Cell Lineage Statistics with Incomplete Population Trees

Arthur Genthon,[1,2,*] Takashi Nozoe,[3,4,5] Luca Peliti,[6] and David Lacoste[2]

[1]*Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*
[2]*Gulliver UMR CNRS 7083, ESPCI Paris, Université PSL, 75005 Paris, France*
[3]*Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo 153-8902, Japan*
[4]*Research Center for Complex Systems Biology, The University of Tokyo, Tokyo 153-8902, Japan*
[5]*Universal Biology Institute, The University of Tokyo, Tokyo 113-0033, Japan*
[6]*Santa Marinella Research Institute, 00052 Santa Marinella, Italy*

Cell lineage statistics is a powerful tool for inferring cellular parameters, such as division rate, death rate, or population growth rate. Yet, in practice such an analysis suffers from a basic problem: how should we treat incomplete lineages that do not survive until the end of the experiment? Here, we develop a model-independent theoretical framework to address this issue. We show how to quantify fitness landscape, survivor bias, and selection for arbitrary cell traits from cell lineage statistics in the presence of death, and we test this method using an experimental data set in which a cell population is exposed to a drug that kills a large fraction of the population. This analysis reveals that failing to properly account for dead lineages can lead to misleading fitness estimations. For simple trait dynamics, we prove and illustrate numerically that the fitness landscape and the survivor bias can in addition be used for the nonparametric estimation of the division and death rates, using only lineage histories. Our framework provides universal bounds on the population growth rate, and a fluctuation-response relation that quantifies the change in population growth rate due to the variability in death rate. Further, in the context of cell size control, we obtain generalizations of Powell's relation that link the distributions of generation times with the population growth rate, and we show that the survivor bias can sometimes conceal the adder property, namely the constant increment of volume between birth and division.

I. INTRODUCTION

Cells are the fundamental object of biology. Sources of cellular heterogeneity are typically blurred in bulk measurements and can only be detected with single-cell analysis. Thanks to advances in experimental techniques (microfluidics, image analyses, sequencing, etc.), single-cell measurements are now widely used for phenotype investigations [1,2]. However, single-cell measurements when performed *in vivo* often take the form of snapshots, which means that important dynamical information contained in lineage trees of cell populations (or equivalently population trees) is often lost or hidden [3].

In fact, even when temporal information about lineages is available, as with time-lapse video microscopy, some practical issues remain. One of them concerns lineages that terminate before the end of the experiment. Lineages can end for various reasons: cells can just stop dividing because of changes in the environment, they can die [4–6], or be flushed away as a result of dilution [7,8]. What should be done in practice with these dead lineages: should they simply be discarded or

should they be included in the analysis and how? Clearly, both options are not equivalent and a wrong methodology could lead to unwanted statistical biases. In this paper, we propose a theoretical framework to address this issue. Our approach is not limited to cell lineages, it could also be applied to other types of lineage trees containing extinct lineages, which appear, for instance, in evolution experiments.

A key idea in lineage analysis is the distinction between the backward sampling that reflects the differences in reproductive success between cells, and the forward sampling that balances these differences [9,10]. We first review this formalism in Sec. II. Then, we generalize this approach to include dead lineages in Sec. III. This framework can be used with any given lineage tree independently of the dynamics that generate the tree and of the cause of early lineage ending, and it is valid at arbitrary finite time.

Based on this framework, in Sec. IV we propose measures of fitness landscape and survivor bias, particularly adapted to cases where the cell traits of interest are correlated with the cause of death. We test these ideas using experimental data of mycobacteria subject to antibiotics [11], an analysis that reveals the importance of correctly accounting for dead lineages. These measures of fitness and survival are very general and can be evaluated for any lineage tree. For some simple models of population dynamics, we show that they can also be used to infer the division and death rates using phenotypic trajectories. These measures are further used to theoretically quantify selection in the presence of

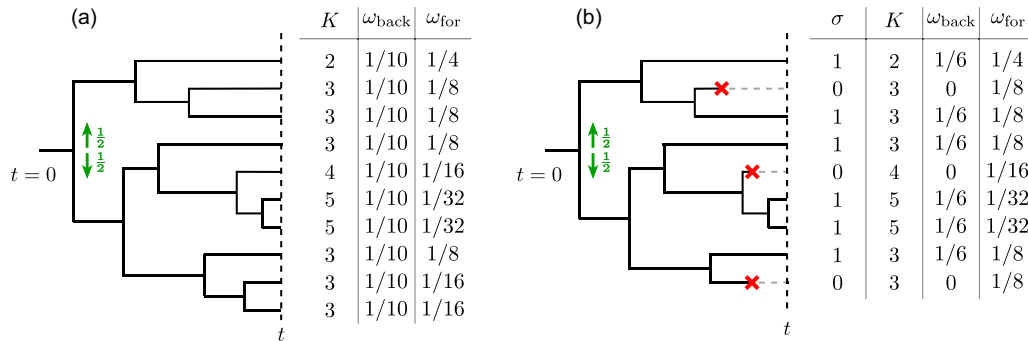*Corresponding author: arthur.genthon@hotmail.fr

FIG. 1. Population trees without and with death. (a) Population tree with no death starting with $N_0 = 1$ cell and ending with $N(t) = 10$ lineages at time t. In the table, the columns indicate the number K of divisions along the lineage and the backward and forward weights. (b) Population tree starting with $N_0 = 1$ cell and ending with $n(\sigma = 0, t) = 3$ dead lineages (red crosses), and $N(t) = n(\sigma = 1, t) = 6$ alive lineages at time t. In the table, the survival status σ takes value 1 for surviving lineages and 0 for dead lineages. The forward probability of survival is computed as $p_{\text{for}}(\sigma = 1, t) = 1 \times 1/4 + 3 \times 1/8 + 2 \times 1/32 = 11/16$.

death and fitness variability, and the death-induced change in selection.

In Sec. V, we explain the analogy between our framework and that of stochastic thermodynamics [12]. From this analysis, we deduce universal bounds on the population growth rate in the presence of death, which we test with experimental data, and we derive a fluctuation-response relation that relates the change in population growth rate to the variability in death rate.

Finally, we apply our formalism to models of cell size control in Sec. VI, where we derive modified versions of Powell's relation [13] in the presence of death. These relations link the backward and forward distributions of interdivision times and imply that cells divide faster under selection. A consequence of death on the statistics of the adder mechanism of cell size control [14] also follows from this approach.

II. THEORETICAL BACKGROUND FOR COMPLETE LINEAGE TREES

In this introductory section, we review the framework proposed in [9,10] to analyze lineage trees in the absence of cell death.

From this framework, it is possible to infer fitness differences associated with distinct states of cellular trait and selection within a growing population from empirical lineage tree data. For instance, time-lapse single-cell measurements provide cellular divisions and trait time series (e.g., fluorescence intensity of reporter protein) in the form of lineage trees. Because phenotypic traits temporally fluctuate, time series branch at division in such data sets, and it is therefore not trivial what statistics correctly report the fitness value of the phenotypic trait.

To resolve this issue, we assign two types of probability weight to each cellular lineage l: the backward weight $\omega_{\text{back}}(l) = N(t)^{-1}$ and the forward weight $\omega_{\text{for}}(l) = N_0^{-1} m^{-K(l)}$, where N_0 and $N(t)$ are the numbers of cells at initial time $t = 0$ and final time t, $K(l)$ is the number of divisions along lineage l, and m is the constant number of daughter cells born at division (including the mother cell). An example of a lineage tree with the corresponding forward and backward weights is represented in Fig. 1(a). The backward

sampling weighs the cells present at final time uniformly, thus leading to an over-representation of cells with above-average reproductive success. Instead, the forward weight represents the probability of choosing a cell lineage descending the tree from one of the ancestor cells at time $t = 0$ and selecting one branch among the m possibilities with equal probability $1/m$ at each division. Doing so, a lineage with high reproductive success is sampled with a lower weight in the forward sampling than in the backward sampling, in such a way as to balance the selection effect. Therefore, the comparison between these two samplings informs the selection undergone by lineages, as detailed below.

Let us now consider a general cell trait $\mathcal{S}$, taking values s, which can represent a phenotype or a genotype, for example. This trait can be any property of a lineage: a snapshot property evaluated at a given time, an averaged trait along lineages, or a trait trajectory, for instance. The discussion above motivates us to compute the forward (backward) probability to pick a lineage with K divisions and with trait value s in a forward (backward) manner. Let $n(K, s, t)$ denote the number of such lineages. We then define

$$p_{\text{back}}(K, s, t) = N(t)^{-1} n(K, s, t), \tag{1}$$

$$p_{\text{for}}(K, s, t) = N_0^{-1} m^{-K} n(K, s, t). \tag{2}$$

The fitness of the exponentially growing population is defined as its growth rate:

$$\Lambda_t = \frac{1}{t} \ln \frac{N(t)}{N_0}, \tag{3}$$

and using the forward and backward distributions, we obtain another representation of population fitness:

$$\Lambda_t = \frac{1}{t} \ln \langle m^K \rangle_{\text{for}}, \tag{4}$$

where $\langle \cdot \rangle_{\text{for}}$ denotes the average with respect to $p_{\text{for}}(K, s, t)$. In this representation, m^K and $p_{\text{for}}(K, s, t)$ are understood as the fitness value of a lineage and its probability distribution prior to selection.

Since the fitness landscape is conventionally defined as the mean fitness conditioned on a genotype or phenotypic trait, it

is natural to define the fitness landscape of cell trait $\mathcal{S}$ as

$$h_t(s) = \frac{1}{t} \ln\langle m^K|s\rangle_{\text{for}}, \tag{5}$$

where $\langle\cdot|s\rangle_{\text{for}}$ denotes the conditional average with respect to $p_{\text{for}}(K,t|s) = p_{\text{for}}(K,s,t)/p_{\text{for}}(s,t)$, with $p_{\text{for}}(s,t) = \sum_K p_{\text{for}}(K,s,t)$. Using Eqs. (1) and (2), together with $p_{\text{back}}(s,t) = \sum_K p_{\text{back}}(K,s,t)$, the fitness landscape of cell trait $\mathcal{S}$ is rewritten as

$$h_t(s) = \Lambda_t + \frac{1}{t} \ln \frac{p_{\text{back}}(s,t)}{p_{\text{for}}(s,t)}. \tag{6}$$

In this alternative form, the difference $h_t(s) - \Lambda_t$ between the fitness value of a cell with trait s and population fitness measures the over- or under-representation of lineages with this trait value s in the growing population as compared to the statistics prior to selection.

To quantitatively evaluate the overall effect of selection on the trait statistics, we define the strength of selection as the gain in mean fitness when "turning on" selection by changing the distribution of trait $\mathcal{S}$ from the forward ensemble, where differences in reproductive success are suppressed by the weighting of the lineages, to the backward ensemble:

$$\Pi_{\mathcal{S}} = \langle h_t\rangle_{\text{back}} - \langle h_t\rangle_{\text{for}}. \tag{7}$$

Because $\Pi_{\mathcal{S}}$ is positively proportional to the Jeffrey divergence [9], a non-negative symmetrical information-theoretic distance between the forward and backward distributions, it is always non-negative. If the fitness landscape is used to determine which values s of a given cell trait $\mathcal{S}$ are selected in a growing population, on the other hand, the strength of selection is a measure of the overall selection acting on trait $\mathcal{S}$, and it can thus be compared to other cell traits $\mathcal{S}'$ to determine which traits are under the strongest selection, i.e., which traits strongly correlate with lineage fitness.

Importantly, we want to emphasize that the notions of fitness and selection introduced here are different from their usual definitions based on division/growth rates. Usually, fitness quantifies the chance to reproduce, and thus measures the amplification over time of the proportion of some cell traits in the population. However, this measure does not account for the already-existing distribution of cell traits prior to selection. The fitness landscape proposed above instead measures the amplification of the forward (or *a priori*) probabilities when cells are competing in a population.

III. A FRAMEWORK TO SAMPLE INCOMPLETE LINEAGE TREES

The main experimental setups used to illustrate our framework are represented in Fig. 2: (a), free growth; and (b), growth in a microchannel [7]. The cell colors are an illustration of the bacterial diversity, and white cells are dead. In the microchannel, the down arrow indicates the medium flux that carries away cells that go out of the channel, so that both biological death and dilution can be present, while no dilution is observed in free growth. The case of free growth does not require cells to be far from each other; they can be strongly coupled or in contact, as happens, for instance, in tissues. Both cases (a) and (b) are represented by a population tree like the
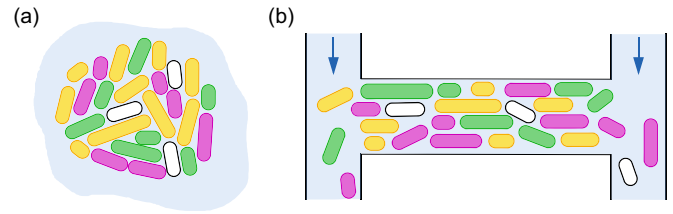

FIG. 2. Cartoons of two experimental setups where lineages can end before the end of the experiment. The cell colors represent cell heterogeneity and the white cells are dead. (a) Free growth in bulk; cells may be far from each other or in contact as in a tissue. (b) Cytometer setup from [7] where cells grow in a chamber open at both ends, and they are evacuated by the flow of growth medium in order to maintain the population constant inside the chamber.

one shown in Fig. 1(b), where the red crosses indicate death in a broad meaning. Such a tree starts with N_0 cells at time $t = 0$ and ends with $N(t)$ living cells at time t. Lineages can either survive up to time t or die before, as indicated by σ, taking the value 1 for alive lineages and 0 for dead lineages. Note that σ refers to the status (dead or alive) of lineages at the final time t, irrespective of the time of death.

When introducing death, the two samplings introduced in Sec. II are modified in the following way. When taking a snapshot of the population at time t, only living lineages are present, and in the backward sampling we sample them uniformly with weights $\omega_{\text{back}}(l) = N(t)^{-1}\delta(\sigma(l) - 1)$, where $\delta(x)$ is a Kronecker delta. On the other hand, starting from $t = 0$ with one of the N_0 initial cells and following the lineages up to time t by choosing with uniform probability $1/m$ one of the m daughter cell born at each division, both dead and living lineages are sampled with the forward weights $\omega_{\text{for}}(l) = N_0^{-1}m^{-K(l)}$, where $K(l)$ is the number of divisions along lineage l up to time t. We give a simple example of how these weights are computed in practice in Fig. 1(b). A major difference with the deathless case immediately appears: some lineages are sampled in the forward manner but not in the backward manner.

For a general cell trait $\mathcal{S}$, the forward and backward probabilities to pick a lineage with K divisions, with survival status σ and trait value s, are modified accordingly:

$$p_{\text{back}}(K,\sigma,s,t) = N(t)^{-1}n(K,\sigma,s,t)\delta(\sigma - 1), \tag{8}$$

$$p_{\text{for}}(K,\sigma,s,t) = N_0^{-1}m^{-K}n(K,\sigma,s,t), \tag{9}$$

where $n(K,\sigma,s,t)$ is the number of such lineages. The number of cells alive in the population at time t is now given by $N(t) = n(\sigma = 1, t) = \sum_{K,s} n(K,s,\sigma = 1, t)$.

The forward probability of survival $p_{\text{for}}(\sigma = 1, t) = \sum_{K,s} p_{\text{for}}(K, s, \sigma = 1, t)$ is a central quantity in this problem. It is a strictly decreasing function of the number of death events, and is unaffected by divisions, therefore it tends to 0 as $t \to \infty$. Indeed, when a lineage with K divisions, associated with a weight $N_0^{-1}m^{-K}$ in the forward sampling, divides in m daughters cells all weighted $N_0^{-1}m^{-(K+1)}$, then the overall weights of the m daughters is equal to that of the mother, and the forward weight is conserved. On the other hand, when a lineage dies, its forward weight disappears from the sum over living lineages that defines $p_{\text{for}}(\sigma = 1, t)$, leading to a

decrease of the latter. We then define the rate of decrease of the forward probability of survival as

$$\Gamma_t = \frac{1}{t} \ln p_{\text{for}}(\sigma = 1, t). \quad (10)$$

Finally, we introduce the forward distribution conditioned on survival:

$$p_{\text{for}}^{\star}(\cdot, t) = p_{\text{for}}(\cdot, t|\sigma = 1). \quad (11)$$

IV. QUANTIFYING FITNESS AND SELECTION IN INCOMPLETE LINEAGE TREES

A. Disentangling fitness and survival

In the absence of death, the measure of fitness defined by Eqs. (5) and (6), which was introduced in [9] and reviewed in Sec. II, depends only on the correlations between the value of the trait and the number of divisions along the lineage. When introducing death, the frequency of the trait in the population is the result of the correlations between this trait and the divisions, together with the correlations between this trait and survival. Therefore, the fitness landscape defined in Eqs. (5) and (6) captures both fitness and survival effects. Ignoring dead lineages and using this measure to quantify selection could thus lead to misleading fitness estimations, as we will demonstrate.

To disentangle the two effects, we define the fitness landscape in the presence of incomplete lineages as the average fitness conditioned on the cell trait for surviving lineages only:

$$h_t^{\star}(s) = \frac{1}{t} \ln \left[\sum_K m^K p_{\text{for}}^{\star}(K, t|s) \right]. \quad (12)$$

We show in Appendix A that this fitness landscape is consistent with notions of fitness proposed previously. In particular, it recovers the historical fitness introduced in [15] for populations undergoing stochastic phenotype switching, while being more general because it is model-independent. When $p_{\text{for}}^{\star}(K, t|s) = p_{\text{for}}^{\star}(K, t)$ for any s, then the fitness landscape is flat and equal to $\Lambda_t - \Gamma_t$. This condition is called the conditional independence of K and s knowing $\sigma = 1$. It is important to note that the conditional independence does not imply, and is not implied by, the regular independence between K and s. This means in particular that a trait s can be correlated to both K and σ, and still have a flat landscape.

As in Sec. II, by expressing $p_{\text{for}}^{\star}(K, t|s)$ in Eq. (12) and using Eqs. (8)–(11), the fitness landscape can be reshaped as

$$h_t^{\star}(s) = \Lambda_t - \Gamma_t + \frac{1}{t} \ln \left[\frac{p_{\text{back}}(s, t)}{p_{\text{for}}^{\star}(s, t)} \right], \quad (13)$$

where the marginalized trait distributions are defined as $p(s, t) = \sum_K p(s, K, t)$ (both backward and forward). By comparing the forward and backward samplings for surviving lineages only, we get rid of the incomplete lineages and thus isolate selection effects from survival effects.

When combining the previous definition Eq. (13) with Eqs. (8)–(11), the fitness landscape can also be written as

$$h_t^{\star}(s) = -\frac{1}{t} \ln \left[\sum_K m^{-K} p_{\text{back}}(K, t|s) \right]. \quad (14)$$

This formulation has a form similar to Eq. (12), but measures the inverse linage fitness m^{-K} averaged with respect to the probability distribution conditioned on the cell trait and posterior to selection.

We finally define the survivor bias as

$$h_t^{\dagger}(s) = h_t(s) - h_t^{\star}(s) = \Gamma_t + \frac{1}{t} \ln \left[\frac{p_{\text{for}}^{\star}(s, t)}{p_{\text{for}}(s, t)} \right], \quad (15)$$

which measures the statistical difference between the ensemble of all lineages and the surviving lineages only. While this definition of survivor bias is not the only possible one, it is reasonable to compare the two lineage ensembles using the forward sampling in order to get rid of selection effects and isolate survival effects. This bias is significant when considering a trait $\mathcal{S}$ which is correlated with survival. One important example is antibiotic resistance, because there, one is interested precisely in traits that appear among surviving bacteria in the presence of antibiotics [11]. We explore this example in the next section.

B. Making sense of antibiotics experiments

To illustrate the importance of the distinction between $h_t^{\star}$, h_t, and $h_t^{\dagger}$ when analyzing experimental data, we use lineage trees of *Mycobacterium smegmatis* exposed to the drug isoniazid (INH) [11]. After a few divisions in the absence of the drug, the population is exposed to the drug over 6 days, and every individual cell lineage is tracked until it is killed or reaches the end of observation [Fig. 3(a)]. In the original paper [11], the authors tested the hypothesis that bacterial persistence, i.e., the survival of a small fraction of cells when exposed to antibiotics which is not linked to a genetically acquired resistance, was linked to the existence of persisters that grow very slowly or do not grow at all before drug exposure, as observed for *E. coli* in [1]. To do so, they compared the preexposure elongation rates of persistent versus nonpersistent cells, and observed no significant difference. They thus concluded that the preexposure elongation rate is not correlated with survival for *M. smegmatis*.

We first aim to reproduce this observation with a formalism adapted to populations with dead lineages, and to further study the correlation between this cell trait and selection. In addition, we exhibit another cell trait that is correlated with selection and survival: the logarithmic cell size at the end of the tracking. Indeed, cell size is biologically insightful because the fitness landscape of cell size in the absence of death has been shown to reveal sources of stochasticity in volume partitioning and in single-cell growth rate [16], and to inform the mechanism of cell size control [17].

Let us first consider the preexposure elongation rate, which is defined by the slope of logarithmic cell size over 1 h before the INH addition. Among the 274 lineages alive or dead by $t = 72$ h (from drug exposure), five lineages have negative preexposure elongation rate, which are removed from the analysis. The three probability distributions $p_{\text{for}}^{\star}$, p_{for}, and p_{back} of the preexposure single-cell growth rate (for lineages extracted at $t = 72$ h) are relatively close to each other, as shown in Fig. 3(b), and equivalently the landscapes $h_t^{\star}$, $h_t^{\dagger}$, and h_t, shown in Fig. 3(d), are all close to flat landscapes. Note that only the dependence of the landscapes on the cell trait is
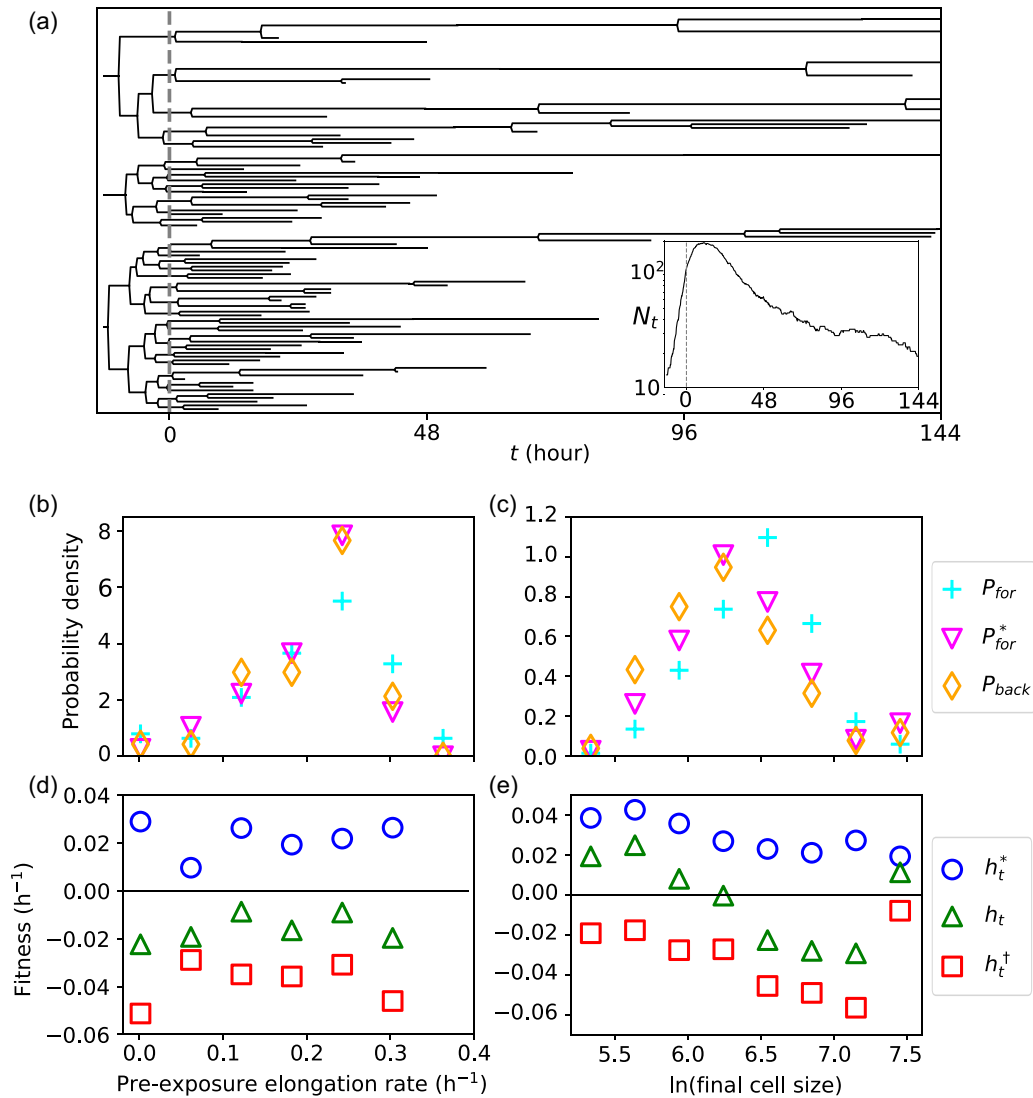
FIG. 3. Analysis of a lineage tree with biological death. (a) Example of lineage tree of individual bacterial cells proliferating and being killed by drug exposure [11]. The population is exposed to INH on a microfluidic device from $t = 0$ h, as indicated by the dashed line. Inset: evolution of the corresponding number N_t of alive cells with time. (b), (d) Distributions and landscapes of the preexposure growth rate (computed 1 h before drug exposure). There are no values at the rightmost bin for $h_t^\star$, h_t, and $h_t^\dagger$ because there are no alive lineages with this value of preexposure growth rate ($p_{\text{for}}^\star = p_{\text{back}} = 0$). (c),(e) Distributions and landscapes of the logarithmic cell size at the end of the lineage (computed at $t = 36$ h for surviving lineages or at the time of death for lineages that died before).

important, and not the offset between the three curves, which only comes from the different constants Λ_t and Γ_t used in the definitions (13) and (15). These observations are consistent with the analysis in [11] [Fig. 2(c), left, in this reference], namely the flatness of the survivor bias $h_t^\dagger$ confirms the independence of the preexposure single-cell elongation rate with the viability of the lineage when exposed to a drug. Further, our framework provides an additional result: the flatness of the fitness landscape $h_t^\star$ indicates the independence between the preexposure single-cell elongation rate and the fitness (reproductive success) of the lineage.

Let us now consider the second cell trait, namely the logarithmic cell size at the end of the lineage. We extracted lineages alive at time $t = 36$ h and those killed before this time, and measured the logarithmic cell size at the end of the

lineages, defined as the extraction time $t = 36$ h for surviving lineages and as the time of death for dead lineages. The corresponding probability distributions p_{for}, $p_{\text{for}}^\star$, and p_{back}, and landscapes $h_t^\star$, h_t, and $h_t^\dagger$ are shown in Figs. 3(c) and 3(e). This analysis reveals two results. All three landscapes are decreasing functions of the final log-size (except for the last point, which is not meaningful due to sampling issues), which indicates that cell size is negatively correlated with both fitness and survival. Moreover, the slope of the decrease of h_t is larger than that of the fitness landscape $h_t^\star$ (or equivalently the difference between p_{back} and p_{for} is more important than the one between p_{back} and $p_{\text{for}}^\star$). Therefore, by looking at h_t instead of $h_t^\star$ one could incorrectly infer a fitness advantage for smaller cells that is larger than in reality, since the decrease of h_t is mainly due to the survivor bias.

We claim that the decrease of the landscapes for the final cell size in Fig. 3(e) truly captures fitness and survival effects, while the variations of the same landscapes for the preexposure growth rate around their mean values in Fig. 3(d) are not meaningful but rather the result of uncorrelated fluctuations. To confirm this point, we plot in Fig. 7 in Appendix B the same landscapes after shuffling the trait values among lineages, thus canceling the correlations between cell trait and number of divisions and between cell trait and survival. The resulting landscapes are significantly flatter than the ones in Fig. 3(e) for final cell size, while they are compatible with the ones in Fig. 3(d) for the preexposure growth rate, which confirms our analysis.

C. Inferring the division and death rates

In the previous sections, we showed that the fitness landscape and the survivor bias provide a universal framework for the analysis of any incomplete lineage tree regardless of its dynamics. If we now introduce a cell trait dynamics, the fitness landscape and the survivor bias can in addition be used for the inference of the division and death rates, under certain conditions detailed below. This is particularly interesting since in general the estimation of the division rate from lineage data is a hard problem [18–20].

More specifically, let us follow a cell trait $\mathcal{S}$ that fluctuates in time along the lineages, with cells dividing and dying at rates $r(s)$ and $\gamma(s)$, respectively. These rates are defined as follows: a cell with trait value s divides (dies) with probability $r(s)dt$ ($\gamma(s)dt$) in a time interval dt. We show in Appendix A that when (i) the cell trait s is unaltered by divisions, and (ii) either (a) the division rate $r(s)$ is only weakly nonlinear or (b) the autocorrelation time of the trait dynamics is large compared to the observation time, then

$$h_t^\star(\bar{s}) = r(\bar{s}), \tag{16}$$

where we defined the time average along a lineage $\overline{f} = t^{-1} \int_0^t dt' f(t')$. This result can be used in the context of stochastic gene expression, for example, where the concentration of expressed proteins is typically continuous at division and fluctuates very little over multiple generations [21].

Moreover, when the autocorrelation time of the trait dynamics is large compared to the observation time, then

$$h_t^\dagger(\bar{s}) = -\gamma(\bar{s}). \tag{17}$$

These relations show that the division (death) rate can be extracted from the fitness landscape (survivor bias) for the time-averaged trait value $\bar{s}$ along lineages. The two important rates can thus be inferred from phenotypic trajectories only when the above conditions hold.

We put this inferring method to the test by simulating clonal cell proliferation and death processes. The cell trait $\mathcal{S}$ is chosen to follow an Ornstein-Uhlenbeck process, and cells divide and die according to division and death rates given by Hill functions. At division, the mother cell produces $m = 2$ daughter cells whose initial states s are the same as that of the mother upon division. A schematic representation of the simulation and the plots of the division and death rates are shown in Figs. 4(a) and 4(b). The details of the simulations can be found in Appendix A 3.
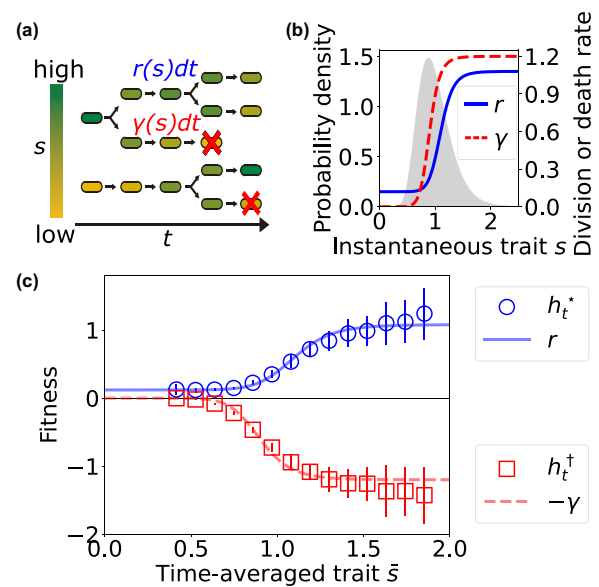

FIG. 4. Inference of division and death rates from fitness landscape and survivor bias of time-averaged trait in the numerical simulation. (a) A schematic of the population model used in the simulation. A cell with the trait s divides with the probability $r(s)dt$ and dies with the probability $\gamma(s)dt$ during dt. (b) Filled curve: Stationary forward probability distribution of the trait s when $\gamma(s) = 0$ (no bias). Blue solid and red dashed curves are the division rate $r(s)$ and death rate $\gamma(s)$, given in Appendix A 3. (c) Fitness landscape $h_t^*(\bar{s})$ and survivor bias $h_t^\dagger(\bar{s})$ of time-averaged trait $\bar{s}$. We choose the parameters so that the population size decreases like what happens in the experimental data we analyzed in Fig. 3, and we choose $t = 4$, which is shorter than the autocorrelation time of s and much shorter than the extinction timescale. The points and the error bars represent means and standard deviations over 100 runs, each starting with $N_0 = 1000$ cells.

The results of the inference are shown in Fig. 4(c), where the autocorrelation time of the trait dynamics was chosen to be larger than the duration of the experiment. As expected, the inference method is successful in this limit, despite the nonlinearity of the rates with respect to the value of the trait.

The effects of the nonlinearity of the division rate and of the autocorrelation time on the inference of the division rate have been investigated in [9] in the absence of death, and conclusions are similar with death. The impact of the autocorrelation time of the trait dynamics on the inference of the death rate is further investigated in Appendix A 3 and Fig. 6, where the same simulation is performed with a shorter autocorrelation time of the trait dynamics.

D. Effect of death on selection

1. Measure of death-induced change in the strength of selection

We reviewed in Sec. II that the degree of dissimilarity between the forward and backward distributions for a given cell trait $\mathcal{S}$ is a natural measure of selection acting on this cell trait in the absence of death. This definition was motivated by the fact that the forward sampling is built to suppress the differences in reproductive success among the lineages that are present in the backward sampling.

To generalize this definition in the presence of death while isolating the strength of selection from survival biases, we compute the strength of selection given in Eq. (7) for surviving lineages only, as we did for the fitness landscape in Sec. IV A:

$$\Pi_{\mathcal{S}} = \langle h_t^{\star} \rangle_{\text{back}} - \langle h_t^{\star} \rangle_{\text{for}}^{\star}. \tag{18}$$

Then, we propose a measure of the death-induced change in the strength of selection:

$$\Delta\Pi_{\mathcal{S}} = \Pi_{\mathcal{S}} - \Pi_{\mathcal{S}}^{\circ}, \tag{19}$$

where the superscript $\circ$ in $\Pi_{\mathcal{S}}^{\circ}$ indicates observables in the absence of death. The sign of $\Delta\Pi_{\mathcal{S}}$ indicates whether death increases or decreases the distance between the forward and backward distributions, i.e., the strength of selection.

Interestingly, the interplay between intrinsic selection and survival is explicit when focusing on trajectories $\mathcal{S}$ of traits subject to a division rate $\gamma(s)$. Indeed, we show in Appendix C that $\Delta\Pi_{\mathcal{S}}$ can be expressed as

$$\Delta\Pi_{\mathcal{S}} = \frac{\text{Cov}_{\text{back}}^{\circ}\left(h_t^{\circ}, p_{\text{surv}}\right)}{\langle p_{\text{surv}} \rangle_{\text{back}}^{\circ}} - \frac{\text{Cov}_{\text{for}}^{\circ}\left(h_t^{\circ}, p_{\text{surv}}\right)}{\langle p_{\text{surv}} \rangle_{\text{for}}^{\circ}}, \tag{20}$$

which depends on the correlations (covariances) between the fitness landscape in the absence of death $h_t^{\circ}(\boldsymbol{s})$, representing the intrinsic selection effect, and the probability of survival up to time t for a trait trajectory $\boldsymbol{s}$ along a lineage:

$$p_{\text{surv}}(\boldsymbol{s}) = \exp\left[-\int_0^t dt' \gamma(s(t'))\right]. \tag{21}$$

2. Illustrative example

We illustrate the possible outcomes of this change in selection strength by considering a simple birth and death process with two states a and b. Cells divide and die with state-dependent division and death rates r and γ. To avoid extinction, we suppose that $r_a \geqslant \gamma_a$ and $r_b \geqslant \gamma_b$, and we start with a large even number N_0 of initial cells, in which phenotypes are equally represented: $N_0(a) = N_0(b) = N_0/2$. Cells cannot switch to the other phenotype, and at division two cells of the same phenotype as the mother are produced. Without loss of generality we suppose that phenotype a divides faster than b: $\Delta r = r_a - r_b > 0$.

To treat this case explicitly, let us introduce the cell trait $\mathcal{S}$, taking the value $s = 1$ for cell state a and $s = 0$ for cell state b. The number of cells in the subpopulation a evolves as $n(a, t) = N_0 \exp[t(r_a - \gamma_a)]/2$ and similarly for the subpopulation b, so the backward probability reads

$$p_{\text{back}}(s, t) = \frac{e^{t(r_a - \gamma_a)}\delta(1 - s) + e^{t(r_b - \gamma_b)}\delta(s)}{e^{t(r_a - \gamma_a)} + e^{t(r_b - \gamma_b)}}. \tag{22}$$

Phenotypes are equally represented in the initial distribution, and the survival probability up to time t for phenotype a is given by $\exp(-t\gamma_a)$ (and similarly for b), therefore the forward distribution reads

$$p_{\text{for}}^{\star}(s, t) = \frac{e^{-t\gamma_a}\delta(1 - s) + e^{-t\gamma_b}\delta(s)}{e^{-t\gamma_a} + e^{-t\gamma_b}}. \tag{23}$$

The fitness landscape $h_t^{\star}(s)$ is obtained by computing the ratio of these two distributions [Eq. (13)]:

$$h_t^{\star}(s) = r_a\delta(1 - s) + r_b\delta(s). \tag{24}$$

Note that this is a simple case where the inference method proposed in Sec. IV C is applicable. Indeed, the fitness landscape exactly recovers the division rate, because the trait is not fluctuating in time.

Finally, the strength of selection is computed using Eq. (18):

$$\Pi_{\mathcal{S}} = \frac{r_a e^{t(r_a - \gamma_a)} + r_b e^{t(r_b - \gamma_b)}}{e^{t(r_a - \gamma_a)} + e^{t(r_b - \gamma_b)}} - \frac{r_a e^{-\gamma_a t} + r_b e^{-\gamma_b t}}{e^{-\gamma_a t} + e^{-\gamma_b t}}. \tag{25}$$

In the absence of death ($\gamma_a = \gamma_b = 0$), the asymptotic value of the strength of selection is controlled by the difference in reproductive rate: $\lim_{t \to \infty} \Pi_{\mathcal{S}}^{\circ} = \Delta r/2$. With death, the asymptotic selection strength results from two processes: the competition for the largest population, controlled by the net offspring production rate $r - \gamma$, which sets the backward trait distribution; and the competition for the smallest death rate γ, which controls the forward trait distribution. From Eq. (25), the possible outcomes for $\Delta\Pi_{\mathcal{S}}$ are

$$\Delta\Pi_{\mathcal{S}} \underset{t \to \infty}{\to} \begin{cases} \Delta r/2 & \text{if } \gamma_a > \gamma_b \text{ and } r_a - \gamma_a > r_b - \gamma_b, \\ -\Delta r/2 & \text{if } \gamma_a > \gamma_b \text{ and } r_a - \gamma_a < r_b - \gamma_b, \\ 0 & \text{if } \gamma_a > \gamma_b \text{ and } r_a - \gamma_a = r_b - \gamma_b, \\ 0 & \text{if } \gamma_a = \gamma_b, \\ -\Delta r/2 & \text{if } \gamma_a < \gamma_b. \end{cases} \tag{26}$$

Death has no effect on selection ($\Delta\Pi_{\mathcal{S}} = 0$) in two situations. First, when death is uniform, $\gamma_a = \gamma_b$, then each covariance in Eq. (20) is null. Second, when $r_a - \gamma_a = r_b - \gamma_b$ (which implies that $\gamma_a > \gamma_b$), meaning that death "shifts" the forward and backward distributions equally, their distance remains constant. In this last case, the covariances in Eq. (20) are equal. Note that the critical birth-death process where the total population is maintained constant, in contrast to the supercritical case where it grows exponentially, is obtained by setting $r_a = \gamma_a$ and $r_b = \gamma_b$. Thus, in the critical case we always have $r_a - \gamma_a = r_b - \gamma_b$, and $\Delta\Pi_{\mathcal{S}} = 0$.

There is only one case where selection is increased by death ($\Delta\Pi_{\mathcal{S}} > 0$): when cells that divide faster also die faster, while keeping a larger division-death balance (i.e., $\gamma_a > \gamma_b$ and $r_a - \gamma_a > r_b - \gamma_b$). In that way, phenotype a remains over-represented in the population like in the absence of death [$p_{\text{back}}(s = 1, t) \to 1$], but because it dies faster, its forward weight vanishes [$p_{\text{for}}^{\star}(s = 1, t) \to 0$, against $p_{\text{for}}^{\circ}(s = 1, t) = 1/2$ at all times in the absence of death]. As a consequence, the distance between the two samplings increases.

On the other hand, selection can be decreased by death ($\Delta\Pi_{\mathcal{S}} < 0$) in two cases. If the balance between division and death favors phenotype b while maintaining a smaller death rate for b (i.e., $r_a - \gamma_a < r_b - \gamma_b$ and $\gamma_a > \gamma_b$), then phenotype b dominates both samplings, which become identical [$p_{\text{back}}(s = 0, t) \to 1$ and $p_{\text{for}}^{\star}(s = 0, t) \to 1$]. On the contrary, when phenotype b has an unfavorable division-death balance and also dies faster (i.e., $\gamma_a < \gamma_b$), it remains under-represented in the population, and also becomes under-represented in the forward sampling, so that the forward and backward distributions become identical [$p_{\text{back}}(s = 0, t) \to 0$ and $p_{\text{for}}^{\star}(s = 0, t) \to 0$]. In these two cases, the strength of selection in the presence of death is null, and so the distance between the two samplings is decreased by death. Note that

this death-induced decrease in the strength of selection can be equivalently obtained by a reduction of fitness (here fitness is understood as division rate, since there is no phenotype switching [22]). Indeed, $\Delta\Pi_{\mathcal{S}} = -\Delta r/2$ can be achieved without death by instead lowering r_a to r_b.

One takeaway message from this analysis is that the strength of selection is more than a difference of growth rates: it instead measures how the intrinsic forward statistics are modified when cells are under selection in a population. In this setting, death rates matter too, and the strength of selection can be increased, decreased, or unaffected by death.

V. NONEQUILIBRIUM THERMODYNAMICS INTERPRETATION

Now that we have motivated our approach by showing the importance of distinguishing between fitness and survival in data analysis, and by proposing a practical use of these two notions for inference purposes, in this section we further analyze our framework from the perspective of stochastic thermodynamics. In the past decade, the thermodynamic structure of population dynamics, and of biological branching trees in general, has been thoroughly investigated [23–28]. These works provided insights into the principles that govern evolution and selection, with a degree of universality similar to that of thermodynamics for physics. In line with these works, we explore here the specific impact of death on these universal principles.

A. Fluctuation theorems

When comparing the probabilities to pick a living lineage with K divisions in the backward and forward samplings, Eqs. (8), (9), and (11), we obtain

$$p_{\text{back}}(K, t) = p_{\text{for}}^{\star}(K, t) e^{K \ln m - t(\Lambda_t - \Gamma_t)}. \tag{27}$$

Note that trait s has been integrated out here, since the exponential bias does not depend on it. This relation is a generalization of the fluctuation theorem obtained for a branching tree without death [9,27], but unlike fluctuation theorems in stochastic thermodynamics, which also compare two probability distributions [12], the forward and backward distributions are not related by time-reversal symmetry. Death modifies the relations in two ways: through the term Γ_t and in the lowering of the population growth rate Λ_t, which is no longer necessarily positive.

Integrating the forward probability in Eq. (27) gives a first integral fluctuation theorem:

$$\langle e^{t\Lambda_t - K \ln m} \rangle_{\text{back}} = 1 - p_{\text{for}}(\sigma = 0, t), \tag{28}$$

which is analogous to a generalization of Jarzynski's equality for absolutely irreversible processes [29], for which time-reversed processes are never observed. Similarly here, dead lineages have a positive weight in the forward sampling but a null one in the backward sampling. Thus, in this analogy, K, Λ_t, and $p_{\text{for}}(\sigma = 0, t)$ play the role of the work, the free energy, and the total statistical weight of absolutely irreversible transitions, respectively. A concrete consequence of this approach is presented in Sec. V C.

Integrating the backward probability in Eq. (27) leads to a second integral fluctuation theorem:

$$\Lambda_t = \frac{1}{t} \ln \langle m^K \rangle_{\text{for}}^{\star} + \Gamma_t, \tag{29}$$

which is a generalization of Eq. (4) in the presence of death. This relation links the population growth rate to the forward statistics of the number of divisions and to the forward probability of survival. This relation clearly shows that an effect of death can be equivalently described by a reduction of fitness: since $\Gamma_t < 0$, the population fitness Λ_t is reduced, which could also be obtained by a reduction of the number of divisions K along the lineages. This death-induced reduction of population fitness is analyzed using the framework of fluctuation-response relations in the next section.

B. Fluctuation-response relation for the population growth rate

When comparing experiments with and without death, it is natural to ask whether it is possible to predict the effect of death on population growth. Along these lines, Yamauchi *et al.* [10] recently showed that for the particular case in which cell death occurs upon division only, and for a small death probability, then at first order the decrease in population growth rate due to death depends on the Kullback-Leibler divergence between the backward and forward distributions of the number of divisions in the absence of death. The Kullback-Leibler (KL) divergence is a positive nonsymmetrical information-theoretic distance between probability distributions:

$$\mathcal{D}_{\text{KL}}(p||q) = \int dx \, p(x) \ln \frac{p(x)}{q(x)} \geqslant 0. \tag{30}$$

In Appendix D, we extend this result and show that for a general death rate γ, the decrease in population growth rate is given by

$$\Lambda_t - \Lambda_t^{\circ} = \frac{1}{t} \ln \langle p_{\text{surv}} \rangle_{\text{back}}^{\circ}, \tag{31}$$

where the survival probability $p_{\text{surv}}(s)$ for a trait trajectory s along a lineage is given by Eq. (21). In particular, if death occurs only at division with probability $1 - 2^{-\epsilon}$, then the survival probability along a lineage with K divisions is $p_{\text{surv}}(K) = 2^{-\epsilon K}$, and for small ϵ we recover the result from [10].

If the death rate is small, i.e., scaled by a small parameter ϵ, the survivor probability in Eq. (31) can be expanded such that

$$\Lambda_t - \left(\Lambda_t^{\circ} - \epsilon \langle \overline{\gamma} \rangle_{\text{back}}^{\circ} \right) = \frac{t}{2} \text{Var}_{\text{back}}^{\circ}(\overline{\gamma}) \epsilon^2 + O(\epsilon^3). \tag{32}$$

When there is no variability in the time-averaged death rate, the population growth rate is simply reduced by the uniform time-averaged death rate: $\Lambda_t = \Lambda_t^{\circ} - \epsilon \overline{\gamma}_0$, with $\overline{\gamma}_0 = \langle \overline{\gamma} \rangle_{\text{back}}^{\circ}$. If now we consider the case where there is variability while keeping the same average $\langle \overline{\gamma} \rangle_{\text{back}}^{\circ}$, then $\Lambda_t - (\Lambda_t^{\circ} - \epsilon \langle \overline{\gamma} \rangle_{\text{back}}^{\circ})$ is the difference between the growth rates with and without variability in death rate, for the same average death rate. This difference is proportional to the variance in death rate. Equation (32) can thus be interpreted as a fluctuation-response

relation, where the response is understood as the gain in population fitness which is caused by an increase of the fluctuations in death rate.

This result can be put in parallel with other fluctuation-response theorems in evolution, like Fisher's fundamental theorem of natural selection, which states in its simplest form that the time-derivative of the population fitness $d\Lambda_t/dt$ is equal to the variance of division rates $\text{Var}(r)$ [22]. Another example of relation of this kind is given in [15], where the time-integrated division rate along a lineage is called historical fitness, $H_t(s) = \int_0^t dt' \, r(s(t')) = t\bar{r}$. In this work, the authors derived the equality between the increase in mean historical fitness $\partial_\beta \langle H_t \rangle$, following the amplification of the differences in division rates $r(s) \to \beta r(s)$, and the variance of historical fitness $\text{Var}(H_t)$.

In contrast to these other fluctuation-response relations, our result shows the importance of fluctuations of death rate in understanding the change in population growth rate.

C. Second-law-like inequalities bounding the population growth rate

An important consequence of fluctuation theorems in stochastic thermodynamics is obtained by computing the KL divergence between the original and time-reversed path probabilities, which gives the second law of thermodynamics. Similarly here, second-law-like inequalities can be deduced from the positivity of the following Kullback-Leibler divergences:

$$\mathcal{D}_{\text{KL}}(p_{\text{back}}||p_{\text{for}}^{\star}) = \langle K \rangle_{\text{back}} \ln m - t(\Lambda_t - \Gamma_t) \geqslant 0, \quad (33)$$

$$\mathcal{D}_{\text{KL}}(p_{\text{for}}^{\star}||p_{\text{back}}) = -\langle K \rangle_{\text{for}}^{\star} \ln m + t(\Lambda_t - \Gamma_t) \geqslant 0. \quad (34)$$

Since the number K of divisions is positive-definite, Eq. (34) implies that $\Lambda_t - \Gamma_t$ is a positive quantity.

When combined, the two inequalities give

$$\frac{\langle K \rangle_{\text{for}}^{\star} \ln m}{t} \leqslant \Lambda_t - \Gamma_t \leqslant \frac{\langle K \rangle_{\text{back}} \ln m}{t}, \quad (35)$$

which provides universal lower and upper bounds for the population growth rate independent of the dynamics. Let us make two comments. First, Eq. (35) involves two bounds instead of one in thermodynamics, again because the forward and backward samplings are not related by a time-reversal symmetry. Second, an important result of Ref. [29] is that the lower bound on the average entropy production is improved due to the irreversible trajectories, going from 0 to a strictly positive quantity. Similarly here, since $\Gamma_t \leqslant 0$, the right part of Eq. (35) implies $t\Lambda_t \leqslant t(\Lambda_t - \Gamma_t) \leqslant \langle K \rangle_{\text{back}} \ln m$, giving a tighter bound on the population growth rate.

We illustrate these inequalities in Fig. 5(a) with experimental data from [7] of *Escherichia coli* cells in a microchannel. The values of $\langle K \rangle_{\text{for}}^{\star}$, $\langle K \rangle_{\text{back}}$, Λ_t, and Γ_t are computed for a single population, maintained approximately constant with 20–40 cells at any time. We see that the relative discrepancies between the curves are decreasing with time from $t = 1300$ min, meaning that the bounds are getting tighter. This is due to the fact that, because of dilution, all cells in the cytometer are likely to have a common ancestor which is close in the past, thus leading to a small variability in the numbers
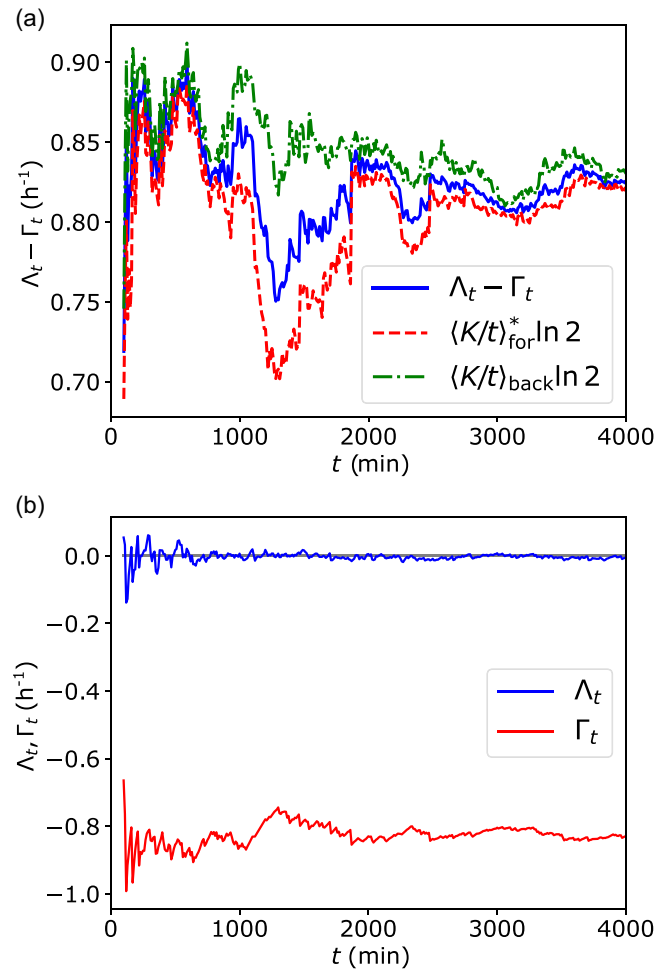

FIG. 5. Analysis of data from [7] for a population maintained constant thanks to dilution. (a) Illustration of Eq. (35), which shows that the bounds for $\Lambda_t - \Gamma_t$ get tighter as time increases. (b) Separate time evolutions of Λ_t and Γ_t, which show the expected convergence of Λ_t toward 0, due to the fact that the population is maintained constant, and a slower convergence of Γ_t toward a steady value, not fully achieved within the reach of the experiment.

of divisions among lineages [30]. We expect each curve to converge to a steady value in the long time limit. Figure 5(b) displays the separate evolution of Λ_t and Γ_t: as expected for constant populations, Λ_t tends to 0 rather quickly, while the convergence of Γ_t is slower.

When studying regulated populations of cells, Eq. (35) further implies inequalities between the forward and backward mean generation times with the population doubling time, as explored in the next section.

VI. EFFECT OF DEATH ON MODELS OF CELL SIZE CONTROL

In recent years, the comparison between the forward and backward samplings of lineages has been used to quantify how the behavior of cells is modified when they are under selection. For example, it has been shown that cells elongate faster [31], are born bigger [17], and end up being smaller in a snapshot [16] when selection is present among lineages.

Importantly, they also divide faster, because cells that are over-represented in the backward sampling because of their higher reproductive success have shorter interdivision times (also called generation times) [27,32,33]. Therefore, selection biases the distribution of cell lifetimes towards small values.

In the next sections, we introduce a noninteracting age-and-size-structured model of cell populations, and we show how this forward-backward bias on generation times is impacted by cell removal due to death or dilution.

A. General model of cell size control with death

We consider a general model of cell size control, where cells are characterized by their age a and size at birth x_b, and divide and die with division and death rates $r(a, x_b)$ and $\gamma(a, x_b)$, respectively. For simplicity, in the following we call γ the death rate but keep in mind that it models any kind of cell removal. Upon division, mother-daughter correlations are described by the transition kernel $\Sigma(x_b|x_b', a')$, which is the probability for the daughter cell to be born with size x_b knowing its mother was born with size x_b' and divided at age a'. Between divisions, cells grow at a rate $\nu(x) = dx/da$, where x is the cell size. This growth rate accounts for the most common growth laws, for example exponential growth $x(a) = x_b \exp(\nu_0 a)$ when $\nu(x) = \nu_0 x$, and linear growth $x(a) = x_b + \nu_1 a$ when $\nu(x) = \nu_1$. Detailed equations for this model can be found in Appendix E 1.

In the absence of death, this model accounts for the most common mechanisms of cell size control: the timer, the sizer, and the adder, where division is triggered by age, size, and increment of volume, respectively [34,35]. Indeed, cell size is determined by age, size at birth, and single-cell growth rate: $x(a, x_b, \nu)$, and the increment of volume since birth is defined as $\Delta = x - x_b$. Therefore, the timer, sizer, and adder models are recovered when setting $r(a, x_b) \equiv \check{r}(a)$, $r(a, x_b) \equiv \hat{r}(x)$, and $r(a, x_b) \equiv \nu(x)\zeta(\Delta)$, respectively, where $\zeta(\Delta)$ is the division rate per unit volume [14]. The complete mappings with the sizer and adder are presented in Appendix E 2. The model is then extended to account for death by adding a general death rate $\gamma(a, x_b)$.

The division and death rates can be linked to the population growth rate Λ_t and to the rate of decrease of the forward probability of survival Γ_t defined in Sec. III. Indeed, we show in Appendix E 3 that when averaging over many population trees, which is equivalent to taking the limit $N_0 \to \infty$ since each initial cell produces an independent tree, the stochastic quantities Λ_t and Γ_t converge to their expected values:

$$\lim_{N_0 \to \infty} \Lambda_t = \overline{\langle (m-1)r - \gamma \rangle_{\text{back}}}, \tag{36}$$

$$\lim_{N_0 \to \infty} \Gamma_t = -\overline{\langle \gamma \rangle_{\text{for}}^\star}, \tag{37}$$

where we used the time-average $\overline{f} = t^{-1} \int_0^t dt' f(t')$.

Finally, this model accounts for the different experimental setups. In microchannels, a uniform dilution rate balancing cell divisions on average is typically assumed [7,13,36]: $\gamma(a, x_b) \equiv \gamma = (m-1)\langle r \rangle_{\text{back}}$, so that the population is maintained constant ($\Lambda_t = 0$). In morbidostats, the cell population is instead maintained constant thanks to a control of the death rate [37].

Note that this model is a particular case of the multitype-age model introduced in [33], where the type here is the newborn size. Therefore, beyond cell size control, the results derived in the next section are equally valid with a general cell type, which could represent a genotypic or phenotypic trait instead of the newborn size.

B. Forward-backward bias for generation times in the presence of death

In the long-time limit, if the population does not go extinct and if it reaches a steady state, the population grows exponentially with a rate $\Lambda = \lim_{t \to \infty} \Lambda_t \geqslant 0$, called the Malthus parameter. In this limit and for cells undergoing binary fission ($m = 2$), the number of cells doubles after a time $T_d = \ln 2/\Lambda$, called the population doubling time, and the forward survival probability is reduced by half after a time $T_s = -\ln 2/\Gamma$, with $\Gamma = \lim_{t \to \infty} \Gamma_t$. Moreover, when a large-deviation principle for the rescaled number of divisions K/t is observed [38,39], then $\lim_{t \to \infty} \langle K \rangle_{\text{back}}/t = \langle \tau \rangle_{\text{back}}^{-1}$, where τ is the generation time, and similarly for the forward average. Such a principle has indeed been derived for the model of cell size control introduced above [40]. Then, in the long time limit, Eq. (35) turns into

$$\langle \tau \rangle_{\text{back}} \leqslant \left(T_d^{-1} + T_s^{-1} \right)^{-1} \leqslant \langle \tau \rangle_{\text{for}}^\star, \tag{38}$$

which generalizes the inequalities between population doubling time and mean generation times known for age-controlled populations in the absence of death [7,13,27]. Note that the inequality $\langle \tau \rangle_{\text{back}} \leqslant \langle \tau \rangle_{\text{for}}^\star$ is not modified by death (except for the conditioning on survival in the forward sampling), which was expected since this inequality results from differences in the selection only. On the other hand, the middle term explicitly depends on the forward survival "halving time", and $T_{\text{eff}} = 1/(T_d^{-1} + T_s^{-1})$ represents an effective population doubling time. When there is no death, $T_{\text{eff}} = T_d$ is the true population doubling time, and when there is no growth, $T_{\text{eff}} = T_s$ is also linked to the population doubling time in microchannel experiments. Indeed, if the population is maintained constant thanks to a dilution rate $\gamma = -\Gamma$ that exactly balances divisions, then $-\Gamma$ is also the rate at which the population would grow in the absence of dilution.

To go beyond average generation times, we next propose a generalization of the bias between generation time distributions, namely Powell's relation. Powell's relation and the Euler-Lotka equation are two important results from the literature on age-controlled populations in steady-state. In a series of papers, Powell derived the statistical bias between the distributions of generation time when measured in population versus single lineages, first for uncorrelated divisions and then in the presence of Markovian correlations [13,41,42]. The Euler-Lotka equation expresses the integral relationship between the population growth rate and the single lineage distribution of generation times [43]. These results have since been generalized in several directions: for models with an age-dependent death rate [44], and more recently, for models with non-Markovian correlations [39], for constant populations with a uniform dilution rate [36], and for unbalanced (non-steady-state) growth [45].

We define the steady-state distributions of generation times conditioned on size at birth as

$$f(\tau|x_b) = \frac{r(\tau, x_b)p(\tau, x_b)}{\int d\tau' \ r(\tau', x_b)p(\tau', x_b)}, \tag{39}$$

where $p(a, x_b)$ is the steady-state joint distribution of age and size at birth, evaluated either in the backward sampling or forward sampling conditioned on survival. This distribution measures the proportion of cells dividing with age τ and size at birth x_b among all cells dividing with size at birth x_b, in either of the samplings. In Appendix F, we derive analytical expressions for $f_{\text{back}}(\tau|x_b)$ and $f_{\text{for}}^\star(\tau|x_b)$, and their relationship:

$$f_{\text{back}}(\tau|x_b) = \frac{m f_{\text{for}}^\star(\tau|x_b)e^{-(\Lambda-\Gamma)\tau}}{Z(x_b)}, \tag{40}$$

where $Z(x_b)$ is a normalizing factor. Since we showed with Eq. (34) that $\Lambda - \Gamma \geqslant 0$, then for any newborn size x_b, the backward distribution is biased towards small division times, where again $\Lambda - \Gamma$ can be interpreted as an effective population growth rate.

A similar but different relation has been derived previously, which compares instead the backward dynamics with death and the forward dynamics without death [33]. We propose in Appendix F 3 a complete comparison between these two relations. We believe that Eq. (40) has the following advantages over that other formulation. First, it compares two distributions evaluated in the same experimental setup, with the same dynamics in the presence of death, which makes sense in light of the definition of fitness landscape proposed in Sec. IV. Second, we obtained an explicit expression for the constant $Z(x_b)$ in terms of the distributions of newborn sizes x_b at birth and at division (see Appendix F).

Finally, in the absence of mother-daughter correlations, i.e., when the correlation kernel is diagonal, $\Sigma(x_b|x_b', a') \equiv \hat{\Sigma}(x_b)$, a simpler version of the relation is obtained: $f_{\text{back}}(\tau) = m f_{\text{for}}^\star(\tau)e^{-(\Lambda-\Gamma)\tau}$, where the newborn size x_b does not appear explicitly anymore. This relation further simplifies for uncorrelated age models without death, in which case the regular Powell's relation is recovered, namely $f_{\text{back}}(\tau) = m f_{\text{for}}(\tau)e^{-\tau\Lambda}$ [7,13].

C. Death can conceal the adder property

The adder is a mechanism of cell size control that postulates that cells divide after growing by a certain amount of volume which is independent of their volume at birth. Experimental evidence of this mechanism for bacteria has been provided in two independent simultaneous articles [14,46]. This property has since been observed for many other organisms [34,47], among which are yeasts [48] and archaea [49]. It is therefore increasingly seen as the most relevant model of cell size.

Within the framework of [14], in the absence of death, the adder mechanism is characterized by the independence of the lineage distribution of added volume Δ_d between birth and division from the newborn volume: $f_{\text{lin}}(\Delta_d|x_b) \equiv f_{\text{lin}}(\Delta_d)$. The lineage distribution is defined along independent single lineages, in a mother machine setup [50]. Importantly, in the absence of death and for large samples, the lineage distribution is equivalent to the forward distribution, which is

also unbiased by the differences in reproductive success by construction [27,38]. Therefore, the adder property also reads $f_{\text{for}}^\circ(\Delta_d|x_b) \equiv f_{\text{for}}^\circ(\Delta_d)$. Recently, evolutionary explanations for the emergence of the adder property have gained interest: the idea is that the adder model ensures a larger population growth rate compared to the sizer in the presence of cell death [51].

Because of its ubiquity in cell biology, and its potential importance in evolution, it appears fundamental to investigate how death might impact the adder property. Similarly, population effects can impact the adder property, which is why in the Supplemental Material of [14], the authors observed the adder property with single lineages in mother machines, to "avoid [...] known bias effects related to the speed of reproduction [13]." In the presence of death, a new bias is introduced which can also blur the adder property for surviving lineages: the survivor bias.

We show in Appendix G that the forward distribution of added volume conditioned on newborn size for surviving lineages is biased with respect to the same distribution in the absence of death as

$$f_{\text{for}}^\star(\Delta_d|x_b) = \frac{f_{\text{for}}^\circ(\Delta_d)}{Y(x_b)} e^{-\Gamma\tau(x_b, \Delta_d) - \int_0^{\Delta_d} d\Delta \frac{\tilde{\gamma}(x_b+\Delta, \Delta)}{v(x_b+\Delta)}}, \tag{41}$$

where $Y(x_b)$ is a normalizing constant, $\tilde{\gamma}(x, \Delta) = \gamma(a(x, \Delta), x_b(x, \Delta))$ is the death rate expressed using the variables volume and increment of volume, and $\tau(x_b, \Delta_d)$ is the generation time for a cell born with volume x_b and which divides with volume $x_b + \Delta_d$, which depends on the growth law $v(x)$. Exponentially growing cells $[v(x) = v_0 x]$ divide after a time $\tau(x_b, \Delta_d) = \ln[(x_b + \Delta_d)/x_b]/v_0$, for example.

Equation (41) can be understood intuitively as follows. The exponential term involving the death rate $\tilde{\gamma}$ is the survival probability for the specific cycle of a cell born with volume x_b and dividing after adding a volume Δ_d. On the other hand, the term $\exp(\Gamma\tau)$ is the forward survival probability introduced in Eq. (10), over the duration $\tau(x_b, \Delta_d)$ of that same cycle, which is an average measure of survival [Eq. (37)]. The bias in Eq. (41) involves the ratio of these two probabilities, and thus compares the survival probability of a cell to its ensemble average. If survival is correlated with the volume at birth and the added volume, via the death rate $\tilde{\gamma}(x_b + \Delta, \Delta)$, then this bias results in a biased conditional statistics of added volume when considering surviving lineages.

Since both terms in the exponential explicitly depend on the size at birth, this demonstrates that the distribution $f_{\text{for}}^\star(\Delta_d|x_b)$ is not independent of the newborn volume in general. Therefore, the adder property would be concealed by analyzing surviving lineages only, even for cells obeying the adder principle in the absence of death. Three simple examples where the independence remains true for surviving lineages are discussed in Appendix G.

Note that in practice, the independence of added volume from newborn volume has often been tested at the level of the first moment only: $\langle\Delta|x_b\rangle_{\text{for}}^\circ = \langle\Delta\rangle_{\text{for}}^\circ$, by checking that the slope of $\langle\Delta|x_b\rangle_{\text{for}}^\circ$ versus x_b is 0 (or equivalently that the slope of $\langle x_d|x_b\rangle_{\text{for}}^\circ$ versus x_b is 1, with x_d the size at division) [46,48,49,52]. Our result then suggests that $\langle\Delta|x_b\rangle_{\text{for}}^\star$ is not independent of x_b in general.

We expect to observe this effect in the new generation of experimental setups which combine population and single

lineage experiments in the same controlled environment, allowing, for example, accurate measurements of the effect of antibiotics at both the population and lineage levels [2].

VII. DISCUSSION

In this paper, we built a general framework to properly account for dead lineages in a statistical analysis of lineage trees. Dead lineages are broadly understood as lineages ending before the end of the experiment, regardless of the cause: biological death, cell removal, dilution, etc. Our framework is based on the one proposed in [9] in the absence of death, which relies on the forward and backward samplings of lineages. In the procedure we proposed, dead lineages are sampled only in the forward manner, given that they do not appear in the population at final time. The statistical description then necessarily involves a new quantity Γ_t to account for the forward weight of surviving lineages, which modifies the relations between forward and backward distributions.

We showed how to quantify fitness and selection in the presence of death, whatever the cause of death might be. We proposed a general measure of the effect of death on selection whose sign indicates if death tends to increase or reduce the difference between the forward and backward statistics. This formalism is well adapted to situations in which a trait of interest is strongly correlated with death. For instance, the apparition of a resistant phenotype is correlated with the application of antibiotics which kills a large fraction of a bacterial population [53]. Similarly, in starved colonies of bacteria, the probability of cell death depends on a maintenance cost of metabolism [54,55], itself related to the growth rate before exposure to antibiotics [4]. Beyond bacteria, in tissues, there is a fine balance of growth and death of cells. When cell death is induced by overcrowding effects, many traits of interest such as cell size are typically correlated with the death rate [5,6]. It is interesting to note that in several places, death is effectively equivalent to a reduction of fitness, as in the reduction of the strength of selection of Eq. (26), in the integral relations of Eqs. (28) and (29), and in the fluctuation-response relation of Eq. (32).

When focusing on models of cell size control, new inequalities between mean generation times, population doubling time, and survival "halving time", and new generalizations of Powell's relation, followed from our approach. These generalizations are useful, in particular in experiments in which sampling individuals from a population can be affected by the survivor bias. In fact, even a standard relation used to characterize the adder model may be hidden if a survivor bias is present.

The present framework generalizes beyond populations of cells, to similar branched trees in other areas of biophysics. In evolutionary biology, lineage trees can be, for instance, phylogenetic trees, which contain dead lineages due to species extinction [56]. Naturally, our framework does not solve a central problem in this field: often we do not know the precise chronology of events from which the forward statistics could be built. As a result, only the backward sampling can be performed, and selection mechanisms and past history of the population need to be inferred together from snapshot data only. Moreover, relations involving a comparison between the two samplings are valid only for surviving lineages, and thus do not take advantage of the large amount of data from dead lineages. These data can be exploited with the tree sampling [17,36,57], and linking the tree and forward/backward samplings in the presence of a general death rate would be a logical next step of this work.

APPENDIX A: FITNESS LANDSCAPE AND SURVIVOR BIAS RECOVER DIVISION AND DEATH RATES

In [15], the authors introduced the notion of historical fitness, which is equal to the time-integrated division rate along a lineage:

$$H_t(\mathbf{s}) = \int_0^t dt'\, r(s(t')), \tag{A1}$$

where s is a phenotypic trait, possibly of high dimension, and $\mathbf{s} = \{s(t')\}_{t' \in [0,t]}$ is a phenotypic trajectory. They considered a simple model of population dynamics with phenotypic fluctuations, where the evolution of the expected number $n(s, t)$ of cells with trait value s at time t follows the equation

$$\partial_t n(s, t) = r(s)n(s, t) + \int ds'\, T(s, s')n(s', t), \tag{A2}$$

where $T(s, s')$ is the rate of phenotype switching from s' to s. Importantly here, the divisions and phenotype switchings are independent events, and newborn cells have the same phenotype as their mother. We call this dynamics the switching-division model. With this model, the authors argued that the measure of selection proposed by Fisher in [22], namely the population variance of the division rate (called fitness or individual fitness in this context), was no longer a good measure of selection. Instead, they shifted the perspective from individual cells to individual trajectories, and showed that the variance of historical fitness was more adequate to measure selection, in the sense of gauging the importance of selective differences for the evolution of the population.

Because the historical fitness is model-dependent and because it relies on the division rate, which is not often measurable, in [9] the authors proposed a new notion of fitness: the fitness landscape $h_t(\mathbf{s})$. The latter is independent of the dynamics of the population, and only relies on phenotypic trajectories: it can thus be evaluated in any branching tree. To show the consistency between the fitness landscape and previous notions of fitness, they proved that, when considering a population governed by the simple model Eq. (A2), the fitness landscape recovers the historical fitness: $h_t(\mathbf{s}) = H_t(\mathbf{s})/t$. When some conditions are met, as detailed in the next section, the historical fitness can be well approximated by the division rate for the time-averaged phenotypic trait: $H_t(\mathbf{s})/t \approx r(\bar{s})$. As a consequence, the fitness landscape can be used to infer the

division rate by simply sampling phenotypic trajectories. This idea has been proven successful numerically in [9].

In this Appendix, we show that the connection between fitness landscape, historical fitness, and division rate holds when adding death to the switching-division model. In addition, we show that the survivor bias defined in the main text can be used in a similar way to infer the death rate.

1. Fitness landscape, historical fitness, and division rate

We consider the switching-division model described by Eq. (A2) in which we add a phenotype-dependent death rate $\gamma(s)$ and show that the link between historical fitness and fitness landscape holds for surviving lineages.

We start from the definition of the fitness landscape under the form of Eq. (12), and for simplicity we consider that the trait $\mathcal{S}$ takes only discrete values $s \in \{s_i\}_{i=1}^n$. Since the division rate only depends on the current phenotype of the cell, and since divisions do not alter the phenotype, the number of divisions on different portions of the trajectory are independent:

$$p_{\text{for}}(K_1, \ldots, K_n, t|\boldsymbol{s}, \sigma = 1) = \prod_{i=1}^n p_{\text{for}}(K_i, t|\boldsymbol{s}, \sigma = 1),$$
(A3)

where K_i is the number of divisions that occurred during the time t_i spent in state s_i (even if this duration is discontinuous), called occupation time. The occupation times sum up as $\sum_{i=1}^n t_i = t$. In each state, the division rate is constant, so that each term in the product is a Poisson distribution:

$$p_{\text{for}}(K_i, t|\boldsymbol{s}, \sigma = 1) = \frac{(r(s_i)t_i)^{K_i}}{K_i!} e^{-r(s_i)t_i}.$$
(A4)

Combining these results, we obtain

$$p_{\text{for}}(K, t|\boldsymbol{s}, \sigma = 1) = \sum_{K_1 + \cdots + K_n = K} p_{\text{for}}(K_1, \ldots, K_n, t|\boldsymbol{s}, \sigma = 1)$$
(A5)

$$= e^{-\sum_{i=1}^n r(s_i)t_i} \sum_{K_1 + \cdots + K_n = K} \prod_{i=1}^n \frac{(r(s_i)t_i)^{K_i}}{K_i!}$$
(A6)

$$= e^{-\sum_{i=1}^n r(s_i)t_i} \frac{\left(\sum_{i=1}^n r(s_i)t_i\right)^K}{K!},$$
(A7)

where we used the multinomial development to obtain the last line.

Finally, plugging Eq. (A7) into Eq. (12) with $m = 2$ leads to

$$h_t^\star(\boldsymbol{s}) = \frac{1}{t} \sum_{i=1}^n r(s_i)t_i$$
(A8)

$$= \frac{1}{t} H_t(\boldsymbol{s}).$$
(A9)

Now, let us detail under which conditions the fitness landscape can be used to infer the division rate. If the division rate is linear, then $\int_0^t dt'\, r(s(t'))/t = r(\bar{s})$ exactly, where $\bar{s} = t^{-1} \int_0^t dt'\, s(t')$, independently of the trait dynamics. Therefore, we have

$$h_t^\star(\bar{s}) = r(\bar{s}).$$
(A10)

The result remains true if the division rate is only weakly nonlinear, even with possibly fast fluctuations of the trait, or if the division rate is significantly nonlinear, but with slow fluctuations of the trait (i.e., when the autocorrelation time of the trait dynamics is smaller than the observation time), as observed in the main text.

Note that this inference method works only when tracking a cell trait that is unaffected by divisions, like with the switching-division model Eq. (A2). With models of cell size control introduced in Appendix E, for example, the cell trait (size, age, etc.) is reset at division, and thus the number of divisions is encoded in the trait trajectory $\boldsymbol{s}$, such that $p_{\text{for}}(K, t|\boldsymbol{s}, \sigma = 1) = \delta(K - K[\boldsymbol{s}])$. Therefore, $h_t^\star(\boldsymbol{s}) = K[\boldsymbol{s}] \ln 2/t$, which does not inform the division rate.

2. Survivor bias and death rate

In this section, we show that, under certain assumptions, the survivor bias given in Eq. (15) can be used to infer the phenotype-dependent death rate $\gamma(s)$.

First let us derive the explicit bias between $p_{\text{for}}^\star(\bar{s})$ and $p_{\text{for}}^\circ(\bar{s})$, where the superscript $^\circ$ indicates quantities from the experiment without death. When comparing two experiments, identical in all points except for the presence of death in one of them only, through a death rate $\gamma(s)$ depending on cell trait s, the expected numbers of lineages following the path $\boldsymbol{s}$ are linked by

$$n(\boldsymbol{s}) = n^\circ(\boldsymbol{s}) \exp\left[-\int_0^t dt'\, \gamma(s(t'))\right].$$
(A11)

Upon division of Eq. (A11) by $n_0^{-1} m^{-K[\boldsymbol{s}]}$ (for simplicity, we consider that $n_0 = n_0^\circ$ here), we obtain

$$p_{\text{for}}(\boldsymbol{s}, \sigma = 1) = p_{\text{for}}^\circ(\boldsymbol{s}) \exp\left[-\int_0^t dt'\, \gamma(s(t'))\right].$$
(A12)

Note that here, no assumption is needed as to whether the cell trait is affected or not by divisions. For simplicity, we considered the case where the trajectory $\boldsymbol{s}$ encodes the number of divisions $K[\boldsymbol{s}]$. If this is not the case, the same result can be obtained by comparing the forward joint probabilities of lineages with path $\boldsymbol{s}$ and K divisions, and then summing over K since the exponential term is independent of the number of divisions.

We then condition the probability on the left-hand side on survival: $p_{\text{for}}(\boldsymbol{s}, \sigma = 1) = p_{\text{for}}^\star(\boldsymbol{s}) p_{\text{for}}(\sigma = 1, t)$, where by definition $p_{\text{for}}(\sigma = 1, t) = \exp(t\Gamma_t)$:

$$p_{\text{for}}^\star(\boldsymbol{s}) = p_{\text{for}}^\circ(\boldsymbol{s}) \exp\left[-\int_0^t dt'\, \gamma(s(t')) - t\Gamma_t\right].$$
(A13)

Now, when the time-integrated death rate can be replaced by the death rate of the time-integrated phenotype along a lineage, $\int_0^t dt'\, \gamma(s(t'))/t \approx \gamma(\bar{s})$, we obtain

$$p_{\text{for}}^\star(\bar{s}) = p_{\text{for}}^\circ(\bar{s}) \exp\left[-t\gamma(\bar{s}) - t\Gamma_t\right].$$
(A14)

Finally, when the forward distribution $p_{\text{for}}(\bar{s})$ of time-averaged phenotypes with death (which includes all lineages, even the dead ones) remains close to the same distribution $p_{\text{for}}^\circ(\bar{s})$ without death, $p_{\text{for}}(\bar{s}) \approx p_{\text{for}}^\circ(\bar{s})$, we plug Eq. (A14)

into Eq. (15) to obtain

$$h_t^\dagger(\bar{s}) = -\gamma(\bar{s}). \tag{A15}$$

Like for the division rate, the hypothesis $\int_0^t dt' \, \gamma(s(t'))/t \approx \gamma(\bar{s})$ is verified either when the death rate is only weakly nonlinear or when the autocorrelation of the trait dynamics is large. The assumption $p_{\text{for}}(\bar{s}) \approx p_{\text{for}}^\circ(\bar{s})$ reflects that death does not induce an important bias on phenotypic averages along trajectories when considering the ensemble of all lineages, both dead and alive. This is true for large autocorrelation of the trait dynamics, independently of the shape of the death rate, as explored in Appendix A 3.

3. Details on simulations and investigation of the effect of autocorrelation time

In this Appendix, we give the details of the numerical simulation used in Sec. IV C, and further investigate the role of the timescale of cell trait fluctuations.

In the simulations used for Fig. 4, we chose an Ornstein-Uhlenbeck process for $\ln s_t$, where $\ln s_{t+\Delta t}$ is randomly sampled from the normal distribution with mean $\mu + e^{-\beta\Delta t}(\ln s_t - \mu)$ and variance $\sigma^2(1 - e^{-2\beta\Delta t})$, with parameters $\Delta t = 0.05$, $\mu = -0.5\ln(1.09)$, $\sigma^2 = \ln(1.09)$, and $\beta = -0.6\ln r_g$. With these parameters, without selection due to $r(s)$ and $\gamma(s)$, s_t follows the log-normal distribution with mean 1.0 and standard deviation 0.3 in the steady state. A cell with the trait value s divides into $m = 2$ cells with the probability of $r(s)\Delta t$ at each time point, where $r(s) = 0.12 + 0.96s^{10}/(s^{10} + 1.1^{10})$. The initial states of two offspring are the same as that of the mother upon division. A cell with trait value s dies with the probability of $\gamma(s)\Delta t$, where $\gamma(s) = 1.2s^{10}/(s^{10} + 0.9^{10})$.

The parameter r_g indicates a strength of autocorrelation of s_t ($0 < r_g < 1$), and $\tau_{\text{corr}} = 1/\beta$ represents the autocorrelation time of the trait dynamics. In Fig. 4(c), we showed the results for final time $t = 4$ and $r_g = 0.8$ corresponding to $\tau_{\text{corr}} = 1/\beta = -1/(0.6\ln r_g) = 7.47 > t$. The timescale of the phenotypic variations is thus larger than the observation time, and therefore phenotypes fluctuate slowly. This implies that the time-averaged division and death rates can be well approximated by the rates for the time-averaged trajectories, and also that the distribution $p_{\text{for}}(\bar{s})$ recovers $p_{\text{for}}^\circ(\bar{s})$ with good precision, as illustrated in Fig. 6(a). The reason is that for slowly fluctuating traits, the shorter length of dead lineages does not bias the statistics, since phenotypes do not deviate more significantly for living lineages than for dead ones on the timescale of the observation. Note that this does not mean that the distribution for surviving lineages $p_{\text{for}}^\star(\bar{s})$ is not biased from $p_{\text{for}}(\bar{s})$. Indeed, because the death rate is an increasing function of the trait, cells are more likely to die with large trait values of the trait, and therefore $p_{\text{for}}^\star(\bar{s})$ is biased toward smaller cells as compared to $p_{\text{for}}(\bar{s})$. In Fig. 6(a) and in the other plots in the following, for simplicity, $p_{\text{for}}^\circ(\bar{s})$ is evaluated by running the simulations in the same way but with $\gamma(s) = 0$ and $m = 1$, given that in the absence of death the forward and lineage distributions are equivalent [27,38]. Figure 6(b) is the same as Fig. 4(c) from the main text, where the inference is successful.

We now run the simulation with the same observation time $t = 4$, but with a lower autocorrelation factor $r_g = 0.2$, corresponding to an autocorrelation time $\tau_{\text{corr}} = 1.04 < t$, leading to fast fluctuations of the trait.

First, in Figs. 6(c) and 6(d), we keep the same death rate as in Figs. 6(a) and 6(b) and in the main text. Because of fast fluctuations, we easily understand why $\int_0^t dt' \, r(s(t'))/t \neq r(\bar{s})$ [and $\int_0^t dt' \, \gamma(s(t'))/t \neq \gamma(\bar{s})$], which is why the inference of the division rate in Fig. 6(d) is slightly less accurate than when $r_g = 0.8$. More importantly, we observe that the survivor bias fails to recover the death rate correctly. This is explained by the significant difference between $p_{\text{for}}(\bar{s})$ and $p_{\text{for}}^\circ(\bar{s})$ observed in Fig. 6(c), induced by fast fluctuations. Indeed, dead lineages are shorter, and can thus explore values of $\bar{s}$ that are inaccessible for living lineages since phenotypes are likely to autoaverage in time.

Second, for completeness, we show in Figs. 6(e) and 6(f) the case of low autocorrelation and flat death rate, to illustrate that the accuracy of the inference of the death rate is truly determined by the trait dynamics, and independent of the shape of the death rate, even if linear, unlike what happens for the division rate. In this case, we observe in Fig. 6(e) that $p_{\text{for}}(\bar{s})$ and $p_{\text{for}}^\circ(\bar{s})$ are still significantly different despite the flat death rate, for the same reason as in the previous paragraph. This translates in a survivor bias that does not recover the flat death rate in Fig. 6(f). Moreover, we observe that $p_{\text{for}}^\star(\bar{s}) = p_{\text{for}}^\circ(\bar{s})$, which is expected from Eq. (A13).

APPENDIX B: SHUFFLED LANDSCAPES IN THE ANTIBIOTICS EXPERIMENT

To evaluate how flat the landscapes $h_t^\star$, $h_t^\dagger$, and h_t are in Fig. 3, we randomly shuffled the trait values and computed the corresponding new landscapes. Let (s_i, K_i, σ_i) denote the triplet of the trait value (logarithmic final cell size or preexposure elongation rate), the number of divisions, and the fate (0 for dead or 1 for alive) of the ith lineage ($i = 1, 2, \ldots, L$). Let $\{\rho(i)\}_{i=1,2,\ldots,L}$ be a random permutation of $i = 1, 2, \ldots, L$. A set of triplets with randomly shuffled traits is thus obtained as $\{(s_{\rho(i)}, K_i, \sigma_i)\}_{i=1,2,\ldots,L}$. We shuffled the trait values 10 000 times, and for each permutation we computed the landscapes. Then, we calculated the mean and the standard deviation of each landscape, shown as the curves and the shaded areas in Fig. 7. As expected, this procedure cuts the correlations between the trait values and the other values, namely the number of divisions and the survival, leading to flat curves. By comparing the actual landscapes with the shuffled landscapes, we confirm our analyses. Indeed, in Fig. 7(a), the dots lie in the shaded area, showing the independence of the preexposure elongation rate with both reproductive success and survival. On the other hand, in Fig. 7(b) for the final logarithmic cell size, the dots are significantly outside of the shaded areas, meaning that the fluctuations of these landscapes really capture fitness and survival effects, rather than random fluctuations.

APPENDIX C: EFFECT OF DEATH ON SELECTION

In this Appendix, we prove the covariance formula for the strength of selection on trajectories, Eq. (20). We start from
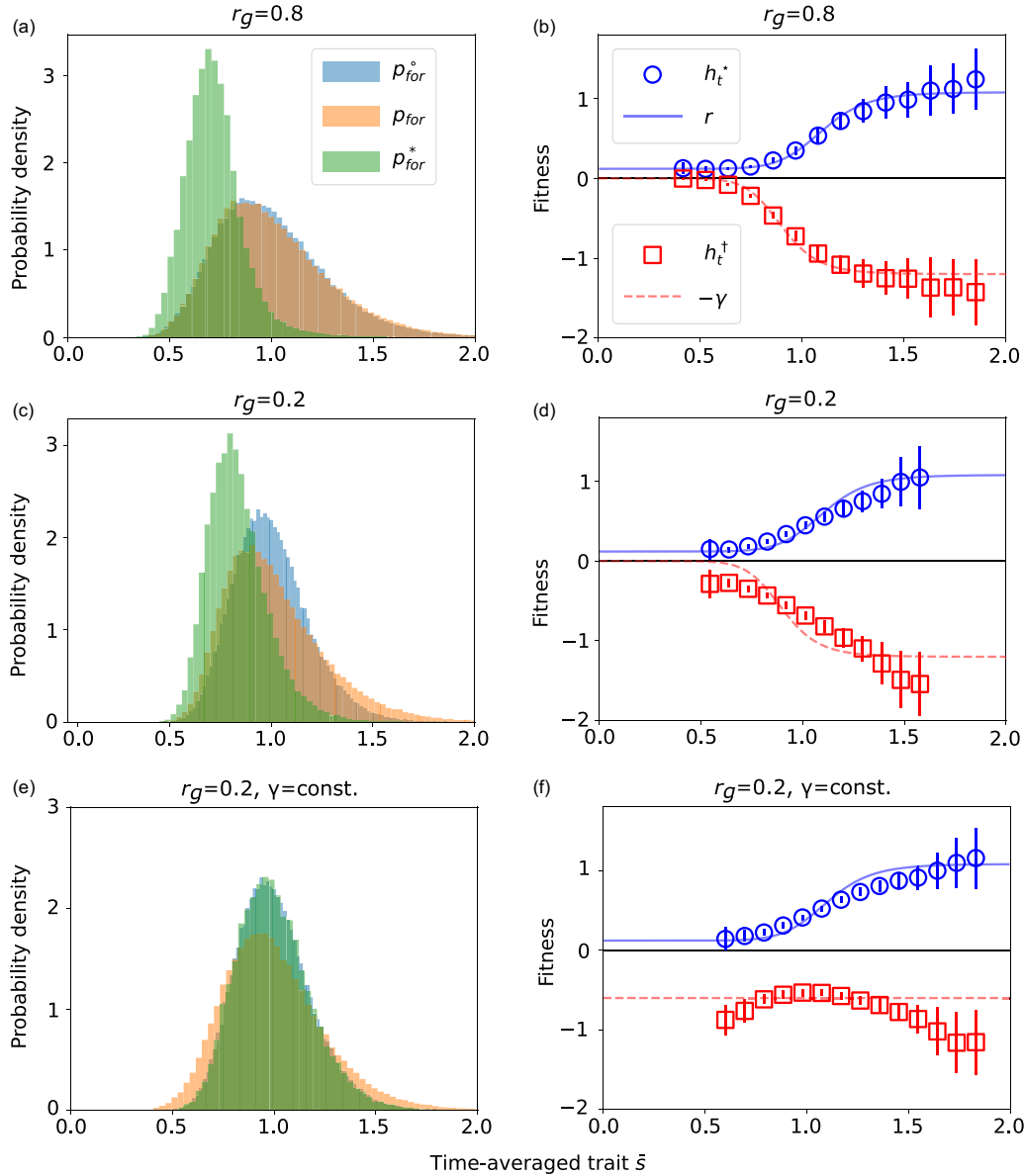
FIG. 6. Influence of autocorrelation factor r_g from simulations on the inference method. Left column: forward distributions (each histogram is evaluated from all 100 runs), right column: comparison between the division rate $r(\bar{s})$ [death rate $\gamma(\bar{s})$] and the fitness landscape $h_t^\star(\bar{s})$ [survivor bias $h_t^\dagger(\bar{s})$]. (a),(b) Large autocorrelation $r_g = 0.8$. (c)–(f) Small autocorrelation $r_g = 0.2$. (a)–(d) Death rate given by $\gamma(s) = 1.2 s^{10}/(s^{10} + 0.9^{10})$. (e),(f) Flat death rate $\gamma(s) = 0.6$. (a)–(f) Division rate $r(s) = 0.12 + 0.96 s^{10}/(s^{10} + 1.1^{10})$ and observation time $t = 4$.

the definition of the strength of selection given in Eq. (18):

$$\Pi_{\mathcal{S}} = \frac{1}{t} \int \mathcal{D}s \, (p_{\text{back}}(s,t) - p_{\text{for}}^\star(s,t)) \ln \left(\frac{p_{\text{back}}(s,t)}{p_{\text{for}}^\star(s,t)} \right), \tag{C1}$$

and we express $p_{\text{for}}^\star$ and p_{back} inside the logarithm using their counterparts in the absence of death. To do so, we integrate Eq. (A11) over trajectories s, and we identify the exponential bias as the survival probability Eq. (21), to obtain

$$n_t = n_t^\circ \, \langle p_{\text{surv}} \rangle_{\text{back}}^\circ. \tag{C2}$$

Then, we divide Eq. (A11) by n_t and use Eq. (C2) to obtain

$$p_{\text{back}}(s) = p_{\text{back}}^\circ(s) \, \frac{p_{\text{surv}}(s)}{\langle p_{\text{surv}} \rangle_{\text{back}}^\circ}. \tag{C3}$$

Similarly, the transformation for the forward distribution reads

$$p_{\text{for}}^\star(s) = p_{\text{for}}^\circ(s) \, \frac{p_{\text{surv}}(s)}{\langle p_{\text{surv}}(s) \rangle_{\text{for}}^\circ}, \tag{C4}$$

which is a rewriting of Eq. (A13).

Doing so, the survival probabilities cancel and we obtain

$$\Pi_{\mathcal{S}} = \frac{1}{t} \int \mathcal{D}s \, (p_{\text{back}}(s,t) - p_{\text{for}}^\star(s,t)) \ln \left(\frac{p_{\text{back}}^\circ(s,t)}{p_{\text{for}}^\circ(s,t)} \right) \tag{C5}$$

$$= \int \mathcal{D}s \, h_t^\circ(s)(p_{\text{back}}(s,t) - p_{\text{for}}^\star(s,t)). \tag{C6}$$
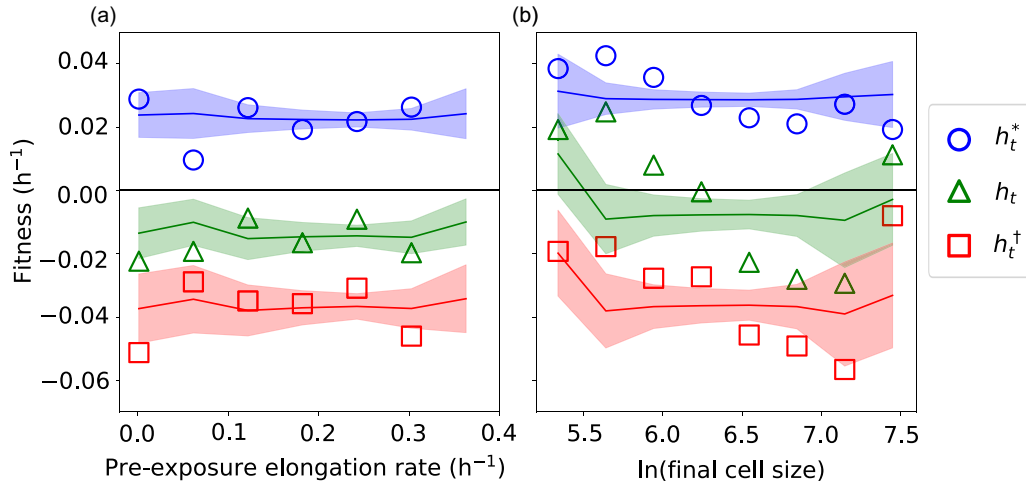
FIG. 7. Landscapes of (a) preexposure elongation rate and (b) final logarithmic cell size, where the curves and the shaded areas show the mean and mean$\pm$standard deviation of each landscape over 10 000 random shuffles of the trait values. See Appendix B for the details on the procedure for random shuffling of trait values.

Then, we subtract $\Pi_{\mathcal{S}}^{\circ} = \int \mathcal{D}\boldsymbol{s}\, h_t^{\circ}(\boldsymbol{s})(p_{\text{back}}^{\circ}(\boldsymbol{s}, t) - p_{\text{for}}^{\circ}(\boldsymbol{s}, t))$ from the relation above:

$$\Delta\Pi_{\mathcal{S}} = \int \mathcal{D}\boldsymbol{s}\, h_t^{\circ}(\boldsymbol{s})\Big[\big(p_{\text{back}}(\boldsymbol{s}, t) - p_{\text{back}}^{\circ}(\boldsymbol{s}, t)\big) - \big(p_{\text{for}}^{\star}(\boldsymbol{s}, t) - p_{\text{for}}^{\circ}(\boldsymbol{s}, t)\big)\Big] \tag{C7}$$

$$= \frac{1}{\langle p_{\text{surv}}(\boldsymbol{s})\rangle_{\text{back}}^{\circ}} \int \mathcal{D}\boldsymbol{s}\, h_t^{\circ}(\boldsymbol{s}) p_{\text{back}}^{\circ}(\boldsymbol{s}, t)\big[p_{\text{surv}}(\boldsymbol{s}) - \langle p_{\text{surv}}(\boldsymbol{s})\rangle_{\text{back}}^{\circ}\big]$$

$$- \frac{1}{\langle p_{\text{surv}}(\boldsymbol{s})\rangle_{\text{for}}^{\circ}} \int \mathcal{D}\boldsymbol{s}\, h_t^{\circ}(\boldsymbol{s}) p_{\text{for}}^{\circ}(\boldsymbol{s}, t)\big[p_{\text{surv}}(\boldsymbol{s}) - \langle p_{\text{surv}}(\boldsymbol{s})\rangle_{\text{for}}^{\circ}\big] \tag{C8}$$

$$= \frac{\text{Cov}_{\text{back}}^{\circ}\big(h_t^{\circ}, p_{\text{surv}}\big)}{\langle p_{\text{surv}}\rangle_{\text{back}}^{\circ}} - \frac{\text{Cov}_{\text{for}}^{\circ}\big(h_t^{\circ}, p_{\text{surv}}\big)}{\langle p_{\text{surv}}\rangle_{\text{for}}^{\circ}}. \tag{C9}$$

APPENDIX D: DEATH-INDUCED DECREASE IN POPULATION GROWTH RATE

In this Appendix, we derive the general fluctuation-response relation Eq. (31) and we give an alternative form in terms of a Kullback-Leibler divergence.

First, taking the logarithm of Eq. (C2) leads to

$$\Lambda_t - \Lambda_t^{\circ} = \frac{1}{t} \ln\langle p_{\text{surv}}\rangle_{\text{back}}^{\circ}. \tag{D1}$$

In particular, if death occurs only at division with probability $1 - 2^{-\epsilon}$, then the survival probability along a lineage with K divisions is $p_{\text{surv}}(K) = 2^{-\epsilon K}$. Then for small ϵ, at first order $\ln\langle p_{\text{surv}}\rangle_{\text{back}}^{\circ} = -\epsilon\langle K\rangle_{\text{back}}^{\circ} \ln 2 = -\epsilon[t\Lambda_t^{\circ} + \mathcal{D}_{\text{KL}}(p_{\text{back}}^{\circ}(K)||p_{\text{for}}^{\circ}(K))]$, and thus we recover the result from [10].

The growth rate difference in Eq. (D1) can be expressed as a distance between the statistics with and without death. To do so, we compute the KL divergence between the backward distributions in the presence and absence of death using Eq. (C3):

$$\mathcal{D}_{\text{KL}}(p_{\text{back}}^{\circ}(\boldsymbol{s})||p_{\text{back}}(\boldsymbol{s})) = \ln\langle p_{\text{surv}}\rangle_{\text{back}}^{\circ} - \langle \ln p_{\text{surv}}\rangle_{\text{back}}^{\circ}, \tag{D2}$$

where $\ln p_{\text{surv}}(\boldsymbol{s}) = -t\overline{\gamma}$ by definition. The death-induced decrease in population growth rate is therefore given by

$$\Lambda_t - \Lambda_t^{\circ} = -\langle\overline{\gamma}\rangle_{\text{back}}^{\circ} + \frac{1}{t}\mathcal{D}_{\text{KL}}\big(p_{\text{back}}^{\circ}(\boldsymbol{s})||p_{\text{back}}(\boldsymbol{s})\big), \tag{D3}$$

where the first term only depends on the dynamics in the absence of death, and the second term is an information-theoretic distance between the statistics with and without death. This relation is similar to recently derived fluctuation-response inequalities valid arbitrarily far from equilibrium which also involve a distance between the statistics before and after a perturbation [28,58].

By comparing Eq. (D3) with Eq. (32) valid for a small death rate $\epsilon\gamma$, the KL divergence is also linked to the variance in the time-integrated death rate:

$$\mathcal{D}_{\text{KL}}\big(p_{\text{back}}^{\circ}(\boldsymbol{s})||p_{\text{back}}(\boldsymbol{s})\big) = \frac{t^2}{2}\text{Var}_{\text{back}}^{\circ}(\overline{\gamma})\epsilon^2 + O(\epsilon^3). \tag{D4}$$

Therefore, this divergence is further interpreted as a gain in population fitness due to the variability in the time-averaged death rate discussed in the main text.

APPENDIX E: DETAILS ON THE MODELS OF CELL SIZE CONTROL

1. General model of cell size control with death

In this Appendix, we give the equations governing the model of cell size control discussed in Sec. VI. The evolution of the expected number $n(a, x_b, K, t)$ of cells of age a, newborn size x_b, and K divisions at time t is given by the following population-balance equation (PBE):

$$\partial_t n(a, x_b, K, t) = -\partial_a n(a, x_b, K, t) - [r(a, x_b) + \gamma(a, x_b)]n(a, x_b, K, t), \tag{E1}$$

$$n(a = 0, x_b, K, t) = m \int da' dx_b' r(a', x_b') \Sigma(x_b | x_b', a') n(a', x_b', K - 1, t), \tag{E2}$$

where the second line is the boundary condition for newborn cells, which accounts for divisions of cells with $K - 1$ divisions that divide a Kth time at time t to give birth to cells of size x_b.

The PBE and the boundary condition, Eqs. (E1) and (E2), can be recast at the level of the expected backward probability distribution by summing over K and using $p_{\text{back}}(a, x_b, t) = n(a, x_b, t)/n(t)$, where $n(t) = \int da dx_b\, n(a, x_b, t)$ is the total expected number of cells at time t:

$$\partial_t p_{\text{back}}(a, x_b, t) = -\partial_a [p_{\text{back}}(a, x_b, t)] - [r(a, x_b) + \gamma(a, x_b) + \Lambda_p(t)]p_{\text{back}}(a, x_b, t), \tag{E3}$$

$$p_{\text{back}}(a = 0, x_b, t) = m \int da' dx_b'\, \Sigma(x_b | x_b', a') r(a', x_b') p_{\text{back}}(a', x_b', t), \tag{E4}$$

where we defined

$$\Lambda_p(t) = \frac{1}{n(t)} \frac{dn}{dt}, \tag{E5}$$

the instantaneous population growth rate. Direct integration of Eq. (E3) over a and x_b shows that, in the presence of death, the instantaneous population growth rate is the net difference between the backward-averaged division and death rates:

$$\Lambda_p(t) = \int da dx_b\, [(m - 1)r(a, x_b) - \gamma(a, x_b)]p_{\text{back}}(a, x_b, t). \tag{E6}$$

To recast the PBE for the expected forward probability, we use $p_{\text{for}}(a, x_b, K, \sigma = 1, t) = n(a, x_b, K, t)m^{-K}n(0)^{-1}$ and then sum over K, which leads to

$$\partial_t p_{\text{for}}(a, x_b, \sigma = 1, t) = -\partial_a [p_{\text{for}}(a, x_b, \sigma = 1, t)] - [r(a, x_b) + \gamma(a, x_b)]p_{\text{for}}(a, x_b, \sigma = 1, t), \tag{E7}$$

$$p_{\text{for}}(a = 0, x_b, \sigma = 1, t) = \int da' dx_b'\, \Sigma(x_b | x_b', a') r(a', x_b') p_{\text{for}}(a', x_b', \sigma = 1, t). \tag{E8}$$

Integrating Eq. (E7) over a and x_b gives

$$\partial_t p_{\text{for}}(\sigma = 1, t) = -p_{\text{for}}(\sigma = 1, t) \int da dx_b\, \gamma(a, x_b) p_{\text{for}}(a, x_b, t | \sigma = 1). \tag{E9}$$

We then define the instantaneous decrease rate of the forward probability of survival:

$$\Gamma_p(t) = \frac{1}{p_{\text{for}}(\sigma = 1, t)} \frac{dp_{\text{for}}(\sigma = 1)}{dt} = -\int da dx_b\, \gamma(a, x_b) p_{\text{for}}(a, x_b, t | \sigma = 1). \tag{E10}$$

Finally, one derives the PBE for the expected forward distribution conditioned on survival: $p_{\text{for}}^\star(a, x_b, t) = p_{\text{for}}(a, x_b, t | \sigma = 1) = p_{\text{for}}(a, x_b, \sigma = 1, t)/p_{\text{for}}(\sigma = 1, t)$:

$$\partial_t p_{\text{for}}^\star(a, x_b, t) = -\partial_a p_{\text{for}}^\star(a, x_b, t) - [r(a, x_b) + \gamma(a, x_b) + \Gamma_p(t)]p_{\text{for}}^\star(a, x_b, t), \tag{E11}$$

$$p_{\text{for}}^\star(a = 0, x_b, t) = \int da' dx_b' r(a', x_b') \Sigma(x_b | x_b', a') p_{\text{for}}^\star(a', x_b', t). \tag{E12}$$

Note that Eqs. (E11) and (E12) can be obtained from the PBE and boundary condition for the backward distribution Eqs. (E3) and (E4) by setting $m = 1$ in the boundary term and replacing $\Lambda_p(t)$ by $\Gamma_p(t)$.

2. Link with the sizer and the adder models

The size x of a cell is determined by its size at birth x_b, its age a, and the single cell growth rate:

$$\frac{dx}{da} = \nu(x), \tag{E13}$$

therefore the model presented in Appendix E1 accounts for the sizer model, where division is triggered by cell size, and for the adder model, where it is triggered by the increment of volume since birth $\Delta = x - x_b$. Let us show explicitly in this Appendix how the sizer and adder PBEs are obtained.

First, we reparametrize the PBE and the boundary condition for the number $\hat{n}(x, x_b, t)$ of cells of size x at time t born with size x_b, with the following change of variables:

$$n(a, x_b, t) = \hat{n}(x(a, x_b), x_b, t)\nu(x(a, x_b)), \tag{E14}$$

so that they read

$$\partial_t \hat{n}(x, x_b, t) = -\partial_x[\nu(x)\hat{n}(x, x_b, t)] - [\hat{r}(x, x_b) + \hat{\gamma}(x, x_b)]\hat{n}(x, x_b, t), \tag{E15}$$

$$\nu(x_b)\hat{n}(x_b, x_b, t)$$
$$= m \int dx' dx_b' \hat{r}(x', x_b') \hat{\Sigma}(x_b|x', x_b') \hat{n}(x', x_b', t), \tag{E16}$$

where we defined the following rates:

$$\hat{r}(x(a, x_b), x_b) = r(a, x_b), \tag{E17}$$

$$\hat{\gamma}(x(a, x_b), x_b) = \gamma(a, x_b), \tag{E18}$$

$$\hat{\Sigma}(x_b|x'(a', x_b'), x_b') = \Sigma(x_b|x_b', a'). \tag{E19}$$

In the classical sizer model, division is triggered by cell size only. To generalize this model to a dynamics with death, we further assume that death is also triggered by cell size only. We thus make the following assumptions:

$$\hat{r}(x, x_b) \equiv \hat{r}(x), \tag{E20}$$

$$\hat{\gamma}(x, x_b) \equiv \hat{\gamma}(x), \tag{E21}$$

$$\hat{\Sigma}(x|x', x_b') \equiv \hat{\Sigma}(x|x'). \tag{E22}$$

We integrate Eq. (E15) over x_b from 0 to x, where the integration of the derivative term yields a boundary term:

$$\int_0^x dx_b \, \partial_x[\nu(x)\hat{n}(x, x_b, t)] = \partial_x[\nu(x)\hat{n}(x, t)] - \nu(x)\hat{n}(x, x, t), \tag{E23}$$

which is expressed by plugging Eq. (E16). Finally, the PBE for the expected number $\hat{n}(x, t) = \int_0^x dx_b \, \hat{n}(x, x_b, t)$ of cells with size x at time t reads

$$\partial_t \hat{n}(x, t) = -\partial_x[\nu(x)\hat{n}(x, t)] - [\hat{r}(x) + \hat{\gamma}(x)]\hat{n}(x, t) + m \int dx' \hat{r}(x') \hat{\Sigma}(x|x')\hat{n}(x', t), \tag{E24}$$

which is the sizer equation with a death term.

For the adder, let us make a change of variable from (x, x_b) to (x, Δ), where $\Delta = x - x_b$ is the added volume since birth. The reparametrization for the expected number $\tilde{n}(x, \Delta, t)$ of cells with size x and added volume Δ at time t is then given by

$$\tilde{n}(x, \Delta(x, x_b), t) = \hat{n}(x, x_b, t), \tag{E25}$$

and we define the rates in the new system of coordinates:

$$\tilde{r}(x, \Delta(x, x_b)) = \hat{r}(x, x_b), \tag{E26}$$

$$\tilde{\gamma}(x, \Delta(x, x_b)) = \hat{\gamma}(x, x_b), \tag{E27}$$

$$\tilde{\Sigma}(x|x', \Delta'(x', x_b')) = \hat{\Sigma}(x|x'). \tag{E28}$$

In the adder model without death, the separation of variables for the division rate:

$$\tilde{r}(x, \Delta) = \nu(x)\zeta(\Delta), \tag{E29}$$

where ζ is the division rate per unit volume, is further assumed to account for the independence between size at birth and added volume between birth and division. Different choices for the death rate $\tilde{\gamma}(x, \Delta)$ are discussed in Appendix G.

3. Population growth and decay rates as average division and death rates

The rates Λ_t and Γ_t introduced in the main text are stochastic quantities, defined for any single realization of the stochastic branching process, and therefore they cannot be expressed explicitly in general. However, when averaging over many population trees, they can be linked to the division and death rates.

Indeed, let us first compute the time-averaged version of the instantaneous rates $\Lambda_p(t)$ and $\Gamma_p(t)$ introduced in the previous section:

$$\Lambda_t^{\text{PBE}} = \overline{\Lambda_p} = \frac{1}{t} \ln\left[\frac{n(t)}{n_0}\right], \tag{E30}$$

$$\Gamma_t^{\text{PBE}} = \overline{\Gamma_p} = \frac{1}{t} \ln p_{\text{for}}(\sigma = 1, t), \tag{E31}$$

where we defined the time-average $\overline{f} = t^{-1} \int_0^t dt' f(t')$. We recall that $n(t)$ and $p_{\text{for}}(\sigma = 1, t)$ are the *expected* number of cells and forward probability of survival.

On the other hand, a given population starting with N_0 cells consists of N_0 independent trees, so that $N(t)/N_0$ is the empirical average of the number of cells produced after a time t by a single initial cell:

$$\frac{N(t)}{N_0} = \frac{1}{N_0} \sum_{i=1}^{N_0} N^i(t), \tag{E32}$$

where $N^i(t)$ is the number of cells at time t coming from ancestor cell i. Thus, in the limit $N_0 \to \infty$, the empirical average $N(t)/N_0$ converges to the true average $n(t)/n_0$, and Λ_t and Γ_t converge to Λ_t^{PBE} and Γ_t^{PBE}. Combining this result with Eqs. (E6) and (E10), we obtain

$$\lim_{N_0 \to \infty} \Lambda_t = \Lambda_t^{\text{PBE}} = \overline{\langle (m-1)r - \gamma \rangle_{\text{back}}}, \tag{E33}$$

$$\lim_{N_0 \to \infty} \Gamma_t = \Gamma_t^{\text{PBE}} = -\overline{\langle \gamma \rangle_{\text{for}}^\star}. \tag{E34}$$

APPENDIX F: GENERALIZED POWELL'S RELATION

In this Appendix, we derive Eq. (40) from the main text, we provide an explicit expression of the constant $Z(x_b)$, and we compare our result to the one obtained in [33]. To do so, we first derive an alternative form of Eq. (40) for the joint distribution of generation time and newborn size.

1. Generalized Powell's relation with joint probabilities

We define the steady-state joint distribution $f(\tau, x_b)$ of generation time τ and newborn size x_b as the ratio of the number of cells born with size x_b and dividing at age τ at a given snapshot time, to the total number of cells dividing in this snapshot, weighted in either the backward sampling or forward sampling conditioned on survival:

$$f(\tau, x_b) = \frac{r(\tau, x_b)p(\tau, x_b)}{\int d\tau' dx_b' \, r(\tau', x_b')p(\tau', x_b')}, \tag{F1}$$

where the joint distributions of age a and newborn size x_b are the steady-state solutions to Eqs. (E3) and (E11) and read

$$p_{\text{back}}(a, x_b) = p_{\text{back}}(0, x_b)\exp\left[-\Lambda a - \int_0^a da' \, (r(a', x_b) + \gamma(a', x_b))\right], \tag{F2}$$

$$p_{\text{for}}^\star(a, x_b) = p_{\text{for}}^\star(0, x_b)\exp\left[-\Gamma a - \int_0^a da' \, (r(a', x_b) + \gamma(a', x_b))\right]. \tag{F3}$$

These distributions are independent of the time of the snapshot in steady-state. Integrating the boundary conditions Eqs. (E4) and (E12) over x_b and using the normalization of kernel Σ, namely $\forall(a', x_b')$, $\int dx_b \, \Sigma(x_b|x_b', a') = 1$, we get that the denominators of Eq. (F1) are given by

$$\int dx_b \, p_{\text{back}}(0, x_b) = m\int d\tau dx_b' r(\tau, x_b')p_{\text{back}}(\tau, x_b'), \tag{F4}$$

$$\int dx_b \, p_{\text{for}}^\star(0, x_b) = \int d\tau dx_b' r(\tau, x_b')p_{\text{for}}^\star(\tau, x_b'). \tag{F5}$$

Then, we identify the distribution of newborn size:

$$\rho^{\text{nb}}(x_b) = \frac{p(0, x_b)}{\int dx_b' \, p(0, x_b')} \tag{F6}$$

for both the forward and backward probabilities.

Finally, combining the results above we obtain

$$f_{\text{back}}(\tau, x_b) = m\rho_{\text{back}}^{\text{nb}}(x_b)r(\tau, x_b)\exp\left[-\Lambda\tau - \int_0^\tau da \, (r(a, x_b) + \gamma(a, x_b))\right], \tag{F7}$$

$$f_{\text{for}}^\star(\tau, x_b) = \rho_{\text{for}}^{\text{nb},\star}(x_b)r(\tau, x_b)\exp\left[-\Gamma\tau - \int_0^\tau da \, (r(a, x_b) + \gamma(a, x_b))\right], \tag{F8}$$

and the generalized Powell's equation reads

$$f_{\text{back}}(\tau, x_b) = m\frac{\rho_{\text{back}}^{\text{nb}}(x_b)}{\rho_{\text{for}}^{\text{nb},\star}(x_b)}f_{\text{for}}^\star(\tau, x_b)e^{-(\Lambda-\Gamma)\tau}. \tag{F9}$$

In the absence of mother-daughter correlations, that is, when $\Sigma(x_b|x_b', a') \equiv \hat{\Sigma}(x_b)$, the newborn distributions are unbiased: $\rho_{\text{back}}^{\text{nb}} = \rho_{\text{for}}^{\text{nb},\star} = \hat{\Sigma}$, which is a direct consequence of the boundary conditions. Therefore, the fraction in Eq. (F9) is canceled, and Eq. (F9) can be integrated over x_b to recover Powell's relation in the presence of death but without correlations: $f_{\text{back}}(\tau) = mf_{\text{for}}^\star(\tau)e^{-(\Lambda-\Gamma)\tau}$.

2. Generalized Powell's relation with conditional probabilities

We now define the distributions $f(\tau|x_b)$ of generation time conditioned on newborn size x_b as

$$f(\tau|x_b) = \frac{r(\tau, x_b)p(\tau, x_b)}{\int d\tau' \, r(\tau', x_b)p(\tau', x_b)} \tag{F10}$$

$$= \frac{f(\tau, x_b)}{\rho^{\text{d}}(x_b)}, \tag{F11}$$

where we identified the distribution of newborn sizes at division:

$$\rho^{\text{d}}(x_b) = \frac{\int d\tau' \, r(\tau', x_b)p(\tau', x_b)}{\int d\tau' dx_b' \, r(\tau', x_b')p(\tau', x_b')}. \tag{F12}$$

The conditioned distributions are thus given by

$$f_{\text{back}}(\tau|x_b) = m \frac{\rho_{\text{back}}^{\text{nb}}(x_b)}{\rho_{\text{back}}^{\text{d}}(x_b)} r(\tau, x_b) \exp\left[-\Lambda\tau - \int_0^\tau da\, (r(a, x_b) + \gamma(a, x_b))\right], \tag{F13}$$

$$f_{\text{for}}^\star(\tau|x_b) = \frac{\rho_{\text{for}}^{\text{nb},\star}(x_b)}{\rho_{\text{for}}^{\text{d},\star}(x_b)} r(\tau, x_b) \exp\left[-\Gamma\tau - \int_0^\tau da\, (r(a, x_b) + \gamma(a, x_b))\right], \tag{F14}$$

and finally, Powell's relation on conditional distributions reads

$$f_{\text{back}}(\tau|x_b) = \frac{m f_{\text{for}}^\star(\tau|x_b) e^{-(\Lambda-\Gamma)\tau}}{Z(x_b)}, \tag{F15}$$

with the normalization constant

$$Z(x_b) = \frac{\rho_{\text{for}}^{\text{nb},\star}(x_b)}{\rho_{\text{back}}^{\text{nb}}(x_b)} \frac{\rho_{\text{back}}^{\text{d}}(x_b)}{\rho_{\text{for}}^{\text{d},\star}(x_b)}. \tag{F16}$$

3. Comparison with Sughiyama's work

In [33], the authors also derived a generalized Powell's relation in the presence of death and mother-daughter correlations. Unlike in our model, they allowed the number of daughter cells at each division to be stochastic and added a probability of death upon division in addition to the death rate between divisions. Moreover, their relation compares the backward distribution in the presence of death to the forward distribution in the absence of death, which we denote $f_{\text{for}}^\circ(\tau|x_b)$. In the simple case where there is no death upon division and where the number of daughter cells is fixed to m, their result reads

$$f_{\text{back}}(\tau|x_b) = \frac{m f_{\text{for}}^\circ(\tau|x_b) e^{-\Lambda\tau - \int_0^\tau da\,\gamma(a,x_b)}}{\tilde{Z}(x_b)}. \tag{F17}$$

We show here that their relation can be recovered from Eq. (F15). In the absence of death, the forward distribution is simply given by

$$f_{\text{for}}^\circ(\tau|x_b) = r(\tau, x_b) \exp\left[-\int_0^\tau da\, r(a, x_b)\right], \tag{F18}$$

which can also be seen from the normalization of f_{for}° in Eq. (F14), which imposes $\rho_{\text{for}}^{\text{nb},\circ}(x_b) = \rho_{\text{for}}^{\text{d},\circ}(x_b)$. Therefore, by comparing Eqs. (F14) and (F18) we obtain

$$f_{\text{for}}^\star(\tau|x_b) = \frac{\rho_{\text{for}}^{\text{nb},\star}(x_b)}{\rho_{\text{for}}^{\text{d},\star}(x_b)} f_{\text{for}}^\circ(\tau|x_b) \exp\left[-\Gamma\tau - \int_0^\tau da\,\gamma(a, x_b)\right]. \tag{F19}$$

Finally, we plug Eq. (F19) into Eq. (F15) to obtain Eq. (F17), with

$$\tilde{Z}(x_b) = Z(x_b) \frac{\rho_{\text{for}}^{\text{d},\star}(x_b)}{\rho_{\text{for}}^{\text{nb},\star}(x_b)} = \frac{\rho_{\text{back}}^{\text{d}}(x_b)}{\rho_{\text{back}}^{\text{nb}}(x_b)}. \tag{F20}$$

We believe that our result, Eq. (F15), has the following advantages over Sughiyama's result. First, it compares two distributions evaluated in the same experimental setup, in the presence of death, which makes sense in light of the definition of fitness proposed in the main text. Second, we obtained explicit expressions for the constants $Z(x_b)$ and $\tilde{Z}(x_b)$, in terms of the distributions of newborn x_b at birth and at division.

Third, in [33] the authors needed the Direction-Time Hypothesis, which imposes that the kernel is independent of the age at division: $\Sigma(x_b|x_b', a') \equiv \Sigma(x_b|x_b')$. In their article, x_b is not the size at birth, but a general phenotypic or genotypic type, within the multitype age model. Importantly, this assumption implies that their model cannot account for the sizer and adder models, since in these models the kernel should describe the partitioning of volume at division $\Sigma(x_b|x')$, where both age a' and newborn size x_b' are needed to compute mother size x', as detailed in Appendix E 2. On the other hand, we did not need this assumption.

APPENDIX G: ADDER PROPERTY CONCEALED BY SURVIVAL

In this Appendix, we derive the bias between the forward distributions of added volume between birth and division Δ_d conditioned on birth size x_b in the presence and absence of death, Eq. (41). This distribution is given by the following change of variable:

$$f_{\text{for}}^\star(\Delta_d|x_b) d\Delta_d = f_{\text{for}}^\star(\tau|x_b) d\tau, \tag{G1}$$

where the distribution of generation time τ conditioned on newborn size x_b is given by Eq. (F14):

$$f_{\text{for}}^\star(\tau|x_b) = \frac{\rho_{\text{for}}^{\text{nb},\star}(x_b)}{\rho_{\text{for}}^{\text{d},\star}(x_b)} \tilde{r}(x(\tau), x(\tau) - x_b)$$

$$\times \exp\left[-\Gamma\tau - \int_0^\tau [\tilde{r}(x(t), x(t) - x_b)\right.$$

$$\left. + \tilde{\gamma}(x(t), x(t) - x_b)]dt\right]. \tag{G2}$$

Using Eqs. (E29) and (G1) with $d\Delta_d/d\tau = \nu(x(\tau))$, we get

$$f_{\text{for}}^\star(\Delta_d|x_b) = \frac{\rho_{\text{for}}^{\text{nb},\star}(x_b)}{\rho_{\text{for}}^{\text{d},\star}(x_b)} \zeta(\Delta_d) \exp\left[-\Gamma\tau(x_b, \Delta_d)\right.$$

$$\left. - \int_0^{\Delta_d} \left[\zeta(\Delta) + \frac{\tilde{\gamma}(x_b + \Delta, \Delta)}{\nu(x_b + \Delta)}\right] d\Delta\right], \tag{G3}$$

where we identify the distribution of added volume in the absence of death:

$$f_{\text{for}}^\circ(\Delta_d) = \zeta(\Delta_d) \exp\left[-\int_0^{\Delta_d} d\Delta\, \zeta(\Delta)\right], \tag{G4}$$

and where we define the normalization constant

$$Y(x_b) = \frac{\rho_{\text{for}}^{\text{d},\star}(x_b)}{\rho_{\text{for}}^{\text{nb},\star}(x_b)}. \tag{G5}$$

Finally, we recovered Eq. (41) from the main text.

To complement this result, let us give several simple examples where the adder property actually remains observable with the forward distribution conditioned on survival. First, when the increment of volume Δ_d between birth and division is fixed, then the distribution of added volume is a Dirac delta function, independent of the size at birth, for every distribution. Second, in the case of a uniform death/dilution rate $\tilde{\gamma}(x, \Delta) \equiv \gamma$, we showed in Eq. (37) that $\Gamma = -\gamma$. Therefore, the two terms in the exponential bias cancel out, as $\int_0^{\Delta_d} d\Delta \; \nu(x_b + \Delta)^{-1} = \tau(x_b, \Delta_d)$. Once again, we find that a constant death rate does not affect quantitatively the observables. Third, the exponential bias becomes independent of the newborn volume when both of its terms are independent of x_b. This is the case when (i) cells grow linearly with a rate $\nu(x) = \nu_1$ such that $\tau(x_b, \Delta_d) \equiv \tau(\Delta_d) = \Delta_d/\nu_1$, and (ii) the death rate obeys a similar separation of variables as the division rate, namely $\tilde{\gamma}(x, \Delta) \equiv \nu(x)\xi(\Delta)$, with $\xi(\Delta)$ the death rate per unit volume. In the second and third cases, note that the independence of the exponential bias from x_b implies that the factor $Y(x_b)$ is also independent of x_b. This is because the normalization of the distribution, $\forall x_b, \; \int_0^\infty d\Delta_d \; f_{\text{for}}^\star(\Delta_d|x_b) = 1$, imposes $Y(x_b) \equiv Y$.

Finally, let us mention that it has also been proposed in [59] to test the adder property on the basis of the value of the Pearson correlation coefficient between mother and daughter properties (typically size at birth), instead of the independence of the lineage distribution of added volume from the newborn volume. This alternative test has been used, for example, in [14] to complement the analysis, leading to the same conclusion that data were best accounted for by the adder model. It remains open to investigate the effect of the survivor bias on the Pearson correlation coefficient.

---

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

[2] S. Bakshi, E. Leoncini, C. Baker, S. J. Cañas-Duarte, B. Okumus, and J. Paulsson, Tracking bacterial lineages in complex and dynamic environments with applications for growth control and persistence, Nat. Microbiol. **6**, 783 (2021).

[3] W. Keil, L. M. Kutscher, S. Shaham, and E. D. Siggia, Long-term high-resolution imaging of developing *C. elegans* larvae with microfluidics, Dev. Cell **40**, 202 (2017).

[4] E. Biselli, S. J. Schink, and U. Gerland, Slower growth of *Escherichia coli* leads to longer survival in carbon starvation due to a decrease in the maintenance rate, Mol. Syst. Biol. **16**, e9478 (2020).

[5] G. T. Eisenhoffer, P. D. Loftus, M. Yoshigi, H. Otsuna, C.-B. Chien, P. A. Morcos, and J. Rosenblatt, Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia, Nature (London) **484**, 546 (2012).

[6] E. Marinari, A. Mehonic, S. Curran, J. Gale, T. Duke, and B. Baum, Live-cell delamination counterbalances epithelial growth to limit tissue overcrowding, Nature (London) **484**, 542 (2012).

[7] 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).

[8] A. Koldaeva, H.-F. Tsai, A. Q. Shen, and S. Pigolotti, Population genetics in microchannels, Proc. Natl. Acad. Sci. USA **119**, e2120821119 (2022).

[9] 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).

[10] S. Yamauchi, T. Nozoe, R. Okura, E. Kussell, and Y. Wakamoto, A unified framework for measuring selection on cellular lineages and traits, eLife **11**, e72299 (2022).

[11] 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).

[12] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).

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

[14] S. Taheri-Araghi, S. Bradde, J. Sauls, N. Hill, P. Levin, J. Paulsson, M. Vergassola, and S. Jun, Cell-size control and homeostasis in bacteria, Curr. Biol. **25**, 385 (2015).

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

[16] A. Genthon, Analytical cell size distribution: lineage-population bias and parameter inference, J. R. Soc., Interface **19**, 20220405 (2022).

[17] P. Thomas, Analysis of cell size homeostasis at the single-cell and population level, Front. Phys. **6**, 64 (2018).

[18] M. Hoffmann and A. Olivier, Nonparametric estimation of the division rate of an age dependent branching process, Stoch. Proc. Appli. **126**, 1433 (2016).

[19] M. Doumic, M. Hoffmann, P. Reynaud-Bouret, and V. Rivoirard, Nonparametric estimation of the division rate of a size-structured population, SIAM J. Numer. Anal. **50**, 925 (2012).

[20] M. Doumic and M. Hoffmann, Individual and population approaches for calibrating division rates in population dynamics: Application to the bacterial cell cycle, *Modeling and Simulation for Collective Dynamics* (World Scientific, 2023), pp. 1–81.

[21] J. Lin and A. Amir, Disentangling Intrinsic and Extrinsic Gene Expression Noise in Growing Cells, Phys. Rev. Lett. **126**, 078101 (2021).

[22] R. A. Fisher, *The Genetical Theory of Natural Selection: A Complete Variorum Edition*, edited by J. H. Bennet (Oxford University Press, Oxford, UK, 2000).

[23] V. Mustonen and M. Lassig, Fitness flux and ubiquity of adaptive evolution, Proc. Natl. Acad. Sci. USA **107**, 4248 (2010).

[24] T. J. Kobayashi and Y. Sughiyama, Fluctuation Relations of Fitness and Information in Population Dynamics, Phys. Rev. Lett. **115**, 238102 (2015).

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

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

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

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

[29] Y. Murashita, K. Funo, and M. Ueda, Nonequilibrium equalities in absolutely irreversible processes, Phys. Rev. E **90**, 042110 (2014).

[30] F. Jafarpour, E. Levien, and A. Amir, Evolutionary dynamics in non-Markovian models of microbial populations, Phys. Rev. E (to be published).

[31] P. Thomas, Single-cell histories in growing populations: relating physiological variability to population growth, bioRxiv:https://doi.org/10.1101/100495.

[32] P. Thomas, Making sense of snapshot data: ergodic principle for clonal cell populations, J. R. Soc. Interface **14**, 20170467 (2017).

[33] Y. Sughiyama, S. Nakashima, and T. J. Kobayashi, Fitness response relation of a multitype age-structured population dynamics, Phys. Rev. E **99**, 012413 (2019).

[34] S. Jun, F. Si, R. Pugatch, and M. Scott, Fundamental principles in bacterial physiology—history, recent progress, and the future with focus on cell size control: a review, Rep. Prog. Phys. **81**, 056601 (2018).

[35] C. Cadart, L. Venkova, P. Recho, M. C. Lagomarsino, and M. Piel, The physics of cell-size regulation across timescales, Nat. Phys. **15**, 993 (2019).

[36] E. Levien, J. Kondev, and A. Amir, The interplay of phenotypic variability and fitness in finite microbial populations, J. R. Soc., Interface **17**, 20190827 (2020).

[37] E. Toprak, A. Veres, J.-B. Michel, R. Chait, D. L. Hartl, and R. Kishony, Evolutionary paths to antibiotic resistance under dynamically sustained drug selection, Nat. Genet. **44**, 101 (2012).

[38] E. Levien, T. GrandPre, and A. Amir, Large Deviation Principle Linking Lineage Statistics to Fitness in Microbial Populations, Phys. Rev. Lett. **125**, 048102 (2020).

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

[40] Y. Sughiyama and T. J. Kobayashi, The explicit form of the rate function for semi-Markov processes and its contractions, J. Phys. A **51**, 125001 (2018).

[41] E. O. Powell, A note on Koch & Schaechter's hypothesis about growth and fission of bacteria, J. Gen. Microbiol. **37**, 231 (1964).

[42] E. O. Powell, Generation times of bacteria: Real and artificial distributions, J. Gen. Microbiol. **58**, 141 (1969).

[43] N. Bacaër, *A Short History of Mathematical Population Dynamics* (Springer London, London, 2011).

[44] J. L. Lebowitz and S. I. Rubinow, A theory for the age and generation time distribution of a microbial population, J. Math. Biol. **1**, 17 (1974).

[45] J. Jędrak, M. Rubin, and A. Ochab-Marcinek, Generalization of Powell's results to population out of steady state, Phys. Rev. E **108**, 024405 (2023).

[46] M. Campos, I. Surovtsev, S. Kato, A. Paintdakhi, B. Beltran, S. Ebmeier, and C. Jacobs-Wagner, A constant size extension drives bacterial cell size homeostasis, Cell **159**, 1433 (2014).

[47] L. Willis and K. C. Huang, Sizing up the bacterial cell cycle, Nat. Rev. Microbiol. **15**, 606 (2017).

[48] I. Soifer, L. Robert, and A. Amir, Single-cell analysis of growth in budding yeast and bacteria reveals a common size regulation strategy, Curr. Biol. **26**, 356 (2016).

[49] Y.-J. Eun, P.-Y. Ho, M. Kim, S. LaRussa, L. Robert, L. D. Renner, A. Schmid, E. Garner, and A. Amir, Archaeal cells share common size control with bacteria despite noisier growth and division, Nat. Microbiol. **3**, 148 (2017).

[50] 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).

[51] S. Hobson-Gutierrez and E. Kussell, Evolutionary advantage of cell size control, arXiv:2303.08042v1 (2023).

[52] M. Priestman, P. Thomas, B. D. Robertson, and V. Shahrezaei, Mycobacteria modify their cell size control under sub-optimal carbon sources, Front. Cell Dev. Biol. **5**, 64 (2017).

[53] G. Lambert and E. Kussell, Quantifying Selective Pressures Driving Bacterial Evolution Using Lineage Analysis, Phys. Rev. X **5**, 011016 (2015).

[54] G. Woronoff, P. Nghe, J. Baudry, L. Boitard, E. Braun, A. D. Griffiths, and J. Bibette, Metabolic cost of rapid adaptation of single yeast cells, Proc. Natl. Acad. Sci. USA **117**, 10660 (2020).

[55] S. J. Schink, E. Biselli, C. Ammar, and U. Gerland, Death rate of *E. coli* during starvation is set by maintenance cost and biomass recycling, Cell Syst. **9**, 64 (2019).

[56] T. Stadler, Recovering speciation and extinction dynamics based on phylogenies, J. Evol. Biol. **26**, 1203 (2013).

[57] S. Nakashima, Y. Sughiyama, and T. J. Kobayashi, Lineage EM algorithm for inferring latent states from cellular lineage trees, Bioinformatics **36**, 2829 (2020).

[58] A. Dechant and S.-i. Sasa, Fluctuation–response inequality out of equilibrium, Proc. Natl. Acad. Sci. USA **117**, 6430 (2020).

[59] A. Amir, Cell Size Regulation in Bacteria, Phys. Rev. Lett. **112**, 208102 (2014).