Diagrammatic expressions for steady-state distribution and static responses in population dynamics

Koya Katayama [1,*] Ryuna Nagayama [1] and Sosuke Ito [1,2]

[1]*Department of Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

[2]*Universal Biology Institute, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

(Received 5 June 2025; accepted 20 January 2026; published 24 March 2026)

One of the fundamental questions in population dynamics is how biological populations respond to environmental perturbations. In population dynamics, the mean fitness and the fraction of a trait in the steady state are important because they indicate how well the trait and the population adapt to the environment. In this study, we examine the parallel mutation–reproduction model, which is one of the simplest models of an evolvable population. As an extension of the Markov chain tree theorem, we derive diagrammatic expressions for the static responses of the mean fitness and the steady-state distribution of the population. For the parallel mutation–reproduction model, we consider self-loops, which represent trait reproduction and are excluded from the Markov chain tree theorem for the linear master equation. To generalize the theorem, we introduce the concept of rooted 0/1 loop forests, which generalize spanning trees with loops. We demonstrate that the weights of rooted 0/1 loop forests yield the static responses of the mean fitness and the steady-state distribution. Our results provide exact expressions for the static responses and the steady-state distribution. Additionally, we discuss approximations of these expressions in cases where reproduction or mutation is dominant. We provide numerical examples to illustrate these approximations and exact expressions. We also demonstrate how our results can be used to design control strategies for harmful populations.

DOI: 10.1103/fc35-47fs

I. INTRODUCTION

Understanding how biological populations respond to environmental changes is fundamental to biology because evolution itself can be viewed as a long-term response to these changes [1,2]. The stability of biological populations in the face of perturbations, such as climate change, is directly linked to their ability to respond to and adapt to shifting environments [3–6]. In practical terms, understanding how viruses and cancers resist treatment and environmental stress remains an urgent challenge in the field of human health [7–12]. Moreover, it has been a key challenge in agriculture to understand how pests acquire tolerance to pesticides [13,14]. Substantial efforts have accordingly been made to quantify how various properties of biological populations respond to external changes.

Two of the most fundamental quantities in population dynamics are the mean fitness and the fraction of a trait. The mean fitness, defined as the growth rate of the total population, reflects how well a population is adapted to its environment. Because the mean fitness is experimentally accessible, the mean fitness is widely discussed not only theoretically [1,2,15–37] but also experimentally [25,34,38–41]. The fraction of a trait indicates how well it adapts to the environment [2,11,12,29], and this fraction is also experimentally accessible [34,42]. These quantities can be discussed theoretically using the parallel mutation–reproduction model [2], which is one of the simplest population genetics models described by a nonlinear master equation. Since then, it has been applied to ecosystems and population dynamics [15–22].

Although the parallel mutation–reproduction model is important, its mathematical treatment is less developed than that of the linear master equation because it is nonlinear. For example, it requires computing the eigenvalues and eigenvectors of a matrix to express the steady-state distribution and the static response of mean fitness, as discussed in Ref. [16]. Here, the steady-state distribution is the fraction of a trait in the steady state, and the static response of mean fitness is the change in mean fitness between two different steady states. Therefore, an analytical discussion may be difficult without considering the eigenvalue problem for a specific matrix. Indeed, many previous studies [15,18–22] have assumed specific functional forms for the fitness landscape or the mutation rate.

In contrast, the steady-state distribution and the static responses in the steady state are well understood for the linear master equation. For example, the steady-state distribution can be described by a canonical distribution if the detailed balance condition is satisfied [43]. Even when the detailed balance condition is not satisfied, the steady-state distribution can be graphically described using the weights of the spanning tree on Markov networks, thanks to the Markov chain tree theorem [44–52] (also known as the matrix tree theorem) for the linear master equation. Recently, the static responses in

*Contact author: koya.katayama@ubi.s.u-tokyo.ac.jp

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

the nonequilibrium steady state have also been discussed in stochastic thermodynamics based on the Markov chain tree theorem [53,54]. Thus, due to the absence of a mathematical treatment similar to that established for the linear master equation in the parallel mutation–reproduction model, it has been difficult to handle analytically the steady-state distribution and the static responses of the mean fitness in the steady state for general cases, despite the importance of these quantities in population dynamics.

To address this issue, we extend the Markov chain tree theorem to the parallel mutation–reproduction model, and derive exact expressions for the steady-state distribution and the responses of the mean fitness in the steady state. To achieve this, we propose a graph called a *rooted 0/1 loop forest*, which is a generalization of a spanning tree. In terms of a loop in a rooted 0/1 loop forest, the effect of self-reproduction can be naturally described. We thus obtain diagrammatic expressions for the steady-state distribution and the static responses of the mean fitness using the weights of rooted 0/1 loop forests, as a generalization of the Markov chain tree theorem. We also discuss approximate expressions that do not use all weights, in cases where mutation or natural selection is dominant. We verify the validity of these approximate expressions and confirm the exact diagrammatic expressions numerically.

This paper is organized as follows. In Sec. II, we introduce the parallel mutation–reproduction model and present some results from Ref. [16]. In Sec. III, we introduce some concepts from graph theory. In particular, we propose a different type of graph called a rooted 0/1 loop forest. In Sec. IV, we present the main results and provide some examples. In Sec. V, we approximate the static responses of the mean fitness and the steady-state distribution in cases where mutation or natural selection is dominant. In Sec. VI, we demonstrate how our results can be used to control the mean fitness of a population.

II. SETUP

Here, we introduce the parallel mutation–reproduction model, as well as quantities such as the steady-state distribution and the static responses of the mean fitness.

A. The parallel mutation–reproduction model

We explain the parallel mutation–reproduction model [2] that we are dealing with in this paper. Suppose that a population consists of N subpopulations, and each subpopulation has a different trait. Each individual in the population reproduces itself and mutates into a different trait [Fig. 1(a)]. Although other factors such as migration or predation may affect the dynamics of the population, we neglect them for simplicity. Let $n_i(t) \geqslant 0$ denote the number of individuals in the subpopulation with trait $i \in \{1, \ldots, N\}$ at time t. In the following, we assume that $n_i(t)$ is large enough to be treated as a continuous number. Let $R_i \in \mathbb{R}$ denote the reproductive rate (or the fitness) of trait i, and $M_{ij} \geqslant 0\,(i \neq j)$ denote the mutation rate from trait j to trait i. As explained in Appendix A, R_i incorporates the net effect of reproduction and death of trait i. In addition, M_{ij} incorporates the combined effect of the reproduction of trait j and the mutation of its descendants. Then, we can write down the equation that describes the dynamics of $n_i(t)$ as follows:

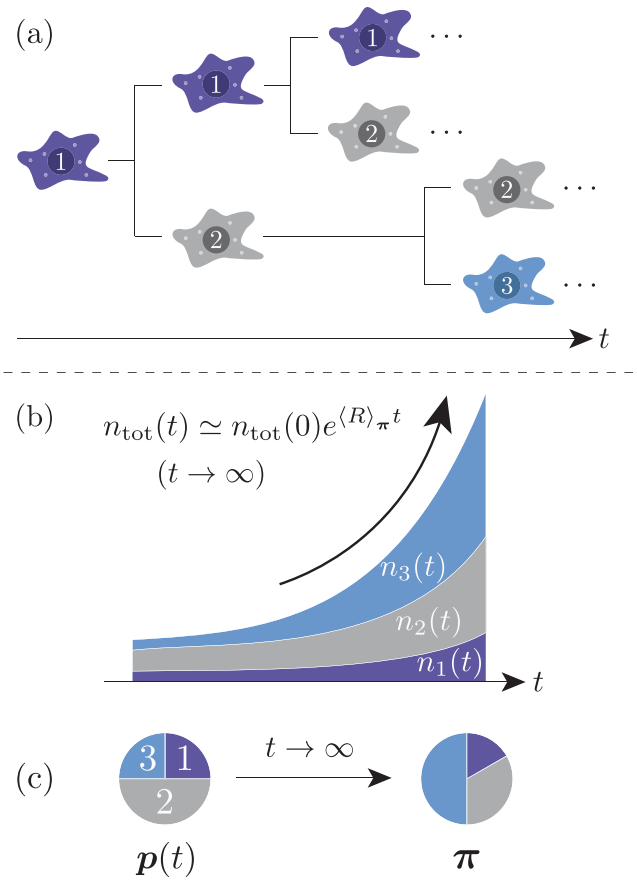

FIG. 1. Schematic illustration of the model. (a) Each individual reproduces, and its offspring may mutate into a different trait. (b) In the long-time limit, the total number of individuals in the population grows exponentially with time, and $\langle R\rangle_{\boldsymbol{\pi}}$ is regarded as the reproductive rate of the total number of the individuals in the population. (c) The distribution of the population $\boldsymbol{p}(t)$ converges to the unique steady state $\boldsymbol{\pi}$ in the limit of $t \to \infty$.

$$\frac{d}{dt}n_i(t) = R_i n_i(t) + \sum_{j(\neq i)} (M_{ij}n_j(t) - M_{ji}n_i(t)). \tag{1}$$

If there is competition for food or predator-prey interaction between individuals, $\{R_i\}_i$ may depend on $\boldsymbol{n}(t) := (n_1(t), \ldots, n_N(t))^\top$. However, in this paper we assume that $\{R_i\}_i$ are independent of $\boldsymbol{n}(t)$. We also assume that all R_i and M_{ij} do not depend directly on time t, which means that the environment is stationary. For simplicity, we rewrite Eq. (1) in the vector form. To do this, we define $M_{ii} := -\sum_{j(\neq i)} M_{ji}$, $\mathsf{M} := (M_{ij}) \in \mathbb{R}^{N\times N}$, and $\mathsf{R} := \mathrm{diag}(R_1, \ldots, R_N) \in \mathbb{R}^{N\times N}$, where $\mathbb{R}^{N\times N}$ denotes the set of all $N \times N$ real matrices. Then, Eq. (1) can be rewritten as

$$\frac{d}{dt}\boldsymbol{n}(t) = (\mathsf{R} + \mathsf{M})\boldsymbol{n}(t). \tag{2}$$

We also assume that the matrix M is irreducible, which means that for any pair of i and j, trait i can mutate into trait j in finite steps.

In this paper, we focus on the *distribution* of the traits rather than the number of them. Defining the total number

of the population as $n_{\rm tot}(t) := \sum_i n_i(t)$, the distribution of the traits is defined as $\boldsymbol{p}(t) := (p_1(t), \ldots, p_N(t))^\top$ with $p_i(t) := n_i(t)/n_{\rm tot}(t)$. From Eq. (1), we obtain the equation that describes the dynamics of $\boldsymbol{p}(t)$ as

$$\frac{d}{dt}\boldsymbol{p}(t) = (\mathsf{R} - \langle R\rangle_{\boldsymbol{p}(t)}\mathsf{I} + \mathsf{M})\boldsymbol{p}(t), \tag{3}$$

where $\langle R\rangle_{\boldsymbol{p}(t)} := \sum_i R_i p_i(t)$ is the *mean fitness* of the population at time t, and $\mathsf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix (see also Appendix B for the derivation). The term $-\langle R\rangle_{\boldsymbol{p}(t)}\boldsymbol{p}(t)$ in Eq. (3) ensures that $\boldsymbol{p}(t)$ remains normalized over time, i.e., $\sum_i p_i(t) = 1$. Here, Eq. (3) is considered a nonlinear master equation because the term $-\langle R\rangle_{\boldsymbol{p}(t)}\boldsymbol{p}(t)$ is a nonlinear function of $\boldsymbol{p}(t)$.

Note that the mean fitness $\langle R\rangle_{\boldsymbol{p}(t)}$ can be interpreted as the reproductive rate of the total number of individuals in the population because the equation

$$\frac{d}{dt}n_{\rm tot}(t) = \langle R\rangle_{\boldsymbol{p}(t)}n_{\rm tot}(t) \tag{4}$$

is obtained by summing Eq. (1) over i.

We note that Eq. (3) reduces to the linear master equation

$$\frac{d}{dt}\boldsymbol{p}(t) = \mathsf{M}\boldsymbol{p}(t), \tag{5}$$

when all reproductive rates are equal because $\langle R\rangle_{\boldsymbol{p}(t)} = R_i$ holds for any i and t, and $\mathsf{R} = \langle R\rangle_{\boldsymbol{p}(t)}\mathsf{I}$ is obtained.

B. Steady-state distribution

We focus on the steady state of the population, which is one of the main topics of this paper. An expression for the steady-state distribution and its properties are obtained by applying the Perron–Frobenius theorem [55] (see also Appendix C).

We clarify what the steady state is in the parallel mutation–reproduction model. Equation (4) implies that the total number of the population could diverge in the limit of $t \to \infty$ due to the reproduction [Fig. 1(b)]. In contrast, the irreducibility of M and the Perron–Frobenius theorem ensure that the distribution of the population $\boldsymbol{p}(t)$ converges to the unique distribution $\boldsymbol{\pi} \in \mathbb{R}^N$, which is the *steady-state distribution* [Fig. 1(c)],

$$\boldsymbol{\pi} := \lim_{t\to\infty}\boldsymbol{p}(t). \tag{6}$$

For the details on the existence and uniqueness of $\boldsymbol{\pi}$, see Appendix C. Equation (6) implies that the mean fitness also converges to the unique value $\langle R\rangle_{\boldsymbol{\pi}}$: $\lim_{t\to\infty}\langle R\rangle_{\boldsymbol{p}(t)} = \langle R\rangle_{\boldsymbol{\pi}}$. This $\langle R\rangle_{\boldsymbol{\pi}}$ can be interpreted as follows. The solution of Eq. (4) is obtained as $n_{\rm tot}(t) = n_{\rm tot}(0)\exp(\int_0^t \langle R\rangle_{\boldsymbol{p}(t')}dt')$. By considering the long-time limit, we obtain

$$\lim_{t\to\infty}\frac{1}{t}\ln\frac{n_{\rm tot}(t)}{n_{\rm tot}(0)} = \langle R\rangle_{\boldsymbol{\pi}}. \tag{7}$$

Roughly speaking, this equality means that the total number of the population $n_{\rm tot}(t)$ increases exponentially as $n_{\rm tot}(t) \simeq n_{\rm tot}(0)e^{\langle R\rangle_{\boldsymbol{\pi}}t}$ in the long-time limit [Fig. 1(b)]. Therefore, we can interpret the mean fitness $\langle R\rangle_{\boldsymbol{\pi}}$ as the reproductive rate of the total number of the individuals in the population in the long-time limit.

Perron–Frobenius theorem also tells us that every component of $\boldsymbol{\pi}$ is strictly positive, i.e., $\pi_i > 0$ for any i (see Appendix C for the details). This fact means that no trait will go extinct in the long-time limit if M is irreducible.

The steady-state distribution $\boldsymbol{\pi}$ can be regarded as the eigenvector of $\mathsf{R}+\mathsf{M}$. In fact, combining Eq. (3) and the fact that $\boldsymbol{\pi}$ is time independent, we obtain

$$(\mathsf{R}+\mathsf{M})\boldsymbol{\pi} = \langle R\rangle_{\boldsymbol{\pi}}\boldsymbol{\pi}, \tag{8}$$

which means that $\boldsymbol{\pi}$ is the eigenvector of $\mathsf{R}+\mathsf{M}$ with eigenvalue $\langle R\rangle_{\boldsymbol{\pi}}$.

The left eigenvector of $\mathsf{R}+\mathsf{M}$ corresponding to the eigenvalue $\langle R\rangle_{\boldsymbol{\pi}}$ also plays an important role in population dynamics. The Perron–Frobenius theorem implies that there exists a unique vector $\boldsymbol{\zeta} \in \mathbb{R}^N$ which satisfies the following equations (see also Appendix C for the details):

$$\boldsymbol{\zeta}^\top(\mathsf{R}+\mathsf{M}) = \langle R\rangle_{\boldsymbol{\pi}}\boldsymbol{\zeta}^\top, \tag{9}$$

$$\boldsymbol{\zeta}^\top\boldsymbol{\pi} = 1, \tag{10}$$

and $\zeta_i > 0$ for any i. In other words, $\boldsymbol{\zeta}$ is the left eigenvector of $\mathsf{R}+\mathsf{M}$ with eigenvalue $\langle R\rangle_{\boldsymbol{\pi}}$ normalized by $\boldsymbol{\pi}$ and whose components are all positive. Note that, in the case of the linear master equation [Eq. (5)], $\boldsymbol{\zeta}$ is the vector whose every component is 1, i.e., $\boldsymbol{\zeta} = \mathbf{1} := (1, \ldots, 1)^\top$. In fact, $\mathbf{1}^\top\mathsf{M} = \mathbf{0}^\top = \mathbf{1}^\top(\langle R\rangle_{\boldsymbol{\pi}}\mathsf{I} - \mathsf{R})$ is satisfied because $\sum_i M_{ik} = 0$ for any k. $\mathbf{1}^\top\boldsymbol{\pi} = \sum_i \pi_i = 1$ is also satisfied. The positivity of the elements of $\mathbf{1}$ is also obviously satisfied.

C. Static responses of the mean fitness

Here, we discuss the static responses of the mean fitness in the steady state, which are also the main topics of this paper. Since Eq. (1) contains two kinds of parameters, R_i and M_{ij}, we can consider two kinds of the static responses $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}} := \partial\langle R\rangle_{\boldsymbol{\pi}}/\partial R_i$ and $\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}}$, which are important quantities that explain the variation in the mean fitness due to changes in the parameters. We note that the general static response of the mean fitness can be expressed as a linear combination of these quantities. Thus, it suffices to consider only $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}}$ and $\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}}$. Indeed, if the reproductive rates and mutation rates depend on a parameter $\theta \in \mathbb{R}$ as $\{R_i(\theta)\}$ and $\{M_{ij}(\theta)\}$, the static response of the mean fitness with respect to θ is given by $\partial_\theta\langle R\rangle_{\boldsymbol{\pi}} = \sum_i \partial_\theta R_i(\theta)\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}} + \sum_{i,j(\neq i)}\partial_\theta M_{ij}(\theta)\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}}$.

The static responses $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}}$ and $\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}}$ are related to the left and right eigenvectors with the eigenvalue $\langle R\rangle_{\boldsymbol{\pi}}$. Interestingly, Ref. [16] shows that the following equations

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial R_i} = \zeta_i\pi_i, \tag{11}$$

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial M_{ij}} = (\zeta_i - \zeta_j)\pi_j, \tag{12}$$

hold for $i \neq j$ (see also Appendix D). While these expressions give the rigorous forms of the static responses of the mean fitness, it is difficult to express these static responses explicitly by using only the given parameters, $\{R_i\}$ and $\{M_{ij}\}$. This is because we need to solve the eigenvalue problem of the matrix $\mathsf{R}+\mathsf{M}$ to obtain the explicit forms of $\zeta_i\pi_i$ and $\zeta_i\pi_j$. However, in this paper, we will give diagrammatic expressions for the static responses of the mean fitness. The diagrammatic

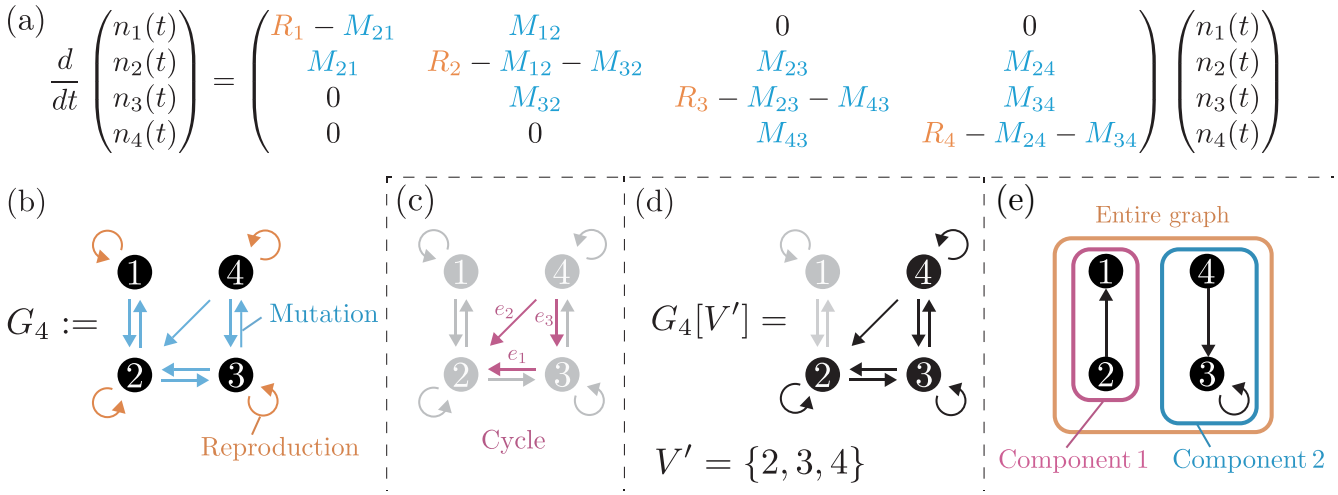

FIG. 2. Example of the basic graph, a cycle, an induced subgraph, and a component. (a) An example of Eq. (2) for a population with four traits. The mutation rates, M_{13}, M_{14}, M_{31}, M_{41}, and M_{42}, are set to zero. (b) The basic graph, denoted by G_4, corresponding to the population dynamics shown in panel (a). At each vertex, there is a loop representing reproduction. Also, if the mutation rate M_{ij} is nonzero, there is a directed edge from vertex j to i. (c) An example of a cycle in the graph shown in panel (b). This cycle (e_1, e_2, e_3) consists of three edges: $e_1 = (2 \leftarrow 3)$, $e_2 = (2 \leftarrow 4)$, and $e_3 = (3 \leftarrow 4)$. Here, a trail is $3 - 2 - 4 - 3$. As this example illustrates, the edges in a cycle do not need to be oriented in the same direction along the cycle. (d) An example of an induced subgraph $G_4[V'] \subseteq G_4$ for $V' = \{2, 3, 4\}$. $G_4[V']$ is obtained from G_4 by removing vertex 1 and the edges with at least one endpoint at vertex 1: $(1 \leftarrow 1)$, $(1 \leftarrow 2)$, and $(2 \leftarrow 1)$. (e) An example of components. The entire graph consists of two components: component 1 and component 2.

expressions include the terms not only $\{R_i\}$ and $\{M_{ij}\}$ but also the mean fitness $\langle R \rangle_{\boldsymbol{\pi}}$.

Note that $\{\partial_{R_i} \langle R \rangle_{\boldsymbol{\pi}}\}$, or equivalently $\{\zeta_i \pi_i\}$, can be regarded as a probability distribution. From Eq. (10) and the positivity of π_i and ζ_i, we obtain $\zeta_i \pi_i > 0$ and $\sum_i \zeta_i \pi_i = 1$. Therefore, $\{\zeta_i \pi_i\}$ is referred to as the *ancestral distribution* [16]. Combining Eq. (11) and the positivity of $\zeta_i \pi_i$, we find $0 < \partial_{R_i} \langle R \rangle_{\boldsymbol{\pi}} < 1$ for any i. Therefore, the mean fitness in the long-term limit is monotonically increasing with respect to any increase in the reproductive rate.

III. GRAPH THEORY

In this section, we explain the graph theory for the parallel mutation–reproduction model. In particular, we newly propose a *rooted* 0/1 *loop forest*, which plays a key role in the main result.

A. Basic graph

We introduce a *basic graph*, which is a directed graph representing the population structure described in Eq. (1). The term "basic graph" is introduced in Ref. [52] for the linear master equation, and we use the same term for the parallel mutation–reproduction model as a generalization of the basic graph in Ref. [52]. The construction of a basic graph, denoted by $G = (V(G), E(G))$, is outlined as follows [see Figs. 2(a) and 2(b) for an example of the dynamics and the corresponding basic graph]. The vertex set of G, denoted by $V(G)$, is defined as the set of the traits in the population, i.e., $V(G) := \{1, \ldots, N\}$. In other words, each vertex of G represents a trait in the population. Next, we define the edge set of G, denoted by $E(G)$, as follows. Let $(i \leftarrow j)$ be a directed edge from vertex j to vertex i. In particular, $(i \leftarrow i)$ represents a loop at vertex i. Using this notation, $E(G)$ is defined as

$$E(G) := \{(i \leftarrow i) \mid i \in V(G)\} \cup \{(i \leftarrow j) \mid M_{ij} \neq 0,\ i \in V(G),\ j(\neq i) \in V(G)\}. \tag{13}$$

A loop at vertex i represents the reproduction of trait i, and a directed edge from vertex j to i represents the mutation from trait j to i. Note that if the mutation from trait j to i never occurs, i.e., $M_{ij} = 0$ holds, then $(i \leftarrow j)$ does not belong to $E(G)$.

B. Cycle

We introduce the concept of a *cycle* [56]. A cycle is a nonempty trail (and is not a loop) in which only the first and last vertices are the same. A sequence of edges $(e_1, \ldots, e_L)$ $(L \geqslant 3)$ is called a cycle if starting from a vertex, it is possible to return to the same vertex by following all the edges $(e_1, \ldots, e_L)$ sequentially without visiting the same vertex twice. Here, "following" an edge e means moving along it either in the given direction or in the opposite direction. Thus, edge directions are not taken into account when defining a cycle, and a loop is not included in a cycle. See Fig. 2(c) for an example.

C. Subgraph and component

We introduce a *subgraph*. A subgraph of a graph G is defined as a graph H such that $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$ [56], and we write $H \subseteq G$.

An *induced subgraph* is a special kind of subgraph. Given a graph G and a subset of its vertex set $V' \subseteq V(G)$, the subgraph of G induced by V', denoted by $G[V']$, is defined as the graph

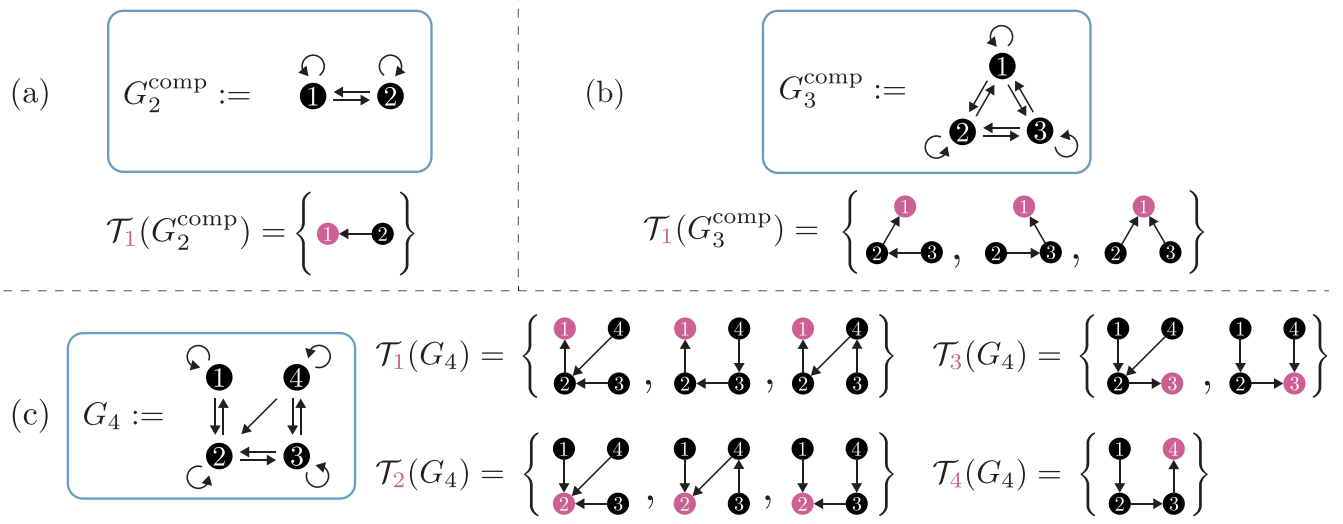

FIG. 3. Examples of basic graphs and their rooted spanning trees. (a) Complete graph on two vertices with loops G_2^{comp} and $\mathcal{T}_1(G_2^{\mathrm{comp}})$. (b) Complete graph on three vertices with loops G_3^{comp} and $\mathcal{T}_1(G_3^{\mathrm{comp}})$. (c) Graph on four vertices with loops G_4, which is not complete. We show all sets of rooted spanning trees.

whose vertex set is V' and whose edge set consists of all edges in $E(G)$ whose endpoints are both in V'. In other words, $G[V']$ is the subgraph obtained from G by removing all vertices in $V(G) \setminus V'$ and all edges with at least one endpoint in $V(G) \setminus V'$. See Fig. 2(d) for an example of an induced subgraph.

A *component* is also a special kind of subgraph. A component of a graph G is a connected subgraph $H \subseteq G$ that is not part of any larger connected subgraph [57] [see Fig. 2(e) for an example]. Roughly speaking, a component of G is an *island* in G.

D. Union of graphs

We explain the union of graphs. Suppose that there are M graphs $H^{(\alpha)}$, $\alpha \in \{1, \ldots, M\}$. The *union* of $\{H^{(\alpha)}\}_\alpha$ is defined as the graph with the union of the vertices and the union of the edges,

$$\bigcup_{\alpha=1}^{M} H^{(\alpha)} := \left(\bigcup_{\alpha=1}^{M} V(H^{(\alpha)}), \bigcup_{\alpha=1}^{M} E(H^{(\alpha)}) \right). \tag{14}$$

We also use the notation $H^{(1)} \cup H^{(2)}$ for the union of $H^{(1)}$ and $H^{(2)}$, which is defined as $H^{(1)} \cup H^{(2)} := (V(H^{(1)}) \cup V(H^{(2)}), E(H^{(1)}) \cup E(H^{(2)}))$. Moreover, when a graph H consists of only one edge e, we sometimes identify H with e. Thus, we also use the notation $(i \leftarrow j) \cup (k \leftarrow l)$ for the union of graphs expressed by the edges $(i \leftarrow j)$ and $(k \leftarrow l)$. For example, $\bigcup_{j(\neq i)}(j \leftarrow j)$ represents the graph that has a loop at every vertex except vertex i.

E. Rooted spanning tree

We explain *rooted spanning trees*, which play a key role in the analysis of the linear master equation. Given a basic graph G and its subgraph H, we call H a spanning tree of G rooted at vertex i if H satisfies all of the following conditions [52]:

(1) H contains no cycles.
(2) H contains no loops.
(3) $V(H) = V(G)$.
(4) H is connected.
(5) All edges in H are directed to vertex i.

If H satisfies all of the above conditions except the fifth condition, then we call H a *spanning tree*. A graph H is simply called a *rooted spanning tree* if it is a spanning tree rooted at a certain vertex. Note that the first and fourth conditions guarantee the existence and uniqueness of the path between any two vertices in H, which allows us to determine the direction of each edge in the fifth condition. In addition to the first, second, third, and fifth conditions above, it is known that a rooted spanning tree can be characterized by the following condition [57]:

(6) H has $|V(G)| - 1$ edges,

where $|V(G)|$ denotes the number of elements in $V(G)$.

Let $\mathcal{T}_i(G)$ denote the set of all spanning trees of G rooted at vertex i. For any $H \in \mathcal{T}_i(G)$, we call vertex i the root of H. See Fig. 3 for examples of rooted spanning trees. Note that the set $\mathcal{T}_i(G)$ is used for a diagrammatic expression for the steady-state distribution in the linear master equation [52].

F. Rooted 0/1 loop forest

We propose a class of graphs called *rooted 0/1 loop forests*, which play a central role in this paper. Given a basic graph G and its subgraph H, H is called a *0/1 loop forest of G rooted at vertex i* if H satisfies all of the following conditions:

(1) H contains no cycles.
(2) $V(H) = V(G)$.
(3) In each component of H that does not contain vertex i, there is exactly one loop.
(4) In the component of H that contains vertex i, there are no loops.
(5) In the component of H that contains vertex i, all edges are directed to vertex i. In each component of H that does not contain vertex i, all edges are directed to the vertex with the loop.

If H satisfies all of the above conditions except the fifth condition for a certain vertex i, we call H a *0/1 loop forest*.

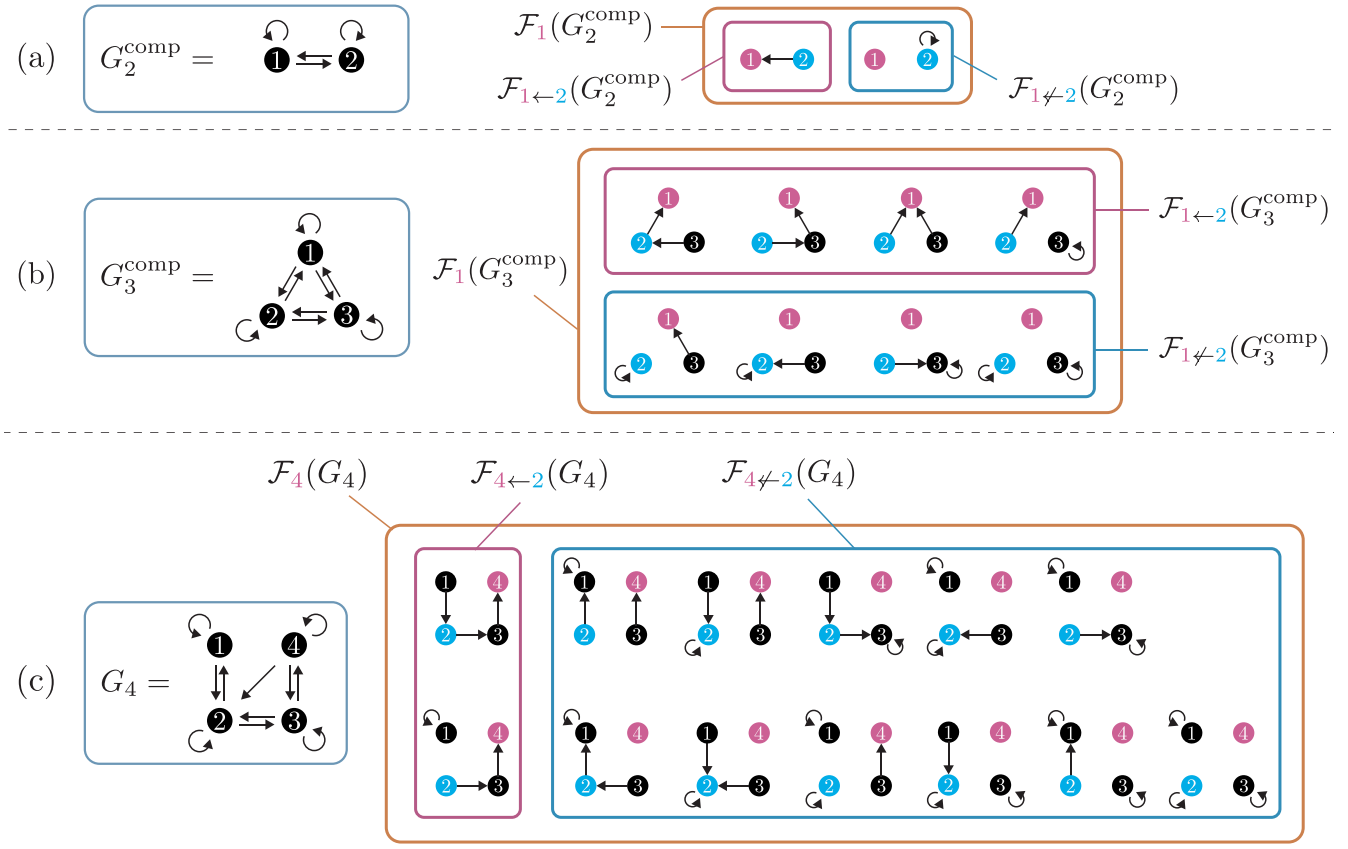

FIG. 4. Examples of rooted 0/1 loop forests. (a) $\mathcal{F}_1(G_2^{\mathrm{comp}})$, $\mathcal{F}_{1\leftarrow 2}(G_2^{\mathrm{comp}})$, and $\mathcal{F}_{1\not\leftarrow 2}(G_2^{\mathrm{comp}})$. (b) $\mathcal{F}_1(G_3^{\mathrm{comp}})$, $\mathcal{F}_{1\leftarrow 2}(G_3^{\mathrm{comp}})$, and $\mathcal{F}_{1\not\leftarrow 2}(G_3^{\mathrm{comp}})$. (c) $\mathcal{F}_4(G_4)$, $\mathcal{F}_{4\leftarrow 2}(G_4)$, and $\mathcal{F}_{4\not\leftarrow 2}(G_4)$. We can check that any rooted 0/1 loop forest of graph G has $|V(G)|-1$ edges, and $\mathcal{F}_i(G)=\mathcal{F}_{i\leftarrow j}(G)\cup\mathcal{F}_{i\not\leftarrow j}(G)$ and $\mathcal{F}_{i\leftarrow j}(G)\cap\mathcal{F}_{i\not\leftarrow j}(G)=\emptyset$ are satisfied.

Thus, a 0/1 loop forest consists of one component with no loops and N_c-1 components with one loop each where N_c is the total number of components. In graph theory, a *forest* refers to a disjoint union of trees with no loops. A 0/1 loop forest is obtained by adding a loop to each component of a forest, except for one component. For a 0/1 loop forest, the root must be chosen from a vertex of a component that has no loops. A graph H is also simply called a *rooted 0/1 loop forest* if it is a 0/1 loop forest rooted at a certain vertex. Note that the first and third conditions allow us to determine the directions of the edges in the fifth condition.

Let $\mathcal{F}_i(G)$ denote the set of all 0/1 loop forests of G rooted at vertex i. For any $H\in\mathcal{F}_i(G)$, we call vertex i the root of H. We call $H\in\mathcal{F}_i(G)$ an (i,j)-connected 0/1 loop forest of G rooted at vertex i if H satisfies the following condition:

(6) Vertex i and vertex j are connected in H.

We call $H\in\mathcal{F}_i(G)$ an (i,j)-disconnected 0/1 loop forest of G rooted at vertex i if H satisfies the following condition:

(7) Vertex i and vertex j are not connected in H.

We define two subsets of $\mathcal{F}_i(G)$, $\mathcal{F}_{i\leftarrow j}(G)$ and $\mathcal{F}_{i\not\leftarrow j}(G)$, as follows. Let $\mathcal{F}_{i\leftarrow j}(G)$ denote the set of all (i,j)-connected 0/1 loop forests of G rooted at vertex i. Let $\mathcal{F}_{i\not\leftarrow j}(G)$ denote the set of all (i,j)-disconnected 0/1 loop forests of G rooted at vertex i. Because all graphs in $\mathcal{F}_i(G)$ can be divided into two subsets based on whether vertex i and vertex j are connected or not, $\mathcal{F}_i(G)=\mathcal{F}_{i\leftarrow j}(G)\cup\mathcal{F}_{i\not\leftarrow j}(G)$ and $\mathcal{F}_{i\leftarrow j}(G)\cap\mathcal{F}_{i\not\leftarrow j}(G)=\emptyset$ are satisfied. We regard $\mathcal{F}_{i\leftarrow i}(G)$ and $\mathcal{F}_{i\not\leftarrow i}(G)$ as $\mathcal{F}_{i\leftarrow i}(G)=\mathcal{F}_i(G)$ and $\mathcal{F}_{i\not\leftarrow i}(G)=\emptyset$, respectively. This is because we consider vertex i and vertex i to be trivially connected.

For any basic graph, we can construct all rooted 0/1 loop forests as discussed in Appendix E. See also Fig. 4 for examples of rooted 0/1 loop forests for specific graphs.

We consider the definition of rooted 0/1 loop forests in a constructive way. Suppose that a rooted 0/1 loop forest $H\in\mathcal{F}_i(G)$ is given. Let N_c be the number of components of H, $V(H^{(\alpha)})$ be the vertex set of the αth component, and $H^{(1)}$ be the component of H that contains vertex i. $H^{(1)}$ is a spanning tree of the induced subgraph $G[V(H^{(1)})]$ rooted at vertex i because $H^{(1)}$ contains neither cycles nor loops, is connected, and all its edges are directed to vertex i. Thus, $H=H^{(1)}$ when $N_c=1$. For any $\alpha\in\{2,\dots,N_c\}$ ($N_c\geqslant 2$), the αth component contains a rooted spanning tree of the induced subgraph $G[V(H^{(\alpha)})]$, denoted by $H^{(\alpha)}$. Each αth component has the unique loop at its root, denoted by $r^{(\alpha)}$. Therefore, when $N_c\geqslant 2$, we can rewrite H as

$$H=H^{(1)}\cup\bigcup_{\alpha=2}^{N_c}\left(H^{(\alpha)}\cup\left(r^{(\alpha)}\leftarrow r^{(\alpha)}\right)\right). \tag{15}$$

Based on this expression, we can show that any rooted 0/1 loop forest of G has $|V(G)|-1$ edges because $\bigcup_{\alpha=1}^{N_c}H^{(\alpha)}$ has $|V(G)|-N_c$ edges and $\bigcup_{\alpha=2}^{N_c}(r^{(\alpha)}\leftarrow r^{(\alpha)})$ has N_c-1 edges.

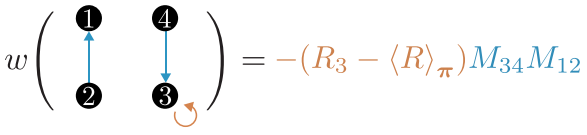

FIG. 5. An example of the weight of a graph. We consider a graph with two directed edges and one loop: $(1 \leftarrow 2)$, $(3 \leftarrow 4)$, and $(3 \leftarrow 3)$. The weights of the edges are given by $\tilde{w}((1 \leftarrow 2)) = M_{12}$, $\tilde{w}((3 \leftarrow 4)) = M_{34}$, and $\tilde{w}((3 \leftarrow 3)) = -(R_3 - \langle R \rangle_{\boldsymbol{\pi}})$. Therefore, the weight of this graph is the product of these values: $-(R_3 - \langle R \rangle_{\boldsymbol{\pi}})M_{34}M_{12}$.

We state a relationship between rooted spanning trees and rooted 0/1 loop forests. $\mathcal{F}_{i \leftarrow j}(G)$ contains all spanning trees of G rooted at vertex i, i.e., $\mathcal{T}_i(G) \subseteq \mathcal{F}_{i \leftarrow j}(G)$ for any vertex i and vertex j. This is because $H \in \mathcal{T}_i(G)$ can be regarded as the rooted 0/1 loop forest for $N_c = 1$, and any vertex i and vertex j are connected for any rooted spanning tree $H \in \mathcal{T}_i(G)$. For example, we can check $\mathcal{T}_i(G) \subseteq \mathcal{F}_{i \leftarrow j}(G)$ by comparing Figs. 3 and 4.

Rooted 0/1 loop forests can be considered as a natural generalization of rooted spanning trees. Each rooted spanning tree represents the influence of the other vertices on a given root. The parallel mutation–reproduction model involves a contribution of reproduction, which can influence the distribution at any vertex. Therefore, each loop can be regarded as representing the influence of the vertex with the loop on any vertex. Consequently, each rooted 0/1 loop forest also represents the influence of the other states on a given root within the parallel mutation–reproduction model.

G. Weight of a graph

We introduce the *weight* of a graph which is applicable to a rooted 0/1 loop forest. This weight of a graph is a generalization of the weight of a rooted spanning tree, which provides a diagrammatic expression for the steady-state distribution for the linear master equation [51,52]. In the main results, we use this weight to obtain diagrammatic expressions for the steady-state distribution and the static responses in the parallel mutation–reproduction model.

We define the weight of a graph H as

$$w(H) := \prod_{e \in E(H)} \tilde{w}(e), \tag{16}$$

where $\tilde{w}(e)$ is the weight of an edge e defined as

$$\tilde{w}(e) := \begin{cases} -(R_i - \langle R \rangle_{\boldsymbol{\pi}}), & \text{if } e = (i \leftarrow i), \\ M_{ij}, & \text{if } e = (i \leftarrow j),\ i \neq j. \end{cases} \tag{17}$$

Thus, the weight of a graph H is given by the product of the weights of all its edges $E(H)$. See Fig. 5 for an example.

IV. MAIN RESULTS

We present the main results, which are diagrammatic expressions for the steady-state distribution and the static responses, and illustrate them with some examples.

A. Generalization of the Markov chain tree theorem

We first state a lemma that plays an important role in the derivation of the main results. The lemma is a generalization of the Markov chain tree theorem [51,52], and provides a diagrammatic expression for $\zeta_j \pi_i$. As discussed in Eqs. (11) and (12), $\zeta_j \pi_i$ characterizes the static responses of the mean fitness.

Lemma 1. Suppose that a basic graph G is given. Then, for any $i \in V(G)$ and any $j \in V(G)$, $\zeta_j \pi_i$ is given by

$$\zeta_j \pi_i = \frac{1}{Z} \sum_{H \in \mathcal{F}_{i \leftarrow j}(G)} w(H), \tag{18}$$

where $Z := \sum_{i \in V(G)} \sum_{H \in \mathcal{F}_i(G)} w(H)$.

See Appendix F for the proof of this lemma.

As shown in Appendix G, the number of elements in $\mathcal{F}_i(G)$, $|\mathcal{F}_i(G)|$, is at most $2(N+1)^{N-2}$ with $N = |V(G)|$. Thus, the number of terms appearing in $\sum_{H \in \mathcal{F}_{i \leftarrow j}(G)} w(H)$ is at most $2(N+1)^{N-2}$ because $|\mathcal{F}_i(G)| \geqslant |\mathcal{F}_{i \leftarrow j}(G)|$. The number of terms appearing in Z is at most $2N(N+1)^{N-2}$.

When all reproductive rates are equal ($\mathsf{R} = \langle R \rangle_{\boldsymbol{p}(t)} \mathsf{I}$), Lemma 1 reduces to the Markov chain tree theorem as follows.

Markov chain tree theorem. Suppose that a basic graph G is given and $R_i = R_j$ holds for any $i \in V(G)$ and any $j \in V(G)$. Then, for any $i \in V(G)$, the steady-state distribution π_i is given by

$$\pi_i = \frac{1}{Z'} \sum_{H \in \mathcal{T}_i(G)} w(H), \tag{19}$$

where $Z' := \sum_{i \in V(G)} \sum_{H \in \mathcal{T}_i(G)} w(H)$.

This theorem is proved from Lemma 1 as follows. When $R_i = R_j$ holds for any pairs (i, j), $\boldsymbol{\zeta} = \mathbf{1}$ holds as mentioned in Sec. II B. Thus, the left-hand side of Eq. (18) reduces to π_i. In this case, $R_i - \langle R \rangle_{\boldsymbol{\pi}} = 0$ also holds for any i. Therefore, if $H \in \mathcal{F}_{i \leftarrow j}(G)$ has at least one loop, the weight of H vanishes, i.e., $w(H) = 0$. Only the graphs without loops contribute to the right-hand side of Eq. (18), and such graph H in $\mathcal{F}_i(G)$ or $\mathcal{F}_{i \leftarrow j}(G) (\subseteq \mathcal{F}_i(G))$ is regarded as a rooted spanning tree $H \in \mathcal{T}_i(G)$ because Eq. (15) implies that a graph $H \in \mathcal{F}_i(G)$ should be a rooted spanning tree if $N_c = 1$, otherwise the graph should contain at least one loop. Thus, we obtain $\sum_{H \in \mathcal{F}_{i \leftarrow j}(G)} w(H) = \sum_{H \in \mathcal{T}_i(G)} w(H)$ and $Z = Z'$. Thus, Lemma 1 [Eq. (18)] reduces to the Markov chain tree theorem [Eq. (19)].

B. Diagrammatic expressions for the static responses

We derive diagrammatic expressions for $\partial_{R_i} \langle R \rangle_{\boldsymbol{\pi}}$ and $\partial_{M_{ij}} \langle R \rangle_{\boldsymbol{\pi}}$ from Lemma 1. First, we discuss a diagrammatic expression for $\partial_{R_i} \langle R \rangle_{\boldsymbol{\pi}}$:

Theorem 1. Suppose that a basic graph G is given. Then, for any $i \in V(G)$, the static response of the mean fitness to a perturbation in the reproductive rate is given by

$$\frac{\partial \langle R \rangle_{\boldsymbol{\pi}}}{\partial R_i} = \frac{1}{Z} \sum_{H \in \mathcal{F}_i(G)} w(H), \tag{20}$$

where $Z = \sum_{i \in V(G)} \sum_{H \in \mathcal{F}_i(G)} w(H)$.

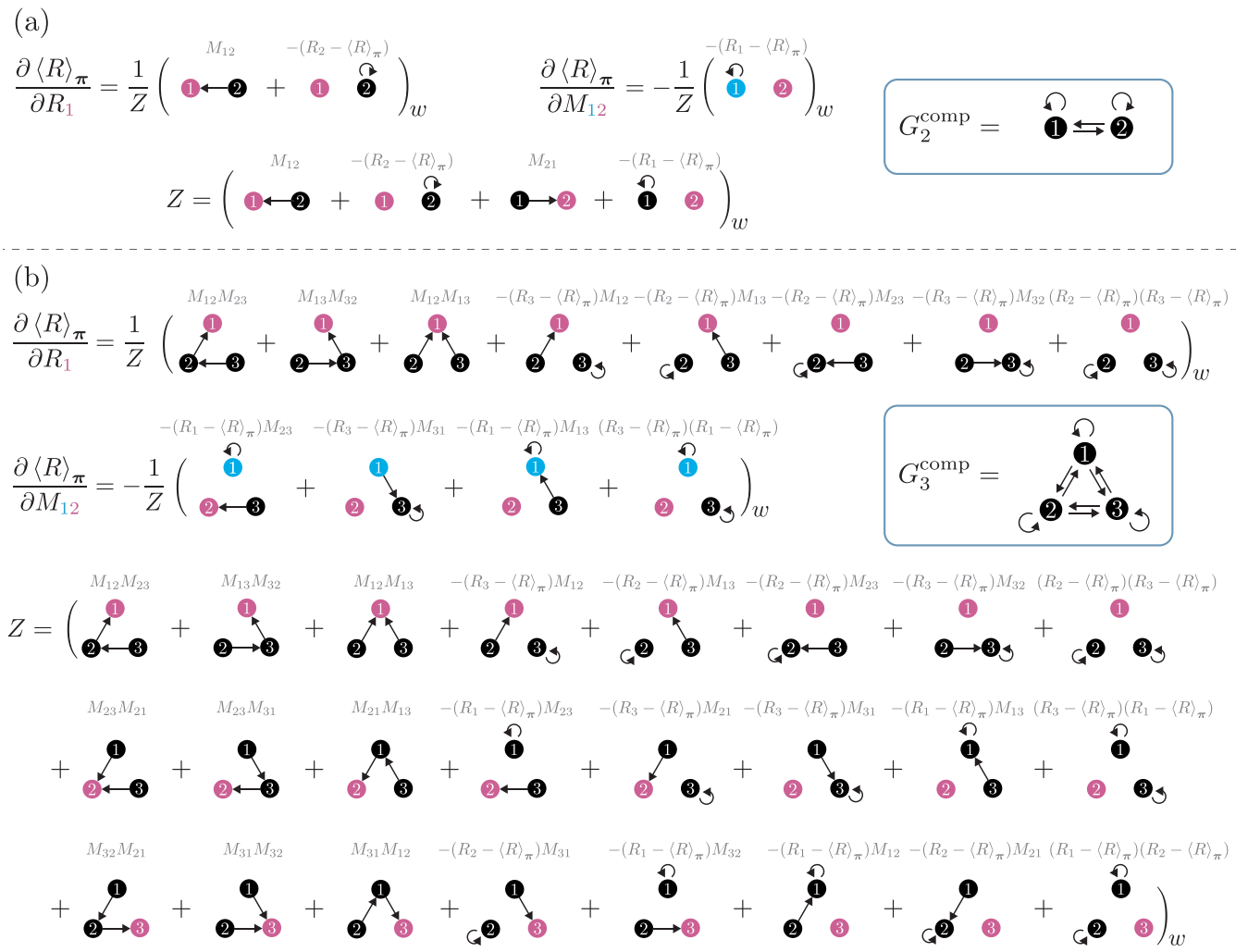

FIG. 6. Examples of Theorems 1 and 2. In this figure, we use the notation $(H + H' + \cdots)_w := w(H) + w(H') + \cdots$. The weight of each graph is shown in gray. (a) Diagrammatic expressions for $\partial_{R_1}\langle R\rangle_{\boldsymbol{\pi}} = \sum_{H\in\mathcal{F}_1(G_2^{\mathrm{comp}})} w(H)/Z$ and $\partial_{M_{12}}\langle R\rangle_{\boldsymbol{\pi}} = -\sum_{H\in\mathcal{F}_{2\leftarrow 1}(G_2^{\mathrm{comp}})} w(H)/Z$ with $Z = \sum_{i=1}^{2}\sum_{H\in\mathcal{F}_i(G_2^{\mathrm{comp}})} w(H)$. (b) Diagrammatic expressions for $\partial_{R_1}\langle R\rangle_{\boldsymbol{\pi}} = \sum_{H\in\mathcal{F}_1(G_3^{\mathrm{comp}})} w(H)/Z$ and $\partial_{M_{12}}\langle R\rangle_{\boldsymbol{\pi}} = -\sum_{H\in\mathcal{F}_{2\leftarrow 1}(G_3^{\mathrm{comp}})} w(H)/Z$ with $Z = \sum_{i=1}^{3}\sum_{H\in\mathcal{F}_i(G_3^{\mathrm{comp}})} w(H)$.

To prove this theorem, we consider $j = i$ in Eq. (18). The left-hand side of Eq. (18) becomes $\zeta_i\pi_i = \partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}}$ as discussed in Eq. (11). The right-hand side coincides with $Z^{-1}\sum_{H\in\mathcal{F}_i(G)} w(H)$ because $\mathcal{F}_{i\leftarrow i}(G) = \mathcal{F}_i(G)$.

This theorem provides an expression for $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}}$ by $\{R_i\}$, $\{M_{ij}\}$, and $\langle R\rangle_{\boldsymbol{\pi}}$. In other words, it allows us to predict the static response of the mean fitness to a perturbation in the reproductive rate using only the physically interpretable and measurable quantities. It can be more useful than the expression $\zeta_i\pi_i$, because it is difficult to interpret the physical meaning only from ζ_i and π_i, which are given by eigenvectors of $\mathsf{R}+\mathsf{M}$. We show two simple examples of Theorem 1 in Fig. 6.

Similarly, we can also obtain a diagrammatic expression for $\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}}$:

Theorem 2. Suppose that a basic graph G is given. Then, for any $i \in V(G)$ and $j \in V(G)$, the static response of the mean fitness to a perturbation in the mutation rate is given by

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial M_{ij}} = -\frac{1}{Z}\sum_{H\in\mathcal{F}_{j\not\leftarrow i}(G)} w(H), \tag{21}$$

where $Z = \sum_{i\in V(G)}\sum_{H\in\mathcal{F}_i(G)} w(H)$.

To prove this theorem, we consider Eqs. (12) and (18). From Eqs. (12) and (18), we obtain

$$\begin{aligned}\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial M_{ij}} &= \frac{1}{Z}\left(\sum_{H\in\mathcal{F}_{j\leftarrow i}(G)} w(H) - \sum_{H\in\mathcal{F}_j(G)} w(H)\right)\\ &= -\frac{1}{Z}\sum_{H\in\mathcal{F}_{j\not\leftarrow i}(G)} w(H),\end{aligned} \tag{22}$$

where we used $\mathcal{F}_j(G) = \mathcal{F}_{j\leftarrow i}(G) \cup \mathcal{F}_{j\not\leftarrow i}(G)$ and $\mathcal{F}_{j\leftarrow i}(G) \cap \mathcal{F}_{j\not\leftarrow i}(G) = \emptyset$.

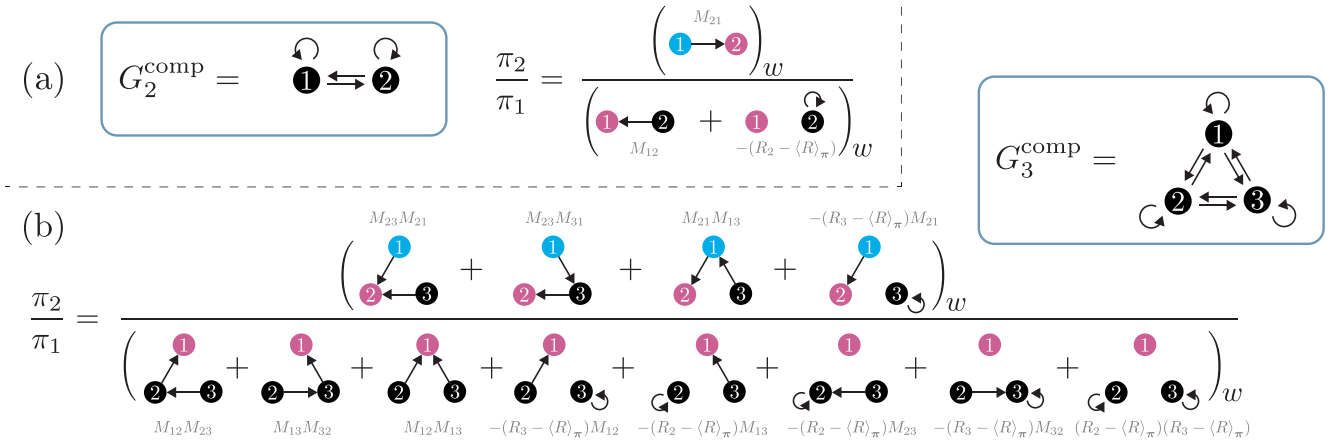

FIG. 7. Examples of Theorem 3. In this figure, we use the notation $(H + H' + \cdots)_w := w(H) + w(H') + \cdots$. The weight of each graph is shown in gray. (a) A diagrammatic expression for $\pi_2/\pi_1 = \left[\sum_{H' \in \mathcal{F}_{2\leftarrow 1}(G_2^{\mathrm{comp}})} w(H')\right] / \left[\sum_{H \in \mathcal{F}_1(G_2^{\mathrm{comp}})} w(H)\right]$. (b) A diagrammatic expression for $\pi_2/\pi_1 = \left[\sum_{H' \in \mathcal{F}_{2\leftarrow 1}(G_3^{\mathrm{comp}})} w(H')\right] / \left[\sum_{H \in \mathcal{F}_1(G_3^{\mathrm{comp}})} w(H)\right]$.

This theorem also provides an expression for $\partial_{M_{ij}} \langle R \rangle_{\boldsymbol{\pi}}$ by physically interpretable and measurable quantities. We also show two simple examples of Theorem 2 in Fig. 6.

Theorems 1 and 2 provide expressions not only for $\partial_{R_i} \langle R \rangle_{\boldsymbol{\pi}}$ and $\partial_{M_{ij}} \langle R \rangle_{\boldsymbol{\pi}}$, but also for general static responses of the mean fitness. This is because, as mentioned in Sec. II C, any static response of the mean fitness can be expressed as a linear combination of $\partial_{R_i} \langle R \rangle_{\boldsymbol{\pi}}$ and $\partial_{M_{ij}} \langle R \rangle_{\boldsymbol{\pi}}$.

C. Diagrammatic expression for the steady-state distribution

Lemma 1 also provides a diagrammatic expression for the ratio of the steady-state distribution:

Theorem 3. Suppose that a basic graph G is given. Then, for any $i \in V(G)$ and $j \in V(G)$, the ratio of the steady-state distribution is given by

$$\frac{\pi_j}{\pi_i} = \frac{\sum_{H' \in \mathcal{F}_{j\leftarrow i}(G)} w(H')}{\sum_{H \in \mathcal{F}_i(G)} w(H)}. \tag{23}$$

To prove this theorem, we give a diagrammatic expression for the ratio of the steady-state distributions π_j/π_i. Substituting Eq. (18) into the right-hand side of $\pi_j/\pi_i = (\zeta_i \pi_j)/(\zeta_i \pi_i)$, we obtain Eq. (23) because $\mathcal{F}_{i\leftarrow i}(G) = \mathcal{F}_i(G)$.

Moreover, we can obtain a diagrammatic expression for $\boldsymbol{\pi}$ itself:

Corollary 1. Suppose that a basic graph G is given. Then, for any $i \in V(G)$, the steady-state distribution is given by

$$\pi_i = \frac{\sum_{H' \in \mathcal{F}_i(G)} w(H')}{\sum_{j \in V(G)} \sum_{H \in \mathcal{F}_{j\leftarrow i}(G)} w(H)}. \tag{24}$$

To prove this corollary, we substitute Eq. (23) into the right-hand side of $\pi_i = \{\sum_{j \in V(G)} (\pi_j/\pi_i)\}^{-1}$.

This corollary also provides an expression for the steady-state distribution by physically interpretable and measurable quantities. See Fig. 7 for some examples of Theorem 3.

V. APPLICATIONS FOR TWO LIMITING CASES

While our results provide exact expressions for the static responses and steady-state distributions, the computation with counting graphs can be complicated. Here, we consider approximations for special situations where either mutation or natural selection is dominant.

A. Mutation-dominant case

We consider the case where the effect of mutation is dominant compared to that of natural selection, and we derive approximate expressions for the steady-state distribution and the static responses of the mean fitness. Let ϵ be a small parameter satisfying $0 < \epsilon \ll \min_{i,j(\neq i)|M_{ij}\neq 0} M_{ij}$. Then, we assume $\max_{i,j} |R_i - R_j| = O(\epsilon)$, where $O(\cdot)$ denotes big O notation. In other words, we assume that mutation rates are much higher than the differences in reproductive rates, or that all reproductive rates are nearly equal; that is, we assume neutral evolution. We call this the mutation-dominant case.

We rewrite this condition so that it is suitable for our diagrammatic expressions. Since $|R_i - \langle R \rangle_{\boldsymbol{\pi}}| \leqslant \max_{j,k} |R_j - R_k|$ holds, we obtain

$$\frac{|R_k - \langle R \rangle_{\boldsymbol{\pi}}|}{M_{ij}} = O(\epsilon) \tag{25}$$

for any i, $j(\neq i)$, and k. This means that the weight of a directed edge between vertices dominates the weight of a loop.

By applying Eq. (25), we approximate $\boldsymbol{\pi}$ in Corollary 1. Let $\mathcal{F}_i^{[n_l]}(G) (\subseteq \mathcal{F}_i(G))$ and $\mathcal{F}_{i\leftarrow j}^{[n_l]}(G) (\subseteq \mathcal{F}_{i\leftarrow j}(G))$ be the sets of graphs in $\mathcal{F}_i(G)$ and $\mathcal{F}_{i\leftarrow j}(G)$ that contain exactly n_l loops, respectively. Equation (25) implies that the fewer loops in H, the heavier the weight $w(H)$. Thus, the leading and second-leading terms in the numerator of π_i are $\sum_{H \in \mathcal{F}_i^{[0]}(G)} w(H) = \sum_{H \in \mathcal{T}_i(G)} w(H)$ and $\sum_{H \in \mathcal{F}_i^{[1]}(G)} w(H)$, respectively. The leading and second-leading terms in the denominator of π_i are $\sum_{j \in V(G)} \sum_{H \in \mathcal{T}_j(G)} w(H) = Z'$ and

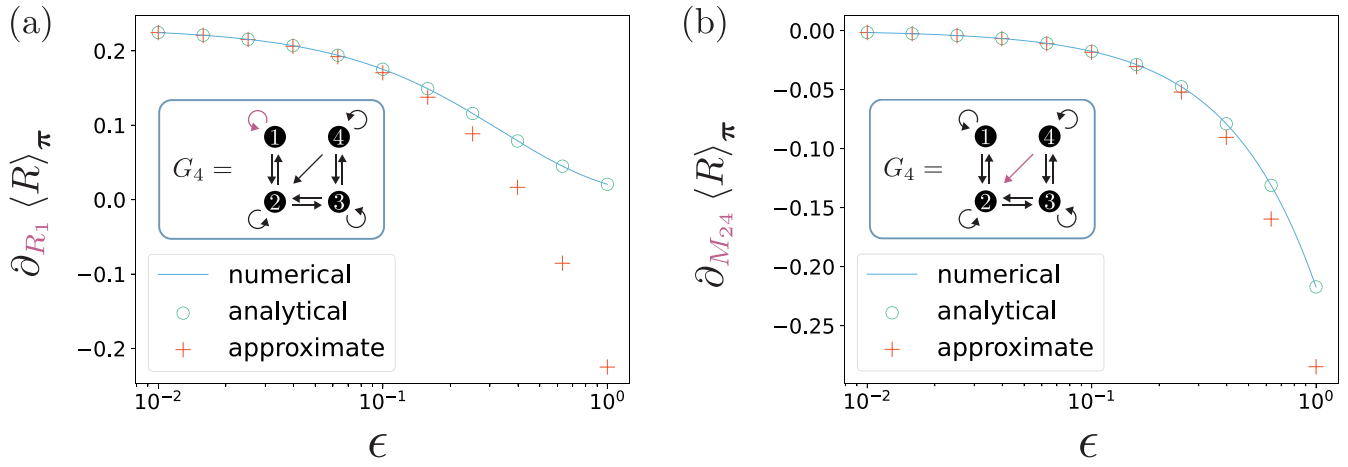

FIG. 8. Numerical, analytical, and approximate evaluations of (a) $\partial_{R_1}\langle R\rangle_{\boldsymbol{\pi}}$ and (b) $\partial_{M_{24}}\langle R\rangle_{\boldsymbol{\pi}}$. The basic graph of the population is given by G_4, shown in the inset. The matrices R and M are given in Eq. (29).

$\sum_{j\in V(G)}\sum_{H\in\mathcal{F}^{[1]}_{j\leftarrow i}(G)} w(H)$, respectively. Thus, we obtain

$$\pi_i = \frac{\sum_{H\in\mathcal{T}_i(G)} w(H)}{Z'} + \frac{\sum_{H\in\mathcal{F}^{[1]}_i(G)} w(H)}{Z'} - \frac{\sum_{j\in V(G)}\sum_{H\in\mathcal{F}^{[1]}_{j\leftarrow i}(G)}\sum_{H'\in\mathcal{T}_i(G)} w(H)w(H')}{(Z')^2} + O(\epsilon^2). \tag{26}$$

The leading term of π_i is $\sum_{H\in\mathcal{T}_i(G)} w(H)/Z' = O(1)$, which is consistent with the Markov chain tree theorem [Eq. (19)].

We also approximate $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}} = \sum_{H\in\mathcal{F}_i(G)} w(H)/Z$ (Theorem 1) in the mutation-dominant case. The leading and second-leading terms in the numerator of $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}}$ are $\sum_{H\in\mathcal{T}_i(G)} w(H)$ and $\sum_{H\in\mathcal{F}^{[1]}_i(G)} w(H)$, respectively. The leading and second-leading terms in Z are $\sum_{i\in V(G)}\sum_{H\in\mathcal{T}_i(G)} w(H) = Z'$ and $\sum_{i\in V(G)}\sum_{H\in\mathcal{F}^{[1]}_i(G)} w(H)$, respectively. Therefore, we have

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial R_i} = \frac{\sum_{H\in\mathcal{T}_i(G)} w(H)}{Z'} + \frac{\sum_{H\in\mathcal{F}^{[1]}_i(G)} w(H)}{Z'} - \frac{\sum_{j\in V(G)}\sum_{H\in\mathcal{F}^{[1]}_j(G)}\sum_{H'\in\mathcal{T}_i(G)} w(H)w(H')}{(Z')^2} + O(\epsilon^2). \tag{27}$$

Note that the leading term of $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}}$ is $\sum_{H\in\mathcal{T}_i(G)} w(H)/Z' = O(1)$, which is identical to that of π_i.

If $R_i = \langle R\rangle_{\boldsymbol{\pi}}$ is satisfied for any i and Eq. (3) is considered to be the linear master equation, $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}} = \pi_i$ holds exactly because the weight of any graph containing at least one loop vanishes, as mentioned below Eq. (19). This fact can also be understood as the situation in which $\zeta_i = 1$ for any i in Eq. (11). Indeed, the left eigenvector for the linear master equation is $\boldsymbol{\zeta} = \mathbf{1}$, and $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}} = \pi_i$ can be confirmed.

Similarly, we discuss an approximate expression for $\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}} = -\sum_{H\in\mathcal{F}_{j\leftarrow i}(G)} w(H)/Z$ (Theorem 2). Let $\mathcal{F}^{[n_1]}_{i\leftarrow j}(G)(\subseteq\mathcal{F}_{i\leftarrow j}(G))$ be the set of graphs in $\mathcal{F}_{i\leftarrow j}(G)$ that contain exactly n_1 loops. Since any graph in $\mathcal{F}_{j\leftarrow i}(G)$ contains at least one loop, the leading term in $\sum_{H\in\mathcal{F}_{j\leftarrow i}(G)} w(H)$ is $\sum_{H\in\mathcal{F}^{[1]}_{j\leftarrow i}(G)} w(H)$. Thus, we obtain

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial M_{ij}} = -\frac{1}{Z'}\sum_{H\in\mathcal{F}^{[1]}_{j\leftarrow i}(G)} w(H) + O(\epsilon^2). \tag{28}$$

Note that the leading term of $\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}}$ is $O(\epsilon)$, whereas that of $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}}$ is $O(1)$.

If $R_i = \langle R\rangle_{\boldsymbol{\pi}}$ is satisfied for any i and Eq. (3) is considered to be the linear master equation, $\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}} = 0$ holds exactly. We can also confirm $\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}} = 0$ because the left eigenvector for the linear master equation is $\boldsymbol{\zeta} = \mathbf{1}$ and $\zeta_i - \zeta_j = 0$ is satisfied in Eq. (12).

In Fig. 8, we evaluate $\partial_{R_1}\langle R\rangle_{\boldsymbol{\pi}}$ and $\partial_{M_{24}}\langle R\rangle_{\boldsymbol{\pi}}$ using three different approaches: numerical calculation, the analytical results (Theorems 1 and 2), and their approximations [Eqs. (27) and (28)]. We consider a population whose basic graph G_4 is shown in Fig. 8. The matrices R and M are given by

$$\mathsf{R} = \epsilon\begin{pmatrix}1&0&0&0\\0&3&0&0\\0&0&4&0\\0&0&0&7\end{pmatrix}, \quad \mathsf{M} = \begin{pmatrix}-2&2&0&0\\2&-9&1&5\\0&7&-7&1\\0&0&6&-6\end{pmatrix}, \tag{29}$$

with $\epsilon > 0$. Since R is parametrized by ϵ, Eq. (25) is satisfied. As shown in Fig. 8, the analytical results agree with the numerical calculations in the whole range of ϵ, and the approximation is valid when $0 < \epsilon \ll 1$.

B. Selection-dominant case

Next, we consider the case where the effect of natural selection dominates that of mutation, and derive approximate expressions for the steady-state distribution and the static responses of the mean fitness. We define the "fittest" trait in the population i_* as $i_* := \mathrm{argmax}_i\{R_i\}$, and assume that this trait i_* is unique. Let ϵ be a small parameter satisfying $0 < \epsilon \ll \min_{j(\neq i_*)}\{R_{i_*} - R_j\}$. We then assume that $M_{ij} = O(\epsilon)$ for any i and $j(\neq i)$. In other words, we assume that the effect of mutation is negligible compared to the differences in reproductive rates. We call this the selection-dominant case.

We rewrite this condition so that it is suitable for our diagrammatic expressions. In our diagrammatic expressions, we use the following inequality:

$$R_i - \langle R\rangle_{\boldsymbol{\pi}} < -M_{ii}, \tag{30}$$

which holds for any i. This inequality [Eq. (30)] always holds and can be derived as follows. Since $\boldsymbol{\pi}$ is the right eigenvector of $\mathsf{R}+\mathsf{M}$ with eigenvalue $\langle R\rangle_{\boldsymbol{\pi}}$, it satisfies $(R_i + M_{ii})\pi_i + \sum_{j(\neq i)} M_{ij}\pi_j = \langle R\rangle_{\boldsymbol{\pi}}\pi_i$. Dividing both sides by $\pi_i > 0$ and rearranging yields $R_i - \langle R\rangle_{\boldsymbol{\pi}} + M_{ii} = -\sum_{j(\neq i)} M_{ij}\pi_j/\pi_i < 0$, which leads to Eq. (30). By setting $i = i_*$ in Eq. (30) and subtracting R_j from both sides, we obtain

$$\langle R\rangle_{\boldsymbol{\pi}} - R_j > R_{i_*} - R_j + M_{i_*i_*} \tag{31}$$

for any $j(\neq i_*)$. Combining this inequality with the assumption $-M_{i_*i_*} = \sum_{j(\neq i_*)} M_{ji_*} = O(\epsilon) \ll R_{i_*} - R_j$, we find $\langle R\rangle_{\boldsymbol{\pi}} - R_j > 0$ for any $j(\neq i_*)$. Moreover, combining Eq. (31) with $M_{ij} = O(\epsilon)$, we obtain

$$0 < \frac{M_{kl}}{\langle R\rangle_{\boldsymbol{\pi}} - R_j} < \frac{M_{kl}}{R_{i_*} - R_j + M_{i_*i_*}} \tag{32}$$

for any $j(\neq i_*)$, k, and $l(\neq k)$. Since $M_{i_*i_*} = O(\epsilon)$ and $M_{kl} = O(\epsilon)$, we obtain

$$\frac{M_{kl}}{\langle R\rangle_{\boldsymbol{\pi}} - R_j} = O(\epsilon) \tag{33}$$

for any $j(\neq i_*)$, k, and $l(\neq k)$. This equation means that the weight of the loop at vertex $j(\neq i_*)$ dominates the weight of a directed edge from vertex l to vertex k.

By applying Eq. (33), we approximate $\boldsymbol{\pi}$. We first approximate π_j/π_{i_*} for each $j(\neq i_*)$. From Theorem 3, we have

$$\frac{\pi_j}{\pi_{i_*}} = \frac{\sum_{H'\in\mathcal{F}_{j\leftarrow i_*}(G)} w(H')}{\sum_{H\in\mathcal{F}_{i_*}(G)} w(H)}. \tag{34}$$

Note that no graph in $\mathcal{F}_{i_*}(G)$ or $\mathcal{F}_{j\leftarrow i_*}(G)$ contains the loop at vertex i_*. Equation (33) implies that the more loops in H, the heavier the weight $w(H)$. Thus, the leading term in the denominator of Eq. (34) is $\sum_{H\in\mathcal{F}_{i_*}^{[N-1]}(G)} w(H)$ with $N = |V(G)|$, and $\mathcal{F}_{i_*}^{[N-1]}(G)$ consists of only one graph $\bigcup_{k\neq i_*}(k\leftarrow k)$. The leading term in the numerator of Eq. (34) is $\sum_{H'\in\mathcal{F}_{j\leftarrow i_*}^{[N-2]}(G)} w(H')$, and $\mathcal{F}_{j\leftarrow i_*}^{[N-2]}(G)$ also consists of only one graph $(j\leftarrow i_*)\cup\bigcup_{k\neq i_*,j}(k\leftarrow k)$. Combining these facts, the approximate form of π_j/π_{i_*} is given by

$$\frac{\pi_j}{\pi_{i_*}} = \frac{M_{ji_*}}{\langle R\rangle_{\boldsymbol{\pi}} - R_j} + O(\epsilon^2). \tag{35}$$

We can also derive an approximate expression for $\boldsymbol{\pi}$ as follows. Since $\pi_{i_*} = (1 + \sum_{j(\neq i_*)}\pi_j/\pi_{i_*})^{-1}$, we have $\pi_{i_*} = (1 + O(\epsilon))^{-1} = 1 + O(\epsilon)$. By multiplying Eq. (35) by $\pi_{i_*} = 1 + O(\epsilon)$, we obtain

$$\pi_j = \frac{M_{ji_*}}{\langle R\rangle_{\boldsymbol{\pi}} - R_j} + O(\epsilon^2) \tag{36}$$

for any $j(\neq i_*)$. This equation implies that, if $M_{ji_*} \neq 0$, it follows that $\pi_j = O(\epsilon)$; otherwise, $\pi_j = O(\epsilon^2)$. Combining Eq. (36) with $\pi_{i_*} = 1 - \sum_{j(\neq i_*)}\pi_j$, we also obtain

$$\pi_{i_*} = 1 - \sum_{j(\neq i_*)} \frac{M_{ji_*}}{\langle R\rangle_{\boldsymbol{\pi}} - R_j} + O(\epsilon^2). \tag{37}$$

We also discuss $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}} = \sum_{H\in\mathcal{F}_i(G)} w(H)/Z$ (Theorem 1). First, we consider $\partial_{R_j}\langle R\rangle_{\boldsymbol{\pi}}$ with $j \neq i_*$. Since some graphs in $\mathcal{F}_j(G)$ contain the loop at vertex i_*, we need to consider $\langle R\rangle_{\boldsymbol{\pi}} - R_{i_*}$. By setting $i = i_*$ in Eq. (30) and using $M_{i_*i_*} = O(\epsilon)$ and $R_{i_*} - \langle R\rangle_{\boldsymbol{\pi}} > 0$, we find $\langle R\rangle_{\boldsymbol{\pi}} - R_{i_*} = O(\epsilon)$. Therefore, the leading terms in $\sum_{H\in\mathcal{F}_j(G)} w(H)$ are $\prod_{k(\neq j)}(\langle R\rangle_{\boldsymbol{\pi}} - R_k)$ and $M_{ki_*}\prod_{l(\neq i_*,j)}(\langle R\rangle_{\boldsymbol{\pi}} - R_l)$ with $k(\neq i_*)$. These terms are $O(\epsilon)$ since they contain $\langle R\rangle_{\boldsymbol{\pi}} - R_{i_*}$ or M_{ki_*}. However, their sum is $O(\epsilon^2)$ because

$$\langle R\rangle_{\boldsymbol{\pi}} - R_{i_*} + \sum_{k(\neq i_*)} M_{ki_*} = O(\epsilon^2), \tag{38}$$

which can be derived by combining

$$R_{i_*}\pi_{i_*} + \sum_{k(\neq i_*)} (M_{i_*k}\pi_k - M_{ki_*}\pi_{i_*}) = \langle R\rangle_{\boldsymbol{\pi}}\pi_{i_*} \tag{39}$$

with $\pi_{i_*} = 1 - O(\epsilon)$, $\pi_k = O(\epsilon)$, and $M_{i_*k} = O(\epsilon)$ for any $k(\neq i_*)$. The leading term in $Z = \sum_{i\in V(G)}\sum_{H\in\mathcal{F}_i(G)} w(H)$ is $\prod_{k(\neq i_*)}(\langle R\rangle_{\boldsymbol{\pi}} - R_k) = O(1)$. Therefore, we find

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial R_j} = O(\epsilon^2) \tag{40}$$

for any $j(\neq i_*)$. Combining this equation with $\partial_{R_{i_*}}\langle R\rangle_{\boldsymbol{\pi}} = 1 - \sum_{j(\neq i_*)}\partial_{R_j}\langle R\rangle_{\boldsymbol{\pi}}$, we also find

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial R_{i_*}} = 1 - O(\epsilon^2). \tag{41}$$

We can also approximate $\partial_{M_{ij}}\langle R\rangle_{\boldsymbol{\pi}} = -\sum_{H\in\mathcal{F}_{j\leftarrow i}(G)} w(H)/Z$ (Theorem 2). First, we consider $\partial_{M_{jk}}\langle R\rangle_{\boldsymbol{\pi}}$ with $j \neq i_*$ and $k \neq i_*, j$. The dominant terms in $\sum_{H\in\mathcal{F}_{k\leftarrow j}(G)} w(H)$ are $\prod_{l(\neq k)}(\langle R\rangle_{\boldsymbol{\pi}} - R_l)$ and $M_{li_*}\prod_{m(\neq i_*,k)}(\langle R\rangle_{\boldsymbol{\pi}} - R_m)$ with $l \neq i_*$. Although these terms are $O(\epsilon)$, their sum is $O(\epsilon^2)$ in the same way as in Eq. (38). Thus, we find

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial M_{jk}} = O(\epsilon^2) \tag{42}$$

for $j \neq i_*$ and $k \neq j, i_*$. Next, we consider $\partial_{M_{i_*j}}\langle R\rangle_{\boldsymbol{\pi}}$ with $j \neq i_*$. The leading terms in $\sum_{H\in\mathcal{F}_{j\leftarrow i_*}(G)} w(H)$ are $\prod_{k(\neq j)}(\langle R\rangle_{\boldsymbol{\pi}} - R_k)$ and $M_{ki_*}\prod_{l(\neq i_*,j)}(\langle R\rangle_{\boldsymbol{\pi}} - R_l)$ with $k(\neq i_*, j)$. The sum of these terms are

$$\left(\langle R\rangle_{\boldsymbol{\pi}} - R_{i_*} + \sum_{k(\neq i_*,j)} M_{ki_*}\right)\prod_{l(\neq i_*,j)}(\langle R\rangle_{\boldsymbol{\pi}} - R_l) = -M_{ji_*}\prod_{l(\neq i_*,j)}(\langle R\rangle_{\boldsymbol{\pi}} - R_l) + O(\epsilon^2), \tag{43}$$

where we used Eq. (38). The leading term in Z is $\prod_{k(\neq i_*)}(\langle R\rangle_{\boldsymbol{\pi}} - R_k)$. Thus, we find

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial M_{i_*j}} = \frac{M_{ji_*}}{\langle R\rangle_{\boldsymbol{\pi}} - R_j} + O(\epsilon^2). \tag{44}$$

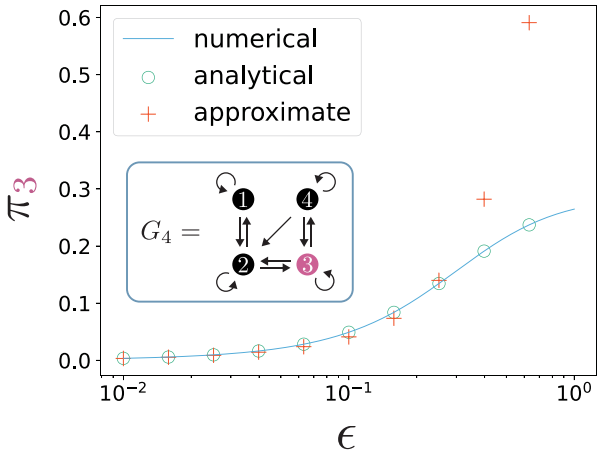

FIG. 9. Numerical, analytical, and approximate evaluations of π_3. The basic graph of the population is given by G_4, shown in the inset. The matrices R and M are given in Eq. (46).

Finally, we consider $\partial_{M_{ji_*}}\langle R\rangle_{\boldsymbol{\pi}}$ with $j \neq i_*$. The leading and second-leading terms in $\sum_{H\in\mathcal{F}_{i_*\leftarrow j}(G)} w(H)$ are $\sum_{H\in\mathcal{F}^{[N-1]}_{i_*\leftarrow j}(G)} w(H)$ and $\sum_{H\in\mathcal{F}^{[N-2]}_{i_*\leftarrow j}(G)} w(H)$ with $N = |V(G)|$, respectively. The leading and second-leading terms in Z are $\sum_{H\in\mathcal{F}^{[N-1]}_{i_*}(G)} w(H)$ and $\sum_{H\in\mathcal{F}^{[N-2]}_{i_*}(G)} w(H)$, respectively. Indeed, from Eq. (40), $\sum_{H\in\mathcal{F}_k(G)} w(H)$ in Z satisfies $\sum_{H\in\mathcal{F}_k(G)} w(H) = Z\partial_{R_k}\langle R\rangle_{\boldsymbol{\pi}} = O(\epsilon^2)$ for $k \neq i_*$. Thus, we obtain

$$\begin{aligned}\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial M_{ji_*}} &= -\frac{\sum_{H\in\mathcal{F}^{[N-1]}_{i_*\leftarrow j}(G)} w(H)+\sum_{H\in\mathcal{F}^{[N-2]}_{i_*\leftarrow j}(G)} w(H)}{\sum_{H\in\mathcal{F}^{[N-1]}_{i_*}(G)} w(H)+\sum_{H\in\mathcal{F}^{[N-2]}_{i_*}(G)} w(H)}+O(\epsilon^2)\\ &= -1+\frac{\sum_{H\in\mathcal{F}^{[N-2]}_{i_*}(G)\setminus\mathcal{F}^{[N-2]}_{i_*\leftarrow j}(G)} w(H)}{\sum_{H\in\mathcal{F}^{[N-1]}_{i_*}(G)} w(H)}+O(\epsilon^2)\\ &= -1+\frac{M_{i_*j}}{\langle R\rangle_{\boldsymbol{\pi}}-R_j}+O(\epsilon^2),\end{aligned} \tag{45}$$

where we used $\mathcal{F}^{[N-1]}_{i_*}(G) = \mathcal{F}^{[N-1]}_{i_*\leftarrow j}(G)$, $\mathcal{F}^{[N-2]}_{i_*\leftarrow j}(G) \subseteq \mathcal{F}^{[N-2]}_{i_*}(G)$, the fact that $\mathcal{F}^{[N-1]}_{i_*}(G)$ consists of one graph $\bigcup_{k(\neq i_*)}(k \leftarrow k)$, and the fact that $\mathcal{F}^{[N-2]}_{i_*}(G) \setminus \mathcal{F}^{[N-2]}_{i_*\leftarrow j}(G)$ consists of one graph $(i_* \leftarrow j)\cup\bigcup_{k(\neq i_*,j)}(k \leftarrow k)$.

In Fig. 9, we evaluate $\boldsymbol{\pi}$ using three different approaches: numerical calculation, the analytical result (Corollary 1), and its approximation [Eq. (36)]. We consider a population whose basic graph G_4 is shown in Fig. 9. The matrices R and M are given by

$$\mathsf{R} = \begin{pmatrix}1&0&0&0\\0&3&0&0\\0&0&4&0\\0&0&0&7\end{pmatrix}, \quad \mathsf{M} = \epsilon\begin{pmatrix}-2&2&0&0\\2&-9&1&5\\0&7&-7&1\\0&0&6&-6\end{pmatrix}, \tag{46}$$

with $\epsilon > 0$. In this example, $i_* = 4$. Since M is parameterized by ϵ, Eq. (33) is satisfied. As shown in Fig. 9, the analytical result agrees with the numerical calculation in the whole range of ϵ, and the approximation is valid when $0 < \epsilon \ll 1$.

VI. APPLICATION FOR COMBINATION THERAPIES

We demonstrate how our results can be used to control harmful populations. Specifically, we consider how to optimally reduce the mean fitness of a harmful population through *combination therapy*, in which multiple kinds of inhibitors are used simultaneously [58–60]. Combination therapies have been developed to prevent harmful populations, such as cancer cells [61,62] or Gram-negative bacteria [63], from acquiring resistance to inhibitors.

Let us consider a harmful population whose mean fitness we aim to reduce. We control the mean fitness of the population by manipulating the concentrations of two kinds of inhibitors, A and B [see Fig. 10(a)]. For cancer cells, A and B can be sabutoclax and minocycline [62], whereas for Gram-negative bacteria, they can be a β lactam and an aminoglycoside [63]. Assume that there are four traits in the population: trait 1 that is sensitive to both A and B, trait 2 that is sensitive to A but resistant to B, trait 3 that is resistant to A but sensitive to B, and trait 4 that is resistant to both A and B [see also Fig. 10(b)]. Here, "sensitive" means that the reproductive rate decreases with increasing inhibitor concentration, while "resistant" means that the reproductive rate is unaffected by inhibitor. Each reproductive rate can be modeled using the generalized Monod equation [64,65]:

$$\begin{aligned}R_1(a,b) &= \frac{\rho_1}{1+\kappa_1^A a+\kappa_1^B b}, \quad R_2(a) = \frac{\rho_2}{1+\kappa_2^A a},\\ R_3(b) &= \frac{\rho_3}{1+\kappa_3^B b}, \quad R_4 = \rho_4.\end{aligned} \tag{47}$$

Here, a and b denote the concentration of A and B, respectively. The constant ρ_i denotes the reproductive rate of trait i in the absence of both A and B, and the positive constants κ_i^A and κ_i^B quantify how strongly A and B inhibit the reproduction of trait i, respectively. We assume that there are no mutations between traits 1 and 4, or between traits 2 and 3, i.e., $M_{14} = M_{41} = M_{23} = M_{32} = 0$. We also assume that the mutation rates associated with resistance to A are much lower than those associated with B. In other words, M_{13}, M_{31}, M_{24}, and M_{42} are of order $O(\epsilon)$ with $0 < \epsilon \ll 1$. Indeed, mutation rates can differ substantially between different drugs in actual experiments [66,67]. The basic graph G corresponding to the population is shown in Fig. 10(c).

Under this scenario, we find the optimal control of a and b. Suppose that we change a and b slightly by δa and δb, respectively. Our goal is to identify the vector $\delta\boldsymbol{a}^* := (\delta a^*, \delta b^*)^\top$ that most significantly reduces the mean fitness, subject to the constraint $\sqrt{(\delta a^*)^2+(\delta b^*)^2} = 1$. The vector $\delta\boldsymbol{a}^*$ is given by

$$\begin{pmatrix}\delta a^*\\ \delta b^*\end{pmatrix} = -\mathcal{N}\begin{pmatrix}\partial_a\langle R\rangle_{\boldsymbol{\pi}}\\ \partial_b\langle R\rangle_{\boldsymbol{\pi}}\end{pmatrix}, \tag{48}$$

with $\mathcal{N} := 1/\sqrt{(\partial_a\langle R\rangle_{\boldsymbol{\pi}})^2+(\partial_b\langle R\rangle_{\boldsymbol{\pi}})^2}$. Indeed, $\delta\boldsymbol{a}^*$ given by Eq. (48) attains the following lower bound of the mean fitness change:

$$\begin{pmatrix}\delta a\\ \delta b\end{pmatrix}\cdot\begin{pmatrix}\partial_a\langle R\rangle_{\boldsymbol{\pi}}\\ \partial_b\langle R\rangle_{\boldsymbol{\pi}}\end{pmatrix} \geqslant -\sqrt{(\partial_a\langle R\rangle_{\boldsymbol{\pi}})^2+(\partial_b\langle R\rangle_{\boldsymbol{\pi}})^2}, \tag{49}$$

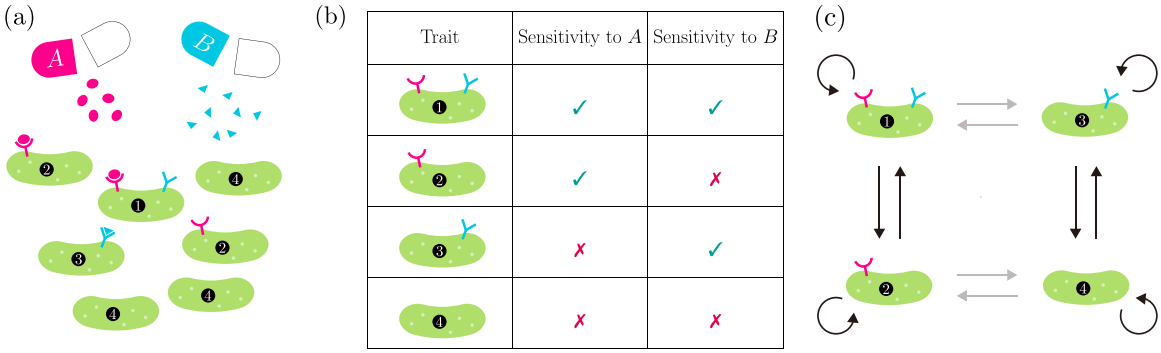

FIG. 10. (a) We manipulate the concentrations of two inhibitors, A and B, in order to reduce the mean fitness of a harmful population. When an inhibitor binds to its corresponding receptor on an individual's surface, its reproduction is inhibited. (b) There are four traits in the population that are classified by sensitivity to A and B. A trait without a receptor for A (or B) is not inhibited by A (or B). (c) Basic graph G for the population. The edges in $\mathcal{E} = \{(1 \leftarrow 3), (3 \leftarrow 1), (2 \leftarrow 4), (4 \leftarrow 2)\}$ are shown in gray to emphasize that their weights are of order $O(\epsilon)$.

where we applied the Cauchy–Schwarz inequality. Therefore, we can obtain the expressions for $\delta\boldsymbol{a}^*$:

$$\begin{aligned}\begin{pmatrix}\delta a^* \\ \delta b^*\end{pmatrix} &\propto \begin{pmatrix}\partial_a R_1 \cdot \partial_{R_1}\langle R\rangle_{\boldsymbol{\pi}} + \partial_a R_2 \cdot \partial_{R_2}\langle R\rangle_{\boldsymbol{\pi}} \\ \partial_b R_1 \cdot \partial_{R_1}\langle R\rangle_{\boldsymbol{\pi}} + \partial_b R_3 \cdot \partial_{R_3}\langle R\rangle_{\boldsymbol{\pi}}\end{pmatrix} \\ &\propto \begin{pmatrix}\partial_a R_1 \sum_{H\in\mathcal{F}_1(G)} w(H) + \partial_a R_2 \sum_{H'\in\mathcal{F}_2(G)} w(H') \\ \partial_b R_1 \sum_{H\in\mathcal{F}_1(G)} w(H) + \partial_b R_3 \sum_{H'\in\mathcal{F}_3(G)} w(H')\end{pmatrix},\end{aligned} \tag{50}$$

where we used Eq. (48) in the first line and Theorem 1 in the last line.

The diagrammatic expression in Eq. (50) allows us to reduce the number of graphs that need to be considered when calculating $\delta\boldsymbol{a}^*$. Since the edges in $\mathcal{E} := \{(1 \leftarrow 3), (3 \leftarrow 1), (2 \leftarrow 4), (4 \leftarrow 2)\}$ have much lower weights, we can obtain two approximations of $\delta\boldsymbol{a}^*$ based on how many of these edges are taken into account in Eq. (50). The first approximation, denoted by $\delta\boldsymbol{a}^{*(1)}$, is obtained from Eq. (50) by including only graphs that do not contain any edges in $\mathcal{E}$ [see Fig. 11(a) for the case of $\mathcal{F}_1(G)$]. The second approximation, denoted by $\delta\boldsymbol{a}^{*(2)}$, is obtained from Eq. (50) by including only graphs that contain at most one edge in $\mathcal{E}$ [see Fig. 11(a)]. Note that if we include graphs containing at most two edges in $\mathcal{E}$, we obtain the exact expression for $\delta\boldsymbol{a}^*$, since every graph in Eq. (50) contains at most two edges in $\mathcal{E}$ [see also Fig. 11(a)]. We emphasize that these approximations, based on the diagrammatic approach, cannot be derived from the conventional expressions for $\partial_{R_i}\langle R\rangle_{\boldsymbol{\pi}}$ [Eq. (11)].

We demonstrate that the second approximation of $\delta\boldsymbol{a}^*$, $\delta\boldsymbol{a}^{*(2)}$, is sufficiently accurate and reduces the number of graphs that need to be considered. To this end, we evaluate the accuracy of $\delta\boldsymbol{a}^{*(1)}$ and $\delta\boldsymbol{a}^{*(2)}$ using the cosine similarity, defined as

$$\mathrm{CS}^{(1)} := \delta\boldsymbol{a}^* \cdot \delta\boldsymbol{a}^{*(1)}, \quad \mathrm{CS}^{(2)} := \delta\boldsymbol{a}^* \cdot \delta\boldsymbol{a}^{*(2)}. \tag{51}$$

In Fig. 11(b), we calculate $\mathrm{CS}^{(1)}$ and $\mathrm{CS}^{(2)}$ for $(a, b) \in [0, 9] \times [0, 9]$. In the calculation, the parameters in Eq. (47) are set as

$$R_1(a, b) = \frac{10}{1 + a + b},$$

$$R_2(a) = \frac{8}{1 + a},$$

$$R_3(b) = \frac{7}{1 + b}, \quad R_4 = 5, \tag{52}$$

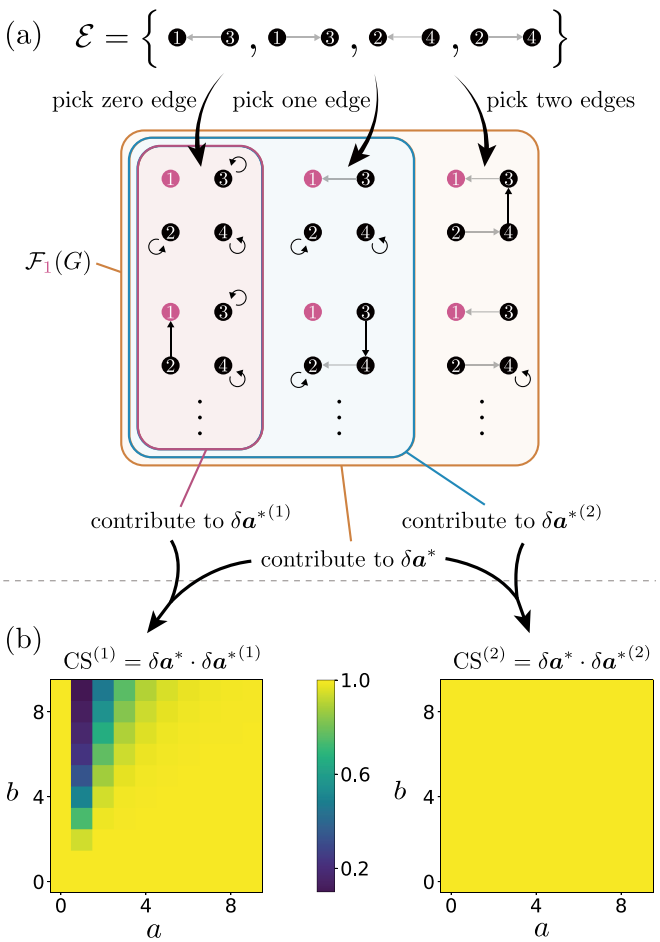

FIG. 11. (a) Illustration of which graphs in $\mathcal{F}_1(G)$ contribute to $\delta\boldsymbol{a}^{*(1)}$, $\delta\boldsymbol{a}^{*(2)}$, and $\delta\boldsymbol{a}^*$. Graphs with no edges in $\mathcal{E}$ contribute to $\delta\boldsymbol{a}^{*(1)}$, those with at most one edge in $\mathcal{E}$ contribute to $\delta\boldsymbol{a}^{*(2)}$, and all graphs in $\mathcal{F}_1(G)$ contribute to $\delta\boldsymbol{a}^*$. (b) Cosine similarity $\mathrm{CS}^{(1)} = \delta\boldsymbol{a}^* \cdot \delta\boldsymbol{a}^{*(1)}$ (left) and $\mathrm{CS}^{(2)} = \delta\boldsymbol{a}^* \cdot \delta\boldsymbol{a}^{*(2)}$ (right).

and mutation rates are set as

$$\mathsf{M} = \begin{pmatrix} -0.2 & 0.1 & 0 & 0 \\ 0.2 & -0.1 & 0 & 0 \\ 0 & 0 & -0.2 & 0.1 \\ 0 & 0 & 0.2 & -0.1 \end{pmatrix} + \epsilon \begin{pmatrix} -0.1 & 0 & 0.5 & 0 \\ 0 & -0.1 & 0 & 0.3 \\ 0.1 & 0 & -0.5 & 0 \\ 0 & 0.1 & 0 & -0.3 \end{pmatrix}, \tag{53}$$

with $\epsilon = 0.001$. As shown in Fig. 11(b), $\delta \boldsymbol{a}^{*(2)}$ is sufficiently accurate, while $\delta \boldsymbol{a}^{*(1)}$ is not. Therefore, it is sufficient to consider only graphs that contain at most one edge in $\mathcal{E}$. In general, the diagrammatic approach enables us to reduce the number of graphs that need to be considered when some mutation rates are much lower than the rest.

VII. SUMMARY AND DISCUSSION

In this paper, we extended the Markov chain tree theorem to the parallel mutation–reproduction model (Lemma 1) and derived diagrammatic expressions for the static responses of the mean fitness (Theorems 1 and 2) and the steady-state distribution (Theorem 3 and Corollary 1). We applied our results to two limiting cases and provided a constructive approach for approximating the static responses of the mean fitness and the steady-state distribution. Specifically, Theorems 1 and 2 enable the prediction of the static responses of the mean fitness based solely on information about the current state of the population. This study focuses primarily on population dynamics, but our results can be applied to various systems described by the parallel mutation–reproduction model. For example, Eq. (1) can be interpreted as the rate equation for deterministic chemical reactions involving self-replication. Therefore, our results could be useful for analyzing the static properties of these systems.

Approximate and exact expressions for the steady-state distribution and the static responses of the mean fitness may both be useful for controlling evolutionary dynamics. Because these expressions are given by the parameters M, R, and $\langle R \rangle_{\boldsymbol{\pi}}$, they provide guidelines for altering the environment to achieve the desired steady state and static responsiveness by changing the mutation rates M and the fitness R. Indeed, we applied our results in the context of combination therapy and reduced the number of graphs required to determine the optimal control of inhibitors. Controlling such evolution and population dynamics will likely be useful from both an epidemiological and a biotechnological perspective.

Similar to how the conventional Markov chain tree theorem and the concept of the spanning tree are used to analyze the static properties of systems in stochastic thermodynamics [52–54,68–76], our result may be useful for investigating the static properties of the parallel mutation–reproduction model. In stochastic thermodynamics, various inequalities have been obtained using the Markov chain tree theorem. It may be possible to find similar inequalities in population dynamics using our generalization of the Markov chain tree theorem.

Our results may be of interest from a graph theory perspective. In order to obtain diagrammatic expressions for the steady-state distribution and the static response of the mean fitness, we must introduce a generalization of a rooted spanning tree, that is a rooted 0/1 loop forest. Since a rooted 0/1 loop forest generalizes a rooted spanning tree, it may be possible to generalize some concepts in graph theory related to spanning trees, such as cycles and cyclic affinities [52], for the parallel mutation–reproduction model. Interestingly, we indeed obtain a generalization of Cayley's formula for the maximum number of rooted 0/1 loop forests in Appendix G.

ACKNOWLEDGMENTS

The authors thank Kohei Yoshimura, Naruo Ohga, and Namiko Mitarai for their suggestive comments. R.N. was supported by JSPS KAKENHI Grant No. 25KJ0931. S.I. was supported by JSPS KAKENHI Grants No. 22H01141, No. 23H00467, and No. 24H00834; JST ERATO Grant No. JPMJER2302; and UTEC-UTokyo FSI Research Grant Program. This research was supported by JSR Fellowship, the University of Tokyo.

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX A: THE DETAILS OF EQ. (1)

Here, we explain the details of Eq. (1). We assume that reproduction and death occur independently, and that mutation to another trait occurs during reproduction. Let r_i and d_i be the probabilities that trait i will reproduce and die in unit time, respectively. Furthermore, let P_{ij} be the probability that the immediate descendants of trait j will have trait i. Note that P_{ij} satisfies $\sum_i P_{ij} = 1$. Then, we can write down the number of trait i at time $t + \Delta t$ as follows:

$$n_i(t + \Delta t) = (1 - d_i \Delta t) n_i(t) + \Delta t \sum_j P_{ij} r_j n_j(t). \tag{A1}$$

The first term represents the individuals with trait i that do not die. The second term represents the mutation to trait i from the other traits $j \neq i$ and the reproduction of trait i. Taking the limit $\Delta t \to 0$ in Eq. (A1) and using $P_{ii} = 1 - \sum_{j(\neq i)} P_{ji}$, we obtain

$$\frac{d}{dt} n_i(t) = (r_i - d_i) n_i(t) + \sum_{j(\neq i)} \left(P_{ij} r_j n_j(t) - P_{ji} r_i n_i(t) \right). \tag{A2}$$

Therefore, defining $R_i := r_i - d_i$ and $M_{ij} := P_{ij} r_j$ $(i \neq j)$ reproduces Eq. (1).

APPENDIX B: DERIVATION OF EQ. (3)

We derive Eq. (3) from Eq. (1). Using Eq. (1), $dp_i(t)/dt$ can be computed as

$$\begin{aligned}\frac{d}{dt}p_i(t) &= \frac{d}{dt}\frac{n_i(t)}{n_{\rm tot}(t)}\\ &= \frac{1}{n_{\rm tot}(t)}\left\{R_i n_i(t) + \sum_{j(\neq i)}\big(M_{ij}n_j(t) - M_{ji}n_i(t)\big)\right\}\\ &\quad - \frac{n_i(t)}{n_{\rm tot}(t)^2}\sum_j\left\{R_j n_j(t)\right.\\ &\qquad\left. + \sum_{k(\neq j)}\big(M_{jk}n_k(t) - M_{kj}n_j(t)\big)\right\}\\ &= R_i p_i(t) + \sum_{j(\neq i)}\big(M_{ij}p_j(t) - M_{ji}p_i(t)\big) - \langle R\rangle_{\boldsymbol{p}(t)}p_i(t).\end{aligned} \tag{B1}$$

In the last equation, we used the definition of mean fitness and $\sum_j\sum_{k(\neq j)}M_{jk}n_k(t) = \sum_j\sum_{k(\neq j)}M_{kj}n_j(t)$. Rewriting Eq. (B1) in the vector form yields Eq. (3).

APPENDIX C: RELATIONSHIP BETWEEN THE STEADY STATE AND THE PERRON–FROBENIUS THEOREM

We show how the Perron–Frobenius theorem provides an expression for the steady state and its mathematical properties in Sec. II B. The statement of the Perron–Frobenius theorem is as follows.

Perron–Frobenius theorem [55]. Let $\mathsf{A}\in\mathbb{R}^{N\times N}$ be an irreducible matrix whose all components are real and nonnegative, and $\lambda_1,\lambda_2,\dots,\lambda_N$ be the eigenvalues of A that are ordered as $\mathrm{Re}\lambda_1 \geqslant \mathrm{Re}\lambda_2\cdots\geqslant\mathrm{Re}\lambda_N$. Then, the following statements hold:

(1) λ_1 is real and positive, and satisfies

$$|\lambda_i| \leqslant \lambda_1, \quad i = 2,\dots,N. \tag{C1}$$

(2) λ_1 is simple, i.e.,

$$\lambda_1 \neq \lambda_i, \quad i = 2,\dots,N. \tag{C2}$$

(3) There exists a right eigenvector $\boldsymbol{u}^{(1)}$ of A with eigenvalue λ_1 such that all components are positive.

To apply the Perron–Frobenius theorem to our model, we define A as $\mathsf{A} := \mathsf{R}+\mathsf{M}+a\mathsf{I}$ with $a = \max_i[-(R_i+M_{ii})]$. For any eigenvalue of A denoted by λ, the corresponding eigenvector $\boldsymbol{u}$ satisfies

$$(\mathsf{R}+\mathsf{M})\boldsymbol{u} = (\lambda - a)\boldsymbol{u}, \tag{C3}$$

which implies that the eigenvalues of $\mathsf{R}+\mathsf{M}$ are shifted by $-a$ from the eigenvalues of A while the eigenvectors of $\mathsf{R}+\mathsf{M}$ are the same as the eigenvectors of A. Since the matrix A is an irreducible matrix whose components are all real and nonnegative, we can apply the Perron–Frobenius theorem to the matrix A. We thus obtain the following corollary:

Corollary C1. Let $\mu_1,\mu_2,\dots,\mu_N$ be the eigenvalues of $\mathsf{R}+\mathsf{M}$ that are ordered as $\mathrm{Re}\mu_1 \geqslant \mathrm{Re}\mu_2\cdots\geqslant\mathrm{Re}\mu_N$. Then, the following statements hold:

(1) μ_1 is real, and satisfies

$$\mathrm{Re}\mu_i < \mu_1, \quad i = 2,\dots,N. \tag{C4}$$

(2) There exists an right eigenvector $\boldsymbol{u}^{(1)}$ of $\mathsf{R}+\mathsf{M}$ with eigenvalue μ_1 such that all components are positive.

The first statement of Corollary C 1 is derived as follows. From the first statement of the Perron–Frobenius theorem, we obtain

$$\mathrm{Re}(\lambda_i - a) \leqslant |\lambda_i| - a \leqslant \lambda_1 - a, \quad i = 2,\dots,N, \tag{C5}$$

where $\lambda_1,\lambda_2,\dots,\lambda_N$ are the eigenvalues of A ordered as $\mathrm{Re}\lambda_1 \geqslant \mathrm{Re}\lambda_2\cdots\geqslant\mathrm{Re}\lambda_N$. Combining Eq. (C5) with Eqs. (C2) and (C3), we have $\mu_i = \lambda_i - a$ for $i = 1,\dots,N$. In particular, we find that $\mu_1 = \lambda_1 - a \in \mathbb{R}$ and Eq. (C4) hold. The second statement is derived by combining Eq. (C3), $\mu_1 = \lambda_1 - a$, and the third statement of the Perron–Frobenius theorem.

From Corollary C.1, we can prove an expression for the steady-state distribution in Sec. II B as follows. For simplicity, we discuss the case where $\mathsf{R}+\mathsf{M}$ is diagonalizable, i.e., there exists an invertible matrix $\mathsf{S}\in\mathbb{C}^{N\times N}$ such that

$$\begin{aligned}\mathsf{R}+\mathsf{M} &= \mathsf{S}\begin{pmatrix}\mu_1 & & \\ & \ddots & \\ & & \mu_N\end{pmatrix}\mathsf{S}^{-1}\\ &= (\boldsymbol{u}^{(1)} \ \cdots \ \boldsymbol{u}^{(N)})\begin{pmatrix}\mu_1 & & \\ & \ddots & \\ & & \mu_N\end{pmatrix}\begin{pmatrix}(\boldsymbol{v}^{(1)})^\top \\ \vdots \\ (\boldsymbol{v}^{(N)})^\top\end{pmatrix}\\ &= \sum_i \mu_i \boldsymbol{u}^{(i)}(\boldsymbol{v}^{(i)})^\top.\end{aligned} \tag{C6}$$

Here, $\boldsymbol{u}^{(i)}$ and $\boldsymbol{v}^{(i)}$ are the right and left eigenvectors of $\mathsf{R}+\mathsf{M}$ with eigenvalue μ_i, respectively. Note that $\boldsymbol{u}^{(i)}$ and $\boldsymbol{v}^{(i)}$ satisfy $(\boldsymbol{v}^{(i)})^\top\boldsymbol{u}^{(j)} = (\mathsf{S}^{-1}\mathsf{S})_{ij} = \delta_{ij}$, where δ_{ij} is the Kronecker delta. Any initial condition $\boldsymbol{n}(0)$ is expanded as $\boldsymbol{n}(0) = \sum_i c_i\boldsymbol{u}^{(i)}$ with $c_i\in\mathbb{C}$. Thus, we obtain

$$\boldsymbol{n}(t) = e^{(\mathsf{R}+\mathsf{M})t}\boldsymbol{n}(0) = \sum_i c_i e^{\mu_i t}\boldsymbol{u}^{(i)}, \tag{C7}$$

where we used $e^{(\mathsf{R}+\mathsf{M})t} = \sum_i e^{\mu_i t}\boldsymbol{u}^{(i)}(\boldsymbol{v}^{(i)})^\top$ which follows from Eq. (C6) and $(\boldsymbol{v}^{(i)})^\top\boldsymbol{u}^{(j)} = \delta_{ij}$. Using Eq. (C7), we obtain the long-time limit of $\boldsymbol{p}(t)$ as

$$\begin{aligned}\lim_{t\to\infty}\boldsymbol{p}(t) &= \lim_{t\to\infty}\frac{\boldsymbol{n}(t)}{\sum_j n_j(t)}\\ &= \lim_{t\to\infty}\frac{\sum_i c_i e^{\mu_i t}\boldsymbol{u}^{(i)}}{\sum_j\sum_i c_i e^{\mu_i t}u_j^{(i)}}\\ &= \frac{\boldsymbol{u}^{(1)}}{\sum_j u_j^{(1)}} =: \boldsymbol{\pi},\end{aligned} \tag{C8}$$

where we used Eq. (C4) in the third equality. Equation (C8) means $\boldsymbol{p}(t)$ converges to $\boldsymbol{\pi} = \boldsymbol{u}^{(1)}/\sum_j u_j^{(1)}$ regardless of the initial condition. Furthermore, the second statement of Corollary C 1 implies that all components of $\boldsymbol{\pi} = \boldsymbol{u}^{(1)}/\sum_j u_j^{(1)}$ are strictly positive, which results in the positivity of $\boldsymbol{\pi}$. If the matrix $\mathsf{R}+\mathsf{M}$ is not diagonalizable, we also obtain a similar expression for $\boldsymbol{n}(t)$ whose dominant term

in the long-time limit is $\boldsymbol{n}(t) \propto e^{\mu_1 t}\boldsymbol{u}^{(1)}$ [77]. Thus, we again obtain Eq. (C8) in general.

We show that the eigenvalue μ_1 is equal to the reproductive rate $\langle R\rangle_{\boldsymbol{\pi}}$. Noting $\boldsymbol{\pi} \propto \boldsymbol{u}^{(1)}$, we find that $\boldsymbol{\pi}$ is the right eigenvector of $\mathsf{R}+\mathsf{M}$ with eigenvalue μ_1. Combining this fact and Eq. (8), we obtain

$$\mu_1 = \langle R\rangle_{\boldsymbol{\pi}}. \tag{C9}$$

We prove that there exists the vector $\boldsymbol{\zeta}$ that satisfies Eqs. (9), (10), and whose every element is positive. Applying the Perron–Frobenius theorem to $\mathsf{A}^\top = (\mathsf{R}+\mathsf{M}+a\mathsf{I})^\top$ with $a = \max_i[-(R_i + M_{ii})]$, we can state that there exists the left eigenvector of $\mathsf{R}+\mathsf{M}$, $(\boldsymbol{v}^{(1)})^\top$, whose all components are strictly positive. Therefore, if we define $\boldsymbol{\zeta} \in \mathbb{R}^N$ as $\boldsymbol{\zeta} := \boldsymbol{v}^{(1)}/\{(\boldsymbol{v}^{(1)})^\top\boldsymbol{\pi}\}$, we find that all components of $\boldsymbol{\zeta}$ are strictly positive, and $\boldsymbol{\zeta}$ satisfies Eq. (10). Since $\boldsymbol{\zeta}$ is the left eigenvector of $\mathsf{R}+\mathsf{M}$ and Eq. (C9) holds, Eq. (9) is also proved.

APPENDIX D: DERIVATION OF EQS. (11) AND (12)

We here derive Eq. (11). Multiplying Eq. (8) by $\boldsymbol{\zeta}^\top$ from the left and using Eq. (10), we obtain

$$\langle R\rangle_{\boldsymbol{\pi}} = \boldsymbol{\zeta}^\top(\mathsf{R}+\mathsf{M})\boldsymbol{\pi}. \tag{D1}$$

By differentiating this equation with respect to R_i, we obtain

$$\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial R_i} = \frac{\partial\boldsymbol{\zeta}^\top}{\partial R_i}(\mathsf{R}+\mathsf{M})\boldsymbol{\pi} + \zeta_i\pi_i + \boldsymbol{\zeta}^\top(\mathsf{R}+\mathsf{M})\frac{\partial\boldsymbol{\pi}}{\partial R_i}. \tag{D2}$$

The first and third terms in Eq. (D2) cancel out as follows:

$$\begin{aligned}
&\frac{\partial\boldsymbol{\zeta}^\top}{\partial R_i}(\mathsf{R}+\mathsf{M})\boldsymbol{\pi} + \boldsymbol{\zeta}^\top(\mathsf{R}+\mathsf{M})\frac{\partial\boldsymbol{\pi}}{\partial R_i}\\
&= \langle R\rangle_{\boldsymbol{\pi}}\left(\frac{\partial\boldsymbol{\zeta}^\top}{\partial R_i}\boldsymbol{\pi} + \boldsymbol{\zeta}^\top\frac{\partial\boldsymbol{\pi}}{\partial R_i}\right)\\
&= \langle R\rangle_{\boldsymbol{\pi}}\frac{\partial}{\partial R_i}(\boldsymbol{\zeta}^\top\boldsymbol{\pi}) = 0.
\end{aligned} \tag{D3}$$

In the second line, we used Eqs. (8) and (9). In the third line, we used Eq. (10). Substituting Eq. (D3) into Eq. (D2), we obtain Eq. (11).

Equation (12) can be obtained in a similar way. Differentiating Eq. (D1) with respect to M_{ij}, we obtain Eq. (12) as follows:

$$\begin{aligned}
\frac{\partial\langle R\rangle_{\boldsymbol{\pi}}}{\partial M_{ij}} &= \frac{\partial\boldsymbol{\zeta}^\top}{\partial M_{ij}}(\mathsf{R}+\mathsf{M})\boldsymbol{\pi} + \zeta_i\pi_j - \zeta_j\pi_j + \boldsymbol{\zeta}^\top(\mathsf{R}+\mathsf{M})\frac{\partial\boldsymbol{\pi}}{\partial M_{ij}}\\
&= \langle R\rangle_{\boldsymbol{\pi}}\frac{\partial}{\partial M_{ij}}(\boldsymbol{\zeta}^\top\boldsymbol{\pi}) + \zeta_i\pi_j - \zeta_j\pi_j\\
&= (\zeta_i - \zeta_j)\pi_j,
\end{aligned} \tag{D4}$$

where we used $\partial M_{jj}/\partial M_{ij} = -1$ and a result similar to Eq. (D3).

APPENDIX E: CONSTRUCTION OF ROOTED 0/1 LOOP FORESTS

We present a method for constructing all rooted 0/1 loop forests of a given basic graph. Given a basic graph G, a graph in $\mathcal{F}_{i\leftarrow j}(G)$ can be constructed using the following procedure (see Fig. 12 for an example).

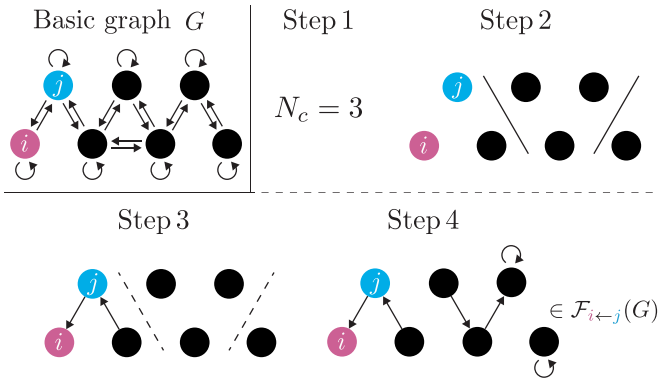

FIG. 12. Illustration of the construction of a rooted 0/1 loop forest in $\mathcal{F}_{i\leftarrow j}(G)$. A basic graph G is shown in the top left. In Step 1, N_c is set to 3. In Step 2, the vertices of G are partitioned into $N_c = 3$ subsets such that vertex i and vertex j belong to the same subset. In Step 3, a spanning tree rooted at vertex i is constructed using the edges of G in the subset containing vertex i. In Step 4, spanning trees rooted at a selected vertex are constructed in the remaining subsets, with a loop attached to each root.

Step 1. Determine the number of components N_c.

Step 2. Partition the vertices of G into N_c subsets such that vertex i and vertex j belong to the same subset.

Step 3. In the subset containing vertex i, construct a spanning tree rooted at vertex i using the edges of G.

Step 4. For each subset that does not contain vertex i, select one vertex. Then, construct a spanning tree rooted at that vertex using the edges of G and attach a loop to the root.

By performing all possibilities for these four procedures, we can construct all rooted forests. Note that Steps 3 and 4 are not always achievable for any partition of vertices.

Similarly, we can construct all rooted forests in $\mathcal{F}_{i\not\leftarrow j}(G)$ and $\mathcal{F}_i(G)$. To construct all rooted forests in $\mathcal{F}_{i\not\leftarrow j}(G)$, replace Step 2 with Step 2', which is described as follows:

Step 2'. Partition the vertices of G into N_c classes such that vertex i and vertex j belong to different classes.

To construct all rooted forests in $\mathcal{F}_i(G)$, replace Step 2 with Step 2", which is described as follows:

Step 2". Partition the vertices of G into N_c classes.

APPENDIX F: DERIVATION OF LEMMA 1

In this Appendix, we present the proof of Lemma 1. For convenience, instead of proving Eq. (18) itself, we prove the following equation obtained by swapping the indices i and j.

$$\zeta_i\pi_j = \frac{1}{Z}\sum_{H\in\mathcal{F}_{j\leftarrow i}(G)} w(H). \tag{F1}$$

The proof in this Appendix is a generalization of the proof of the Markov chain tree theorem in Sec. 2.3.3 of Ref. [78].

The outline of the proof is as follows. In Appendix F 1, we express $\zeta_i\pi_j$ in terms of the matrix $-(\mathsf{R} - \langle R\rangle_{\boldsymbol{\pi}}\mathsf{I} + \mathsf{M})$. In Appendix F 2, we rewrite the expression for $\zeta_i\pi_j$ derived in Appendix F 1 as a sum of several contributions. We show that each contribution corresponds to a graph. In Appendix F 3, we give the necessary and sufficient conditions for rooted 0/1 loop forests that are suitable for our proof. In Appendix F 4,

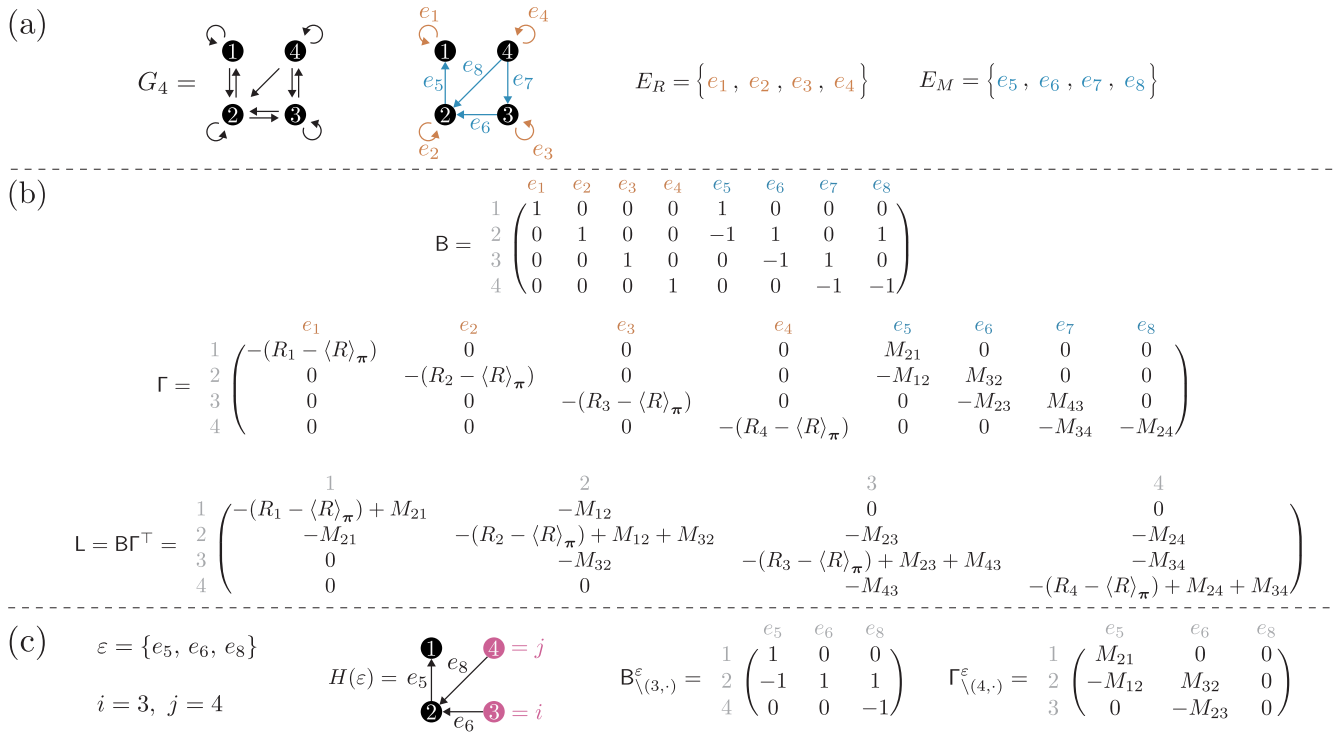

FIG. 13. (a) An example of E_R and E_M for the basic graph G_4. (b) An example of B, Γ, and L for the basic graph G_4. Equation (F7) can be confirmed. (c) An example of $H(\varepsilon)$, $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$, and $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ for the basic graph G_4.

we derive the necessary conditions for the contribution to be nonzero. These conditions allow us to restrict the range of the summation in Appendix F 2. We also show that these necessary conditions are related to some conditions for rooted 0/1 loop forests given in Appendix F 3. In Appendix F 5, we transform each matrix corresponding to a contribution into block diagonal matrix. This makes it easy to calculate each contribution. Each block of the block diagonal matrix represents a component of the graph corresponding to each contribution. In Appendix F 6, we express each contribution in terms of the weight of a graph by using the block diagonalization given in Appendix F 5. In Appendix F 7, we express $\zeta_i\pi_j$ in terms of the weights of rooted 0/1 loop forests by combining the above results.

1. An expression for $\zeta_i\pi_j$

We derive an expression for $\zeta_i\pi_j$ in terms of the matrix $\mathsf{L} := -(\mathsf{R} - \langle R\rangle_{\boldsymbol{\pi}}\mathsf{I} + \mathsf{M})$. The adjugate matrix $\mathrm{adj}(\mathsf{L})$ is defined as $(\mathrm{adj}(\mathsf{L}))_{ji} := (-1)^{i+j}\det(\mathsf{L}_{\setminus(i,j)})$, where $\mathsf{L}_{\setminus(i,j)}$ is the matrix obtained by removing the ith row and jth column from L. The adjugate matrix $\mathrm{adj}(\mathsf{L})$ satisfies $\mathsf{L}\,\mathrm{adj}(\mathsf{L}) = \det(\mathsf{L})\mathsf{I}$. Note that $\det(\mathsf{L}) = 0$ because Eq. (8) implies that L has zero eigenvalue. Combining these equations, we obtain $\mathsf{L}\,\mathrm{adj}(\mathsf{L}) = \mathsf{O}$ where O is the zero matrix. This means that every column of $\mathrm{adj}(\mathsf{L})$ is in the kernel $\ker(\mathsf{L})$. Equation (8) implies $\boldsymbol{\pi} \in \ker(\mathsf{L})$. Moreover, the eigenvalue $\langle R\rangle_{\boldsymbol{\pi}}$ of the matrix $\mathsf{R} + \mathsf{M}$ is simple, and the corresponding eigenvector is $\boldsymbol{\pi}$ as discussed in Eqs. (C4) and (C9). This fact implies that $\ker(\mathsf{L})$ is spanned by $\boldsymbol{\pi}$, i.e., $\ker(\mathsf{L}) = \{c\boldsymbol{\pi} \mid c \in \mathbb{C}\}$. Thus, each column of $\mathrm{adj}(\mathsf{L})$ is proportional to $\boldsymbol{\pi}$. Similarly, the adjugate matrix $\mathrm{adj}(\mathsf{L}^\top)$ also satisfies $\mathsf{L}^\top\mathrm{adj}(\mathsf{L}^\top) = \det(\mathsf{L})\mathsf{I}$, and we also obtain $\mathsf{L}^\top\mathrm{adj}(\mathsf{L}^\top) = \mathsf{O}$ and $\ker(\mathsf{L}^\top) = \{c\boldsymbol{\zeta} \mid c \in \mathbb{C}\}$. Thus, each column of $\mathrm{adj}(\mathsf{L}^\top) = (\mathrm{adj}(\mathsf{L}))^\top$ is proportional to $\boldsymbol{\zeta}$. Combining these facts, we find $\mathrm{adj}(\mathsf{L}) \propto \boldsymbol{\pi}\boldsymbol{\zeta}^\top$. Since Eq. (10) gives $\mathrm{tr}(\boldsymbol{\pi}\boldsymbol{\zeta}^\top) := \sum_i(\boldsymbol{\pi}\boldsymbol{\zeta}^\top)_{ii} = \boldsymbol{\zeta}^\top\boldsymbol{\pi} = 1$, we obtain $\boldsymbol{\pi}\boldsymbol{\zeta}^\top = \mathrm{adj}(\mathsf{L})/\mathrm{tr}(\mathrm{adj}(\mathsf{L}))$. Therefore, we can rewrite $\zeta_i\pi_j = (\boldsymbol{\pi}\boldsymbol{\zeta}^\top)_{ji}$ as

$$\zeta_i\pi_j = \frac{(\mathrm{adj}(\mathsf{L}))_{ji}}{\mathrm{tr}(\mathrm{adj}(\mathsf{L}))} = (-1)^{i+j}\frac{\det\left(\mathsf{L}_{\setminus(i,j)}\right)}{\sum_k \det\left(\mathsf{L}_{\setminus(k,k)}\right)}. \tag{F2}$$

2. An expression for $\det(\mathsf{L}_{\setminus(i,j)})$ by generalizations of incidence matrix and weighted incidence matrix

We introduce generalizations of the incidence matrix and the weighted incidence matrix to express L as the product of these matrices. Since these matrices represent the structure of the basic graph, this fact allows us to apply methods from graph theory to compute $\mathsf{L}_{\setminus(i,j)}$ in Eq. (F2). To introduce these two matrices, we define two sets of edges E_R and E_M as follows [see also Fig. 13(a) for an example].

$$E_R := \{(i \leftarrow i) \mid i \in V(G)\}, \tag{F3}$$

$$E_M := \{(i \leftarrow j) \mid i < j,\ M_{ij} \neq 0 \text{ or } M_{ji} \neq 0,\ i \in V(G),\ j \in V(G)\}. \tag{F4}$$

Note that the element of E_M is not necessarily in $E(G)$, whereas an element of E_R is in $E(G)$ because $(i \leftarrow j)$ is in E_M when $i < j$, $M_{ij} = 0$, and $M_{ji} \neq 0$. By arbitrarily labeling the edges in $E_R \cup E_M$ as $E_R \cup E_M = \{e_1, \ldots, e_{|E_R|+|E_M|}\}$, we define the *generalized incidence matrix* $\mathsf{B} \in \mathbb{R}^{N\times(|E_R|+|E_M|)}$ and

the *generalized weighted incidence matrix* $\Gamma \in \mathbb{R}^{N\times(|E_R|+|E_M|)}$ as

$$(\mathsf{B})_{i\alpha} := \begin{cases} \delta_{i\mathrm{t}(e_\alpha)}, & \text{if } e_\alpha \in E_R, \\ \delta_{i\mathrm{t}(e_\alpha)} - \delta_{i\mathrm{s}(e_\alpha)}, & \text{if } e_\alpha \in E_M, \end{cases} \tag{F5}$$

$$(\Gamma)_{i\alpha} := \begin{cases} -(R_i - \langle R\rangle_{\boldsymbol{\pi}})\delta_{i\mathrm{t}(e_\alpha)}, & \text{if } e_\alpha \in E_R, \\ M_{\mathrm{s}(e_\alpha)\mathrm{t}(e_\alpha)}\delta_{i\mathrm{t}(e_\alpha)} - M_{\mathrm{t}(e_\alpha)\mathrm{s}(e_\alpha)}\delta_{i\mathrm{s}(e_\alpha)}, & \text{if } e_\alpha \in E_M, \end{cases} \tag{F6}$$

respectively, where $N = |V(G)|$, and $\mathrm{s}(e)$ and $\mathrm{t}(e)$ are the source and target vertices of an edge $e = (\mathrm{t}(e) \leftarrow \mathrm{s}(e))$, respectively [see also Fig. 13(b) for an example]. Note that while the element of the conventional incidence matrix corresponding to a loop is set to zero, the element of the generalized incidence matrix $(\mathsf{B})_{i\alpha}$ can be nonzero even if the edge e_α is a loop.

Using B and Γ, we can express L as follows [see also Fig. 13(b) for an example]:

$$\mathsf{L} = \mathsf{B}\Gamma^\top. \tag{F7}$$

This equation can be verified by direct calculation as follows:

$$\begin{aligned}(\mathsf{B}\Gamma^\top)_{ij} &= \sum_{\alpha|e_\alpha\in E_R\cup E_M} (\mathsf{B})_{i\alpha}(\Gamma)_{j\alpha} \\ &= -\sum_{e_\alpha\in E_R} (R_j - \langle R\rangle_{\boldsymbol{\pi}})\delta_{i\mathrm{t}(e_\alpha)}\delta_{j\mathrm{t}(e_\alpha)} \\ &\quad + \sum_{e_\alpha\in E_M} \big(M_{\mathrm{s}(e_\alpha)\mathrm{t}(e_\alpha)}\delta_{j\mathrm{t}(e_\alpha)}\delta_{i\mathrm{t}(e_\alpha)} \\ &\quad + M_{\mathrm{t}(e_\alpha)\mathrm{s}(e_\alpha)}\delta_{j\mathrm{s}(e_\alpha)}\delta_{i\mathrm{s}(e_\alpha)}\big) \\ &\quad - \sum_{e_\alpha\in E_M} \big(M_{\mathrm{s}(e_\alpha)\mathrm{t}(e_\alpha)}\delta_{j\mathrm{t}(e_\alpha)}\delta_{i\mathrm{s}(e_\alpha)} \\ &\quad + M_{\mathrm{t}(e_\alpha)\mathrm{s}(e_\alpha)}\delta_{j\mathrm{s}(e_\alpha)}\delta_{i\mathrm{t}(e_\alpha)}\big) \\ &= \delta_{ij}\{-(R_j - \langle R\rangle_{\boldsymbol{\pi}}) - M_{ii}\} + (1-\delta_{ij})(-M_{ij}) \\ &= \mathsf{L}_{ij}.\end{aligned}$$

We refer to L as the *generalized Laplacian matrix* because the conventional Laplacian matrix for the weighted graph can be written as the product of the incidence matrix and the weighted incidence matrix [56]. The matrix $\mathsf{L}_{\setminus(i,j)}$ is also given by

$$\mathsf{L}_{\setminus(i,j)} = \mathsf{B}_{\setminus(i,\cdot)}(\Gamma_{\setminus(j,\cdot)})^\top, \tag{F8}$$

where $\mathsf{B}_{\setminus(i,\cdot)}$ and $\Gamma_{\setminus(i,\cdot)}$ are the matrices obtained by removing the ith row from B and Γ, respectively.

We obtain an expression for $\det(\mathsf{L}_{\setminus(i,j)})$. Let $\mathsf{B}^\varepsilon_{\setminus(i,\cdot)} \in \mathbb{R}^{(N-1)\times(N-1)}$ and $\Gamma^\varepsilon_{\setminus(i,\cdot)} \in \mathbb{R}^{(N-1)\times(N-1)}$ be the matrices obtained by removing all columns except the selected $N-1$ columns corresponding to the set of the edges ε from $\mathsf{B}_{\setminus(i,\cdot)}$ and $\Gamma_{\setminus(i,\cdot)}$, respectively [see also Fig. 13(c) for an example]. Applying the Cauchy–Binet formula to Eq. (F8), we obtain

$$\det\left(\mathsf{L}_{\setminus(i,j)}\right) = \sum_\varepsilon \det\left(\mathsf{B}^\varepsilon_{\setminus(i,\cdot)}\right)\det\left(\Gamma^\varepsilon_{\setminus(j,\cdot)}\right), \tag{F9}$$

where the sum over ε is taken over all possible subsets of $E_R \cup E_M$ consisting of $N-1$ edges.

As shown below, each term in the summation in Eq. (F9) corresponds to a graph. For each edge set ε in Eq. (F9), we define the corresponding graph $H(\varepsilon)$ as the graph whose edge set is ε and whose vertices are the endpoints of the edges in ε [see also Fig. 13(c) for an example]. We then interpret the right-hand side of Eq. (F9) as the summation of the contributions from all such graphs $H(\varepsilon)$. We note that $H(\varepsilon)$ is not necessarily a subgraph of G.

3. Conditions for a rooted 0/1 loop forest

We introduce some conditions imposed on a graph. Some of these conditions are used to characterize a rooted 0/1 loop forest as shown in Eqs. (F10) and (F11). Moreover, some of these conditions imposed on $H(\varepsilon)$ are related to the conditions for $\det(\mathsf{B}^\varepsilon_{\setminus(i,\cdot)})\det(\Gamma^\varepsilon_{\setminus(j,\cdot)}) \neq 0$ in Eq. (F9). Therefore, these conditions connect $H(\varepsilon)$ to a rooted 0/1 loop forest.

We now introduce the following notations, C_1, C_2, $\mathrm{C}_3(i)$, $\mathrm{C}_4(i)$, $\mathrm{C}_5(i)$, $\mathrm{C}_6(i,j)$, and C_7 for various conditions imposed on a graph H.

(1) H satisfies C_1. $\overset{\text{def}}{\iff}$ H contains no cycles.

(2) H satisfies C_2. $\overset{\text{def}}{\iff}$ $V(H) = V(G)$.

(3) H satisfies $\mathrm{C}_3(i)$. $\overset{\text{def}}{\iff}$ In each component of H that does not contain vertex i, there is exactly one loop.

(4) H satisfies $\mathrm{C}_4(i)$. $\overset{\text{def}}{\iff}$ In the component of H that contains vertex i, there are no loops.

(5) H satisfies $\mathrm{C}_5(i)$. $\overset{\text{def}}{\iff}$ In the component of H that contains vertex i, all edges are directed to vertex i. In each component of H that does not contain vertex i, all edges are directed to the vertex with the loop.

(6) H satisfies $\mathrm{C}_6(i,j)$. $\overset{\text{def}}{\iff}$ Vertex i and vertex j are connected in H.

(7) H satisfies C_7. $\overset{\text{def}}{\iff}$ H has $|V(G)|-1$ edges.

Using these notations, the definition of $\mathcal{F}_{j\leftarrow i}(G)$ given in Sec. III F can be rewritten as follows:

$$\begin{aligned}&H \in \mathcal{F}_{j\leftarrow i}(G) \overset{\text{def}}{\iff} H \subseteq G \text{ and} \\ &\quad H \text{ satisfies } \mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3(j), \mathrm{C}_4(j), \mathrm{C}_5(j), \text{ and } \mathrm{C}_6(i,j).\end{aligned} \tag{F10}$$

The condition C_2 can be replaced with C_7 as follows:

$$\begin{aligned}&H \in \mathcal{F}_{j\leftarrow i}(G) \iff H \subseteq G \text{ and} \\ &\quad H \text{ satisfies } \mathrm{C}_1, \mathrm{C}_3(j), \mathrm{C}_4(j), \mathrm{C}_5(j), \mathrm{C}_6(i,j), \text{ and } \mathrm{C}_7.\end{aligned} \tag{F11}$$

The equivalent expression for $\mathcal{F}_{j\leftarrow i}(G)$ in Eq. (F11) is used later to prove Lemma 1, rather than its definition in Eq. (F10).

The statement in Eq. (F11) can be proved as follows. As mentioned below Eq. (15), any H in $\mathcal{F}_{j\leftarrow i}(G)$ defined by Eq. (F10) has $|V(G)|-1$ edges. In other words, if H satisfies all the conditions in Eq. (F10), then it also satisfies C_7, and thus it satisfies all the conditions in Eq. (F11). Next, we show that if H satisfies all the conditions in Eq. (F11), then it also satisfies all the conditions in Eq. (F10). Let N_c be the number of components of H. Since $\mathrm{C}_3(j)$ is satisfied, H can be expressed as $H = H^{(1)} \cup \bigcup_{\alpha=2}^{N_\mathrm{c}}(H^{(\alpha)} \cup (r^{(\alpha)} \leftarrow r^{(\alpha)}))$, where $H^{(1)}$ is the component containing vertex j, $r^{(\alpha)}$ is the vertex that has the unique loop in the αth component, and $H^{(\alpha)}$ is the graph obtained by removing the loop

$(r^{(\alpha)} \leftarrow r^{(\alpha)})$ from the αth component for $\alpha \geqslant 2$. Since C_1, $C_3(j)$, and $C_4(j)$ are satisfied, each $H^{(\alpha)}$ is connected and contains neither cycles nor loops. Therefore, $H^{(\alpha)}$ is a rooted spanning tree for any α, and $|E(H^{(\alpha)})| = |V(H^{(\alpha)})| - 1$ is satisfied. Using this equation and C_7, we obtain $|V(G)| - 1 = |E(H)| = |E(H^{(1)})| + \sum_{\alpha=2}^{N_c}(|E(H^{(\alpha)})| + 1) = |V(H^{(1)})| - 1 + \sum_{\alpha=2}^{N_c} |V(H^{(\alpha)})| = |V(H)| - 1$. Because H is a subgraph of G, $|V(G)| = |V(H)|$ leads to $V(H) = V(G)$, that is the condition C_2. Therefore, H satisfying all the conditions in Eq. (F11) satisfies all the conditions in Eq. (F10). From the above, the necessary and sufficient conditions in Eq. (F10) have been proved.

4. Conditions for $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) \neq 0$

We consider the conditions for $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) \neq 0$ because ε satisfying $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) \neq 0$ only contributes to $\sum_{\varepsilon} \det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)})$ in Eq. (F9). Here, we show that $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ if the corresponding graph $H(\varepsilon)$ does not satisfy any of C_1, $C_3(j)$, $C_4(j)$, and $C_6(i, j)$. Note that $H(\varepsilon)$ always satisfies C_7 since ε consists of $N - 1$ edges with $N = |V(G)|$. This fact allows us to restrict the sum to those ε such that the corresponding graph $H(\varepsilon)$ satisfies these conditions, which are related to $\mathcal{F}_{j\leftarrow i}(G)$.

a. C_1

We show that $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ if $H(\varepsilon)$ does not satisfy C_1. If $H(\varepsilon)$ does not satisfy C_1, then $H(\varepsilon)$ contains a cycle. Here, we introduce the notation $E(H(\varepsilon)) = \varepsilon = \{e_{\alpha_1}, \ldots, e_{\alpha_{N-1}}\}$ where e_{α_k} corresponds to the kth column of $\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)}$ (or $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$). Suppose that $H(\varepsilon)$ contains a cycle $(e'_1, \ldots, e'_L)$ where $e'_k \in E(H(\varepsilon))$ for $k = 1, \ldots, L$ and $L \geqslant 3$. Note that a cycle does not contain a loop, and the target vertex of the edge e'_k is different from the source vertex of the edge e'_k ($s(e'_k) \neq t(e'_k)$). We introduce $-e := (s(e) \leftarrow t(e))$ as the reverse of the edge $e = (t(e) \leftarrow s(e))$. By appropriately considering the reverse of some edges in the cycle $(e'_1, \ldots, e'_L)$, we obtain a *directed cycle* $\mathcal{C} = (\bar{e}'_1, \ldots, \bar{e}'_L)$ satisfying $s(\bar{e}'_{k+1}) = t(\bar{e}'_k)$ for any $k = 1, \ldots, L$, where $s(\bar{e}'_{L+1}) := s(\bar{e}'_1)$. Using the directed cycle $\mathcal{C}$, we define the vector $\boldsymbol{x} = (x_1, \ldots, x_{N-1})^{\top} \in \mathbb{R}^{N-1}$ as

$$x_k := \begin{cases} 1, & \text{if } e_{\alpha_k} \in \mathcal{C}, \\ -1, & \text{if } -e_{\alpha_k} \in \mathcal{C}, \\ 0, & \text{otherwise.} \end{cases} \tag{F12}$$

Therefore, the following relation for the directed cycle and the incidence matrix [57],

$$\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)}\boldsymbol{x} = \boldsymbol{0}, \tag{F13}$$

is satisfied [see also Fig. 14(a) for an example] because the contributions of loops do not exist in $\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)}\boldsymbol{x}$, i.e., $x_k = 0$ if $e_{\alpha_k} \in E_R$, and we only consider the contribution by the conventional incidence matrix for $e_{\alpha_k} \in E_M$. We can also verify Eq. (F13) directly using Eqs. (F5) and (F12), and $t(\bar{e}'_k) = s(\bar{e}'_{k+1})$ for $k = 1, \ldots, L$. Equation (F13) implies that the columns of $\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)}$ are linearly dependent and thus $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)}) = 0$. Therefore, $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ if $H(\varepsilon)$ does not satisfy C_1.

b. $C_3(j)$

We show that $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ if $H(\varepsilon)$ does not satisfy $C_3(j)$ and satisfies C_1. We assume that $H(\varepsilon)$ satisfies C_1, since otherwise $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ already holds, as discussed in Appendix F 4 a. Therefore, we assume that the component does not contain cycles.

There are two cases where $C_3(j)$ is not satisfied. The first case is that $H(\varepsilon)$ contains a component with at least two loops, where this component does not contain vertex j. The second case is that $H(\varepsilon)$ contains a component without loops, where this component does not contain vertex j. We show $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ in both cases.

First, we consider the first case. We arbitrarily choose such a component and let V' and E' denote the set of vertices and the set of edges for this component, respectively. We can show $|E'| - 1 \geqslant |V'|$ because the number of edges in a component without cycles and with $n_l (\geqslant 2)$ loops is $|E'| = |V'| - 1 + n_l$. Here, the number of edges that are not loops for a component without cycles is $|V'| - 1$ because the component without cycles becomes a spanning tree for V' with $|V'| - 1$ edges when we remove the loops from the component. Using $|E'| - 1 \geqslant |V'|$, we can show that the columns of $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$ are linearly dependent as follows. Note that all rows corresponding to V' are present in $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$ because the component we are now considering does not contain vertex j. We arbitrarily label the edges in E' as $E' = \{e^{(1)}, \ldots, e^{(|E'|)}\}$. Let $\boldsymbol{w}^{(k)} \in \mathbb{R}^{|V(G)|-1}$ be the column of $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$ corresponding to $e^{(k)}$ [see also Fig. 14(b) for examples of $\boldsymbol{w}^{(k)}$]. By the definition of Γ in Eq. (F6), $(\Gamma)_{i\alpha} \neq 0$ if vertex i is the target vertex or the source vertex of e_α. Therefore, the element of $\boldsymbol{w}^{(k)}$ can be nonzero only if the corresponding node is in V'. In other words, each column of $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$ corresponding to an edge in E' can have nonzero elements only in the rows corresponding to V'. Combining this fact with $|E'| - 1 \geqslant |V'|$, we find that $\boldsymbol{w}^{(k)}$ has at most $|E'| - 1$ nonzero elements in the rows corresponding to V'. Thus, the $|E'|$ vectors, $\boldsymbol{w}^{(1)}, \ldots, \boldsymbol{w}^{(|E'|)}$, are linearly dependent. As a result, the columns of $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$ are linearly dependent, and hence $\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ [see also Fig. 14(b) for an example]. Thus, we conclude $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ in the first case.

Next, we consider the second case. Let V' and E' denote the set of vertices and the set of edges for this component, respectively. We have $|E'| = |V'| - 1$ because the component does not contain loops and cycles, and the component is a spanning tree for V' with $|V'| - 1$ edges. Using $|E'| = |V'| - 1$, we can show that the rows of $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$ are linearly dependent as follows. Note that all rows corresponding to V' are present in $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$. These rows can have nonzero elements only in the columns corresponding to E'. Combining this fact with $|E'| = |V'| - 1$, we find that each row of $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$ corresponding to an element of V' can have at most $|V'| - 1$ nonzero elements in the columns corresponding to E'. Thus, the $|V'|$ rows of $\Gamma^{\varepsilon}_{\backslash(j,\cdot)}$ corresponding to the set V' are linearly dependent. Therefore, we conclude $\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$, and thus $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ in the second case.

Combining the results of the two cases, we conclude that $\det(\mathsf{B}^{\varepsilon}_{\backslash(i,\cdot)})\det(\Gamma^{\varepsilon}_{\backslash(j,\cdot)}) = 0$ holds when $H(\varepsilon)$ does not satisfy $C_3(j)$ and satisfies C_1.

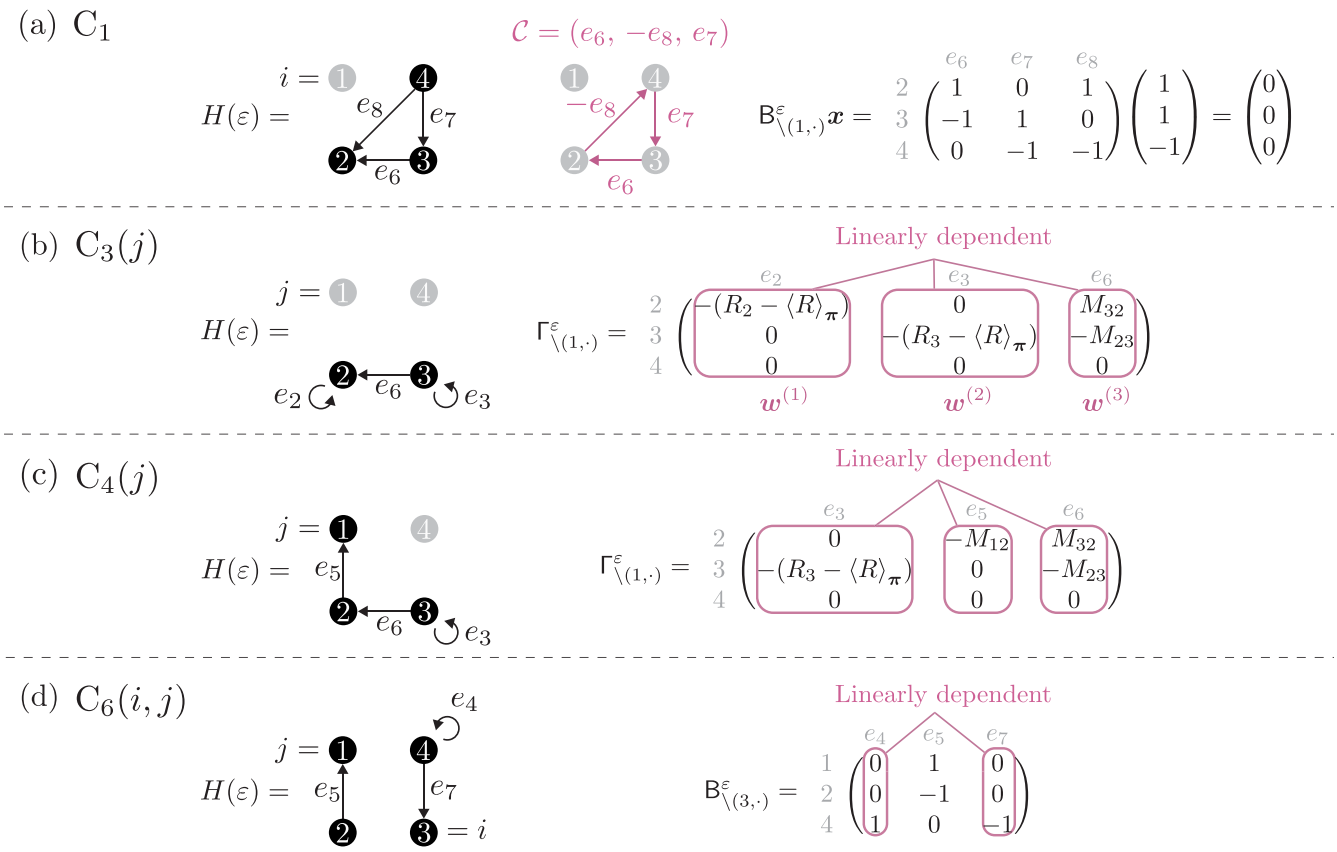

FIG. 14. Examples of graphs that do not satisfy C_1, $C_3(j)$, $C_4(j)$, or $C_6(i, j)$. (a) This $H(\varepsilon)$ does not satisfy C_1 because it has a cycle (e_6, e_8, e_7). A directed cycle $\mathcal{C} = (e_6, -e_8, e_7)$ is obtained by reversing the direction of e_8. The vector $\boldsymbol{x}$ defined in Eq. (F12) is $\boldsymbol{x} = (1, 1, -1)^\top$ because $e_6,\ e_7 \in \mathcal{C}$, and $-e_8 \in \mathcal{C}$. Then, $\mathsf{B}^{\varepsilon}_{\setminus(1,\cdot)}\boldsymbol{x} = \boldsymbol{0}$ holds. (b) This $H(\varepsilon)$ does not satisfy $C_3(1)$ because it contains two loops, e_2 and e_3, in the component that does not contain $j = 1$. Then, the three columns in $\Gamma^{\varepsilon}_{\setminus(1,\cdot)}$, $\boldsymbol{w}^{(1)} := (-(R_2 - \langle R\rangle_{\boldsymbol{\pi}}), 0, 0)^\top$, $\boldsymbol{w}^{(2)} := (0, -(R_3 - \langle R\rangle_{\boldsymbol{\pi}}), 0)^\top$, and $\boldsymbol{w}^{(3)} := (M_{32}, -M_{23}, 0)^\top$ are linearly dependent. (c) This $H(\varepsilon)$ does not satisfy $C_4(1)$ because it has a loop, e_3, in the component containing $j = 1$. Then, the three columns in $\Gamma^{\varepsilon}_{\setminus(1,\cdot)}$ are linearly dependent. (d) This $H(\varepsilon)$ does not satisfy $C_6(3, 1)$ because $i = 3$ and $j = 1$ belong to different components. Then, two columns in $\mathsf{B}^{\varepsilon}_{\setminus(3,\cdot)}$ are linearly dependent.

c. $C_4(j)$

We show that $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}) = 0$ if $H(\varepsilon)$ does not satisfy $C_4(j)$ and satisfies C_1. If $H(\varepsilon)$ does not satisfy $C_4(j)$, then $H(\varepsilon)$ contains a component with at least one loop and vertex j. Let V' and E' be the set of vertices and the set of edges for this component, respectively. We can show $|E'| \geqslant |V'|$ because the component with $n_l(\geqslant 1)$ loops and without cycles, the number of the edges is $|E'| = |V'| - 1 + n_l$. Using $|E'| \geqslant |V'|$, we show that the columns of $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ are linearly dependent as follows. Note that the component contains vertex j, while $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ does not have the row corresponding to vertex j. Therefore, each column of $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ corresponding to an edge in E' can have nonzero elements only in the rows corresponding to $V' \setminus \{j\}$. Since $|E'| \geqslant |V'|$ implies $|E'| - 1 \geqslant |V' \setminus \{j\}|$, the $|E'|$ columns of $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ corresponding to E' can have at most $|E'| - 1$ nonzero elements in the rows corresponding to $V' \setminus \{j\}$. Thus, these $|E'|$ columns of $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ corresponding to the set E' are linearly dependent [see also Fig. 14(c) for an example]. Therefore, we conclude $\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}) = 0$, and thus $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}) = 0$ if $H(\varepsilon)$ does not satisfy $C_4(j)$ and satisfies C_1.

d. $C_6(i, j)$

We show that $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}) = 0$ if $H(\varepsilon)$ does not satisfy $C_6(i, j)$ and satisfies C_1, $C_3(j)$, and $C_4(j)$. We assume that $H(\varepsilon)$ satisfies C_1, $C_3(j)$, and $C_4(j)$, since otherwise $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}) = 0$ already holds. Since $C_6(i, j)$ is trivially satisfied when $j = i$, we consider only the case $j \neq i$. If $H(\varepsilon)$ does not satisfy $C_6(i, j)$ with $j \neq i$, then vertex i and vertex j belong to different components. We focus on the component that contains vertex i and does not contain vertex j. Let V' and E' be the set of vertices and the set of edges for this component, respectively. Since C_1, $C_3(j)$, and $C_4(j)$ are satisfied, this component has a loop and no cycles. We can show $|E'| = |V'|$ because the component is connected and contains a loop and no cycles. Using $|E'| = |V'|$, we can show that the columns of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ are linearly dependent as follows. The $|E'|$ columns of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ corresponding to E' can have nonzero elements only in the rows corresponding to $V' \setminus \{i\}$, since $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ does not have the row corresponding to vertex i. Since $|E'| = |V'|$ implies $|E'| - 1 = |V' \setminus \{i\}|$, each column corresponding to an edge in E' can have at most $|E'| - 1$ nonzero elements in the rows corresponding to

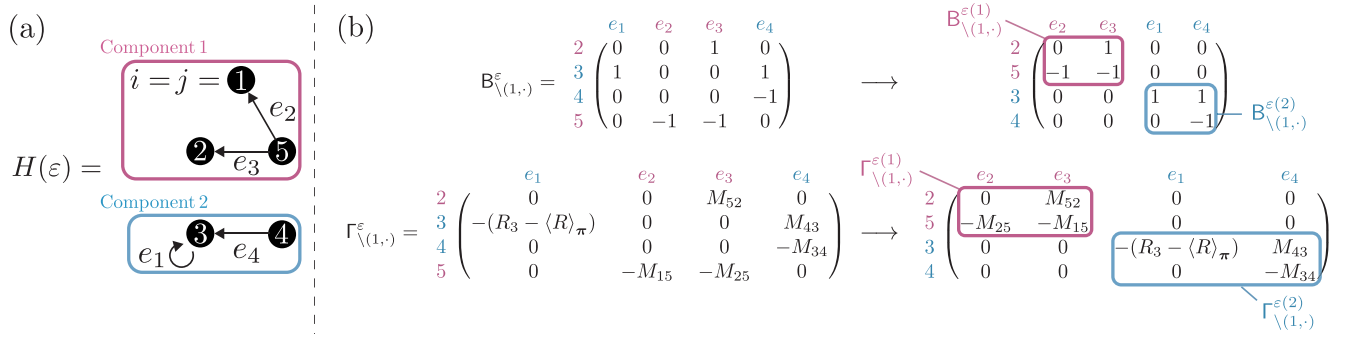

FIG. 15. An example of block diagonalization of $\mathsf{B}^{\varepsilon}_{\setminus(1,\cdot)}$ and $\Gamma^{\varepsilon}_{\setminus(1,\cdot)}$. (a) This $H(\varepsilon)$ consists of two components. (b) By appropriately swapping the rows and columns, $\mathsf{B}^{\varepsilon}_{\setminus(1,\cdot)}$ is transformed into a block diagonal matrix, and $\Gamma^{\varepsilon}_{\setminus(1,\cdot)}$ is also transformed into a block diagonal matrix. Each block corresponds to each component of the graph $H(\varepsilon)$.

$V' \setminus \{i\}$. Therefore, the $|E'|$ columns of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ corresponding to the set E' are linearly dependent [see also Fig. 14(d) for an example]. Therefore, we conclude $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}) = 0$, and thus $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}) = 0$ if $H(\varepsilon)$ does not satisfy $\mathrm{C}_6(i,j)$ and satisfies C_1, $\mathrm{C}_3(j)$, and $\mathrm{C}_4(j)$.

e. Another expression for Eq. (F9)

Combining the discussion above, we can restrict the summation in Eq. (F9) as

$$\det\left(\mathsf{L}_{\setminus(i,j)}\right) = \sum_{\substack{\varepsilon | H(\varepsilon) \text{ satisfies } \mathrm{C}_1,\mathrm{C}_3(j),\\ \mathrm{C}_4(j),\mathrm{C}_6(i,j), \text{ and } \mathrm{C}_7}} \det\left(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}\right), \tag{F14}$$

where the summation over ε is taken over all subsets of $E_R \cup E_M$ such that the corresponding graph $H(\varepsilon)$ satisfies C_1, $\mathrm{C}_3(j)$, $\mathrm{C}_4(j)$, $\mathrm{C}_6(i,j)$, and C_7.

5. Block diagonalization of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$

We show that $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ can be transformed into block diagonal matrices if $H(\varepsilon)$ satisfies all the conditions C_1, $\mathrm{C}_3(j)$, $\mathrm{C}_4(j)$, $\mathrm{C}_6(i,j)$, and C_7, which are necessary conditions for $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}) \neq 0$. The block diagonalization makes it easy to compute nonzero values of $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)})$.

If $H(\varepsilon)$ satisfies C_1, $\mathrm{C}_3(j)$, $\mathrm{C}_4(j)$, $\mathrm{C}_6(i,j)$, and C_7, we can introduce a procedure that transforms $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ into block diagonal matrices by swapping rows and columns as follows:

$$\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)} \to \begin{pmatrix} \mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)} & & \\ & \ddots & \\ & & \mathsf{B}^{\varepsilon(N_c)}_{\setminus(i,\cdot)} \end{pmatrix}, \tag{F15}$$

$$\Gamma^{\varepsilon}_{\setminus(j,\cdot)} \to \begin{pmatrix} \Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)} & & \\ & \ddots & \\ & & \Gamma^{\varepsilon(N_c)}_{\setminus(j,\cdot)} \end{pmatrix} \tag{F16}$$

(see also Fig. 15 for an example). Note that $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ or $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ corresponds to the αth component $H(\varepsilon)^{(\alpha)}$, where $H(\varepsilon)$ satisfies C_1, $\mathrm{C}_3(j)$, $\mathrm{C}_4(j)$, $\mathrm{C}_6(i,j)$, and C_7. Thus, N_c is the number of components in $H(\varepsilon)$. If $i = j$, the rows and columns are swapped in the same way for both matrices $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$. If $i \neq j$, the columns are swapped in the same way for both matrices $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$, and the rows except the rows corresponding to i and j are swapped in the same way for both matrices $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$. The row corresponding to j for the matrix $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ is placed in a row corresponding to the first block $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$, and the row corresponding to i for the matrix $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ is placed in the same row as the row corresponding to j for the matrix $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$. Thus, the row corresponding to i for the matrix $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ is placed in a row corresponding to the first block $\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}$. The first component corresponding to $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ contains the vertices i and j because of the condition $\mathrm{C}_6(i,j)$. Similarly, the first component corresponding to $\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}$ contains the vertices j and i.

First, we discuss a procedure for the block diagonalization of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ in Eq. (F15) more precisely. Let $V(H(\varepsilon)^{(\alpha)})$ be the set of vertices for the αth component corresponding to $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ or $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$. Let $E(H(\varepsilon)^{(\alpha)})$ be the set of edges for the αth component corresponding to $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ or $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$. For the first component, we have $|E(H(\varepsilon)^{(\alpha)})| = |V(H(\varepsilon)^{(\alpha)})| - 1$ because C_1 and $\mathrm{C}_4(i)$ imply that the first component is a spanning tree. We swap the $|V(H(\varepsilon)^{(1)})| - 1$ rows of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ corresponding to $V(H(\varepsilon)^{(1)}) \setminus \{i\}$ with the top $|V(H(\varepsilon)^{(1)})| - 1$ rows, preserving their relative order in $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$. We also swap the columns of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ corresponding to $E(H(\varepsilon)^{(1)})$ with the first $|E(H(\varepsilon)^{(1)})|$ columns, preserving their relative order in $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$. Note that the rows of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ corresponding to $V(H(\varepsilon)^{(1)}) \setminus \{i\}$ can have nonzero elements only in the columns corresponding to $E(H(\varepsilon)^{(1)})$, and the columns of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ corresponding to $E(H(\varepsilon)^{(1)})$ can have nonzero elements only in the rows corresponding to $V(H(\varepsilon)^{(1)}) \setminus \{i\}$. Because $|V(H(\varepsilon)^{(1)}) \setminus \{i\}| = |V(H(\varepsilon)^{(1)})| - 1 = |E(H(\varepsilon)^{(1)})|$, we obtain a $|E(H(\varepsilon)^{(1)})| \times |E(H(\varepsilon)^{(1)})|$ matrix, that is $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$. For the αth component with $\alpha \neq 1$, we have $|V(H(\varepsilon)^{(\alpha)})| = |E(H(\varepsilon)^{(\alpha)})|$ because C_1 and $\mathrm{C}_3(i)$ imply that the αth component contains no cycles and contains a loop with $\alpha \neq 1$. Note that $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ contains all the rows corresponding to $V(H(\varepsilon)^{(\alpha)})$ since the αth component does not contain vertex i with $\alpha \neq 1$. We swap the rows corresponding to $V(H(\varepsilon)^{(\alpha)})$ with the rows in the range from $(\sum_{\alpha'=1}^{\alpha-1} |V(H(\varepsilon)^{(\alpha')})| + 1)$th row to $\sum_{\alpha'=1}^{\alpha} |V(H(\varepsilon)^{(\alpha')})|$th row, preserving their relative order in $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$. We also swap the columns corresponding to

$E(H(\varepsilon)^{(\alpha)})$ with the columns in the same range, preserving their relative order in $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$. This swapping results in a $|E(H(\varepsilon)^{(\alpha)})| \times |E(H(\varepsilon)^{(\alpha)})|$ square matrix, which is $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ with $\alpha \neq 1$. Note that the rows of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ corresponding to $V(H(\varepsilon)^{(\alpha)})$ can have nonzero elements only in the columns corresponding to $E(H(\varepsilon)^{(\alpha)})$, and the columns of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ corresponding to $E(H(\varepsilon)^{(\alpha)})$ can have nonzero elements only in the rows corresponding to $V(H(\varepsilon)^{(\alpha)})$ with $\alpha \neq 1$. Thus, the block diagonalization of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ in Eq. (F15) is obtained in the above procedure. Note that this swapping can be done sequentially from $\alpha = 1$ to $\alpha = N_{\mathrm{c}}$ because $V(H(\varepsilon)^{(\alpha)}) \cap V(H(\varepsilon)^{(\beta)}) = \emptyset$ and $E(H(\varepsilon)^{(\alpha)}) \cap E(H(\varepsilon)^{(\beta)}) = \emptyset$ hold for any α and $\beta(\neq \alpha)$.

Second, we discuss a procedure for the block diagonalization of $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ in Eq. (F16). By swapping the rows of $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$, the row corresponding to the vertex i is placed in the same relative position as the row corresponding to the vertex j in $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$, and the other rows are also placed in the same relative positions as the other rows in $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$. After swapping the rows, we additionally do the same swapping for the rows and columns of $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ as we did for the rows and columns of $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ in the block diagonalization. Comparing the definition of B in Eq. (F5) with that of Γ in Eq. (F6), we find that $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ can have nonzero elements only at the same positions where $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ can have nonzero elements, except for the rows corresponding to vertices i and j. Note that the row of $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ corresponding to i can have nonzero elements only in the columns corresponding to $E(H(\varepsilon)^{(1)})$. Therefore, this swapping provides the block diagonalization of $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ in Eq. (F16).

The block diagonalizations in Eqs. (F15) and (F16) provides

$$\left|\det\left(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}\right)\right| = \prod_{\alpha=1}^{N_{\mathrm{c}}}\left|\det\left(\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}\right)\right|, \tag{F17}$$

because swapping rows or columns of a matrix does not affect the absolute value of its determinant, and block diagonalization factorizes the determinant.

We consider the sign when the absolute value symbols are removed from both sides of Eq. (F17). Note that swapping two rows multiplies the determinant by -1. By swapping two rows many times, we replace the row corresponding to i in the matrix $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ with the same position of the row corresponding to j in the matrix $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$ while preserving the relative order of the other rows in the matrices $\Gamma^{\varepsilon}_{\setminus(j,\cdot)}$ and $\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}$. If the swapping is only between adjacent rows, this replacement can be achieved by swapping $|i-j|-1$ times for $i \neq j$ and by swapping 0 times for $i = j$. After the replacement, we obtain the block diagonalizations in Eqs. (F15) and (F16) using the same swapping procedure. When the same swapping is performed on two matrices, the product of the determinants of the two matrices does not change sign even if the sign of the determinant of each matrix changes due to the swapping. Thus, we conclude that the sign change occurs only by swapping $|i-j|-1$ times for $i \neq j$ and does not occur for $i = j$. Because the factor of the determinant for swapping $|i-j|-1$ times is $(-1)^{|i-j|-1}$, we obtain

$$\det\left(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}\right) = \left\{(-1)^{|i-j|}(\delta_{ij}-1)+\delta_{ij}\right\}\prod_{\alpha=1}^{N_{\mathrm{c}}}\det\left(\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}\right). \tag{F18}$$

6. Expression for $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)})$ in terms of the weight of a graph

We express $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)})$ in Eq. (F18) in terms of the weight of a graph. Here we introduce the graph $\widetilde{H}(\varepsilon)$, which is defined as the graph obtained from $H(\varepsilon)$ by changing the edge directions according to the condition $\mathrm{C}_5(j)$. Thus, $\widetilde{H}(\varepsilon)$ satisfies all of the conditions for a rooted 0/1 loop forest in Eq. (F11), except for the condition $\widetilde{H}(\varepsilon) \subseteq G$. Note that $\widetilde{H}(\varepsilon)$ is not necessarily a subgraph of G because $\widetilde{H}(\varepsilon)$ can contain an edge $(k \leftarrow l)$ such that $k < l$, $M_{kl} = 0$, and $M_{lk} \neq 0$, and then $E(\widetilde{H}(\varepsilon)) \not\subseteq E(G)$. Therefore, $\widetilde{H}(\varepsilon)$ is not necessarily a rooted 0/1 loop forest. However, the weight of $\widetilde{H}(\varepsilon)$ is related to the weight of a rooted 0/1 loop forest.

To compute $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)})$, we consider each $\det(\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(\alpha)}_{\setminus(i,\cdot)})$ in Eq. (F18) individually. We show that $\det(\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)})$ can be expressed by the weight of $\widetilde{H}(\varepsilon)^{(\alpha)}$, which is the αth component of $\widetilde{H}(\varepsilon)$. Thus, we find that $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)})$ can be expressed by the weight of $\widetilde{H}(\varepsilon)$. We will discuss the case $\alpha \neq 1$, the case $\alpha = 1$ with $i = j$, and the case $\alpha = 1$ with $i \neq j$ one by one, and summarize the result.

a. Expression for $\det(\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)})$ *for* $\alpha \neq 1$

First, we label the vertices and edges of the αth component $H(\varepsilon)^{(\alpha)}$ with $\alpha \neq 1$. Here, we use the notation $\star$ for the labeled vertices $V(H(\varepsilon)^{(\alpha)}) = \{1^{\star}, \ldots, |E(H(\varepsilon)^{(\alpha)})|^{\star}\}$. We also introduce the total order $\prec$ (or $\succ$), such that $1^{\star} \prec 2^{\star} \prec \cdots \prec |E(H(\varepsilon)^{(\alpha)})|^{\star}$ (or $|E(H(\varepsilon)^{(\alpha)})|^{\star} \succ \cdots \succ 2^{\star} \succ 1^{\star}$). We label the vertices and edges, and change the edge directions according to the following steps [see also Fig. 16(a) for an example].

Step 1. Label the vertex containing the loop as vertex $1^{\star}$. Change the direction of all edges to point toward vertex $1^{\star}$. Because of the conditions C_1 and $\mathrm{C}_3(j)$, these processes are uniquely determined.

Step 2. Repeat the following process from $k^{\star} = 1^{\star}$ until all the vertices are labeled. If we can find a vertex that is the source vertex of an edge whose target vertex is $k^{\star}$, label this vertex as vertex $(k+1)^{\star}$. If we cannot find a vertex that is the source vertex of an edge whose target vertex is $k^{\star}$, find the source vertex of an edge whose target vertex is $l^{\star}$ with $l^{\star} \prec k^{\star}$ and label this vertex as vertex $(k+1)^{\star}$.

Step 3. Label the loop as $e_{1^{\star}}$. Label the edge from vertex $k^{\star}$ to vertex $l^{\star}(\prec k^{\star})$ as $e_{k^{\star}}$ for each $k^{\star} \in \{2^{\star}, \ldots, |E(H(\varepsilon)^{(\alpha)})|^{\star}\}$.

Note that this procedure is not unique, but its nonunique nature does not violate the validity of the later discussions. Note also that $E(\widetilde{H}(\varepsilon)^{(\alpha)}) = \{e_{k^{\star}} | k^{\star} = 1^{\star}, \ldots, |E(H(\varepsilon)^{(\alpha)})|^{\star}\}$ because the edges in $E(H(\varepsilon)^{(\alpha)})$ are inverted in this procedure according to the direction corresponding to the condition

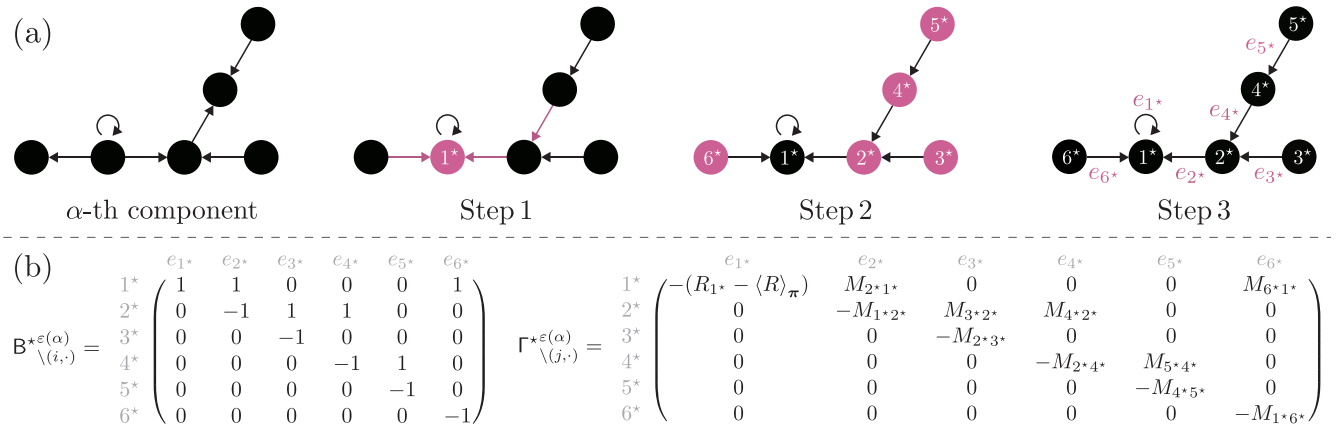

FIG. 16. (a) An example of labeling the vertices and edges with $\alpha \neq 1$. Step 1: Label the vertex with the loop as $1^\star$ and change the direction of the edges so that they point toward $1^\star$. Step 2: Starting from $l^\star = 1^\star$, label the other vertices. Step 3: Label the loop as $e_{1^\star}$, and the edge from $k^\star$ to $l^\star(\prec k^\star)$ as $e_{k^\star}$ for $k^\star \in \{2^\star, \ldots, |E(H(\varepsilon)^{(\alpha)})|^\star\}$. (b) Upper triangular matrices $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ obtained by the procedure in panel (a).

$\mathrm{C}_5(j)$, and $\widetilde{H}(\varepsilon)$ is defined as the graph obtained from $H(\varepsilon)$ by changing the edge directions according to the condition $\mathrm{C}_5(j)$. This procedure is used to define the upper triangle matrices $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$, which are the generalized incidence matrix and the generalized weighted incidence matrix corresponding to $\widetilde{H}(\varepsilon)^{(\alpha)}$, respectively. The matrices $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ for $\alpha \neq 1$ are obtained by applying the transformation into $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ according to the procedure, respectively. Let $\sigma(k)$ be the vertex $k^\star$ for indices $k = 1, \ldots, |E(H(\varepsilon)^{(\alpha)})|$. We rearrange the rows and columns of $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ so that the kth row corresponds to vertex $\sigma(k)$ and the lth column corresponds to $e_{\sigma(l)}$. Additionally, we also multiply the columns of $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ corresponding to the edge $e_{k^\star}$ by -1 if the edge $e_{k^\star}$ is inverted from the corresponding edge in $H(\varepsilon)^{(\alpha)}$ in Step 1. This is because inverting the edge $e \to -e$ gives the factor -1 for the elements of $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ in Eqs. (F5) and (F6) as $\delta_{i\mathrm{t}(-e)} - \delta_{i\mathrm{s}(-e)} = (-1)(\delta_{i\mathrm{t}(e)} - \delta_{i\mathrm{s}(e)})$ and $M_{\mathrm{s}(-e)\mathrm{t}(-e)}\delta_{i\mathrm{t}(-e)} - M_{\mathrm{t}(-e)\mathrm{s}(-e)}\delta_{i\mathrm{s}(-e)} = (-1)(M_{\mathrm{s}(e)\mathrm{t}(e)}\delta_{i\mathrm{t}(e)} - M_{\mathrm{t}(e)\mathrm{s}(e)}\delta_{i\mathrm{s}(e)})$ for $e \in E_M$.

We show that $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ are upper triangular matrices. Here, we define $R_{k^\star}$ and $M_{k^\star l^\star}$ as $R_{k^\star} = R_i$ and $M_{k^\star l^\star} = M_{ij}$ where $i \in V(G)$ and $j \in V(G)$ are the vertices before relabeling, corresponding to $k^\star$ and $l^\star$, respectively. The diagonal elements of $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ are given by $(\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)})_{11} = 1$, $(\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)})_{kk} = -1$ for $k \geqslant 2$, $(\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)})_{11} = -(R_{\sigma(1)} - \langle R\rangle_{\boldsymbol{\pi}})$, and $(\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)})_{kk} = -M_{\mathrm{t}(e_{\sigma(k)})\mathrm{s}(e_{\sigma(k)})}$ for $k \geqslant 2$ [see also Fig. 16(b) for an example]. Since the vertex with the loop is labeled $1^\star$ and the loop is labeled $e_{1^\star}$, the first column of $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ have nonzero elements only in the first row, and these nonzero elements are given by $(\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)})_{11} = 1$ and $(\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)})_{11} = -(R_{\sigma(1)} - \langle R\rangle_{\boldsymbol{\pi}})$. Since the labeling in Step 3 implies $\mathrm{s}(e_{k^\star}) = k^\star$ for $k^\star \in \{2^\star, \ldots, |E(H(\varepsilon)^{(\alpha)})|^\star\}$, we obtain $(\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)})_{kk} = -1$ and $(\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)})_{kk} = -M_{\mathrm{t}(e_{\sigma(k)})\mathrm{s}(e_{\sigma(k)})}$ for $k \geqslant 2$. Since the vertices are labeled in ascending order, we find that $\mathrm{s}(e_{k^\star}) = k^\star$ and $\mathrm{t}(e_{k^\star}) \prec k^\star$ for $k^\star \in \{2^\star, \ldots, |E(H(\varepsilon)^{(\alpha)})|^\star\}$. Therefore, the kth column of $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ can have nonzero elements for the lth row and the kth row, where the lth row corresponds to the row for vertex $\mathrm{t}(e_{k^\star})$ and thus $l < k$ is satisfied. Thus, $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ are upper triangular matrices.

Based on the transformations of $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ into $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$, we provide the expression for $\det(\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)})$ in terms of the weight of $\widetilde{H}(\varepsilon)^{(\alpha)}$. Note that this transformation can be achieved by applying the same procedure to both $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$, since their rows and columns are arranged in the same way. Because the transformation only changes the signs of the determinants of $\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ in the same way, the transformation does not change the value of the product of the determinants, i.e., $\det(\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}) = \det(\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)})\det(\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)})$. Moreover, the determinants of $\mathsf{B}^{\star\varepsilon(\alpha)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(\alpha)}_{\setminus(j,\cdot)}$ are given by the products of their diagonal elements, since they are upper triangular matrices. Combining these facts, we obtain

$$\det\left(\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon(\alpha)}_{\setminus(j,\cdot)}\right) = -(R_{\sigma(1)} - \langle R\rangle_{\boldsymbol{\pi}}) \prod_{k=2}^{|E(H(\varepsilon)^{(\alpha)})|} M_{\mathrm{t}(e_{\sigma(k)})\mathrm{s}(e_{\sigma(k)})}. \tag{F19}$$

Note that $-(R_{\sigma(1)} - \langle R\rangle_{\boldsymbol{\pi}}) = -(R_{1^\star} - \langle R\rangle_{\boldsymbol{\pi}})$ is the weight of the loop $e_{1^\star}$, and $M_{\mathrm{t}(e_{\sigma(k)})\mathrm{s}(e_{\sigma(k)})} = M_{\mathrm{t}(e_{k^\star})\mathrm{s}(e_{k^\star})}$ is the weight of edge $e_{k^\star}$ for $k^\star \in \{2^\star, \ldots, |E(H(\varepsilon)^{(\alpha)})|^\star\}$. Combining these facts with the expression $E(\widetilde{H}(\varepsilon)^{(\alpha)}) = \{e_{k^\star}|k^\star = 1^\star, \ldots, |E(H(\varepsilon)^{(\alpha)})|^\star\}$, we find

$$\det\left(\mathsf{B}^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon(\alpha)}_{\setminus(i,\cdot)}\right) = w\left(\widetilde{H}(\varepsilon)^{(\alpha)}\right). \tag{F20}$$

b. Expression for $\det(\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(1)}_{\setminus(i,\cdot)})$

For $i = j$, we introduce the procedure to obtain $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)} = \Gamma^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$, which are upper triangular matrices corresponding to $\widetilde{H}(\varepsilon)^{(1)}$. First, we label the vertices and edges of $H(\varepsilon)^{(1)}$, and change the edge directions according to the following steps [see also Fig. 17(a) for an example].

Step 1. Label vertex i as vertex $1^\star$. Change the direction of all edges so that they point toward vertex $1^\star$.

Step 2. Repeat the process introduced in Step 2 in Appendix F 6 a from $k^\star = 1^\star$ until all the vertices are labeled.

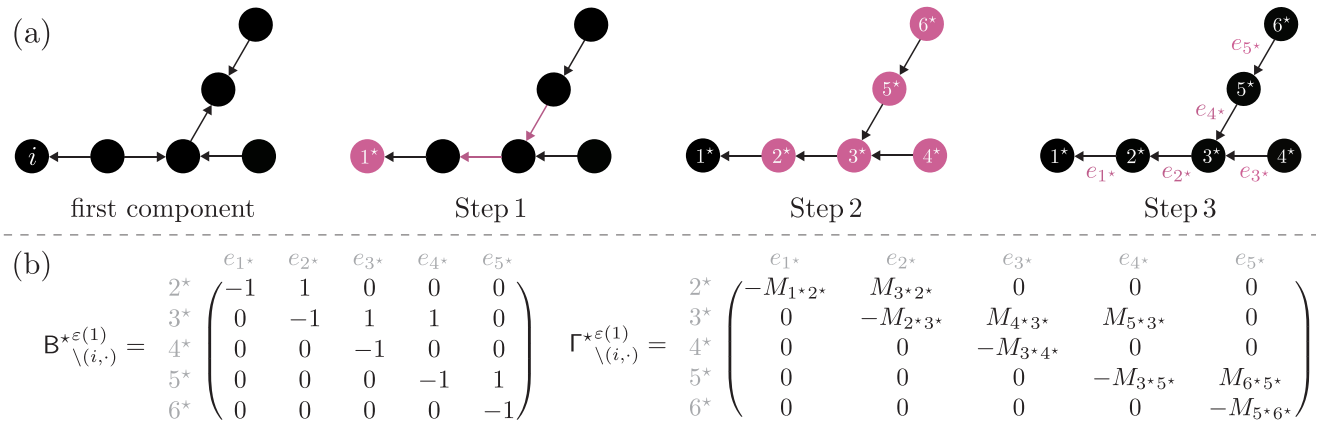

FIG. 17. (a) An example of labeling the vertices and edges with $\alpha = 1$. Step 1: Label vertex i as $1^\star$ and change the direction of the edges so that they point toward $1^\star$. Step 2: Starting from vertex $1^\star$, label the other vertices. Step 3: Label the edge from $(k+1)^\star$ to $l^\star(\prec (k+1)^\star)$ as $e_{k^\star}$. (b) Upper triangular matrices $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ obtained by the procedure in panel (a).

Step 3. Label the edge from vertex $(k+1)^\star$ to $l^\star(\prec (k+1)^\star)$ as $e_{k^\star}$ for each $k^\star \in \{1^\star, \ldots, |E(H(\varepsilon)^{(1)})|^\star\}$.
Note that $E(\widetilde{H}(\varepsilon)^{(1)}) = \{e_{k^\star} | k^\star = 1^\star, \ldots, |E(H(\varepsilon)^{(1)})|^\star\}$ because the edges in $E(H(\varepsilon)^{(1)})$ are inverted in this procedure according to the direction corresponding to the condition $\mathrm{C}_5(j) = \mathrm{C}_5(i)$. Then, we rearrange the rows and columns of $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(1)}_{\setminus(i,\cdot)}$ so that the kth row corresponds to vertex $(k+1)^\star$, and lth column corresponds to edge $e_{l^\star}$. Additionally, we also multiply the columns of $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(1)}_{\setminus(i,\cdot)}$ corresponding to the edge $e_{k^\star}$ by -1 if the edge $e_{k^\star}$ is inverted from the corresponding edge in $H(\varepsilon)^{(1)}$ in Step 1. Through this process, we obtain $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ from $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(1)}_{\setminus(i,\cdot)}$, respectively. For the same reason as in the case of $\alpha \neq 1$, $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ are upper triangular matrices, and their diagonal elements are given by $(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kk} = -1$ and $(\Gamma^{\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kk} = -M_{\mathrm{t}(e_{\sigma(k)})\mathrm{s}(e_{\sigma(k)})}$ [see also Fig. 17(b) for an example].

We then express $\det(\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(1)}_{\setminus(i,\cdot)})$ in terms of the weight of $\widetilde{H}(\varepsilon)^{(1)}$. Note that this transformation can be achieved by applying the same procedure to both $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(1)}_{\setminus(i,\cdot)}$. For the same reason as in the case of $\alpha \neq 1$, we obtain $\det(\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(1)}_{\setminus(i,\cdot)}) = \det(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})\det(\Gamma^{\star\varepsilon(1)}_{\setminus(i,\cdot)})$. Furthermore, since the matrices $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ are upper triangular, their determinants are equal to the product of their diagonal elements. Combining these facts with the expression $E(\widetilde{H}(\varepsilon)^{(1)}) = \{e_{k^\star} | k^\star = 1^\star, \ldots, |E(H(\varepsilon)^{(1)})|^\star\}$, we obtain

$$\det\left(\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon(1)}_{\setminus(i,\cdot)}\right) = \prod_{k=1}^{|E(H(\varepsilon)^{(1)})|} M_{\mathrm{t}(e_{\sigma(k)})\mathrm{s}(e_{\sigma(k)})} = w(\widetilde{H}(\varepsilon)^{(1)}). \tag{F21}$$

c. *Expression for* $\det(\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)})$ *with* $j \neq i$

For $j \neq i$, we first introduce the procedure for transforming $\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}$ into an upper triangular matrix $\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)}$. We label the vertices and edges of $H(\varepsilon)^{(1)}$ and change the edge directions according to the following steps [see also Fig. 18(a) for an example].

Step 1. Label vertex j as vertex $1^\star$. Change the direction of all the edges so that they point toward vertex $1^\star$.

Step 2. Find a path from vertex i to vertex j. Following the path, label the vertices in ascending order $1^\star, 2^\star, \ldots, (L+1)^\star$, where vertex i is labeled as vertex $(L+1)^\star$. Repeat the process introduced in Step 2 in Appendix F 6 a from $k^\star = (L+1)^\star$ until all the vertices are labeled. Because of $\mathrm{C}_6(i,j)$, these processes are possible.

Step 3. Label the edge from vertex $(k+1)^\star$ to $m^\star(\prec (k+1)^\star)$ as $e_{k^\star}$ for each $k^\star \in \{1^\star, \ldots, |E(H(\varepsilon)^{(1)})|^\star\}$.
Note that $E(\widetilde{H}(\varepsilon)^{(1)}) = \{e_{k^\star} | k^\star = 1^\star, \ldots, |E(H(\varepsilon)^{(1)})|^\star\}$ because the edges in $E(H(\varepsilon)^{(1)})$ are inverted in this procedure according to the direction corresponding to the condition $\mathrm{C}_5(j)$.

The matrices $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)}$ for $j \neq i$ are obtained by applying the transformations of $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}$ according to the procedure, respectively. We rearrange the rows and columns of $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}$ so that the kth row corresponds to vertex $(k+1)^\star$ and the lth column corresponds to $e_{l^\star}$. Additionally, we also multiply the columns of $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}$ corresponding to the edge $e_{k^\star}$ by -1 if the edge $e_{k^\star}$ is inverted from the corresponding edge in $H(\varepsilon)^{(1)}$ in Step 1. As a result, $\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)}$ becomes an upper triangular matrix whose diagonal elements are given by $(\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)})_{kk} = -M_{\mathrm{t}(e_{\sigma(k)})\mathrm{s}(e_{\sigma(k)})}$ [see also Fig. 18(b) for an example]. This fact can be shown in the same way as in Appendix F 6 a.

Here, $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ is not an upper triangular matrix in general [see also Fig. 18(b) for an example]. For $k \geqslant L+1$ and $l < k$, $(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kk} = -1$ and $(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kl} = 0$ hold where L is the number of edges in the path from vertex i to vertex j [see also Fig. 18(b) for an example]. These facts can be shown as follows. The rows and columns in $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}$ are arranged in the same way, except that the row for vertex j in $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ is in the same position as the row for vertex i in $\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}$. The row corresponding to vertex i is in the Lth row in $\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)}$ because in Step 2, we label the vertices on the path from vertex j to vertex i as $1^\star, \ldots, (L+1)^\star$ and the row corresponding to vertex j is missing in $\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)}$. Therefore, the Lth row of $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ corresponds to vertex j. Since the target vertex of $e_{1^\star}$ is vertex j, we find $(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})_{L1} = 1$, which means that $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ is not an upper triangular matrix if $L \geqslant 2$. Note that the rows and

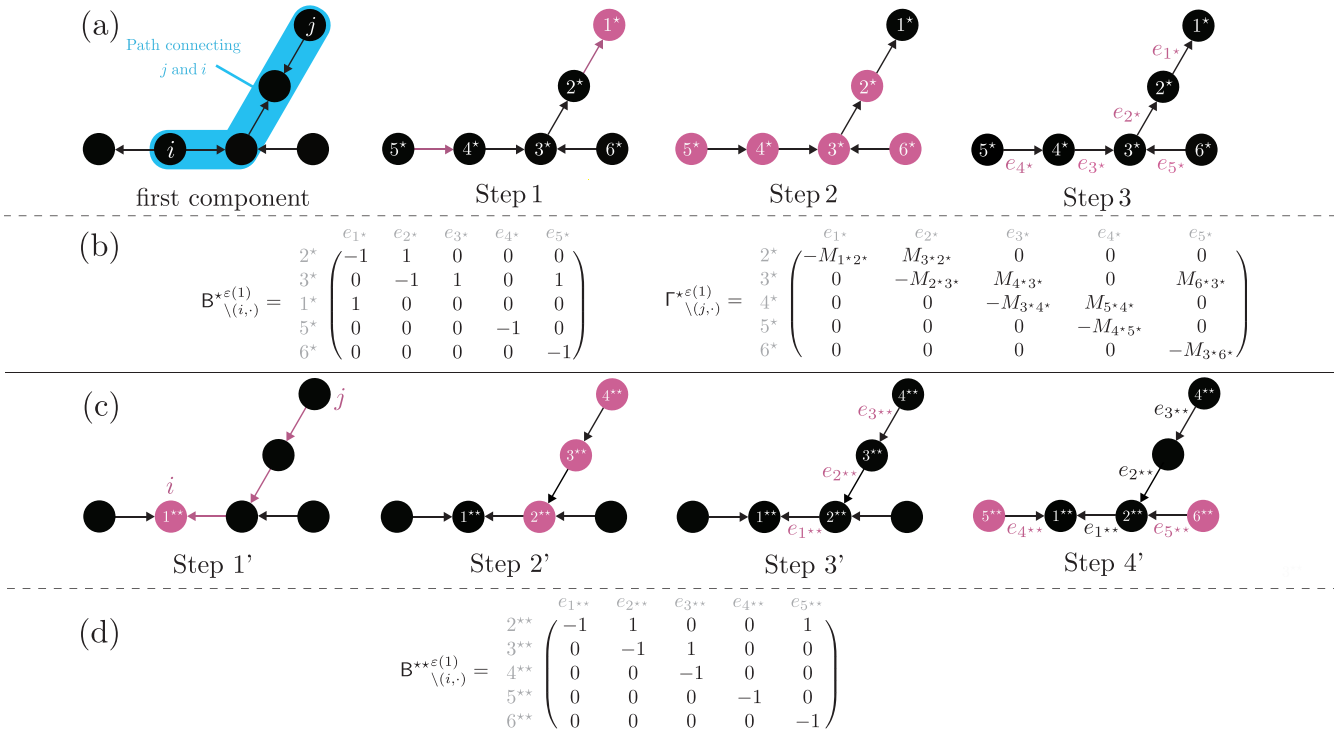

FIG. 18. (a) An example of labeling the vertices and edges with $\alpha = 1$ and $j \neq i$. Step 1: Label vertex j as $1^\star$ and change the direction of the edges so that they point toward $1^\star$. Step 2: Starting from vertex $1^\star$, label the other vertices. Step 3: Label the edge from $(k+1)^\star$ to $m^\star(\prec (k+1)^\star)$ as $e_{k^\star}$. (b) The procedure in panel (a) yields an upper triangular matrix $\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)}$, while the yielding matrix $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ is not an upper triangular matrix. Here, the number of edges in the path from vertex i to vertex j is $L = 3$. (c) An example of labeling the vertices and edges. Step 1': Label vertex i as $1^{\star\star}$ and change the direction of the edges so that they point toward $1^{\star\star}$. Step 2': Starting from vertex $1^{\star\star}$, label the vertices in the path from j to i. Step 3': Label the edge from $(k+1)^{\star\star}$ to $m^{\star\star}(\prec (k+1)^{\star\star})$ as $e_{k^{\star\star}}$. Step 4': Label the remaining vertices and edges according to the procedure in panel (a). (d) The procedure in panel (c) yields an upper triangular matrix $\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)}$

columns from the $(L+1)$th to the $|E(H(\varepsilon)^{(1)})|$th in $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ are arranged in the same order as in $\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)}$. Thus, as in the case of $j = i$ in Appendix F 6 b, we have $(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kk} = -1$ and $(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kl} = 0$ for $k \geqslant L+1$ and $l < k$.

We consider the transformation of $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ into an upper triangular matrix. We relabel the vertices and edges of $\widetilde{H}(\varepsilon)^{(1)}$. First, we relabel the vertices and edges of $H(\varepsilon)^{(1)}$. Here, we use the notation ** for the relabeled vertices $V(H(\varepsilon)^{(\alpha)}) = \{1^{\star\star}, \ldots, |E(H(\varepsilon)^{(\alpha)})|^{\star\star}\}$. We also introduce the total order $\prec$ (or $\succ$), such that $1^{\star\star} \prec 2^{\star\star} \prec \cdots \prec |E(H(\varepsilon)^{(\alpha)}|^{\star\star}$ (or $|E(H(\varepsilon)^{(\alpha)}|^{\star\star} \succ \cdots \succ 2^{\star\star} \succ 1^{\star\star}$). We relabel the vertices and edges, and change the edge directions according to the following steps [see also Fig. 18(c) for an example],

Step 1'. Label vertex i as vertex $1^{\star\star}$. Change the direction of all edges so that they point toward vertex $1^{\star\star}$.

Step 2'. Find a path from vertex j to vertex i. Following the path, label the vertices in ascending order $1^{\star\star}, 2^{\star\star}, \ldots, (L+1)^{\star\star}$, where vertex j is labeled as vertex $(L+1)^{\star\star}$.

Step 3'. Label the edge from vertex $(k+1)^{\star\star}$ to $m^{\star\star}(\prec (k+1)^{\star\star})$ as $e_{k^{\star\star}}$ for each $k^{\star\star} \in \{1^{\star\star}, \ldots, L^{\star\star}\}$.

Step 4'. Label the remaining vertex $k^\star$ as $k^{\star\star}$ and the remaining edge $e_{k^\star}$ as $e_{k^{\star\star}}$, respectively, for each $k^\star \in \{(L+1)^\star, \ldots, |E(H(\varepsilon)^{(1)})|^\star\}$.

We rearrange the rows and columns of $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ so that the kth row corresponds to vertex $k^{\star\star}$ and the lth row corresponds to $e_{l^{\star\star}}$. Additionally, we also multiply the columns of $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ corresponding to the edge $e_{k^{\star\star}}$ by -1 if the edge $e_{k^{\star\star}}$ is inverted from the corresponding edge in $\widetilde{H}(\varepsilon)^{(1)}$ in Step 1'. Let $\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)}$ denote the resulting matrix.

We show that $\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)}$ is an upper triangular matrix whose diagonal elements are given by $(\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kk} = -1$ [see also Fig. 18(d) for an example]. As implied in step 4', this transformation does not affect the $(L+1)$th through the $|E(H(\varepsilon)^{(1)})|$th rows and columns. Therefore, $\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)}$ satisfies $(\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kk} = -1$ and $(\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kl} = 0$ for $k \geqslant L+1$ and $l < k$, as $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ does. Furthermore, we show that $(\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kk} = -1$ and $(\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kl} = 0$ for $k \in \{1, \ldots, L\}$ and $l < k$ as follows. Since the source vertex of $e_{k^{\star\star}}$ is vertex $(k+1)^{\star\star}$ corresponding to the kth row for $k \in \{1, \ldots, L\}$, we obtain $(\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kk} = -1$. Since $e_{1^{\star\star}}, \ldots, e_{L^{\star\star}}$ forms the path from vertex $(L+1)^{\star\star}$ to vertex $1^{\star\star}$, $e_{l^{\star\star}} = (l^{\star\star} \leftarrow (l+1)^{\star\star})$ holds for all $l^{\star\star} \in \{1^{\star\star}, \ldots, L^{\star\star}\}$. Therefore, $(\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)})_{kl} = 0$ also holds for $k \in \{1, \ldots, L\}$ and $l < k$.

We show that the transformation of $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ into $\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)}$ reverses the sign of $\det(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})$, i.e.,

$$\det\left(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}\right) = -\det\left(\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)}\right). \tag{F22}$$

In Step 1', we invert L edges, $e_{1^\star}, \ldots, e_{L^\star}$. This procedure corresponds to multiplying the first L columns in $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ by -1. In Step 2', we label the vertices $2^\star, \ldots, L^\star$ as $L^{\star\star}, \ldots, 2^{\star\star}$, respectively. This procedure rearranges the first to the $(L-1)$th rows in $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ in reverse order, which can be achieved by swapping two rows $(L-2)/2$ times (or $(L-1)/2$ times) if L is even (or odd). In Step 3', we relabel the edges $e_{1^\star}, \ldots, e_{L^\star}$ as $e_{L^{\star\star}}, \ldots, e_{1^{\star\star}}$. This procedure rearranges the first to the Lth columns in $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ in reverse order, which can be achieved by swapping two columns $L/2$ times (or $(L-1)/2$ times) if L is even (or odd). Overall, the transformation of $\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)}$ into $\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)}$ reverses the sign of $\det(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})$ exactly $2L-1$ times, regardless of whether L is odd or even. Therefore, this procedure multiplies $\det(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})$ by $(-1)^{2L-1} = -1$, as shown in Eq. (F22).

We now express $\det(\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)})$ in terms of the weight of $\widetilde{H}(\varepsilon)^{(1)}$. Note that $\det(\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}) = \det(\mathsf{B}^{\star\varepsilon(1)}_{\setminus(i,\cdot)})\det(\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)})$ holds because the same procedure is applied to $\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}$. Combining this equality with Eq. (F22), we obtain

$$\begin{aligned}\det\left(\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}\right) &= -\det\left(\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)}\right)\\ &= -\prod_{k=1}^{|E(H(\varepsilon)^{(1)})|} M_{\mathrm{t}(e_{\sigma(k)})\mathrm{s}(e_{\sigma(k)})}\\ &= -w\big(\widetilde{H}(\varepsilon)^{(1)}\big),\end{aligned} \tag{F23}$$

where we used the fact that $\mathsf{B}^{\star\star\varepsilon(1)}_{\setminus(i,\cdot)}$ and $\Gamma^{\star\varepsilon(1)}_{\setminus(j,\cdot)}$ are upper triangular matrices. The third equality follows from $E(\widetilde{H}(\varepsilon)^{(1)}) = \{e_{k^\star}|k^\star = 1^\star, \ldots, |E(H(\varepsilon)^{(1)})|^\star\}$.

d. Expression for $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)})$ *in terms of the weight of* $\widetilde{H}(\varepsilon)$

We now express $\det(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)})\det(\Gamma^{\varepsilon}_{\setminus(j,\cdot)})$ for $H(\varepsilon)$ which satisfies C_1, $\mathrm{C}_3(j)$, $\mathrm{C}_4(j)$, $\mathrm{C}_6(i,j)$, and C_7 in terms of the weight of $\widetilde{H}(\varepsilon)$. Here, $\widetilde{H}(\varepsilon)$ is expressed as $\widetilde{H}(\varepsilon) = \bigcup_{\alpha=1}^{N_\mathrm{c}} \widetilde{H}(\varepsilon)^{(\alpha)}$, where N_c is the number of components in $\widetilde{H}(\varepsilon)$. Combining the result in Eq. (F21) for $i=j$ with the result in Eq. (F23) for $i \neq j$, we obtain an expression

$$\det\left(\mathsf{B}^{\varepsilon(1)}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon(1)}_{\setminus(j,\cdot)}\right) = (2\delta_{ij}-1)w\big(\widetilde{H}(\varepsilon)^{(1)}\big). \tag{F24}$$

By combining Eqs. (F18) and (F20) with (F24), we obtain

$$\det\left(\mathsf{B}^{\varepsilon}_{\setminus(i,\cdot)}\right)\det\left(\Gamma^{\varepsilon}_{\setminus(j,\cdot)}\right) = \left\{\delta_{ij} + (1-\delta_{ij})(-1)^{|i-j|}\right\} w\big(\widetilde{H}(\varepsilon)\big), \tag{F25}$$

where we used $w(\widetilde{H}(\varepsilon)) = \prod_{\alpha=1}^{N_\mathrm{c}} w(\widetilde{H}(\varepsilon)^{(\alpha)})$, which follows from $\widetilde{H}(\varepsilon) = \bigcup_{\alpha=1}^{N_\mathrm{c}} \widetilde{H}(\varepsilon)^{(\alpha)}$.

7. Expression for $\zeta_i\pi_j$ in terms of the weights of rooted 0/1 loop forests

We now derive Eq. (F1), which is an expression for $\zeta_i\pi_j$ in terms of the weights of rooted 0/1 loop forests. To derive Eq. (F1), we express $\det(\mathsf{L}_{\setminus(i,j)})$ in Eq. (F2) in terms of the weights of rooted 0/1 loop forests.

We express $\det(\mathsf{L}_{\setminus(i,j)})$ in terms of the weights of graphs. By substituting Eq. (F25) into Eq. (F14), we obtain

$$\det\left(\mathsf{L}_{\setminus(i,j)}\right) = \left\{\delta_{ij} + (1-\delta_{ij})(-1)^{|i-j|}\right\} \sum_{H\in\widetilde{\mathcal{F}}_{j\leftarrow i}(G)} w(H), \tag{F26}$$

where $\widetilde{\mathcal{F}}_{j\leftarrow i}(G)$ is the set of all $\widetilde{H}(\varepsilon)$ satisfying C_1, $\mathrm{C}_3(j)$, $\mathrm{C}_4(j)$, $\mathrm{C}_5(j)$, $\mathrm{C}_6(i,j)$, and C_7. We can restrict the summation in Eq. (F26) to graphs $H(\in \widetilde{\mathcal{F}}_{j\leftarrow i}(G))$ such that $\tilde{w}(e) \neq 0$ holds for every edge $e \in E(H)$ that is not a loop. This is because the weight $w(H)$ also vanishes, and thus such H does not contribute to $\det(\mathsf{L}_{\setminus(i,j)})$ if $H \in \widetilde{\mathcal{F}}_{j\leftarrow i}(G)$ contains an edge e that is not a loop and whose weight vanishes. Therefore, we rewrite Eq. (F26) as

$$\det\left(\mathsf{L}_{\setminus(i,j)}\right) = \left\{\delta_{ij} + (1-\delta_{ij})(-1)^{|i-j|}\right\} \sum_{H\in\widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G)} w(H), \tag{F27}$$

where $\widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G)$ is the set of all graphs $H \in \widetilde{\mathcal{F}}_{j\leftarrow i}(G)$ such that $\tilde{w}(e) \neq 0$ holds for every edge $e \in E(H)$ that is not a loop.

We prove the identity $\widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G) = \mathcal{F}_{j\leftarrow i}(G)$, which allows us to replace $\widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G)$ by $\mathcal{F}_{j\leftarrow i}(G)$ in Eq. (F27). Suppose that H belongs to $\widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G)$. Here, $H(\in \widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G))$ satisfies C_1, $\mathrm{C}_3(j)$, $\mathrm{C}_4(j)$, $\mathrm{C}_5(j)$, $\mathrm{C}_6(i,j)$, and C_7. Since every edge in H that is not a loop has nonzero weight, it belongs to $E(G)$ by definition Eq. (13). Furthermore, every loop in H belongs to $E(G)$, and every vertex of H belongs to $V(G)$. Therefore, we have $H \subseteq G$. Combining these facts, we find that H satisfies all the conditions in Eq. (F11), which means $H \in \mathcal{F}_{j\leftarrow i}(G)$. Next, suppose that H belongs to $\mathcal{F}_{j\leftarrow i}(G)$. We consider that the edge directions of H are changed such that every nonloop edge $e = (k \leftarrow l)$ satisfies $k < l$. The resulting graph is $H(\varepsilon)$ for a set of edges $\varepsilon \in E_R \cup E_M$ that satisfies C_1, $\mathrm{C}_3(j)$, $\mathrm{C}_4(j)$, $\mathrm{C}_6(i,j)$, and C_7. We consider that the edge directions of $H(\varepsilon)$ are changed according to $\mathrm{C}_5(j)$. The resulting graph is $\widetilde{H}(\varepsilon)$ corresponding to $H(\varepsilon)$. Note that this graph $\widetilde{H}(\varepsilon)$ is nothing but the original graph $H \in \mathcal{F}_{j\leftarrow i}(G)$ because the edge directions are recovered by changing the edge directions according to $\mathrm{C}_5(j)$. Thus, we find that $H = \widetilde{H}(\varepsilon) \in \widetilde{\mathcal{F}}_{j\leftarrow i}(G)$. Moreover, since every edge in H that is not a loop has nonzero weight, we obtain $H \in \widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G)$. Therefore, $\mathcal{F}_{j\leftarrow i}(G) \subseteq \widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G)$ and $\widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G) \subseteq \mathcal{F}_{j\leftarrow i}(G)$ are proved, and thus $\widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G) = \mathcal{F}_{j\leftarrow i}(G)$.

Finally, we derive Eq. (F1). Since $\widetilde{\mathcal{F}}^{+}_{j\leftarrow i}(G) = \mathcal{F}_{j\leftarrow i}(G)$ holds as shown above, we can rewrite Eq. (F27) as

$$\det\left(\mathsf{L}_{\setminus(i,j)}\right) = \left\{\delta_{ij} + (1-\delta_{ij})(-1)^{|i-j|}\right\} \sum_{H\in\mathcal{F}_{j\leftarrow i}(G)} w(H). \tag{F28}$$

Substituting this equality into Eq. (F2), we obtain

$$\zeta_i\pi_j = \frac{1}{Z}\sum_{H\in\mathcal{F}_{j\leftarrow i}(G)} w(H) \tag{F29}$$

with $Z := \sum_{i\in V(G)}\sum_{H\in\mathcal{F}_i(G)} w(H)$, where we used $\{\delta_{ij} + (1-\delta_{ij})(-1)^{|i-j|}\}(-1)^{i+j} = 1$ and $Z =$

$\sum_k \sum_{H\in\mathcal{F}_{k\leftarrow k}(G)} w(H)$. This equation is Eq. (F1), i.e., Lemma 1.

APPENDIX G: MAXIMUM NUMBER OF ROOTED 0/1 LOOP FORESTS

We determine the maximum number of rooted 0/1 loop forests for a given basic graph. The number of graphs in $\mathcal{F}_i(G)$ is bounded by

$$|\mathcal{F}_i(G)| \leqslant 2(N+1)^{N-2}, \tag{G1}$$

with $N = |V(G)|$. Since $\mathcal{F}_{i\leftarrow j}(G)$ and $\mathcal{F}_{i\not\leftarrow j}(G)$ are subsets of $\mathcal{F}_i(G)$, the maximum number of them is also bounded by $2(N+1)^{N-2}$. Note that the number of rooted spanning trees with N vertices is bounded by N^{N-2}, i.e., $|\mathcal{T}_i(G)| \leqslant N^{N-2}$ holds with $N = |V(G)|$, which is known as Cayley's formula [57].

The inequality (G1) is proved as follows. To derive the upper bound on $|\mathcal{F}_i(G)|$, we consider the *complete graph* with N vertices G_N^{comp}, which is defined as a basic graph where each vertex has a loop and there are edges in both directions between each pair of vertices. The set of 0/1 loop forests $\mathcal{F}_i(G_N^{\mathrm{comp}})$ satisfies $\mathcal{F}_i(G) \subseteq \mathcal{F}_i(G_N^{\mathrm{comp}})$ for any basic graph G with N vertices. Therefore, we obtain

$$|\mathcal{F}_i(G)| \leqslant \left|\mathcal{F}_i\left(G_N^{\mathrm{comp}}\right)\right| = \sum_{H\in\mathcal{F}_i(G_N^{\mathrm{comp}})} 1. \tag{G2}$$

If we compare the right-hand side of this inequality with the right-hand side of Eq. (F28) when $i = j$, we notice that $\sum_{H\in\mathcal{F}_i(G_N^{\mathrm{comp}})} 1$ can be obtained by replacing every $R_k - \langle R\rangle_{\boldsymbol{\pi}}$ with -1 and every M_{kl} with 1 for $k \neq l$, because this replacement turns every $w(H)$ in Eq. (F28) to 1. Because $M_{kk} = -\sum_{l(\neq k)} M_{lk}$ is replaced with $-(N-1)$, applying this replacement transforms $\mathsf{L} = -(\mathsf{R} - \langle R\rangle_{\boldsymbol{\pi}}\mathsf{I} + \mathsf{M})$ into $\tilde{\mathsf{L}} \in \mathbb{R}^{N\times N}$, which is defined as $\tilde{L}_{kl} := N$ if $k = l$, and $\tilde{L}_{kl} := -1$ if $k \neq l$. Then, we obtain

$$\begin{aligned}
|\mathcal{F}_i(G)| \leqslant \det(\tilde{\mathsf{L}}_{\setminus(i,i)}) &= \begin{vmatrix} N & -1 & \cdots & \cdots & -1 \\ -1 & N & -1 & \cdots & -1 \\ \vdots & -1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -1 \\ -1 & -1 & \cdots & -1 & N \end{vmatrix} \\
&= \begin{vmatrix} 2 & -1 & \cdots & \cdots & -1 \\ 2 & N & -1 & \cdots & -1 \\ \vdots & -1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -1 \\ 2 & -1 & \cdots & -1 & N \end{vmatrix} \\
&= \begin{vmatrix} 2 & 0 & \cdots & \cdots & 0 \\ 2 & N+1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 2 & 0 & \cdots & 0 & N+1 \end{vmatrix} \\
&= 2(N+1)^{N-2}.
\end{aligned}$$

In the third line, we added the columns from the second to the $(N-1)$th column to the first column. In the fourth line, we added half of the first column, $(1,\ldots,1)^\top \in \mathbb{R}^{N-1}$, to each of the columns from the second to the $(N-1)$th column.

[1] R. A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, 1930).

[2] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Harper & Row, New York, 1970).

[3] O. E. Sala *et al.*, Global biodiversity scenarios for the year 2100, Science **287**, 1770 (2000).

[4] M. R. Gardner and W. R. Ashby, Connectance of large dynamic (cybernetic) systems: Critical values for stability, Nature (London) **228**, 784 (1970).

[5] R. M. May, Will a large complex system be stable? Nature (London) **238**, 413 (1972).

[6] S. Allesina and S. Tang, Stability criteria for complex ecosystems, Nature (London) **483**, 205 (2012).

[7] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, Bacterial persistence as a phenotypic switch, Science **305**, 1622 (2004).

[8] B. B. Aldridge, M. Fernandez-Suarez, D. Heller, V. Ambravaneswaran, D. Irimia, M. Toner, and S. M. Fortune, Asymmetry and aging of mycobacterial cells lead to variable growth and antibiotic susceptibility, Science **335**, 100 (2012).

[9] Y. Wakamoto, N. Dhar, R. Chait, K. Schneider, F. Signorino-Gelo, S. Leibler, and J. D. McKinney, Dynamic persistence of antibiotic-stressed mycobacteria, Science **339**, 91 (2013).

[10] J. Maltas, A. Huynh, and K. B. Wood, Dynamic collateral sensitivity profiles highlight opportunities and challenges for optimizing antibiotic treatments, PLoS Biol. **23**, e3002970 (2025).

[11] P. V. Markov, M. Ghafari, M. Beer, K. Lythgoe, P. Simmonds, N. I. Stilianakis, and A. Katzourakis, The evolution of SARS-CoV-2, Nat. Rev. Microbiol. **21**, 361 (2023).

[12] C. Furusawa and K. Kaneko, Global relationships in fluctuation and response in adaptive evolution, J. R. Soc. Interface **12**, 20150482 (2015).

[13] R. van den Bosch and V. Stern, The integration of chemical and biological control of arthropod pests, Annu. Rev. Entomol. **7**, 367 (1962).

[14] G. P. Georghiou, The evolution of resistance to pesticides, Annu. Rev. Ecol. Syst. **3**, 133 (1972).

[15] E. Baake, M. Baake, and H. Wagner, Ising quantum chain is equivalent to a model of biological evolution, Phys. Rev. Lett. **78**, 559 (1997).

[16] J. Hermisson, O. Redner, H. Wagner, and E. Baake, Mutation–selection balance: Ancestry, load, and maximum principle, Theor. Popul Biol. **62**, 9 (2002).

[17] E. Baake and H.-O. Georgii, Mutation, selection, and ancestry in branching models: A variational approach, J. Math. Biol. **54**, 257 (2007).

[18] D. B. Saakian, A new method for the solution of models of biological evolution: Derivation of exact steady-state distributions, J. Stat. Phys. **128**, 781 (2007).

[19] D. B. Saakian, O. Rozanova, and A. Akmetzhanov, Dynamics of the eigen and the Crow-Kimura models for molecular evolution, Phys. Rev. E **78**, 041908 (2008).

[20] E. Muñoz, J.-M. Park, and M. W. Deem, Solution of the Crow-Kimura and eigen models for alphabets of arbitrary size by Schwinger spin coherent states, J. Stat. Phys. **135**, 429 (2009).

[21] A. S. Bratus, A. S. Novozhilov, and Y. S. Semenov, Linear algebra of the permutation invariant Crow–Kimura model of prebiotic evolution, Math. Biosci. **256**, 42 (2014).

[22] Y. S. Semenov and A. S. Novozhilov, Exact solutions for the selection–mutation equilibrium in the Crow–Kimura evolutionary model, Math. Biosci. **266**, 1 (2015).

[23] L. Euler, Recherches générales sur la mortalité et la multiplication du genre humain, Mémoires de l'académie des sciences de Berlin **16**, 144 (1767).

[24] A. J. Lotka, Relation between birth rates and death rates, Science **26**, 21 (1907).

[25] E. O. Powell, Growth rate and generation time of bacteria, with special reference to continuous culture, Microbiology **15**, 492 (1956).

[26] S. Pigolotti, Generalized Euler-Lotka equation for correlated cell divisions, Phys. Rev. E **103**, L060402 (2021).

[27] W. J. Ewens, An interpretation and proof of the fundamental theorem of natural selection, Theor. Popul. Biol. **36**, 167 (1989).

[28] S. A. Frank, The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis, Evolution **51**, 1712 (1997).

[29] M. Eigen, J. McCaskill, and P. Schuster, The molecular quasi-species, Adv. Chem. Phys. **75**, 149 (1989).

[30] Y. Sughiyama, T. J. Kobayashi, K. Tsumura, and K. Aihara, Pathwise thermodynamic structure in population dynamics, Phys. Rev. E **91**, 032120 (2015).

[31] H. Miyahara, Steady-state thermodynamics for population dynamics in fluctuating environments with side information, J. Stat. Mech. (2022) 013501.

[32] S. Leibler and E. Kussell, Individual histories and selection in heterogeneous populations, Proc. Natl. Acad. Sci. USA **107**, 13183 (2010).

[33] T. J. Kobayashi and Y. Sughiyama, Fluctuation relations of fitness and information in population dynamics, Phys. Rev. Lett. **115**, 238102 (2015).

[34] T. Nozoe, E. Kussell, and Y. Wakamoto, Inferring fitness landscapes and selection on phenotypic states from single-cell genealogical data, PLoS Genet. **13**, e1006653 (2017).

[35] A. Genthon and D. Lacoste, Universal constraints on selection strength in lineage trees, Phys. Rev. Res. **3**, 023187 (2021).

[36] R. García-García, A. Genthon, and D. Lacoste, Linking lineage and population observables in biological branching processes, Phys. Rev. E **99**, 042413 (2019).

[37] A. Genthon and D. Lacoste, Fluctuation relations and fitness landscapes of growing cell populations, Sci. Rep. **10**, 11889 (2020).

[38] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa, Interdependence of cell growth and gene expression: Origins and consequences, Science **330**, 1099 (2010).

[39] P. Wang, L. Robert, J. Pelletier, W. L. Dang, F. Taddei, A. Wright, and S. Jun, Robust growth of *Escherichia coli*, Curr. Biol. **20**, 1099 (2010).

[40] G. Lambert and E. Kussell, Quantifying selective pressures driving bacterial evolution using lineage analysis, Phys. Rev. X **5**, 011016 (2015).

[41] M. Hashimoto, T. Nozoe, H. Nakaoka, R. Okura, S. Akiyoshi, K. Kaneko, E. Kussell, and Y. Wakamoto, Noise-driven growth rate gain in clonal cellular populations, Proc. Natl. Acad. Sci. USA **113**, 3251 (2016).

[42] R. E. Lenski, Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations, ISME J. **11**, 2181 (2017).

[43] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992), Vol. 1.

[44] J. Moon, *Counting Labelled Trees*, Canadian Mathematical Monographs (Canadian Mathematical Congress, Montreal, 1970).

[45] S. Chaiken, A combinatorial proof of the all minors matrix tree theorem, SIAM J. Alg. Disc. Meth. **3**, 319 (1982).

[46] A. Cayley, Note sur une formule pour la reversion des séries, J. Reine Angew. Math. **52**, 276 (1856).

[47] J. J. Sylvester, On the change of systems of independent variables, Q. J. Pure Appl. Math. **1**, 42, 126 (1857).

[48] C. W. Borchardt, Ueber eine der Interpolation entsprechende Darstellung der Eliminations-Resultante, J. Reine Angew. Math. **57**, 111 (1860).

[49] G. Kirchhoff, Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird, Ann. Phys. **148**, 497 (1847).

[50] J. C. Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed. (Clarendon Press, Oxford, 1892).

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

[52] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, Rev. Mod. Phys. **48**, 571 (1976).

[53] J. A. Owen, T. R. Gingrich, and J. M. Horowitz, Universal thermodynamic bounds on nonequilibrium response with biochemical applications, Phys. Rev. X **10**, 011066 (2020).

[54] G. Fernandes Martins and J. M. Horowitz, Topologically constrained fluctuations and thermodynamics regulate nonequilibrium response, Phys. Rev. E **108**, 044113 (2023).

[55] S. U. Pillai, T. Suel, and S. Cha, The Perron–Frobenius theorem: Some of its applications, IEEE Signal Process. Mag. **22**, 62 (2005).

[56] C. Godsil and G. F. Royle, *Algebraic Graph Theory* (Springer Science & Business Media, New York, 2001), Vol. 207.

[57] D. B. West, *Introduction to Graph Theory* (Prentice Hall, Upper Saddle River, 2001), Vol. 2.

[58] M. Baym, L. K. Stone, and R. Kishony, Multidrug evolutionary strategies to reverse antibiotic resistance, Science **351**, aad3292 (2016).

[59] M. Tyers and G. D. Wright, Drug combinations: A strategy to extend the life of antibiotics in the 21st century, Nat. Rev. Microbiol. **17**, 141 (2019).

[60] J. Molina-Hernández, J. A. Cuesta, B. Pascual-Escudero, S. Ares, and P. Catalán, Optimization of sequential therapies to maximize extinction of resistant bacteria through collateral sensitivity, arXiv:2510.01808.

[61] R. B. Mokhtari, T. S. Homayouni, N. Baluch, E. Morgatskaya, S. Kumar, B. Das, and H. Yeger, Combination therapy in combating cancer, Oncotarget **8**, 38022 (2017).

[62] B. A. Quinn *et al.*, Pancreatic cancer combination therapy using a BH3 mimetic and a synthetic tetracycline, Cancer Res. **75**, 2305 (2015).

[63] P. D. Tamma, S. E. Cosgrove, and L. L. Maragakis, Combination therapy for treatment of infections with Gram-negative bacteria, Clin. Microbiol. Rev. **25**, 450 (2012).

[64] T.-C. Chou and P. Talalay, Generalized equations for the analysis of inhibitions of Michaelis-Menten and higher-order kinetic systems with two or more mutually exclusive and nonexclusive inhibitors, Eur. J. Biochem. **115**, 207 (1981).

[65] R. Peña-Miller, A. Fuentes-Hernandez, C. Reding, I. Gudelj, and R. Beardmore, Testing the optimality properties of a dual antibiotic treatment in a two-locus, two-allele model, J. R. Soc. Interface **11**, 20131035 (2014).

[66] H. L. David, Probability distribution of drug-resistant mutants in unselected populations of *Mycobacterium tuberculosis*, Appl. Microbiol. **20**, 810 (1970).

[67] I. L. Bergval, A. R. Schuitema, P. R. Klatser, and R. M. Anthony, Resistant mutants of *Mycobacterium tuberculosis* selected in vitro do not reflect the in vivo mechanism of isoniazid resistance, J. Antimicrob. Chemother. **64**, 515 (2009).

[68] C. Maes and K. Netočný, Heat bounds and the blowtorch theorem, Annales Henri Poincaré, **14**, 1193 (2013).

[69] M. Polettini and M. Esposito, Effective thermodynamics for a marginal observer, Phys. Rev. Lett. **119**, 240601 (2017).

[70] F. Khodabandehlou, C. Maes, and K. Netočný, Trees and forests for nonequilibrium purposes: An introduction to graphical representations, J. Stat. Phys. **189**, 41 (2022).

[71] S. Dal Cengio, V. Lecomte, and M. Polettini, Geometry of nonequilibrium reaction networks, Phys. Rev. X **13**, 021040 (2023).

[72] H.-M. Chun and J. M. Horowitz, Trade-offs between number fluctuations and response in nonequilibrium chemical reaction networks, J. Chem. Phys. **158**, 174115 (2023).

[73] S. Liang, P. De Los Rios, and D. M. Busiello, Thermodynamic bounds on symmetry breaking in linear and catalytic biochemical systems, Phys. Rev. Lett. **132**, 228402 (2024).

[74] P. E. Harunari, S. Dal Cengio, V. Lecomte, and M. Polettini, Mutual linearity of nonequilibrium network currents, Phys. Rev. Lett. **133**, 047401 (2024).

[75] C. Floyd, A. R. Dinner, and S. Vaikuntanathan, Learning to control non-equilibrium dynamics using local imperfect gradients, arXiv:2404.03798.

[76] C. Floyd, A. R. Dinner, A. Murugan, and S. Vaikuntanathan, Limits on the computational expressivity of non-equilibrium biophysical processes, Nat. Commun. **16**, 7184 (2025).

[77] P. J. Antsaklis and A. N. Michel, *Linear Systems* (Birkhäuser, Boston, 2006).

[78] F. Avanzini *et al.*, Methods and conversations in (post) modern thermodynamics, SciPost Phys. Lect. Notes 80 (2024).