

# Special Function Methods for Bursty Models of Transcription

Gennady Gorin<sup>1</sup> and Lior Pachter<sup>2,\*</sup>

<sup>1</sup>Division of Chemistry and Chemical Engineering, California Institute of Technology,  
Pasadena, CA, 91125

<sup>2</sup>Division of Biology and Biological Engineering & Department of Computing and  
Mathematical Sciences, California Institute of Technology, Pasadena, CA, 91125

\*Address correspondence to Lior Pachter ([lpachter@caltech.edu](mailto:lpachter@caltech.edu))

March 31, 2020

## Abstract

We explore a Markov model used in the analysis of gene expression, involving the bursty production of pre-mRNA, its conversion to mature mRNA, and its consequent degradation. We demonstrate that the integration used to compute the solution of the stochastic system can be approximated by the evaluation of special functions. Furthermore, the form of the special function solution generalizes to a broader class of burst distributions. In light of the broader goal of biophysical parameter inference from transcriptomics data, we apply the method to simulated data, demonstrating effective control of precision and runtime. Finally, we suggest a non-Bayesian approach to reducing the computational complexity of parameter inference to linear order in state space size and number of candidate parameters.

## Background

Recent improvements in transcriptomics and fluorescence microscopy methods have enabled the rapid and accurate quantification of mRNA on a transcriptome-wide scale with single-molecule precision [1–6]. Simultaneous advances in biophysical and statistical modeling have enabled the effective discrimination of gene expression models and the determination of physical parameters from these data. The estimation of underlying parameters relies on the ability to compute the distribution of molecules for a proposed set of parameters. The Chemical Master Equation (CME) is the standard modeling framework for low-copy single-molecule kinetics, treating such systems with Markov chains traversing state spaces of integer molecule counts [7–9]. However, solutions are available only for a relatively small set of models [8,10]. Furthermore, the existence of a closed-form solution does not guarantee its computational tractability.

Currently popular approaches to solving the CME can be roughly divided into three categories: simulation, matrix, and analytical methods. Simulation methods, such as the Gillespie stochastic simulation algorithm [11,12], are easily implemented and parallelized; the sample statistics of numerous realizations asymptotically approach the statistics of the underlying process, although

the speed of approach varies. Matrix methods, such as finite state projection [13] or multi-finite buffers [14], rely on matrix exponentiation or eigenvalue calculation to directly solve a truncation of the infinite-dimensional CME system; however, barring convenient symmetries, these methods require a characteristic running time of roughly  $O(n^3)$ , where  $n$  is the state space size. Finally, analytical methods directly solve the underlying system of ordinary differential equations (ODEs), e.g. using a generating function representation [8] or a convenient basis [15], and can be run in  $O(n)$  time.

Due to lower computational complexity, these analytical methods are highly relevant to the determination of biophysical parameters from high-dimensional, multimodal data, such as that available by modern transcriptomics and proteomics methods. Recent findings suggest that the use of joint data can provide substantial improvements to model and parameter estimation [16], motivating the development of more efficient solvers for the CME. Current chemistries can quantify spliced and unspliced mRNA molecules [3, 17], as well as surface proteins [18, 19]. The following multimodal models have analytical CME solutions, as well as drawbacks limiting their direct application to biological data.

1. Combination of Poissonian solutions [15, 20]: cannot be applied to proteomics, and does not explicitly model multistate genes.
2. Constitutive mRNA and protein production [21]: exact solution, but applies poorly to eukaryotic systems due to prevalence of multistate genes.
3. Telegraph mRNA and protein production [22, 23]: perturbative solution that relies on time-scale separation between mRNA and protein lifetimes, and inapplicable to a large fraction of eukaryotic genes.
4. Multi-state gene solutions with a single product [24–26]: exact solution, but does not provide information regarding downstream gene products. Current sequencing methods cannot be easily integrated with DNA accessibility testing.
5. Bursty mRNA production and isomerization [27]: exact solution, but relies on numerical integration and uses a fairly simple burst model.

A recent method [17] uses joint distributions of spliced and unspliced mRNA to perform short-time extrapolation on the cell landscape, motivating a more detailed treatment using stochastic biophysics. Motivated by that work, we propose a semi-analytical method for the evaluation of joint distributions resulting from the bursty transcription model [27], which describes a large fraction of mammalian genes [28–32]. Furthermore, we apply these models to parameter estimation, and discuss their applications to a set of burst size distributions that have not been previously solved.

## Methods

We follow previous literature [27] in implementing a Markov model for production, isomerization, and degradation of mRNA (**Figure 1a**). A single gene locus undergoes transcriptional bursting at a rate of  $k_i$ , producing  $B$  nascent mRNA transcripts (pre-mRNA) per burst, with  $P(B = \rho) = \alpha_\rho$ . The nascent transcripts are isomerized to mature mRNA.  $B$  is a random variable; if the underlying gene expression follows a two-state telegraph model with short bursts of finite size,  $B$  is drawn from

a geometric distribution [33]. The reactions are modeled as a Poisson processes with constant rates, which enables their representation using a homogeneous continuous-time Markov chain (CTMC).  $P(n, m, t)$ , the law of this CTMC model, yields the probability of finding  $n$  nascent and  $m$  mature molecules at time  $t$ .

The full set of CME ODEs is as follows:

$$\begin{aligned} \frac{dP(n, m, t)}{dt} = & k_i \left( \sum_{\rho=0}^n \alpha_{\rho} P(n - \rho, m, t) - P(n, m, t) \right) \\ & + \beta((n + 1)P(n + 1, m - 1, t) - nP(n, m, t)) \\ & + \gamma((m + 1)P(n, m + 1, t) - mP(n, m, t)) \end{aligned} \quad (1)$$

Using the probability-generating functions (PGF)  $G(x, y, t) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} x^n y^m P(n, m, t)$  and  $F(x) = \sum_{\rho=0}^{\infty} \alpha_{\rho} x^{\rho}$ , the CME recurrence relation may be cast into the form of a single partial differential equation (PDE):

$$\frac{\partial G}{\partial t} = k(F(x) - 1)G + \beta(y - x) \frac{\partial G}{\partial x} + \gamma(1 - y) \frac{\partial G}{\partial y} \quad (2)$$

subject to the initial condition  $G(x, y, 0) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} x^n y^m P(n, m, 0)$  and the normalization condition  $G(1, 1, t) = 1$ . Introducing the transformations  $x = 1 + u$ ,  $y = 1 + v$ , and  $G = e^{\phi}$  results in the following PDE:

$$\frac{\partial \phi}{\partial t} = k(M(u) - 1) + \beta(v - u) \frac{\partial \phi}{\partial u} + \gamma v \frac{\partial \phi}{\partial v} \quad (3)$$

such that  $M(u) = F(1 + u)$ . The solution of the PDE at time  $t$  is expressed by the following integral:

$$\phi(u, v, t) = k_i \int_0^t [M(U(s)) - 1] ds + \phi(U(t), V(t), 0) \quad (4)$$

Per the method of characteristics,  $V(s) = ve^{-\gamma s}$ ,  $U(s) = vfe^{-\gamma s} + (u - vf)e^{-\beta s}$  whenever  $\gamma \neq \beta$  and  $e^{-\gamma s}(u + \gamma vs)$  otherwise, where  $f \equiv \frac{\beta}{\beta - \gamma}$ . Finally, the PGF  $G$  is recovered by exponentiating  $\phi$ . We follow the approach of Bokes [21, 27] in evaluating the PGF for  $x, y$  around the complex unit circle, interpreting these values as the two-dimensional discrete Fourier transform (DFT), or characteristic function (CF) values, of the original probability distribution, and converting them to the discrete domain by application of the inverse discrete Fourier transform (IDFT). This method has time complexity  $O(\mathcal{N} \log \mathcal{N})$ , where  $\mathcal{N}$  is the state space size, such that  $\mathcal{N} = \max n \times \max m$  of interest (**Figure 1b**). For systems with relatively low copy numbers up to  $\approx 100$ , where CME modeling is necessary,  $N \sim 100 \times 100$ , requiring on the order of 10,000 evaluations of the integral  $\int_0^t [M(U(s)) - 1] ds$ .

The model with a geometric burst size distribution of mean  $b$  requires the evaluation of  $\int_0^t \frac{bU}{1-bU} ds$ . This integral does not have a closed-form solution, and must be treated using repeated numerical quadrature. However, an approximation to the integral can be computed by decomposing the integrand into an integrable power series. Any expression in the form of  $\frac{X}{1-X}$  is amenable to an expansion in powers of  $X = bU$ . In the region  $|X| > 1$ , the Laurent expansion  $-\sum_{i=0}^{\infty} X^{-i}$  is available. The intuitive choice of the complementary Taylor expansion  $\sum_{i=0}^{\infty} X^i$ , which is valid for  $|X| < 1$ , is inappropriate for integration across the boundary  $|X| = 1$ : the approximation

diverges and the integral of the expansion ceases to be identical to the original integral. Instead, we leverage the form of  $U$ , and note that  $Re(U) < 0$  for all nontrivial choices of  $u, v$ . Therefore, we utilize the Taylor expansion about  $-1$ , which is valid for  $|X + 1| < 2$ ; the form of the series is  $-\sum_{i=0}^{\infty} 2^{-i-1}(1 + X)^i - 1/2$ . As shown in the illustration of their shared domain of convergence (**Figure 1c**), it is possible to select the appropriate approximation based solely on a threshold for the real-valued  $|U|$ , which simplifies the computation.

Thus,  $X$  is decomposed into multiple approximation domains  $\{S_j\}$ , such that  $|X|$  evaluated at the boundary  $\partial S_j$  is  $\alpha$ , the threshold choice, and successive domains alternate in having  $|X|$  strictly greater or less than  $\alpha$  (**Figure 1d**). As discussed in the **Supplementary Note**, the form of  $U$  guarantees that  $|\{S_j\}| \leq 4$ ; at most two Laurent and two Taylor approximations are necessary. Examination of the expansions shows that both can be expressed as  $\sum_i \Omega_{j,i} U^i$ . If  $|U(s)| \geq \alpha \forall s \in S_j$ , the Laurent approximation is appropriate, and  $\Omega_{j,i} = -b^{-i}$ . For a Laurent order of approximation  $N_L, i \in \{0, -1, -2, \dots, -N_L\}$ . Conversely, if  $|U(s)| \leq \alpha \forall s \in S_j$ , the Taylor approximation is appropriate. For a Taylor order of approximation  $N_T$ , binomial expansion of  $(1 + X)^i$  yields  $\Omega_{j,i} = \sum_{k=i}^{N_T} b^i 2^{-k-1} \binom{k}{i}$ . The resulting approximation  $\sum_i \Omega_{j,i} U^i$  has  $i \in \{1, 2, \dots, N_T\}$ .

Finally, the full integrand  $\frac{bU}{1-bU}$  is approximately  $\sum_j \sum_i \Omega_{j,i} U^i$ . Therefore, the sought integral  $\int_0^t \frac{bU}{1-bU} ds$  can be computed using the truncated power series  $\sum_j \sum_i \Omega_{j,i} \int_{S_j} U^i ds$ , where each expansion is only integrated over its appropriate domain of convergence  $S_j$ . The details of computation are provided in the **Supplementary Note**, and the integrals  $\int U^i ds$  are given in **Table I**. Numerical routines to evaluate the exponential integral and the Gaussian hypergeometric function are readily available; however, they are not necessarily optimized for speed. We discuss the approximation schema used to make them practical for large-scale computation in the **Supplementary Note**.

Furthermore, the same approach can be used for other burst distributions. We consider a degenerate distribution (a gene locus that produces  $b$  transcripts per burst), a uniform distribution (a gene locus equally probably to produce any number of transcripts between  $a$  and  $b$ ) [34,35], and a shifted geometric distribution (a gene locus guaranteed to produce at least one transcript per burst, e.g. due to the inhibitor being removed by an advancing RNA polymerase). We find that the approximate solutions to these systems can also be expressed in the form  $\sum_j \sum_i \Omega_{j,i} \int_{S_j} U^i ds$ , as shown in **Table II**. Equivalently, as long as numerical routines are available to compute  $\int U^i ds$  for  $i \in \mathbb{Z}$ , a broad array of burst distributions can be computed simply by determining the appropriate integration limits (domains where the expansions converge) and computing the coefficients  $\Omega_{j,i}$ .

## Results and Discussion

We have presented an approximation for the Chemical Master Equation solution of bursty pre-mRNA production and its conversion to mature mRNA. We explored several burst distributions discussed in previous studies and explored an extension to a polymerase-inhibitor interaction model. The CME solutions can be found via the computation of  $\sum_i \Omega_i \int U^i ds$  for the finite-support distributions and  $\sum_j \sum_i \Omega_{j,i} \int_{S_j} U^i ds$  for the infinite-support distributions. The analytical solutions of  $\int U^i ds$  are given in **Table I**, whereas the combinatorial weights for specific burst distributions are given in **Table II**.

The series form of the solution enables the modulation of approximation order for computational facility (**Figure 2a**). The control of method precision and runtime motivates the development of

adaptive methods that determine a broad parameter domain using a low-fidelity approximation, then refine it using higher-order or quadrature-based methods.

The purpose of the current investigation is the development of a unified framework for the computation of CME solutions for a variety of burst models, as well as the determination of analytical solutions for the approximations. To maintain generality, we do not emphasize a particular implementation of the underlying special functions, but presuppose the availability of efficient implementations of the incomplete gamma and Gaussian hypergeometric function. Nevertheless, as a proof of concept, we develop a case study to benchmark the performance of the degenerate case  $\beta = \gamma$ . Furthermore, we discuss several considerations for implementation and evaluation of the special functions (**Supplementary note**).

In light of the motivating broader goal of parameter estimation, we use the algorithm to compute likelihood (Kullback-Leibler divergence) landscapes for simulated data [11] with a geometric burst size distribution and  $b = 19$ ,  $k_i = 2.5$ ,  $\beta = \gamma = 1$  (**Figure 2b**). The landscapes produced by the approximation method (shown for  $N_L = N_T = 7$ ) closely follow those produced via numerical integration. Repeating this analysis for a range of approximation orders allows benchmarking the method. Over the entire domain shown in **Figure 2b**, the quality of approximation can be easily controlled by modulating the Taylor approximation order (**Figure 2c**). The runtime is largely a function of the Laurent approximation order (**Figure 2d**), due to its explicit reliance on the computation of special functions. We particularly note that the commercial adaptive quadrature method used for benchmarking [36] provides poor control of runtime.

The procedure for regenerating the discrete distributions from generating functions presents certain problems for inference. As shown in **Figure 2a**, the result of the IDFT is not guaranteed to be a probability distribution; the IDFT enforces  $\sum_k \pi_k = 1$ , but does not enforce  $\pi_k \geq 0 \forall k$ . However, the properties of the Markov chain ostensibly guarantee that  $\pi_k \geq 0$ , with the inequality becoming strict at equilibrium. For the computation of divergence, we treat this problem in an ad hoc manner, by setting  $\pi_k \leq 0$  to a small float near machine epsilon. A natural, and potentially valuable, extension of this method is the development of transformations using non-negative, non-Fourier basis functions.

An alternative approach is available, and yields faster performance at the expense of interpretability in the Bayesian framework. Instead of computing Kullback-Leibler divergence in the probability domain, it is possible to compute a measure of distance between the characteristic functions of proposed and observed distributions, or even their corresponding cumulants (logarithms). This approach provides two advantages. Firstly, the roundoff and computational expense of repeated exponentiation and logarithm operations is eliminated. Secondly, the overall computational complexity of an inference procedure that uses the Fourier transform method and samples  $\mathcal{M}$  candidate parameters is  $O(\mathcal{M}\mathcal{N} \log \mathcal{N})$ . Performing the entire analysis in the Fourier domain requires only a single Fourier transform to determine the empirical characteristic function, reducing the computational complexity to  $O(\mathcal{N} \log \mathcal{N} + \mathcal{M}\mathcal{N})$ , equivalent to  $O(\mathcal{M}\mathcal{N})$  in the practical limit of large  $\mathcal{M}$ . This approach has been explored for use in goodness-of-fit testing and model selection [37,38], but only rarely for parameter inference [39]. However, we anticipate that the computational advantages may outweigh the incompatibility with Bayesian inference, similarly to the recent interest in using nonparametric Kolmogorov and Wasserstein distances for parameter inference [40–42]. Further, we note that optimization of the characteristic function uses information about the entire distribution, potentially overcoming identifiability issues observed using other computationally inexpensive non-Bayesian approaches, such as the method of moments [16]. Since characteristic function methods

have primarily been used for analysis of (continuous) stable distributions [39, 43, 44] and their performance for inference from discrete-valued random variable observations has not, to our knowledge, been systematically explored, signifying a substantial lacuna. Thus, this approach is a natural next step for optimizing inference from large datasets.

Our discussion of parameter estimation only touched upon inference from steady-state data, which is relevant for fixed-cell experiments that produce information about molecule distributions without a natural time coordinate, such as those available via scRNA-seq [45] and smFISH [46, 47]. However, experimental methods with temporal information are available [33, 48–51]. Given live-cell data, where cell identities are tracked across time, it is straightforward to extend this method to compute the probability of transitioning from an initial state to any other state, and thus compute the full likelihood of a time-series (**Supplementary Note: Addenda**). Repeating this process for all observed cells, assuming their trajectories are independent, and summing the log-likelihoods of their time-series yields a joint likelihood for the observations of the entire experiment [52–54]. Furthermore, given fixed-cell data, where only the *population-level statistics* are tracked across time, it is likewise straightforward to compute the probability of transitioning from one copy-number distribution to another, and use it for likelihood computation [55] (**Supplementary Note: Addenda**).

Finally, technical challenges in single-cell transcriptomics, such as sparsity of sampling in sequencing [56] and noise in fluorescence microscopy [57], have resulted in alternative competing explanations for qualitative features of observed biomolecule distributions, such as heavy-tailed laws [25, 27] and apparent dropouts [58–61]. We anticipate that intrinsic degeneracies, as well as aleatory effects, in mapping from a model parameter space to an observable space preclude the unambiguous identification of underlying biophysical schema: the presence of parameter equivalence classes, even in inference of simple models, is well-characterized [62–65]. Nevertheless, we also anticipate that the development of analytical solutions, as well as numerical solvers, for a diversity of transcriptional mechanisms, sampling behaviors, and multimodal observables will aid in making inference sufficiently robust for design and extrapolation. For example, as a natural extension, it is straightforward to calculate the laws for observed pre-mRNA and mRNA copy numbers by computing the distributions under an arbitrary sampling schema. This approach enables the natural integration of experimental noise in the same framework as the underlying transcriptional and molecular stochasticity, enabling the simultaneous inference of experimental and physiological parameters.

## Acknowledgments

The DNA, pre-mRNA, and mature mRNA used in **Figure 1a** are derivatives of the DNA Twemoji by Twitter, Inc., used under CC-BY 4.0. The routine for computing the Taylor approximation coefficient  $\Omega_{j,i}$  uses a function by Ben Barrowes [66], translated from the FORTRAN original by Zhang and Jin [67]. The routine for computing the Taylor series approximation to the exponential integral  $E_1(z)$  is a heavily modified version of a function by Ben Barrowes [66], translated from the FORTRAN original by Zhang and Jin [67]. G.G. and L.P. were partially funded by NIH U19MH114830.

## Data and Code Availability

The algorithm for the degenerate system is available at [https://github.com/pachterlab/GP\\_2020](https://github.com/pachterlab/GP_2020), along with MATLAB codes to reproduce **Figure 2** and **Supplementary Figure 1**.

## References

- [1] Xu H, Skinner SO, Sokac AM, Golding I. Stochastic Kinetics of Nascent RNA. *Physical Review Letters*. 2016;117(12):128101.
- [2] Lee S, Zhang AY, Su S, Ng AP, Holik AZ, Asselin-Labat ML, et al. Covering all your bases: incorporating intron signal from RNA-seq data. *Bioinformatics*; 2018. Available from: <http://biorxiv.org/lookup/doi/10.1101/352823>.
- [3] Shah S, Takei Y, Zhou W, Lubeck E, Yun J, Eng CHL, et al. Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH. *Cell*. 2018 Jul;174(2):363–376.e16. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867418306470>.
- [4] Eng CHL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature*. 2019 Apr;568(7751):235–239. Available from: <http://www.nature.com/articles/s41586-019-1049-y>.
- [5] Erhard F, Baptista MAP, Krammer T, Hennig T, Lange M, Arampatzi P, et al. scSLAM-seq reveals core features of transcription dynamics in single cells. *Nature*. 2019 Jul;571(7765):419–423. Available from: <http://www.nature.com/articles/s41586-019-1369-y>.
- [6] Wissink EM, Vihervaara A, Tippens ND, Lis JT. Nascent RNA analyses: tracking transcription and its regulation. *Nature Reviews Genetics*. 2019 Aug; Available from: <http://www.nature.com/articles/s41576-019-0159-6>.
- [7] Phillips R. *Physical biology of the cell*. 2nd ed. New York, NY: Garland Science; 2013.
- [8] Gardiner C. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. 3rd ed. Springer; 2004.
- [9] McQuarrie DA. Stochastic approach to chemical kinetics. *Journal of Applied Probability*. 1967 Nov;4(3):413–478. Available from: [https://www.cambridge.org/core/product/identifier/S002190020002547X/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S002190020002547X/type/journal_article).
- [10] Erdi P, Lente G. *Stochastic chemical kinetics: theory and (mostly) systems biological applications*. Springer complexity. New York: Springer; 2014. OCLC: ocn881399164.
- [11] Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*. 1976 Dec;22(4):403–434. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0021999176900413>.
- [12] Gillespie DT. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*. 1977 Dec;81(25):2340–2361. Available from: <https://pubs.acs.org/doi/abs/10.1021/j100540a008>.

- [13] Munsky B, Khammash M. The finite state projection algorithm for the solution of the chemical master equation. *The Journal of Chemical Physics*. 2006;124(4):044104.
- [14] Cao Y, Terebus A, Liang J. Accurate Chemical Master Equation Solution Using Multi-Finite Buffers. *Multiscale Modeling & Simulation*. 2016 Jan;14(2):923–963. Available from: <http://epubs.siam.org/doi/10.1137/15M1034180>.
- [15] Jahnke T, Huisinga W. Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of Mathematical Biology*. 2006 Dec;54(1):1–26. Available from: <http://link.springer.com/10.1007/s00285-006-0034-x>.
- [16] Munsky B, Li G, Fox ZR, Shepherd DP, Neuert G. Distribution shapes govern the discovery of predictive models for gene regulation. *Proceedings of the National Academy of Sciences*. 2018;115(29):7533–7538.
- [17] La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. *Nature*. 2018 Aug;560(7719):494–498. Available from: <http://www.nature.com/articles/s41586-018-0414-6>.
- [18] Stoeckius M. CITE-seq: Large scale simultaneous measurement of epitopes and transcriptomes in single cells; 2017. Available from: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE100866>.
- [19] Peterson VM, Zhang KX. The dynamics of cellular response to therapeutic perturbation using multiplexed quantification of the proteome and transcriptome at single-cell resolution; 2017. Available from: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE100501>.
- [20] Vastola JJ. Solving the chemical master equation for monomolecular reaction systems analytically: a Doi-Peliti path integral view. *arXiv:1911.00978 [q-bio]*. 2019 Nov;ArXiv: 1911.00978. Available from: <http://arxiv.org/abs/1911.00978>.
- [21] Bokes P, King JR, Wood ATA, Loose M. Exact and approximate distributions of protein and mRNA levels in the low-copy regime of gene expression. *Journal of Mathematical Biology*. 2012 Apr;64(5):829–854. Available from: <http://link.springer.com/10.1007/s00285-011-0433-5>.
- [22] Shahrezaei V, Swain PS. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*. 2008 Nov;105(45):17256–17261. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.0803850105>.
- [23] Veerman F, Marr C, Popovi N. Time-dependent propagators for stochastic models of gene expression: an analytical method. *Journal of Mathematical Biology*. 2018 Aug;77(2):261–312. Available from: <http://link.springer.com/10.1007/s00285-017-1196-4>.
- [24] Zhou T, Zhang J. Analytical Results for a Multistate Gene Model. *SIAM Journal on Applied Mathematics*. 2012 Jan;72(3):789–818. Available from: <http://epubs.siam.org/doi/10.1137/110852887>.
- [25] Ham L, Brackston RD, Stumpf MPH. Extrinsic noise and heavy-tailed laws in gene expression. *Systems Biology*; 2019. Available from: <http://biorxiv.org/lookup/doi/10.1101/623371>.



- [26] Ham L, Schnoerr D, Brackston RD, Stumpf MPH. Exactly solvable models of stochastic gene expression. *Systems Biology*; 2020. Available from: <http://biorxiv.org/lookup/doi/10.1101/2020.01.05.895359>.
- [27] Singh A, Bokes P. Consequences of mRNA Transport on Stochastic Variability in Protein Levels. *Biophysical Journal*. 2012 Sep;103(5):1087–1096. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0006349512007904>.
- [28] Bahar Halpern K, Tanami S, Landen S, Chapal M, Szlak L, Hutzler A, et al. Bursty Gene Expression in the Intact Mammalian Liver. *Molecular Cell*. 2015 Apr;58(1):147–156. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276515000507>.
- [29] Golan-Lavi R, Giacomelli C, Fuks G, Zeisel A, Sonntag J, Sinha S, et al. Coordinated Pulses of mRNA and of Protein Translation or Degradation Produce EGF-Induced Protein Bursts. *Cell Reports*. 2017 Mar;18(13):3129–3142. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2211124717303315>.
- [30] Dar RD, Razooky BS, Singh A, Trimeloni TV, McCollum JM, Cox CD, et al. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences*. 2012 Oct;109(43):17454–17459. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.1213530109>.
- [31] Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science*. 2011 Apr;332(6028):472–474. Available from: <http://www.sciencemag.org/cgi/doi/10.1126/science.1198817>.
- [32] Amrhein L, Harsha K, Fuchs C. A mechanistic model for the negative binomial distribution of single-cell mRNA counts. *Bioinformatics*; 2019. Available from: <http://biorxiv.org/lookup/doi/10.1101/657619>.
- [33] Golding I, Paulsson J, Zawilski SM, Cox EC. Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell*. 2005 Dec;123(6):1025–1036. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867405010378>.
- [34] Be'er S, Heller-Algazi M, Assaf M. Effect of reaction-step-size noise on the switching dynamics of stochastic populations. *Physical Review E*. 2016 May;93(5):052117. Available from: <https://link.aps.org/doi/10.1103/PhysRevE.93.052117>.
- [35] Kuwahara H, Arold ST, Gao X. Beyond initiation-limited translational bursting: the effects of burst size distributions on the stability of gene expression. *Integrative Biology*. 2015 Dec;7(12):1622–1632. Available from: <https://academic.oup.com/ib/article/7/12/1622/5199173>.
- [36] MATLAB R2019b. The MathWorks, Inc.; 2019. Available from: <https://www.mathworks.com/products/matlab.html>.
- [37] Lee S, Meintanis SG, Jo M. Inferential procedures based on the integrated empirical characteristic function. *AStA Advances in Statistical Analysis*. 2019 Sep;103(3):357–386. Available from: <http://link.springer.com/10.1007/s10182-018-00335-z>.

- [38] Jimnez-Gamero MD, Batsidis A, Alba-Fernndez MV. Fourier methods for model selection. *Annals of the Institute of Statistical Mathematics*. 2016 Feb;68(1):105–133. Available from: <http://link.springer.com/10.1007/s10463-014-0491-8>.
- [39] Bee M, Trapin L. A characteristic function-based approach to approximate maximum likelihood estimation. *Communications in Statistics - Theory and Methods*. 2018 Jul;47(13):3138–3160. Available from: <https://www.tandfonline.com/doi/full/10.1080/03610926.2017.1348523>.
- [40] Sommerfeld M, Munk A. Inference for empirical Wasserstein distances on finite spaces. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2018 Jan;80(1):219–238. Available from: <http://doi.wiley.com/10.1111/rssb.12236>.
- [41] Bernton E, Jacob PE, Gerber M, Robert CP. On parameter estimation with the Wasserstein distance. arXiv:170105146 [math, stat]. 2019 May;ArXiv: 1701.05146. Available from: <http://arxiv.org/abs/1701.05146>.
- [42] Gyrfi L, Vajda I, van der Meulen E. Minimum Kolmogorov distance estimates of parameters and parametrized distributions. *Metrika*. 1996 Dec;43(1):237–255. Available from: <http://link.springer.com/10.1007/BF02613911>.
- [43] Yu J. Empirical Characteristic Function Estimation and Its Applications. *Econometric Reviews*. 2004 Dec;23(2):93–123. Available from: <http://www.tandfonline.com/doi/abs/10.1081/ETC-120039605>.
- [44] Feuerverger A, McDunnough P. On the Efficiency of Empirical Characteristic Function Procedures. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1981 Sep;43(1):20–27. Available from: <http://doi.wiley.com/10.1111/j.2517-6161.1981.tb01143.x>.
- [45] Kolodziejczyk A, Kim JK, Svensson V, Marioni J, Teichmann S. The Technology and Biology of Single-Cell RNA Sequencing. *Molecular Cell*. 2015 May;58(4):610–620. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276515002610>.
- [46] Femino AM, Fay FS, Fogarty K, Singer RH. Visualization of Single RNA Transcripts in Situ. 1998;280:7.
- [47] Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nature Methods*. 2008 Oct;5(10):877–879. Available from: <http://www.nature.com/articles/nmeth.1253>.
- [48] Garcia H, Tikhonov M, Lin A, Gregor T. Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning. *Current Biology*. 2013 Nov;23(21):2140–2145. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0960982213011135>.
- [49] Specht EA, Braselmann E, Palmer AE. A Critical and Comparative Review of Fluorescent Tools for Live-Cell Imaging. *Annual Review of Physiology*. 2017 Feb;79(1):93–117. Available from: <http://www.annualreviews.org/doi/10.1146/annurev-physiol-022516-034055>.

- [50] Schofield JA, Duffy EE, Kiefer L, Sullivan MC, Simon MD. TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding. *Nature Methods*. 2018 Mar;15(3):221–225. Available from: <http://www.nature.com/articles/nmeth.4582>.
- [51] Herzog VA, Reichholf B, Neumann T, Rescheneder P, Bhat P, Burkard TR, et al. Thiol-linked alkylation of RNA to assess expression dynamics. *Nature Methods*. 2017 Dec;14(12):1198–1204. Available from: <http://www.nature.com/articles/nmeth.4435>.
- [52] Daigle BJ, Roh MK, Petzold LR, Niemi J. Accelerated maximum likelihood parameter estimation for stochastic biochemical systems. *BMC Bioinformatics*. 2012 Dec;13(1):68. Available from: <http://link.springer.com/10.1186/1471-2105-13-68>.
- [53] Corrigan AM, Tunnacliffe E, Cannon D, Chubb JR. A continuum model of transcriptional bursting. *eLife*. 2016 Feb;5:e13051. Available from: <https://elifesciences.org/articles/13051>.
- [54] Golightly A, Wilkinson DJ. Bayesian parameter inference for stochastic biochemical network models using particle Markov chain Monte Carlo. *Interface Focus*. 2011 Dec;1(6):807–820. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3262293/>.
- [55] Poovathingal SK, Gunawan R. Global parameter estimation methods for stochastic biochemical systems. *BMC Bioinformatics*. 2010;11(1):414. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-414>.
- [56] Grn D, Kester L, van Oudenaarden A. Validation of noise models for single-cell transcriptomics. *Nature Methods*. 2014 Jun;11(6):637–640. Available from: <http://www.nature.com/articles/nmeth.2930>.
- [57] Sgouralis I, Press S. An Introduction to Infinite HMMs for Single-Molecule Data Analysis. *Biophysical Journal*. 2017 May;112(10):2021–2029. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0006349517304435>.
- [58] Qiu P. Embracing the dropouts in single-cell RNA-seq data. *Bioinformatics*; 2018. Available from: <http://biorxiv.org/lookup/doi/10.1101/468025>.
- [59] Svensson V. Droplet scRNA-seq is not zero-inflated. *Bioinformatics*; 2019. Available from: <http://biorxiv.org/lookup/doi/10.1101/582064>.
- [60] Andrews T, Hemberg M. False signals induced by single-cell imputation. *F1000Research*. 2019;7:1740.
- [61] Li WV, Li JJ. An accurate and robust imputation method scImpute for single-cell RNA-seq data. *Nature Communications*. 2018 Dec;9(1):997. Available from: <http://www.nature.com/articles/s41467-018-03405-7>.
- [62] Weinreb C, Wolock S, Tusi BK, Socolovsky M, Klein AM. Fundamental limits on dynamic inference from single-cell snapshots. *Proceedings of the National Academy of Sciences*. 2018 Mar;115(10):E2467–E2476. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1714723115>.

- [63] Weber L, Raymond W, Munsky B. Identification of gene regulation models from single-cell data. *Physical Biology*. 2018 May;15(5):055001. Available from: <http://stacks.iop.org/1478-3975/15/i=5/a=055001?key=crossref.af8aa9f8142bbe6e351c9f9e3fca4287>.
- [64] Cinquemani E. Identifiability and Reconstruction of Biochemical Reaction Networks from Population Snapshot Data. *Processes*. 2018 Aug;6(9):136. Available from: <http://www.mdpi.com/2227-9717/6/9/136>.
- [65] Gorin G, Wang M, Golding I, Xu H. Stochastic simulation and statistical inference platform for visualization and estimation of transcriptional kinetics. *PLOS ONE*. 2020 Mar;15(3):e0230736. Available from: <https://dx.plos.org/10.1371/journal.pone.0230736>.
- [66] Barrowes B. Computation of Special Functions. MATLAB Central File Exchange; 2005. Available from: <https://www.mathworks.com/matlabcentral/fileexchange/6218-computation-of-special-functions>.
- [67] Zhang S, Jin J. *Computation of Special Functions*. 1st ed. New York: Wiley; 1996.

## Figures

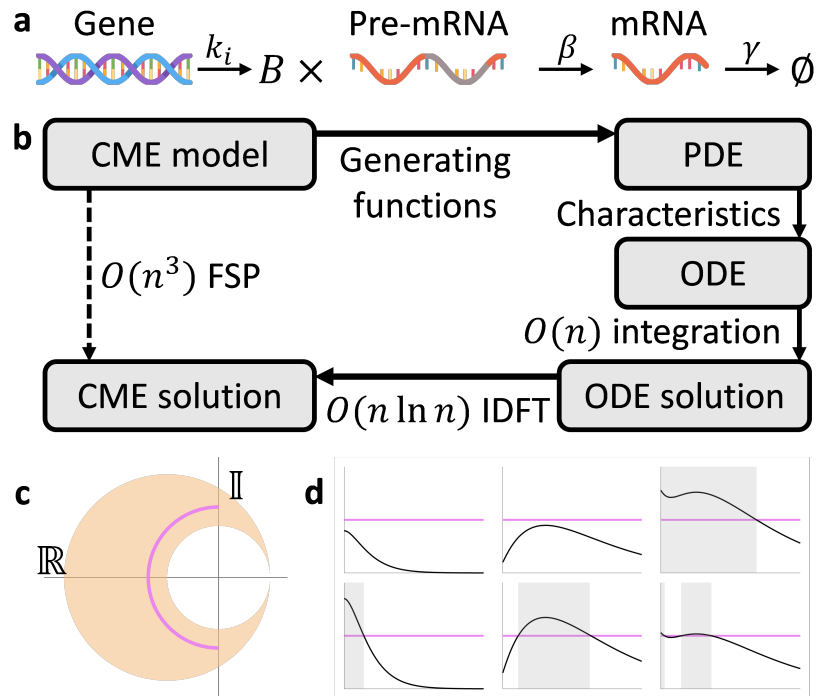


Figure 1: (a) Schema of modeled physiology ( $k_i$ : burst frequency;  $B$ : burst size drawn from discrete distribution on  $\mathbb{N}$ ;  $\beta$ : pre-mRNA splicing rate;  $\gamma$ : mRNA degradation rate). (b) Outline of the solution procedure. (c) Taylor/Laurent approximation criterion (orange: approximations common region of convergence; purple: threshold value of  $|U|$ ). (d) Sample shapes of  $|U|$  and their partitions (black curve:  $|U|$ ; purple: threshold value of  $|U|$ ; grey: Laurent approximation regions).

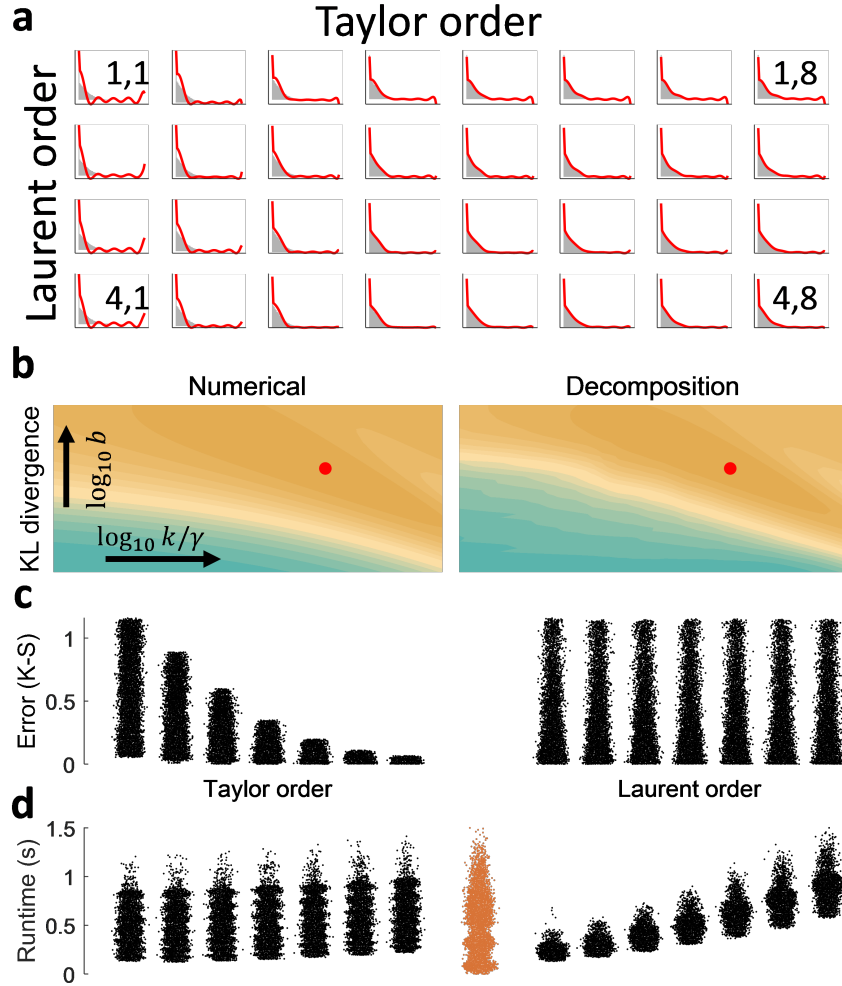


Figure 2: (a) Comparison of marginal mature mRNA copy-number distributions for a range of approximation orders ( $\#, \#$  tuple and subplot location: Laurent, Taylor approximation order; grey: histogram from  $10^5$  Gillespie simulations; red line: distribution calculated from approximation). (b) Likelihood landscape for a set of simulated steady-state data with  $\gamma = \beta$ , calculated over  $50 \times 50$  trial parameter combinations (numerical: quadrature-based computation; decomposition: expansion-based computation; abscissa:  $\log_{10} k_i/\gamma \in [-1, 1]$ ; ordinate:  $\log_{10} b \in (0, 2]$ ; orange: low divergence; teal: high divergence, red point: ground truth). (c) Kolmogorov-Smirnov error between quadrature- and expansion-based joint distributions for parameter sets in (b), calculated for combinations of Taylor and Laurent orders  $\in \{1, \dots, 7\} \times \{1, \dots, 7\}$  (black point: single parameter set; uniform jitter added). (d) Joint distribution calculation times, determined over the domain in (b) and approximation orders in (c) (black point: single parameter set computed using expansions; orange point: single parameter set computed using numerical quadrature; uniform jitter added).

## Tables

		$\gamma = \beta$	$\gamma \neq \beta$
$U$		$e^{-\gamma s}(u + \gamma vs)$	$U(s) = vfe^{-\gamma s} + (u - vf)e^{-\beta s}$
Taylor	$v = 0$	$-\frac{u^i}{i\gamma}e^{-i\gamma s}$	$-\frac{u^i}{i\beta}e^{-i\beta s}$
	$v \neq 0$	$-\frac{1}{\gamma^i}(\frac{v}{i})^i i! \sum_{j=0}^i \frac{j^i}{j!} \left[ (\frac{u}{v} + \gamma s)^j e^{-i\gamma s} \right]$	$\times \sum_{j=0}^i \binom{i}{j} (\frac{vf}{u-vf})^j \frac{1}{j-i\beta} e^{[j(\beta-\gamma)-i\beta]s}$
Laurent	$v = 0$	$\frac{u^{-i}}{i\gamma} e^{i\gamma s}$	$\frac{u^{-i}}{i\beta} e^{i\beta s}$
	$v \neq 0$	$-\frac{e^{-iu/v}}{\gamma v} (u + \gamma vs)^{1-i} E_i \left( -\frac{i}{v} (u + \gamma vs) \right)$	$\frac{(vf-u)^\rho (vf)^{-i-\rho}}{(\gamma-\beta)^\rho \left[ \frac{u-vf}{vf} e^{-(\beta-\gamma)s} \right]^\rho} \times {}_2F_1(i, -\rho; -\rho + 1; \frac{u-vf}{vf} e^{-(\beta-\gamma)s})$

**Table I.** Integrals of  $U^i$  for various approximations and levels of degeneration.

	Burst distribution			
	$b$ -step	Uniform	Geometric	Shifted geometric
$\frac{1}{k_i} \frac{\partial \phi}{\partial s}$	$(1 - U)^b - 1$	$\frac{1}{n} \sum_{i=a}^b (1 + U)^i - 1$	$\frac{bU}{1-bU}$	$\frac{bU}{1+(1-b)U}$
$\Omega_{j,i}$	$\binom{b}{i}$	$\frac{1}{n} [ \binom{b+1}{i+1} - \binom{a}{i+1} ]$	$b^i \sum_{k=i}^{N_T} \frac{1}{2^{k+1}} \binom{k}{i} - b^{-i}$	$b(b-1)^{i-1} \sum_{k=i}^{N_T} \frac{1}{2^{k+1}} \binom{k}{i} - b(b-1)^{-i-1}$
$i$	$1, \dots, b$	$1, \dots, b$	$1, \dots, N_T$ $0, -1, \dots, -N_L$	$1, \dots, N_T$ $0, -1, \dots, -N_L$
$U$	$\mathbb{C}$	$\mathbb{C}$	$ U  < \frac{1+\sqrt{3}}{2b}$ $ U  > \frac{1+\sqrt{3}}{2b}$	$ U  < \frac{1+\sqrt{3}}{2(b-1)}$ $ U  > \frac{1+\sqrt{3}}{2(b-1)}$

**Table II.** Integrand, expansion coefficients, summation indices, and expansion domain thresholds associated with approximating the CME solutions for four burst distributions.

# Special Function Methods for Bursty Models of Transcription

## Supplementary Note

Gennady Gorin, Lior Pachter

### Contents

<b>Burst generating functions and their expansions</b>	<b>2</b>
Geometric distribution, $\mu = b$ . . . . .	2
Shifted geometric distribution, $\mu = b$ . . . . .	2
Degenerate ( $b$ -step) . . . . .	3
Uniform on $[a, b]$ . . . . .	3
<b>Special function solutions to <math>\int U^i ds</math></b>	<b>3</b>
Degenerate case: $U(s; u, v) = e^{-\gamma s}(u + \gamma vs)$ . . . . .	3
Taylor expansion . . . . .	3
Laurent expansion . . . . .	4
Non-degenerate case: $U(s; u, v) = vfe^{-\gamma s} + (u - vf)e^{-\beta s}$ . . . . .	4
Taylor expansion . . . . .	4
Laurent expansion . . . . .	5
<b>Considerations for the numerical evaluation of <math>\phi(u, v)</math></b>	<b>5</b>
Numerical stability . . . . .	5
Binomial coefficients . . . . .	5
Overflow and underflow for large $b$ . . . . .	5
Special function evaluation . . . . .	6
Exponential integral . . . . .	6
<b>Domain decomposition: geometrically-distributed bursts</b>	<b>6</b>
Selection of threshold $\alpha$ . . . . .	7
Degenerate case: $\beta = \gamma$ . . . . .	7
General case: $\gamma_e \neq \gamma_c$ . . . . .	8
Root computation . . . . .	8
<b>Addenda</b>	<b>9</b>
Solutions for arbitrary initial conditions . . . . .	9
Negative binomial burst generating function . . . . .	9
<b>References</b>	<b>9</b>

arXiv:2003.12919v1 [stat.ME] 29 Mar 2020



## Burst generating functions and their expansions

Let  $\phi$  be the logarithm of the probability generating function (PDF)  $G$ . To determine  $\phi$ , it is necessary to integrate  $\frac{d\phi}{dt} = k_i(M(U(s)) - 1)$ , where  $M(U(s))$  is the factorial-moment generating function (FMGF) of the burst distribution. This function is represented as  $M(U) = F(1 + U)$  where  $F(x) = \sum_{\rho=0}^{\infty} P(B = \rho)x^\rho$  is the PGF of the burst distribution.

In what follows we use  $\mu$  to denote the mean burst size for a single-parameter burst distribution.

### Geometric distribution, $\mu = b$

As discussed by Singh and Bokes [1], the geometric distribution with mean burst size  $b$  has the probability mass function (PMF)  $P(B = \rho) = p(1 - p)^\rho$ , where  $p = \frac{1}{1+b}$  and  $\rho = 0, 1, 2, \dots$ . The resulting PGF is  $F(x) = E[x^\rho] = \sum_{\rho=0}^{\infty} P(B = \rho)x^\rho = \sum_{i=0}^{\infty} p(1 - p)^i x^i = p \sum_{i=0}^{\infty} [(1 - p)x]^i = \frac{p}{1 - (1-p)x}$ . This is exact for  $|x| \leq 1$  and extends to  $x \in \mathbb{C}$  by analytical continuation. Using the definition of  $p$ ,  $F(x) = \frac{1}{1+b} \frac{1}{1 - \frac{b}{1+b}x} = \frac{1}{1+b-bx}$ . The transformed PGF  $M(u) = F(1 + u) = \frac{1}{1+b-b(1+u)} = \frac{1}{1-bu}$ . Finally,  $M(u) - 1 = \frac{1}{1-bu} - 1 = \frac{bu}{1-bu}$ , which recapitulates previous work.

Defining  $X = bu$ , we have that  $M(u) - 1 = \frac{X}{1-X}$ . This expression has the well-known Laurent expansion  $-\sum_{i=0}^{\infty} \frac{1}{X^i}$  for all  $|X| > 1$ . Since  $b$  is real-valued,  $M(u) - 1 = \sum_{i=0}^{\infty} \Omega_i u^{-i}$ , where  $\Omega_i = -b^{-i}$ . This expansion can be truncated at order  $N_L$ , yielding  $M(u) - 1 \approx \sum_{i=0}^{N_L} \Omega_i u^{-i}$ .

Within the region  $|X+1| < 2$ , the function  $\frac{X}{1-X}$  has the Taylor expansion  $\sum_{i=1}^{\infty} \frac{(1+X)^i}{2^{i+1}} - \frac{1}{2}$ . Expanding the binomial yields  $\sum_{i=1}^{\infty} 2^{-i-1} \sum_{j=0}^i \binom{i}{j} X^j = \sum_{i=1}^{\infty} \frac{1}{2^{i+1}} \sum_{j=1}^i \binom{i}{j} X^j = \sum_{i=1}^{\infty} \frac{1}{2^{i+1}} \sum_{j=1}^i \binom{i}{j} (bu)^j$ .

This expansion can be truncated at order  $N_T$ , yielding  $M(u) - 1 \approx \sum_{i=1}^{N_T} \frac{b^i}{2} (\sum_{j=i}^{N_T} \frac{1}{2^j} \binom{j}{i}) u^i = \sum_{i=1}^{N_T} \Omega_i u^i$ , where  $\Omega_i = \frac{b^i}{2} \sum_{j=i}^{N_T} \frac{1}{2^j} \binom{j}{i}$ .

Equivalently,  $\Omega_i = b^i [1 - 2^{-N_T-2} \binom{N_T+1}{i} {}_2F_1(1, N_T + 2; -i + N_T + 2; \frac{1}{2})]$ .

### Shifted geometric distribution, $\mu = b$

A shifted geometric distribution with mean burst size  $b$  has the PMF  $P(B = \rho) = p(1 - p)^{\rho-1}$ , where  $p = \frac{1}{b}$  and  $\rho = 1, 2, 3, \dots$ . The resulting PGF over the relevant support is  $F(x) = E[x^\rho] = \sum_{\rho=1}^{\infty} P(B = \rho)x^\rho = \sum_{\rho=1}^{\infty} p(1 - p)^{\rho-1} x^\rho = p \sum_{\rho=0}^{\infty} (1 - p)^\rho x^{\rho+1} = px \sum_{\rho=0}^{\infty} [(1 - p)x]^\rho = \frac{px}{1 - (1-p)x}$ . Using the definition of  $p$ ,  $F(x) = \frac{x}{b} \frac{1}{1 - \frac{b-1}{b}x} = \frac{x}{b - (b-1)x}$ . The transformed PGF  $M(u) = F(1 + u) = \frac{1+u}{b - (b-1)(1+u)} = \frac{1+u}{b-b+1-bu+u} = \frac{1+u}{1+(1-b)u}$ . Finally,  $M(u) - 1 = \frac{1+u-1-(1-b)u}{1+(1-b)u} = \frac{bu}{1+(1-b)u}$ .

Within the region  $|u| > \frac{1}{|b-1|}$ , the function  $M(u) - 1$  has the Laurent expansion  $-\frac{b}{b-1} \sum_{i=0}^{\infty} \frac{1}{(b-1)^i u^i} = \sum_{i=0}^{\infty} \Omega_i u^{-i}$ , where  $\Omega_i = -\frac{b}{(b-1)^{i+1}}$ . This expansion can be truncated at order  $N_L$ , yielding  $M(u) - 1 \approx \sum_{i=0}^{N_L} \Omega_i u^{-i}$ .

Defining the intermediate variable  $a = \frac{1}{b-1}$  and noting  $a + 1 = \frac{b}{b-1}$ ,  $M(u) - 1 = \frac{b}{b-1} \frac{(b-1)u}{1+(1-b)u} = -\frac{b}{b-1} \frac{(1-b)u}{1+(1-b)u} = -\frac{b}{b-1} \frac{u}{u - \frac{1}{1-b}} = -\frac{b}{b-1} \frac{u}{u-a}$ . The Taylor expansion of  $\frac{u}{u-a}$  about  $-a$ , within the

domain  $|u + a| < 2|a|$ , is  $-\sum_{i=1}^{\infty} \frac{1}{2^{i+1}a^i} (a + u)^i + \frac{1}{2} = -\sum_{i=1}^{\infty} \frac{1}{2^{i+1}a^i} \sum_{j=0}^i \binom{i}{j} u^j a^{i-j} + \frac{1}{2}$   
 $= -\sum_{i=1}^{\infty} \frac{1}{2^{i+1}} \sum_{j=0}^i \binom{i}{j} u^j a^{-j} - \frac{1}{2} = -\sum_{i=1}^{\infty} \frac{1}{2^{i+1}} \sum_{j=1}^i \binom{i}{j} u^j a^{-j}$ . Therefore,  $M(u) - 1$   
 $= \frac{b}{2(b-1)} \sum_{i=1}^{\infty} \frac{1}{2^i} \sum_{j=1}^i \binom{i}{j} u^j (b-1)^j$ . This expansion can be truncated at order  $N_T$ , yielding

$M(u) - 1 \approx \sum_{i=1}^{N_T} \Omega_i u^i$ , where  $\Omega_i = \frac{b}{2} (b-1)^{i-1} \sum_{j=i}^{N_T} \frac{1}{2^j} \binom{j}{i}$ .

Equivalently,  $\Omega_i = b(b-1)^{i-1}[1 - 2^{-N_T-2} \binom{N_T+1}{i} {}_2F_1(1, N_T+2; -i+N_T+2, \frac{1}{2})]$ .

### Degenerate ( $b$ -step)

The degenerate burst distribution that yields  $b$  pre-mRNA products with every burst has PMF  $P(B = \rho) = \delta_{\rho b}$ , where  $\delta_{ij}$  is the Kronecker delta. The resulting PGF is  $F(x) = E[x^\rho] = \sum_{\rho=0}^{\infty} P(B = \rho)x^\rho = x^b$ . Therefore,  $M(u) - 1 = (1+u)^b - 1$ .

Expanding the binomial yields the expression  $M(u) - 1 = \sum_{i=0}^b \binom{b}{i} u^i - 1 = \sum_{i=1}^b \binom{b}{i} u^i = \sum_{i=1}^b \Omega_i u^i$ , where  $\Omega_i = \binom{b}{i}$ .

### Uniform on $[a, b]$

The uniform distribution on  $[a, b]$  has PMF  $P(B = \rho) = \frac{1}{n} I(\rho \in [a, b])$ , where  $I(\cdot)$  is the indicator function and  $n = b - a + 1$ . The resulting PGF over the relevant support is  $F(x) = E[x^\rho] = \frac{1}{n} \sum_{\rho=a}^b x^\rho$ . Therefore,  $M(u) - 1 = \frac{1}{n} \sum_{i=a}^b (1+u)^i - 1$ .

Expanding the binomial yields  $M(u) - 1 = \frac{1}{n} \sum_{i=a}^b \sum_{j=0}^i \binom{i}{j} u^j - 1 = \frac{1}{n} \sum_{i=a}^b \sum_{j=1}^i \binom{i}{j} u^j$ . Assuming  $a > 0$ , reversing the order of summation yields  $M(u) - 1 = \sum_{j=1}^b \frac{W_1(j) - W_2(j)}{n(j+1)} u^j$ , where  $W_1(j) = \frac{\Gamma(b+2)}{\Gamma(b-j+1)}$  for all  $j$  and  $W_2(j) = 0$  for  $j \in [a, b]$  and  $\frac{\Gamma(a+1)}{\Gamma(a-j)}$  otherwise. Therefore,  $M(u) - 1 = \sum_{i=1}^b \Omega_i u^i$ , where  $\Omega_i = \frac{W_1(i) - W_2(i)}{n(i+1)}$ . Equivalently,  $\Omega_i = \frac{1}{n} [\binom{b+1}{i+1} - \binom{a}{i+1}]$ .

### Special function solutions to $\int U^i ds$

As explored in the section ”**Burst generating functions and their expansions,**” the FCGFs of the burst models explored here can be represented in the common form  $M(u) - 1 \approx \sum_i \Omega_i u^i$ , where  $i \in \mathbb{Z}$ . Given  $U(s)$ , a characteristic solution representing the dynamics downstream of the gene locus, the integral  $\frac{\phi}{k_i} = \int [M(U(s)) - 1] ds$  can be approximated as  $\sum_i \Omega_i \int U(s)^i ds$  wherever the expansion holds. The determination of the appropriate domains is treated in the section ”**Domain Decomposition.**”

**Degenerate case:**  $U(s; u, v) = e^{-\gamma s}(u + \gamma v s)$

#### Taylor expansion

From standard identities [2], the Taylor expansion of order  $i \in \mathbb{Z}$  yields  $T_i(s; u, v) = \int e^{-i\gamma s}(u + \gamma v s)^i ds = -\frac{e^{iu/v} v^i}{\gamma^{i+1}} \Gamma(i+1, \frac{i}{v}(u + \gamma v s))$ , where  $\Gamma(a, z) = \int_z^\infty t^{a-1} e^{-t} dt$  is the upper incomplete Gamma function [2]. The evaluation of  $\Gamma(a, z)$  for arbitrary complex arguments is computationally nontrivial. However, for  $n \in \mathbb{N}$ ,  $\Gamma(i+1, z) = i! e^{-z} e_n(z)$ , where  $e_n(z) = \sum_{k=0}^n \frac{z^k}{k!}$ .

Therefore,  $T_i = -\frac{e^{iu/v} v^i}{\gamma^{i+1}} i! e^{-\frac{i}{v}(u + \gamma v s)} \sum_{k=0}^i \frac{1}{k!} (\frac{i}{v}(u + \gamma v s))^k = -\frac{e^{-i\gamma s} v^i}{\gamma^{i+1}} i! \sum_{k=0}^i \frac{i^k}{k!} (\frac{u}{v} + \gamma s)^k$ . The resulting definite integral from  $s_1$  to  $s_2$  is simply  $-\frac{v^i i!}{\gamma^{i+1}} \sum_{k=0}^i \frac{i^k}{k!} [(\frac{u}{v} + \gamma s_2)^k e^{-i\gamma s_2} - (\frac{u}{v} + \gamma s_1)^k e^{-i\gamma s_1}]$ , which is directly computable without the use of special function routines.

This expression for  $T_i$  is inappropriate for the degenerate case  $v = 0$ , corresponding to the marginal with respect to the pre-mRNA. In that case, the definite integral reduces to  $-\frac{u^i}{\gamma^i} [e^{-i\gamma s_2} - e^{-i\gamma s_1}]$ .

## Laurent expansion

For  $k = -i$ , where the order of Laurent expansion  $k \in \mathbb{Z}$ , the identity  $L_k(s; u, v) = \int e^{i\gamma s} (u + \gamma vs)^{-i} ds = -\frac{e^{-iu/v} v^{-i}}{\gamma(-i)^{1-i}} \Gamma(1-i, -\frac{i}{v}(u + \gamma vs))$  holds.

$L_k = \frac{e^{-iu/v} (-1)^i}{\gamma v^i i^{1-i}} \Gamma(1-i, -\frac{i}{v}(u + \gamma vs))$ . The direct computation of  $\Gamma(1-i, z)$  is computationally non-trivial, and finite power series expansions are unavailable for this purpose. However, the following identity holds:

$\Gamma(1-i, z) = z^{1-i} E_i(z)$ , where  $E_i(z)$  is the generalized exponential integral of order  $i$ . Further,  $E_i(z) = \frac{(-z)^{i-1}}{(i-1)!} E_1(z) + \frac{e^{-z}}{(i-1)!} \sum_{k=0}^{i-2} (i-k-2)! (-z)^k$ , where  $E_1(z) = \int_z^\infty \frac{e^{-t}}{t} dt$  is the first-order complex exponential integral.

Therefore,  $L_k = \frac{e^{-iu/v} (-1)^i}{\gamma v^i i^{1-i}} [-\frac{i}{v}(u + \gamma vs)]^{1-i} E_i(-\frac{i}{v}(u + \gamma vs)) = -\frac{e^{-iu/v}}{\gamma v} [u + \gamma vs]^{1-i} E_i(-\frac{i}{v}(u + \gamma vs))$ . Using the relation for  $E_i(z)$ ,  $E_i(-\frac{i}{v}(u + \gamma vs)) = \frac{1}{\Gamma(i)} [\frac{i}{v}(u + \gamma vs)]^{i-1} E_1(z) + \frac{1}{\Gamma(i)} \exp(\frac{i}{v}(u + \gamma vs)) \sum_{k=0}^{i-2} (i-k-2)! (\frac{i}{v}(u + \gamma vs))^k$ , where  $z = -\frac{i}{v}(u + \gamma vs)$ .

This yields  $L_k = -\frac{e^{-iu/v}}{\gamma v \Gamma(i)} (\frac{i}{v})^{i-1} E_1(z) - \frac{e^{i\gamma s}}{\gamma v \Gamma(i)} \sum_{k=0}^{i-2} \Gamma(i-k-1) (\frac{i}{v})^k (u + \gamma vs)^{k+1-i}$ . Finally, the definite integral  $L_k(s_2; u, v) - L_k(s_1; u, v)$  is

$$-\frac{e^{-iu/v}}{\gamma v \Gamma(i)} (\frac{i}{v})^{i-1} [E_1(z_2) - E_1(z_1)] - \frac{1}{\gamma v \Gamma(i)} \sum_{k=0}^{i-2} \Gamma(i-k-1) (\frac{i}{v})^k [(u + \gamma vs_2)^{k+1-i} e^{i\gamma s_2} - (u + \gamma vs_1)^{k+1-i} e^{i\gamma s_1}], \text{ where } z_l = -\frac{i}{v}(u + \gamma vs_l).$$

Equivalently, this is

$$-\frac{e^{-iu/v}}{\gamma v \Gamma(i)} (\frac{i}{v})^{i-1} [E_1(z_2) - E_1(z_1)] - \frac{e^{i\gamma s_1}}{\gamma v \Gamma(i)} \sum_{k=0}^{i-2} \Gamma(i-k-1) (\frac{i}{v})^k [(u + \gamma vs_2)^{k+1-i} e^{i\gamma(s_2-s_1)} - (u + \gamma vs_1)^{k+1-i}].$$

As in the Taylor expansion, this expression is inappropriate for the degenerate case  $v = 0$ . In that case, the definite integral reduces to  $\frac{1}{iu^i \gamma} (e^{i\gamma s_2} - e^{i\gamma s_1})$ .

**Non-degenerate case:**  $U(s; u, v) = v f e^{-\gamma s} + (u - v f) e^{-\beta s}$

As throughout,  $f$  is defined as  $\frac{\beta}{\beta - \gamma}$ .

## Taylor expansion

$$\begin{aligned} \text{For a Taylor expansion of order } i \in \mathbb{Z}, T_i(s; u, v) &= \int [v f e^{-\gamma s} + (u - v f) e^{-\beta s}]^i ds \\ &= \int (A e^{-\gamma s} + B e^{-\beta s})^i ds. \text{ Expanding the binomial term yields } T_i \\ &= \int \sum_{j=0}^i \binom{i}{j} (A e^{-\gamma s})^j (B e^{-\beta s})^{i-j} ds = \int \sum_{j=0}^i \binom{i}{j} A^j B^{i-j} e^{-j\gamma s} e^{-(i-j)\beta s} ds \\ &= \int \sum_{j=0}^i \binom{i}{j} A^j B^{i-j} e^{[-j\gamma - (i-j)\beta]s} ds = \sum_{j=0}^i \binom{i}{j} A^j B^{i-j} \int e^{[-j\gamma - (i-j)\beta]s} ds \\ &= \sum_{j=0}^i \binom{i}{j} A^j B^{i-j} \frac{1}{-j\gamma - (i-j)\beta} e^{[-j\gamma - (i-j)\beta]s} = \sum_{j=0}^i \binom{i}{j} (v f)^j (u - v f)^{i-j} \frac{1}{-j\gamma - (i-j)\beta} e^{[-j\gamma - (i-j)\beta]s}. \end{aligned}$$

The resulting definite integral, which can be computed directly, is:

$$\begin{aligned} &\sum_{j=0}^i \binom{i}{j} (v f)^j (u - v f)^{i-j} \frac{1}{-j\gamma - (i-j)\beta} [e^{[-j\gamma - (i-j)\beta]s_2} - e^{[-j\gamma - (i-j)\beta]s_1}] \\ &= (u - v f)^i \sum_{j=0}^i \binom{i}{j} (\frac{v f}{u - v f})^j \frac{1}{j(\beta - \gamma) - i\beta} [e^{[j(\beta - \gamma) - i\beta]s_2} - e^{[j(\beta - \gamma) - i\beta]s_1}]. \\ &= \frac{1}{\beta - \gamma} (u - v f)^i \sum_{j=0}^i \binom{i}{j} (\frac{v f}{u - v f})^j \frac{1}{j - i f} [e^{[j(\beta - \gamma) - i\beta]s_2} - e^{[j(\beta - \gamma) - i\beta]s_1}] \end{aligned}$$

In the degenerate case  $v = 0$ , the antiderivative is simply  $\int [u e^{-\beta s}]^i ds = u^i \int e^{-i\beta s} ds = -\frac{u^i}{i\beta} e^{-i\beta s}$ , with the definite integral  $-\frac{u^i}{i\beta} [e^{-i\beta s_2} - e^{-i\beta s_1}]$ .

## Laurent expansion

For a Laurent expansion of order  $i \in \mathbb{Z}$ ,  $L_i(s; u, v) = \int [v f e^{-\gamma s} + (u - v f) e^{-\beta s}]^{-i} ds$   
 $= \int (A e^{-\gamma s} + B e^{-\beta s})^{-i} ds$ . Considering the integrand:

$l(s; u, v) = (A e^{-\gamma s} + B e^{-\beta s})^{-i} = \frac{1}{(A e^{-\gamma s} + B e^{-\beta s})^i} = \frac{1}{(A e^{-\gamma s} + B e^{-\beta s})^i} = \frac{1}{(1 + \frac{B}{A} e^{-(\beta-\gamma)s})^i A^i e^{-i\gamma s}}$ . Defining  
 $z = -\frac{B}{A} e^{-(\beta-\gamma)s}$ , this yields  $l = \frac{1}{(1-z)^i A^i e^{-i\gamma s}}$ . Furthermore,  $e^{-i\gamma s} = [e^{-(\beta-\gamma)s}]^\rho$  for  $\rho = \frac{-(\beta-\gamma)}{-i\gamma} = \frac{\beta-\gamma}{i\gamma}$ . Therefore,  $z^\rho = (-\frac{B}{A} e^{-(\beta-\gamma)s})^\rho = (-\frac{B}{A})^\rho e^{-i\gamma s}$  and  $A^i e^{-i\gamma s} = A^i (-\frac{B}{A})^{-\rho} z^\rho$ . This, in turn,  
yields  $l = (-\frac{B}{A})^\rho A^{-i} (1-z)^{-i} z^{-\rho}$ .

Using variable substitution,  $\int_{s_1}^{s_2} l(s; u, v) ds = \int_{z_1}^{z_2} l(z; u, v) dz$ , where  $z_i = -\frac{B}{A} e^{-(\beta-\gamma)s_i}$ .  $\frac{dz}{ds} = -(\beta - \gamma)z$ , yielding  $dz = -(\beta - \gamma)^{-1} z^{-1} ds$ .

The solution to  $\int_{z_1}^{z_2} l(z; u, v) dz$  is given by the incomplete beta function  $B(z; \mu_1, \mu_2)$  [2]:

$r = (-\frac{B}{A})^\rho A^{-i} (\gamma - \beta)^{-1} [B(z_2; \mu_1, \mu_2) - B(z_1; \mu_1, \mu_2)]$ ,  $\mu_1 = -\rho$ , and  $\mu_2 = 1 - i$ . This solution is, in principle, exact. However, methods for the computation of  $B(z; \mu_1, \mu_2)$  for arbitrary complex arguments  $z$  are unavailable [2], presumably due to the ubiquity of the function in the field of statistical computation, which only requires evaluation of  $B(x; \mu_1, \mu_2)$  for  $x$  on the real line. Therefore, it is more practical to use the Gaussian hypergeometric function representation  $B(z; \mu_1, \mu_2)$   
 $= \frac{z^{\mu_1}}{\mu_1} {}_2F_1(1 - \mu_2, \mu_1; \mu_1 + 1; z)$ , yielding the desired result:

$L_i(s; u, v) = (-\frac{B}{A})^\rho A^{-i} (\gamma - \beta)^{-1} \rho^{-1} [z_2^{-\rho} {}_2F_1(i, -\rho; -\rho + 1; z_2) - z_1^{-\rho} {}_2F_1(i, -\rho; -\rho + 1; z_1)]$ .

In the degenerate case  $v = 0$ , the antiderivative is simply  $\int [(u e^{-\beta s})^{-i} ds = u^{-i} \int e^{i\beta s} ds = \frac{1}{i\beta u^i} e^{i\beta s}$ ,  
with the definite integral  $\frac{1}{i\beta u^i} [e^{i\beta s_2} - e^{i\beta s_1}]$ .

If a robust evaluation routine for  ${}_2F_1$  is available, it is also straightforward to compute the Taylor terms  $T_i(s; u, v) = L_{-i}(s; u, v)$  without explicitly performing the summation described above.

## Considerations for the numerical evaluation of $\phi(u, v)$

### Numerical stability

#### Binomial coefficients

The expressions for  $\Omega_i$  described in the section "Burst generating functions and their expansions" require the computation of binomial coefficients. For the geometric and shifted geometric distributions of burst sizes, which have binomial coefficients on the order of  $\binom{N_T}{n}$  (for small integer  $N_T$ ), calculating the coefficients *via* the log-gamma function  $\ln \Gamma$  is recommended, per  $\binom{m}{n} = \exp(\ln \Gamma(m + 1) - \ln \Gamma(n + 1) - \ln \Gamma(m - n + 1))$ . For the degenerate and uniform distributions (with modest to large  $b$ ), calculating the coefficients *via*  $\ln \Gamma$  is required for practical implementation.

#### Overflow and underflow for large $b$

Although the approximation  $\sum_i \Omega_i \int U(s)^i ds$  converges to the desired generating function  $\frac{\phi}{k_i}$ , the series as written is not ideal for numerical evaluation by the independent computation of  $\Omega_i$  and  $\int U(s)^i ds$ . In particular, the Taylor coefficients for the geometric and shifted geometric distribution scale with  $b^i$  and  $(b-1)^i$  respectively. Similarly, their Laurent coefficients scale with  $b^{-i}$  and  $(b-1)^{-i}$ . Since  $b$  may be fairly large (up to 300) for mammalian systems [3], the computation of its powers can produce overflow or underflow problems for Taylor and Laurent coefficients, respectively.

Conversely, the Taylor integrals for the degenerate system scale with  $\frac{1}{i^i}$ , and the Laurent ones scale with  $i^i$ .

In case of the degenerate system, issues can be ameliorated somewhat by computing  $\frac{\Omega_i}{i^i}$  and  $i^i \int U(s)^i ds$ , which yields a  $(\frac{b}{i})^i$  term and balances the growth behavior.

It is not clear whether an analogous approach is available for the non-degenerate case, although solutions based on Stirling's approximation ( $i! \sim i^i \sqrt{2\pi i} e^{-i}$ ) may be viable for larger  $i$ .

## Special function evaluation

### Exponential integral

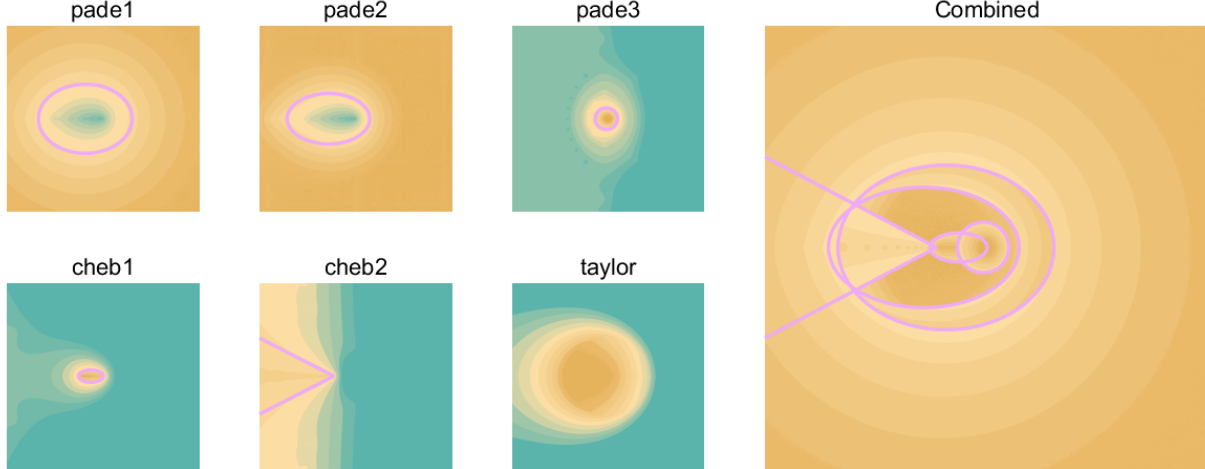
The solution for the Laurent expansion of the degenerate system, which only involves the evaluation of a parameter-free exponential integral  $E_1(z)$ , lends itself to optimization. We found that the following combination of approximations yields no more than  $10^{-8}$  relative error with respect to the MATLAB function `expint` throughout the entire complex plane [4]. In this enumeration,  $x \equiv \text{Re}(z)$  and  $y \equiv \text{Im}(z)$ . For benchmarking,  $\tau = T_M/T_A$ , where  $T_M$  is the runtime of the built-in function and  $T_A$  is the runtime of each approximation routine. The functions were run over a  $1000 \times 1000$  point grid for  $x, y \in [-35, 35]$ .

1. Exterior region  $[x/17+0.3824]^2+(y/13)^2 > 1$ : 6th order Padé approximation (Luke p.110-112 [5], table 4.3): `pade1`.  $\tau = 45.7$ .
2. Exterior region  $[((x+10)/15)^2+(y/9.5)^2 > 1]$ : 10th order Padé approximation (Luke p.110-112 [6], table 4.3): `pade2`.  $\tau = 40.4$ .
3. Elliptic region  $[((x+0.65)/4.05)^2+(y/4)^2 < 1]$ : 10th order Padé approximation (Luke p.107-108 [6], table 4.2): `pade3`.  $\tau = 33.8$ .
4. Elliptic region  $[((x+4.5)/4.5)^2+(y/2.3)^2 < 1]$ : 20th order Chebyshev approximation (Luke p.104 [6], table 4.1): `cheb1`.  $\tau = 15.4$ .
5. Radial region  $[x < -8 \cap |y| < (-x-8)0.5294]$ : 20th order Chebyshev approximation (Luke p.105 [6], table 4.1): `cheb2`.  $\tau = 12.6$ .
6. Annular region (all other values of  $z$ ): 55th order series approximation (Zhang and Jin [7]): `taylor`.  $\tau = 7.3$ .

The resulting performance is shown in Supplementary Figure 1. Teal denotes error above  $10^{-8}$ ; brown denotes error below  $10^{-8}$ ; the purple lines are the approximation regions. The combined approximation produced  $\tau = 24.1$ , with maximum error of  $10^{-7.9}$ .

## Domain decomposition: geometrically-distributed bursts

The Taylor and Laurent approximations are only valid over some domains. For example, the Laurent approximation to  $\frac{b\tilde{u}(s)}{1-b\tilde{u}(s)}$  is only valid for  $|b\tilde{u}(s)| \equiv |U(s)| > 1$ . We split the domain of integration into multiple sub-domains, and integrate each one separately. The domain boundaries are values of  $s$  ( $s_1, s_2, s_3 \dots$ ) where  $|U(s)| = \alpha$ , such that  $\alpha$  is within the domain of convergence for both Taylor and Laurent approximations.



Supplementary Figure 1: Exponential integral approximations.

The number of domains depends upon the number of times the function  $|U(s)|$  crosses  $\alpha$ . From integration from 0 to  $t$ , where the function crosses  $\alpha$  at  $s_1, s_2, s_3, \dots, s_i$ , the number of domains is  $i + 1$ . By the mean value theorem, for differentiable  $|U(s)|$ ,  $i \leq j + 1$ , where  $j$  is the number of points such that  $\frac{d|U(s)|}{ds} = 0$ .

In the present section, we demonstrate that *at most* four domains are necessary; equivalently,  $\frac{d|U(s)|}{ds} = 0$  for at most *two*  $s \geq 0$ .

### Selection of threshold $\alpha$

For the geometric distribution of burst sizes, the region of convergence for the Taylor expansion is a circle of radius 2 centered at -1, while the region of convergence for the Laurent expansion is a circle of radius 1 centered at 0. For the relevant case of  $Re(U) \leq 0$ , the minimum distance between the boundaries of the regions of convergence is achieved at  $Re(U) = 0$ . The corresponding region of convergence is  $(i, \sqrt{3}i)$ . As shown in **Figure 1c**, we select  $\alpha$  equidistant from the two regions' boundaries, as a minimax strategy. This threshold for  $|U|$  corresponds to  $\alpha = \frac{1+\sqrt{3}}{2b}$ .

For the shifted geometric distribution, the region of convergence for the Taylor expansion is a circle of radius  $\frac{2}{b-1}$  centered at  $-\frac{1}{b-1}$ , while the region of convergence for the Laurent expansion is a circle of radius  $\frac{1}{b-1}$  centered at 0. The minimum distance between the boundaries is again achieved at  $Re(U) = 0$ , defining the region of convergence  $(\frac{1}{b-1}i, \frac{\sqrt{3}}{b-1}i)$ . Therefore, the minimax optimal threshold corresponds to  $\alpha = \frac{1+\sqrt{3}}{2(b-1)}$ .

### Degenerate case: $\beta = \gamma$

In the degenerate case where the export rate is equal to the degradation rate,  $U(s) = e^{-\gamma s}(u + \gamma v s)$ .  
 $|U|^2 = e^{-2\gamma s}(u + \gamma v s)(\bar{u} + \gamma \bar{v} s)$   
 $= e^{-2\gamma s}(u\bar{u} + \gamma^2 v\bar{v} s^2 + \gamma[u\bar{v} + \bar{u}v]s)$   
 $= e^{-2\gamma s}(\gamma^2 |v|^2 s^2 + \gamma[u\bar{v} + \bar{u}v]s + |u|^2)$ .

Note that  $|U(s)| = 0$  only if  $U(s) = 0$ . This may occur if  $u = v = 0$  or  $u = s = 0$  which are trivial

edge cases. The alternative case  $u = -\gamma vs$  is not relevant since  $\gamma, s > 0$  and both  $u$  and  $v$  are only evaluated in the negative real half-plane. Thus, all nontrivial combinations of  $u, v, s$  yield a strictly positive  $|U(s)|$  whose behavior is one-to-one with that of  $|U(s)|^2$ .

$$\begin{aligned} & \text{Differentiation yields } \frac{d|U|^2}{ds} \\ &= e^{-2\gamma s} [2\gamma^2 |v|^2 s + \gamma [u\bar{v} + \bar{u}v]] - 2\gamma e^{-2\gamma s} [\gamma^2 |v|^2 s^2 + \gamma [u\bar{v} + \bar{u}v] s + |u|^2] \\ &= e^{-2\gamma s} [(-2\gamma^3 |v|^2) s^2 + (2\gamma^2 |v|^2 - 2\gamma^2 [u\bar{v} + \bar{u}v]) s + (\gamma [u\bar{v} + \bar{u}v] - 2\gamma |u|^2)] \\ &= 2\gamma e^{-2\gamma s} [(-\gamma^2 |v|^2) s^2 + (\gamma |v|^2 - \gamma [u\bar{v} + \bar{u}v]) s + (\frac{1}{2} [u\bar{v} + \bar{u}v] - |u|^2)] \\ &= 2\gamma e^{-2\gamma s} [As^2 + Bs + C], \end{aligned}$$

where  $A = -\gamma^2 |v|^2 \leq 0$ ,  $B = \gamma (|v|^2 - [u\bar{v} + \bar{u}v])$ , and  $C = \frac{1}{2} [u\bar{v} + \bar{u}v] - |u|^2$

For  $u = 0$  or  $v = 0$ , explicit solutions to  $|U| = \alpha$  can be directly computed using the Lambert W function. For  $u, v \neq 0$ , a generalized Lambert W function is required for an analytical solution [8, 9]. Although approximation methods are available [10, 11], they are not generally necessary here. Instead, a numerical scheme can be used to compute coarse estimates for roots constrained by the zeroes of  $\frac{d|U|^2}{ds}$ . As long as the roots are within  $\Delta\alpha$  of  $\alpha$ , such that  $(\alpha - \Delta\alpha, \alpha + \Delta\alpha)$  is within the radius of convergence for both expansions, the approximation is valid.

### General case: $\gamma_e \neq \gamma_c$

An analogous demonstration may be provided for the general case where the export rate is not equal to the degradation rate.

As elsewhere,  $f \equiv \frac{\beta}{\beta - \gamma}$ .  $U(s) = v f e^{-\gamma s} + (u - v f) e^{-\beta s}$ .

$$\begin{aligned} |U|^2 &= [v f e^{-\gamma s} + (u - v f) e^{-\beta s}] [\bar{v} f e^{-\gamma s} + (\bar{u} - \bar{v} f) e^{-\beta s}] \\ &= |v f|^2 e^{-2\gamma s} + |u - v f|^2 e^{-2\beta s} + [v f (\bar{u} - \bar{v} f) + \bar{v} f (u - v f)] e^{-(\gamma + \beta) s} \\ &= |v f|^2 e^{-2\gamma s} + |u - v f|^2 e^{-2\beta s} + f [v (\bar{u} - \bar{v} f) + \bar{v} (u - v f)] e^{-(\gamma + \beta) s} \\ &= |v f|^2 e^{-2\gamma s} + |u - v f|^2 e^{-2\beta s} + f [v \bar{u} - |v|^2 f + \bar{v} u - |v|^2 f] e^{-(\gamma + \beta) s} \\ &= f^2 |v|^2 e^{-2\gamma s} + |u - v f|^2 e^{-2\beta s} + f ([\bar{u}v + u\bar{v}] - 2|v|^2 f) e^{-(\gamma + \beta) s}. \end{aligned}$$

$$|u - v f|^2 = (u - v f)(\bar{u} - \bar{v} f) = (|u|^2 + f^2 |v|^2 - f[\bar{u}v + u\bar{v}]).$$

Therefore,  $|U|^2 = f^2 |v|^2 e^{-2\gamma s} + (|u|^2 + f^2 |v|^2 - f[\bar{u}v + u\bar{v}]) e^{-2\beta s} + f([\bar{u}v + u\bar{v}] - 2|v|^2 f) e^{-(\gamma + \beta) s}$ .

$$\begin{aligned} & \text{Differentiation yields } \frac{d|U|^2}{ds} \\ &= -2\gamma f^2 |v|^2 e^{-2\gamma s} - 2\beta (|u|^2 + f^2 |v|^2 - f[\bar{u}v + u\bar{v}]) e^{-2\beta s} - (\gamma + \beta) f([\bar{u}v + u\bar{v}] - 2|v|^2 f) e^{-(\gamma + \beta) s} \\ &= 0 \text{ whenever } C e^{-2\gamma s} + A e^{-2\beta s} + B e^{-(\gamma + \beta) s} = 0, \end{aligned}$$

Where  $C = -2\gamma f^2 |v|^2 < 0$ ,  $A = -2\beta (|u|^2 + f^2 |v|^2 - f[\bar{u}v + u\bar{v}]) < 0$ , and

$B = -(\gamma + \beta) f([\bar{u}v + u\bar{v}] - 2|v|^2 f) \in \mathbb{R}$ . Multiplying both sides by  $e^{2\gamma s} (> 0)$ ,

$$0 = C + A e^{-2(\beta - \gamma)s} + B e^{-(\beta - \gamma)s}$$

$$= A e^{2(\gamma - \beta)s} + B e^{(\gamma - \beta)s} + C$$

Making the substitution  $z = e^{(\gamma - \beta)s}$ ,

$$0 = Az^2 + Bz + C$$

Note that two roots exist, and the solutions  $z^*$  may be converted back to the time domain per  $s^* = \frac{1}{\gamma - \beta} \ln z^*$ . Only real positive extrema  $s^*$  are pertinent to the domain partition; all others may be discarded.

### Root computation

Apart from the degenerate cases described above, where the roots can be computed *via* logarithms or the Lambert W function, the determination of domain boundaries requires the use of numerical

routines. Specifically, we perform 20 iterations of the Newton-Raphson method to find locations where  $|U|^2 = \alpha^2$ . If the routine fails to converge, we switch to the bisection method and search in the domain bracketed by extrema of  $|U|^2$ ; as discussed above, the locations of these extrema are known analytically. If the root is not bracketed (i.e., the rightmost extremum at the location  $s^*$  has a value  $y^* > \alpha^2$ ), we use the heuristic bracket  $(s^*, s^* + 10/\min(\beta, \gamma))$ , which suffices due to the fast (exponential) decay of the function  $|U|^2$ .

## Addenda

### Solutions for arbitrary initial conditions

As described by Singh and Bokes [1], the generating function at time  $t$  for an arbitrary initial distribution is given by  $\phi(u, v, t) = k_i \int_0^t (M(U(s)) - 1) ds + \phi(U(t), V(t), 0)$ , where  $\phi(u, v, 0)$  is the factorial cumulant generating function of the initial distribution. If  $n(t=0) = n_0$  and  $m(t=0) = m_0$ , the corresponding generating function is  $G(x, y, t=0) = \sum_n \sum_m x^n y^m P(n, m, t=0) = x^{n_0} y^{m_0}$ . Its logarithm  $\phi(u, v, 0)$  is simply  $n_0 \ln(1+u) + m_0 \ln(1+v)$ . To find the value of this contribution, all that remains is to compute  $U(t)$  and  $V(t)$ , where  $U(s)$  is given above and  $V(s) = ve^{-\gamma s}$ . This method can be trivially extended to an arbitrary initial distribution of molecules by summing over the appropriate  $n$  and  $m$  and using the observed initial copy-number distribution as the initial condition  $P(n, m, t=0)$ .

### Negative binomial burst generating function

A negative binomial (NB) burst distribution has  $P(B = \rho) = \binom{\rho+a-1}{\rho} (1-p)^a p^\rho$ , where  $p = \frac{b}{1+b}$  and the mean burst size is  $ba$ . The resulting generating function is  $F(x) = (\frac{1-p}{1-px})^a$ . Substituting  $M(u) = F(1+u)$  yields  $M(u) - 1 = \frac{1}{(1-bu)^a} - 1$ , which reduces to the geometric distribution for  $a = 1$ . Defining  $X = (bu)^{-1}$ , the first term can be rewritten as  $\frac{X^a}{X^a(1-X^{-1})^a} = \frac{X^a}{(X-1)^a}$ . This form affords the Taylor expansion in the form of  $\sum_{i=0}^{\infty} X^{i+a} (-1)^{i-a} \binom{-a}{i}$ , and an analogous Laurent expansion. Therefore, it is, in principle, possible to compute the CME solutions through a series of integrals of the form  $\int U^{i+a} ds$ ,  $i \in \mathbb{Z}$ , which are given by  $\Gamma(i+a, \cdot)$  for the degenerate case and  ${}_2F_1(i+a, \cdot; \cdot; \cdot)$  otherwise.

Previous discussions of the NB model treat it as a natural generalization of the geometric model, but do not motivate its use from physical arguments. Therefore, we do not develop the solution to the model in further detail.

## References

- [1] Singh A, Bokes P. Consequences of mRNA Transport on Stochastic Variability in Protein Levels. *Biophysical Journal*. 2012 Sep;103(5):1087–1096. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0006349512007904>.
- [2] Abramowitz M, Stegun I, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 9th ed. United States National Bureau of Standards; 1970.



- [3] Dar RD, Razooky BS, Singh A, Trimeloni TV, McCollum JM, Cox CD, et al. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences*. 2012 Oct;109(43):17454–17459. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.1213530109>.
- [4] MATLAB R2019b. The MathWorks, Inc.; 2019. Available from: <https://www.mathworks.com/products/matlab.html>.
- [5] Luke YL. *Mathematical Functions and their Approximations*. New York, San Francisco, London: Academic Press; 1975.
- [6] Luke YL. *The Special Functions and Their Approximations, Vol 2*. London ; New York: Academic Press; 1969.
- [7] Zhang S, Jin J. *Computation of Special Functions*. 1st ed. New York: Wiley; 1996.
- [8] Scott TC, Mann RB. General Relativity and Quantum Mechanics: Towards a Generalization of the Lambert W Function. *arXiv:math-ph/0607011*. 2006 Sep;ArXiv: math-ph/0607011. Available from: <http://arxiv.org/abs/math-ph/0607011>.
- [9] Maignan A, Scott TC. Fleshing out the generalized Lambert W function. *ACM Communications in Computer Algebra*. 2016 Aug;50(2):45–60. Available from: <http://dl.acm.org/citation.cfm?doid=2992274.2992275>.
- [10] Scott TC, Fee G, Grotendorst J. Asymptotic series of generalized Lambert W function. *ACM Communications in Computer Algebra*. 2014 Jan;47(3/4):75–83. Available from: <http://dl.acm.org/citation.cfm?doid=2576802.2576804>.
- [11] Castle P. Taylor series for generalized Lambert W functions. *arXiv:180109904 [math]*. 2018 Jan;ArXiv: 1801.09904. Available from: <http://arxiv.org/abs/1801.09904>.