A Markovian Arrival Stream Approach to Stochastic Gene Expression in Cells

Brian Fralix* Mark Holmes† Andreas Löpker‡

June 18, 2021

Abstract

We analyse a generalisation of the stochastic gene expression model studied recently in Fromion et al. (*SIAM Journal of Applied Mathematics*, 2013) and Robert (*Probability Surveys*, 2019) that keeps track of the production of both mRNA and protein molecules, using techniques from the theory of point processes, as well as ideas from the theory of matrix-analytic methods. In the model we consider, both the activity of a gene and the creation of mRNA are modelled with an arbitrary Markovian Arrival Process governed by finitely many phases, and each mRNA molecule during its lifetime gives rise to protein molecules in accordance to a Poisson process. This modification is important, as Markovian Arrival Processes can be used to approximate many types of point processes on the nonnegative real line, meaning this framework allows us to further relax our assumptions on the overall process of transcription.

Keywords: infinite-server queues, Markov arrival process, matrix analytic methods, stochastic gene expression.
2020 MSC: 60G55, 60J28, 60K25, 92C40

1 Introduction

In this paper we are interested in analysing stochastic models of gene expression, which is the process through which genes within a cell produce both messenger RNA (mRNA) and protein molecules. Gene expression models date back to the 1970s with the works of Rigney and Schieve [25], Berg [4], and Rigney [23, 24], with later work occurring in the 1990s, see e.g. Peccoud and Ycart [20].

Over the last 15-20 years, interest in creating/analysing suitable models of gene expression has increased considerably within the molecular biology community, with contributions being made by researchers from many different disciplines within not only the natural sciences, but also engineering and the mathematical sciences. In order to properly place more recent contributions to this research area in proper context, it will help to first briefly introduce the classical stochastic gene expression model: we tersely follow the description of the classical three-stage model of gene expression found in the recent article of Robert [26].

*School of Mathematical and Statistical Sciences, Clemson University (bfralix@clemson.edu).

†School of Mathematics and Statistics, The University of Melbourne (holmes.m@unimelb.edu.au). MH's research is supported by Future Fellowship FT160100166 from the Australian Research Council.

‡Department of Computer Science and Mathematics, Dresden University of Applied Sciences (lopker@htw-dresden.de).

In many mathematical analyses of protein production in cells, the production process of both mRNA and protein molecules by a fixed gene (a fixed segment of DNA) in a cell is considered. This process is often described via the evolution through time of a random vector $(I(t), M(t), P(t))$, where $I(t) \in \{0, 1\}$ describes the state of the gene at time t, and $M(t)$ and $P(t)$ denote the number of mRNA and protein molecules, respectively, present in the cell at time t. When $I(t) = 1$ the gene is said to be *active* at time t; otherwise the gene is *inactive* at this time when $I(t) = 0$. When the gene is active, an RNA polymerase may bind to it and start to copy the gene via sequential production of nucleotides: once this elongation process is complete, a new messenger RNA (mRNA) molecule is formed (so $M(t)$ increases by 1 if this process completes at time t). Molecular biologists refer to the overall process of creating new mRNA molecules from segments of DNA as *transcription*. In many stochastic models of gene expression, it is assumed that the elongation time associated with the production of each mRNA molecule is negligible, meaning all such elongation times are set equal to zero, but models have been created recently that allow for such elongation times to be both random and nonzero.

Once an mRNA molecule has been created, a ribosome may in turn bind with it and move along its chain of nucleotides to build a protein (so $P(t)$ increases by 1) via a chain of amino acids: this process is referred to as *translation*. Again, in many models of gene expression it is assumed that the amount of time it takes a ribosome to create a protein molecule, once it manages to bind to the mRNA molecule is negligible, but recent models have been created that allow for such production times, or elongation times to be both random and nonzero.

Both proteins and mRNA molecules are assumed to have finite, random lifetimes (at the end of such a lifetime $M(t)$ or $P(t)$ decreases by 1). Due to the relatively low numbers of both mRNA and protein molecules within a typical cell, many researchers believe the process of gene expression can be modelled reasonably well with continuous-time, discrete-state stochastic processes: see e.g. Paulsson [19]. We refer readers seeking a textbook-level discussion on stochastic gene expression to Chapter 6 of Bressloff [7]: another text that could be of interest to such readers is Anderson and Kurtz [1].

In many previous studies, as explained in [19], production times are assumed to be negligible, and both mRNA and protein lifetimes are assumed to be exponential with certain rates. In the work of Shahrezaei and Swain [27] the authors study the joint distribution of $M(t)$ and $P(t)$ while further assuming the gene is always active: their analysis involves setting up the Kolmogorov forward equations associated with the continuous-time Markov chain that captures the dynamics of this model, then solving the equations with the method of characteristics under the further assumption that the lifetime of a protein molecule is typically much longer than the lifetime of an mRNA molecule (this imprecise statement is made precise in their paper). This method results in an expression for the joint generating function of $M(t)$ and $P(t)$ that is in terms of a hypergeometric function. A similar analysis is also used in [27] to study the joint distribution of $I(t), M(t)$, and $P(t)$, again under the assumption that the lifetime of a protein molecule is typically much longer than the lifetime of an mRNA molecule. A few years later, in the work of Bokes et al. [5] the authors found the joint generating function of the mRNA and protein levels for the case where the gene is always active, without having to assume mRNA lifetimes are much smaller than protein lifetimes.

Aspects of the above simple model are problematic. In particular, modelling all mRNA lifetimes, protein lifetimes, and elongation times as exponentially distributed random variables is not ideal due to the sequential nature of these physical building processes. There have been many generalisations of the above simple model. For example, the model that is the primary focus of [26] (which is a generalisation of that from Fromion et al. [8]) allows for elongation times and lifetimes with general distributions, with families of Poisson processes corresponding to possible binding times. Bindings

occur at these possible binding times only if they occur while the gene is active (for the first kind of binding above) and during the lifetime of the mRNA molecule (for the second kind of binding above) respectively. It should also be noted here that in the work of Jansen and Pfaffelhuber [12] the authors model gene expression via random time changes of independent Poisson processes, from which they derive the mean and variance of both mRNA and protein levels, when the gene alternates between being active and inactive in accordance to a two-state CTMC, and when the lifetimes of each mRNA and each protein are exponentially distributed. It is notable that in their analysis, elongation times are assumed to be generally distributed. Special types of elongation time distributions are also incorporated in the work of Pendar et al. [21].

In [8, 26], the authors advocate that the three-stage model of stochastic gene expression can be fruitfully analysed with the theory of Poisson random measures, but what differentiates their approach from the approach given in [12] is their use of ideas that are heavily related to the theory of infinite-server queueing systems fed by Poisson arrival processes. One major advantage of this approach is that it easily allows researchers to remove the exponential assumption in many crucial places within the model: more particularly, from this viewpoint one can derive many interesting quantities (e.g. means and variances of mRNA and protein counts) even when all elongation times, mRNA lifetimes and protein lifetimes are allowed to be generally distributed.

Our objective is to complement the approach found in [8, 26] by pointing out that many other established results/tools from the applied probability community can be used to analyse, rather thoroughly, various generalizations of the classical stochastic gene expression model, which allow for the gene to vary between activity and inactivity in accordance to an environment process (called a *Markovian Arrival Process*, or MAP) that is much more general than Markov-modulated Poisson process governed by a two-state continuous-time Markov chain (CTMC). Our methods will involve a combination of both point process theory and ideas from a subject often referred to in applied probability as *matrix-analytic methods*. These methods were originally designed to find computable quantities that describe the behavior of queueing systems where interarrival times and/or service times are no longer exponentially distributed. In particular, it is a well-established fact that many types of stochastic systems having parameters that vary with time due to the behavior of some background environment process can be studied with methods from the matrix-analytic framework, meaning this theory could potentially provide the molecular biology community with useful tools for deriving many results that go further beyond the results discussed in this article.

We should mention that our work seems to also be related to the work of Herbach [10]. In [10], the author analyses the steady-state distribution of the number of mRNA molecules present in a cell, where these molecules are created in accordance to a Markov-modulated Poisson process, and where each mRNA molecule's lifetime is exponentially distributed, independently of everything else. This Markov-modulated Poisson process considered in [10] has an arbitrary, but finite, number of states, and in the models studied in [10], mRNA lifetimes are exponentially distributed, and the time between the instant a RNA polymerase binds with a DNA strand and the creation of a mRNA molecule is assumed to be negligible. Another recent, somewhat related study is that of Jia [13], which focuses on other aspects of transcription when mRNA molecules are created in accordance with a Markov-modulated Poisson process, and each mRNA molecule has an exponentially distributed lifetime that is independent of everything else.

It should also be noted at this point that recently, in the work of Horowitz and Kulkarni [11] the authors used Batch Markovian Arrival Processes (BMAP) to study a somewhat related model, where bursts of mRNA molecules are modelled with a BMAP. In their setting, the authors also argue that it is reasonable to model protein production with a BMAP where proteins are created in batches in accordance to a MAP, for the case where mRNA lifetimes are significantly shorter than protein

lifetimes. The authors of [11] focus more on studying rare events associated with their model, but unlike their model, our model is a straight generalization of the classical three-stage model of gene expression, where the activity/inactivity of the gene is modelled using a MAP, and all elongation times and lifetimes are generally distributed.

Readers should note that MAPs can theoretically be used to approximate many types of point processes on the nonnegative real line—see Asmussen and Koole [2]—so we feel as though this could be an important step towards analyzing more general models of stochastic gene expression.

2 Model

In this paper we analyse an important extension of the three-stage model of gene expression from [26]. Here we assume RNA polymerase bind with an active gene at random points of a so-called *Markovian Arrival Process* (MAP)—see below—having finite phase set S, meaning the transcription process of the gene is governed by the MAP. Associated with each created mRNA molecule is its elongation time (the time it takes for the mRNA molecule to be created from the RNA polymerase that binds with the gene) and its lifetime: the elongation time is generally distributed with CDF L_1, and the lifetime is generally distibuted with cdf F. Furthermore, during the lifetime of an mRNA molecule, ribosomes bind to it in accordance to a homogeneous Poisson process having rate k_2. Each created protein has associated with it both an elongation time (the time it takes a ribosome that binds with the mRNA molecule to produce the protein molecule) and a lifetime: the elongation time of a protein is generally distributed with CDF L_2, while the lifetime of the protein is generally distributed with CDF G. Throughout we assume that all elongation times and lifetimes are independent of each other, as well as the MAP that governs when transcription is initiated.

Our main results are formulae describing the (joint) distribution of $M(t)$ and $P(t)$, and their moments. To be more precise, for $x \in S$ we will let $\mathbb{P}_x$ denote a probability measure under which this process evolves from initial conditions $M(0) = P(0) = 0$ and $I(0) = x$. Then for each $t \in \mathbb{R}_+ := [0, \infty)$, each $m \in \mathbb{Z}_+ := \{0, 1, 2, \dots\}$, and each $n \in \mathbb{Z}_+$, we define the matrix $\mathbf{J}(t, m, n)$ as

$$\mathbf{J}(t, m, n) := \left[\mathbb{P}_x(M(t) = m, P(t) = n, I(t) = y)\right]_{x,y \in S}. \tag{1}$$

For fixed x this gives the joint *probability mass function* (PMF) of $(M(t), P(t), I(t))$. Letting $\mathbf{I}$ denote the $|S| \times |S|$ identity matrix, and $\mathbf{0}$ the $|S| \times |S|$ zero matrix, readers should note that for each $m, n \in \mathbb{Z}_+$,

$$\mathbf{J}(0, m, n) = \begin{cases} \mathbf{I}, & m = n = 0; \\ \mathbf{0}, & \text{otherwise.} \end{cases}$$

A more compact description of the joint distribution of $(M(t), P(t), I(t))$ can be made through the use of *probability generating functions* (PGFs). For each $t \in \mathbb{R}_+$, and each $z_1, z_2 \in \mathbb{D} := \{z \in \mathbb{C} : |z| \leq 1\}$, we define the matrix

$$\hat{\mathbf{J}}(t, z_1, z_2) := \left[\mathbb{E}_x[z_1^{M(t)} z_2^{P(t)} \mathbb{1}_{\{I(t)=y\}}]\right]_{x,y \in S}, \tag{2}$$

where $\mathbb{E}_x$ denotes expectation with respect to $\mathbb{P}_x$, and $\mathbb{1}_A$ denotes the indicator of A, where $\mathbb{1}_A = 1$ if the event A occurs, and $\mathbb{1}_A = 0$ otherwise. We will also be interested in studying the matrices $\mathbf{C}_{m,n}(t)$, where for each integer $m, n \in \mathbb{Z}_+$,

$$\mathbf{C}_{m,n}(t) := \left[\mathbb{E}_x[M(t)^m P(t)^n \mathbb{1}_{\{I(t)=y\}}]\right]_{x,y \in S}. \tag{3}$$

These matrices $\{\mathbf{C}_{m,n}(t)\}_{m,n\geq 0}$ yield important information about various moments and cross-moments associated with $M(t)$ and $P(t)$.

Our main results include formulae for $\mathbf{J}$, $\hat{\mathbf{J}}$, and $\mathbf{C}$ in Theorems 2.4, 2.5, and 2.6, respectively. One could in principle use either Theorem 2.4 or Theorem 2.5 to prove Theorem 2.6 but instead we show how Theorem 2.6 follows more easily from an argument analogous to the arguments used to derive Theorems 2.4 and 2.5. We also illustrate how our results simplify for the case where the gene is assumed to always be active. For this special case, it is possible to derive the joint PGF of $M(t)$ and $P(t)$ explicitly, and from this joint PGF the reader will see how Theorem 2.5 is not so surprising.

2.1 Markovian Arrival Processes

A MAP is a continuous-time Markov chain $\{(N(t), I(t)); t \geq 0\}$ having state space $\mathbb{Z}_+ \times S$ and transition rate matrix $\mathbf{Q}$, with S being a finite set. We assume without loss of generality that $S := \{0, 1 \ldots, n\}$ for some integer $n \geq 0$, so the state space can be ordered lexicographically, and we order the states in this way when we construct the rows and columns of $\mathbf{Q}$. Here $N(t) \in \mathbb{Z}_+$ is a counting process (in our gene expression model $N(t)$ represents the number of mRNA polymerase bindings that occur in the interval $(0, t]$) and $I(t) \in S$ denotes the phase that the system is in at time t, which in our model corresponds to the state of the gene at time t (e.g. in the three-stage model of gene expression analysed in [26] $S = \{0, 1\}$, where $I(t) = 0$ means that the gene is inactive). The transition rates of such a process are governed by a pair of $|S| \times |S|$ matrices, $\mathbf{K}_0$ and $\mathbf{K}_1$. The off-diagonal entries of $\mathbf{K}_0$ (are non-negative and) correspond to transition rates associated with phase transition instants that are not counted by the counting process. Similarly, the off-diagonal entries of $\mathbf{K}_1$ correspond to transition rates associated with phase transitions that are counted by the counting process, and the diagonal entries of $\mathbf{K}_1$ correspond to instants at which the counting process increases by one unit, without there being a change in the phase process. In queueing models, transition instants corresponding to rates from $\mathbf{K}_1$ often represent arrival times of customers to a queueing system, which is where the name *Markovian Arrival Process* comes from. Finally, the diagonal entries of $\mathbf{K}_0$ are then chosen so that each row sum of $\mathbf{Q}$ is zero, and due to $\mathbb{Z}_+ \times S$ being lexicographically ordered, $\mathbf{Q}$ can be expressed as follows:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{K}_0 & \mathbf{K}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{K}_0 & \mathbf{K}_1 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_0 & \mathbf{K}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K_0} & \mathbf{K}_1 & \ddots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K}_0 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}.$$

We refer to $\{(N(t), I(t)); t \geq 0\}$ as the $(\mathbf{K}_0, \mathbf{K}_1)$-MAP.

Example 2.1. The simple model of mRNA production with $S = \{0, 1\}$ above can be represented via

$$\mathbf{K}_0 = \begin{pmatrix} -a & a \\ \lambda & -(\rho + \lambda) \end{pmatrix}, \qquad \mathbf{K}_1 = \begin{pmatrix} 0 & 0 \\ 0 & \rho \end{pmatrix}.$$

where ρ is the rate at which RNA polymerase bindings occur with the (active) gene, a is the rate at which the gene (when inactive) becomes active, and λ is the rate at which the gene (when active) becomes inactive.

MAPs play a prominent role in what is known in applied probability as the theory of matrix-analytic methods: a textbook-level treatment of these topics is given in e.g. Latouche and Ramaswami [17] and He [9]. One way to construct a MAP is to do so by thinking of it as being governed by a finite collection of independent, homogeneous Poisson processes, where each Poisson process is associated with a particular element from $\mathbf{Q}$ (i.e. from $\mathbf{K}_0$ or $\mathbf{K}_1$), where

$$\mathbf{K}_0 := [k_0(z,w)]_{z,w\in S}, \qquad \mathbf{K}_1 := [k_1(z,w)]_{z,w\in S}.$$

Each off-diagonal rate $k_0(z,w)$ has associated with it a homogeneous Poisson process $\{A^{(0)}_{z,w}(t); t \geq 0\}$ with rate $k_0(z,w)$, and each element $k_1(z,w)$ has associated with it a homogeneous Poisson process $\{A^{(1)}_{z,w}(t); t \geq 0\}$ with rate $k_1(z,w)$. We assume all of these Poisson processes are independent of each other. These processes provide all of the randomness needed to construct $\{(N(t), I(t)); t \geq 0\}$ (from a given initial state). The evolution of $I(\cdot)$ is determined by the processes $\mathcal{I} = \{A^{(i)}_{z,w} : i \in \{0,1\}, z, w \in S, z \neq w\}$: when the current state is $I(t) = \ell$ the next transition of $I(\cdot)$ occurs at the next firing time among the processes $A^{(\cdot)}_{\ell,\cdot} \in \mathcal{I}$ (i.e. those with $z = \ell$), and the new state is the $w \neq \ell$ whose Poisson process fired at this time. In particular, $\{I(t); t \geq 0\}$ is a CTMC having generator matrix $\mathbf{K}_0 + \mathbf{K}_1$. Readers interested in a more rigorous construction that shows how a collection of independent, homogeneous Poisson processes can be used to govern CTMC are referred to Chapter 9 of Brémaud [6].

The $N(t)$ process can be constructed similarly from the full set of Poisson processes (or from the $I(t)$ process and the $A^{(1)}$ Poisson processes). For each $t \geq 0$,

$$N(t) := \sum_{z,w\in S} \int_{(0,t]} \mathbb{1}_{\{I(s-)=z\}} A^{(1)}_{z,w}(ds)$$

so that $N(t)$ counts the number of firings (up to time t) of $A^{(1)}_{z,\cdot}$ processes that occur while I is in the corresponding state z, summed over $z \in S$. Readers should interpret each integral encountered in this paper as a Lebesgue-Stieltjes integral.

Example 2.2. In Example 2.1, $(I(t))_{t\geq 0}$ is a 2-state CTMC with generator $\begin{pmatrix} -a & a \\ \lambda & -\lambda \end{pmatrix}$, and $N(t) = \int_{(0,t]} \mathbb{1}_{\{I(s-)=1\}} A^{(1)}_{1,1}(ds)$.

To illustrate the flexibility of this class of processes, consider the following example.

Example 2.3. Let $S = \{0,1,2\}$ and

$$\mathbf{K}_0 = \begin{pmatrix} -a_1 & a_1 & 0 \\ \lambda_- & -(\rho_1 + \lambda_- + \lambda_+) & \lambda_+ \\ b_0 & b_1 & -(b_0 + b_1 + \rho_2 + \rho_{2,1}) \end{pmatrix}, \qquad \mathbf{K}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \rho_1 & 0 \\ 0 & \rho_{2,1} & \rho_2 \end{pmatrix}.$$

An interpretation of this MAP in terms of a mRNA production model could be as follows: When the gene is inactive there are no RNA polymerase bindings. There are two "active" phases. The (inactive) gene becomes active-1 at rate a_1, and when active-1 the gene becomes inactive at rate λ_- and active-2 with rate λ_+ (for each such transition no RNA polymerase bindings occur). Such bindings do occur at rate ρ_1 while the gene is active-1 and at rate $\rho_2 + \rho_{2,1}$ while the gene is active-2. The rate ρ_2 corresponds to bindings while the gene stays as active-2. The rate $\rho_{2,1}$ corresponds to transitions where a binding occurs with the active-2 gene and the gene immediately becomes active-1. While active-2, the gene may also become active-1 (at rate b_1) or inactive (at rate b_0) via transitions that do not produce an RNA molecule.

2.2 Main Results

The main results of this paper will address the following three-stage model of stochastic gene expression, which behaves in accordance to the following rules:

(A1) the transcription process within a cell is assumed to be governed by a MAP $\{(N(t), I(t)); t \geq 0\}$ having finite phase set S and matrices $\mathbf{K}_0$ and $\mathbf{K}_1$, where each point of N corresponds to an instant where an RNA polymerase binds with the (active) gene to begin producing an mRNA molecule,

(A2) for each such binding, the amount of time it takes to create an mRNA molecule is random with CDF L_1, and independent of everything else (let $(D^{(i)})_{i\in\mathbb{N}}$ denote these (i.i.d.$\sim L_1$) elongation times),

(A3) the lifetime of each mRNA is assumed to be generally distributed with CDF F, and independent of everything else (let $(B^{(i)})_{i\in\mathbb{N}}$ denote these (i.i.d.$\sim F$) lifetimes),

(A4) while an mRNA exists, it initiates the creation of proteins in accordance to a homogeneous Poisson process having rate $k_2 > 0$, independent of everything else,

(A5) when an mRNA initiates the creation of a protein, it takes a random amount of time having CDF L_2 to actually create the protein, with this creation time being independent of everything else,

(A6) each protein exists in the cell for a random amount of time, having CDF G, independent of everything else.

We also assume throughout that the cdfs F, G, L_1, and L_2 are all proper, and correspond to nonnegative random variables: saying F satisfies this criteria means $\lim_{x\uparrow 0} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$. In fact, in order to guarantee the random vector $(I(t), M(t), P(t))$ converges in distribution as $t \to \infty$ to a non-degenerate limit we will also need to assume the random variables associated with L_1, F, L_2, and G have finite means.

It is important to point out that the MAP only determines the instants at which a RNA polymerase binds with the (active) gene. For instance, if the gene switches from being active to being inactive while an mRNA is currently going through its elongation phase, that elongation phase is not affected in any way, nor is the lifetime following that phase. This may or may not be realistic, but this is how elongation phases behave in both [12] and [26].

Before we begin stating our mathematical results, we first need to introduce some additional notation, much of which is standard. Given a proper cdf F defined on $\mathbb{R}$, we let $\overline{F}$ denote its corresponding tail function, meaning $\overline{F}(x) := 1 - F(x)$ for each $x \in \mathbb{R}$. Recall also that the CDF of the sum of two independent nonnegative random variables with CDFs F and G is given by the convolution $F \star G$, which is defined as

$$(F \star G)(t) = \int_{[0,t]} G(t-s)dF(s), \quad t \geq 0.$$

As we proceed through this paper, we will require an understanding of the distribution of the number of proteins present in the system at time $s + t$ that were produced by an RNA polymerase binding event that occurred at time s. This distribution depends on t but not s, so without loss of generality we consider the case $s = 0$ in what follows. Associated with an RNA polymerase that binds with

the (active) gene at time zero is the (elongation) time $D_0 \sim L_1$ that it takes to create an mRNA molecule, and $B_0 \sim F$, the lifetime of that mRNA molecule. Finally, let $P(\{0\}, t)$ denote the number of proteins created by that mRNA molecule that are present in the cell at time t. We assume that this "bind at time 0" process is defined on some space with probability measure $\mathcal{P}_0$ and associated expectation $\mathcal{E}_0$.

Recall the definition (1), and note that $\mathbf{J}(0, m, n) = \mathbf{I}\mathbb{1}_{\{m=n=0\}}$, where $\mathbf{I}$ denotes the identity matrix. Our first main result shows that the matrices $\mathbf{J}(t, m, n)$ satisfy a relatively simple recursion.

Theorem 2.4. *Assume (A1)-(A6). For Lebesgue almost every $t > 0$, the following statements hold:*

(a) For each integer $n \geq 0$,

$$\frac{\partial}{\partial t}\mathbf{J}(t,0,n) = \mathbf{K}_0\mathbf{J}(t,0,n) + \overline{L}_1(t)\mathbf{K}_1\mathbf{J}(t,0,n) + \sum_{k=0}^{n} \mathcal{P}_0(D_0 + B_0 \leq t, P(\{0\},t) = n-k)\mathbf{K}_1\mathbf{J}(t,0,k).$$

(b) For each integer $m \geq 1$, and each integer $n \geq 0$,

$$\begin{aligned}\frac{\partial}{\partial t}\mathbf{J}(t,m,n) = \mathbf{K}_0\mathbf{J}(t,m,n) &+ \overline{L}_1(t)\mathbf{K}_1\mathbf{J}(t,m,n) \\ &+ \sum_{k=0}^{n} \mathcal{P}_0(D_0 \leq t, D_0 + B_0 > t, P(\{0\},t) = n-k)\mathbf{K}_1\mathbf{J}(t,m-1,k) \\ &+ \sum_{k=0}^{n} \mathcal{P}_0(D_0 + B_0 \leq t, P(\{0\},t) = n-k)\mathbf{K}_1\mathbf{J}(t,m,k).\end{aligned}$$

Our second main result concerns the corresponding probability generating function. Recall (2) and note that $\hat{\mathbf{J}}(0, z_1, z_2) = \mathbf{I}$. Define the function $h : [0,\infty) \times \mathbb{D} \times \mathbb{D} \to \mathbb{C}$ (see (4)) as

$$h(t, z_1, z_2) := \mathcal{E}_0\left[z_1^{\mathbb{1}\{D_0 \leq t, D_0 + B_0 > t\}} z_2^{P(\{0\},t)}\right].$$

Theorem 2.5. *Assume (A1)-(A6). Then the matrices $(\hat{\mathbf{J}}(t, z_1, z_2))_{t>0}$ satisfy the following: for each $z_1, z_2 \in \mathbb{D}$, for Lebesgue almost every $t > 0$ we have,*

$$\frac{\partial}{\partial t}\hat{\mathbf{J}}(t, z_1, z_2) = (\mathbf{K}_0 + h(t, z_1, z_2)\mathbf{K}_1)\hat{\mathbf{J}}(t, z_1, z_2).$$

Readers should note that Theorems 2.4 and 2.5 are analogous in form to results found in Ramaswami and Neuts [22], which addresses the marginal distributions of the queue-length process of an infinite-server queue where customers arrive to the queue in accordance to a phase-type renewal process. An extension of the results found in [22] to the case where batches of arrivals occur in accordance to a Markovian arrival process can be found in Masuyama and Takine [18].

We next address the problem of calculating the $\mathbf{C}_{m,n}(t)$ matrices, for each real $t \geq 0$ and each pair of integers $m, n \geq 0$. Theoretically these moments can be derived from their joint PGF by taking derivatives. We present a more direct approach via point process theory which yields a recursive scheme that in principle can be used to find all moments. These matrices can be stated in terms of

matrix exponentials: recall that we can associate with each square matrix $\mathbf{M}$ its matrix exponential $\mathbf{e}^{\mathbf{M}}$, which is defined as

$$\mathbf{e}^{\mathbf{M}} = \sum_{m=0}^{\infty} \frac{1}{m!}\mathbf{M}^m$$

with $\mathbf{M}^0 \equiv \mathbf{I}$. We also follow the convention throughout that all empty sums we encounter are equal to zero. Define $q_{j,k} : [0,\infty) \to [0,\infty)$, for each $j \in \{0,1\}$ and each integer $k \geq 0$ as follows:

$$q_{j,k}(t) := \mathcal{E}_0\big[\mathbb{1}^j_{\{D_0 \leq t, D_0+B_0>t\}} P(\{0\},t)^k\big].$$

Theorem 2.6. *Assume (A1)-(A6). Then for each integer $m \geq 0$, and each integer $n \geq 0$,*

$$\mathbf{C}_{m,n}(t) = \mathbf{e}^{(\mathbf{K}_0+\mathbf{K}_1)t}\mathbb{1}_{\{m=n=0\}} + \sum_{j=0}^{m-1}\sum_{k=0}^{n}\binom{m}{j}\binom{n}{k}\int_0^t \mathbf{e}^{(\mathbf{K}_0+\mathbf{K}_1)(t-s)} q_{1,n-k}(s)\mathbf{K}_1\mathbf{C}_{j,k}(s)ds$$
$$+ \sum_{k=0}^{n-1}\binom{n}{k}\int_0^t \mathbf{e}^{(\mathbf{K}_0+\mathbf{K}_1)(t-s)} q_{0,n-k}(s)\mathbf{K}_1\mathbf{C}_{m,k}(s)ds.$$

In particular, the matrices $\mathbf{C}_{1,0}(t)$ and $\mathbf{C}_{0,1}(t)$ are as follows:

$$\mathbf{C}_{1,0}(t) = \int_0^t \mathbf{e}^{(\mathbf{K}_0+\mathbf{K}_1)(t-s)}(\overline{L_1 \star F}(s) - \overline{L_1}(s))\mathbf{K}_1\mathbf{e}^{(\mathbf{K}_0+\mathbf{K}_1)s}ds$$
$$\mathbf{C}_{0,1}(t) = \int_0^t \mathbf{e}^{(\mathbf{K}_0+\mathbf{K}_1)(t-s)} q_{0,1}(s)\mathbf{K}_1\mathbf{e}^{(\mathbf{K}_0+\mathbf{K}_1)s}ds.$$

These 'moment matrices' satisfy a simpler recursive structure than the matrices we studied in Theorems 2.4 and 2.5, thanks to $(\mathbf{K}_0 + \mathbf{K}_1)t$ and its derivative with respect to t being commutative with respect to matrix multiplication.

3 Proofs of the main results

In this Section we prove Theorems 2.4-2.6. A tool that we will make use of in all three proofs is the Campbell-Mecke formula [3], which will be stated below (see also Theorem 3.1.9 on page 121 of [3]). In the meantime we present a lemma that quantifies the relationship between D_0, B_0, and $P(\{0\},t)$. In order to simplify the statement of this lemma, we define the function $r : \mathbb{R}^2_+ \to \mathbb{R}$ as

$$r(a,b) := k_2\int_a^b (\overline{L_2 \star G}(t-x) - \overline{L_2}(t-x))dx.$$

Henceforth, by convention "Poisson with mean 0" means almost surely equal to 0.

Lemma 3.1. *The following statements are true.*

$$h(t,z_1,z_2) = z_1\int_{[0,t]} \overline{F}(t-u)e^{-(1-z_2)r(u,t)}dL_1(u)$$
$$+ \int_{[0,t]}\int_{[0,t-u]} e^{-(1-z_2)r(u,u+v)}dF(v)dL_1(u) + \overline{L_1}(t), \tag{4}$$

$$\mathcal{P}_0\big(D_0+B_0 \leq t, P(\{0\},t) = k\big) = \int_{[0,t]}\int_{[0,t-u]} \frac{\big(r(u,u+v)\big)^k e^{-r(u,u+v)}}{k!}dF(v)dL_1(u), \tag{5}$$

$$\mathcal{P}_0\big(D_0 \leq t, D_0+B_0 > t, P(\{0\},t) = k\big) = \int_{[0,t]} \overline{F}(t-u)\frac{\big(r(u,t)\big)^k e^{-r(u,t)}}{k!}dL_1(u), \tag{6}$$

and

$$q_{j,k}(t) = \int_{[0,\infty)} \int_{[0,\infty)} \left(\mathbb{1}_{\{u+v>t\}}\right)^j \mu_{k,d}\big((u \wedge t, (u+v) \wedge t]\big) dF(v) dL_1(u) \tag{7}$$

where $\mu_{k,d}((a,b])$ is the kth moment of a Poisson random variable having mean $r(a,b)$.

Proof. First, note that a protein initiated at time x in the $\mathcal{P}_0$-setting is present at time t if its elongation time ($\sim L_2$) is less than $t-x$, and its elongation time plus lifetime ($\sim L_2 \star G$) is greater than $t-x$. Thus, conditional on both D_0 and B_0, $P(\{0\},t)$ is a Poisson random variable with mean $\lambda_t(D_0,B_0)$, where

$$\lambda_t(a,b) := k_2 \int_{\min(a,t)}^{\min(a+b,t)} (\overline{L_2 \star G}(t-x) - \overline{L_2}(t-x))dx = r\big(\min(a,t), \min(a+b,t)\big).$$

Thus, for functions f_1, f_2, where $f_1 : [0,\infty)\times[0,\infty) \to [0,\infty)$ is $\mathcal{B}([0,\infty))\otimes\mathcal{B}([0,\infty))$-measurable, and $f_2 : \mathbb{Z}_+ \to [0,\infty)$, by conditioning on both B_0 and D_0, we can write

$$\begin{aligned}\mathcal{E}_0\big[f_1(D_0,B_0)f_2(P(\{0\},t))\big] &= \mathcal{E}_0\Big[f_1(D_0,B_0)\mathcal{E}_0\big[f_2(P(\{0\},t))\big|D_0,B_0\big]\Big] \\ &= \int_0^\infty \int_0^\infty f_1(u,v)\phi_2(t,u,v)dF(v)dL_1(u),\end{aligned}$$

where $\phi_2(t,u,v)$ is the expected value of $f_2(N^*)$, where $N^* \sim$Pois$(\lambda_t(u,v))$.

From this observation Formulas (4), (5), (6), and (7) can be derived quickly, so we will only spend time establishing (4): since $\{D_0 > t\}$ and $\{D_0 + B_0 \le t\}$ are disjoint events we have

$$\mathcal{E}_0\big[z_1^{\mathbb{1}_{\{D_0\le t, D_0+B_0>t\}}} z_2^{P(\{0\},t)}\big] = \mathcal{E}_0\big[(z_1\mathbb{1}_{\{D_0\le t, D_0+B_0>t\}} + \mathbb{1}_{\{D_0>t\}} + \mathbb{1}_{\{D_0+B_0\le t\}})z_2^{P(\{0\},t)}\big].$$

Condition on D_0, B_0 as above and recall that the probability generating function for a Poisson(λ) random variable is $\phi(z) = e^{-\lambda(1-z)}$ to see that this is equal to

$$z_1 \int_{[0,t]} \overline{F}(t-u)e^{-(1-z_2)r(u,t)}dL_1(u) \quad + \overline{L_1}(t) \quad + \int_{[0,t]}\int_{[0,t-u]} e^{-(1-z_2)r(u,u+v)}dF(v)dL_1(u).$$

which proves the claim (4). □

In order to ensure that the Palm distributions we need can be constructed, we assume throughout that all random elements exist on an underlying probability space $(\Omega,\mathcal{F},\mathbb{P})$, where we further assume Ω is a Polish space and $\mathcal{F}$ is its associated Borel σ-field. Readers should not view this as an issue, as all random elements we will encounter can be constructed on such a space: recall that a countable product space of Polish spaces is still a Polish space if we endow the countable product space with a suitable metric that generates the product topology (see e.g. Theorem 14.8 on page 275 of Klenke [16]) and this topology will be fine enough to ensure all random elements we encounter are appropriately measurable with respect to the Borel σ-field generated by the topology. In fact, the topology itself will never be used in our analysis, its only role is to provide enough structure on the underlying space to ensure Palm distributions (to be defined below) exist.

Given a simple point process $\{\mathcal{A}(t); t \ge 0\}$—recall that a point process is simple if none of its points overlap with probability one—with points found in $[0,\infty)$, we associate with it its mean measure μ, which is defined on the Borel σ-field $\mathcal{B}([0,\infty))$ as

$$\mu(B) := \mathbb{E}[\mathcal{A}(B)], \quad B \in \mathcal{B}([0,\infty)).$$

We assume throughout that μ is σ-finite on $([0,\infty), \mathcal{B}([0,\infty)))$.

Under these assumptions, we can define the *Campbell measure* C on the measurable space $(\Omega \times [0,\infty), \mathcal{F} \otimes \mathcal{B}([0,\infty)))$ as the unique measure on this space that satisfies the following property: for each measurable rectangle $A \times B \in \mathcal{F} \otimes \mathcal{B}([0,\infty))$,

$$C(A \times B) = \mathbb{E}[\mathbb{1}_A \mathcal{A}(B)].$$

From this Campbell measure, it is possible to show—see Chapter 10, pages 83 and 84 of Kallenberg [14] that there exists a μ-almost surely unique collection of probability measures $\{\mathcal{P}_t\}_{t\geq 0}$, known as *Palm distributions*, that satisfy, for each measurable rectangle $A \times B \in \mathcal{F} \otimes \mathcal{B}([0,\infty))$,

$$C(A \times B) = \int_B \mathcal{P}_s(A)\mu(ds).$$

These Palm distributions, or Palm probabilities, can be constructed in a manner so that they form a probability kernel on $\mathcal{F} \times [0,\infty)$. Throughout we let $\mathcal{E}_s$ be the expectation associated with $\mathcal{P}_s$.

The following result is known as the Campbell-Mecke formula, which can be proven using a standard approximation procedure, coupled with Dynkin's Theorem.

Theorem 3.1. *(*Campbell-Mecke Formula*) Let $\{\mathcal{A}(s); s \geq 0\}$ be a simple point process on $[0,\infty)$ having mean measure μ. For each nonnegative $\mathcal{F} \otimes \mathcal{B}([0,\infty))$-measurable stochastic process $\{X(s); s \geq 0\}$ we have that for each $t \geq 0$,*

$$\mathbb{E}\left[\int_0^t X(s)\mathcal{A}(ds)\right] = \int_0^t \mathcal{E}_s[X(s)]\mu(ds).$$

The Palm probability $\mathcal{P}_s(A)$ can intuitively be interpreted as the probability of the event A, given that $\{\mathcal{A}(t); t \geq 0\}$ has a single point at time s. In general, Palm probabilities can be difficult to calculate explicitly, but the Palm probabilities associated with a Poisson process take on a very simple form. The following result, attributed to Silvnyak and Mecke, can be found on page 130 of [3]. It says that *conditioning* a Poisson process on having a point at time s (note that this is an event of probability 0) is equivalent to *adding* a point at time s to the unconditioned process.

Theorem 3.2. *Let $\{\mathcal{A}(t); t \geq 0\}$ be a Poisson process on $[0,\infty)$, with Palm probabilities $\{\mathcal{P}_t\}_{t\geq 0}$. Then under the measure $\mathcal{P}_s$, the law of $\{\mathcal{A}(t); t \geq 0\}$ is (for all $s \in [0,\infty)$ outside of a Borel set having μ-measure zero) simply the law of $\{\mathcal{A}_s(t); t \geq 0\}$ under $\mathbb{P}$, where for each $s \geq 0$,*

$$\mathcal{A}_s = \mathcal{A} + \delta_s$$

with $\delta_s(\cdot)$ being the measure on $([0,\infty), \mathcal{B}([0,\infty)))$ associated with a single unit mass located at time s.

We now have the tools we need to prove our main results. We define the *conception time* of any mRNA or protein molecule as the time of the binding of the RNA polymerase (that led to the creation of the molecule) with the (active) gene. Throughout our derivations, for each $t \geq 0$, and each $B \in \mathcal{B}([0,\infty))$, $M(B,t)$ (resp. $P(B,t)$) represents the number of mRNA (resp. protein) molecules present at time t with conception time in B. The elongation time and lifetime associated to the first arrival in the MAP are denoted by $D^{(1)}$ and $B^{(1)}$ respectively. In each case our proof starts with the

observation that (for particular choices of f_1, f_2) we have

$$f_1(M(t))f_2(P(t)) \tag{8}$$
$$= f_1(0)f_2(0)\mathbb{1}_{\{N(t)=0\}} + \sum_{z,w\in S}\int_{(0,t]} \mathbb{1}_{\{N(s-)=0,I(s-)=z\}} f_1\big(\mathbb{1}_{\{D^{(1)}\le t-s, D^{(1)}+B^{(1)}>t-s\}} + M((s,t],t)\big)$$
$$\times f_2\big(P(\{s\},t) + P((s,t],t)\big)A^{(1)}_{z,w}(ds).$$

We then multiply both sides by $\mathbb{1}_{\{I(t)=y\}}$ and use the particular form of f_1, f_2 to write $f_i(n+m)$ in terms of $f_i(k)$ etc. (in our first two proofs, each f_i factorizes, and in our third proof we use the binomial expansion). We take expectations of both sides and apply both Theorems 3.1 and 3.2 to the right hand side. Exploiting independence allows us to factorize the resulting Palm expectation. Simplifying, expressing in matrix form, and differentiating with respect to t completes the proof.

Proof of Theorem 2.4. For each $t > 0$, and each phase $y \in S$, using (8) with $f_1(m) = \mathbb{1}_{\{m=0\}}$ and $f_2(m) = \mathbb{1}_{\{m=n\}}$ we obtain,

$$\begin{aligned}\mathbb{1}_{\{M(t)=0,P(t)=n,I(t)=y\}} &= \mathbb{1}_{\{N(t)=0,I(t)=y\}}\mathbb{1}_{\{n=0\}}\\
&+ \sum_{z,w\in S}\int_{(0,t]} \mathbb{1}_{\{N(s-)=0,I(s-)=z,D^{(1)}>t-s,M((s,t],t)=0,P((s,t],t)=n,I(t)=y\}}A^{(1)}_{z,w}(ds)\\
&+ \sum_{k=0}^{n}\sum_{z,w\in S}\int_{(0,t]} \mathbb{1}_{\{N(s-)=0,I(s-)=z,D^{(1)}+B^{(1)}\le t-s,P(\{s\},t)=n-k\}}\\
&\qquad\times \mathbb{1}_{\{M((s,t],t)=0,P((s,t],t)=k,I(t)=y\}}A^{(1)}_{z,w}(ds).\end{aligned}$$

This equality follows from the following observations. Suppose first that $n = 0$: in order for the event $\{M(t) = 0, P(t) = 0, I(t) = y\}$ to happen, either (a) $\{N(t) = 0, I(t) = y\}$, meaning no mRNA are conceived in $(0, t]$ and $I(t) = y$; (b) an mRNA is conceived in $(0, t]$, and the first mRNA conceived in $(0, t]$ is still undergoing its elongation phase at time t; (c) an mRNA is conceived in $(0, t]$, the first mRNA conceived in $(0, t]$ dies before time t, the number of proteins present in the system at time t created by the first mRNA conceived is 0. This equality can be understood in a similar manner for the case where $n \ge 1$.

After taking expectations of both sides, while applying the Campbell-Mecke formula to e.g. the process

$$X_t(s) := \mathbb{1}_{\{N(s-)=0,I(s-)=z,D^{(1)}>t-s,M((s,t],t)=0,P((s,t],t)=n,I(t)=y\}},$$

on the right-hand-side, we find that for each $x \in S$, and $t \geq 0$

$$\begin{aligned}
&\mathbb{P}_x(M(t) = 0, P(t) = n, I(t) = y) \\
&= \mathbb{P}_x(N(t) = 0, I(t) = y)\mathbb{1}_{\{n=0\}} \\
&\quad + \sum_{z,w \in S} \int_0^t \mathcal{P}_{x,s}^{(z,w)}\Big(N(s-) = 0, I(s-) = z, D^{(1)} > t - s, \\
&\qquad\qquad M((s,t],t) = 0, P((s,t],t) = n, I(t) = y\Big) k_1(z,w)ds \\
&\quad + \sum_{k=0}^{n} \sum_{z,w \in S} \int_0^t \mathcal{P}_{x,s}^{(z,w)}\Big(N(s-) = 0, I(s-) = z, D^{(1)} + B^{(1)} \leq t - s, \\
&\qquad\qquad P(\{s\},t) = n - k, M((s,t],t) = 0, P((s,t],t) = k, I(t) = y\Big) k_1(z,w)ds,
\end{aligned}$$

where $\{\mathcal{P}_{x,s}^{(z,w)}\}_{s \geq 0}$ denotes the family of Palm distributions associated with the Poisson process $A_{z,w}^{(1)}$ (when the underlying probability measure is $\mathbb{P}_x$). Next, observe that the event $\{N(s-) = 0, I(s-) = z\}$ depends only on the behaviour of all points on the interval $(0, s)$, the event $\{M((s,t],t) = 0, P((s,t],t) = n, I(t) = y\}$ depends only on the behaviour of the point processes in the set $(s,t]$ (while simultaneously keeping in mind that $I(s) = w$) and the events $\{D^{(1)} > t - s\}$ and $\{D^{(1)} + B^{(1)} \leq t - s\}$ do not depend on the point process at all for fixed s.

Applying now Theorem 3.2 yields, for each $t \geq 0$ and Lebesgue almost-all $s \in [0,t]$,

$$\begin{aligned}
&\mathcal{P}_{x,s}^{(z,w)}\Big(N(s-) = 0, I(s-) = z, D^{(1)} > t - s, M((s,t],t) = 0, P((s,t],t) = n, I(t) = y\Big) \\
&= \mathbb{P}_x\big(N(s) = 0, I(s) = z\big)\mathcal{P}_0\big(D_0 > t - s\big)\mathbb{P}_w\big(M(t-s) = 0, P(t-s) = n, I(t-s) = y\big)
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{P}_{x,s}^{(z,w)}\Big(N(s-) = 0, I(s-) = z, D^{(1)} + B^{(1)} \leq t - s, \\
&\qquad P(\{s\},t) = n - k, M((s,t],t) = 0, P((s,t],t) = k, I(t) = y\Big) \\
&= \mathbb{P}_x\big(N(s) = 0, I(s) = z\big)\mathcal{P}_0\big(D_0 + B_0 \leq t - s, P(\{0\}, t-s) = n - k\big) \\
&\quad \times \mathbb{P}_w\big(M(t-s) = 0, P(t-s) = k, I(t-s) = y\big).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\mathbb{P}_x(M(t) = 0, P(t) = n, I(t) = y) \\
&= \mathbb{P}_x(N(t) = 0, I(t) = y)\mathbb{1}_{\{n=0\}} \\
&\quad + \sum_{z,w \in S} \int_0^t \overline{L}_1(t-s)\mathbb{P}_x\big(N(s) = 0, I(s) = z\big)k_1(z,w)\mathbb{P}_w\big(M(t-s) = 0, P(t-s) = n, I(t-s) = y\big)ds \\
&\quad + \sum_{k=0}^{n} \sum_{z,w \in S} \int_0^t \mathbb{P}_x\big(N(s) = 0, I(s) = z\big)\mathcal{P}_0\big(D_0 + B_0 \leq t - s, P(\{0\}, t-s) = n - k\big)k_1(z,w) \\
&\qquad\qquad \times \mathbb{P}_w\big(M(t-s) = 0, P(t-s) = k, I(t-s) = y\big)ds.
\end{aligned}$$

Observe also that from the Kolmogorov forward equations associated with the CTMC $\{(N(t), I(t)); t \geq 0\}$, the probability $\mathbb{P}_x(N(t) = 0, I(t) = y)$ is the element found in row x, column y of the matrix exponential $\mathbf{e}^{\mathbf{K}_0 t}$. Expressing the above system in matrix form and changing variables to $u = t - s$ we obtain

$$\begin{aligned}
\mathbf{J}(t,0,n) &= \mathbf{e}^{\mathbf{K}_0 t}\mathbb{1}_{\{n=0\}} + \int_0^t \overline{L}_1(t-s)\mathbf{e}^{\mathbf{K}_0 s}\mathbf{K}_1\mathbf{J}(t-s,0,n)ds \\
&\quad + \sum_{k=0}^{n}\int_0^t \mathcal{P}_0\big(D_0 + B_0 \leq t-s, P(\{0\}, t-s) = n-k\big)\mathbf{e}^{\mathbf{K}_0 s}\mathbf{K}_1\mathbf{J}(t-s,0,k)ds \\
&= \mathbf{e}^{\mathbf{K}_0 t}\Bigg[\mathbb{1}_{\{n=0\}} + \int_0^t \overline{L}_1(u)\mathbf{e}^{-\mathbf{K}_0 u}\mathbf{K}_1\mathbf{J}(u,0,n)du \\
&\qquad + \sum_{k=0}^{n}\int_0^t \mathcal{P}_0\big(D_0 + B_0 \leq u, P(\{0\}, u) = n-k\big)\mathbf{e}^{-\mathbf{K}_0 u}\mathbf{K}_1\mathbf{J}(u,0,k)du\Bigg].
\end{aligned}$$

Taking derivatives of both sides with respect to t establishes part (a). A similar set of equalities can be found for the case where $m \geq 1$, and $n \geq 0$: for each $t > 0$, $y \in S$,

$$\begin{aligned}
\mathbb{1}_{\{M(t)=m,P(t)=n,I(t)=y\}} &= \sum_{z,w\in S}\int_{(0,t]} \mathbb{1}_{\{N(s-)=0,I(s-)=z,D^{(1)}>t-s,M((s,t],t)=m,P((s,t],t)=n,I(t)=y\}}A^{(1)}_{z,w}(ds) \\
&\quad + \sum_{k=0}^{n}\sum_{z,w\in S}\int_{(0,t]} \mathbb{1}_{\{N(s-)=0,I(s-)=z,D^{(1)}\leq t-s,D^{(1)}+B^{(1)}>t-s,P(\{s\},t)=n-k\}} \\
&\qquad\qquad \times \mathbb{1}_{\{M((s,t],t)=m-1,P((s,t],t)=k,I(t)=y\}}A^{(1)}_{z,w}(ds) \\
&\quad + \sum_{k=0}^{n}\sum_{z,w\in S}\int_{(0,t]} \mathbb{1}_{\{N(s-)=0,I(s-)=z,D^{(1)}+B^{(1)}\leq t-s,P(\{s\},t)=n-k\}} \\
&\qquad\qquad \times \mathbb{1}_{\{M((s,t],t)=m,P((s,t],t)=k,I(t)=y\}}A^{(1)}_{z,w}(ds).
\end{aligned}$$

Again, taking expectations of both sides and applying both Theorems 3.1 and 3.2 to the right-hand-

side gives

$$
\begin{aligned}
&\mathbb{P}_x(M(t)=m,P(t)=n,I(t)=y)\\
&=\sum_{z,w\in S}\int_0^t \overline{L}_1(t-s)\mathbb{P}_x\big(N(s)=0,I(s)=z\big)k_1(z,w)\mathbb{P}_w\big(M(t-s)=m,P(t-s)=n,I(t-s)=y\big)ds\\
&\quad+\sum_{k=0}^{n}\sum_{z,w\in S}\int_0^t \mathbb{P}_x\big(N(s)=0,I(s)=z\big)\mathcal{P}_0\big(D_0\le t-s,D_0+B_0>t-s,P(\{0\},t-s)=n-k\big)\\
&\qquad\qquad\times k_1(z,w)\mathbb{P}_w\big(M(t-s)=m-1,P(t-s)=k,I(t-s)=y\big)ds\\
&\quad+\sum_{k=0}^{n}\sum_{z,w\in S}\int_0^t \mathbb{P}_x\big(N(s)=0,I(s)=z\big)\mathcal{P}_0\big(D_0+B_0\le t-s,P(\{0\},t-s)=n-k\big)\\
&\qquad\qquad\times k_1(z,w)\mathbb{P}_w\big(M(t-s)=m,P(t-s)=k,I(t-s)=y\big)ds.
\end{aligned}
$$

Expressing in matrix form and changing variables gives

$$
\begin{aligned}
\mathbf{J}(t,m,n)=\mathbf{e}^{\mathbf{K}_0 t}\Bigg[&\int_0^t \overline{L}_1(u)\mathbf{e}^{-\mathbf{K}_0 u}\mathbf{K}_1\mathbf{J}(u,m,n)ds\\
&+\sum_{k=0}^{n}\int_0^t \mathcal{P}_0\big(D_0\le u,D_0+B_0>u,P(\{0\},u)=n-k\big)\mathbf{e}^{-\mathbf{K}_0 u}\mathbf{K}_1\mathbf{J}(u,m-1,k)ds\\
&+\sum_{k=0}^{n}\int_0^t \mathcal{P}_0\big(D_0+B_0\le u,P(\{0\},u)=n-k\big)\mathbf{e}^{-\mathbf{K}_0 u}\mathbf{K}_1\mathbf{J}(u,m,k)ds\Bigg].
\end{aligned}
$$

Again, taking partial derivatives with respect to t establishes (b), proving Theorem 2.4. □

We now turn our attention to the probability generating function $\hat{\mathbf{J}}$. The idea behind the proof we give for Theorem 2.5 is analogous to the idea we used to derive Theorem 2.4. Again, this appears to be closely related to the argument from pages 68 and 69 of [17] for the PH/G/∞ queue, except there they do not think in terms of point processes.

Proof of Theorem 2.5. Observe that for each $y\in S$ and $t>0$, using (8) with $f_1(m)=z_1^m$ and $f_2(m)=z_2^m$ we obtain

$$
\begin{aligned}
z_1^{M(t)}z_2^{P(t)}\mathbb{1}_{\{I(t)=y\}}=\mathbb{1}_{\{N(t)=0,I(t)=y\}}+\sum_{z,w\in S}\int_{(0,t]}&\mathbb{1}_{\{N(s-)=0,I(s-)=z\}}z_1^{\mathbb{1}_{\{D^{(1)}\le t-s,D^{(1)}+B^{(1)}>t-s\}}}\\
&z_2^{P(\{s\},t)}z_1^{M((s,t],t)}z_2^{P((s,t],t)}\mathbb{1}_{\{I(t)=y\}}A^{(1)}_{z,w}(ds).
\end{aligned}
$$

Taking expectations, and applying Theorems 3.1 and 3.2 yields, for each $x,y\in S$,

$$
\begin{aligned}
&\mathbb{E}_x[z_1^{M(t)}z_2^{P(t)}\mathbb{1}_{\{I(t)=y\}}]\\
&=\mathbb{P}_x(N(t)=0,I(t)=y)\\
&\quad+\sum_{z,w\in S}\int_0^t \mathbb{P}_x(N(s)=0,I(s)=z)k_1(z,w)h(t-s,z_1,z_2)\mathbb{E}_w[z_1^{M(t-s)}z_2^{P(t-s)}\mathbb{1}_{\{I(t)=y\}}]ds.
\end{aligned}
$$

In matrix form this is

$$\begin{aligned}\hat{\mathbf{J}}(t,z_1,z_2) &= \mathbf{e}^{\mathbf{K}_0 t} + \int_0^t h(t-s,z_1,z_2)\mathbf{e}^{\mathbf{K}_0 s}\mathbf{K}_1\hat{\mathbf{J}}(t-s,z_1,z_2)ds \\ &= \mathbf{e}^{\mathbf{K}_0 t}\left[1 + \int_0^t h(u,z_1,z_2)\mathbf{e}^{-\mathbf{K}_0 u}\mathbf{K}_1\hat{\mathbf{J}}(u,z_1,z_2)du\right],\end{aligned}$$

and taking the partial derivative of both sides with respect to t gives

$$\frac{\partial}{\partial t}\hat{\mathbf{J}}(t,z_1,z_2) = (\mathbf{K}_0 + h(t,z_1,z_2)\mathbf{K}_1)\hat{\mathbf{J}}(t,z_1,z_2)$$

proving the claim. □

It remains to prove Theorem 2.6.

Proof of Theorem 2.6. Using (8) with $f_1(k) = k^m$ and $f_2(k) = k^n$ gives

$$\begin{aligned}M(t)^m P(t)^n \mathbb{1}_{\{I(t)=y\}} &= \mathbb{1}_{\{m=n=0\}}\mathbb{1}_{\{N(t)=0,I(t)=y\}} \\ &\quad + \sum_{z,w\in S}\int_{(0,t]} \mathbb{1}_{\{N(s-)=0,I(s-)=z\}}\left(\mathbb{1}_{\{D^{(1)}\le t-s, D^{(1)}+B^{(1)}>t-s\}} + M((s,t],t)\right)^m \\ &\qquad \times \left(P(\{s\},t) + P((s,t],t)\right)^n \mathbb{1}_{\{I(t)=y\}} A^{(1)}_{z,w}(ds)\end{aligned}$$

After expanding the integrand with two applications of the binomial theorem, taking expectations and applying Theorems 3.1 and 3.2, we observe that for each $x, y \in S$,

$$\begin{aligned}&\mathbb{E}_x[M(t)^m P(t)^n \mathbb{1}_{\{I(t)=y\}}] \\ &= \mathbb{P}_x(N(t)=0, I(t)=y)\mathbb{1}_{\{m=n=0\}} \\ &\quad + \sum_{z,w\in S}\int_0^t \mathbb{P}_x\big(N(s)=0, I(s)=z\big)\sum_{j=0}^m\sum_{k=0}^n \binom{m}{j}\binom{n}{k}\mathcal{E}_s\left[\mathbb{1}^{m-j}_{\{D^{(1)}\le t-s, D^{(1)}+B^{(1)}>t-s\}} P(\{s\},t)^{n-k}\right] \\ &\qquad \times k_1(z,w)\mathbb{E}_w\left[M(t-s)^j P(t-s)^k \mathbb{1}_{\{I(t-s)=y\}}\right]ds.\end{aligned}$$

In matrix form (and following a substitution $u = t - s$) this is

$$\begin{aligned}\mathbf{C}_{m,n}(t) = \mathbf{e}^{\mathbf{K}_0 t}\Bigg[&\mathbb{1}_{\{m=n=0\}} + \sum_{j=0}^{m-1}\sum_{k=0}^n \binom{m}{j}\binom{n}{k}\int_0^t q_{1,n-k}(u)\mathbf{e}^{-\mathbf{K}_0 u}\mathbf{K}_1\mathbf{C}_{j,k}(u)du \\ &+ \sum_{k=0}^n \binom{n}{k}\int_0^t q_{0,n-k}(u)\mathbf{e}^{-\mathbf{K}_0 u}\mathbf{K}_1\mathbf{C}_{m,k}(u)du\Bigg].\end{aligned}$$

Differentiating with respect to t and using the fact that $q_{0,0}(t) \equiv 1$ (in the case $k = n$ in the last sum) we obtain

$$\frac{\partial}{\partial t}\mathbf{C}_{m,n}(t) = (\mathbf{K}_0 + \mathbf{K}_1)\mathbf{C}_{m,n}(t) + \sum_{j=0}^{m-1}\sum_{k=0}^n \binom{m}{j}\binom{n}{k} q_{1,n-k}(t)\mathbf{K}_1\mathbf{C}_{j,k}(t) + \sum_{k=0}^{n-1}\binom{n}{k} q_{0,n-k}(t)\mathbf{K}_1\mathbf{C}_{m,k}(t).$$

We can use this equation to express $\mathbf{C}_{m,n}(t)$ in terms of the $\mathbf{C}_{j,k}$ functions, for each (j,k) satisfying $0 \le j \le m, 0 \le k \le n$, and $j + k < m + n$, but when we solve this equation we will first assume throughout that the distribution functions L_1, L_2, F, and G are all continuous on $[0,\infty)$. Setting

$$\mathbf{F}(t) := \sum_{j=0}^{m-1}\sum_{k=0}^{n}\binom{m}{j}\binom{n}{k}q_{1,n-k}(t)\mathbf{K}_1\mathbf{C}_{j,k}(t) + \sum_{k=0}^{n-1}\binom{n}{k}q_{0,n-k}(t)\mathbf{K}_1\mathbf{C}_{m,k}(t)$$

we can see that

$$\frac{\partial}{\partial t}\mathbf{C}_{m,n}(t) = (\mathbf{K}_0 + \mathbf{K}_1)\mathbf{C}_{m,n}(t) + \mathbf{F}(t)$$

and this can be solved using standard methods. Indeed, for each $t > 0$,

$$\mathbf{e}^{-(\mathbf{K}_0+\mathbf{K}_1)t}\frac{\partial}{\partial t}\mathbf{C}_{m,n}(t) - \mathbf{e}^{-(\mathbf{K}_0+\mathbf{K}_1)t}(\mathbf{K}_0 + \mathbf{K}_1)\mathbf{C}_{m,n}(t) = \mathbf{e}^{-(\mathbf{K}_0+\mathbf{K}_1)t}\mathbf{F}(t).$$

Since $\mathbf{e}^{\mathbf{A}t}\mathbf{A} = \mathbf{A}\mathbf{e}^{\mathbf{A}t}$ we have by the product rule and the fundamental theorem of calculus, that for each $t > 0$,

$$\mathbf{e}^{-(\mathbf{K}_0+\mathbf{K}_1)t}\mathbf{C}_{m,n}(t) - \mathbf{I}\mathbb{1}_{\{m=n=0\}} = \int_0^t \mathbf{e}^{-(\mathbf{K}_0+\mathbf{K}_1)s}\mathbf{F}(s)ds$$

i.e.

$$\mathbf{C}_{m,n}(t) = \mathbf{e}^{(\mathbf{K}_0+\mathbf{K}_1)t}\mathbb{1}_{\{m=n=0\}} + \int_0^t \mathbf{e}^{(\mathbf{K}_0+\mathbf{K}_1)(t-s)}\mathbf{F}(s)ds.$$

Replacing $\mathbf{F}(s)$ with the expression it represents completes the derivation. Finally, to handle the case where L_1, L_2, F, and G are arbitrary, we can make use of the fact that each of these distribution functions can be interpreted as the weak limit of a sequence of phase-type distributions: see e.g. Chapter 3 of Kelly [15]. □

References

[1] Anderson, D.F., and Kurtz, T.G. (2015). *Stochastic Analysis of Biochemical Systems*. Springer International Publishing, Switzerland.

[2] Asmussen, S., and Koole, G. (1993). Marked point processes as limits of Markovian arrival streams. *Journal of Applied Probability* **30**, 365-372.

[3] Baccelli, F., Blaszczyszyn, B., and Karray, M. (2020). *Random Measures, Point Processes, and Stochastic Geometry*. INRIA, 2020, hal-02460214.

[4] Berg, O.G. (1978). A model for the statistical fluctuations of protein numbers in a microbial population. *Journal of Theoretical Biology* **71**, 587-603.

[5] Bokes, P., King, J.R., Wood, A.T.A., and Loose, M. (2012). Exact and approximate distribution of protein and mRNA levels in the low-copy regime of gene expression. *Journal of Mathematical Biology* **64**, 829-854.

[6] Brémaud, P. (1999). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer, New York.

[7] Bressloff, P.C. (2014). *Stochastic Processes in Cell Biology*. Springer International Publishing, Switzerland.

[8] Fromion, V., Leoncini, E., and Robert, P. (2013). Stochastic gene expression in cells: a point process approach. *SIAM Journal on Applied Mathematics* **73**, 195-211.

[9] He, Q. (2014). *Fundamentals of Matrix-Analytic Methods*. Springer Science+Business Media, New York.

[10] Herbach, U. (2019). Stochastic gene expression with a multistate promoter: breaking down exact distributions. *SIAM Journal on Applied Mathematics* **79**, 1007-1029.

[11] Horowitz, J.M., and Kulkarni, R.V. (2017). Stochastic gene expression conditioned on large deviations. *Physical Biology* **14**, 03LT01.

[12] Jansen, M., and Pfaffelhuber, P. (2015). Stochastic gene expression with delay. *Journal of Theoretical Biology* **364**, 355-363.

[13] Jia, C. (2020). Kinetic foundation of the zero-inflated negative binomial model for single-cell RNA sequencing data. *SIAM Journal on Applied Mathematics* **80**, 1336-1355.

[14] Kallenberg, O. (1983). *Random Measures*. Akademie-Verlag, Berlin.

[15] Kelly, F.P. (1979). *Reversibility and Stochastic Networks*. Wiley.

[16] Klenke, A. (2014). *Probability Theory: A Comprehensive Course*. Springer-Verlag, London.

[17] Latouche, G., and Ramaswami, V. (1999). *Matrix-Analytic Methods in Stochastic Modeling*. ASA-SIAM publications, Philadelphia, PA, USA.

[18] Masuyama, M., and Takine, T. (2002). Analysis of an infinite-server queue with batch Markovian arrival streams. *Queueing Systems* **42**, 269-296.

[19] Paulsson, J. (2005). Models of stochastic gene expression. *Physics of Life Reviews* **2**, 157-175.

[20] Peccoud, J. and Ycart, B. (1995). Markovian modeling of gene-product synthesis. *Theoretical Population Biology* **48**, 222-234.

[21] Pendar, H., Platini, T., and Kulkarni, R.V. (2013). Exact protein distributions for stochastic models of gene expression using partitioning of Poisson processes. *Physical Review E* **87**, 042720.

[22] Ramaswami, V., and Neuts, M. (1980). Some explicit formulas and computational methods for infinite-server queues with phase-type arrivals. *Journal of Applied Probability* **17**, 498-514.

[23] Rigney, D.R. (1979). Stochastic model of constructive protein levels in growing and dividing bacterial cells. *Journal of Theoretical Biology* **76**, 453-480.

[24] Rigney, D.R. (1979). Note on the kinetics and stochastics of induced protein synthesis as influenced by various models of messenger RNA degradation. *Journal of Theoretical Biology* **79**, 247-257.

[25] Rigney, D.R., and Schieve, W.C. (1977). Stochastic model of linear, continuous protein synthesis in bacterial populations. *Journal of Theoretical Biology* **69**, 761-766.

[26] Robert, P. (2019). Mathematical models of gene expression. *Probability Surveys* **16**, 277-332.

[27] Shahrezaei, V., and Swain, P.S. (2008). Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences* **105**, 17256-17261.