Probalistic and piecewise deterministic models in biology

Bertrand Cloez, Renaud Dessalles, Alexandre Genadot, Florent Malrieu, Aline Marguet, Romain Yvinec

PROBABILISTIC AND PIECEWISE DETERMINISTIC MODELS IN BIOLOGY

BERTRAND CLOEZ*, RENAUD DESSALLES†, ALEXANDRE GENADOT‡, FLORENT MALRIEU§, ALINE MARGUET⊥, AND ROMAIN YVINEC♯

Abstract. We present recent results on Piecewise Deterministic Markov Processes (PDMPs), involved in biological modeling. PDMPs, first introduced in the probabilistic literature by [30], are a very general class of Markov processes and are being increasingly popular in biological applications. They also give new interesting challenges from the theoretical point of view. We give here different examples on the long time behavior of switching Markov models applied to population dynamics, on uniform sampling in general branching models applied to structured population dynamic, on time scale separation in integrate-and-fire models used in neuroscience, and, finally, on moment calculus in stochastic models of gene expression.

Introduction

The piecewise deterministic Markov processes (denoted PDMPs) were first introduced in the literature by Davis [30, 31]. The key point of his work was to endow the PDMP with rather general tools, similar as such that already exist for diffusion processes. Indeed, PDMPs form a family of non-diffusive càdlàg Markov processes, involving a deterministic motion punctuated by random jumps. The motion of the PDMP $\{X(t)\}_{t\geq 0}$ depends on three local characteristics, namely the jump rate λ, the deterministic flow φ and the transition measure Q according to which the location of the process at the jump time is chosen. The process starts from x and follows the flow $\varphi(x,t)$ until the first jump time T_1 which occurs either spontaneously in a Poisson-like fashion with rate $\lambda(\varphi(x,t))$ or when the flow $\varphi(x,t)$ hits the boundary of the state-space. In both cases, the location of the process at the jump time T_1, denoted by $Z_1 = X(T_1)$, is selected by the transition measure $Q(\varphi(x,T_1),\cdot)$ and the motion restarts from this new point as before. This fully describes a piecewise continuous trajectory for $\{X(t)\}_{t\geq 0}$ with jump times $\{T_k\}$ and post jump locations $\{Z_k\}$, and which evolves according to the flow φ between two jumps.

Since the seminal work of Davis, PDMPs have been heavily studied from the theoretical perspective, we may refer for instance the readers to [27, 2, 1, 57, 7, 9] among many others.

From the applied point of view, and more precisely in biological applications, these processes are sometimes referred as hybrid, and are especially appealing for their ability to capture both continuous (deterministic) dynamics and discrete (probabilistic or deterministic) events. First applications date back at least to the studies of the cell cycle model [53, 52], and more recent works, to name just a few, in neurobiology [63, 35, 35], cell population and branching models [4, 22, 60, 36, 51, 20], gene expression models [72, 56, 33], food contaminants [14, 12] or multiscale chemical reaction network models [28, 3, 50, 46]. See also [47, 16, 66] for selected reviews on hybrid models in biology.

The paper is organized as follows. In Section 1, we present long time behavior results for switched flows in population dynamics. In Section 2, we present many-to-one formulas and description of the trait of an individual uniformly sampled in a general branching model applied to structured population dynamic. In Section 3, we

*MISTEA, INRA, Montpellier SupAgro, Univ. Montpellier, France.
†INRA, UR MaIAGE, domaine de Vilvert, Jouy en Josas, France.
‡Institut de Mathématiques de Bordeaux, Inria Bordeaux-Sud-Ouest, France.
§Laboratoire de Mathématiques et Physique Théorique (UMR CNRS 7350), Fédération Denis Poisson (FR CNRS 2964), Université François-Rabelais, Parc de Grandmont, 37200 Tours, France.
⊥CMAP, École Polytechnique, Université Paris-Saclay, Palaiseau, France.
♯Physiologie de la Reproduction et des Comportements, Institut National de la Recherche Agronomique (INRA) UMR85, CNRS-Université François-Rabelais UMR7247, IFCE, Nouzilly, 37380 France

present limit theorems and time scale separation in integrate-and-fire models used in neuroscience. Finally, in Section 4, we derive asymptotic moments formulas in a stochastic model of gene expression.

1. Long time behavior of some PDMP for population dynamics

In this section, we present recent results on long time behavior and exponential convergence of some PDMP used in population dynamics, in particular, for modeling population growth in a varying environment.

We consider processes $(X_t, I_t)_{t\geq 0}$, where the first component X_t has continuous paths in a continuous space E, namely a subset of $\mathbb{R}^d$, for some integer d, and the second component I_t is a pure jump process on a finite state space F or in $\mathbb{N}$. The continuous variable represents the number of individuals (or more precisely the population density or the relative abundance) of a population and the discrete variable models the population environment. In a fixed environment $I_t = i \in F$, we assume that $(X_t)_{t\geq 0}$ evolves as the solution of an ordinary differential equation, that is

$$\forall t \geq 0\,, \quad \partial_t X_t = F^{(i)}(X_t), \tag{1.1}$$

where $F^{(i)}$ is some smooth function. For instance, the oldest and maybe the most famous model to represent the evolution of a population is the choice $E = \mathbb{R}_+$ and

$$F^{(i)} : x \mapsto a_i x, \quad a_i \in \mathbb{R}. \tag{1.2}$$

When there is no variability in the environment, solutions are given by exponential functions and there is either explosion or extinction according to the sign of a_i. We develop in Subsection 1.1 this toy model when the environment varies. Almost all properties (almost-sure convergence, moments...) can be derived and it gives a first hint to understand more complicated models such as multi-dimensional linear models or non-linear models. These generalizations are described in Subsection 1.2, and 1.3, respectively. For non-linear population models, maybe one of the most famous models is given by the Lotka-Volterra equations (also known as the predator-prey equations), with state-space $E = \mathbb{R}_+^2$ and

$$F^{(i)}(x, y) = \begin{pmatrix} \alpha_i x(1 - a_i x - b_i y) \\ \beta_i x(1 - c_i x - d_i y) \end{pmatrix}, \quad a_i, b_i, c_i, d_i, \alpha_i, \beta_i \in \mathbb{R}_+.$$

To end the introduction of this section, let us mention that all of these models are particular instance of a class of switching Markov model, see for instance [9, 71] for general references. Some other examples of application are developed in [57, 38, 61].

1.1. Malthus model.

Let us consider, in this subsection, that $(I_t)_{t\geq 0}$ is an irreducible continuous time Markov chain on a finite state space F. We denote by ν its unique invariant probability measure. The process $(X_t)_{t\geq 0}$ will be the solution of

$$\forall t \geq 0,\ \partial_t X_t = a_{I_t} X_t.$$

This means that $F^{(i)}$ is given by (1.2). Thus, we have

$$\forall t \geq 0,\ X_t = X_0 e^{\int_0^t a_{I_s} ds}.$$

The almost-sure behavior of this process is easy to understand. Indeed, by the ergodic theorem, we have that

$$\lim_{t\to+\infty} \frac{1}{t} \int_0^t a_{I_s} ds = \sum_{i\in F} a_i \nu(\{i\}). = \nu(a).$$

Then, we have the following dichotomy: either $\nu(a) > 0$ and $(X_t)_{t\geq 0}$ tends almost-surely to infinity at an exponential rate, or $\nu(a) < 0$ and $(X_t)_{t\geq 0}$ tends to 0. The second statement can be understood as the extinction of the population, even if, in contrast with classical birth and death processes, $(X_t)_{t\geq 0}$ never hits 0; that is the extinction time (or the hitting time of 0) is almost-surely infinite.

The convergence of moments of $(X_t)_{t\geq 0}$ is more tricky than the almost-sure convergence. Indeed, one cannot use a dominated convergence theorem or a Jensen inequality (which is in the wrong way). Nevertheless, we have

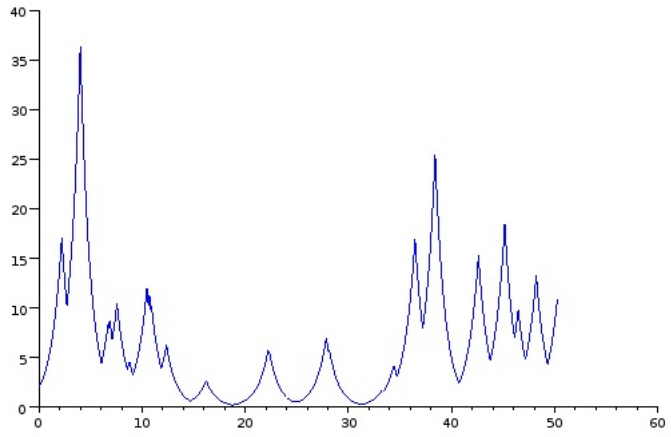


FIGURE 1. Trajectory of $(X_t)_{t\geq 0}$ in the Malthus model.

Lemma 1.1 (Convergence of moments)**.** *If $\nu(a) > 0$ then*

$$\forall q > 0, \quad \lim_{t\to+\infty} \mathbb{E}[X_t^q] = +\infty.$$

If $\nu(a) < 0$ then there exists $p > 0$ such that

$$\forall q \leq p, \quad \lim_{t\to+\infty} \mathbb{E}[X_t^q] = 0.$$

The proof which follows comes from [6]; see also [32].

Proof. If we denote by A the generator of the process $(I_t)_{t\geq 0}$ and μ_0 the vector corresponding to its initial law, then the Feynman-Kac formula states that, for every $p > 0$,

$$\mathbb{E}[X_t^p] = \mathbb{E}[X_0^p]\mathbb{E}\left[e^{\int_0^t pa_{I_s} ds}\right] = \mathbb{E}[X_0^p]\mu_0 e^{t(A+pD)}\mathbf{1},$$

where it is the classical exponential of matrices, D is the diagonal matrix such that $D_{i,i} = a_i$ and $\mathbf{1}$ is the vector made of ones. Now, classical results on matrices (including Perron-Frobenius Theorem) give the existence of two constants $c_p, C_p > 0$ such that

$$c_p e^{\lambda_p t} \leq \mu_0 e^{t(A+pD)}\mathbf{1} \leq C_p e^{\lambda_p t},$$

where λ_p is the (unique) eigenvalue of largest real part of the matrix $(A + pD)$. It then remains to understand the sign of λ_p to understand the moment behaviour. By definition of λ_p, there exists v_p such that

$$v_p(A + pD) = \lambda_p v_p.$$

Moreover, for $p = 0$, we have $\lambda_0 = 0$ and $v_0 = \nu$. Again classical results on matrices ensure that $p \mapsto \lambda_p$ and $p \mapsto v_p$ are smooth and then differentiating in p, we obtain

$$v_p D + \partial_p v_p(A + pD) = v_p \partial_p \lambda_p + \lambda_p \partial_p v_p.$$

In particular, for $p = 0$ and multiplying by $\mathbf{1}$ by the right, we have $\nu(a) = (\partial_p \lambda_p)_{|p=0}$. As a consequence, if $\nu(a) > 0$, then $p \mapsto \lambda_p$ is increasing in a neighborhood of 0 and for all $p > 0$ small enough we have $\lambda_p > 0$. Thus, $\mathbb{E}[X_t^p]$ tends to infinity for p small enough, and then for all $p > 0$ by the Jensen inequality. In contrary, if $\nu(a) < 0$ then there exists $p > 0$ such that $\lim_{t\to+\infty} \mathbb{E}[X_t^p] = 0$ and also for all $q < p$, again by Jensen inequality. □

Understanding the long time behavior of moments is primordial when we are interested in the convergence to some invariant law more general than δ_0. Indeed, moments provide Lyapunov-type functionals ensuring that the process returns frequently in a compact-set.

1.2. **Others linear models.** From a modeling point of view, the results of the previous section are easily understandable. Indeed, it is enough that the environment has in mean a positive effect to ensure an exponential growth, while if the environment is, in mean, not favorable, then the population goes to extinct.

However, note that this interpretation only holds because the growth parameter is well-chosen. Indeed, let us first consider a discrete time analogous of the Malthus model: the sequence $(Y_n)_{n\geq 0}$ satisfying

$$\forall n \geq 0, \quad Y_{n+1} = \Theta_n Y_n,$$

where $(\Theta_n)_n$ is a sequence of i.i.d. non-negative random variables. Here Θ_n represents the mean number of children per individuals at generation n (all individuals have the same number of children which depends on the environment). If $(\Theta_n)_{n\geq 0}$ is a constant sequence then $Y_n = \Theta_1^n Y_0$ and a dichotomy occurs with respect to a critical value $\Theta_1 = 1$. When $(\Theta_n)_{n\geq 0}$ is a sequence of i.i.d random variable, then there is also a dichotomy but the critical value (to be smaller or larger than 1) is no longer $\mathbb{E}[\Theta_1]$ but $\exp(\mathbb{E}[\ln(\Theta_1)])$. Indeed, we have

$$\forall n \geq 0, \quad Y_n = Y_0 \prod_{k=0}^{n-1} \Theta_k = X_0 e^{\sum_{k=0}^{n-1} \ln(\Theta_k)}.$$

This is exactly the same phenomenon for Galton-Watson chain in random environment, see for instance [5] and [42, Section 2.9.2].

There is a similar (but more complex, maybe) problem in upper dimension. Let us consider that $E = \mathbb{R}^2$, $F = \{0, 1\}$, and $(I_t)_{t\geq 0}$ jumps from 0 to 1 and from 1 to 0 with the same rate $\lambda > 0$, and

$$F^{(0)} : v \mapsto \begin{pmatrix} -1 & 4 \\ -1/4 & -1 \end{pmatrix} \cdot v, \quad F^{(1)} : v \mapsto \begin{pmatrix} -1 & -1/4 \\ 4 & -1 \end{pmatrix} \cdot v. \tag{1.3}$$

Then the solutions of the equation $\partial_t \begin{pmatrix} x \\ y \end{pmatrix} = F^{(1)} \begin{pmatrix} x \\ y \end{pmatrix}$ are given by

$$\forall t \geq 0, \quad \begin{cases} x(t) = e^{-t}(\cos(t)x(0) - \sin(t)y(0)/4) \\ y(t) = e^{-t}(4\sin(t)x(0) + \cos(t)y(0)). \end{cases} \tag{1.4}$$

It is easy to see that there exists some constants $C > 0$ such that for all $t \geq 0$,

$$\|(x(t), y(t))\| \leq Ce^{-t}\|(x(0), y(0))\|.$$

This property is also satisfied by the solutions of $\partial_t \begin{pmatrix} x \\ y \end{pmatrix} = F^{(0)} \begin{pmatrix} x \\ y \end{pmatrix}$. However, we have the following theorem.

Theorem 1.2 ([8])**.** *Let $(X_t)_{t\geq 0}$ be the solution of Equation* (1.1) *with flows defined in* (1.3)*. There exists $\beta_c > 0$ such that if $\lambda < \beta_c$ then*

$$\lim_{t\to+\infty} \|X(t)\| = 0,$$

and if $\lambda > \beta_c$ then $\lim_{t\to+\infty} \|X(t)\| = +\infty$. Moreover, the speed of convergence is exponential.

For this type of results for more general vector fields, see for instance [8, 57, 54]. Let us briefly explain this phenomenon. One can see that, whatever the initial conditions of the solutions (x, y) of (1.4), the map $t \mapsto \|(x(t), y(t))\|$ is not decreasing; see for instance Figure (2).

Then, heuristically, we see that if I jumps sufficiently fast (before the decay of the distance) then the distance may grow. However, if I is too slow, then $(X_t)_{t\geq 0}$ essentially follows one of the two flows and goes rapidly to 0.

Switching linear models can have even more unpredictable behavior than the previous one. Indeed, in [54], the author exhibits two different linear functions $F^{(0)}, F^{(1)}$ on $\mathbb{R}^2$ and two critical values $\beta_2 > \beta_1$ such that the switched process $(X_t)_{t\geq 0}$ tends to infinity for at least one $\lambda \in (\beta_1, \beta_2)$ and to 0 if $\lambda \in (0, \beta_1) \cup (\beta_2, +\infty)$.

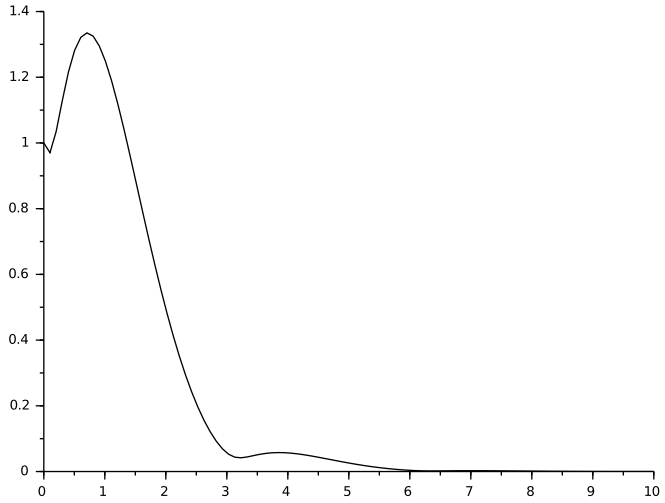


FIGURE 2. Graph of $t \mapsto \|(x(t), y(t))\| = e^{-t}\sqrt{15\sin^2(t)+1}$ for (1.4) when $(x_0, y_0) = (1, 0)$.

1.3. **A general theorem for convergence.** In this subsection, we expose a condition similar to Subsection 1.1 for ensuring convergence to an equilibrium (not necessary δ_0). As pointed out in the previous section, we have to measure precisely the growth of each underlying flow. So we assume that for all $i \in F$, there exists $\rho(i) \in \mathbb{R}$ such that

$$\langle x - y, F^{(i)}(x) - F^{(i)}(y)\rangle \leq -\rho(i)\|x - y\|^2. \tag{1.5}$$

The parameter $\rho(i)$ measures the contraction of a solution $\dot{u} = F^{(i)}(u)$ with respect to the norm $\|\cdot\|$. Indeed, for two solutions, u^1 and u^2, we have

$$\forall t \geq 0, \quad \|u^1(t) - u^2(t)\|^2 \leq e^{-\rho(i)t}\|u^1(0) - u^2(0)\|^2.$$

Note that the next result does not rely on a specific norm, but it has to be the same for all flows. A sufficient condition to prove (1.5) is given by the next lemma.

Lemma 1.3. *Let $G : \mathbb{R}^d \to \mathbb{R}^d$ be a C^1 map and let $\mathbf{J}_x G$ be its Jacobian matrix at the point $x \in \mathbb{R}^d$. If for all $u, x \in \mathbb{R}^d$,*

$$\langle u, \mathbf{J}_x G \cdot u\rangle \leq -\rho\|u\|^2,$$

then, for all $x, y \in \mathbb{R}^d$,

$$\langle x - y, G(x) - G(y)\rangle \leq -\rho\|x - y\|^2. \tag{1.6}$$

In particular, if $G = \nabla V$, for some C^2 function V, then (1.6) holds under the classical condition that V is strongly convex with constant ρ; namely for all $x \in \mathbb{R}^2$,

$$Hess_x V \geq \rho I,$$

in the sense of quadratic form.

Proof. It is classic to derive the previous bounds using the mean value theorem for the following map

$$\varphi : t \mapsto \frac{\langle x - y, G(x) - G(ty + (1-t)x)\rangle}{\|x - y\|^2}.$$

□

We can now state the main result of this subsection.

Theorem 1.4 ([24]). *Assume that $(I_t)_{t\geq 0}$ is an irreducible Markov process on a finite state space F with invariant measure ν, $(X_t)_{t\geq 0}$ satisfies* (1.1)*, and that Eq.* (1.5) *holds. If*

$$\sum_{i\in F} \rho(i)\nu(i) > 0,$$

then (X_t, I_t) converges in law to a unique invariant law.

The proof of this theorem is given in [24, 7]. The proof of [24] is based on a generalization (developed in [44]) of the classical Harris Theorem [43], which does not hold in general for PDMP (see for instance the introduction of [7]). One of the main point of the proof is the construction of a Lyapunov function V, in the sense that there exist, $C, \gamma, K > 0$ such that

$$\forall t \geq 0,\ \mathbb{E}[V(X_t, I_t)] \leq Ce^{-\gamma t}V(X_0, I_0) + K. \tag{1.7}$$

The construction of V and the control of its moments are based on Lemma 1.1.

Note that another proof of the previous theorem is given in [7]. Nevertheless, the proof of [24] takes the advantage to be generalizable in many contexts (dependent environment, non-deterministic underlying dynamics, asymptotic assumptions...). See [24] for details.

Closely related articles [10, 58, 59] deal with a fluctuating Lotka-Volterra model. They show that random switching between two environments both favorable to the same specie can lead to the extinction of this specie or coexistence of both species.

As Theorem 1.4, their proofs are based on the construction of a Lyapunov function. Nevertheless instead of controlling the exit from a compact set of $\mathbb{R}$ to infinity, this Lyapunov function is used to control the exit from a compact set included in $(0,\infty)^2$ to the axes $\{(0, y)|y \geq 0\}$ or $\{(x, 0,)|x \geq 0\}$. A complete description of this model is given in [58]. Note however that their counter-intuitive result is similar to the one of Theorem 1.2. It is based on the fact that, when I jumps sufficiently fast (namely λ being large in Theorem 1.2), then X is close to follow the following deterministic dynamics:

$$\forall t \geq 0, \quad \partial_t X_t = \sum_{i\in F} \nu(i)F^{(i)}(X).$$

Finally, we end this section mentioning that such hybrid framework can also be applied to models where the population is represented by a discrete variable, and/or the environment fluctuates continuously. One can cite for instance [26], where the author studies a model of interaction between a population of insects and a population of trees. Another example is the chemostat model (a type of bioreactor), introduced in [29] and studied in [25, 19, 18, 23]. Lyapunov functions are again a key theoretical tool to study long time behavior for such models. However, we point out that extinction events may occur in discrete population models, and that the interplay between the discrete and continuous components makes the problem more difficult than standard absorbing time studies in purely discrete population models. Nevertheless, it was proved in [23, Theorem 3.4] that extinction event is almost-surely finite and has furthermore exponential tails. This generalizes part of a result of [25].

2. Uniform sampling in a branching structured population

In this section, which is a short version of [60], we give a characterization of the trait of a uniformly sampled individual among the population at time t in a structured branching population.

2.1. Introduction.

We consider a branching Markov process where each individual u is characterized by a trait $(X_t^u,\ t \geq 0)$ which dynamic follows an $\mathcal{X}$-valued Markov process, where $\mathcal{X} := \widetilde{\mathcal{X}} \times \mathbb{R}_+$ and $\widetilde{\mathcal{X}} \subset (\mathbb{R}_+)^d$ for some $d \geq 1$. For any $x \in \mathcal{X}$, the last coordinate corresponds to a time coordinate and we denote by $\widetilde{x} \in \widetilde{\mathcal{X}}$ the vector of the first d coordinates. We assume that this trait influences the lifecycle of each individual (in terms of lifetime, number of descendants and inheritance of the trait). Then, an interesting problem consists in the characterization of the trait of a "typical" individual in the population. This is the question we address here.

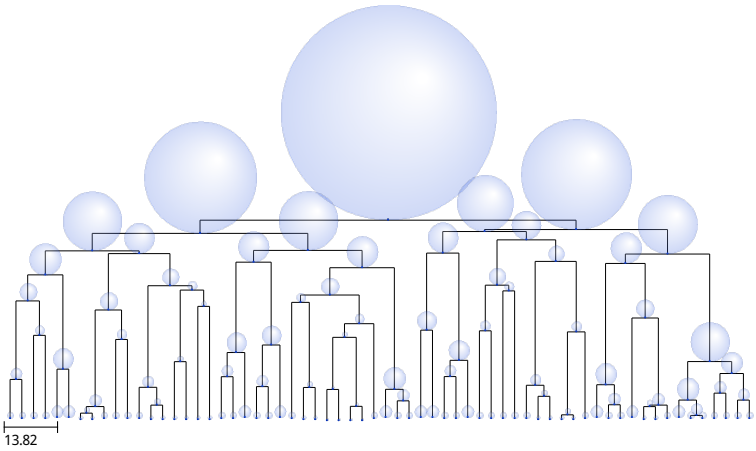


FIGURE 3. Descending genealogy from an individual with size 1 until time $T = 50$ of a size-structured population. Each cell grows exponentially at rate 0.01 and divide at rate $B(x) = x$.

Let us describe the process in more details. We consider a strongly continuous contraction semi-group with associated infinitesimal generator $\mathcal{G} : \mathcal{D}(\mathcal{G}) \subset \mathcal{C}_b(\mathcal{X}) \to \mathcal{C}_b(\mathcal{X})$, where $\mathcal{C}_b(\mathcal{X})$ denotes the space of continuous bounded functions from $\mathcal{X}$ to $\mathbb{R}$. Then, each individual u has a trait $(X_t^u,\ t \geq 0)$ which evolves as a Markov process defined as the unique $\mathcal{X}$-valued càdlàg solution of the martingale problem associated with $(\mathcal{G}, \mathcal{D}(\mathcal{G}))$. An individual with trait x dies at an instantaneous rate $B(x)$, where B is a continuous function from $\mathcal{X}$ to $\mathbb{R}_+$. It is replaced by $A_u(x)$ children, where $A_u(x)$ is a $\mathbb{N}$-valued random variable with distribution $(p_k(x), k \geq 0)$. For convenience, we assume that $p_1(x) \equiv 0$ for all $x \in \mathcal{X}$. The trait of the descendants at birth depends on the trait of the mother at death. For all $k \in \mathbb{N}$, let $P^{(k)}(x, \cdot)$ be the probability measure on $\mathcal{X}^k$ corresponding to the trait distribution at birth of the k descendants of an individual with trait x. We denote by $P_j^{(k)}(x, \cdot)$ the jth marginal distribution of $P^{(k)}$ for all $k \in \mathbb{N}$ and $j \leq k$.

Finally, we denote by $\mathcal{M}_P(\mathcal{X})$ the set of point measures on $\mathcal{X}$. Following Fournier and Méléard [39], we work in $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_P(\mathcal{X}))$, the state of càdlàg measure-valued Markov processes. For any $Z \in \mathbb{D}(\mathbb{R}_+, \mathcal{M}_P(\mathcal{X}))$, we write:

$$Z_t = \sum_{u \in V_t} \delta_{X_t^u},\ t \geq 0,$$

the measure-valued process describing the dynamic of the population where V_t denotes the set of all individuals in the population at time t. We will denote by N_t the cardinality of V_t and by:

$$m(x, s, t) = \mathbb{E}\left(N_t \middle| Z_s = \delta_x\right),$$

its expected value, for $0 \leq s \leq t$.

An example of such a process is the size-structured population where each individual grows exponentially fast. For more details we refer the reader to [36] or [13]. Then, at rate $B(X_t^u)$, the individual u splits at time t in two daughter cells of size $X_t^u/2$. Figure 3 is a realization of such a process. The diameter of each circle corresponds to the size of the individual at division and the length of the branch represents the lifetime of each individual. In particular, we notice the link between the lifetime and the size of each individual: the bigger the cell is, the shorter its lifetime is.

In Section 2.2, we give the chosen assumptions on the model to ensure that our process is well-defined. Then, in Section 2.3, we describe the process corresponding to the trait of a "typical" individual in the population. Finally, in Section 2.4, we explain why this process corresponds to the trait of a uniformly sampled individual in a large population approximation.

2.2. **Definition and existence of the structured branching process.** To define rigorously the structured branching process on $\mathbb{R}_+$, we need to ensure that almost surely the number of individuals in the population does not blow up in finite time.

As explained in the previous description of the process, the lifetime of each individual depends on the dynamic of the trait through the division rate B. One way of dealing with this dependency is to assume that the division rate is upper bounded by some constant $\overline{B}$. Then, in the case of a binary division, the expected number of individuals in the population at time t is upper bounded by the expected value of a Yule process with birth rate $\overline{B}$ at time t which is equal to $\exp(\overline{B}t)$ and which is bounded on compact sets. In the case of an unbounded division rate, the same reasoning applies if we can control the excursions of the dynamic of the trait. Thus, we consider two sets of hypotheses: the first one controls what happens regarding divisions (in term of rate of division and of mass creation) and the second one controls the dynamic of the trait between divisions.

Assumption 1. *We consider the following assumptions:*

(1) *There exist* $b_1, b_2 \geq 0$ *and* $\gamma \geq 0$ *such that for all* $x \in \mathcal{X}$,
$$B(x) \leq b_1 |x|^\gamma + b_2.$$

(2) *There exists* $\underline{x} \in \widetilde{\mathcal{X}}$ *such that for all* $x \in \mathcal{X}$, $k \in \mathbb{N}$:
$$\int_{\mathcal{X}^k} \left(\sum_{i=1}^k \widetilde{y}_i \right) P^{(k)}(x, dy_1 \dots dy_k) \leq \widetilde{x} \vee \underline{x}, \text{ componentwise.}$$

(3) *There exists* $\overline{m} \geq 0$ *such that for all* $x \in \mathcal{X}$,
$$m(x) := \sum_k k p_k(x) \leq \overline{m}.$$

The first point controls the division rate. The second point means that we consider a fragmentation process with a possibility of mass creation at division when the mass is small enough. In particular, clones are allowed in the case of bounded traits and bounded number of descendants and any finite type branching structured process can be considered.

As explained before, we make a second assumption to control the behavior of traits between divisions.

Assumption 2. *There exist* $c_1, c_2 \geq 0$ *such that for all* $x \in \mathcal{X}$*:*
$$\mathcal{G} h_\gamma(x) \leq c_1 h_\gamma(x) + c_2,$$
where γ *is defined in Assumption* 1 *and for* $x \in (\mathbb{R}_+)^d$, $h_\gamma(x) = |x|^\gamma = \left(\sum_{i=1}^d x_i \right)^\gamma$.

This assumption ensures in particular that the trait does not blow up in finite time in the case of a deterministic dynamic for the trait. Assumptions 1(1) and 2 are linked via the parameter γ which controls the balance between the growth of the population and the dynamic of the trait. In particular, if Assumption 1 is satisfied for $\gamma = 0$, the division rate is bounded and Assumption 2 is always satisfied.

Under Assumptions 1 and 2, the previously described measure-valued process is well-defined as the strongly unique solution of a stochastic differential equation.

The proof of existence relies on a recursive construction of a solution via the sequence of jumps time of the population. Then, we prove that this sequence is unique conditionally to the Poisson point measure determining the jumps in the population and to the stochastic flows corresponding to the dynamic of the trait. Finally, using Assumptions 1 and 2, we prove that the number of individuals in the population does not blow up in finite time. This is detailed in [60] (Theorem 2.3.).

2.3. **The trait of sampled individuals at a fixed time : Many-to-One formula.** In order to characterize the trait of a uniformly sampled individual, the spinal approach ([21],[55]) consists in following a "typical" individual in the population whose behavior summarizes the behavior of the entire population. In this section, we give the dynamic of the trait of a typical individual: it is a Markov process called the auxiliary process.

From now on, we assume that for all $x \in \mathcal{X}$, $t \geq 0$ and $s \leq t$, $m(x,s,t) \neq 0$. With a slight abuse of notation, we denote by X_s^u the trait of the unique ancestor living at time s of u.

Let f be a non-negative measurable function and $x \in \mathcal{X}$. The idea is to notice that the following operator:

$$P_{r,s}^{(t)} f(x) := \frac{\mathbb{E}\left[\sum_{u \in V_t} f\left(X_s^u\right) \middle| Z_r = \delta_x\right]}{m(x,r,t)}, \tag{2.1}$$

is a conservative (non-homogeneous) semi-group. Then, the auxiliary process is defined as the time-inhomogeneous Markov process with associated family of semi-groups $\left(P_{r,s}^{(t)}, r \leq s \leq t\right)$. It describes the dynamic of the trait of a "typical" individual in the population in the sense that it satisfies a so-called "Many-to-One" formula. This formula is true under Assumptions 1, 2 and under the following technical assumptions:

Assumption 3. *There exists a function C such that for all $j \leq k,\ j,k \in \mathbb{N}$ and $0 \leq s \leq t$, we have:*

$$\sup_{x \in \mathcal{X}} \sup_{s \in [0,t]} \int_{\mathcal{X}} \frac{m(y,s,t)}{m(x,s,t)} P_j^{(k)}(x,dy) \leq C(t),\ \forall t \geq 0.$$

This assumption tells us that we control uniformly in x the benefit or the penalty of a division.

Assumption 4. *For all $t \geq 0$ and $x \in \mathcal{X}$ we have:*

- *$s \mapsto m(x,s,t)$ is differentiable on $[0,t]$ and its derivative is continuous on $[0,t]$,*
- *for all $f \in \mathcal{D}(\mathcal{A}) := \{f \in \mathcal{D}(\mathcal{G})$ s.t. $m(\cdot,s,t)f \in \mathcal{D}(\mathcal{G})\ \forall t \geq 0,\ \forall s \leq t\}$, $s \mapsto \mathcal{G}(m(\cdot,s,t)f)(x)$ is continuous,*

We can now state the Many-to-One formula.

Theorem 2.1. *Under Assumptions 1, 2, 3 and 4, for all $t \geq 0$, for all $x_0 \in \mathcal{X}$ and for all non-negative measurable functions $f : \mathcal{X} \to \mathbb{R}_+$, we have:*

$$\mathbb{E}_{\delta_{x_0}}\left[\sum_{u \in V_t} f\left(X_s^u\right)\right] = m(x_0,0,t)\mathbb{E}_{x_0}\left[f\left(Y_s^{(t)}\right)\right], \tag{2.2}$$

where $\left(Y_s^{(t)}, s \leq t\right)$ is an inhomogeneous-Markov process with infinitesimal generator $\left(\mathcal{A}_s^{(t)}, \mathcal{D}(\mathcal{A})\right)$ given for $f \in \mathcal{D}(\mathcal{A})$ and $x \in \mathcal{X}$ by:

$$\mathcal{A}_s^{(t)} f(x) = \widehat{\mathcal{G}}_s^{(t)} f(x) + \widehat{B}_s^{(t)}(x) \int_{\mathcal{X}} \left(f(y) - f(x)\right) \widehat{P}_s^{(t)}(x,dy),$$

where:

$$\widehat{\mathcal{G}}_s^{(t)} f(x) = \frac{\mathcal{G}\left(m(\cdot,s,t)f\right)(x) - f(x)\,\mathcal{G}\left(m(\cdot,s,t)\right)(x)}{m(x,s,t)},$$

$$\widehat{B}_s^{(t)}(x) = B(x) \int_{\mathcal{X}} \frac{m(y,s,t)}{m(x,s,t)} m(x,dy),$$

$$\widehat{P}_s^{(t)}(x,dy) = \frac{m(y,s,t)}{m(x,s,t)} m(x,dy) \left(\int_{\mathcal{X}} \frac{m(y,s,t)}{m(x,s,t)} m(x,dy)\right)^{-1}.$$

Comments on the proof. To prove the Many-to-One formula (2.2), we first show that the family of semi-groups $\left(P_{r,s}^{(t)}, r \leq s \leq t\right)$ is uniquely defined as the unique solution of an integro-differential equation (see Lemma 3.2 in [60]). In particular, we need Assumption 3 to prove the uniqueness. Then, the infinitesimal generator $\left(\left(\mathcal{A}_s^{(t)}\right)_{s \leq t}, \mathcal{D}(\mathcal{A})\right)$ of the auxiliary process is obtained by differentiation of the semi-group $\left(P_{r,s}^{(t)}, r \leq s \leq t\right)$ using Assumption 4.

Remark 2.2. *We can also prove a pathwise version of* (2.2) *using a monotone class argument (Theorem 3.1. in* [60]*).*

Unlike previous works on this subject ([41], [45], [22]), the existence of our auxiliary process does not rely on the existence of spectral elements for the mean operator of the branching process. In particular, we can apply this result to models with a varying environment. The dynamic of this auxiliary process heavily depends on the comparison between $m(x,s,t)$ and $m(y,s,t)$, for $x,y \in \mathcal{X}$. It emphasizes several bias due to growth of the population. First, the auxiliary process jumps more than the original process, if jumping is beneficial in terms of number of descendants. This phenomenon of time-acceleration also appears for examples in [21], [55] or [45]. Moreover, the reproduction law favors the creation of a large number of descendant as in [4] and the non-neutrality favors individuals with an "efficient" trait at birth in terms of number of descendants. Finally, a new bias appears on the dynamic of the trait because of the combination of the random evolution of the trait and non-neutrality. Indeed, if the dynamic of the trait is deterministic, we have $\widehat{\mathcal{G}}_s^{(t)} f(x) = \mathcal{G} f(x)$.

2.4. **Ancestral lineage of a uniform sampling at a fixed time in a large population.** The Many-to-One formula (2.2) gives the law of the trait of a typical individual in the population. But so far, we did not prove that it corresponds to the trait of a uniformly sampled individual. In particular, we have to take into account that the number of individuals in the population is stochastic and depends on the dynamic of the trait. Using the law of large numbers, we can approximate the number of individuals in a population starting for n individuals, divided by n, by the mean number of individual in the population. That is why we now look at the ancestral lineage of a uniform sampling in a large population.

It only makes sense to speak of a uniformly sampled individual at time t if the population does not become extinct before time t. For all $t \geq 0$, let $\Omega_t = \{N_t > 0\}$ denote the event of survival of the population. Let $\nu \in \mathcal{M}_P(\mathcal{X})$ be such that:

$$\mathbb{P}_\nu(\Omega_t) > 0.$$

We set

$$\nu_n := \sum_{i=1}^{n} \delta_{X_i},$$

where X_i are i.i.d. random variables with distribution ν. For $t \geq 0$, we denote by $U(t)$ the random variable with uniform distribution on V_t conditionally on Ω_t and by $\left(X_s^{U(t)},\ s \leq t\right)$ the process describing the trait of a sampling along its ancestral lineage. If X is a stochastic process, we denote by X^ν the process with initial distribution $\nu \in \mathcal{M}_P(\mathcal{X})$.

Theorem 2.3. *Under Assumptions 1,2, 3 and 4, for any $t \geq 0$, the following convergence in law in $\mathbb{D}([0,t],\mathcal{X})$ holds:*

$$X_{[0,t]}^{U(t),\nu_n} \underset{n\to+\infty}{\longrightarrow} Y_{[0,t]}^{(t),\pi_t},\ \textit{where } \pi_t(dx) = \frac{m(x,0,t)\nu(dx)}{\int_{\mathcal{X}} m(x,0,t)\nu(dx)}.$$

Remark 2.4. *If we start with n individuals with the same trait x, we obtain:*

$$\mathbb{E}\left[F\left(X_{[0,t]}^{U(t),\nu_n}\right)\right] \underset{n\to+\infty}{\longrightarrow} \mathbb{E}_x\left[F\left(Y_{[0,t]}^{(t)}\right)\right].$$

Therefore, the auxiliary process describes exactly the dynamic of the trait of a uniformly sampled individual in the large population limit, if all the starting individuals have the same trait. If the initial individuals have different traits at the beginning, the large population approximation of a uniformly sampled individual is a linear combination of the auxiliary process.

3. Averaging for some integrate-and-fire models

3.1. **The model.** We examine the following slow-fast hybrid version of integrate-and-fire models [17], used in mathematical neuroscience. In such a setting, X represents the membrane potential of a neural cell which is increasing until it reaches some threshold c, corresponding to the time where a nerve impulse is triggered, and then the potential is reset to some lower value. More precisely, the process $(X(t), t \in [0,T])$, with T some time horizon, obeys the following dynamic:

(1) **Initial state:** At time $T_0^* = 0$, the process starts at $X(T_0^*) = \xi_0$, a random variable with support included in (m,c) where $\{c\}$ is considered as a boundary and $m < c$ is some real.

(2) **First jumping time:** Let Y be a continuous time Markov chain valued in a countable space $\mathcal{Y}$. This chain starts at $Y(0) = \zeta$, a $\mathcal{Y}$-valued random variable. The first hitting time of the boundary occurs at time T_1^* defined as

$$T_1^* = \inf\left\{ t > 0 \ ; \ \xi_0 + \int_{T_0^*}^t \alpha(Y(s))F(X(s))ds = c \right\},$$

where α is a positive measurable function such that $\alpha(\mathcal{Y})$ is bounded from above and F is a positive continuous function.

(3) **Piecewise deterministic motion:** For $t \in [T_0^*, T_1^*)$, we set

$$X(t) = \xi_0 + \int_{T_0^*}^t \alpha(Y(s))F(X(s))ds.$$

The dynamic of X is thus continuous here, and given by the differential equation:

$$\frac{dX}{dt}(t) = \alpha(Y(t))F(X(t)), \quad X(0) = \xi_0.$$

(4) **Jumping measure:** Then, at time T_1^{*-}, the process is constrained to stay inside (m,c) by jumping according to the Y-dependent measure $\mu_{Y_{T_1^{*-}}}$ whose support is included in (m,c):

$$\forall A \text{ Borel subset of } (m,c), \quad \mathbb{P}(X(T_1^*) \in A) = \mu_{Y_{T_1^{*-}}}(A).$$

(5) **And so on:** Go back to step 1 in replacing T_0^* by T_1^* and ξ_0 by $\xi_1 = X(T_1^*)$.

This is not hard to see that the process (X,Y) is a piecewise deterministic Markov process. Moreover, due to the presence of a boundary c, this is also a constrained Markov process. Our aim is, in this quite simple situation, to describe the behavior of the process X at the boundary c, when the dynamic of the underlying process Y is infinitely accelerated.

3.2. **Acceleration and averaged process.** From now on, we assume that the underlying process of celerities, that is the continuous time Markov chain Y, has a fast dynamic, by introducing a (small) timescale parameter ε such that

$$\forall t \geq 0, \quad Y_\varepsilon(t) = Y\left(t/\varepsilon\right).$$

In the same time, to insure a limiting behavior, we assume that Y is positive recurrent with intensity matrix $Q = (q_{zy})_{z,y\in\mathcal{Y}}$ and invariant probability measure π, considered as a row vector. For convenience, let us also define by $\alpha(V)$ the diagonal matrix such that $\mathrm{diag}(\alpha(V)) = \{\alpha(y) \ ; \ y \in \mathcal{Y}\}$.
As ε goes to zero, the process Y_ε converges towards the stationary state associated to Y in the sense that, by the ergodic theorem,

$$\forall t \geq 0, \forall y \in \mathcal{Y}, \quad \lim_{\varepsilon\to 0} \mathbb{P}(Y_\varepsilon(t) = y) = \pi(y).$$

Therefore, as ε goes to zero, the process X_ε, defined as X by replacing Y by Y_ε, should have its dynamic averaged with respect to the measure π. The behavior of the limiting process away from the boundary is indeed not hard to describe, see for example [49] and references therein. Here and later on, $\mathbb{D}\left([0,\tau],\mathbb{R}\right)$ is the space of càdlàg functions from $[0,\tau]$ to $\mathbb{R}$, with τ a time horizon. For technical reasons, we assume that

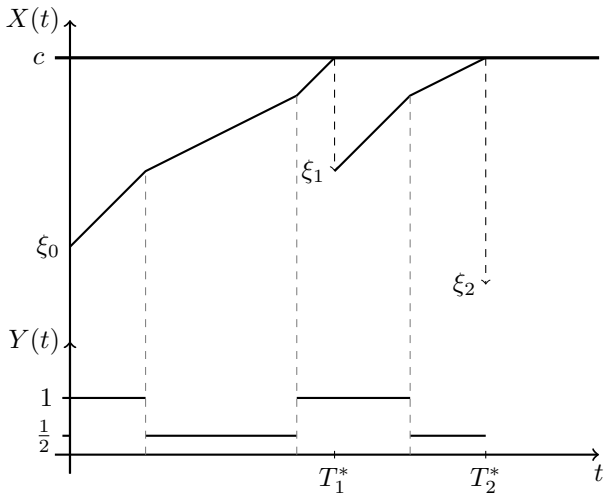


FIGURE 4. A trajectory of X, in a piecewise linear case, $dX(t)/dt = Y(t)$, with Y switching between $1/2$ and 1.

- $\frac{1}{F}$ is integrable over (m, c),
- $\alpha(\mathcal{Y})$ is bounded from above,
- $\mathbb{E}(\sup_{\varepsilon\in(0,1]} p^*_\varepsilon(T))$ is finite, where $p^*_\varepsilon(T)$ is the counting measure at the boundary for the process X_ε.

Proposition 3.1. *Assume that ξ_0 is deterministic. Then, for any $\eta > 0$, the process X_ε converges in law towards a process $\bar{X}$ on $\mathbb{D}\left(\left[0, \frac{c-\xi_0}{\max\alpha(\mathcal{Y})} - \eta\right], \mathbb{R}\right)$ defined as:*

$$\bar{X}(t) = \xi_0 + \sum_{y\in\mathcal{Y}} \alpha(y)\pi(y) \int_0^t F(\bar{X}(s))ds.$$

Proposition 3.1 gives the behavior of the limiting process away from the boundary: celerities are averaged against the measure π. But what happens at boundary? This is what we want to characterize. The following result holds.

Theorem 3.2. *The process X_ε converges in law in $\mathbb{D}([0,T],\mathbb{R})$ towards a process $\bar{X}$ such that for any sufficiently regular function $f : (m, c) \to \mathbb{R}$ the process*

$$f(\bar{X}(t)) - f(\xi_0) - \sum_{y\in\mathcal{Y}} \alpha(y)\pi(y) \int_0^t f'(\bar{X}(s))F(\bar{X}(s))ds - \int_0^t \int_m^c [f(u) - f(\bar{X}(s^-))]\bar{\mu}(du)\bar{p}^*(ds) \tag{3.1}$$

for $t \in [0,T]$, defined a martingale, with $\bar{p}^*$ the counting measure at the boundary for $\bar{X}$. The averaging measure at the boundary $\bar{\mu}$ is defined by

$$\bar{\mu}(du) = \sum_{y\in\mathcal{Y}} \mu_y(du)\pi^*(y)$$

where, for $y \in \mathcal{Y}$, π^* is given by

$$\pi^*(y) = \frac{\pi(y)\alpha(y)}{\sum_{y'\in\mathcal{Y}} \pi(y')\alpha(y')}.$$

We can read in (3.1) the behavior of the averaged process $\bar{X}$. Away from the boundary, the dynamic of $\bar{X}$ is the one of the averaged flow, as indeed described in Proposition 3.1. At the boundary, the averaged jump

measure is not given by the averaged version of the jumping measure, that is

$$\bar{\mu}(du) \neq \sum_{y \in \mathcal{Y}} \mu_y(du)\pi(y),$$

but the values of the celerity at the boundary are taken into account through the presence of α. There is a balance between the celerity intensities and the invariant probabilities. For example, consider a piecewise linear motion whose speed switches at rate $1/\varepsilon$ between 1 and 2. When, ε goes to zero, the process will spend as much time increasing with celerities 1 and 2, but when increasing with celerity 2, the process obviously increases twice more than with celerity one in a same time window, and then approaches the boundary twice more rapidly, and therefore will have in fact twice more chance to hit the boundary with this celerity.
Note that because of the separation of variables in the form of the flow, the value of the process $\bar{X}$ at the boundary does not appear in the expression of π^*, as it could be the case in more general situations. Theorem such as 3.2 can be derived using the Prohorov program : prove the tightness of the family $\{X_\varepsilon, \varepsilon \in (0,1)\}$ and then identify the limit. Tightness for constrained Markov processes, as the process described in the present section, has an interesting and rich literature, see [48] and references therein. We refer to [40] for more details about the proof of Theorem 3.2.

4. Stochastic models of protein production with cell division and gene replication

The production of proteins is the main mechanism of the prokaryotic cells (like bacteria). These functional molecules represent about 3.6×10^6 molecules of 2000 different types and compose about half of the dry mass of the cell [15] and this pool needs to be duplicated in the cell cycle time. In total, it is estimated that 67% of the energy of the cell is dedicated to the protein production [67]. The protein production is a two steps process by which the genetic information is first transcribed into an intermediate molecule, the messenger-RNA (mRNA), and then translated into a protein. Experimental studies have shown that this production is subject to a large variability [37, 62, 68]. This heterogeneity can either be explained by the stochastic nature of the production mechanism (both transcription and translation are due to random encounter between molecules) or the effect of cell scale phenomenon like the DNA replication, cell division, or fluctuation in the resources.

In particular, the extensive experimental study of Taniguchi et al. (2010) [68] measured the variability of protein concentration for more than 1000 different proteins in *E. coli*. For each of them, the authors have measured the average protein concentration, the average mRNA concentration and their lifetime. They interpret different sources of proteins, either from the protein production mechanism itself (i.e. the transcription of mRNAs and the translation of proteins) or from external factors, like the gene replication, division or the sharing of common resources in the production.

One can interpret the experimental results in the light of classical stochastic models of protein production [11, 65, 64] that predict the variability of protein production. But these classical models do not consider important phenomenon that may have an impact on the production: they usually do not consider the replication of the gene at some point in the cell cycle and also do not take into account the random assignment of each protein in either of the two daughter cells at division. Moreover, they only consider the number of mRNAs and proteins without any explicit notion of volume, and therefore the notion of concentration is unclear in this case. Since the experimental study only considers concentrations, it would be difficult to have quantitative comparisons between the variances predicted by these models and the ones experimentally measured.

We therefore propose here a more realistic model of protein production that takes into account all the basic features that can be expected for the production of a type of protein inside a cell cycle: the transcription, the translation, the gene replication and the division. We also explicitly represent the number of mRNAs and proteins in the growing volume of the bacteria, so that we can consider the concentration of each of these entities in the cell. We will then estimate the theoretical contribution of these features to the protein variance through mathematical analysis. We will also make a biological interpretation of these results with respect to experimental measures.

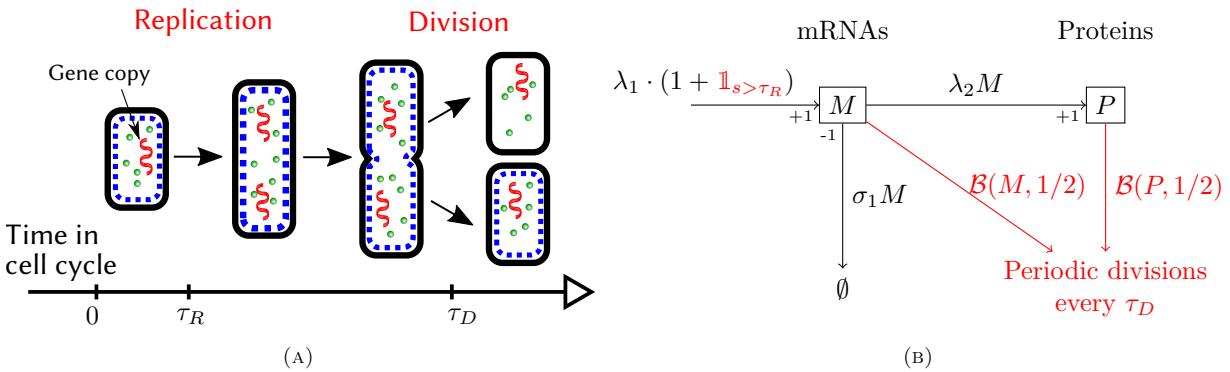


FIGURE 5. Model with gene replication and division. **(A)** The concept of cell cycle in the model. **(B)** Model of production for one protein.

4.1. **Model and Results.** Our model presented here is adapted from a simple gene constitutive model that does not consider any gene regulation (it is a particular case of the model presented in [64]). Like this classical model, our model is gene centered (it represents the production of only one protein), and has two stages (transcription and translation) where both the number of mRNAs, M, and proteins, P, are explicitly represented. Each event (creation or degradation of an mRNA, creation of a protein, etc.) is supposed to occur at random times. The time intervals are considered as exponentially distributed. In addition to this classical framework, we introduce a notion of cell cycle represented by cell growth, division and gene replication. Periodically two events occurs: the replication at a deterministic time τ_R in the cell cycle that doubles the mRNA production rate and the division after a time τ_D starting a new cell cycle. We can then consider the evolution of one type of mRNA and protein through a cell lineage (see Figure 5a).

Overall, the model represents five aspects that intervene in protein production (see Figure 5b):

mRNAs: Messenger-RNAs are transcribed at rate λ_1 before the replication and at rate $2\lambda_1$ after gene replication. When transcription happens, the number M of mRNAs increases by 1. As in the previous model, each mRNA has a lifetime of rate σ_1 (so the global mRNA degradation rate is $\sigma_1 M$).

Proteins: Each mRNA can be translated into a protein at rate λ_2 (so the global protein production rate is $\lambda_2 M$). The number of proteins P is increased by 1. As the protein lifetime is usually of the order of several cell cycles, we consider that the proteins do not degrade.

Gene replication: After a deterministic time τ_R after each division (with $\tau_R < \tau_D$), the gene replication occurs. The gene responsible for the mRNA transcription is replicated, hence doubling the mRNA transcription rate until next division.

Division: Divisions occur periodically after deterministic time interval τ_D. After each division, we only follow one of the two daughter cells (see Figure 5a) so that each mRNA and each protein have a probability 1/2 to be in this next considered cell. The result is a *binomial sampling*: for instance, with M mRNAs just before division, the number of mRNAs after division follows a binomial distribution $\mathcal{B}(n, p)$ with parameters $n = M$ and $p = 1/2$. Moreover, as there is only one copy of the gene in the newborn cell, the mRNA transcription rate is anew set to λ_1 until the next gene replication.

Volume growth: The volume $V(s)$ considers the entire volume of the cell at any moment s of the cell cycle and it increases as the cell grows. One considers the volume growth of the cell as deterministic. In real life experiments, a bacteria volume globally grows exponentially (see [70]) and approximately doubles its volume at the time of division τ_D. As a consequence, in our model, for a time s in the cell cycle, then the volume grows as

$$V(s) = V_0 2^{s/\tau_D}$$

with V_0 the typical size of a cell at birth.

From this, if M_s and P_s denote respectively the number of mRNAs and proteins at time s, the concentrations are

$$\frac{M_s}{V(s)} \text{ and } \frac{P_s}{V(s)}.$$

Our main results are the analytical expressions for the distribution of the number of mRNAs and the first moments of the number of proteins at any time s of the cell cycle, when the system is at equilibrium (ie after many cell divisions).

Proposition 4.1. *At equilibrium, the mRNA number M_s at time s in the cell cycle follows a Poisson distribution of parameter*

$$x_s = \frac{\lambda_1}{\sigma_1}\left[1 - \frac{e^{-(s+\tau_D-\tau_R)\sigma_1}}{2-e^{-\tau_D\sigma_1}} + \mathbb{1}_{s\geq\tau_R}\left(1-e^{-(s-\tau_R)\sigma_1}\right)\right].$$

The proof of this proposition uses the framework of Marked Point Poisson Processes to determine the mRNA distribution at any time of the cell cycle knowing the random variable M_0, the number of mRNA at birth. Then, the distribution of M_0 can be determined as the equilibrium stipulates that the distribution of M is the same at time 0 and just after the next division. The proof will be given in [34] (in preparation).

Proposition 4.2. *At any time $s \in [0, \tau_R[$ before replication, the mean and the variance of P_s are given as a function of the moments of (P_0, M_0) by:*

$$\begin{aligned}
\mathbb{E}[P_s] &= \mathbb{E}[P_0] + \lambda_2\frac{\lambda_1}{\sigma_1}s + \left(x_0 - \frac{\lambda_1}{\sigma_1}\right)\frac{1-e^{-\sigma_1 s}}{\sigma_1},\\
\mathbb{V}\mathrm{ar}[P_s] &= \mathbb{V}\mathrm{ar}[P_0] + 2\lambda_2\frac{1-e^{-\sigma_1 s}}{\sigma_1}\mathbb{C}\mathrm{ov}[P_0, M_0]\\
&\quad \left(\lambda_2\frac{1-e^{-\sigma_1 s}}{\sigma_1}\right)^2 x_0 + x_0\frac{\lambda_2}{\sigma_1}\left(1 - e^{-\sigma_1 s} + \frac{\lambda_2}{\sigma_1}\left[1 - e^{-\sigma_1 s}\left(e^{-\sigma_1 s} + 2s\sigma_1\right)\right]\right) +\\
&\quad + \frac{\lambda_1\lambda_2}{\sigma_1^2}\left[s\sigma_1 - 1 + e^{-\sigma_1 s} + 2\frac{\lambda_2}{\sigma_1}\left(\sigma_1 s\left(1+e^{-\sigma_1 s}\right) - 2\left(1-e^{-\sigma_1 s}\right)\right)\right].
\end{aligned}$$

with x_0 defined in Proposition 4.1.

Moreover, we can show a similar result for $s \in [\tau_R, \tau_D[$, a time after gene replication. Then, at equilibrium, we have that the distribution of the couple (M, P) after division is the same as at the beginning of the cell cycle. This allows us to determine the explicit values for $\mathbb{E}[P_0]$, $\mathbb{V}\mathrm{ar}[P_0]$ and $\mathbb{C}\mathrm{ov}[P_0, M_0]$. These complementary results and the proof of Proposition 4.2 will be available in [34] (in preparation).

4.2. **Discussion.** In order to have realistic values for the parameters λ_1, σ_1, λ_2 and τ_R, we fix them to correspond to the average production of each protein measured in Taniguchi et al. (2010) [68]. It gives groups of parameters that represent a large spectrum of different genes. Then, using Propositions 4.1 and 4.2, as well as the additional results, we can compute the concentration for every gene at any time of the cell cycle.

An example of the evolution of one protein concentration is shown in Figure 6. It appears that the mean concentration at a given time s of the cell cycle $\mathbb{E}[P_s/V(s)]$ is not constant during the cell cycle. The curve of $\mathbb{E}[P_s/V(s)]$ fluctuates around 2% of the global average protein production (denoted by μ_p in the graph). Experimental measures of average protein expression during the cell cycle show similar results: for instance, the expression of genes at different positions on the chromosome have been measured in [69]; it shows a similar profile during the cell cycle and depicts a fluctuation also around 2% of the global average (see Figure 1.d and Figure S6.b of the article).

Overall, for all the genes considered, the effect of the cell cycle (the fluctuation of $\mathbb{E}[P_s/V(s)]$ around μ_p) is small compared to the standard deviation at each moment of the cell cycle $\sqrt{\mathbb{V}\mathrm{ar}[P_s/V(s)]}$ (the coloured areas in Figure 6). We have determined the range of parameters where, on the contrary, the fluctuations due to the cell cycle would be higher than the inherent variability of the system. It appears that such regime is possible

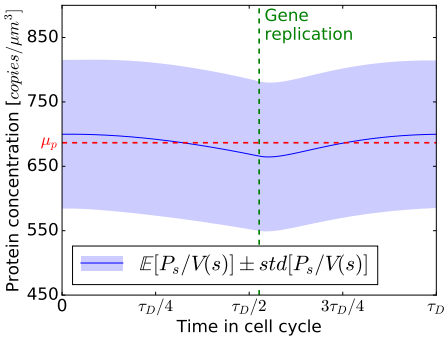
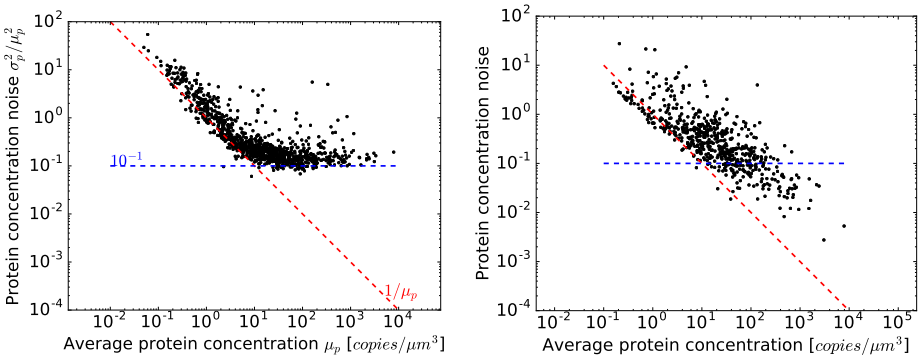


FIGURE 6. Profile example of the gene Adk, the central line represents the mean protein concentration over the cell cycle, and the coloured areas represent the standard-deviation in the models.



(A) Coefficient of variation in Taniguchi et al. (2010)

(B) Coefficient of variation in our model

FIGURE 7. Coefficient of variation of each proteins. **(A)** Each dot represents the coefficient of variation (CV) of the concentration for a particular protein (the CV is defined as σ_p^2/μ_p^2 with σ_p^2 the variance and μ_p the mean) experimentally measured in [68]. It clearly appears two regimes: the first one (for $\mu_p < 10$) shows a CV inversely proportional to the mean; the second regime (for $\mu_p > 10$) shows a CV independent of the average production. **(B)** Same graph obtained with the proteins of our model. Even if the CV is roughly inversely proportional to the average production, there is no lower bound limit for highly expressed proteins.

only for highly active gene (high λ_1) and for mRNAs not very active (low λ_2) and which exist for short periods of time (high σ_1). Biologically, such regime seems unrealistic because the cost of mRNA production would be very high.

In Figure 7, we compare the experimental results obtained in the article of Taniguchi et al. (2010) [68] and our analytical expressions obtained in our model, for the global coefficient of variation (CV) of protein concentration as a function of the average concentration. In the experimental measures, it clearly appears two regimes for the CV of protein concentration: for low expressed proteins (mean protein concentration < 10), the CV roughly scales inversely with the average concentration; for genes with higher protein production (mean protein concentration > 10), the CV becomes independent of the average protein production level, the plateau

is around 10^{-1}. In our model, the global tendency of the CV approximately inversely scales the average protein concentration, and there is no lower bound limit, as in the experiments.

We have therefore shown through this model that both DNA replication and division can not explain the second regime observed in Figure 7a. Yet the authors of the experimental study propose another possible origin for the noise observed in highly produced proteins; they propose that it is due to fluctuations in the availability of RNA-polymerases and ribosomes. These two macro-molecules are shared among the productions of all the different proteins because they respectively participate to the transcription of mRNAs and the translation of proteins. Since they are present in limited quantities, fluctuations in their availability could indeed have repercussions on the protein noise. We will present in [34] an extension of the model presented here that takes into account this sharing of ribosomes and RNA-polymerases in the production of all proteins of the bacteria, and the impact it has on the protein noise.

Acknowledgements Bertrand Cloez would like to thank Romain Yvinec for his enthusiastic and stimulating organization of the "Journées MAS 2016" and Florent Malrieu for many interesting discussions on the subject of his section. The research of Bertrand Cloez was partially supported by the Agence Nationale de la Recherche PIECE 12-JS01-0006-01 and by the Chaire "Modélisation Mathématique et Biodiversité " . The research of Aline Marguet was partially supported by the Chaire "Modélisation Mathématique et Biodiversité". Renaud Dessalles would like to thank Philippe Robert and Vincent Fromion for their support.

References

[1] R. Azaïs, J.-B. Bardet, A. Génadot, N. Krell, and P.-A. Zitt. Piecewise deterministic Markov process – recent results. *ESAIM: Proceedings*, 44:276–290, 2014.
[2] Y. Bakhtin and T. Hurth. Invariant densities for dynamical systems with random switching. *Nonlinearity*, 25:2937, 2012.
[3] K. Ball, T. G. Kurtz, L. Popovic, and G. Rempala. Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Probab.*, 16(4):1925–1961, 2006.
[4] V. Bansaye, J.-F. Delmas, L. Marsalle, and V. C. Tran. Limit theorems for Markov processes indexed by continuous time Galton-Watson trees. *Ann. Appl. Probab.*, 21(6):2263–2314, 2011.
[5] V. Bansaye and V. Vatutin. On the survival probability for a class of subcritical branching processes in random environment. *Bernoulli*, 23(1):58–88, 2017.
[6] J.-B. Bardet, H. Guérin, and F. Malrieu. Long time behavior of diffusions with Markov switching. *ALEA Lat. Am. J. Probab. Math. Stat.*, 7:151–170, 2010.
[7] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt. Quantitative ergodicity for some switched dynamical systems. *Electron. Commun. Probab.*, 17(56), 2012.
[8] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt. On the stability of planar randomly switched systems. *Ann. Appl. Probab.*, 24(1):292–311, 2014.
[9] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt. Qualitative properties of certain piecewise deterministic Markov processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(3):1040–1075, 2015.
[10] M. Benaïm and C. Lobry. Lotka Volterra in fluctuating environment or "how switching between beneficial environments can make survival harder". *Ann. Appl. Probab.*, 26(6):3754–3785, 2016.
[11] O. G. Berg. A model for the statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.*, 71(4):587–603, 1978.
[12] P. Bertail, S. Clémençon, and J. Tressou. A storage model with random release rate for modeling exposure to food contaminants. *Math. Biosci. Eng.*, 5(1):35–60, 2008.
[13] J. Bertoin. Compensated fragmentation processes and limits of dilated fragmentations. *Ann. Prob.*, 44(2):1254–1284, 2016.
[14] F. Bouguet. Quantitative speeds of convergence for exposure to food contaminants. *ESAIM Probab. Stat.*, 19:482–501, 2015.
[15] H. Bremer and P. P. Dennis. Modulation of chemical composition and other parameters of the cell at different exponential growth rates. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. ASM Press, second edition, 1996.
[16] P. C. Bressloff. *Stochastic processes in cell biology*. Springer, 2014.
[17] A. Burkitt. Averaging for some simple constrained markov processes. *Biological cybernetics*, 95(1):1–19, 2006.
[18] F. Campillo and C. Fritsch. Weak convergence of a mass-structured individual-based model. *Appl. Math. Opt.*, 72(1):37–73, 2014.
[19] F. Campillo, M. Joannides, and I. Larramendy-Valverde. Stochastic modeling of the chemostat. *Ecol. Model.*, 222(15):2676–2689, 2011.
[20] N. Champagnat and H. Benoit. Moments of the frequency spectrum of a splitting tree with neutral poissonian mutations. *Electron. J. Probab.*, 21, 2016.

[21] B. Chauvin and A. Rouault. KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees. *Probab. Theory Related Fields*, 80(2):299–314, 1988.
[22] B. Cloez. Limit theorems for some branching measure-valued processes. *arXiv:1106.0660*, 2011.
[23] B. Cloez and C. Fritsch. Gaussian approximations for chemostat models in finite and infinite dimensions. *J. Math. Biol.*, 2017. In Press.
[24] B. Cloez and M. Hairer. Exponential ergodicity for Markov processes with random switching. *Bernoulli*, 21(1):505–536, 2015.
[25] P. Collet, S. Martínez, S. Méléard, and J. San Martín. Stochastic models for a chemostat and long-time behavior. *Adv. in Appl. Probab.*, 45(3):822–836, 2013.
[26] M. Costa. A piecewise deterministic model for a prey-predator community. *Ann. Appl. Probab.*, 26(6):3491–3530, 2016.
[27] O.L.V. Costa and F. Dufour. Stability and ergodicity of piecewise deterministic Markov processes. *Proc. IEEE Conf. Decis. Control*, 47:1525–1530, 2008.
[28] A. Crudu, A. Debussche, A. Muller, and O. Radulescu. Convergence of stochastic gene networks to hybrid piecewise deterministic processes. *Ann. Appl. Probab.*, 22:1822–1859, 2012.
[29] K. S. Crump and W.-S. C. O'Young. Some stochastic features of bacterial constant growth apparatus. *Bull. Math. Biol.*, 41(1):53–66, 1979.
[30] M. H. A. Davis. Piecewise-deterministic markov processes: a general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B*, 46(3):353–388, With discussion, 1984.
[31] M. H. A. Davis. *Markov models and optimization*, volume 49. Monographs on Statistics and Applied Probability, Chapman & Hall, 1993.
[32] B. de Saporta and J.-F. Yao. Tail of a linear diffusion with Markov switching. *Ann. Appl. Probab.*, 15(1B):992–1018, 2005.
[33] R. Dessalles, V. Fromion, and P. Robert. A stochastic analysis of autoregulation of gene expression. *J. Math. Biol.*, 2017. In Press.
[34] R. Dessalles, V. Fromion, and P. Robert. Stochastic models for gene expression with cell division, gene replication and protein production interactions. 2017. In preparation.
[35] S. Ditlevsen and E. Löcherbach. Multi-class oscillating systems of interacting neurons. *Stoch. Proc. Appl.*, 127(6):1840–1869, 2017.
[36] M. Doumic, M. Hoffmann, N. Krell, and L. Robert. Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *Bernoulli*, 21(3):1760–1799, 2015.
[37] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.
[38] J. Fontbona, H. Guérin, and F. Malrieu. Long time behavior of telegraph processes under convex potentials. *Stoch. Proc. Appl.*, 126(10):3077–3101, 2016.
[39] N. Fournier and S. Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.*, 14(4):1880–1919, 2004.
[40] A. Genadot. Averaging for some simple constrained markov processes. *arXiv:1606.07577*.
[41] H.-O. Georgii and E. Baake. Supercritical multitype branching processes: the ancestral types of typical individuals. *Adv. Appl. Probab.*, 35(4):1090–1110, 2003.
[42] P. Haccou, P. Jagers, and V. A. Vatutin. *Branching processes: variation, growth, and extinction of populations*, volume 5 of *Cambridge Studies in Adaptive Dynamics*. Cambridge University Press, Cambridge; IIASA, Laxenburg, 2007.
[43] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011.
[44] M. Hairer, J. C. Mattingly, and M. Scheutzow. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields*, 149(1-2):223–259, 2011.
[45] R. Hardy and S. C. Harris. A spine approach to branching diffusions with applications to L_p-convergence of martingales. In *Séminaire de Probabilités XLII*, pages 281–330. Springer, 2009.
[46] B. Hepp, A. Gupta, and M. Khammash. Adaptive hybrid simulations for multiscale stochastic reaction networks. *J. Chem. Phys.*, 142(3):034118, 2015.
[47] P. Kouretas, K. Koutroumpas, J. Lygeros, and Z. Lygerou. *Stochastic Hybrid Systems*, chapter Stochastic hybrid modeling of biochemical processes. Taylor & Francis, 2006.
[48] T. Kurtz. *Recent Advances in Stochastic Calculus*, chapter Martingale problems for constrained Markov processes. Springer Verlag, New York, 1990.
[49] T. Kurtz. *Applied Stochastic Analysis*, chapter Averaging for martingale problems and stochastic approximation., pages 186–209. Springer, 1992.
[50] T. G. Kurtz and H.-W. Kang. Separation of time-scales and model reduction for stochastic reaction networks. *Ann. Appl. Probab.*, 23(2):529–583, 2013.
[51] A. Lambert. The contour of splitting trees is a Lévy process. *Ann. Probab.*, 38(1):348–395, 2010.
[52] A. Lasota and M. C. Mackey. Cell division and the stability of cellular populations. *J. Math. Biol.*, 38(3):241–261, 1999.

[53] A. Lasota, M. C. Mackey, and J. Tyrcha. The statistical dynamics of recurrent biological events. *J. Math. Biol.*, 30(8):775–800, 1992.
[54] S. D. Lawley, J. C. Mattingly, and M. C. Reed. Sensitivity to switching rates in stochastically switched ODEs. *Commun. Math. Sci.*, 12(7):1343–1352, 2014.
[55] R. Lyons, R. Pemantle, and Y. Peres. Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Prob.*, 23(3):1125–1138, 1995.
[56] M. C. Mackey, M. Tyran-Kamińska, and R. Yvinec. Dynamic behavior of stochastic gene expression models in the presence of bursting. *SIAM J. Appl. Math.*, 73(5):1830–1852, 2013.
[57] F. Malrieu. Some simple but challenging Markov processes. *Ann. Fac. Sci. Toulouse Math. (6)*, 24(4):857–883, 2015.
[58] F. Malrieu and T. Hoa Phu. Lotka-Volterra with randomly fluctuating environments: a full description. *arXiv:1607.04395*, 2016.
[59] F. Malrieu and P.-A. Zitt. On the persistence regime for Lotka-Volterra in randomly fluctuating environments. *arXiv:1601.08151*, 2016.
[60] A. Marguet. Uniform sampling in a structured branching population. *arXiv:1609.05678*, 2016.
[61] P. Monmarché. Piecewise deterministic simulated annealing. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(1):357–398, 2016.
[62] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden. Regulation of noise in the expression of a single gene. *Nat. Genet.*, 31(1):69–73, 2002.
[63] K. Pakdaman, M. Thieullen, and G. Wainrib. Fluid limit theorems for stochastic hybrid systems with application to neuron models. *Adv. in Appl. Probab.*, 42(3):761–794, 2010.
[64] J. Paulsson. Models of stochastic gene expression. *Phys. Life Rev.*, 2(2):157–175, 2005.
[65] D. R. Rigney and W. C. Schieve. Stochastic model of linear, continuous protein synthesis in bacterial populations. *J. Theor. Biol.*, 69(4):761–766, 1977.
[66] R. Rudnicki and M. Tyran-Kamińska. *Semigroups of Operators -Theory and Applications.*, volume 113, chapter Piecewise deterministic Markov processes in biological models. Springer Proceedings in Mathematics & Statistics, 2015.
[67] J. B. Russell and G. M. Cook. Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol. Rev.*, 59(1):48–62, 1995.
[68] 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(5991):533–538, 2010.
[69] N. Walker, P. Nghe, and S. J. Tans. Generation and filtering of gene expression noise by the bacterial cell cycle. *BMC Biol.*, 14:11, 2016.
[70] P. Wang, L. Robert, J. Pelletier, W. L. Dang, F. Taddei, A. Wright, and S. Jun. Robust growth of Escherichia coli. *Curr. Biol.*, 20(12):1099–1103, 2010.
[71] G. G. Yin and C. Zhu. *Hybrid switching diffusions*, volume 63 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2010.
[72] R. Yvinec, C. Zhuge, J. Lei, and M. C. Mackey. Adiabatic reduction of a model of stochastic gene expression with jump Markov process. *J. Math. Biol.*, 68(5):1051–1070, 2014.