Bull Math Biol
DOI 10.1007/s11538-013-9884-8

Society for Mathematical Biology

ORIGINAL ARTICLE

Laplacian Dynamics on General Graphs

Inomzhon Mirzaev · Jeremy Gunawardena

Received: 13 May 2013 / Accepted: 23 July 2013

Abstract In previous work, we have introduced a "linear framework" for time-scale separation in biochemical systems, which is based on a labelled, directed graph, G, and an associated linear differential equation, $dx/dt = \mathcal{L}(G) \cdot x$, where $\mathcal{L}(G)$ is the Laplacian matrix of G. Biochemical nonlinearity is encoded in the graph labels. Many central results in molecular biology can be systematically derived within this framework, including those for enzyme kinetics, allosteric proteins, G-protein coupled receptors, ion channels, gene regulation at thermodynamic equilibrium, and protein post-translational modification. In the present paper, in response to new applications, which accommodate nonequilibrium mechanisms in eukaryotic gene regulation, we lay out the mathematical foundations of the framework. We show that, for any graph and any initial condition, the dynamics always reaches a steady state, which can be algorithmically calculated. If the graph is not strongly connected, which may occur in gene regulation, we show that the dynamics can exhibit flexible behavior that resembles multistability. We further reveal an unexpected equivalence between deterministic Laplacian dynamics and the master equations of continuous-time Markov processes, which allows rigorous treatment within the framework of stochastic, single-molecule mechanisms.

Keywords Time-scale separation · Linear framework · Graph Laplacian · Matrix-Tree theorem · Gene regulation · Nonequilibrium mechanisms · Markov process · Master equation

I. Mirzaev
Applied Mathematics Graduate Program, University of Colorado, Boulder, CO, USA
e-mail: mirzaev@colorado.edu

J. Gunawardena (✉)
Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA
e-mail: jeremy@hms.harvard.edu

Published online: 10 September 2013

1 Introduction

Time-scale separation is one of the few systematic methods for analytically reducing the complexity of a system. In this procedure, a subsystem is assumed to be operating sufficiently fast relative to the rest of the system that it may be taken to have reached a steady state. The *fast variables* within the subsystem can then be eliminated, in favor of the remaining *slow variables*. In previous work, we introduced a mathematical framework for undertaking time-scale separation in biochemical systems (Thomson and Gunawardena 2009a, 2009b; Gunawardena 2012). It is based on a directed graph, G, with labeled edges and no self-loops (hereafter, a "graph"), as in Fig. 1. In the applications, the vertices of the graph represent the fast variables in the subsystem and the edges represent the biochemical interactions between these variables. Such a graph gives rise to a dynamical system by treating each edge as a chemical reaction under mass-action kinetics, with the corresponding label as the rate constant. Since each edge has only one source vertex, the corresponding reaction is first order and the dynamics is linear, giving rise to the matrix equation,

$$\frac{dx}{dt} = \mathcal{L}(G) \cdot x. \tag{1}$$

$\mathcal{L}(G)$ is called the Laplacian matrix of G. Here, $x = (x_1, \ldots, x_n)^t$ is a column vector of concentrations at each vertex, $1, \ldots, n$, of G and the notation x^t denotes the transpose of a vector x.

$\mathcal{L}(G)$ has the following structure (Fig. 1). If there is an edge in G from vertex j to vertex i, $j \to i$, let the label on this edge be e_{ij}, which should be a positive number with units of time^{-1}. It there is no such edge, let $e_{ij} = 0$. Note the reversal of the indices, which is necessary because we use column vectors. The use of positive labels is not a restriction, as a negative label amounts to an edge in the opposite direction. It follows that

$$\mathcal{L}(G)_{ij} = \begin{cases} e_{ij} & \text{if } i \neq j, \\ -\sum_{v \neq j} e_{vj} & \text{if } i = j. \end{cases} \tag{2}$$

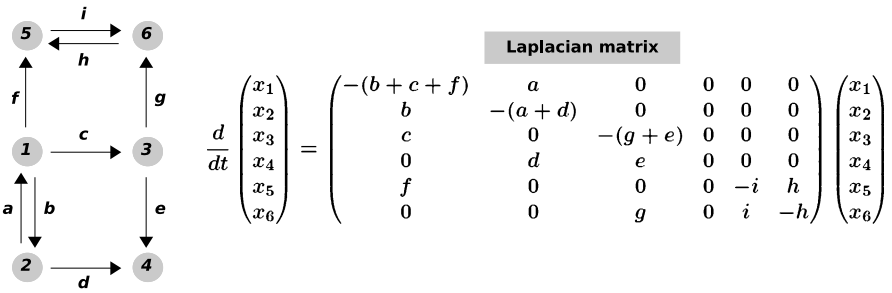


Fig. 1 Laplacian dynamics on a graph. On the left, a labeled, directed graph with indices $1, \ldots, 6$ and edges labeled, $a, \ldots, i$. Note that the graph is not strongly connected. On the right, the corresponding Laplacian dynamics, as in Eq. (1), with the Laplacian matrix of the graph

Laplacian matrices were first introduced by Gustav Kirchhoff in his pioneering study of electrical networks (Kirchhoff 1847) and they have been widely studied under different guises, as discussed below.

The key to applying such a "linear framework" to nonlinear biochemical systems lies in the edge labels, e_{ij}, which can be algebraic expressions involving the concentrations of slow variables and the rate constants of biochemical reactions. In effect, the linear framework trades nonlinear dynamics with simple rate constants for linear dynamics with complex labels. This allows steady states to be calculated and the fast variables to be eliminated, as in the example discussed below.

Many well-known calculations in molecular and systems biology, which have been undertaken by distinct, *ad hoc* methods, can be systematically derived using this linear framework, as explained in Gunawardena (2012). These include the Michaelis–Menten and King–Altman procedures in enzyme kinetics (Michaelis and Menten 1913; King and Altman 1956; Cornish-Bowden 1995); the Monod–Wyman–Changeux and Koshland–Némethy–Filmer formulas in protein allostery (Monod et al. 1965; Koshland et al. 1966) and models derived from these to study G-protein coupled receptors (Lean et al. 1980; Kenakin 2005) and ligand-gated ion channels (Colquhoun 2006); and the Ackers–Johnson–Shea (Ackers et al. 1982) and related formulas based on the "thermodynamic formalism" (Bintu et al. 2005a, 2005b; Segal and Widom 2009; Sherman and Cohen 2012), which have been used to study gene regulation in bacteria (Setty et al. 2003; Kuhlman et al. 2007), yeast (Gertz et al. 2009) and flies (Janssens et al. 2006; Zinzen et al. 2006; Segal et al. 2008; He et al. 2010). More recently, the linear framework has been used to analyze the steady-state behavior of post-translational modification systems (Thomson and Gunawardena 2009a, 2009b; Xu and Gunawardena 2012; Dasgupta et al. 2013).

To introduce the linear framework and provide background for the rest of the paper, the classical Michaelis–Menten formula provides a convenient example. Michaelis and Menten considered the reaction scheme

$$S + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} P + E, \tag{3}$$

in which k_1, k_2, and k_3 are the rate constants for mass-action kinetics. They made the time-scale separation that the enzyme forms E and ES were fast variables, while substrate S and product P were slow. (The historical details are somewhat different; see Cornish-Bowden 1995.) To apply the linear framework, the following labeled, directed graph G is constructed,

$$\textcircled{E} \underset{k_2+k_3}{\overset{k_1[S]}{\rightleftarrows}} \textcircled{ES} \tag{4}$$

in which the vertices represent the fast variables, E and ES, the edges describe the transitions between these forms and the labels encode the reactions, along with the contributions of the slow variables. It can easily be checked that, with this labeling, the Laplacian dynamics on G recapitulates the actual dynamics of E and ES coming from the reactions in Eq. (3). At steady state, the fluxes must balance, so that

$$k_1[S][E] = (k_2 + k_3)[ES], \tag{5}$$

which gives one equation between the two fast variables. This example is so simple that it is not necessary to construct the graph G in order to deduce Eq. (5) but, in the absence of G, the linearity of this steady-state calculation is less obvious. A second equation between the fast variables comes from observing that enzyme is neither created nor destroyed and must therefore obey the conservation law

$$[E] + [ES] = E_{\text{tot}}. \tag{6}$$

Using Eqs. (5) and (6) to solve for $[E]$ and $[ES]$, we find that

$$[E] = \frac{E_{\text{tot}} K_M}{K_M + [S]} \quad \text{and} \quad [ES] = \frac{E_{\text{tot}}[S]}{K_M + [S]}, \tag{7}$$

where $K_M = (k_2 + k_3)/k_1$ is the Michaelis–Menten constant. Equation (7) allows the fast variables to be eliminated in favour of the conserved total, E_{tot}, and the quantities derived from the labels, such as $[S]$ and K_M. The rate of the reaction can now be calculated as

$$\frac{d[P]}{dt} = k_3[ES] = \frac{V_{\max}[S]}{K_M + [S]},$$

where $V_{\max} = k_3 E_{\text{tot}}$ is the maximal reaction velocity. This is the well-known Michaelis–Menten formula, in which only the slow variables appear.

The main difference between this calculation and other applications of the framework is that the graph G can be far more complicated and the steady state can no longer be determined by simply balancing two fluxes. Instead, it becomes necessary to calculate x such that $dx/dt = 0$ or, equivalently, x is in the kernel of the Laplacian: $\ker \mathcal{L}(G) = \{x \mid \mathcal{L}(G) \cdot x = 0\}$. In many applications, it has been found that G is *strongly connected*. That is, given any two distinct vertices, i and j, there is a path in G of directed edges from i to j,

$$i = i_1 \to i_2 \to \cdots \to i_k = j.$$

(Since i and j are arbitrary, there must also be such a path from j to i.) Note that the graph in Eq. (4) is strongly connected. When G is strongly connected, it can be shown that the kernel of the Laplacian is one- dimensional,

$$\dim \ker \mathcal{L}(G) = 1. \tag{8}$$

This result has been central to previous applications of the linear framework. It is proved in Thomson and Gunawardena (2009a) and reappears in the course of the treatment here as Proposition 3. It has the following implication. The Laplacian dynamics may be started with arbitrary concentrations at each vertex, so that there are as many degrees of freedom initially as there are vertices in the graph. Nevertheless, if a steady state is reached, all these degrees of freedom are lost except one. This is true no matter how complex the graph, provided only that it is strongly connected. It is this collapse in the degrees of freedom which allows elimination of the fast variables.

To obtain formulas, it is necessary to identify a basis element in $\ker \mathcal{L}(G)$. A canonical element, $\rho \in \ker \mathcal{L}(G)$, may be algorithmically calculated from the graph

structure using the Matrix-Tree theorem, which is stated below as Theorem 1. The ρ_i are sums of products of labels (i.e.: polynomials in the labels) that are determined by the structure of G. For G in Eq. (4), if E and ES have the indices 1 and 2, respectively, then $\rho_1 = k_2 + k_3$ and $\rho_2 = k_1[S]$.

Any steady state, $x \in \ker \mathcal{L}(G)$, can now be determined in terms of ρ. According to Eq. (8), $x = \lambda\rho$ where λ is a scalar. This gives n equations, $x_i = \lambda\rho_i$, for the $n+1$ unknown quantities, $x_1, \ldots, x_n$, and λ. The remaining equation comes from the fact that the total concentration of material on the graph is conserved, so that

$$x_1 + \cdots + x_n = x_{\text{tot}} \tag{9}$$

which is the equivalent of Eq. (6) above. Substituting $x_i = \lambda\rho_i$, gives

$$\lambda = \frac{x_{\text{tot}}}{\rho_1 + \cdots + \rho_n},$$

from which the x_i can be calculated as

$$x_i = \frac{x_{\text{tot}}\rho_i}{\rho_1 + \cdots + \rho_n}. \tag{10}$$

This is the equivalent of Eq. (7): the x_i have been eliminated in terms of the conserved quantity x_{tot} and the quantities in the labels, which appear through the ρ_i. The various formulas and models cited above in different areas of biology, in which time-scale separation has been used to simplify systems and calculate properties of interest, can all be deduced from Eq. (10) (Gunawardena 2012).

Up to now, the focus of studies on the linear framework has been on steady-state behavior in strongly-connected graphs and several foundational questions have not been addressed. First, it has been generally assumed that no matter what initial condition is chosen, the Laplacian dynamics eventually reaches a steady state. While this seems intuitively plausible, the dynamical behavior of the Laplacian has not been rigorously analysed. Second, in some application areas, such as gene regulation or ion channel activity, the system is stochastic in nature, with the vertices of the graph representing the states of individual molecules, rather than populations of molecules. The relationship between the macroscopic, deterministic Laplacian dynamics in Eq. (1) and stochastic dynamics has been treated in an *ad hoc* manner, but has not been rigorously clarified. The present paper reveals an unexpectedly close relationship between these two apparently dissimilar formulations.

A further important issue has emerged in recent application of the framework to gene regulation away from thermodynamic equilibrium (Ahsendorf et al. 2013). This becomes particularly significant in eukaryotes where dissipative mechanisms play a central role, as reviewed in the Discussion. In contrast to previous applications of the linear framework, nonstrongly connected graphs emerge in this new context. For these, the kernel of the Laplacian need no longer be one-dimensional, and we show here that this can lead to significant qualitative differences in steady-state behavior from the strongly-connected case.

The present paper is intended to address these foundational issues. Applications of the results to gene regulation will appear elsewhere (Ahsendorf et al. 2013).

Two further points are relevant to our general approach. First, the Laplacian dynamics in Eq. (1) is linear and the behavior of linear systems is well understood. For instance, there are standard methods in linear algebra for calculating solutions of the steady-state matrix equation $\mathcal{L}(G) \cdot x = 0$. However, these methods use determinants, which have terms with alternating signs, and it is not immediately clear that the resulting solutions have the important property of *positivity* (Thomson and Gunawardena 2009b, Supplementary Information). That is, if the rate constants and total concentrations are all positive, it is important to confirm that quantities like ρ_i above are also positive. The Matrix-Tree theorem guarantees positivity. In general, the interplay between dynamics, algebra, and combinatorics has been highly informative and that approach is continued here.

Second, matrices of Laplacian type have arisen in both fundamental graph theory (Merris 1994; Chung 1997; Chebotarev and Agaev 2002) and in various application areas in physics, engineering, computer science, and economics (Schnakenberg 1976; Pecora and Carroll 1998; Nishikawa and Motter 2010; Chen 1971; Olfati-Saber et al. 2007; Chebotarev and Agaev 2009; Bott and Mayberry 1954). While some of the algebraic results proved here can be found in these other literatures, the present treatment is necessary for several reasons. Laplacian dynamics, in the form of the first-order biochemistry represented by Eq. (1), has not been discussed previously. Determining the long-time stability of this dynamics is an important goal of the present paper. Laplacian matrices have also been defined for many different types of graphs (undirected, unlabeled, etc.) under different notations, conventions, and normalizations. The treatment given here builds upon previous work (Thomson and Gunawardena 2009a, 2009b; Gunawardena 2012), and follows uniform conventions that are grounded in biochemical kinetics. Other sources often focus on particular kinds of graphs, such as strongly-connected ones, and address questions of interest to other domains. No restrictions are placed here on the graphs being studied and the focus is on issues of interest to biology. Finally, literature sources that are both complete and accessible to a biological audience have not been found. The treatment given here is self-contained and elementary, relying largely on well-known results in linear algebra. We cite relevant sources when these are known to us but must apologize in advance if we have inadvertently missed others. We would be pleased to know of these. We hope that the present paper will provide a helpful resource and encourage further exploitation of the linear framework in different areas of biology.

2 Eigenvalues of the Laplacian

We outline some conventions before embarking on the main results. In applications of the linear framework, the labels can be algebraic expressions involving actual biochemical rate constants and concentrations of slow variables, as in Eq. (4). It can be helpful to work over a rational function field $\mathbb{R}(a_1, \ldots, a_m)$ in these other parameters, rather than over $\mathbb{R}$; see, for instance, Thomson and Gunawardena (2009a). Here, for simplicity, we leave these issues to specific applications and work throughout over $\mathbb{R}$. Labels are therefore taken to be uninterpreted symbols which represent positive real numbers and have units of time^{-1}. The word "graph" will always mean *non-empty,*

labelled, directed graph with no self-loops. The symbol n will be used generically for the number of vertices in the graph under discussion. I will denote the identity matrix. Within a matrix multiplication, 1 will denote the all-ones vector; the context should make clear whether this is a row vector or a column vector. A subscript will be added, as in I_n or 1_n, to specify the dimension, when needed.

We recall some well-known facts about linear systems (Hirsch and Smale 1974). Let A be a $n \times n$ matrix with entries in the real numbers, $A_{ij} \in \mathbb{R}$. The linear matrix equation $dx/dt = A \cdot x$ has the solution

$$x(t) = \exp(At) \cdot x(0),$$

where $x(0)$ is the initial condition and $\exp(X)$ is the matrix exponential

$$\exp(X) = I + X + \frac{X^2}{2} + \frac{X^3}{3.2} + \cdots + \frac{X^m}{m!} + \cdots . \tag{11}$$

The exponential can be calculated from the Jordan normal form of A, from which it follows that the solution, $x_i(t)$, is a linear combination, over the complex numbers, of terms of the form

$$t^{\mu} \exp(\lambda t). \tag{12}$$

Here, μ is a nonnegative integer, $\mu = 0, 1, \ldots$, and λ is an eigenvalue of A. A positive power of t in Eq. (12) indicates that the eigenvalue λ is repeated. Note that λ may be complex, $\lambda \in \mathbb{C}$. If so, because A is a real matrix, the complex conjugate, $\bar{\lambda} = \mathrm{Re}(\lambda) - i\,\mathrm{Im}(\lambda)$, is also an eigenvalue of A and the linear combination of terms from Eq. (12) still yields a real function.

Our concern here is what happens to $x(t)$ in the limit as $t \to \infty$. Does the linear system tend to a steady-state or oscillate or go off to infinity? If λ is an eigenvalue with $\mathrm{Re}(\lambda) = a$ and $\mathrm{Im}(\lambda) = b$, then

$$\exp(\lambda t) = \exp(at)\big(\cos(bt) + i\,\sin(bt)\big).$$

Hence, if all the eigenvalues of A satisfy $\mathrm{Re}(\lambda) < 0$, the negative exponential $\exp(at)$ in each term from Eq. (12) will kill both the corresponding t^{μ}, no matter how large μ is, and the $\cos(bt)$ and $\sin(bt)$ functions, so that $x = 0$ is the unique steady state and $x(t) \to 0$ as $t \to \infty$. This corresponds to the well-known eigenvalue criterion for stability of a dynamical system (Hirsch and Smale 1974).

On the other hand, if none of the eigenvalues are zero and at least one of them satisfies $\mathrm{Re}(\lambda) \geq 0$, then the system will either oscillate or go off to infinity. The corner case occurs when one or more of the eigenvalues is zero, for which a more delicate analysis is required.

This is particularly relevant to the Laplacian matrix of a graph. We note first a consequence of the conservation law in Eq. (9). Although that was written for the steady-state concentrations, the total concentration does not change over time, so that

$$\frac{d}{dt}\big(x_1(t) + \cdots + x_n(t)\big) = 0.$$

This follows from Eq. (2), which implies that

$$1 \cdot \mathcal{L}(G) = 0. \tag{13}$$

Equation (13) already indicates that $\mathcal{L}(G)$ is not of full rank and, therefore, that zero is an eigenvalue.

To know more about the Laplacian eigenvalues, recall Geršgorin's theorem for the columns of a matrix A (Horn and Johnson 1985). The ith Geršgorin disc is the circular area in the complex plane that is centered on the ith diagonal element and whose radius is the sum of the absolute values of the remaining elements in the ith column:

$$\left\{ z \in \mathbb{C} \mid |z - A_{ii}| \leq \sum_{j \neq i} |A_{ij}| \right\}.$$

Geršgorin's theorem states that each eigenvalue occurs in some Geršgorin disc.

Proposition 1 *For any graph G, if λ is an eigenvalue of $\mathcal{L}(G)$, then either $\lambda = 0$ or $\mathrm{Re}(\lambda) < 0$.*

Proof It follows from Eq. (2) that each Geršgorin disc of $\mathcal{L}(G)$ is centered on a negative real number and has the same radius in absolute value. Hence, for any eigenvalue λ, either $\lambda = 0$ or $\mathrm{Re}(\lambda) < 0$, as required. □

The limiting behavior of Laplacian dynamics therefore hinges on the zero eigenvalue. If this eigenvalue is repeated, then Eq. (12) may yield terms of the form t^{μ}, which go to infinity with t. To illustrate the possibilities, consider the following two matrices:

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \tag{14}$$

in each of which the eigenvalue 0 is repeated three times. For A, $\exp(At) = I$, so that $x_A(t) = x(0)$ and the system remains forever in steady state at its initial condition. For B, it is easy to check by direct calculation using Eq. (11) that

$$\exp(Bt) = \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}, \tag{15}$$

so that $x_B(t) \to \infty$ unless $x(0)_2 = x(0)_3 = 0$.

Recall that the *algebraic multiplicity* of the eigenvalue λ, $\mathrm{alg}(\lambda)$, is the number of times that λ occurs as a repeated root of the characteristic polynomial or, equivalently, the total size of all the Jordan blocks associated to λ in the Jordan normal form. In contrast, the *geometric multiplicity* of λ, $\mathrm{geo}(\lambda)$, is simply the number of Jordan blocks associated to λ or, equivalently, the dimension of the λ eigenspace. Evidently, $\mathrm{alg}(\lambda) \geq \mathrm{geo}(\lambda)$ and, since A and B are already in Jordan normal form, it is easy to

check that

$$\mathrm{alg}_A(0) = 3, \quad \mathrm{geo}_A(0) = 3 \quad \text{and} \quad \mathrm{alg}_B(0) = 3, \quad \mathrm{geo}_B(0) = 1.$$

It is the discrepancy between the two multiplicities of B that leads to the unstable behavior of $x_B(t)$, as confirmed by the following result.

Proposition 2 *Suppose that the real matrix A has the same eigenvalue distribution as a Laplacian matrix, as given by Proposition 1. If $\mathrm{alg}_A(0) = \mathrm{geo}_A(0)$, then the solution of $dx/dt = A \cdot x$ tends to a steady state as $t \to \infty$ for any initial condition, while if $\mathrm{alg}_A(0) > \mathrm{geo}_A(0)$ then there is some initial condition for which $x(t) \to \infty$.*

Proof The proof follows from standard results (Hirsch and Smale 1974) but for the sake of completeness an outline is given here. Consider first the case in which a basis has been chosen so that A is in Jordan normal form,

$$A = \begin{pmatrix} \boxed{A_1} & 0 & \dots & 0 \\ 0 & \boxed{A_2} & 0 & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \boxed{A_k} \end{pmatrix}.$$

A has diagonal Jordan blocks, $A_1, \dots, A_k$, each corresponding to some eigenvalue of A. Since matrices can be multiplied in blocks, $\exp(At)$ is also block diagonal, with the corresponding blocks $\exp(A_1 t), \dots, \exp(A_k t)$. If A_i corresponds to an eigenvalue λ with $\mathrm{Re}(\lambda) < 0$ and u is any initial condition for A_i, then $\exp(A_i t) \cdot u \to 0$ as $t \to \infty$. It follows that the behavior of the solution $x(t) = \exp(At) \cdot x(0)$ as $t \to \infty$ depends only on the blocks corresponding to the eigenvalue 0.

If A_i is such a Jordan block then it has the form of matrix B in Eq. (14), with 0 on the diagonal, 1 on the super-diagonal and 0 elsewhere, but it can be of any dimension. If $\mathrm{alg}_A(0) = \mathrm{geo}_A(0)$, then each of the blocks corresponding to the eigenvalue 0 has dimension 1. By permuting the basis, if necessary, these blocks can be combined into a single composite block, which now has the form of matrix A in Eq. (14). In this case, as above, $\exp(At) = I$ and $x(t) \to c$ as $t \to \infty$. Evidently, $c_i = x_i(0)$ if i is an index in the zero-eigenvalue block and $c_i = 0$ if i is an index outside this block. It also follows that $A \cdot c = 0$, so that c is a steady state vector at which $dc/dt = 0$.

Conversely, if $\mathrm{alg}_A(0) > \mathrm{geo}_A(0)$ then some block corresponding to the eigenvalue 0 has dimension greater than 1. Assume that this block looks like matrix B in Eq. (14); the argument is identical for any other size of block. The exponential of this block is given in Eq. (15). Let i be the index of the first row and j be the index of the last column of this block within matrix A and choose the initial condition $x(0)$ to be the unit vector in coordinate j. Then $(\exp(At) \cdot x(0))_i = t^3$, so that $x(t) \to \infty$ as $t \to \infty$, as required.

Finally, if A is any matrix, there is an alternative basis in which it has Jordan normal form. That is, there is an invertible matrix P, such that $P^{-1} \cdot A \cdot P$ is in Jordan normal form. Since $\exp(P^{-1} \cdot A \cdot Pt) = P^{-1} \cdot \exp(At) \cdot P$, the conclusions reached for the Jordan normal form apply also to A. □

To apply Proposition 2 to Laplacian matrices, it is necessary to exploit the structure of the underlying graph, to which we now turn.

3 Zero Multiplicities for Strongly Connected Graphs

If G is a graph, vertex i is said to *ultimately reach* vertex j, denoted $i \Rightarrow j$, if either $i = j$ or there is a path of directed edges from i to j

$$i = i_1 \to i_2 \to \cdots \to i_k = j.$$

The relation of ultimate reaching is the reflexive, transitive closure of the edge relation. Vertex i is said to be *strongly connected* to vertex j, denoted $i \approx j$, if $i \Rightarrow j$ and $j \Rightarrow i$. Strong connectivity is the symmetric closure of ultimate reaching. It is an equivalence relation on the set of vertices, whose equivalence classes are the *strongly connected components* (SCCs) of G. A graph G is strongly connected, in the sense described in Sect. 1, if all of its vertices lie in a single SCC.

Proposition 3 *If G is a strongly connected graph, then* $\dim\ker \mathcal{L}(G) = 1$.

Proof The kernel of $\mathcal{L}(G)$ is the column null space, which has the same dimension as the row null space. To show that the latter has dimension one, it is sufficient, since $1 \cdot \mathcal{L}(G) = 0$, to show that if $v \cdot \mathcal{L}(G) = 0$ then $v_1 = \cdots = v_n$. Let v_p be the smallest component of v, so that $v_i \geq v_p$ for all i. Let $U \subseteq \{1, \ldots, n\}$ be the set of those indices i for which $v_i = v_p$. Note that $U \neq \emptyset$ since $p \in U$. If $m \in U$, then

$$\big(v \cdot \mathcal{L}(G)\big)_m = \sum_{k \neq m} (v_k - v_m)\mathcal{L}(G)_{km} = 0.$$

Each term in the sum is a product of two nonnegative quantities, but $\mathcal{L}(G)_{km} > 0$ when $m \to k$. Hence, $v_k = v_m = v_p$ whenever there is an edge $m \to k$, and so U is closed under outgoing edges. Since G is strongly connected, $U = \{1, \ldots, n\}$ and $v_1 = \cdots = v_n$, as required. □

Proposition 3 is a simplified version of (Thomson and Gunawardena 2009a, Lemma 1), which holds over a rational function field. Proposition 3 originated in the biological literature in the work of Feinberg and Horn in Chemical Reaction Network Theory (Feinberg and Horn 1977, Appendix), although the Laplacian interpretation came later; see Gunawardena (2012) for details. It has been independently proved in several other literatures; see, for instance, Olfati-Saber and Murray (2004, Theorem 1) and the commentary on this in Chebotarev (2010).

Since $\ker \mathcal{L}(G)$ is the eigenspace for the zero eigenvalue, Proposition 3 shows that when G is strongly connected, the geometric multiplicity of the zero eigenvalue of its Laplacian is one.

The algebraic multiplicity of the zero eigenvalue requires more work. We first set up some basic infrastructure that will be crucial for later results.

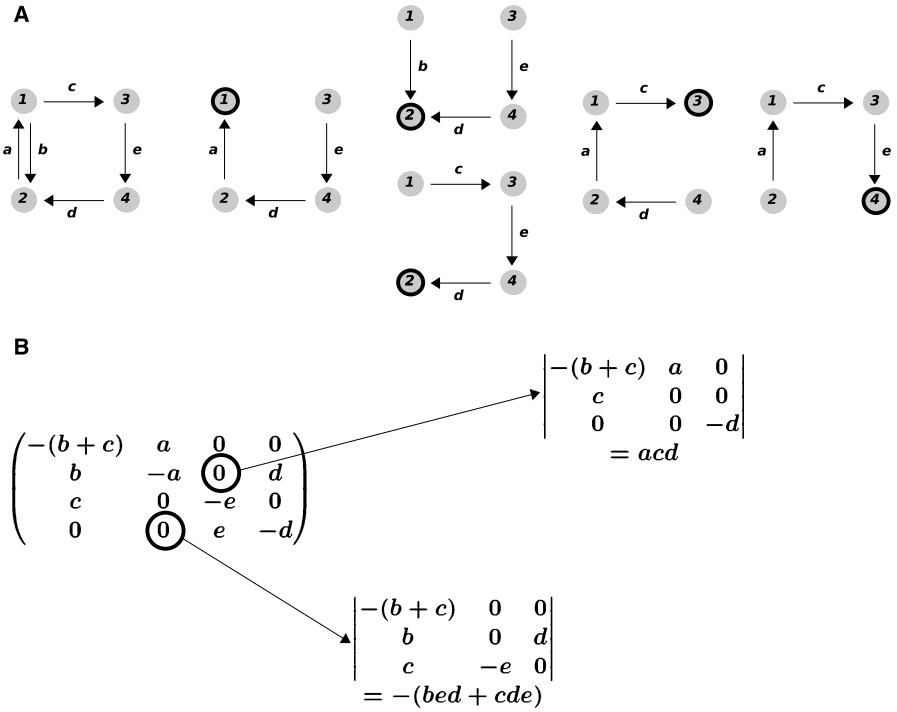


Fig. 2 Spanning trees and the Matrix-Tree theorem. (**A**) *On the left*, a strongly-connected graph. *On the right*, the spanning trees rooted at each vertex, with each root outlined in *black*. (**B**) The Laplacian matrix of the graph in (**A**) with two of its minors, $\mathcal{L}(G)_{(42)}$ and $\mathcal{L}(G)_{(23)}$, calculated, illustrating the Matrix-Tree formula in Eq. (16) of Theorem 1. The determinants are placed beneath the spanning trees in (**A**) that give the corresponding values

For any graph G, a subgraph T is a *spanning tree* of G if T reaches each vertex of G and is connected and acyclic as an undirected graph. T is *rooted* at $i \in G$ if i is the only vertex of T with no edges leaving it, $i \nrightarrow j$. This implies that any nonroot vertex has exactly one edge leaving it, for otherwise there would be an additional root or an undirected cycle (Fig. 2(A)). Let $\Theta_i(G)$ denote the set of spanning trees of G rooted at i. Note that in a general graph there may be no spanning trees rooted at a particular vertex, i, so that $\Theta_i(G) = \emptyset$.

Given a $n \times n$ matrix A, recall that $A_{(ij)}$ is the *minor* given by the determinant of the $(n-1) \times (n-1)$ matrix obtained from A by removing the ith row and jth column.

Theorem 1 (Matrix-Tree Theorem (Tutte 1948)) *If G is any graph with n vertices then the minors of its Laplacian are given by*

$$\mathcal{L}(G)_{(ij)} = (-1)^{n+i+j-1} \sum_{T \in \Theta_j(G)} \left(\prod_{k \xrightarrow{a} l \in T} a \right). \tag{16}$$

In other words, up to sign, the (ij)th minor of the Laplacian is the sum, over all spanning trees rooted at j, of the product of the labels on the edges of each tree (Fig. 2(B)). In view of its fundamental importance, a proof of Theorem 1 is given in the Appendix.

Results of this kind for labeled, directed graphs appear to have been first proved by Tutte (1948); see Moon (1970, Chap. 5) for more background on the mathematical literature. Later authors working in different fields have often rediscovered this result without being aware of Tutte's precedence. For example, the paper of Bott and Mayberry (1954) contains a version of Theorem 1, which has been used in economics (Whitin 1954), quantum field theory (Jaffe 1965), and statistical mechanics (Schnakenberg 1976). In the biological literature, versions of Theorem 1 were rediscovered by King and Altman in enzyme kinetics (King and Altman 1956) and by Hill in analyzing nonequilibrium biological systems (Hill 1966, 1985). Several more powerful variants of the Matrix-Tree theorem are now known, such as the all-minors Matrix-Tree theorem and the Matrix Forest theorem; see Chebotarev and Agaev (2002) for more details.

Lemma 1 *A graph G is strongly connected if, and only if, it has a spanning tree rooted at each vertex.*

Proof Suppose G is strongly connected and choose any vertex i of G. If i is the only vertex, then i itself is a spanning tree rooted at i. If not, suppose inductively that a tree rooted at i has been constructed and that j is a vertex of G that is not in the tree. Since G is strongly connected, $j \Rightarrow i$ and there is some directed path from j to i. Let k be the first vertex along the path, starting from j, that is a vertex in the tree. The subpath from j to k may be added to the tree to form a larger tree in which j is now a vertex. Hence, by induction, there is a tree rooted at i that reaches every vertex. Now suppose that G is a graph. If it has only one vertex, it is strongly connected by definition, so suppose that it has two distinct vertices i and j. Since there is a spanning tree rooted at i, j must be a vertex on that tree and so $j \Rightarrow i$. Similarly, $i \Rightarrow j$. Hence, $i \approx j$ and G is strongly connected. □

If A is a matrix, recall that its *adjugate matrix*, $\mathrm{adj}(A)$ is the transpose of the matrix of *cofactors*, or signed minors (Horn and Johnson 1985):

$$\mathrm{adj}(A)_{ij} = (-1)^{i+j} A_{(ji)}. \tag{17}$$

Note the reversal of the indices: the ijth entry of the adjugate is, up to sign, the (ji)th minor of A. The adjugate appears in two well known-formulas. The first is the Laplace expansion for the determinant (Horn and Johnson 1985),

$$\mathrm{adj}(A) \cdot A = A \cdot \mathrm{adj}(A) = \det(A) \cdot I. \tag{18}$$

If A is the Laplacian matrix of a graph G, then by Proposition 3, $\det \mathcal{L}(G) = 0$, and so, by Eq. (17), the columns of $\mathrm{adj}(\mathcal{L}(G))$ are elements of $\ker \mathcal{L}(G)$. Let $\rho^G \in \mathbb{R}^n_{>0}$ denote the column vector given by the spanning trees rooted at each vertex, through

the formula in Eq. (16):

$$\rho_i^G = \sum_{T\in\Theta_i(G)} \Big(\prod_{k\xrightarrow{a} l\in T} a \Big).$$

Using Eqs. (16) and (17), the adjugate can be expressed in terms of ρ^G as

$$\operatorname{adj}\big(\mathcal{L}(G)\big)_{ij} = (-1)^{i+j}(-1)^{n+j+i-1}\rho_i^G = (-1)^{n-1}\rho_i^G. \tag{19}$$

It follows that each of the columns of $\operatorname{adj}(\mathcal{L}(G))$ is, up to sign, equal to ρ^G. If, moreover, G is strongly connected, then it follows from Lemma 1 that $\rho_i^G > 0$ for all i. In particular, $\rho^G \neq 0$. Since $\dim\ker\mathcal{L}(G) = 1$ by Proposition 3, this gives a basis for $\ker\mathcal{L}(G)$.

Corollary 1 *If G is a strongly connected graph, then $\ker\mathcal{L}(G) = \langle\rho^G\rangle$ and $\rho_i^G > 0$ for all i.*

Corollary 1 provides the *positivity* mentioned in Sect. 1. If the graph is strongly connected and the rate constants and concentrations that make up the labels in any application are positive, then ρ^G has only positive entries.

A second use of the adjugate, which will lead us to the algebraic multiplicity of the zero eigenvalue, is the formula for the derivative of a determinant (Magnus and Neudecker 1988, Theorem 8.1),

$$\frac{d}{dt}\det A(t) = \operatorname{Tr}\Big(\operatorname{adj}\big(A(t)\big)\cdot\frac{dA(t)}{dt}\Big),$$

where $A(t)$ is any differentiable matrix function of a single variable. This formula can be applied, in particular, to the characteristic polynomial of A,

$$\chi_A(t) = \det(A - tI).$$

Evaluating the derivative of $\chi_A(t)$ at $t = 0$ yields

$$\begin{aligned}\frac{d}{dt}\chi_A(t)\big|_{t=0} &= \operatorname{Tr}\Big(\operatorname{adj}(A-tI)\cdot\frac{d}{dt}(A-tI)\Big)\Big|_{t=0} = \operatorname{Tr}\big(\operatorname{adj}(A-tI)\cdot(-I)\big)\big|_{t=0}\\ &= -\operatorname{Tr}\big(\operatorname{adj}(A)\big).\end{aligned} \tag{20}$$

Proposition 4 *If G is a strongly-connected graph, then $\operatorname{alg}_{\mathcal{L}(G)}(0) = 1$.*

Proof If $\operatorname{alg}_{\mathcal{L}(G)}(0) = k$, where $k \geq 1$, then, by definition of the algebraic multiplicity, t^k is a factor of $\chi_{\mathcal{L}(G)}(t)$ but t^{k+1} is not:

$$\chi_{\mathcal{L}(G)}(t) = t^k f(t) \quad \text{where } f(0)\neq 0.$$

Accordingly,

$$\frac{d}{dt}\chi_{\mathcal{L}(G)}(t)\big|_{t=0}\begin{cases}\neq 0 & \text{if } k=1,\\ =0 & \text{if } k>1.\end{cases} \tag{21}$$

According to Eq. (20), the derivative in question is given by

$$-\big(\mathrm{adj}_{11}\big(\mathcal{L}(G)\big)+\cdots+\mathrm{adj}_{nn}\big(\mathcal{L}(G)\big)\big).$$

By Eq. (19),

$$\mathrm{adj}_{ii}=(-1)^{n-1}\rho_i^G,$$

where, by Corollary 1, $\rho_i^G>0$ for all i. Hence,

$$\frac{d}{dt}\chi_{\mathcal{L}(G)}(t)\big|_{t=0}=(-1)^n\big(\rho_1^G+\cdots+\rho_n^G\big)\neq 0.$$

It follows from Eq. (21) that $\mathrm{alg}_{\mathcal{L}(G)}(0)=1$, as required. □

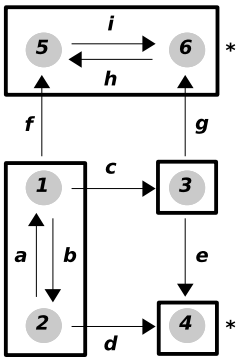
We have established that for a strongly-connected graph, $\mathrm{alg}_{\mathcal{L}(G)}(0)=\mathrm{geo}_{\mathcal{L}(G)}(0)=1$, so that the conditions of Proposition 2 are fulfilled. The extension to general graphs can now be undertaken through a structural decomposition of G into its SCCs.

4 Zero Multiplicities for General Graphs

Let $[i]$ denote the SCC containing i. Given two SCCs, $[i]$ and $[j]$, $[i]$ is said to *precede* $[j]$, denoted $[i]\preceq[j]$, if $i'\Rightarrow j'$ for some $i'\in[i]$ and some $j'\in[j]$. It is easy to check that this is well defined. The relation of precedes is a partial order on the SCCs: it is clearly reflexive and transitive and if $[i]\preceq[j]$ and $[j]\preceq[i]$ then i and j are mutually linked by directed paths, so $[i]=[j]$, and the relation is also antisymmetric. It follows that we can distinguish *terminal* SCCs, which are those $[i]$ such that, if $[i]\preceq[j]$ then $[i]=[j]$.

Suppose now, without any loss of generality, that the vertices of G are renamed as follows. Such a renaming amounts to a permutation of the standard basis in $\mathbb{R}^n$. First, choose an ordering of the SCCs, $C_1,\ldots,C_p$, such that, if $C_i\preceq C_j$, then $i\leq j$. Since $\preceq$ is a partial order, this can always be done. As a consequence, if there are q terminal SCCs, then these are, in some order, $C_{p-q+1},\ldots,C_p$, while if $i<p-q+1$ then C_i is nonterminal. Let c_i denote the size of SCC C_i, $c_i=\#C_i$, so that $c_1+\cdots+c_p=n$. Let d_i denote the partial sums of this sequence, $d_i=\sum_{1\leq j\leq i}c_j$ for $1\leq i\leq p$, with $d_0=0$. Now rename the vertices in C_i in any order, using only the indices $d_{i-1}+1,\ldots,d_{i-1}+c_i=d_i$. By construction, the Laplacian matrix of G in this new basis is in block lower-triangular form.

$$\mathcal{L}(G)=\left(\begin{array}{ccc:ccc}\boxed{L_1} & 0 & 0 & 0 & \ldots & 0\\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots\\ * & * & \boxed{L_{p-q}} & 0 & \ldots & 0\\ \hdashline * & \ldots & * & \boxed{L_{p-q+1}} & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ * & \ldots & * & 0 & 0 & \boxed{L_p}\end{array}\right)\tag{22}$$



$$\left(\begin{array}{ccc:ccc} -(b+c+f) & a & 0 & 0 & 0 & 0 \\ b & -(a+d) & 0 & 0 & 0 & 0 \\ c & 0 & -(g+e) & 0 & 0 & 0 \\ \hdashline 0 & d & e & 0 & 0 & 0 \\ f & 0 & 0 & 0 & -i & h \\ 0 & 0 & g & 0 & i & -h \end{array}\right)$$

Fig. 3 Graph decomposition. *On the left*, the graph in Fig. 1 is shown with its four strongly-connected components enclosed in *boxes*. The two terminal SCCs are marked by $*$. Note that the ordering of the vertices already follows the renaming scheme described in the text. *On the right*, the Laplacian matrix of the graph is shown in block lower-triangular form, as in Eq. (22), with *the dashed lines* separating the blocks corresponding to nonterminal SCCs from those corresponding to terminal SCCs

In this matrix, an asterisk, $*$, denotes a block which can be nonzero. The blocks in boxes on the main diagonal, denoted $L_1, \ldots, L_p$, are the submatrices defined by restricting $\mathcal{L}(G)$ to the vertices of the corresponding SCCs, $C_1, \ldots, C_p$. Each L_i is a square submatrix of size $c_i \times c_i$. The block lower-triangular form of $\mathcal{L}(G)$ itself comes from the fact that, if there is an edge from a vertex in C_i to a vertex in a different block C_j, then, by construction, $i < j$. The blocks in the bottom right-hand segment demarcated by the dashed lines, $L_{p-q+1}, \ldots, L_p$, correspond to the terminal SCCs. Since there can be no edges coming out of a terminal SCC, this segment is block diagonal.

It will be helpful to consider each submatrix L_i as a matrix in its own right, so that its rows and columns are indexed from $1, \ldots, c_i$ and not from $d_{i-1}+1, \ldots, d_i$. As is customary, we use L_i to denote either the submatrix of $\mathcal{L}(G)$ or the matrix in its own right. In a similar vein, it will be helpful to consider the subset of vertices $C_i \subseteq \{1, \ldots, n\}$ as a graph, with vertices $1, \ldots, c_i$ in place of $d_{i-1}+1, \ldots, d_i$, and precisely those edges that are found between the latter vertices in G. Edges that enter or leave the SCC are disregarded. Here, we use C_i^* for the resulting graph, to distinguish it from the subset C_i.

The behavior of L_i depends crucially on whether C_i is a terminal or non-terminal SCC (Fig. 3). Because there are no edges emerging from any terminal SCC, it is easy to see that

$$L_i = \mathcal{L}\big(C_i^*\big) \quad \text{when } p-q+1 \le i \le p. \tag{23}$$

According to Corollary 1, each such L_i has an associated $\rho^{C_i^*} \in R^{c_i}_{>0}$ that is a basis element for $\ker L_i$. We can extend this to a vector $\bar{\rho}^{C_i^*} \in \mathbb{R}^n_{>0}$ by setting all components with indices outside C_i to zero:

$$\big(\bar{\rho}^{C_i^*}\big)_j = \begin{cases} (\rho^{C_i^*})_{j-d_{i-1}} & \text{if } d_{i-1}+1 \le j \le d_i, \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

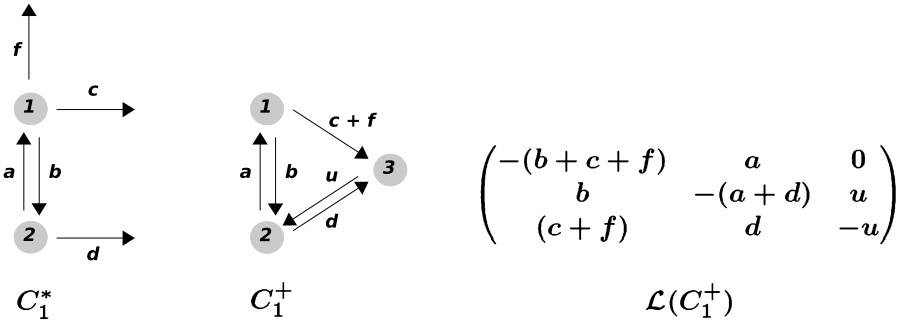


Fig. 4 Graph construction for Lemma 2. *On the left*, the nonterminal strongly connected component of the graph in Fig. 3 denoted C_1^*, shown with the edges leading out of the SCC. *In the middle*, the graph C_1^+, constructed by adding a new vertex $c_1{+}1 = 3$ along with labeled edges to and from that vertex as described in the text. *On the right*, the Laplacian matrix of C_1^+ in which L_1 is the $c_1 \times c_1$ sub-matrix obtained by omitting the $(c_1{+}1)$-st row and the $(c_1{+}1)$-st column

Because any two distinct SCCs are completely disjoint, it is evident that these q vectors in $\mathbb{R}^n_{>0}$ are linearly independent. Moreover, because of the block structure of the Laplacian matrix in Eq. (22), not only is $L_i \cdot \rho^{C_i^*} = 0$ but also $\mathcal{L}(G) \cdot (\bar{\rho}^{C_i^*}) = 0$, so that

$$\ker \mathcal{L}(G) \supseteq \left\langle \bar{\rho}^{C^*_{p-q+1}}, \ldots, \bar{\rho}^{C^*_p} \right\rangle. \tag{25}$$

This almost determines the structure of $\ker \mathcal{L}(G)$. The final step requires an understanding of the L_i corresponding to non-terminal SCCs.

Equation (23) no longer holds for nonterminal SCCs, for which $i < p - q + 1$, because the edges that emerge from each nonterminal SCC ensure that $1 \cdot L_i \neq 0$, as would be required of a Laplacian matrix (Fig. 3). Specifically,

$$(1 \cdot L_i)_j = -\left(\sum_{j \xrightarrow{a}} a\right) \leq 0, \tag{26}$$

where the sum is taken over all edges from vertex j that leave C_i^* for some other SCC. However, we can repair L_i by introducing a new vertex into C_i^*, say vertex $c_i{+}1$, to form the new graph C_i^+ (Fig. 4). For every vertex $j \in C_i^*$ for which $(1 \cdot L_i)_j < 0$, introduce the new edge $j \xrightarrow{\alpha} c_i{+}1$, with the label $\alpha = -(1 \cdot L_i)_j$. This choice is intended to precisely cancel out the discrepancy from zero in Eq. (26). Finally, add one new edge $c_i{+}1 \xrightarrow{u} j$ from the new vertex back to one of those vertices j which has a new edge to $c_i{+}1$. The label, u, on this edge may be chosen arbitrarily. Figure 4 illustrates this procedure. Since C_i^* is strongly connected and the new vertex $c_i{+}1$ is reversibly linked to a vertex in C_i^*, it follows that the new graph C_i^+ is also strongly connected. By Theorem 1 and Lemma 1, the minors of $\mathcal{L}(C_i^+)$ are nonzero. But, by construction, the $c_i \times c_i$ sub-matrix of $\mathcal{L}(C_i^+)$ obtained by removing the $(c_i{+}1)$-st row and the $(c_i{+}1)$-st column is precisely L_i. Hence, $\det(L_i) = \mathrm{adj}(\mathcal{L}(C_i^+))_{n+1,n+1}$, so that, by Eq. (19), $\det(L_i) = (-1)^n \rho^{C_i^+}_{n+1} \neq 0$. If follows that L_i is invertible.

Lemma 2 *If a graph G is placed in the form required for Eq. (22) then any diagonal block for a nonterminal SCC is invertible.*

With Lemma 2 in hand, the rest is easy. Recall that q is the number of terminal strongly-connected components of G.

Proposition 5 *For any graph G,*

$$\ker \mathcal{L}(G) = \left\langle \bar{\rho}^{C^*_{p-q+1}}, \ldots, \bar{\rho}^{C^*_p} \right\rangle, \tag{27}$$

$\dim \ker \mathcal{L}(G) = q$ *and* $\mathrm{geo}_{\mathcal{L}(G)}(0) = q$.

Proof We already know from Eq. (25) that the $\ker \mathcal{L}(G)$ contains the span of the vectors in Eq. (27). If $x \in \ker \mathcal{L}(G)$, so that $\mathcal{L}(G) \cdot x = 0$, it follows from the block structure of the matrix in Eq. (22) and Lemma 2 that $x_i = 0$ for $1 \leq i \leq p - q$. Since the bottom segment of the matrix is block diagonal, x must be in the span of the vectors in Eq. (27). This proves Eq. (27) and the other assertions follow immediately. □

Proposition 6 *For any graph G,* $\mathrm{alg}_{\mathcal{L}(G)}(0) = q$.

Proof From the block structure of the matrix in Eq. (22), the characteristic polynomial of $\mathcal{L}(G)$ can be written as the product

$$\chi_{\mathcal{L}(G)}(t) = \prod_{i=1}^{p} \chi_{L_i}(t).$$

For $1 \leq i \leq p - q$, L_i is invertible by Lemma 2, so t does not divide $\chi_{L_i}(t)$. For $p - q + 1 \leq i \leq p$, each L_i is the Laplacian of a strongly-connected graph, so, by Proposition 4, t divides χ_{L_i} but t^2 does not. Hence, t^q divides $\chi_{\mathcal{L}(G)}(t)$ but t^{q+1} does not, so that $\mathrm{alg}_{\mathcal{L}(G)}(0) = q$, as required. □

In the biological literature, $\dim \ker \mathcal{L}(G)$ was first determined for an arbitrary graph in the work of Feinberg and Horn in Chemical Reaction Network Theory (Feinberg and Horn 1977, Appendix). The description of the basis elements of $\ker \mathcal{L}(G)$ using the Matrix-Tree theorem is given in Gunawardena (2012) but the use of Lemma 2, which greatly simplifies the argument, is new. Proposition 5 appears in Agaev and Chebotarev (2000). Proposition 6 is deduced in Chebotarev and Agaev (2002) using a more powerful variant of the Matrix-Tree theorem.

We can finally appeal to Proposition 2 to deduce the result towards which we have been working.

Theorem 2 *If G is any graph, then the solution of the Laplacian dynamics $dx/dt = \mathcal{L}(G) \cdot x$ converges to a steady state as $t \to \infty$, for any initial condition $x(0) \in \mathbb{R}^n$.*

5 Calculating Steady States

Theorem 2 leads naturally to the question of how the steady state to which the Laplacian dynamics tends depends on the initial condition from which the dynamics started.

Suppose that A is a $n \times n$ matrix and the dimension of its column null space (or, equivalently, its kernel) is q. Let R be a $n \times q$ matrix whose columns are a basis for the column null space. Dually, let L be a $q \times n$ matrix whose rows are a basis for the row null space (which has the same dimension q). Suppose further that L and R can be chosen so that the following equations hold:

$$A \cdot R = 0, \qquad L \cdot A = 0, \qquad L \cdot R = I_q. \tag{28}$$

Lemma 3 *Suppose that A is a matrix for which Eq. (28) holds and that, when started from the initial condition $x(0)$, the solution of $dx/dt = A \cdot x$ tends to a steady state c as $t \to \infty$. Then $c = R \cdot L \cdot x(0)$.*

Proof The solution of $dx/dt = A \cdot x$ is given by $x(t) = \exp(At) \cdot x(0)$, so that, by Eq. (11),

$$x(t) = \big(I + A \cdot B(t)\big) \cdot x(0),$$

where $B(t)$ is some matrix function of t. It follows from Eq. (28) that $L \cdot x(t) = L \cdot x(0)$. Letting $t \to \infty$, we find that $L \cdot c = L \cdot x(0)$. Since c is a steady state, $dc/dt = A \cdot c = 0$, and c is in the column null space. Hence, $c = R \cdot d$ for some vector d. Therefore, by Eq. (28) again, $L \cdot x(0) = L \cdot c = L \cdot (R \cdot d) = d$, so that $R \cdot L \cdot x(0) = R \cdot d = c$, as required. □

The decomposition of $\mathcal{L}(G)$ in Eq. (22) offers a way to find matrices L and R that satisfy Eq. (28). It is convenient to work in terms of the segments in Eq. (22) that separate the vertices of G into nonterminal and terminal SCCs. Let u be the total number of vertices of G that are not in terminal SCCs. In the language of the previous section, $u = d_{p-q}$. Consider the block decomposition of L, $\mathcal{L}(G)$ and R corresponding to the dashed lines in Eq. (22),

$$L = \big(X \mid U\big), \qquad \mathcal{L}(G) = \left(\begin{array}{c|c} N & 0 \\ \hline B & T \end{array}\right), \qquad R = \begin{pmatrix} Y \\ V \end{pmatrix}, \tag{29}$$

in which X is $q \times u$, U is $q \times (n-u)$, N is $u \times u$, 0 is $u \times (n-u)$, B is $(n-u) \times u$, T is $(n-u) \times (n-u)$, Y is $u \times q$ and V is $(n-u) \times q$. Here, N, 0, B, and T correspond to the four segment blocks shown in Eq. (22).

The block T is itself block diagonal and, by Eq. (23), each subblock corresponds to the Laplacian matrix of the associated terminal SCC. The special case when G itself is strongly connected is instructive. Then T has only a single subblock and is all of $\mathcal{L}(G)$. We know from Corollary 1 that the column vector $\bar{\rho}^G$ is the basis of $\ker \mathcal{L}(G)$, so if we take R to be the $n \times 1$ matrix with this vector as its only column, then $\mathcal{L}(G) \cdot R = 0$. We also know that $1 \cdot \mathcal{L}(G) = 0$, so we can take L to be the $1 \times n$

matrix, which has 1_n as its only row. The only issue with this choice of L and R is that

$$L \cdot R = 1 \cdot \bar{\rho}^G = \bar{\rho}_1^G + \cdots + \bar{\rho}_n^G \neq 1,$$

so that the last requirement in Eq. (28) is not satisfied. However, this is easily remedied by normalizing R to the total amount in the column, $1 \cdot \bar{\rho}^G$.

For a general graph G, Proposition 5 provides a specific basis for the column null space. Keeping in mind the need for normalization, let R be the $n \times q$ matrix whose jth column is given by

$$R_{*j} = \left(\frac{1}{1 \cdot \bar{\rho}^{C^*_{p-q+j}}} \right) \bar{\rho}^{C^*_{p-q+j}}. \tag{30}$$

With this choice of R, any row of R whose index is not in a terminal SCC is zero, so that $Y = 0$, and each column of V has nonzero components only for those indices which are in the corresponding terminal SCC. As for L, Eq. (28) requires that $L \cdot R = I_n$, which by block multiplication requires that $U \cdot V = I_q$. Accordingly, let U be the $q \times (n - u)$ matrix given by first transposing V and then replacing each nonzero element of V by 1. It is then easy to see that, because of the normalisation of R, $U \cdot V = I_q$. Moreover, by Eq. (13) applied to each terminal SCC, $U \cdot T = 0$. It remains to determine X. Equation (28) requires that $L \cdot \mathcal{L}(G) = 0$, which by block multiplication requires that $X \cdot N + U \cdot B = 0$. We can now appeal once again to Lemma 2, which tells us that N is invertible. Hence, we can solve uniquely for X to obtain $X = -U \cdot B \cdot N^{-1}$. In summary, Eq. (28) is satisfied if

$$Y = 0, \qquad U \cdot T = 0, \qquad U \cdot V = I, \qquad X = -U \cdot B \cdot N^{-1}. \tag{31}$$

We deduce from Lemma 3 the following linear relationship between initial conditions and their resulting steady states.

Theorem 3 *Let G be any graph. If matrices L and R are defined by the procedure above then, for any initial condition $x(0)$, the Laplacian dynamics $dx/dt = \mathcal{L}(G) \cdot x$ tends to the steady state $R \cdot L \cdot x(0)$ as $t \to \infty$.*

The consequences of this relationship between initial conditions and steady states are explored in the next section.

6 Plasticity Arising from Nonstrong Connectivity

If G is strongly connected, then Corollary 1 says that $\dim \ker \mathcal{L}(G) = 1$ and that $\rho_i^G > 0$ at every vertex i. Furthermore, the decomposition in Eq. (29) has $q = 1$, $u = 0$, so that $L = U = 1_n$, $\mathcal{L}(G) = T$ and $R = V = \rho^G/(1 \cdot \rho^G)$. Theorem 3 then implies that initial condition $x(0)$ leads to the steady state

$$R \cdot L \cdot x(0) = R \cdot \big(1 \cdot x(0)\big) = \left(\frac{1 \cdot x(0)}{1 \cdot \rho^G} \right) \rho^G. \tag{32}$$

The numerator of the scalar factor in Eq. (32) is the conserved quantity, or total amount of x present initially, $1 \cdot x(0) = x_{\mathrm{tot}}$, and Eq. (32) corresponds to Eq. (10) in the Introduction. It does not matter on which vertices the initial concentration of material is placed. No matter how the material is distributed, the steady state that is reached depends only on the total amount that was present. Furthermore, changes in the values of the graph labels may alter the balance of material between different vertices, but does not cause a qualitative change in the steady-state behavior.

In the general case, where G may no longer be strongly connected, it follows from Theorem 3 that, if $z = L \cdot x(0) \in \mathbb{R}^q$, then the steady state corresponding to $x(0)$ is given by the linear combination of the columns of R whose coefficients are the entries in z. That is, using Eq. (30),

$$R \cdot L \cdot x(0) = z_1 \left(\frac{\bar{\rho}^{C^*_{p-q+1}}}{1 \cdot \bar{\rho}^{C^*_{p-q+1}}} \right) + \cdots + z_q \left(\frac{\bar{\rho}^{C^*_p}}{1 \cdot \bar{\rho}^{C^*_p}} \right). \tag{33}$$

Since $L \cdot \mathcal{L}(G) = 0$, the rows of the $q \times n$ matrix L provide a set of q conservation laws for the Laplacian dynamics, which generalize the single conservation law, $1 \cdot \mathcal{L}(G) = 0$ for the case of a strongly-connected graph. The coefficients $z_1, \ldots, z_q$ in Eq. (33) are the conserved quantities. Since the columns of R are non-zero only on vertices which lie in terminal SCCs, $(R \cdot L \cdot x(0))_i = 0$ for $1 \leq i \leq u$. For a non-strongly connected graph, no matter what initial condition is chosen, the resulting steady state is zero on some vertex, indeed, on any vertex that is not in a terminal SCC. Moreover, if k is a vertex in the ith terminal SCC, C_{p-q+i}, then because the SCCs are pairwise disjoint, the steady state picks out just one of the summands in the linear combination in Eq. (33):

$$\left(R \cdot L \cdot x(0)\right)_k = \left(\frac{z_i}{1 \cdot \bar{\rho}^{C^*_{p-q+i}}} \right) \left(\bar{\rho}^{C^*_{p-q+i}}\right)_k,$$

for $d_{p-q+i-1} + 1 \leq k \leq d_{p-q+i}$. This has a similar form to Eq. (32), with the conserved quantity z_i appearing in the numerator of the scalar factor.

The conserved quantities, however, have a more complicated algebraic form, which can be determined by substituting Eq. (31) in Eq. (29) to yield

$$L = U \cdot \left(-B \cdot N^{-1} \;\middle|\; I_{n-u} \right). \tag{34}$$

It is convenient in working with Eq. (34) to partition the initial condition $x(0)$ into $x^N(0) \in \mathbb{R}^u$ and $x^T(0) \in \mathbb{R}^{n-u}$, corresponding to the indices in nonterminal and terminal SCCs, respectively, so that

$$x(0) = \left(x^N(0) \;\middle|\; x^T(0) \right)^t.$$

The conservation laws can now be seen to be built up as follows. The vector $x^N(0) \in \mathbb{R}^u$ is first transformed using the matrix N^{-1} to yield a new vector $N^{-1} \cdot x^N(0) \in \mathbb{R}^u$. The matrix B encodes those edges in the graph which link nonterminal SCCs to terminal SCCs and the term $-B \cdot N^{-1} \cdot x^N(0)$ has the effect of transferring

the transformed vector, $N^{-1} \cdot x^N(0)$, into a set of values on the terminal SCCs. According to Eq. (34), these values are added to that part of the initial condition that was already on the vertices in the terminal SCCs, to form

$$-B \cdot N^{-1} \cdot x^N(0) + x^T(0).$$

Finally, the matrix U calculates the total value that is now present on each terminal SCC. Recalling the notation used in Eq. (24), the ith conserved quantity in Eq. (33), which corresponds to the ith terminal SCC, C_{p-q+i}, is given by

$$z_i = \sum_{k=d_{p-q+i-1}}^{d_{p-q+i}} \left(-B \cdot N^{-1} \cdot x^N(0) + x^T(0)\right)_k.$$

The expression on the right-hand side gives the corresponding conservation law for the Laplacian dynamics:

$$\frac{d}{dt}\left(\sum_{k=d_{p-q+i-1}}^{d_{p-q+i}} \left(-B \cdot N^{-1} \cdot x^N(t) + x^T(t)\right)_k\right) = 0.$$

Let us use the example in Fig. 3 to illustrate this calculation. The matrices B and N^{-1} are given by

$$\begin{pmatrix} 0 & d & e \\ f & 0 & 0 \\ 0 & 0 & g \end{pmatrix} \quad \text{and}$$

$$\left(\frac{1}{\Delta}\right)\begin{pmatrix} (a+d)(e+g) & a(e+g) & 0 \\ b(e+g) & (b+c+f)(e+g) & 0 \\ c(a+d) & ac & a(c+f)+d(b+c+f) \end{pmatrix},$$

respectively, where $\Delta = \det N = -(e+g)(a(c+f)+d(b+c+f))$. The matrix $(-B \cdot N^{-1} \mid I_{n-u})$ in Eq. (34) is then given by

$$\left(\frac{1}{-\Delta}\right)\begin{pmatrix} ce(a+d)+bd(e+g) & ace+d(b+c+f)(e+g) & e(a(c+f)+d(b+c+f)) & 1 & 0 & 0 \\ f(a+d)(e+g) & af(e+g) & 0 & 0 & 1 & 0 \\ cg(a+d) & acg & g(a(c+f)+d(b+c+f)) & 0 & 0 & 1 \end{pmatrix},$$

while the prefactor matrix U is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

As explained above, U adds up the contributions within each terminal SCC, so that L is given by

$$\left(\frac{1}{-\Delta}\right)\begin{pmatrix} ce(a+d)+bd(e+g) & ace+d(b+c+f)(e+g) \\ f(a+d)(e+g)+cg(a+d) & af(e+g)+acg \\ e(a(c+f)+d(b+c+f)) & 1 & 0 & 0 \\ g(a(c+f)+d(b+c+f)) & 0 & 1 & 1 \end{pmatrix}.$$

The two conservation laws can now be read off from the two rows of L.

There is also the general conservation law given by Eq. (13): $1 \cdot \mathcal{L}(G) = 0$. Using this together with Eq. (29), it follows that $1_u \cdot N + 1_{n-u} \cdot B = 0$, so that $1_u = -1_{n-u} \cdot B \cdot N^{-1}$. By construction of U, $1_q \cdot U = 1_{n-u}$, so that $1_q \cdot X = 1_q \cdot (-U \cdot B \cdot N^{-1}) = 1_u$ and so $1_q \cdot L = 1_n$. For the example in Fig. 3, the general conservation law comes about by adding the two rows of L, as can be readily checked.

The basis vectors appearing in Eq. (33), $\bar{\rho}^{C_3^*}$ and $\bar{\rho}^{C_4^*}$, corresponding to the two terminal SCCs, can also be calculated, following the prescription in Eq. (24), to give

$$(0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0)^t \quad \text{and} \quad (0 \quad 0 \quad 0 \quad 0 \quad h \quad i)^t,$$

respectively.

These calculations reveal, in marked contrast to the strongly-connected case, that the steady state may vary qualitatively depending on the initial condition and the values of the graph labels. For example, suppose that $a = b = c = f = g = 1$ and $e = d = \epsilon$, where $\epsilon \ll 1$, so that the edges leading into the first terminal SCC (Fig. 3) have low rates compared to the other edges, and suppose that the initial condition has nothing on the terminal SCCs, so that $x(0)_4 = x(0)_5 = x(0)_6 = 0$. Then

$$z_1 = \frac{2\epsilon(1+\epsilon)x(0)_1 + \epsilon(4+3\epsilon)x(0)_2 + \epsilon(2+3\epsilon)x(0)_3}{(1+\epsilon)(2+3\epsilon)},$$
$$z_2 = \frac{(1+\epsilon)(2+\epsilon)x(0)_1 + (2+\epsilon)x(0)_2 + (2+3\epsilon)x(0)_3}{(1+\epsilon)(2+3\epsilon)},$$

so that, as ϵ gets smaller, the steady state is dominated by the contribution from the second terminal SCC and

$$x(0) \to \sim \big(1 \cdot x(0)\big)\big(0 \quad 0 \quad 0 \quad 0 \quad \tfrac{h}{h+i} \quad \tfrac{i}{h+i}\big)^t, \tag{35}$$

where $1 \cdot x(0) = x(0)_1 + x(0)_2 + x(0)_3$. On the other hand, if $a = b = c = d = e = 1$ and $f = g = \epsilon$, so that the edges leading into the second terminal SCC have low rates compared to the other edges, then with the same initial conditions,

$$z_1 = \frac{(3+\epsilon)x(0)_1 + (1+(2+\epsilon)(1+\epsilon))x(0)_2 + (3+2\epsilon)x(0)_3}{(1+\epsilon)(3+2\epsilon)},$$
$$z_2 = \frac{2\epsilon(2+\epsilon)x(0)_1 + \epsilon(2+\epsilon)x(0)_2 + \epsilon(3+2\epsilon)x(0)_3}{(1+\epsilon)(3+2\epsilon)}.$$

As ϵ gets smaller, the steady state is dominated by the contribution from the first terminal SCC and

$$x(0) \to\sim \big(1 \cdot x(0)\big)(0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0)^t. \tag{36}$$

Equations (35) and (36) exhibit very different steady states to which the same initial condition tends, depending on the relative values of the graph labels. The nonoverlapping nature of the SCCs exacerbates the sharpness of the transition, with only vertices 5 and 6 playing a role in Eq. (35) and only vertex 4 playing a role in Eq. (36). In applications of the linear framework, the labels may be algebraic expressions, which include rate constants or the concentrations of variables outside the graph, such as slow components, as in Eq. (4) in the Introduction. Hence, changes in the larger system of which the graph is a part may lead to different steady-state behaviors. This kind of plasticity is not the same as multistability, since it is the conserved quantities z_1 and z_2 whose values are being altered, but the biological effects may appear to be similar.

7 Laplacian Dynamics as a Stochastic Master Equation

Laplacian dynamics was introduced in Eq. (1) as a form of linear biochemistry in which the quantities of interest are the concentrations of chemical species. However, at the molecular level, chemical interactions are fundamentally stochastic. Several of the applications of the linear framework, such as to gene regulation or to ion channel activity, are in contexts where single molecules are being analyzed and where single-molecule data can be acquired as to probabilities of occurrence. Given a graph G, we turn now from thinking of macroscopic concentrations of molecules transforming from vertex i to vertex j to thinking of an individual molecule as it changes from vertex i to vertex j. (It may help to think of the molecule changing from state i to state j.) Let $u_i(t)$ be the probability of finding a molecule at vertex i at time t, considered as a continuous time, finite-state space Markov process. How do we calculate $u_i(t)$?

This question is best answered in a broader context. Suppose that $X(t)$ is any continuous-time Markov process on finitely many states, $1, \ldots, n$, which is time homogeneous. That is, we assume that we are given the conditional probability for finding X in state i at time $t + \Delta t$, given that it was in state j at time t, denoted $\Pr(X(t + \Delta t) = i \mid X(t) = j)$, and that this probability does not depend on t. We assume further that X is sufficiently well behaved that we can determine the infinitesimal *transition rate* from j to $i \neq j$, which we will denote e_{ij},

$$e_{ij} = \lim_{\Delta t \to 0} \frac{\Pr(X(t + \Delta t) = i \mid X(t) = j)}{\Delta t}. \tag{37}$$

This rate is necessarily nonnegative because probabilities are nonnegative and it has units of time^{-1}. The conditional probability can be rewritten in terms of the transition rate in a form that is equivalent to Eq. (37),

$$\Pr\big(X(t + \Delta t) = i \mid X(t) = j\big) = e_{ij}\Delta t + o(\Delta t). \tag{38}$$

Here, the little o notation is used to stand for any function $f(\Delta t)$ for which $f(\Delta t)/\Delta t \to 0$ as $\Delta t \to 0$. As is customary in Markov process theory, the diagonal term e_{jj} is calculated differently by using the fact that the process must be in some state while the total probability must normalize to 1. Hence, provided Δt is sufficiently small,

$$\Pr\big(X(t+\Delta t) = j \mid X(t) = j\big) = 1 - \left(\sum_{v \neq j} e_{vj}\right)\Delta t + o(\Delta t). \tag{39}$$

The transition rates lead to a *master equation*, or Kolmogorov forward equation that describes the time evolution of X. The master equation formulation of Markov processes is developed by van Kampen (1992, Chap. 4). Let $X_i(t) = \Pr(X(t) = i)$. We need to calculate $X_i(t + \Delta t)$. There are two possibilities. Either the process is in state $j \neq i$ at time t, with probability $X_j(t)$, and undergoes a transition in time Δt from j to i with probability given by Eq. (38), or the process is already in state i, with probability $X_i(t)$, and remains there over time Δt with probability given by Eq. (39). Accordingly,

$$X_i(t+\Delta t) = \sum_{j \neq i} (e_{ij}\Delta t) X_j(t) + \left(1 - \left(\sum_{v \neq i} e_{vi}\right)\Delta t\right) X_i(t) + o(\Delta t).$$

Hence,

$$\frac{X_i(t+\Delta t) - X_i(t)}{\Delta t} = \sum_{j \neq i} e_{ij} X_j(t) - \left(\sum_{v \neq i} e_{vi}\right) X_i(t) + \frac{o(\Delta t)}{\Delta t}. \tag{40}$$

The right-hand side of Eq. (40) is suggestive. Let us construct a graph, G_X, associated to the Markov process X, by taking its vertices to be the states of X, placing an edge $j \to i$ if, and only if, $e_{ij} \neq 0$ and giving the edge the label e_{ij}. As required, the graph has no self-loops. The Laplacian matrix, $\mathcal{L}(G_X)$, is then given by Eq. (2). Substituting in Eq. (40) and taking the limit as $\Delta t \to 0$, we find that

$$\frac{dX_i(t)}{dt} = \sum_{j \neq i} \mathcal{L}(G_X)_{ij} X_j(t) + \mathcal{L}(G_X)_{ii} X_i(t) = \big(\mathcal{L}(G_X) \cdot X\big)_i.$$

Theorem 4 *Let X be any continuous time, finite-state space Markov process that is time homogeneous, for which transition rates may be determined by Eq. (37). The master equation of X is identical to Laplacian dynamics on the graph G_X:*

$$\frac{dX(t)}{dt} = \mathcal{L}(G_X) \cdot X.$$

The answer to our initial question is now clear. Given a graph G in the linear framework, the label $\mathcal{L}(G)_{ij}$ for $i \neq j$, when it is nonzero, is a macroscopic rate constant with units of time^{-1}. This means that, if $x_j(t)$ is the concentration of material at vertex j at time t, then, in any sufficiently small time interval Δt, a concentration of material given by $(\mathcal{L}(G)_{ij}\Delta t)x_j(t) + o(\Delta t)$ is transferred from vertex j to vertex i. It

follows that the probability that a molecule at j is transferred to i in the time interval Δt is given by

$$\frac{(\mathcal{L}(G)_{ij}\Delta t)x_j(t)}{x_j(t)} = \mathcal{L}(G)_{ij}\Delta t + o(\Delta t). \tag{41}$$

Note that this does not depend on t, so that the Markov process underlying the Laplacian dynamics is time homogeneous. Comparing Eq. (41) with Eq. (38), we see that this Markov process has exactly the same transition rates as the labels on the graph.

Corollary 2 *Let G be any graph and suppose that the Laplacian dynamics on G is interpreted stochastically to yield a continuous time Markov process on the vertices. If $u_i(t)$ denotes the probability of finding a molecule at vertex i at time t, then $u(t)$ obeys the same Laplacian dynamics*:

$$\frac{du(t)}{dt} = \mathcal{L}(G) \cdot u.$$

Theorem 4 reveals that deterministic Laplacian dynamics and stochastic master equations share an unexpected identity. The graph Laplacian is not introduced in van Kampen's text (van Kampen 1992), which is particularly associated with the master equation approach, but makes an occasional appearance in the probability theory literature when the relevant forward equation is discussed (Kelly 2011). A graphical approach to continuous-time Markov processes was introduced in the literature on nonequilibrium statistical mechanics (Hill 1966; Schnakenberg 1976) but, so far as we are aware, the identity expressed in Theorem 4 was never formulated.

8 Discussion

The main impetus for the present paper comes from recent application of the linear framework to gene regulation away from thermodynamic equilibrium (Ahsendorf et al. 2013). Gene regulation has been widely studied by using the timescale separation that transcription-factor binding to DNA is at thermodynamic equilibrium, while transcription is a slow process that averages over the equilibrium states. This method originates in the work of Ackers, Johnson, and Shea on phage lambda, (Ackers et al. 1982). They used equilibrium statistical mechanics and partition functions to eliminate the equilibrium states in favor of the concentrations of the transcription factors. The "thermodynamic formalism" has been revamped in recent years (Bintu et al. 2005a, 2005b; Segal and Widom 2009; Sherman and Cohen 2012) and used to analyze gene regulation in bacteria (Setty et al. 2003; Kuhlman et al. 2007), yeast (Gertz et al. 2009), and flies (Janssens et al. 2006; Zinzen et al. 2006; Segal et al. 2008; He et al. 2010).

In contrast to bacteria, however, eukaryotic organisms utilize many nonequilibrium mechanisms for regulating gene expression. Among these are DNA methylation, nucleosome remodelling, post-translational modification (PTM) of histone tails, PTM of transcription factors and transcriptional coregulators, such as the Mediator complex, phosphorylation of the carboxy-terminal domain of RNA Pol II and

DNA uncoiling by helicases. Such dissipative mechanisms cannot be accommodated within the equilibrium thermodynamic formalism, although attempts have been made to model some of them by equilibrium approximations (Raveh-Sadka et al. 2009; Mirny 2010).

Reliance on equilibrium models may be helpful initially but is not satisfactory as a fundamental approach. A key issue is that systems at equilibrium satisfy the condition of *detailed balance* (Lewis 1925), which originates in the time-reversal symmetry of the basic laws of physics (Tolman 1938). As shown by Hopfield in his classic study of kinetic proofreading, detailed balance is a biochemical constraint that prevents certain kinds of functionality from being achieved (Hopfield 1974). Evolution has found energy-dependent, nonequilibrium mechanisms that overcome this constraint, thereby enabling, for instance, the high levels of fidelity achieved by transcription and translation (Hopfield 1974; Zaher and Green 2009). Without incorporating nonequilibrium mechanisms, the full functionality of gene regulation cannot be explored.

The linear framework allows nonequilibrium mechanisms to be accommodated with no greater difficulty than equilibrium mechanisms. For example, it has already been used to model the nonequilibrium mechanism of post-translational modification (Thomson and Gunawardena 2009a, 2009b; Xu and Gunawardena 2012; Dasgupta et al. 2013). This has now been extended to gene regulation so as to include the nonequilibrium mechanisms mentioned above (Ahsendorf et al. 2013). The linear framework provides a strict generalisation of the thermodynamic formalism. If a system is capable of reaching thermodynamic equilibrium, then detailed balance leads to a simpler elimination procedure than the Matrix Tree theorem (Gunawardena 2012).

The results of the present paper are not only required for applying the linear framework to gene regulation, they also suggest new avenues for investigation.

One consequence of incorporating nonequilibrium mechanisms is that energy-dependent enzymatic reactions, such as phosphorylation, become effectively irreversible in typical cellular contexts. This cannot happen at equilibrium, where every reaction is reversible. For example, the yeast phosphate-response gene *PHO5* is regulated by both transcription-factor binding and nonequilibrium nucleosome remodeling and has been studied in quantitative detail (Kim and O'Shea 2008; Lam et al. 2008). The thermodynamic formalism was not applicable but an *ad hoc* model was constructed in which nucleosome assembly and disassembly are treated as irreversible reactions and detailed balance is not assumed (Kim and O'Shea 2008). This example is readily incorporated within the linear framework. The resulting graph is strongly connected and the corresponding gene regulation function can be algorithmically calculated using Corollary 1 (Ahsendorf et al. 2013).

Interestingly, with the introduction of irreversible reactions, there appears to be no fundamental reason why the resulting graphs have to be strongly connected. It is, of course, difficult to assert definitely that a graph derived from experimental knowledge is not strongly connected, as a single new edge may suffice to make it so. However, nonstrongly connected graphs already appear to be of value in analysing certain gene regulatory systems, such as those involved in epigenetic inheritance (Ahsendorf et al. 2013). Here, it is essential to understand Laplacian dynamics in full generality and to know how the steady states of an arbitrary graph are related to the graph structure, as described by Theorem 3.

We have shown in this paper that the steady-state behavior of nonstrongly connected graphs is qualitatively different from that of strongly-connected graphs. It can exhibit considerable plasticity. A surprising amalgam of two distinct regulatory modes has been revealed in yeast (Tirosh and Barkai 2008; Cairns 2009), where many genes incorporate, to varying degrees, the features of two extreme examples (Cairns 2009). One extreme consists of constitutively expressed genes with "open" promoters, which tend to have nucleosome-free regions and lack TATA boxes, while the other extreme consists of inducible genes with "covered" promoters, which have TATA boxes and dynamically-shifting nucleosomes. At present there is insufficient quantitative analysis of different genes to know whether this amalgam can give rise to flexible responses due to nonstrong connectivity of the underlying graphs, but it remains an intriguing possibility for subsequent investigation.

Gene regulation is a single-molecule process—there is only one gene present—and stochasticity must be taken into account at the outset. Theorem 4 shows that this can be done with complete rigour while retaining the familiar biochemical intuition for molecules in solution. It may prove helpful in future work to exploit Theorem 4 to translate the machinery of Markov process theory, as used to calculate averages and higher moments of probability distributions, into the combinatorial language of graphs and trees. Such translation between different fields often leads to new insights and may provide new tools for analysing fluctuations in gene expression.

It is of considerable topical importance to develop new methods for analysing gene regulation. The ENCODE project has recently provided a wealth of new data about the processes at work on DNA, including non-equilibrium processes such as histone PTMs (Stamatoyannopoulos 2012). The linear framework provides a starting point from which such data can be formally represented in mathematical structures—labeled, directed graphs—from which the regulatory behavior of an individual gene can be inferred, without relying on the constraining assumption of thermodynamic equilibrium. The results of the present paper provide the mathematical foundation for such biological analysis.

Acknowledgements We thank Arthur Jaffe for historical remarks on citations (Bott and Mayberry 1954; Jaffe 1965), David Perkinson for information about the proof in the Appendix, and an anonymous reviewer for helpful comments. The work undertaken here was supported by the United States NSF under Grant 0856285.

Appendix

We include for completeness a proof of the Matrix-Tree theorem as stated in Theorem 1. There are many versions of this and the one given here is adapted from David Perkinson's handout,[1] which is itself adapted from the proof in Tutte's book (Tutte 2001). This has the merits of being elementary, concise, and transparent. Let G be a graph on the vertices $1, \ldots, n$ and let the symbols e_{ij} denote the labels on the edges: $j \to i$, if, and only, if $e_{ij} \neq 0$. The entries of $\mathcal{L}(G)$ are then given by Eq. (2).

[1] http://people.reed.edu/~davidp/412/handouts/matrix-tree.pdf.

The proof of Theorem 1 is in two steps, first for the special case of a principal minor, which is the essential part of the argument. Some additional notation will be helpful. If A is any (not necessarily square) matrix let $A_{[u,v]}$ denote the matrix obtained from A by removing row u and column v. It is easier to work with the negative Laplacian, $L = -\mathcal{L}(G)$, to keep control of the signs. Since the minor is the determinant of a $(n-1) \times (n-1)$ matrix, this change incurs a sign of $(-1)^{n-1}$, so we can rewrite Eq. (16) for the jth principal minor as

$$\det L_{[j,j]} = \sum_{T \in \Theta_j(G)} \left(\prod_{k \to l \in T} e_{lk} \right), \tag{42}$$

which is what we want to prove.

Proof The argument does not depend on j, so take $j = 1$ to simplify the notation. The determinant can be written in the standard form as

$$\det L_{[1,1]} = \sum_{\pi \in S_{n-1}} \operatorname{sign}(\pi) L_{\pi(2)2} \cdots L_{\pi(n)n}, \tag{43}$$

where $\pi : \{2, \ldots, n\} \to \{2, \ldots, n\}$ is a permutation on $\{2, \ldots, n\}$, S_{n-1} is the symmetric group of all such permutations and $\operatorname{sign}(\pi) = \pm 1$ is the sign of the permutation. A permutation can be uniquely expressed as a product of disjoint cycles, where the cycle $(i_1 i_2 \cdots i_k)$, with $k > 1$, is the permutation that sends i_v to i_{v+1} for $v < k$ and cycles around to send i_k to i_1. A permutation may also fix an element by sending j to j and this is denoted by a degenerate cycle, (j). For example,

i	2	3	4	5	6	7
$\pi(i)$	4	5	7	3	6	2

, $\quad \pi = (247)(35)(6)$.

A nonzero summand in Eq. (43), corresponding to the permutation π, can be written in terms of nonzero symbols $e_{ij} \neq 0$ as follows. As mentioned above, π is a product of disjoint cycles. Each cycle gives rise to a product of nonzero symbols,

$$(i_1 i_2 \cdots i_k) \quad \text{gives } (-1)^k e_{i_2 i_1} e_{i_3 i_2} \cdots e_{i_k i_{k-1}} e_{i_1 i_k}, \tag{44}$$

where the signs arise because L is the negative of the Laplacian matrix defined by Eq. (2). Each fixed point of π, $\pi(j) = j$, introduces into the overall product the sum of all the nondiagonal symbols in column j of L,

$$\sum_{v \neq j} e_{vj}, \tag{45}$$

of which at least one symbol e_{vj} must be nonzero in order for the summand itself to be nonzero. Each cycle of odd length contributes a sign of -1 through Eq. (44), while each cycle of even length contributes a sign of -1 to $\operatorname{sign}(\pi)$. The resulting overall product, when expanded out, is therefore a sum of monomials of the form

$$(-1)^t e_{k_2 2} e_{k_3 3} \cdots e_{k_n n}, \tag{46}$$

where each symbol is nonzero, $e_{k_i i} \neq 0$, and t is the number of cycles in π.

Monomials of the form in Eq. (46) have a natural interpretation in terms of the graph G. They correspond to subgraphs of G in which there is exactly one outgoing edge from the vertices $2, \ldots, n$ and no outgoing edge from vertex 1. Lets call these *special* subgraphs. Given a special subgraph γ, it has an associated monomial e_γ which is the product of the symbols on all the edges and it is evident that this correspondence defines a bijection between special subgraphs and the monomials in Eq. (46).

Special subgraphs have a distinctive structure: If we start at any vertex other than 1, then there is a unique path leading away from this vertex. Because G has only finitely many vertices, this path must either intersect itself to form a cycle of edges or reach vertex 1 and stop. It follows that each connected component of a special subgraph either contains an unique cycle or is a tree rooted at 1. (Recall that the connected components of a directed graph are the disjoint pieces into which the graph falls if edge directions are ignored.) Of course, the cycle components may also contain tree-like spurs that lead to its unique cycle. Equation (44) shows that each cycle of edges in G may be uniquely identified with a permutation cycle.

Suppose given a permutation $\pi \in S_{n-1}$ and a special subgraph γ. If all the cycles of π appear as the cycles of cycle components of γ, then the symbols of any other edges in γ, including those of any other cycle components of γ, can be constructed in π by using the elements fixed by π, since these introduce, through the terms in Eq. (45), all the symbols needed for these remaining edges. It follows that e_γ is one of the monomials that appear when π is expanded into a sum of monomials through Eq. (46). Conversely, if e_γ appears in the monomial expansion of π then the cycles of π must occur among the cycles of cycle components of γ.

Given a special subgraph γ, we can now determine the multiplicity of e_γ in the determinant in Eq. (43). Suppose that γ has s cycle components, where $s > 0$. Any permutation $\pi \in S_{n-1}$ which uses exactly t of these cycles gives rise to e_γ with multiplicity $(-1)^t$ according to Eq. (46). There are $\binom{s}{t}$ ways of choosing t cycle components out of s. Considering all possibilities for t from $t = 0$ to $t = s$, the total multiplicity of e_γ is

$$\sum_{t=0}^{s} \binom{s}{t} (-1)^t = (1-1)^s = 0. \tag{47}$$

This amazing cancellation is the heart of the Matrix-Tree theorem.

The only remaining possibility is that γ has no cycle components, so that $s = 0$. But then γ is a spanning tree rooted at 1. Moreover, e_γ only appears as a monomial in the expansion of the identity permutation, whose monomial expansion is given by

$$\prod_{j=2}^{n} \left(\sum_{v \neq j} e_{vj} \right). \tag{48}$$

Evidently, e_γ appears in Eq. (48) with multiplicity 1. Of course, Eq. (48) gives rise to monomials other than those coming from spanning trees rooted at 1 but these other monomials disappear through cancellation because of Eq. (47). We are left with exactly the sum of monomials over rooted spanning trees in Eq. (42). □

To complete the proof of Theorem 1, it remains to deal with a non-principal minor, $\det L_{[i,j]}$, which contributes the sign $(-1)^{i+j}$ to Eq. (16). This has nothing to do with Laplacian matrices and depends only on the fact that $1 \cdot \mathcal{L}(G) = 0$. Some further notation will be helpful. Let A be a $n \times (n-1)$ matrix and let A_{i*} denote its i-th row. We want to consider various $(n-1) \times (n-1)$ matrices made from these rows and we will write these matrices by listing their rows in order. For instance, the $(n-1) \times (n-1)$ submatrix of A in which the first row is omitted, is written $[A_{2*}|A_{3*}|\cdots|A_{n*}]$. We will also use the convention that, in an ordered sequence, a bar over a term signifies its absence, so we can also write this same submatrix as $[\overline{A_{1*}}|A_{2*}|\cdots|A_{n*}]$.

Lemma 4 *If A is any $n \times (n-1)$ matrix for which $1 \cdot A = 0$, then*

$$\det\big[\,\overline{A_{1*}}|A_{2*}|\cdots|A_{n*}\big] = (-1)^{i-1}\det\big[A_{1*}|A_{2*}|\cdots|\overline{A_{i*}}|\cdots|A_{n*}\big].$$

Proof Let $B = [A_{1*} + A_{i*}|A_{2*}|\cdots|\overline{A_{i*}}|\cdots|A_{n*}]$. By construction, $1 \cdot B = 0$, so that $\det B = 0$. Because the determinant is multilinear,

$$\det\big[A_{1*}|A_{2*}|\cdots|\overline{A_{i*}}|\cdots|A_{n*}\big] = (-1)\det\big[A_{i*}|A_{2*}|\cdots|\overline{A_{i*}}|\cdots|A_{n*}\big].$$

The determinant is also antisymmetric for interchange of rows, so that

$$\det\big[A_{i*}|A_{2*}|\cdots|\overline{A_{i*}}|\cdots|A_{n*}\big] = (-1)\det\big[A_{2*}|A_{i*}|\cdots|\overline{A_{i*}}|\cdots|A_{n*}\big].$$

It takes $i-2$ successive row interchanges to bring A_{i*} back to where it originally was in the list, from which the result follows. □

Now let A be the $n \times (n-1)$ matrix obtained from $L = -\mathcal{L}(G)$ by removing the jth column. Then, using Lemma 4 twice, we see that

$$\begin{aligned}
\det L_{[i,j]} &= \det\big[A_{1*}|\cdots|\overline{A_{i*}}|\cdots|A_{n*}\big] \\
&= (-1)^{i-1}\det\big[\overline{A_{1*}}|\cdots|A_{i*}|\cdots|A_{n*}\big] \\
&= (-1)^{i+j-2}\det\big[A_{1*}|\cdots|\overline{A_{j*}}|\cdots|A_{n*}\big] \\
&= (-1)^{i+j}\det L_{[j,j]}.
\end{aligned}$$

Combining this with Eq. (42) yields Eq. (16) and proves Theorem 1.

References

Ackers, G. K., Johnson, A. D., & Shea, M. A. (1982). Quantitative model for gene regulation by lambda phage repressor. *Proc. Natl. Acad. Sci. USA*, *79*, 1129–1133.

Agaev, R. P., & Chebotarev, P. Y. (2000). The matrix of maximum out forests of a digraph and its applications. *Autom. Remote Control*, *61*, 1424–1450.

Ahsendorf, T., Wong, F., Eils, R., & Gunawardena, J. (2013, in preparation). A framework for modelling eukaryotic gene regulation that accommodates non-equilibrium mechanisms.

Bintu, L., Buchler, N. E., Garcia, G. G., Gerland, U., Hwa, T., Kondev, J., Kuhlman, T., & Phillips, R. (2005a). Transcriptional regulation by the numbers: applications. *Curr. Opin. Genet. Dev.*, *15*, 125–135.

Bintu, L., Buchler, N. E., Garcia, G. G., Gerland, U., Hwa, T., Kondev, J., & Phillips, R. (2005b). Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.*, *15*, 116–124.

Bott, R., & Mayberry, J. P. (1954). Matrices and trees. In O. Morgenstern (Ed.), *Economic activity analysis* (pp. 391–400). New York: Wiley

Cairns, B. R. (2009). The logic of chromatin architecture and remodelling at promoters. *Nature*, *461*, 193–198.

Chebotarev, P. (2010). Comment on 'Consensus and cooperation in networked multi-agent systems'. *Proc. IEEE*, *98*, 1353–1354.

Chebotarev, P., & Agaev, R. (2002). Forest matrices around the Laplacian matrix. *Linear Algebra Appl.*, *356*, 253–274.

Chebotarev, P. Y., & Agaev, R. P. (2009). Coordination in multiagent systems and Laplacian spectra of digraphs. *Autom. Remote Control*, *70*, 469–483.

Chen, W. K. (1971). Applied graph theory. In *Applied mathematics and mechanics*, Amsterdam: North-Holland.

Chung, F. R. K. (1997). *Spectral graph theory*. *Regional conference series in mathematics: Vol. 92*. Providence: Am. Math. Soc.

Colquhoun, D. (2006). The quantitative analysis of drug-receptor interactions: a short history. *Trends Pharmacol. Sci.*, *27*, 149–157.

Cornish-Bowden, A. (1995). *Fundamentals of enzyme kinetics* (2nd ed.). London: Portland Press.

Dasgupta, T., Croll, D. H., Owen, J. A., Vander Heiden, M. G., Locasale, J. W., Alon, U., Cantley, L. C., & Gunawardena, J. (2013). A fundamental trade off in covalent switching and its circumvention in glucose homeostasis. Submitted.

Feinberg, M., & Horn, F. (1977). Chemical mechanism structure and the coincidence of the stoichiometric and kinetic subspace. *Arch. Ration. Mech. Anal.*, *66*, 83–97.

Gertz, J., Siggia, E. D., & Cohen, B. A. (2009). Analysis of combinatorial cis-regulation in synthetic and genomic promoters. *Nature*, *457*, 215–218.

Gunawardena, J. (2012). A linear framework for time-scale separation in nonlinear biochemical systems. *PLoS ONE*, *7*, e36321.

He, X., Samee, M. A. H., Blatti, C., & Sinha, S. (2010). Thermodynamics-based models of transcriptional regulation by enhancers: the roles of synergistic activation, cooperative binding and short-range repression. *PLoS Comput. Biol.*, *6*, e1000935.

Hill, T. L. (1966). Studies in irreversible thermodynamics IV. Diagrammatic representation of steady state fluxes for unimolecular systems. *J. Theor. Biol.*, *10*, 442–459.

Hill, T. L. (1985). *Cooperativity theory in biochemistry: steady-state and equilibrium systems*. *Springer series in molecular biology*. New York: Springer.

Hirsch, M. W., & Smale, S. (1974). *Differential equations, dynamical systems and linear algebra. Pure and applied mathematics*. San Diego: Academic Press.

Hopfield, J. J. (1974). Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. USA*, *71*, 4135–4139.

Horn, R. A., & Johnson, C. A. (1985). *Matrix analysis*. Cambridge: Cambridge University Press.

Jaffe, A. (1965). Divergence of perturbation theory for bosons. *Commun. Math. Phys.*, *1*, 127–149.

Janssens, H., Hou, S., Jaeger, J., Kim, A. R., Myasnikova, E., Sharp, D., & Reinitz, J. (2006). Quantitative and predictive model of transcriptional control of the drosophila melanogaster even skipped gene. *Nat. Genet.*, *38*, 1159–1165.

van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry*. Amsterdam: Elsevier.

Kelly, F. P. (2011). *Reversibility and stochastic networks*. Cambridge: Cambridge University Press.

Kenakin, T. (2005). New concepts in drug discovery: collateral efficacy and permissive antagonism. *Nat. Rev. Drug Discov.*, *4*, 919–927.

Kim, H. D., & O'Shea, E. K. (2008). A quantitative model of transcription factor-activated gene expression. *Nat. Struct. Mol. Biol.*, *15*, 1192–1198.

King, E. L., & Altman, C. (1956). A schematic method of deriving the rate laws for enzyme-catalyzed reactions. *J. Phys. Chem.*, *60*, 1375–1378.

Kirchhoff, G. (1847). Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird. *Ann. Phys. Chem.*, *72*, 497–508.

Koshland, D. E., Némethy, G., & Filmer, D. (1966). Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry*, *5*, 365–385.

Kuhlman, T., Zhang, Z., Saier, M. H. Jr., & Hwa, T. (2007). Combinatorial transcriptional control of the lactose operon of Escherichia coli. *Proc. Natl. Acad. Sci. USA*, *104*, 6043–6048.

Lam, F. H., Steger, D. J., & O'Shea, E. K. (2008). Chromatin decouples promoter threshold from dynamic range. *Nature*, *453*, 246–250.

Lean, A. D., Stadel, J. M., & Lefkowitz, R. J. (1980). A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled β-adrenergic receptor. *J. Biol. Chem.*, *255*, 7108–7117.

Lewis, G. N. (1925). A new principle of equilibrium. *Proc. Natl. Acad. Sci. USA*, *11*, 179–183.

Magnus, J. R., & Neudecker, H. (1988). *Matrix differential calculus with applications in statistics and econometrics*. Chichester: Wiley

Merris, R. (1994). Laplacian matrices of graphs: a survey. *Linear Algebra Appl.*, *198*, 143–176.

Michaelis, L., & Menten, M. (1913). Die kinetik der Invertinwirkung. *Biochem. Z.*, *49*, 333–369.

Mirny, L. (2010). Nucleosome-mediated cooperativity between transcription factors. *Proc. Natl. Acad. Sci. USA*, *107*(22), 534–539.

Monod, J., Wyman, J., & Changeux, J. P. (1965). On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.*, *12*, 88–118.

Moon, J. W. (1970). *Counting Labelled Trees*. *Canadian mathematical monographs: Vol. 1*. Ottawa: Canadian Mathematical Society.

Nishikawa, T., & Motter, A. E. (2010). Network synchronization landscape reveals compensatory structures, quantization, and the positive effect of negative interactions. *Proc. Natl. Acad. Sci. USA*, *107*(10), 342–347.

Olfati-Saber, R., Fax, J. A., & Murray, R. M. (2007). Consensus and cooperation in networked multi-agent systems. *Proc. IEEE*, *95*, 215–233.

Olfati-Saber, R., & Murray, R. M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control*, *49*, 1520–1533.

Pecora, L. M., & Carroll, T. L. (1998). Master stability functions for synchronized coupled systems. *Phys. Rev. Lett.*, *80*, 2109–2112.

Raveh-Sadka, T., Levo, M., & Segal, E. (2009). Incorporating nucleosomes into thermodynamic models of transcription regulation. *Genome Res.*, *19*, 1480–1496.

Schnakenberg, J. (1976). Network theory of microscopic and macroscopic behaviour of master equation systems. *Rev. Mod. Phys.*, *48*, 571–586.

Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U., & Gaul, U. (2008). Predicting expression patters from regulatory sequence in Drosophila segmentation. *Nature*, *451*, 535–540.

Segal, E., & Widom, J. (2009). From DNA sequence to transcriptional behaviour: a quantitative approach. *Nat. Rev. Genet.*, *10*, 443–456.

Setty, Y., Mayo, A. E., Surette, M. G., & Alon, U. (2003). Detailed map of a cis-regulatory input function. *Proc. Natl. Acad. Sci. USA*, *100*, 7702–7707.

Sherman, M. S., & Cohen, B. A. (2012). Thermodynamic state ensemble models of cis-regulation. *PLoS Comput. Biol.*, *8*, e1002407.

Stamatoyannopoulos, J. (2012). What does our genome encode? *Genome Res.*, *22*, 1602–1611.

Thomson, M., & Gunawardena, J. (2009a). The rational parameterisation theorem for multisite post-translational modification systems. *J. Theor. Biol.*, *261*, 626–636.

Thomson, M., & Gunawardena, J. (2009b). Unlimited multistability in multisite phosphorylation systems. *Nature*, *460*, 274–277.

Tirosh, I., & Barkai, N. (2008). Two strategies for gene regulation by promoter nucleosomes. *Genome Res.*, *18*, 1084–1091.

Tolman, R. C. (1938). *The principles of statistical mechanics*. Oxford: Clarendon Press.

Tutte, W. T. (1948). The dissection of equilateral triangles into equilateral triangles. *Proc. Camb. Philol. Soc.*, *44*, 463–482.

Tutte, W. T. (2001). *Graph theory*. *Encyclopedia of mathematics and its applications: Vol. 21*. Cambridge: Cambridge University Press.

Whitin, T. M. (1954). An economic application of 'Matrices and trees'. In O. Morgenstern (Ed.), *Economic activity analysis* (pp. 401–418). New York: Wiley

Xu, Y., & Gunawardena, J. (2012). Realistic enzymology for post-translational modification: zero-order ultrasensitivity revisited. *J. Theor. Biol.*, *311*, 139–152.

Zaher, H. S., & Green, R. (2009). Fidelity at the molecular level: lessons from protein synthesis. *Cell*, *136*, 746–762.

Zinzen, R. P., Senger, K., Levine, M., & Papatsenko, D. (2006). Computational models for neurogenic gene expression in the Drosophila embryo. *Curr. Biol.*, *16*, 1358–1365.