Exponential trajectories, cell size fluctuations, and the adder property in bacteria follow from simple chemical dynamics and division control

Parth Pratim Pandey,[1,*,†] Harshant Singh,[1,*] and Sanjay Jain [1,2,‡]

[1]*Department of Physics and Astrophysics, University of Delhi, Delhi 110007, India*
[2]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

Experiments on steady-state bacterial cultures have uncovered several quantitative regularities at the system level. These include, first, the exponential growth of cell size with time and the balanced growth of intracellular chemicals between cell birth and division, which are puzzling given the nonlinear and decentralized chemical dynamics in the cell. We model a cell as a set of chemical populations undergoing nonlinear mass action kinetics in a container whose volume is a linear function of the chemical populations. This turns out to be a special class of dynamical systems that generically has attractors in which all populations grow exponentially with time at the same rate. This explains exponential balanced growth of bacterial cells without invoking any regulatory mechanisms and suggests that this could be a robust property of protocells as well. Second, we consider the hypothesis that cells commit themselves to division when a certain internal chemical population reaches a threshold of N molecules. We show that this hypothesis leads to a simple explanation of some of the variability observed across cells in a bacterial culture. In particular, it reproduces the adder property of cell size fluctuations observed recently in *E. coli*; the observed correlations among interdivision time, birth volume, and added volume in a generation; and the observed scale of the fluctuations (CV $\approx$ 10–30%) when N is between 10 and 100. Third, upon including a suitable regulatory mechanism that optimizes the growth rate of the cell, the model reproduces the observed bacterial growth laws including the dependence of the growth rate and ribosomal protein fraction on the medium. Thus, the models provide a framework for unifying diverse aspects of bacterial growth physiology under one roof. They also suggest new questions for experimental and theoretical enquiry.

I. INTRODUCTION

The simplest cells, bacteria, exhibit several generic phenomena that were discovered decades ago but still require explanation. One such phenomenon is that in steady-state bacterial cultures the size of an individual bacterial cell grows exponentially with time between birth and division. This was observed early on [1,2] and its existence in various organisms has been the subject of debate [3,4]. Recently, detailed single-cell experiments [5–11] have confirmed this property for many bacterial species in steady-state cultures. Intracellular molecular populations have also been observed to grow exponentially in *Escherichia coli* cells within a generation [12]. Since bacterial cells divide, the range of size over which exponential trajectories are seen is limited to a factor of 2 (or similar), thereby potentially permitting alternate fits to the data. However, the exponential function fits quite well and is therefore at least a very good approximation. In other single-celled organisms such as *Schyzosaccharomyces pombe* a clear departure from exponential trajectories is seen [4,13], so this is not a property that can be taken for granted. Why

most bacteria exhibit this property remains an unanswered question.

In linear autocatalytic systems [14–16], exponential growth is not surprising. The asymptotic rate of exponential growth of chemical populations in such systems is the largest eigenvalue of the matrix defining the dynamical system. However, the chemical dynamics of the intracellular molecular species in a bacterial cell is highly nonlinear and therefore exponential trajectories are surprising. In this paper, we provide an explanation for how cells can exhibit exponential trajectories of intracellular chemicals and cell volume in spite of the nonlinearity of the dynamics. We show that ordinary differential equations representing mass-action-based nonlinear chemical dynamics in an expanding container whose volume is linearly dependent on the constituent populations have a special scaling property, or "quasilinearity." Such nonlinear systems naturally have exponentially growing trajectories as attractors.

Another unexplained phenomenon is "balanced growth" [17], the remarkable coordination between thousands of intracellular chemicals so that all of them (on average) double in the same time. This is required for self-replication: A cell at birth must grow in such a way that division, when it happens, produces two daughters identical to itself; hence the mother at division must have twice of everything as the daughter at birth. In spite of their decentralized dynamics characterized by reaction-specific catalysts and rate constants, how do thousands of chemicals conspire to double at the same time? It

_____

*These authors contributed equally to this work.
†Present address: Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.
‡Corresponding author: jain@physics.du.ac.in

is normally supposed that regulatory mechanisms involving checkpoints and feedback loops are responsible for this coordination. We argue that this sophistication is not needed. The exponential trajectories that we find as generic attractors in autocatalytic systems have all chemical populations increasing at the *same* exponential rate even when no regulation is present. Thus, their ratios are automatically constant in time once they reach the attractor. The genericity and robustness of these attractors suggests that balanced growth could have been a property of protocells at the origin of life. The protocell literature has been concerned with this question [18–23] and has been an inspiration for this work, though the solution presented in this paper is different from those advocated there. At a mathematical level, we show that the growth rate as well as the ratios of chemicals in the attractor are determined from a nonlinear generalization of the eigenvalue equation of a matrix.

Another phenomenon that has consistently received attention but is still not understood is the origin and scale of variation in interdivision time and cell size, reviewed in Refs. [24–27]. Genetically identical bacterial cells subjected to the same steady environment exhibit a phenotypic variation in both these quantities. In the steady state, the volume of the cells at birth and the time between birth and division have steady distributions with a constant mean and standard deviation (s.d.) and a coefficient of variation, CV = s.d./mean, in the range of 10–30%. The mean size of the cell at birth can be changed by close to a factor of 10 by choosing different environmental conditions; its CV remains between 10% and 30% [9,11,28]. The CV of the molecular population of an abundant intracellular chemical across cells in a bacterial culture (the extrinsic CV) is also 10–30% [29–31]. It is believed that the source of variation lies in the process that controls cell division; however, neither the molecular basis of this process, even for *E. coli*, nor the origin of the scale of CV is fully understood. In this paper, we explore, through a mechanistic mathematical model, the hypothesis [32] that division is controlled by the intracellular population (not concentration) of a certain molecule reaching a fixed threshold $\sim N$, a parameter of the model. Then N controls both the average size and the fluctuations in size and interdivision time. We provide an explicit expression for the average cell volume V in terms of N and the other molecular parameters of the model. Fluctuations arise because molecular production is stochastic and hence the interdivision time, which under this hypothesis is the first passage time for this molecule to reach its threshold, is stochastic. The CV of the interdivision time is $\sim 1/\sqrt{N}$; hence, the CV of cell volume and of other molecular populations is also $\sim 1/\sqrt{N}$. We note that this hypothesis therefore explains the observed range 0.1–0.3 of the CV of all these quantities provided N lies between 10 and 100. We compare the theoretical distributions of size, interdivision time, and added volume with the experimental data of Ref. [11] to constrain the range of N for *E. coli*.

Size fluctuations in *E. coli* have been recently observed [10,11] to satisfy the "adder" property [33,34], wherein the volume added by a cell in each generation is independent of its birth volume in that generation. This property has been explored in various phenomenological models [8,10,11,33–35] and implies specific correlations between the cell volume

at birth and division and between interdivision time and birth volume. It has also been shown to arise in mechanistic models, a simple linear one [36] as well as a biologically more realistic model [37]. We show that the adder property appears robustly for a wide class of nonlinear mechanistic models under the hypothesis that division is triggered when a molecular population reaches a threshold value, generalizing the results of Ref. [36]. We indicate circumstances where fluctuations depart from the adder.

Fluctuations of the growth rate across cells in a steady-state culture have been measured in the literature and CV values ranging from 10% to 40% have been reported [11,12,38,39]. Our models suggest that the physical origin of fluctuations in the growth rate is different from that of interdivision time and cell size. The simplest and most natural models that we consider exhibit a smaller CV (1–5%) than that reported in the experimental literature, though larger values can also be accounted for. Our models also reproduce the observed crossover in the CV of an intracellular molecular population as a function of the mean population $\bar{X}$ (CV $\sim \bar{X}^{-1}$ for small $\bar{X}$ and constant for large $\bar{X}$) [30].

Growth laws of bacterial composition [40–43] constitute another class of generic phenomena in bacteria which describe how the cellular steady-state growth rate and ribosomal protein fraction in the cell depend upon the medium. These bacterial growth laws have received a substantial theoretical attention [43–50] and are much better understood compared to the phenomena mentioned in the previous paragraphs. A particular model that we discuss in detail reproduces the bacterial growth laws in addition to the other phenomena mentioned above. Thus, this work provides possible explanations and a unified understanding of a number of generic properties of bacteria, including exponential trajectories in the steady state, balanced growth of cells, the growth laws of bacterial composition, as well as the scale and correlations of the fluctuations of several cellular variables. The work makes predictions that can be experimentally tested.

A. Organization of the paper

In Sec. II, we describe the general class of mathematical models of cells that we consider. In Sec. III, a specific nonlinear model is considered and results based on numerical simulations are presented that reproduce the features mentioned above. In addition to intrinsic stochasticity in chemical dynamics, we also discuss here the effect of stochasticity in the partitioning of chemicals at division and stochasticity in the threshold value of the division trigger. In Sec. IV, we derive analytically several of the results presented in Sec. III. In particular, we provide an analytic understanding of the robustness (or universality) of the distributions of cell size, intracellular populations and interdivision time, and the nonuniversality of the distribution of the cellular growth rate. Section V discusses the mathematical and physical basis for exponential trajectories and balanced growth and shows that these are generic properties of mass action chemical dynamics in expanding containers whose volume is a linear function of the internal chemical populations. This section also extends the analytical results of Sec. IV to a large class of models. Section VI discusses the unification of the bacterial growth

laws with the other cellular properties. Finally, Sec. VII gives a detailed summary of the model assumptions and results and discusses future experimental and theoretical directions. The Supplemental Material [51] contains additional figures and proofs. Some readers may benefit by turning directly to the summary in Sec. VII for a quick guide to the main results including equations and figures.

II. A GENERAL MODEL OF A GROWING-DIVIDING CELL

We model the cell as a container with $n+1$ chemical species, whose molecular populations at time t are denoted $X_i(t)$, $i = 1, 2, \ldots, n$, and $Z(t)$. Z is the species triggering cell division and $\mathbf{X} = (X_1, \ldots, X_n)$ represents all the other chemicals in the cell. The population dynamics is given by

$$dX_i/dt \equiv \dot{X}_i = f_i(\mathbf{X}), \quad i = 1, 2, \ldots, n, \tag{1a}$$

$$dZ/dt \equiv \dot{Z} = h(\mathbf{X}, Z), \tag{1b}$$

where the functions f_i and h encapsulate the kinetics of the chemical reactions in the cell. The Z population is assumed to have a negligible effect on the $\mathbf{X}$ dynamics; hence, the f_i are independent of Z. The volume of the cell is assumed to depend linearly on the X_i:

$$V(\mathbf{X}) = \sum_{i=1}^{n} v_i X_i, \tag{2}$$

where the v_i are constant parameters of the model. V depends on time through the X_i. A particular case of this assumption, usually made in most models, is that V is proportional to the mass of protein in the cell. Again, the contribution of Z to V is assumed to be negligible. Explicit examples of the functions f_i and h will be discussed later.

In this work, we will primarily consider the case where the functions f_i are *homogeneous degree-1* functions of the molecular populations, i.e., for any $\beta > 0$,

$$f_i(\beta\mathbf{X}) = \beta f_i(\mathbf{X}), \quad i = 1, 2, \ldots, n. \tag{3}$$

We will refer to n-dimensional dynamical systems defined by (1a) which satisfy (3) as *class-I* dynamical systems. This class includes linear systems as well as a large class of nonlinear dynamical systems. In fact, we show later that *any* nonlinear chemical dynamics based on mass-action kinetics becomes class-I when considered in a container whose volume is given by (2). Such systems, especially if the underlying chemical network is autocatalytic, have attractors that show exponential growth of all the chemicals and the volume with time, as will be discussed later. We will see that departures from the class-I property result in departures from exponentially growing trajectories, examples of which are the models in Ref. [37] that display both exponential and nonexponential growth.

A. Division control

Starting from some initial condition for the $n+1$ populations, the populations are evolved according to (1). The function h is assumed to be positive, and hence Z increases under (1). When Z reaches the threshold Z_c, the cell divides. The two daughter cells are assumed to be identical, and hence each gets half of every chemical in the cell. Since in the

model we track one of the two daughters, at division all $n+1$ populations are replaced by half their values. This defines the state of the daughter at birth in the next generation and with that initial condition for the populations the above procedure is iterated.

B. Resetting of Z after triggering; delay between triggering and division

The above defines the simplest version of the model. It is useful, however, to consider two generalizations motivated by empirical observations. The first involves a resetting of the Z population after it triggers division. In the above scheme, Z_b, the value of Z in a daughter at birth, is always $Z_c/2$. However, specific biochemical processes may cause a degradation of the triggering molecule on a short timescale after triggering of division. In order to model this, we assume for simplicity that when Z reaches Z_c it triggers division and is also instantaneously reset to a value $Z_r \leqslant Z_c$. Z_r is another parameter of the model; $Z_r = 0$ implies complete degradation of the triggering molecule; $Z_r = Z_c$ implies no degradation. Several properties that we discuss (e.g., the adder property) hold for the entire range of values $0 \leqslant Z_r \leqslant Z_c$. The second generalization is that division follows triggering (and resetting of Z, if any) after a time delay τ_1, which is another parameter of the model. When $\tau_1 = 0$, then division immediately follows triggering and resetting; hence, $Z_b = Z_r/2$. When $\tau_1 > 0$, we assume that after triggering (and instantaneous resetting of Z to Z_r) all $n+1$ populations continue to evolve via (1) for a fixed time τ_1, whereupon cell division takes place and all are halved. These generalizations are motivated by evidence that the replication of DNA is initiated when a certain protein reaches a threshold population, and this chemical degrades soon after the initiation of replication to avoid multiple replication rounds. Cell division follows the initiation of DNA replication after the lapse of a certain time.

To summarize, the population dynamics in the model is a combination of continuous dynamics described by (1) and discrete events. The $\mathbf{X}$ variables evolve via (1a) from their initial values up to the time of cell division; only at division do they experience a discrete change: They are halved. The Z variable evolves via (1b) from its initial value up to the point where it reaches the threshold Z_c. There it is instantaneously reset to Z_r. Subsequently, it again evolves via (1b) for a fixed time τ_1. That defines the time of cell division. At that time, Z is halved (along with the X_i). That brings the cell to the beginning of the next generation and a new set of initial values of the $n+1$ populations, whereupon the procedure is iterated.

Note that the dynamics of Z explicitly depends upon the X sector through $h(\mathbf{X}, Z)$. On the other hand, Z does not affect the short timescale dynamics of the X sector since f_i are assumed to be independent of Z. However, Z exercises a discrete control over the X variables through the triggering of division, which in effect determines how large the X_i can grow before cell division occurs and they halve. The functions f_i contain all the interactions among the X chemicals and encapsulate most of the complexity of the intracellular dynamics.

The above formulation describes the deterministic version of the model. We also consider various sources of stochasticity. One is the intrinsic stochasticity in the chemical dynamics

resulting from the populations being non-negative integers instead of continuous variables and each reaction having a certain probability of occurring. Then Eq. (1) is replaced by its stochastic counterpart. Further, at division the chemicals may not partition equally into both daughters. Also Z_c, the threshold value of of Z at which division is triggered, may vary from generation to generation. The specific implementation of these sources of stochasticity in the model will be discussed later.

C. Key differences with respect to earlier work

(i) The origin of exponential trajectories and balanced growth of bacterial cells has remained a puzzle; we explain the origin as a consequence of the ubiquitous class-I property of the dynamics. For class-I systems, we introduce a useful geometric construct in phase space, the "curve of balanced growth," which is often the attractor for the deterministic dynamics and a useful reference point for stochastic perturbations. (ii) We extend the analytical proof of applicability of the adder property to a much larger class of models than before. The division control dynamics above (based on a time-keeper molecule Z reaching a population threshold Z_c) has been considered earlier in Ref. [36] in a simple model where Z is the only variable (the X sector is absent) and in Ref. [37] in a more realistic model where the X sector contains various proteins, messenger ribonucleic acid (mRNA) molecules, and genes (including ribosomes and RNA polymerase). They find that the adder property arises from this division control mechanism when the volume grows exponentially and the Z dynamics is treated stochastically. They have proved this analytically for the simple model in Ref. [36] and established the result numerically in the more complicated model of Ref. [37] in the exponential phase. We analytically prove the adder property for all models (with some caveats discussed later) defined at the deterministic level by Eqs. (1)–(3) and the property $h(\beta \mathbf{X}, Z) = \beta h(\mathbf{X}, Z)$, when the Z dynamics is treated stochastically. The model in Ref. [36] and the basic version of the model in Ref. [37] are special cases of this class. Note that our result implies that the adder property holds irrespective of the detailed form of functions f_i and h and the number of variables therein (with some caveats), as long as they satisfy the homogeneous degree-1 property. This helps in understanding the extent of universality of the adder property with respect to the dynamics. We also show that the property does not depend on the constant reset value Z_r. Furthermore, when the function h is independent of Z, we show analytically for this entire class of models that the distribution of the added volume is the same as the one obtained in Ref. [36] with a CV of $1/\sqrt{N}$, where $N \equiv Z_c - (Z_r/2)$. By comparing with the experimental distributions of Ref. [11] for *E. coli*, we find N to be between 20 and 60. We also discuss some new mechanisms for the departure from the adder property. (iii) We explain the behavior of the protein number fluctuations (the dependence of the CV on the mean population of the protein) observed in Ref. [30] for the same large class of models. Essentially, the so-called extrinsic noise is nothing but the fluctuation of the first passage time of the time-keeper molecule Z, which affects the CV of all protein populations equally and sets the scale to be $\sim N$ for the mean protein number at which the observed crossover occurs. (iv) We provide a possible

geometric interpretation of the origin of fluctuations in the growth rate, namely, perturbations orthogonal to the curve of balanced growth in phase space.

III. A SPECIFIC EXAMPLE WITH $3 + 1$ CHEMICAL SPECIES

We now consider a concrete example in which the X sector is described by the precursor-transporter-ribosome (PTR) model discussed in Ref. [50] containing $n = 3$ species. Later on, we will argue that several of the results of the PTR model hold for the general model described above with a very broad class of functions f_i and h. Here $X_1 = P$ is a coarse-grained population variable representing the total number of amino acid molecules (precursors or monomers out of which proteins are constructed), $X_2 = T$ is the total number of transporter and other metabolic enzyme molecules in the cell responsible for making P from the food molecules available in the extracellular medium, and $X_3 = R$ is the total number of ribosomes in the cell (catalysts for making T and R from the P molecules). The equations of the X sector (or PTR sector) are

$$\frac{dP}{dt} = f_1(P, T, R) = K_P T - k \frac{RP}{V}, \tag{4a}$$

$$\frac{dT}{dt} = f_2(P, T, R) = K_T \frac{RP}{V} - d_T T, \tag{4b}$$

$$\frac{dR}{dt} = f_3(P, T, R) = K_R \frac{RP}{V} - d_R R, \tag{4c}$$

where

$$V = v_P P + v_T T + v_R R. \tag{5}$$

The constants $K_P, K_T, K_R, k, d_T, d_R, v_P, v_T,$ and v_R are parameters of the model. The rate constant K_P represents the efficiency of metabolic enzymes T in transporting and producing P from external food. It is an increasing function of the external food concentration $[F]$ and the quality of the food source q (the number of P molecules produced per food molecule transported). k represents ribosomal catalytic efficiency and is the peptide elongation rate per unit concentration of P. d_T and d_R are degradation rates. $v_P, v_T,$ and v_R define the contributions of the individual species to the cellular volume V. It is convenient to parametrize the rate constants K_T and K_R as follows:

$$K_T = \frac{f_T k}{m_T}, \quad K_R = \frac{f_R k}{m_R}, \quad f_T + f_R = 1. \tag{6}$$

Here f_T and f_R represent the fractions of ribosomes engaged in making T and R, respectively, and m_T and m_R are the number of amino acid residues (or units of P) in a T molecule and ribosome respectively. It follows that the mass of the PTR cell (in units of mass of a P molecule) is $M = P + m_T T + m_R R$. For more details, see Ref. [50]. Note that by virtue of (5), the PTR sector is a class-I dynamical system, i.e., $f_i(\beta P, \beta T, \beta R) = \beta f_i(P, T, R)$ for $i = 1, 2, 3$.

In the Z sector, the function h is defined by

$$\frac{dZ}{dt} = h(P, T, R) = K_Z \frac{RP}{V}, \tag{7}$$

where K_Z is a constant. This form assumes that the production of the molecule that triggers cell division mirrors the growth
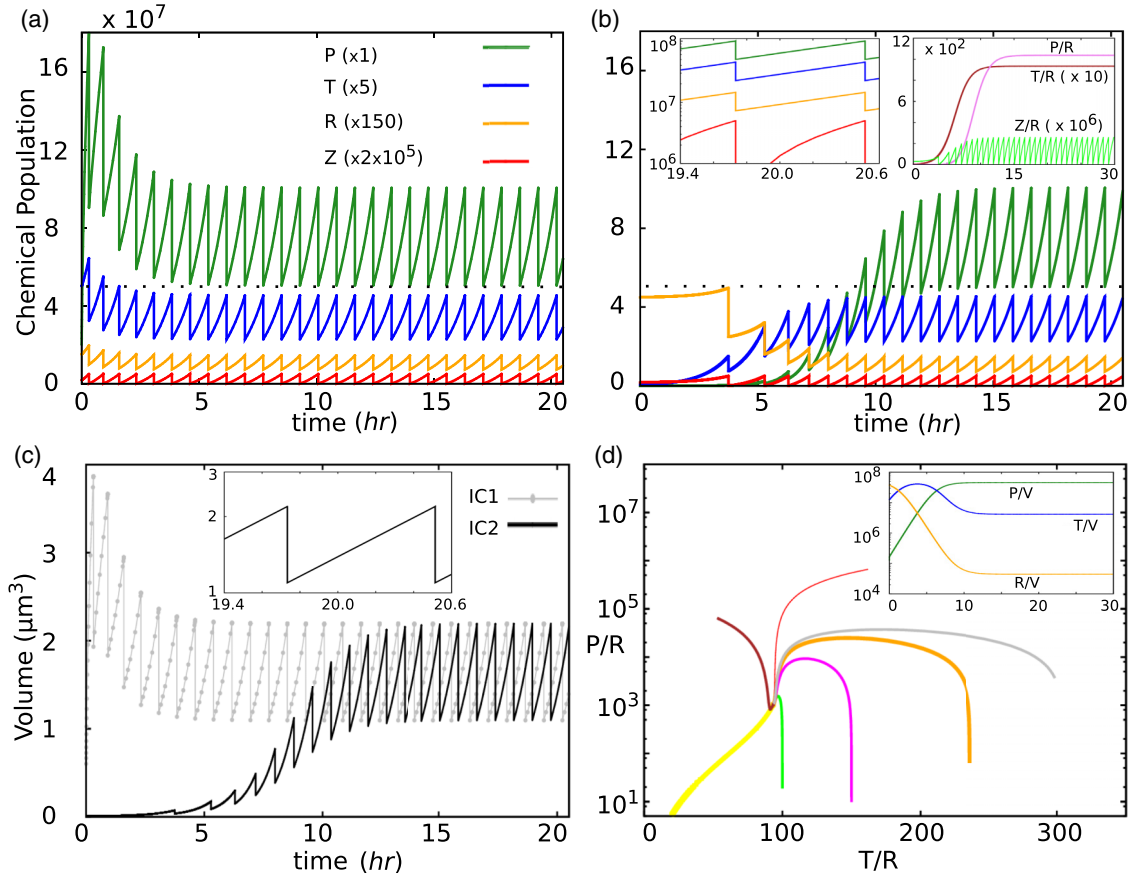
FIG. 1. Plot of the chemical populations P, T, R, Z and volume V vs time t for deterministic trajectories. Parameters of the model are $K_P = 500$ h^{-1}, $k = 10^{-3}$ h^{-1} $(\mu m)^3$, $d_T = 0.1$ h^{-1}, $d_R = 0$, $m_T = 400$, $m_R = 10000$, $v_P = v_T = v_R = 2 \times 10^{-8}$ $(\mu m)^3$, $f_R = 0.19377$, $K_Z = 10^{-11}$ h^{-1} $(\mu m)^3$, $Z_c = 25$, $Z_r = 0$, $\tau_1 = 0$. (a) The trajectory of a cell with the initial conditions IC1: $P(0) = 2 \times 10^7$, $T(0) = 10^7$, $R(0) = 10^5$, $Z(0) = 9$, and (b) IC2: $P(0) = 10^3$, $T(0) = 10^5$, $R(0) = 3 \times 10^5$, $Z(0) = 9$. The population values in panels (a) and (b) have been multiplied by the factors mentioned in the legend of panel (a) to bring them on the same figure. The dynamics converges to a steady state (a periodic solution or limit cycle attractor). The chemical populations in the daughter cell at birth in the steady state are $P_b = 5.023 \times 10^7$, $T_b = 4.529 \times 10^6$, $R_b = 4.844 \times 10^4$ [P_b is indicated by black dotted lines in panels (a) and (b)]. The period, or interdivision time $\tau = 0.781$ h. (c) The volume of the cell, V, defined by Eq. (5), as a function of time for the two initial conditions, IC1 and IC2. In this and subsequent figures, the unit of the time axis is hr and the volume axis is $(\mu m)^3$. First inset of panel (b) and inset of panel (c): Semilog plots of P, T, R, Z, and V vs time in the steady-state with IC2. The plots of P, T, R, and V are piecewise linear and have the same slope 0.888 h^{-1} (slope of the natural logarithm of the quantity vs time), showing that they all increase exponentially with the same rate in the growth phase of the limit cycle attractor. The slope agrees with formula (10c). Second inset of panel (b) plots the ratios of chemical populations as a function of t on the trajectory starting from IC2. The ratios of populations of the PTR sector become constant on the limit cycle attractor and the constant values agree with Eqs. (10a) and (10b). (d) Simulated phase portrait of trajectories from diverse initial conditions projected on to the two-dimensional space of T/R and P/R showing that the trajectories converge to the same attractor independent of initial conditions. Inset of panel (d): Time dependence of concentrations of P, T, R on the trajectory starting from IC2, becoming constant on the limit cycle attractor.

of the cell [this term has the same form as the growth term of T and R in (4)]. Note that in (7) $h(\mathbf{X}, Z)$ is independent of Z and also satisfies $h(\beta \mathbf{X}) = \beta h(\mathbf{X})$. We will later (in Sec. IV D) give results when h has a nontrivial Z dependence arising from autoregulation of Z (e.g., cooperativity). The PTR model [50] has been independently extended by Ref. [52] to include a sector whose dynamics is given by (7) and that triggers division upon reaching a population threshold. Their treatment differs from ours in various respects, in particular, that the new sector affects the local dynamics of the PTR sector through the volume. They have also observed the adder property of the cell volume in numerical simulations. Apart from that commonality, Ref. [52] and the present work explore different aspects of the models.

A. Deterministic version; numerical results

The rules mentioned above fully specify the dynamical system once the parameters of the PTR sector ($K_P, k, d_T, d_R, m_T, m_R, f_R, v_P, v_T, v_R$), the Z sector (K_Z, Z_c, Z_r), and τ_1 are specified. We refer to this as the PTRZ model. Its deterministic dynamics is discussed in this subsection and stochastic dynamics in the next. For simplicity, we consider the case where $\tau_1 = 0$ ($\tau_1 > 0$ is discussed later). Thus, when Z reaches Z_c, the cell divides without any delay; i.e., the populations P, T, R are halved, and the birth value of Z is simultaneously set to $Z_r/2$. Subsequently, the populations follow Eqs. (4) and (7) until Z again reaches Z_c and the procedure is repeated. Simulations are shown in Fig. 1. Most parameters of the PTR sector are chosen in the ballpark of realistic
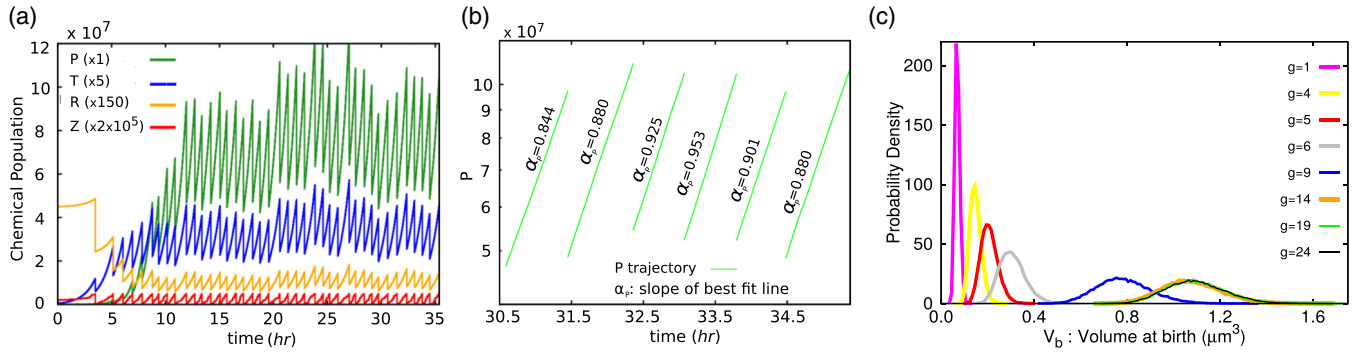
FIG. 2. Effect of stochasticity in chemical dynamics of the $PTRZ$ cell. Parameter values are as in Fig. 1. (a) A stochastic trajectory of P, T, R, Z starting from IC2 (see legend of Fig. 1). (b) P vs time t in a semilog plot for five consecutive generations in the statistical steady state. The slope α_P of the best fit straight line ($\ln P$ vs t) varies from generation to generation. (c) Time evolution of the volume distribution of 30 000 cells at birth. The distribution at the beginning of the gth generation (or just after g divisions) is shown for various values of g. In this case the initial distribution of V is a δ function at $V = 0.06$ $(\mu\text{m})^3$, since all the trajectories start at $t = 0$ from the initial condition IC3: $P(0) = T(0) = R(0) = 10^6, Z = 0$. The distributions at $g = 19$ and 24 are indistinguishable, suggesting a stationary state.

values [42,43,53,54] and reproduce known experimental data on *E. coli* within a factor of order unity (namely, the values of P, T, R and the growth rate). The values of ν_T and ν_R have been set equal to ν_P for simplicity; we are not aware of experiments that measure the ν_i independently. The value of f_R has been chosen so as to maximize the steady-state growth rate of the PTR cell with the above-mentioned choice of parameters; this procedure reproduces the bacterial growth laws as discussed in Ref. [50]. In the Z sector, the value of Z_c and Z_r are chosen to get the scale of size fluctuations in the experimentally observed ballpark (it will be seen that only the combination $Z_c - Z_r/2$ matters). K_Z, an unknown parameter of the model is chosen so that the steady-state birth volume of the cell turns out in the correct ballpark, ≈ 1 $(\mu\text{m})^3$. Our conclusions are robust to independent variation of each parameter (within collective limits as discussed later).

As evident from Fig. 1, after a transient that depends on initial conditions, the dynamics converges to a steady state. This steady state is a limit cycle attractor in which the just born daughter cell grows to twice its size (the populations P, T, R and the cell volume double) and then divides into two halves, bringing the system back to the same daughter state. It is seen that in the attractor, between birth and the next division (i.e., within a single generation of the cell) the chemicals of the X sector grow exponentially with time at the same rate [inset of Fig. 1(b)]. The same is true of V [inset of Fig. 1(c)], a consequence of (5). Because of exponential growth, the ratios T/R and P/R remain constant in time throughout the attractor. Hence, the concentrations of the X-sector chemicals also remain constant in the attractor [inset of Fig. 1(d)]. During the transient period before the attractor is reached, the ratios are not constant. For all the initial conditions considered (with $0 \leqslant Z(0) < Z_c$), the trajectory converges to the same steady state [Fig. 1(d)], suggesting that the limit cycle is a stable attractor with a wide basin of attraction. When the parameters are varied within a wide range (specified later; Sec. IV A), the same dynamical behavior is observed: initial transient leading to a limit cycle attractor characterized by exponential trajectories in the growth phase and constant ratios of chemicals.

B. PTRZ with stochastic chemical dynamics; numerical results

In Fig. 2(a), we show a typical trajectory of the PTRZ cell following stochastic chemical dynamics. The difference from the deterministic version is that the evolution of P, T, R, and Z is no longer given by the differential equations (4) and (7) but by a stochastic version of (4) and (7). The populations are now non-negative integers and are updated according to the probabilities of the chemical reactions, the latter being proportional to the corresponding terms in the differential equations. The rest of the dynamical rules are the same as before [starting from any given initial condition, the populations P, T, R, and Z are evolved with time until $Z = Z_c$, at which point P, T, and R are halved (and rounded off to the nearest integer value); Z is set to $Z_r/2$ and the procedure is iterated]. No new physical parameter is introduced in the stochastic dynamics. The algorithm used for numerical simulations is mentioned in Computational Methods below.

1. Stochastic steady state

We observe in Fig. 2(a) that after a few rounds of growth and division the trajectory of the cell reaches a statistical steady state, a stochastic version of the deterministic steady state seen in Figs. 1(a) and 1(b). In the stochastic steady state, the daughter cells produced at the beginning of each cycle are no longer identical from generation to generation (in terms of the values of P, T, R, and V_b), as was the case in the deterministic simulation, and τ, the interdivision time, also varies from cycle to cycle. The latter variation arises because the first passage time of Z to reach Z_c varies from cycle to cycle. Figure 2(b) shows that in the statistical steady state the population growth of the X-sector chemicals within a generation can be approximated by an exponentially growing trajectory with the effective growth rate varying from generation to generation.

We performed 30 000 stochastic simulations for the PTRZ model, each with the same initial condition as described in the caption of Fig. 2(c). The cellular variables were tracked for each trajectory. The distribution of the cell volume at birth, V_b, across the 30 000 trajectories at different generations is shown
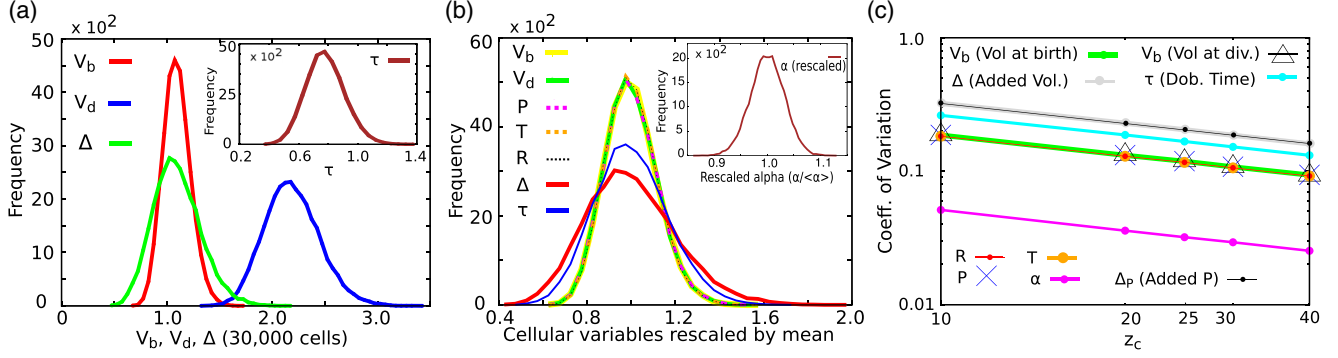
FIG. 3. Distribution of volume, interdivision time, growth rate, and chemical populations in the PTRZ model. This figure shows the frequency distributions of various cellular variables, measured across 30 000 cells at the steady state. Data were compiled from each trajectory in the 30th generation to ensure that it had reached the statistical steady state. (a) Distributions of volume at birth V_b, volume at division V_d, added volume $\Delta \equiv V_d - V_b$, and (in inset) interdivision time τ. (b) Frequency distributions of various quantities rescaled by their respective means. The x axis for the curve denoted by V_b in the legend stands for $V_b/\langle V_b \rangle$, and similarly for the other curves. The distributions of rescaled P_b, T_b, R_b, V_b, and V_d are essentially indistinguishable. The distribution of rescaled α is shown in the inset of panel (b). (c) CV of $\Delta, \tau, V_b, V_d, \alpha, P_b, T_b, R_b$, and $\Delta_P \equiv P_d - P_b$ as functions of Z_c on a log-log plot. The CV of Δ and Δ_P are indistinguishable from each other (and shown in the topmost line), as are those of V_b, V_d, P_b, T_b, and R_b (third line from top). The slope of the linear fit is consistent with CV $\propto N^{-1/2}$ for all the quantities. The vertical separation between the lines gives the ratios of the CV of various quantities and they are consistent with what is expected when the adder property holds, Eq. (8). Parameter values are as in Fig. 1 except in panel (c), where Z_c takes a range of values.

in Fig. 2(c). The distribution becomes stationary after a few generations. The same steady-state distribution is obtained by (a) starting from a different initial distribution of the 30 000 cells and (b) sampling over different generations of a single trajectory after the initial transient period. The distributions of V_d, P_b, T_b, R_b, τ, and α are also found to converge to their respective steady-state distributions independent of the initial distribution of cellular configurations (α is the "growth rate" of the cell in a generation, defined as the slope of the best straight line fit of $\ln V(t)$ versus t in that generation). This indicates that the statistical steady state is a stable attractor of the dynamics characterized by fixed distributions of the cellular variables.

2. Distributions of cell size, interdivision time and growth rate, correlations, and the adder property

The steady-state distributions are shown in Fig. 3. The averages of the steady-state distributions are $\langle P \rangle = 5.012 \times 10^7$, $\langle T \rangle = 4.518 \times 10^6$, $\langle R \rangle = 4.833 \times 10^4$, $\langle V_b \rangle = 1.094 \ \mu\text{m}^3$, $\langle V_d \rangle = 2.190 \ \mu\text{m}^3$, $\langle \Delta \rangle = 1.096 \ \mu\text{m}^3$, $\langle \tau \rangle = 0.782$ h, $\langle \alpha \rangle = 0.888 \ \text{h}^{-1}$. These are close to the deterministic values in Fig. 1. The model predictions of the distributions of the above quantities rescaled by their means (except the rescaled α distribution) are largely independent of all the PTRZ model parameters (and hence independent of growth rate, ratios of chemicals, etc.) and of K_Z when $\tau_1 = 0$. They only depend upon the quantity $N = Z_c - (1/2)Z_r$. (See Fig. S1 [51] for a completely different parameter set giving the same rescaled distributions.) The reason for this extraordinary robustness will be discussed later. We find that the CV of the Δ, V_b, V_d, and τ distributions are consistent with

$$\sqrt{3}\text{CV}(V_d) = \sqrt{3}\text{CV}(V_b) = \sqrt{3}\ln 2\text{CV}(\tau)$$

$$= \text{CV}(\Delta) = \frac{1}{\sqrt{Z_c}}. \tag{8}$$

The first three relations are expected to hold if the cell volume grows exponentially with a small variation in the growth rate and the system satisfies the adder property [11]. The last equality relates the fluctuations to the parameters Z_c and Z_r of the present model. This connects the CV of a macroscopic quantity Δ to a potentially microscopic quantity N, the change in the population of the Z molecule from birth to division. The CV of P_b, P_d, and $\Delta_P \equiv P_d - P_b$ also satisfy (8), as do the corresponding quantities for T and R.

Using the simulation data of the 30 000 stochastic cells, we determine the correlations between the cell size at birth (V_b) and division (V_d), the added volume (Δ), and the interdivision time (τ) in the steady state (Fig. 4). The correlations show that the model exhibits the adder property, because the mean added volume Δ between birth and division is independent of the birth size. These are the kinds of correlations observed in experiments with *E. coli* [10,11]. Taheri-Araghi *et al.* [11] have also measured the distribution of Δ for a fixed V_b at different values of V_b (binned) for *E. coli*. They find that not just the mean of Δ but the entire distribution to be independent of V_b, which is a strong version of the adder property. For the PTRZ model in Fig. S2, we exhibit this "conditional" distribution of Δ for different fixed values of V_b (in a bin). The distributions collapse onto each other, showing that the model exhibits this property.

Adder property is independent of Z_r. We remark that in the present model the adder property of cell volume is observed for the entire range of the reset parameter $Z_r, 0 \leqslant Z_r \leqslant Z_c$. This is also shown analytically in Sec. IV D.

3. Intracellular chemical populations: Distribution, adder property, and dependence of CV on the mean population

Each population P_b, T_b, R_b at birth has its own mean and CV, but the populations rescaled by their respective means
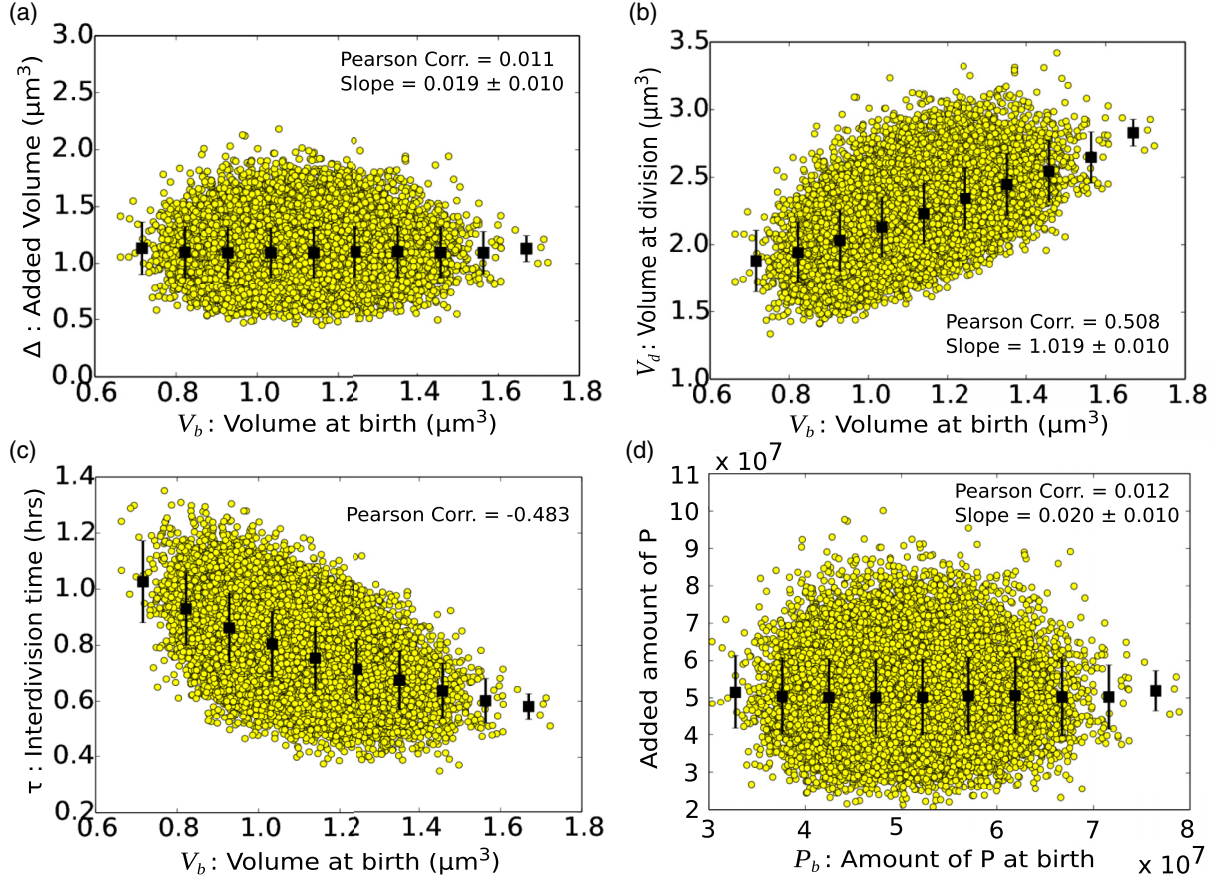
FIG. 4. Dynamics of the PTRZ cell model displays the adder property for cell size and intracellular chemical populations. Each data point in the scatter plots refers to one trajectory out of 30 000. The two axes show two cellular properties pertaining to the 30th generation. For example, in panel (a) the x axis shows the volume of a cell at birth and the y axis shows the added volume of the cell in the same generation. The black squares and vertical lines show the mean and s.d. of Δ within a bin of V_b. The lack of correlation between Δ and V_b is evidence of the adder property. The slope ($\pm$ standard error) of the line of binned data is 0.019 ± 0.010 (we expect zero slope for adder). (b) The volume of the cell at division, V_d, is positively correlated with the birth volume, with slope 1.019 ± 0.010 for the curve of V_d averaged over bins of V_b (for the adder, expect slope unity). (c) The interdivision time τ is negatively correlated with the birth volume, with the correlation coefficient between the two variables rescaled by their respective means being -0.483 (expect -0.5 for adder). (d) The increment in P from birth to division is uncorrelated with P_b, showing that the chemical population P also exhibits the adder property.

have the same distribution which matches the rescaled V_b and V_d distributions [Fig. 3(b)]. This is because the P, T, R are all strongly correlated with each other (see Fig. S3), with their ratios dominated by the deterministic dynamics. By virtue of (5), they are also strongly correlated to V. We remark that strong correlations between high copy number proteins are also observed experimentally [30]. Further, universal fluctuations in protein populations within cells have been reported in Ref. [31].

We find that in the PTR model with stochastic chemical dynamics, the chemical populations P, T, and R, like the volume, also exhibit the adder property [see Fig. 4(d) for P and Fig. S4 for T, R]. The reason for these populations showing the same distribution and the adder property is discussed later in Sec. IV D, where we give an analytical derivation. We also later identify circumstances wherein other chemical populations in the cell depart from the adder.

The crossover behavior of protein number fluctuations. A slight extension of the present model explains the crossover

behavior of the CV of protein levels as a function of their mean levels observed in Ref. [30]. To see this, we introduce another variable Q in the X sector, with $\dot{Q} = K_Q RP/V$. This has the same form as the production of other proteins in (4). The system now has five variables, *PTRQZ*. By dialing K_Q, we can control the absolute value of Q in the steady state. We perform stochastic simulations (with stochasticity in the chemical dynamics of all five molecules *PTRQZ*) using different values of K_Q and determine the mean value of Q at birth ($\bar{Q}$) and CV of Q for each. The result (Fig. 5) shows a universal $1/\bar{Q}$ behavior for small $\bar{Q}$ and a plateau for large $\bar{Q}$, as observed in Ref. [30].

An intrinsic source of "extrinsic" noise. In the model, there is a simple explanation of the above behavior. The $(CV)^2$ due to intrinsic fluctuations is $\approx 1/\bar{Q}$ and due to division control is $1/(3Z_c)$ [from Eq. (8) using Q in place of V_b]. For small $\bar{Q}$ ($\bar{Q} \ll Z_c$), the intrinsic fluctuations dominate; for large $\bar{Q}$ ($\bar{Q} \gg 3Z_c$), the $1/(3Z_c)$ term which is independent of $\bar{Q}$ dominates. A protein population has intrinsic fluctuations from
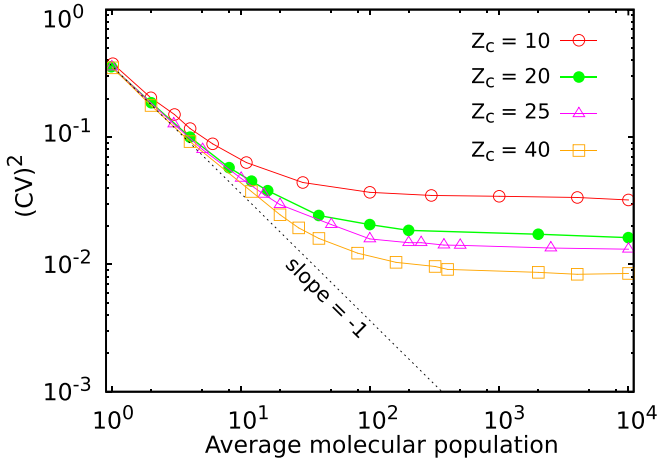
FIG. 5. Dependence of CV of a protein population on its mean value. Simulation results of the PTRZ model with an additional protein population Q as discussed in Sec. III B 3 are shown. A crossover from a $1/\bar{Q}$ behavior of $(CV)^2$ to a constant value is seen at $\bar{Q} = 3Z_c$ and $(CV)^2 = 1/(3Z_c)$.

stochasticity in production and degrading processes [29,55]. The variation not explained by intrinsic fluctuations is referred to as "extrinsic" [56]. Stochasticity in the division time has been recognized as a source of extrinsic noise leading to this crossover behavior [57]. In the present model, both the saturating value of $(CV)^2 [\simeq 1/(3Z_c)]$ and the scale of $\bar{Q}$ at which the crossover occurs ($\bar{Q} \simeq 3Z_c$) are governed by the division threshold Z_c (or in general by $N = Z_c - Z_r/2$). The so-called extrinsic noise in the division time is, in the present model, ultimately a consequence of intrinsic fluctuations in the Z population, which causes the fluctuations in the first passage time of Z reaching its threshold Z_c and hence determines a lower bound on cell-to-cell variation of size and of intracellular chemical populations.

4. Comparison with experimental distributions

We compare the distributions of certain quantities rescaled by their means as predicted by the model with the experimentally observed distributions for *E. coli* [11] in Fig. 6. As mentioned earlier, the distributions of rescaled Δ, τ, V_d predicted by the model are robust and depend only on a single
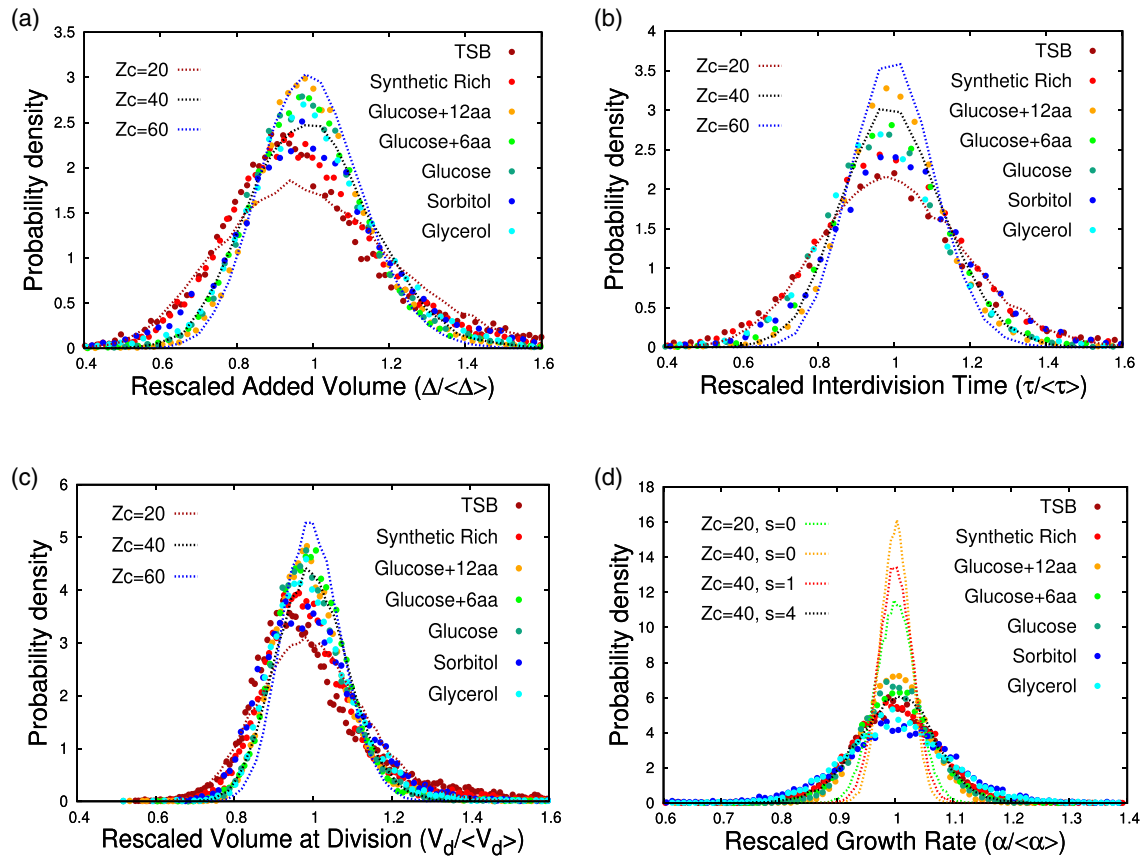

FIG. 6. Comparison of steady-state distributions predicted by PTRZ model with experimental data. Panels (a)–(d) respectively show the distributions of Δ, τ, V_d, and α rescaled by their means. Experimental values of various quantities for *E. coli* in seven media were obtained from data of Ref. [11] to produce the binned histograms. The different media are represented by dots of different color. Model simulations are done with parameter values as in Fig. 1, except Z_c, which takes values 20, 40, and 60. All model histograms (dotted lines) are obtained from an ensemble of 30 000 cells. In all simulations, the intrinsic stochasticity in the dynamics of all four chemicals P, T, R, and Z is present. Partitioning stochasticity is absent in panels (a)–(c) and in two curves of panel (d) ($s = 0$; see Sec. III B 5 for definition), but is present in two simulations shown in panel (d) ($s = 1, 4$). Threshold stochasticity is absent in all simulations shown ($s' = 0$).

parameter $N = Z_c - Z_r/2$. Figures 6(a)–6(c) show that the experimental data points for all three quantities largely are between model curves corresponding to $N = 20$ and $N = 60$, when the only source of stochasticity in the model is the intrinsic dynamics of the populations. ($N = Z_c$ in Fig. 6 since $Z_r = 0$.) This places an experimental constraint on a key parameter of the model, N. Note that in the real cell there could be other upstream sources of stochasticity in Z production, such as the stochasticity in its mRNA production and degradation. These have not been taken into account in our simple model. They would also contribute to the CV of the first passage time distribution, thereby allowing larger values of N to be consistent with the data.

It is seen in Fig. 6(d) that in the range $N \geqslant 20$ the model predicts a much narrower width for the α distribution than is experimentally observed [compare the $Z_c = 20, 40; s = 0$ curves for the model in Fig. 6(d) with those of the data]. While Ref. [11] reports a CV ranging from 6% to 11% depending upon the medium in the experimental data, the model with only intrinsic stochasticity ($s = 0$, see below) predicts a CV of the α distribution to be only 3.6% at $N = 20$, 2.5% at $N = 40$ and lower for larger values of N. We recall that the α distribution in the model is not as universal as the distributions of Δ, τ, V_d, and depends on other model parameters including the parameter values in the PTR sector (Fig. S1 [51]). We show below that its width is sensitive to and increases significantly with the inclusion of partitioning stochasticity in the model. In other words, the model allows room to get the experimentally observed width of the α distribution also, without destroying its agreement with the other distributions. It is also worth mentioning that there is a significant variation in the CV of the α distribution reported by different groups, suggesting that this quantity is sensitive to experimental details and perhaps needs to be measured with a higher degree of control.

5. Effect of other sources of stochasticity on the distributions and the adder property

Partitioning stochasticity. In addition to the intrinsic stochasticity in the chemical population dynamics, consider a second source of stochasticity, the uneven partitioning of chemicals between the two daughter cells during cell division. Thus, at division, instead of replacing X_i by $X_i/2$ we replace it by X_i' drawn (independently for each i and each generation) from a Gaussian distribution with mean $= X_i/2$ and standard deviation $s\sqrt{X_i}/2$, where $s \geqslant 0$ is a parameter characterizing the strength of partitioning stochasticity [58]. $s = 1$ corresponds to the symmetric binomial partitioning wherein each molecule can independently go to either daughter with equal probability (the Gaussian X_i' distribution approximates the binomial distribution $P(x) = \binom{M}{x} p^x q^{M-x}$ with $M = X_i$ and $p = q = 1/2$). $s > 1$ implies that fluctuations are stronger than the binomial case, a situation that can arise if molecule clusters are partitioned instead of single molecules [58]. The result for the α distribution with both sources of stochasticity (intrinsic + partitioning) is shown in Fig. 6(d) and for the other distributions in Fig. S5 [51]. It is seen that when partitioning stochasticity is included [e.g., $Z_c = 40, s = 4$ in Fig. 6(d)], the model reproduces a width of the α distribution

comparable to the experimental data (CV of α is 6.8% for these parameter values). Figure S5 shows that the inclusion of partitioning stochasticity has a relatively small effect on the width of Δ, τ, V_d distributions in this regime and that the adder property continues to hold in the model at these strengths of the partitioning stochasticity.

Stochasticity in the threshold value of Z. The biochemical mechanism implementing the trigger when Z reaches Z_c is expected to have its own stochasticity. Thus, the value of Z at which division is triggered need not be precisely Z_c in every cell, but could vary from cell to cell and generation to generation around Z_c. We implement this threshold stochasticity by drawing the threshold (now denoted Z_c') from a Gaussian distribution with mean Z_c and standard deviation s', independently for each cell in every generation. The parameter $s' \geqslant 0$ characterizes the strength of the threshold stochasticity. We ran simulations of the PTRZ model with two sources of stochasticity, the intrinsic stochasticity of the chemical dynamics of the populations P, T, R, Z and the threshold stochasticity implemented as above. In these simulations, the local dynamics of the populations is as before, except that the value of Z at which division is triggered, Z_c', is chosen from the above mentioned Gaussian distribution in each generation. Immediately after triggering, Z is reset to Z_r, and at division, all populations are halved. We consider two cases (a) Z_r is a fixed number, the same for all generations, and (b) Z_r is a fixed fraction of Z_c' (the reset value of Z after triggering is a fixed fraction of the value of Z at triggering).

Threshold stochasticity modifies the distributions of Δ, τ, V_d; however, the modification remains small as long as $s' \ll \sqrt{Z_c}$ as the dominant contribution comes from intrinsic stochasticity in this regime. When $s' > \sqrt{Z_c}$, threshold stochasticity is the dominant contributor to these distributions. However, while it modifies these distributions, in case (a) it does not affect the adder property, which survives even when $s' > \sqrt{Z_c}$. This is shown in Fig. S6A [51]. In fact, it can be shown analytically that when threshold stochasticity is the only stochasticity present (the intrinsic stochasticity is turned off), the PTRZ model exhibits the adder property in the steady state, provided that Z_r is independent of Z_c'. However, when Z_r is correlated with Z_c', as is true for case (b), we lose the adder property (see Fig. S6B [51]. This is because the added volume from one trigger to the next depends upon the reset value Z_r after the first trigger, while V_b of a generation depends on the value of Z_c' in the previous generation. Thus, the added volume becomes correlated with V_b when Z_r is correlated with Z_c'.

Unlike partitioning stochasticity, threshold stochasticity does not contribute significantly to the width of the α distribution even though it affects the Δ, τ, V_d distributions. In fact, when threshold stochasticity is the only stochasticity present, the latter distributions have a finite CV but the CV of the α distribution is zero. The reason for this will be given in Sec. IV E.

Another kind of stochasticity present in the cell is that the position of the septum is not necessarily at the middle of the dividing cell. This directly affects the distribution of V_b and enhances its CV, which is found to be larger than the CV of V_d [11]. The effect of this stochasticity is not considered in the present paper.

IV. ANALYTICAL DERIVATION OF THE RESULTS OF THE MODEL

In this section, we provide explanations of the results of the PTRZ model presented in the previous section. In particular, we discuss the exponential growth of populations and V, and give analytic expressions for the average interdivision timescale, intracellular concentrations, and the cell size. We also explain the origin of the adder property and the shape of the Δ and τ distributions. We explain why the α distribution behaves so differently from the Δ, τ, V_d distributions under different types of stochasticity and what factors contribute to it. These results will be generalized to a much wider class of models than the PTRZ model in Sec. V.

A. The PTR sector, exponential growth, and the growth rate

At the deterministic level, all the features of the steady-state solution in Fig. 1 can be understood by considering an exponential ansatz for the trajectories of the chemical populations:

$$P(t) = P_b e^{\mu t}, \quad T(t) = T_b e^{\mu t}, \quad R(t) = R_b e^{\mu t}, \quad (9)$$

with four unknown constants P_b, T_b, R_b (representing the populations at birth of the cell in the steady state), and μ, the growth rate of the PTR cell. It is easy to see [50] by substituting (9) into Eq. (4) that (9) is a solution of (4) only if the ratios of the populations and μ are fixed in terms of the parameters. Specifically,

$$\psi_T \equiv \frac{T}{R} = \frac{T_b}{R_b} = \frac{m_R}{m_T} \frac{f_T}{f_R} \frac{(\mu + d_R)}{(\mu + d_T)}, \quad (10a)$$

$$\psi_P \equiv \frac{P}{R} = \frac{P_b}{R_b} = \frac{m_R}{f_R} \frac{(\mu + d_R)}{\mu} \left[\frac{K_P f_T}{m_T(\mu + d_T)} - 1 \right], \quad (10b)$$

$$\mu = \frac{\beta - \sqrt{\beta^2 - 4\alpha'\gamma}}{2\alpha'}. \quad (10c)$$

Here $\alpha' = 1 - \epsilon_1$, $\beta = a + b + \epsilon_2$, $\gamma = ab$; $a = \nu f_T - d_T$, $b = \rho f_R - d_R$; $\nu = K_P/m_T$, $\rho = k/(m_R v_P)$; $\epsilon_1 = \frac{1}{m_T} \frac{v_T}{v_P} f_T + \frac{1}{m_R} \frac{v_R}{v_P} f_R$, $\epsilon_2 = \frac{1}{m_T} \frac{v_T}{v_P} f_T d_R + \frac{1}{m_R} \frac{v_R}{v_P} f_R d_T$. If P_b, T_b, R_b satisfy (10a) and (10b) and μ is given by (10c), then (9) is a solution of (4). That an exponentially growing trajectory is a solution of the nonlinear equations (4) is a consequence of the fact that (4) defines a class-I system.

Not only is (9) a solution of (4), it is a stable attractor of the dynamics. In Fig. 1, we have considered the PTR sector in conjunction with the Z sector which truncates the growth of PTR at discrete times [at the point of division, Eq. (4) is effectively suspended]. However, if there is no such truncation and (4) together with (5) is the sole dynamics of PTR, then the exponentially growing solution (9) happens to be an asymptotic attractor of the dynamics (see Fig. S9 [51] for numerical evidence with different initial conditions and parameter sets). Thus, starting from an arbitrary initial condition, eventually the system approaches the exponential trajectory (9) with μ given by (10c) and ratios of populations given by (10a) and (10b).

When division control via the Z sector is included in the dynamics, we get the behavior shown in Fig. 1 with a limit cycle attractor. In the growth phase of this attractor (the period after birth and just before division), the populations again grow exponentially following (9), with the exponential growth rate of P, T, R, V matching the formula (10c), and the ratios of chemicals in the attractor matching (10a) and (10b). We have verified this numerically for diverse initial conditions [with $P(0), T(0), R(0) > 0$, and $0 \leqslant Z(0) < Z_c$] and diverse parameter sets for which the right-hand side of (10c) is positive (when the latter is negative, there is no exponential growth). The numerical work suggests that given a fixed set of parameters [with a positive right-hand side of (10c)], for arbitrary physical initial conditions the system always settles down in a limit cycle attractor similar to the one described in Fig. 1 such that the trajectory between birth and division in every cycle is described by (9) and (10) (numerical simulations have been done for non-negative values of the rate constants and other parameters of the model). The positivity of the right-hand side of (10c) seems to be the only collective requirement on the parameters of the PTRZ model for this kind of dynamical behavior to arise.

B. Curve of balanced growth, interdivision time, and concentrations

By construction, the dynamics of the PTR sector does not depend upon Z, except for the fact that at certain discrete times (when Z approaches Z_c) the three populations are halved. Therefore, it is useful to consider the projection of the dynamics onto the three-dimensional space with coordinates (P, T, R). (Since we are dealing with populations, we only consider the positive octant.) Geometrically, Eqs. (10a) and (10b) define a straight line passing through the origin of this three-dimensional space whose angles with the three coordinate axes are fixed by the parameters. We refer to this line in the three-dimensional space as the curve of balanced growth (CBG) for this system. If the initial point of a trajectory lies on this line, the three populations grow exponentially according to (9) with the rate μ given by (10c), and their ratios remain constant in time and are given by (10a) and (10b). Since the division process halves the three populations, they remain on the CBG after division. Thus, the CBG is an invariant manifold of the deterministic dynamics. Since numerically we find that starting from arbitrary initial values the ratios approach those given by (10a) and (10b), this means that the stable attractor of the dynamics lies on the CBG. In fact, the steady state is a limit cycle lying on the CBG characterized by repeated rounds of exponential growth of populations from birth to division with growth rate μ until the populations double, followed by halving of the populations. The interdivision timescale on this limit cycle is therefore given by

$$\tau = \ln 2 / \mu, \quad (11)$$

with μ given by (10c). Note that the concentrations of three chemicals, given by $[P] \equiv P/V$, $[T] \equiv T/V$, and $[R] \equiv R/V$, are the same everywhere on the CBG and hence are constant on the limit cycle. This is because V is a linear function of the populations (5). Thus, for example, $V = R(v_P \psi_P + v_T \psi_T + v_R)$; hence, R/V is completely expressed in terms of the ratios ψ_P and ψ_T, which are constant on the CBG. Thus, the growth rate or interdivision timescale and all intensive

quantities pertaining to the PTR sector in the steady state of the deterministic dynamics are completely determined by the parameters of the PTR sector of the model.

C. The Z sector and cell size

The Z sector determines the absolute size scale of the cell by fixing one extensive quantity pertaining to the PTR sector. In the deterministic steady state, since P, T, R satisfy (9), we can write (7) as

$$\dot{Z} = Ce^{\mu t}, \tag{12}$$

where $C = \frac{K_Z R_b P_b}{V_b}$. C is an extensive quantity of the PTR sector (homogeneous degree 1 in the populations) and can be written in terms of V_b and the intensive quantities:

$$C = K_Z\left(R_b P_b / V_b^2\right)V_b = K_Z\psi_P(v_P\psi_P + v_T\psi_T + v_R)^{-2}V_b. \tag{13}$$

The requirement that in the steady state Z must also complete its cycle in the doubling time τ fixes C and hence the size of the cell. Equation (12) has the solution

$$Z(t) = Z_b + \lambda(t), \quad \lambda(t) \equiv B(e^{\mu t} - 1), \quad B \equiv \frac{C}{\mu}, \tag{14}$$

where Z_b is the value of Z at $t = 0$. Note that unlike P, T, and R, Z does not increase exponentially because it is reset from Z_c to a value $Z_r \neq Z_c$. If there is no resetting ($Z_r = Z_c$), then (14) implies that Z also increases exponentially.

When $\tau_1 = 0$, then Z equals Z_c at $t = \tau$, and $Z_b = Z_r/2$. Substituting this in Eq. (14) gives $B = N \equiv Z_c - Z_r/2$. This fixes the absolute size scale of the cell and the absolute populations. In particular, we get $R_b = N\mu(v_P\psi_P + v_T\psi_T + v_R)/(K_Z\psi_P)$, from which P_b, T_b can be obtained by multiplying by ψ_P, ψ_T. Further, using (13) we get

$$V_b = \frac{N\mu}{K_Z}\frac{(v_P\psi_P + v_T\psi_T + v_R)^2}{\psi_P} \equiv A. \tag{15}$$

This is an explicit formula for the cell size at birth in terms of the model parameters [ψ_T and ψ_P being given by (10a) and (10b)]. The numerical values obtained in deterministic simulations (e.g., Fig 1) agree with this formula (as do the absolute populations).

When $\tau > \tau_1 > 0$, instead of $B = N$ we get $B = Ne^{\mu\tau_1}$ (see Supplemental Material, Sec. S2 A, for the derivation [51]). Then it follows that

$$V_b = Ae^{\mu\tau_1}, \tag{16}$$

where A is defined in Eq. (15). This expression contains the exponential factor $e^{\mu\tau_1} = e^{\mu(C+D)}$ obtained by Donachie [59]. The $e^{\mu\tau_1}$ factor in the expression for V_b arises because of the exponential dependence of $\dot{Z}$ on time, Eq. (12), which is a consequence of the exponential growth of the X-sector populations in the steady state. The latter, in turn, is a consequence of the homogeneous degree one nature of the functions f_i defining the X-sector dynamics, which we argue later is quite universal and not restricted to the PTR model.

The PTRZ model has been defined above for $\tau > \tau_1$. To extend it for higher growth rates, one needs multiple origins of replication [60]. Adapting the work of Ref. [35] to the present model, it can be shown that the formula (16) holds for higher growth rates as well. Details will be presented elsewhere.

The Schaechter *et al.* growth law of cell size [40], a strong version of which has been established in Ref. [61], states that the average cell volume in different growth conditions depends exponentially on the growth rate μ. Note that while Eq. (16) contains the exponential factor $e^{\mu\tau_1}$, it cannot be construed as equivalent to this growth law. This is because (16) has, in addition to the exponential factor $e^{\mu\tau_1}$, the prefactor A, which itself has a complicated μ dependence. This factor depends upon the ratios of various chemicals in the steady state [which, in turn, depend upon μ; see (10)] and the contribution of each chemical to the volume of the cell (the constants v_i). In general, the prefactor A (unlike the $e^{\mu\tau_1}$ factor which is much more universal) depends upon the details of the X and Z sectors—the actual form of the functions f_i, h, regulatory mechanisms acting in the X sector (which are contained in the functions f_i), parameter values, etc. A more general expression for A in terms of the functions f_i, h, and v_i is given later, Eq. (34). However, comparing the theoretical prediction (16) or (34) with experiments, in particular reproducing the growth law [40,61], is a task for the future.

D. Stochastic dynamics and the adder property

The full stochastic dynamics of the PTRZ model whose numerical results were presented earlier is difficult to treat analytically. However, we can make some approximations to obtain partial results. As observed numerically in the statistical steady state, the growth of PTR sector populations in a given generation could be approximated by an exponential fit [Fig. 2(b). Moreover, since the actual numbers of P, T, and R were large, their relative fluctuations around their average trajectories were small. Thus, as an approximation we ignore the stochasticity in the chemical dynamics of P, T, and R and only consider the stochasticity in Z. For simplicity, we assume that in the steady state the PTR dynamics is deterministic and lies on the curve of balanced growth discussed earlier. Thus $P(t)$, $T(t)$, and $R(t)$ are assumed to be given by (9), their population ratios and the growth rate being constant and fixed by (10). However, due to the fluctuation in the time taken by Z to reach Z_c, the exponential factor by which they grow varies from generation to generation and in each generation they start from a different point on the CBG at birth (i.e., the absolute scale of P_b, T_b, R_b varies from generation to generation). Clearly, this approximation assumes that there is no fluctuation in the growth rate α and hence we cannot hope to obtain the α distribution from this approach [α, defined as the slope of $\ln V$ versus t, equals μ in this approximation, whose value is given by the right-hand side of (10c)]. However, we can obtain analytic expressions for the τ and Δ distributions.

1. Derivation of the adder property and distributions of τ and Δ

In order to get the probability distribution of the interdivision time, we need to consider the stochastic version of the differential equation (1b), or more specifically, (7). As a consequence of the assumptions in the previous paragraph, we can use (9), and hence (7) reduces to (12). This means that the probability that Z increases by unity in the small time interval $(t, t + \delta t)$ is given by $\dot{Z}\delta t$ where $\dot{Z}$ is given by (12), and the probability that it remains unchanged is $1 - \dot{Z}\delta t$. A

given trajectory starts with fixed values of $P, T, R, V,$ and Z at birth, denoted $P_b, T_b, R_b, V_b,$ and Z_b and hence a fixed value of $C = \frac{K_Z R_b P_b}{V_b}$. These values can change from generation to generation, but since in the present approximation $P_b, T_b,$ and R_b are assumed to always lie on the CBG, the ratio C/V_b is constant for all generations and given by (13). This justifies a crucial assumption, namely, that $C \propto V_b$ or equivalently $\dot{Z} \propto V_b e^{\mu t}$, made in Ref. [36] where the first passage time problem of a division triggering chemical based on (12) has been discussed. Consider the ensemble of trajectories with a fixed value of $P_b, T_b, R_b, Z_b,$ and hence fixed V_b and C. Since the right-hand side of (12) is a fixed function of time, the stochastic process in Z is the one-dimensional inhomogeneous Poisson process (if the right-hand side had been independent of t it would have been the standard Poisson process). Given $Z = Z_b$ at $t = 0$, the probability that the first passage time for it to reach Z_c is between τ and $\tau + d\tau$ is given by $\mathcal{P}(\tau) d\tau$, with (see Supplemental Material Sec. S2 B for a derivation [51])

$$\mathcal{P}(\tau) = \frac{\lambda^{N-1} e^{-\lambda}}{(N-1)!} \frac{d\lambda}{d\tau}, \tag{17}$$

where $\lambda = \lambda(\tau) = B(e^{\mu\tau} - 1)$ and $N = Z_c - Z_b = Z_c - Z_r/2$.

In a given trajectory of the ensemble, the volume increment between birth and division is given by $\Delta = V_b(e^{\mu\tau} - 1) = (V_b/B)\lambda(\tau)$. This being proportional to λ, the probability $\mathcal{R}(\Delta) d\Delta$ that a given trajectory in this ensemble has a volume increment in the range Δ to $\Delta + d\Delta$ is obtained from (17)

$$\mathcal{R}(\Delta) = a^N \frac{\Delta^{N-1} e^{-a\Delta}}{(N-1)!}, \quad a = \frac{B}{V_b} = \frac{C}{\mu V_b} = \frac{N}{A}, \tag{18}$$

where A is given by (15). From this, it follows that

$$\langle \Delta \rangle \equiv \int_0^\infty d\Delta \, \mathcal{R}(\Delta) \Delta = \frac{N}{a}, \tag{19}$$

which is the same as the deterministic value of V_b (15). The distribution of rescaled Δ, $u \equiv \Delta/\langle\Delta\rangle$ is thus a Γ distribution given by

$$P(u) = N^N \frac{u^{N-1}}{(N-1)!} e^{-Nu}. \tag{20}$$

The adder property follows from (18). Note that $N = Z_c - Z_r/2$ is a constant parameter of the model that does not change from generation to generation as long as Z_c, Z_r do not change from generation to generation. The only other parameter that determines the shape of $\mathcal{R}(\Delta)$ in Eq. (18) is a, which is independent of the value of V_b in any generation. The latter is the case because on the CBG, where (12) holds, a is a constant irrespective of the scale of V_b ($a = C/(\mu V_b)$ with C given by (13)). Thus, a depends only on the *ratios* of populations of the PTR sector, and on the CBG these ratios are the same irrespective of the scale. This proves the strong version of the adder property: not just the mean but the entire distribution of Δ is independent of V_b. The distribution of u only depends upon N and on no other parameter of the PTRZ model. This explains the extraordinary robustness of the u distribution mentioned in Sec. III B 2. From (20), it follows that CV of $\Delta = \sigma_u = \frac{1}{\sqrt{N}}$, which derives the last equality in Eq. (8).

The above distributions and the adder property were obtained in Ref. [36] from the assumptions that the cell volume grows exponentially with time [$V(t) = V_b e^{\mu t}$] and that the rate of growth of the time-keeper protein Z is proportional to $V(t)$. The first assumption is equivalent to a linear model for cell size growth, $\dot{V} = \mu V$. The present work obtains the exponential dependence of V on t from a nonlinear dynamical model of the cell whose volume is defined in terms of the three chemical populations. As discussed above, it also justifies the second assumption ($\dot{Z} \propto V$). In other words, in the present work Eq. (12) with $C \propto V_b$ is not an assumption but a consequence of more basic dynamics, wherein the cell is attracted to the CBG. We show later in Sec. V C that this holds for a much larger class of models in which the X sector has n chemicals where the functions f_i and h in Eq. (1) are nonlinear functions arising from mass action kinetics. Thus, the present work generalizes the results of Ref. [36] to a very large class of nonlinear models.

We note that the present model implies that the populations of the X sector also satisfy the adder property. Their increments between birth and division, $\Delta_i \equiv X_{id} - X_{ib}$, have the same distribution as Δ given by (18), with $a = B/X_b$. This follows from the fact that like V, the X_i also grow exponentially, and hence $\Delta_i = X_{ib}(e^{\mu\tau} - 1) = (X_{ib}/B)\lambda(\tau)$.

In the above derivation of the Δ and τ distributions and the adder property, it has been assumed that P, T, R have no stochasticity and lie on the CBG. When stochasticity in $P, T,$ and R is included, they no longer lie on the CBG. In fact, in a given generation, the best-fit value of the growth rate is not the same for all three populations, and hence (7) does not reduce to (12). Nevertheless, the analytically derived distributions (20) and (17) compare well with our numerical simulations of PTRZ model where stochasticity in $P, T,$ and R is included; see Figs. 7(a) and 7(b). Further, as discussed in Sec. III B 2 and shown in Fig. 4, the adder property also appears. This suggests that the approximation of treating $P, T,$ and R as deterministic and on the CBG is a good approximation to the full stochastic dynamics.

2. When does the adder property arise?

The adder property and the Δ and τ distributions have been derived above from three assumptions: (i) In a given generation, the cell volume obeys $V(t) = V_b e^{\mu t}$, where V_b is the volume at birth in that generation (taken to be at $t = 0$). (ii) A molecular population Z which starts at some value Z_b triggers division at time τ when it reaches a threshold value Z_c, where $Z_c - Z_b$ is a positive constant (the same for all generations). (iii) The dynamics of Z between birth and division is the stochastic version of (12) in which the constant C is such that C/V_b is a constant for all generations. In fact, the same can be derived from weaker assumptions that do not require V and $\dot{Z}$ to be exponential functions of time. Consider the case where Z and some quantity Y have their rates of increase between birth and division to be proportional to each other,

$$\dot{Z} = H(t), \quad \dot{Y} = KH(t), \tag{21}$$

where the function $H(t)$ and the constant K could depend upon the generation. If now we make the Z dynamics the
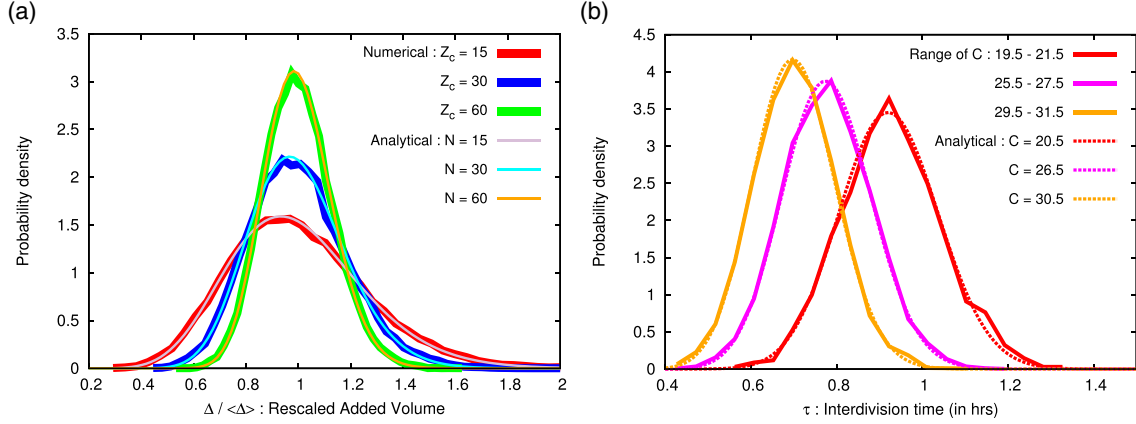
FIG. 7. Analytic Δ, τ distributions Eqs. (20) and (17) agree with simulations of PTRZ model. (a) Rescaled Δ distribution (20) compared with numerical simulations of PTRZ model for different values of N. Data from all bins of V_b have been pooled together in the numerical curves, since the Δ distribution is independent of V_b (see Fig. S2). (b) Conditional τ distribution obtained from numerical simulations of PTRZ model compared with Eq. (17). Since the analytical distribution is conditional on the value of C, for comparison the numerical data was binned in ranges of C indicated in the legend. The derivation of Eqs. (20) and (17) assumes stochasticity in Z dynamics only; P, T, and R are assumed to lie on the CBG and obey deterministic dynamics. In the PTRZ simulations, there is stochasticity in the chemical dynamics of Z as well as P, T, and R. The parameters for the PTRZ model are the same as in Fig. 1, except that $Z_c = 15, 30, 60$ in panel (a) and $Z_c = 30$ in panel (b).

stochastic version of $\dot{Z} = H(t)$, while retaining the deterministic dynamics of Y, $\dot{Y} = KH(t)$ and also retaining assumption (ii) above, the distribution (17) still arises, with $\lambda \equiv \int_0^{\tau} dt' H(t')$ (see Supplemental Material, Sec. S2 B [51]). In other words, (17) gives the conditional distribution of interdivision time, valid for those generations in which the function $H(t)$ is the same. Now $\Delta \equiv \Delta_Y \equiv Y(\tau) - Y(0) = \int_0^{\tau} dt' K H(t') = K\lambda$, and hence the distribution of Δ is again given by (18) with $a = 1/K$. This distribution is independent of the functional form of $H(t)$ and only depends upon K. In other words, $H(t)$ appearing in Eq. (21) does not have to be an exponential function of time as in Eq. (12). Thus, if K is independent of $Y_b = Y(0)$, then so is $\mathcal{R}(\Delta_Y)$, and therefore Y displays the adder property. This argument shows that not only the adder property for Y but also the shape of the distribution (18) for Δ_Y arises when these assumptions hold [namely, (ii), (21) with K independent of the generation, and the stochastic dynamics of Z].

A biochemical scenario in which Eq. (21) can arise was presented by Sompayrac and Maaloe [32] and emphasized recently in the context of the adder property [34,35]. In that scenario, the Z molecule is on the same operon as another molecule A whose concentration is held constant in the cell through regulatory mechanisms (autorepression). This implies that the increase in cell volume is proportional to the number of molecules of Z produced, thereby realizing Eq. (21) with $H(t)$ being an unspecified function of t. Then, together with assumption (ii) and the stochastic dynamics of Z, the adder property of cell volume and the distributions (17) and (20) follow as argued above.

The present model presents an alternative mechanism for realizing Eq. (21) which does *not* require a regulatory mechanism. Here this property arises because the population dynamics has an attractor lying on the CBG in which both $\dot{Z}$ and $\dot{V}$ are proportional to $e^{\mu t}$. We argue later that the latter is a very generic property of chemical dynamics in self-expanding containers. As a consequence, it predicts that

the adder property of the volume should be accompanied by the same for intracellular populations (the variable Y could be V or any of the X_i).

Cooperativity in the Z dynamics does not spoil the adder property. We now consider the generalization in which the function $h(\mathbf{X}, Z)$ in Eq. (1b) has a nontrivial Z dependence. In particular, we consider the generalization of (7):

$$\frac{dZ}{dt} = h(P, T, R, Z) = K_Z \frac{RP}{V} r(Z), \quad r(Z) = \frac{Z^{h'}}{K^{h'} + Z^{h'}},$$
(22)

where K and h' are constants. $r(Z)$ represents a positive autoregulation of Z production, with h' being a Hill coefficient measuring the strength of cooperativity in the Z dynamics. Since Z does not appear in the functions f_i, in the deterministic dynamics the PTR variables converge to same ratios as before (the CBG is unaffected by this change) and the growth rate μ in the steady state is also the same and given by (10). Under the approximation that PTR sector is treated deterministically (discussed at the beginning of Sec. IV D), Eq. (12) is therefore replaced by

$$\dot{Z} = Ce^{\mu t} r(Z),$$
(23)

where C is given by (13) as before. In order to determine the first passage time distribution and the added volume distribution, we now need to consider the stochastic version of (23), where the right-hand side has a nontrivial Z dependence. By considering the master equation following from (23), we show analytically in the Supplemental Material, Sec. S2 C [51], that the presence of a nontrivial autoregulation changes the shape of the Δ distribution [which is no longer given by (18)], but the adder property *remains intact*. This is also borne out by numerical simulations shown in Figs. S7A and S7B, conducted on the PTRZ model with intrinsic stochasticity in all four population variables and autoregulation implemented in the Z sector via $r(Z)$. Both positive and negative autoregulation of Z preserve the adder property.

However, the adder property is *not* preserved if the regulatory function r is a function of the concentration of Z, $[Z] \equiv Z/V$, instead of the population Z (see Figs. S7C, S7D, and text in Supplemental Material, Sec. S2 C [51]).

Departure from adder. After the brief excursion in the last two paragraphs to the case where h depends nontrivially on Z, we return to case where h is independent of Z [where the Z dynamics is given by (7)] and the distributions (17) and (18) hold along with the adder property. Since the Δ distribution (18) depends upon N, one way of losing the adder property is to have N correlated with V_b. The part of assumption (ii) that $Z_c - Z_b$ is the same in all generations ensures that N is not correlated with V_b. A weaker assumption that N varies from generation to generation but is uncorrelated with V_b would still give the adder property, though the distribution for Δ would in general change and depend upon the distribution of N. An example of this is given in Sec. III B 5 where threshold stochasticity is considered. We discussed two cases: (a) where Z_c varies randomly from generation to generation but $Z_b = Z_r/2$ is fixed; this corresponds to the weaker assumption (ii) that preserves the adder property (see Fig. S6A) because N is still uncorrelated with V_b; and (b) where Z_c varies randomly from generation to generation and Z_r of a generation is correlated with the Z_c of that generation; this spoils the adder property (see Fig. S6B) because this correlates the Z_b (and hence N) of a given generation with the V_b of the same generation. Since the triggering event at the threshold and the degradation of Z must be implemented by biochemical mechanisms, it is possible to imagine both kinds of scenarios.

Departure from the adder property for an intracellular chemical population while cell size exhibits the adder property. Up to now, we have seen that in the PTRZ model, the adder property for V is accompanied by the adder property for the chemical populations P, T, and R. However, in general, it is not necessary that all molecules in the cell exhibit the adder property when the cell size does. We now discuss an example where P, T, R, and V exhibit the adder property while another molecule Q in the cell does not. Consider the PTRZ model defined by Eqs. (4), (5), and (7) augmented with another molecule Q in the X sector whose population dynamics is given by

$$\dot{Q} = K_Q R P r([Q])/V. \tag{24}$$

Here $r([Q])$ is a function of the concentration of Q, $[Q] \equiv Q/V$, that represents auto-regulation of Q production. For example, $r([Q]) = [Q]^{h'}/(K^{h'} + [Q]^{h'})$ represents auto-enhancement and $r([Q]) = K^{h'}/(K^{h'} + [Q]^{h'})$ represents autorepression (with $h' > 0$ being a Hill coefficient and K being another constant). In the model, we allow two sources of stochasticity: (i) the intrinsic stochasticity in the chemical dynamics of the molecules P, T, R, Z, and Q and (ii) the partitioning stochasticity at division as discussed in Sec. III B 5. Figure S8 displays a simulation where V exhibits the adder property but Q does not. It is seen that the increment in Q, $\Delta Q \equiv Q_d - Q_b$ (where Q_b and Q_d, are, respectively, the population values at birth and division) increases with Q_b in the case of autoenhancement and decreases with Q_b in the case of autorepression. This behavior occurs because partitioning stochasticity causes a departure from the curve of balanced growth as discussed in the next subsection.

E. Origin of the α distribution

As is evident from the above, under the hypothesis that the first passage time of a molecule to reach a threshold controls cell division, the distributions of τ, Δ, and V_d are closely related and their CVs are governed by N or equivalently Z_c as described by (8). One can now ask for what controls the distribution of α in this setting. Note that α is obtained for a given trajectory by fitting the observed trajectory of cell volume to the exponential form $V(t) \sim e^{\alpha t}$. Thus α in any generation is the slope of the best fit straight line to the $\ln V$ vs t curve from birth to division. In the analysis of the previous subsection where analytic expressions for the τ and Δ distributions have been derived, the X sector populations P, T, and R are assumed to obey their deterministic dynamics and lie on their CBG. On the CBG, the deterministic growth trajectory of P, T, R is given by (9), and hence the value of α is fixed and equal to μ given by (10c). Thus, the analysis of the previous subsection *assumes* that the α distribution has zero width and is therefore just the Dirac δ function $\delta(\alpha - \mu)$. That the analysis produces nontrivial distributions of Δ, τ, V_d while assuming that the α distribution has zero width suggests that any explanation of the origin of a nontrivial α distribution must invoke phenomena not included in that analysis.

In our numerical simulations of the PTRZ model, we included the stochastic fluctuations of the P, T, and R populations. There we found the α distribution to have a finite width [Fig. 6(d), although smaller than the one experimentally observed in Ref. [11]. We noticed that partitioning stochasticity increased the width of the α distribution to reach the width observed in Ref. [11] as the strength s of partitioning stochasticity was increased. However, threshold stochasticity did not contribute to α width. Furthermore, while the rescaled Δ, τ, V_d distributions depended only on the parameter N, the rescaled α distribution was found to be dependent on the other parameters of the PTRZ model as well (Fig. S1). These observations suggest that (a) intrinsic stochasticity and (b) departure from the CBG resulting in a mixing of a pure exponential (given by rate μ) with other functions of time contribute to the width in α in these models.

The role of intrinsic stochasticity in producing a nonzero width is straightforward. A population variable that is a smooth exponential function of time in a deterministic simulation would be a jagged function when intrinsic stochasticity is switched on, to which an exponential fit would give a different value of growth rate in different trials. The volume, being a linear combination of chemical populations, will consequently also have a variation in α in different generations.

To explain how the departure from the CBG is an independent source of variation of α, we consider two models in the same broad class defined in Sec. II but linear and much simpler than PTRZ, one containing only one population variable in the X sector (the XZ model) and the other containing two (the XYZ model). These are defined by

$$\text{XZ model: } \dot{X} = \mu X, \quad \dot{Z} = K_Z X; \quad V \equiv vX, \tag{25a}$$

$$\text{XYZ model: } \dot{X} = c_1 Y, \quad \dot{Y} = c_2 X, \quad \dot{Z} = K_Z X; \tag{25b}$$

$$V \equiv v_1 X + v_2 Y. \tag{25c}$$

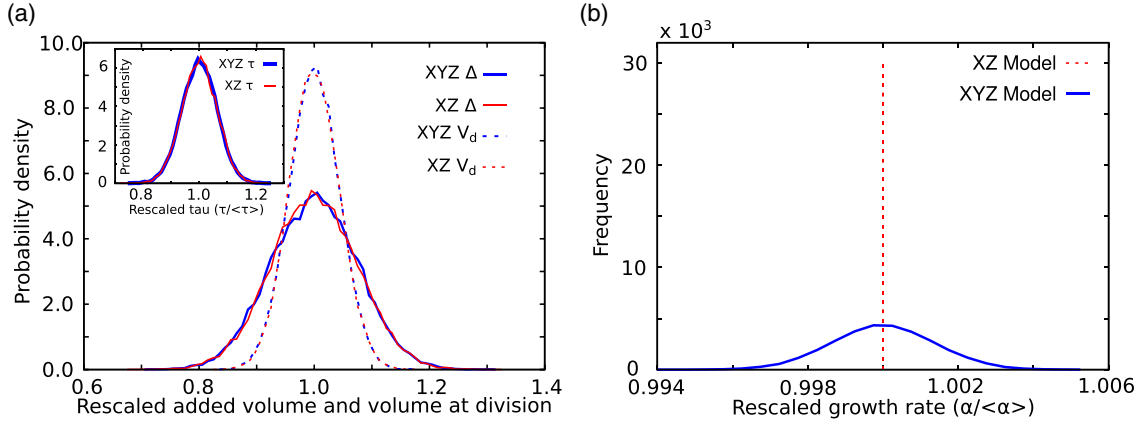FIG. 8. Distributions for XZ and XYZ models with only partitioning stochasticity. (a) Rescaled Δ, V_d distributions (τ distributions in the inset). (b) Rescaled α distribution. Distributions are made from 30 000 generations. Note that the α distribution has zero width for the XZ model but nonzero width for the XYZ model, which explains how being thrown off the CBG contributes to the width of the α distribution (see discussion in main text). In these simulations, intrinsic stochasticity due to chemical dynamics as well as threshold stochasticity are both absent, but partitioning stochasticity is present with strength $s = 1$. XZ model parameters: $\mu = 10$, $K_Z = 10^{-3}$, $v = 1$, $s = 1$. XYZ model parameters: $c_1 = 10$, $c_2 = 1$, $K_Z = 10^{-3}$, $v_1 = v_2 = 1$, $s = 1$. For both models, $Z_c = 40$, $Z_r = 20$, $\tau_1 = 0$.

Here μ, K_Z, v, c_1, c_2, v_1, and v_2 are constants. When these models are simulated with only intrinsic stochasticity in the chemical dynamics (partitioning stochasticity and threshold stochasticity absent), they produce the same rescaled Δ, τ, V_d distributions as the PTRZ model and their own specific nonzero width α distributions (see Fig. S1). This is in keeping with the discussion above. However, now consider simulations of both these models treating the population dynamics of all chemicals deterministically and implementing only partitioning stochasticity with some strength s as described in Sec. III B 5 (intrinsic stochasticity and threshold stochasticity are absent). The resulting distributions are shown in Fig. 8. The Δ, and τ, V_d distributions are different from what we get when only intrinsic stochasticity is present and all three have nonzero width for both models. But the α distribution has zero width for the XZ model and a nonzero width for the XYZ model. This difference between the two models shows that departure from the CBG contributes to the α distribution. For, in the XZ model, the phase space of the X sector coincides with the CBG (both are one-dimensional) and there can be no departure from the CBG: Any initial condition $X(0)$ leads to the trajectory $X(t) = X(0)e^{\mu t}$, which always leads to $\alpha = \mu$. Whereas in the XYZ model the XY sector phase space is two-dimensional and for a general initial condition $[X(0), Y(0)]$ the trajectory of the system is a linear superposition of two exponentials: $X(t) = ae^{\mu t} + be^{-\mu t}$ and $Y(t) = a'e^{\mu t} + b'e^{-\mu t}$, where $\mu = (c_1 c_2)^{1/2}$ and a, b, a', b' are linear combinations of $X(0), Y(0)$. Thus, $V(t)$ is also in general a superposition of two exponentials and fitting it to a single exponential will yield a value of α that will depend upon $X(0), Y(0)$. In this linear example, the dynamics is governed by the matrix $A = \begin{pmatrix} 0 & c_1 \\ c_2 & 0 \end{pmatrix}$, which has two eigenvalues $\pm\mu$. The eigenvector corresponding to $+\mu$ is the attractor of the dynamics and is the CBG [it is the line in the XY plane which passes through the origin and has the slope $Y/X = (c_2/c_1)^{1/2}$]. On this line, the growth is pure exponential with $\alpha = \mu$. But even if the system has reached this line before division, partitioning

stochasticity will throw it off this line after division, because the X and Y populations of the daughter cell are chosen independently of each other at partition and will no longer have the same ratio as before. Thus, in the next generation, the trajectory will again be a superposition of two exponentials, resulting in a value of $\alpha \neq \mu$.

The above argument makes it clear how being thrown off the CBG contributes to the width of the α distribution, because in every generation the departure of $X(0), Y(0)$ (birth coordinates of the cell) from the CBG will be random due to partitioning stochasticity and the extent of that departure will govern how different the fitted value of α in that generation is from μ, whereas in the XZ model since there can be no departure from the CBG partitioning stochasticity always gives $\alpha = \mu$. This explains why in the PTRZ model we found the width of the α distribution to grow with s: Larger s means a greater departure from the CBG and hence a greater departure from a pure exponential solution. This also explains why when only threshold stochasticity is present (intrinsic stochasticity and partitioning stochasticity absent) the width of the α distribution is zero. This is because once the system reaches the CBG attractor, threshold stochasticity by itself cannot throw it off the CBG because it only operates in the Z sector and does not affect the ratios of populations in the X sector. In view of the above discussion, it is not surprising that the α distribution depends upon parameter values and details of the models.

The mixing of exponentials as a possible origin of the α distribution has also been discussed in Ref. [62] in the context of a linear model. Our explanation, in terms of the cell being thrown off the CBG by partitioning and intrinsic stochasticity covers both linear and nonlinear dynamics.

The above discussion also clarifies the origin of the departure from the adder property of an autoregulated chemical Q discussed in the previous subsection. Since the cell is thrown off the CBG, the growth rate of Q is no longer μ but can be affected by Q_b because of autoregulation. Hence, it departs

from the adder. PTR also has a growth rate different from μ, but being large and unaffected by Q, does not show a significant departure from the adder property (and hence V also does not).

V. THE ORIGIN AND CONSEQUENCES OF EXPONENTIAL GROWTH

A. Exponential growth arises in class-I dynamical systems

We saw that the PTR sector defined by (4) and (5) was characterized by the exponential growth of chemicals in the deterministic steady state in the growth phase [insets of Figs. 1(a), 1(c), S9A, and S9B], given by Eq. (9). We now argue that this is a general property of a large class of chemical systems which naturally arise in cellular dynamics, the class-I systems. Consider a set of n chemicals whose dynamics is given by (1a), with f_i satisfying (3). For such systems, an exponential ansatz

$$X_i(t) = X_i(0)e^{\mu t} \quad \text{for all} \quad i = 1, \ldots, n \qquad (26)$$

is generically a solution of the dynamics. Substituting Eq. (26) in Eq. (1a), and using Eq. (3) to write

$$
\begin{aligned}
f_i(X_1(0)&e^{\mu t}, X_2(0)e^{\mu t}, \ldots, X_n(0)e^{\mu t}) \\
&= e^{\mu t} f_i(X_1(0), X_2(0), \ldots, X_n(0)),
\end{aligned}
\qquad (27)
$$

we get:

$$\mu X_i(0) = f_i\big(X_1(0), X_2(0), \ldots, X_n(0)\big), \quad \text{for } i = 1, 2, \ldots, n. \qquad (28)$$

The t dependence has canceled out from both sides because of the class-I nature of the f_i. This shows that for class-I systems (26) is a solution of the dynamics if and only if $\mathbf{X}(0)$ and μ satisfy (28). Equation (28) is a set of n equations for the $n + 1$ constants μ and $X_i(0)$ appearing in Eq. (26). Assuming $X_n(t) > 0$, we can define the ratios of populations $\psi_i(t) \equiv X_i(t)/X_n(t)$, $i = 1, 2, \ldots, n - 1$, $\psi_n \equiv 1$. It follows that ψ_i satisfy the differential equation $d\psi_i/dt = f_i(\boldsymbol{\psi}) - \psi_i f_n(\boldsymbol{\psi})$, where $\boldsymbol{\psi}$ denotes the vector $(\psi_1, \psi_2, \ldots, \psi_{n-1}, 1)$. Under the ansatz (26) ψ_i are time independent, $\psi_i(t) = \psi_i(0) = X_i(0)/X_n(0)$. Dividing both sides of (28) by $X_n(0)$ and using (3), we get

$$
\begin{aligned}
\mu \psi_i &= f_i(\psi_1, \psi_2, \ldots, \psi_{n-1}, 1) \quad \text{for } i = 1, 2, \ldots, n - 1, \\
\mu &= f_n(\psi_1, \psi_2, \ldots, \psi_{n-1}, 1).
\end{aligned}
\qquad (29)
$$

For class-I systems, (28) and (29) are equivalent. Equation (29) is a set of n (in general, nonlinear) equations for the n unknowns μ and $\psi_1, \ldots, \psi_{n-1}$. Generically these n equations provide a solution for the $n - 1$ independent ratios of chemicals ψ_i ($i = 1, \ldots, n - 1$) as well as μ in terms of the parameters appearing in the functions f_i (see Supplemental Material, Sec. S2 D [51] for additional remarks and qualifications). The PTR model is a particular case of $n = 3$ for which the explicit solution is given in Eq. (10). The fixed values of the ratios ψ_i so obtained define a straight line passing through the origin of the n-dimensional phase space Γ of the variables $\mathbf{X}$ in the direction of the vector $\boldsymbol{\psi}$. The implication of the above analysis is that for any initial condition $\mathbf{X}(0)$ lying on this line, the trajectory of system will satisfy (26); in other words, all the populations $X_i(t)$ will grow exponentially with the same rate μ preserving their ratios. Therefore, this line

can be referred to as a *curve of balanced growth* (CBG) for the system (a trajectory that starts on the CBG remains on the CBG with constant ratios of populations). Such a curve does not exist in general for systems that are not class I.

As discussed earlier in Secs. IV A and IV B and Fig. S9, in the case of the PTR dynamics defined by (4) and (5), an exponentially growing trajectory lying on the CBG is not just a solution but also a stable attractor of the dynamics. We have found this property to be true in simulations of many other class-I chemical systems representing cellular dynamics with different forms of the functions f_i. This includes autocatalytic systems with regulation involving Hill functions (see Eq. (24) and models in Ref. [63]), another with a network of a thousand chemical species constructed along the lines of Ref. [64], and other complicated class-I autocatalytic systems (details to be presented elsewhere). On the other hand, without the class-I property, even simple systems do not have an attractor satisfying (26). For example, we can consider the PTR model with the same Eq. (4) but with a modified Eq. (5) such that the volume is not a linear function of the populations. Let us assume that V is proportional to $S^{3/2}$ where S is the surface area of the cell (motivated by spherical shaped cells), and let us take S to be proportional to T, the number of transporter molecules. Then the asymptotic trajectory is not exponential (see Fig. S10). The reintroduction of the division process does not produce the exponential behavior in this non-class-I system. Similar behavior is observed in several other non-class-I systems we have studied. In this context, we remark that in Ref. [37] where the dynamics is class-I, exponential growth is observed, and where it departs from exponential growth it is not class-I (in the latter case, the rate of protein production becomes constant, leading to linear growth).

Thus, we find the class-I condition (3) to be an important condition for the existence of attractors with exponential growth. Note that linear functions [$f_i(\mathbf{X}) = \sum_{j=1}^{n} A_{ij} X_j$ with constant A_{ij}] are a special case of class-I and here the asymptotic attractor is exponential: $\mathbf{X}(t) = \mathbf{X}^{(\mu)} e^{\mu t}$ where μ is the eigenvalue of the matrix $A = (A_{ij})$ with the largest real part and $\mathbf{X}^{(\mu)}$ is the corresponding eigenvector. In this case, $\boldsymbol{\psi}$ is proportional to $\mathbf{X}^{(\mu)}$. Class-I systems are in general nonlinear (e.g., the PTR model above) but share the feature of exponential solutions with linear systems. Equation (28) can be considered a generalization of the eigenvalue equation $A\mathbf{X} = \mu\mathbf{X}$ to the nonlinear case, which determines the ratios of the $X_i(0)$ and the "eigenvalue" μ.

B. When does the class-I property arise?

The class-I property of f_i will always arise for well-stirred mass action kinetics, however complicated the latter may be, when the volume of the container is a linear function of the chemical populations. This can be seen as follows.

Let us consider a well-stirred chemical reactor of fixed volume V containing n chemical species whose populations are given by X_i and concentrations by $x_i = X_i/V$. The law of mass action implies that the dynamics of the concentrations is given by the set of nonlinear equations

$$\dot{x}_i = g_i(\mathbf{x}), \quad i = 1, 2, \ldots, n, \qquad (30)$$

where g_i are some nonlinear functions of the arguments that depend upon the set of chemical reactions that take place in the system. For example, if there is a chemical reaction of the kind $A + 2B \rightarrow C$, then $\dot{x}_C$ will contain a term of the kind $k x_A x_B^2$. It is important to note that the right-hand side of (30) is a function of the concentrations of the chemicals, not their populations.

Now assume the container is expanding, with the volume having a time dependence $V = V(t)$. Then the above equation would be modified to

$$\dot{x}_i = g_i(\mathbf{x}) - \frac{x_i}{V}\frac{dV}{dt}, \tag{31}$$

with the second term reflecting the effect of dilution due to expansion. Let us ask for the dynamical equations in terms of the populations $X_i = V x_i$. In a fixed (constant) volume, $\dot{X}_i = V\dot{x}_i = V g_i(\mathbf{x}) = V g_i(\mathbf{X}/V)$. In an expanding volume $\dot{X}_i = V\dot{x}_i + \dot{V}x_i = V g_i(\mathbf{X}/V) - x_i\dot{V} + \dot{V}x_i = V g_i(\mathbf{X}/V)$. Thus, in terms of the population variables, the dynamics in the expanding container does not contain any extra term and is given by

$$\dot{X}_i = V g_i(\mathbf{X}/V). \tag{32}$$

Now suppose the volume of the container is a function of the populations, $V = V(\mathbf{X})$; i.e., it depends upon time only through $\mathbf{X}(t)$. This feature would modify the nature of the dependence of the right-hand side of (32) on $\mathbf{X}$:

$$\dot{X}_i = V(\mathbf{X})g_i\left(\frac{\mathbf{X}}{V(\mathbf{X})}\right) \equiv f_i(\mathbf{X}). \tag{33}$$

A particularly interesting situation arises when $V(\mathbf{X})$ is a linear function of the populations given by (2), as is possibly true for bacterial cells. [This can happen, for example, due to osmotic pressure. If we assume that water enters or leaves the cell on a short timescale compared to the timescales of the dynamics (32) to maintain the total concentration of solute inside equal to that outside the cell x_{ext}, then $\frac{\sum_i X_i}{V} = x_{\text{ext}}$ or $V = \frac{1}{x_{\text{ext}}}\sum_i X_i$. This is a particular case of (2) where the sum over i includes all the chemical species in the bulk (interior) of the cell.] Then it follows that f_i satisfies the class-I property because (2) implies that $V(\beta\mathbf{X}) = \beta V(\mathbf{X})$; then $f_i(\beta\mathbf{X}) = V(\beta\mathbf{X})g_i(\frac{\beta\mathbf{X}}{V(\beta\mathbf{X})}) = \beta V(\mathbf{X})g_i(\frac{\mathbf{X}}{V(\mathbf{X})}) = \beta f_i(\mathbf{X})$.

Thus, the class-I property (3) follows from well-stirred mass action kinetics in an expanding container Eq. (31) and the assumption that V is itself a homogeneous degree-1 function of the molecular populations [in particular its special case (2) that it is a linear function of the populations]. Note that the $g_i(\mathbf{x})$ can be highly nonlinear functions of their arguments and so will in general $f_i(\mathbf{X})$ be; nevertheless, the f_i will satisfy the condition (3). This property of well-stirred mass action kinetics is hidden when the dynamics is formulated in terms of concentrations but is apparent when formulated in terms of populations which are extensive quantities.

Note that the dynamical system (32) in terms of extensive variables is not fully specified until V is specified as a function of time or as a function of $\mathbf{X}$. Similarly, the dynamical system (31) in terms of intensive variables is not fully specified until $\dot{V}/V$ is specified as a function of time or as a function of $\mathbf{x}$. The choice (2) specifies both dynamical systems completely where the constants v_i are treated as parameters of the system.

With this choice the concentrations satisfy the constraint $\sum_i v_i x_i = 1$, obtained by dividing both sides of (2) by V. It can be easily seen that this constraint is preserved by the time evolution under (31). Thus, there are only $n - 1$ independent intensive variables, which can be taken to be the $(n-1)$-independent x_i or the $n - 1$ ψ_i (each set can be expressed in terms of the other set).

It is important to note that the specification (2) allows us to find the steady-state growth rate as a function of the parameters. Specification (2) implies that the f_i appearing in (33) satisfy the class-I property and this leads to (28) under the ansatz (26) from which both μ and the ψ_i (or x_i) can be determined. The $\mathbf{x}$ so obtained is a fixed point of (31). The μ and $\mathbf{x}$ so obtained depend, among other things, on the parameters v_i appearing in Eq. (2) [see, e.g., (10c)]. If, instead of specifying V as a function of $\mathbf{X}$, we had simply replaced $\dot{V}/V$ by μ in (31), we could still solve for the fixed point of (31) in terms of μ and other parameters appearing in the functions g_i but we would not be able to solve for μ in terms of the parameters.

C. Consequences of the class-I property

The class-I property allows us to generalize the results of the PTRZ model to a much more general class of models described in Sec. II. In this subsection, we discuss the generalizations and the assumptions under which they hold.

1. Exponential growth of size, expressions for the average birth volume, and interdivision timescale

Exponential trajectories. Most mathematical models of cellular dynamics are formulated in terms of concentrations, with particular choices of the functions g_i in Eq. (30). For cells in a steady-state culture, it is usually *assumed* that the volume grows exponentially; consequently a term $-\mu x_i$ is added to account for dilution due to volume expansion. However, the exponential dependence of V on t is, *a priori*, a puzzling fact given the nonlinear nature of cellular dynamics. The discussion above provides an explanation of that and also explains why the exponential growth property is so generic and independent of the form of g_i. As remarked in Sec. V A, we find exponentially growing chemical trajectories (26) as attractors of the dynamics for a wide range of systems irrespective of the form of f_i when the f_i are class-I. In all these cases, f_i were derived from physically motivated g_i via Eqs. (33) and (2). Then, since the volume of the container is a linear function of the populations (2), it follows that the container size in the attractor also grows exponentially with the same rate: $V(t) \sim e^{\mu t}$. Thus, the ubiquity of exponential growth of V is a consequence of the facts the f_i are class I irrespective of the form of g_i and that such systems possess exponential solutions which are often the attractors of the dynamics.

This explanation also exposes a physical property that is required for exponential trajectories that has not been so far recognized, namely the linear dependence (2) of V on the populations [or, more generally, $V(\beta\mathbf{X}) = \beta V(\mathbf{X})$]. This suggests that the biophysical origins of the assumption (2) need to be explored further. In particular, the constants v_i, which also affect the steady-state growth rate of the cell (and

hence cellular fitness), should be determined experimentally and estimated theoretically. This also provides a possible explanation of the departure from exponential trajectories observed in certain eukaryotic unicellular organisms [4,13]. The departure may be in part a consequence of the violation of (2) caused by other structural features of the cell such as the cytoskeleton.

Expression for τ. Let us consider the consequences of this for the general model described in Sec. II. We assume, for the rest of this section and the next section, that (a) V is a linear function of the populations and the functions f_i satisfy the class-I property, and (b) that an exponential solution (26) is the attractor of the X-sector dynamics for the initial conditions of interest. In discussing averages, we ignore all forms of stochasticity and treat the dynamics as deterministic. Then, the X_i variables flow toward the CBG defined by the vector $\boldsymbol{\psi}$ whose components, the ratios of populations, as well as μ, satisfy (28). This flow is not affected by the change of Z because $\dot{X}_i$ is independent of Z, except at discrete points when all the X_i are halved. The latter interruption does not disturb the flow towards the CBG because $\dot{\psi}_i$ are unaffected by it [as mentioned above, (1a) and (3) imply that $\dot{\psi}_i = f_i(\boldsymbol{\psi}) - \psi_i f_n(\boldsymbol{\psi})$, which is invariant under change of scale of the X]. The dynamics of ratios of chemicals and their concentrations (31) therefore does not see any discontinuity at division. Thus, the X variables reach the CBG in the presence of the division process. This will be true as long as division changes all the X_i by the same scale (not necessarily half), and irrespective of whether division is triggered by the Z dynamics as in the present case or by some other process. Thus, in particular it is guaranteed that in the attractor of the growth-division dynamics, the cell volume and all populations X_i will grow exponentially with time between birth and division. Furthermore in the steady state, since the variables must grow by a factor of 2 between birth and division (assuming X_i are halved at division), the interdivision time is given by $\tau = \ln 2/\mu$. Note that τ only depends upon the parameters of the X sector and not the Z sector.

Expression for V_b. Now let us further assume that (c) the function $h(\mathbf{X}, Z)$ in Eq. (1b) is independent of Z and satisfies the class-I property, $h(\beta\mathbf{X}) = \beta h(\mathbf{X})$ [as is the case in the PTRZ model, Eq. (7)]. When the X have reached their CBG, we can write $X_i(t) = X_{ib}e^{\mu t} = X_{nb}\psi_i e^{\mu t}$ or $\mathbf{X}(t) = X_{nb}\boldsymbol{\psi} e^{\mu t}$, where X_{ib} is the value of X_i at a birth of the cell (taken to be $t = 0$ for convenience). Then $\dot{Z} = h(\mathbf{X}) = Ce^{\mu t}$, with the last step following from (c), and $C = X_{nb}h(\boldsymbol{\psi})$. This is the same as (12), except that the expression for C is now more general. The argument below (12) and in the Supplemental Material Sec. S2 A [51] applies as before, fixing the absolute scale of the X_{ib} and V_b. In particular, $X_{nb} = N\mu e^{\mu\tau_1}/h(\boldsymbol{\psi})$ and since $V = \sum_{i=1}^{n} v_i X_i = X_{nb} \sum_{1=1}^{n} v_i \psi_i$, we get

$$V_b = Ae^{\mu\tau_1} = \frac{N\mu}{h(\boldsymbol{\psi})} \sum_{i=1}^{n} v_i \psi_i e^{\mu\tau_1}. \tag{34}$$

This is the expression for cell size at birth in terms of the parameters for the general model described in Sec. II when the conditions (a)–(c) hold. Equations (15) and (16) for the PTRZ model are special cases of this formula. Note that μ

and ψ_i are determined in terms of the parameters appearing in the f_i through (28).

2. Balanced growth of chemicals

In order to replicate itself, a growing cell must solve a high-dimensional coordination problem between its chemicals. The amount of each chemical in the mother cell at division must be double that in the daughter cell at birth (assuming that upon division a daughter gets half of every chemical from its mother). Otherwise, daughters at birth in successive generations will not be identical. How does the cell manage to double all of its chemicals (and there are thousands of them) at the same time? This is a puzzle because each chemical is produced and consumed in a specific set of reactions that have their own specific rate constants, varying from reaction to reaction. However, if the exponential solution (26) is an attractor of the dynamics (1a), the problem is automatically solved because in the attractor each chemical has the same growth rate μ and ratios of chemicals are preserved in time. It is remarkable that class-I systems describing cells seem to have such attractors. This is not true for systems that are not class I (for an example, see Fig. S10), which must solve their coordination problems by other means (to be discussed elsewhere).

3. Genericity of the adder property and the Δ and τ distributions

So far in this subsection, we have discussed the consequences of the class-I property for the general system (1) at the deterministic level (without stochasticity). Now consider the inclusion of intrinsic stochasticity in the chemical dynamics of Z, while still treating the dynamics of the X-sector chemicals as deterministic and also ignoring all other sources of stochasticity such as partitioning stochasticity, threshold stochasticity, etc. The adder property and the Δ and τ distributions discussed in Sec. IV D are derived from three key ingredients: (i) $V(t) = V_b e^{\mu t}$, (ii) cell division occurs at the time when Z first reaches Z_c, whereupon it is reset to a fixed value, and (iii) the dynamics of Z is the stochastic version of (12) in which C is an extensive quantity proportional to the chemical populations and hence to the volume of the cell. In the class of models described in Sec. II, property (ii) is taken for granted (it is part of the definition). Property (i) follows from assumptions (a) and (b) mentioned above (in Sec. V C 1). As noted above, when assumption (c) holds, then on the CBG attractor of the X-sector deterministic dynamics, (12) also holds with C and V_b both proportional to X_{nb} and hence to each other. Thus, (iii) also holds provided we treat the Z dynamics as stochastic. This proves that under conditions (a)–(c) when we treat the X-sector dynamics as deterministic and the Z dynamics as stochastic the general model described in Sec. II displays the adder property and the Δ and τ distributions given in Sec. IV D [Eqs. (17) to (20)] arise with $C = h(\mathbf{X}_b)$ and $B = C/\mu$. Further, the populations $\mathbf{X}_b$ (rescaled by their means) also show the same distribution as V_b and exhibit the adder property.

When the X-sector dynamics are also treated stochastically, and other sources of stochasticity such as partitioning stochasticity and stochasticity of Z_c are included, we expect that the model will exhibit a behavior similar to that discussed for the PTRZ model in Sec. III B, e.g., a nontrivial α

distribution will arise, the adder property for the cell volume will continue to hold unless Z_r and Z_c are correlated, the CV of molecular populations as a function of the mean $\bar{X}$ will decline as $\bar{X}$ increases for small $\bar{X} \ll N$ and flatten out for $\bar{X} \gg N$, etc.

VI. UNIFIED DESCRIPTION OF BACTERIAL CELL SIZE BEHAVIOR AND GROWTH LAWS

So far we have discussed issues related to the size of bacterial cells in a growth culture. The models considered here have a modular character, in that there are two sectors, the Z sector which is concerned with the triggering of division and the X sector which contains all other chemicals in the cell. We have seen that certain properties of the models are independent of the details of the X sector, as long as certain broad conditions are met (class-I nature of dynamics, linear dependence of V on populations, etc.). The average cell volume at birth depends upon various cellular parameters (15), (16), and (34). It depends on the X-sector parameters through the growth rate μ, the ratios of chemicals ψ_i, and the coefficients v_i, and on the Z-sector parameter N/K_Z [or $N/h(\boldsymbol{\psi})$ in general]. It also has the exponential factor $e^{\mu \tau_1}$. The overall dependence of V on μ is complicated in light of the fact that the ratios ψ_i also depend upon the same parameters of the X sector that μ depends upon [through a solution of (28)] and this would be different for different models. On the other hand, the fluctuations in V and the populations X_i, in particular their CV, as well as the CV of τ, are governed by a single parameter N and are largely independent of the details of what happens in the X sector. Experiments seem to constrain the parameter N between 20 and 60, as discussed in Sec. III B 4. The properties of exponential growth and the adder property for V and X_i are also largely independent of the details of the X sector.

We now mention another aspect of bacterial growth physiology that pertains to the effect of the medium on the growth rate and composition of the cells, the bacterial growth laws. These are summarized in three empirically derived equations:

$$\mu = \mu_\infty \frac{[F]}{C_1 + [F]}, \tag{35a}$$

$$\Phi_R = \Phi_R^{\min} + \frac{\mu}{\kappa_t}, \tag{35b}$$

$$\Phi_R = \Phi_R^{\max} - \frac{\mu}{\kappa_n}. \tag{35c}$$

The first [65] describes the dependence of the steady-state growth rate on the concentration $[F]$ of a limiting food resource in the external medium. The second [40–43] describes how the fraction of total protein in the cell that is ribosomal protein, Φ_R, increases as μ is increased by improving the nutritional quality of the limiting food resource. The third [43] describes how Φ_R increases as μ is *decreased* by adding antibiotics to the medium that diminish the translational efficiency of ribosomes or by producing mutants that specifically target the translational efficiency. The six constants $\mu_\infty, C_1, \Phi_R^{\min}, \Phi_R^{\max}, \kappa_t$, and κ_n are phenomenological constants [43] whose values are extracted from experiment.

In Ref. [50], the above three laws were derived from the PTR model, which is the X sector of the PTRZ model discussed above. There are two classes of macromolecules represented in the PTR model, T and R, made up of m_T and m_R units of P respectively, and the ribosomal fraction can be defined as $\Phi_R = m_R R/(m_T T + m_R R) = m_R/(m_T \psi_T + m_R)$. To derive the growth laws (35), the expression (10c) for the steady-state growth rate was maximized with respect to f_R, keeping all other parameters fixed. This implements an underlying regulatory mechanism in the cell that regulates the fraction of ribosomes that are engaged in making ribosomal protein. The value of f_R that maximizes the right-hand side of (10c) was substituted in Eq. (10a) to get an expression for ψ_T (and hence Φ_R) and in Eq. (10c) to get an expression for μ. This yielded Eqs. (35) for the growth laws together with expressions for the six phenomenological constants in terms of the parameters of the PTR sector. For the parameter values given in Fig. 1, the predictions of the model agree with the experimental data up to factors of order unity.

It is clear that the above procedure and its results are unaffected by the presence of the Z sector in the PTRZ model. The modular structure of the PTRZ model and the fact that the division process does not cause any discontinuity in the dynamics of the intensive variables of the X sector ensures that the PTRZ model also reproduces the same equations for the bacterial growth laws. Setting f_R to a particular value (that optimizes μ) affects the values of μ and the ratios ψ_T and ψ_R, and hence the value of V_b in Eq. (16), but does not affect the existence of exponential growth. In particular, (12) still holds, albeit with values of C and μ given by the above procedure. Thus the consequences of the Z-sector dynamics are also unaffected. The PTRZ model thus provides a unified explanation of the bacterial growth laws together with fluctuations of size, interdivision time, growth rate, and intracellular molecular populations as well as the adder property.

The PTR sector can be augmented [63] by introducing other chemical species whose dynamics model the regulation of f_R mechanistically instead of using the optimization procedure. Standard regulatory mechanisms are consistent with the general form (30) and do not alter the class-I nature of the X sector. The feature of exponentially growing trajectories remains intact in such models. The results above about cell size and fluctuations will therefore also hold for such models.

VII. SUMMARY AND DISCUSSION

We have presented a class of mathematical models that explain a number of observed properties of bacterial cells and make testable predictions. The work also introduces concepts that may help further theoretical analysis and identify new experiments.

A. Summary of assumptions and results

The models assume that the cell can be described in terms of the intracellular chemical populations X_i, $i = 1, \ldots, n$, whose growth dynamics, at the deterministic level, is given by coupled ordinary differential equations (1a). The cell volume V is assumed to be a linear function of the chemical populations, Eq. (2). The models have further structure to

describe the control of cell division, but already, even without that structure, certain key properties of the cell, summarized in the next paragraph, are fixed by the X sector itself. These are governed by the character of the functions f_i appearing in Eq. (1a).

Class-I property: Exponential trajectories, growth rate, interdivision timescale, and intensive quantities. The models provide an explanation of the observation that cell size and intracellular molecular populations grow exponentially with time between birth and division in steady-state bacterial cultures. Exponential growth is shown (in Sec. V A) to be a consequence of the class-I property [Eq. (3)] of the functions f_i. This property in turn is a general consequence of mass action kinetics in expanding containers (see Sec. V B) when the cell volume V depends linearly on the populations [Eq. (2)] and could therefore hold for a large class of models applicable to bacteria. When the class-I property holds, we find that the populations are typically attracted to a curve of balanced growth (CBG) in which their ratios $\psi_i \equiv X_i/X_n$ are constant, and the populations themselves grow exponentially in time [Eq. (26)] with a constant growth rate μ. This explains the phenomenon of balanced growth [17] in bacterial cells. This also explains why V grows exponentially in the steady state. The ratios ψ_i and μ are determined in terms of the constant parameters of the model as solutions to a set of (typically algebraic) equations [Eqs. (28)], which constitute a nonlinear analog of the eigenvalue equation of a matrix. Hence, all the intensive quantities in the steady state including the concentrations $x_i \equiv X_i/V$ of the chemicals are fixed. The interdivision timescale τ is fixed by the reciprocal of μ. For example, if the division is symmetric (i.e., the two daughters are identical), then $e^{\mu\tau} = 2$ since all chemicals should double in quantity between birth and division in the steady state; hence, $\tau = \ln 2/\mu$. Explicit solutions for μ and ψ_i [Eq. (10)] are presented for a nonlinear model with $n = 3$ populations, the PTR model [50] defined by Eqs. (4). We also show examples of non-class-I models where the growth trajectory of the individual cell is not exponential (Sec. V A, Fig. S10).

Bacterial growth laws. The PTR model also reproduces the bacterial growth laws of composition [40–43] when regulation is implemented through an optimization procedure [50] or by introducing additional molecular species [63] while preserving its class-I property and exponential trajectories (Sec. VI).

Division control: Absolute size and populations. While intensive quantities and the interdivision timescale in the steady state are fixed by the previous assumptions, the absolute scale of cell size and populations requires a specification of the division control mechanism. In this work, we assume that the cell commits itself to division when one of the chemical populations (denoted by Z and separated out from the X_i for convenience) reaches a threshold Z_c, with division following commitment after a possible delay τ_1. The dynamics of Z depends upon the X_i through (1b). For simplicity, we assume that Z contributes negligibly to its own dynamics, to the continuous time dynamics of the X_i, and to V. Immediately after it reaches its threshold Z_c, we allow Z to be reset to a value $Z_r \leqslant Z_c$. Upon division, all the populations are halved. With these assumptions, division is governed by the nature of the function h appearing in Eq. (1b). In this paper, we have primarily considered the consequences of h also satisfying

the class-I property like the f_i [Eq. (3)]. Then we can get a general formula for the birth volume V_b [Eq. (34)] in terms of h and a specific formula in the case of the PTRZ model [Eq. (16)] when a specific form [given by (7)] of h is chosen. The absolute populations are also fixed [expressions given above (34) and (15)].

Stochasticity in chemical dynamics of Z: Analytic distributions of τ and Δ; the adder property for cell size and chemical populations. The preceding three paragraphs summarize results for deterministic versions of the models and therefore pertain to averages across cells in the cultures. In order to understand cell-to-cell variation in steady-state cultures, we include various sources of stochasticity. Our stochastic results are obtained for the case $\tau_1 = 0$. It is convenient to distinguish four sources of stochasticity that we have considered (which have distinct physical consequences): (A) Intrinsic stochasticity in the chemical dynamics of Z, (B) intrinsic stochasticity in the chemical dynamics of the X-sector chemicals, (C) partitioning stochasticity, and (D) stochasticity in the value of the threshold Z_c. Intrinsic stochasticity results from the fact that each chemical reaction is a molecular event with a certain probability. In particular, this makes the interdivision time, which by assumption is the first passage time for Z to reach Z_c, a stochastic quantity. Our implementation of (C) and (D) is defined in Sec. III B 5. We have investigated different types of stochasticity in isolation and in suitable combinations. When type A is the only stochasticity present, we obtain some analytic results for general class-I systems, namely, the distributions of τ (17), and of added volume Δ (18), both conditional on the birth configuration of the cell. The conditional Δ distribution is independent of V_b, thereby proving the adder property. In fact, the rescaled Δ, $u \equiv \Delta/\langle\Delta\rangle$, has a distribution (20) that depends on only one parameter, $N = Z_c - Z_r/2$, and is independent of the details of the functions f_i, h and other parameters (Sec. V C 3; Fig. S1). Molecular populations also satisfy the adder property and their rescaled increments have the same distribution as u. However, the distribution of growth rate α has zero width when only type A stochasticity is present (discussed in the beginning of Sec. IV D).

Other sources of stochasticity: Robustness of size and τ distributions; crossover of population distributions; origin of the growth rate distribution; departure from the adder property. We have studied stochasticities of types B, C, and D with numerical simulations of the nonlinear PTRZ model [defined by Eqs. (4), (5), and (7)] and two linear models, the XZ and XYZ models [Eq. (25)]. We find, in the parameter ranges considered, that the rescaled distributions of Δ, τ, V_d, and V_b obtained with purely A-type stochasticity are robust to the inclusion of B-type stochasticity (Figs. S1 and 7) and also to the inclusion of the C- and D-type stochasticities (Sec. III B 5; Figs. S5), provided their strength is not too large. The robustness of the rescaled size and interdivision time distributions to parameter values (except N) and other sources of stochasticity is one of the striking results of this work. The CV of all these distributions is proportional $1/\sqrt{N}$ when only the A- and B-type stochasticities are included [Fig. 2(c)], with proportionality constants given by Eq. (8). The rescaled distributions of the populations X_i are also robust, provided the population mean $\langle X_i \rangle$ is not too small. We find a crossover behavior of the

CV of X_i as a function of $\langle X_i \rangle$ (Fig. 5), with CV $\sim \langle X_i \rangle^{-1/2}$ for $\langle X_i \rangle \ll N$ and CV constant for $\langle X_i \rangle \gg N$. The adder property of the volume and populations is also found to be robust to the introduction of the BCD-type stochasticities (Figs. 4, S4, S5D, and S6A). However, the adder property is lost with D-type stochasticity if Z_c and Z_r are correlated with each other (Fig. S6B). The distribution of growth rate α acquires a nonzero width upon the introduction of type-B stochasticity (Fig. 6) depends upon parameters other than N (Fig. S1) and is strongly influenced by type-C stochasticity (Fig. 6). We show that the origin of the α distribution is in part a consequence of the fact that type-B and type-C stochasticities throw the populations off the CBG and thereby cause a mixing of a pure exponential function of time (having a rate μ) with other functions of time (including exponentials with rates other than μ).

Comparison with experiment. At the level of averages, the models reproduce the exponential trajectory of the volume and intracellular populations as seen in *E. coli* experiments, as well as the bacterial growth laws of composition. At the level of fluctuations, they reproduce the adder property of cell volume fluctuations in *E. coli*. Experimental distributions of Δ, τ, V_d [11] constrain the parameter N of the models to be between 20 and 60 (Sec. III B 4). The models find a much narrower distribution of α than the one observed in Ref. [11] when only the intrinsic stochasticity of chemical dynamics is included and require a significant strength (perhaps unnaturally large) of partitioning stochasticity to reach the observed width (Sec. III B 5). This discrepancy needs to be explored further, both theoretically and experimentally. The dependence of the α distribution in our model on various parameters (strength of partitioning stochasticity and other parameters indicated in Fig. S1C) might be experimentally testable. It would be interesting to explore the relationship between the mechanisms contributing to the α distribution in the present model to those discussed in other works [66,67]. Another task for the future is comparing the observed correlations between α and the other cellular variables [68] with the predictions from our models. The models reproduce the observed crossover behavior [30] of the CV of a molecular population as a function of its mean population $\bar{X}$: CV $\sim \bar{X}^{-1/2}$ for $\bar{X} \ll N$ and constant for $\bar{X} \gg N$ (Sec. III B 3).

B. Discussion

A key question is the following: Which molecular population in the cell does Z correspond to? One possible candidate is the DnaA molecule, which is known to initiate deoxyribonucleic acid (DNA) replication and has been suggested as an upstream trigger for cell division (for reviews, see Refs. [69–72]). DNA replication in *E. coli* is known to be initiated when a certain number, believed to be between 20 and 30, of active [attached to adenosine triphosphate (ATP)] molecules of DnaA bind to sites on the DNA molecule at *oriC*, a region of DNA at the origin of replication. Soon after the initiation of replication, the DnaA is deactivated by other enzymes; the complex of active DnaA falls apart to prevent multiple rounds of initiation. Each initiation is followed first by the replication of DNA (referred to as the C period in the bacterial growth cycle) and then by the separation of the chromosomes into two halves of the cell and cell division (referred to as the D period). After division, the daughter cells have a smaller number of active DnaA molecules bound to the above-mentioned sites, and this number grows in the period between birth and the initiation of replication (known as the B period). Various aspects of the dynamics of DnaA have been modeled mathematically [71,73–76].

In the context of the present model, it is tempting to identify Z with the number of active DnaA molecules in the initiation complex bound to *oriC*. Then present experiments with DnaA suggest that the parameter Z_c of the model should be between 20 and 30, Z_r should be essentially zero (the initiation complex dissociates after triggering replication), and $\tau_1 = C + D \sim 1$ h. It is interesting that this identification leads to a value of N between 20 and 30, which overlaps with the range 20–60 obtained from a completely independent experimental constraint, the spread in the distributions of Δ, τ, and V_d. In other words, this identification provides a natural explanation of why the CVs of cell size, interdivision time, and the large intracellular molecular populations are in the ballpark of 20% [Eq. (8) with $N \sim 25$]. The model predicts that the average cell size increases when the rate of production of the Z population is lowered [V_b is inversely proportional to K_Z or $h(\boldsymbol{\psi})$, Eqs. (15) and (34)]. This is consistent with the empirical observation that cell size increases when the production of DnaA is impaired [77]. We note that the fact that cooperativity in the Z dynamics leaves the adder property of the volume intact (shown in Sec. IV D 2 and Supplemental Material Sec. S2 C [51]) is an encouraging sign for the above interpretation of Z, as such cooperativity is known to exist for the active-DnaA molecules bound at *oriC*.

However, the model also has problems with respect to the above interpretation. Note that in the previous paragraph, Z is identified not with the total number of DnaA protein molecules in the cell, or even the total number of active (ATP-bound) DnaA molecules in the cell, but with the number of active DnaA molecules *bound to the DNA molecule at oriC*. A problem with the present model is that a production term for Z such as in Eq. (7), where $\dot{Z}$ is an extensive quantity, may be appropriate for a chemical produced in the bulk of the cell, but a chemical species localized in space (such as the set of active-DnaA molecules bound to the DNA at *oriC*) would need further justification or a mechanism not provided by the present model. Further, in this interpretation, it is not clear how to account for the synchrony observed in the initiation of replication at multiple origins [78]. The resolution of the above-mentioned problems might lie in the fact that *oriC* and the *dnaA* gene are close by on the DNA molecule and spatial proximity effects need to be taken into account. It also needs to be noted that the dynamics of active-DnaA molecules bound to the origin of DNA replication is affected by many factors in the cell, including autoregulation of DnaA production, binding of DnaA to a large number of sites other than *oriC* on the DNA, and sequestering of the binding sites by other enzymes. Another proposed model [79] is that initially in the cell cycle DnaA binds other sites on the DNA and the replication-initiating event is the binding of active DnaA to a site in *oriC* that triggers cooperative binding of active DnaA on a relatively short timescale to *oriC*. In this interpretation, Z_c would correspond to the effective number of sites that DnaA

binds to before the cooperative binding event occurs. In the light of the above caveats, the question as to whether the Z population in the present model or its extension corresponds to some subpopulation of the DnaA protein in the cell remains an open question to be investigated further.

Note that the adder property and the Δ distribution (18) [and consequently Eq. (8) relating the CV of Δ to $1/\sqrt{N}$] follow from very general assumptions [Eq. (21) and assumptions mentioned below it; or the class-I property and assumptions mentioned in Sec. V C 3]. They and many other properties discussed above do not depend upon other details of the Z dynamics. Thus, the models could apply to other candidates for the Z molecule than DnaA. Another candidate for Z is the FtsZ protein, which forms a ring around the cell that constricts and causes cell division and degrades after cell division (for reviews, see Refs. [80,81]). However, the number of FtsZ molecules required for ring formation seems to be large (in the range $\approx 10^3$–10^4, suggesting that N is in the range $\approx 10^3$–10^4); in which case stochasticity in the production of FtsZ alone would not account for the observed CV of cell size and division time (the model predicts CV $\sim 1/\sqrt{N}$, which would be too low). The present model ignores stochasticity in the mRNA population corresponding to the Z protein; its inclusion would enhance the CV of the cellular variables in Eq. (8) beyond $\frac{1}{\sqrt{N}}$ and in particular improve the case for FtsZ. Further, in the cell, stochasticity in both DnaA and FtsZ can simultaneously contribute to the stochasticity in interdivision time and cell size.

In sum, while there are two good candidates (DnaA and FtsZ) for the time-keeper molecule, we believe that more work, both experimental and theoretical, is needed to establish whether the observed stochasticity in the interdivision time and cell size is a consequence of their populations (or appropriate subpopulations) having to reach a critical threshold. Of course, there may be other molecular candidates in the cell for triggering division that we are not aware of.

The models discussed in the present work predict that the adder property of the cell volume is accompanied by the adder property for coarse-grained intracellular molecular pools, such as the pool of amino acids in the cell or the pool of metabolic enzymes, whose copy numbers are $\gg N$ and that the distribution of such a molecular population rescaled by its mean is the same as the rescaled volume distribution. Chemical population distributions have been measured and found to be universal [31] but the adder property for molecular pools has not been reported to our knowledge. Susman *et al.* [62] have reported a departure from the adder property for two individual protein populations in single-cell trajectories of *E. coli* but the adder behavior of larger molecular pools (predicted by our models) remains an open empirical question. The models also predict that in the presence of threshold stochasticity (Z_c being distributed over a range of values from cell to cell) the adder property is lost if the reset value Z_r is correlated with Z_c. It would be interesting if this could be tested by suitable mutations of *E. coli*, or by studying bacteria in which the adder property is absent.

An experimental question that this work draws attention to is the role of osmosis and other biophysical mechanisms in understanding exponential trajectories of cells. In Sec. V B,

we argued that the linear dependence of V on the intracellular populations, Eq. (2), is a crucial requirement for exponential trajectories. As mentioned in that section such an assumption might be justified if cells equalize the osmotic pressure of solute inside and outside the cell. Cells actually maintain an osmotic pressure difference between their interior and the exterior. It is important to test whether the linearity assumption is valid and also to measure the coefficients v_i. It is worth noting that these coefficients affect the steady-state growth rate μ [see, e.g., Eq. (10c) and the expressions below it] and thus contribute to cellular fitness. A departure from the linear dependence of V on intracellular molecular populations, caused by other structural elements such as the cytoskeleton, may explain the departure from exponential trajectories observed in certain eukaryotic organisms.

At a mathematical level, this work suggests that a certain class of dynamical systems, class-I systems [defined by (1a) and (3)], are both generic and analytically useful for modeling bacterial cells. The present study shows that some physical properties (summarized above, and including averages and fluctuations) are universal for all such systems in that they do not depend upon the model and (many of) its parameters, and it identifies some that are not. For this class, the steady states are exponentially growing trajectories whose growth rate μ is a solution of a nonlinear version of the eigenvalue equation, Eq. (28). When the functions f_i are algebraic μ is given implicitly in terms of the model parameters as a solution to a set of algebraic equations. In this work, we have only investigated some properties in a few examples of class-I systems; other mathematical properties and examples need to be investigated. Since exponential growth appears in many areas where the underlying dynamics is nonlinear, it is quite possible that these systems find applications elsewhere.

For any fixed environment, μ is a measure of the organism's fitness. Having it as a function of the system parameters allows us to describe the fitness landscape including the neutral directions, valleys, and hills in parameter space. Thus, it specifies the evolutionary paths in that environment. This class of models could be relevant for the study of evolution because they provide μ as a function of the parameters. It is worth mentioning that for the class of models studied here, μ is independent of the parameters of the division control sector [e.g., the function h in Eq. (1b), and the thresholds Z_c and Z_r], which are therefore neutral directions of variation as far as steady-state fitness is concerned. This is a consequence of a kind of modularity in the model implicit in the assumption that the dynamics of the X sector is independent of Z and is influenced by Z only through division. It would be interesting to empirically explore the extent of, or departure from, such a modularity in real cells.

Since the models discussed here are dynamical, they capture not just the steady state but also the transients. Thus, they could also be useful in exploring those regulatory mechanisms that affect or seek to optimize performance over the transients. Extension of these models may find applications in modeling the stationary phase [82] and antibiotic environments [83] where the net cell population growth goes to zero with a balance between cell growth and death. It would also be interesting to use these models to make contact with other allometric modeling approaches [84] that seek to understand

how the amounts of cellular components like ribosomes and proteins depend upon cell size across diverse bacteria.

VIII. COMPUTATIONAL METHODS

All numerical simulations were done in C programming language. Numerical solution of the ODEs for the mathematical models were done using the CVODE solver library of the SUNDIALS (Suite of Nonlinear and Differential/Algebraic Equation Solvers) package [85] and the adaptive Runge-Kutta (RK5) method. The stochastic simulations were done using the τ leaping method [86,87]. In particular, the trapezoidal variant of adaptive implicit *tau* leaping was used [88,89].

ACKNOWLEDGMENTS

[1] E. F. Adolph and S. Bayne-Jones, Growth in size of microorganisms measured from motion pictures II: *Bacillus megatherium*, J. Cell. Comp. Physiol. **1**, 409 (1932).

[2] M. Schaechter, J. P. Williamson, J. R. Hood Jr., and A. L. Koch, Growth, cell, and nuclear divisions in some bacteria, J. Gen. Microbiol. **29**, 421 (1962).

[3] S. Cooper, Distinguishing between linear and exponential cell growth during the division cycle: Single-cell studies, cell-culture studies, and the object of cell-cycle research, Theor. Biol. Med. Model. **3**, 10 (2006).

[4] J. Mitchison, Growth during the cell cycle, Int. Rev. Cytol. **226**, 165 (2003).

[5] M. Godin, F. F. Delgado, S. Son, W. H. Grover, A. K. Bryan, A. Tzur, P. Jorgensen, K. Payer, A. D. Grossman, M. W. Kirschner, and S. R. Manalis, Using buoyant mass to measure the growth of single cells, Nat. Methods **7**, 387 (2010).

[6] 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).

[7] M. Mir, Z. Wang, Z. Shen, M. Bednarz, R. Bashira, I. Golding, S. G. Prasanth, and G. Popescu, Optical measurement of cycle-dependent cell growth, Proc. Natl. Acad. Sci. USA **108**, 13124 (2011).

[8] M. Osella, E. Nugent, and M. C. Lagomarsino, Concerted control of *Escherichia coli* cell division, Proc. Natl. Acad. Sci. USA **111**, 3431 (2014).

[9] S. Iyer-Biswas, C. S. Wright, J. T. Henry, K. Lo, S. Burov, Y. Lin, G. E. Crooks, S. Crosson, A. R. Dinner, and N. F. Scherer, Scaling laws governing stochastic growth and division of single bacterial cells, Proc. Natl. Acad. Sci. USA **111**, 15912 (2014).

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

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

[12] N. Brenner, E. Braun, A. Yoney, L. Susman, J. Rotella, and H. Salman, Single-cell protein dynamics reproduce universal fluctuations in cell populations, Eur. Phys. J. E **38**, 1 (2015).

[13] J.-B. Nobs and S. J. Maerkl, Long-term single cell analysis of *S. pombe* on a microfluidic microchemostat array, PloS One **9**, e93466 (2014).

[14] C. N. Hinshelwood, On the chemical kinetics of autosynthetic systems, J. Chem. Soc., 745 (1952).

[15] S. Jain and S. Krishna, Autocatalytic Sets and the Growth of Complexity in an Evolutionary Model, Phys. Rev. Lett. **81**, 5684 (1998).

[16] S. Iyer-Biswas, G. E. Crooks, N. F. Scherer, and A. R. Dinner, Universality in Stochastic Exponential Growth, Phys. Rev. Lett. **113**, 028101 (2014).

[17] A. Campbell, Synchronization of cell division, Bacteriol. Rev. **21**, 263 (1957).

[18] T. Ganti, *Chemoton Theory*, Vol. 1 Theoretical Foundations of Fluid Machineries (Kluwer Academic/Plenum Publishers, New York, 2003).

[19] R. Serra, T. Carletti, and I. Poli, Synchronization phenomena in surface-reaction models of protocells, Artificial Life **13**, 123 (2007).

[20] T. Carletti, R. Serra, I. Poli, M. Villani, and A. Filisetti, Sufficient conditions for emergent synchronization in protocell models, J. Theor. Biol. **254**, 741 (2008).

[21] F. Mavelli and K. Ruiz-Mirazo, Stochastic simulations of minimal self-reproducing cellular systems, Philos. Trans. R. Soc. B **362**, 1789 (2007).

[22] F. Mavelli and K. Ruiz-Mirazo, Theoretical conditions for the stationary reproduction of model protocells, Integr. Biol. **5**, 324 (2013).

[23] R. Serra and M. Villani, *Modelling Protocells: The Emergent Synchronization of Reproduction and Molecular Replication* (Springer, Berlin, 2017).

[24] M. Osella, S. J. Tans, and M. C. Lagomarsino, Step by step, cell by cell: Quantification of the bacterial cell cycle, Trends Microbiol. **25**, 250 (2017).

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

[26] 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).

[27] P.-Y. Ho, J. Lin, and A. Amir, Modeling cell size regulation: From single-cell-level statistics to molecular mechanisms and population-level effects, Annu. Rev. Biophys. **47**, 251 (2018).

[28] A. S. Kennard, M. Osella, A. Javer, J. Grilli, P. Nghe, S. J. Tans, P. Cicuta, and M. Cosentino Lagomarsino, Individuality and universality in the growth-division laws of single *E. coli* cells, Phys. Rev. E **93**, 012408 (2016).

[29] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Stochastic gene expression in a single cell, Science **297**, 1183 (2002).

[30] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells, Science **329**, 533 (2010).

[31] H. Salman, N. Brenner, C.-k. Tung, N. Elyahu, E. Stolovicki, L. Moore, A. Libchaber, and E. Braun, Universal Protein Fluctuations in Populations of Microorganisms, Phys. Rev. Lett. **108**, 238105 (2012).

[32] L. Sompayrac and O. Maaløe, Autorepressor model for control of DNA replication, Nat. New Biol. **241**, 133 (1973).

[33] W. Voorn, L. Koppes, and N. Grover, Mathematics of cell division in *Escherichia coli*: Comparison between sloppy-size and incremental-size kinetics, Curr. Top. Mol. Gen. **1**, 187 (1993).

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

[35] P.-Y. Ho and A. Amir, Simultaneous regulation of cell size and chromosome replication in bacteria, Front. Microbiol. **6**, 662 (2015).

[36] K. R. Ghusinga, C. A. Vargas-Garcia, and A. Singh, A mechanistic stochastic framework for regulating bacterial cell division, Sci. Rep. **6**, 30229 (2016).

[37] J. Lin and A. Amir, Homeostasis of protein and mRNA concentrations in growing cells, Nat. Commun. **9**, 4496 (2018).

[38] D. J. Kiviet, P. Nghe, N. Walker, S. Boulineau, V. Sunderlikova, and S. J. Tans, Stochasticity of metabolism and growth at the single-cell level, Nature (London) **514**, 376 (2014).

[39] M. Wallden, D. Fange, E. G. Lundius, Ö. Baltekin, and J. Elf, The synchronization of replication and division cycles in individual *E. coli* cells, Cell **166**, 729 (2016).

[40] M. Schaechter, O. Maaloe, and N. Kjeldgaard, Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella Typhimurium*, J. Gen. Microbiol. **19**, 592 (1958).

[41] O. Maaløe, Regulation of the protein-synthesizing machinery: Ribosomes, tRNA, factors, and so on, in *Biological Regulation and Development*, edited by R. F. Goldberger (Plenum Press, New York, 1979), pp. 487–542.

[42] H. Bremer and P. P. Dennis, Modulation of chemical composition and other parameters of the cell by growth rate, in *Escherichia coli and Salmonella*, edited by F. C. Neidhardt (ASM Press, Washington, D.C., 1996), pp. 1553–1569.

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

[44] D. Molenaar, R. van Berlo, D. de Ridder, and B. Teusink, Shifts in growth strategies reflect tradeoffs in cellular economics, Mol. Syst. Biol. **5**, 323 (2009).

[45] M. Scott, S. Klumpp, E. M. Mateescu, and T. Hwa, Emergence of robust growth laws from optimal regulation of ribosome synthesis, Mol. Syst. Biol. **10**, 747 (2014).

[46] A. Maitra and K. A. Dill, Bacterial growth laws reflect the evolutionary importance of energy efficiency, Proc. Natl. Acad. Sci. USA **112**, 406 (2015).

[47] A. Y. Weiße, D. A. Oyarzún, V. Danos, and P. S. Swain, Mechanistic links between cellular trade-offs, gene expression, and growth, Proc. Natl. Acad. Sci. USA **112**, E1038 (2015).

[48] E. Bosdriesz, D. Molenaar, B. Teusink, and F. J. Bruggeman, How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization, FEBS J. **282**, 2029 (2015).

[49] N. Giordano, F. Mairet, J.-L. Gouzé, J. Geiselmann, and H. de Jong, Dynamical allocation of cellular resources as an optimal control problem: Novel insights into microbial growth strategies, PLoS Comput. Biol. **12**, e1004802 (2016).

[50] P. P. Pandey and S. Jain, Analytic derivation of bacterial growth laws from a simple model of intracellular chemical dynamics, Theor. Biosci. **135**, 121 (2016).

[51] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.101.062406 which has two sections, the first containing ten figures (S1-S10) and the second containing four subsections of text (A-D).

[52] F. Bertaux, J. Von Kügelgen, S. Marguerat, and V. Shahrezaei, A unified coarse-grained theory of bacterial physiology explains the relationship between cell size, growth rate and pro-teome composition under various growth limitations, https://www.biorxiv.org/content/10.1101/078998v1.

[53] R. Milo, P. Jorgensen, U. Moran, G. Weber, and M. Springer, Bionumbers: The database of key numbers in molecular and cell biology, Nucl. Acids Res. **38**, D750 (2010).

[54] B. D. Bennett, E. H. Kimball, M. Gao, R. Osterhout, S. J. Van Dien, and J. D. Rabinowitz, Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*, Nat. Chem. Biol. **5**, 593 (2009).

[55] M. Thattai and A. Van Oudenaarden, Intrinsic noise in gene regulatory networks, Proc. Natl. Acad. Sci. USA **98**, 8614 (2001).

[56] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, Proc. Natl. Acad. Sci. USA **99**, 12795 (2002).

[57] M. Soltani, C. A. Vargas-Garcia, D. Antunes, and A. Singh, Intercellular variability in protein levels from stochastic expression and noisy cell cycle processes, PLoS Comput. Biol. **12**, e1004972 (2016).

[58] D. Huh and J. Paulsson, Non-genetic heterogeneity from stochastic partitioning at cell division, Nat. Genet. **43**, 95 (2011).

[59] W. D. Donachie, Relationship between cell size and time of initiation of DNA replication, Nature (London) **219**, 1077 (1968).

[60] S. Cooper and C. E. Helmstetter, Chromosome replication and the division cycle of *Escherichia coli* Br, J. Mol. Biol. **31**, 519 (1968).

[61] F. Si, D. Li, J. T. Sauls, O. Azizi, C. Sou, A. B. Schwartz, M. J. Erickstad, Y. Jun, X. Li, and S. Jun, Invariance of initiation mass

and predictability of cell size in *Escherichia coli*, Curr. Biol. **27**, 1278 (2017).

[62] L. Susman, M. Kohram, H. Vashistha, J. T. Nechleba, H. Salman, and N. Brenner, Individuality and slow dynamics in bacterial growth homeostasis, Proc. Natl. Acad. Sci. USA **115**, E5679 (2018).

[63] P. Sharma, P. P. Pandey, and S. Jain, Modeling the cost and benefit of proteome regulation in a growing bacterial cell, Phys. Biol. **15**, 046005 (2018).

[64] C. Furusawa and K. Kaneko, Zipf's Law in Gene Expression, Phys. Rev. Lett. **90**, 088102 (2003).

[65] J. Monod, The growth of bacterial cultures, Annu Rev. Microbiol. **3**, 371 (1949).

[66] R. Pugatch, Greedy scheduling of cellular self-replication leads to optimal doubling times with a log-Frechet distribution, Proc. Natl. Acad. Sci. USA **112**, 2611 (2015).

[67] D. De Martino, F. Capuani, and D. M. Andrea, Growth against entropy in bacterial metabolism: The phenotypic trade-off behind empirical growth rate distributions in *E. coli*, Phys. Biol. **13**, 036005 (2016).

[68] J. Grilli, C. Cadart, G. Micali, O. Matteo, and M. C. Lagomarsino, The empirical fluctuation pattern of *E. coli* division control, Front. Microbiol. **9**, 1451 (2018).

[69] A. Kornberg and T. A. Baker, *DNA Replication* (W. H. Freeman, New York, 1992).

[70] W. Messer, The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication, FEMS Microbiol. Rev. **26**, 355 (2002).

[71] W. D. Donachie and G. W. Blakely, Coupling the initiation of chromosome replication to cell size in *Escherichia coli*, Curr. Opin. Microbiol. **6**, 146 (2003).

[72] K. Skarstad and T. Katayama, Regulating DNA replication in bacteria, Cold Spring Harbor Perspect. Biol. **5**, a012922 (2013).

[73] S. T. Browning, M. Castellanos, and M. L. Shuler, Robust control of initiation of prokaryotic chhromosome replication: Essential considerations for a minimal cell, Biotechnol. Bioeng. **88**, 575 (2004).

[74] J. Atlas, E. Nikolaev, S. Browning, and M. Shuler, Incorporating genome-wide DNA sequence information into a dynamic whole-cell model of *Escherichia coli*: Application to DNA replication, IET Syst. Biol. **2**, 369 (2008).

[75] M. A. Grant, C. Saggioro, U. Ferrari, B. Bassetti, B. Sclavi, and M. C. Lagomarsino, DnaA and the timing of chromosome replication in *Escherichia coli* as a function of growth rate, BMC Syst. Biol. **5**, 201 (2011).

[76] J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V. Gutschow, J. M. Jacobs, B. Bolival, N. Assad-Garcia, J. I. Glass, and M. W. Covert, A whole-cell computational model predicts phenotype from genotype, Cell **150**, 389 (2012).

[77] A. Løbner-Olesen, K. Skarstad, F. G. Hansen, K. von Meyenburg, and E. Boye, The DnaA protein determines the initiation mass of *Escherichia coli* k-12, Cell **57**, 881 (1989).

[78] S. Kirsten, E. Boye, and T. Altung, Timing of initiation of chromosome replication in individual *Escherichia coli* cells, EMBO J. **5**, 1711 (1986).

[79] F. G. Hansen, B. B. Christensen, and T. Altung, The intiator titration model: Computer simulation of chromosome and mini-chromosome control, Res. Microbiol. **142**, 161 (1991).

[80] W. Margolin, FtsZ and the division of prokaryotic cells and organelles, Nat. Rev. Mol. Cell Biol. **6**, 862 (2005).

[81] H. P. Erickson, D. E. Anderson, and M. Osawa, FtsZ in bacterial cytokinesis: Cytoskeleton and force generator all in one, Microbiol. Mol. Biol. Rev. **74**, 504 (2010).

[82] Y. Himeoka and K. Kaneko, Theory for Transitions between Exponential and Stationary Phases: Universal Laws for Lag Time, Phys. Rev. X **7**, 021049, (2017).

[83] S. Chib, S. Das, S. Venkatesan, A. S. N. Seshasayee, and M. Thattai, Using stochastic cell division and death to probe minimal units of cellular replication, New J. Phys. **20**, 035004 (2018).

[84] C. P. Kempes, L. Wang, J. P. Amend, J. Doyle, and T. Hoehler, Evolutionary tradeoffs in cellular composition across diverse bacteria, ISME J. **10**, 2145 (2016).

[85] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward, Sundials: Suite of nonlinear and differential/algebraic equation solvers, ACM T. Math. Software **31**, 363 (2005).

[86] D. T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems, J. Chem. Phys. **115**, 1716 (2001).

[87] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, Stiffness in stochastic chemically reacting systems: The implicit τ-leaping method, J. Chem. Phys. **119**, 12784 (2003).

[88] Y. Cao and L. Petzold, Trapezoidal tau-leaping formula for the stochastic simulation of biochemical systems, *Proceedings of International Conference on Foundations of Systems Biology in Engineering (FOSBE 2005)*, Santa Barbara, California, August 7-10, 2005, pp. 149–152.

[89] Y. Cao, D. T. Gillespie, and L. R. Petzold, Efficient step size selection for the τ-leaping simulation method, J. Chem. Phys. **124**, 044109 (2006).