.

However, real transition networks are typically very sparse, meaning the total bounds $\sigma^{\mathrm{tot}}_{\pm}$ could be very large as many possible transitions do not exist. We therefore also estimate the network sparsity ξ, which gives the approximate probability that a search should return a successful connection. We estimate ξ with

$$\widehat{\xi} = \frac{N_{trans}}{N^2_{states}} + \left(\xi_0 - \frac{N_{trans}}{N^2_{states}}\right) \exp\left(-\frac{N_{\texttt{DNEB}}}{N_{thresh}}\right), \tag{61}$$

the number of found transition states, minima, and `DNEB` searches, respectively, and N_{thresh} is a hyperparameter controlling the influence of empirical data, here set to 5000.

The first use of ξ is simply to multiply the total bounds $\sigma^{\mathrm{tot}}_{\pm}$, giving the reduced bounds $\xi\sigma^{\mathrm{tot}}_{\pm}$ and projected change $\xi\sigma^{\mathrm{tot}}$. We can also use the sparsity to estimate the number of further connections. Assuming independence, the probability of finding m connections is simply $\xi^m(1-\xi)$, giving an expected number of connections $\langle m\rangle = \xi/(1-\xi)$. We combine this expression with the average changes to give our last investigated bounds $\sigma^{\xi}_{\pm}$, defined as

$$\sigma^{\xi}_{+} \equiv \left(\frac{\xi}{1-\xi}\right) \frac{\sum_{\delta \mathrm{C}^{\mathcal{A}}_{\mathcal{B}} > 0} \delta^{ml}_{\mathcal{XY}} \mathrm{C}^{\mathcal{A}}_{\mathcal{B}}}{\sum_{\delta \mathrm{C}^{\mathcal{A}}_{\mathcal{B}} > 0} 1}, \tag{62}$$

and similarly for σ^{ξ}_{-} with the constraint $\delta \mathrm{C}^{\mathcal{A}}_{\mathcal{B}} < 0$. Finally, we note that only using the first order derivative to predict changes in a nonlinear quantity is clearly inaccurate when the predicted changes in the argument, here $\widehat{\mathrm{k}}^U$, are very large. However, higher derivatives would require multiple linear solves, which imposes a prohibitive computational cost.

IX. VALIDATION TESTS ON LJ_{13} AND LJ_{38}

To test the sampling protocol described in this work, an 'exactly known' rate matrix was constructed using the LJ_{13} and LJ_{38} energy landscapes from the Cambridge Landscape Database.[?] As is well known, LJ_{13} has a single funnel landscape, whilst LJ_{38} has a double funnel landscape[? ? ? ?] with two competing morphologies, corresponding to two free energy minima at low temperatures.

As the LJ_{13} cluster has a single funnel landscape, the definition of $\mathcal{A}$ and $\mathcal{B}$ is somewhat arbitrary. The $\mathcal{B}$ basin was taken as the global potential energy minimum plus one connected state, and $\mathcal{A}$ was taken as the highest minimum plus one connected state. Alternative choices were also considered. With only one state in each basin the kinetics are dominated by a single pathway and converged to within numerical accuracy in just two iterations. For larger basins the direct transition dominated, again giving extremely rapid convergence. With two states in each basin a more useful test system was obtained, so we focus on these results.

The LJ_{13} landscape is well conditioned even for $\beta = 10$ in Lennard-Jones units, meaning that we can set $\mathcal{P} = \mathcal{I}$, i.e. with no renormalisation. To simulate `DNEB` connection attempts for a pair (i, j), Dijkstra's shortest path algorithm[?] as implemented in `scipy`[?] was applied to an unweighted, undirected graph created from the reference rate matrix. The use of an unweighted graph is to simulate the fact that the `DNEB` algorithm has an energy penalty for the length of the path, and is unlikely to find the fastest path between two distant states in one iteration.

The 'sampled' rate matrix initially contained all $\mathcal{A} \to \mathcal{A}$ and $\mathcal{B} \to \mathcal{B}$ connections, a single known $\mathcal{B} \to \mathcal{A}$ path and, for each state on the path, the result of a simulated single ended saddle search, which returned at most four connecting states. To create an initial $\mathcal{B} \to \mathcal{A}$ path we used the simulated `DNEB` routine described above; the initial sampled set contained around 20 states.

To give a high data resolution in figure **??** the sensitivity analysis was used to identify, each cycle, the two most sensitive state pairs that had not been previously sampled. If these simulated searches returned no new results, the next two most sensitive pairs were considered, and so on, until at least one new transition state was found. In practice, as discussed above, it is likely beneficial to sample many more pairs simultaneously; a detailed investigation of the optimal deployment will be the subject of future work.

At the top of figure **??** we present the results from this test procedure with the convergence metrics described above. The metrics exhibit a large number of positive 'spikes' in the committor probability, due to the possibility of a high transition rate to $\mathcal{A}$. Such large sensitivities will invariably be sampled in the next cycle, whereupon if they are found to not exist the prediction returns to essentially the same value as before. To regularise the out-

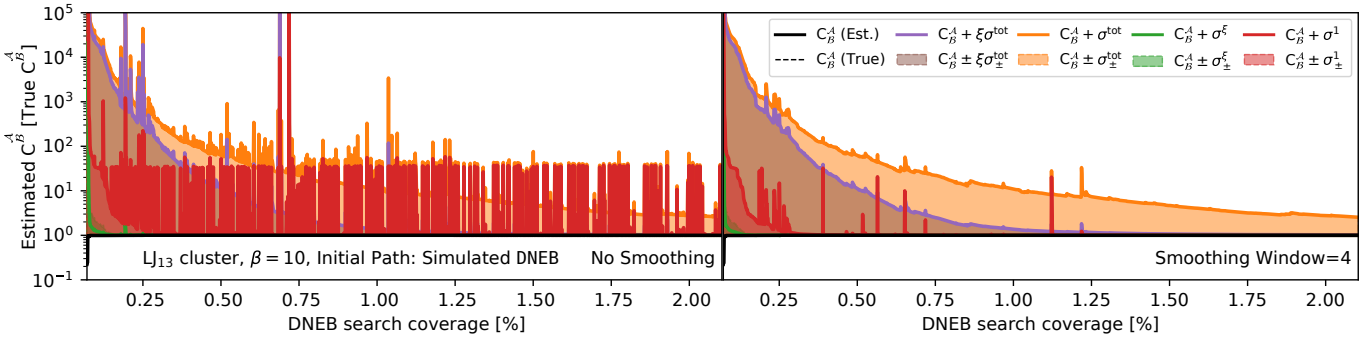

FIG. 1. Convergence tests on LJ_{13} at $\beta = 10$, where $\mathcal{B}$ is the global potential energy minimum plus a single connected state, and $\mathcal{A}$ is the highest minimum plus a single connected state. The initial pathway was chosen to be that with the fewest jumps, simulating a `DNEB` search. The `DNEB` search coverage is approximated as the ratio of attempted `DNEB` searches to total possible connections, which is fairly high in these toy examples as no attempt to filter admissible transitions was made.

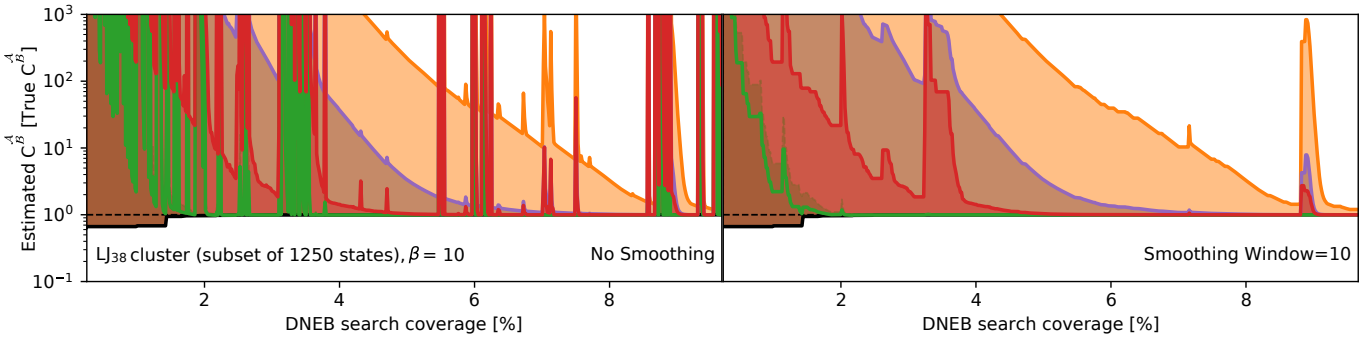

FIG. 2. Convergence test on a subset of the interbasin committor probability for the LJ_{38} landscape, with two competing morphologies, without any renormalisation, leaving a subset of 1250 states. $C^{\mathcal{A}}_{\mathcal{B}}$ is defined between the lowest minima in each funnel, with $\beta = 10$.

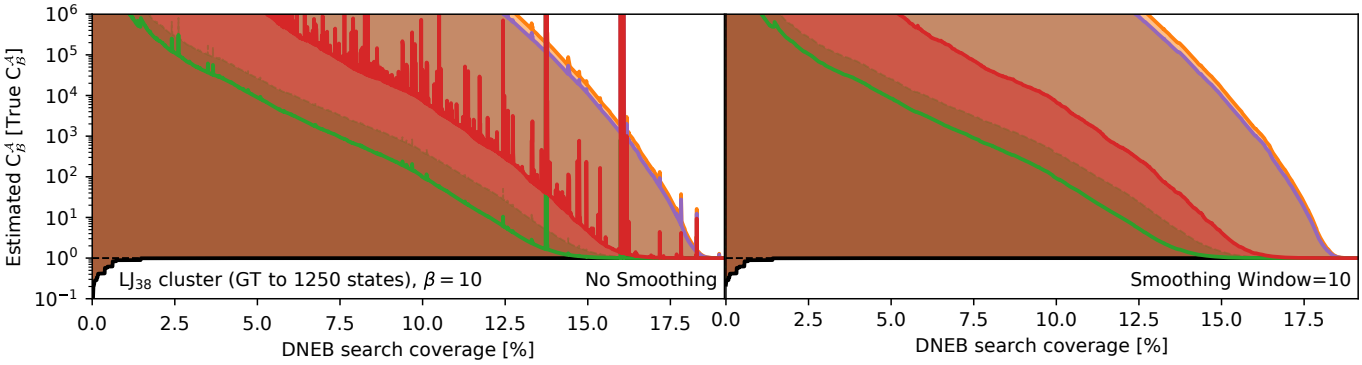

FIG. 3. Convergence test on a subset of the interbasin committor probability for LJ_{38} landscape, using the graph transformation method to yields a renormalised system with 1250 states, but a much higher connection density. As before, $C^{\mathcal{A}}_{\mathcal{B}}$ is defined between the lowest minima in each funnel, with $\beta = 10$.

put against these isolated fluctuations we applied a short window filter, typically the minimum absolute bound values $|\sigma^{\mathrm{tot},1,\phi}_{\pm}|$ over a small number of cycles. This smoothing is reasonable because the spikes occur much less frequently as more data is accumulated, the averaging window is very short, and the predictions immediately before and after a spike are largely unchanged.

As shown in the figures, this smoothing significantly enhances the interpretability of the sampling data, with the majority of spikes filtered out. Sampling the most sensitive connections quickly yields the correct branching probability, but much more sampling is required to be confident of convergence; in particular the total sensitivity $\sigma^{\mathrm{tot}}_{+}$ is very slow to converge, but the sparsity estimator $\hat{\xi}$ implies convergence is reached earlier. The individual sensitivities converge much more quickly, as

expected. After more sampling cycles the confidence intervals contain all changes in the estimates of $C_{\mathcal{B}}^{\mathcal{A}}$, and the projected values are also stable.

The two funnel landscape of the LJ_{38} cluster gives natural definitions for the $\mathcal{A}$ and $\mathcal{B}$ basins. The $\mathcal{B}$ region was chosen to be the global minimum truncated octahedron, whilst the $\mathcal{A}$ region was chosen to be the lowest minimum in the icosahedral funnel. The minimum free energy path for $\beta = 10$ (again in LJ units) has 28 intermediate states with a large effective barrier height, representing a realistic test of the sensitivity framework developed here for a rare event. The sampling was initially restricted to a artificially truncated subnetwork of 1250 minima that contained the minimum energy pathway, meaning that again $\mathcal{P} = \mathcal{I}$. We note that the interconversion rates for the different morphologies are usually summed over sets of local minima in each funnel.[? ? ? ?] However, this step is not important for the present benchmarking, where we can consider the convergence of any chosen rate constant as a function of the database. As for the LJ_{13} example, a useful convergence measure can be obtained when applying a small smoothing window.

Importantly, the single transition bounds $\sigma_{\pm}^{1,\phi}$ converge much more quickly than the total expectation bounds $\sigma_{\pm}^{\mathrm{tot}}$, indicating that the existing landscape requires a new pathway to significantly change the branching probability. This result demonstrates how multiple convergence metrics can help inform measures of convergence; future work will incorporate these ideas within a Bayesian framework.

Finally, we also applied the same methodology to the full LJ_{38} system, using the graph transformation method to yields $\mathcal{P} \subset \mathcal{I}$, with 1250 states in total. Whilst the effective rank is the same as the truncated example, the resulting KTN has a much higher density of connections due to the greater number of pathways, meaning that convergence is much slower. In all of these toy examples we have made no attempt to filter the possibility of a connection existing between a given pair of minima based on knowledge of the existing KTN nor structural or energetic properties of the minima pair under consideration, meaning that our metrics consider all possible interstate connections. Optimal strategies for applying the GT renormalisation and estimating the probability that a given minima pair will yields a new connection will be the subject of another contribution.

X. CONCLUSIONS

In this contribution we have developed a linear algebra formulation for calculating waiting times and rates corresponding to a kinetic transition network, to develop a tractable scheme for judging the convergence with respect to sampling. We have first provided expressions for the observables within a hierarchy of approximations, starting with a steady state assumption for intervening minima, and local equilibrium for the reactants. We establish the equivalence of the resulting matrix/vector representations and formulae previously derived by considering sums over pathways directly. The extension to exact rates and waiting times further enables us to connect results based on a local equilibrium in the reactant space to the quasi-stationary distribution, which corresponds to the limiting case for an absorbing boundary at the products. The linear algebra results are fully consistent with previous results, and provide additional insight, as well as a more efficient way to compute some of the properties of interest. In particular, we derive formulae to estimate the sensitivity of branching matrices, committor probabilities, and hence rates when new network connections are hypothesised. These sensitivity measures can be used to direct the sampling strategy to converge the database if stationary points with respect to rates.

To test convergence we have applied the sensitivity calculations to existing databases for two atomic clusters bound by the Lennard-Jones potential, namely LJ_{13} with a single funnel landscape corresponding to efficient relaxation to the global minimum, and LJ_{38}, where two morphologies compete to give a double funnel landscape that causes broken ergodicity, and structural interconversion constitutes a rare event. Starting from small subsets of each database, we use the sensitivity indices to propose new searches for connections between local minima, and simulate their discovery using the known connectivity of the full databases. In each case the sensitivities and bounds on the committor probability initially exhibit large fluctuations, which decrease as the sampling progresses. The bounds converge rapidly to the correct values, demonstrating that this framework should provide a powerful tool for constructing and converging kinetic transition networks in new systems. This approach is generally applicable throughout molecular and condensed matter science, and we envisage future applications to problems ranging from atomic and molecular clusters, through biophysics, to condensed matter.

XI. ACKNOWLEDGEMENTS

DJW gratefully acknowledges financial support from the Engineering and Physical Sciences Research Council. TDS and DJW would like to thank the Institute of Pure and Applied Mathematics at the University of California, Los Angeles for providing a stimulating environment during their long program on "Complex Processes in High Dimensional Energy Landscapes".

Appendix A: Glossary of Useful Formulae

Here we summarise various results that are useful for the derivations. Note that the identities for $\mathbf{B}$ and $\mathbf{G}$ matrices do not hold for the corresponding $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{G}}$ versions.

Compound branching probabilities and renormalised waiting time:

$$\begin{aligned}
\mathbf{B}_{\mathcal{A}\leftarrow\mathcal{B}} &= \mathbf{B}_{\mathcal{A}\mathcal{B}} + \mathbf{B}_{\mathcal{A}\mathcal{I}}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{B}} \equiv \mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}},\\
\mathbf{B}_{\mathcal{B}\leftarrow\mathcal{B}} &= \mathbf{B}_{\mathcal{B}\mathcal{B}} + \mathbf{B}_{\mathcal{B}\mathcal{I}}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{B}} \equiv \mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}},\\
\mathbf{G}_{\mathcal{B}\leftarrow\mathcal{B}} &= (\mathbb{I}_{\mathcal{B}} - \mathbf{B}_{\mathcal{B}\leftarrow\mathcal{B}})^{-1} = (\mathbb{I}_{\mathcal{B}} - (\mathbf{B}_{\mathcal{B}\mathcal{B}} + \mathbf{B}_{\mathcal{B}\mathcal{I}}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{B}}))^{-1} \equiv \mathbf{G}^{\mathcal{I}}_{\mathcal{B}} \qquad \text{(A1)}\\
\mathbf{1}_{\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{B}} + \mathbf{1}_{\mathcal{B}}\mathbf{D}_{\mathcal{B}}^{-1} &\equiv \mathbf{1}_{\mathcal{B}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]^{-1}.
\end{aligned}$$

Here $\mathbf{B}_{\mathcal{A}\leftarrow\mathcal{B}}$ is the probability corresponding to all possible paths that leave $\mathcal{B}$ and reach $\mathcal{A}$ via any number of steps in $\mathcal{I}$ without returning to $\mathcal{B}$; $\mathbf{B}_{\mathcal{B}\leftarrow\mathcal{B}}$ is the probability corresponding to all possible paths that leave $\mathcal{B}$ and return to $\mathcal{B}$ via any number of steps in $\mathcal{I}$ without reaching $\mathcal{A}$. The $\mathbf{G}$ matrices sum over all paths consisting of any number of steps defined by simple or compound branching matrices. Hence $\mathbf{G}_{\mathcal{B}\leftarrow\mathcal{B}}$ is the sum of probabilities for all non-reactive paths. The quantities identified with an $\mathcal{I}$ superscript also correspond to the values used in graph transformation renormalisation after all the intervening minima in the $\mathcal{I}$ region are removed.

Identities involving $\mathbf{G}$ matrices:

$$\begin{aligned}
&\mathbf{G}_{\mathcal{X}} = (\mathbb{I}_{\mathcal{X}} - \mathbf{B}_{\mathcal{X}\mathcal{X}})^{-1}, \quad \mathbf{G}_{\mathcal{X}} = \mathbb{I}_{\mathcal{X}} + \mathbf{G}_{\mathcal{X}}\mathbf{B}_{\mathcal{X}\mathcal{X}}, \quad \mathbf{G}_{\mathcal{X}} = \mathbb{I}_{\mathcal{X}} + \mathbf{B}_{\mathcal{X}\mathcal{X}}\mathbf{G}_{\mathcal{X}}, \quad \mathbf{B}_{\mathcal{X}\mathcal{X}}\mathbf{G}_{\mathcal{X}} = \mathbf{G}_{\mathcal{X}}\mathbf{B}_{\mathcal{X}\mathcal{X}},\\
&\partial_\zeta\left(\widetilde{\mathbf{G}}_{\mathcal{X}}\right)\Big|_{\zeta=0} = \mathbf{G}_{\mathcal{X}}\partial_\zeta\left(\widetilde{\mathbf{B}}_{\mathcal{X}\mathcal{X}}\right)\Big|_{\zeta=0}\mathbf{G}_{\mathcal{X}}. \qquad \text{(A2)}
\end{aligned}$$

$\mathbf{G}$ matrix derivative for simple branching matrices:

$$\partial_\zeta\left(\widetilde{\mathbf{G}}_{\mathcal{X}}\right)\Big|_{\zeta=0} = \mathbf{G}_{\mathcal{X}}\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{G}_{\mathcal{X}} - \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{G}_{\mathcal{X}} = (\mathbf{G}_{\mathcal{X}} - \mathbb{I}_{\mathcal{X}})\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{G}_{\mathcal{X}} = \mathbf{G}_{\mathcal{X}}\mathbf{B}_{\mathcal{X}\mathcal{X}}\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{G}_{\mathcal{X}}. \qquad \text{(A3)}$$

Committor probabilities:

$$\mathbf{C}^{\mathcal{A}}_{\mathcal{B}} = \mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}} = \mathbf{1}_{\mathcal{B}}\left[\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1} \quad \text{so} \quad C^a_b = \left[\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}\right]_{ab}, \quad C^A_b = \left[\mathbf{1}_{\mathcal{B}}\left[\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]^{-1}\right]_b, \quad \text{and} \quad C^{b'}_b = \left[\mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}\right]_{b'b} \qquad \text{(A4)}$$

for minima $b, b' \in \mathcal{B}$ and $a \in \mathcal{A}$.

Identities related to conservation of probability:

$$\begin{aligned}
\mathbf{1}_{\mathcal{X}} = \sum_{\mathcal{Y}\in A,B,I} \mathbf{1}_{\mathcal{Y}}\mathbf{B}_{\mathcal{Y}X} \Rightarrow \sum_{\mathcal{Y}\neq\mathcal{X}} \mathbf{1}_{\mathcal{Y}}\mathbf{B}_{\mathcal{Y}\mathcal{X}}\mathbf{G}_{\mathcal{X}} &= \mathbf{1}_{\mathcal{X}},\\
\mathbf{1}_{\mathcal{A}}\mathbf{B}_{\mathcal{A}\leftarrow\mathcal{B}} + \mathbf{1}_{\mathcal{B}}\mathbf{B}_{\mathcal{B}\leftarrow\mathcal{B}} &= \mathbf{1}_{\mathcal{B}},\\
\mathbf{1}_{\mathcal{A}}\mathbf{B}_{\mathcal{A}\leftarrow\mathcal{B}}\mathbf{G}_{\mathcal{B}\leftarrow\mathcal{B}} &= \mathbf{1}_{\mathcal{B}}\\
\mathbf{1}_{\mathcal{A}}\mathbf{B}_{\mathcal{A}\leftarrow\mathcal{B}}\mathbf{G}_{\mathcal{B}} &= \mathbf{1}_{\mathcal{B}}\mathbf{G}^{-1}_{\mathcal{B}\leftarrow\mathcal{B}}\mathbf{G}_{\mathcal{B}} \qquad \text{(A5)}
\end{aligned}$$

Appendix B: Derivation of equation (??)

The expected waiting time for a transition from any state $z \in \mathcal{Z}$ to any another state in $\mathcal{Z}$ via an arbitrary number of steps between states in $\mathcal{I}$, as presented in (**??**), is given by the z component of

$$\begin{aligned}
&\mathbf{1}_{\mathcal{Z}}\frac{\partial}{\partial\zeta}\left(\widetilde{\mathbf{B}}^{\mathcal{I}}_{\mathcal{Z}\mathcal{Z}}\right)\Big|_{\zeta=0} = \mathbf{1}_{\mathcal{Z}}\frac{\partial}{\partial\zeta}\left(\widetilde{\mathbf{B}}_{\mathcal{Z}\mathcal{Z}} + \widetilde{\mathbf{B}}_{\mathcal{Z}\mathcal{I}}\widetilde{\mathbf{G}}_{\mathcal{I}}\widetilde{\mathbf{B}}_{\mathcal{I}\mathcal{Z}}\right)\Big|_{\zeta=0}\\
=&\mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{Z}}\mathbf{D}_{\mathcal{Z}}^{-1} + \mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}} + \mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}}\frac{\partial}{\partial\zeta}\left(\widetilde{\mathbf{G}}_{\mathcal{I}}\right)\Big|_{\zeta=0}\mathbf{B}_{\mathcal{I}\mathcal{Z}} + \mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}}\mathbf{D}_{\mathcal{Z}}^{-1}\\
=&\mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{Z}}\mathbf{D}_{\mathcal{Z}}^{-1} + \mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}} + \mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}} + \mathbf{1}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}}\mathbf{D}_{\mathcal{Z}}^{-1}\\
=&(\mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{Z}}\mathbf{D}_{\mathcal{Z}}^{-1} + \mathbf{1}_{\mathcal{I}}\mathbf{B}_{IY})\mathbf{D}_{\mathcal{Z}}^{-1} + \mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}} + \mathbf{1}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}}\\
=&\mathbf{1}_{\mathcal{Z}}\mathbf{D}_{\mathcal{Z}}^{-1} + (\mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{I}} + \mathbf{1}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{I}})\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}} = \mathbf{1}_{\mathcal{Z}}\mathbf{D}_{\mathcal{Z}}^{-1} + \mathbf{1}_{\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{Z}} \equiv \mathbf{1}_{\mathcal{Z}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{Z}}]^{-1}. \qquad \text{(B1)}
\end{aligned}$$

The sum of path weights out of every component of $\mathcal{Z}$ is unity because $\mathbf{1}_{\mathcal{Z}}\mathbf{B}_{\mathcal{Z}\mathcal{Z}} = \mathbf{1}_{\mathcal{Z}}$, so we obtain the average escape times from the above construction.

The derivative $\widetilde{\mathbf{G}}_{\mathcal{I}}$, used above, can be obtained by differentiating the series form $\sum_{n=0}^{\infty}\left(\widetilde{\mathbf{B}}_{\mathcal{I}\mathcal{I}}\right)^n$, or using

$$\frac{\partial}{\partial\zeta}\widetilde{\mathbf{G}}_{\mathcal{I}}\Big|_{\zeta=0}\mathbf{G}_{\mathcal{I}}^{-1} = -\mathbf{G}_{\mathcal{I}}\frac{\partial}{\partial\zeta}\widetilde{\mathbf{G}}_{\mathcal{I}}^{-1}\Big|_{\zeta=0}, \tag{B2}$$

and $\widetilde{\mathbf{G}}_{\mathcal{I}}^{-1} = \mathbb{I}_{\mathcal{I}} - \widetilde{\mathbf{B}}_{\mathcal{I}\mathcal{I}}$, to obtain $\frac{\partial}{\partial\zeta}\widetilde{\mathbf{G}}_{\mathcal{I}}|_{\zeta=0} = \mathbf{G}_{\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}} - \mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}} = (\mathbf{G}_{\mathcal{I}} - \mathbb{I}_{\mathcal{I}})\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}} = \mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{G}_{\mathcal{I}}$ (see Appendix **??** for a summary of useful relations between these quantities).

Appendix C: Derivation of matrix expression for exact waiting time

Writing equation (**??**) in the more compact form $\dot{\mathbf{P}}_{\mathcal{I}\cup\mathcal{B}} = \mathbf{M}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}$ and a row vector of ones of dimension corresponding to $\mathcal{I}\cup\mathcal{B}$ as $\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}$ we can show

$$\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{M} = -\mathbf{1}_{\mathcal{A}}\left(\mathbf{K}_{\mathcal{A}\mathcal{I}}, \mathbf{K}_{\mathcal{A}\mathcal{B}}\right), \tag{C1}$$

and hence

$$P(\tau \in [t, t+\mathrm{d}t]) = -\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\dot{\mathbf{P}}_{\mathcal{I}\cup\mathcal{B}}\,\mathrm{d}t = -\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{M}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}\,\mathrm{d}t = \mathbf{1}_{\mathcal{A}}\left(\mathbf{K}_{\mathcal{A}\mathcal{I}}, \mathbf{K}_{\mathcal{A}\mathcal{B}}\right)\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}\,\mathrm{d}t. \tag{C2}$$

Integrating the master equation (**??**) formally to give $\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(t) = \exp\left(\mathbf{M}t\right)\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0)$ and performing the integrals in equation (**??**) gives

$$\langle\tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle = -\frac{\mathbf{1}_{\mathcal{A}}\left(\mathbf{K}_{\mathcal{A}\mathcal{I}}, \mathbf{K}_{\mathcal{A}\mathcal{B}}\right)\mathbf{M}^{-2}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0)}{\mathbf{1}_{\mathcal{A}}\left(\mathbf{K}_{\mathcal{A}\mathcal{I}}, \mathbf{K}_{\mathcal{A}\mathcal{B}}\right)\mathbf{M}^{-1}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0)}. \tag{C3}$$

Using equation (**??**) we can simplify the above expression, since

$$\begin{aligned} \mathbf{1}_{\mathcal{I}\cup\mathcal{B}} &= -\mathbf{1}_{\mathcal{A}}\left(\mathbf{K}_{\mathcal{A}\mathcal{I}}, \mathbf{K}_{\mathcal{A}\mathcal{B}}\right)\mathbf{M}^{-1} \\ \text{we have} \qquad \mathbf{1}_{\mathcal{A}}\left(\mathbf{K}_{\mathcal{A}\mathcal{I}}, \mathbf{K}_{\mathcal{A}\mathcal{B}}\right)\mathbf{M}^{-1}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0) &= -\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0) = -1 \end{aligned} \tag{C4}$$

for the given initial conditions. Hence

$$\begin{aligned} \langle\tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle &= -\mathbf{1}_{\mathcal{I}\cup\mathcal{B}}\mathbf{M}^{-1}\mathbf{P}_{\mathcal{I}\cup\mathcal{B}}(0) \\ &= -\left(\mathbf{1}_{\mathcal{I}}, \mathbf{1}_{\mathcal{B}}\right)\begin{bmatrix}\mathbf{K}_{\mathcal{I}\mathcal{I}} - \mathbf{D}_{\mathcal{I}} & \mathbf{K}_{\mathcal{I}\mathcal{B}} \\ \mathbf{K}_{\mathcal{B}\mathcal{I}} & \mathbf{K}_{\mathcal{B}\mathcal{B}} - \mathbf{D}_{\mathcal{B}}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{0}_{\mathcal{I}} \\ \mathbf{P}_{\mathcal{B}}(0)\end{bmatrix} \\ &= -\left(\mathbf{1}_{\mathcal{I}}, \mathbf{1}_{\mathcal{B}}\right)\begin{bmatrix}(\mathbf{B}_{\mathcal{I}\mathcal{I}} - \mathbb{I}_{\mathcal{I}})\mathbf{D}_{\mathcal{I}} & \mathbf{B}_{\mathcal{I}\mathcal{B}}\mathbf{D}_{\mathcal{B}} \\ \mathbf{B}_{\mathcal{B}\mathcal{I}}\mathbf{D}_{\mathcal{I}} & (\mathbf{B}_{\mathcal{B}\mathcal{B}} - \mathbb{I}_{\mathcal{B}})\mathbf{D}_{\mathcal{B}}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{0}_{\mathcal{I}} \\ \mathbf{P}_{\mathcal{B}}(0)\end{bmatrix} \\ &= \left(\mathbf{1}_{\mathcal{I}}, \mathbf{1}_{\mathcal{B}}\right)\left\{\begin{bmatrix}\mathbf{G}_{\mathcal{I}}^{-1} & -\mathbf{B}_{\mathcal{I}\mathcal{B}} \\ -\mathbf{B}_{\mathcal{B}\mathcal{I}} & \mathbf{G}_{\mathcal{B}}^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{D}_{\mathcal{I}} & \mathbf{0}_{\mathcal{I}\mathcal{B}} \\ \mathbf{0}_{\mathcal{B}\mathcal{I}} & \mathbf{D}_{\mathcal{B}}\end{bmatrix}\right\}^{-1}\begin{bmatrix}\mathbf{0}_{\mathcal{I}} \\ \mathbf{P}_{\mathcal{B}}(0)\end{bmatrix} \\ &= \left[\mathbf{1}_{\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}, \mathbf{1}_{\mathcal{B}}\mathbf{D}_{\mathcal{B}}^{-1}\right]\begin{bmatrix}\mathbf{G}_{\mathcal{I}}^{-1} & -\mathbf{B}_{\mathcal{I}\mathcal{B}} \\ -\mathbf{B}_{\mathcal{B}\mathcal{I}} & \mathbf{G}_{\mathcal{B}}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{0}_{\mathcal{I}} \\ \mathbf{P}_{\mathcal{B}}(0)\end{bmatrix}, \end{aligned} \tag{C5}$$

where $\mathbf{G}_{\mathcal{X}}^{-1} = \mathbb{I} - \mathbf{B}_{\mathcal{X}\mathcal{X}}$.

To proceed, we consider the linear system

$$\begin{bmatrix}\mathbf{G}_{\mathcal{I}}^{-1} & -\mathbf{B}_{\mathcal{I}\mathcal{B}} \\ -\mathbf{B}_{\mathcal{B}\mathcal{I}} & \mathbf{G}_{\mathcal{B}}^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{x}_{\mathcal{I}} \\ \mathbf{x}_{\mathcal{B}}\end{bmatrix} = \begin{bmatrix}\mathbf{0}_{\mathcal{I}} \\ \mathbf{P}_{\mathcal{B}}(0)\end{bmatrix} \quad \Rightarrow \langle\tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle = \mathbf{1}_{\mathcal{I}}\mathbf{D}_{\mathcal{I}}^{-1}\mathbf{x}_{\mathcal{I}} + \mathbf{1}_{\mathcal{B}}\mathbf{D}_{\mathcal{B}}^{-1}\mathbf{x}_{\mathcal{B}}, \tag{C6}$$

We can solve for $\mathbf{x}_{\mathcal{I}}$ and $\mathbf{x}_{\mathcal{B}}$ analytically by first expanding the matrix vector product as

$$\begin{aligned} &\mathbf{G}_{\mathcal{I}}^{-1}\mathbf{x}_{\mathcal{I}} - \mathbf{B}_{\mathcal{I}\mathcal{B}}\mathbf{x}_{\mathcal{B}} = \mathbf{0}_{\mathcal{I}}, \quad \Rightarrow \mathbf{x}_{\mathcal{I}} = \mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{B}}\mathbf{x}_{\mathcal{B}}, \\ &\mathbf{G}_{\mathcal{B}}^{-1}\mathbf{x}_{\mathcal{B}} - \mathbf{B}_{\mathcal{B}\mathcal{I}}\mathbf{x}_{\mathcal{I}} = \mathbf{P}_{\mathcal{B}}(0), \quad \Rightarrow \mathbf{x}_{\mathcal{B}} = \left[\mathbf{G}_{\mathcal{B}}^{-1} - \mathbf{B}_{\mathcal{B}\mathcal{I}}\mathbf{G}_{\mathcal{I}}\mathbf{B}_{\mathcal{I}\mathcal{B}}\right]^{-1}\mathbf{P}_{\mathcal{B}}(0). \end{aligned} \tag{C7}$$

We then recover the equation (**??**) directly.

Appendix D: Derivation of exact passage time expressions

Here we follow §?? and consider the waiting time associated with a sum over all paths that start in $\mathcal{B}$ and reach any minimum in $\mathcal{A}$ via any number of steps in the $\mathcal{I}$ region with returns to $\mathcal{B}$ allowed. The path weight can be factorised into a component to account for non-reactive paths starting and finishing in $\mathcal{B}$ via any number of steps in $\mathcal{I}$, and then a reactive path from $\mathcal{B}$ to $\mathcal{A}$ via the $\mathcal{I}$ region:

$$\langle \tau^*_{\mathcal{A}\leftarrow B}\rangle = \partial_\zeta \left(\mathbf{1}_\mathcal{A}\widetilde{\mathbf{B}}^\mathcal{I}_{\mathcal{A}\mathcal{B}}\widetilde{\mathbf{G}}^\mathcal{I}_\mathcal{B}\right)\Big|_{\zeta=0}\mathbf{P}_\mathcal{B}(0). \tag{D1}$$

This formulation gives the average waiting time for each component in $\mathcal{B}$ because the path weights sum to unity in each case: from equation (??) we have $\mathbf{1}_\mathcal{A}\mathbf{B}^\mathcal{I}_{\mathcal{A}\mathcal{B}}\mathbf{G}^\mathcal{I}_\mathcal{B} = \mathbf{1}_\mathcal{B}$. Then from equation (??) we have

$$\partial_\zeta \mathbf{1}_\mathcal{A}\widetilde{\mathbf{B}}^\mathcal{I}_{\mathcal{A}\mathcal{B}}\big|_{\zeta=0} + \partial_\zeta \mathbf{1}_\mathcal{B}\widetilde{\mathbf{B}}^\mathcal{I}_{\mathcal{B}\mathcal{B}}\big|_{\zeta=0} = \mathbf{1}_\mathcal{B}[\mathbf{D}^\mathcal{I}_\mathcal{B}]^{-1}.$$

We now use the chain rule, the derivative $\partial_\zeta\widetilde{\mathbf{G}}^\mathcal{I}_\mathcal{B}\big|_{\zeta=0} = \mathbf{G}^\mathcal{I}_\mathcal{B}\partial_\zeta\widetilde{\mathbf{B}}^\mathcal{I}_{\mathcal{B}\mathcal{B}}\big|_{\zeta=0}\mathbf{G}^\mathcal{I}_\mathcal{B}$, and $\mathbf{1}_\mathcal{A}\mathbf{B}^\mathcal{I}_{\mathcal{A}\mathcal{B}}\mathbf{G}^\mathcal{I}_\mathcal{B} = \mathbf{1}_\mathcal{B}$ from conservation of probability (Appendix ??) to rewrite (??) as

$$\begin{aligned}
\langle \tau^*_{\mathcal{A}\leftarrow B}\rangle &= \left[\mathbf{1}_\mathcal{A}\partial_\zeta\widetilde{\mathbf{B}}^\mathcal{I}_{\mathcal{A}\mathcal{B}}\big|_{\zeta=0} + \mathbf{1}_\mathcal{A}\mathbf{B}^\mathcal{I}_{\mathcal{A}\mathcal{B}}\mathbf{G}^\mathcal{I}_\mathcal{B}\partial_\zeta\widetilde{\mathbf{B}}^\mathcal{I}_{\mathcal{B}\mathcal{B}}\big|_{\zeta=0}\right]\mathbf{G}^\mathcal{I}_\mathcal{B}\mathbf{P}_\mathcal{B}(0)\\
&= \left[\mathbf{1}_\mathcal{A}\partial_\zeta\widetilde{\mathbf{B}}^\mathcal{I}_{\mathcal{A}\mathcal{B}}\big|_{\zeta=0} + \mathbf{1}_\mathcal{B}\partial_\zeta\widetilde{\mathbf{B}}^\mathcal{I}_{\mathcal{B}\mathcal{B}}\big|_{\zeta=0}\right]\mathbf{G}^\mathcal{I}_\mathcal{B}\mathbf{P}_\mathcal{B}(0)\\
&= \mathbf{1}_\mathcal{B}[\mathbf{D}^\mathcal{I}_\mathcal{B}]^{-1}\mathbf{G}^\mathcal{I}_\mathcal{B}\mathbf{P}_\mathcal{B}(0),
\end{aligned} \tag{D2}$$

in agreement with (??), where $\mathbf{1}_\mathcal{B}[\mathbf{D}^\mathcal{I}_\mathcal{B}]^{-1} = \mathbf{1}_\mathcal{B}\mathbf{D}^{-1}_\mathcal{B} + \mathbf{1}_\mathcal{I}\mathbf{D}^{-1}_\mathcal{I}\mathbf{G}_\mathcal{I}\mathbf{B}_{\mathcal{I}\mathcal{B}}$.

Appendix E: Initial condition in $\mathcal{B}^\mathcal{I}$ to reproduce exact waiting time for an arbitrary initial condition in $\mathcal{I}\cup\mathcal{B}$

Following (??), the exact waiting time for any initial condition in $\mathcal{I}\cup\mathcal{B}$ reads

$$\langle \tau^*_{\mathcal{A}\leftarrow\mathcal{I}\cup\mathcal{B}}\rangle \equiv \begin{bmatrix}\mathbf{1}_\mathcal{I}\\ \mathbf{1}_\mathcal{B}\end{bmatrix}^\top \begin{bmatrix}\mathbf{D}_\mathcal{I}-\mathbf{K}_{\mathcal{I}\mathcal{I}} & -\mathbf{K}_{\mathcal{I}\mathcal{B}}\\ -\mathbf{K}_{\mathcal{B}\mathcal{I}} & \mathbf{D}_\mathcal{B}-\mathbf{K}_{\mathcal{B}\mathcal{B}}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{P}_\mathcal{I}(0)\\ \mathbf{P}_\mathcal{B}(0)\end{bmatrix}, \tag{E1}$$

which is identical to (??) except now $\mathbf{P}_\mathcal{I}(0)\neq\mathbf{0}_\mathcal{I}$. Recall the exact $\mathbf{P}_\mathcal{I}(0)=\mathbf{0}_\mathcal{I}$ result (??) or (??) can be written $\langle \tau^*_{\mathcal{A}\leftarrow\mathcal{B}}\rangle = \mathbf{1}_\mathcal{B}\left[\mathbf{D}^\mathcal{I}_\mathcal{B}\right]^{-1}\mathbf{G}^\mathcal{I}_\mathcal{B}\mathbf{P}_\mathcal{B}(0)$. Rearranging (??) gives the analogous result

$$\begin{aligned}
\langle \tau^*_{\mathcal{A}\leftarrow\mathcal{I}\cup\mathcal{B}}\rangle &= \mathbf{1}_\mathcal{B}\left[\mathbf{D}^\mathcal{I}_\mathcal{B}\right]^{-1}\mathbf{G}^\mathcal{I}_\mathcal{B}\left[\mathbf{P}_\mathcal{B}(0)+\mathbf{B}_{\mathcal{B}\mathcal{I}}\mathbf{G}_\mathcal{I}\mathbf{P}_\mathcal{I}(0)\right] + \mathbf{1}_\mathcal{I}\mathbf{D}^{-1}_\mathcal{I}\mathbf{G}_\mathcal{I}\mathbf{P}_\mathcal{I}(0)\\
&= \mathbf{1}_\mathcal{B}\left[\mathbf{D}^\mathcal{I}_\mathcal{B}\right]^{-1}\mathbf{G}^\mathcal{I}_\mathcal{B}\overline{\mathbf{P}}_\mathcal{B}(0)\\
\text{with}\qquad \overline{\mathbf{P}}_\mathcal{B}(0) &= \mathbf{P}_\mathcal{B}(0) + \left[\mathbf{K}_{\mathcal{B}\mathcal{I}} + \left[\mathbf{G}^\mathcal{I}_\mathcal{B}\right]^{-1}\mathbf{D}^\mathcal{I}_\mathcal{B}\left(\mathbf{z}_\mathcal{B}\otimes\mathbf{1}_\mathcal{I}\right)\right]\mathbf{D}^{-1}_\mathcal{I}\mathbf{G}_\mathcal{I}\mathbf{P}_\mathcal{I}(0).
\end{aligned} \tag{E2}$$

where $\mathbf{z}_\mathcal{B}$ is *any* vector that satisfies $\mathbf{1}_\mathcal{B}\mathbf{z}_\mathcal{B}=1$. This freedom is not surprising as this term accounts for all paths that decay to $\mathcal{A}$ without passing through $\mathcal{B}$. We thus see that the exact waiting time for arbitrary initial condition in $\mathcal{I}\cup\mathcal{B}$ can be recovered in the GT renormalised Markov chain with an initial condition $\mathbf{P}_{\mathcal{B}^\mathcal{I}}(0)=\overline{\mathbf{P}}_\mathcal{B}(0)$. Requiring $\mathbf{1}_\mathcal{B}\overline{\mathbf{P}}_\mathcal{B}(0)=1$ for a probability distribution in $\mathcal{B}^\mathcal{I}$ yields

$$\begin{aligned}
\mathbf{1}_\mathcal{B}\overline{\mathbf{P}}_\mathcal{B}(0) &= \mathbf{1}_\mathcal{B}\mathbf{P}_\mathcal{B}(0) + \left[\mathbf{1}_\mathcal{B}\mathbf{B}_{\mathcal{B}\mathcal{I}}\mathbf{G}_\mathcal{I} + \left(\mathbf{1}_\mathcal{B}\left[\mathbf{G}^\mathcal{I}_\mathcal{B}\right]^{-1}\mathbf{D}^\mathcal{I}_\mathcal{B}\mathbf{z}_\mathcal{B}\right)\mathbf{1}_\mathcal{I}\mathbf{D}^{-1}_\mathcal{I}\mathbf{G}_\mathcal{I}\right]\mathbf{P}_\mathcal{I}(0) = 1\\
\Rightarrow \mathbf{1}_\mathcal{B}\left[\mathbf{G}^\mathcal{I}_\mathcal{B}\right]^{-1}\mathbf{D}^\mathcal{I}_\mathcal{B}\mathbf{z}_\mathcal{B} &= \frac{1-\mathbf{1}_\mathcal{B}\mathbf{P}_\mathcal{B}(0)-\mathbf{1}_\mathcal{B}\mathbf{B}_{\mathcal{B}\mathcal{I}}\mathbf{G}_\mathcal{I}\mathbf{P}_\mathcal{I}(0)}{\mathbf{1}_\mathcal{B}\mathbf{D}^{-1}_\mathcal{I}\mathbf{G}_\mathcal{I}\mathbf{P}_\mathcal{I}(0)}.
\end{aligned} \tag{E3}$$

As $\mathbf{1}_\mathcal{B}\mathbf{z}_\mathcal{B}=1$, $\mathbf{z}_\mathcal{B}$ can be thought of as a probability distribution in $\mathcal{B}^\mathcal{I}$, meaning $\mathbf{1}_\mathcal{B}\left[\mathbf{G}^\mathcal{I}_\mathcal{B}\right]^{-1}\mathbf{D}^\mathcal{I}_\mathcal{B}\mathbf{z}_\mathcal{B}$ is the flux to $\mathcal{A}$ for $\mathbf{P}_{\mathcal{B}^\mathcal{I}}=\mathbf{z}_\mathcal{B}$ by comparison with the evolution equation (??). Similarly, the denominator on the right can be interpreted as the waiting time for escape to $\mathcal{A}$ for paths that never pass through $\mathcal{B}^\mathcal{I}$, while the numerator is the corresponding total probability for such paths.

Finally, $\mathbf{B}_{\mathcal{B}\mathcal{I}}\mathbf{G}_\mathcal{I}\mathbf{P}_\mathcal{I}(0)$ is probability mass from $\mathcal{I}$ that passes through $\mathcal{B}$ before reaching $\mathcal{A}$, the product of the compound probability $\mathbf{G}_\mathcal{I}\mathbf{P}_\mathcal{I}(0)$ of all paths that stay in $\mathcal{I}$ multiplied by the branching probability $\mathbf{B}_{\mathcal{B}\mathcal{I}}=\mathbf{K}_{\mathcal{B}\mathcal{I}}\mathbf{D}^{-1}_\mathcal{I}$ into $\mathcal{B}$. The exact GT result for the waiting time shows the initial distribution in $\mathcal{B}^\mathcal{I}$ is $\mathbf{P}_\mathcal{B}(0)$ plus a renormalised branching probability for probability in $\mathcal{I}$, namely with effective rates $\mathbf{K}_{\mathcal{B}\mathcal{I}}\to\mathbf{K}_{\mathcal{B}\mathcal{I}}+\left[\mathbf{G}^\mathcal{I}_\mathcal{B}\right]^{-1}\mathbf{D}^\mathcal{I}_\mathcal{B}\mathbf{w}_\mathcal{B}\otimes\mathbf{1}_\mathcal{I}$, which accounts for the flux into $\mathcal{A}$ from paths that do not enter $\mathcal{B}^\mathcal{I}$.

Appendix F: Derivation of steady state transition rate and waiting time

The steady state approximation assumes that the intervening region is in local equilibrium with $\dot{\mathbf{P}}_{\mathcal{I}} = \mathbf{0}$ on the timescale of transitions between $\mathcal{A}, \mathcal{B}$. This assumption is reasonable when $\mathcal{A}, \mathcal{B}$ are sufficiently metastable. The approximation of a steady state in $\mathcal{I}$ yields the equality

$$\mathbf{D}_{\mathcal{I}}\mathbf{P}_{\mathcal{I}} = \left[\mathbb{I}_{\mathcal{I}} - \mathbf{B}_{\mathcal{I}\mathcal{I}}\right]^{-1}\left[\mathbf{B}_{\mathcal{I}\mathcal{A}}\mathbf{D}_{\mathcal{A}}\mathbf{P}_{\mathcal{A}} + \mathbf{B}_{\mathcal{I}\mathcal{B}}\mathbf{D}_{\mathcal{B}}\mathbf{P}_{\mathcal{B}}\right]. \tag{F1}$$

In addition, the original SS DPS derivation assumes that both $\mathcal{A}$ and $\mathcal{B}$ are populated with local restricted equilibrium distributions

$$\mathbf{P}_{\mathcal{A}}(t) \to \widehat{\boldsymbol{\pi}}_{\mathcal{A}}\mathrm{P}_{\mathcal{A}}(t), \quad \mathbf{P}_{\mathcal{B}}(t) \to \widehat{\boldsymbol{\pi}}_{\mathcal{B}}\mathrm{P}_{\mathcal{B}}(t). \tag{F2}$$

We therefore obtain a reduced evolution equation for $\mathbf{P}_{\mathcal{A}}$ (with the obvious analogue for $\mathcal{B}$)

$$\begin{aligned}\dot{\mathrm{P}}_{\mathcal{A}} &= \mathbf{1}_{\mathcal{A}}\left[\mathbf{B}_{\mathcal{A}\mathcal{A}} - \mathbb{I}_{\mathcal{A}}\right]\mathbf{D}_{\mathcal{A}}\widehat{\boldsymbol{\pi}}_{\mathcal{A}}\mathrm{P}_{\mathcal{A}} + \mathbf{1}_{\mathcal{A}}\mathbf{B}_{\mathcal{A}\mathcal{B}}\mathbf{D}_{\mathcal{B}}\widehat{\boldsymbol{\pi}}_{\mathcal{B}}\mathrm{P}_{\mathcal{B}} + \mathbf{1}_{\mathcal{A}}\mathbf{B}_{\mathcal{A}\mathcal{I}}\mathbf{D}_{\mathcal{I}}\mathbf{P}_{\mathcal{I}} \\ &= \mathbf{1}_{\mathcal{A}}\left[\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{A}} - \mathbb{I}_{\mathcal{A}}\right]\mathbf{D}_{\mathcal{A}}\widehat{\boldsymbol{\pi}}_{\mathcal{A}}\mathrm{P}_{\mathcal{A}} + \mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}\mathbf{D}_{\mathcal{B}}\widehat{\boldsymbol{\pi}}_{\mathcal{B}}\mathrm{P}_{\mathcal{B}},\end{aligned} \tag{F3}$$

where the branching probabilities $\mathbf{B}^{\mathcal{I}}_{\mathcal{X}\mathcal{Y}}$ are defined in (**??**). The steady state $\mathcal{A} \leftarrow \mathcal{B}$ rate into $\mathcal{A}$ from $\mathcal{B}$ can be read directly from (**??**), giving

$$k^{SS}_{\mathcal{A}\leftarrow\mathcal{B}} = \mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}}_{\mathcal{A}\mathcal{B}}\mathbf{D}_{\mathcal{B}}\widehat{\boldsymbol{\pi}}_{\mathcal{B}} = \mathbf{C}^{\mathcal{A}}_{\mathcal{B}}\mathbf{D}_{\mathcal{B}}\widehat{\boldsymbol{\pi}}_{\mathcal{B}}. \tag{F4}$$

The last equality uses the committor definition (**??**) to recover (**??**).

As noted in §**??**, the above expression for the rate associates a waiting time for each path that only accounts for the escape time from the initial $\mathcal{B}$ minimum.

Appendix G: First step analysis for equivalence of (??) and (??)

In the state space after removal of all $\mathcal{I}$ minima and $b' \neq b$ in $\mathcal{B}$ the renormalised waiting time and branching probabilities, denoted τ^{F}_b, $P^{\mathrm{F}}_{\mathcal{A}b}$ and P^{F}_{bb} in previous work,[?] correspond to steps from b to any $a \in \mathcal{A}$ or back to b. $\mathcal{T}_{\mathcal{A}b}$ is then obtained as

$$\mathcal{T}_{\mathcal{A}b} = \tau^{\mathrm{F}}_b P^{\mathrm{F}}_{\mathcal{A}b}(1 + 2P^{\mathrm{F}}_{bb} + 3(P^{\mathrm{F}}_{bb})^2 + \ldots) = \tau^{\mathrm{F}}_b P^{\mathrm{F}}_{\mathcal{A}b}/(1 - P^{\mathrm{F}}_{bb})^2 = \tau^{\mathrm{F}}_b/P^{\mathrm{F}}_{\mathcal{A}b}. \tag{G1}$$

From the first-step relation[?]

$$\mathcal{T}_{\mathcal{A}b_1} = \tau^{\mathcal{I}}_{b_1} + \sum_{b\in\mathcal{B}} \mathcal{T}_{\mathcal{A}b}[\mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}]_{bb_1}, \tag{G2}$$

adding and subtracting $\mathcal{T}_{\mathcal{A}b_1}$ on the right we find

$$\mathcal{T}_{\mathcal{A}b_1} = \tau^{\mathcal{I}}_{b_1} + \sum_{b\in\mathcal{B}}[\mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}]_{bb_1}(\mathcal{T}_{\mathcal{A}b} - \mathcal{T}_{\mathcal{A}b_1}) + \sum_{b\in\mathcal{B}}[\mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}]_{bb_1}\mathcal{T}_{\mathcal{A}b_1} \quad \text{so} \quad C^{\mathcal{A}}_{b_1}\mathcal{T}_{\mathcal{A}b_1} = \tau^{\mathcal{I}}_{b_1} + \sum_{b\in\mathcal{B}} C^{b}_{b_1}(\mathcal{T}_{\mathcal{A}b} - \mathcal{T}_{\mathcal{A}b_1}).$$

Hence $\mathcal{T}_{\mathcal{A}b_1} = \tau^{\mathcal{I}}_{b_1}/C^{\mathcal{A}}_{b_1}$ if $\mathcal{T}_{\mathcal{A}b}$ is the same for all b, and the two rate formulations in (**??**) and (**??**) agree.

For completeness we now demonstrate how the first-step relation is encoded in the matrix formulation that is the principal representation of the present contribution. With $\mathcal{T}_{\mathcal{A}b} = \left[\mathbf{1}_{\mathcal{B}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]^{-1}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]_b$ we have

$$\begin{aligned}\sum_{b\in\mathcal{B}}\mathcal{T}_{\mathcal{A}b}[\mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}]_{bb_1} &= \left[\mathbf{1}_{\mathcal{B}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]^{-1}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\mathbf{B}^{\mathcal{I}}_{\mathcal{B}\mathcal{B}}\right]_{b_1} = \left[\mathbf{1}_{\mathcal{B}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]^{-1}\left(\mathbf{G}^{\mathcal{I}}_{\mathcal{B}} - \mathbb{I}_{\mathcal{B}}\right)\right]_{b_1} \\ &= \left[\mathbf{1}_{\mathcal{B}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]^{-1}\mathbf{G}^{\mathcal{I}}_{\mathcal{B}}\right]_{b_1} - \left[\mathbf{1}_{\mathcal{B}}[\mathbf{D}^{\mathcal{I}}_{\mathcal{B}}]^{-1}\right]_{b_1} = \mathcal{T}_{\mathcal{A}b_1} - \tau^{I}_{b_1}, \quad \text{as required.}\end{aligned} \tag{G3}$$

Appendix H: Sensitivity to direct transitions

1. Sensitivity to direct $\mathcal{P} \to \mathcal{P}$ transitions

For the case $(l, m) \in (\mathcal{P}, \mathcal{P})$ we have to consider sensitivity to both the $l \leftarrow m$ perturbation and the $m \leftarrow l$ perturbation. If we define the 'one-sided' difference operator $\delta^{m\leftarrow l}_{\mathcal{P}\mathcal{P}}$ as only accounting for the $l \to m$ perturbation, from detailed balance we can write

$$\delta^{ml}_{\mathcal{P}\mathcal{P}}\mathrm{C}^{\mathcal{A}}_{\mathcal{B}} = \delta^{m\leftarrow l}_{\mathcal{P}\mathcal{P}}\mathrm{C}^{\mathcal{A}}_{\mathcal{B}} + \phi^{ml}\delta^{l\leftarrow m}_{\mathcal{P}\mathcal{P}}\mathrm{C}^{\mathcal{A}}_{\mathcal{B}}. \tag{H1}$$

The rate matrix modification from (**??**) reads

$$[\delta^{ml}_{\mathcal{PP}}\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}]_{ij} = \mathrm{k}^U\left(\delta_{im}\delta_{jl} + (\phi^{ml} - \delta_{lm})\delta_{jm}\delta_{il}\right), \tag{H2}$$

with all other modifications zero. This gives branching probability modifications of

$$\text{general form}: \qquad \delta^{ml}_{\mathcal{PP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XP}} = \delta^{m\leftarrow l}_{\mathcal{PP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XP}} + \phi^{ml}\delta^{l\leftarrow m}_{\mathcal{PP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XP}}. \tag{H3}$$

$$\text{So} \qquad [\delta^{m\leftarrow l}_{\mathcal{PP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}]_{ij} = \mathrm{k}^U\left(\delta_{im} - [\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}]_{il}\right)[\mathbf{D}^{-1}_{\mathcal{P}}]_{ll}\delta_{jl}, \tag{H4}$$

$$[\delta^{m\leftarrow l}_{\mathcal{PP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}]_{ij} = -\mathrm{k}^U[\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}]_{il}[\mathbf{D}^{-1}_{\mathcal{P}}]_{ll}\delta_{jl}, \tag{H5}$$

$$\delta^{m\leftarrow l}_{\mathcal{PP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}} = 0, \tag{H6}$$

where we emphasise that $\delta^{m\leftarrow l}_{\mathcal{PP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XP}}$ is the modification only for $l \to m$. Note the index swap in the second term of (**??**) to capture the reverse $m \to l$ contribution with rate $\mathrm{k}^U\phi^{ml}$.

The 'one-sided' sensitivity of the total branching probability can then be written

$$\begin{aligned}\delta^{m\leftarrow l}_{\mathcal{PP}}\mathrm{C}^{\mathcal{A}}_{\mathcal{B}} &= \mathbf{1}_{\mathcal{A}}\delta^{m\leftarrow l}_{\mathcal{PP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}\mathbf{x} + \mathbf{y}\delta^{m\leftarrow l}_{\mathcal{PP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}\mathbf{x} \\ &= \mathrm{k}^U\left([\mathbf{y}]_m - \left[\mathbf{y}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}} + \mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}\right]_l\right)[\mathbf{D}^{-1}_{\mathcal{P}}\mathbf{x}]_l, \\ &= \mathrm{k}^U\left([\mathbf{y}]_m - [\mathbf{y}]_l\right)[\mathbf{D}^{-1}_{\mathcal{P}}\mathbf{x}]_l,\end{aligned} \tag{H7}$$

where the last in equality uses equation (**??**), which implies that $\mathbf{y}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}} + \mathbf{1}_{\mathcal{A}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}} = \mathbf{y}$.

2. Sensitivity to direct $\mathcal{B} \to \mathcal{P}$ transitions

For $(l, m) \in (\mathcal{B}, \mathcal{P})$ we have the rate and branching probability matrix modifications

$$[\delta^{ml}_{\mathcal{PB}}\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}}]_{ij} = \mathrm{k}^U\delta_{im}\delta_{jl}, \tag{H8}$$

$$[\delta^{ml}_{\mathcal{PB}}\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XP}}]_{ij} = 0, \tag{H9}$$

$$[\delta^{ml}_{\mathcal{PB}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}}]_{ij} = \mathrm{k}^U\left(\delta_{im} - [\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}}]_{il}\right)[\mathbf{D}^{-1}_{\mathcal{B}}]_{ll}\delta_{jl}, \tag{H10}$$

$$[\delta^{ml}_{\mathcal{PB}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XP}}]_{ij} = -\mathrm{k}^U\phi^{ml}[\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{XP}}]_{im}[\mathbf{D}^{-1}_{\mathcal{P}}]_{mm}\delta_{jm}, \tag{H11}$$

where $\mathcal{X} = \mathcal{A}, \mathcal{P}$, and thus the final sensitivity, with $l \in \mathcal{B}, m \in \mathcal{P}$, is

$$\begin{aligned}\delta^{ml}_{\mathcal{PB}}\mathrm{C}^{\mathcal{A}}_{\mathcal{B}} =& \mathrm{k}^U\left([\mathbf{y}]_m - [\mathbf{y}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}}]_l\right)[\mathbf{D}^{-1}_{\mathcal{B}}]_{ll} \\ &- \mathrm{k}^U\phi^{ml}[\mathbf{y}]_m[\mathbf{D}^{-1}_{\mathcal{P}}\mathbf{x}]_m\end{aligned} \tag{H12}$$

3. Sensitivity to direct $\mathcal{P} \to \mathcal{A}$ transitions

For $(l, m) \in (\mathcal{P}, \mathcal{A})$ we have the rate matrix modifications

$$[\delta^{ml}_{\mathcal{AP}}\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}]_{ij} = \mathrm{k}^U\delta_{im}\delta_{jl}, \tag{H13}$$

$$[\delta^{ml}_{\mathcal{AP}}\mathbf{K}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PA}}]_{ij} = \mathrm{k}^U\phi^{ml}\delta_{il}\delta_{jm}, \tag{H14}$$

giving in turn

$$[\delta^{ml}_{\mathcal{AP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}]_{ij} = \mathrm{k}^U\left(\delta_{im} - [\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{AP}}]_{il}\right)[\mathbf{D}^{-1}_{\mathcal{P}}]_{ll}\delta_{jl}, \tag{H15}$$

$$[\delta^{ml}_{\mathcal{AP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}]_{ij} = -\mathrm{k}^U[\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PP}}]_{il}[\mathbf{D}^{-1}_{\mathcal{P}}]_{ll}\delta_{jl}, \tag{H16}$$

$$[\delta^{ml}_{\mathcal{AP}}\mathbf{B}^{\mathcal{I}\setminus\mathcal{P}}_{\mathcal{PB}}]_{ij} = 0 \tag{H17}$$

and thus the final sensitivity, with $l \in \mathcal{P}, m \in \mathcal{A}$, of

$$\delta^{ml}_{\mathcal{AP}}\mathrm{C}^{\mathcal{A}}_{\mathcal{B}} = \mathrm{k}^U\left(1 - [\mathbf{y}]_l\right)[\mathbf{D}^{-1}_{\mathcal{P}}\mathbf{x}]_l, \tag{H18}$$

which we note is independent of $m \in \mathcal{A}$, as expected due to the form of the committor vector $\mathbf{C}^{\mathcal{A}}_{\mathcal{B}}$ defined in (**??**), with a sum over all $\mathcal{A}$ states.