



Algebraic formulas for first-passage times of Markov processes in the linear framework

Kee-Myoung Nam^{1,2} · Jeremy Gunawardena^{1,3} 

Received: 9 March 2025 / Accepted: 22 August 2025

© The Author(s) 2025

Abstract

The linear framework is an approach to analysing biochemical systems based on directed graphs with labelled edges. When applied to individual molecular systems, graph vertices correspond to system states, directed edges to transitions, and edge labels to transition rates. Such a graph specifies the infinitesimal generator of a continuous-time Markov process. The master equation of this Markov process, which describes the forward evolution of vertex probabilities, is a linear differential equation, after which the framework is named, whose operator is the Laplacian matrix of the graph. The Matrix-Tree theorem, when applied to this Laplacian matrix, allows the steady-state probabilities of the Markov process to be expressed as rational algebraic functions of the transition rates. This capability gives algebraic access to problems that have otherwise been treated by approximations or numerical simulations, and enables theorems to be proved about biochemical systems that rise above their underlying molecular complexity. Here, we extend this capability from the steady state to the transient regime. We use the All-Minors Matrix-Tree theorem to express the moments of the conditional first-passage time distribution, and the corresponding splitting probabilities, as rational algebraic functions of the transition rates. This extended capability brings many new biological problems within the scope of the linear framework.

Keywords Labelled directed graphs · Laplacian matrix · Rational algebraic functions · First-passage times · Matrix-Tree theorems

✉ Jeremy Gunawardena
jeremy@hms.harvard.edu ; jeremy.gunawardena@upf.edu

Kee-Myoung Nam
keemyoung.nam@yale.edu

¹ Department of Systems Biology, Harvard Medical School, 200 Longwood Ave., Boston, MA 02115, USA

² Department of Molecular, Cellular and Developmental Biology, Yale University, 260 Whitney Ave., New Haven, CT 06511, USA

³ Department of Medicine and Life Sciences, Pompeu Fabra University, Dr. Aiguader 80, Barcelona 08003, Spain

1 Introduction

In previous work, we introduced a graph-theoretic “linear framework” for analysing biochemical systems (Gunawardena 2012; Mirzaev and Gunawardena 2013); for a recent overview, see Nam et al. (2022). The framework is based on finite, simple, directed graphs with labelled edges and no self-loops. Given such a graph G , the vertices of G , $\mathcal{V}(G) = \{1, \dots, N\}$, typically represent biochemical components; the directed edges, $\mathcal{E}(G) \subseteq \mathcal{V}(G) \times \mathcal{V}(G)$, denoted $i \rightarrow j$, represent reactions between the corresponding components; and the edge labels, denoted $\ell(i \rightarrow j)$, represent positive reaction rates, with units of $(\text{time})^{-1}$. The edge labels may include terms that describe interactions between the graph and its environment, such as the concentration of a binding ligand, and this allows typical biochemical nonlinearities to be accommodated.

A linear framework graph gives rise to a linear dynamics, from which the framework gets its name. The simplest way to describe this dynamics is to consider each edge as a chemical reaction obeying mass-action kinetics with the corresponding edge label as the rate constant. This yields a system of differential equations for the concentrations of the vertices, $\mathbf{x}(t) = (x_1(t), \dots, x_N(t))^T$, where T denotes transpose. Since an edge has only one source vertex, these equations are linear and can be expressed in matrix form as

$$\frac{d\mathbf{x}}{dt} = \mathcal{L}(G) \mathbf{x}. \quad (1)$$

Here, $\mathcal{L}(G)$ is the *Laplacian matrix* of the graph G (Chung 1997). Under appropriate scalings, Laplacian matrices correspond to discrete versions of the Laplacian differential operator (Chung 1997), so that Eqn. 1 may be thought of as a discretised diffusion equation. Since material is neither created nor destroyed by the dynamics, there is a conservation law,

$$x_1(t) + \dots + x_N(t) = x_{\text{tot}}.$$

This equation corresponds to the column sums of $\mathcal{L}(G)$ being zero, $\mathbf{1}^T \mathcal{L}(G) = \mathbf{0}^T$, where $\mathbf{0}$ and $\mathbf{1}$ denote the all-zero and all-one column vectors of the appropriate dimension, respectively. (Synthesis and degradation can be accommodated (Gunawardena 2012; Mirzaev and Bortz 2015) but are not considered here.) If we unwrap the definition of $\mathcal{L}(G)$, we see that the matrix entries are given by,

$$\mathcal{L}(G)_{i,j} = \begin{cases} 0 & \text{if } i \neq j \text{ and } j \nrightarrow i \\ \ell(j \rightarrow i) & \text{if } i \neq j \text{ and } j \rightarrow i \\ -\sum_{k: j \rightarrow k} \ell(j \rightarrow k) & \text{if } i = j. \end{cases} \quad (2)$$

An example linear framework graph and its associated Laplacian matrix are shown in Fig. 1A.

Two biochemical settings have been studied within the linear framework. In the first setting, the system is a macroscopic bulk mixture of biochemical components interacting through chemical reactions. Here, the variables correspond to the concentrations of the components, which evolve deterministically over time, as described by

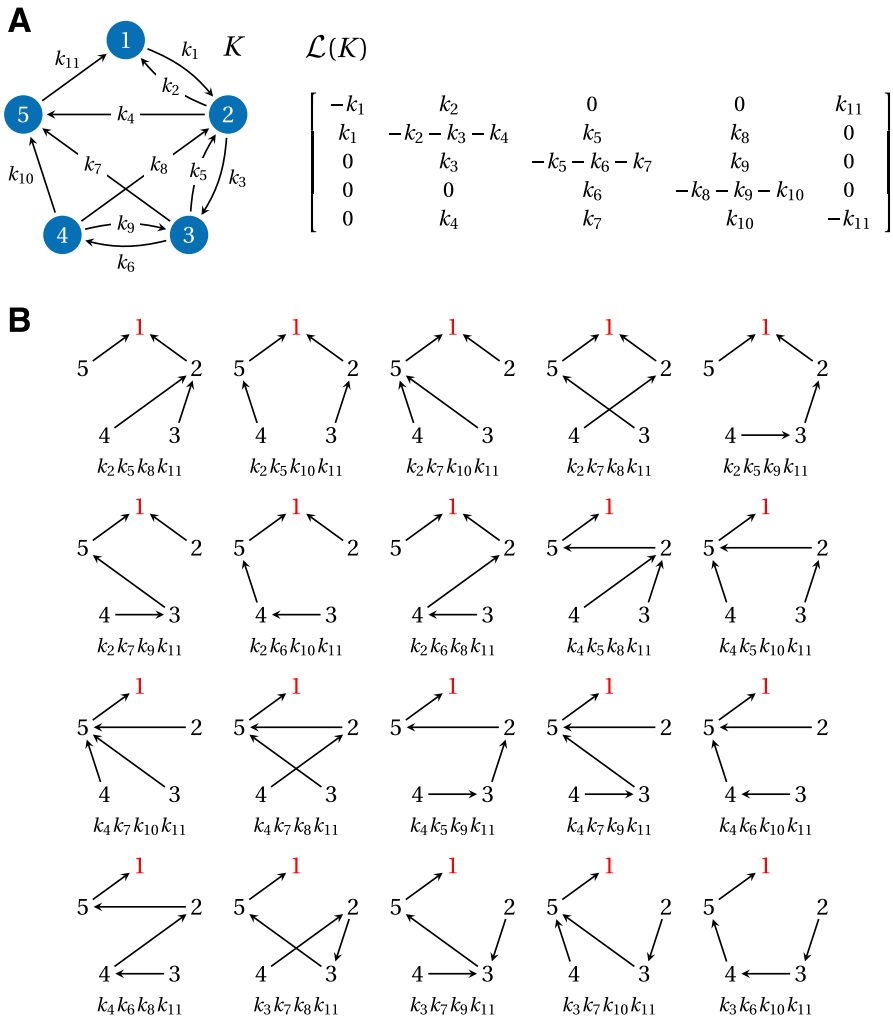


Fig. 1 (A) An example linear framework graph, K , on five vertices and its Laplacian matrix $\mathcal{L}(K)$. This graph also appears in Nam and Gunawardena (2023, Fig. 2) but with a different labelling. (B) The 20 spanning trees of K that are rooted at vertex 1 (red), and their corresponding products of edge labels

the Laplacian dynamics in Eqn. 1. A typical example is a network of enzymes and substrates; the linear framework was originally introduced to analyse such systems (Gunawardena 2012). For related applications in the macroscopic setting, see Dasgupta et al. (2014), Nam et al. (2020), Yordanov and Stelling (2018), (2020), and for a review, see Gunawardena (2014).

In the second setting, the system is a microscopic one, typically consisting of an individual macromolecule, such as a gene. The macromolecule evolves stochastically among a collection of states, partly through interactions with external components in the surrounding thermal bath. The vertices of the graph represent the states of the

macromolecule, the edges represent transitions among these states, and the edge labels represent transition rates. The components in the environment are often assumed to be present in such abundance, in thermodynamic “reservoirs”, that their concentrations can be treated as constant during interactions with the graph. In this case, the edge labels are constants. This specification gives rise to a finite-state, continuous-time, time-homogeneous Markov process, $X(t)$. Each variable, $x_i(t)$, in Eqn. 1 corresponds to the probability of the system occupying state i at time t ; and there is an edge $i \rightarrow j$ if, and only if, the infinitesimal transition rate from i to j is positive,

$$\lim_{h \rightarrow 0} \frac{\Pr[X(t+h) = j \mid X(t) = i]}{h} > 0, \quad (3)$$

and this positive rate is the corresponding edge label, $\ell(i \rightarrow j)$. Any finite-state, time-homogeneous Markov process for which the infinitesimal transition rates are well-defined can be described by a linear framework graph in this way (Mirzaev and Gunawardena 2013). (In some probability textbooks (Anderson 1991; Norris 1997), the infinitesimal generator of the Markov process is $\mathcal{L}(G)^T$ and is called the “ Q -matrix.”) Gene regulation is a biological context in which the microscopic interpretation of the linear framework has been particularly useful (Ahrendorf et al. 2014; Estrada et al. 2016; Biddle et al. 2019; Park et al. 2019; Martinez-Corral et al. 2024; Nasser et al. 2024); for a review, see Wong and Gunawardena (2020).

In the microscopic interpretation, the Laplacian dynamics in Eqn. 1 is the *master equation*, or *Kolmogorov forward equation*, associated with $X(t)$ (van Kampen 2007), which describes the deterministic time evolution of the probability distribution on the states in the process. It is an interesting fact that both the macroscopic and microscopic contexts can be described by the same Laplacian dynamics in Eqn. 1. In this paper, we will focus on the microscopic setting, so that $x_i(t)$ will denote probability rather than concentration, with $x_{\text{tot}} = 1$. We will accordingly be studying Markov processes from a graph-theoretic perspective.

Up to now, the focus of the linear framework has been on the steady state, which is appropriate for the timescale separations that are frequently used in analysing biochemical systems (Gunawardena 2014). It follows from Eqn. 1 that the steady state, \mathbf{x}^* , lies in the kernel of $\mathcal{L}(G)$. A basis for $\ker \mathcal{L}(G)$ can be described in terms of the edge labels of G by the Matrix-Tree theorem (MTT) (Gunawardena 2012; Mirzaev and Gunawardena 2013); see Theorem 1. When G is strongly connected, meaning that there is a directed path of edges from each vertex to every other vertex in G , then (Gunawardena 2012),

$$\dim \ker \mathcal{L}(G) = 1.$$

The MTT shows that a canonical basis vector, $\boldsymbol{\rho}(G) \in \ker \mathcal{L}(G)$, is given by

$$\rho_i(G) = \sum_{F \in \Phi_{\{i\}}(G)} \left(\prod_{j \rightarrow k \in F} \ell(j \rightarrow k) \right), \quad (4)$$

where $\Phi_{\{i\}}(G)$ is the set of *spanning trees* of G that are *rooted* at i . A spanning tree, F , of G is a subgraph of G which includes each vertex (“spanning”), has no cycles if edge directions are ignored (“tree”), and has exactly one outgoing edge from each vertex except one, which has no outgoing edges; this special vertex is the *root* of F . For instance, the graph in Fig. 1A, which is strongly connected, has 20 spanning trees that are rooted at the vertex 1, each of which contributes a product of four edge labels to the sum in Eqn. 4 (Fig. 1B). We can obtain the steady-state probability of vertex i from Eqn. 4 by normalising to the total,

$$x_i^* = \frac{\rho_i(G)}{\rho_1(G) + \cdots + \rho_N(G)}. \quad (5)$$

The case of a general, non-strongly connected graph is discussed in Mirzaev and Gunawardena (2013). Importantly, Eqns. 4 and 5 are symbolic in the edge labels and do not require the numerical values of the transition rates to be known.

Since the Laplacian dynamics in Eqn. 1 is linear, the steady state is an eigenvector for the zero eigenvalue. This eigenvector can be readily obtained from $\mathcal{L}(G)$ but the resulting determinant formulas contain alternating signs. The MTT is important because it accounts for the remarkable cancellations that allow each $\rho_i(G)$ to be expressed as a *manifestly positive* polynomial in the edge labels (Eqn. 4). (The distinction here is that although the polynomial $a^2 - 2ab + b^2 = (a - b)^2$ is always positive when a and b are positive and unequal, it is not a sum of monomials with positive coefficients, which is “manifestly” positive.) Indeed, Eqn. 4 yields a sum of monomials whose coefficients are all positive integers, in this case the same integer, 1. Eqn. 5 then shows that steady-state probabilities are also rational algebraic functions of the edge labels with the same property. Positivity would be expected for the functional dependency of the probability on the edge labels but the manifest (integral) positivity is an additional mathematical feature. It arises because the steady-state probabilities of a Markov process are not merely probabilities. They are also governed by an underlying combinatorial structure, derived from the spanning trees of the corresponding linear framework graph, as expressed by the MTT (Eqn. 4). Manifest positivity also arises in other contexts; see Williams (2022) for more discussion. The existence within the steady state of an underlying combinatorial structure, in the form of spanning trees, from which the manifest positivity arises, is a mathematical feature that has been central to the linear framework.

Another significant consequence of the MTT that emerges from Eqn. 4 is that steady-state probabilities can be calculated irrespective of the thermodynamic context. If the graph can reach a steady state of thermodynamic equilibrium, so that *detailed balance*, or the *cycle condition*, is satisfied, then Eqn. 5 reduces to the standard prescription of equilibrium statistical mechanics, with the denominator corresponding to the partition

function for the grand canonical ensemble (Gunawardena 2012; Ahsendorf et al. 2014; Estrada et al. 2016). However, Eqns. 4 and 5 hold equally well in a steady state away from thermodynamic equilibrium. The linear framework thereby provides a restricted context in which non-equilibrium statistical mechanics is exactly solvable. This capability has been exploited in several recent applications (Estrada et al. 2016; Biddle et al. 2019; Park et al. 2019; Martinez-Corral et al. 2024); for a review in the context of gene regulation, see Wong and Gunawardena (2020).

The use of graph-theoretic methods in stochastic thermodynamics goes back to the work of Hill (1966) and Schnakenberg (1976) and has been resurrected more recently; see, for example, Andrieux and Gaspard (2007); Murugan et al. (2014); Owen and Horowitz (2023); Dal Cengio et al. (2023); Qureshi et al. (2023). For biochemical systems, the graph theory approach of the linear framework allows appropriate molecular details to be left unspecified, by imposing purely topological requirements, while the rational algebraic access through Eqn. 5, which does not require parameter values to be known, has allowed rigorous theorems to be proved that rise above the underlying molecular complexity (Wong et al. 2018; Biddle et al. 2021; Martinez-Corral et al. 2024).

The linear framework has been successful in analysing systems at steady state but it had long been thought that it could not be extended to the transient regime, during which the system is relaxing to a steady state. It seemed that this regime could only be accessed through the eigenvalues of $\mathcal{L}(G)$. The main contribution of the present paper is to show that this is not the case. It is possible to express certain transient quantities, such as the moments of the *first-passage time* (FPT) distribution, as manifestly positive rational algebraic functions of the edge labels. This capability opens up a large class of new problems that can potentially be addressed in a similar way to the applications described above.

The FPT from one vertex, i , to another, j , is the random variable measuring the time taken by a trajectory of the associated Markov process to reach j for the first time, when started from i . FPTs have been used to quantify a variety of observables in molecular biophysics: the completion time of an enzymatic turnover (Kou et al. 2005; Shaevitz et al. 2005; Kolomeisky and Fisher 2007; Chemla et al. 2008; Garai et al. 2009; Bel et al. 2010; Moffitt et al. 2010; Moffitt and Bustamante 2014), the probability and speed with which an enzyme can choose between correct and incorrect reaction pathways (Banerjee et al. 2017; Cui and Mehta 2018; Mallory et al. 2019; Wang et al. 2021), the speed with which a gene alternates between different regulatory states (Lammers et al. 2020; Alamos et al. 2023; Lammers et al. 2023), the time taken by a regulated molecule to cross an abundance threshold (Dal Co et al. 2017; Ghusinga et al. 2017; Gupta et al. 2018; Ham et al. 2024), and the rate of templated copolymer growth (Qureshi et al. 2023; Guntoro et al. 2025). However, despite their broad usefulness in biology, FPTs have usually been calculated numerically or by *ad hoc* methods for special cases (Bel et al. 2010), or by more general methods for which closed-form solutions are not readily found (Shaevitz et al. 2005; Chemla et al. 2008). Even when rational functions emerge in such treatments, their manifest positivity may not always be visible (Bel et al. 2010). The use of Terrell Hill's method for calculating mean FPTs (Hill 1988), as in Qureshi et al. (2023); Guntoro et al. (2025), is an exception to which we will return below.

In this paper, we introduce a systematic graph-theoretic treatment of FPTs for Markov processes, and thereby extend the scope of the linear framework from the steady state to the transient regime. To do this, we exploit a generalisation of the MTT, the *All-Minors Matrix-Tree theorem* (AMMTT), stated here as Theorem 1. The AMMTT uses *spanning forests*, which generalise the spanning trees of the classical MTT described above. The AMMTT enables us to provide closed-form solutions for the moments of any FPT distribution, either unconditional or conditional, along with related quantities such as splitting probabilities, as manifestly positive rational algebraic functions of the edge labels.

The AMMTT has been used to derive formulas for mean FPTs (Chebotarev 2007; Pitman and Tang 2018) and for splitting probabilities (Pitman and Tang 2018) in discrete-time Markov chains, but not, to our knowledge, for higher moments and conditional FPTs of the continuous-time Markov processes studied here.

Generalisations of the classical work on FPTs of Hill (1988) and of Kac (1947) can be shown to follow from the results presented here. For reasons of space, we discuss these connections in a sequel to the present paper (Nam and Gunawardena 2025). The main application of our results so far has been to the analysis of a tradeoff between accuracy and speed in the CRISPR–Cas9 system, which will also be presented elsewhere (Nam et al. 2025).

We begin with some preliminary definitions and results; a summary of the various symbols and notation that we use is also provided in Table 1. The AMMTT is introduced in Theorem 1. We deduce an important consequence of this in Proposition 2, which shows how the inverse of a principal submatrix of the Laplacian matrix may be calculated in terms of spanning forests. This technical result is a key ingredient in what follows. We then focus on the special case of graphs in which the target vertices for FPTs are the “terminal” vertices in the graph, in a sense explained below. This special case allows a differential equation to be formulated for the time evolution of the probabilities of reaching the target vertices (Eqn. 34). The Laplace transform of this *adjoint master equation* allows the moments of the appropriate FPT distribution to be calculated in a standard way (Eqn. 42). This moment calculation requires the inverse of a principal submatrix of the Laplacian matrix (Eqn. 44), which Proposition 2 describes in terms of spanning forests (Eqn. 47). It is then straightforward to deduce our main results: Theorem 4 for the splitting probabilities, Theorem 5 for the moments of the conditional FPT to a specific target vertex, and Theorem 6 for the moments of the unconditional FPT to any target vertex. We then explain how FPTs may be calculated in a general graph by modifying the graph so that it falls into the special case on which the previous analysis was based. Finally, we work through some examples, of pipeline graphs (Fig. 4) and of a butterfly graph (Fig. 5), both taken from the literature, to illustrate how the methods introduced here can be used in practice.

The work described here was begun in K-MN’s Harvard PhD thesis (Nam 2021), under the supervision of JG. The present paper uses a more direct proof strategy and fills in some details that were omitted in Nam (2021). An overview of this material, without proofs but with examples, appeared in Nam and Gunawardena (2023), which should be consulted for further background.

2 Results

2.1 Background and preliminaries

2.1.1 Graphs and spanning forests

We begin by establishing some notation and terminology. See the Table in the Appendix for a summary. We will continue to use the notation in the Introduction, with decorations where necessary, as in $i \rightarrow_G j$, to specify which graph, G , is intended. We assume that all graphs are connected when edge directions are ignored, so that they do not break up into separated pieces.

Given $i \in \mathcal{V}(G)$, let $V_{\text{out}}(i) \subseteq \mathcal{V}(G)$ denote the set of vertices to which i has an edge, $V_{\text{out}}(i) = \{j \in \mathcal{V}(G) : i \rightarrow j\}$, and let $V_{\text{in}}(i) \subseteq \mathcal{V}(G)$ denote the set of vertices which have an edge to i , $V_{\text{in}}(i) = \{j \in \mathcal{V}(G) : j \rightarrow i\}$.

Let $A, B \subseteq \mathcal{V}(G) = \{1, \dots, N\}$ be vertex subsets. We denote the size of A by $\#A$ and the complement of A by $\bar{A} = \mathcal{V}(G) \setminus A$. Given a matrix $\mathbf{M} \in \mathbb{R}^{N \times N}$, we denote by $\mathbf{M}_{[A, B]} \in \mathbb{R}^{\#A \times \#B}$ the submatrix consisting of the rows indexed by A and the columns indexed by B . We use \mathbf{I} to denote the identity matrix, leaving its dimensions to be inferred from context.

Let H be a subgraph of G . The *weight* of H , denoted $w(H)$, is the product of the edge labels in H ,

$$w(H) = \prod_{i \rightarrow j \in H} \ell(i \rightarrow j).$$

We follow the standard convention that empty sums evaluate to zero and empty products evaluate to one. It follows that the weight of an edgeless graph is one. If \mathcal{H} is a set of subgraphs of G , then $w(\mathcal{H})$ is the sum of the weights of the subgraphs in \mathcal{H} ,

$$w(\mathcal{H}) = \sum_{H \in \mathcal{H}} w(H).$$

A subgraph F of G is a *spanning forest* of G if it contains every vertex in G (“spanning”), it contains no cycles of edges when edge directions are ignored (“forest”), and each vertex has at most one outgoing edge (which orients the forest). Since F has no cycles, it must have at least one vertex with no outgoing edges. Such vertices are the *roots* of F ; let $\emptyset \neq \mathcal{R}(F) \subseteq \mathcal{V}(G)$ denote the set of all roots of F . If $\mathcal{R}(F) = \{i\}$, then F is a *spanning tree rooted at i* , as defined in the Introduction. Note that a spanning forest F is a disjoint union of $\#\mathcal{R}(F)$ trees, and must have exactly $N - \#\mathcal{R}(F)$ edges. We denote the set of all spanning forests of G rooted at A by $\Phi_A(G)$.

Given two vertices, $i, j \in \mathcal{V}(G)$, we say that i *leads to* j , denoted $i \rightsquigarrow j$, if there is a directed path of edges from i to j : $i = i_1 \rightarrow \dots \rightarrow i_k = j$. There is always a trivial path from any vertex to itself, so that $i \rightsquigarrow i$. Furthermore, if F is a spanning forest of G , then any $i \in \mathcal{V}(G)$ has a unique directed path in F , with no repeated vertices, to exactly one root, namely the root of the tree containing i .

To accompany the vertex subsets $V_{\text{out}}(i)$ and $V_{\text{in}}(i)$ coming from edges adjoining $i \in \mathcal{V}(G)$, as defined above, we denote by $\mathcal{V}_{\text{out}}(i)$ and $\mathcal{V}_{\text{in}}(i)$ the corresponding vertex subsets for directed paths that lead from i , $\mathcal{V}_{\text{out}}(i) = \{j \in \mathcal{V}(G) : i \rightsquigarrow j\}$, and lead to i , $\mathcal{V}_{\text{in}}(i) = \{j \in \mathcal{V}(G) : j \rightsquigarrow i\}$, respectively. Note that $i \notin V_{\text{out}}(i)$ but $i \in \mathcal{V}_{\text{out}}(i) \neq \emptyset$; similarly, $i \notin V_{\text{in}}(i)$ but $i \in \mathcal{V}_{\text{in}}(i) \neq \emptyset$. It is possible that one of $V_{\text{out}}(i)$ or $V_{\text{in}}(i)$ is empty but not both, unless G consists solely of the vertex i .

Given $A, B \subseteq \mathcal{V}(G)$ with $\#A = \#B$, define $\Phi_{B \rightarrow A}(G) \subseteq \Phi_A(G)$ as the set of spanning forests of G rooted at A in which each vertex in B has a path to a *distinct* root in A . So, $\Phi_{\{1,2\} \rightarrow \{3,4\}}(G)$ is the set of all spanning forests of G rooted at $\{3, 4\}$ in which either $1 \rightsquigarrow 3, 2 \rightsquigarrow 4$ or $1 \rightsquigarrow 4, 2 \rightsquigarrow 3$. Since every root has a trivial path to itself, it follows that,

$$\Phi_{A \rightarrow A}(G) = \Phi_A(G) \quad \text{for all } A \subseteq \mathcal{V}(G). \quad (6)$$

If $A = \{j\}$, then each $F \in \Phi_A(G)$ is a spanning tree and every vertex has a path to j . Hence, $\Phi_{\{i\} \rightarrow \{j\}}(G) = \Phi_{\{j\}}(G)$ for all $i, j \in \mathcal{V}(G)$. A spanning forest $F \in \Phi_{B \rightarrow A}(G)$ induces a bijection, $\chi_F : B \rightarrow A$, where $\chi_F(i) \in A$ is the unique root to which there is a path from i in F . Since a root always has a path to itself, $\chi_F(i) = i$ for all $i \in A \cap B$. In particular, if $A = B$, then $\chi_F : A \rightarrow A$ is the identity for all $F \in \Phi_{A \rightarrow A}(G) = \Phi_A(G)$ (Eqn. 6).

We say that i is *strongly connected to* j if $i \rightsquigarrow j$ and $j \rightsquigarrow i$. Strong connectedness is an equivalence relation on $\mathcal{V}(G)$, and the corresponding equivalence classes are the *strongly connected components* (SCCs) of G . We denote the set of SCCs of G by $\mathcal{C}(G)$. Given $C, D \in \mathcal{C}(G)$, we say that C *precedes* D , denoted $C \prec D$, if $i \rightsquigarrow j$ for some $i \in C$ and some $j \in D$. Since C and D are each strongly connected, this implies that $i \rightsquigarrow j$ for any $i \in C$ and any $j \in D$. In particular, if $C \prec D$ and $D \prec C$, then $C = D$. Hence, the relation $C \leq D$, defined as $C = D$ or $C \prec D$, defines a partial order on $\mathcal{C}(G)$. The maximal elements under this partial order are the *terminal SCCs* of G , collectively denoted by $T(G) \subseteq \mathcal{C}(G)$. Note that $V_{\text{out}}(i) = \emptyset$ if, and only if, $\{i\}$ is a singleton terminal SCC.

The terminal SCCs determine the minimal number of roots in a spanning forest. Suppose that G has T terminal SCCs, C_1, \dots, C_T . Let $\mathcal{Q}(G) = C_1 \cup \dots \cup C_T$ denote all the vertices within the terminal SCCs, and let $\mathcal{N}(G) = \mathcal{V}(G) \setminus \mathcal{Q}(G)$ be all the vertices within the non-terminal SCCs. Since there can be no directed paths between terminal SCCs, any spanning forest F of G must have at least one root in each terminal SCC. Hence, $\#\mathcal{R}(F) \geq T$. If G is strongly connected, so that $T = 1$, then it is clear that the preceding inequality is sharp. If G is not strongly connected, then, because G is connected, it must be that $\mathcal{N}(G) \neq \emptyset$. Each $i \in \mathcal{N}(G)$ must lead to some $j \in \mathcal{Q}(G)$ and, because G is connected, every $j \in \mathcal{Q}(G)$ must have some vertex in $\mathcal{N}(G)$ that leads to it. It follows that, provided $A \subseteq \mathcal{V}(G)$ contains at least one vertex from each terminal SCC, so that $A \cap C_k \neq \emptyset$ for $1 \leq k \leq T$, then there is a spanning forest F with $\mathcal{R}(F) = A$. In particular, there is always a spanning forest with exactly T roots, so that the inequality $\#\mathcal{R}(F) \geq T$ is sharp.

2.1.2 The All-Minors Matrix-Tree theorem

The Matrix-Tree theorem (MTT) allows the steady-state probabilities of a graph to be expressed in terms of the edge labels, as shown by Eqns. 4 and 5. Here, we introduce the more general All-Minors Matrix-Tree theorem (AMMTT), which we use to express FPTs in the same way. This class of theorems reveals that Laplacian matrices, $\mathcal{L}(G)$, have remarkable properties: the determinants of submatrices, or the minors, of $\mathcal{L}(G)$ exhibit extensive sign cancellations and can be expressed in terms of the weights of spanning forests of G . We saw this for steady-state probabilities and spanning trees in Eqn. 4.

A bijection, $\sigma : A \rightarrow A$, is a *permutation* on A , whose sign, $\text{sgn}(\sigma)$, is given by $(-1)^p$, where p is the number of transpositions in any decomposition of σ into a product of transpositions (Herstein 1975). Recall that $\mathcal{L}(G)_{[\overline{B}, \overline{A}]}$ is the submatrix of $\mathcal{L}(G)$ in which the rows indexed by B and the columns indexed by A have been removed. If $\#A = \#B$, then the determinant of such a submatrix is a *minor* of $\mathcal{L}(G)$.

Theorem 1 (All-Minors Matrix-Tree theorem) *Let G be a graph, and let $\emptyset \neq A, B \subsetneq \mathcal{V}(G)$ be proper, non-empty vertex subsets of the same size, $\#A = \#B = k$. Write A and B in ascending order as $A = \{a_1 < \dots < a_k\}$ and $B = \{b_1 < \dots < b_k\}$ and let $\eta : A \rightarrow B$ be the bijection $\eta(a_i) = b_i$ for $i = 1, \dots, k$. Then,*

$$\det \mathcal{L}(G)_{[\overline{B}, \overline{A}]} = (-1)^{N-k+\sum_{a \in A} a + \sum_{b \in B} b} \sum_{F \in \Phi_{B \rightarrow A}(G)} \text{sgn}(\chi_F \circ \eta) w(F), \quad (7)$$

where $\chi_F : B \rightarrow A$ is the bijection defined after Eqn. 6.

If we take $A = \{i\}$ and $B = \{j\}$, the AMMTT reduces to the MTT (Mirzaev and Gunawardena 2013), from which Eqn. 4 is derived. The MTT has a long history that goes back to the 19th century work of Kirchhoff—see Mirzaev and Gunawardena (2013) for details—but the first report of the AMMTT is due to Fiedler and Sedláček (1958) in the Czech mathematical literature; subsequent proofs were given by Chaiken (1982) and Moon (1994).

As an example of Theorem 1 in action, consider the graph, K , in Fig. 1A, whose Laplacian matrix, $\mathcal{L}(K)$, is also shown there. The submatrix $\mathcal{L}(K)_{[\overline{\{2,4\}}, \overline{\{1,4\}}]}$ is given by

$$\begin{bmatrix} k_2 & 0 & k_{11} \\ k_3 - k_5 - k_6 - k_7 & 0 & \\ k_4 & k_7 & -k_{11} \end{bmatrix}. \quad (8)$$

According to Theorem 1, the minor, $\det \mathcal{L}(K)_{[\overline{\{2,4\}}, \overline{\{1,4\}}]}$, may be calculated in terms of the spanning forests in $\Phi_{\{2,4\} \rightarrow \{1,4\}}(K)$, which are enumerated in Fig. 2A. Since vertex 4 is common to both $\{1, 4\}$ and $\{2, 4\}$, the map $\chi_F : \{2, 4\} \rightarrow \{1, 4\}$ is the same for all $F \in \Phi_{\{2,4\} \rightarrow \{1,4\}}(K)$, and this map is also the inverse of the map $\eta : \{1, 4\} \rightarrow \{2, 4\}$ defined in Theorem 1 (Fig. 2B). It follows that $\text{sgn}(\chi_F \circ \eta) = 1$ for all $F \in \Phi_{\{2,4\} \rightarrow \{1,4\}}(K)$. Furthermore, $N - k = 5 - 2 = 3$ and $(1 + 4) + (2 + 4) = 11$

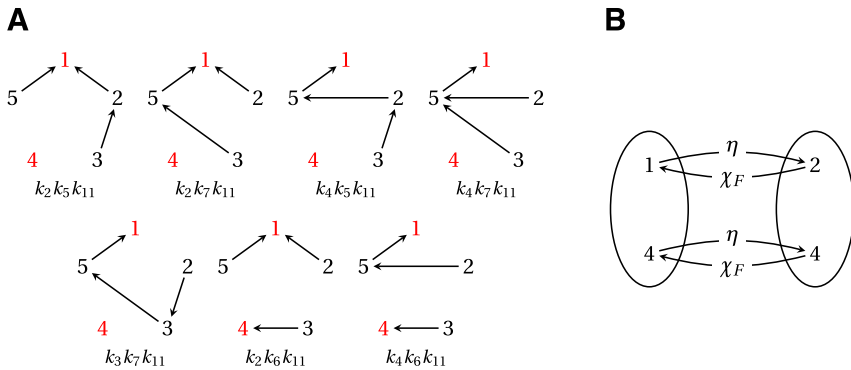


Fig. 2 (A) The seven spanning forests in $\Phi_{[2,4] \rightarrow \{1,4\}}(K)$, where K is the example graph in Fig. 1 A, with their corresponding weights. Roots are shown in red. (B) Schematic of the maps between $\{1, 4\}$ and $\{2, 4\}$ for each forest F in panel A. In this case, χ_F is the same for all F and is the inverse of η

combine to yield a sign of $(-1)^{14} = 1$. It follows from Theorem 1 that,

$$\det \mathcal{L}(K)_{[\overline{[2,4]}, \overline{[1,4]}}] = k_{11} (k_2 k_5 + k_2 k_7 + k_4 k_5 + k_4 k_7 + k_3 k_7 + k_2 k_6 + k_4 k_6),$$

which may easily be confirmed from Eqn. 8.

As a further example of how Theorem 1 works, consider the extreme case in which $A = \{i\}$ and $B = \{j\}$ with $i \neq j$, so that $k = N - 1$. Then,

$$\mathcal{L}(G)_{[\overline{B}, \overline{A}]} = \mathcal{L}(G)_{j,i},$$

so that (Eqn. 2),

$$\det \mathcal{L}(G)_{[\overline{B}, \overline{A}]} = \begin{cases} \ell(i \rightarrow j) & \text{if } i \rightarrow j \\ 0 & \text{if } i \nrightarrow j. \end{cases} \quad (9)$$

To see what Theorem 1 says in this context, note that a spanning forest in $\Phi_{B \rightarrow A}(G)$ must contain only a single edge, which must evidently be $i \rightarrow j$, should the edge exist. Hence, if $i \nrightarrow j$, then $\Phi_{B \rightarrow A}(G) = \emptyset$ and Theorem 1 correctly gives the second case in Eqn. 9. If, however, $i \rightarrow j$ does exist in G , then Theorem 1 says that $\det \mathcal{L}(G)_{[\overline{B}, \overline{A}]}$ is given by

$$(-1)^{N-(N-1)+\sum_{a \in A} a + \sum_{b \in B} b} \operatorname{sgn}(\chi_F \circ \eta) \ell(i \rightarrow j).$$

We now examine how the two signs work out. First, note that the parity of $1 + \sum_{a \in A} a + \sum_{b \in B} b = 1 + 2(1 + \dots + N) - i - j$ is the same as that of $1 - i - j$. As for the sign of $\chi_F \circ \eta$, assume that $i < j$ and let us write A and B out as a $2 \times (N - 1)$ array of entries,

$$\begin{aligned} A : & 1, \dots, i-1, \underline{i}, i+1, \dots, j-1, j, j+1, \dots, N \\ B : & 1, \dots, i-1, i, i+1, \dots, j-1, \underline{j}, j+1, \dots, N, \end{aligned}$$

where \underline{i} and \underline{j} signify their absence from the array. It is clear that χ_F takes i to j and is the identity elsewhere, while from the array above, we see that η takes m to $m - 1$ if $i < m \leq j$ and is the identity elsewhere. It follows that $\chi_F \circ \eta$ is the permutation,

$$\left(\begin{array}{cccccccc} 1 & \cdots & i-1 & i+1 & i+2 & \cdots & j-1 & j & j+1 & \cdots & N \\ 1 & \cdots & i-1 & j & i+1 & \cdots & j-2 & j-1 & j+1 & \cdots & N \end{array} \right),$$

which can be written in cycle notation as $(j, j-1, j-2, \dots, i+1)$. This cycle has length $j-i$, and therefore can be written as a product of $j-i-1$ transpositions, so that

$$\operatorname{sgn}(\chi_F \circ \eta) = (-1)^{j-i-1}.$$

Accordingly, the total sign is

$$(-1)^{1-i-j+j-i-1} = (-1)^{-2i} = 1.$$

Theorem 1 therefore tells us that, if $i \rightarrow j$ does exist in G , then $\det \mathcal{L}(G)_{[\overline{B}, \overline{A}]} = \ell(i \rightarrow j)$, which agrees with the first case in Eqn. 9. We hope these examples illustrate the workings of Theorem 1.

2.1.3 Preliminary results

In this section, we state and prove some basic results that follow from Theorem 1, upon which we will rely throughout the paper.

We need some machinery to accomplish this. It will be convenient to define, for any finite, ordered set, $S = \{s_1 < \cdots < s_k\}$, the ordering bijection,

$$\theta_S : \{1, \dots, k\} \rightarrow S, \quad \theta_S(i) = s_i.$$

This bijection is helpful to keep track of the rows and columns of submatrices, which acquire a different indexing to those of the original matrix. Let $A \subseteq \mathcal{V}(G)$ be a subset with $\#A = k$, and let $i, j \in \{1, \dots, k\}$. Then the (i, j) -th entry of the $k \times k$ submatrix, $\mathcal{L}(G)_{[A, A]}$, corresponds to the $(\theta_A(i), \theta_A(j))$ -th entry of $\mathcal{L}(G)$,

$$(\mathcal{L}(G)_{[A, A]})_{i, j} = \mathcal{L}(G)_{\theta_A(i), \theta_A(j)}. \quad (10)$$

Alternatively, if $i, j \in A$, then we can invert the correspondence in Eqn. 10 to see that,

$$\mathcal{L}(G)_{i, j} = (\mathcal{L}(G)_{[A, A]})_{\theta_A^{-1}(i), \theta_A^{-1}(j)}. \quad (11)$$

We now consider what happens when we re-index the vertices of G by a permutation, $\sigma : \mathcal{V}(G) \rightarrow \mathcal{V}(G)$. What this means is that the edge, $j \rightarrow_G i$, becomes the edge

$\sigma(j) \rightarrow_{G^\sigma} \sigma(i)$ and retains the same label, $\ell(j \rightarrow_G i) = \ell(\sigma(j) \rightarrow_{G^\sigma} \sigma(i))$. It then follows from Eqn. 2 that,

$$\mathcal{L}(G)_{i,j} = \mathcal{L}(G^\sigma)_{\sigma(i),\sigma(j)}, \quad (12)$$

so that the rows and columns of $\mathcal{L}(G)$ are permuted by σ to form $\mathcal{L}(G^\sigma)$. If we have the same subset $A \subseteq \mathcal{V}(G)$ as above, it will give rise under the permutation σ to the subset $\sigma(A) \subseteq \mathcal{V}(G^\sigma)$. If $i, j \in \sigma(A)$, then applying Eqn. 11 to G^σ , we see that,

$$\mathcal{L}(G^\sigma)_{i,j} = (\mathcal{L}(G^\sigma)_{[\sigma(A),\sigma(A)]})_{\theta_{\sigma(A)}^{-1}(i),\theta_{\sigma(A)}^{-1}(j)}. \quad (13)$$

Now let us return to choosing $i, j \in \{1, \dots, k\}$ and consider the sequence of bijections,

$$\{1, \dots, k\} \xrightarrow{\theta_A} A \xrightarrow{\sigma} \sigma(A) \xrightarrow{\theta_{\sigma(A)}^{-1}} \{1, \dots, k\},$$

which amount to a form of conjugation of σ by θ_A . Let us call the composition of these bijections $\tau_{\sigma,A} : \{1, \dots, k\} \rightarrow \{1, \dots, k\}$,

$$\tau_{\sigma,A} = \theta_{\sigma(A)}^{-1} (\sigma (\theta_A)). \quad (14)$$

Applying Eqns. 10, 12 and 13 in sequence, we then see that,

$$\begin{aligned} (\mathcal{L}(G)_{[A,A]})_{i,j} &= \mathcal{L}(G)_{\theta_A(i),\theta_A(j)} \\ &= \mathcal{L}(G^\sigma)_{\sigma(\theta_A(i)),\sigma(\theta_A(j))} \\ &= (\mathcal{L}(G^\sigma)_{[\sigma(A),\sigma(A)]})_{\tau_{\sigma,A}(i),\tau_{\sigma,A}(j)}. \end{aligned} \quad (15)$$

The impact of the permutation σ on spanning forests is less intricate. If $A \subseteq \mathcal{V}(G)$ and $F \in \Phi_A(G)$, then applying σ to the vertices of F yields a re-indexed forest, $\sigma^*(F) \in \Phi_{\sigma(A)}(G^\sigma)$, with the same weight, $w(\sigma^*(F)) = w(F)$. If we have a second subset, $B \subseteq \mathcal{V}(G)$, with $\#B = \#A$, then the map, $\sigma^* : \Phi_A(G) \rightarrow \Phi_{\sigma(A)}(G^\sigma)$, which is evidently a bijection, restricts to a bijection $\sigma^* : \Phi_{B \rightarrow A}(G) \rightarrow \Phi_{\sigma(B) \rightarrow \sigma(A)}(G^\sigma)$. It follows that,

$$w(\Phi_{B \rightarrow A}(G)) = w(\Phi_{\sigma(B) \rightarrow \sigma(A)}(G^\sigma)). \quad (16)$$

Finally, we review some linear algebra. To derive the main results in this paper, we shall primarily invoke Theorem 1 in combination with Cramer's rule. Recall that, if \mathbf{M} is an $N \times N$ matrix, then the *adjugate* of \mathbf{M} , which we denote by $\text{adj } \mathbf{M}$, is the $N \times N$ matrix with entries,

$$(\text{adj } \mathbf{M})_{i,j} = (-1)^{i+j} \det \mathbf{M}_{[\overline{\{j\}}, \overline{\{i\}}]}. \quad (17)$$

Cramer's rule then states that,

$$\mathbf{M}(\text{adj } \mathbf{M}) = (\text{adj } \mathbf{M})\mathbf{M} = (\det \mathbf{M})\mathbf{I}. \quad (18)$$

It follows that, when $\det \mathbf{M} \neq 0$, $\mathbf{M}^{-1} = (\text{adj } \mathbf{M})/(\det \mathbf{M})$.

This preparation leads us to one of the central technical results of the paper.

Proposition 2 *Let G be a graph and let $\emptyset \neq A \subsetneq \mathcal{V}(G)$ be a subset of k vertices, where $0 < k < N$. Then $\mathcal{L}(G)_{[\overline{A}, \overline{A}]}$ is invertible if, and only if, $\Phi_A(G)$ is non-empty. Furthermore, if $\mathcal{L}(G)_{[\overline{A}, \overline{A}]}$ is invertible, the inverse is given by,*

$$\left(\mathcal{L}(G)_{[\overline{A}, \overline{A}]}\right)_{i,j}^{-1} = -\frac{w(\Phi_{A \cup \{\theta_{\overline{A}}(j)\} \rightarrow A \cup \{\theta_{\overline{A}}(i)\}}(G))}{w(\Phi_A(G))},$$

for $i, j \in \{1, \dots, N - k\}$.

Proof The first statement follows by setting $A = B$ in Eqn. 7 and recalling Eqn. 6. To prove the second statement, we first show that it holds when $A = \{N - k + 1, \dots, N\}$, in which case $\theta_{\overline{A}}(i) = i$ for all $i = 1, \dots, N - k$. Then Eqn. 18 says that,

$$\left(\mathcal{L}(G)_{[\overline{A}, \overline{A}]}\right)^{-1} = \left(\frac{1}{\det \mathcal{L}(G)_{[\overline{A}, \overline{A}]}}\right) \text{adj} \left(\mathcal{L}(G)_{[\overline{A}, \overline{A}]}\right), \quad (19)$$

and Eqn. 17 tells us that,

$$\text{adj} \left(\mathcal{L}(G)_{[\overline{A}, \overline{A}]}\right)_{i,j} = (-1)^{i+j} \det \mathcal{L}(G)_{[\overline{A \cup \{j\}}, \overline{A \cup \{i\}}]}. \quad (20)$$

At this point, we can appeal to Theorem 1 to get the minor on the right as,

$$(-1)^{N-k-1+i+j+2\sum_{a \in A} a} \sum_{F \in \Phi_{A \cup \{j\} \rightarrow A \cup \{i\}}(G)} \text{sgn}(\chi_F \circ \eta) w(F). \quad (21)$$

Since $A = \{N - k + 1, \dots, N\}$ and $i, j < N - k + 1$, the bijection $\eta : A \cup \{i\} \rightarrow A \cup \{j\}$ takes i to j and is the identity elsewhere,

$$\eta(q) = \begin{cases} j & \text{if } q = i \\ q & \text{if } q \neq i. \end{cases}$$

To determine the signs in Eqn. 21, the parity term contributes $(-1)^{N-k-1+i+j}$. As for the permutation term, for any forest $F \in \Phi_{A \cup \{j\} \rightarrow A \cup \{i\}}(G)$, χ_F must take j to i and be the identity elsewhere. Hence, $\chi_F \circ \eta$ is the identity on $A \cup \{i\}$ for each F . Therefore, the minor given by Eqn. 21 simplifies to,

$$\det \mathcal{L}(G)_{[\overline{A \cup \{j\}}, \overline{A \cup \{i\}}]} = (-1)^{N-k-1+i+j} w(\Phi_{A \cup \{j\} \rightarrow A \cup \{i\}}(G)).$$

We can combine this with Eqn. 20 to get,

$$\text{adj} \left(\mathcal{L}(G)_{[\overline{A}, \overline{A}]}\right)_{i,j} = (-1)^{N-k-1} w(\Phi_{A \cup \{j\} \rightarrow A \cup \{i\}}(G)).$$

As for $\det \mathcal{L}(G)_{[\bar{A}, \bar{A}]}$, Theorem 1 tells us that,

$$\det \mathcal{L}(G)_{[\bar{A}, \bar{A}]} = (-1)^{N-k} w(\Phi_A(G)).$$

Incorporating these calculations into Eqn. 19, we obtain the required conclusion for the case where $A = \{N - k + 1, \dots, N\}$.

We now consider the case of a general subset, A , which may not be indexed as previously. Write A and \bar{A} in ascending order, as $A = \{a_1 < \dots < a_k\}$ and $\bar{A} = \{\bar{a}_1 < \dots < \bar{a}_{N-k}\}$, and let $B = \{N - k + 1, \dots, N\}$. Let G^σ be the graph obtained by re-indexing the vertices according to the permutation that maps the current situation onto the one considered previously,

$$\sigma = \begin{pmatrix} \bar{a}_1 & \dots & \bar{a}_{N-k} & a_1 & \dots & a_k \\ 1 & \dots & N-k & N-k+1 & \dots & N \end{pmatrix}.$$

We can now make use of the preparation we did above and apply Eqn. 15 to see that,

$$\left(\mathcal{L}(G)_{[\bar{A}, \bar{A}]} \right)_{i,j} = \left(\mathcal{L}(G^\sigma)_{[\sigma(\bar{A}), \sigma(\bar{A})]} \right)_{\tau_{\sigma, \bar{A}}(i), \tau_{\sigma, \bar{A}}(j)}, \quad (22)$$

where we have used the permutation defined in Eqn. 14 for \bar{A} ,

$$\tau_{\sigma, \bar{A}} = \theta_{\sigma(\bar{A})}^{-1} (\sigma(\theta_{\bar{A}})) : \{1, \dots, N-k\} \rightarrow \{1, \dots, N-k\}.$$

Since $\sigma(A) = B$ and $\sigma(\bar{A}) = \bar{B} = \{1, \dots, N-k\}$, it follows that $\theta_{\sigma(\bar{A})}$ is just the identity on $\{1, \dots, N-k\}$. Furthermore, $\sigma(\bar{a}_i) = i$, which means that,

$$\theta_{\bar{A}}(i) = \bar{a}_i = \sigma^{-1}(i) \quad \text{for } i = 1, \dots, N-k. \quad (23)$$

It follows that, on \bar{A} , $\sigma = \theta_{\bar{A}}^{-1}$. Putting these pieces together, it then follows from Eqn. 14 that $\tau_{\sigma, \bar{A}}$ is the identity on $\{1, \dots, N-k\}$. Hence, Eqn. 22 becomes,

$$\left(\mathcal{L}(G)_{[\bar{A}, \bar{A}]} \right)_{i,j} = \left(\mathcal{L}(G^\sigma)_{[\bar{B}, \bar{B}]} \right)_{i,j},$$

so that we may, in fact, just equate these two matrices, $\mathcal{L}(G)_{[\bar{A}, \bar{A}]} = \mathcal{L}(G^\sigma)_{[\bar{B}, \bar{B}]}$. Since $B = \{N - k + 1, \dots, N\}$, we can now appeal to the first part of the proof to deduce that,

$$\left(\mathcal{L}(G^\sigma)_{[\bar{B}, \bar{B}]} \right)_{i,j}^{-1} = - \frac{w(\Phi_{B \cup \{j\} \rightarrow B \cup \{i\}}(G^\sigma))}{w(\Phi_B(G^\sigma))}.$$

It follows using Eqns. 16 and 23 that,

$$w(\Phi_{A \cup \{\theta_{\bar{A}}(j)\} \rightarrow A \cup \{\theta_{\bar{A}}(i)\}}(G)) = w(\Phi_{\sigma(A) \cup \{\sigma(\theta_{\bar{A}}(j))\} \rightarrow \sigma(A) \cup \{\sigma(\theta_{\bar{A}}(i))\}}(G^\sigma))$$

$$= w(\Phi_{B \cup \{j\} \rightarrow B \cup \{i\}}(G^\sigma)),$$

and that,

$$w(\Phi_A(G)) = w(\Phi_{\sigma(A)}(G^\sigma)) = w(\Phi_B(G^\sigma)).$$

Putting all these pieces together, we finally obtain,

$$\begin{aligned} \left(\mathcal{L}(G)_{[\bar{A}, \bar{A}]}\right)_{i,j}^{-1} &= \left(\mathcal{L}(G^\sigma)_{[\bar{B}, \bar{B}]}\right)_{i,j}^{-1} \\ &= -\frac{w(\Phi_{B \cup \{j\} \rightarrow B \cup \{i\}}(G^\sigma))}{w(\Phi_B(G^\sigma))} \\ &= -\frac{w(\Phi_{A \cup \{\theta_A^{-1}(j)\} \rightarrow A \cup \{\theta_A^{-1}(i)\}}(G))}{w(\Phi_A(G))}, \end{aligned}$$

as required. \square

The following elementary result on spanning forests will also be helpful in what follows. Choose $A \subseteq \mathcal{V}(G)$ and $i, j \in \bar{A}$ such that the set, $\Phi_{A \cup \{i\} \rightarrow A \cup \{j\}}(G)$, of spanning forests rooted at $A \cup \{j\}$ with a path from i to j is non-empty, and let F be such a spanning forest. If G contains an edge $j \rightarrow z$ for some other root $z \in A$, then adjoining this edge to F yields a spanning forest that is rooted at A and contains a path from i to z . That is, the new spanning forest lies in $\Phi_{(A \setminus \{z\}) \cup \{i\} \rightarrow A}(G)$. Conversely, if $F \in \Phi_{(A \setminus \{z\}) \cup \{i\} \rightarrow A}(G)$, then F must contain a path from i to z . Since i is itself not a root, this path must possess at least one edge; if the last edge on this path is $j \rightarrow z$, then $j \notin A$ and omitting this edge creates a forest in $\Phi_{A \cup \{i\} \rightarrow A \cup \{j\}}(G)$. These procedures are evidently inverse to each other.

Lemma 3 *Given $A \subseteq \mathcal{V}(G)$, $i \in \bar{A}$ and $z \in A$, the map,*

$$\bigcup_{j \in V_{\text{in}}(z) \setminus A} \Phi_{A \cup \{i\} \rightarrow A \cup \{j\}}(G) \rightarrow \Phi_{(A \setminus \{z\}) \cup \{i\} \rightarrow A}(G),$$

that adjoins the edge $j \rightarrow z$ is a bijection.

It follows immediately from Lemma 3 that, under the same assumptions,

$$\sum_{j \in V_{\text{in}}(z) \setminus A} w(\Phi_{A \cup \{i\} \rightarrow A \cup \{j\}}(G)) \ell(j \rightarrow z) = w(\Phi_{(A \setminus \{z\}) \cup \{i\} \rightarrow A}(G)). \quad (24)$$

2.2 FPTs in graphs with singleton terminal SCCs

With the above preliminaries in hand, we can embark on calculating FPTs. This will require several steps. We will derive the basic time-evolution equation, use the Laplace transform to calculate moment-like quantities and then obtain formulas for the splitting probabilities and the moments of the conditional FPT distribution.

2.2.1 Definitions and notation

To simplify the argument, we will assume that each terminal SCC is a singleton, which we will refer to as a *terminal vertex*. Let Z denote the set of terminal vertices. More general situations can be reduced to this special case, as we will explain subsequently.

Let $X(\cdot)$ be the Markov process associated with G , and choose some starting vertex $i \in \mathcal{V}(G)$ and some terminal vertex $z \in Z$. We will argue somewhat intuitively here, to avoid getting into measure-theoretic complexities; see Anderson (1991); Norris (1997); Serfozo (2009) for the rigorous underpinnings. We consider stochastically generating a large finite ensemble, E , of unbounded trajectories of $X(\cdot)$ each starting from i , so that $X(0) = i$. With probability one, each trajectory will reach some terminal vertex in Z and stay there from then on. Let $A_{z,t} \subseteq E$ be the subset of those trajectories that have reached z by time t . The probability of reaching z from i by time t is then given by,

$$p_{i,z}(t) = \Pr[X(t') = z \text{ for some } 0 \leq t' \leq t \mid X(0) = i] \approx \frac{\#A_{z,t}}{\#E}, \quad (25)$$

where the approximation becomes exact in the limit of an infinite ensemble. Since $A_{z,t}$ clearly enlarges with t , $p_{i,z}(t)$ must be non-decreasing, $p_{i,z}(t_1) \leq p_{i,z}(t_2)$ if $t_1 \leq t_2$, so that it resembles a cumulative distribution function. However, it is not a valid cumulative distribution function over t because it is not correctly normalised: the quantity $\lim_{t \rightarrow \infty} p_{i,z}(t)$, which we denote for convenience as $p_{i,z}(\infty)$, is not 1 but, rather, the *splitting probability*, $\pi_{i,z}$, of eventually reaching z from i . Let $B_z \subseteq E$ be the subset of trajectories that eventually reach z . Then,

$$\pi_{i,z} = \Pr[X(t) = z \text{ for some } t \geq 0 \mid X(0) = i] \approx \frac{\#B_z}{\#E}. \quad (26)$$

For sufficiently large t , $A_{z,t} = B_z$, so it follows from Eqn. 25 that $p_{i,z}(\infty) = \pi_{i,z}$. This splitting probability is one of the quantities that we want to calculate.

The other quantities of interest arise from the conditional FPT distribution, which we can approach through the probability, $p_{c,i,z}(t)$, of reaching z from i by time t , conditioned on reaching z eventually. For this probability to be well-defined, it is necessary that $i \in \mathcal{V}_{\text{in}}(z)$, so that $B_z \neq \emptyset$. With that proviso,

$$\begin{aligned} p_{c,i,z}(t) &= \Pr[X(t') = z \text{ for some } 0 \leq t' \leq t \mid \\ &\quad X(0) = i \text{ and } X(u) = z \text{ for some } u > 0] \\ &\approx \frac{\#A_{z,t}}{\#B_z}. \end{aligned} \quad (27)$$

It follows from Eqns. 26 and 27 that, as long as $i \in \mathcal{V}_{\text{in}}(z)$,

$$p_{c,i,z}(t) = \frac{p_{i,z}(t)}{\pi_{i,z}}. \quad (28)$$

In contrast to $p_{i,z}(t)$, $p_{c,i,z}(t)$ is a cumulative distribution function that is correctly normalised with respect to t , with $p_{c,i,z}(\infty) = 1$. Its time derivative, $dp_{c,i,z}(t)/dt$, gives the probability density of the conditional FPT distribution over t . We want to calculate the moments of this distribution.

We also consider the probability to reach any terminal vertex in Z . Let $A_t \subseteq E$ be those trajectories that have reached Z by time t . Then we can define,

$$p_{i,Z}(t) = \Pr[X(t') \in Z \text{ for some } 0 \leq t' \leq t \mid X(0) = i] \approx \frac{\#A_t}{\#E}. \quad (29)$$

Evidently, $p_{i,Z}(\infty) = 1$, so that $p_{i,Z}(t)$ is also a bona fide cumulative distribution function, whose corresponding probability density is given by $dp_{i,Z}(t)/dt$. It is also clear that $A_t = \bigcup_{z \in Z} A_{z,t}$ and that this union is disjoint, so that $\#A_t = \sum_{z \in Z} \#A_{z,t}$. It follows from Eqns. 26 and 27 that,

$$p_{i,Z}(t) = \sum_{z \in Z: i \in \mathcal{V}_{\text{in}}(z)} p_{c,i,z}(t) \pi_{i,z}. \quad (30)$$

Instead of working directly with $p_{c,i,z}(t)$ to calculate FPTs, it is easier to work with $p_{i,z}(t)$, for which we can find a system of differential equations for $i \in \mathcal{V}_{\text{in}}(z)$. From there, via the Laplace transform, we can determine the splitting probabilities, $\pi_{i,z}$, the moments of the conditional FPT to each terminal vertex, and the moments of the FPT to any terminal vertex.

We need some further notation before embarking on this calculation. Let G be a graph with $T \geq 1$ terminal SCCs. We can assume, without loss of generality, that the vertices are indexed so that the terminal SCCs are as follows,

$$\mathcal{T}(G) = \{\{N - T + 1\}, \dots, \{N\}\},$$

so that $Z = \{N - T + 1, \dots, N\}$. If $i \in \mathcal{V}(G)$, then it must have a path to some terminal vertex. Let $\mathcal{Z}(i) \subseteq Z$ denote the subset of those terminal vertices,

$$\mathcal{Z}(i) = \mathcal{V}_{\text{out}}(i) \cap Z \neq \emptyset. \quad (31)$$

$\mathcal{Z}(i) = \{i\}$ if, and only if, i is itself terminal.

Let us now choose $z \in Z$. It will be held fixed for much of the argument that follows. The vertices that lead to z , excluding z itself, which form the subset $\mathcal{V}_{\text{in}}(z) \setminus \{z\}$, must all be non-terminal. We can, again without loss of generality, index these non-terminal vertices as $1, \dots, N_z$, for some $N_z \leq N - T$. Note that this indexing depends on the choice of $z \in Z$. It follows that,

$$\mathcal{V}_{\text{in}}(z) = \{1, \dots, N_z, z\}. \quad (32)$$

It will be helpful to let $\mathcal{V}_{\text{in}}^-(z)$ denote just the non-terminal vertices in $\mathcal{V}_{\text{in}}(z)$, $\mathcal{V}_{\text{in}}^-(z) = \mathcal{V}_{\text{in}}(z) \setminus \{z\}$. Let us also choose $i \in \mathcal{V}_{\text{in}}(z)$, so that $z \in \mathcal{Z}(i)$. The vertex i will also be held fixed for much of the following argument.

2.2.2 The basic time-evolution equation

With that preparation, we can now derive a system of differential equations for $p_{i,z}(t)$. First note that, because $X(\cdot)$ is time-homogeneous, if we shift the time axis by h , so that trajectories start at $X(h) = i$, then the probability of reaching z by time $t + h$ is also given by $p_{i,z}(t)$. Recall from the elementary properties of Markov processes that the probability that $X(\cdot)$ remains at i up to time t is given by $e^{-\lambda_i t}$, where λ_i is the total exit rate from i ,

$$\lambda_i = \sum_{j \in V_{\text{out}}(i)} \ell(i \rightarrow j).$$

The probability that $X(\cdot)$ has left i by time t is then $1 - e^{-\lambda_i t}$. If we choose $0 < h < t$ sufficiently small, then the probability that $X(\cdot)$ has taken one transition from i by time h is $\lambda_i h$, to first order in h . The probability that it has taken k transitions by time h can then be no less than order k in h . Accordingly, to first order in h , we may assume that $X(\cdot)$ has taken no more than one transition from i . There are then two mutually exclusive possibilities. First, $X(\cdot)$ remains at i up to time h , with probability $e^{-\lambda_i h}$, and then reaches z within the remaining time from h to t . Because of the time-homogeneity mentioned above, the latter probability is just $p_{i,z}(t - h)$. Second, $X(\cdot)$ leaves i by time h , which it does with probability $1 - e^{-\lambda_i h}$, and takes the edge $i \rightarrow j$, which it does with probability $\ell(i \rightarrow j)/\lambda_i$, and then reaches z from j within the remaining time from h to t , which, as before, occurs with probability $p_{j,z}(t - h)$. Putting these parts together, we see that,

$$p_{i,z}(t) = e^{-\lambda_i h} p_{i,z}(t - h) + (1 - e^{-\lambda_i h}) \sum_{j \in V_{\text{out}}(i)} \left(\frac{\ell(i \rightarrow j)}{\lambda_i} \right) p_{j,z}(t - h) + o(h),$$

where the “little o” notation, $o(h)$, denotes some function $f(h)$ which is of order higher than 1 in h , so that $\lim_{h \rightarrow 0} f(h)/h = 0$. If we subtract $p_{i,z}(t - h)$ from both sides, divide through by h , and let $h \rightarrow 0$, then, keeping in mind that $\lim_{h \rightarrow 0} (1 - e^{-\lambda_i h})/h = \lambda_i$, we obtain,

$$\frac{d}{dt} p_{i,z}(t) = \sum_{j \in V_{\text{out}}(i)} \ell(i \rightarrow j) p_{j,z}(t) - \lambda_i p_{i,z}(t). \quad (33)$$

We could use Eqn. 28 to rewrite Eqn. 33 as a system of equations for the conditional probabilities, $p_{c,i,z}(t)$, but this would also involve the splitting probabilities, $\pi_{i,z}$, which are themselves unknown. This illustrates why it is preferable to work with the system of equations for the unconditional probabilities, $p_{i,z}(t)$, in Eqn. 33.

We can rewrite Eqn. 33 in terms of the underlying graph G by defining the vector of probabilities to reach z by time t , $\mathbf{p}_z(t) = (p_{1,z}(t), \dots, p_{N_z,z}(t))^T \in \mathbb{R}^{N_z}$, where the indices run over the non-terminal vertices of G that lead to z , i.e., $i \in \mathcal{V}_{\text{in}}^-(z)$. (Note that, if $i = z$, then $p_{i,z}(t) = p_{z,z}(t) = 1$ for all $t \geq 0$.) Now, let $\mathbf{p}_z^+(t) \in \mathbb{R}^{N_z+1}$

denote the augmented vector,

$$\mathbf{p}_z^+(t) = (p_{1,z}(t), \dots, p_{N_z,z}(t), p_{z,z}(t))^T = (p_{1,z}(t), \dots, p_{N_z,z}(t), 1)^T.$$

Then, rewriting Eqn. 33, we see that,

$$\frac{d}{dt} \mathbf{p}_z(t) = \left(\mathcal{L}(G)^T \right)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}(z)]} \mathbf{p}_z^+(t), \quad (34)$$

where the submatrix of $\mathcal{L}(G)^T$ has size $N_z \times (N_z + 1)$. Eqn. 34 is the *adjoint master equation* of the Markov process (Iyer-Biswas et al. 2016; van Kampen 2007), so named because its operator is the transpose to that in the master equation (Eqn. 1).

2.2.3 The Laplace transform

The calculation of moments is most readily done by using the Laplace transform of the probability density. Let $u_{i,z}(t)$ be the time derivative of $p_{i,z}(t)$,

$$u_{i,z}(t) = \frac{dp_{i,z}(t)}{dt}. \quad (35)$$

As noted above, $u_{i,z}(t)$ is generally not the density of a probability distribution over t because it is not correctly normalised, but it can serve the same purpose here. Differentiating Eqn. 34 with respect to t yields a similar adjoint time-evolution equation for $\mathbf{u}_z(t) = (u_{1,z}(t), \dots, u_{N_z,z}(t))^T$,

$$\frac{d}{dt} \mathbf{u}_z(t) = \left(\mathcal{L}(G)^T \right)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \mathbf{u}_z(t), \quad (36)$$

where we have used the fact that $u_{z,z}(t) = 0$ for all $t \geq 0$ to drop the final column in the right-hand matrix, which now has size $N_z \times N_z$.

Now, let $\tilde{u}_{i,z}(s) = \mathcal{L}\{u_{i,z}(t)\}$ denote the Laplace transform of $u_{i,z}(t)$,

$$\tilde{u}_{i,z}(s) = \mathcal{L}\{u_{i,z}(t)\} = \int_0^\infty e^{-st} u_{i,z}(t) dt.$$

Applying the Laplace transform to both sides of Eqn. 36 and using the well-known property that it converts differentiation by t into multiplication by s , we get a linear algebraic equation for $\tilde{\mathbf{u}}_z(s) = (\tilde{u}_{1,z}(s), \dots, \tilde{u}_{N_z,z}(s))^T$,

$$s\tilde{\mathbf{u}}_z(s) - \mathbf{u}_z(0) = \left(\mathcal{L}(G)^T \right)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \tilde{\mathbf{u}}_z(s). \quad (37)$$

To determine the initial condition vector, $\mathbf{u}_z(0)$, note that $p_{i,z}(0) = 0$ if $i \neq z$, and $p_{z,z}(0) = 1$. Substituting into Eqn. 33 for $t = 0$, we obtain,

$$u_{i,z}(0) = \left. \frac{dp_{i,z}(t)}{dt} \right|_{t=0} = \begin{cases} \ell(i \rightarrow z) & \text{if } i \rightarrow z \\ 0 & \text{otherwise.} \end{cases} \quad (38)$$

It follows that $\mathbf{u}_z(0)$ contains those entries in the z -th column of $\mathcal{L}(G)^T$ that correspond to the vertices in $\mathcal{V}_{\text{in}}^-(z)$, which we will denote by \mathbf{v} ,

$$\mathbf{v} = \mathbf{u}_z(0) = \left(\mathcal{L}(G)^T \right)_{[\mathcal{V}_{\text{in}}^-(z), \{z\}]} . \quad (39)$$

Therefore, we can now rewrite Eqn. 37 as,

$$\tilde{\mathbf{u}}_z(s) = \left(s\mathbf{I} - \left(\mathcal{L}(G)^T \right)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \right)^{-1} \mathbf{v} . \quad (40)$$

Proposition 2 tells us that the inverse of $\mathcal{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]}$ exists if, and only if, there exists a spanning forest of G rooted at $\mathcal{V}_{\text{in}}^-(z)$. Since $\mathcal{V}_{\text{in}}^-(z)$ contains all the terminal vertices, Z , and every vertex in G leads to some terminal vertex, such a spanning forest must exist (§2.1.1). Because the determinant is a continuous function of the entries of a matrix, Eqn. 40 is well-defined within some neighbourhood of $s = 0$.

For notational convenience in what follows, we introduce the negative transpose of the Laplacian matrix, as

$$\mathbf{L}(G) = -\mathcal{L}(G)^T, \quad (41)$$

so that we can rewrite Eqn. 40 as,

$$\tilde{\mathbf{u}}_z(s) = \left(s\mathbf{I} + \mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \right)^{-1} \mathbf{v} . \quad (42)$$

2.2.4 Calculating moments

With the Laplace transform $\tilde{\mathbf{u}}_z(s)$ at our disposal, we can now start to calculate moments. Although $u_{i,z}(t)$ is not a bona fide probability density over t , as noted above, we can still define the quantity corresponding to its r -th moment,

$$\mu_{i,z}^{(r)} = \int_0^\infty t^r u_{i,z}(t) dt = (-1)^r \left. \frac{d^r \tilde{u}_{i,z}(s)}{ds^r} \right|_{s=0}, \quad (43)$$

where the second equality comes from the well-known property of the Laplace transform that differentiation by s corresponds to multiplication by t . Introducing the vector

$\mu_z^{(r)} = (\mu_{1,z}^{(r)}, \dots, \mu_{N_z,z}^{(r)})^T$ and using Eqn. 42, we can write,

$$\mu_z^{(r)} = (-1)^r \frac{d^r}{ds^r} \left(s\mathbf{I} + \mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \right)^{-1} \Big|_{s=0} \mathbf{v}.$$

Recall that, if \mathbf{M} is a square matrix that does not depend on s , then (Bernstein 2009),

$$\frac{d^r}{ds^r} (s\mathbf{I} + \mathbf{M})^{-1} = (-1)^r r! (s\mathbf{I} + \mathbf{M})^{-(r+1)}.$$

It follows that,

$$\mu_z^{(r)} = r! \left(\mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \right)^{-(r+1)} \mathbf{v}. \quad (44)$$

Eqn. 44 immediately reveals how $\mu_z^{(r)}$ can be calculated by appealing to Proposition 2.

Before doing that, it will be helpful for what follows (see Theorem 6) to allow greater generality in the choice of the submatrix of $\mathbf{L}(G)$ that we use. Suppose that we choose $U \subseteq \mathcal{V}(G)$ to be any subset of vertices that contains all the terminal vertices but does not contain any of the non-terminal vertices that lead to z ,

$$\begin{aligned} Z &\subseteq U \subseteq \overline{\mathcal{V}_{\text{in}}^-(z)} \\ \mathcal{V}_{\text{in}}^-(z) &\subseteq \overline{U} \subseteq \overline{Z}. \end{aligned} \quad (45)$$

Because U contains the terminal SCCs (in this case, terminal vertices) of G , it follows from what was said in §2.1.1 that $\Phi_U(G) \neq \emptyset$. Hence, Proposition 2 tells us that the submatrix $\mathbf{L}(G)_{[\overline{U}, \overline{U}]}$ is invertible. Moreover, suppose that $i > N_z$ and $1 \leq j \leq N_z$, so that $j \in \mathcal{V}_{\text{in}}^-(z)$. Then there can be no edge $i \rightarrow j$: if there were, then, because of our choice of indices, $i \in \mathcal{V}_{\text{in}}^-(z) = \{1, \dots, N_z, z\}$. But, since $i > N_z$, this means that $i = z$, which is terminal, so there can be no edge $z \rightarrow j$, which yields a contradiction. It follows that this submatrix has a block upper-triangular form,

$$\mathbf{L}(G)_{[\overline{U}, \overline{U}]} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}, \quad (46)$$

where $\mathbf{A} = \mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]}$. Since we know that the block matrix in Eqn. 46 is invertible, we can apply the block matrix inversion formula (Bernstein 2009) to obtain,

$$\left(\mathbf{L}(G)_{[\overline{U}, \overline{U}]} \right)^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1}) \\ \mathbf{0} & \mathbf{C}^{-1} \end{bmatrix}.$$

It follows that, provided $i, j \in \mathcal{V}_{\text{in}}^-(z)$,

$$\left(\mathbf{L}(G)_{[\overline{U}, \overline{U}]} \right)^{-1}_{i,j} = \mathbf{A}^{-1}_{i,j} = \left(\mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \right)^{-1}_{i,j}.$$

If we consider Proposition 2 with $A = U$, then, in view of Eqn. 45, $\{1, \dots, N_z\} \subseteq \overline{U}$, so that the ordering map, $\theta_{\overline{U}}: \{1, \dots, \#\overline{U}\} \rightarrow \overline{U}$, is the identity, $\theta_{\overline{U}}(q) = q$, so long as $q \in \mathcal{V}_{\text{in}}^-(z) = \{1, \dots, N_z\}$. Accordingly, we can apply Proposition 2 to conclude that,

$$\left(\mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]}\right)_{i,j}^{-1} = \frac{w(\Phi_{U \cup \{i\} \rightarrow U \cup \{j\}}(G))}{w(\Phi_U(G))}, \quad (47)$$

where $i, j \in \mathcal{V}_{\text{in}}^-(z)$ and U is chosen as in Eqn. 45.

2.2.5 Splitting probabilities and conditional FPT moments

We can now calculate the quantities of interest. First, if we set $r = 0$ in Eqn. 43 and recall Eqn. 35, then, for all $i \in \mathcal{V}_{\text{in}}^-(z)$, we have,

$$\mu_{i,z}^{(0)} = \int_0^\infty u_{i,z}(t) dt = p_{i,z}(\infty) - p_{i,z}(0) = \pi_{i,z}. \quad (48)$$

Therefore, we can use Eqn. 44 to write,

$$\pi_{i,z} = \sum_{j \in \mathcal{V}_{\text{in}}^-(z)} \left(\mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]}\right)_{i,j}^{-1} v_j.$$

Making use of Eqns. 39 and 47, we then see that,

$$\pi_{i,z} = \sum_{j \in \mathcal{V}_{\text{in}}(z)} \left(\frac{w(\Phi_{U \cup \{i\} \rightarrow U \cup \{j\}}(G))}{w(\Phi_U(G))} \right) \ell(j \rightarrow z),$$

where we recall that $\mathcal{V}_{\text{in}}(z)$ is the subset of vertices with edges into z . We can now apply Eqn. 24, with $A = U$, to see that,

$$\pi_{i,z} = \frac{w(\Phi_{(U \setminus \{z\}) \cup \{i\} \rightarrow U}(G))}{w(\Phi_U(G))}.$$

This leads to the first of our main results.

Theorem 4 *Let G be a graph whose terminal SCCs are single vertices and let Z be the set of terminal vertices. Choose $z \in Z$. Then, for any subset $U \subseteq \mathcal{V}(G)$ such that $Z \subseteq U \subseteq \overline{\mathcal{V}_{\text{in}}^-(z)}$ and any $i \in \mathcal{V}_{\text{in}}^-(z)$, the splitting probability from i to z is given by,*

$$\pi_{i,z} = \frac{w(\Phi_{(U \setminus \{z\}) \cup \{i\} \rightarrow U}(G))}{w(\Phi_U(G))}. \quad (49)$$

Proof Here, $T = \#Z$. We are at liberty to re-index the vertices of G so that $Z = \{N - T + 1, \dots, N\}$ and, for the chosen $z \in Z$, $\mathcal{V}_{\text{in}}(z) = \{1, \dots, N_z, z\}$. The preceding calculation then yields Eqn. 49. \square

The denominator in Eqn. 49 is the weight of all spanning forests of G rooted at U , while the numerator is the weight of the subset of such spanning forests in which there is a path from i to z . Therefore, this ratio is a unitless number that must lie between 0 and 1, as expected for a probability. Moreover, since $i \in \mathcal{V}_{\text{in}}^-(z)$, there is at least one spanning forest that contributes to the numerator, which implies that $\pi_{i,z} \neq 0$.

Having calculated the splitting probability, we now turn to the moments of the conditional FPT distribution, which we denote by $\mu_{c,i,z}^{(r)}$. It follows from Eqn. 28 that,

$$\mu_{c,i,z}^{(r)} = \frac{\mu_{i,z}^{(r)}}{\pi_{i,z}}. \quad (50)$$

In view of Eqn. 48, we see that $\mu_{c,i,z}^{(0)} = 1$. For the higher moments, we see from Eqn. 44 that,

$$\mu_z^{(r)} = r! \left(\mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \right)^{-r} \mu_z^{(0)},$$

where Eqn. 48 tells us that $\mu_z^{(0)} = (\pi_{1,z}, \dots, \pi_{N_z,z})^T$. Expanding out the matrix multiplications, we see that,

$$\mu_{i,z}^{(r)} = r! \sum_{(i_1, \dots, i_r)} \left(\prod_{j=0}^{r-1} \left(\mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \right)^{-1}_{i_j, i_{j+1}} \right) \pi_{i_r, z},$$

where we have set $i_0 = i$ and the sum runs over $(i_1, \dots, i_r) \in (\mathcal{V}_{\text{in}}^-(z))^r$. It then follows from Eqn. 50 that,

$$\mu_{c,i,z}^{(r)} = \frac{r!}{\pi_{i,z}} \sum_{(i_1, \dots, i_r)} \left(\prod_{j=0}^{r-1} \left(\mathbf{L}(G)_{[\mathcal{V}_{\text{in}}^-(z), \mathcal{V}_{\text{in}}^-(z)]} \right)^{-1}_{i_j, i_{j+1}} \right) \pi_{i_r, z}. \quad (51)$$

Note that this equation is valid even when $r = 0$. In this case, the sum runs over the empty Cartesian product, $(\mathcal{V}_{\text{in}}^-(z))^0$, and the term corresponding to the sole 0-tuple contains an empty product, which evaluates to one. Furthermore, $\pi_{i_r, z} = \pi_{i_0, z} = \pi_{i,z}$. Therefore, we obtain $\mu_{c,i,z}^{(0)} = 1$, as before.

We can now deduce our second main result. As with Theorem 4, it is convenient to state this in slightly greater generality, in terms of a subset U that satisfies Eqn. 45.

Theorem 5 *Let G be a graph whose terminal SCCs are single vertices and let Z be the set of terminal vertices. Choose $z \in Z$. Then, for any subset $U \subseteq \mathcal{V}(G)$ such that $Z \subseteq U \subseteq \mathcal{V}_{\text{in}}^-(z)$ and $i \in \mathcal{V}_{\text{in}}^-(z)$, the r -th moment of the conditional FPT from i to z is given by*

$$\mu_{c,i,z}^{(r)} = r! \sum_{(i_1, \dots, i_r) \in \overline{U}^r} \left(\prod_{j=0}^{r-1} \frac{w(\Phi_{U \cup \{i_j\} \rightarrow U \cup \{i_{j+1}\}}(G))}{w(\Phi_U(G))} \right) \left(\frac{w(\Phi_{(U \setminus \{z\}) \cup \{i_r\} \rightarrow U}(G))}{w(\Phi_{(U \setminus \{z\}) \cup \{i\} \rightarrow U}(G))} \right),$$

where $i_0 = i$.

Proof We undertake a similar re-indexing as in the proof of Theorem 4 and make use of the calculations above. We can rewrite the quantity in Eqn. 51 by using Eqn. 47 to get,

$$\mu_{c,i,z}^{(r)} = r! \sum_{(i_1, \dots, i_r)} \left(\prod_{j=0}^{r-1} \frac{w(\Phi_{U \cup \{i_j\} \rightarrow U \cup \{i_{j+1}\}}(G))}{w(\Phi_U(G))} \right) \left(\frac{\pi_{i_r, z}}{\pi_{i, z}} \right),$$

where $(i_1, \dots, i_r) \in (\mathcal{V}_{\text{in}}^-(z))^r$. We can then use Theorem 4 to rewrite the ratio of splitting probabilities to get,

$$\mu_{c,i,z}^{(r)} = r! \sum_{(i_1, \dots, i_r)} \left(\prod_{j=0}^{r-1} \frac{w(\Phi_{U \cup \{i_j\} \rightarrow U \cup \{i_{j+1}\}}(G))}{w(\Phi_U(G))} \right) \left(\frac{w(\Phi_{(U \setminus \{z\}) \cup \{i_r\} \rightarrow U}(G))}{w(\Phi_{(U \setminus \{z\}) \cup \{i\} \rightarrow U}(G))} \right).$$

This is the quantity in the statement of the Theorem, except that the summation is taken over $(i_1, \dots, i_r) \in (\mathcal{V}_{\text{in}}^-(z))^r$. To see that we can convert the sum to one that runs over \overline{U}^r , note that the r -fold product in the numerator of the first parenthetical term,

$$\prod_{j=0}^{r-1} w(\Phi_{U \cup \{i_j\} \rightarrow U \cup \{i_{j+1}\}}(G)),$$

is nonzero if, and only if, there are paths

$$i_0 = i \rightsquigarrow i_1, \quad i_1 \rightsquigarrow i_2, \quad \dots, \quad i_{r-1} \rightsquigarrow i_r.$$

Furthermore, the weight in the numerator of the second parenthetical term is nonzero if, and only if, $i_r \rightsquigarrow z$. Therefore, the summand corresponding to (i_1, \dots, i_r) is zero if at least one vertex in the r -tuple does not have a path to z . In view of Eqn. 45, any elements in \overline{U} that are not in $\mathcal{V}_{\text{in}}^-(z)$ do not contribute to the sum. Hence, we are at liberty to take $(i_1, \dots, i_r) \in \overline{U}^r$, as stated. \square

The r -th moment of the conditional FPT, $\mu_{c,i,z}^{(r)}$, as given by the formula in Theorem 5, has units of $(\text{time})^r$. It is instructive to see how this arises on the right-hand side of the formula. Recall that a spanning forest F has $N - \#\mathcal{R}(F)$ edges. Hence, the second parenthetical term in the right-hand sum is unitless. In the first term, the weight in the numerator incurs one fewer edge label, and corresponding unit of $(\text{time})^{-1}$, than the weight in the denominator, so that the ratio has units of time. It follows that the right-hand side as a whole has units of $(\text{time})^r$, as required.

We also note that the formula for $\mu_{c,i,z}^{(r)}$ in Theorem 5 shares with that for the steady-state probabilities in Eqn. 5 the property of being a manifestly positive rational algebraic function of the edge labels.

As the mean is often of interest, it may be helpful to extricate this as a separate result. Under the same conditions as Theorem 5,

$$\mu_{c,i,z}^{(1)} = \sum_{j \in \bar{U}} \left(\frac{w(\Phi_{U \cup \{i\} \rightarrow U \cup \{j\}}(G))}{w(\Phi_U(G))} \right) \left(\frac{w(\Phi_{(U \setminus \{z\}) \cup \{j\} \rightarrow U}(G))}{w(\Phi_{(U \setminus \{z\}) \cup \{i\} \rightarrow U}(G))} \right). \quad (52)$$

2.2.6 The unconditional FPT to Z

It follows readily from Eqn. 30 that the r -th moment of the FPT distribution to reach any of the terminal vertices in Z from i is given by,

$$\mu_{i,Z}^{(r)} = \sum_{z \in \mathcal{Z}(i)} \mu_{c,i,z}^{(r)} \pi_{i,z}, \quad (53)$$

where we have made use of $\mathcal{Z}(i) \subseteq Z$, as introduced in Eqn. 31. Eqn. 53 allows us to deduce our third main result.

Theorem 6 *Let G be a graph whose terminal SCCs are single vertices and let Z be the set of terminal vertices. The r -th moment of the FPT from $i \in \bar{Z}$ to Z is given by*

$$\mu_{i,Z}^{(r)} = r! \sum_{(i_1, \dots, i_r) \in \bar{Z}^r} \left(\prod_{j=0}^{r-1} \frac{w(\Phi_{Z \cup \{i_j\} \rightarrow Z \cup \{i_{j+1}\}}(G))}{w(\Phi_Z(G))} \right),$$

where $i_0 = i$.

Proof Let us take $U = Z$ and apply Theorems 4 and 5 to Eqn. 53 to obtain the following formula for $\mu_{i,Z}^{(r)}$,

$$r! \sum_{z \in \mathcal{Z}(i)} \sum_{(i_1, \dots, i_r) \in \bar{Z}^r} \left(\prod_{j=0}^{r-1} \frac{w(\Phi_{Z \cup \{i_j\} \rightarrow Z \cup \{i_{j+1}\}}(G))}{w(\Phi_Z(G))} \right) \left(\frac{w(\Phi_{(Z \setminus \{z\}) \cup \{i_r\} \rightarrow Z}(G))}{w(\Phi_Z(G))} \right).$$

The r -fold product within the inner sum does not depend on z , so we can rearrange the formula to give,

$$r! \sum_{(i_1, \dots, i_r) \in \bar{Z}^r} \left(\prod_{j=0}^{r-1} \frac{w(\Phi_{Z \cup \{i_j\} \rightarrow Z \cup \{i_{j+1}\}}(G))}{w(\Phi_Z(G))} \right) \left(\sum_{z \in \mathcal{Z}(i)} \frac{w(\Phi_{(Z \setminus \{z\}) \cup \{i_r\} \rightarrow Z}(G))}{w(\Phi_Z(G))} \right).$$

As in the proof of Theorem 5, the r -fold product of weights in the numerator of the first parenthetical term is nonzero if, and only if, there are paths

$$i_0 = i \rightsquigarrow i_1, \quad i_1 \rightsquigarrow i_2, \quad \dots, \quad i_{r-1} \rightsquigarrow i_r. \quad (54)$$

Furthermore, the numerator in the second parenthetical term is nonzero if, and only if, $i_r \rightsquigarrow z$, so that $z \in \mathcal{Z}(i_r)$. Because \rightsquigarrow is transitive, Eqn. 54 tells us that $i \rightsquigarrow i_r$. If $z' \in \mathcal{Z}(i_r)$, so that $i_r \rightsquigarrow z'$, then it follows that $i \rightsquigarrow z'$, so that $z' \in \mathcal{Z}(i)$. Hence, $\mathcal{Z}(i_r) \subseteq \mathcal{Z}(i)$ for any r -tuple for which the corresponding summand above is nonzero. Therefore, we can rewrite the second parenthetical expression as,

$$\sum_{z \in \mathcal{Z}(i_r)} \frac{w(\Phi_{(Z \setminus \{z\}) \cup \{i_r\} \rightarrow Z}(G))}{w(\Phi_Z(G))}.$$

But now we see from Theorem 4 that this expression is nothing other than the total sum of splitting probabilities from i_r ,

$$\sum_{z \in \mathcal{Z}(i_r)} \pi_{i_r, z} = 1,$$

from which the result follows. \square

In a similar fashion to the formula in Theorem 5, the formula in Theorem 6 has the required units of (time) ^{r} and is a manifestly positive rational algebraic function of the edge labels.

As before, it may be helpful to extricate from Theorem 6 the mean FPT from i to Z , which is given by,

$$\mu_{i, Z}^{(1)} = \sum_{j \in \overline{Z}} \frac{w(\Phi_{Z \cup \{i\} \rightarrow Z \cup \{j\}}(G))}{w(\Phi_Z(G))}. \quad (55)$$

2.2.7 A single terminal vertex

The case of a single terminal vertex, z , is often encountered and it may be helpful to see what the results above yield in this case. Evidently, $\mathcal{V}_{\text{in}}(z) = \mathcal{V}(G)$. Taking $U = \{z\}$, as we would expect, Theorem 4 implies that the splitting probability from every vertex is one,

$$\pi_{i, z} = \frac{w(\Phi_{\{i\} \rightarrow \{z\}}(G))}{w(\Phi_{\{z\}}(G))} = 1, \quad (56)$$

since every spanning tree in $\Phi_{\{z\}}(G)$ has a path from i to z . It is also clear that

$$\mu_{c, i, z}^{(r)} = \mu_{i, z}^{(r)} = \mu_{i, \{z\}}^{(r)},$$

where the first equality follows from Eqn. 50 and the second from Eqn. 53. We can summarise this as follows.

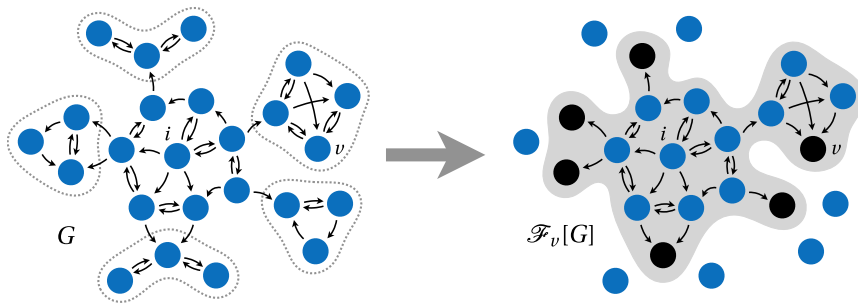


Fig. 3 The edge removal procedure for FPT calculations on general graphs. The graph, G , on the left has five terminal SCCs, shown within dashed ovals, and two non-terminal SCCs. The source vertex i and the target vertex v are labelled. Note that v is in a terminal SCC. The edge removal procedure described in the text yields the graph, $\mathcal{F}_v[G]$, in the shaded region on the right. As implied by Lemma 8, the terminal SCCs of $\mathcal{F}_v[G]$ are all singletons, and are shown in black

Proposition 7 *Let G be a graph with one terminal vertex, z . Then the r -th moment of the FPT from $i \in \mathcal{V}(G)$ to z is given by*

$$\mu_{i,z}^{(r)} = r! \sum_{(i_1, \dots, i_r) \in (\mathcal{V}(G) \setminus \{z\})^r} \left(\prod_{j=0}^{r-1} \frac{w(\Phi_{\{i_j, z\} \rightarrow \{i_{j+1}, z\}}(G))}{w(\Phi_{\{z\}}(G))} \right),$$

where $i_0 = i$.

As before, it may be helpful to extricate from Proposition 7 the result for the mean FPT from i to z ,

$$\mu_{i,z}^{(1)} = \sum_{j \in \mathcal{V}(G) \setminus \{z\}} \frac{w(\Phi_{\{i, z\} \rightarrow \{j, z\}}(G))}{w(\Phi_{\{z\}}(G))}. \quad (57)$$

2.3 FPTs on general graphs

Our main results above, Theorems 4, 5 and 6, were formulated for graphs whose terminal SCCs are singletons. Here, we show how these results can be extended to more general situations through a process of removing an appropriate set of edges.

Let G be a connected graph with any number of terminal SCCs, each of which may contain any number of vertices, and let $X(\cdot)$ be the associated Markov process. Let $v \in \mathcal{V}(G)$ be any vertex, which need not lie within a terminal SCC. We want to calculate the splitting probability and conditional FPT to v from some vertex, $i \in \mathcal{V}_{\text{in}}(v)$. To determine these quantities, first define a modified graph in which the outgoing edges from v are removed, along with the outgoing edges from any vertex in a terminal SCC, unless that SCC contains v itself. More precisely, the outgoing edges are removed from the set of vertices,

$$\{v\} \cup \left(\bigcup_{C \in \mathcal{T}(G), v \notin C} C \right).$$

This procedure is illustrated in Fig. 3. As shown in the figure, it may yield a graph with several connected components. Let $\mathcal{F}_v[G]$ denote the connected component of the modified graph that contains v , as shown with a shaded background in Fig. 3.

Lemma 8 *The terminal SCCs of $\mathcal{F}_v[G]$ are all singletons. Furthermore, if $i \in \mathcal{V}_{in}(v)$, then the trajectory probabilities from i defined in Eqn. 25 are identical in G and in $\mathcal{F}_v[G]$, so that $p_{i,v}^G(t) = p_{i,v}^{\mathcal{F}_v[G]}(t)$.*

Proof The edge-removal procedure only removes edges, so that $\mathcal{V}(\mathcal{F}_v[G]) \subseteq \mathcal{V}(G)$ and the SCCs of G can only break up into SCCs of $\mathcal{F}_v[G]$. If $C \in \mathcal{C}(G) \setminus \mathcal{T}(G)$ and $C \not\ni v$, then the procedure does nothing to it. Since C is non-terminal, it must have edges leaving it that are not removed by the procedure, so C remains a non-terminal SCC of $\mathcal{F}_v[G]$. If $C \in \mathcal{T}(G)$ and $C \not\ni v$, then the procedure disintegrates it so that only those vertices $j \in C$ with an incoming edge from outside C are retained in $\mathcal{F}_v[G]$. Each such j becomes a singleton terminal SCC, $\{j\} \in \mathcal{T}(\mathcal{F}_v[G])$. Finally, if $C \ni v$, then the procedure removes the outgoing edges from v and thereby breaks up the strongly connected structure of C so that v becomes a terminal SCC, $\{v\} \in \mathcal{T}(\mathcal{F}_v[G])$. Accordingly, the only terminal SCCs of $\mathcal{F}_v[G]$ are singletons.

For the second claim, we will argue as we did for Eqn. 25. Given $i \in \mathcal{V}_{in}(v)$, consider stochastically generating a large finite ensemble of unbounded trajectories of $X(\cdot)$ starting from i . With probability one, each of these trajectories will either reach v at some time or it will not. If a trajectory reaches v for the first time, truncate it at that time and extend it indefinitely while remaining at v . If a trajectory does not reach v , it will eventually reach a terminal SCC, $C \in \mathcal{T}(G)$, which does not contain v . In this case, truncate the trajectory at the first vertex $j \in C$ that is reached and extend it indefinitely while remaining at j . From the construction above, we see that $\{j\} \in \mathcal{T}(\mathcal{F}_v[G])$. Each truncated trajectory is a trajectory of the Markov process associated with $\mathcal{F}_v[G]$ and this truncation process gives rise to a bona fide ensemble of trajectories in this latter Markov process. As the sizes of the two ensembles go to infinity, the probabilities of the original and truncated trajectories become identical, from which the second claim follows. \square

In view of Lemma 8, the splitting probabilities and the moments of the conditional FPT distribution from i to v in G can now be calculated by applying Theorems 4 and 5, respectively, to $\mathcal{F}_v[G]$.

3 Examples

In a prequel to the present paper, we summarised the main results discussed above without providing proofs and also worked through some example calculations for *pipeline graphs* (Nam and Gunawardena 2023, §2.4). These graphs consist of sequences of vertices with edges only between consecutive vertices (Fig. 4). The corresponding

Markov processes have been widely studied in the context of single-molecule enzyme kinetics (Shaevitz et al. 2005; Kou et al. 2005; Chemla et al. 2008; Garai et al. 2009; Moffitt et al. 2010; Moffitt and Bustamante 2014). Pipeline graphs are particularly amenable to spanning forest enumeration, and we showed in Nam and Gunawardena (2023) how our results provide a systematic way to recover findings in the literature. Here, we will examine three examples, two pipeline graphs (Fig. 4) and a butterfly graph (Fig. 5), to illustrate in more detail how our formulas can be worked out in practice.

3.1 Pipeline graphs

Consider first the pipeline graph in Fig. 4A, consisting of a pair of vertices linked by reversible edges, leading to a terminal vertex. This graph has been used to model a single catalytic cycle of a Michaelis–Menten enzyme mechanism (Shaevitz et al. 2005; Kou et al. 2005; Moffitt et al. 2010; Moffitt and Bustamante 2014), with the vertices representing enzyme states and the edges $1 \rightarrow 2$ corresponding to substrate binding, $2 \rightarrow 1$ to substrate unbinding and $2 \rightarrow 3$ to catalysis. The FPT from 1 to 3 is then a measure of the catalytic rate. Since there is only a single terminal vertex, we can use Proposition 7 to calculate the corresponding moments, $\mu_{1,3}^{(r)}$, where

$$Z = \{3\} \quad \text{and} \quad \mathcal{V}_{\text{in}}^-(3) = \{1, 2\}.$$

We will omit the name of the graph in what follows, as it is clear from the context.

According to Proposition 7, we need to calculate a sum over the r -tuples,

$$(i_1, \dots, i_r) \in \{1, 2\}^r,$$

using spanning trees and forests drawn from the sets shown in Fig. 4A. There is only a single spanning tree rooted at 3, which contributes a weight of $k_1 k_3$. As for the forests, the subsets required have the form $\Phi_{\{i_j, 3\} \rightarrow \{i_{j+1}, 3\}}$, in which there is a path from i_j to i_{j+1} in the forest. Each such forest has only a single edge, and it follows from Fig. 4A that the forest weights are as given in the following table,

	$i_{j+1} = 1$	$i_{j+1} = 2$
$i_j = 1$	$k_2 + k_3$	k_1
$i_j = 2$	k_2	k_1

(58)

The r -tuple, (i_1, \dots, i_r) , provides a code for reading off the corresponding product of forest weights, $w(\Phi_{\{i_0, 3\} \rightarrow \{i_1, 3\}}) \cdots w(\Phi_{\{i_{r-1}, 3\} \rightarrow \{i_r, 3\}})$, in the formula in Proposition 7.

For the mean FPT, we see from Eqn. 58 that,

$$\mu_{1,3}^{(1)} = \sum_{i_1 \in \{1, 2\}} \frac{w(\Phi_{\{1, 3\} \rightarrow \{i_1, 3\}})}{w(\Phi_{\{3\}})} = \frac{(k_2 + k_3) + k_1}{k_1 k_3}. \quad (59)$$

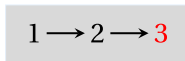
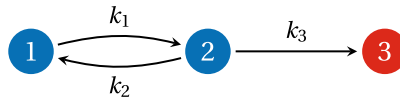
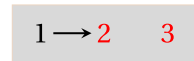
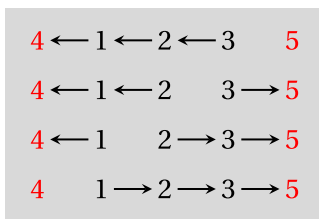
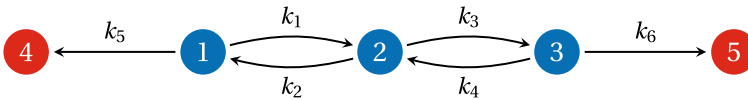
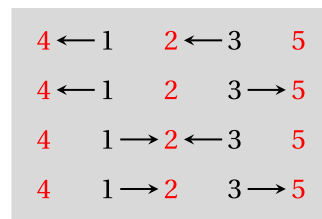
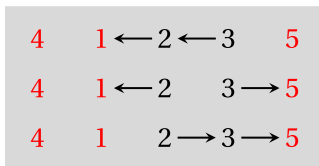
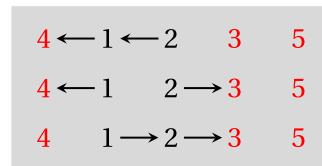
A $\Phi_{\{3\}}$  $\Phi_{\{1,3\}}$  $\Phi_{\{2,3\}}$ **B** $\Phi_{\{4,5\}}$  $\Phi_{\{2,4,5\}}$  $\Phi_{\{1,4,5\}}$  $\Phi_{\{3,4,5\}}$

Fig. 4 (A) A pipeline graph on three vertices, as discussed further in the text. The spanning forests are shown below. Red signifies a root vertex. (B) A pipeline graph on five vertices, as discussed further in the text, with the relevant spanning forests shown below. Red signifies a root vertex

This formula is intuitively reasonable: if either k_1 or k_3 become comparatively large, the mean FPT approaches the reciprocal of the other rate; if k_2 becomes comparatively large, the mean FPT increases in proportion.

For the second moment, we see from Eqn. 58 that,

$$\begin{aligned} \mu_{1,3}^{(2)} &= 2 \sum_{(i_1, i_2) \in \{1,2\}^2} \left(\prod_{j=0}^1 \frac{w(\Phi_{\{i_j,3\} \rightarrow \{i_{j+1},3\}})}{w(\Phi_{\{3\}})} \right) \\ &= \frac{2((k_2 + k_3)^2 + (k_2 + k_3)k_1 + k_1k_2 + k_1^2)}{(k_1k_3)^2}. \end{aligned} \quad (60)$$

The higher moments may be calculated in a similar way. It follows from Eqns. 59 and 60 that the quantity $\mu_{1,3}^{(2)} - 2(\mu_{1,3}^{(1)})^2$, which is similar to the variance, simplifies to

$-2/(k_1 k_3)$. This quantity is independent of k_2 , which may be helpful in experimentally testing whether an observed system can be reasonably described by the graph in Fig. 4A. It is plausible that similar polynomial simplifications exist between the higher moments, but this question lies outside the scope of the present paper.

We now consider the pipeline graph in Fig. 4B with five vertices, of which 4 and 5 are both terminal. Here,

$$Z = \{4, 5\} \quad \text{and} \quad \mathcal{V}_{\text{in}}^-(4) = \mathcal{V}_{\text{in}}^-(5) = \{1, 2, 3\},$$

and we will take the source vertex to be the central vertex 2. This graph was used by Terrell Hill to illustrate another method for calculating FPTs (Hill 1988), to which we will return in the Discussion, and we make use of a similar pipeline graph with 23 vertices in our analysis of the CRISPR–Cas9 system (Nam and Gunawardena 2025). Hill’s method permits calculation of the splitting probabilities, $\pi_{2,4}$ and $\pi_{2,5}$, and the mean unconditional FPT to any terminal vertex, $\mu_{2,\{4,5\}}^{(1)}$, but not the mean conditional FPTs to specific terminal vertices, $\mu_{c,2,4}^{(1)}$ and $\mu_{c,2,5}^{(1)}$, which Hill calculated by “a separate stochastic argument (omitted).” In contrast, we can calculate all these quantities with our results, using the spanning forests shown in Fig. 4B.

Hill used a different notation, with the Markov transition rates given in the following table,

$$\begin{array}{|c|c|c|c|c|c|} \hline k_1 & k_2 & k_3 & k_4 & k_5 & k_6 \\ \hline \gamma' & \gamma & \alpha & \alpha' & \delta & \beta \\ \hline \end{array}, \quad (61)$$

along with the symbols (Hill 1988, Eqn. 11),

$$\begin{aligned} A &= \beta + \alpha' = k_4 + k_6 \\ B &= \delta + \gamma' = k_1 + k_5 \\ \Sigma &= \gamma A + \alpha B + AB \\ &= k_2 (k_4 + k_6) + k_3 (k_1 + k_5) + (k_4 + k_6) (k_1 + k_5). \end{aligned} \quad (62)$$

To calculate the splitting probabilities, we can use Theorem 4, with $U = Z = \{4, 5\}$, and the spanning forests in Fig. 4B, which give,

$$\pi_{2,4} = \frac{w(\Phi_{\{2,5\} \rightarrow \{4,5\}})}{w(\Phi_{\{4,5\}})} = \frac{k_2 k_5 (k_4 + k_6)}{k_2 k_5 (k_4 + k_6) + k_3 k_6 (k_1 + k_5)} = \frac{\gamma \delta A}{\gamma \delta A + \alpha \beta B},$$

and similarly (or by noting that $\pi_{2,4} + \pi_{2,5} = 1$),

$$\pi_{2,5} = \frac{w(\Phi_{\{2,4\} \rightarrow \{4,5\}})}{w(\Phi_{\{4,5\}})} = \frac{k_3 k_6 (k_1 + k_5)}{k_2 k_5 (k_4 + k_6) + k_3 k_6 (k_1 + k_5)} = \frac{\alpha \beta B}{\gamma \delta A + \alpha \beta B}.$$

These formulas can be inferred from Hill (1988, Eqns. 12 and 13) by following the prescription given by Hill for determining the “fraction of walks that end with absorption” at a terminal vertex.

To calculate the mean unconditional FPT to Z , we can use the simplified version of Theorem 6 in Eqn. 55 with $Z = \{4, 5\}$, which gives,

$$\mu_{2,\{4,5\}}^{(1)} = \sum_{j \in \{1,2,3\}} \frac{w(\Phi_{\{2,4,5\} \rightarrow \{j,4,5\}})}{w(\Phi_{\{4,5\}})}.$$

The spanning forests needed for this formula are shown in Fig. 4B. Recalling that $\Phi_{X \rightarrow X} = \Phi_X$, we see that,

$$\begin{aligned} w(\Phi_{\{2,4,5\} \rightarrow \{1,4,5\}}) &= k_2 (k_4 + k_6) \\ w(\Phi_{\{2,4,5\}}) &= (k_1 + k_5) (k_4 + k_6) \\ w(\Phi_{\{2,4,5\} \rightarrow \{3,4,5\}}) &= k_3 (k_1 + k_5) \\ w(\Phi_{\{4,5\}}) &= k_2 k_5 (k_4 + k_6) + k_3 k_6 (k_1 + k_5). \end{aligned} \quad (63)$$

It then follows that,

$$\mu_{2,\{4,5\}}^{(1)} = \frac{k_2 (k_4 + k_6) + (k_1 + k_5) (k_4 + k_6) + k_3 (k_1 + k_5)}{k_2 k_5 (k_4 + k_6) + k_3 k_6 (k_1 + k_5)} = \frac{\Sigma}{\gamma \delta A + \alpha \beta B}.$$

which recovers Hill (1988, Eqn. 13).

To calculate the mean conditional FPT from 2 to 4, we can use Theorem 5 with $U = Z = \{4, 5\}$, which yields,

$$\mu_{c,2,4}^{(1)} = \sum_{i_1 \in \{1,2,3\}} \left(\frac{w(\Phi_{\{2,4,5\} \rightarrow \{i_1,4,5\}})}{w(\Phi_{\{4,5\}})} \right) \left(\frac{w(\Phi_{\{i_1,5\} \rightarrow \{4,5\}})}{w(\Phi_{\{2,5\} \rightarrow \{4,5\}})} \right).$$

We need some more spanning forests to calculate this, in addition to those given in Eqn. 63. Examining Fig. 4B, we see that,

$$\begin{aligned} w(\Phi_{\{1,5\} \rightarrow \{4,5\}}) &= k_5 (k_2 k_4 + k_2 k_6 + k_3 k_6) \\ w(\Phi_{\{2,5\} \rightarrow \{4,5\}}) &= k_2 k_5 (k_4 + k_6) \\ w(\Phi_{\{3,5\} \rightarrow \{4,5\}}) &= k_2 k_4 k_5. \end{aligned}$$

It follows that,

$$\begin{aligned} \mu_{c,2,4}^{(1)} &= \frac{(k_4 + k_6) (k_2 k_4 + k_2 k_6 + k_3 k_6) + (k_1 + k_5) (k_4 + k_6)^2 + k_3 k_4 (k_1 + k_5)}{(k_4 + k_6) (k_2 k_5 (k_4 + k_6) + k_3 k_6 (k_1 + k_5))} \\ &= \frac{\gamma A^2 + \alpha \beta A + B A^2 + \alpha \alpha' B}{A(\gamma \delta A + \alpha \beta B)}, \end{aligned}$$

which recovers the corresponding formula in Hill (1988, Eqn. 14). The formula for $\mu_{c,2,5}^{(1)}$ may be calculated in a similar way. As noted above, Hill's method, in contrast to ours, does not extend to these conditional FPTs, which he determined separately, nor to higher moments than the mean, which he did not calculate. We will provide

such generalisations of Hill's approach in a sequel to the present paper (Nam and Gunawardena 2025).

Pipeline graphs have relatively simple structures, which lend themselves to these calculations (Nam and Gunawardena 2023). However, enumerating spanning trees or forests rapidly becomes intractable as the graph becomes larger or more complex. Methods have emerged within the linear framework for dealing with this combinatorial explosion and they are explained further in Nam et al. (2022); Nam and Gunawardena (2023). It is also possible to exploit symmetries in the graph, as the next example shows.

3.2 A butterfly graph

Fig. 5A shows a *butterfly graph*, consisting of two structurally identical “wings” that intersect at a single vertex. Butterfly graphs were defined in Wong et al. (2018) to formalise John Hopfield's concept of discrimination by kinetic proofreading (Hopfield 1974). They have been widely used in studies of discrimination (Murugan et al. 2014; Banerjee et al. 2017; Mallory et al. 2019; Çetiner and Gunawardena 2022), although more often as a diagram than as a mathematical object, as we use them here. One wing of the butterfly graph describes a discriminatory mechanism when it is processing the “right” substrate, while the other wing describes the mechanism when it is processing the “wrong” substrate. The structural symmetry reflects the assumption that the mechanism behaves in an identical way with regard to either substrate except for the transition rates. The edge labels may therefore be different between the two wings.

The graph K in Fig. 5A is taken from the paper of Banerjee et al. (2017), where it was used to model the discriminatory process during stepwise incorporation of monomers in mRNA translation and in DNA replication. For convenience, we use the same notation for the edge labels in K as in Banerjee et al. (2017) but we have followed our integer convention for vertices in place of the names used in Banerjee et al. (2017).

Banerjee *et al.* considered the ratio of splitting probabilities,

$$\eta = \frac{\pi_{1,6}(K)}{\pi_{1,7}(K)},$$

as a measure of accuracy in translation or replication. Theorem 4 tells us that,

$$\eta = \frac{w(\Phi_{\{1,7\} \rightarrow \{6,7\}}(K))}{w(\Phi_{\{1,6\} \rightarrow \{6,7\}}(K))}. \quad (64)$$

Enumerating the required spanning forests is considerably simplified by exploiting the symmetry of the butterfly graph. Let K_R and K_W be the subgraphs of K corresponding to the “right” and “wrong” wings of K , respectively (Fig. 5A). We can think of K as arising from the union of K_R and K_W by identifying them at the common vertex 1. We previously introduced the following notation (Wong et al. 2018),

$$K = K_W \oplus_1 K_R,$$

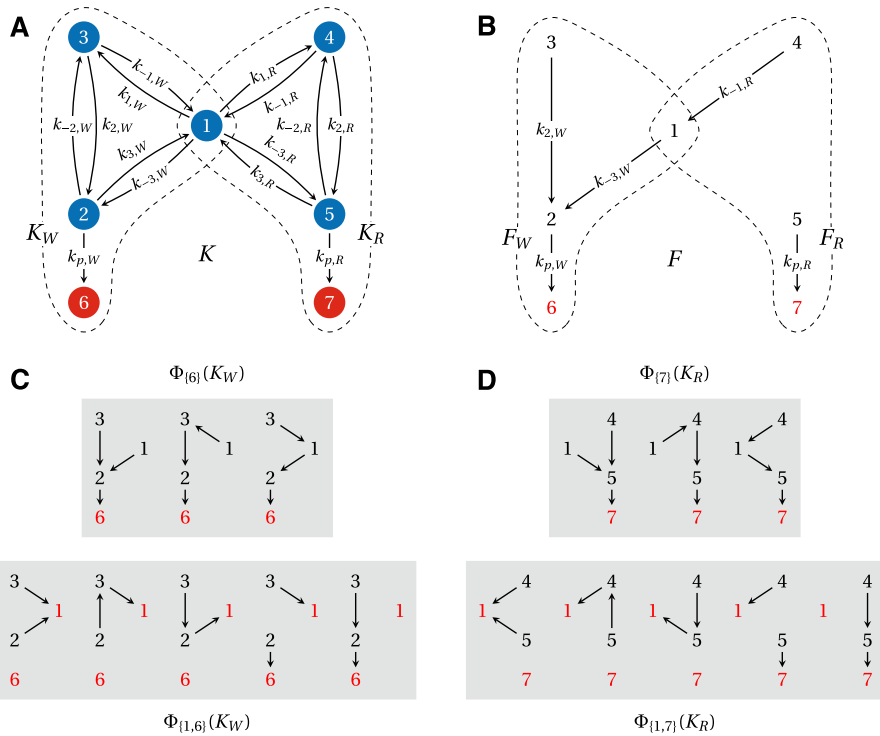


Fig. 5 (A) A butterfly graph, K , adapted from Banerjee et al. (2017, Fig. 1C and D), as discussed further in the text. K_R and K_W denote the “right” and “wrong” subgraphs, respectively. (B) A spanning forest $F \in \Phi_{\{6,7\}}(K)$, showing how it splits into a spanning tree, F_W , of K_W joined to a spanning forest, F_R , of K_R , as described in the text. (C, D) The spanning trees (top) and spanning forests (bottom) of K_W rooted at 1 and 6 and of K_R rooted at 1 and 7, respectively. Root vertices are in red font

for this joining operation on graphs, where the subscript on the operator \oplus specifies the common identified vertex. Now consider a spanning forest, $F \in \Phi_{\{6,7\}}(K)$. If 1 leads to 6 in F , so that $F \in \Phi_{\{1,7\} \rightarrow \{6,7\}}(K)$, as shown in Fig. 5B, then the other vertices of F in K_W must also lead to 6 in F . This means that the subgraph of F in K_W is a spanning tree of K_W rooted at 6, denoted $F_W \in \Phi_{\{6\}}(K_W)$. As for the vertices of K that lie in K_R , each such vertex must either lead to 6 or to 7 in F . If a vertex leads to 6, it must do so via 1. Therefore, the subgraph of F in K_R must be a spanning forest rooted at 1 and 7, denoted $F_R \in \Phi_{\{1,7\}}(K_R)$. We see that,

$$F = F_W \oplus_1 F_R.$$

A similar argument works when 1 leads to 7 in F , so that $F \in \Phi_{\{1,6\} \rightarrow \{6,7\}}(K)$. It follows that the spanning forests of K can be readily constructed from the spanning forests of K_W and K_R , which are, of course, structurally identical.

The weight of a graph that is obtained by the joining operation is easily seen to be the product of the weights of the corresponding subgraphs,

$$w(F) = w(F_W) w(F_R).$$

Summing over spanning forests, we see that,

$$w(\Phi_{\{1,7\} \rightarrow \{6,7\}}(K)) = w(\Phi_{\{6\}}(K_W)) w(\Phi_{\{1,7\}}(K_R)),$$

and by a similar argument,

$$w(\Phi_{\{1,6\} \rightarrow \{6,7\}}(K)) = w(\Phi_{\{1,6\}}(K_W)) w(\Phi_{\{7\}}(K_R)).$$

It follows from Eqn. 64 that,

$$\eta = \frac{w(\Phi_{\{6\}}(K_W)) w(\Phi_{\{1,7\}}(K_R))}{w(\Phi_{\{1,6\}}(K_W)) w(\Phi_{\{7\}}(K_R))},$$

where

$$\begin{aligned} w(\Phi_{\{6\}}(K_W)) &= k_{p,W} (k_{2,W} k_{-3,W} + k_{1,W} k_{2,W} + k_{-1,W} k_{-3,W}) \\ &= k_{p,W} (k_{2,W} (k_{-3,W} + k_{1,W}) + k_{-1,W} k_{-3,W}) \\ w(\Phi_{\{1,6\}}(K_W)) &= k_{-1,W} k_{3,W} + k_{-1,W} k_{-2,W} + k_{2,W} k_{3,W} + k_{-1,W} k_{p,W} + k_{2,W} k_{p,W} \\ &= (k_{-1,W} + k_{2,W}) (k_{3,W} + k_{p,W}) + k_{-1,W} k_{-2,W} \\ w(\Phi_{\{7\}}(K_R)) &= k_{p,R} (k_{2,R} k_{-3,R} + k_{1,R} k_{2,R} + k_{-1,R} k_{-3,R}) \\ &= k_{p,R} (k_{2,R} (k_{-3,R} + k_{1,R}) + k_{-1,R} k_{-3,R}) \\ w(\Phi_{\{1,7\}}(K_R)) &= k_{-1,R} k_{3,R} + k_{-1,R} k_{-2,R} + k_{2,R} k_{3,R} + k_{-1,R} k_{p,R} + k_{2,R} k_{p,R} \\ &= (k_{-1,R} + k_{2,R}) (k_{3,R} + k_{p,R}) + k_{-1,R} k_{-2,R}. \end{aligned}$$

The reader may check that this expression for η is identical to that given by Banerjee et al. (2017, Eqn. S9), which was calculated using Mathematica (Oleg Igoshin, private communication). The methods introduced here allow these calculations to be done by hand and reveal how the structure of the formula arises from the linear framework graph.

4 Discussion

The linear framework emerged in Gunawardena (2012) from the fundamental realisation that the steady-state vector of a Laplacian dynamical system, as described by Eqn. 1, may be expressed as manifestly positive rational algebraic functions of the edge labels of the corresponding graph (Eqn. 5), by virtue of the Matrix-Tree theorem (Eqn. 4). This algebraic access to the steady state has been invaluable because it has yielded insights without having to estimate the numerical values of the transition

rates, which, in the biological context, are usually unknown. The underlying graph also enables particular system features to be specified, while allowing the structure and the topology of the rest of the graph to remain arbitrary. This has enabled general theorems to be proved that rise above the ever-present molecular complexity that confronts us in biology. The impact of these capabilities has been reviewed in Nam et al. (2022).

The present paper has extended these capabilities to the transient regime in the microscopic setting, where the graph describes the infinitesimal generator of a Markov process and the Laplacian dynamics gives the corresponding master equation. The All-Minors Matrix-Tree theorem (Theorem 1) plays the key role, by showing how the inverse of a principal submatrix of the Laplacian may be calculated in terms of spanning forests (Proposition 2). The splitting probabilities to reach a set of target vertices (Theorem 4), the moments of the conditional FPT to a specific target vertex (Theorem 5) and the moments of the unconditional FPT to any target vertex (Theorem 6) may then be expressed as manifestly positive rational algebraic functions of the transition rates.

These results are timely. Timescale separation, in which a sub-system is assumed to have reached a steady state, has been a convenient approximation for analysing biochemical systems ever since Michaelis and Menten first introduced the method in 1910 (Gunawardena 2014). This has proved to be a good approximation for *in vitro* biochemistry but whether it remains appropriate within living cells is less clear (Wong and Gunawardena 2020). Moreover, as noted in the Introduction, a variety of single-molecule methods have made the transient regime experimentally accessible and FPTs have been widely used to analyse the resulting data (Kou et al. 2005; Shaevitz et al. 2005; Kolomeisky and Fisher 2007; Chemla et al. 2008; Garai et al. 2009; Bel et al. 2010; Moffitt et al. 2010; Moffitt and Bustamante 2014; Banerjee et al. 2017; Cui and Mehta 2018; Mallory et al. 2019; Wang et al. 2021; Lammers et al. 2020; Alamos et al. 2023; Lammers et al. 2023; Dal Co et al. 2017; Ghusinga et al. 2017; Gupta et al. 2018; Ham et al. 2024). Indeed, the most immediate application of the results presented here has been to analysing the interplay between specificity and speed in the CRISPR–Cas9 system, which we report on elsewhere (Nam et al. 2025).

Previous analyses of FPTs in the biological context have not yielded the kinds of closed-form mathematical formulas found in Theorems 4, 5 and 6. Accordingly, it has been difficult to extricate general principles and the underlying mathematical structure of FPTs as rational algebraic functions has been obscured. Our results make clear that, in this respect, the transient regime shares the same mathematical character as the steady state.

Notwithstanding this similarity, many interesting questions remain as to the distinctions between these two regimes. One such question concerns the role of energy expenditure. As noted in the Introduction, the formula for steady-state probabilities in Eqn. 5 reduces to the standard prescription of equilibrium statistical mechanics when the graph satisfies the *cycle condition* and thereby reaches a steady state of thermodynamic equilibrium (Gunawardena 2012; Ahsendorf et al. 2014; Estrada et al. 2016). This reduction has been particularly useful in characterising what we have called *Hopfield barriers*, or the limits to information processing at thermodynamic equilibrium (Estrada et al. 2016; Martinez-Corral et al. 2024). These limits have been

studied at steady state but an interesting problem emerges as to whether similar limits may exist in the transient regime. In other words, if a system is relaxing to a steady state of thermodynamic equilibrium, is its capability to process information fundamentally limited, compared to when it is relaxing to a non-equilibrium steady state? Preliminary calculations, on which we hope to report subsequently, suggest that this may be the case. However, such barriers seem to arise in the transient regime for quite different mathematical reasons to their appearance at steady state. For instance, the rational functions for splitting probabilities, in Theorem 4, and for FPT moments, in Theorems 5 and 6, do not appear to simplify in any obvious way when the cycle condition is satisfied, as is the case for steady-state probabilities. It remains a challenging open problem to understand how thermodynamic equilibrium may be influencing the mathematical structure of the rational algebraic formulas introduced here.

The most obvious mathematical distinction between the steady state and the transient regime, in the treatment we have provided here, is that the rational algebraic descriptions of the former arise from spanning trees, while those of the latter arise from spanning forests. This distinction, however, may not be a fundamental one. The pioneering biophysicist, Terrell Hill, suggested a way to express mean unconditional FPTs and splitting probabilities in terms of the steady-state probabilities of a modified Markov process (Hill 1988). (We revisited some of Hill's calculations in the examples treated above.) Similarly, Mark Kac pointed out that the mean recurrence times for a discrete-time Markov chain can be related to its steady-state probabilities (Kac 1947), and a similar result holds for continuous-time Markov processes (Serfozo 2009). These findings suggest a more systematic relationship between the transient and steady-state regimes and between spanning forests and spanning trees. We explore this further in a sequel to the present paper, in which we prove generalisations of the work of Hill and Kac (Nam and Gunawardena 2025).

The results presented here bring the transient regime of Markov processes within the same rational algebraic perspective that the linear framework previously developed for steady states. We hope this will encourage others to build upon the foundations laid here.

Acknowledgements We thank two anonymous reviewers for their helpful suggestions and members of the Gunawardena lab for their comments and feedback. K-MN and JG were supported by the US National Institutes of Health award R01GM122928.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data Availability This paper has no associated data.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

Table 1 Summary of the notation used in the paper

General mathematics

$\mathbf{0}, \mathbf{1}$	Zero or all-ones matrix or vector
\mathbf{I}	Identity matrix
$\#A$	Size of the set A
$\mathbf{M}_{[A,B]}$	Submatrix of \mathbf{M} with the rows indexed by A and columns indexed by B
$\theta_A : \{1, \dots, k\} \rightarrow A$	The ordering bijection on $A = \{a_1 < \dots < a_k\}$, $\theta_A(i) = a_i$ for $i = 1, \dots, k$
$\mathcal{L}\{f\}$	Laplace transform of f
<i>Graphs</i>	
G	Graph
$\mathcal{V}(G)$	Vertices of G , taken to be $\{1, \dots, N\}$ unless otherwise specified
N	Number of vertices in G , unless otherwise specified
$\mathcal{E}(G)$	Edges of G
$i \rightarrow j$, $i \rightarrow_G j$	Edge from vertex i to vertex j in G
$\ell(i \rightarrow j)$, $\ell(i \rightarrow_G j)$	Label on edge $i \rightarrow j$ in G , with units of $(\text{time})^{-1}$
\overline{A}	Complement of the vertex subset $A \subseteq \mathcal{V}(G)$
$\mathcal{L}(G)$	Laplacian matrix of G (Eqn. 2)
$\mathbf{L}(G)$	$-\mathcal{L}(G)^T$ (Eqn. 41)
λ_i	Sum of outgoing edge labels from i
$V_{\text{out}}(i)$	Vertices to which there is an edge from i , $\{j \in \mathcal{V}(G) : i \rightarrow j\}$
$V_{\text{in}}(i)$	Vertices from which there is an edge to i , $\{j \in \mathcal{V}(G) : j \rightarrow i\}$
$i \rightsquigarrow j$, $i \rightsquigarrow_G j$	Existence of a directed path of edges from i to j in G
$\mathcal{V}_{\text{out}}(i)$	Vertices to which there is a directed path of edges from i , $\{j \in \mathcal{V}(G) : i \rightsquigarrow j\}$
$\mathcal{V}_{\text{in}}(i)$	Vertices from which there is a directed path of edges to i , $\{j \in \mathcal{V}(G) : j \rightsquigarrow i\}$
$\mathcal{V}_{\text{in}}^-(i)$	$\mathcal{V}_{\text{in}}(i) \setminus \{i\}$
$w(\cdot)$	The weight, or product of edge labels, of a subgraph, or the sum of those for a set of subgraphs
$\mathcal{C}(G)$	Strongly connected components (SCCs) of G
$\mathcal{T}(G)$	Terminal SCCs of G
T	Number of terminal SCCs in G

Table 1 continued

G^σ	Graph obtained from G by re-indexing the vertices according to a permutation, $\sigma : \mathcal{V}(G) \rightarrow \mathcal{V}(G)$
$\tau_{\sigma,A}$	Given $A \subseteq \mathcal{V}(G)$ and a permutation $\sigma : \mathcal{V}(G) \rightarrow \mathcal{V}(G)$, the permutation $\theta_{\sigma(A)}^{-1}(\sigma(\theta_A))$ on $\{1, \dots, \#A\}$ (Eqn. 14)
Z	If G is a graph with terminal vertices (singleton terminal SCCs), the union $\bigcup \mathcal{T}(G)$; taken to be $\{N - T + 1, \dots, N\}$ without loss of generality
$\mathcal{Z}(i)$	Subset of terminal vertices z such that $i \rightsquigarrow z$ (Eqn. 31)
N_z	Number of non-terminal vertices that lead to the terminal vertex z
<i>Spanning trees and forests</i>	
$\mathcal{R}(F)$	Set of roots of a spanning forest F
$\Phi_A(G)$	Set of spanning forests of G rooted at $A \subseteq \mathcal{V}(G)$
$\Phi_{B \rightarrow A}(G)$	Assuming $\#A = \#B$, the set of spanning forests of G rooted at $A \subseteq \mathcal{V}(G)$ in which each $b \in B \subseteq \mathcal{V}(G)$ has a path to a distinct root in A
$\chi_F : B \rightarrow A$	Given a spanning forest $F \in \Phi_{B \rightarrow A}(G)$, the bijection $B \rightarrow A$ that maps each $b \in B$ to the root in A to which it has a path (Theorem 1)
$\eta : A \rightarrow B$	Given vertex subsets $A, B \subseteq \mathcal{V}(G)$ of the same size, the bijection $A \rightarrow B$ that maps the i -th vertex in A to the i -th vertex in B (Theorem 1)
<i>Markov processes</i>	
$X(\cdot)$	Markov process associated with G
$p_{i,z}(t)$	Probability that $X(\cdot)$ reaches the terminal vertex z by time t (Eqn. 25)
$\pi_{i,z}$	Splitting probability from i to z in G (Eqn. 26)
$p_{c,i,z}(t)$	Probability that $X(\cdot)$ reaches the terminal vertex z by time t , conditioned on z being reached (Eqn. 27)
$p_{i,Z}(t)$	Probability that $X(\cdot)$ reaches some terminal vertex by time t (Eqn. 29)
$u_{i,z}(t)$	Density corresponding to $p_{i,z}(t)$ (Eqn. 35)
$\mu_{i,z}^{(r)}, \mu_{c,i,z}^{(r)}, \mu_{i,Z}^{(r)}$	r -th moments of $p_{i,z}(t)$ (Eqn. 43), $p_{c,i,z}(t)$ (Eqn. 50), and $p_{i,Z}(t)$ (Eqn. 53)

References

- Gunawardena J (2012) A linear framework for time-scale separation in nonlinear biochemical systems. *PLoS ONE* 7:36321. <https://doi.org/10.1371/journal.pone.0036321>
- Mirzaev I, Gunawardena J (2013) Laplacian dynamics on general graphs. *Bull Math Biol* 75:2118–2149. <https://doi.org/10.1007/s11538-013-9884-8>
- Nam K-M, Martinez-Corral R, Gunawardena J (2022) The linear framework: using graph theory to reveal the algebra and thermodynamics of biomolecular systems. *Interface Focus* 12:20220013. <https://doi.org/10.1098/rsfs.2022.0013>
- Chung F (1997) *Spectral Graph Theory*. American Mathematical Society, Providence, RI, USA
- Mirzaev I, Bortz DM (2015) Laplacian dynamics with synthesis and degradation. *Bull Math Biol* 77:1013–45. <https://doi.org/10.1007/s11538-015-0075-7>
- Nam K-M, Gunawardena J (2023) The linear framework II: using graph theory to analyse the transient regime of Markov processes. *Front Cell Dev Biol* 11:1233808. <https://doi.org/10.3389/fcell.2023.1233808>
- Dasgupta T, Croll DH, Owen JA, Vander Heiden MG, Locasale JW, Alon U, Cantley LC, Gunawardena J (2014) A fundamental trade-off in covalent switching and its circumvention by enzyme bifunctionality in glucose homeostasis. *J Biol Chem* 289:13010–13025. <https://doi.org/10.1074/jbc.M113.546515>
- Nam K-M, Gyori BM, Amethyst SV, Bates DJ, Gunawardena J (2020) Robustness and parameter geography in post-translational modification systems. *PLoS Comput Biol* 16:1007573. <https://doi.org/10.1371/journal.pcbi.1007573>
- Yordanov P, Stelling J (2018) Steady-state differential dose response in biological systems. *Biophys J* 114:723–36. <https://doi.org/10.1016/j.bpj.2017.11.3780>
- Yordanov P, Stelling J (2020) Efficient manipulation and generation of Kirchhoff polynomials for the analysis of non-equilibrium biochemical reaction networks. *J R Soc Interface* 17:20190828. <https://doi.org/10.1098/rsif.2019.0828>
- Anderson WJ (1991) *Continuous-Time Markov Chains: An Applications-Oriented Approach*. Springer, New York, NY, USA
- Norris JR (1997) *Markov Chains*. Cambridge University Press, Cambridge, United Kingdom
- Ahsendorf T, Wong F, Eils R, Gunawardena J (2014) A framework for modelling gene regulation which accommodates non-equilibrium mechanisms. *BMC Biol* 12:102. <https://doi.org/10.1186/s12915-014-0102-4>
- Estrada J, Wong F, DePace A, Gunawardena J (2016) Information integration and energy expenditure in gene regulation. *Cell* 166:234–244. <https://doi.org/10.1016/j.cell.2016.06.012>
- Biddle JW, Nguyen M, Gunawardena J (2019) Negative reciprocity, not ordered assembly, underlies the interaction of Sox2 and Oct4 on DNA *eLife* 8:41017. <https://doi.org/10.7554/eLife.41017>
- Park J, Estrada J, Johnson G, Vincent BJ, Ricci-Tam C, Bragdon MDJ, Shulgina Y, Cha A, Wunderlich Z, Gunawardena J, DePace AH (2019) Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity. *eLife* 8, 41266. <https://doi.org/10.7554/eLife.41266>
- Martinez-Corral R, Nam K-M, DePace AH, Gunawardena J (2024) The Hill function is the universal Hopfield barrier for sharpness of input-output responses. *Proc Natl Acad Sci USA* 121:2318329121. <https://doi.org/10.1073/pnas.2318329121>
- Nasser J, Nam K-M, Gunawardena J (2024) A mathematical model clarifies the ABC Score formula used in enhancer-gene regulation. *Elife* 14:105740. <https://doi.org/10.7554/eLife.105740.1>
- Wong F, Gunawardena J (2020) Gene regulation in and out of equilibrium. *Annu Rev Biophys* 49:199–226. <https://doi.org/10.1146/annurev-biophys-121219-081542>
- van Kampen NG (2007) *Stochastic Processes in Physics and Chemistry*, 3rd edn. North Holland, Amsterdam, Netherlands
- Gunawardena J (2014) Time-scale separation: Michaelis and Menten's old idea, still bearing fruit. *FEBS J* 281:473–88. <https://doi.org/10.1111/febs.12532>
- Williams LK (2022) Combinatorics of hopping particles and positivity in Markov chains. *LMS Newsletter* 500:50–59
- Hill TL (1966) Studies in irreversible thermodynamics IV. Diagrammatic representation of steady state fluxes for unimolecular systems. *J Theoret Biol* 10:442–59. [https://doi.org/10.1016/0022-5193\(66\)90137-8](https://doi.org/10.1016/0022-5193(66)90137-8)
- Schnakenberg J (1976) Network theory of microscopic and macroscopic behaviour of master equation systems. *Rev Mod Phys* 48:571–86. <https://doi.org/10.1103/RevModPhys.48.571>

- Andrieux D, Gaspard P (2007) Fluctuation theorem for currents and Schnakenberg network theory. *J Stat Phys* 127:107–31. <https://doi.org/10.1007/s10955-006-9233-5>
- Murugan A, Huse DA, Leibler S (2014) Discriminatory proofreading regimes in nonequilibrium systems. *Phys Rev X* 4:021016. <https://doi.org/10.1103/PhysRevX.4.021016>
- Owen JA, Horowitz JM (2023) Size limits the sensitivity of kinetic schemes. *Nat Commun* 14:1280. <https://doi.org/10.1038/s41467-023-36705-8>
- Dal Cengio S, Lecomte V, Poletti M (2023) Geometry of nonequilibrium reaction networks. *Phys Rev X* 13:021040. <https://doi.org/10.1103/PhysRevX.13.021040>
- Qureshi B, Juritz J, Poulton JM, Beersing-Vasquez A, Ouldrige TE (2023) A universal method for analyzing copolymer growth. *J Chem Phys* 158:104906. <https://doi.org/10.1063/5.0133489>
- Wong F, Dutta A, Chowdhury D, Gunawardena J (2018) Structural conditions on complex networks for the Michaelis-Menten input-output response. *Proc Natl Acad Sci USA* 115:9738–9743. <https://doi.org/10.1073/pnas.1808053115>
- Biddle JW, Martinez-Corral R, Wong F, Gunawardena J (2021) Allosteric conformational ensembles have unlimited capacity for integrating information. *Elife* 10:65498. <https://doi.org/10.7554/eLife.65498>
- Kou SC, Cherayil BJ, Min W, English BP, Xie XS (2005) Single-molecule Michaelis-Menten equations. *J Phys Chem B* 109:19068–19081. <https://doi.org/10.1021/jp051490q>
- Shaevitz JW, Block SM, Schnitzer MJ (2005) Statistical kinetics of macromolecular dynamics. *Biophys J* 89:2277–2285. <https://doi.org/10.1529/biophysj.105.064295>
- Kolomeisky AB, Fisher ME (2007) Molecular motors: a theorist's perspective. *Annu Rev Phys Chem* 58:675–695. <https://doi.org/10.1146/annurev.physchem.58.032806.104532>
- Chemla YR, Moffitt JR, Bustamante C (2008) Exact solutions for kinetic models of macromolecular dynamics. *J Phys Chem B* 112:6025–6044. <https://doi.org/10.1021/jp076153r>
- Garai A, Chowdhury D, Chowdhury D, Ramakrishnan TV (2009) Stochastic kinetics of ribosomes: single motor properties and collective behavior. *Phys Rev E* 80:011908. <https://doi.org/10.1103/PhysRevE.80.011908>
- Bel G, Minsky B, Nemenman I (2010) The simplicity of completion time distributions for common complex biochemical processes. *Phys Biol* 7:016003. <https://doi.org/10.1088/1478-3975/7/1/016003>
- Moffitt JR, Chemla YR, Bustamante C (2010) Mechanistic constraints from the substrate concentration dependence of enzymatic fluctuations. *Proc Natl Acad Sci USA* 107:15739–15744. <https://doi.org/10.1073/pnas.1006997107>
- Moffitt JR, Bustamante C (2014) Extracting signal from noise: kinetic mechanisms from a Michaelis-Menten-like expression for enzymatic fluctuations. *FEBS J* 281:498–517. <https://doi.org/10.1111/febs.12545>
- Banerjee K, Kolomeisky AB, Igoshin OA (2017) Elucidating interplay between speed and accuracy in biological error correction. *Proc Natl Acad Sci USA* 114:5183–5188. <https://doi.org/10.1073/pnas.1614838114>
- Cui W, Mehta P (2018) Identifying feasible operating regimes for early T-cell recognition: The speed, energy, accuracy trade-off in kinetic proofreading and adaptive sorting. *PLoS ONE* 13:0202331. <https://doi.org/10.1371/journal.pone.0202331>
- Mallory JD, Kolomeisky AB, Igoshin OA (2019) Trade-offs between error, speed, noise, and energy dissipation in biological processes with proofreading. *J Phys Chem B* 123:4718–4725. <https://doi.org/10.1021/acs.jpcc.9b03757>
- Wang Q, Yang J, Zhong Z, Vanegas JA, Gao X, Kolomeisky AB (2021) A general theoretical framework to design base editors with reduced bystander effects. *Nat Commun* 12:6529. <https://doi.org/10.1038/s41467-021-26789-5>
- Lammers NC, Kim YJ, Zhao J, Garcia HG (2020) A matter of time: Using dynamics and theory to uncover mechanisms of transcriptional bursting. *Curr Opin Cell Biol* 67:147–157. <https://doi.org/10.1016/j.ccb.2020.08.001>
- Alamos S, Reimer A, Westrum C, Turner MA, Talledo P, Zhao J, Luu E, Garcia HG (2023) Minimal synthetic enhancers reveal control of the probability of transcriptional engagement and its timing by a morphogen gradient. *Cell Syst* 14:1–17. <https://doi.org/10.1016/j.cels.2022.12.008>
- Lammers NC, Flamholz AI, Garcia HG (2023) Competing constraints shape the nonequilibrium limits of cellular decision-making. *Proc Natl Acad Sci USA* 120:2211203120. <https://doi.org/10.1073/pnas.2211203120>
- Dal Co A, Lagomarsino MC, Caselle M, Osella M (2017) Stochastic timing in gene expression for simple regulatory strategies. *Nucleic Acids Res* 45:1069–1078. <https://doi.org/10.1093/nar/gkw1235>

- Ghusinga KR, Dennehy JJ, Singh A (2017) First-passage time approach to controlling noise in the timing of intracellular events. *Proc Natl Acad Sci USA* 114:693–698. <https://doi.org/10.1073/pnas.1609012114>
- Gupta S, Varennes J, Korswagen HC, Mugler A (2018) Temporal precision of regulated gene expression. *PLoS Comput Biol* 14:1006201. <https://doi.org/10.1371/journal.pcbi.1006201>
- Ham L, Coomer MA, Öcal K, Grima R, Stumpf MPH (2024) A stochastic vs deterministic perspective on the timing of cellular events. *Nat Commun* 15:5286. <https://doi.org/10.1038/s41467-024-49624-z>
- Guntoro JEB, Qureshi BJ, Ouldrige TE (2025) The interplay of heterogeneity and product detachment in templated polymer copying. *J Chem Phys* 162:055103. <https://doi.org/10.1063/5.0245687>
- Hill TL (1988) Interrelations between random walks on diagrams (graphs) with and without cycles. *Proc Natl Acad Sci USA* 85:2879–2883. <https://doi.org/10.1073/pnas.85.9.2879>
- Chebotarev P (2007) A graph theoretic interpretation of the mean first passage times [arXiv:math/0701359](https://arxiv.org/abs/math/0701359)
- Pitman J, Tang W (2018) Tree formulas, mean first passage times and Kemeny's constant of a Markov chain. *Bernoulli* 24:1942–1972. <https://doi.org/10.3150/16-BEJ916>
- Kac M (1947) On the notion of recurrence in discrete stochastic processes. *Bull Amer Math Soc* 53:1002–1010
- Nam K-M, Gunawardena J (2025) Algebraic formulas for first-passage times of Markov processes in the linear framework: generalising the work of Hill and Kac. In preparation
- Nam K-M, Padula S, Gunawardena J (2025) Interplay between specificity and speed in target recognition by CRISPR–Cas9. In preparation
- Nam K-M (2021) Algebraic approaches to molecular information processing. PhD thesis, Harvard University
- Herstein IN (1975) *Topics in Algebra*, 2nd edn. John Wiley & Sons, New York, NY, USA
- Fiedler M, Sedláček J (1958) O w -basích orientovaných grafu. *Časopis pro pěstování matematiky* 83:214–225
- Chaiken S (1982) A combinatorial proof of the all minors matrix tree theorem. *SIAM J Alg Disc Meth* 3:319–329. <https://doi.org/10.1137/0603033>
- Moon JW (1994) Some determinant expansions and the matrix-tree theorem. *Disc Math* 124:163–171. [https://doi.org/10.1016/0012-365X\(92\)00059-Z](https://doi.org/10.1016/0012-365X(92)00059-Z)
- Serfozo R (2009) *Basics of Applied Stochastic Processes*. Springer, Heidelberg, Germany
- Iyer-Biswas S, Zilman A (2016) First-passage processes in cellular biology. In: *Advances in Chemical Physics*, John Wiley & Sons, Ltd, Hoboken, NJ, USA. <https://doi.org/10.1002/9781119165156.ch5>
- Bernstein DS (2009) *Matrix Mathematics*. Princeton University Press, Princeton, NJ, USA
- Wong F, Amir A, Gunawardena J (2018) Energy-speed-accuracy relation in complex networks for biological discrimination. *Phys Rev E* 98:012420. <https://doi.org/10.1103/PhysRevE.98.012420>
- Hopfield JJ (1974) Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci USA* 71:4135–4139. <https://doi.org/10.1073/pnas.71.10.4135>
- Murugan A, Huse DA, Leibler S (2014) Discriminatory proofreading regimes in nonequilibrium systems. *Phys Rev X* 4:021016. <https://doi.org/10.1103/PhysRevX.4.021016>
- Çetiner U, Gunawardena J (2022) Reformulating non-equilibrium steady states and generalized Hopfield discrimination. *Phys Rev E* 106:064128. <https://doi.org/10.1103/PhysRevE.106.064128>