A GEOMETRIC ANALYSIS OF FAST-SLOW MODELS FOR STOCHASTIC GENE EXPRESSION

NIKOLA POPOVIĆ, CARSTEN MARR, AND PETER S. SWAIN

Abstract. Stochastic models for gene expression frequently exhibit dynamics on several different scales. One potential time-scale separation is caused by significant differences in the lifetimes of mRNA and protein; the ratio of the two degradation rates gives a natural small parameter in the resulting Chemical Master Equation, allowing for the application of perturbation techniques. Here, we develop a framework for the analysis of a family of 'fast-slow' models for gene expression that is based on geometric singular perturbation theory. We illustrate our approach by giving a complete characterisation of a standard two-stage model which assumes transcription, translation, and degradation to be first-order reactions. In particular, we present a systematic expansion procedure for the probability-generating function that can in principle be taken to any order in the perturbation parameter, allowing for an approximation of the corresponding propagator probabilities to that same order. For illustrative purposes, we perform this expansion explicitly to first order, both on the fast and the slow time-scales; then, we combine the resulting asymptotics into a composite fast-slow expansion that is uniformly valid in time. In the process, we extend, and prove rigorously, results previously obtained by Shahrezaei and Swain [50] and Bokes *et al.* [8, 9]. We verify our asymptotics by numerical simulation, and we explore its practical applicability and the effects of a variation in the system parameters and the time-scale separation. Focussing on biologically relevant parameter regimes that induce translational bursting, as well as those in which mRNA is frequently transcribed, we find that the first-order correction can significantly improve the steady-state probability distribution. Similarly, in the time-dependent scenario, inclusion of the first-order fast asymptotics results in a uniform approximation for the propagator probabilities that is superior to the slow dynamics alone. Finally, we discuss the generalisation of our geometric framework to models for regulated gene expression that involve additional stages.

1. Introduction

Gene expression in prokaryotic and eukaryotic organisms alike can be a highly stochastic process, which complicates the modelling of gene regulatory networks; see, *e.g.*, [49] and the references therein. Yet, stochasticity should be included if models are to describe accurately the dynamics of gene expression when the abundance of the involved species is low. While stochastic fluctuations can be either extrinsic or intrinsic, we will focus on intrinsic fluctuations here, *i.e.*, on those generated by the random timing of chemical reactions. (A more complete discussion of the relationship between the two types of fluctuation can be found in [10, 48].) In both scenarios, the aim is the derivation – or accurate approximation – of the probability distributions that describe the numbers of mRNA and protein which are synthesised over time. These distributions are obtained as solutions of the

Date: March 29, 2015.

2000 *Mathematics Subject Classification.* 34E05, 92C40, 34E15, 60J25, 34C45, 37N25.

Key words and phrases. Stochastic gene expression; Chemical Master Equation; Two-stage model; Generating function; Propagator probabilities; Asymptotic expansion; Geometric Singular Perturbation Theory.

Nikola Popović is grateful to Peter De Maesschalck and Ramon Grima for their careful reading of drafts of the manuscript and numerous helpful suggestions, as well as to Peter Szmolyan for stimulating discussions. Moreover, the authors acknowledge grant support from the Moray Endowment Fund, as well as from MAXIMATHS, an initiative by the School of Mathematics at the University of Edinburgh aimed at maximising the impact of mathematics in science and engineering. Finally, the authors thank three anonymous referees for valuable comments which greatly improved the original manuscript.

Chemical Master Equation (CME), which constitutes the accepted mathematical description of reaction processes in general and of gene expression in particular; see, *e.g.*, [21, 22, 50] and the references therein. However, as the CME can be solved exactly only in special cases [25, 28, 33], there is a definite need for approximate solution techniques.

In this article, we develop a systematic, perturbative framework for the approximation of probability distributions in stochastic gene expression under the assumption that the reaction dynamics occurs on two fundamentally different time-scales; specifically, it is assumed that the degradation of mRNA is much faster than that of protein. The resulting scale separation between mRNA and protein is well-documented in many (microbial) organisms that include bacteria [6, 60] and yeast [50, 57], where the scales typically differ by about an order of magnitude. (It is, however, by no means generic: thus, it has been found recently that the two scales are often comparable in mammalian cells [45].) The presence of a singular perturbation parameter – the inverse of the ratio of lifetimes of protein and mRNA – allows for the application of perturbative techniques; specifically, our analysis is based on geometric singular perturbation theory [17, 29].

The application of (singular) perturbation techniques in biological modelling has a long and distinguished history, starting with the seminal article by Segel and Slemrod [46], where such techniques were first popularised in the context of the quasi-steady-state approximation (QSSA). The classical method of matched asymptotic expansions [32] can provide rapid insight into the asymptotics of singularly perturbed differential equation models, and has been applied widely in mathematical biology; see, *e.g.*, [39] for examples and references. An alternative, more geometric approach, which was pioneered by Fenichel [17] and popularised by Jones [29], is based on the well-developed theory of dynamical systems [2, 47, 59], allowing for a visual interpretation and rigorous justification of the resulting asymptotic expansions in terms of invariant manifolds, and their foliations, in phase space. (Intuitively, the former correspond to the slow dynamics, while the latter describe the fast component of the flow.)

We illustrate our approach by characterising completely the multiple-scale ('fast-slow') dynamics of a two-stage model for stochastic gene expression which was, to the best of our knowledge, first proposed in [53]; see also [8, 9, 30, 50, 54] and the references therein. As that model has been studied extensively, we mention four relevant publications here; in particular, we emphasise the article by Shahrezaei and Swain [50], which provided the motivation for our study. However, while they proposed a perturbative approximation akin to ours, they merely derived the leading-order asymptotics of the resulting protein distribution given zero mRNA initially; moreover, they neglected any transient dynamics in their analysis. (Here, we give a mathematically rigorous justification of their results, and we extend them substantially in the process.) More recently, Bokes *et al.* [8] employed a combination of analytical, asymptotic, and numerical techniques to approximate the probability-generating function for the joint distribution of mRNA and protein in the above model in a number of asymptotic regimes; however, they only studied the system at steady state. In the follow-up article [9], the same authors then focussed on the asymptotic regime considered here; while they did formulate the 'inner' ('fast') equations, with the aim of matching them to the 'outer' ('slow') ones, they only did so to zeroth order in the perturbation parameter. Furthermore, they did not obtain closed-form expressions for the resulting propagator probabilities of observing certain numbers of mRNA and protein at a point in time, given some initial numbers thereof; rather, they integrated their leading-order equations numerically. Finally, in [42], the authors invoked certain partitioning properties of Poisson processes to map regulatory networks onto appropriately defined reduced models, which allowed them to obtain both time-dependent and stationary closed-form expressions for the generating function in the two-stage model considered here. While their approach seems to be equally applicable to generalised models for stochastic gene expression, like ours, no expressions – exact or approximate – are given for the resulting probability distributions.

Thus, our results represent a three-fold advance over previous studies of the standard two-stage model: first, we develop a systematic approximation procedure for the corresponding propagator probabilities that can in principle be performed algorithmically, and to any order in the perturbation parameter; moreover, our approach yields asymptotic formulae in closed form for these propagators, unlike in [8, 9, 42]. Second, the resulting formulae systematically account for contributions both from the fast ('transient') and the slow ('long-term') dynamics, in contrast to [50]; in particular, the former are vital for the accurate approximation of propagator probabilities in a variety of biologically relevant parameter regimes, as discussed in detail in Section 5. (For demonstrative purposes, we restrict ourselves to deriving explicitly the first-order asymptotics here.) Third, and again in contrast to [8, 50], our approach allows for arbitrary initial numbers of both mRNA and protein. The reader is referred to Section 5, where the practical implications of these advances are explored numerically and where, moreover, their biological significance is assessed and interpreted.

This article is organised as follows: in Section 2, we introduce the two-stage model for gene expression studied here. In Section 3, which is aimed at a mathematically inclined readership, we construct our geometric framework for the perturbative approximation of an appropriately defined generating function. In Section 4, we derive first-order expansions for the corresponding probability distributions on the fast and the slow time-scales, which we then combine into a 'composite' fast-slow expansion that is uniformly valid in time. In Section 5, we verify our results numerically, and we interpret them from a practitioners' point of view; crucially, we show that our uniform propagator significantly outperforms the slow asymptotics alone, both to zeroth and to first order. Then, in Section 6, we summarise and discuss our findings, and we present potential topics for future research. Additional material, and mathematical detail, has been relegated to an Online Supplement: in Section A, we give a brief overview of geometric singular perturbation theory; in Section B, we quote asymptotic formulae for the propagator probabilities under the assumption that mRNA and protein numbers are not necessarily zero initially; Section C contains the mathematical proofs which underlie the asymptotics developed in the main text, while Section D lists the resulting formulae for the marginal probability distribution of protein in tabular form, for the reader's convenience.

2. Two-stage model for gene expression

In this section, we briefly introduce the standard two-stage model for unregulated gene expression; see, *e.g.*, [8, 9, 50, 53] for details. Then, we outline how the corresponding CME can be analysed via the method of characteristics [61].

2.1. **Background.** A cartoon illustration of the reaction kinetics underlying the two-stage model considered in the present article – which is a widely accepted representation of constitutive, or unregulated, gene expression – can be found in Figure 1(a), while the corresponding reaction scheme is sketched in Figure 1(b): under the assumption that the promoter region of the modelled gene is always active, one requires only two stochastic variables, namely, the numbers of mRNA and protein; that assumption, though simplistic, frequently seems to be reasonable in practice [7, 62]. (In fact, the two-stage model is capable of complex dynamics such as translational bursting, *i.e.*, of bursts in protein synthesis that seem to be typical of gene expression in bacteria and yeast; see [9, 30, 37, 50], as well as Section 5.2 below.) An additional simplification is achieved by the assumption that both transcription and translation, as well as the degradation of mRNA and protein, can be modelled as first-order chemical reactions; we denote the corresponding transcription and translation rates by ν_0 and ν_1, respectively, and we write d_0 and d_1 for the respective degradation rates of mRNA and protein, as illustrated in Figure 1(a).

2.2. **Chemical Master Equation (CME).** As in [18, 50], we introduce the new dimensionless parameters $a = \frac{\nu_0}{d_1}$, $b = \frac{\nu_1}{d_0}$, and $\gamma = \frac{d_0}{d_1}$; moreover, we rescale time with d_1 to obtain a non-dimensionalised time variable τ. Then, the propagator $P_{mn|m_0n_0}$ – *i.e.*, the probability of observing

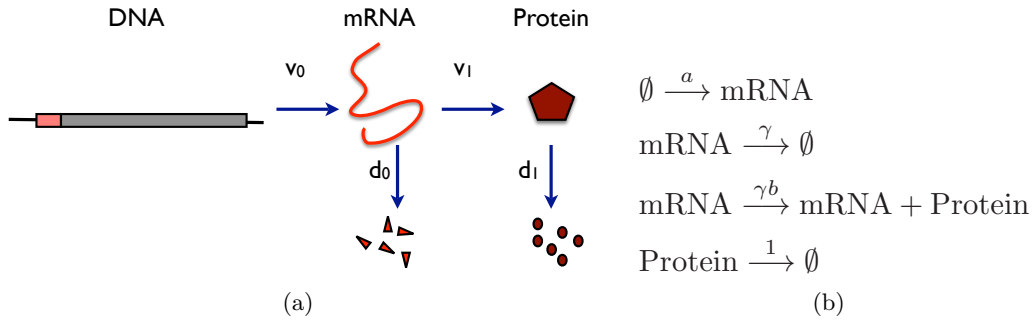


FIGURE 1. The two-stage model for unregulated gene expression. Panel 1(a) illustrates the reaction kinetics in which transcription, translation, and degradation are modelled as first-order processes. Correspondingly, ν_0 and ν_1 denote the probabilities per unit time (or rates) of transcription and translation, respectively, while the degradation rates of mRNA and protein are given by d_0 and d_1, respectively. In panel 1(b), the underlying reaction scheme is sketched; the rates ν_j and d_j $(j = 1, 2)$ have been replaced with three parameters $a = \frac{\nu_0}{d_1}$, $b = \frac{\nu_1}{d_0}$, and $\gamma = \frac{d_0}{d_1}$ after non-dimensionalisation.

m mRNA and n protein molecules at time τ, given m_0 and n_0 of each, respectively, at time zero – satisfies the non-dimensional CME

$$\begin{aligned}(1)\quad \frac{d}{d\tau}P_{mn|m_0n_0} &= a\big(P_{m-1,n|m_0n_0} - P_{mn|m_0n_0}\big) + \gamma bm\big(P_{m,n-1|m_0n_0} - P_{mn|m_0n_0}\big)\\ &\quad + \gamma\big[(m+1)P_{m+1,n|m_0n_0} - mP_{mn|m_0n_0}\big] + \big[(n+1)P_{m,n+1|m_0n_0} - nP_{mn|m_0n_0}\big],\end{aligned}$$

with m, n, m_0, $n_0 \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$; cf. also [50, Equation (1)] and [9, Equation (4)]. For convenience of notation, we will henceforth write $P_{mn} \equiv P_{mn|00}$ when $m_0 = 0 = n_0$.

Given our assumption that $\gamma \gg 1$ in Equation (5), *i.e.*, that the degradation rate of mRNA is much larger than that of protein, it is natural to introduce $\varepsilon = \gamma^{-1}$ as a (small) perturbation parameter. Correspondingly, we may interpret $P_{mn|m_0n_0}(\tau, \varepsilon)$ as a function of both τ and ε.

2.3. **Probability-generating function.** Our analysis relies crucially on the probability-generating function that is induced by the propagator probabilities $P_{mn|m_0n_0}$; see [38] for a recent exposition. In the context of the CME, Equation (1), that function is defined as

$$(2)\qquad F(z', z, \tau, \varepsilon) = \sum_{m,n=0}^{\infty} P_{mn|m_0n_0}(\tau, \varepsilon)(z')^m z^n,$$

where $z', z \in \mathbb{C}$ [38, Section 10.4]. We remark that the domain of definition of F must contain any pairs $(z', z) \in \mathbb{C}^2$ for which $|z'|, |z| \leq 1$, as well as that the series expansion in (2) is uniformly and absolutely convergent on that domain. For future reference, we note that the probabilities $P_{mn|m_0n_0}$ can be retrieved from F via

$$(3)\qquad P_{mn|m_0n_0}(\tau, \varepsilon) = \frac{1}{m!}\frac{1}{n!}\frac{\partial^m}{\partial(z')^m}\frac{\partial^n}{\partial z^n}F(z', z, \tau, \varepsilon)\Big|_{(z',z)=(0,0)},$$

as well as that F satisfies the normalisation condition

$$(4)\qquad F(1, 1, \tau, \varepsilon) = \lim_{(z',z)\to(1^-,1^-)} F(z', z, \tau, \varepsilon) = \sum_{m,n=0}^{\infty} P_{mn|m_0n_0}(\tau, \varepsilon) = 1.$$

Finally, introducing the new variables $u = z' - 1$ and $v = z - 1$, we find that F solves the first-order linear partial differential equation

$$\frac{\partial F}{\partial \tau} + \gamma[u - b(1+u)v]\frac{\partial F}{\partial u} + v\frac{\partial F}{\partial v} = auF; \tag{5}$$

see also [50, Equation (2)]. In particular, since $P_{mn|m_0n_0}(0,\varepsilon) = \delta_{mm_0}\delta_{nn_0}$ (irrespective of the value of ε), the function F must satisfy the condition

$$F(u,v,0,\varepsilon) = (1+u)^{m_0}(1+v)^{n_0} \tag{6}$$

for $\tau = 0$. (Here, δ_{jk} denotes the Kronecker delta, with $\delta_{jj} = 1$ and $\delta_{jk} = 0$ for any $j \neq k$.)

***Remark* 1.** Throughout this article, we will interchangeably consider F to be a function either of (z', z) or of (u, v), as required. While our use of notation is thus not entirely precise, its meaning should be clear in context. □

2.4. **'Characteristic' equations.** As is well-known [8, 50], the partial differential Equation (5) can be solved via the method of characteristics; see, *e.g.*, [61, Chapter 2]. Denoting by r the distance along a characteristic curve whose initial point is located at $(u_0, v_0) \in \mathbb{R}^2$ for $\tau = 0$, we obtain the following 'characteristic' system of ordinary differential equations from (5):

$$\frac{d\tau}{dr} = 1, \tag{7a}$$

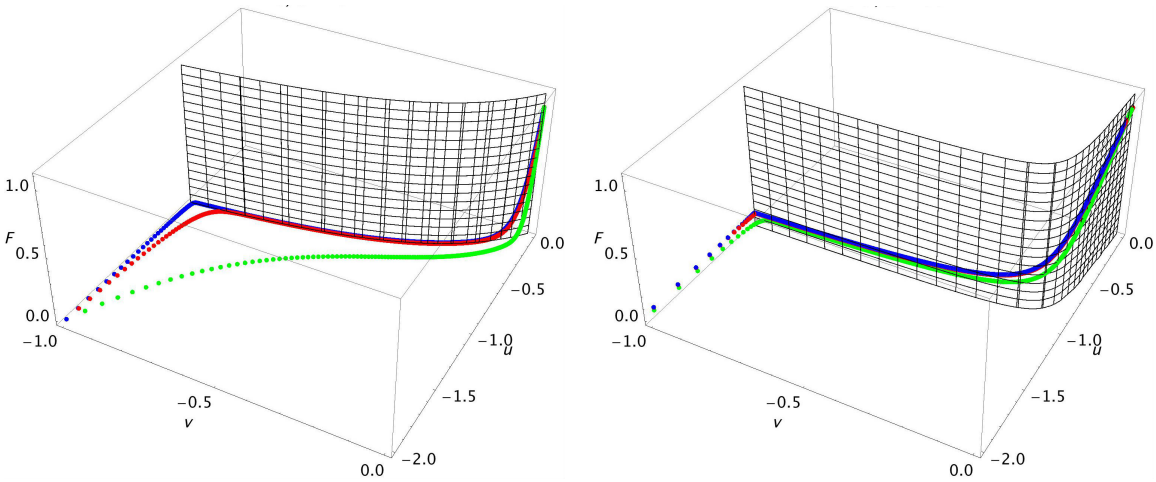
$$\frac{du}{dr} = \frac{1}{\varepsilon}[u - b(1+u)v], \tag{7b}$$

$$\frac{dv}{dr} = v, \tag{7c}$$

$$\frac{dF}{dr} = auF. \tag{7d}$$

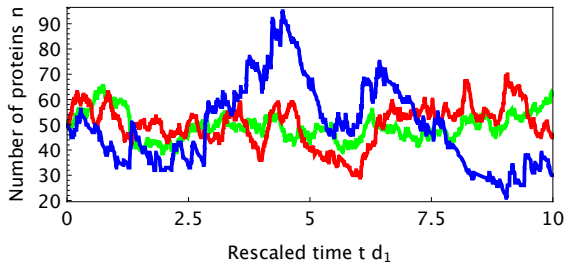
Equation (7b) becomes a linear non-autonomous equation for u after substitution of the exact solution for $v(v_0, r) = v_0\mathrm{e}^r$ from (7c), and can hence be solved by introduction of an integrating factor. However, the resulting integral cannot be evaluated in closed form; cf. also [50, Supporting Information, Equation (26)]. The existence of an integral-form solution to (7) necessarily follows from the fact that the scheme in Figure 1(b) contains only first-order reactions; see, *e.g.*, [24, 25] and the references therein for details. While the CME has been solved exactly in [28] for certain such ('unimolecular') schemes, we emphasise that their results do not apply in the case of catalytic production which underlies the two-stage model studied in this article; cf. [28, Section 6].

In the following, we will exploit the presence of the perturbation parameter ε in Equation (7) to apply the perturbative technique known as geometric singular perturbation theory [17, 29]; a brief overview of the latter can be found in Section A of the Online Supplement. (A related approach for the approximation of the probability-generating function F is developed in [9, Section 4].) In fact, the resulting scale separation can be made evident by solving Equation (7) numerically, and is confirmed by simulation, via Gillespie's stochastic simulation algorithm (SSA) [20], of the underlying CME, Equation (1). The geometry of the former is illustrated in Figures 2(a) and 2(b), for varying values of ε and two regimes for the non-dimensional parameters a and b; one observes convergence of u to an invariant manifold in (backward) fast time t, with (v, F) nearly constant for ε sufficiently small, at which stage the slow flow on that manifold takes over, with (u, v, F) tending to some steady state thereon. The time evolution of protein numbers in the two regimes is shown in Figures 2(c) and 2(d); throughout, one finds that the fast-slow structure of the model becomes more pronounced with decreasing ε, as is to be expected.

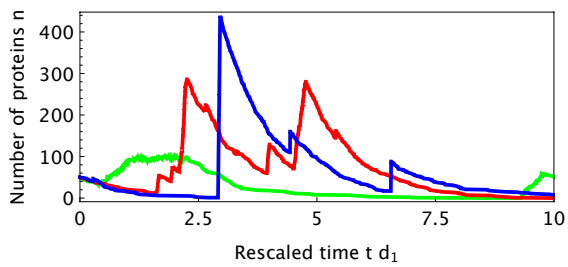


(a) $a = 20$, $b = 2.5$.

(b) $a = 0.5$, $b = 100$.



(c) $a = 20$, $b = 2.5$.



(d) $a = 0.5$, $b = 100$.

FIGURE 2. Fast-slow dynamics of the two-stage model for gene expression in the two parameter regimes considered here, with $\varepsilon = 1$ (dotted green), $\varepsilon = 0.1$ (dotted red), and $\varepsilon = 0.01$ (dotted blue). Panels 2(a) and 2(b) illustrate the geometry of Equation (8); the spacing of the dots indicates the time-scale of the flow, with larger intervals corresponding to faster motion along trajectories, while gridded surfaces represent 'critical manifolds' for Equation (8). Typical time series of protein – as obtained from stochastic simulation – are displayed in panels 2(c) and 2(d).

3. Geometric framework

In this section, we construct our geometric framework for the analysis of the characteristic system, Equation (7). In particular, we derive rigorous asymptotics for the characteristic coordinates u and v, as well as for the probability-generating function F, in the perturbation parameter ε .

3.1. **Preliminaries.** First, we verify that Equation (7) represents a singularly perturbed fast-slow system in standard form which satisfies the requirements of Fenichel's geometric singular perturbation theory [17, 27, 29]; cf. Section A of the Online Supplement for details. Since $r(0) = 0$, it follows from (7a) that $r \equiv \tau$, where τ equals the non-dimensionalised time introduced in Section 2.2; hence, we may rewrite Equation (7) with τ as the independent variable, thus obtaining the three-dimensional 'slow system'

$$\varepsilon \dot{u} = u - b(1+u)v, \tag{8a}$$

$$\dot{v} = v, \tag{8b}$$

$$\dot{F} = auF. \tag{8c}$$

Here, the overdot denotes differentiation with respect to τ. Recalling the definition of the small perturbation parameter $\varepsilon(=\gamma^{-1})$, we conclude that u is a fast variable, whereas v and F are slow variables; hence, in the notation of Section A, we have $k=1$ and $l=2$ as well as $\mathbf{f}=f(u,(v,F),\varepsilon)\equiv f(u,v)=u-b(1+u)v$ and $\mathbf{g}=\mathbf{g}(u,(v,F),\varepsilon)\equiv\mathbf{g}(u,(v,F))=(v,auF)$. (In particular, we remark that neither $\mathbf{f}$ nor $\mathbf{g}$ depend on the parameter ε in our case.) Correspondingly, introducing a new fast time $t=\frac{\tau}{\varepsilon}$ in Equation (8), we find the 'fast system'

$$u' = u - b(1+u)v, \tag{9a}$$

$$v' = \varepsilon v, \tag{9b}$$

$$F' = \varepsilon auF. \tag{9c}$$

(For future reference, we note that the (u,v)-subsystem decouples both in (8) and in (9).)

Solving the relation $f(u,v)=0=u-b(1+u)v$ for u, we obtain

$$u = U_0(v) = \frac{bv}{1-bv} \tag{10}$$

for the (two-dimensional) 'critical manifold' $\mathcal{S}_0$; see also [50, Equation (5)]. While $\mathcal{S}_0$ is *a priori* unbounded in (u,v), the definition of the probability-generating function F in (2) assumes $(z',z)\in[-1,1]^2$ and, correspondingly, $(u,v)\in[-2,0]^2$; thus, we may restrict ourselves to a closed and bounded (compact) subset $\mathcal{D}\subset[-2,0]^2$ of $\mathbb{R}^2$. (We remark that $\mathcal{D}$ certainly has to contain the point $(u,v)=(-1,-1)$; recall Equation (3).) It then follows that the representation in Equation (10) is well-defined in the v-regime that is relevant to us, as the singularity at $v=b^{-1}$ of the function U_0 is excluded by our definition of $\mathcal{D}$.

Next, rewriting the representation for $\mathcal{S}_0$ in (10) in terms of v, we find

$$\frac{\partial f}{\partial u}\Big|_{\mathcal{S}_0} = 1-bv\Big|_{v=\frac{u}{b(1+u)}} = \frac{1}{1+u} \neq 0$$

for the linearisation of f about the critical manifold $\mathcal{S}_0$; here, we note that $u\neq-1$ on $\mathcal{S}_0$, by Equation (10). Hence, the manifold $\mathcal{S}_0$ is normally hyperbolic; specifically, since $v\in[-2,0]$ implies $u\in\big[-\frac{2b}{1+2b},0\big]\subset[-1,0]$, and since $\frac{\partial f}{\partial u}\big|_{\mathcal{S}_0}>0$ for $u>-1$, $\mathcal{S}_0$ is normally repelling in the (u,v)-regime considered here. In particular, it follows that $k_{\mathrm{s}}=0$ and $k_{\mathrm{u}}=1$, in the notation of Section A of the Online Supplement, which, together with $l=2$, implies the existence of a three-dimensional unstable manifold $\mathcal{W}^{\mathrm{u}}(\mathcal{S}_0)$ for $\mathcal{S}_0$ in that regime.

The above observations, in combination with Theorems A.1 and A.2 of the Online Supplement, yield the following result on the persistence of the critical manifold $\mathcal{S}_0$, as well as of its associated unstable manifold $\mathcal{W}^{\mathrm{u}}(\mathcal{S}_0)$, under the flow that is induced by (8):

Proposition 3.1. *Let $\varepsilon\in[0,\varepsilon_0]$, with $\varepsilon_0>0$ sufficiently small, let $(u,v)\in\mathcal{D}\subset[-2,0]^2$, as above, and let $K\in\mathbb{N}$ be arbitrary, but fixed. Then, the following statements hold:*

(i) *The manifold $\mathcal{S}_0$ perturbs to a locally invariant, $\mathcal{C}^K$-smooth slow manifold $\mathcal{S}_\varepsilon$ that is $\mathcal{O}(\varepsilon)$-close to $\mathcal{S}_0$.*

(ii) *The manifold $\mathcal{S}_\varepsilon$ can be written as the graph of a ($\mathcal{C}^K$-smooth) function $U(v,\varepsilon)$, with*

$$u = U(v,\varepsilon) = \sum_{k=0}^{K} U_k(v)\varepsilon^k + \mathcal{O}(\varepsilon^{K+1}); \tag{11}$$

here, $U_k(v)=\frac{\eta_k(bv)}{(1-bv)^{2k+1}}$ for $k=1,\dots,K$, with $\eta_k(bv)=\sum_{j=1}^{k}\eta_{kj}(bv)^j$ a polynomial of degree k in bv. In particular, $U_0(v)=\frac{bv}{1-bv}$, cf. Equation (10), while

$$U_1(v) = \frac{bv}{(1-bv)^3}. \tag{12}$$

(iii) *The flow on $\mathcal{S}_\varepsilon$ is given by*

$$\dot{v} = v, \tag{13a}$$

$$\dot{F} = aU(v,\varepsilon)F, \tag{13b}$$

to any order in ε.

(iv) *The line $\ell : \{(u,v,F) = (0,0,F_0)\}$, with F_0 non-negative and constant, is contained in $\mathcal{S}_\varepsilon$ for any $\varepsilon \in [0,\varepsilon_0]$; in particular, any point on ℓ is an equilibrium state of* (8).

(v) *The manifold $\mathcal{W}^{\mathrm{u}}(\mathcal{S}_0)$ perturbs to a locally invariant, $\mathcal{C}^K$-smooth unstable manifold $\mathcal{W}^{\mathrm{u}}(\mathcal{S}_\varepsilon)$ for $\mathcal{S}_\varepsilon$ that is $\mathcal{O}(\varepsilon)$-close to $\mathcal{W}^{\mathrm{u}}(\mathcal{S}_0)$.*

Proposition 3.1(i) implies, in particular, that $\mathcal{S}_\varepsilon$ is a regular perturbation of $\mathcal{S}_0$; correspondingly, the series expansion for $U(v,\varepsilon)$ in Equation (11) can be taken to arbitrary (fixed) order in ε, with coefficient functions U_j that are, in fact, $\mathcal{C}^\infty$-smooth for $v \in [-2,0]$, by (12). Moreover, given (iii), the 'reduced flow' on $\mathcal{S}_0$ then perturbs to the corresponding flow on $\mathcal{S}_\varepsilon$ in a regular fashion. Finally, by Proposition 3.1(iv), the generating function F can be made to satisfy the normalisation condition, Equation (4), to any order in ε, as $(0,0,1) \in \ell$.

***Remark* 2.** The slow manifold $\mathcal{S}_\varepsilon$ – or, rather, its representation in (11) – is unique only up to exponentially small terms in ε, *i.e.*, terms of the form $\mathcal{O}(\mathrm{e}^{-\frac{\kappa}{\varepsilon}})$, where κ is some positive constant. Mathematically, this non-uniqueness is caused by the 'cut-off' that needs to be applied to $\mathcal{S}_0$ outside of $\mathcal{D} \subset \mathbb{R}^2$ to enforce its boundedness; see [17, 29] for details. □

In the remainder of this section, we will approximate the slow flow on the manifold $\mathcal{S}_\varepsilon$, as well as the fast dynamics on its unstable manifold $\mathcal{W}^{\mathrm{u}}(\mathcal{S}_\varepsilon)$, to first order in ε. By Proposition 3.1, both components of the flow can be obtained by regular perturbation off the respective singular dynamics that is obtained in the limit as $\varepsilon \to 0$; hence, we may assume asymptotic series expansions for u and F, which will need to satisfy the fast and the slow Equations (8) and (9), respectively, order-by-order in ε. (An alternative procedure would involve a transformation of the governing equations into normal form (Fenichel) coordinates [29]; however, as that approach seems more involved computationally, we do not pursue it here.)

3.2. **Fast dynamics.** In this subsection, we derive the asymptotics of u and F on the fast t-scale, *i.e.*, under the flow that is induced by Equation (9).

Our analysis is simplified by the observation that Equation (9b) can be solved explicitly, with $v(v_0,t,\varepsilon) = v_0\mathrm{e}^{\varepsilon t}$, which immediately yields an expansion for v to any order in ε; in particular, $v(v_0,t,\varepsilon) = v_0 + \varepsilon v_0 t + \mathcal{O}(\varepsilon^2)$. From Proposition 3.1(v), in combination with standard results on the smooth dependence of solutions of ordinary differential equations on their initial conditions and parameters [1], it then follows that u and F admit asymptotic expansions of the form

$$u(u_0,v_0,t,\varepsilon) = \sum_{k=0}^{K} \mathcal{U}_k(u_0,v_0,t)\varepsilon^k + \Delta\mathcal{U}(u_0,v_0,t,\varepsilon), \quad \text{with } \Delta\mathcal{U} = \mathcal{O}(\varepsilon^{K+1}), \quad \text{and} \tag{14a}$$

$$F(u_0,v_0,t,\varepsilon) = \sum_{k=0}^{K} \mathcal{F}_k(u_0,v_0,t)\varepsilon^k + \Delta\mathcal{F}(u_0,v_0,t,\varepsilon), \quad \text{with } \Delta\mathcal{F} = \mathcal{O}(\varepsilon^{K+1}), \tag{14b}$$

for any $K \in \mathbb{N}$. (Here, we note that any dependence of F on its initial value F_0 may be eliminated in (14b), as $F(u_0,v_0,0,\varepsilon) = (1+u_0)^{m_0}(1+v_0)^{n_0}$ is a function of u_0 and v_0 only, by Equation (6).)

For illustrative purposes, we restrict ourselves to the case where $K = 1$ in (14): substituting into Equations (9a) and (9c) and collecting like powers of ε in the resulting equations, we find that the

coefficient functions $\mathcal{U}_j$ and $\mathcal{F}_j$ $(j = 0, 1)$ solve the recursive (linear) system

$$\mathcal{U}_0' = \mathcal{U}_0 - b(1+\mathcal{U}_0)v_0, \quad \text{with } \mathcal{U}_0(0) = u_0; \tag{15a}$$

$$\mathcal{F}_0' = 0, \quad \text{with } \mathcal{F}_0(0) = (1+u_0)^{m_0}(1+v_0)^{n_0}; \tag{15b}$$

$$\mathcal{U}_1' = \mathcal{U}_1 - b\mathcal{U}_1 v_0 - b(1+\mathcal{U}_0)v_0 t, \quad \text{with } \mathcal{U}_1(0) = 0; \tag{15c}$$

$$\mathcal{F}_1' = a\,\mathcal{U}_0\mathcal{F}_0, \quad \text{with } \mathcal{F}_1(0) = 0. \tag{15d}$$

Here, the prime denotes differentiation with respect to t, as before, and we have suppressed any dependence on the latter, as well as on the initial values u_0 and v_0, for convenience of notation.

Solving Equations (15a) and (15b), we obtain

$$\mathcal{U}_0(u_0, v_0, t) = \frac{bv_0}{1-bv_0} + \Big(u_0 - \frac{bv_0}{1-bv_0}\Big)\mathrm{e}^{(1-bv_0)t} \tag{16}$$

and

$$\mathcal{F}_0(u_0, v_0, t) = (1+u_0)^{m_0}(1+v_0)^{n_0} \tag{17}$$

for u and F, respectively, to lowest order in ε. (We remark that the $(\mathcal{U}_0, \mathcal{F}_0)$-subsystem in (15) is equivalent to the so-called 'layer problem' that is found in the limit as $\varepsilon \to 0^+$ in Equation (9); cf. Section A of the Online Supplement.)

Finally, we derive the first-order correction in ε on this fast t-scale: substituting the above expressions for $\mathcal{U}_0$ and $\mathcal{F}_0$ into (15c) and (15d) and solving, we have

$$\mathcal{U}_1(u_0, v_0, t) = \frac{bv_0}{(1-bv_0)^3}\big[1 - \mathrm{e}^{(1-bv_0)t}\big] + \frac{bv_0}{(1-bv_0)^2}t - \frac{bv_0}{2}\Big(u_0 - \frac{bv_0}{1-bv_0}\Big)t^2\mathrm{e}^{(1-bv_0)t} \tag{18}$$

and

$$\mathcal{F}_1(u_0, v_0, t) = \frac{a}{1-bv_0}\left\{bv_0 t - \Big(u_0 - \frac{bv_0}{1-bv_0}\Big)\big[1 - \mathrm{e}^{(1-bv_0)t}\big]\right\} \tag{19}$$

for the coefficients $\mathcal{U}_1$ and $\mathcal{F}_1$ in (14).

3.3. **Slow dynamics.** Next, we consider the asymptotics of u and F on the slow τ-scale, to first order in ε.

We observe that, by Proposition 3.1(ii), an expansion for u is provided by the representation for the slow manifold $\mathcal{S}_\varepsilon$ postulated in Equation (11); it then follows immediately that $u(v, \varepsilon) = U_0(v) + \varepsilon U_1(v) + \mathcal{O}(\varepsilon^2)$, where U_0 and U_1 are defined as in Equations (10) and (12), respectively.

Similarly, Proposition 3.1(iv) implies that F admits a regular asymptotic expansion on $\mathcal{S}_\varepsilon$, to any order in ε. Hence, dividing (13b) formally by (13a) to rewrite Equation (13) with v as the independent variable, making the Ansatz $F(v, \varepsilon) = \sum_{k=0}^{K} F_k(v)\varepsilon^k + \Delta F(v, \varepsilon)$, with $\Delta F = \mathcal{O}(\varepsilon^{K+1})$, substituting in the corresponding expressions for U_0 and U_1, and collecting like powers of ε in the resulting equation, we obtain the system

$$\frac{dF_0}{dv} = \frac{ab}{1-bv}F_0, \tag{20a}$$

$$\frac{dF_1}{dv} = \frac{ab}{1-bv}\left[F_1 + \frac{1}{(1-bv)^2}F_0\right] \tag{20b}$$

for the coefficient functions F_j $(j = 0, 1)$. We note that any free constants arising in the solution of the above equations will be fixed by the requirement that the slow (outer) asymptotics of F agrees with the fast (inner) asymptotics determined in the previous subsection on some 'domain of overlap,' as is also the case in asymptotic matching [32]. (Geometrically speaking, one hence needs to describe the flow on the unstable manifold $\mathcal{W}^{\mathrm{u}}(\mathcal{S}_\varepsilon)$ in a neighbourhood of the slow manifold $\mathcal{S}_\varepsilon$ to the appropriate order in ε.)

Solving Equation (20a), we find that the generating function F defined in (2) satisfies

$$F_0(v) = \frac{C_0}{(1-bv)^a} \tag{21}$$

on the slow τ-scale, to lowest order in ε. Here, C_0 denotes a constant that is as yet undetermined. (We remark that the expression in (21) agrees with the solution of the corresponding so-called 'reduced problem' which is obtained in the limit as $\varepsilon \to 0^+$ in Equation (8), as does the leading-order approximation U_0 for u; recall (10).)

It remains to determine the coefficient function F_1: substituting the (known) expressions for U_0, U_1, and F_0 into Equation (20b) and solving, we obtain

$$F_1(v) = \frac{a}{2}\frac{C_0}{(1-bv)^{a+2}} + \frac{C_1}{(1-bv)^a}, \tag{22}$$

where C_1 is again a free constant.

4. Asymptotic analysis

In this section, we derive the first-order asymptotics (in ε) of the 'inverse characteristic transformation' – which expresses the initial points (u_0, v_0) of characteristic curves for the partial differential Equation (5) in terms of the coordinates u and v – as well as of the generating function F. Then, we deduce corresponding expansions for the propagator probabilities $P_{mn|m_0n_0}$ of observing m mRNAs and n proteins at some point in time, given m_0 and n_0 of each initially, respectively. Finally, we approximate the resulting marginal probability distributions of protein, to first order in ε.

To that end, we will first study the fast and the slow dynamics separately, as is also frequently done in matched asymptotics [32]. Second, we will 'match' the resulting 'inner' and 'outer' expansions by requiring that their respective first-order truncations agree on some domain of overlap between the two scales. Third, we will construct composite expansions which are hence uniformly valid in time, both on the fast and the slow time-scales, up to an $\mathcal{O}(\varepsilon^2)$-error.

While we only consider explicitly the first-order asymptotics of $P_{mn|m_0n_0}$, we emphasise that the perturbative procedure outlined here can be extended, in a systematic fashion, to arbitrary order in ε. For the sake of exposition, we will again restrict ourselves to the case where $m_0 = 0 = n_0$ in $P_{mn|m_0n_0}$; the general case of m_0, $n_0 \in \mathbb{N}$, which is significantly more involved algebraically, is discussed in Section B of the Online Supplement.

4.1. **Inverse characteristic transformation.**

In this subsection, we derive asymptotic expansions for the transformation $(u, v) \mapsto (u_0, v_0)$, to first order in ε.

4.1.1. *Inner ('fast') asymptotics.* Combining Equations (16) and (18), noting that $v_0 = v(1-\varepsilon t) + \mathcal{O}(\varepsilon^2)$, expanding the resulting expression in ε, and solving for u_0, we find

$$\begin{aligned}(u_0, v_0)(u, v, t, \varepsilon) = \Bigg(&\frac{bv}{1-bv} + \varepsilon\frac{bv}{(1-bv)^3} + \left[u - \frac{bv}{1-bv} - \varepsilon\frac{bv}{(1-bv)^3}\right]\mathrm{e}^{-(1-bv)t} \\ &- \varepsilon\frac{bv}{(1-bv)^2}t - \varepsilon\frac{1}{2}bv\Big(u - \frac{bv}{1-bv}\Big)t^2\mathrm{e}^{-(1-bv)t} + \mathcal{O}(\varepsilon^2), v - \varepsilon vt + \mathcal{O}(\varepsilon^2)\Bigg)\end{aligned} \tag{23}$$

on the fast t-scale. (By the Implicit Function Theorem, Equation (23) defines a $\mathcal{C}^K$-diffeomorphism for any $K \in \mathbb{N}$ and ε positive, but sufficiently small – and, in fact, even a $\mathcal{C}^\infty$-diffeomorphism.)

4.1.2. *Outer ('slow') asymptotics.* Evaluation of the first-order truncation of the asymptotics of the invariant slow manifold $\mathcal{S}_\varepsilon$, as given in Equation (11), at $v = v_0$, in combination with the fact that $v_0(v,\tau) = v\mathrm{e}^{-\tau}$, directly implies that the transformation $(u,v) \mapsto (u_0,v_0)$ reads

$$(u_0,v_0)(u,v,\tau,\varepsilon) = \left(\frac{bv\mathrm{e}^{-\tau}}{1-bv\mathrm{e}^{-\tau}} + \varepsilon\frac{bv\mathrm{e}^{-\tau}}{(1-bv\mathrm{e}^{-\tau})^3} + \mathcal{O}(\varepsilon^2), v\mathrm{e}^{-\tau}\right) \tag{24}$$

on the slow τ-scale.

4.1.3. *Uniform ('fast-slow') asymptotics.* Finally, we combine the inner and outer asymptotics of u_0, as given in Equations (23) and (24), respectively, into a composite asymptotic series that is uniformly valid both on the fast and the slow time-scales.

We note that the terms which are common to the two expansions are obtained as $\big(\frac{bv}{1-bv} + \varepsilon\frac{bv}{(1-bv)^3} - \varepsilon\frac{bv}{(1-bv)^2}t, v - v\varepsilon t\big)$ up to an $\mathcal{O}(\varepsilon^2)$-error, as is best seen by expanding Equation (24) in terms of t. Thus, we find

$$\begin{aligned}(u_0,v_0)(u,v,\tau,t,\varepsilon) = \left(\frac{bv\mathrm{e}^{-\tau}}{1-bv\mathrm{e}^{-\tau}} + \varepsilon\frac{bv\mathrm{e}^{-\tau}}{(1-bv\mathrm{e}^{-\tau})^3} + \left[u - \frac{bv}{1-bv} - \varepsilon\frac{bv}{(1-bv)^3}\right]\mathrm{e}^{-(1-bv)t}\right.\\ \left. - \varepsilon\frac{1}{2}bv\Big(u - \frac{bv}{1-bv}\Big)t^2\mathrm{e}^{-(1-bv)t} + \mathcal{O}(\varepsilon^2), v\mathrm{e}^{-\tau}\right)\end{aligned} \tag{25}$$

for the inverse characteristic transformation $(u_0,v_0)(u,v,\tau,t,\varepsilon)$. We emphasise that the limit as $t \to \infty$ – or, equivalently, as $\tau \to \infty$ – in Equation (25) is well-defined, since $1-bv \in [1,1+2b]$ when $v \in [-2,0]$: while the manifold $\mathcal{S}_\varepsilon$ is repelling for $(u,v) \in \mathcal{D}$, recall Section 3.1, (u_0,v_0) is obtained by solving the characteristic system in (7) in 'backward time,' given some fixed choice of (u,v).

Equation (25) implies, in particular, that $u_0(u,v)$ can be written as the sum of a τ-dependent contribution, which lies on the invariant slow manifold $\mathcal{S}_\varepsilon$ and which is represented by the first two terms therein, and of a t-dependent remainder that arises through the fast flow towards $\mathcal{S}_\varepsilon$. In other words, for any coordinate pair $(u,v) \in \mathcal{D}$, the initial point (u_0,v_0) of the corresponding characteristic curve converges to the point $(U(v_0,\varepsilon),v_0)$ on the slow manifold $\mathcal{S}_\varepsilon$ as $t \to \infty$, with τ fixed, *i.e.*, after the initial (fast) transient has decayed. (Consequently, Equation (25) reduces to (24) in that limit; similarly, substituting $\tau = \varepsilon t$ into (25) and expanding in ε, one finds agreement with Equation (23) for $\varepsilon \to 0$, with t fixed.)

***Remark* 3.** Since $1-bv \in [1,1+2b]$ for $v \in [-2,0]$, it follows that $\mathrm{e}^{-(1-bv)t} = \mathrm{e}^{-(1-bv)\frac{\tau}{\varepsilon}} = \mathcal{O}(\mathrm{e}^{-\frac{1}{\varepsilon}})$ for $\tau = \mathcal{O}(1)$ in Equation (25). The corresponding terms are hence exponentially small (in ε), and can be safely neglected; in fact, since the manifold $\mathcal{S}_\varepsilon$ itself is unique only up to exponentially small terms, recall Remark 2, it would certainly be inconsistent to retain them in (25) when τ is large, as one can always choose a representative for $\mathcal{S}_\varepsilon$ for which they cancel. □

4.2. **Generating function.** In this subsection, we derive the first-order asymptotics (in ε) of the generating function F.

4.2.1. *Inner (fast) asymptotics.* To describe the fast asymptotics of the generating function F up to an $\mathcal{O}(\varepsilon^2)$-error, we consider the expansion for $F(u_0,v_0,t,\varepsilon)$ from Equation (14b), truncated at first order in ε. Applying the transformation defined in (23) to eliminate any dependence on (u_0,v_0) from $\mathcal{F}_1$, cf. Equation (19), expanding the result in ε, and retaining first-order terms, we find

$$\begin{aligned}F(u,v,t,\varepsilon) &= 1 + \varepsilon\frac{a}{1-bv}\left\{bvt + \Big(u - \frac{bv}{1-bv}\Big)\big[1-\mathrm{e}^{-(1-bv)t}\big]\right\} + \mathcal{O}(\varepsilon^2)\\ &= \mathcal{F}_0(u,v,t) + \varepsilon\mathcal{F}_1(u,v,t) + \mathcal{O}(\varepsilon^2).\end{aligned} \tag{26}$$

We emphasise that, since $F(u,v,t,\varepsilon)\big|_{(u,v)=(1,0)} = 1+\varepsilon a u[1-\mathrm{e}^{-(1-bv)t}]+\mathcal{O}(\varepsilon^2)$, the marginal distribution of mRNA – which is defined as $P_m = \sum_{n=0}^{\infty} P_{mn}$ – satisfies

$$P_m^{\infty}(\varepsilon) = \begin{cases} 1-\varepsilon a & \text{when } m=0,\\ \varepsilon a & \text{when } m=1,\\ 0 & \text{when } m\geq 2 \end{cases}$$

in the stationary limit as $t\to\infty$; clearly, the above formulae are consistent with the well-known exact Poissonian distribution for P_m^{∞} [8, Equation (32)] for ε sufficiently small, up to an $\mathcal{O}(\varepsilon^2)$-error.

For future reference, we note that the validity of the fast asymptotics derived in this subsection is restricted to $\varepsilon t = \mathcal{O}(1)$, as the expansions for u_0, v_0, and F in Equations (23) and (26) become inconsistent when $t=\mathcal{O}(\varepsilon^{-1})$ due to the presence of 'secular' terms [32, 55] therein.

4.2.2. *Outer (slow) asymptotics.* The initial condition in the leading-order approximation F_0 for F, Equation (21), can be determined by a 'matching' argument, which fixes the constant C_0: we note that $\lim_{t\to\infty}\mathcal{F}_0(u,v,t)$ must equal $\lim_{\tau\to 0^+}F_0(v,\tau)$ for any $(u,v)\in\mathcal{D}$; cf. also [9, Section 4]. Here, $F_0(v,\tau)\equiv F(U_0(v),v,\tau,0)$ corresponds to the leading-order term in an expansion for the generating function F of the form

$$F(u,v,\tau,\varepsilon) = \sum_{k=0}^{K} F_k(u,v,\tau)\varepsilon^k + \Delta F(u,v,\tau,\varepsilon), \quad \text{with } \Delta F = \mathcal{O}(\varepsilon^{K+1}), \tag{27}$$

whose existence is guaranteed by Proposition 3.1(iv) and standard smoothness results for ordinary differential equations [1]. Hence, Equation (17) implies that $F_0(v_0)=1$ in Equation (21), which gives $C_0=(1-bv_0)^a$. Rewriting the resulting expression as a function of v and τ, via $v_0(v,\tau)=v\mathrm{e}^{-\tau}$, we conclude that

$$F_0(v,\tau) = \Big(\frac{1-bv\mathrm{e}^{-\tau}}{1-bv}\Big)^a; \tag{28}$$

see also [50, Equation (7)].

To fix the constant C_1 in Equation (22), we match the truncated first-order slow expansion $F_0+\varepsilon F_1$ to its fast counterpart $\mathcal{F}_0+\varepsilon\mathcal{F}_1$; in fact, expanding the coefficient function F_0 in terms of the fast time t and comparing the two resulting expansions, one finds that the relation $\frac{a}{1-bv}\big(u-\frac{bv}{1-bv}\big) = \frac{a}{2}\frac{1}{(1-bv_0)^2}+\frac{C_1}{(1-bv_0)^a}$ must hold. (Here, we emphasise that the presence of a u-dependent term is due to the structure of the inverse characteristic transformation, Equation (23), which induces a t-independent such term in (26).) Substituting into (22) and eliminating any v_0-dependence therein, we obtain

$$F_1(u,v,\tau) = a\bigg\{\frac{1}{2}\bigg[\frac{1}{(1-bv)^2}-\frac{1}{(1-bv\mathrm{e}^{-\tau})^2}\bigg]+\frac{1}{1-bv}\Big(u-\frac{bv}{1-bv}\Big)\bigg\}\Big(\frac{1-bv\mathrm{e}^{-\tau}}{1-bv}\Big)^a. \tag{29}$$

The first-order asymptotics (in ε) of the generating function F is now obtained by substituting the expressions for F_0 and F_1 from Equations (28) and (29), respectively, into the expansion in (27); recalling that $m_0=0=n_0$, we find

$$F(u,v,\tau,\varepsilon) = \bigg\{1+\varepsilon\frac{a}{2}\bigg[\frac{1}{(1-bv)^2}-\frac{1}{(1-bv\mathrm{e}^{-\tau})^2}\bigg]+\varepsilon\frac{a}{1-bv}\Big(u-\frac{bv}{1-bv}\Big)\bigg\}\Big(\frac{1-bv\mathrm{e}^{-\tau}}{1-bv}\Big)^a+\mathcal{O}(\varepsilon^2) \tag{30}$$

for ε positive, but sufficiently small, and $\tau\gg\varepsilon$. We remark that Equation (30) satisfies the normalisation condition $F(0,0,\tau,\varepsilon)=1$ to the order $\mathcal{O}(\varepsilon^2)$, as required by Equation (4). Moreover, in the stationary limit as $\tau\to\infty$, F reduces to $F^{\infty}(u,v,\varepsilon)\equiv\lim_{\tau\to\infty}F(u,v,\tau,\varepsilon) = (1-bv)^{-a}\big\{1+\varepsilon\frac{a}{1-bv}\big[u-\frac{1}{2}\frac{(bv)^2}{1-bv}\big]\big\}+\mathcal{O}(\varepsilon^2)$. Finally, since $F(u,v,\tau,\varepsilon)\big|_{(u,v)=(1,0)} = 1+\varepsilon a u+\mathcal{O}(\varepsilon^2)$, the marginal

mRNA distribution obtained from (30) again agrees with the asymptotics of the corresponding exact distribution, up to an $\mathcal{O}(\varepsilon^2)$-error.

We emphasise that the leading-order generating function $F(u,v,\tau,0)$ is independent of u on this slow τ-scale, as both the flow on the critical manifold $\mathcal{S}_0$ and the corresponding stationary limit of $F(u,v,t,0)$ on the fast t-scale are u-independent. (The latter assertion, while trivial in the case of $m_0 = 0 = n_0$ considered here, can be shown to be valid for any m_0, $n_0 \in \mathbb{N}_0$; cf. Section B of the Online Supplement.) However, no such claim can be made at first order in ε: clearly, the expansion for F in Equation (30) depends not only on v, τ, and ε, but also on u. This u-dependence is introduced by the requirement that the fast and the slow expansions for F agree to first order in ε on some domain of overlap, which implies that the free constant C_1 in the general solution for F_1 has to account for the contribution from the 'boundary layer' coefficient $\mathcal{F}_1$; in other words, mere knowledge of the flow induced by Equation (8) on the slow manifold $\mathcal{S}_\varepsilon$ is insufficient.

That same argument can be generalised to any order (in ε): the higher-order asymptotics of F on the slow τ-scale – as encoded in the coefficient functions F_j ($j \in \mathbb{N}$) – can be derived by observing that the flow on the slow manifold $\mathcal{S}_\varepsilon$ is a regular perturbation of the singular (reduced) dynamics that is obtained in the limit as $\varepsilon \to 0$, as was done in Section 3.3. However, contributions from the fast flow will be introduced via the free constants that arise in the solution: while these constants must be τ-independent, they will typically depend on u (as well as on v), as one cannot expect the u-dependence of the coefficient functions $\mathcal{F}_j$ to have subsided in the large-t limit once the inverse characteristic transformation $(u_0,v_0) \mapsto (u,v)$ has been applied. (Correspondingly, the point $(u,v) = (-1,-1)$ at which, by Equation (3), F and its derivatives need to be evaluated, will not generically lie on $\mathcal{S}_\varepsilon$.) The resulting dependence of F on both u and v implies that the distribution P_{mn} will, in general, be non-zero when $m \geq 1$ at higher orders in ε; specifically, we conjecture that $P_{mn}(\tau,\varepsilon) = \mathcal{O}(\varepsilon^m)$ will hold for any $m \in \mathbb{N}_0$.

4.2.3. *Uniform ('fast-slow') asymptotics.* Finally, a uniform expansion for the generating function F may be obtained by combining the fast and the slow asymptotics given in Equations (26) and (30), respectively, taking into account that the corresponding common terms are given by $1+\varepsilon\frac{a}{1-bv}\big(bvt+u-\frac{bv}{1-bv}\big)$, up to an $\mathcal{O}(\varepsilon^2)$-error. Thus, we find

$$F(u,v,\tau,t,\varepsilon) = F_0(v,\tau) + \varepsilon\big[F_1(u,v,\tau) + \mathcal{F}_1'(u,v,t)\big] + \mathcal{O}(\varepsilon^2), \tag{31}$$

where the coefficient functions F_j ($j = 0,1$) are defined as in Equation (30) and where, moreover, $\mathcal{F}_1'(u,v,t) = -\frac{a}{1-bv}\big(u-\frac{bv}{1-bv}\big)\mathrm{e}^{-(1-bv)t}$. The resulting composite (fast-slow) expansion for F in τ and t is uniformly valid for $t \in [0,t_*]$, with $t_* > 0$ arbitrary, but fixed – or, equivalently, for $\tau \in [0,\tau_*]$, where $\tau_* = \varepsilon t_*$, with ε positive and sufficiently small; see, *e.g.*, [32, 55].

4.3. **Probability distributions.** In this subsection, we derive the first-order asymptotics (in ε) of the probability distributions $P_{mn}(t,\varepsilon)$ and $P_{mn}(\tau,\varepsilon)$ on the fast and the slow time-scales, respectively, as well as of the corresponding marginal distributions of protein. Then, we deduce a composite expansion for the marginal protein distribution $P_n(\tau,t,\varepsilon)$ that is uniformly valid in time.

4.3.1. *Inner (fast) asymptotics.* Recalling Equation (3), rewritten in terms of the translated coordinates u and v and the fast time t and evaluated at $m_0 = 0 = n_0$, we find

$$P_{mn}(t,\varepsilon) = \frac{1}{m!}\frac{1}{n!}\frac{\partial^m}{\partial u^m}\frac{\partial^n}{\partial v^n}\big[\mathcal{F}_0(u,v,t) + \varepsilon\mathcal{F}_1(u,v,t)\big]\Big|_{(u,v)=(-1,-1)} + \mathcal{O}(\varepsilon^2) \tag{32}$$

for the relationship between F and P_{mn}. (Here, we allow for $m,n \in \mathbb{N}_0 = \mathbb{N}\cup\{0\}$, as before.) We emphasise that the relation in (32) is, in fact, well-defined to arbitrary order in ε, which is due to the differentiability of the expansion for F in Equation (14b) with respect to both u and v, in combination with the regularity of the inverse characteristic transformation $(u,v) \mapsto (u_0,v_0)$; cf. Section 4.1. Moreover, well-known uniqueness properties of asymptotic expansions [32] imply that

the derivatives $\frac{\partial^m}{\partial u^m}\frac{\partial^n}{\partial v^n}F$ may be approximated by differentiation of Equation (14b) with respect to (u,v), for any $1 \leq m \leq M$ and $1 \leq n \leq N$ with $M, N \in \mathbb{N}$.

In particular, the first-order asymptotics of the probability distribution $P_{mn}(t,\varepsilon)$ can now be derived by differentiating the corresponding expansion for F, as given in Equation (26), repeatedly with respect to u and v, and by evaluating the result at $(u,v) = (-1,-1)$:

Proposition 4.1. *Let $m, n \in \mathbb{N}_0$, let $\varepsilon \in [0, \varepsilon_0]$, with $\varepsilon_0 >$ sufficiently small, and assume that $t \ll \varepsilon^{-1}$, i.e., let $\varepsilon t = \mathcal{O}(1)$. Then, the probability distribution $P_{mn}(t,\varepsilon) \equiv P_{mn|00}(t,\varepsilon)$ is given by*

$$P_{00}(t,\varepsilon) = 1 - \varepsilon\frac{a}{1+b}\Big\{bt + \frac{1}{1+b}\big[1 - \mathrm{e}^{-(1+b)t}\big]\Big\} + \mathcal{O}(\varepsilon^2) \tag{33}$$

for $m = 0 = n$ and by

$$\begin{aligned} P_{0n}(t,\varepsilon) &= \frac{\varepsilon}{\Gamma(n+2)}ab^n t^{n+1} \\ &\quad\times\Big\{{}_1F_1\big(n+1;n+2;-(1+b)t\big)t - \frac{n+1}{1+b}\big[{}_1F_1\big(n+1;n+2;-(1+b)t\big) - \mathrm{e}^{-(1+b)t}\big]\Big\} + \mathcal{O}(\varepsilon^2) \end{aligned} \tag{34}$$

for $m = 0$ and any $n \in \mathbb{N}$. When $m = 1$,

$$P_{1n}(t,\varepsilon) = \frac{\varepsilon}{\Gamma(n+2)}ab^n t^{n+1}{}_1F_1\big(n+1;n+2;-(1+b)t\big) + \mathcal{O}(\varepsilon^2) \tag{35}$$

for arbitrary $n \in \mathbb{N}_0$; here, Γ and ${}_1F_1$ denote the standard Gamma function [3, Section 6] *and the confluent hypergeometric function* [3, Section 13]*, respectively. For $m \geq 2$, $P_{mn}(t,\varepsilon) \equiv 0$ to the order considered here.*

Proposition 4.1 affirms, in particular, that $P_{mn}(t,\varepsilon) = 0$ to leading order in ε unless $m = 0 = n$, as $F(u,v,t,0) \equiv 1$ when $m_0 = 0 = n_0$; for a detailed discussion, the reader is referred to Section B of the Online Supplement, where the large-t asymptotics of the leading-order fast propagator $P_{mn|m_0n_0}(t,0)$ is studied for arbitrary values of m_0 and n_0. For future reference, we emphasise that the statement of Proposition 4.1 is only valid for $t \ll \varepsilon^{-1}$, as the expansion for F in (26) breaks down when $\varepsilon t = \mathcal{O}(1)$; in other words, the limit as $t \to \infty$ in $P_{mn}(t,\varepsilon)$ is undefined at first order in ε. Finally, it follows from Equations (33) through (35) that $P_{0n}(0,\varepsilon) = \delta_{0n}$, with $n \in \mathbb{N}_0$, while $P_{1n}(0,\varepsilon) = 0$ throughout, which is consistent with our assumption that $P_{mn|m_0n_0}(0,\varepsilon) = \delta_{mm_0}\delta_{nn_0}$.

***Remark* 4.** The fact that $P_{mn}(t,\varepsilon) \equiv 0$ for $m \neq 0, 1$ in Proposition 4.1 agrees with findings reported in Section B.B of [42], where only transitions from 0 to 1 mRNAs were considered in the two-stage model for gene expression studied here, as well as with [9, Section 3]. (In [42, Section B.C], the resulting generating function was expressed in terms of a confluent hypergeometric function ${}_1F_1$, which also appears in Proposition 4.1.) Related results were obtained in Section 2 of [8], where the stationary generating function of the joint distribution of mRNA and protein was represented as a Kummer function [3, Section 13]. □

Finally, we approximate the fast marginal distribution of protein $P_n(t,\varepsilon) = \sum_{m=0}^{\infty} P_{mn}(t,\varepsilon)$, up to an $\mathcal{O}(\varepsilon^2)$-error: as non-zero contributions to P_n are only obtained for $m = 0$ and $m = 1$ to that order, by Proposition 4.1, it follows that $P_n(t,\varepsilon) = P_{0n}(t,\varepsilon) + P_{1n}(t,\varepsilon) + \mathcal{O}(\varepsilon^2)$, which gives

$$P_n(t,\varepsilon) = \begin{cases} 1 - \varepsilon\frac{ab}{1+b}\Big\{t - \frac{1}{1+b}\big[1 - \mathrm{e}^{-(1+b)t}\big]\Big\} + \mathcal{O}(\varepsilon^2) & \text{for } n = 0, \\ \frac{\varepsilon}{(n+1)!}ab^n t^{n+1}\Big\{(t+1){}_1F_1\big(n+1;n+2;-(1+b)t\big) \\ \quad -\frac{n+1}{1+b}\big[{}_1F_1\big(n+1;n+2;-(1+b)t\big) - \mathrm{e}^{-(1+b)t}\big]\Big\} + \mathcal{O}(\varepsilon^2) & \text{for } n \in \mathbb{N}, \end{cases} \tag{36}$$

after some simplification; see also Section B.2.1 of the Online Supplement, where the more general case of $n_0 \neq 0$ is considered. (Alternatively, the above expansion can be obtained by evaluating the

first-order generating function F, as given in Equation (26), at $u = 0$, and by differentiating the result repeatedly with respect to v.)

4.3.2. *Outer (slow) asymptotics.* Next, we consider the asymptotics of the distribution P_{mn} on the slow time-scale τ: by Equation (3), we have

$$P_{mn}(\tau,\varepsilon) = \frac{1}{m!}\frac{1}{n!}\frac{\partial^m}{\partial u^m}\frac{\partial^n}{\partial v^n}\big[F_0(v,\tau)+\varepsilon F_1(u,v,\tau)\big]\Big|_{(u,v)=(-1,-1)} + \mathcal{O}(\varepsilon^2) \tag{37}$$

to first order in ε, with $m,n\in\mathbb{N}_0$, as before. Moreover, we remark that the relation in (37) is again well-defined, as the asymptotic expansion for the generating function F in Equation (27) can be differentiated arbitrarily often with respect to u and v [1, 32]; recall Section 4.3.1.

As was done there, we may hence make use of the first-order approximation for F obtained in the previous subsection to approximate the distribution $P_{mn}(\tau,\varepsilon)$ on the slow τ-scale:

Proposition 4.2. *Let $n\in\mathbb{N}_0$, let $\varepsilon\in[0,\varepsilon_0]$, with $\varepsilon_0>0$ sufficiently small, and assume that $\tau\gg\varepsilon$. Then, the probability distribution $P_{mn}(\tau,\varepsilon)\equiv P_{mn|00}(\tau,\varepsilon)$ is given by*

$$\begin{aligned} P_{0n}(\tau,\varepsilon) &= \frac{\Gamma(a+n)}{\Gamma(n+1)\Gamma(a)}\Big(\frac{b}{1+b}\Big)^n\Big(\frac{1+b\mathrm{e}^{-\tau}}{1+b}\Big)^a \\ &\quad\times\Bigg\{{}_2F_1\big(-n,-a;1-a-n;\tfrac{1+b}{\mathrm{e}^{\tau}+b}\big) - \frac{\varepsilon}{2}\frac{a}{(1+b)^2}\sum_{k=0}^{n}\frac{\Gamma(n+1)}{\Gamma(n-k+1)}\frac{\Gamma(a+n-k)}{\Gamma(a+n)} \\ &\qquad\times(k+1)\bigg[1+\Big(\frac{1+b}{\mathrm{e}^{\tau}+b}\Big)^{k+2}\mathrm{e}^{2\tau}\bigg]{}_2F_1\big(-n+k,-a;1-a-n+k;\tfrac{1+b}{\mathrm{e}^{\tau}+b}\big)\Bigg\} + \mathcal{O}(\varepsilon^2) \end{aligned} \tag{38}$$

when $m=0$ and by

$$\begin{aligned} P_{1n}(\tau,\varepsilon) &= \varepsilon\frac{\Gamma(a+n)}{\Gamma(n+1)\Gamma(a)}\Big(\frac{b}{1+b}\Big)^n\Big(\frac{1+b\mathrm{e}^{-\tau}}{1+b}\Big)^a\frac{a}{1+b}\sum_{k=0}^{n}\frac{\Gamma(n+1)}{\Gamma(n-k+1)}\frac{\Gamma(a+n-k)}{\Gamma(a+n)} \\ &\qquad\times{}_2F_1\big(-n+k,-a;1-a-n+k;\tfrac{1+b}{\mathrm{e}^{\tau}+b}\big) + \mathcal{O}(\varepsilon^2) \end{aligned} \tag{39}$$

when $m=1$; here, ${}_2F_1$ denotes the standard hypergeometric function [3, Section 15]. *For $m\geq 2$, $P_{mn}(\tau,\varepsilon)\equiv 0$ to the order considered here.*

We remark that the leading-order contribution $P_n(\tau,0)\equiv P_{0n}(\tau,0)$ was previously derived in [50, Equation (9)], as well as that the expression for P_{1n} in (39) is consistent with [8, Equation (13)].

Moreover, we emphasise that $P_{mn}(\tau,0)\equiv 0$ for any $m\in\mathbb{N}$, *i.e.*, that the marginal distribution of mRNA peaks at zero to leading order in ε, as was noted already in [50]; see also [9, Equations (9) and (10)]. Specifically, the bivariate distribution P_{0n} equals the marginal protein distribution P_n when $\varepsilon=0$, as $P_n(\tau,0)=\sum_{m\in\mathbb{N}_0}P_{mn}(\tau,0)=P_{0n}(\tau,0)$. At first order in ε, we have $P_n(\tau,\varepsilon)=P_{0n}(\tau,\varepsilon)+P_{1n}(\tau,\varepsilon)+\mathcal{O}(\varepsilon^2)$ for the marginal distribution of protein on the slow τ-scale, recall the discussion towards the end of Section 4.3.1, which implies

$$\begin{aligned} P_n(\tau,\varepsilon) &= \frac{\Gamma(a+n)}{\Gamma(n+1)\Gamma(a)}\Big(\frac{b}{1+b}\Big)^n\Big(\frac{1+b\mathrm{e}^{-\tau}}{1+b}\Big)^a \\ &\quad\times\Bigg({}_2F_1\big(-n,-a;1-a-n;\tfrac{1+b}{\mathrm{e}^{\tau}+b}\big) + \frac{\varepsilon}{2}\frac{a}{(1+b)^2}\sum_{k=0}^{n}\frac{\Gamma(n+1)}{\Gamma(n-k+1)}\frac{\Gamma(a+n-k)}{\Gamma(a+n)} \\ &\quad\times\bigg\{2(1+b)-(k+1)\bigg[1+\Big(\frac{1+b}{\mathrm{e}^{\tau}+b}\Big)^{k+2}\mathrm{e}^{2\tau}\bigg]\bigg\}{}_2F_1\big(-n+k,-a;1-a-n+k;\tfrac{1+b}{\mathrm{e}^{\tau}+b}\big)\Bigg) + \mathcal{O}(\varepsilon^2), \end{aligned} \tag{40}$$

for any $n\in\mathbb{N}_0$. (Again, the expression in (40) can equivalently be derived directly from the first-order slow asymptotics of F; recall Equation (30).)

4.3.3. *Stationary limit.* The result of Proposition 4.2 simplifies substantially in the stationary limit as $\tau \to \infty$ in (38):

Corollary 4.1. *Let $m, n \in \mathbb{N}_0$, and let $\varepsilon \in [0, \varepsilon_0]$, with $\varepsilon_0 > 0$ sufficiently small. Then, the stationary probability distribution $P_{mn}^{\infty}(\varepsilon) := \lim_{\tau\to\infty} P_{mn}(\tau, \varepsilon)$ is given by*

$$P_{0n}^{\infty}(\varepsilon) = \frac{\Gamma(a+n)}{\Gamma(n+1)\Gamma(a)}\Big(\frac{b}{1+b}\Big)^n\Big(1-\frac{b}{1+b}\Big)^a\bigg\{1-\frac{\varepsilon}{2}\bigg[\frac{(a+n+1)(a+n)}{(1+b)^2(a+1)}+a\bigg]\bigg\}+\mathcal{O}(\varepsilon^2) \tag{41}$$

when $m = 0$ and by

$$P_{1n}^{\infty}(\varepsilon) = \varepsilon\frac{\Gamma(a+n)}{\Gamma(n+1)\Gamma(a)}\Big(\frac{b}{1+b}\Big)^n\Big(1-\frac{b}{1+b}\Big)^a\frac{a+n}{1+b}+\mathcal{O}(\varepsilon^2) \tag{42}$$

when $m = 1$, for arbitrary $n \in \mathbb{N}_0$. For $m \geq 2$, $P_{mn}(\varepsilon) \equiv 0$ to the order considered here.

Given Corollary 4.1, it is straightforward to determine the stationary marginal protein distribution $P_n^{\infty}(\varepsilon)$, by taking the limit as $\tau \to \infty$ in Equation (40):

$$\begin{aligned} P_n^{\infty}(\varepsilon) = \frac{\Gamma(a+n)}{\Gamma(n+1)\Gamma(a)}\Big(\frac{b}{1+b}\Big)^n\Big(1-\frac{b}{1+b}\Big)^a \\ \times\bigg[1-\varepsilon\frac{(a+1)ab^2-2(a+1)bn+n(n-1)}{2(a+1)(1+b)^2}\bigg]+\mathcal{O}(\varepsilon^2). \end{aligned} \tag{43}$$

4.3.4. *Uniform (fast-slow) asymptotics.* Finally, we derive a composite first-order expansion (in ε) for the marginal protein distribution P_n that is uniformly valid both on the fast and the slow time-scales. We have the following result:

Proposition 4.3. *Let $\varepsilon \in [0, \varepsilon_0]$, with $\varepsilon_0 > 0$ sufficiently small, let $t_* > 0$ be arbitrary, but fixed, and let $\tau_* = \varepsilon t_*$. Up to an $\mathcal{O}(\varepsilon^2)$-error, the uniform marginal protein distribution $P_n(\tau, t, \varepsilon)$ is then given by*

$$\begin{aligned} P_n(\tau,t,\varepsilon) = P_n(\tau,\varepsilon)+\varepsilon a\frac{b^n}{(1+b)^{n+2}}\big[n-b-(1+b)t\big]+\frac{\varepsilon}{(n+1)!}ab^nt^{n+1} \\ \times\bigg\{{}_1F_1\big(n+1;n+2;-(1+b)t\big)t-\frac{1}{1+b}\big[(n-b){}_1F_1\big(n+1;n+2;-(1+b)t\big)-(n+1)\mathrm{e}^{-(1+b)t}\big]\bigg\}, \end{aligned} \tag{44}$$

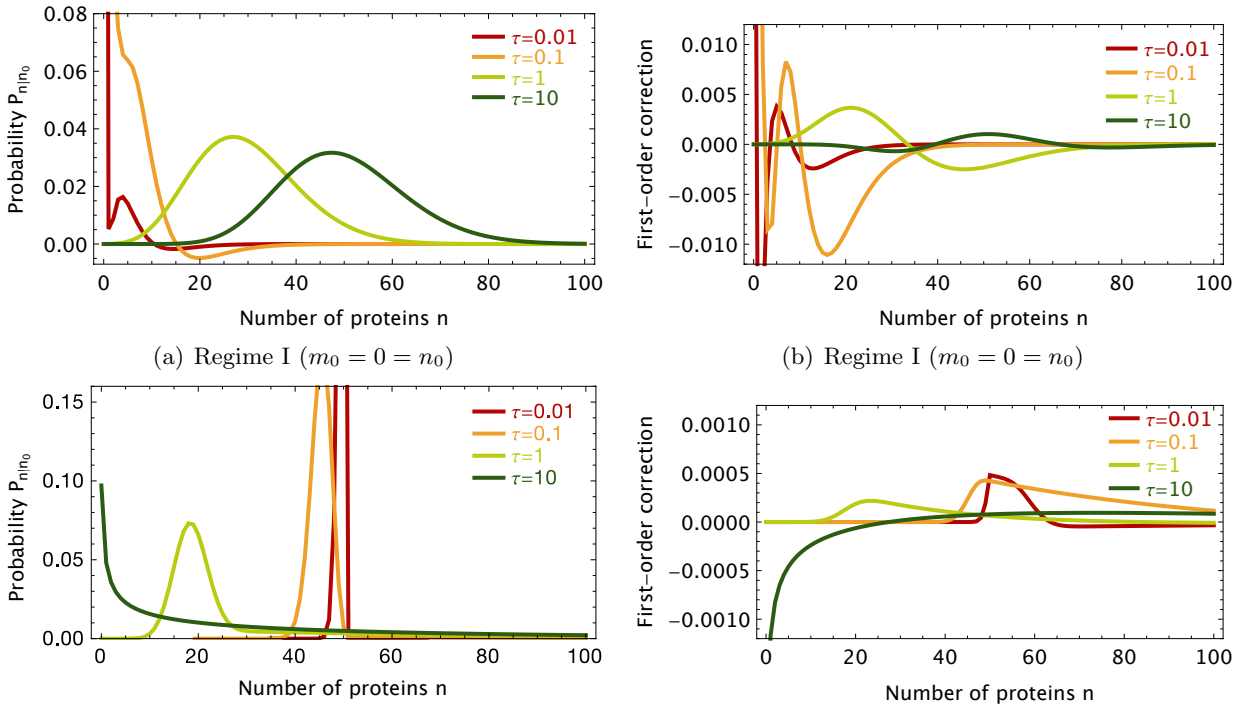
for any $t \in [0, t_*]$ and $\tau \in [0, \tau_*]$. (Here, $P_n(\tau, \varepsilon)$ is defined as in Equation (40).)

The proof of Proposition 4.3 is based on repeated differentiation, with respect to u and v, of the corresponding uniform expansion, Equation (31), for the probability-generating function F; recall Equations (3) and (32). Details can be found in Section C of the Online Supplement.

Since $\Gamma(n+1, z) = z^n\mathrm{e}^{-z}\big[1+\mathcal{O}(z^{-1})\big]$, by [3, Equation 6.5.32], it follows that $\lim_{z\to\infty}\big[(b-n+z)\Gamma(n+1,z)-z^{n+1}\mathrm{e}^{-z}\big]=0$ and, hence, that $P_n(\tau, t, \varepsilon)$ reduces to its slow counterpart $P_n(\tau, \varepsilon)$ in the large-t limit, as required. Moreover, Equation (43) above then implies that the large-τ limit in Equation (44) is well-defined; in other words, we may take the time intervals $[0, t_*]$ and $[0, \tau_*]$ on which the uniform (fast-slow) distribution $P_n(\tau, t, \varepsilon)$ is valid to be unbounded.

The first-order asymptotics of the uniform propagator $P_{n|n_0}(\tau, t, \varepsilon)$ is illustrated in Figure 3 for varying values of $\tau = \varepsilon t$, as well as of $n_0 \in \mathbb{N}$. (We remark that, for $n_0 = 0$, $P_n \equiv P_{n|0}$ is approximated in Equation (44), while the corresponding result for general $n_0 \in \mathbb{N}$ can be found in Proposition B.7 of the Online Supplement.) Throughout, we consider the two parameter regimes proposed in [50], with either $a = 20$ and $b = 2.5$ or $a = 0.5$ and $b = 100$, which will henceforth be labelled 'regime I' and 'regime II,' respectively; moreover, we set $\gamma = 10$ in both cases.

In Figure 3(a), we note a pronounced shift (with time) towards increasing protein numbers, as the large value of a in regime I implies high levels of mRNA synthesis. By contrast, in regime II,



(a) Regime I ($m_0 = 0 = n_0$)

(b) Regime I ($m_0 = 0 = n_0$)

(c) Regime II ($m_0 = 0$, $n_0 = 50$)

(d) Regime II ($m_0 = 0$, $n_0 = 50$)

FIGURE 3. The uniform propagator $P_{n|n_0}(\tau, t, \varepsilon)$, as well as the corresponding first-order correction alone, for $\gamma = 10$, $m_0 = 0$, and varying values of n_0. Here, $a = 20$ and $b = 2.5$ in panels 3(a) and 3(b) ('regime I'), while $a = 0.5$ and $b = 100$ in panels 3(c) and 3(d) ('regime II'); moreover, $P_{n|n_0}(\tau, t, \varepsilon)$ is described in Propositions 4.3 and B.7 for $n_0 = 0$ and $n_0 \in \mathbb{N}$, respectively. We note that, in regime II, the distribution evolves towards $n = 0$ due to the corresponding low value of a, whereas the large-τ limit in regime I appears near-symmetric, with a peak close to $ab(= 50)$. The negative probabilities observed for small τ in panel 3(a) indicate that $P_{n|n_0}$ is inconsistent for $\gamma = 10$ in regime I; cf. Section 5 below for details.

mRNA is synthesised only infrequently due to a being small; hence, the initial sharp peak seen at $n_0(= 50)$ in Figure 3(c) rapidly abates to zero. An in-depth discussion of the dependence of $P_{n|n_0}$ on the two parameters a and b is given in Section 5 below. Finally, in Figures 3(b) and 3(d), the first-order correction alone – which can be expressed as $\lim_{\varepsilon\to 0+}\left[P_{n|n_0}(\tau, t, \varepsilon) - P_{0n|0n_0}(\tau, 0)\right]\varepsilon^{-1}$ – is depicted; here, $P_{0n|0n_0}(\tau, 0)$ is defined as in Proposition B.2. In both parameter regimes, one confirms the convergence of the propagator $P_{n|n_0}$ to the corresponding stationary distribution P_n^∞ with increasing τ.

5. Verification and application

In this section, we present a numerical verification of the asymptotics derived in Section 4; then, we discuss the practical applicability thereof. Specifically, we compare the accuracy of the asymptotic series expansions for the marginal protein distributions $P_{n|n_0}$ and P_n^∞ obtained in this article with a stochastic simulation of the CME, Equation (1). The latter relies on Gillespie's stochastic simulation algorithm (SSA) [20], which was implemented using the software package StochKit [51]; again, we focus predominantly on the two parameter regimes I ($a = 20$ and $b = 2.5$) and II ($a = 0.5$ and $b = 100$) which were also studied in [50].

While we find that the first-order correction to $P_{n|n_0}$ gives an improvement over the leading-order asymptotics in many parameter regimes both at steady state and in the time-dependent case, the accuracy of the resulting approximation will depend crucially on the singular perturbation parameter $\varepsilon = \gamma^{-1}$ and the non-dimensionalised system parameters (a, b), as specified in detail below. (The dependence on the respective initial numbers m_0 and n_0 of mRNA and protein is less pronounced, and is discussed briefly in the Online Supplement.)

5.1. **Stationary protein distribution.** To investigate systematically the significance of the first-order correction in $\varepsilon(=\gamma^{-1})$ for the asymptotics of the protein distribution P_n, Equation (44), we compared the corresponding stationary limit P_n^∞, as approximated to zeroth and to first order in γ^{-1} in Corollary 4.1, with a stochastic simulation of the underlying two-stage model; recall Figure 1(a).

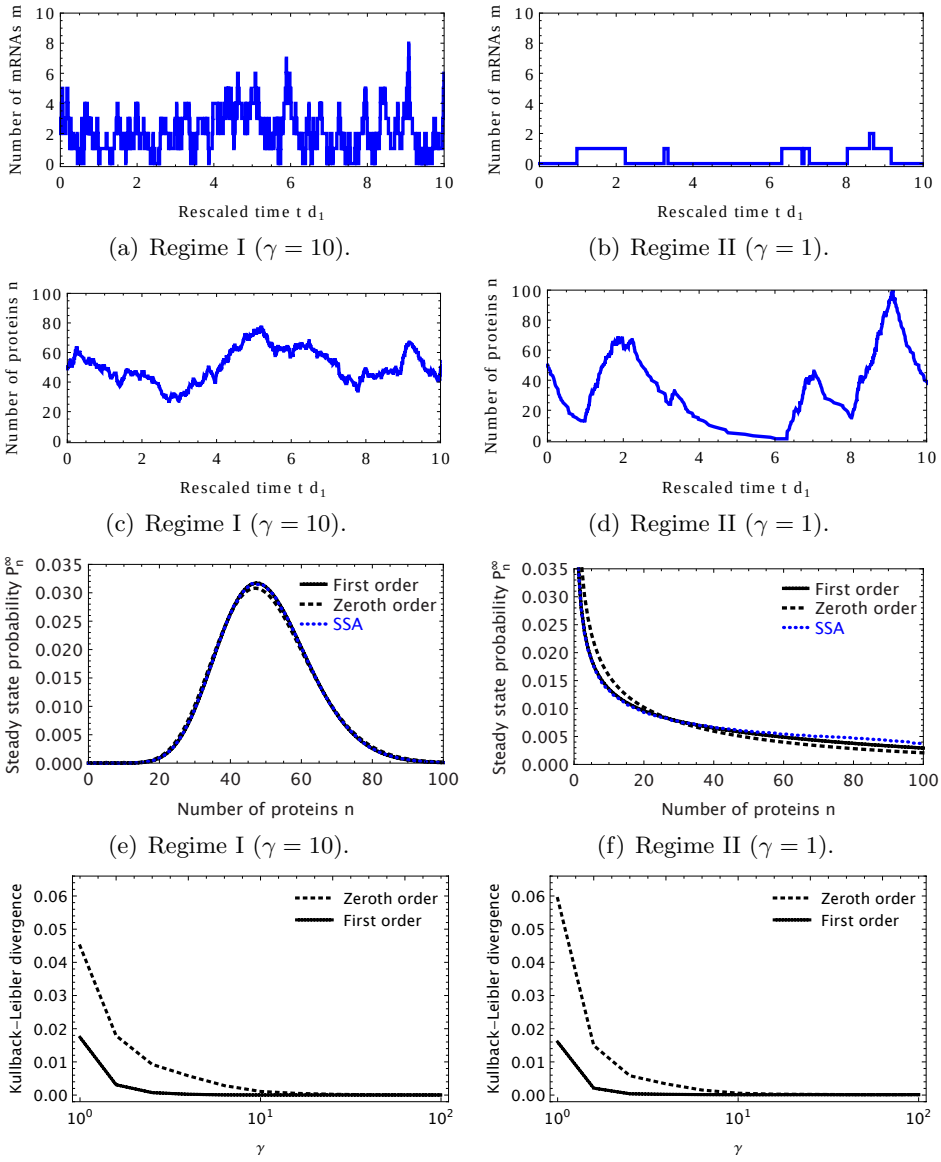
We first considered regimes I and II with $\gamma = 10$ and $\gamma = 1$, respectively; representative time series of both mRNA and protein are presented in Figures 4(a) and 4(c) and in Figures 4(b) and 4(d), respectively, while the corresponding protein distributions at steady state are shown in Figures 4(e) and 4(f). These distributions were generated from $2 \cdot 10^6$ samples each, which were taken from simulated trajectories at uncorrelated time points; to ensure non-correlation, we determined the typical de-correlation time of the system – as the time where the autocorrelation of the trajectories drops to 0.5 – to be about $1d_1$ in regime I, and about $2d_2$ in regime II.

As is obvious from Figure 4(b), the small value of a – the average number of mRNA molecules synthesised during a protein lifetime – in regime II implies that only very few mRNAs are typically generated. However, as b is large in that case, each such event results in a rapid 'burst' in protein numbers; cf. Figure 4(d). We note the negativity of the first-order correction in that regime at low protein levels, which is consistent with bursting, *i.e.*, with a bias towards large numbers of protein that is correctly accounted for by our asymptotics. (In fact, since the mean time between bursts is $\nu_0^{-1} = d_1 a^{-1}$, one may expect a burst to appear for $\tau = \mathcal{O}(1)$, as is also observed in Figure 3(d); the reader is referred to [37, 50] for a detailed discussion of translational bursting in the two-stage model for gene expression studied here.) By contrast, in regime I, a is moderately large, while b – the mean burst size – is small; hence, considerable numbers of mRNA are synthesised throughout, as seen in Figure 4(a). Consequently, the fast-slow structure of the system seems less pronounced, which is confirmed by Figure 4(c); cf. also Figure 2 above.

Finally, in Figures 4(g) and 4(h), the (averaged) Kullback-Leibler divergence – or 'relative entropy' [14] – between the simulated steady-state distribution and our perturbative approximation for P_n^∞ is shown for varying γ. (Specifically, we evaluated the Kullback-Leibler divergence as $\sum_n \mathcal{P}_n(\gamma) \log_2 \frac{\mathcal{P}_n(\gamma)}{\mathcal{Q}_n(\gamma)}$, where $\mathcal{P}_n$ and $\mathcal{Q}_n$ denote the 'true' distribution obtained via SSA and the first-order expansion quoted in Corollary 4.1, respectively; to ensure the numerical stability of our implementation, we extracted the distribution at steady state from $2 \cdot 10^6$ uncorrelated samples and then determined the Kullback-Leibler divergence up to the 99th percentile of that simulated distribution, disregarding terms with absolute simulated probabilities below γ^{-2}.) In both regimes, our first-order expansion consistently outperforms the leading-order approximation alone for 'moderately large' values of $1 < \gamma < 10$, whereas no discernible difference is seen when $\gamma > 10$. (We remark that the Kullback-Leibler divergence of the leading-order approximation is lower in regime I than it is in regime II when $\gamma = \mathcal{O}(1)$, while the situation is reversed at first order, in agreement with Figure 1D of [50].)

To assess more generally the dependence of our perturbative approximation for the stationary probability distribution P_n^∞, Equation (43), on the values of the dimensionless parameters $a = \frac{\nu_0}{d_1}$ and $b = \frac{\nu_1}{d_0}$, we performed a series of numerical experiments, the outcome of which is summarised in Figure 5.

The inconsistency of the first-order expansion for P_n^∞, as observed in some parameter regimes in Figure 5(a), agrees with a result known as Pawula's Theorem [44, Section 4.3], whereby asymptotic



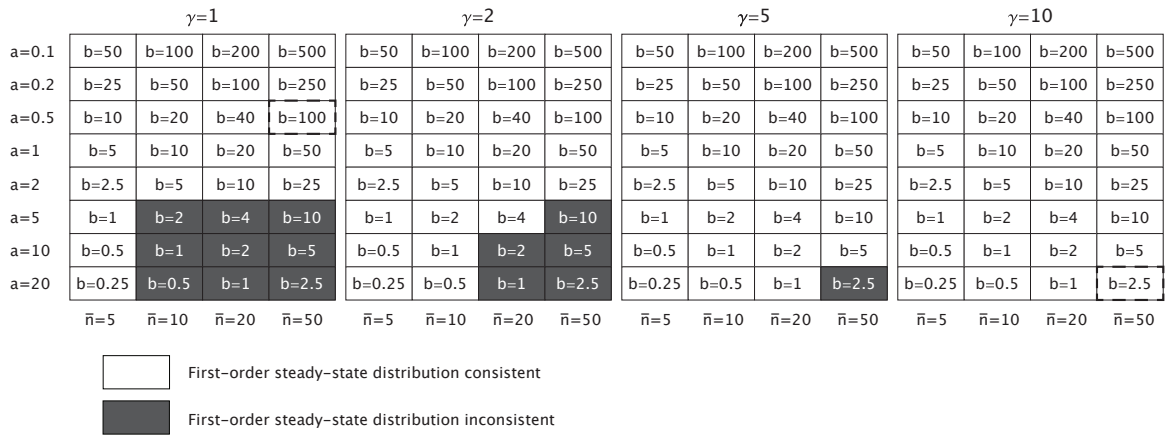
(a) Regime I ($\gamma = 10$). (b) Regime II ($\gamma = 1$).

(c) Regime I ($\gamma = 10$). (d) Regime II ($\gamma = 1$).

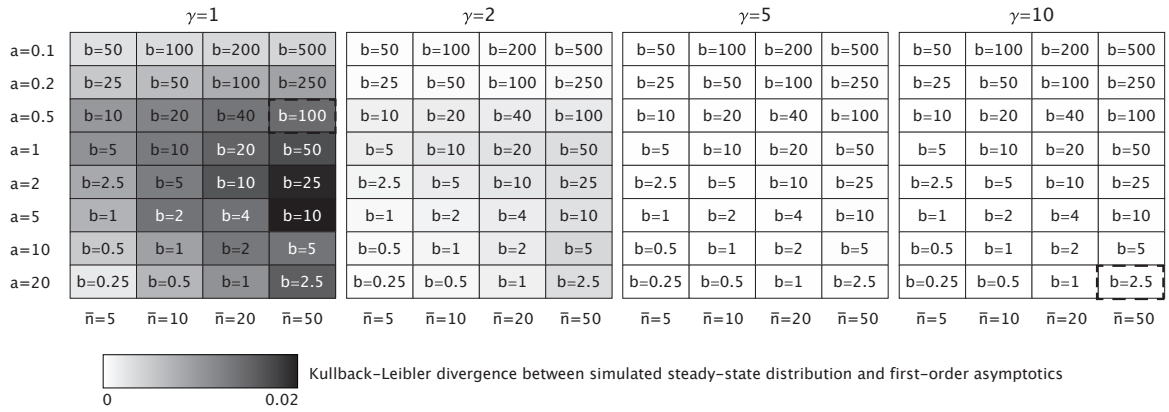
(e) Regime I ($\gamma = 10$). (f) Regime II ($\gamma = 1$).

(g) Regime I. (h) Regime II.

FIGURE 4. Comparison of the first-order steady-state distribution $P_n^\infty(\varepsilon)$ with stochastic simulation. Panels 4(a) and 4(b) illustrate representative mRNA time series in regime I ($a = 20$, $b = 2.5$) and regime II ($a = 0.5$, $b = 100$) with $\gamma = 10$ and $\gamma = 1$, respectively; the corresponding protein series are displayed in panels 4(c) and 4(d). We note pronounced protein bursts and low levels of mRNA expression in regime II, as opposed to regime I. In panels 4(e) and 4(f), we compare the resulting steady-state protein distributions (solid blue) with the asymptotics to zeroth order (dashed black) and to first order (solid black). The first-order correction clearly improves the predicted protein distribution in both parameter regimes. Finally, panels 4(g) and 4(h) show the Kullback-Leibler divergence of our perturbative approximation at steady state for varying values of γ: the asymptotics to first order outperforms the zeroth-order approximation when $1 < \gamma < 10$ in both regimes I and II, while there appears to be no significant difference for $\gamma > 10$.



(a) Consistency of P_n^∞.



(b) Accuracy of P_n^∞.

FIGURE 5. Consistency and accuracy of the first-order steady-state distribution $P_n^\infty(\varepsilon)$. While our $\mathcal{O}(\gamma^{-1})$-correction can lead to inconsistent (negative) protein probabilities in some parameter regimes, as seen in panel 5(a), these probabilities can be safely set to zero, as they are within the $\mathcal{O}(\gamma^{-2})$-error incurred by Equation (43). The Kullback-Leibler divergence between simulated steady-state distributions and the first-order asymptotics of P_n^∞ is shown in panel 5(b); comparison with the corresponding zeroth-order approximation reveals that the expansion in (43) is closer to simulation throughout. (Parameter regimes I and II, with $\gamma = 10$ and $\gamma = 1$, respectively, are indicated by dashed-edged squares.)

expansions for probability distributions do not necessarily satisfy the non-negativity conditions required of the 'full' distributions; cf. [24] for a recent application. (For completeness, we note that the leading-order approximation $P_n(\tau, 0)$ is, in fact, a distribution in its own right, by the normalisation condition in Equation (4).)

The accuracy of our first-order asymptotics is assessed in Figure 5(b): for each parameter triple (a, b, γ), we approximated the distribution at steady state by averaging over 10^2 simulation runs of length $400d_1^{-1}$. Sampling protein numbers in time steps of $2d_1^{-1}$, which we verified to exceed the typical de-correlation time of the system, we then calculated the Kullback-Leibler divergence between the resulting numerical distribution – truncated at its 99th percentile – and the expansion in Equation (43). (Here, ν_0 and ν_1 denote the rates of transcription and translation, respectively,

with d_0 and d_1 the respective degradation rates of mRNA and protein; recall Section 2.2.) These experiments also support our expectation that the first-order expansion for P_n^∞ significantly outperforms the zeroth-order approximation over a wide range of values of a and b provided $1 < \gamma < 10$ is at most 'moderately large.' Corresponding parameter regimes have been observed experimentally in a variety of organisms: in bacteria and yeast, representative γ-values often seem to range between about 1 and 10 [6, 50, 57], while moderately large values of b (between about 5 and 20) have been reported in [37, 53, 60]. (In fact, for budding yeast, experimental data presented in [50, Figure 4], suggests that γ is reliably larger than 1, but smaller than 10, in a vast majority of genes.) Regimes in which b is large, while a and γ are small, on the other hand, seem to be relevant for protein expression in mammalian cells, as reported in [45] for mouse fibroblasts.

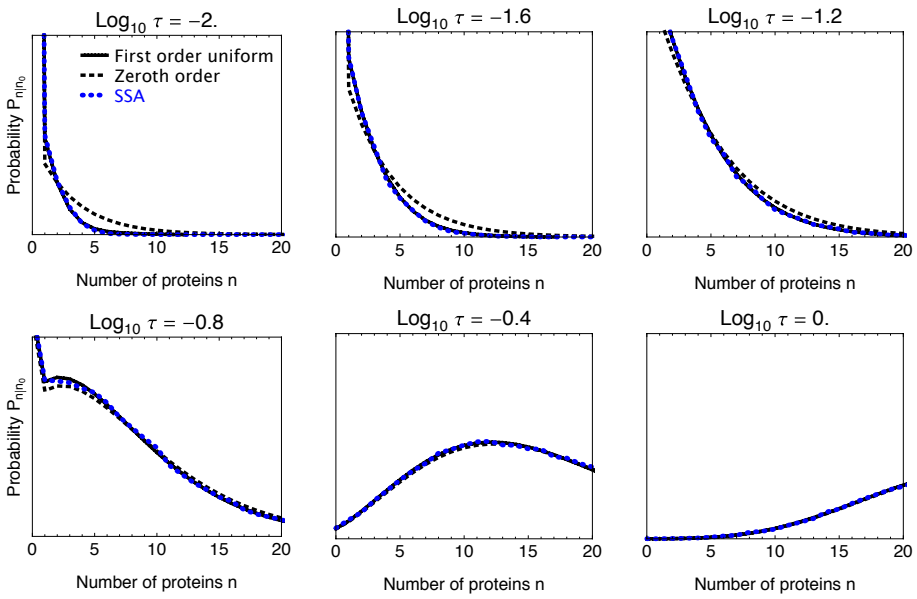
Finally, and as indicated already by Figure 4, the distinction between the first-order expansion for P_n^∞ and the zeroth-order asymptotics becomes insignificant in an increasing number of parameter regimes as γ is increased. In fact, our perturbative approximation, Equation (41), is most probably not convergent, being an asymptotic series in $\varepsilon(=\gamma^{-1})$. Hence, inclusion of the first-order correction in the expansion will not necessarily improve its accuracy uniformly in γ, *i.e.*, the optimal truncation for $\gamma = \mathcal{O}(10)$ may well involve the calculation of higher-order terms in the series. (The issue is almost certainly exacerbated by the fact that the coefficients in an asymptotic series can increase in magnitude through repeated differentiation, as is done here.) The corresponding optimal truncation point can potentially be determined by considering the Gevrey properties [4, 12] of the probability-generating function F. While we expect a 'truncation to the least term' to be optimal, in accordance with standard theory [32], a detailed investigation is beyond the scope of this work.

5.2. **Uniform (fast-slow) propagation.** By definition, the steady-state analysis presented thus far does not account for any fast (transient) dynamics: to evaluate the significance of the latter to first order in γ^{-1}, we considered the uniform marginal probability distribution of protein, as given in Proposition 4.3. Throughout, our focus is on the case where $n_0 = 0$ in $P_{n|n_0}$, which was studied in detail in Section 4.

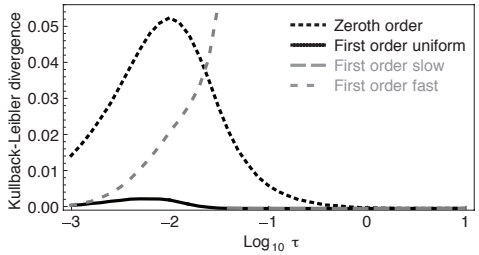
In Figure 6, a simulated marginal protein distribution is compared at different points in time both with the zeroth-order slow approximation and the uniform (composite) first-order distribution $P_n(\tau, t, \varepsilon)$ in regime I, for $\gamma = 100$; specifically, given initial numbers of mRNA and protein – which are, in our case, both assumed to be zero – are propagated in time $\tau = \varepsilon t$ according to Equation (44). As is seen in Figure 6(a), the uniform distribution achieves excellent agreement between simulation and asymptotics throughout. The corresponding Kullback-Leibler divergence is shown in Figure 6(b): again, our uniform first-order expansion for P_n clearly outperforms the zeroth-order approximation derived in [50]. In particular, the peak in the divergence which is seen to zeroth order for small times is eliminated, which implies that the contribution from the fast dynamics – *i.e.*, from the fast marginal distribution $P_n(t, \varepsilon)$ defined in (36) – cannot be neglected in regime I; cf. also [50, Figure 2C]. (The slight increase in divergence observed at about $\log_{10}\tau = -2$ is probably due to the fact that the corresponding τ-value roughly represents the point in time at which the dynamics 'switches' between the fast and the slow scales.)

However, the agreement between $P_n(\tau, t, \varepsilon)$ and the simulated probability distribution deteriorates for $\gamma = \mathcal{O}(10)$ in regime I (data not shown), which is due to a breakdown of the underlying assumption of a scale separation between mRNA and protein lifetimes in that regime for moderately large γ; recall Figure 3(a). (Mathematically, the uniform expansion in (44), while asymptotically correct to $\mathcal{O}(\varepsilon^2)$, can become numerically inaccurate unless ε is sufficiently small, as the contribution from $P_n(\tau, \varepsilon)$ therein contains higher-order terms in εt when considered on the fast t-scale.)

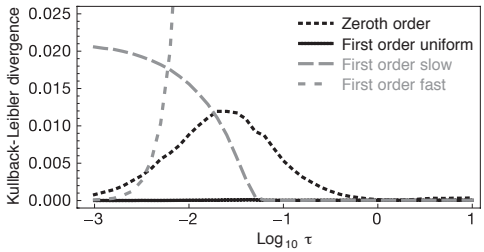
Analogous results have been obtained in regime II, where we have additionally assumed $n_0 = 50$, as in Figure 3 above. (The relevant expansion for $P_{n|n_0}$ can be found in Section B of the Online Supplement.) However, since the agreement between simulation and asymptotics is excellent – to the point of the respective curves being entirely indistinguishable down to $\gamma = \mathcal{O}(1)$ – we have opted not



(a) Regime I ($\gamma = 100$).



(b) Regime I ($\gamma = 100$).



(c) Regime II ($\gamma = 10$).

FIGURE 6. Comparison of the uniform (fast-slow) propagator $P_{n|n_0}(\tau, t, \varepsilon)$ with stochastic simulation. In panel 6(a), we compare the time evolution of the simulated probability distribution P_n (solid blue) in regime I for $\gamma = 100$ – and zero mRNAs and proteins initially – with the zeroth-order slow approximation (dotted) and the uniform distribution (solid). As shown in panel 6(b), the peak in the Kullback-Leibler divergence between simulation and asymptotics that is observed for the zeroth-order slow distribution alone (dashed) is eliminated by considering the uniform propagator (solid), which is equally superior to the first-order fast (dashed grey) and slow (long-dashed grey) asymptotics alone. Panel 6(c) illustrates the corresponding Kullback-Leibler divergence in regime II for $\gamma = 10$ and $n_0 = 50$; again, the uniform propagator outperforms the slow asymptotics, both to zeroth and to first order.

to include a detailed comparison here. The corresponding Kullback-Leibler divergence is presented in Figure 6(c); we note that, unlike in regime I, inclusion of the fast flow does not dramatically lower the divergence initially, as compared to the zeroth-order approximation alone. Still, Figure 6(c) also implies that the uniform first-order expansion for the propagator $P_{n|n_0}$ is superior to the zeroth-order asymptotics in regime II for larger times and, hence, that it can significantly improve the description of translational bursting in the two-stage model for gene expression studied here.

This reduced significance of the transient dynamics in regime II can be motivated geometrically: as is evident from Figure 2(b) above, the point $(u, v) = (-1, -1)$ at which the derivatives of the

generating function F are evaluated, by the definition of $P_{n|n_0}$, lies close to the critical manifold $\mathcal{S}_0$; in other words, the fast (layer) flow is almost instantaneous. (Correspondingly, at first order in ε, the slow asymptotics alone – while inferior – remains valid in regime II for short times, unlike in regime I; cf. again Figures 6(b) and 6(c). It seems plausible that a similar reasoning will apply whenever b is large (but finite) even if $\gamma = \mathcal{O}(1)$; here, we remark that γb – which, according to the reaction scheme in Figure 1(b), denotes the non-dimensionalised rate of translation – can possibly be interpreted as an effective perturbation parameter in such regimes. Consequently, the validity of the uniform propagator $P_{n|n_0}(\tau, t, \varepsilon)$ is then likely to extend to moderate γ-values for which the assumption of a scale separation between mRNA and protein lifetimes becomes blurred.

Finally, it can be shown that the contribution from the transient (fast) dynamics becomes less significant in both regimes I and II with increasing $\gamma > 100$, as is to be expected: as mRNA lifetime decreases, the system is ever more accurately described in terms of the slow protein dynamics alone (data not shown). A similar observation was made in Section 4 of [9], where the corresponding inner and outer solutions to a pair of reduced CMEs were matched numerically in their domain of overlap for varying values of $\varepsilon(= \gamma^{-1})$.

6. Discussion and Outlook

In this article, we have developed a systematic procedure for approximating the propagator probabilities $P_{mn|m_0n_0}$ – *i.e.*, the probabilities of observing m mRNAs and n proteins, given m_0 of the former and n_0 of the latter initially – in a standard two-stage model for unregulated gene expression [8, 50, 53]. Our approximation is given by a perturbative (asymptotic) series expansion in terms of ε, the inverse of the 'large' (dimensionless) ratio γ of the degradation rates of mRNA and protein; the presence of the 'small' parameter ε in the (singularly perturbed) 'characteristic' system of ordinary differential equations in (7) that is induced by the partial differential Equation (5) for the probability-generating function F allows for the application of Fenichel's geometric singular perturbation theory [17, 29], as shown in Section 3. Corresponding expansions for $P_{mn|m_0n_0}$ can then be obtained by repeated differentiation of the resulting asymptotic series for F. In Section 4, we have implemented our approach to first order in ε, under the additional simplifying assumption that no mRNAs or proteins are present initially. (A detailed study of the general case, with arbitrary initial numbers of mRNA and protein, can be found in the Online Supplement.) As in classical matched asymptotics [32], we have derived inner and outer expansions for F, and we have subsequently matched them in some overlap domain between the fast and the slow scales. Subsequently, we have formulated a composite fast-slow expansion which incorporates both components of the flow, and we have deduced a corresponding uniform approximation for the marginal protein distribution $P_n(\tau, t, \varepsilon)$. The resulting asymptotic procedure is constructive and can, of course, be extended to arbitrary order in ε; in the parlance of geometric singular perturbation theory, both the slow manifold $\mathcal{S}_\varepsilon$ and the fast foliation which yields its unstable manifold $\mathcal{W}^{\mathrm{u}}(\mathcal{S}_\varepsilon)$ will then need to be expanded to the appropriate order. (While the requisite generalisation is conceptually straightforward, the algebra involved in the calculation of higher-order terms may become too cumbersome to allow for the analytical evaluation of arbitrary derivatives of F.) Perhaps most significantly, in Section 5, we have shown that the uniform first-order asymptotics derived here yields a marked improvement over the zeroth-order slow approximation alone in a number of biologically relevant parameter regimes, including in two specific regimes considered in [50], both in the stationary and the time-dependent scenarios.

In sum, our results thus represent a three-fold extension of those reported in [8, 50], where the correction to F due to the fast (transient) flow in t was neglected and where, moreover, only the leading-order slow expansion (in τ) was derived under the assumption that no mRNA is present initially, whereas our asymptotics is valid for any choice of m_0 and n_0; see the Online Supplement for details. (We remark that these generalised initial conditions are still deterministic, as they correspond to Dirac-δ functions that are centred on m_0 and n_0, respectively; an extension of our

results to initial (stochastic) distributions can be accomplished by superposition, as was also done in [28, Section 2.2].) Finally, and in contrast to [9, 42], we have obtained closed-form asymptotic formulae for the propagator probabilities $P_{mn|m_0n_0}$, to first order in ε.

A substantial limitation of the procedure developed in this article is due to the fact that the method of characteristics [61] can, in general, only be applied to first-order partial differential equations. Our approach will hence *a priori* be restricted to reaction schemes that consist solely of first-order processes; in other words, they must not contain nonlinear kinetics. However, we emphasise that it does allow for the treatment of (auto)catalytic production as well as of 'splitting' reactions, both of which had to be excluded in [28]; recall the discussion at the end of Section 2. We remark that one could attempt a 'direct' perturbative solution of the partial differential equation for the probability-generating function F in the fully nonlinear case; potentially relevant asymptotic techniques can, *e.g.*, be found in [61, Chapter 9]. However, to the best of our knowledge, no analogue of geometric singular perturbation theory is available for infinite-dimensional dynamical systems, such as those defined by Equations (1) and (5). Hence, the applicability of the method of characteristics seems indispensable in the context of our approach.

A vast array of methodologies – both analytical and numerical – have been developed for the approximate solution of the CME; we refer the reader to [24] for a recent review, and comparison, of some of these. Thus, ideas from singular perturbation theory have been applied by Mastny *et al.* [36] directly to the CME, without the detour via the probability-generating function F; given the correspondence between the latter and the resulting probability distributions, their approach may be equivalent to ours when restricted to first-order reaction processes. A different perturbative technique, which was proposed in [35], does target generating functions; however, their perturbation parameter is given by the inverse of the (large) reaction volume. Finally, in recent work by Thomas *et al.* [54], an accurate stochastic reduction of biochemical networks with time-scale separation was developed under the additional assumption that the probability distribution allows for a Gaussian (or 'linear noise') approximation. It remains to be seen if and how these alternative approaches can be related to ours.

A conceptually straightforward extension of the results obtained in this article can be achieved by considering more complex, and, hence, potentially biologically more relevant models for regulated gene expression that account for additional stages in the process. Thus, a three-stage model has been suggested in which the promoter region of the gene of interest can transition between two states, one active and one inactive. Transcription can only occur if promoter is active, allowing, in particular, for transcriptional bursting which is characterised by transient periods of activity that lead to rapid mRNA synthesis; cf. [9, 37, 50] and the references therein. A geometric analysis of that generalised model will be the topic of an upcoming publication [41]: while we expect the framework developed in the present article to remain applicable, the added dimension may limit the extent to which analytical expressions can be obtained for the resulting probability distributions.

The three-stage model for gene expression can be augmented further; specifically, a four-stage model has been proposed [26, 40, 52] which includes a prolonged inactive ('refractory') state following periods of active transcription. Other extensions and generalisations [23, 43] may be feasible both biologically and mathematically, and will be explored in future research. Proceeding as in [25] – where the corresponding generating function was made to satisfy a pair of first-order equations – one could thus investigate the applicability of our geometric framework to models that include regulatory circuits, such as feedback from protein to DNA ('autoregulation') and to mRNA ('post-transcriptional regulation'). (While the CME is solved exactly in [25] for a reaction scheme that involves a regulatory feedback loop, such solutions are exceedingly rare in general, which makes the development of efficient approximation techniques, such as the one presented in this article, all the more relevant.) Finally, one could consider systems with variable reaction rates; the resulting multiple-parameter singular perturbation problems are bound to display interesting dynamics.

Since a direct comparison with well-converged stochastic simulation will arguably be intractable in such generalised models, a verification of the resulting asymptotics will almost certainly have to rely on sophisticated numerical techniques, such as those developed in [11, 13, 16, 58]. (As is evident from Section 5, a straightforward implementation of SSA on the basis of the StochKit package [52] was sufficient for the standard two-stage model considered here.)

Another natural-seeming generalisation concerns the reverse asymptotic regime of short protein lifetimes, in which the ratio γ of the degradation rates of mRNA and protein itself represents the small perturbation parameter. Preliminary analysis indicates the presence of lines of non-hyperbolic singularities in that regime, which precludes the application of geometric singular perturbation theory. A loss of normal hyperbolicity can often be remedied via the desingularisation technique known as 'blow-up' [15, 31]; however, since the large-γ asymptotics obtained in this article remains valid for $\gamma = \mathcal{O}(1)$ in a number of relevant parameter regimes, cf. Section 5.1, the practical value of an in-depth study of the small-γ regime may appear questionable [19].

On a related note, we emphasise that the asymptotic analysis performed in the present article assumes the system parameters a and b to be constant, *i.e.*, γ-independent, as the asymptotically small perturbation parameter is naturally given by the ratio γ^{-1} of the lifetimes of mRNA and protein in our case. A perturbation in a and b instead of, or in addition to, γ could be considered biologically realistic, and may in fact lead to refined asymptotics in specific parameter regimes. Again, the requisite analysis is left for future work.

A phenomenon that has attracted much recent attention is the transition from unimodal to bimodal behaviour in the stochastic modelling of biological networks; see again [25] and the references therein for details. The case of 'noise-induced bistability,' whereby a probability distribution can be bimodal even if the corresponding deterministic (rate) equations do not admit bistable solutions, is particularly intriguing in this context. While the two-stage model for gene expression does not allow for bimodality in the large-γ regime considered in this article, bimodal distributions have been documented in the generalised three-stage model [50]. It would be interesting to see if our approach can shed additional light on the transition to multimodality in that model and, in particular, on the question whether the former can occur at higher orders in the perturbation parameter.

In the long term, we envision that the geometric framework developed in this article – and, in particular, the proposed generalisation to multi-stage models for gene expression – will enable life scientists to obtain reliable estimates for molecular parameters from single-cell, time-lapse microscopy data of fluorescence reporters or fusion proteins; the availability of such data is increasing rapidly, particularly due to the adoption of microfluidic techniques [5, 34]. Experimentally, the expression of a given protein over time is monitored by repeated measurement of the light intensity of a fluorescent reporter. After segmentation and quantification of the cell signal, a time series for the fluorescent intensity is deduced in intervals of the measurement frequency.

Previous approaches for the inference of parameter likelihoods from time series data have relied on various *ad hoc* approximations. Thus, Harper *et al.* [26] introduced stochastic differential equations and the corresponding propagator probabilities to describe the protein dynamics during active and inactive periods of DNA expression. The likelihood of the complete time series was calculated by introducing reversible jumps between these two states; finally, an intricate Markov chain Monte Carlo model determined optimal parameters and jump points for given time series. By contrast, Suter and collaborators [52] employed a fully stochastic description of a two-stage model to describe gene expression in the active state. The transition between active and inactive states was assumed to be random, resulting in a so-called telegraph signal model. In a two-step procedure, the authors first estimated optimal parameters for each time series via an expectation-maximisation algorithm, and then found optimal paths for the states using a Viterbi algorithm [56].

Closed-form asymptotic formulae for the propagator probabilities, such as the ones derived in this article, would allow for the direct estimation of control parameters, as well as for the inference of

molecular mechanisms, from experimental data, thus reducing the dependence of the fitting process on fine-tuned and computationally intensive algorithms. In particular, the uniform (fast-slow) propagation introduced here should prove useful, as contributions from the fast (transient) dynamics may be significant in numerous biologically relevant parameter regimes. Ultimately, we thus hope that our framework will assist life scientists in comparing competing multiple-scale models for gene expression – and the error incurred by them – in a systematic, accurate, and efficient manner.

References

[1] H. Amann, *Ordinary differential equations. An introduction to nonlinear analysis*, translated from the German by Gerhard Metzen, de Gruyter Studies in Mathematics **13**, Walter de Gruyter & Co., Berlin, 1990.

[2] D. K. Arrowsmith and C. M. Place, *An introduction to dynamical systems*, Cambridge University Press, Cambridge, 1990.

[3] M. Abramowitz and I. A. Stegun (Eds.), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards, Applied Mathematics Series **55**, Dover Publications, New York, 1972.

[4] W. Balser, *From divergent power series to analytic functions. Theory and application of multisummable power series*, Lecture Notes in Mathematics **1582**, Springer-Verlag, Berlin, 1994.

[5] M.R. Bennett and J. Hasty, *Microfluidic devices for measuring gene network dynamics in single cells*, Nat. Rev. Genet. **10**(9), 628–638, 2009.

[6] J. Bernstein, A. Khodursky, P. Lin, S. Lin-Chao, and S. Cohen, *Global analysis of mRNA decay and abundance in* Escherichia coli *at single-gene resolution using two-color fluorescent DNA microarrays*, Proc. Natl. Acad. Sci. USA **99**(15), 9697-9702, 2002.

[7] T. Blasi, F. Buettner, M. Strasser, S. Linnarsson, C. Marr, and F. J. Theis, *Cell-cycle effects obscure true kinetics of gene expression in single cell transcriptomics data*, in preparation, 2013.

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

[9] P. Bokes, J. R. King, A. T. A. Wood, and M. Loose, *Multiscale stochastic modelling of gene expression*, J. Math. Biol. **65**(3), 493–520, 2012.

[10] C. G. Bowsher and P. S. Swain, *Identifying sources of variation and the flow of information in biochemical networks*, Proc. Natl. Acad. Sci. USA **109**(20), E1320–E1328, 2012.

[11] Y. Cao, D. T. Gillespie, and L. R. Petzold, *The slow-scale stochastic simulation algorithm*, J. Chem. Phys. **122**(1), 014116, 2005.

[12] M. Canalis-Durand, J. P. Ramis, R. Schäfke, and Y. Sibuya, *Gevrey solutions of singularly perturbed differential equations*, J. Reine Angew. Math. **518**, 95–129, 2000.

[13] S. L. Cotter, K. C. Zygalakis, I. G. Kevrekidis, and R. Erban, *A constrained approach to multiscale stochastic simulation of chemically reacting systems*, J. Chem. Phys. **135**(9), 094102, 2011.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, second edition, Wiley Series in Telecommunications and Signal Processing, Wiley-Interscience [John Wiley & Sons], Hoboken, New Jersey, 2006.

[15] F. Dumortier and R. Roussarie, *Canard cycles and center manifolds*, Mem. Amer. Math. Soc. **121**(577), 1996.

[16] R. Erban, I. G. Kevrekidis, D. Adalsteinsson, and T. C. Elston, *Gene regulatory networks: A coarse-grained, equation-free approach to multiscale computation*, J. Chem. Phys. **124**(8), 084106, 2006.

[17] N. Fenichel, *Geometric singular perturbation theory for ordinary differential equations*, J. Differential Equations **31**(1), 53–98, 1979.

[18] N. Friedman, L. Cai, and X. S. Xie, *Linking stochastic dynamics to population distribution: an analytical framework of gene expression*, Phys. Rev. Lett. **97**(16), 168302, 2006.

[19] J. García-Martínezm F. González-Candelas, and J. Pérez-Ortín, *Common gene expression strategies revealed by genome-wide analysis in yeast*, Genome Biol. **8**, R222, 2007.

[20] D. T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem. **81**(25), 2340–2361, 1977.

[21] D. T. Gillespie, *A rigorous derivation of the chemical master equation*, Physica A **188**(1–3), 404–425, 1992.

[22] D. T. Gillespie, *Deterministic limit of stochastic chemical kinetics*, J. Phys. Chem. **113**(6), 1640–1644, 2009.

[23] I. Golding, J. Paulsson, S.M Zawilski, and E. C. Cox, *Real-time kinetics of gene activity in individual bacteria*, Cell **123**(6), 1025–1036, 2005.

[24] R. Grima, *Construction and accuracy of partial differential equation approximations to the chemical master equation*, Phys. Rev. E **84**(5), 056109, 2011.

[25] R. Grima, D. Schmidt, and T. J. Newman, *Steady-state fluctuations of a genetic feedback loop: an exact solution*, J. Chem. Phys. **137**(3), 035104, 2012.

[26] C. V. Harper, B. Finkenstädt, D. J. Woodcock, S. Friedrichsen, S. Semprini, L. Ashall, D. G. Spiller, J. J. Mullins, D. A. Rand, J. R. E. Davis, and M. R. H. White, *Dynamic analysis of stochastic transcription cycles*, PLoS Biol. **9**(4), e1000607, 2011.
[27] G. Hek, *Geometric singular perturbation theory in biological practice*, J. Math. Biol. **60**(3), 347–386, 2010.
[28] T. Jahnke and W. Huisinga, *Solving the chemical master equation for monomolecular systems analytically*, J. Math. Biol. **54**(1), 1–26, 2007.
[29] C. K. R. T. Jones, *Geometric singular perturbation theory*, in Dynamical systems (Montecatini Terme, 1994), Lecture Notes in Math. **1609**, 44–118, Springer-Verlag, Berlin, 1995.
[30] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins, *Stochasticity in gene expression: from theories to phenotypes*, Nat. Rev. Genet. **6**(6), 451–464, 2005.
[31] M. Krupa and P. Szmolyan, *Extending geometric singular perturbation theory to non-hyperbolic points—fold and canard points in two dimensions*, SIAM J. Math. Anal. **33**(2), 286–314, 2001.
[32] P. A. Lagerstrom, *Matched Asymptotic Expansions: Ideas and Techniques*, Applied Mathematical Sciences **76**, Springer-Verlag, New York, 2010.
[33] I. J. Laurenzi, *An analytical solution of the stochastic master equation for reversible bimolecular reaction kinetics*, J. Chem. Phys. **113**(8), 3315–3322, 2000.
[34] J.C. Locke and M. B. Elowitz, *Using movies to analyse gene circuit dynamics in single cells*, Nat. Rev. Microbiol. **7**(5), 383–392, 2009.
[35] M. M. Mansour, C. Van den Broeck, G. Nicolis, and J. W. Turner, *Asymptotic properties of markovian master equations*, Ann. Phys. **131**(2), 283–313, 1981.
[36] E. A. Mastny, E. L. Haseltine, and J. B. Rawlings, *Two classes of quasi-steady-state model reductions for stochastic kinetics*, J. Chem. Phys. **127**(9), 094106, 2007.
[37] H. McAdams and A. Arkin, *Stochastic mechanisms in gene expression*, Proc. Natl. Acad. Sci. USA **94**(3), 814–819, 1997.
[38] P. Mukhopadhyay, *An Introduction to the Theory of Probability*, World Scientific Publishing Co., Singapore, 2011.
[39] J. D. Murray, *Mathematical Biology I: An Introduction*, Interdisciplinary Applied Mathematics **17**, Third edition. Springer-Verlag, New York, 2008.
[40] J. Peccoud and B. Ycart, *Markovian modeling of gene-product synthesis*, Theor. Popul. Biol. **48**(2), 222–234, 1995.
[41] N. Popović, C. Marr, and P. S. Swain, *A dynamical systems framework for parameter inference in stochastic gene expression*, in preparation, 2015.
[42] H. Pendar, T. Platini, and R. V. Kulkarni, *Exact protein distributions for stochastic models of gene expression using partitioning of Poisson processes*, Phys. Rev. E **87**(4), 042720, 2013.
[43] J. M. Raser and E. K. O'Shea, *Control of stochasticity in eukaryotic gene expression*, Science **304**(5678), 1811-1814, 2004.
[44] H. Risken, *The Fokker-Planck equation: methods of solution and applications*, second edition, Springer Series in Synergetics **18**, Springer-Verlag, Berlin, 1989.
[45] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, *Global quantification of mammalian gene expression control*, Nature **473**(7347), 337–342, 2011. Corrigendum: Nature **495**(7439), 126–127, 2013.
[46] L. A. Segel and M. Slemrod, *The quasi-steady-state assumption: a case study in perturbation*, SIAM Rev. **31**(3), 446–477, 1989.
[47] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, Studies in Nonlinearity, Addison-Wesley Publishing Company, Reading, Massachusetts, 1994.
[48] P. Swain, M. B. Elowitz, and E. D. Siggia, *Intrinsic and extrinsic contributions to stochasticity in gene expression*, Proc. Natl. Acad. Sci. USA **99**(20), 12795–12800, 2002.
[49] V. Shahrezaei and P. S. Swain, *The stochastic nature of biochemical networks*, Curr. Opin. Biotechnol. **19**, 369–374, 2008.
[50] V. Shahrezaei and P. S. Swain, *Analytical distributions for stochastic gene expression*, Proc. Natl. Acad. Sci. USA **105**(45), 17256–17261, 2008.
[51] K. R. Sanft, S. Wu, M. Roh, J. Fu, R. K. Lim, and L. R. Petzold, *StochKit2: software for discrete stochastic simulation of biochemical systems with events*, Bioinformatics **27**(17), 2457–2458, 2011.
[52] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef, *Mammalian genes are transcribed with widely different bursting kinetics*, Science **332**(6028), 472–474, 2011.
[53] M. Thattai and A. van Oudenaarden, *Intrinsic noise in gene regulatory networks*, Proc. Natl. Acad. Sci. USA **98**(15), 151588–151598, 2001.
[54] P. Thomas, A. V. Straube, and R. Grima, *The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions*, BMC Systems Biology **6**(1), 39, 2012.

[55] F. Verhulst, *Methods and Applications of Singular Perturbations: Boundary Layers and Multiple Timescale Dynamics*, Texts in Applied Mathematics **50**, Springer-Verlag, New York, 2005.
[56] A. Viterbi, *Error bounds for convolutional codes and an asymptotically optimum decoding algorithm*, IEEE Trans. IInform. Theory **13**(2),260–269, 1967.
[57] Y. Wang, C. Liu, J. Storey, R. Tibshirani, D. Herschlag, and P. Brown, *Precision and functional specificity in mRNA decay*, Proc. Natl. Acad. Sci. USA **99**(9), 5860–5865, 2002.
[58] E. Weinan, D. Liu, and E. Vanden-Eijnden, *Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates*, J. Chem. Phys. **123**(19), 194107, 2005.
[59] S. Wiggins, *Introduction to applied nonlinear dynamical systems and chaos*, second edition, Texts in Applied Mathematics **2**, Springer-Verlag, New York, 2003.
[60] J. Yu, J. Xiao, X. Ren, K. Lao, and X. Xie, *Probing gene expression in live cells, one protein molecule at a time*, Science **311**(5767), 1600–1603, 2006.
[61] E. Zauderer, *Partial differential equations of applied mathematics*, third edition, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], Hoboken, New Jersey, 2006.
[62] C. J. Zopf, K. Quinn, J. Zeidman, and N. Maheshri, *Cell-Cycle Dependence of Transcription Dominates Noise in Gene Expression*, PLoS Comput. Biol. **9**(7), 1–12, 2013.

University of Edinburgh, School of Mathematics and Maxwell Institute for Mathematical Sciences, James Clerk Maxwell Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom

Helmholtz Zentrum München - German Research Center for Environmental Health, Institute of Computational Biology, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany

University of Edinburgh, SynthSys–Systems and Synthetic Biology, CH Waddington Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JD, United Kingdom