

# Generic Temperature Response of Large Biochemical Networks

Julian B. Voits <sup>1</sup> and Ulrich S. Schwarz <sup>1,2,\*</sup>

<sup>1</sup>*Institute for Theoretical Physics, Heidelberg University, Heidelberg, Germany*

<sup>2</sup>*BioQuant-Center for Quantitative Biology, Heidelberg University, Heidelberg, Germany*



(Received 25 March 2024; accepted 9 October 2025; published 13 November 2025)

Biological systems are remarkably susceptible to relatively small temperature changes. The most obvious example is fever, when a modest rise in body temperature of only few Kelvin has strong effects on our immune system and how it fights pathogens. Another very important example is climate change, when even smaller temperature changes lead to dramatic shifts in ecosystems. Although it is generally accepted that the main effect of an increase in temperature is the acceleration of biochemical reactions according to the Arrhenius equation, it is not clear how it affects large biochemical networks with complicated architectures. For developmental systems such as fly and frog, it has been shown that the system response to temperature deviates in a characteristic manner from the linear Arrhenius plot of single reactions, but a rigorous explanation has not been given yet. Here we use a graph-theoretical interpretation of the mean first-passage times of a biochemical master equation to give a statistical description. We find that in the limit of large system size and if the network has a bias towards a target state, then the Arrhenius plot is generically quadratic, in excellent agreement with numerical simulations for large networks as well as with experimental data for developmental times in fly. We also discuss under which conditions this generic response can be violated, for example for linear chains, which have only one spanning tree.

DOI: [10.1103/nyr3-5my3](https://doi.org/10.1103/nyr3-5my3)

## I. INTRODUCTION

High fever has a dramatic effect on our body, but from the physics point of view, it is only a modest change: increasing the human body temperature by three degrees is less than one percent on the absolute temperature scale [1]. An increase of the same amount due to global warming would most likely result in an extensive loss of biodiversity [2,3]. As illustrated by these examples, complex biological systems are remarkably susceptible to changes in temperature. The explanation for these sensitive responses to temperature was already given in the 19th century by Arrhenius, who suggested that the rates of all chemical reactions exponentially increase with rising temperature [4]. This can be verified by an Arrhenius plot, in which the logarithm of the kinetic rate decreases linearly as a function of the inverse temperature  $1/T$ . This insight has been confirmed over and over again for single chemical reactions and has led to many important advances over the last century [5]. Moreover, it has been made more rigorous by the theories by Kramers [6,7] and Eyring [8,9] on potential barriers as transition states in chemical reactions.

Biochemical systems do not behave differently from chemical systems in this regard, except that they are typically

limited by protein denaturation at high temperatures [10]. A large body of experimental work exists on temperature effects in biochemical networks, ranging from intracellular temperature effects [11–23] to the impact of temperature on growth and development [24–33]. We refer to Ref. [34] for a comparison of more than 1000 studies on biological temperature responses. From this body of experimental work, the picture arises that the system's response typically does not show the Arrhenius form of single reactions. However, a fundamental theoretical understanding of this striking observation is missing and it is an open question if such networks show a generic response to temperature changes. Yet this question is key if one aims at understanding biochemical systems such as the immune system or ecosystems from a theoretical point of view.

There is one subject area for which the investigation of temperature effects has been relatively systematic and comprehensive, and that is the case of biochemical oscillators [35–46]. Here, the general picture has emerged that they often come with compensation mechanisms, which can assure that their functioning is unaltered within the physiological temperature range. A notable example for this general observation are circadian clocks, whose function is to instruct the organism about upcoming changes due to the diurnal cycle. Because temperature changes are one of the main consequences of changing solar input, circadian clocks are typically temperature compensated, to ensure reliable time measurements [39–43]. Another important example of temperature compensation is chemotaxis of swimming organisms such as *E. Coli*, which have to reliably find food sources despite temperature changes and gradients in their environment [21].

\*Contact author: [schwarz@thphys.uni-heidelberg.de](mailto:schwarz@thphys.uni-heidelberg.de)

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

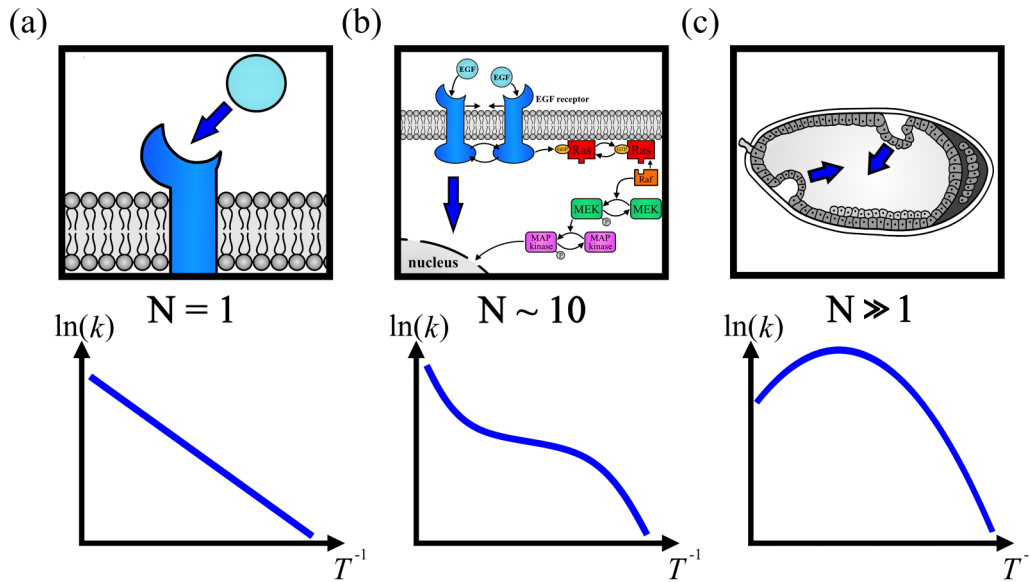


FIG. 1. The Arrhenius plot depends on network size. (a) The temperature dependence of the rate of a single biochemical reaction, such as ligand-receptor binding, usually follows the Arrhenius equation, resulting in a linear relationship in the Arrhenius plot (logarithmic rates against inverse temperature). (b) For networks of intermediate size, such as the MAPK-signaling pathway shown here, the temperature response is more complex and depends on microscopic details. (c) In large biochemical networks, such as those governing the development from a *Drosophila* fertilized egg to a larva, a quadratic dependence emerges in the Arrhenius plot.

Although the existence of temperature compensation for biochemical oscillators and chemotaxis proves the relevance of temperature for biochemical networks, it does not instruct us about its generic effect, exactly because in these cases it is compensated by specific mechanisms, typically by the action of proteins that have evolved for that purpose.

Developmental systems seem to have evolved less temperature compensation and therefore appear to be better model systems to study the effect of temperature changes on large biochemical networks [47,48]. Measurements of the prepupal development time of *Drosophila* under different temperature conditions were published almost 100 years ago by Bliss [25]. The results clearly showed a concave curve in the Arrhenius plot, as opposed to the straight line usually found for single reactions. About a decade later, Powsner obtained similar results for the embryonic, larval and pupal phases of the development of *Drosophila*, already hypothesizing that this may be a consequence of the complexity of the underlying network [26]. More recently, Crapse *et al.* published a study in which the embryonic development of *Drosophila* was subdivided into 12 phases [24]. The authors measured the temperature dependence of the completion time of each of them. On this level, a concave temperature response emerged as well, which could be accurately described by a quadratic fit in the Arrhenius plot. In the same study, similar results were also found for *Xenopus laevis* embryos. The authors also performed simulations for a linear sequence of Poisson processes with statistically distributed parameters, but the results could not fully reproduce the shapes of the experimental curves [24]. However, this approach disregards the fact that the underlying biochemical networks are bound to be highly nonlinear, with complicated feedback mechanisms and network motifs [49]. Even more recently, Rombouts *et al.* confirmed the quadratic shape in the Arrhenius plot of the development rates of

*Xenopus* embryos and also observed it for the development of zebra fish (*Danio rerio*) embryos, while for *Caenorhabditis elegans* the rate shows a more biphasic behavior [48]. In general, although there has been some work on understanding the temperature dependence of development rates from first principles [50–54], the main body of work has remained empirical [55–60].

The general question addressed in this work is illustrated in Fig. 1. The temperature dependence of a single reaction is usually described well by the Arrhenius equation, resulting in a linear relationship in the Arrhenius plot [Fig. 1(a)]. However, when considering complex biological processes, e.g., the whole mitogen-activated protein kinase (MAPK) signaling pathway triggered by epidermal growth factors (EGF), that transmits extracellular signals to the nucleus and regulates differentiation [61], the linear relation is lost and microscopic details might govern the response [Fig. 1(b)]. In the limit of very large systems, such as whole organisms such as *Drosophila* developing from a fertilized egg to a larva [62], the overall rate usually approaches a quadratic dependence [Fig. 1(c)]. Here we provide a rigorous mathematical derivation of why large biochemical networks typically show a quadratic shape in the Arrhenius plot. Our proof is based on calculating first-passage times using a graph-theoretical decomposition of the network into spanning trees and forests. Our result is valid for all network architectures, including feedback loops, as long as several spanning trees stay relevant in the limit of large system size. Thus, our work solves the longstanding question of what the generic temperature response of large and complicated biochemical networks is. To support our analytical findings, we have developed a numerical procedure that generates networks of given sizes and with a bias towards an endpoint, and then uses the graph-theoretical interpretation of mean first-passage times to generate their

distribution. This numerical procedure confirms the generic temperature response predicted by our theory. We also show that our results are in excellent agreement with experimental data from the developmental systems of the fly. Finally we discuss the cases in which the generic response does not occur, including linear chains, which have only one spanning tree. Our approach paves the way to deal with smaller biochemical systems and more specific temperature responses, including protein denaturation.

## II. RESULTS

### A. Meaning of mean first-passage time

A general framework to mathematically describe a chemical reaction network is to express it as a time-dependent probability distribution  $p_i(t)$  on  $N_v$  discrete states, where without loss of generality we denote states 1 and  $N_v$  as start and target states, respectively. For a linear chain, we would have  $N = N_v - 1$  reactions, each with forward and backward directions, while in general, the system could contain many loops. In the following, we will first use  $N$  as main variable; for the numerical part coming later, it is more convenient to use  $N_v = N + 1$ .

The time evolution of  $p_i$  is governed by the master equation [63,64]:

$$\dot{p}_i(t) = \sum_{j=1}^{N+1} (k_{ji}p_j(t) - k_{ij}p_i(t)), \quad (1)$$

or in vectorial notation:

$$\dot{\vec{p}}(t) = K\vec{p}(t), \quad (2)$$

where  $k_{ij}$  is the transition rate from state  $i$  to state  $j$  and

$$K_{ij} = \begin{cases} k_{ji} & \text{if } i \neq j \\ -\sum_m k_{im} & \text{if } i = j \end{cases}. \quad (3)$$

Realistic biochemical networks typically exhibit a large number of intertwined reactions, meaning that  $N \gg 1$  and Eq. (1) then describes a large system of coupled ordinary differential equations, for which it is difficult to find a full mathematical solution. Another important limitation is that often not all connections and/or rates might be known. Together, this raises the question if one can make progress without explicitly solving the complete system.

Indeed, certain features of the system described by Eq. (1) can be computed without the need for a complete solution [65]. A notable example are first-passage times (also called first-hitting times or exit times) [66,67] and in particular, the mean first-passage time (MFPT), which characterizes the typical completion time of the process of interest. The main example in this work is a developmental process seen as the consequence of a complex network of biochemical reactions. Developmental systems are very large and complex and besides biochemistry also involve many spatial processes, such as cytoskeletal rearrangements [68]. However, experimental measurements of heat generation in zebra fish embryos combined with modeling of biochemical networks suggest that even in this case, the rate-limiting steps are in the biochemical control system for the cell cycle [69], which in itself is again

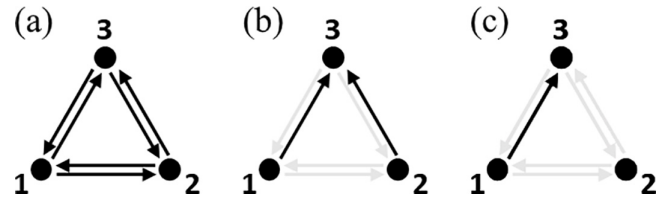


FIG. 2. An illustration of the graph-theoretical decomposition of networks into spanning trees and forests used throughout this work. (a) The full directed graph on three vertices. (b) A spanning tree rooted at 3. (c) A spanning forest of two trees rooted at 2 and 3.

a large and complex biochemical network [70]. To give an instructive example of a transparent biochemical network that is active during development and that can be described with the master equation from Eq. (1), one can think of the mitogen-activated protein kinase (MAPK) signaling pathway triggered by epidermal growth factors (EGF) shown in Fig. 1(b). In this case, the MFPT describes the typical time for the signal to reach from the plasma membrane to the nucleus, where gene expression is changed. Here we ask how such MFPTs can be calculated for this and more complex networks, and how they depend on temperature.

### B. Solution with graph theory

#### 1. Some concepts from graph theory

It is often a useful approach to represent networks described by a master equation as weighted and directed graphs, i.e., vertices and directed edges (pairs of vertices), where the edges have a weight attributed to them [71]. Then the vertices are the states, the directed edges are the possible jumps and their weights are the jumping rates  $k_{ij}$  from vertex  $i$  to vertex  $j$ . In the following, graphs are always understood as weighted and directed. Also, the rates  $k_{ij}$  are simply referred to as edges, for the ease of notation, setting  $k_{ij} = 0$  if there is no transition from  $i$  to  $j$ .

A tree rooted at a vertex  $i$  is a cycle-free graph where there is a directed path from any other vertex to  $i$ . Note that this requires  $i$  to be an endpoint because any outgoing edge from  $i$  would create a cycle. It is then also clear that the root of a tree is unique. A disjoint union of trees is called a forest. A spanning tree (forest) is a subgraph of a given graph that contains all its vertices and is a tree (forest). This is illustrated in Fig. 2.

The (in-degree) Laplacian matrix  $L$  of a graph is defined as:

$$L_{ij} := \begin{cases} -k_{ij} & \text{if } i \neq j \\ \sum_{m \neq i} k_{im} & \text{if } i = j \end{cases}, \quad (4)$$

so the  $i$ th entry on the main diagonal is the sum of all incoming edges and  $L_{ij}$  for  $i \neq j$  are the negative transition rates from  $i$  to  $j$ .

#### 2. Graph-theoretical interpretation of FPT moments

The standard textbook approach to find first-passage times (FPTs) is to evaluate them in Laplace space [66,67,72]. However, this becomes increasingly cumbersome for larger systems. Here, we instead use a solution that can be derived

with graph theory [73–75]. Consider the master equation on a finite space  $i = 1, \dots, N + 1$  with time-independent rates  $k_{ij}$ . Then the formal solution to Eq. (2) is given by:

$$\bar{p}(t) = \exp(Kt)\bar{p}_0, \quad (5)$$

with  $\exp$  denoting the matrix exponential and  $\bar{p}_0 = \bar{p}(0)$ . Since  $1 - p_{N+1} = \sum_{i=1}^N p_i$ , Eq. (5) can be restricted to only the first  $N$  components. In particular,  $K$  is then no longer singular.

Asking for the FPT to reach state  $N + 1$ , this state can be assumed to be absorbing, i.e.,  $k_{N+1i} = 0$  for all  $i$ , because we are only interested in the first time that it is reached. The FPT density  $f(t)$  then follows from:

$$p(\tau \leq t) = p_{N+1}(t) \quad (6)$$

$$\begin{aligned} \Rightarrow f(t) &= \dot{p}_{N+1}(t) \\ &= \hat{e}_{N+1}^T K \bar{p}(t) \\ &= \hat{e}_{N+1}^T K \exp(Kt) \bar{p}_0, \end{aligned} \quad (7)$$

where  $\hat{e}_i$  denotes the  $i$ th unit vector. We define the FPT density to reach  $N + 1$  starting at state  $i$

$$\begin{aligned} f_i(t) &:= \hat{e}_{N+1}^T K \exp(Kt) \hat{e}_i \\ &= \hat{e}_i^T \exp(K^T t) K^T \hat{e}_{N+1}, \end{aligned} \quad (8)$$

where the second equality uses the symmetry of the inner product. In vector form, this now reads:

$$\begin{aligned} \vec{f} &= \exp(K^T t) K^T \hat{e}_{N+1} \\ &= \exp(K^T t) \vec{f}_0, \end{aligned} \quad (9)$$

where  $\vec{f}_0 = (k_{1,N+1}, \dots, k_{N,N+1})^T$ . A distribution of this shape is called phase-type distribution [76]. The moments are obtained by integration by parts:

$$\begin{aligned} \langle \tau^n \rangle &= \int dt t^n \vec{f} \\ &= -n(K^T)^{-1} \int dt t^{n-1} \exp(K^T t) \vec{f}_0 \\ &= -n(K^T)^{-1} \langle \tau^{n-1} \rangle. \end{aligned} \quad (10)$$

Equivalently, we have

$$K^T \langle \tau^n \rangle = -n \langle \tau^{n-1} \rangle. \quad (11)$$

Here  $\langle \tau^n \rangle$  is the  $n$ th moment of the FPT when starting at state  $i$ . Iterative application of the formula above yields:

$$\langle \tau^n \rangle = n!((-K^T)^{-1})^n \vec{1}, \quad (12)$$

where  $\vec{1} = (1, 1, \dots, 1)^T$ . Note that

$$(-K^T)_{ij=1,\dots,N} = \begin{cases} -k_{ij} & \text{if } i \neq j \\ \sum_{m=1}^{N+1} k_{im} & \text{if } i = j \end{cases} \quad (13)$$

can be interpreted as the submatrix obtained by deleting the  $(N + 1)$ th row and column of the Laplace matrix of the weighted and directed graph identifying the states with vertices and the edges with the transitions between them, where  $k_{ij}$  is the weight for the edge going from  $i$  to  $j$ .

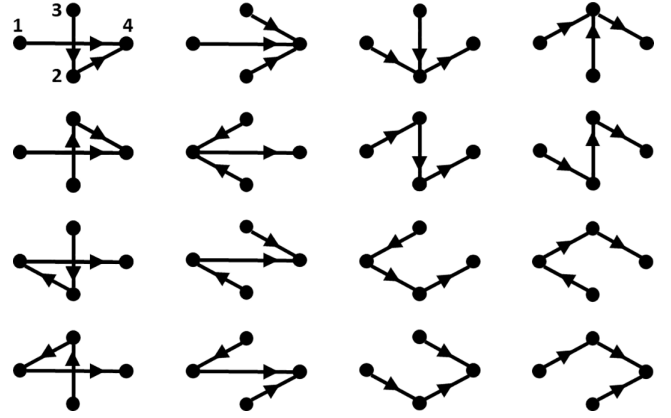


FIG. 3. All spanning trees of the complete graph on four vertices rooted at 4. The graph-theoretical interpretation of the sum in Eq. (14) consists in finding the spanning trees. The vertex labels are for all graphs as indicated for the first one.

By a generalization of Kirchhoff's theorem to weighted and directed graphs [77], the determinant of  $-K^T$  is given by:

$$\det(-K^T) = \sum_{\mathcal{T}_{[N+1]}} w(\mathcal{T}), \quad (14)$$

where  $\mathcal{T}_{[N+1]}$  denotes the spanning trees of the corresponding graph rooted at  $N + 1$  and their weights are defined as  $w(\mathcal{T}) := \prod_{k_{ij} \in E(\mathcal{T})} k_{ij}$ , with  $E(\mathcal{T})$  being the edge set of  $\mathcal{T}$ . Figure 3 shows  $\mathcal{T}_{[4]}$  for the complete graph on four vertices as an example.

$(-K^T)^{-1}$  can be expressed as:

$$(-K^T)^{-1}_{ij} = \frac{(-1)^{i+j}}{\det(-K^T)} M_{ji}, \quad (15)$$

where  $M_{ji}$  is the determinant of the  $(j, i)$  minor of  $-K^T$ , which can also be expressed as a sum with a graph-theoretical interpretation [73]:

$$M_{ji} = (-1)^{i+j} \sum_{\mathcal{F}_{[j,N+1]}^{i \rightarrow j}} w(\mathcal{F}), \quad (16)$$

where  $\mathcal{F}_{[j,N+1]}^{i \rightarrow j}$  are the spanning forests with two trees of the graph, one rooted at  $j$  and containing  $i$  and the other one rooted at  $N + 1$ . Figure 4 shows  $\mathcal{F}_{[i,4]}^{1 \rightarrow i}$  for the complete graph on four vertices.

This yields

$$(-K^T)^{-1}_{ij} = \frac{\sum_{\mathcal{F}_{[j,N+1]}^{i \rightarrow j}} w(\mathcal{F})}{\sum_{\mathcal{T}_{[N+1]}} w(\mathcal{T})}, \quad (17)$$

and therefore, Eq. (12) becomes [73]:

$$\langle \tau^n \rangle = n! \sum_{j_0, \dots, j_n=1}^N \frac{\prod_{m=1}^n \sum_{\mathcal{F}_{[j_m, N+1]}^{j_{m-1} \rightarrow j_m}} w(\mathcal{F})}{(\sum_{\mathcal{T}_{[N+1]}} w(\mathcal{T}))^n} \hat{e}_{j_0}. \quad (18)$$

The mean in particular reads [73,74]:

$$\langle \tau \rangle = \frac{\sum_{j=1}^N \sum_{\mathcal{F}_{[j,N+1]}^{1 \rightarrow j}} w(\mathcal{F})}{\sum_{\mathcal{T}_{[N+1]}} w(\mathcal{T})}. \quad (19)$$



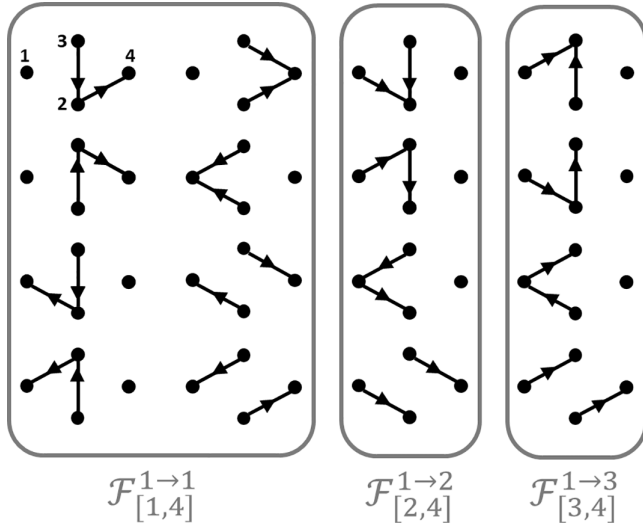


FIG. 4. All spanning forests of the complete graph on four vertices rooted at 4. The graph-theoretical interpretation of the summands in Eq. (16) consists in finding the two-tree spanning forests  $\mathcal{F}_{[1,4]}^{1 \rightarrow 1}$ ,  $\mathcal{F}_{[2,4]}^{1 \rightarrow 2}$ , and  $\mathcal{F}_{[3,4]}^{1 \rightarrow 3}$  of the complete graph on four vertices rooted at 4. The vertex labels are for all graphs as indicated for the first one.

Note that similar graph-theoretical counting schemes also exist for the steady-state distributions of ergodic networks [78,79]. In Appendix A, we show how this formula can be used to derive MFPTs for two example networks for which the results are also known from Laplace transforms.

Overall, the graph-theoretical approach results in three possible approaches to the FPTs of a homogeneous master equation. The first one consists in solving the matrix exponential as given in Eq. (9) or, alternatively, the corresponding differential equation:

$$\dot{\vec{f}} = K^T \vec{f}, \quad (20)$$

with  $\vec{f}_0$  as above. This is numerically feasible for systems of moderate size. If no full analytic solution is available, the moments of the FPT can be computed by either inverting  $-K^T$  algebraically and using Eq. (12) or by solving the combinatorial problem in Eq. (18).

First-passage time distributions can be broad, especially when considering the statistics of rare events, and higher moments are therefore often of interest [66]. For processes with a clear bias towards a final state, the distribution of the completion time tends to be more narrow and its mean characterizes it sufficiently well [80], which is likely to be the case for most developmental systems. Indeed, the development times of *Drosophila* embryos vary only by a few percent [81].

### C. Effect of temperature

Temperature now enters the description via the rate constants. By the Arrhenius equation [4], the temperature dependence of the  $k_{ij}$  can be described as follows:

$$k_{ij} = A_{ij} e^{-\frac{E_{ij}}{k_B T}} = A_{ij} e^{-\frac{E_{ij} [\frac{J}{mol}]}{RT}}, \quad (21)$$

where  $T$  denotes the temperature,  $k_B$  Boltzmann's constant,  $A_{ij}$  a (temperature-independent) prefactor and  $E_{ij}$  is the

activation energy. For large systems such as developmental ones, one often expresses  $E_{ij}$  in  $\frac{J}{mol}$ . Then  $k_B$  has to be replaced by the universal gas constant  $R$ .

Since most biological organisms have evolutionarily adapted to a relatively fixed thermal environment, it is reasonable to rephrase the Arrhenius equation in terms of the rate constant at a reference temperature  $T_0$ , introducing  $k_{ij}^0 = k_{ij}(T = T_0)$ .  $T_0$  describes the temperature at which the organism usually operates. For instance, for endothermic animals, the body temperature would serve as the natural reference point, while for exothermic animals, it would be the typical temperature of the environment. The  $k_{ij}^0$  can be seen as the standard rates, and a crucial assumption for the following analysis is that these standard rates are independent of the activation energies  $E_{ij}$ . The Arrhenius equation can then be expressed in terms of these parameters as:

$$k_{ij} = A_{ij} e^{-\frac{E_{ij}}{k_B T}} = k_{ij}^0 e^{-\Delta\beta E_{ij}}, \quad (22)$$

where  $\Delta\beta := \frac{1}{k_B T} - \frac{1}{k_B T_0}$ .

### D. Partition sums and generating functions

We note that the number of spanning trees on a complete graph on  $N + 1$  vertices is given by the Cayley formula [82] as  $(N + 1)^{N-1}$ , so it grows faster than exponentially with  $N$ . Even though many of the rates vanish, for sufficiently complex networks one can expect the numerator and the denominator of Eq. (19) to be the sum of many such graphs. This motivates the introduction of a partition sumlike quantity  $Z$  summing over the spanning trees:

$$Z_{\mathcal{T}} = \sum_{\mathcal{T}_{[N+1]}} w(\mathcal{T}), \quad (23)$$

and for the two-tree spanning forests:

$$Z_{\mathcal{F}} = \sum_{j=1}^N \sum_{\mathcal{F}_{[j,N+1]}^{1 \rightarrow j}} w(\mathcal{F}), \quad (24)$$

so that Eq. (19) reads in terms of these two quantities:

$$\langle \tau \rangle = \frac{Z_{\mathcal{F}}}{Z_{\mathcal{T}}}. \quad (25)$$

Letting the sums in Eqs. (23) and (24) run only over the nonvanishing trees, all involved rates satisfy  $k_{ij} > 0$ , and one can parametrize them as  $k_{ij}(a) = e^{aX_{ij}}$  such that  $k_{ij}(a = 1) = e^{X_{ij}}$  is the rate of interest. Applying this parametrization, one obtains for  $Z_{\mathcal{T}}$ :

$$\begin{aligned} Z_{\mathcal{T}}(a) &= \sum_{\mathcal{T}_{[N+1]}} \prod_{k_{ij} \in E(\mathcal{T})} k_{ij} \\ &= \sum_{\mathcal{T}_{[N+1]}} e^{a \sum_{k_{ij} \in E(\mathcal{T})} X_{ij}} \\ &= \sum_{\mathcal{T}_{[N+1]}} e^{aX_{\mathcal{T}}}, \end{aligned} \quad (26)$$

defining  $X_{\mathcal{T}} := \sum_{k_{ij} \in E(\mathcal{T})} X_{ij}$ , which is the sum over all  $X_{ij}$  of a given tree. Now consider the Taylor series of  $\ln Z_{\mathcal{T}}$  around

$a = 0$ :

$$\ln Z_{\mathcal{T}}(a) = \sum_{n=0}^{\infty} \frac{\partial^n \ln Z_{\mathcal{T}}}{\partial^n a} \Big|_{a=0} \frac{a^n}{n!}. \quad (27)$$

The zeroth-order coefficient is

$$\ln Z_{\mathcal{T}}(a=0) = \ln \left( \sum_{\mathcal{T} \in \mathcal{T}_{[N+1]}} 1 \right) = \ln |\mathcal{T}|, \quad (28)$$

i.e., it is the logarithm of the total number of spanning trees. For the first order, one finds:

$$\frac{\partial \ln Z_{\mathcal{T}}}{\partial a} \Big|_{a=0} = \frac{\sum_{\mathcal{T} \in \mathcal{T}_{[N+1]}} X_{\mathcal{T}}}{|\mathcal{T}|} = \langle X_{\mathcal{T}} \rangle_{\mathcal{T}}, \quad (29)$$

i.e., the mean of  $X_{\mathcal{T}}$  over the trees. For the second order, one obtains:

$$\begin{aligned} \frac{\partial^2 \ln Z_{\mathcal{T}}}{\partial a^2} \Big|_{a=0} &= \frac{\sum_{\mathcal{T} \in \mathcal{T}_{[N+1]}} X_{\mathcal{T}}^2}{|\mathcal{T}|} - \left( \frac{\sum_{\mathcal{T} \in \mathcal{T}_{[N+1]}} X_{\mathcal{T}}}{|\mathcal{T}|} \right)^2 \\ &= \langle X_{\mathcal{T}}^2 \rangle_{\mathcal{T}} - \langle X_{\mathcal{T}} \rangle_{\mathcal{T}}^2 \\ &= \sigma_{X_{\mathcal{T}}}^2, \end{aligned} \quad (30)$$

i.e., the variance of  $X_{\mathcal{T}}$  over the trees. These results are no coincidence: Eq. (26) is precisely the cumulant generating function for  $X_{\mathcal{T}}$  such that one finds in general:

$$\frac{\partial^n \ln Z_{\mathcal{T}}}{\partial a^n} \Big|_{a=0} = \kappa_n^{X_{\mathcal{T}}}. \quad (31)$$

Hence, one can express  $\ln Z_{\mathcal{T}}$  as:

$$\ln Z_{\mathcal{T}}(a) = \sum_{n=0}^{\infty} \kappa_n^{X_{\mathcal{T}}} \frac{a^n}{n!}, \quad (32)$$

with the convention that  $\kappa_0^{X_{\mathcal{T}}} := \ln |\mathcal{T}|$ . Completely analogously for  $\ln Z_{\mathcal{F}}$  one gets:

$$\ln Z_{\mathcal{F}}(a) = \sum_{n=0}^{\infty} \kappa_n^{X_{\mathcal{F}}} \frac{a^n}{n!}, \quad (33)$$

with  $\kappa_0^{X_{\mathcal{F}}} := \ln |\mathcal{F}|$ . For the rates of interest ( $a = 1$ ), the logarithm of Eq. (25) becomes:

$$\begin{aligned} \ln \langle \tau \rangle &= \ln Z_{\mathcal{F}}(a=1) - \ln Z_{\mathcal{T}}(a=1) \\ &= \sum_{n=0}^{\infty} \frac{\kappa_n^{X_{\mathcal{F}}} - \kappa_n^{X_{\mathcal{T}}}}{n!}. \end{aligned} \quad (34)$$

Note that Eq. (34) is completely general and, as such, could be applied, for example, to the MAPK system shown in Fig. 1(b). The cumulants are taken over the spanning trees and the expression is precise if all  $k_{ij}$  are known for any  $N$ . Also, our treatment so far is not restricted to biochemical networks. In principle, it can also be applied to other settings described by a master equation.

### E. Limit of large system size

We now are in a position to study the limit of large networks ( $N \gg 1$ ). Because in this case not all details can be known, at the same time we switch to a statistical description

for the reaction rates  $k_{ij}$ . To describe the effect of temperature, we consider biochemical networks with rates that have an Arrhenius-like temperature dependence, that is to say,  $X_{ij} = -\Delta\beta E_{ij} + \ln k_{ij}^0$ , where  $E_{ij}$  and  $\ln k_{ij}^0$  are assumed to be independent random variables. For that case, using that

$$\begin{aligned} \kappa_n^{X_{\mathcal{T}}} &= \kappa_n^{-\Delta\beta E_{\mathcal{T}} + \sum_{k_{ij} \in E(\mathcal{T})} \ln k_{ij}^0} \\ &= (-\Delta\beta)^n \kappa_n^{E_{\mathcal{T}}} + \kappa_n^{\sum_{k_{ij} \in E(\mathcal{T})} \ln k_{ij}^0}, \end{aligned} \quad (35)$$

where  $E_{\mathcal{T}} := \sum_{k_{ij} \in E(\mathcal{T})} E_{ij}$ , one obtains:

$$\ln \langle \tau \rangle = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} (\kappa_n^{E_{\mathcal{F}}} - \kappa_n^{E_{\mathcal{T}}}) \Delta\beta^n + \text{const.} \quad (36)$$

Note that backward rates do not fully enter in the spanning trees (cf. examples in Appendix A), so the activation energies on the trees and on the two-tree forests are different in general. The result so far requires to know the cumulants  $\kappa_n^{E_{\mathcal{T}}}$  and  $\kappa_n^{E_{\mathcal{F}}}$  to get the global rate constant.

If now  $E_{\mathcal{T}}$  and  $E_{\mathcal{F}}$  can be sufficiently well approximated by normal distributions,  $E_{\mathcal{T}} \sim \mathcal{N}(\langle E \rangle_{\mathcal{T}}, \sigma_{\mathcal{T}}^2)$  and  $E_{\mathcal{F}} \sim \mathcal{N}(\langle E \rangle_{\mathcal{F}}, \sigma_{\mathcal{F}}^2)$ , then all but the first two cumulants vanish, resulting in the following expression:

$$\ln \langle \tau \rangle = (\langle E \rangle_{\mathcal{T}} - \langle E \rangle_{\mathcal{F}}) \Delta\beta + \frac{\sigma_{\mathcal{F}}^2 - \sigma_{\mathcal{T}}^2}{2} \Delta\beta^2 + \text{const.} \quad (37)$$

yielding a quadratic dependence in the Arrhenius plot (since  $\ln \langle \tau \rangle = -\ln k$ ).

$E_{\mathcal{T}} = \sum_{k_{ij} \in E(\mathcal{T})} E_{ij}$  and  $E_{\mathcal{F}} = \sum_{k_{ij} \in E(\mathcal{F})} E_{ij}$  are the sums of  $N$  and  $N-1$  independent random variables, respectively. Since  $N \gg 1$ , the central limit theorem suggests that this is a valid assumption as long as the higher cumulants vanish fast enough compared to the differences of the first two terms, requiring sufficiently well-behaved convergence to the normal distribution. The simplest case is that the  $E_{ij}$  already follow a normal distribution. However, any distribution, including the uniform one, is sufficient, as long as it satisfies the conditions required for the central limit theorem. Then one obtains:

$$\begin{aligned} \ln \langle \tau \rangle &= (N \langle E_{ij} \rangle_{\mathcal{T}} - (N-1) \langle E_{ij} \rangle_{\mathcal{F}}) \Delta\beta \\ &\quad + \frac{(N-1) \sigma_{\mathcal{F}, E_{ij}}^2 - N \sigma_{\mathcal{T}, E_{ij}}^2}{2} \Delta\beta^2 + \text{const.} \end{aligned} \quad (38)$$

This constitutes our main theory result: the Arrhenius plot for large and complex biochemical networks is expected to show a quadratic dependence on the inverse temperature.

### F. Comparison to simulated networks

To illustrate the convergence predicted by our theory, random networks on  $N_v = 10$  to  $N_v = 1000$  vertices displaying some basic characteristics of the biochemical network of a developmental process were generated (see Appendix C for details) [83]. In particular, they were designed to have a relatively large distance between the initial vertex 0 and the final vertex  $N_v = N+1$  as well as a blocklike structures with many connections and irreversible transitions between them, reflecting the checkpoints present in development. Two small networks with  $N_v = 30$  and  $N_v = 50$  are shown as examples in Figs. 5(a) and 5(b), respectively.

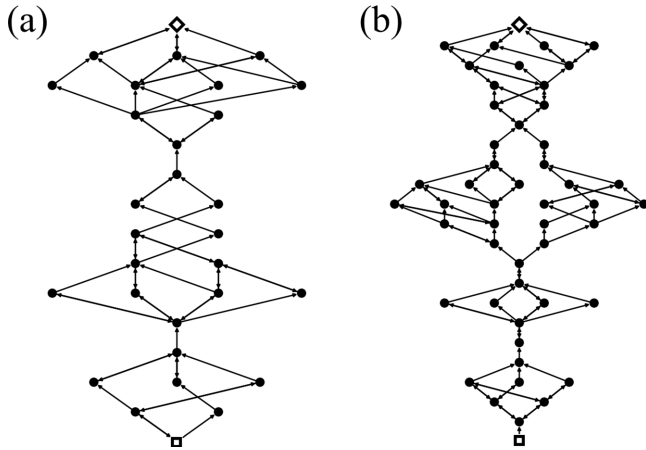


FIG. 5. Randomly generated networks with features that are typical for development of sizes (a)  $N_v = 30$  and (b)  $N_v = 50$  vertices. The initial vertices are represented as squares and the final vertices as diamonds. The networks are biased towards the final vertices and contain compact elements with many connections.

The spanning trees and two-tree spanning forests of these base networks were then drawn at random. A standard procedure to generate random spanning trees is given by Wilson's algorithm [84]. While removing a random edge from a spanning tree will result in a spanning forest, this procedure does not yield all spanning forests with the same probability. In order to ensure an unbiased sample, a specific algorithm was devised for the generation of two-tree spanning forests. In Appendix D, the steps needed for this algorithm are explained and its functionality is illustrated for a simple example network.

The activation energies were assigned randomly to the edges. The activation energies were drawn at random from the intervals  $[60 \frac{\text{kJ}}{\text{mol}}, 100 \frac{\text{kJ}}{\text{mol}}]$  for the forward rates and from  $[20 \frac{\text{kJ}}{\text{mol}}, 60 \frac{\text{kJ}}{\text{mol}}]$  for the backward rates. These represent typical values of activation energies of biochemical reactions listed in the literature [85].

In each case, 10000 spanning trees and at least 10000 two-tree spanning forests (note that the number of generated forests is not predefined to ensure an unbiased sample, as explained in Appendix D) and the total activation energies for each of them were calculated. The total activation energies for  $N_v = 10$  are shown in Fig. 6 as histograms. One observes that there are just a few spanning trees and also a moderate number of two-tree spanning forests, so that the distribution is not reminiscent of a Gaussian yet. The other histograms, along with fits to normal distributions, are shown in Fig. 7. As expected, the histograms of the total activation energies approach a normal distribution as  $N_v$  increases.

Then, Eq. (37) can be applied to obtain the Arrhenius plot for the global rate. The corresponding plots for networks of size 20, 100, and 1000 are shown in Fig. 7. The convergence of the total energy distributions to a Gaussian is observed as the system size  $N_v$  increases. For the parameters chosen here, the distribution appears already well described by a Gaussian for  $N_v \sim 100$ . The second derivative, i.e., twice the quadratic term as a function of  $N_v$  is shown in Fig. 8. One observes that as the system size increases, so does the quadratic term, in a

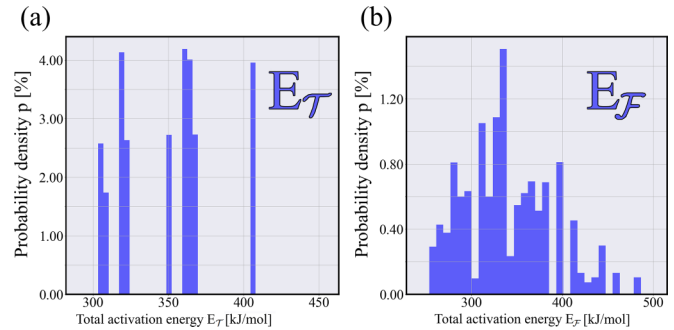


FIG. 6. Histograms of the total activation energies of (a) spanning trees  $E_T$  and (b) two-tree spanning forests  $E_F$ , i.e., the sum of all the activation energies along their edges, on randomly generated graphs on  $N_v = 10$  vertices.

statistical manner. This proves numerically that this kind of networks generically lead to a quadratic Arrhenius plot in the limit of large system size.

### G. Comparison to experimental data

Our theoretical prediction is in striking agreement with experimental data for developmental systems. In a work published in 2021, Crapse *et al.* conducted a study on the development rates of *Drosophila* embryos at different temperatures [24]. To that end, the embryonic phase of *Drosophila* was subdivided into 12 stages and the rates to reach each stage were measured under different temperature conditions. The 14th nuclear division in the zygote was chosen as the starting point (stage A) and the first breath was chosen as the end point (stage K). The cumulative rates starting with the 14th nuclear division to four different stages (B,E,H,K in the notation of the original authors) are shown in Fig. 9 as an example, together with a quadratic fit performed in python via the `numpy.polyfit` command. The data were taken from the supplementary material of the original publication, which tabulated the measured times for the entire sample and all stages. The fitting parameters are listed in Appendix B. Note that these do not provide insights into the microscopic structure of the network without further assumptions, since the coefficients of the quadratic functions contain the differences of the first two moments along the trees and forests as well as terms containing  $\frac{1}{k_B T_0}$ . Moreover, the residuals of the quadratic fit were computed to analyze whether there are any systematic deviations that hint towards a deviation from the quadratic dependence.

For the beginning of stage B (begin of cellularization), no robust statement can be made about the goodness of the quadratic fit. Since this is the first stage following the starting point, it may well be that the assumption of a sufficiently large underlying biochemical network is simply not valid. However, as more steps accumulate, such as in the rates to stages E, H, and K, the quadratic shape of the data points becomes more apparent. The residuals do not point to a systematic deviation, suggesting that higher-order terms at best play a minor role. A quadratic fit of the data was also suggested by Crapse *et al.* [24], although without providing an explanation on fundamental grounds, as done here.

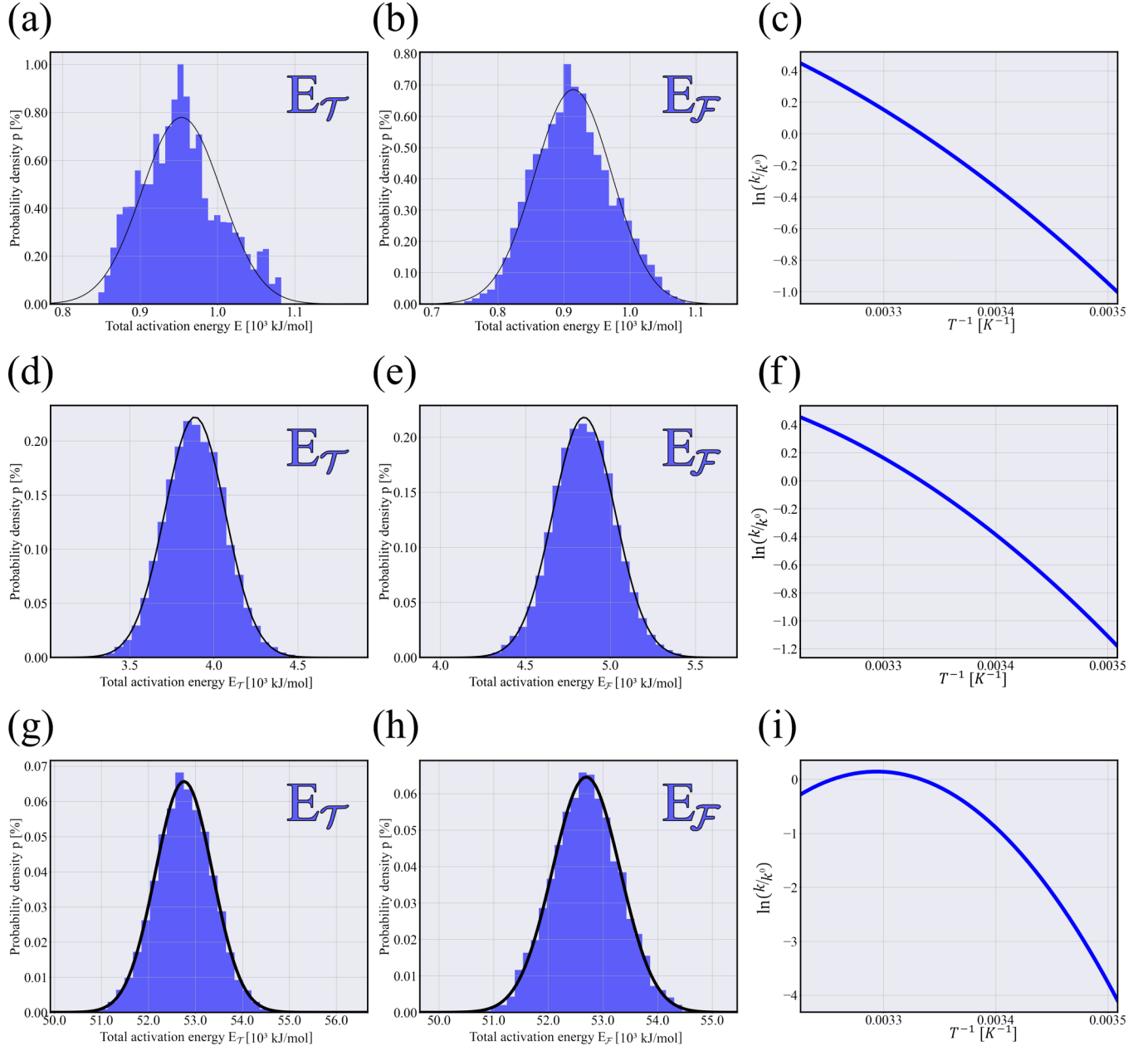


FIG. 7. Histograms and Arrhenius plots for simulated random networks of increasing sizes. (a)–(c)  $N_v = 20$ . (d)–(f)  $N_v = 100$ . (g)–(i)  $N_v = 1000$ . (a), (d), (g) Spanning trees. (b), (e), (h) Spanning forests. (c), (f), (i) Resulting Arrhenius plots. The fitted parameters are listed in Appendix B.

Measurements of the temperature dependence of developmental rates of *Drosophila* were already performed by Bliss [25] in 1926 for the prepupal development and by Powsner [26] in 1935 for the embryonic, pupal, and larval stages. Their results are shown in Figs. 10 and 11, respectively, also with a quadratic fit to their data points. Both articles list their data in a table and show them in a diagram, albeit together with piecewise linear functions as fit instead of a global quadratic function. One sees that in all cases, the quadratic fits predicted by our theory are in excellent agreement with the experimental data. Also the residuals indicate only a minor contributions from terms of higher order, except for the larval development rates from the data by Powsner [Figs. 11(b) and 11(c)], where most likely some additional effect contributes.

Figure 12 shows the quadratic terms of the quadratic fit for the rates to the different stages defined by Crapse *et al.* [24]. Consistent with the simulation data shown in Fig. 8, the quadratic term tends to increase as more steps accumulate. However, a slight decrease is observed for the rates to the final stages. This is most likely due to the lack of data points at the extreme ends of the temperature spectrum, as most embryos did not survive long enough under these conditions. Consequently, this also results in greater uncertainties for these values.

Note that the maximal rate, i.e., the peak in the diagrams, is at  $T_{\text{peak}} \sim 30^\circ\text{C}$ . Expecting the mean total activation energies to be higher for the trees than for the forests, and vice-versa for the variances, Eq. (37) predicts the maximum to be attained at



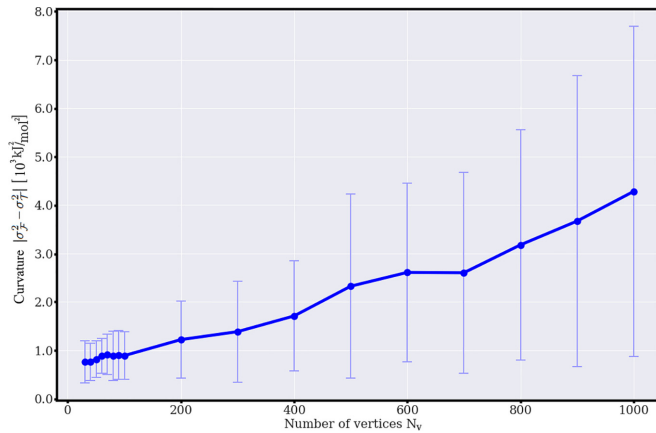


FIG. 8. The absolute value of the quadratic term in the Arrhenius plot  $\sigma_F^2 - \sigma_T^2$  versus number of vertices  $N_v$ . For every value  $N_v$ , 100 networks of this size were generated and the distribution of the total activation energies of the trees and two tree spanning forests were computed. The error bars indicate the standard deviations.

$\Delta\beta < 0$ , so  $T_{\text{peak}} > T_0$ , and one would in fact assume  $T_0 \sim 20^\circ\text{C}$ .

## H. Deviations from the generic quadratic response

Both our theory and computer simulations demonstrate that large networks show a generic Arrhenius plot with a quadratic form. We now discuss deviations from this generic response that might arise because of our mathematical assumptions. We start with networks that have correlations between their activation energies and show that the generic response persists as long as such correlations are sufficiently dilute in large networks. We then discuss small networks and demonstrate that they easily can be designed to give nonlinear Arrhenius plots. We finally discuss the case of linear chains, which shows that even large networks can deviate, because they contain only few spanning trees.

### 1. Correlations

Our theory assumes that activation energies on different edges are independent and that their summations over spanning trees and forests satisfy the conditions of the central limit theorem. In real biochemical networks, the master equation description contains repeated reaction channels due to multiplicity, so the corresponding activation energies are in fact correlated. As an instructive example, consider an autorepressive production of species  $X$ :

$$\dot{X} = \frac{k}{1 + \left(\frac{X}{C}\right)^n}, \quad (39)$$

where  $k$  is the production rate in the absence of  $X$ ,  $C$  is a reference concentration, and  $n$  is the Hill coefficient. In a chemical master equation, the stochastic variable is the copy number  $m$  of  $X$  (for multiple species, it is the product space of all molecule counts). This yields a linear chain of reaction steps with rates  $k_m = \frac{k}{1 + (\frac{m}{c})^n}$ , with  $c$  being an appropriately rescaled constant (see Fig. 13).

Since the Arrhenius equation gives  $k \sim e^{-\Delta\beta E_A}$ , all these edges inherit the same activation energy  $E_A$ , because they are different instances of the same microscopic reaction. A more precise bookkeeping of total activation energy should therefore treat the energy of a spanning tree or forest as a weighted sum of distinct reaction types:

$$E_{\mathcal{T}} = \sum_{k_{ij}} n_{ij} E_{ij}, \quad (40)$$

where  $n_{ij} \in \mathbb{N}$  counts how many times that same reaction appears in the tree or forest. Although the state space of such master equations can be infinite in principle, in real biological systems the number of copies is effectively bounded due to resource limitations and hence it can be safely assumed that the first two moments of  $n_{ij}$  stay finite. Note that the multiplicities  $n_{ij}$  can themselves be treated as independent random variables and consequently  $n_{ij} E_{ij}$  are independent random variables again. As long as there are still sufficiently many distinct, effectively independent reactions, Eq. (40) is still the sum of many independent random variables and the central limit theorem retains its validity. This argument also applies to the possibility that the same network motif appears several times in the network, as known for example from the feed-forward chain appearing in the developmental network of *Caenorhabditis elegans* [86]. Again this case could be remedied by introducing multiplicities  $n_{ij}$ . Only if the network was highly correlated throughout, would our generic prediction break down.

### 2. Small networks

As an example for a small network, consider the  $N_v = 5$  network depicted in Fig. 14(a). This network has two spanning trees rooted at vertex 4 [see Fig. 14(b)]:

$$Z_{\mathcal{T}} = (k_1 + k_2)k_3k_4k_5, \quad (41)$$

and four two-tree spanning forests, where one tree is rooted at vertex 4 and the other contains vertex 0 [see Fig. 14(c)]:

$$Z_{\mathcal{F}} = k_1k_4(k_3 + k_5) + k_2k_3(k_4 + k_5). \quad (42)$$

Thus, the MFPT from 0–4 reads:

$$\langle \tau \rangle = \frac{Z_{\mathcal{F}}}{Z_{\mathcal{T}}} = \frac{1}{k_1 + k_2} \left( \frac{k_1}{k_3} + \frac{k_2}{k_4} \right) + \frac{1}{k_5}, \quad (43)$$

which can give rise to the dependence in the Arrhenius plot shown in Fig. 14(d) for realistic values for the rate constants and activation energies.

The resulting temperature response no longer resembles the linear slope of the Arrhenius equation; instead, it is strongly curved, reaching a minimum around  $T = 35^\circ\text{C}$  and a maximum around  $T = 18^\circ\text{C}$ . This implies that the global rate would actually decrease with increasing temperature within the temperature range typically relevant to biological systems. This example illustrates that the temperature dependence in medium-sized networks depends on microscopic details and may generally be more complex.

An interesting case are networks with a single rate-limiting step, i.e., a single reaction that proceeds significantly slower than the rest of the reaction network. In this case, one might

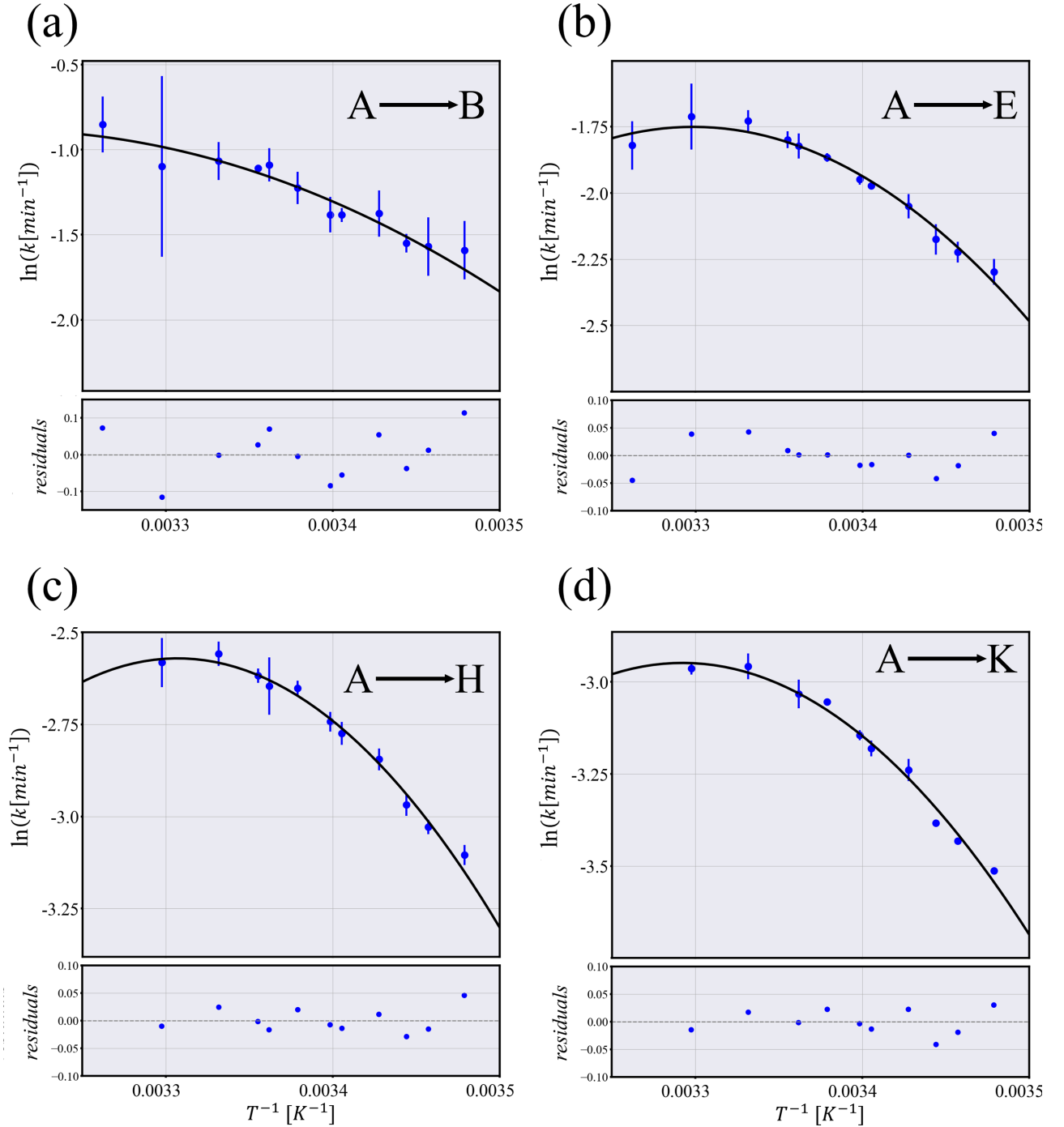


FIG. 9. Arrhenius plot for the rates to the different stages in the development of *Drosophila melanogaster* embryos, starting at the 14th nuclear division (A) adapted from [24] with quadratic fit to data. The residuals to the fit are shown below the main diagram. The labeling of the stages follows the convention from the original authors: (a) rate from A to B (beginning of cellularization),  $R^2 = 0.9587$ ; (b) rate from A to E (development of horizontal posterior midgut),  $R^2 = 0.9909$ ; (c) rate from A to H (full germ band retraction),  $R^2 = 0.9952$ ; (d) rate from A to K (first breath),  $R^2 = 0.9868$ .

naively expect that the Arrhenius equation for this reaction dominates and that the global response becomes linear again. However, this is not necessarily the case. Consider the minimal example shown in Fig. 15. Observe that:

$$\langle \tau \rangle = \frac{Z_{\mathcal{F}}}{Z_{\mathcal{T}}} = \frac{k + r + \gamma}{k\gamma} = \left(1 + \frac{r}{k}\right) \frac{1}{\gamma} + \frac{1}{k}. \quad (44)$$

If the step from 1–2 is rate limiting, i.e.,  $\gamma \ll k, r$ , then:

$$\langle \tau \rangle \approx \left(1 + \frac{r}{k}\right) \frac{1}{\gamma}, \quad (45)$$

not just  $\frac{1}{\gamma}$  as one could naively expect. So, the ratio  $\frac{r}{k} \sim e^{\beta \Delta E}$  also contributes even if  $\gamma$  is many orders of magnitude below  $k$  and  $r$ . The reason for this is that temperature also impacts

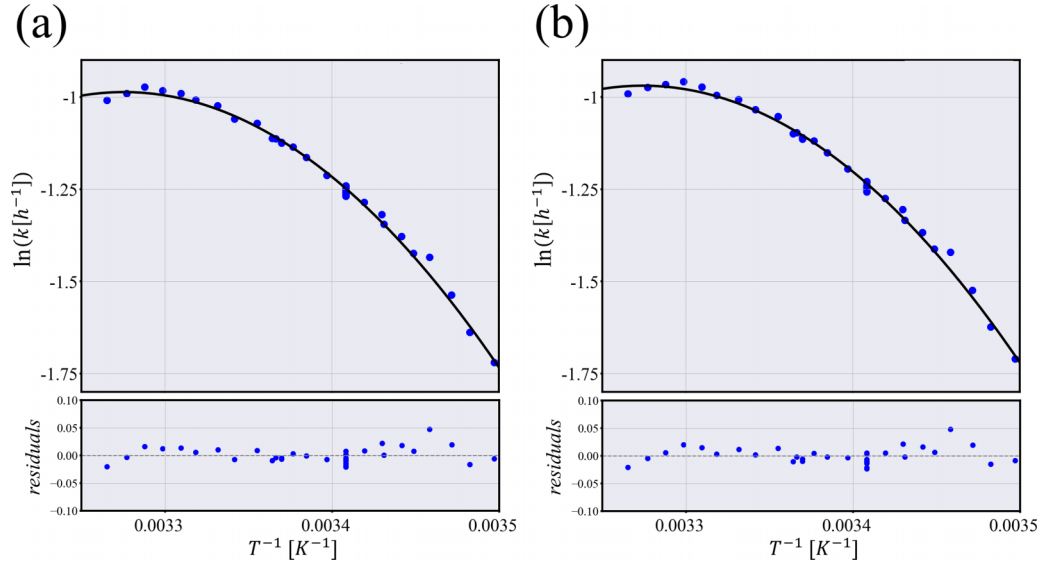


FIG. 10. Logarithmic developmental rate to the pupal stage of *Drosophila melanogaster* adapted from [25] with quadratic fit. The residuals to the fit are shown below the main diagram. (a) For male specimen,  $R^2 = 0.9950$ , (b) for female specimen,  $R^2 = 0.9950$ .

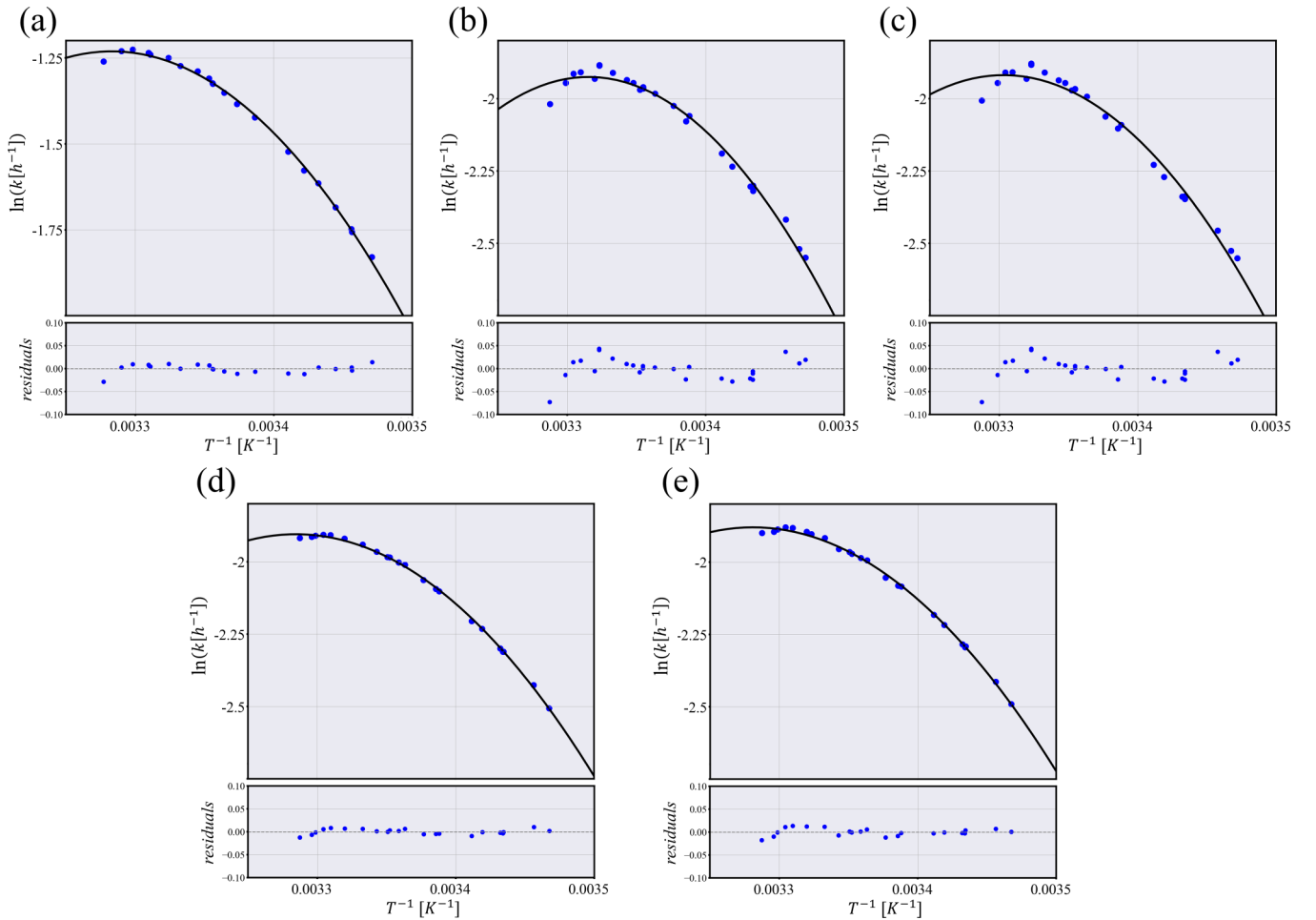


FIG. 11. Logarithmic developmental rate for the embryonic, pupal, and larval stage of *Drosophila melanogaster* [26] versus inverse temperature with quadratic fit. The residuals to the fit are shown below the main diagram. (a) Embryonic development,  $R^2 = 0.9976$ ; (b) male larval development,  $R^2 = 0.9838$ ; (c) female larval development,  $R^2 = 0.9808$ ; (d) male pupal development,  $R^2 = 0.9990$ ; (e) female pupal development,  $R^2 = 0.9981$ .

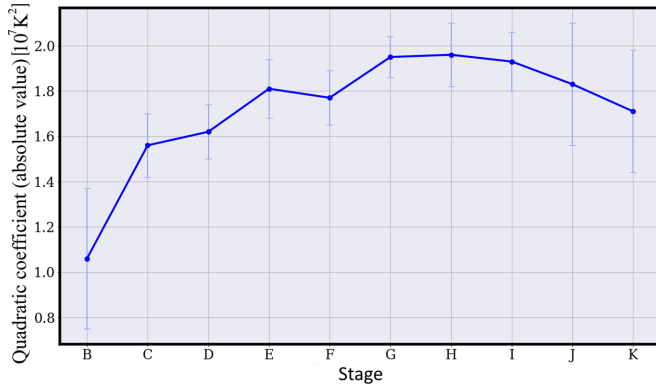


FIG. 12. The absolute value of the quadratic coefficient from the quadratic fit to the data from [24] with the last cleavage as starting point to later stages in development.

the equilibrium distribution. If  $\gamma$  is the rate-limiting step, one can assume that the states 0 and 1 are effectively in their equilibrium configuration, meaning that the flow from 0–1 must be equal to the flow from 1–0. Denoting the probability to be in state 1 as  $p$ , this means:

$$k(1 - p) = rp \Rightarrow p = \frac{k}{k + r}. \quad (46)$$

The crucial point is that the system can only jump to 2 via 1, and the rate must therefore be multiplied by the probability  $p$  to be in 1. Hence, one gets:

$$\langle \tau \rangle = (p\gamma)^{-1} = \left(1 + \frac{r}{k}\right) \frac{1}{\gamma}, \quad (47)$$

which is the same result as before. One can therefore conclude that separation of time scales is not sufficient to obtain a pure Arrhenius dependence.

### 3. Linear chain

The generic quadratic behavior predicted here can be violated not only for small networks, but also for large networks, which contain only few spanning trees. The simplest possible example is a linear chain of irreversible reactions, i.e.,  $k_{i,j} = k_i \delta_{i+1,j}$ , as illustrated in Fig. 16.

Then the corresponding graph is already a tree and thus its own unique spanning tree:

$$\sum_{\mathcal{T}_{[N+1]}} w(\mathcal{T}) = \prod_{i=1}^N k_i. \quad (48)$$

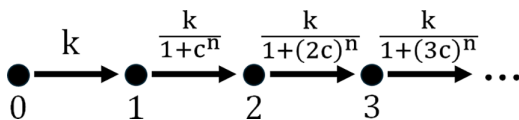


FIG. 13. Describing the autorepressive reaction in Eq. (39) as a master equation results in a linear chain of reactions where the rates all scale with the same reaction rate  $k$ .

Also, for the spanning forests it is easy to see that there is either exactly one or no combination:

$$\sum_{\mathcal{F}_{[j,N+1]}^{i \rightarrow j}} w(\mathcal{F}) = \frac{1}{k_j} \prod_{i=1}^N k_i \delta_{i \leq j}. \quad (49)$$

Then,

$$(-K^T)^{-1} = \begin{pmatrix} k_1^{-1} & k_2^{-1} & \dots & k_N^{-1} \\ 0 & k_2^{-1} & \dots & k_N^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & k_N^{-1} \end{pmatrix}. \quad (50)$$

One finds for the first two moments:

$$\langle \tau_i \rangle = \sum_{j=i}^N \frac{1}{k_j}, \quad (51)$$

$$\langle \tau_i^2 \rangle = 2 \sum_{j=1}^N \frac{1}{k_j} \sum_{l=j}^N \frac{1}{k_l} = \left( \sum_{j=1}^N \frac{1}{k_j} \right)^2 + \sum_{j=1}^N \frac{1}{k_j^2}, \quad (52)$$

which yields the standard deviation:

$$\sigma_{\tau_1} = \left( \sum_{j=1}^N \frac{1}{k_j^2} \right)^{\frac{1}{2}} \quad (53)$$

and the coefficient of variation (CV) reads:

$$\text{CV}_{\tau_1} = \frac{\sigma_{\tau_1}}{\langle \tau_1 \rangle} = \left( \frac{\sum_{j=1}^N \frac{1}{k_j^2}}{(\sum_{j=1}^N \frac{1}{k_j})^2} \right)^{\frac{1}{2}}, \quad (54)$$

meaning that  $\text{CV}_{\tau_1} \rightarrow 0$  for  $N \rightarrow \infty$  if most of the  $k_i$  are of comparable sizes [87], so the relative variation vanishes in the limit of large chains.

If  $k_i = k \forall i$ , one finds in general:

$$\langle \tau^n \rangle_i = \frac{1}{k^n} \frac{(N + n - i)!}{(N - i)!}, \quad (55)$$

which are precisely the moments for an  $(N - i + 1)$ -dimensional Erlang distribution ( $\Gamma$  distribution with non-negative integer as shape parameter) with  $\lambda = \frac{1}{k}$ :

$$f^{\tau_i}(t) = \frac{k^{N-i+1}}{(N - i + 1)!} t^{N-i} e^{-kt}, \quad (56)$$

as it should be since this is precisely the distribution that arises from summing up  $N - i$  iid exponentially distributed random variables.

Note that the coefficient of variation (CV) reads:

$$\text{CV}_{\tau_1} = \frac{\sigma_{\tau_1}}{\langle \tau_1 \rangle} = \frac{1}{\sqrt{N}} \xrightarrow{N \rightarrow \infty} 0, \quad (57)$$

so the relative deviation from the mean becomes increasingly small as the system size increases.

Moreover, the mean first-passage time for the whole process,  $\langle \tau \rangle = \langle \tau_0 \rangle$ , given in Eq. (51) may be written in the limit of large  $N$  as:

$$\langle \tau \rangle = N \langle k_j^{-1} \rangle = N \langle k_j^{0-1} e^{\Delta \beta E_j} \rangle, \quad (58)$$

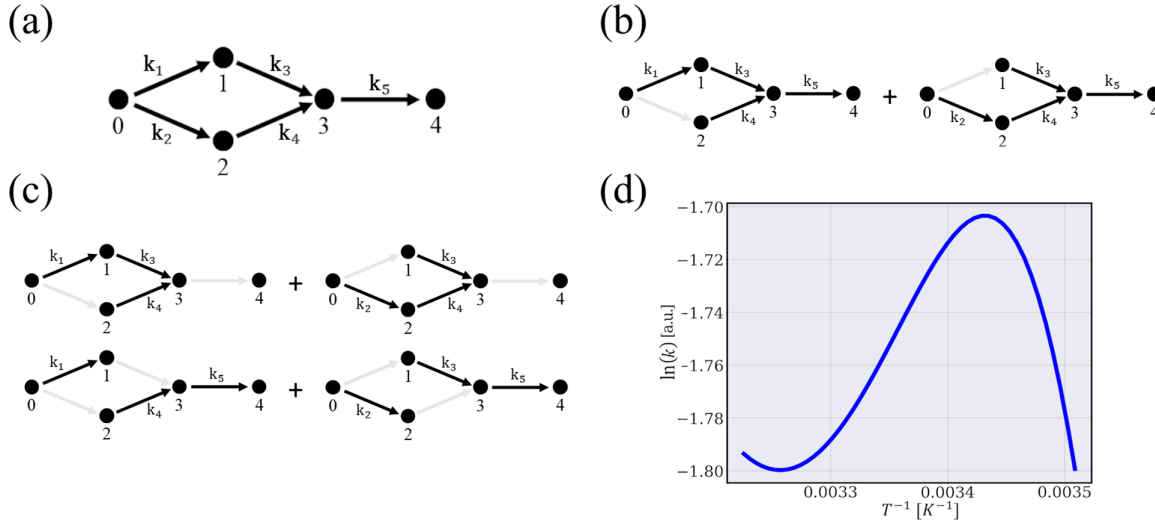


FIG. 14. Example network on five vertices with a strong deviation from an Arrhenius-like response. (a) Network, (b) spanning trees, (c) spanning forests, (d) Arrhenius plot for the ratio  $(k_1 : k_2 : k_3 : k_4 : k_5) = (2 : 2 : 5 : 0.1 : 2)$  at  $T_0 = 300$  K and  $E_1 = 20 \frac{\text{kJ}}{\text{mol}}$ ,  $E_2 = 100 \frac{\text{kJ}}{\text{mol}}$ ,  $E_3 = 30 \frac{\text{kJ}}{\text{mol}}$ ,  $E_4 = 20 \frac{\text{kJ}}{\text{mol}}$ , and  $E_5 = 80 \frac{\text{kJ}}{\text{mol}}$ .

writing  $k_j$  as in Eq. (22). Assuming  $k_j^0$  and  $E_j$  to be independent, one finds:

$$\ln \langle \tau \rangle = \ln \langle e^{\Delta \beta E_j} \rangle + \ln N + \langle k_j^{-1} \rangle \quad (59)$$

$$= \ln \langle e^{\Delta \beta E_j} \rangle + \text{const.} \quad (60)$$

Compare this to Eq. (36): for the linear chain, there was only a single spanning tree, so  $\kappa_1^{E_T} = E_T$  and  $\kappa_n^{E_T} = 0$  for  $n \geq 0$ . The total activation energies of the forests  $\mathcal{F}_i$  for  $i = 1, \dots, N$  read:

$$E_{\mathcal{F}_i} = \sum_{\substack{j=1 \\ j \neq i}}^N E_j = E_T - E_i. \quad (61)$$

One then obtains for the cumulant-generating function of  $E_{\mathcal{F}}$ :

$$\sum \frac{\Delta \beta^n}{n!} \kappa_n^{E_{\mathcal{F}}} = \ln \langle e^{\Delta \beta E_{\mathcal{F}}} \rangle \quad (62)$$

$$= \ln e^{\Delta \beta E_T} + \ln \langle e^{-\Delta \beta E_i} \rangle \quad (63)$$

$$= \Delta \beta E_T + \sum_{n=1}^{\infty} \frac{\Delta \beta^n}{n!} \kappa_n^{E_i}. \quad (64)$$

Inserting this into Eq. (36) yields:

$$\ln \langle \tau \rangle = \Delta \beta E_T + \sum_{n=1}^{\infty} \frac{\Delta \beta^n}{n!} \kappa_n^{E_i} - \Delta \beta E_T + \text{const.} \quad (65)$$

$$= \ln \langle e^{\Delta \beta E_j} \rangle + \text{const.}, \quad (66)$$

so indeed the same result as Eq. (60).

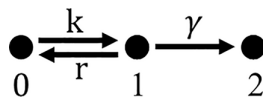


FIG. 15. A minimal example to show that a network with a rate-limiting step does not necessarily yield a global Arrhenius equation dominated by this step.

The total activation energy of the forests only differs from the total activation energy of the trees by a single  $E_i$  and only the cumulant generating function of  $E_j$  but not a sum over many activation energies remains. Therefore, there is no longer a statistical limit  $N \rightarrow \infty$  for the distribution and the expression will depend on the distribution of  $E_i$ . If  $E_i$  is itself normal distributed, the quadratic dependence in Eq. (37) is recovered. But choosing a uniform distribution  $E_i \sim U([E_{\min}, E_{\max}])$ , one obtains:

$$\langle e^{\Delta \beta E_j} \rangle = \frac{1}{\Delta E} \int_{E_{\min}}^{E_{\max}} dE_j e^{\Delta \beta E_j} \quad (67)$$

$$= \frac{1}{\Delta E \Delta \beta} (e^{\Delta \beta E_{\max}} - e^{\Delta \beta E_{\min}}), \quad (68)$$

with  $\Delta E := E_{\max} - E_{\min}$ . Then Eq. (60) reads:

$$\begin{aligned} \ln \langle \tau \rangle &= -\ln \Delta E \Delta \beta + \Delta \beta E_{\min} + \ln (e^{\Delta \beta \Delta E} - 1) \\ &= \Delta \beta \frac{E_{\max} + E_{\min}}{2} + \sum_{n=2}^{\infty} \frac{\Delta E^n}{n!} \frac{B_n}{n} \Delta \beta^n, \end{aligned} \quad (69)$$

where  $B_n$  is the  $n$ th Bernoulli number. This corresponds to the simulations performed by Crapse *et al.* [24], involving a

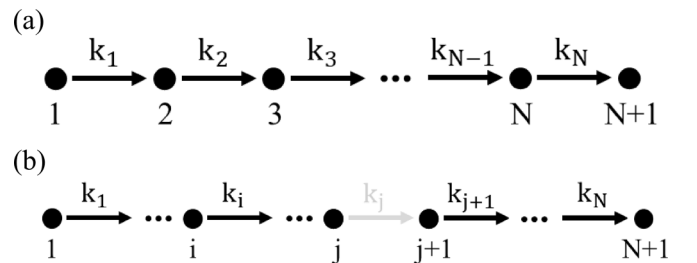


FIG. 16. (a) A linear chain of  $N$  reactions with reaction rates  $k_{i,j} = k_i \delta_{i+1,j}$ . (b) Spanning forests of the linear chain. For  $i \leq j$ , the only corresponding spanning forest is the one without the  $k_j$  edge. For  $i > j$ , there is no such tree.



chain of 1000 reactions with activation energies drawn from a uniform distribution. The authors found that the results of their simulation could not reproduce the quadratic shape that they found describing their experimental data and therefore discarded the possibility that it was a result of the complexity of the network. In light of the analytical theory developed here, that choice seems unfortunate because the linear chain is too simplistic to allow for the large system limit. Additionally, if they had chosen  $E_i \sim \mathcal{N}(\langle E \rangle, \sigma^2)$ , it would have resulted in a quadratic shape.

### III. DISCUSSION

Here we have shown that the generic temperature response of large biochemical networks is quadratic in the Arrhenius plot if the network has an imbalance of activation energies in forward versus backward directions. This seems reasonable for developmental processes and is supported by evidence from *in vitro* studies with enzymes from *Xenopus* showing a substantial imbalance in the activation energies of the cyclin synthesis and degradation rates [48].

In order to make general statements on such networks, we employed a statistical description, which does not require knowledge of all activation energies. Instead, we only make assumptions about their distribution, which is a reasonable description for a large system size  $N$ , where not all degrees of freedom can be known. Our discussion was based on a general expression of the MFPT for arbitrary networks, Eq. (19), which follows from graph theory [74] and has been derived here in a pedagogical manner. We then rewrote this formula in the language of statistical physics. The coefficients appearing in the Taylor series of the logarithmic MFPT can be interpreted in terms of the cumulants of the distribution of the sums along the spanning trees and two-tree forests. If this sum can be sufficiently well approximated by a Gaussian distribution, this yields a quadratic dependence in the Arrhenius plot, as described by Eq. (37). This is our main result, namely the quadratic response as the natural description for complex biochemical networks. This result need not apply for a network which has similar distributions for the activation energies for forward and backward directions, but this is unlikely for developmental systems.

Being the result of evolution, the assumption that the activation energies can be described as independent from each other needs to be treated with care. We first note that correlations arising from copies numbers in chemical master equations are not a fundamental limitation of our theory, because this effect can be described by multiplicities in the relevant sums. We also note that duplication of the same motif would not matter for the same reason. However, it could well be that two network motifs have been duplicated and then drifted apart during evolution. To date, no experimental evidence seems to be available to determine how strong such correlations might be in real developmental networks. However, since there are many different reactions involved in a process such as the development of an embryo, some of which are purely biochemical (like the phosphorylations occurring in cell cycle control), some of which are more biophysical in nature (like the growth of microtubules), it seems reasonable to assume that most activation energies are indeed independent

of each other, especially when contained in different modules of the overall network. Furthermore, generalizations of the central limit theorem exist that allow for a weak degree of dependence, e.g., correlations decaying with distance in the network (see, for example, Ref. [88]). In the future, it might be interesting to experimentally determine all activation energies in a network of interest and then to predict the resulting temperature dependence using the formalism presented here.

The theory was validated numerically by drawing spanning trees and two-tree spanning forests from a random graph on  $N_v = 1000$  vertices. Individual activation energies were drawn from homogeneous distributions, also in order to avoid using a Gaussian distribution from the start. We then computed the histograms of their total activation energies, which approach a normal distribution as expected from the central limit theorem (Fig. 7). The global rate constant in an Arrhenius plot is then of quadratic shape, and this quadratic dependence becomes more pronounced for larger  $N_v$ , as predicted by the theory (Fig. 7).

Crapse *et al.* [24] suggested that the observed deviation from a pure Arrhenius-equation was due to the biochemical reactions not following an ideal Arrhenius-behavior, rather than a result of the complexity of the system. They reached that conclusion based on the simulation of a linear chain of 1000 reactions described by an Arrhenius-equation, which failed to reproduce the observed quadratic response in the Arrhenius plot. They also conducted an experiment measuring the conversion of NAD<sup>+</sup> to NADH via GAPDH catalysis and found that this reaction does not follow a pure Arrhenius-like dependence. However, closer inspection of their simulations shows that the activation energies were chosen uniformly in an energy range from  $20 \frac{\text{kJ}}{\text{mol}}$ – $100 \frac{\text{kJ}}{\text{mol}}$  and the prefactors independently at  $T_0 = 295$  K. Due to the absence of backward rates and branching in this model, this leads to a different series (see Sec. II H), for which the uniform distribution yields infinitely many nonvanishing terms. If the activation energies were chosen from a normal distribution, this would yield a quadratic dependence also on this level. In contrast, our results are more general and require fewer assumptions. Moreover, the measurement of the temperature dependence of the rate of the NAD<sup>+</sup> to NADH reaction seems to be well described by a linear fit in the Arrhenius-plot up to at least 35 °C. Indeed, for higher temperatures, the denaturation of the involved enzymes significantly decreases the reaction rate. The measurements of the development rates only reached about 33 °C, though. Therefore, this can most likely not explain the quadratic temperature dependence observed over the entire range.

The theory developed so far is of course no longer valid above the melting temperature of the proteins. Denaturation substantially slows the reaction rates, and this is not described by the Arrhenius equation [85,89]. This could explain the curvature of the Arrhenius plots seen at small values of the inverse temperature, in particular for temperatures of about 30 °C and above, where we observe that the high-temperature data points (low inverse temperature) in Figs. 9–11 all lie below the quadratic curve. In the measurement by Powsner on the pupal development rates [Fig. 11(b)], the decrease is most notable, and it is not unlikely that this indicates that the denaturation threshold for a protein that is important for the developmental progression is surpassed at these temperatures.

On the other hand, however, this observation also suggests that the larger part of the curve should not be affected by protein denaturation, and that our theory indeed captures the generic behavior of this system. Future work could extend our theory to include the effects of protein denaturation, for example, by introducing it as an additional state in the network.

It should be mentioned that Powsner already noticed in his work that the combination of a few reactions, each described by an Arrhenius equation, does not necessarily yield a global Arrhenius equation and hypothesized that the observed temperature response may be a consequence of the complexity of the underlying biochemical network. However, he did not give a mathematical description as provided here.

The observation that the development rates can be well described by a quadratic fit in the Arrhenius plot does not seem to be limited to the fly as discussed above. Rombouts *et al.* used it to fit the timing of the early cell cycles of *Caenorhabditis elegans*, *Caenorhabditis briggsae*, *Danio rerio*, *Xenopus tropicalis*, and *Xenopus laevis* [48], and Aquilanti *et al.* found that a quadratic dependence in the Arrhenius plot is a good description for data on the respiration rate of *Camellia Japonica* leaves [90].

Our result is a statement about large biochemical networks. Due to the exponential nature of the Arrhenius equation, it is reasonable to assume that this effect dominates more subtle temperature dependencies, such as the dependence of the diffusion constant on temperature. According to the Einstein relation, the diffusion constant  $D$  is linear in  $T$  and a change from  $T = 300$  K to  $T = 310$  K increases  $D$  by a mere 3%, while chemical rates often double or triple under a change of  $\Delta T = 10$  K [91]. Therefore, it seems justified to compare the theory to data on developmental processes, without claiming to capture every nuance of the effect of temperature on these complex systems. If needed, the effect of temperature on the diffusion constant can be included by decomposing binding rates into diffusion and reaction parts [92].

The discussion in this work was limited to the mean of the first-passage time, although the formalism in principle covers all higher moments, too. The focus on the first moment seems justified by the observation that the coefficient of variation (standard deviation divided by mean) for the development times of *Drosophila* embryos is in the order of a few percent across different temperatures [81]. This results most likely from a strong forward bias for the rates of a developmental network with many steps such as cell cycle checkpoints and cell division being practically irreversible. It is well known that this tends to yield sharp FPT distributions around the mean in the limit of large networks [80] (compare also the example in Sec. IIH).

The framework developed here was motivated by describing the temperature dependence of complex networks where not all degrees of freedom are known. In this work, we focused on development, but as mentioned in Sec. I, there exist other complex biological systems with interesting temperature dependence, in particular in the contexts of fever and climate change. Our results are relatively general statements about the mean first-passage time of master equations and are therefore not *per se* limited to biochemical networks and temperature effects. The algorithm used to obtain the distributions of total activation energies along trees and two-tree forests also pro-

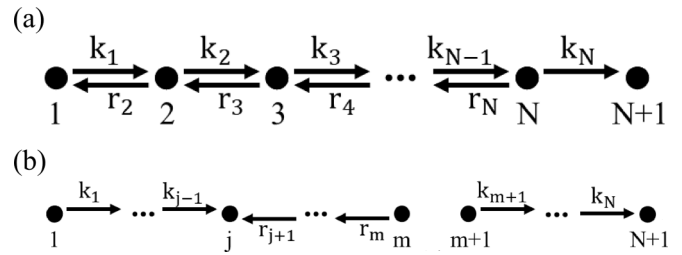


FIG. 17. (a) A one-step master equation from state 1 to state  $N + 1$ , the latter one being absorbing. (b) Counting the two-tree spanning forests rooted at  $N + 1$   $\mathcal{F}_{[j, N+1]}^{1 \rightarrow j}$  for this network.

vides a method for calculating the MFPT for large networks. It may be worthwhile to explore how this compares to obtaining the MFPT from simulations of stochastic dynamics [93], and whether it can also be applied to study other network properties related to graph theory.

The idea to describe large reaction networks from a coarse-grained perspective with appropriate approximations has been discussed elsewhere [94], but neither the mean first-passage time nor the possibility of describing the rate constants on a statistical basis as proposed here seem to be explored so far. First-passage time problems are ubiquitous in biology and biochemistry [95,96] and it is very interesting how far one can get without knowing all the details of the system under consideration.

## ACKNOWLEDGMENTS

J.B.V. thanks the German Academic Scholarship Foundation (Studienstiftung des Deutschen Volkes) for support. We also acknowledge support by the Max Planck School Matter to Life funded by the German Federal Ministry of Education and Research (BMBF) in collaboration with the Max Planck Society.

## DATA AVAILABILITY

The data that support the findings of this article are openly available [83].

## APPENDIX A: OBTAINING THE MFPT BY COUNTING

To illustrate the application of the graph-theoretical approach used here, we derive the MFPTs for two examples, which also can be obtained by applying Laplace transforms [80].

### 1. One-step master equation

Consider the one-step master equation with forward rates  $k_i$  and backward rates  $r_i$  as depicted in Fig. 17. Note that there is only one spanning tree with all flow directed towards  $N_v$ , namely the one consisting only of the forward rates, meaning that:

$$\sum_{\mathcal{T}_{[N+1]}} w(\mathcal{T}) = \prod_{i=1}^N k_i. \quad (\text{A1})$$

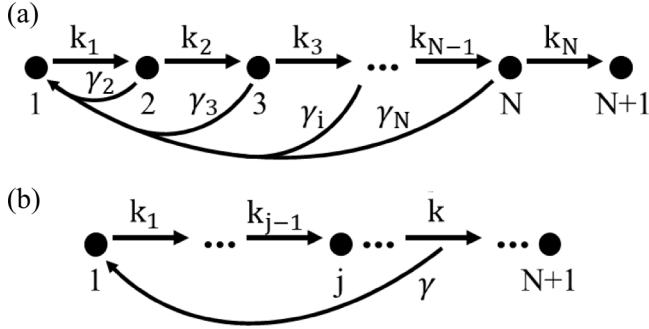


FIG. 18. (a) A simple KPR network on  $N + 1$  states consisting of a chain of forward rates  $k_i$  and reset rates  $\gamma_i$ . (b) Counting the  $\mathcal{F}_{[j,N+1]}^{1 \rightarrow j}$  for this network.

The  $\mathcal{F}_{[j,N+1]}^{1 \rightarrow j}$  are the chains which start off towards the right and change their direction at  $j$  and possibly also at a later vertex  $m$ . This is shown in Fig. 17(b). Therefore, one gets:

$$\sum_{j=1}^N \sum_{\mathcal{F}_{[j,N+1]}^{1 \rightarrow j}} w(\mathcal{F}) = \sum_{m=1}^N \sum_{j=1}^m \prod_{\alpha=1}^{j-1} k_{\alpha} \prod_{\beta=l+1}^m r_{\beta} \prod_{\gamma=m+1}^N k_{\gamma}. \quad (\text{A2})$$

By Eq. (19), the MFPT is thus given by:

$$\langle \tau \rangle = \sum_{i=1}^N \sum_{l=1}^i \frac{1}{k_l} \prod_{m=l+1}^i \frac{r_m}{k_m}. \quad (\text{A3})$$

## 2. Simple kinetic proofreading

One can also derive the MFPT for a simple kinetic proofreading (KPR) scheme shown in Fig. 18 by correct counting. Again, there is just one spanning tree, namely the one containing all  $k_i$ . So one has again:

$$\sum_{\mathcal{T}_{[N+1]}} w(\mathcal{T}) = \prod_{i=1}^N k_i. \quad (\text{A4})$$

The  $\mathcal{F}_{[j,N+1]}^{1 \rightarrow j}$  are the graphs with forward rates up to some vertex  $j$  and either  $k_i$  or  $\gamma_i$  for any  $i > j$ . This is illustrated in Fig. 18(b). Therefore, one gets:

$$\sum_{j=1}^N \sum_{\mathcal{F}_{[j,N+1]}^{1 \rightarrow j}} w(\mathcal{F}) = \sum_{j=1}^N \prod_{i=1}^{j-1} k_i \prod_{m=j+1}^N (k_m + \gamma_m) \quad (\text{A5})$$

and obtains for the MFPT:

$$\langle \tau \rangle = \sum_{j=1}^N \frac{1}{k_j} \prod_{m=j+1}^N \left( 1 + \frac{\gamma_m}{k_m} \right). \quad (\text{A6})$$

For the case that  $k_i = k$  and  $\gamma_i = \gamma$ , this expression simplifies to:

$$\langle \tau \rangle = \sum_{j=1}^N \frac{1}{k} \left( 1 + \frac{\gamma}{k} \right)^{N-j} = \frac{\left( 1 + \frac{\gamma}{k} \right)^N - 1}{\gamma}. \quad (\text{A7})$$

TABLE I. Fitting parameters for the normal distribution in Fig. 7.

Figures	$N_v$	$E_{\mathcal{T}} \left[ \frac{\text{kJ}}{\text{mol}} \right]$	$E_{\mathcal{F}} \left[ \frac{\text{kJ}}{\text{mol}} \right]$
7(a) and 7(b)	20	$953.74 \pm 51.22$	$914.26 \pm 58.23$
7(d) and 7(e)	100	$4888.04 \pm 179.66$	$4845.02 \pm 183.05$
7(g) and 7(h)	1.000	$52756.76 \pm 607.74$	$52696.46 \pm 618.11$

## APPENDIX B: FITTING PARAMETERS FOR THE FIGURES IN THE MAIN TEXT

The parameters for the fitted normal distributions for the distribution of the total activation energies of the spanning trees and spanning forests in Fig. 7 are listed in Table I.

Quadratic function of the shape  $\ln k = aT^{-2} + bT^{-1} + c$  were fitted to the data in Figs. 9–11. The corresponding parameters and the standard errors are listed in Table II.

## APPENDIX C: GENERATING THE TOTAL ACTIVATION ENERGY DISTRIBUTIONS

The base graphs for the simulations in Sec. II E were generated randomly, with the intent to display fundamental characteristics of a biochemical network involved in a developmental process. Specifically:

- (i) There must be a path from any vertex to the final vertex  $N + 1$ ;
- (ii) The distance from vertex 0 to  $N + 1$  should be relatively large.

For these reasons, the standard Erdős-Rényi model, which creates edges with a fixed probability, was deemed inappropriate. Instead, the base graph was constructed and the total activation energies of the trees and two-tree spanning forests were generated using the following procedure:

(i) A function was defined to generate a random skeleton graph on  $n$  vertices, organized into a hierarchical structure with the following properties:

- (1) The first layer contains a single source vertex, and the final layer contains a single target vertex.
- (2) The remaining vertices are assigned to intermediate layers with a probability  $p^k$ , where  $k$  is the number of

TABLE II. Fitting parameters for the quadratic fit of the data from [24–26] shown in the main text in Figs. 9–11, respectively.

Figure	$a \left[ 10^7 \text{K}^2 \right]$	$b \left[ 10^5 \text{K} \right]$	$c \left[ 10^2 \ln \text{h}^{-1} \right]$
9(a)	$-1.06 \pm 0.31$	$0.68 \pm 0.21$	$-1.10 \pm 0.36$
9(b)	$-1.81 \pm 0.13$	$1.19 \pm 0.09$	$-1.98 \pm 0.15$
9(c)	$-1.96 \pm 0.14$	$1.30 \pm 0.09$	$-2.17 \pm 0.16$
9(d)	$-1.71 \pm 0.27$	$1.13 \pm 0.19$	$-1.88 \pm 0.31$
10(a)	$-1.48 \pm 0.06$	$0.97 \pm 0.04$	$-1.60 \pm 0.07$
10(b)	$-1.47 \pm 0.06$	$0.97 \pm 0.04$	$-1.59 \pm 0.07$
11(a)	$-1.74 \pm 0.07$	$1.14 \pm 0.04$	$-1.88 \pm 0.08$
11(b)	$-2.62 \pm 0.18$	$1.73 \pm 0.12$	$-2.89 \pm 0.21$
11(c)	$-2.36 \pm 0.22$	$1.56 \pm 0.15$	$-2.59 \pm 0.25$
11(d)	$-1.80 \pm 0.04$	$1.18 \pm 0.03$	$-1.96 \pm 0.06$
11(e)	$-1.75 \pm 0.07$	$1.15 \pm 0.04$	$-1.91 \pm 0.08$

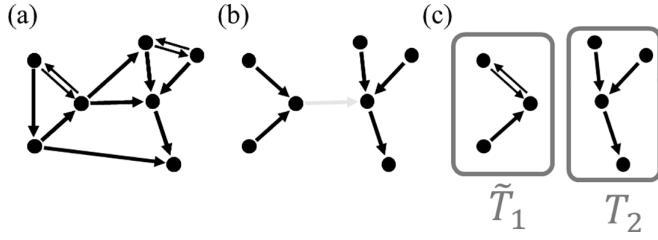


FIG. 19. Part one of the algorithm. The first steps in the algorithm to get a random two-tree forests  $\mathcal{F}$  of graph  $G$ : (a) an example graph  $G$ , (b) find a random spanning tree of  $G$  and remove an edge from it to obtain two components, (c) add to the tree without vertex  $N + 1$  any  $(y, x) \in E(G)$ , if  $(x, y)$  is on the tree.

vertices already in the current layer. If not added to the current layer, the vertex is assigned to the next layer.

(3) Each vertex is connected by an outgoing edge to a randomly chosen vertex in the next layer and to an incoming edge from a randomly chosen vertex in the previous layer.

This protocol ensures that the first vertex acts as a source and the final vertex as a target.

(ii) The function was then used to generate the main graph with  $N$  vertices, divided into  $n$  clusters:

(1) The  $N$  vertices were randomly partitioned into  $n$  nonempty sets.

Each cluster was generated as a skeleton graph for the vertices in one partition set. To also include reversible steps within clusters, a back rate was added to each edge with a probability  $q$ .

(2) A skeleton graph on  $n$  vertices was then generated to represent the interconnections between clusters. In this skeleton, each vertex was replaced by a cluster, ensuring that the source and target vertices of the cluster corresponded to the edges in the skeleton graph.

The resulting graph has a global source and a global target vertex that are relatively far apart. Additionally, it contains reversible steps within clusters and irreversible steps between clusters.

(iii) For the reaction rates, an  $N \times N$  matrix was generated, with entries  $E_{ij}$  interpreted as activation energies for the  $k_{ij}$  rates:

(1) Entries above the main diagonal ( $i < j$ ) represent forward rates and entries below the main diagonal ( $i > j$ ) represent backward rates.

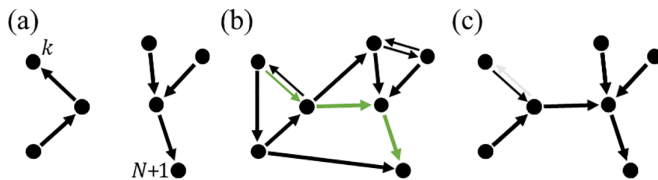


FIG. 20. Part two of the algorithm. Illustration to show that any two-tree spanning forest is the subgraph of some graph constructed as in Fig. 19: (a) choose any two-tree spanning forest, (b) choose a path from  $k$  to  $N + 1$  in  $G$ , (c) add path to forest and remove opposite edges.

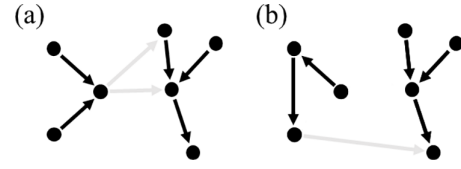


FIG. 21. Part three of the algorithm. The initial partitions are degenerate; some subgraphs are induced by more trees than others: (a) the same subgraph is induced by removing an edge of two different spanning trees, (b) a subgraph that can only be obtained by edge removal from a single spanning tree.

(2) The activation energies were drawn from uniform probability distributions. For the forward rates from  $[60 \frac{\text{kJ}}{\text{mol}}, 100 \frac{\text{kJ}}{\text{mol}}]$  and for the backward rates from  $[20 \frac{\text{kJ}}{\text{mol}}, 60 \frac{\text{kJ}}{\text{mol}}]$ .

Uniform distributions were chosen as a parsimonious alternative to Gaussian-like distributions, reflecting the typical range of activation energies in biochemical reactions [24,85].

(iv) To study the behavior for large  $N$ , 100 graphs were generated for  $N + 1 = 10$  to  $N + 1 = 100$  in steps of 10, and for  $N + 1 = 200$  to  $N + 1 = 1000$  in steps of 100.

(1) For all cases,  $n = N/10$  and  $p = q = 0.5$  were chosen.

(2) The distributions of total activation energies for spanning trees and two-tree spanning forests were computed.

(3) 10000 spanning trees and at least 10000 two-tree spanning forests were sampled for each case.

The results of these simulations are presented in the main text.

#### APPENDIX D: ALGORITHM TO COMPUTE THE MEAN FIRST-PASSAGE TIME IN GRAPH-THEORETICAL FRAMEWORK

To evaluate expressions such as Eq. (25) for large networks, it is necessary to consider all spanning trees and two-tree spanning forests. As discussed in the main text, the number of such graphs grows very rapidly with the size of the network. Except in cases where the network has a simple, recurrent structure (such as the examples provided in Appendix A), obtaining all these graphs quickly becomes impractical. Therefore, a probabilistic approach is more feasible to obtain a representative sample of the required spanning trees and two-tree spanning forests. Note that Eq. (25) is given by:

$$\langle \tau \rangle = \frac{|\mathcal{F}|}{|\mathcal{T}|} \frac{\frac{1}{|\mathcal{F}|} \sum_{\mathcal{F}} w(\mathcal{F})}{\frac{1}{|\mathcal{T}|} \sum_{\mathcal{T}} w(\mathcal{T})} = \frac{|\mathcal{F}|}{|\mathcal{T}|} \frac{\langle w(\mathcal{F}) \rangle_{\mathcal{F}}}{\langle w(\mathcal{T}) \rangle_{\mathcal{T}}}, \quad (\text{D1})$$

where  $\sum_{\mathcal{T}}$  and  $\sum_{\mathcal{F}}$  denote sums over all appropriate spanning trees and two-tree spanning forests, respectively. An algorithm to evaluate this expression thus has three objectives:

- (i) Sample  $\mathcal{T}$  at random and evaluate  $\langle w(\mathcal{T}) \rangle$ ;
- (ii) Sample  $\mathcal{F}$  at random and evaluate  $\langle w(\mathcal{F}) \rangle$ ;
- (iii) Determine the ratio  $\frac{|\mathcal{F}|}{|\mathcal{T}|}$ .



*Objective 1.* Sampling random trees  $\mathcal{T}$  rooted at vertex  $N + 1$  is a well-known problem in graph theory [71]. The most efficient solution is Wilson's algorithm [84], which is based on a loop-erased random walk.

*Objective 2.* Sampling random two-tree forests  $\mathcal{F}$ , where one tree is rooted at vertex  $N + 1$  and the other one is rooted at any another vertex and contains the initial vertex 1, involves multiple steps to ensure that all possible forests are drawn with equal probability.

Start by drawing a random spanning tree  $\mathcal{T}$  from the full graph (e.g., using Wilson's algorithm) and remove a random edge  $(r, v)$  from the path from 1 to  $N + 1$  in  $\mathcal{T}$ . While the resulting graph is a two-tree spanning forest with  $r$  and  $N + 1$  as roots of the two trees, this procedure does not yield all possible two-tree spanning forests. This is immediately clear when considering the example of a one-step master equation as discussed in Appendix A: the (in this case unique) spanning tree does not contain any back rates, which are, however, present on most two-tree spanning forest. To take this into account, for any edge  $(x, y)$  on the tree without vertex  $N + 1$ ,  $(y, x)$  has to be added to the tree, if  $(y, x) \in E(G)$ . Call the resulting subgraph  $\tilde{T}_1$ . The other component  $T_2$  is a tree rooted at  $N + 1$  and will indeed be one of the two trees of the forest. These first steps are illustrated in Fig. 19.

Any two-tree spanning forest  $\mathcal{F}$  on  $G$  is a subgraph of some  $\tilde{T}_1 \cup T_2$  obtained this way. To see this, let  $T'_k$  and  $T'_{N+1}$  be the two trees in  $\mathcal{F}$  with roots  $k$  and  $N + 1$ , respectively. Choose a path from  $k$  to  $N + 1$  in  $G$ , follow it and add all its edges to  $T'_k \cup T'_{N+1}$  until the path reaches a vertex in  $T'_{N+1}$  for the first time. Remove any edges that point in opposite direction to the edges that were added. The resulting graph will then be a spanning tree of  $G$ . If  $\tilde{T}_1 \cup T_2$  is constructed from it identifying the last edge added from the path as  $(r, v)$ , one has  $T'_k \subseteq \tilde{T}_1$  and  $T'_{N+1} = T_2$ . This construction is illustrated in Fig. 20.

Next, identify all two-tree spanning forests in  $\tilde{T}_1 \cup T_2$ . This requires finding all vertices that can serve as roots for a spanning tree of  $\tilde{T}_1$ . These vertices are precisely those reachable from  $r$  in  $\tilde{T}_1$ ,  $R = \{r_0, \dots, r_m\}$  with  $r_0 := r$ , and can be found using basic breadth-first or depth-first search algorithms [97]. Construct the spanning tree  $T_{1,r_i} \subseteq \tilde{T}_1$  for each  $i = 0, \dots, m$ . Note that there is exactly one such tree for each  $r_i$  because the graph obtained from turning all edges of  $\tilde{T}_1$  into undirected edges is a (undirected) tree and a given root uniquely defines its orientation.

Now each  $T_{1,r_i} \cup T_2$  yields a two-tree spanning forest and any two-tree spanning forest can be obtained from such a construction. But there is still one caveat: so far, these forests do not constitute a uniformly drawn sample because some forests are more likely to be obtained than others.

The first problem is that the same  $\tilde{T}_1 \cup T_2$  partition may be obtained from different spanning trees, as is shown in Fig. 21.

The degeneracy  $d$  is precisely the number of directed edges from  $R$  to  $T_2$  and the probability of obtaining this partition is proportional to the degeneracy. To ensure the same probability for all partitions, one should accept a partition with probability  $P(\text{accept } \tilde{T}_1 \cup T_2) = d^{-1}$ .

The second problem arises from the fact that the number of  $r_i$  is not the same for different  $\tilde{T}_1$ , as shown in Fig. 22. Indeed, if only one  $T_{1,r_i} \subseteq \tilde{T}_1$  were chosen, the probability

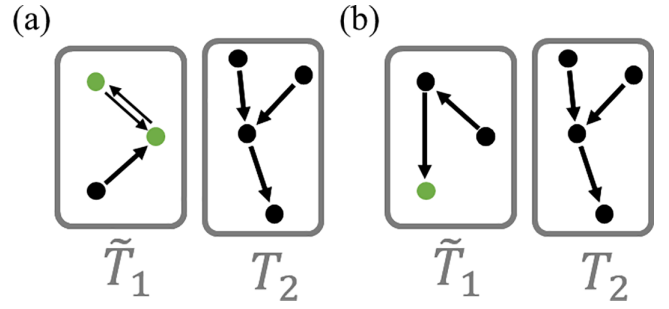


FIG. 22. Part four of the algorithm. The partitions can contain a different number of potential roots, which has to be accommodated for to ensure a uniformly-drawn sample. (a) A partition where  $\tilde{T}_1$  contains two roots (green vertices). (b) A partition where  $\tilde{T}_1$  contains one root (green vertex).

to choose  $\tilde{T}_1$  in the first place should be proportional to  $|R|$ . However, achieving this for general networks would require accepting it with a probability of  $\frac{|R|}{N}$ , which is too small to be an efficient method for generating spanning forests. Since one is only interested in sampling the mean, it is more practical to keep **all** the  $T_{1,r_i}$  for  $i = 0, \dots, m$  instead.

Note that if one only seeks to evaluate Eq. (D1), there is a third possibility that is even simpler: one may directly compute:

$$\langle w(\mathcal{F}) \rangle_{\mathcal{F}} = \frac{\sum_{m=1}^M \frac{|R_m|}{d_m} w(\mathcal{F}_m)}{\sum_{m=1}^M \frac{|R_m|}{d_m}}, \quad (\text{D2})$$

where  $M$  is the number of two-tree forests that were drawn,  $\mathcal{F}_m$  the  $m$ th forest and  $|R_m|$  and  $d_m$  the number of its roots and its degeneracy, respectively.

The essential steps to get the two-tree spanning forests and the weights are summarized as a pseudocode (Algorithm 1).

However, the objective of the simulations performed for this work was to obtain the probability distributions of the total activation energies along the two-tree spanning forests, which requires a uniformly drawn base sample. Therefore, the slightly more complicated and wasteful procedure explained before was used.

*Objective 3.* Finding  $\frac{|\mathcal{F}|}{|\mathcal{T}|}$  can be integrated in the part where the two-tree spanning forests are generated. Note that any spanning tree has as many edges that can be removed, as there are edges on the path from 1 to  $N + 1$ . Once an edge is removed, there are  $|R|$  two-tree spanning forests obtained from this tree. However, these are obtained from  $d$  different spanning trees, and one has to correct for this again. In total, the ratio can be estimated as:

$$\frac{|\mathcal{F}|}{|\mathcal{T}|} = \frac{1}{M} \sum_{m=1}^M L_m \frac{|R_m|}{d_m}, \quad (\text{D3})$$

where  $L_m$  is the length of the path from 1 to  $N + 1$  on the  $m$ th tree. Observe that Eq. (D2) can now be written as:

$$\langle w(\mathcal{F}) \rangle_{\mathcal{F}} = \frac{1}{M} \frac{|\mathcal{T}|}{|\mathcal{F}|} \frac{\sum_{m=1}^M L_m \frac{|R_m|}{d_m}}{\sum_{m=1}^M \frac{|R_m|}{d_m}} \sum_{m=1}^M L_m \frac{|R_m|}{d_m} w(\mathcal{F}_m), \quad (\text{D4})$$



ALGORITHM 1. Find random spanning forests of two trees of the directed graph  $G$ .

---

**Input:** A directed graph  $G$   
**Output:** Random two-tree spanning forest of  $G$ , one tree rooted in  $N + 1$  and the other one rooted at any other vertex together with the weight of this choice  $|R|/d$

---

```

1   $v_{\text{end}} \leftarrow \max(\text{vertices of } G);$ 
2   $T \leftarrow \text{directed\_spanning\_tree}(G, v_{\text{end}});$ 
3   $\text{edges} \leftarrow \text{list of edges of } T;$ 
4   $\text{random\_edge} \leftarrow \text{choose a random edge from edges};$ 
5   $T \leftarrow T \setminus \text{random\_edge};$ 
6   $r_0 \leftarrow \text{source vertex of random\_edge};$ 
7   $T_{\text{undirected}} \leftarrow T.\text{to\_undirected}();$ 
8   $C \leftarrow \text{connected components of } T_{\text{undirected}};$ 
9   $\text{trees} \leftarrow [T_{\text{undirected}}.\text{subgraph}(c).\text{copy}() \text{ for } c \text{ in } C];$ 
10  $T_1, T_2 \leftarrow \text{trees}[0], \text{trees}[1];$ 
11 if  $v_{\text{end}} \in T_1.\text{vertices}()$  then
12    $T_1, T_2 \leftarrow T_2, T_1;$ 
13  $T_{1,\text{directed}} \leftarrow \text{inherit\_directions}(G, T_1);$ 
14  $R \leftarrow \text{descendants}(T_{1,\text{directed}}, r_0);$ 
15  $d \leftarrow 0;$ 
16 for  $u \in T_2.\text{vertices}()$  do
17   for  $v \in \text{end\_vertices}_1()$  do
18     if  $G.\text{has\_edge}(v, u)$  then
19        $d \leftarrow d + 1;$ 
20  $r \leftarrow \text{choose a random root from } R;$ 
21  $T_{1,\text{oriented}} \leftarrow \text{direct\_towards\_root}(T_1, r);$ 
22  $T_{2,\text{oriented}} \leftarrow \text{direct\_towards\_root}(T_2, v_{\text{end}});$ 
23  $F \leftarrow \text{compose}(T_{1,\text{oriented}}, T_{2,\text{oriented}});$ 
24 return  $F, \frac{|R|}{d};$ 

```

---

and inserting this into Eq. (D1) yields:

$$\langle \tau \rangle = \frac{1}{M} \frac{\sum_{m=1}^M L_n \frac{|R_n|}{d_n}}{\sum_{m=1}^M \frac{|R_n|}{d_n}} \frac{\sum_{m=1}^M L_m \frac{|R_m|}{d_m} w(\mathcal{F}_m)}{\langle w(\mathcal{T}) \rangle_{\mathcal{T}}}. \quad (\text{D5})$$

So, this method does not even require to explicitly calculate  $\frac{|\mathcal{F}|}{|\mathcal{T}|}$ . To illustrate that the sampling algorithm for the two-tree spanning forests works, it was applied to the one-step master equation from Appendix A. The two-tree spanning forests were already characterized before. They consist of all choices of  $j < N + 1$  and  $j \leq m < N - 1$  corresponding to the forest with back rates from  $m$  to  $j$  and forward rates otherwise [compare also Fig. 17(b)]. So for a fixed number of back rates  $l = m - j$ , there are exactly  $N - l$  choices for  $j$  and hence,  $N - l$  corresponding forests. In total, this means there are  $\sum_{l=0}^N (N - l) = \frac{N(N+1)}{2}$  forests. So, if one draws uniformly a two-tree spanning forest, the probability  $p(l)$  that it has  $l$  back rates is given by:

$$p(l) = \frac{2}{N(N+1)}(N - l). \quad (\text{D6})$$

The exact distribution for  $p(l)$  was compared to the one obtained by sampling with the algorithm for a one-step master equation on 100 vertices, and one observes that the simulated distribution indeed converges to the exact solution, confirming that the algorithm yields a uniform sample (see Fig. 23)

The convergence of the MFPT obtained via the algorithm to the correct value is shown in Fig. 24(a). The exact value is given by Eq. (A3).

An advantage of this algorithm over simulating the stochastic dynamics of the network is that it allows a sample of spanning trees and two-tree spanning forests to be generated once, enabling the computation of the mean first-passage time for different rates without regenerating the sample. This facilitates studying the impact of changing rates without the need to redo the entire simulation each time. To illustrate this, one sample of spanning forests of 50176 spanning forests (from 1000 draws) was generated and the values of the MFPT were calculated for  $k_i = 1$  from  $r_i = 0$  to  $r_i = 0.9$  for all  $i$ . The result is again compared to the exact solution given in Eq. (A3). This is shown in Fig. 24(b).

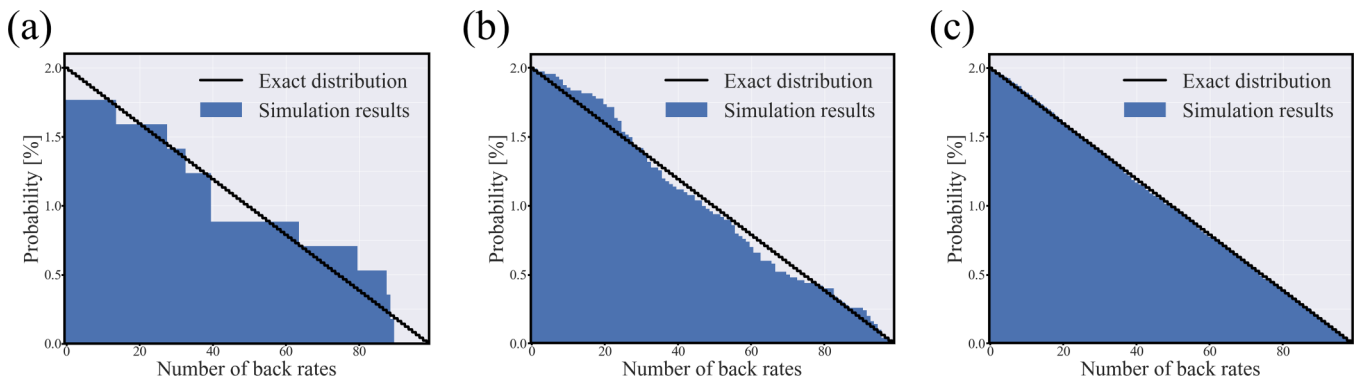


FIG. 23. Distribution of the number of back rates on the two-tree spanning forests for samples of increasing size  $n$  generated by the algorithm. The base graph is a linear graph on 100 vertices with forward and backward rates (one-step master equation). The black curve shows the exact distribution from the analytical formula. (a)  $n = 10$ . (b)  $n = 100$ . (c)  $n = 1000$ .

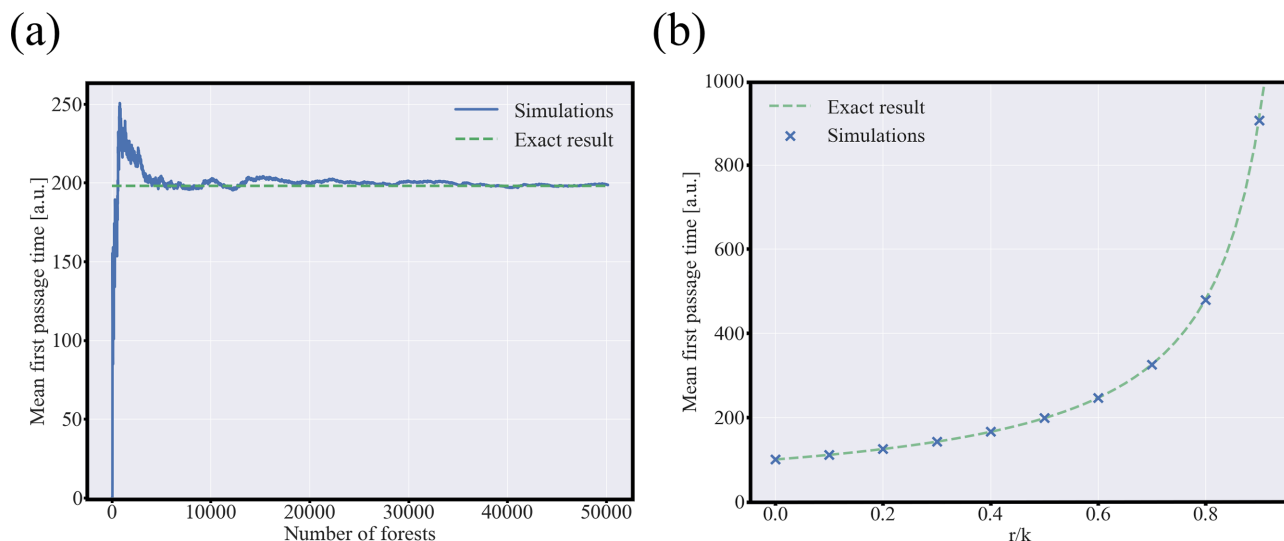


FIG. 24. The algorithm applied to the one-step master equation. (a) MFPT obtained by the algorithm for an increasing number of two-tree spanning forests compared to the theoretical value for a one-step master equation on 100 vertices with forward rates  $k_i = 1$  and backward rates  $r_i = 0.5$  for all  $i$ . (b) MFPT from same sample of two-tree spanning forests for different values of the back rate  $r$  compared to the exact result.

- [1] C. M. Blatteis, Fever: Pathological or physiological, injurious or beneficial? *J. Therm. Biol.* **28**, 1 (2003).
- [2] IPCC, *Synthesis Report of the IPCC Sixth Assessment Report (ar6)* (IPCC, Geneva, Switzerland, 2023).
- [3] F. García *et al.*, Changes in temperature alter the relationship between biodiversity and ecosystem functioning, *Proc. Natl. Acad. Sci. USA* **115**, 10989 (2018).
- [4] S. Arrhenius, Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren, *Z. Phys. Chem.* **4**, 226 (1889).
- [5] S. Logan, The origin and status of the Arrhenius equation, *J. Chem. Educ.* **59**, 279 (1982).
- [6] H. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, *Physica* **7**, 284 (1940).
- [7] P. Hänggi, P. Talkner, and M. Borkovec, Reaction-rate theory: Fifty years after Kramers, *Rev. Mod. Phys.* **62**, 251 (1990).
- [8] H. Eyring, The activated complex in chemical reactions, *J. Chem. Phys.* **3**, 107 (1935).
- [9] M. Evans and M. Polanyi, Some applications of the transition state method to the calculation of reaction velocities, especially in solution, *Trans. Faraday Soc.* **31**, 875 (1935).
- [10] J. C. Bischof and X. He, Thermal stability of proteins, *Ann. N.Y. Acad. Sci.* **1066**, 12 (2006).
- [11] M. L. Begasse *et al.*, Temperature dependence of cell division timing accounts for a shift in the thermal limits of *C. Elegans* and *C. Briggsae*, *Cell Rep.* **10**, 647 (2015).
- [12] L. Fujise *et al.*, Cell cycle dynamics of cultured coral endosymbiotic microalgae (symbiodinium) across different types (species) under alternate light and temperature conditions, *J. Eukaryotic Microbiol.* **65**, 505 (20018).
- [13] C. Rieder and H. Maiato, Stuck in division or passing through: Review what happens when cells cannot satisfy the spindle assembly checkpoint, *Dev. Cell* **7**, 637, (2004).
- [14] A. Moore *et al.*, Effects of temperature shift on cell cycle, apoptosis and nucleotide pools in CHO cell batch cultures, *Cytotechnology* **23**, 47 (1997).
- [15] T. Hayashida *et al.*, The CCGF/TAE250 gene is mutated in thermosensitive G1 mutants of the BHK21 cell line derived from golden hamster, *Gene*, **41**, 247 (1994).
- [16] P. Rao and J. Engelberg, Hela cells: Effects of temperature on the life cycle, *Science* **148**, 1092 (1965).
- [17] J. Siskin, L. Morasca, and S. Kibby, Effects of temperature on the kinetics of the mitotic cycle of mammalian cells in culture, *Exp. Cell Res.* **39**, 103 (1965).
- [18] I. Watanabe and S. Okada, Effects of temperature on growth rate of cultured mammalian cells (L5178Y), *J. Cell Biol.* **32**, 309 (1967).
- [19] H. Falahati *et al.*, Temperature-induced uncoupling of cell cycle regulators, *Dev Biol.* **470**, 147 (2021).
- [20] H. J. Evans and J. R. K. Savage, The effect of temperature on mitosis and on the action of colchicine in root meristem cells of *Vicia faba*, *Exp. Cell Res.* **18**, 51 (1959).
- [21] O. Oleksiuk *et al.*, Thermal robustness of signaling in bacterial chemotaxis, *Cell* **145**, 312 (2011).
- [22] M. Vanoni, M. Vai, and G. Frascotti, Effects of temperature on the yeast cell cycle analyzed by flow cytometry, *Cytometry* **5**, 530 (1984).
- [23] V. Zachleder *et al.*, Cell cycle arrest by supraoptimal temperature in the alga *Chlamydomonas reinhardtii*, *Cells* **8**, 1237 (2019).
- [24] J. Crapse *et al.*, Evaluating the Arrhenius equation for developmental processes, *Mol. Syst. Biol.* **2021**, e9895 (2021).
- [25] C. Bliss, Temperature characteristics for prepupal development in drosophila melanogaster, *J. Gen. Physiol.* **9**, 467 (1926).
- [26] L. Powsner, The effects of temperature on the durations of the developmental stages of drosophila melanogaster, *Physiolog. Zoology* **8**, 474 (1935).

- [27] D. Ratkowsky *et al.*, Relationship between temperature and growth rate of bacterial cultures, *ASM J. Bacteriol.* **149**, 1 (1982).
- [28] F. Johnson, H. Eyring, and B. Stover, *The Theory of Rate Processes in Biology and Medicine* (John Wiley and Sons, New York, 1974).
- [29] J. M. Membré *et al.*, Temperature effect on bacterial growth rate: Quantitative microbiology approach including cardinal values and variability estimates to perform growth simulations on/in food, *Int. J. Food Microbiol.* **100**, 179 (2005).
- [30] H. Thormar, Effect of temperature on the reproduction rate of tetrahymena pyriformis, *Exp. Cell. Res.* **28**, 269 (1962).
- [31] F. A. M. Alberghina, Growth regulation in *Neurospora crassa* effects of nutrients and of temperature, *Arch. Microbiol.* **89**, 83 (1973).
- [32] T. Yoshida *et al.*, Temperature effects on the egg development time and hatching success of three acartia species (*Copepoda: Calanoida*) from the Strait of Malacca, *Zool. Stud.* **51**, 644 (2012).
- [33] J. Choi *et al.*, Probing and manipulating embryogenesis via nanoscale thermometry and temperature control, *Proc. Nat. Acad. Sci. USA* **117**, 14636 (2020).
- [34] A. Dell, S. Pawara, and V. Savage, Systematic variation in the temperature dependence of physiological and ecological traits, *Proc. Natl. Acad. Sci. USA* **108**, 10591 (2011).
- [35] J. Hastings and B. Sweeney, On the mechanism of temperature independence in a biological clock, *Proc. Nat. Acad. Sci. USA* **43**, 804 (1957).
- [36] P. Ruoff *et al.*, Modeling temperature compensation in chemical and biological oscillators, *Chronobiol. Int.* **14**, 499 (1997).
- [37] P. Ruoff, M. Vinsjevik, and L. Rensing, Temperature compensation in biological oscillators: A challenge for joint experimental and theoretical analysis, *Comments Theor. Biol.* **5**, 361 (2000).
- [38] P. Ruoff *et al.*, Temperature dependency and temperature compensation in a model of yeast glycolytic oscillations, *Biophys. Chem.* **106**, 179 (2003).
- [39] D. Virshup and D. Forger, Keeping the beat in the rising heat, *Cell* **137**, 602 (2009).
- [40] J. Leloup and A. Goldbeter, Temperature compensation of circadian rhythms: Control of the period in a model for circadian oscillations of the per protein in *Drosophila*, *Chronobiol. Int.* **14**, 511 (1997).
- [41] C. Hong and J. Tyson, A proposal for temperature compensation of the circadian rhythm in *Drosophila* based on dimerization of the per protein, *Chronobiol. Int.* **14**, 521 (1997).
- [42] C. Pittendrigh, On temperature independence in the clock system controlling emergence time in *Drosophila*, *Proc. Nat. Acad. Sci. USA* **40**, 1018 (1954).
- [43] P. Kida, M. Younga, and E. Siggia, Temperature compensation and temperature sensation in the circadian clock, *Proc. Nat. Acad. Sci. USA* **112**, E6284 (2015).
- [44] C. Vibe, The temperature response of the medaka segmentation clock and its link to robustness in embryonic patterning, Ph.D. thesis, University of Heidelberg, Germany, 2021.
- [45] K. Valeur and R. degli Agosti, Simulations of temperature sensitivity of the peroxidase-oxidase oscillator, *Biophys. Chem.* **99**, 259 (2002).
- [46] P. Ruoff, Introducing temperature-compensation in any reaction kinetic oscillator model, *J. Interdiscipl. Cycle Res.* **23**, 92 (1992).
- [47] J. Bourn and M. Dorrity, Degrees of freedom: Temperature's influence on developmental rate, *Curr. Opin. Gen. Dev.* **85**, 102155 (2024).
- [48] J. Rombouts Jr., F. Tavella, A. Vandervelde, C. Phong, J. E. Ferrell, Q. Yang, and L. Gelens, Mechanistic origins of temperature scaling in the early embryonic cell cycle, *Nat. Commun.* **16**, 8045 (2025).
- [49] J. J. Tyson, K. C. Chen, and B. Novak, Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell, *Curr. Opin. Cell Biol.* **15**, 221 (2003).
- [50] D. Ratkowsky, J. Olley, and T. Ross, Unifying temperature effects on the growth rate of bacteria and the stability of globular protein, *J. Theor. Biol.* **233**, 351 (2005).
- [51] R. Schoolfield, P. Sharpe, and C. Magnuson, Non-linear regression of biological temperature-dependent rate models based on absolute reaction-rate theory, *J. Theor. Biol.* **88**, 719 (1981).
- [52] T. Wagner *et al.*, Modeling insect development rates: A literature review and application of a biophysical model, *Ann. Entomol. Soc. Am.* **77**, 208 (1984).
- [53] J. Knies and J. Kingsolver, Erroneous Arrhenius: Modified Arrhenius model best explains the temperature dependence of ectotherm fitness, *Am. Nat.* **176**, 227 (2010).
- [54] S. Iyer-Biswas *et al.*, Scaling laws governing stochastic growth and division of single bacterial cells, *Proc. Nat. Acad. Sci. USA* **111**, 15912 (2014).
- [55] B. Quinn, A critical review of the use and performance of different function types for modeling temperature-dependent development of arthropod larvae, *J. Therm. Biol.* **63**, 65 (2017).
- [56] F. Rebaudo and V. Rabhi, Modeling temperature-dependent development rate and phenology in insects: Review of major developments, challenges, and future directions, *Entomol. Exp. Appl.* **166**, 607 (2018).
- [57] B. Régnier, J. Legrand, and F. Rebaudo, Modeling temperature-dependent development rate in insects and implications of experimental design, *Environ. Entomol.* **51**, 132 (2022).
- [58] X. Yin *et al.*, A nonlinear model for crop development as a function of temperature, *Agric. For. Meteorol.* **77**, 1 (1995).
- [59] M. Zwietering *et al.*, Modeling of bacterial growth as a function of temperature, *Appl. Environ. Microbiol.* **57**, 1094 (1991).
- [60] F. Guerrero, J. M. Blanco, and V. Rodríguez, Temperature-dependent development in marine copepods: A comparative analysis of models, *J. Plankton Res.* **16**, 95 (1994).
- [61] W. Müller-Esterl, *Biochemie* (Springer Spektrum, Berlin, 2018).
- [62] W. Müller, *Developmental Biology* (Springer Science & Business Media, Berlin, 1996).
- [63] T. Gillespie, A rigorous derivation of the chemical master equation, *Physica A* **188**, 404 (1992).
- [64] D. Schnoerr, G. Sanguinetti, and R. Grima, Approximation and inference methods for stochastic biochemical kinetics—a tutorial review, *J. Phys. A: Math. Theor.* **50**, 093001 (2017).
- [65] K. Nam, R. Martinez-Corral, and J. Gunawardena, The linear framework: Using graph theory to reveal the algebra and thermodynamics of biomolecular systems, *Interface Focus* **12**, 20220013 (2022).
- [66] S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, Cambridge, 2001).
- [67] N. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 2004).

- [68] T. Kanesaki, C. M. Edwards, U. S. Schwarz, and J. Grosshans, Dynamic ordering of nuclei in syncytial embryos: A quantitative analysis of the role of cytoskeletal networks, *Integr. Biol.* **3**, 1112 (2011).
- [69] J. Rodenfels, K. M. Neugebauer, and J. Howard, Heat oscillations driven by the embryonic cell cycle reveal the energetic costs of signaling, *Dev. Cell* **48**, 646 (2019).
- [70] J. E. Ferrell, T. Y.-C. Tsai, and Q. Yang, Modeling the cell cycle: Why do certain circuits oscillate? *Cell* **144**, 874 (2011).
- [71] R. Lyons and Y. Peres, *Probability on Trees and Networks* (Cambridge University Press, Cambridge, 2017), Vol. 42.
- [72] J. Honerkamp, *Stochastic Dynamical Systems: Concepts, Numerical Methods, Data Analysis* (John Wiley & Sons, New York, 1996).
- [73] K. Nam, Algebraic approaches to molecular information processing, Ph.D. thesis, Harvard University, 2021.
- [74] K. Nam and J. Gunawardena, The linear framework II: Using graph theory to analyze the transient regime of Markov processes, *Front. Cell Dev. Biol.* **11**, 1233808 (2023).
- [75] F. Khodabandehlou, C. Maes, and K. Netočný, Trees and forests for nonequilibrium purposes: An introduction to graphical representations, *J. Stat. Phys.* **189**, 41 (2022).
- [76] M. Bladt and B. Nielsen, *Matrix-Exponential Distributions in Applied Probability* (Springer, Berlin, 2017), Vol. 81.
- [77] P. D. Leenheer, An elementary proof of a matrix tree theorem for directed graphs, *SIAM Rev.* **62**, 716 (2020).
- [78] T. Hill, Studies in irreversible thermodynamics IV: Diagrammatic representation of steady state fluxes for unimolecular systems, *J. Theor. Biol.* **10**, 442 (1966).
- [79] R. Zia and B. Schmittmann, Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states, *J. Stat. Mech.: Theory Exp.* (2007) P07012.
- [80] G. Bell, B. Munsky, and I. Nemenman, The simplicity of completion time distributions for common complex biochemical processes, *Phys. Biol.* **7**, 016003 (2009).
- [81] J. Chong, C. Amourda, and T. Saunders, Temporal development of *Drosophila* embryos is highly robust across a wide temperature range, *J. R. Soc. Interface* **15**, 20180304 (2018).
- [82] A. Cayley, A theorem on trees, *Quart. J. Math.* **23**, 376 (1889).
- [83] J. B. Voits and U. S. Schwarz, Code and Data for ‘The Generic Temperature Response of Large Biochemical Networks’, Github (2025), <https://github.com/JVoits/networks>.
- [84] D. B. Wilson, Generating random spanning trees more quickly than the cover time, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery (ACM), New York, 1996), pp. 296–303.
- [85] J. Lepock, Measurement of protein stability and protein denaturation in cells using differential scanning calorimetry, *Methods* **35**, 117 (2005).
- [86] C. K. Ewe, E. M. Sommermann, J. Kenchel, S. E. Flowers, M. F. Maduro, P. M. Joshi, and J. H. Rothman, Feedforward regulatory logic controls the specification-to-differentiation transition and terminal cell fate during *Caenorhabditis elegans* endoderm development, *Development* **149**, dev200337 (2022).
- [87] One can find mathematical counterexamples otherwise, e.g.,  $k_i = i^2 k$ , which converges to  $CV = \frac{2}{\sqrt{10}}$ .
- [88] R. C. Bradley, *Introduction to Strong Mixing Conditions* (Kendrick Press, Heber City, 2007).
- [89] A. Mateus *et al.*, Thermal proteome profiling in bacteria: Probing protein state *in vivo*, *Mol. Syst. Biol.* **14**, e8242 (2018).
- [90] V. Aquilanti *et al.*, Temperature dependence of chemical and biophysical rate processes: Phenomenological approach to deviations from Arrhenius law, *Chem. Phys. Lett.* **498**, 209 (2010).
- [91] M. Elias, G. Wieczorek, S. Rosenne, and D. S. Tawfik, The universality of enzymatic rate–temperature dependency, *Trends Biochem. Sci.* **39**, 1 (2014).
- [92] S. Gomez Melo, D. Wörthmüller, P. Gönczy, N. Banterle, and U. S. Schwarz, Grand canonical brownian dynamics simulations of adsorption and self-assembly of SAS-6 rings on a surface, *J. Chem. Phys.* **158**, 085102 (2023).
- [93] D. T. Gillespie, Stochastic simulation of chemical kinetics, *Annu. Rev. Phys. Chem.* **58**, 35 (2007).
- [94] D. Kulasiri and R. Kosarwal, *Chemical Master Equation for Large Biological Networks: State-Space Expansion Methods Using AI* (Springer, Singapore, 2021).
- [95] N. Polizzi, M. J. Therien, and D. Beratan, Mean first-passage times in biology, *Isr. J. Chem.* **56**, 816 (2017).
- [96] L. Ham *et al.*, A stochastic vs deterministic perspective on the timing of cellular events, *Nat. Commun.* **15**, 54286 (2024).
- [97] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT Press, Cambridge, 2022).