

PIECEWISE-DETERMINISTIC MARKOV PROCESSES AS LIMITS OF MARKOV JUMP PROCESSES

UWE FRANZ,* *University of Franche-Comté*

VOLKMAR LIEBSCHER,** *Ernst Moritz Arndt University Greifswald*

STEFAN ZEISER,*** *Kinesis Pharma BV*

Abstract

A classical result about Markov jump processes states that a certain class of dynamical systems given by ordinary differential equations are obtained as the limit of a sequence of scaled Markov jump processes. This approach fails if the scaling cannot be carried out equally across all entities. In the present paper we present a convergence theorem for such an unequal scaling. In contrast to an equal scaling the limit process is not purely deterministic but still possesses randomness. We show that these processes constitute a rich subclass of piecewise-deterministic processes. Such processes apply in molecular biology where entities often occur in different scales of numbers.

Keywords: Piecewise-deterministic Markov process; stochastic model of gene regulation; limit theorem; Skorokhod space

2010 Mathematics Subject Classification: Primary 60F15

Secondary 60J28; 92C40

1. Introduction

In molecular biology there is a growing interest in a stochastic description of the underlying chemical reactions. In particular, this is the case in functional genomics where experimental findings provided evidence that gene expression is rather a stochastic than a deterministic process [3], [4], [9].

Gillespie [7], [8] showed that the dynamics of discrete entities in a well-stirred physical or biological system is well described by Markov jump processes. Moreover, if the number of each species is high, such systems can often be described by ordinary differential equations. This fact is substantiated by the work of T. G. Kurtz, who showed that, under an equal scaling of the particle numbers, the corresponding Markov jump processes can be approximated by ordinary differential equations; see [10], [11], and [12]. However, if the scaling cannot be made equal, such an approach may fail. A case-by-case study of unequal scaling can be found in [1]. Important examples of such situations can be found in molecular biology, in particular in genetics where genes normally occur in two-digit numbers but the number of the gene products, i.e. the mRNAs and the proteins, can achieve hundreds or thousands. Although the dynamical behaviour of gene regulatory networks can often be well described by systems of ordinary differential equations, some of the features of the stochastic nature of such systems

Received 14 September 2010; revision received 14 February 2012.

* Postal address: Faculty of Mathematics of Besançon, University of Franche-Comté, Route de Gray 16, 25 030 Besançon cedex, France. Email address: uwe.franz@univ-fcomte.fr

** Postal address: Faculty of Mathematics and Sciences, Ernst Moritz Arndt University Greifswald, Walther-Rathenau-Straße 47, 17487 Greifswald, Germany. Email address: volkmar.liebscher@uni-greifswald.de

*** Postal address: Kinesis Pharma BV, Lage Mosten 29, 4822 NK Breda, The Netherlands.

Email address: stefan.zeiser@kinesis-pharma.com

get lost. An alternative approach to describe systems possessing species appearing in high and low numbers is to use a class of piecewise-deterministic Markov processes (PDMPs) [5]. In several recently published works gene regulatory networks were modelled by such a class of stochastic processes; see [15], [16], and [17].

PDMPs were introduced by M. H. Davis in a more general setting as a particular class of stochastic processes in 1984. They form a family of Markov processes which move deterministically interrupted by random jumps. The motion of a PDMP is determined by three parameters (v, λ, Q) , which are sometimes called the local characteristics. The vector field v determines the deterministic movement of the PDMP. Starting from an initial point in the state space, the motion of the process follows the corresponding flow until the first jump time is reached. This jump time is triggered by a transition rate function λ . The distribution at the jump time is given by the transition kernel Q . After the jump the process restarts in a deterministic way until the next jump time is reached.

In the present work we will show that such an approach can be justified in a rigorous way. To this end, we will use an alternative representation of PDMPs. We use a certain time-change representation which is similar to that used by Kurtz in his approximation results. However, now the reactants are partitioned into two subpopulations, described by discrete and continuous variables. The slow reactions causing changes of the discrete variable are triggered by independent Poisson processes, whereas the fast reactions causing temporal changes of the continuous variables are given as deterministic solutions of integral equations. Thus, in contrast to the representation used in [12], now, besides terms involving Poisson processes, there occur terms describing purely deterministic motion too. Moreover, since even the limiting process is discontinuous, we need a mode of convergence different from uniform convergence. For our purpose, the Skorokhod J_1 -topology proves useful and strong enough to also derive convergence of path properties.

As an example, we consider a small autocatalytic genetic network. The system consists of an inducible gene with one binding site in its promoter, and of proteins which are present as monomers. The gene is activated by binding of a translated protein to the binding site in the regulatory region of the gene. In a recently published paper [17] we have modelled such a system by a Markov jump process as well as by a PDMP. For both systems, we showed by numerical means that the expression of the gene expires in finite time. Moreover, in [15] we proved that the point mass in zero is indeed the only stationary distribution of the system modelled by a PDMP. In the present work we show that the system fulfills the conditions of the presented convergence theorem. This way the numerical results from [17] and the mathematical results from [15] find their deeper justification in convergence of *any* reasonable path property.

The paper is structured as follows. In Section 2 we give a formal definition of a class of PDMPs which covers, among other possibilities, our model for certain two-time-scale gene regulatory networks. In Section 3 we present an alternative description of PDMPs given by the abovementioned time-change representation. Section 4 contains the convergence theorem for partially scaled Markov jump processes to a PDMP. In Section 5, we discuss the autocatalytic network example mentioned above. All proofs and some additional lemmas can be found in Appendices A and B.

2. PDMPs

To define the class of PDMPs, we follow the construction of Davis [5] in a somewhat simpler setting. We will mention the few simplifications below. The state space and the dynamics of a PDMP $X = (X_t)_{t \geq 0}$ are defined as follows. Let M be a discrete set. Furthermore, $d \in \mathbb{N}$

together with $E \subset \mathbb{R}^d$ will be fixed. The state space of X is given by

$$S = M \times E. \tag{2.1}$$

Let S be provided with the Borel σ -algebra $\mathcal{B}(S)$. The state space S can be considered as a disjoint union of distinct E , labelled by an element of M . Later, we assume without loss of generality that $M \subset \mathbb{N}_0$ such that $S \subset \mathbb{R}^{d+1}$ safely. We will call a set $\{i\} \times E$ an arm or a branch of the state space.

The deterministic motion of the process on a particular branch $\{i\} \times E$ is specified by a vector field given by locally Lipschitz continuous maps

$$v^{(i)}: \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad i \in M.$$

Throughout this paper, we will assume that, for all these vector fields, there is no blow up in finite time. Thus, the corresponding flow $\Phi_i(z_0, t)$ is uniquely determined by the differential equation $dz/dt = v^{(i)}(z)$ with $z(0) = z_0$ for all $t \geq 0$. That is, $\Phi_i(z_0, 0) = z_0$ and

$$\frac{d}{dt} \Phi_i(z_0, t) = v^{(i)}(\Phi_i(z_0, t)).$$

Jumps of the PDMP from one branch to another are governed by a transition rate function

$$\lambda: S \rightarrow [0, \infty),$$

which is assumed to be positive and measurable. The function λ is a measure for the jump intensity of the process, i.e. it determines the distribution of the waiting times of the process in a particular branch. The post jump location is determined by a transition kernel

$$Q: S \times \mathcal{B}(S) \rightarrow [0, 1].$$

It specifies the distribution of the process at the instant when a jump has occurred. We assume that $Q(x, \{x\}) = 0$ for all $x \in S$ so that each jump is a ‘proper’ jump.

With these ingredients we are ready to describe the temporal evolution of the process. Let the state of a random variable $X_t: \Omega \rightarrow S$ at any time instant $t \in [0, \infty)$ be described by a two-dimensional vector $X_t = (I_t, Z_t)$. The sample paths of the PDMP X with above local characteristics $(v^{(i)}, \lambda, Q)$ are constructed as follows. At time $t = 0$ the process X starts at some $x_0 = (i_0, z_0)$. Consider a random variable τ_1 (the waiting time till the first jump) having distribution

$$P(\tau_1 > t) = \exp\left(-\int_0^t \lambda(i_0, \Phi_{i_0}(z_0, u)) du\right).$$

Then, the S -valued random variable (I_{τ_1}, Z_{τ_1}) is selected independently having distribution

$$Q((i_0, \Phi_{i_0}(z_0, \tau_1)), \cdot).$$

In this way, the process X_t is defined up to the first jump time, and its trajectory reads

$$X_t = (I_t, Z_t) = \begin{cases} (i_0, \Phi_{i_0}(z_0, t)), & t \in [0, \tau_1), \\ (I_{\tau_1}, Z_{\tau_1}), & t = \tau_1. \end{cases}$$

In our applications, we may assume that the concentrations of the continuous entities (the Z -part of X) do not change at a jump time. This implies that $Z_{\tau_1} = \Phi_{i_0}(z_0, \tau_1)$. Starting from X_{τ_1} , the

sojourn time $\tau_2 - \tau_1$ in the branch $\{I_{\tau_1}\} \times E$ and the post-jump location are selected in a similar way. Proceeding in this way we obtain a piecewise-deterministic trajectory with random jump times τ_1, τ_2, \dots . Let $N_t = \sum_{i=1}^{\infty} \mathbf{1}_{\{t \geq \tau_i\}}$ denote the random variable giving the numbers of jumps in $[0, t]$. To ensure that the process is uniquely determined for all $t > 0$ almost surely (a.s.), we suppose that $N_t < \infty$ for all t a.s. Hence, the explosion time $\tau_{\infty} = \lim_{i \rightarrow \infty} \tau_i$ is infinite a.s. The so-constructed process is a PDMP.

Definition 2.1. A stochastic process $(X_t)_{t \geq 0}$ on a probability space (Ω, \mathcal{F}, P) and state space S defined in (2.1) is called a PDMP if its dynamics are given by the three local characteristics $(v^{(i)}, i \in M; \lambda; Q)$ and constructed as described above.

The above defined PDMPs constitute a subclass of PDMPs defined by Davis in a more general setting in which the set E may differ among the branches and the integral curves may also hit the boundary of the sets E_i ; see [5]. The latter is excluded by us through the nonexplosion condition for the vector fields $v^{(i)}$. In [5] it was shown that the defined processes possess the strong Markov property and a characterisation of the extended generator was given.

3. PDMP as a time-change representation

In the following we describe the subclass of PDMPs from the previous section using a certain time-change representation. It is similar to that used in [12] for a Markov jump process. In contrast thereto, our representation also contains a deterministic integral. Again, the state space S is given by (2.1). In the following sections we require the set M to be finite. Furthermore, we identify M with a subset of natural numbers such that $S \subseteq \mathbb{N}_0 \times \mathbb{R}^d \subset \mathbb{R}^{d+1}$.

Let $e_k \in \mathbb{Z} \times \mathbb{R}^d, 1 \leq k \leq J, J \geq 1$, be transition vectors of the form

$$e_k = (m_k, 0, \dots, 0),$$

where $m_k \in \mathbb{Z}$. These transition vectors are fixed jump sizes between discrete elements of the set M . Furthermore, we need J transition rate functions r_k , and J independent Poisson processes Π_k which trigger the random times of the discrete transitions. The deterministic movement between the jumps is triggered by a vector field $V: S \rightarrow \mathbb{R}^{d+1}$ such that $V(i, z) = (0, v^{(i)}(z))$ for all $i \in m$.

Let (Ω, \mathcal{F}, P) be a probability space. Throughout the paper, let $\Pi_k: \Omega \rightarrow \mathbb{N}_0, 1 \leq k \leq J$, be independent right-continuous Poisson processes with rate 1.

Definition 3.1. Let (Ω, \mathcal{F}, P) be a probability space, $\{e_k\}_{1 \leq k \leq J}$ be $(d + 1)$ -dimensional real vectors, $r_k: S \rightarrow \mathbb{R}, 1 \leq k \leq J$, be nonnegative locally Lipschitz functions, $V: S \rightarrow \mathbb{R}^{d+1}$ be a locally Lipschitz function, and let $x_0 \in S$ be fixed.

A stochastic process $X = (X_t)_{t \geq 0}$ with $X_t: \Omega \rightarrow S$ is said to obey the time-change representation with respect to $(\Pi_k)_{k=1}^J$ determined by $(e_k)_{1 \leq k \leq J}, (r_k)_{1 \leq k \leq J}$, and V if it solves the equation

$$X_t = x_0 + \sum_{k=1}^J e_k \Pi_k \left(\int_0^t r_k(X_s) ds \right) + \int_0^t V(X_s) ds \quad \text{a.s.} \tag{3.1}$$

for all $t \geq 0$.

Observe that, if X is given by the time-change representation, the vector field V cannot possess explosions and $\int_0^t r_k(X_s) ds$ has to stay finite for all t . Thus, the random variable N_t

which counts the number of jumps in the time interval $[0, t]$ is finite. The stochastic process in the above definition starts in the fixed state x_0 and moves deterministically determined by the vector field V as long as t has not reached the first jump time of any of the Poisson processes Π_k , $1 \leq k \leq J$. As soon as it has reached the smallest jump time, the respective Poisson process jumps and the deterministic movement, determined by the vector field V , changes. Hence, X exhibits behaviour very similar to a PDMP. Indeed, the main goal of this section is to show that X is a PDMP. Observe that we just use the structure (3.1) for this, i.e. S could be a more general subspace of \mathbb{R}^{d+1} than $M \times E$.

We can regard the integral terms inside the Poisson processes in (3.1) as time transformations and denote them sometimes by

$$\beta_k(t) = \int_0^t r_k(X_s) ds, \quad 1 \leq k \leq J.$$

Such a representation determines $(X_t)_{t \geq 0}$ uniquely.

Proposition 3.1. *If $X = (X_t)_{t \geq 0}$ solves the random time change representation (3.1) then this solution is unique, i.e. if $\tilde{X} = (\tilde{X}_t)_{t \geq 0}$ is another solution to (3.1), it holds that*

$$X_t = \tilde{X}_t \quad \text{a.s. for all } t \geq 0.$$

As already mentioned, the aim of the present section is to show that a process solving the time-change representation (3.1) is a PDMP.

Theorem 3.1. *Let $X = (X_t)_{t \geq 0}$ be the solution of the time-change representation (3.1). It holds that*

- (i) X possesses the strong Markov property,
- (ii) X is a PDMP.

With a little more effort we can show that a PDMP given by a time-change representation fulfills the Feller property as well. However, we omit this result since we do not need it for our purposes.

4. Convergence of pure jump processes to PDMPs

In this section we formulate the convergence theorem, which states that, under appropriate scaling, a family of Markov jump processes converges pathwise to a certain type of PDMP. Our method of proof relies on the fact that both the family of Markov jump processes and the PDMP obey a time-change representation with respect to the same Poisson processes. This technique has been used in [10], but with deterministic limits only.

The members of the one-parameter family of Markov processes $\{(X_t^{(N)})_{t \geq 0} : N \geq 1\}$ are of the form

$$X_t^{(N)} = x_0^{(N)} + \sum_{k=1}^K e_k^{(N)} \Pi_k \left(\int_0^t r_k^{(N)}(X_s^{(N)}) ds \right), \tag{4.1}$$

where the Π_k are independent Poisson processes, the $e_k^{(N)}$ are $(d + 1)$ -dimensional real vectors, and the $r_k^{(N)} : \mathbb{R}^{d+1} \rightarrow \mathbb{R}_+$ are real functions for $k = 1, \dots, K$. In contrast to the convergence results of Kurtz we do not scale the transition vectors e_k equally. That is, we scale one part of the original Markov jump process by $1/N$ and leave the other part unscaled. Only the

transitions triggered by reaction channels including reactants in high numbers are scaled by the parameter N . Hence, we need to separate the class of transition vectors $e_{1 \leq k \leq K}^{(N)}$ into a set

$$\{e_k^{(N)} : 1 \leq k \leq J\}, \tag{4.2}$$

which converges to the vectors $\{e_k \in \mathbb{R}^{d+1} : 1 \leq k \leq J\}$, and into a set

$$\{e_k^{(N)} : J + 1 \leq k \leq K\}, \tag{4.3}$$

where the $N e_k^{(N)}$ converge to the vectors $\{e_k \in \mathbb{R}^{d+1} : J + 1 \leq k \leq K\}$. We call transitions induced by elements of (4.2) big jumps, and transitions induced by elements of (4.3) small jumps. We want to conclude that the sequence $(X_t^{(N)})_{N \geq 1}$ of scaled Markov jump processes converges a.s. to a PDMP given as the solution of

$$X_t = x_0 + \sum_{k=1}^J e_k \Pi_k \left(\int_0^t r_k(X_s) ds \right) + \int_0^t V(X_s) ds \tag{4.4}$$

with the special vector field $V(x) = \sum_{k=J+1}^K e_k r_k(x)$.

Before we specify the kind of convergence, we collect several attributes in a global condition which is assumed to be fulfilled throughout the rest of the paper.

Condition 4.1. (PDMP conditions.) *Let $(X_t^{(N)})_{t \geq 0}$, $N \geq 1$, be Markov jump processes, and let $(X_t)_{t \geq 0}$ be a PDMP solving (4.1) and (4.4), respectively. For $x_0^{(N)}$, x_0 , $r_k^{(N)}$, r_k , $e_k^{(N)}$, and e_k , $k = 1, \dots, K$, the following statements are assumed to hold.*

- (i) $\lim_{N \rightarrow \infty} x_0^{(N)} = x_0$.
- (ii) $e_k^{(N)} \rightarrow e_k$ as $N \rightarrow \infty$ for $k = 1, \dots, J$.
- (iii) $N e_k^{(N)} \rightarrow e_k$ as $N \rightarrow \infty$ for $k = J + 1, \dots, K$.
- (iv) $r_k^{(N)} \rightarrow r_k$ as $N \rightarrow \infty$ for $k = 1, \dots, J$ uniformly on compacts.
- (v) $(1/N)r_k^{(N)} \rightarrow r_k$ as $N \rightarrow \infty$ for $k = J + 1, \dots, K$ uniformly on compacts.
- (vi) r_k is locally Lipschitz on \mathbb{R}^{d+1} for all $k = 1, \dots, K$.
- (vii) *There is a locally bounded function $c: \mathbb{R}^{d+1} \rightarrow \mathbb{R}_+$ such that, for all $x \in E$, all $L, N \in \mathbb{N}$, and every sequence $(k_l)_{l \in \mathbb{N}} \subset \{J + 1, \dots, K\}^{\mathbb{N}}$ with $x + \sum_{l=1}^L e_{k_l}^{(N)} \in E$, it holds that*

$$\frac{1}{N} r_k^{(N)} \left(x + \sum_{l=1}^L e_{k_l}^{(N)} \right) \leq c(x) \left(\frac{L}{N} + 1 \right).$$

Let us briefly comment on the statements of Condition 4.1. Statement (i) means convergence of the initial conditions. Statement (ii) is the scaling of the large jumps, whereas (iii) describes the scaling of the small jumps. Likewise, (iv) and (v) describe the scaling of the transition rates of the large and small jumps, respectively. Statement (vi) is used for uniqueness of the process and (vii) is a uniform nonexplosion condition on $(X_t^{(N)})$.

In the following we will show that the family of scaled Markov jump processes converges to a PDMP a.s. In contrast to [10] we cannot hope for uniform convergence since the limiting process still has jumps and uniform convergence does not allow for flexible jumps. In our opinion, the correct mode of convergence is given by the Skorokhod or J_1 -metric [14] on the

space D of càdlàg paths (i.e. right continuous functions from $[0, \infty)$ to \mathbb{R}^{d+1} with left limits everywhere). Let us describe this metric briefly.

Let Λ be the family of Lipschitz continuous functions $\lambda: [0, \infty) \rightarrow [0, \infty)$ which are strictly increasing and onto. For those functions, we declare

$$\gamma(\lambda) = \sup_{s>t \geq 0} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right|.$$

Now, for two càdlàg functions $x, y \in D$, and $\lambda \in \Lambda$ and $u \geq 0$, we set

$$d(x, y, \lambda, u) = \sup_{t \geq 0} \frac{\|x(t \wedge u) - y(\lambda(t) \wedge u)\|}{1 + \|x(t \wedge u) - y(\lambda(t) \wedge u)\|}.$$

We can define the Skorokhod distance of x and y as

$$d_S(x, y) = \inf_{\lambda \in \Lambda} \left(\max \left\{ \gamma(\lambda), \int_0^\infty e^{-u} d(x, y, \lambda, u) du \right\} \right).$$

With these preparations we are ready to state the convergence theorem for PDMPs.

Theorem 4.1. *Let $(X_t^{(N)})_{t \geq 0}$, $N \in \mathbb{N}$, be a sequence of Markov jump processes, and let $(X_t)_{t \geq 0}$ be a PDMP obeying the PDMP conditions. Then, a.s.,*

$$X_t^{(N)} \rightarrow X_t.$$

in the Skorokhod topology as N tends to ∞ .

The proof of the above theorem shows that X_t is continuous at fixed $t \geq 0$ a.s. Hence, we immediately obtain the following result.

Corollary 4.1. *Let $(X_t^{(N)})_{t \geq 0}$ and $(X_t)_{t \geq 0}$ be as in Theorem 4.1. Then it holds for all $t \geq 0$ that*

$$X_t^{(N)} \rightarrow X_t \quad \text{P-a.s.}$$

as N tends to ∞ .

Let us remark that, if $X_t^{(N)}$ and X_t are not necessarily defined on the same probability space, convergence in distribution with respect to the Skorokhod topology still holds.

5. Example: an autocatalytic network

As an example, we will investigate a small gene regulatory system consisting of an inducible gene and its gene product. Consider a gene g which is transcribed into mRNA, which in turn is translated into protein p . In general, transcription and translation takes place in several intermediate steps. In the following we will ignore these individual processes and consider only the gene g and its gene product p . A protein which regulates the transcription of its own encoding gene or others is called a transcription factor. If the transcription factor p activates transcription of its own encoding gene g , the gene g and the transcription factor p constitute a positive feedback loop. Such a system is called an autocatalytic genetic network. We assume that the gene g possesses only one binding site in its regulatory region to which p can bind, and that the transcription factor p binds as a monomer to this binding site.

Let g_0 and g_1 denote the unoccupied and the occupied states of the promoter. Since p acts as an activator, g_0 and g_1 also denote the inactive and the active states of the gene, respectively. Then, such an autocatalytic network can be described by the following four reaction channels:



The first reaction channel describes activation of the gene. The activation rate is given as the product of the association rate constant κ_1 and the concentration of the transcription factor p . Hence, if there is no gene product or any other external stimulus present, no synthesis will take place. On the other hand, dissociation of p from the binding site of the promoter, described by the second reaction channel, is assumed to take place at a constant rate κ_{-1} . The third and fourth reaction channels describe production and degradation of the transcription factor p , respectively, where the symbol \emptyset denotes the decay products. If the gene is activated, protein is synthesised at a rate α , whereas it is degraded at a constant rate γ independently of the gene state.

Such autocatalytic systems typically convert a graded signal of inducer amounts into a binary response. By this, experimentalists mean that the cells are observed to be essentially in either of two distinct states. For example, in [13] it was shown that the positive feedback loop of the ComK system in *Bacillus subtilis* responds to increasing inducer concentrations in such a binary way. A similar expression behaviour was observed for the synthetic positive feedback system of rtTA in *Saccharomyces cerevisiae* [2]. There, in contrast to the constitutive system, the autocatalytic system also exhibits a binary response on increasing inducer concentrations.

We define a one-parameter family of Markov jump processes $X^{(N)} = (X_t^{(N)})_{t \geq 0}$, $N \in \mathbb{N}$, as follows. The parameter N measures the scale of the system; we can think of it as the volume of the cell. $X^{(N)}$ takes values in the state space $S = \{0, 1\} \times (1/N)\mathbb{N}$. Writing $X_t^{(N)} = (X_1^{(N)}(t), X_2^{(N)}(t))$ the random variable $X_1(t) \in \{0, 1\}$ denotes the state of the gene g at time t , and $X_2(t) \in (1/N)\mathbb{N}$ denotes the concentration of the transcription factor p present at time t . According to the reaction scheme (5.1) there are five possible transitions. Together with the corresponding transition vectors they read

$$\begin{aligned} \left(0, \frac{n}{N}\right) &\mapsto \left(1, \frac{n-1}{N}\right), & e_1^{(N)} &= \left(1, -\frac{1}{N}\right), \\ \left(1, \frac{n}{N}\right) &\mapsto \left(0, \frac{n+1}{N}\right), & e_2^{(N)} &= \left(-1, \frac{1}{N}\right), \\ \left(1, \frac{n}{N}\right) &\mapsto \left(1, \frac{n+1}{N}\right), & e_3^{(N)} &= \left(0, \frac{1}{N}\right), \\ \left(1, \frac{n}{N}\right) &\mapsto \left(1, \frac{n-1}{N}\right), & e_4^{(N)} &= \left(0, -\frac{1}{N}\right), \\ \left(0, \frac{n}{N}\right) &\mapsto \left(0, \frac{n-1}{N}\right), & e_5^{(N)} &= \left(0, -\frac{1}{N}\right). \end{aligned}$$

Observe that the transition vectors which cause the turning on and off of the gene g converge to nonzero constant vectors as $N \rightarrow \infty$. Such jumps are called big jumps. On the other hand,

the components of the transition vectors describing formation and degradation of the gene product p become smaller and smaller. Jumps induced by these vectors are called small jumps. Given that the process is in state $(i, n/N)$ at time t , the following scaling of the transition rates is sensible:

$$\begin{aligned}
 r_1\left(i, \frac{n}{N}\right) &= \begin{cases} \kappa_1 \frac{n}{N} & \text{if } i = 0, \\ 0 & \text{otherwise,} \end{cases} & r_2\left(i, \frac{n}{N}\right) &= \begin{cases} \kappa_{-1} & \text{if } i = 1, \\ 0 & \text{otherwise,} \end{cases} \\
 r_3\left(i, \frac{n}{N}\right) &= \begin{cases} N\alpha & \text{if } i = 1, \\ 0 & \text{otherwise,} \end{cases} & r_4\left(i, \frac{n}{N}\right) &= \begin{cases} N\gamma \frac{n}{N} & \text{if } i = 1, \\ 0 & \text{otherwise,} \end{cases} \\
 r_5\left(i, \frac{n}{N}\right) &= \begin{cases} N\gamma \frac{n}{N} & \text{if } i = 0, \\ 0 & \text{otherwise.} \end{cases}
 \end{aligned}$$

The one-parameter family $(X_t^{(N)})_{N>1}$ describing the autocatalytic system is defined as the solution of the following time-change representation:

$$X_t^{(N)} = x_0^{(N)} + \sum_{k=1}^5 e_k^{(N)} \Pi_k \left(\int_0^t r_k^{(N)}(X_s^{(N)}) ds \right). \tag{5.2}$$

The transition vectors $e_k^{(N)}$ and rate functions $r_k^{(N)}$, $1 \leq k \leq K$, fulfill the PDMP conditions. According to the scaling in the PDMP conditions, the limit transition vectors read

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad e_4 = e_5 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}. \tag{5.3}$$

Furthermore, with $z = n/N$, the limit transition functions read

$$r_1(i, z) = \begin{cases} \kappa_1 z & \text{if } i = 0, \\ 0 & \text{otherwise,} \end{cases} \quad r_2(i, z) = \begin{cases} \kappa_{-1} & \text{if } i = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{5.4a}$$

$$r_3(i, z) = \begin{cases} \alpha & \text{if } i = 1, \\ 0 & \text{otherwise,} \end{cases} \quad r_4(i, z) = \begin{cases} \gamma z & \text{if } i = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{5.4b}$$

$$r_5(i, z) = \begin{cases} \gamma z & \text{if } i = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{5.4c}$$

By Theorem 4.1, the one-parameter Markov jump processes given by (5.2) converge to the PDMP given as the solution of

$$X_t = x_0 + \sum_{k=1}^2 e_k \Pi_k \left(\int_0^t r_k(X_s) ds \right) + \sum_{k=3}^5 e_k \int_0^t r_k(X_s) ds,$$

where the transition vectors and the transition functions are given by (5.3) and (5.4), respectively.

A more detailed analysis of the behaviour of the limit process can be found in [15] and [17]. In [17], we also analysed the influence of dimerisation on the expression behaviour of the above autocatalytic network by numerical means. In [15], we investigated the response of different autocatalytic networks on different initial inducer amounts and different expression durations.

Furthermore, we have shown that the stationary distribution of the above-discussed network is the point mass at zero. As a consequence, the proteins become extinct in contrast to a purely deterministic system which converges to its positive stationary point. Note that we need convergence in the path space D to transfer this extinction statement in the right manner to the approximating processes.

Appendix A. Proofs of the results for the time-change representation

In this appendix we present the proofs of the results from Section 3. For these proofs, note that after a jump time τ the process X moves again deterministically till it reaches the next jump time $\tau' > \tau$. In that situation, we call a process $Y^{(\tau)} = (Y_t^{(\tau)})_{t \geq 0}$ the deterministic part of X after the jump time τ if it solves the equation

$$Y_t^{(\tau)} = X_\tau + \int_0^t V(Y_s^{(\tau)}) ds.$$

Owing to the uniqueness of solution for ordinary differential equations, it follows that $X_{\tau+t}$ and Y_t coincide in the time interval $[0, \tau' - \tau)$.

Proof of Proposition 3.1. Let \tilde{X} solve the equation

$$\tilde{X}_t = x_0 + \sum_{k=1}^J e_k \Pi_k \left(\int_0^t r_k(\tilde{X}_s) ds \right) + \int_0^t V(\tilde{X}_s) ds,$$

and denote by $(\tau_i)_{i \geq 0}$ and $(\tilde{\tau}_i)_{i \geq 0}$ the jump times of X and \tilde{X} , respectively.

We assumed that V is locally Lipschitz. Thus, if $0 \leq t \leq \tau_1 \wedge \tilde{\tau}_1$, uniqueness of the solution of the ordinary differential equation implies that $X_t = \tilde{X}_t$. Since X_t and \tilde{X}_t coincide in $[0, \tau_1 \wedge \tilde{\tau}_1)$, we find that $\tau_1 = \tilde{\tau}_1$. Thus, it holds that $X_t = \tilde{X}_t$ for all $t \in [0, \tau_1]$ too. Now we proceed by induction. Suppose that equality holds up to a jump time $\tau_i, i > 1$. Consider the deterministic parts of X and \tilde{X} solving the equations

$$Y_t^{(\tau_i)} = X_{\tau_i} + \int_0^t V(Y_s^{(\tau_i)}) ds, \quad \tilde{Y}_t^{(\tau_i)} = X_{\tau_i} + \int_0^t V(\tilde{Y}_s^{(\tau_i)}) ds,$$

for $t \geq 0$. Again, we see by the Picard–Lindelöf theorem that $Y_t^{(\tau_i)}$ and $\tilde{Y}_t^{(\tau_i)}$ are identical and X and \tilde{X} are identical on $[\tau_i, \tau_{i+1} \wedge \tilde{\tau}_{i+1})$. This gives $\tau_{i+1} = \tilde{\tau}_{i+1}$. The principle of induction completes the proof.

Now let $(X_t)_{t \geq 0}$ be given by a time-change representation. To show that X is strong Markov, we still need some further preparation. Let $u = (u_1, \dots, u_J) \in \mathbb{R}_+^J$ be a fixed real vector. The stopped process at $u, (X_t^{(u)})_{t \geq 0}$, is defined as

$$X_t^{(u)} = x_0 + \sum_{k=1}^J e_k \Pi_k \left(\int_0^t r_k(X_s) ds \wedge u_k \right) + \int_0^t V(X_s) ds. \tag{A.1}$$

Observe that the Poisson processes Π_k are stopped, not X . The Poisson processes $\Pi_k, 1 \leq k \leq J$, define a multiparameter filtration $\mathcal{F}^\Pi = (\mathcal{F}_u^\Pi)_{u \in \mathbb{R}_+^J}$, where

$$\mathcal{F}_u^\Pi = \sigma(\{\Pi_k(w) : 0 \leq w \leq u_k, k = 1, \dots, J\} \cup \mathcal{N}_P)$$

and \mathcal{N}_P denotes all the null sets of P .

Let $\zeta_k : \Omega \rightarrow \mathbb{R}_+, 1 \leq k \leq J$, be random variables. Then we call the random variable $\zeta = \{\zeta_1, \dots, \zeta_J\} : \Omega \rightarrow \mathbb{R}_+^J$ a multivariate stopping time with respect to \mathcal{F}^Π if

$$\{\zeta \leq \mathbf{u}\} = \{\zeta_1 \leq u_1, \dots, \zeta_J \leq u_J\} \in \mathcal{F}_{\mathbf{u}}^\Pi$$

for all $\mathbf{u} = (u_1, \dots, u_J) \in \mathbb{R}_+^J$. The σ -algebra of the ζ -past is defined by

$$\mathcal{F}_\zeta = \{A \in \mathcal{F} : A \cap \{\zeta_k \leq u_k, 1 \leq k \leq J\} \in \mathcal{F}_{\mathbf{u}}^\Pi \text{ for all } \mathbf{u} \in \mathbb{R}_+^J\}.$$

Let $\zeta : \Omega \rightarrow [0, \infty)$ be a random variable, and let $\Pi : \Omega \rightarrow \mathbb{N}_0$ be a Poisson process. The randomly shifted Poisson process $(\Pi^{(\zeta)}(u))_{u \geq 0}$ is defined by

$$\Pi^{(\zeta)}(u) = \Pi(\zeta + u) - \Pi(\zeta). \tag{A.2}$$

For multivariate stopping times, these processes have nice properties.

Lemma A.1. *Let $\zeta = (\zeta_1, \dots, \zeta_J)$ be a multivariate stopping time, and let $\{\Pi_k^{(\zeta_k)} : 1 \leq k \leq J\}$ be the associated randomly shifted Poisson processes. Then \mathcal{F}_ζ and $\{\Pi_k^{(\zeta_k)} : 1 \leq k \leq J\}$ are independent.*

Furthermore, $\{\Pi_k^{(\zeta_k)} : k = 1, \dots, J\}$ are again J independent Poisson processes.

Proof. We show that, for all bounded $f : \mathbb{N}^{K \times M} \rightarrow \mathbb{R}$ and each $A \in \mathcal{F}_\zeta$, it holds that

$$\begin{aligned} & \mathbb{E}(\mathbf{1}_A f(\Pi_1^{(\zeta_1)}(t_1^{(1)}), \dots, \Pi_1^{(\zeta_1)}(t_M^{(1)}), \dots, \Pi_K^{(\zeta_K)}(t_1^{(K)}), \dots, \Pi_K^{(\zeta_K)}(t_M^{(K)}))) \\ &= \mathbb{P}(A) \mathbb{E}(f(\Pi_1(t_1^{(1)}), \dots, \Pi_1(t_M^{(1)}), \dots, \Pi_K(t_1^{(K)}), \dots, \Pi_K(t_M^{(K)}))). \end{aligned}$$

For $n \geq 1$, define $\zeta^{(n)} = \lceil n\zeta \rceil / n$. Then, $\zeta^{(n)} = (1/n)(l_1^{(n)}, \dots, l_K^{(n)})$ for an $(l_1^{(n)}, \dots, l_K^{(n)}) \in \mathbb{N}_0^K$, and $\zeta^{(n)}$ is again a multivariate stopping time with $\zeta^{(n)} \searrow \zeta$ componentwise as n tends to ∞ . If n is large enough, $\Pi_k^{(\zeta^{(n)})}(t)$ is pointwise equal to $\Pi_k^{(\zeta_k)}(t)$ by the right continuity of Π_k . Together with the stationarity of increments, we obtain

$$\begin{aligned} & \mathbb{E}(\mathbf{1}_A f(\Pi_1^{(\zeta_1^{(n)})}(t_1^{(1)}), \dots, \Pi_1^{(\zeta_1^{(n)})}(t_M^{(1)}), \dots, \Pi_K^{(\zeta_K^{(n)})}(t_1^{(K)}), \dots, \Pi_K^{(\zeta_K^{(n)})}(t_M^{(K)}))) \\ &= \sum_{(l_1, \dots, l_K) \in \mathbb{N}_0^K} \mathbb{E}(\mathbf{1}_A \mathbf{1}_{\{\zeta^{(n)} = (1/n)(l_1, \dots, l_K)\}} f(\Pi_1^{(\zeta_1^{(n)})}(t_1^{(1)}), \dots, \Pi_1^{(\zeta_1^{(n)})}(t_M^{(1)}), \dots, \\ & \hspace{15em} \Pi_K^{(\zeta_K^{(n)})}(t_1^{(K)}), \dots, \Pi_K^{(\zeta_K^{(n)})}(t_M^{(K)}))) \\ &= \sum_{(l_1, \dots, l_K) \in \mathbb{N}_0^K} \mathbb{E}(\mathbf{1}_{A \cap \{\zeta^{(n)} = (1/n)(l_1, \dots, l_K)\}} \mathbb{E}(f(\Pi_1^{(l_1/n)}(t_1^{(1)}), \dots, \Pi_1^{(l_1/n)}(t_M^{(1)}), \dots, \\ & \hspace{15em} \Pi_K^{(l_K/n)}(t_1^{(K)}), \dots, \Pi_K^{(l_K/n)}(t_M^{(K)}))) \\ &= \mathbb{P}(A) \mathbb{E}(f(\Pi_1(t_1^{(1)}), \dots, \Pi_1(t_M^{(1)}), \dots, \Pi_K(t_1^{(K)}), \dots, \Pi_K(t_M^{(K)}))). \end{aligned}$$

By applying Lebesgue’s dominated convergence theorem we obtain

$$\begin{aligned} & \mathbb{E}(\mathbf{1}_A f(\Pi_1^{(\zeta)}(t_1^{(1)}), \dots, \Pi_1^{(\zeta)}(t_M^{(1)}), \dots, \Pi_K^{(\zeta)}(t_1^{(K)}), \dots, \Pi_K^{(\zeta)}(t_M^{(K)}))) \\ &= \lim_{n \rightarrow \infty} \mathbb{E}(\mathbf{1}_A f(\Pi_1^{(\zeta_1^{(n)})}(t_1^{(1)}), \dots, \Pi_1^{(\zeta_1^{(n)})}(t_M^{(1)}), \dots, \Pi_K^{(\zeta_K^{(n)})}(t_1^{(K)}), \dots, \Pi_K^{(\zeta_K^{(n)})}(t_M^{(K)}))) \\ &= \mathbb{P}(A) \mathbb{E}(f(\Pi_1(t_1^{(1)}), \dots, \Pi_1(t_M^{(1)}), \dots, \Pi_K(t_1^{(K)}), \dots, \Pi_K(t_M^{(K)}))). \end{aligned}$$

For the proof of Theorem 3.1, we need two more technical lemmas. First, we show that, for any random time τ , the stopped process and the nonstopped process are equal on the set

$$W(\tau) = \left\{ \int_0^\tau r_k(X_s) ds \leq u_k, k = 1, \dots, J \right\} \tag{A.3}$$

for any real vector $\mathbf{u} = (u_1, \dots, u_J)$ up to time τ .

Lemma A.2. *Let $\mathbf{u} = (u_1, \dots, u_J) \in \mathbb{R}_+^J$, and let the processes $(X_t)_{t \geq 0}$ and $(X_t^{(\mathbf{u})})_{t \geq 0}$ be solutions to (3.1) and (A.1). Then, for each random variable τ , it holds that*

$$\int_0^\tau r_k(X_s) ds \leq u_k \iff \int_0^\tau r_k(X_s^{(\mathbf{u})}) ds \leq u_k$$

for all $k = 1, \dots, J$.

Furthermore, on the set $W(\tau)$ defined by (A.3), the processes $(X_t)_{t \geq 0}$ and $(X_t^{(\mathbf{u})})_{t \geq 0}$ are indistinguishable for all $t \leq \tau$.

Proof. Assume that, for a real positive vector $\mathbf{u} = (u_1, \dots, u_J)$ and a random time τ , it holds that $\int_0^\tau r_k(X_s^{(\mathbf{u})}) ds \leq u_k$ for all $k = 1, \dots, J$ but $\int_0^\tau r_k(X_s) ds > u_k$ for at least one k . Let \tilde{k} denote the smallest of those k . Consider the random variable

$$T = \sup \left\{ t' : \int_0^{t'} r_{\tilde{k}}(X_s) ds = u_{\tilde{k}} \right\}.$$

It holds that $T < \tau$ and $X_s = X_s^{(\mathbf{u})}$ for all $0 \leq s \leq T$. Since both processes are right continuous, there exists an $\varepsilon > 0$ with $T + \varepsilon < \tau$ such that $X_s = X_s^{(\mathbf{u})}$ for all $0 \leq s \leq T + \varepsilon$. This implies that

$$\int_0^\tau r_{\tilde{k}}(X_s^{(\mathbf{u})}) ds \geq \int_0^{T+\varepsilon} r_{\tilde{k}}(X_s^{(\mathbf{u})}) ds = \int_0^{T+\varepsilon} r_{\tilde{k}}(X_s) ds > u_{\tilde{k}},$$

contradicting the initial assumption. The reverse direction is obvious.

Recall the definition

$$\beta_k(t) = \int_0^t r_k(X_s) ds \quad \text{for } k = 1, \dots, J.$$

Corollary A.1. *Let τ be a stopping time of $(X_t)_{t \geq 0}$. Then $\tilde{\beta}(\tau) = (\beta_k(\tau))_{k=1}^J \in \mathbb{R}_+^J$ is a multivariate stopping time with respect to \mathcal{F}^Π .*

Proof. Assume first that τ is countably valued, $\tau(\Omega) = \{t_n : n \in \mathbb{N}\}$, and $\mathbf{u} \in [0, \infty)^J$. For $n \in \mathbb{N}$, we find that

$$\{\beta_k(\tau) \leq u_k, k = 1, \dots, J\} \cap \{\tau = t_n\} = \left\{ \int_0^{t_n} r_k(X_s) ds \leq u_k, k = 1, \dots, J \right\} \cap \{\tau = t_n\}.$$

Since τ is a stopping time, $\{\tau = t_n\}$ is measurable with respect to $(X_s)_{0 \leq s \leq t_n}$. By Lemma A.2 we obtain

$$\{\beta_k(\tau) \leq u_k, k = 1, \dots, J\} \in \mathcal{F}_\mathbf{u}^\Pi,$$

which proves the assertion in the case of countably valued τ .

For general τ , we see that, for $\tau_n = 2^{-n} \lceil 2^n \tau \rceil$, $\{\beta_k(\tau_n) \leq u_k, k = 1, \dots, J\} \in \mathcal{F}_\mathbf{u}^\Pi$. Observe that $\mathcal{F}_\mathbf{u}^\Pi$ is the canonical σ -algebra generated by the process $(\sum_{k=1}^J (1/k) \Pi_k(u_k t))_{t \geq 0}$

at time 1. The filtration of the latter is right continuous; see, e.g. [6]. This shows that the multiparameter filtration $(\mathcal{F}_u^\Pi)_{u \in \mathbb{R}_+^J}$ is right continuous too. Now continuity of β_k and $\tau_n \searrow n \rightarrow \infty \tau$ imply the above result for general τ also.

Recall that the rate functions r_k are nonnegative, meaning that the time transformations β_k are nondecreasing. However, a stronger result is true.

Lemma A.3. *Let $X = (X_t)_{t \geq 0}$ solve the time-change representation (3.1), let $0 < \tau_1$ be the first jump time of X , and let $Y = (Y_t)_{t \geq 0}$ be the deterministic part of X starting at x_0 at time 0. Furthermore, let k_1 denote the index of that Poisson process which jumps at time $\beta_{k_1}(\tau_1)$. Then the time transformation $\beta_{k_1}^{(Y)}$ derived from Y ,*

$$\beta_{k_1}^{(Y)}(t) = \int_0^t r_{k_1}(Y_s) ds,$$

is strictly increasing at τ_1 a.s.

Proof. Since the rate functions r_k are nonnegative, the time transformations $\beta_k^{(Y)}$ are non-decreasing. Now, by the continuity of r_k , the set

$$D_k = \bigcup_{\substack{I \subseteq \mathbb{R}_+ \text{ interval} \\ r_k(Y_s) = 0 \text{ for all } s \in I}} I$$

on which $\beta_k^{(Y)}$ is not strictly increasing consists of at most countably many intervals. If $\beta_{k_1}^{(Y)}$ is not strictly increasing at τ_1 then τ_1 would be the leftmost point of a left closed interval within D_{k_1} . For arbitrary k , let D_k^0 be the set of all leftmost points in D_k . That is,

$$D_k^0 = \left\{ t \geq 0: r_k(Y_t) = 0, \int_{t-\varepsilon}^t r_k(Y_s) ds > 0 \text{ for all } \varepsilon > 0, \text{ there exists } \delta > 0 \right. \\ \left. \text{such that } \int_t^{t+\delta} r_k(Y_s) ds = 0 \right\}.$$

Let $\eta_1^{(k)}$ be the first jump time of the Poisson process Π_k . Observe that $\tilde{D}_k^0 = \beta_k^{(Y)}(D_k^0)$ is still at most countable. Thus, by continuity of the exponential distribution, the probability that $\tau_1 \in D_{k_1}^0$ can be estimated as follows:

$$P(\tau_1 \in D_{k_1}^0) = P(\eta_1^{(k_1)} \in \tilde{D}_{k_1}^0) \leq \sum_{k=1}^J P(\eta_1^{(k)} \in \tilde{D}_k) = 0.$$

Proof of Theorem 3.1. (i) For any stopping time $T > 0$, we obtain

$$X_{T+t} = x_0 + \sum_{k=1}^J e_k \Pi_k \left(\int_0^{T+t} r_k(X_s) ds \right) + \int_0^{T+t} V(X_s) ds \\ = X_T + \sum_{k=1}^J e_k \left(\Pi_k \left(\int_0^T r_k(X_s) ds + \int_T^{T+t} r_k(X_s) ds \right) - \Pi_k \left(\int_0^T r_k(X_s) ds \right) \right) \\ + \int_T^{T+t} V(X_s) ds.$$

Set $\tau_k = \int_0^T r_k(X_s) ds$ for $k = 1, \dots, J$. Then $\tau = (\tau_1, \dots, \tau_J)$ is a multivariate stopping

time for \mathcal{F}^Π . With the notation for $\Pi_k^{(\tau)}$ given in (A.2) we obtain

$$X_{T+t} = X_T + \sum_{k=1}^J e_k \Pi_k^{(\tau_k)} \left(\int_0^t r_k(X_{T+s}) ds \right) + \int_0^t V(X_{T+s}) ds.$$

Now X_T is \mathcal{F}_T -measurable. By Lemma A.1, the distribution of $(X_{T+t})_{t \geq 0}$ conditional on $\sigma(X_s : 0 \leq s \leq T)$ is equal to the distribution of $(X_t)_{t \geq 0}$ conditional on X_T .

(ii) We know that $X = (X_t)_{t \geq 0}$ is strong Markov. Therefore, it is enough to look at the distribution up to the first jump time τ_1 . Let X be given by (3.1) starting in the fixed state $x_0 = (i_0, z_0) \in S$. Let $\eta_1^{(k)}$ denote the first jump times of the Poisson processes Π_k , $1 \leq k \leq J$, let $(Y_t)_{t \geq 0}$ be the deterministic part of X starting at x_0 at time 0, and let $\beta_k^{(Y)}(t)$, $k = 1, \dots, J$, be the time transformations of the Poisson process Π_k with respect to Y . As long as $t < \tau_1$, it holds that $\Pi_k(\beta_k^{(Y)}(t)) = 0$ and the process moves deterministically given by

$$Y_t = x_0 + \int_0^t V(Y_s) ds. \tag{A.4}$$

The distribution of τ_1 computes to

$$P(\tau_1 > t) = P(\beta_k^{-1}(\eta_1^{(k)}) > t, 1 \leq k \leq J) = \exp\left(-\sum_{k=1}^J \beta_k(t)\right). \tag{A.5}$$

Thus, up to the first jump time the process evolves like a PDMP.

The jump time of X is distributed with rate function $\lambda(i_0, \Phi_{i_0}(z_0, t)) = \sum_{k=1}^J \beta_k(t)$ and the flow $\Phi_{i_0}(z_0, t)$ is determined by the vector field V through (A.4).

At the jump time the process chooses a new state according to a discrete distribution. We show that, for $j = 1, \dots, J$,

$$P(X_{\tau_1} = Y_{\tau_1} + e_j \mid \tau_1) = \frac{r_j(Y_{\tau_1})}{\sum_{k=1}^J r_k(Y_{\tau_1})}. \tag{A.6}$$

Observe that we showed in the proof of Lemma A.3 that, a.s., $r_k(Y_{\tau_1}) > 0$ for at least one $k = 1, \dots, J$. From (A.5) we derive the density $f_{\tau_1}(t) = (\sum_{k=1}^J r_k(Y_t)) \exp(-\sum_{k=1}^J \beta_k(t))$. Thus, for $a < b$, we find that

$$\begin{aligned} P(X_{\tau_1} = Y_{\tau_1} + e_j, \tau_1 \in [a, b]) &= \int_a^b P(t < \beta_k^{-1}(\eta_1^{(k)}), k \neq j) f_{\beta_j^{-1}(\eta_1^{(j)})}(t) dt \\ &= \int_a^b \exp\left(-\sum_{k \neq j} \beta_k(t)\right) r_j(Y_t) e^{-\beta_j(t)} dt \\ &= \int_a^b \frac{r_j(Y_t)}{\sum_{k=1}^J r_k(Y_t)} f_{\tau_1}(t) dt. \end{aligned}$$

This shows (A.6).

Appendix B. Proof of the convergence theorem

Now we derive the almost-sure uniform convergence of $X^{(N)}$ to X in the absence of jumps in the limiting process X . To this end, we consider Markov jump processes $Y^{(N)}$ with small

jumps only and their corresponding limit process Y :

$$Y_t^{(N)} = Y_0^{(N)} + \sum_{k=J+1}^K e_k^{(N)} \Theta_k^{(N)} \left(\int_0^t r_k^{(N)}(Y_s^{(N)}) ds \right) \tag{B.1}$$

and

$$Y_t = Y_0 + \sum_{k=J+1}^K e_k \int_0^t r_k(Y_s) ds. \tag{B.2}$$

The counting processes $\Theta_k^{(N)}$ will be derived from parts of the trajectories of Π_k . Thus, they obey a ‘strong law of large numbers (SLLN) property’. More precisely, we require the following.

Definition B.1. (SLLN property.) A family of counting processes $(\Theta_k^{(N)}(t))_{t \geq 0}$ fulfills a uniform SLLN property if

- (i) there exists a random variable M such that, for all $t \geq 0$, $N \in \mathbb{N}$, and $k = 1, \dots, K$, it holds that

$$\Theta_k^{(N)}(t) \leq Mt \quad \text{a.s.},$$

- (ii) there exists a random variable \tilde{M} such that, for all $T > 0$, $t \geq 0$, $\tilde{t} \in [0, T]$, $N \in \mathbb{N}$, and $k = 1, \dots, K$, it holds that

$$\Theta_k^{(N)}(t + \tilde{t}) - \Theta_k^{(N)}(\tilde{t}) \leq 1 + \tilde{M}t \quad \text{a.s.},$$

- (iii) for all $m > 0$, it holds a.s. that

$$\lim_{t \rightarrow \infty} \sup_{N,k} \sup_{a \leq m} \left| \frac{\Theta_k^{(N)}(at)}{t} - a \right| = 0.$$

Note that the constant sequence formed from a Poisson process possesses the SLLN property. Hence, for example, we can consider (B.1) and (B.2) under the choice $\Theta_k^{(N)} = \Pi_k$, yielding the processes $X^{(N)}$ and X . up to the time of the first big jump in one of the Poisson processes Π_k , $k = 1, \dots, J$.

Now we can derive bounds on Y .

Lemma B.1. *There a.s. exist random variables C_1 and C_2 such that, for all $N \in \mathbb{N}$, $t \in [0, T]$, and $k \in \{J + 1, \dots, K\}$,*

$$\max \left(\|Y_t^{(N)}\|, \frac{1}{N} r_k^{(N)}(Y_t^{(N)}) \right) \leq C_1 e^{C_2 t}.$$

Proof. From the sublinear growth condition of the PDMP conditions (Condition 4.1(vii)) and condition (ii) of the SLLN property of $\Theta^{(N)}$, we obtain, for all $k \in \{J + 1, \dots, K\}$,

$$\begin{aligned} \frac{1}{N} r_k^{(N)}(Y_t^{(N)}) &= \frac{1}{N} r_k^{(N)} \left(Y_0^{(N)} + \sum_{k=J+1}^K e_k^{(N)} \Theta_k^{(N)} \left(\int_0^t r_k^{(N)}(Y_s^{(N)}) ds \right) \right) \\ &\leq c(Y_0^{(N)})((K - J) + 1) + c(Y_0^{(N)})\tilde{M} \int_0^t \sum_{k=J+1}^K \frac{1}{N} r_k^{(N)}(Y_s^{(N)}) ds. \end{aligned}$$

Summation of both sides yields

$$\sum_{k=J+1}^K \frac{1}{N} r_k^{(N)}(Y_t^{(N)}) \leq c(Y_0^{(N)})(1 + (K - J))(K - J) + c(Y_0^{(N)})\tilde{M}(K - J) \int_0^t \sum_{k=J+1}^K \frac{1}{N} r_k^{(N)}(Y_s^{(N)}) ds.$$

Introduce two random variables $C'_1 = \sup_{N \rightarrow \infty} c(Y_0^{(N)})(1 + (K - J))(K - J)$ and $C_2 = \sup_{N \rightarrow \infty} c(Y_0^{(N)})\tilde{M}(K - J)$. Then Gronwall's lemma yields

$$\sum_{k=J+1}^K \frac{1}{N} r_k^{(N)}(Y_t^{(N)}) \leq C'_1 e^{C_2 t}$$

for all $N \in \mathbb{N}$, $k \in \{J + 1, \dots, K\}$, and $t \in [0, T]$.

Furthermore, again using condition (ii) of the SLLN property of $\Theta^{(N)}$, we find that

$$\begin{aligned} \|Y_t^{(N)}\| &\leq \|Y_0^{(N)}\| + \sum_{k=J+1}^K \|Ne_k^{(N)}\| \frac{1}{N} \Theta_k^{(N)} \left(\int_0^t r_k^{(N)}(Y_s^{(N)}) ds \right) \\ &\leq \sup_N \|Y_0^{(N)}\| + (K - J) \sup_{N,k} \|Ne_k^{(N)}\| \left(\frac{1}{N} + \tilde{M} \frac{C'_1}{C_2} (e^{C_2 t} - 1) \right). \end{aligned}$$

From above we find, with an appropriate choice of the random variable C_1 , e.g.

$$C_1 = \sup_N \|Y_0^{(N)}\| + (K - J) \sup_N \frac{1}{N} \|Ne_k^{(N)}\| \left(1 + \tilde{M} \frac{C'_1}{C_2} \right),$$

that $|Y_t^{(N)}| \leq C_1 e^{C_2 t}$ uniformly for all $N \in \mathbb{N}$ and $t \in [0, T]$.

With the estimate from the previous lemma we obtain uniform convergence on compacta of the rate functions of the discrete processes to those of the PDMP.

Corollary B.1. *For all $T > 0$, it holds a.s. that, for $k = 1, \dots, J$,*

$$r_k^{(N)}(Y_t^{(N)}) - r_k(Y_t^{(N)}) \rightarrow 0 \quad \text{as } N \rightarrow \infty$$

uniformly in $t \in [0, T]$, and, for $k = J + 1, \dots, K$,

$$\frac{1}{N} r_k^{(N)}(Y_t^{(N)}) - r_k(Y_t^{(N)}) \rightarrow 0 \quad \text{as } N \rightarrow \infty$$

uniformly in $t \in [0, T]$.

Proof. From Lemma B.1 we know that $\|Y_t^{(N)}\| \leq C_1 e^{C_2 T}$ for some random constants C_1 and C_2 . The assertion follows from the uniform convergence of $r_k^{(N)}$ and $(1/N)r_k^{(N)}$ on compacta.

The following result is a reformulation of Kurtz's convergence theorem; see Theorem 2.16 of [12].

Lemma B.2. Let $(Y_t^{(N)})_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ be the solutions of (B.1) and (B.2), respectively. Suppose that $Y_0^{(N)} \rightarrow Y_0$ as $N \rightarrow \infty$, and that conditions (ii) and (iii) of the SLLN property for counting processes, and (iii)–(vii) of the PDMP conditions hold. Then, a.s., it holds that, for all $T > 0$,

$$\|Y_t^{(N)} - Y_t\| \rightarrow 0$$

uniformly in $t \in [0, T]$ as $N \rightarrow \infty$.

We are now in a position to prove the convergence theorem. Our aim is to show that the process $X^{(N)}$ has essentially the same big jump times (up to order $o(1)$) and the same jump types as X .

Proof of Theorem 4.1. Let $(\tau_i)_{i \geq 1}$ be the jump times of the PDMP solving (4.4), and let $(\tau_i^{(N)})_{i \geq 1}$ be the times of the ‘big’ jumps of the Markov jump processes solving (4.1). Thus, these times are ordered increasingly and

$$\begin{aligned} \{\tau_i : i \geq 1\} &= \left\{ t > 0 : X_t \neq \lim_{\varepsilon \downarrow 0} X_{t-\varepsilon} \right\}, \\ \{\tau_i^{(N)} : i \geq 1\} &= \left\{ t > 0 : X_t^{(N)} - \lim_{\varepsilon \downarrow 0} X_{t-\varepsilon}^{(N)} \in \{e_k^{(N)} : 1 \leq k \leq J\} \right\}. \end{aligned}$$

Furthermore, denote by k_i and $k_i^{(N)}$ the indices of the Poisson processes which cause the jumps at τ_i and $\tau_i^{(N)}$ of X and $X^{(N)}$, respectively.

We proceed by induction over $i \geq 1$. The induction hypothesis states that, a.s.,

- (a) $\tau_i^{(N)} \rightarrow \tau_i$ as $N \rightarrow \infty$,
- (b) $\int_0^{\tau_i^{(N)}} r_k^{(N)}(X_s^{(N)}) ds \rightarrow \int_0^{\tau_i} r_k(X_s) ds$ as $N \rightarrow \infty$ for all $k = 1, \dots, J$,
- (c) $X_{\tau_i^{(N)}}^{(N)} \rightarrow X_{\tau_i}$ as $N \rightarrow \infty$,

for all $i = 1, 2, \dots$. For $i = 1$, we derive this by considering processes Y and $Y^{(N)}$ solving

$$\begin{aligned} Y_t^{(N)} &= x_0^{(N)} + \sum_{k=J+1}^K e_k^{(N)} \Pi_k \left(\int_0^t r_k^{(N)}(Y_s^{(N)}) ds \right), \\ Y_t &= x_0 + \sum_{k=J+1}^K e_k \int_0^t r_k(Y_s) ds. \end{aligned}$$

At first, we consider the time transformations $\beta_k^{(N)}$ and β_k defined by

$$\beta_k^{(N)}(t) = \int_0^t r_k^{(N)}(Y_s^{(N)}) ds \quad \text{and} \quad \beta_k(t) = \int_0^t r_k(Y_s) ds$$

for $k = 1, \dots, J$. Observe also that $\beta_k(t) = \int_0^t r_k(X_s) ds$ as long as $t \leq \tau_1$. A similar formula is valid for $\beta_k^{(N)}$. It now holds that

$$|\beta_k^{(N)}(t) - \beta_k(t)| \leq \int_0^t |r_k^{(N)}(Y_s^{(N)}) - r_k(Y_s^{(N)})| ds + \int_0^t |r_k(Y_s^{(N)}) - r_k(Y_s)| ds.$$

Conditions (ii) and (iii) of the uniform SLLN property for counting processes ensures that we can apply Corollary B.1 and Lemma B.2. From the former we deduce that

$$\int_0^t |r_k^{(N)}(Y_s^{(N)}) - r_k(Y_s^{(N)})| ds \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

From Lemma B.2 we derive via the Lipschitz continuity of r_k that

$$\int_0^t |r_k(Y_s^{(N)}) - r_k(Y_s)| ds \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

Thus, for all $T > 0$, we conclude that it holds uniformly for $t \in [0, T]$ that

$$\beta_k^{(N)}(t) \rightarrow \beta_k(t) \quad \text{as } N \rightarrow \infty. \tag{B.3}$$

Let $\eta_1^{(k)}$ be the first jump times of the Poisson processes Π_k , and let k_1 be the index of the Poisson process which causes the jump of X at τ_1 . We know that $\beta_k^{(N)}(\tau_1)$ converges to $\beta_k(\tau_1)$ as N tends to ∞ . Specifically, $\beta_{k_1}^{(N)}(\tau_1) \rightarrow \eta_1^{(k_1)}$. Moreover, since k_1 causes the jump at τ_1 , it must hold for all $k \neq k_1$ that

$$\eta_1^{(k)} > \beta_k(\tau_1) = \lim_{N \rightarrow \infty} \beta_k^{(N)}(\tau_1).$$

Hence, $k_1^{(N)}$ is equal to k_1 for large enough N . Since we know from Lemma A.3 that $\beta_{k_1}(t)$ is strictly increasing at τ_1 , we deduce for the pseudo inverse $\hat{\beta}_k$ of β_k that

$$\hat{\beta}_{k_1}^{(N)}(\eta_1^{(k_1)}) \rightarrow \hat{\beta}_{(k_1)}(\eta_1^{(k_1)}) = \tau_1 \quad \text{and} \quad \hat{\beta}_k^{(N)}(\eta_1^{(k)}) \rightarrow \hat{\beta}_k(\eta_1^{(k)}) > \tau_1 \quad \text{for } k \neq k_1$$

as N tends to ∞ . Thus, $\tau_1^{(N)} \rightarrow \tau_1$ as N tends to ∞ , proving (a).

To prove (b), consider the following estimate:

$$\begin{aligned} & \left| \int_0^{\tau_1^{(N)}} r_k^{(N)}(X_s^{(N)}) ds - \int_0^{\tau_1} r_k(X_s) ds \right| \\ & \leq \left| \int_0^{\tau_1^{(N)}} (r_k^{(N)}(X_s^{(N)}) - r_k(X_s)) ds \right| + \left| \int_{\tau_1^{(N)}}^{\tau_1} r_k(X_s) ds \right|. \end{aligned}$$

Since X_t coincides with Y_t , and $X_t^{(N)}$ coincides with $Y_t^{(N)}$ for $t < \tau_1$ and $t < \tau_1^{(N)}$, respectively, we can apply property (B.3). Together with $\tau_i^{(N)} \rightarrow \tau_i$ we conclude that both terms converge to 0 as N tends to ∞ , proving (b).

To prove (c), consider the following estimation:

$$\|X_{\tau_1^{(N)}}^{(N)} - X_{\tau_1}\| \leq \|(X_{\tau_1^{(N)}}^{(N)} - X_{\tau_1^{(N)-}}^{(N)}) - (X_{\tau_1} - X_{\tau_1-})\| + \|X_{\tau_1^{(N)-}}^{(N)} - X_{\tau_1-}\|.$$

Choosing N large enough, we can assume that $k_1^{(N)} = k_1$. Since $\tau_1^{(N)}$ and τ_1 are the first jump times of $X_t^{(N)}$ and X_t , respectively, it holds that $\Pi_{k_1^{(N)}}$ and Π_{k_1} are equal to 1 where all other Poisson processes Π_k , $k \neq k_1$, are equal to 0. Thus,

$$\|X_{\tau_1^{(N)}}^{(N)} - X_{\tau_1}\| \leq \|e_{k_1^{(N)}}^{(N)} - e_{k_1}\| + \|Y_{\tau_1^{(N)}}^{(N)} - Y_{\tau_1}\|.$$

With (i) and (ii) of the PDMP conditions we obtain (c) and the induction is initialised.

For the induction step, assume that (a)–(c) hold up to the i th jump, $i > 1$. We consider the processes $(Z_t^{(N)})_{t \geq 0}$ and $(Z_t)_{t \geq 0}$ satisfying the equations

$$Z_t^{(N)} = X_{\tau_i^{(N)}}^{(N)} + \sum_{k=J+1}^K e_k^{(N)} \left(\Pi_k \left(\int_0^t r_k^{(N)}(Z_s^{(N)}) ds + \int_0^{\tau_i^{(N)}} r_k^{(N)}(X_s^{(N)}) ds \right) - \Pi_k \left(\int_0^{\tau_i^{(N)}} r_k^{(N)}(X_s^{(N)}) ds \right) \right)$$

and

$$Z_t = X_{\tau_i} + \sum_{k=J+1}^K e_k \int_0^t r_k(Z_s) ds.$$

By induction hypothesis (b), we know that $\int_0^{\tau_i^{(N)}} r_k^{(N)}(X_s^{(N)}) ds \rightarrow \int_0^{\tau_i} r_k(X_s) ds$ for all $1 \leq k \leq J$ as $N \rightarrow \infty$. Since

$$\Theta_k^{(N)}(t) = \Pi_k \left(t + \int_0^{\tau_i^{(N)}} r_k^{(N)}(X_s^{(N)}) ds \right) - \Pi_k \left(\int_0^{\tau_i^{(N)}} r_k^{(N)}(X_s^{(N)}) ds \right)$$

are counting processes following the SLLN of Definition B.1, we can apply Lemma B.2. It follows that $Z_t^{(N)} \rightarrow Z_t$ uniformly on compacts as $N \rightarrow \infty$.

By similar arguments as for $i = 1$ we find that $k_{i+1}^{(N)} = k_{i+1}$ eventually, and $\tau_{i+1}^{(N)} - \tau_i^{(N)} \rightarrow \tau_{i+1} - \tau_i$ as N tends to ∞ . Since we already know that $\tau_i^{(N)} \rightarrow \tau_i$ as $N \rightarrow \infty$, we have shown (a). From this, (b) and (c) follow as in the proof of the initial hypothesis for induction.

To prove convergence in the J_1 -topology, we choose $T > 0$, $T \notin \{\tau_i : i \in \mathbb{N}\}$, and N so large that all $k_i^{(N)} = k_i$ for all i with $\tau_i \leq T$ as well as $\tau_{i_T}^{(N)} > T$, where τ_{i_T} is the first jump time beyond T . Then we construct $\lambda_N \in \Lambda_\infty$ from $\lambda_N(\tau_i) = \tau_i^{(N)}$ for all i with $\tau_i \leq T$ and $\lambda_N(t) = t$ for $t \geq T$ by linear interpolation. Clearly, since $\tau_i^{(N)} \rightarrow \tau_i$ we obtain $\gamma(\lambda_N) \rightarrow 0$ as $N \rightarrow \infty$.

In the next step we will show that with the above-defined time dilations λ_N applied to the processes $(X_{\tau_i^{(N)}}^{(N)})_{N \geq 1}$, $X_{\tau_i^{(N)}}^{(N)}$ will converge uniformly on each time interval $[\tau_i, \tau_{i+1})$. Thereto, let $\zeta_i^{(N)} = (\tau_{i+1}^{(N)} - \tau_i^{(N)}) / (\tau_{i+1} - \tau_i)$, and consider the processes

$$Z_t^{(N)} := X_{\tau_i^{(N)} + \zeta_i^{(N)} \cdot t}^{(N)}$$

on the time interval $[0, \tau_{i+1} - \tau_i)$. Then, by a linear change of the variable, $Z_t^{(N)}$ solves the equation

$$Z_t^{(N)} = X_{\tau_i^{(N)}}^{(N)} + \sum_{k=J+1}^K e_k^{(N)} \left(\Pi_k \left(\zeta_i^{(N)} \int_0^t r_k^{(N)}(Z_s^{(N)}) ds + \int_0^{\tau_i^{(N)}} r_k^{(N)}(X_s^{(N)}) ds \right) - \Pi_k \left(\int_0^{\tau_i^{(N)}} r_k^{(N)}(X_s^{(N)}) ds \right) \right).$$

Again, as $\lim_{N \rightarrow \infty} (\tau_{i+1}^{(N)} - \tau_i^{(N)}) / (\tau_{i+1} - \tau_i) = 1$, Lemma B.2 is applicable and yields

$$\sup_{t \in [\tau_i, \tau_{i+1})} |X_{\lambda_N(t)}^{(N)} - X_t| = \sup_{t \in [0, \tau_{i+1} - \tau_i)} |Y_t^{(N)} - Y_t| \rightarrow 0 \quad \text{as } N \rightarrow \infty \text{ for } i = 1, \dots, i_T.$$

In exactly the same way we derive

$$\sup_{t \in [\tau_{i_0}, T]} |X_{\lambda_N(t)}^{(N)} - X_t| \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

Consequently,

$$\sup_{t \in [0, T]} |X_{\lambda_N(t)}^{(N)} - X_t| \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

As $T > 0$ is arbitrary, we obtain $d_S(X^{(N)}, X) \rightarrow 0$ as $N \rightarrow \infty$. This proves the assertion.

Acknowledgements

We would like to thank the anonymous referees, whose comments helped to improve the paper.

References

- [1] BALL, K., KURTZ, T. G., POPOVIC, L. AND REMPALA, G. (2006). Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Prob.* **16**, 1925–1961.
- [2] BECKSEI, A. AND SERRANO, L. (2000). Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593.
- [3] BLAKE, W. J., KÆRN, M., CANTOR, C. R. AND COLLINS, J. J. (2003). Noise in eukaryotic gene expression. *Nature* **422**, 633–637.
- [4] BUNDSCHUH, R., HAYOT, F. AND JAYAPRAKASH, C. (2003). Fluctuations and slow variables in genetic networks. *Biophys. J.* **84**, 1606–1615.
- [5] DAVIS, M. H. A. (1984). Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. R. Statist. Soc. B* **46**, 353–388.
- [6] DOOB, J. L. (1953). *Stochastic Processes*. John Wiley, New York.
- [7] GILLESPIE, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–434.
- [8] GILLESPIE, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361.
- [9] HUME, D. A. (2000). Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. *Blood* **96**, 2323–2328.
- [10] KURTZ, T. G. (1970). Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Prob.* **7**, 49–58.
- [11] KURTZ, T. G. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Prob.* **8**, 344–356.
- [12] KURTZ, T. G. (1980). Relationships between stochastic and deterministic population models. In *Biological Growth and Spread* (Proc. Conf. Heidelberg, 1979; Lecture Notes Biomath. **38**), Springer, Berlin, pp. 449–467.
- [13] MAAMAR, H. AND DUBNAU, D. (2005). Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. *Mol. Microbiol.* **56**, 615–624.
- [14] SKOROHOD, A. V. (1956). Limit theorems for stochastic processes. *Theory Prob. Appl.* **1**, 261–290.
- [15] ZEISER, S., FRANZ, U. AND LIEBSCHER, V. (2009). Autocatalytic genetic networks modeled by piecewise-deterministic Markov processes. *J. Math. Biol.* **60**, 207–246.
- [16] ZEISER, S., FRANZ, U., MÜLLER, J. AND LIEBSCHER, V. (2009). Hybrid modeling of noise reduction by a negatively autoregulated system. *Bull. Math. Biol.* **71**, 1006–1024.
- [17] ZEISER, S., FRANZ, U., WITTICH, O. AND LIEBSCHER, V. (2008). Simulation of genetic networks modelled by piecewise deterministic Markov processes. *IET Systems Biol.* **2**, 113–135.