

Continuous Rate Modelling of bacterial stochastic size dynamics.

Cesar Nieto¹, Cesar Vargas-Garcia², and Juan Pedraza^{1,†}

¹Universidad de los Andes

²Corporacion Colombiana de Investigacion Agropecuaria

†Correspondence: jmpedraza@uniandes.edu.co

Abstract

Bacterial division is an inherently stochastic process. However, theoretical tools to simulate and study the stochastic transient dynamics of cell-size are scarce. Here, we present a general theoretical approach based on the Chapman-Kolmogorov formalism to describe these stochastic dynamics including continuous growth and division events as jump processes. Using this approach, we analyze the effect of different sources of noise on the dynamics of the size distribution. Oscillations in the distribution central moments were found as consequence of the discrete translation invariance of the system with period of one doubling time, these oscillations are found in both the central moments of the size distribution and the auto-correlation function and do not disappear including stochasticity on division times or size heterogeneity on the population but only after include noise in either growth rate or septum position.

Introduction

Recent experiments involving time-lapse microscopy [1], single-cell tracking [2, 3], and gene tagging [4] have revealed how cell size stochasticity, and division events play an important role in the random fluctuations of bio-molecular concentrations [5–8]. This with important consequences for phenotype variability and cell heterogeneity over a clonal population of microorganisms [9].

Methods approaching the bacterial size control include the discrete stochastic maps (DSMs) [10]. These models define the known as *division strategy* as a map that takes cell size at birth s_b to a targeted cell size at division s_d trough a deterministic function $s_d = f(s_b)$ plus stochastic fluctuations that have to be fitted from experiment. DSMs, however, are unable to reproduce cell size transient dynamics at arbitrary infinitesimal time intervals without further extension.

To solve this continuous dynamics of the distributions describing stochastic processes usually the Chapman-Kolmogorov formalism (CK) is used [11]. CK solutions corresponds to the distributions of all the possible stochastic hybrid trajectories at a given time. Among the processes that can be modeled by CK it can be included continuous size growth and division as a jump process requiring the definition of a continuous rate of division. In fact, these models are also known as Continuous Rate Models (CRM) [10]. Recently, [12] proposed a power-law function to explain observations in *E. coli* bacteria cells. Instead, [13] suggested a convoluted function of size and cycle progression is required. [14] proposed a deconvoluted version by introducing division as a multistep process where the occurrence rate of these steps is a function of the size. Despite these recent attempts,

there is a lack of a complete formalism describing the phenomena related to bacterial division.

Here, we propose the CK formalism as a framework for studying single-cell transient dynamics. We present how to overcome the non-locality of the division jumps and how to model the division steps. These steps being analogous to the experimental accumulation of FtsZ to trigger the division [15, 16]. These equations are solved using both simulations and numerical methods. Analytical expressions are also presented in the cases where it was possible. We also present modifications to the division rates to define multiple division strategies. Then, we will show how stochasticity on division influence the size distribution dynamics and how this dynamics changes considering additional sources of noise like a distribution of initial sizes with finite variance, the noise in septum position and in cell-to-cell growth rate. We discuss how this approach could be coupled to simulations of gene expression.

Theoretical details

The population balance Equation

Consider the distribution $p(s; t)$ of sizes s at a given time t solving the Chapman-Kolmogorov equations (CKE). This distribution is related to the histograms of bacterial populations. Some attempts have described the dynamics of this distributions including effects due to the increasing of population number and the mother-daughter correlations [17–19]. In our framework, we consider a cell population with an asymptotically large number. After each division, only one of the descendants is tracked such as the number of bacteria in the population remains constant in similar way observed in experiments such as those using *Mother Machine* micro-fluids [3]. Then, $p(s; t)$ can be normalized:

$$\int p(s, t) ds = 1 \quad (1)$$

To describe the distribution dynamics, let an individual cell grow in size s as per

$$ds = g(s, t) dt, \quad s(0) = s_0, \quad (2)$$

where $g(s, t)$ is the size change per unit time. Some studies have considered the constant growth [17, 20, 21] we will focus on exponential growth rate ($g(s) = \mu s$) with μ being the growth rate and s_0 is the initial cell size. Since (2) defines a deterministic process, the change in the size distribution $p(s, t|s_0)$ conditioned on initial size s_0 can be obtained by solving

$$\frac{\partial}{\partial t} p(s, t|s_0) = -\frac{\partial}{\partial s} [g(s, t)p(s, t|s_0)]. \quad (3)$$

Expression (3) is known as the CKE in its differential version (dCKE), and its solution for this deterministic process is given by

$$p(s, t|s_0) = \delta(s - s(t)), \quad (4)$$

where $s(t)$ is the solution of equation (2).

Considering division as a jump process switching, at a time t , from cell size s' to size s at rates $W(s|s', t)$, equation (3) can be written as

$$\frac{\partial}{\partial t} p(s, t|s_0) = \frac{\partial}{\partial s} [g(s, t)p(s, t|s_0)] + \int ds' [W(s|s', t)p(s', t|s_0) - W(s'|s, t)p(s, t|s_0)] \quad (5)$$

If perfect symmetric splitting is considered through the condition $(\delta(s' - 2s))$ and $|W(s'|s, t)| = h(s, s')$, the division rates $W(s|s', t)$ can be written as:

$$W(s|s', t) = \delta(s' - 2s)h(s, s'). \quad (6)$$

Some studies have explored the particular case where $h(s) = ks$ with k being a constant and discarding the dependence upon s' [17, 22]. This, resulting on:

$$\frac{\partial p}{\partial t}(s, t) = -\frac{\partial(g(s)p(t, s))}{\partial s} - (ks)p(t, s) + (k2s)p(t, 2s). \quad (7)$$

Which is also known as the Population Balance Equation for a fixed population number [17]. Some theoretical methods like moment closure were used in past studies to solve (7) but stable solutions were not found already [17].

$$\frac{\partial}{\partial t} p(s, n; t|x_0) = \underbrace{\frac{\partial}{\partial s} [g(s)p(s, n; t|x_0)]}_{\text{Drift by growth}} + \underbrace{[k(2s)p(2s, n-1; t|x_0) - k(s)p(s, n; t|x_0)]}_{\text{Jumps by divisions}}, \quad (9)$$

where $x_0 = (s_0, n_0)$.

The inclusion of the new variable n breaks the non-locality of the operator $W(s|s', t)$ that makes (5) hard to solve. Instead $W(s, n|s', n-1, t)$ performs jumps between independent subspaces that can be merged together later with marginal sums.

Using this new variable n , exponential growth and $s(0) = s_0$ equation (9) has closed solutions

$$p(s, n; t) = \delta(s - s(n, t))P_n(t), \quad (10)$$

where $P_n(t)$ is the probability to get divided n times at time t . We will present its associated equation later. $s(n, t)$ corresponds to the bacterial size after n divisions at a time t . To explain how to solve this size, let us consider the size $s(1, t)$ at time t after one division. If $t_1 < t$ is the time when division occurs, this size is:

$$s(1, t) = \frac{s_0 e^{\mu t_1}}{2} e^{\mu(t-t_1)}. \quad (11)$$

So, having the general sequence of division times $0 < t_1 < t_2 <$

The Chapman-Kolmogorov equation including division events.

To overcome these instabilities, other studies reparametrize the size distribution adding an additional variable: the number of divisions $n \in \{0, 1, 2, \dots\}$ [22]. In this case, the probability distribution is now $p(s, n; t)$ with transition rates satisfying:

$$W(s', n'|s, n) = \delta(s' - \frac{s}{2})\delta_{n', n+1}h(s', s). \quad (8)$$

This is, after division, not only the size is halved but the number of divisions n , increase by one unit. In the particular case where $h(s, s') = ks$, the associated CKE is:

$\dots < t_{n-1} < t_n < t$, $s(n, t)$ satisfies:

$$\begin{aligned} s(n, t) &= s_0 \left(\prod_{i=1}^n \frac{e^{\mu(t_i-t_{i-1})}}{2} \right) e^{\mu(t-t_n)} \\ &= \frac{s_0}{2^n} \exp \left(\mu \sum_{i=1}^n (t_i - t_{i-1}) + \mu(t - t_n) \right) \\ &= \frac{s_0}{2^n} e^{\mu(t-t_0)} = \frac{s_0 e^{\mu t}}{2^n}, \end{aligned} \quad (12)$$

where we used $t_0 = 0$ and the telescopic properties of the sum.

Using this result (12) and (10), (9) is separated to the system of equations:

$$\begin{aligned} p(s; t) &= \sum_{n=0}^{\infty} p(s, n; t) = \sum_{n=0}^{\infty} \delta \left(s - \frac{s_0 e^{\mu t}}{2^n} \right) P_n(t) \\ \frac{dP_0}{dt} &= -ks_0 e^{\mu t} P_0 \\ &\vdots \\ \frac{dP_n}{dt} &= -\frac{ks_0 e^{\mu t}}{2^n} P_n + \frac{ks_0 e^{\mu t}}{2^{n-1}} P_{n-1} \\ &\vdots, \end{aligned} \quad (13)$$

defining the dynamics of every $P_n(t)$.

The division strategy

Focusing on the jump process between the state $n - 1$ to n , or simply, one division, we can modify the space $n \in 0, 1, 2, \dots$ to $n \in \{0, 1\}$ and truncate (13) to:

$$\begin{aligned} \frac{dP_0}{dt} &= -h(s)P_0 \\ \frac{dP_1}{dt} &= h(s)P_0, \end{aligned} \quad (14)$$

Where a general size dependent splitting rate function $h(s)$ can be used. This system can be integrated under the initial conditions $P_0(0) = 1$ and $P_1(0) = 0$. Thus, the probability $P_1(t)$ —to simplify notation, $P(t)$ —that the cell divides in the time interval $(0, t)$ evolves according to:

$$P(t) = 1 - \exp \left(- \int_0^t h(s(t')) dt' \right). \quad (15)$$

In the particular case of $h(s)$, proportional to the size, this is, $h(s) = ks$, assuming exponential growth as well, the integration on time has to be done using the implicit formula:

$$h(t) = ks(t) = ks_0 e^{\mu t}. \quad (16)$$

Once $P(t)$ is obtained, the probability density function $\rho(t)$ for the time of division can be obtained as:

$$\rho(t) = \frac{dP(t)}{dt}. \quad (17)$$

A transformation of variables, allows us to get the distribution of sizes at division $\rho(s_d)$:

$$\rho(s_d) = \rho(t(s_d)) \frac{dt}{ds_d}, \quad (18)$$

where, if we assume exponential growth $t(s_d) = \frac{1}{\mu} \ln \left(\frac{s_d}{s_0} \right)$, then, $\frac{dt}{ds_d} = \frac{1}{\mu s_d}$. Using this $\rho(s_d)$, one can calculate the mean size at division by integrating:

$$\langle s_d \rangle = \int_{s_b}^{\infty} s_d \rho(s_d) ds_d. \quad (19)$$

Hence, we can calculate the mean added size per cell cycle $\langle \Delta \rangle = \langle s_d \rangle - s_b$ as a function of the size at birth s_b . This relationship defines the division strategy.

The multi-step Single division

In the general case, division does not correspond to a single jump process, instead, division occurs once bacteria have reached some goal steps M . If the occurrence rate of these steps is proportional to the cell size s by the constant k_d , the probability $P_m(t)$ of having done $m < M$ steps at time t can be modelled following:

$$\begin{aligned} \frac{dP_0}{dt} &= -k_d s(t) P_0 \\ &\vdots \\ \frac{dP_m}{dt} &= k_d s(t) P_{m-1} - k_d s(t) P_m \\ &\vdots \\ \frac{dP_M}{dt} &= k_d s(t) P_{M-1}. \end{aligned} \quad (20)$$

P_M is the probability of reaching the target steps M or equivalently, the probability of a division event to occur. Once the division event happens, the process is reset to zero steps and size is halved.

Using this $P_M(t)$ and the growth regime (4), if the procedure (18) is followed, the probability density $\rho(s_d|s_b)$ of size at division s_d given the size at birth s_b in a cell cycle, satisfies:

$$\rho(s_d|s_b) = \left(\frac{k_d}{\mu} \right)^M \frac{(s_d - s_b)^{(M-1)}}{(M-1)!} \exp \left(- \frac{k_d}{\mu} (s_d - s_b) \right) \quad (21)$$

Defining the added size before division $\Delta = s_d - s_b$, from , we observe that $\langle \Delta \rangle = \langle s_d \rangle - s_b$ is independent on the size at birth s_b and is related to the growth rate μ , the objective steps M and the step occurrence rate k_d per size unit, satisfying:

$$\langle \Delta \rangle = M \frac{\mu}{k_d} \quad (22)$$

The solution of the CKE including multiple divisions

Using a similar procedure as (13) now with the additional variable m , the probability of have a size s , have done m division steps and n divisions up to time t can be written now as:

$$p(s, n, m; t) = \delta(s - s(n, t)) P_{m,n} \quad (23)$$

Cell size $s(n, t)$ follows, again, the equation (12). While, the probability $P_{m,n}(t)$ of have done m division steps and n division events up to time t , can be estimated through the master equation system:

$$\begin{aligned} \frac{dP_{0,0}}{dt} &= -k_d s_0 e^{\mu t} P_{0,0} \\ \frac{dP_{1,0}}{dt} &= k_d s_0 e^{\mu t} P_{0,0} - k_d s_0 e^{\mu t} P_{1,0} \\ &\vdots \\ \frac{dP_{m,n}}{dt} &= \frac{s_0 e^{\mu t}}{2^n} P_{m-1,n} - \frac{s_0 e^{\mu t}}{2^n} P_{m,n} \\ &\vdots \\ \frac{dP_{M,n}}{dt} &= k_d \frac{s_0 e^{\mu t}}{2^n} P_{M-1,n} - k_d \frac{s_0 e^{\mu t}}{2^n} P_{M,n} \\ \frac{dP_{0,n+1}}{dt} &= k_d \frac{s_0 e^{\mu t}}{2^n} P_{M,n} - k_d \frac{s_0 e^{\mu t}}{2^{n+1}} P_{0,n+1} \\ &\vdots, \end{aligned} \quad (24)$$

where we showed the selection rule defining the divisions as jumps between states $(M, n-1)$ to $(0, n)$ and the division steps as jumps between states (m, n) to $(m+1, n)$. These jumps happening at rate $h = k_d s(n, t)$ with $s(n, t)$ following the equation (12).

Numerical estimation of size dynamics

Solution of (24) can be obtained using different methods. Analytically, one can start from the initial conditions:

$$P_{m,n}(t=0) = \delta_{n,0} \delta_{m,0}, \quad (25)$$

with $\delta_{i,j}$, the Kronecker delta. Hence, $P_{m,n}(t)$ can be obtained knowing $P_{m-1,n}(t)$ with $m \in 1, \dots, M$ and $P_{0,n}(t)$, can be estimated from $P_{M,n-1}(t)$. Both, using, using the closed recurrence expression:

$$\begin{aligned} P_{m,n}(t) &= \frac{k_d s_0}{2^n} \exp \left[-\frac{k s_0}{\mu 2^n} e^{\mu t} \right] \int_0^t K_n(t') P_{m-1,n}(t') dt' \\ P_{0,n}(t) &= \frac{k_d s_0}{2^{n-1}} \exp \left[-\frac{k s_0}{\mu 2^n} e^{\mu t} \right] \int_0^t K_n(t') P_{M,n-1}(t') dt' \\ \text{with} \\ K_n(\tau) &= \exp \left[\mu \tau + \frac{k s_0}{\mu 2^n} e^{\mu \tau} \right] \\ P_{0,0}(t) &= \exp \left[-\frac{k_d s_0}{\mu} (e^{\mu t} - 1) \right] \end{aligned} \quad (26)$$

The Finite State Projection Algorithm

In general, in (24), while the number of possible steps $m \in \{0, 1, \dots, M\}$ has finite cardinality, the number of possible divisions $n \in \{0, 1, \dots\}$ is infinite. Thus, making impossible the complete solution of (24) using methods like Matrix exponential. As we explained before [22], this infinite set can be projected into a finite set using the known Finite State Projection (FSP) algorithm [23]. Using this approach, the number of equations in (24) are truncated up to a maximum divisions N and the number of possible division states is now finite. Hence, known methods for solving these finite systems can be used to estimate size dynamics during infinitesimal periods of time.

From $P_{m,n}(t)$, the size distribution $\rho(s|s_0)$ given the starting size s_0 and the moments of this size distribution: the mean size $\langle s \rangle$ and the variance $\text{var}(s) = \langle s^2 \rangle - \langle s \rangle^2$, can be estimated from the equations:

$$\begin{aligned} \rho(s|s_0) &= \sum_{n=0}^N \delta \left(s - s_0 \frac{e^{\mu t}}{2^n} \right) P_n \\ \langle s \rangle &= \sum_{n=0}^N s_0 \frac{e^{\mu t}}{2^n} P_n \\ \text{var}(s) &= \sum_{n=0}^N \left(\left(s_0 \frac{e^{\mu t}}{2^n} \right)^2 - \langle s \rangle^2 \right) P_n, \end{aligned} \quad (27)$$

with $P_n = \sum_{m=0}^M P_{m,n}$ and $\delta(x)$, the Dirac-delta distribution.

The computing of these moments was done considered that all cells began at initial size $s(0) = s_0$, this is, $\rho(s_0) = \delta(s_0 - s(0))$. However, if a general density function $\rho(s_0)$ is considered, the size distribution $\rho(s)$ is a convolution of solutions of (27):

$$\rho(s) = \int \rho(s|s_0) \rho(s_0) ds_0 \quad (28)$$

Stochastic simulation of size dynamics

Let the single step process shown in (14). While (14) was presented to modeling the division as a single step process, in general, these equations are also valid for a division step. Setting

explicitly the dependence $h = k_d s$, the system describing the single step process is now:

$$\begin{aligned} \frac{dP_0}{dt} &= -k_d s_0 e^{\mu t} P_0 \\ \frac{dP_1}{dt} &= k_d s_0 e^{\mu t} P_0. \end{aligned} \quad (29)$$

If $P_0(0) = 1$ and $P_1(0) = 0$, $P_1(t)$, or simply $P(t)$, has solution:

$$\begin{aligned} P(t) &= 1 - \exp \left(- \int_0^t k_d s(t') dt' \right) \\ &= 1 - \exp \left[-s_0 \frac{k_d}{\mu} (e^{\mu t} - 1) \right], \end{aligned} \quad (30)$$

while the associated density function is:

$$\begin{aligned} \rho(t) &= \frac{dP(t)}{dt} \\ &= s_0 \frac{k_d}{\mu} \exp \left[\mu t - s_0 \frac{k_d}{\mu} (e^{\mu t} - 1) \right]. \end{aligned} \quad (31)$$

The main idea behind the stochastic simulation algorithm is to generate random time events τ_s distributed as (31). Following the Gillespie's method [24], we generate a random number r uniformly distributed in the interval $(0, 1)$ and from the cumulative function (30), τ_s is obtained matching $P(t)$ and r thus, solving for t :

$$\tau_s = \frac{1}{\mu} \ln \left[1 - \frac{\mu}{s_0 k_d} \ln(r) \right], \quad (32)$$

where we take advantage of the fact that $1 - r$ is distributed as r . This τ_s is the time to the occurrence of the next division step.

Additional details to model the cell division

The asymmetric splitting

The main assumption proposing (12) is that after each division, the cell size is perfectly halved. However, this is not the case in a realistic situation. Experimentally, some stochastic fluctuations in the septum position are found. In some growth conditions, this noise can be as high as 5% [25, 26].

Considering again the size at time t after one division, if the division occurred at time $t_1 < t$. If the size is not perfectly halved but multiplied by a random variable b_1 , centered on 0.5 and also known as division ratio, the size at time t is now:

$$s(1, t) = s_0 e^{\mu(t - t_0)} b_1 e^{\mu(t - t_1)} \quad (33)$$

If the sequence of division ratios $\{b_1, b_2, \dots, b_n\}$ is known, the size after n division at time t is given by:

$$s(n, t) = s_0 e^{\mu t} \prod_{k=1}^n b_k. \quad (34)$$

Theoretically, b_k can be approximated to a beta distributed variable centered on 0.5 with variance fitted from experiments.

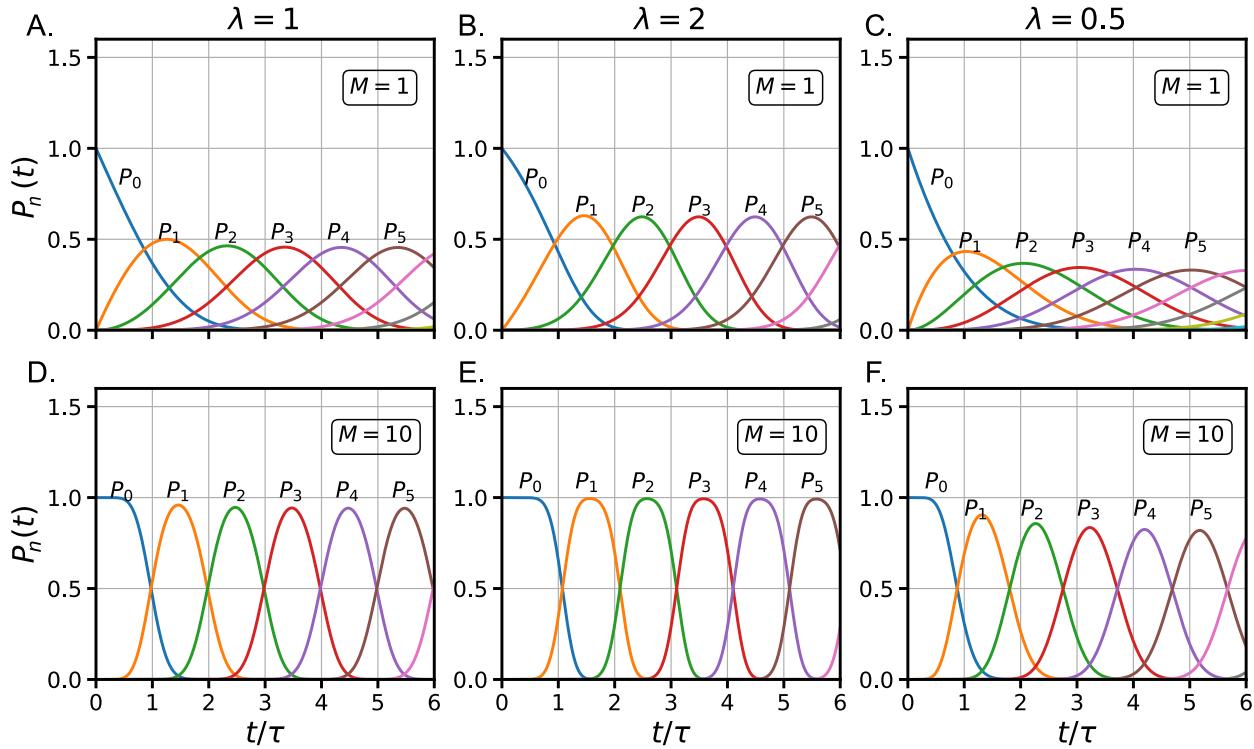


Figure 1: Time dynamics of the first P_n for different division strategies and division steps (M) with $\rho(s_b) = \delta(s_b - \bar{s}_b)$.

Cell-to-cell noise in growth rate

Other important stochastic variable is the cell-to-cell growth rate [27, 28]. This fluctuation can be as high as 10% [3]. We can assume that after a division k , a new growth rate μ_k is chosen randomly from a distribution centered on μ and cell grows during that cycle with rate μ_k . If symmetric splitting is considered again, the size at time t after n divisions is now:

$$s(n, t) = \frac{s_0}{2^n} \left(\prod_{k=0}^{n-1} \exp(\mu_k(t_{k+1} - t_k)) \right) \exp(\mu_n(t - t_n)) \quad (35)$$

Numerically, we modeled these μ_k as a gamma distributed variable centered on the mean growth rate μ , with the experimental variance and no correlation with past cycles. This last assumption might be incorrect and these correlation between cycles can be studied in deeper studies [26].

A general CKE including additional sources of noise

A general CKE, assuming exponential growth, can be modeled with growth rate distributed with given distribution $\rho(\mu)$:

$$\frac{\partial p(s; t)}{\partial t} = \int \left[\frac{\partial}{\partial s} (\mu s p(s; t)) \right] \rho(\mu) d\mu + \int h(s|s') p(s'; t) ds' - h(s) p(s; t). \quad (36)$$

The division rate $h(s|s')$ in general depends on hidden variables like the number steps but can be estimated, at least numerically [29–31]. In the case of single step division, this rate can be written as $h(s|s') = k s \rho(s, s')$ with $\int \rho(s, s') ds' = \rho(s)$ and $\rho(s)$ being the distribution of size at birth. In the therm $\rho(s, s')$ non-symmetric division can also be considered. $h(s)$, in the last therm, is obtained from $h(s) = \int h(s'|s) ds'$. We already do not have a closed solution to (36) and numerical methods could be hard to implement.

Different division strategies

Depending on the mapping $s_d = f(s_b)$, or traditionally, between the relationship between added size $\Delta = s_d - s_b$ and s_b , three main division strategies have been defined for exponentially growing bacteria: the *timer*, *adder* and *sizer* strategies [32]. Differing from the slope of Δ vs s_b for timer this slope is +1; -1 in sizer strategy and for null-valued for adder.

The adder strategy, observed for instance, in *E. coli* and *B. subtilis* [3], is considered the most common strategy in bacteria. In some bacterial populations, however, division strategies with intermediate slopes for Δ vs. s_b have also been observed [14, 32]. This has led to the definition of the *timer-like* strategy, for slopes

between 0 and 1, and of the *sizer-like* strategy, for slopes between -1 and 0 .

These deviations from the adder can be obtained if the step rate (h) is not proportional to the size but proportional to a power (λ) of the size [14]. This is:

$$h = k_d s^\lambda \quad (37)$$

A multi-step process similar to that described by (24) can be proposed. As was explained in past studies [14], if division is triggered by the occurrence of M steps happening at rate (37), the distribution of size at division s_d given the size at birth s_b is [14]:

$$\rho(s_d|s_b) = \left(\frac{k_d s_d^{\lambda-1}}{\mu} \right) \exp \left[-\frac{k_d}{\mu \lambda} (s_d^\lambda - s_b^\lambda) \right] \frac{\left(\frac{k_d}{\mu \lambda} (s_d^\lambda - s_b^\lambda) \right)^{M-1}}{(M-1)!} \quad (38)$$

Timer strategy (Added size has a slope of -1 with size at birth) is obtained if $\lambda \rightarrow 0$ and sizer strategy (Added size is proportional to the size at birth) is obtained if $\lambda \rightarrow \infty$. The adder is obtained when $\lambda = 1$.

Considering the non-linear step rate given by (37) and a similar method like (31), the general stochastic time is given by:

$$\tau_s(\lambda) = \frac{1}{\lambda \mu} \ln \left[1 - \frac{\lambda \mu}{k_d s_0^\lambda} \ln(r) \right] \quad (39)$$

The mean cell size at birth

The main variables defining the mean cell size are the growth rate μ , the number of division steps M and the division steps occurrence rate k_d . If adder strategy is considered ($\lambda = 1$), the mean added size $\langle \Delta \rangle$ follows the relationship (22).

Since there is already a discussion about the nature of k_d , the inference of its actual value is not straightforward. For the adder, this k_d can be inferred from their mean added size using (22) and by observing that this $\langle \Delta \rangle$ is independent on the size at birth s_b . In different division strategies ($\lambda \neq 1$), $\langle \Delta \rangle$ is now function of s_b . Now, the typical size as explained in past studies [14, 17], \bar{s}_b , is the size at birth that is perfectly doubled by the division strategy:

$$\bar{s}_b = s_b : \langle s_d \rangle (s_b) = 2s_b \quad (40)$$

This, since after division, the cell, with $s_d = 2\bar{s}_b$, splits on a half and the size of its offspring is \bar{s}_b again.

In general, \bar{s}_b , is dependent on k_d , μ and λ and is one of the variables that can be measured most easily if we assume that this \bar{s}_b is actually the mean size at birth in a steady growing cell population. Hence, other variables like k_d can be estimated from this \bar{s}_b using (40) and root-finding algorithms.

Illustrating examples

To observe the dynamics of the probability of get divided n times at time t , we present, in Figure 1, time trends of some $P_n(t)$ s for different λ s and M s with initial condition $P_{n,m}(0) = \delta_{n,0} \delta_{m,0}$.

A numerical analysis of the behavior in Figure 1 allows us to find that, in the limit of $t \rightarrow \infty$, the distribution of P_n s satisfies

$$\lim_{t \rightarrow \infty} \|P_n(t) - P_{n+1}(t + \tau)\| = 0. \quad (41)$$

Which implies asymptotic invariance of the system under translation on, simultaneously, $n \rightarrow n+1$ and $t \rightarrow t+\tau$ over the P_n . Since $\frac{s_b e^{\mu t}}{2^n}$ also satisfies this invariance, we expect $\rho(s|t, s_b)$ to show periodic properties in the limit $t \rightarrow \infty$. This periodicity was already discussed in some theoretical papers [33–35]. This convergence is shown in Appendix A.

Three different scenarios in size dynamics can be explored using our formalism. Figure 2 a. shows the mean size dynamics obtained from a simulation of 5000 cells, all of them with the same starting size and beginning with zero division steps or equivalently, starting from their most recent division. We assumed that they have the same growth rate and get split perfectly symmetric. Simulation was done using the stochastic times (32) and numerical estimation was done solving the master equation (24), both of them, having perfect adjustment to each other. Ten examples of cell cycles were plotted on the background to observe how variable the distributions are. The dynamics of this variability, quantified by the coefficient of variation $C_v^2(s) = \frac{\text{var}(s)}{\langle s \rangle^2}$, is also shown in Figure 2.e. As main effect observed, we can highlight the oscillations in both $\langle s \rangle$ and $C_v^2(s)$ with period equal to τ , the doubling time. The oscillations in the $C_v^2(s)$ present their peaks just when bacteria are dividing on average and their valleys when bacteria are growing. The dynamics of the size distribution can be seen in supplemental video 1.

In Figure 2. b., the mean dynamics corresponds to cells with an initial distribution with finite variance ($C_v^2(s, t=0) = 0.02$). This distribution was assumed to be Gamma distribution since it is well defined from its mean size and the $C_v^2(s)$. Simulations, on the other hand, were modified by using random initial sizes, numerical estimation was done by performing the convolution (28) and approaching the integral to a numerical Riemann sum. Dynamics on cell size variability are also presented in Figure 2. f. Similar oscillations were found in both the mean and the variability but with less amplitude than the first case. The dynamics of this distribution appears in supplemental video 2.

A third scenario consists on the assumption that bacteria do not split perfectly on a half but in a beta distributed independent stochastic variable centered on 0.5 and with a given variability $C_v^2 = 0.002$. Growth rate could be considered stochastic as well with variability set to $C_v^2 = 0.02$. The dynamics of the mean size (Fig 1.c.) and its variability (Fig 1.g.) are presented. We plotted ten cycles in the background of Figure 2.c. to show some examples of typical single-cell size dynamics. Since this noise is not considered in equation (24), the numerical approach is hard to compute and thus not presented in Figure 2. The distribution dynamics can be seen in supplementary video 3.

We also estimate the division strategy using both simulations and numerical estimations. Data from stochastic simulations can be obtained using the stochastic division times (39) and exponential growth (4) while the trends in added size and its variability can be obtained from the distribution of size at division (38) being both dependent on the exponent λ . In Figure 2 d. we present the mean added size δ as function of the mean size at birth s_b for three different λ . These λ were chosen to represent three of the most important division strategies: timer-like ($0 < \lambda < 1$, where we choose $\lambda = 0.5$) with its characteristic positive slope on Δ vs s_b , Adder ($\lambda = 1$) with no correlation between Δ and s_b and sizer-like ($1 < \lambda < \infty$, where we choose $\lambda = 2$) with a negative slope in Δ vs s_b . Fluctuations over these trends are also shown in Figure 2

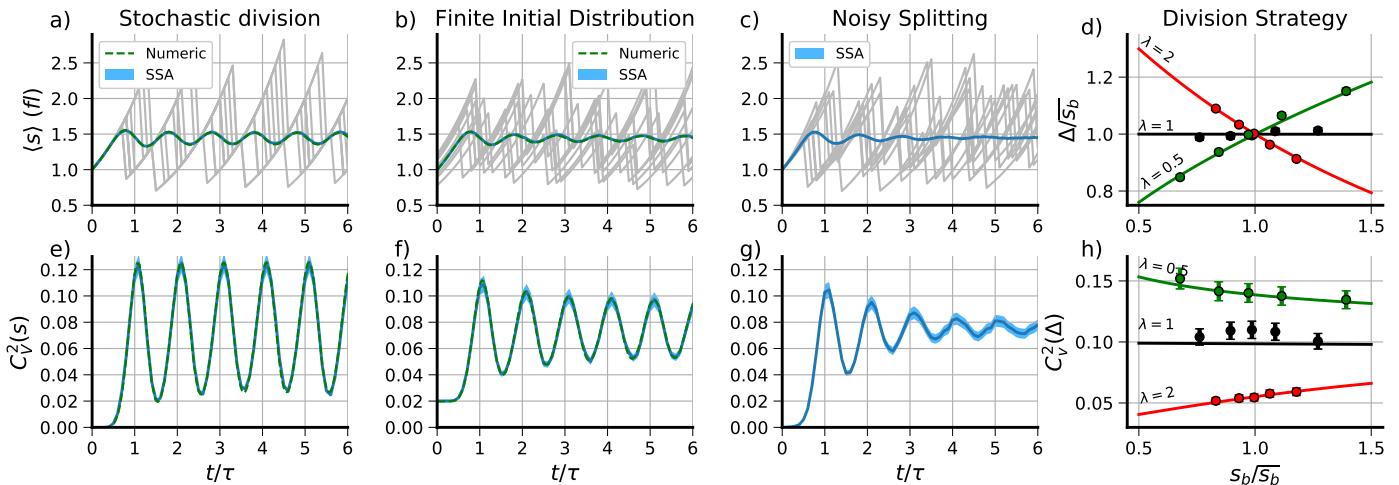


Figure 2: Main properties of bacterial cell division explored using PyEcoLib. a) Mean cell size $\langle s \rangle$ and e) Its variability $C_v^2(s)$ along the time considering only stochastic division. b) Mean cell size and f) Its variability along the time considering both stochastic division and an initial size distribution with finite variance. c) Mean cell size and g) Its variability along the time considering stochastic division and noise in cell-to-cell growth rate and septal position. d) mean added size Δ vs the size at birth s_b and h) the fluctuations $C_v^2(\Delta)$ vs s_b for different division strategies. Timer-like ($\lambda = 0.5$), adder ($\lambda = 1$) and sizer-like ($\lambda = 2$). Simulations (dots) and numerical estimations (lines) are shown. $M = 10$ division steps were considered in all cases.

h. where it can be seen that sizer-like shows positive correlation in $C_v^2(\Delta)$ vs s_b , adder strategy shows no-correlation and timer-like shows a negative correlation. To understand better the properties of the robust oscillations on the dynamics of the central moments presented in Figure 3.a and 3.b., we studied the auto-correlation function of the size. This auto-correlation $\gamma(t')$ is defined through the formula:

$$\gamma(t') = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{\langle [(s(t) - \langle s(t) \rangle)(s(t + t') - \langle s(t + t') \rangle)] \rangle}{\sigma(t)\sigma(t + t')} dt, \quad (42)$$

with $\sigma(t)$ being the standard deviation of the size at time t and $\langle x \rangle$ is the mean value of the random variable x

How the auto-correlation dynamics changes along the time is presented on Figure 3. a. where we present four different cases: a single division step, which size dynamics have been presented already [22]. This auto-correlation decays exponentially to zero. By increasing the division steps, for instance to 10 steps, oscillations appear around a decaying trend. For a division almost deterministic, for instance 50 steps, these oscillations have higher amplitude around zero. When noise on both growth rate and septum position is considered, these oscillations are damped in the same way found in size dynamics, converging to zero. This asymptotic decorrelation let the distribution reach a stationary distribution (at $t = 10\tau$) with fixed moments which is presented in Figure 3. b. for $M = 10$ and the noises explained above.

Discussion

In this article, we present a theoretical scheme to estimate the stochastic dynamics of the cell size for a population of constant number. This approach, based on the Chapman-Kolmogorov formalism, assumed that the population number remains constant along the time. Although the framework can be extended to a exponentially growing population, simulation could be unstable since

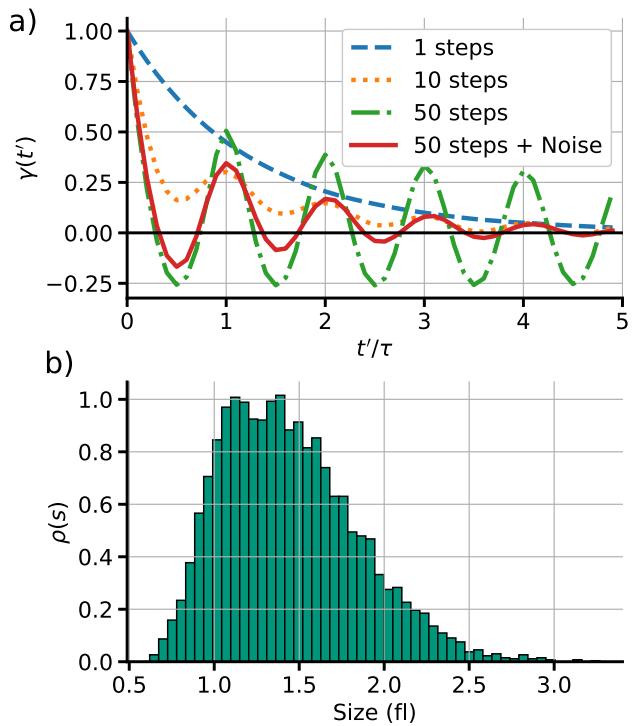


Figure 3: a) Cell size auto-correlation $\gamma(t')$ for different time periods t' for four different division conditions: one division step (dashed, blue line), 10 division steps (dotted, orange line), 50 division steps (green, dash-dot line) and 50 division steps including noise in growth rate and septum position (red continuous line). b) Simulation of stationary state of the histogram of bacterial size with all the noise sources considered.

the number of cells in these populations grows exponentially on time. We think to consider this case of CKE in future studies.

Additionally to the estimation of size dynamics, our framework can be also used for: Simulate most of the division strategies found in *E. coli*: timer-like, adder and sizer-like [10, 32]. Estimate the distribution of division times, size at division and size at birth. Noise in both septal ring placing and cell-to-cell growth rate can be considered as well [26].

As is shown in Figure 2, oscillations in both, the mean size $\langle s \rangle$ and $C_v^2(s)$, were found. When only stochasticity on division times is considered, these oscillations are maintained over an arbitrary long period of time having lower amplitude when an initial size distribution with finite variance is considered. Some damping occurs if other sources of noise like the cell-to-cell growth rate variability and septum position are added.

This robustness on the oscillations can be understood as result of the asymptotic periodic properties of the probability P_n of have n divisions at time n . These probabilities are invariant under the simultaneous transformation $n \rightarrow n + 1$ and $t \rightarrow t + \tau$ originating the oscillations. This invariance under this transformation, is broken when other noise sources are considered.

These properties found in size dynamics can be also obtained using the classical Discrete Stochastic Maps. A clear correspondence between DSM and our model can be found when the adder strategy is considered. In this case, the stochastic map between size at birth s_b and size at division is:

$$s_d = s_b + \Delta + \epsilon, \quad (43)$$

with ϵ being an independent random variable with zero mean and a distribution fitted from experiments. If exponential growth is considered, the cell-cycle duration τ_d , as random variable, can be obtained from (43):

$$\tau_d = \frac{1}{\mu} \ln \left[1 + \frac{\Delta}{s_b} \left(1 + \frac{\epsilon}{\Delta} \right) \right] \quad (44)$$

where, using (22), some analogy to (32) can be found. Using these times with parameters fitted from the data, similar oscillations in both size trends and auto-correlation can be obtained.

The main difference with DSMs is found where deviations from the adder are considered. Using the CRMs, the fluctuations (ϵ) in the division strategy (43) will depend on the size at birth unlike the DSM where these fluctuations are not related to any other variable. Some preliminary observations on the dependence of the fluctuations on added size with the size at birth in sizer-like division in *E. coli* have been reported [14, 36] but further observations are needed.

Including cell size stochasticity to gene expression can be an important tool to understand the origin of the fluctuations in molecule concentration. Some efforts have already been done to understand these effects in simple regulatory networks [37–39] but our formalism can improve this study to more complex gene regulatory architectures. Other effects such as the division strategy, the noise in growth rate and the asymmetric cell splitting can be also studied using our framework.

Appendix

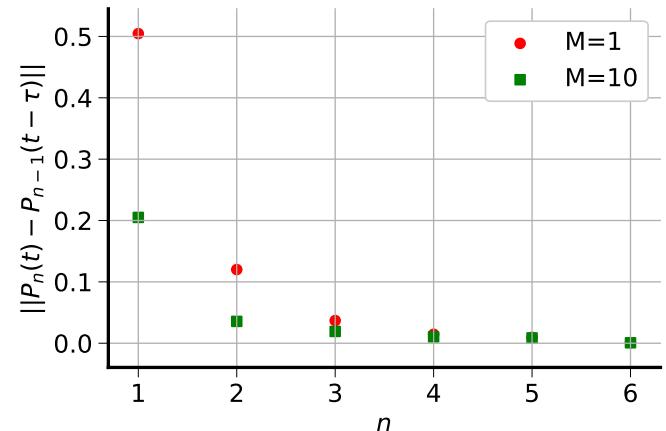


Figure A1: Distance between the probabilities $P_n(t)$ and $P_{n-1}(t - \tau)$ for two different division steps (red dots: $M = 1$) (green squares: $M = 10$).

A The asymptotic periodicity of the system

To check the property (41), we calculate numerically the distance between the probabilities $P_n(t)$ and $P_{n-1}(t - \tau)$ using the expression:

$$\|P_n(t) - P_{n-1}(t - \tau)\| = \int |P_n(t) - P_{n-1}(t - \tau)| dt, \quad (A1)$$

the results are presented in Figure A1 where we can check how this distance decays asymptotically to zero as n increases.

B Size dynamics for different growth conditions and division rates

Past studies suggested that the periodicity under translations in τ and n is an exclusive property of the exponential growth [33]. Thus, if other growth conditions are considered, the symmetry under these translations is broken and the robust oscillations are no expected.

We simulated the size dynamics of different possible growth and division rates in similar way explained in past studies [17]. Thus, we define the growth law as (2), where exponential growth is defined as $g(s, t) = \mu s$ and the linear growth is given by $g(s, t) = \mu$ with μ being a constant.

On the other hand, the division rate is defined by the function h as (6). In figure A2, we compare two division rates for the linear growth, one of these are $h = k$ and the other is $h = ks$ with k a constant. These rates define the occurrence of a given division steps. In figure A2 we considered $M = 10$ steps to trigger the division.

In Figure A2.a. we present the dynamics of the mean cell size $\langle s \rangle$ along the time with some single trajectories in the background (gray lines) presented to compare them to the mean trend. While in Figure A2.c. we present its variability $C_v^2(s)$ along the time for for a linear growth $g = \mu$ and constant division rate $h = k$. b) $\langle s \rangle$ vs t

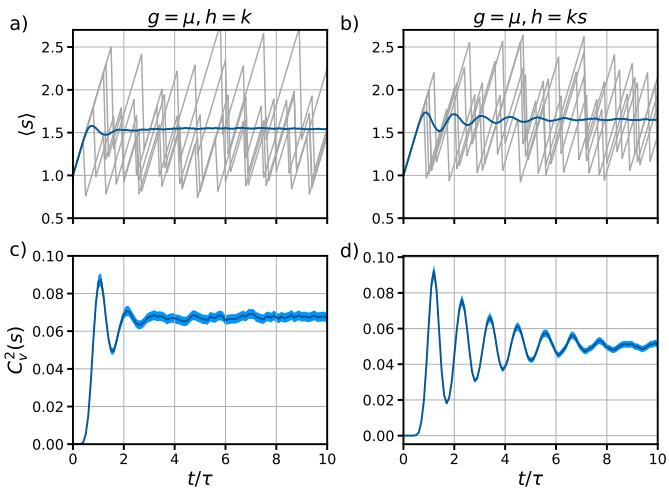


Figure A2: Mean cell size $\langle s \rangle$ and its variability $C_v^2(s)$ along the time for different growth and splitting rates. a) $\langle s \rangle$ vs t and c) $C_v^2(s)$ vs t both for a linear growth $g = \mu$ and constant division rate $h = k$. b) $\langle s \rangle$ vs t and d) $C_v^2(s)$ vs t both for a linear growth $g = \mu$ and a division rate proportional to the cell size $h = ks$.

and d) $C_v^2(s)$ vs t both for a linear growth $g = \mu$ and a division rate proportional to the cell size $h = ks$. Perfectly symmetric splitting and no-noise in growth rate is considered.

References

- [1] 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*, vol. 166, no. 3, pp. 729–739, 2016.
- [2] 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*, vol. 159, no. 6, pp. 1433–1446, 2014.
- [3] 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," *Current biology*, vol. 25, no. 3, pp. 385–391, 2015.
- [4] J. B. Andersen, C. Sternberg, L. K. Poulsen, S. P. Bjørn, M. Givskov, and S. Molin, "New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria," *Applied and environmental microbiology*, vol. 64, no. 6, pp. 2240–2246, 1998.
- [5] C. Nieto-Acuña, J. C. Arias-Castro, C. Vargas-García, C. Sánchez, and J. M. Pedraza, "Correlation between protein concentration and bacterial cell size can reveal mechanisms of gene expression," *Physical Biology*, vol. 17, no. 4, p. 045002, 2020.
- [6] 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*, vol. 329, no. 5991, pp. 533–538, 2010.
- [7] D. F. Anderson and T. G. Kurtz, *Stochastic analysis of biochemical systems*. Springer, 2015.
- [8] J. M. Raser and E. K. O'shea, "Noise in gene expression: origins, consequences, and control," *Science*, vol. 309, no. 5743, pp. 2010–2013, 2005.
- [9] W. J. Blake, G. Balázs, M. A. Kohanski, F. J. Isaacs, K. F. Murphy, Y. Kuang, C. R. Cantor, D. R. Walt, and J. J. Collins, "Phenotypic consequences of promoter-mediated transcriptional noise," *Molecular cell*, vol. 24, no. 6, pp. 853–865, 2006.
- [10] P.-Y. Ho, J. Lin, and A. Amir, "Modeling cell size regulation: From single-cell-level statistics to molecular mechanisms and population-level effects," *Annual review of biophysics*, vol. 47, pp. 251–271, 2018.
- [11] C. W. Gardiner *et al.*, *Handbook of stochastic methods*, vol. 3. Springer Berlin, 1985.
- [12] P. Wang, L. Robert, J. Pelletier, W. L. Dang, F. Taddei, A. Wright, and S. Jun, "Robust growth of *escherichia coli*," *Current biology*, vol. 20, no. 12, pp. 1099–1103, 2010.
- [13] M. Osella, E. Nugent, and M. C. Lagomarsino, "Concerted control of *escherichia coli* cell division," *Proceedings of the National Academy of Sciences*, vol. 111, no. 9, pp. 3431–3435, 2014.
- [14] C. Nieto, J. Arias-Castro, C. Sánchez, C. Vargas-García, and J. M. Pedraza, "Unification of cell division control strategies through continuous rate models," *Physical Review E*, vol. 101, no. 2, p. 022401, 2020.
- [15] F. Si, G. Le Treut, J. T. Sauls, S. Vadia, P. A. Levin, and S. Jun, "Mechanistic origin of cell-size control and homeostasis in bacteria," *Current Biology*, vol. 29, no. 11, pp. 1760–1770, 2019.
- [16] K. Sekar, R. Rusconi, J. T. Sauls, T. Fuhrer, E. Noor, J. Nguyen, V. I. Fernandez, M. F. Buffing, M. Berney, S. Jun, *et al.*, "Synthesis and degradation of ftsz quantitatively predict the first cell division in starved bacteria," *Molecular systems biology*, vol. 14, no. 11, p. e8623, 2018.
- [17] N. Totis, C. Nieto, A. Küper, C. Vargas-García, A. Singh, and S. Waldherr, "A population-based approach to study the effects of growth and division rates on the dynamics of cell size statistics," *IEEE Control Systems Letters*, vol. 5, no. 2, pp. 725–730, 2020.
- [18] J. Lin and A. Amir, "From single-cell variability to population growth," *Physical Review E*, vol. 101, no. 1, p. 012401, 2020.
- [19] P. Thomas, "Making sense of snapshot data: ergodic principle for clonal cell populations," *Journal of The Royal Society Interface*, vol. 14, no. 136, p. 20170467, 2017.
- [20] B. van Brunt, A. Almalki, T. Lynch, and A. Zaidi, "On a cell division equation with a linear growth rate," *The ANZIAM Journal*, vol. 59, no. 3, pp. 293–312, 2018.
- [21] C. A. Vargas-García, K. R. Ghusinga, and A. Singh, "Cell size control and gene expression homeostasis in single-cells," *Current opinion in systems biology*, vol. 8, pp. 109–116, 2018.
- [22] C. A. Nieto-Acuña, C. A. Vargas-García, A. Singh, and J. M. Pedraza, "Efficient computation of stochastic cell-size transient dynamics," *BMC bioinformatics*, vol. 20, no. 23, pp. 1–6, 2019.
- [23] B. Munsky and M. Khammash, "The finite state projection algorithm for the solution of the chemical master equation," *The Journal of chemical physics*, vol. 124, no. 4, p. 044104, 2006.
- [24] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *Journal of computational physics*, vol. 22, no. 4, pp. 403–434, 1976.
- [25] 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," *Proceedings of the National Academy of Sciences*, vol. 111, no. 45, pp. 15912–15917, 2014.
- [26] S. Modi, C. A. Vargas-García, K. R. Ghusinga, and A. Singh, "Analysis of noise mechanisms in cell-size control," *Biophysical journal*, vol. 112, no. 11, pp. 2408–2418, 2017.
- [27] 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*, vol. 514, no. 7522, pp. 376–379, 2014.
- [28] S. Vadia and P. A. Levin, "Growth rate and cell size: a re-examination of the growth law," *Current opinion in microbiology*, vol. 24, pp. 96–103, 2015.
- [29] C. A. Vargas-García and A. Singh, "Elucidating cell size control mechanisms with stochastic hybrid systems," in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 4366–4371, IEEE, 2018.
- [30] M. Doumic, A. Olivier, and L. Robert, "Estimating the division rate from indirect measurements of single cells," *arXiv preprint arXiv:1907.05108*, 2019.
- [31] M. Xia, C. D. Greenman, and T. Chou, "Pde models of adder mechanisms in cellular proliferation," *SIAM Journal on Applied Mathematics*, vol. 80, no. 3, pp. 1307–1335, 2020.
- [32] J. T. Sauls, D. Li, and S. Jun, "Adder and a coarse-grained approach to cell size homeostasis in bacteria," *Current opinion in cell biology*, vol. 38, pp. 38–44, 2016.
- [33] P. Bokes and A. Singh, "Cell volume distributions in exponentially growing populations," in *International Conference on Computational Methods in Systems Biology*, pp. 140–154, Springer, 2019.

- [34] E. Bernard, M. Doumic, and P. Gabriel, "Cyclic asymptotic behaviour of a population reproducing by fission into two equal parts," *arXiv preprint arXiv:1609.03846*, 2016.
- [35] O. Diekmann, H. J. Heijmans, and H. R. Thieme, "On the stability of the cell size distribution," *Journal of Mathematical Biology*, vol. 19, no. 2, pp. 227–248, 1984.
- [36] C. Nieto, J. Arias-Castro, C. Vargas-Garcia, C. Sanchez, and J. M. Pedraza, "Noise signature in added size suggests bacteria target a commitment size to enable division," *bioRxiv*, 2020.
- [37] C. A. N. Acuna, C. A. V. Garcia, and J. M. Pedraza, "Stochasticity in bacterial division control: Preliminary consequences for protein concentration," *bioRxiv*, p. 826867, 2019.
- [38] 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 computational biology*, vol. 12, no. 8, p. e1004972, 2016.
- [39] P. Thomas, G. Terradot, V. Danos, and A. Y. Weiße, "Sources, propagation and consequences of stochasticity in cellular growth," *Nature communications*, vol. 9, no. 1, pp. 1–11, 2018.