Research

Author for correspondence:
Christopher E. Miles
e-mail: chris.miles@uci.edu

Incorporating spatial diffusion into models of bursty stochastic transcription

Christopher E. Miles

Department of Mathematics, Center for Complex Biological Systems, University of California, Irvine, CA, USA

CEM, 0000-0001-5494-403X

The dynamics of gene expression are stochastic and spatial at the molecular scale, with messenger RNA (mRNA) transcribed at specific nuclear locations and then transported to the nuclear boundary for export. Consequently, the spatial distributions of these molecules encode their underlying dynamics. While mechanistic models for molecular counts have revealed numerous insights into gene expression, they have largely neglected now-available subcellular spatial resolution down to individual molecules. Owing to the technical challenges inherent in spatial stochastic processes, tools for studying these subcellular spatial patterns are still limited. Here, we introduce a spatial stochastic model of nuclear mRNA with two-state (telegraph) transcriptional dynamics. Observations of the model can be concisely described as following a spatial Cox process driven by a stochastically switching partial differential equation. We derive analytical solutions for spatial and demographic moments and validate them with simulations. We show that the distribution of mRNA counts can be accurately approximated by a Poisson-beta distribution with tractable parameters, even with complex spatial dynamics. This observation allows for efficient parameter inference demonstrated on synthetic data. Altogether, our work adds progress towards a new frontier of subcellular spatial resolution in inferring the dynamics of gene expression from static snapshot data.

1. Introduction

Gene expression at the molecular scale is stochastic [1,2]. The consequences of this variability span development and disease [3–5]. Over the past decades, a vast body of research has evolved on constructing and analysing increasingly intricate biophysical models to disentangle the sources and functions of gene expression stochasticity [6–13]. More recently, these mechanistic models have also seen uptake and success in revealing insights from messenger RNA (mRNA) count distributions from imaging and sequencing transcriptomics technologies [14–21].

Intertwined with stochasticity, gene expression is also an inherently spatial process at the subcellular scale [22]. After transcription at distinct locations within the highly structured nucleus, mRNA must be transported through the nuclear interior to be exported through nuclear pores [23]. This spatial transport and export of mRNA into the cytoplasm is followed by translation into proteins and serves as a fundamental regulation of gene expression [24–27]. Imaging technologies now give access to observing these subcellular spatial processes at unprecedented resolution. For instance, single-molecule fluorescence *in situ* hybridization (smFISH) [28,29] provides spatial locations of individual mRNA molecules within nuclei and cytoplasm for multiple genes [30,31]. State-of-the-art analysis of this subcellular spatial data is largely phenomenological [32–34] and challenging to associate with biophysical mechanisms. Models that do mechanistically account for nuclear export largely do so by treating the nucleus as a homogeneous compartment [35–40] and do not readily incorporate fine-grained spatial features that influence mRNA dynamics. For instance,

nuclear geometries [41] and transcription site locations [42] vary per cell even for the same gene, both of which shape the timescale of mRNA export. Perhaps even more importantly, spatial locations of mRNA encode the underlying dynamics of their production and degradation, information that is discarded by considering only counts. In summary, important subcellular spatial details are readily available from imaging, but the ability to incorporate them into current mechanistic modelling machinery is lacking. Motivated by this gap between data and theory, this work pursues the advancement of stochastic models of gene expression to the next frontier of subcellular spatial resolution.

The slow progress towards faithful subcellular spatial models of gene expression is an outcome of the immense challenges involved. Beyond the staggering complexities of the spatial organization of the nucleus, even simple spatial stochastic models have considerable technical obstacles facing their uptake [43]. For non-spatial models, the pursuit is a scalar stochastic quantity described by a distribution. The addition of space increases the complexity dramatically with both stochastic numbers and spatial locations of interest. From a computational perspective, this hurdle cannot be overstated. For instance, consider a two-dimensional smFISH image discretized into $N \times N$ pixels. There is a temptation to use any of the zoo of techniques that have enjoyed success for non-spatial models, including generating functions [44], finite-state projections [16,45], or neural networks [46,47]. However, one would seemingly need to perform these sometimes already costly or complex calculations for all pixels in the image, each with a distribution for counts. Even for a modest $N \sim 100$, solving for N^2 distributions for a single image becomes computationally prohibitive. Therefore, the analysis of spatial stochastic models for gene expression must be carefully considered to ensure that it remains computationally tractable enough to be associated with data.

To address the computational challenges of spatial stochasticity, we use a spatial point process formulation of the problem [48]. Specifically, we build upon our recent work [49] that investigates inference of the dynamics in a model of nuclear mRNA that undergoes spatial stochastic birth, death, and diffusion. The resulting description is a spatial Poisson process with an intensity that is the solution of a deterministic partial differential equation (PDE). This description allows for the rapid evaluation of a mechanistic likelihood that encodes both the stochastic number and positions of the particles simultaneously, including the important complication of heterogeneities [15,41]. With knowledge of the diffusion coefficient (measurable through live-cell tracking [50]), both the birth rate λ and death rate γ can be recovered from steady-state snapshot spatial data. This is in contrast to the non-spatial birth-death process, where only the ratio λ/γ can be recovered, highlighting the value of spatial information even for purely demographic inquiries.

The mathematics of our previous work hinges on a Poissonian birth process corresponding to a constitutive gene. The constitutive model neglects the bursty behaviour of transcription [51] and consequently fails to reconcile the dispersion (variance relative to the mean) of RNA counts seen in real data [52–54]. In this work, we extend the point process framework to a more realistic telegraph model [55] of transcriptional activity that stochastically switches on and off. The introduction of particle correlations creates a technical obstacle unaddressed with the previous work's machinery. Here, we follow [48] and use the spatial Poisson representation of the process [56,57] that yields a stochastic partial differential equation (SPDE) for the intensity of a Poisson process, making it a Cox process. Motivated by the success of the Poisson representation in deriving valuable analytical results for non-spatial models [58–61], the focus of this work is interrogating exact and approximate analytical results on this SPDE-driven point process model for spatial bursty gene expression.

The outline of the article is as follows. We first introduce the formulation of the spatial stochastic model of nuclear mRNA with stochastically switching transcription. Using the spatial Poisson representation, we show that observations of the model follow a Cox process whose intensity is described by a stochastically switching PDE. For a one-dimensional model, we analytically compute spatial and demographic moments for the point process, including verification by comparison with stochastic simulations. The value of these moments is limited, but we identify that the full distribution of counts can be well approximated by a Poisson-beta distribution with analytically tractable parameters that encode the spatial effects. This distributional approximation is shown to extend to more realistic variations of the model, including a semi-reflecting boundary that approximates pores on the nuclear surface, and more realistic cell shapes with heterogeneous interiors. The results culminate into a proof-of-concept demonstration of inference on synthetic data with heterogeneous cell shapes and gene locations. Altogether, the work paves a clearer path towards mechanistic model-based inference of stochastic gene expression that faithfully incorporates subcellular spatial dynamics.

2. Model

The model considered here is a spatial stochastic description of individual molecules of mRNA in the nucleus of a cell, shown schematically in figure 1. The nuclear geometry is encoded in a domain Ω with boundary $\partial\Omega$. The mRNA molecules are created through transcription at a fixed gene spatial location $z \in \Omega$, at a stochastically varying rate based on a discrete promoter state. The transcription rate follows a dichotomous noise process [62], switching between an 'on' rate λ and an 'off' rate λ_{off}, at rates α, β, respectively. Once transcribed, mRNA diffuses in the nuclear geometry with diffusivity D until they face two possible outcomes: degradation at rate γ or exportation at the nuclear boundary. Nuclear export rates are encoded in a parameter κ discussed in further detail later. Heuristically, $\kappa = 0$ corresponds to no exportation (a purely reflecting boundary) and $\kappa = \infty$ describes the instantaneous export of RNA at the nuclear boundary (purely absorbing). The model observations correspond to the spatial positions of the RNA $x_1, \ldots, x_n$ within the nucleus, where both positions and numbers are stochastically evolving. Throughout, only steady-state dynamics $t \to \infty$ are considered, acknowledging the neglect of transient effects from cell-cycle dynamics [13,39,63]. For illustrative purposes throughout the article, the domain is considered to be one spatial dimension with $\Omega = [-R, R]$, with R crudely interpreted as the nuclear radius.

The rest of the manuscript pursues statistical descriptions of observations from this model. We reiterate the challenges for emphasis. Both the number and locations of the molecules are stochastic. Moreover, the temporally correlated nature of the birth
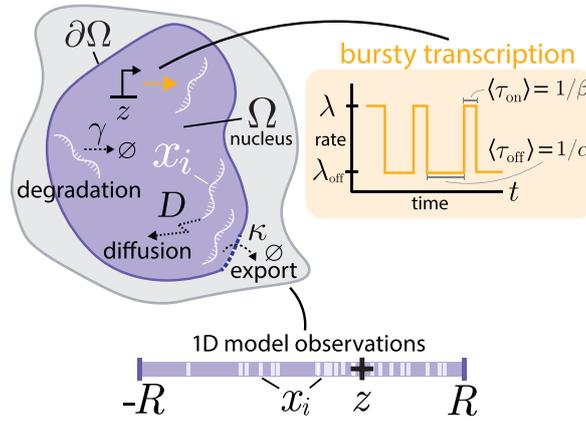
Figure 1. Schematic of the nuclear mRNA model. Transcription occurs at a spatial location z at a rate that stochastically switches between λ and λ_{off} at rates α, β. After transcription, mRNA diffuse with diffusivity D until degradation at rate γ or export at the boundary, controlled by parameter κ. Throughout the article, we primarily consider a one-dimensional spatial model with $\Omega = [-R, R]$.

process induces correlations between the molecules, so they may not be considered independently. These challenges are alleviated by the discovery that the model enjoys a straightforward spatial Poisson representation [48,56,57,64]. With more details shown in appendix A, we show that the particle locations follow the spatial Poisson process:

$$x_1, \dots, x_n \sim \text{Poiss}(u(x)), \tag{2.1}$$

where the intensity $u(x)$ corresponds to the steady-state distributional solution $u(x, t)$ of the SPDE:

$$\partial_t u = D\nabla^2 u - \gamma u + \Lambda(t)\delta(x - z). \tag{2.2}$$

In equation (2.2), $\delta(x)$ is the Dirac delta function and $\Lambda(t)$ is the continuous-time (asymmetric) dichotomous noise process, switching between values $\{\lambda_{\text{off}}, \lambda\}$, at rates $\{\alpha, \beta\}$, summarized by:

$$\Lambda(t): \ \lambda_{\text{off}} \underset{\alpha}{\overset{\beta}{\leftrightarrows}} \lambda. \tag{2.3}$$

The SPDE equation (2.2) can be written as a stochastically switching PDE [65]:

$$\partial_t u = D\nabla^2 u - \gamma u + \lambda_{\text{off}}\delta(x - z)$$

$$\beta \uparrow\downarrow \alpha$$

$$\partial_t u = D\nabla^2 u - \gamma u + \lambda\delta(x - z). \tag{2.4}$$

The stochastic nature of the intensity $u(x, t)$ in the Poisson process (equation (2.1)) makes the observations $x_1, \dots, x_n$ a spatial Cox process, a doubly stochastic Poisson process. We defer a discussion of boundary conditions for now.

It is worthwhile to note that the contrast to the purely Poissonian birth (constitutive gene) case. If transcripts are produced at a constant rate λ, the Poisson process (equation (2.1)) remains but now with a purely deterministic intensity:

$$\partial_t u = D\nabla^2 u - \gamma u + \lambda\delta(x - z).$$

In this Poissonian birth scenario, the particles are entirely independent from each other. Therefore, the stochastic intensity underlying the point process can be attributed to arising from the correlated nature of the particles.

The Poisson representation provides clarity in encoding the locations and stochastic counts concisely. However, the resulting Poisson point process with an intensity driven by an SPDE is not immediately illuminating. This leads us to pursue calculating emergent properties from this formulation.

3. Model analysis

3.1. Mean behaviour

Our starting point for analysis of the model is in computing moments for both spatial positions and molecular counts. The mean behaviour is straightforward to compute. Taking averages of equation (2.2) with respect to the realizations of $\Lambda(t)$ gives the deterministic PDE for the average $\langle u(x, t) \rangle$:

$$\partial_t \langle u \rangle = D\nabla^2 \langle u \rangle - \gamma \langle u \rangle + \frac{\alpha\lambda + \beta\lambda_{\text{off}}}{\alpha + \beta}\delta(x - z). \tag{3.1}$$

The mean of the counts for the Cox process (equation (2.1)) is the integrated mean intensity [66]:

$$\langle n(t) \rangle = \int_\Omega \langle u(x, t) \rangle \mathrm{d}x. \tag{3.2}$$
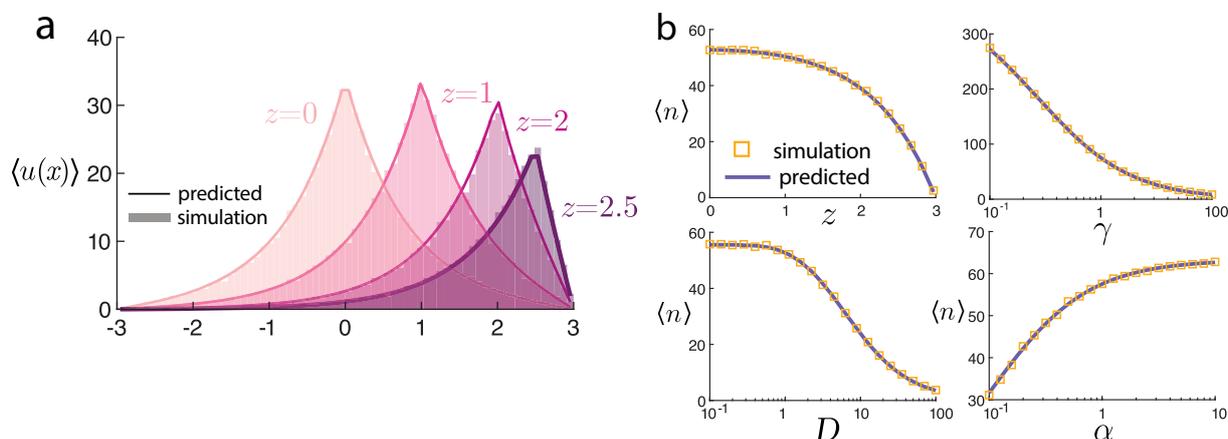
Figure 2. Mean steady-state behaviour of the model. (a) Mean positions $\langle u(x) \rangle$ from equation (3.4) varying source locations z. (b) Mean number of particles $\langle n \rangle$ from equation (3.5) for various parameter sweeps. Simulations and predicted values closely agree.

For ease of digesting the resulting formulae, we take $\lambda_{\text{off}} = 0$, $\kappa = \infty$, and $\Omega = [-R, R]$ for now. In the steady-state limit $t \to \infty$ equation (3.1) becomes the boundary value problem:

$$0 = D\partial_{xx}\langle u \rangle - \gamma \langle u \rangle + \frac{\alpha\lambda}{\alpha + \beta}\delta(x - z),$$

$$\langle u(-R) \rangle = 0, \ \langle u(R) \rangle = 0. \tag{3.3}$$

Calling $\rho = \alpha/(\alpha + \beta)$, the fraction of time the transcription activity is on, the solution of equation (3.3) is:

$$\langle u(x) \rangle = \frac{\lambda\rho}{2\sqrt{\gamma D}}\operatorname{csch}\left(2R\sqrt{\gamma/D}\right)\left[\cosh\left(\sqrt{\gamma/D}(2R - |x - z|)\right) - \cosh\left(\sqrt{\gamma/D}(x + z)\right)\right]. \tag{3.4}$$

By equation (3.2), the total average number of molecules can then be computed by integrating equation (3.4):

$$\langle n \rangle = \frac{\lambda\rho}{\gamma}\left[1 - \cosh\left(z\sqrt{\gamma/D}\right)\operatorname{sech}\left(R\sqrt{\gamma/D}\right)\right].$$

$$:= \frac{\lambda\rho}{\gamma}S. \tag{3.5}$$

The resulting mean (equation (3.5)) deserves interpretation. The dimensionless scale factor $S \in [0, 1]$ reflects the level of export out of the boundaries. As $S \to 1$, the purely non-spatial mean is recovered and space plays no role in the molecular counts. Otherwise, S decreases with any factor that increases the overall flux out of the boundary: faster diffusion, smaller domains or z closer to the boundary. S is an increasing function of γ, suggesting an interpretation of S as the probability a molecule is exported *before* it is degraded.

Verification of the predicted means (equations (3.4) and (3.5)) can be seen in figure 2. In panel (*a*), the mean spatial position is shown for varying gene site location z, and all other parameters are fixed. As the gene site shifts closer to the boundary, the overall intensity level decreases. This is further highlighted in panel (*b*), where the average total number of molecules goes to zero as the gene site approaches the boundary. The mean $\langle n \rangle$ also decreases with the degradation rate γ and diffusivity D, but increases with α. The predicted means and stochastic simulations show near-perfect agreement. Further details on the stochastic simulation can be found in appendix B.

3.2. Variance of the molecular counts

The mean behaviour of the telegraph model is effectively indistinguishable from Poissonian production with the lumped parameter λp as the effective production rate. We anticipate that higher order moments do not bear this equivalence.

The variance for the total number of particles of the Cox process (equation (2.1)) can be computed by:

$$\operatorname{var}(n) = \langle n \rangle + \int_{\Omega}\operatorname{var}(u(x))\mathrm{d}x. \tag{3.6}$$

For the constitutive case, the intensity is deterministic, and $\operatorname{var}(u(x)) = 0$ so that $\operatorname{var}(n) = \langle n \rangle$ and a Poisson distribution is recovered. For the bursty process, $u(x, t)$ evolves stochastically (equation (2.2)), so fluctuations in the intensity also manifest in contributing super-Poissonian variance to the molecular counts. The mean $\langle n \rangle$ was computed in the previous section, so the determination of the variance of the counts is left to determine the variance of the intensity.

Continuing with purely absorbing boundaries $\kappa = \infty$ in one spatial dimension $\Omega = [-R, R]$, we compute $\int_{\Omega}\operatorname{var}(u)\mathrm{d}x$ after a lengthy calculation. The result is a doubly infinite sum:

$$\int_{\Omega}\operatorname{var}(u)\mathrm{d}x = \psi\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{\mu_{2m-1} + \mu_{2n-1} + 2\omega}{(\mu_{2m-1} + \omega)(\mu_{2n-1} + \omega)(\mu_{2m-1} + \mu_{2n-1})} \times \frac{16\sin\left(\frac{\pi(2m-1)(R+z)}{2R}\right)\sin\left(\frac{\pi(2n-1)(R+z)}{2R}\right)}{\pi^2(2m-1)(2n-1)}, \tag{3.7}$$
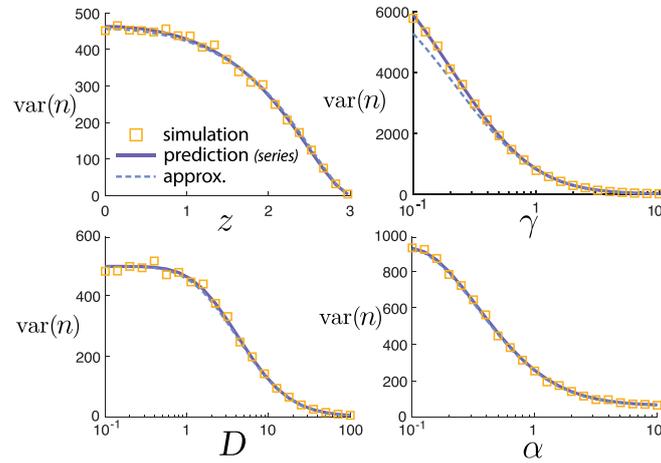
Figure 3. Variance of the number of molecules var(n) for various parameter sweeps. Simulations and predicted values from the infinite series (dashed line) solution (3.7) closely agree. Moreover, the approximate scaling value from the Poisson-beta distribution (equation (3.10)) (solid line) is nearly (but not exactly) identical over the parameter ranges.

where $\psi = \lambda^2 \alpha \beta / (\alpha + \beta)^2$, $\omega = \alpha + \beta$, and $\mu_m := -\gamma - \pi^2 D m^2 / (4R^2)$. The calculation largely follows [67,68] on the stochastic cable equation. The key observation is to note that the SPDE solution to equation (2.2) can formally be written as:

$$u(x, t) = \int_0^t G(x, t - s) \Lambda(s) \mathrm{d}s, \tag{3.8}$$

where $G(x, t)$ is the appropriate Green's function with a textbook series solution [69]. This Green's function is combined with moments of dichotomous noise $\Lambda(t)$ to yield this series solution for the integrated variance of the intensity. Further details of the calculation can be found in appendix C.

To the best of our knowledge and `Mathematica`'s abilities to simplify symbolic expressions, the infinite series (equation (3.7)) does not afford an elementary expression. It can be evaluated straightforwardly numerically. A comparison between the variance var(n) predicted by the series solution can be seen in figure 3. The parameters D, γ, α, and $z \to R$ all yield a decrease in the variance of molecular counts. Furthermore, there is close agreement between the series prediction (equation (3.7)) and the stochastic simulations.

3.3. Poisson-beta distributional approximation

The difficulty in computing and digesting this series solution lends little hope to the direction of generalizing this machinery to more complex set-ups. Moreover, these moments are computed over realizations of the process in the same domain. Since the cell shape and gene site vary with each observation, it is not clear how moments may be directly connected to data with these heterogeneities. Instead, we identify a simple approximate description that will turn out to be surprisingly useful and accurate.

We motivate the approximation by reminding the reader that the mean molecular counts in equation (3.5) is $S\lambda\alpha/(\alpha + \beta)$, with the terms multiplying S interpreted as the non-spatial mean. One could arrive at this same answer by rescaling the parameters of the non-spatial process $\lambda \to S\lambda/\gamma$, $\alpha \to S\alpha/\gamma$, and $\beta \to S\beta/\gamma$. The division of γ arises from the lack of identifiability of a single parameter (time scale) in steady state. The factor of S will be the basis of the approximation.

The non-spatial variance is [7,61]:

$$\mathrm{var}(n)_{\text{non-spat}} = \langle n \rangle_{\text{non-spat}} + \frac{\alpha\beta\lambda^2}{(\alpha + \beta)^2 \gamma(\alpha + \beta + \gamma)}. \tag{3.9}$$

If we follow the same line of reasoning and rescale parameters $\lambda \to S\lambda/\gamma$, $\alpha \to S\alpha/\gamma$ and $\beta \to S\beta/\gamma$ to account for spatial effects, this suggests:

$$\mathrm{var}(n) \approx \langle n \rangle + \frac{S^2 \alpha\beta\lambda^2}{S(\alpha + \beta)^2 \gamma(\alpha + \beta + \gamma)}. \tag{3.10}$$

Values for this expression are plotted alongside the series solution in figure 3. Although this predicted value for the variance is distinct from the series prediction, the values are nearly indistinguishable over all parameters considered.

The non-spatial variance (equation (3.9)) corresponds to a Poisson-beta distribution for the molecular counts, parameterized by three values $\tilde{\alpha}, \tilde{\beta}, \tilde{\lambda}$:

$$p(n; \tilde{\alpha}, \tilde{\beta}, \tilde{\lambda})_{\text{PB}} = \text{Poiss}(n; \tilde{\lambda}t) \wedge_t \text{Beta}(t; \tilde{\alpha}, \tilde{\beta}) = \int_0^1 \frac{\tilde{\mu}t}{n!} e^{-\tilde{\mu}t} \frac{t^{\tilde{\alpha}-1}(1-t)^{\tilde{\beta}-1}}{B(\tilde{\alpha}, \tilde{\beta})} \mathrm{d}t = \frac{\tilde{\mu}^n}{n!} \frac{(\tilde{\alpha})_n}{(\tilde{\alpha} + \tilde{\beta})_n} {}_1F_1(\tilde{\alpha} + n, \tilde{\beta} + n; -\tilde{\lambda}), \tag{3.11}$$

where $\wedge_t$ denotes the mixture distribution with respect to t, ${}_1F_1$ is the confluent hypergeometric function, and $(c)_n = c(c + 1) \cdots (c + n - 1)$ is the Pochhammer symbol.
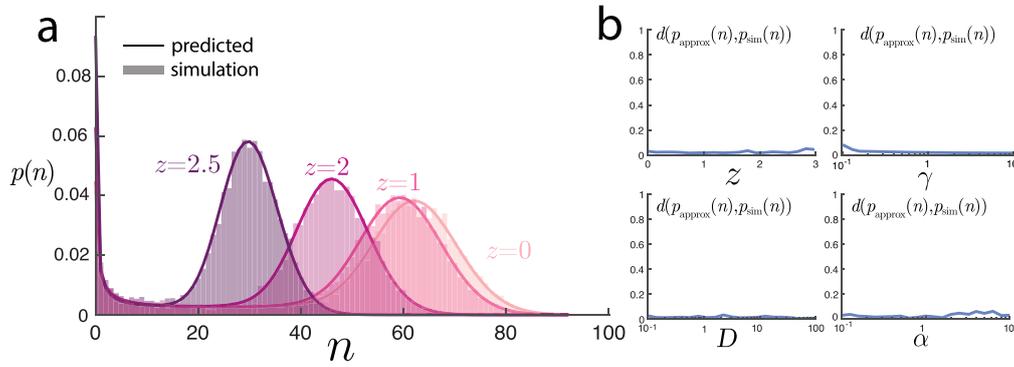
Figure 4. Comparison of the Poisson-beta distributional predictions with simulations. (a) For the same set-ups shown in 2 a with varying z, the predicted full distribution of the number of molecules closely agrees with simulations. (b) Over the parameter sweeps in figures 2 and 3, the Jensen–Shannon distance (between 0 and 1) between the predicted Poisson-beta distribution of n and the values from simulations are consistently small, highlighting the broad applicability of the approximation.

The spatial variance approximation (equation (3.10)) is recovered with the choice of $\tilde\lambda = S\lambda/\gamma$, $\tilde\alpha = S\alpha/\gamma$, and $\tilde\beta = S\beta/\gamma$. Although this expression was obtained by heuristic scaling arguments of the moments, we can plot the full distribution with these parameter choices against the values from stochastic simulations. Figure 4 shows these comparisons with remarkable agreement. In figure 4a, we demonstrate the predicted distribution of n for the same set-up as figure 2a and near-identical agreement is seen for all values of n.

One should be sceptical about the validity of this approximation. To investigate this, we computed the Jensen–Shannon divergence $d_{JS}(p_1, p_2) := [d_{KL}(p_1|\frac{p_1+p_2}{2}) + d_{KL}(p_2|\frac{p_1+p_2}{2})]/2$ between the predicted Poisson-beta distribution and the empirical distribution from simulations. This divergence takes values between 0 and 1, and across all parameter ranges tested shown in figure 4, the values were on the order of ≈ 0.01, suggesting remarkable agreement.

3.4. Generalizing to semi-absorbing boundaries

The Poisson-Beta distribution provides an accurate prediction for the full distribution of molecular counts for purely absorbing boundaries and only requires a single deterministic PDE solution. To demonstrate this approach's surprisingly broader applicability, we first extend the model to account for more realistic nuclear export. Consider a Robin boundary condition for the SPDE (2.2):

$$D\nabla u(x,t) \cdot \boldsymbol{n}(x)\big|_{x\in\partial\Omega} = \kappa u(x,t)\big|_{x\in\partial\Omega}, \tag{3.12}$$

where $\boldsymbol{n}$ is the outward normal vector to the boundary $\partial\Omega$. Such a boundary condition can arise from homogenizing the surface of the nucleus with absorbing pores [70,71]. We interpret κ as controlling the kinetics of nuclear export [27]. With these boundary conditions, the PDE for the steady-state mean intensity $\langle u(x)\rangle(x)$ (equation (3.3)) now becomes:

$$0 = D\partial_{xx}\langle u\rangle - \gamma\langle u\rangle + \frac{\alpha\lambda}{\alpha+\beta}\delta(x-z),$$

$$D\langle\partial_x u(-R)\rangle = \kappa\langle u(-R)\rangle, \quad -D\langle\partial_x u(R)\rangle = \kappa\langle u(R)\rangle. \tag{3.13}$$

This can again be solved directly, and by equation (3.2) the mean number of molecules can be computed from the integral over the mean intensity, yielding a similar result to the $\kappa = \infty$ case (equation (3.5)):

$$\langle n\rangle = \int_{-R}^{R}\langle u(x)\rangle \mathrm{d}x = \frac{\lambda\rho}{\gamma}S_\kappa, \tag{3.14}$$

where S_κ is a slightly unwieldy but straightforward to compute expression

$$S_\kappa = \frac{\sqrt{\gamma D}-\left(\sqrt{\gamma D}+\kappa\right)e^{2R\sqrt{\frac{\gamma}{D}}}+\kappa e^{\sqrt{\frac{\gamma}{D}}(R-z)}+\kappa e^{\sqrt{\frac{\gamma}{D}}(R+z)}-\kappa}{-\sqrt{\gamma D}+\left(\sqrt{\gamma D}+\kappa\right)e^{2R\sqrt{\frac{\gamma}{D}}}+\kappa}. \tag{3.15}$$

This scaling factor has the same interpretation: $S_\kappa \in [0,1]$ and can be understood as the probability of export before degradation. As $\kappa \to 0$, $S_\kappa \to 0$, since no export occurs. On the other hand, as $\kappa \to \infty$, $S_\kappa \to S$, the fully absorbing limit defined in equation (3.5). In figure 5a,b, we show close agreement between stochastic simulations and the predicted mean behaviour. As κ gets smaller, the number of molecules increases, and the intensity profile flattens out. In stochastic simulations, κ is interpreted in the partially reflected sense [72], where larger κ encodes a higher probability of exit. Further simulation details can be found in appendix B.

We now consider the same parameter scaling for a Poisson-beta distribution, $\tilde\lambda = S_\kappa\lambda/\gamma$, $\tilde\alpha = S_\kappa\alpha/\gamma$, and $\tilde\beta = S_\kappa\beta/\gamma$. In figure 5c,d, we show the predicted variance and Fano factor $F = \mathrm{var}(n)/\langle n\rangle$. As κ increases, both the variance and Fano factor decrease, in agreement with experimental findings of [25] that show slowing down export (lower κ) leads to an increase in the nuclear mRNA Fano factor for several genes. Moreover, over this same range of κ, the Poisson-Beta prediction provides a remarkably accurate prediction for the full distribution of molecular counts, shown in figure 5e. This predicted distribution required only the computation
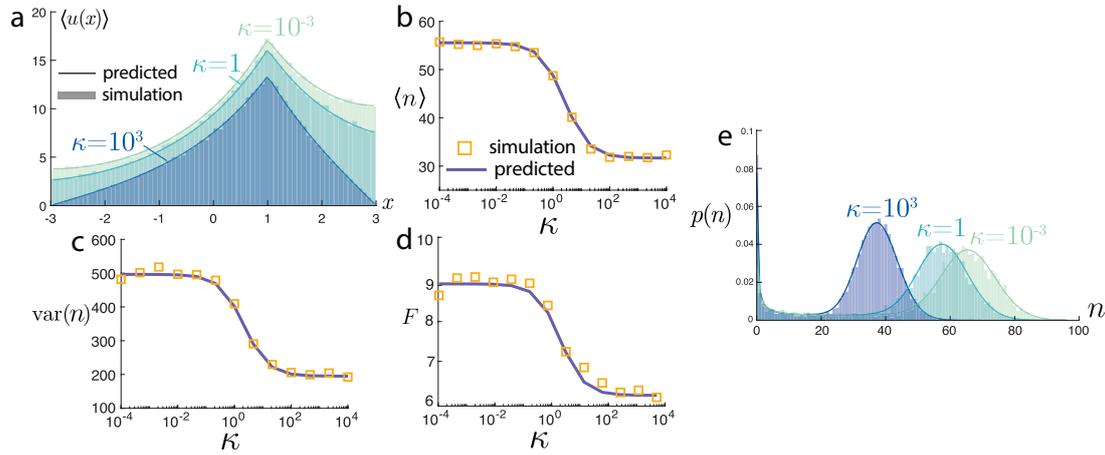
Figure 5. Influence of the export rate parameter κ. (a) Mean positions $\langle u(x) \rangle$ for varying levels of κ from (3.13). (b) Mean number from (3.14), (c) variance, and (d) Fano factor $F = \text{var}(n)/\langle n \rangle$ of molecular counts all decrease with κ. (e) The Poisson-beta predicted distribution shows close agreement with stochastic simulations for various κ.

of the deterministic PDE for the mean behaviour (equation (3.14)). By contrast, the full variance calculation would have required tedious computations, ultimately likely to also result in an unwieldy infinite series solution. Thus, the finite κ scenario highlights the use of the Poisson-beta approximation.

3.5. Fano factor interpretation of spatial effects

We are now equipped with the ability to study the nuclear mRNA model's statistical behaviour for various parameters and model variants. Before continuing to generalize the model, we take a brief interlude to highlight the value of an explicit spatial model in interpreting molecular counts of nuclear mRNA. In figure 6, we compare the Fano factor for the explicit spatial model:

$$F_{spat} = \frac{\text{var}(n)}{\langle n \rangle} = 1 + \frac{\lambda \beta S}{(\alpha + \beta)(\gamma + S(\alpha + \beta))}, \tag{3.16}$$

to the nonspatial variant which neglects nuclear export:

$$F_{non-spat} = 1 + \frac{\lambda \beta}{(\alpha + \beta)(\gamma + (\alpha + \beta))}. \tag{3.17}$$

In figure 6a, we see that the largest deviation between the Fano factors of the spatial and non-spatial models occurs when exportation occurs most frequently: a gene site near the boundary ($z \approx R$) and fast export (κ large). As the gene site moves to the interior of the nucleus or exportation slows, the two models more closely agree in their prediction. Notably, the Fano factor for the spatial model is always smaller than that of the non-spatial. Taken together, this reduction can be straightforwardly understood by noting that export removes molecules from the count and therefore reduces the overall fluctuations. This point is supported in figure 6b, where larger values of the degradation rate γ are shown to lead to the smallest deviation between the spatial and non-spatial models. This is expected, as larger γ means the outcome of transcripts is dominated by degradation, and export plays less of a role in their dynamics. The interplay between export and the transcriptional state is less easy to predict. In figure 6c,d, we see that slow transcriptional switching dynamics (small $\omega = \alpha + \beta$) leads to the biggest deviation between spatial and non-spatial variants. Moreover, the deviation is relatively robust to the fraction of time the transcription state is on, $p = \alpha/(\alpha + \beta)$, except when $p \approx 1$ and production approaches a constitutive Poissonian process. Although these deviations between the non-spatial and spatial models may be easily understood and predicted, we emphasize their quantitative importance. If one were to fit a non-spatial telegraph model to nuclear molecular counts with some Fano factor to infer the mechanisms of transcriptional activity [73], the underlying dynamics may be incorrectly recovered and could lead to erroneous results. It is worth noting that transcriptomic measurements are more commonly total cellular counts, including the nucleus and cytoplasm. Assuming that degradation rates differ between these regions, the discussed spatial processes would still impact the Fano factor for the total count, but further investigation of this extended model is deferred to future work.

3.6. Two-dimensional cell with spatial heterogeneities

We make one last note about the applicability of findings towards more realistic set-ups. Undoubtedly, mRNA transport in the nucleus is not one dimensional, nor a spatially homogeneous process. The nucleus is crowded with various factors that are known to modulate mRNA motion [23,50]. Moreover, the geometry of the nucleus itself plays a role in the export process because the mRNA must be transported to the boundary to be exported. In this last demonstration, we highlight the ability to handle these challenges within the currently presented framework.

In figure 7, we show a synthetically generated, two-dimensional cell geometry. The nuclear interior is assumed to be heterogeneous, modelled by a spatially dependent diffusion coefficient $D(x)$. There are inherent technical challenges in interpreting
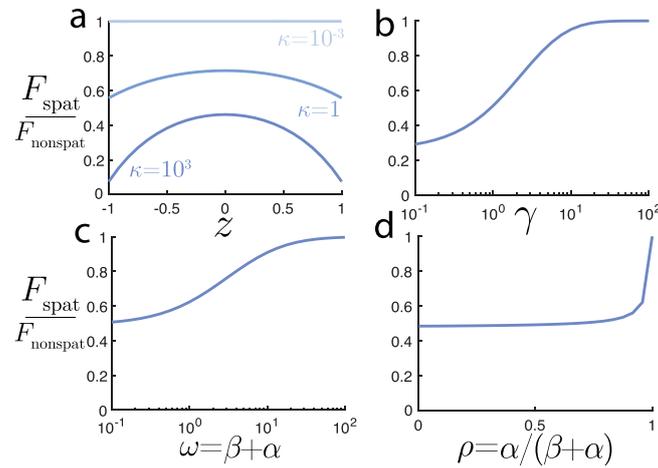
Figure 6. Comparison of the Fano factors for spatial and non-spatial models. (a) Source locations with the largest κ and closest to the boundary have the largest deviation between spatial and non-spatial models. (b) The degradation rate γ and (c) transcriptional switching timescale ω both increase the agreement between the spatial and non-spatial Fano factors. (d) The proportion of time in the 'on' state, p has a relatively weak influence on the Fano factor ratio, except for nearly constitutive $p \approx 1$.
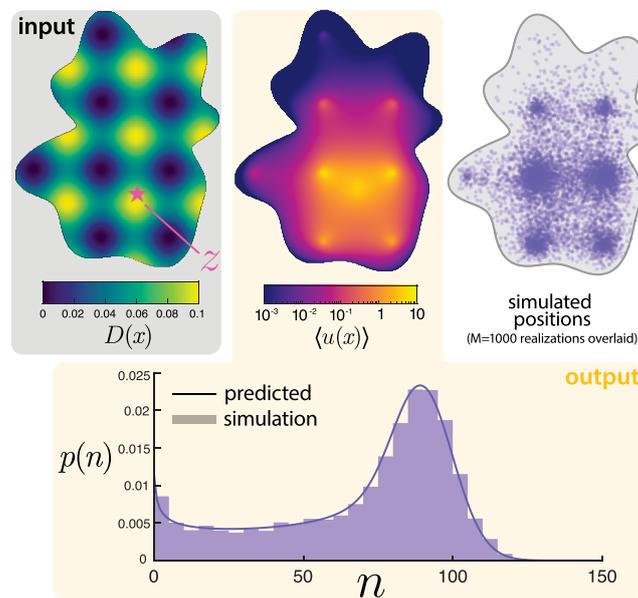

Figure 7. Generalizing the model to more realistic scenarios. The diffusivity $D(x)$ is taken to be spatially varying (in the Itô sense) in a randomly generated two-dimensional domain. The mean number of particles $\langle u(x) \rangle$ is solved numerically from the PDE (3.18) and shown. The total mean is then used to specify the parameters of a Poisson-beta distribution for the total number of particles n, which shows close agreement with stochastic simulations.

spatially dependent diffusivities [74,75]. Here, we take an Itô interpretation with no justification beyond being the most straightforward to implement. In this vein, we also consider purely absorbing boundaries ($\kappa = \infty$). How can the distribution of mRNA counts and positions in this spatially heterogeneous domain be predicted? Based on the findings thus far, both of these distributions can seemingly be accurately approximated by a single deterministic PDE solve for the mean intensity. Motivated by this observation, we use MATLAB's PDE Toolbox to solve the deterministic equation

$$0 = \nabla^2[D(x)\langle u(x) \rangle] - \gamma\langle u \rangle + \lambda p \delta(x - z),$$
$$\langle u(x) \rangle\big|_{x \in \partial\Omega} = 0. \tag{3.18}$$

The choice of Itô interpretation manifests in the location of the derivative with respect to the diffusivity. The solution of this PDE for $\langle u(x) \rangle$ is shown in figure 7 alongside $M = 1000$ superimposed stochastic realizations of the particle positions. The prediction for positions and stochastic simulations demonstrate excellent agreement. Qualitatively, the molecules tend to get localized in regions of low diffusivity. Next, we turn to quantitative predictions. Motivated by the one-dimensional answers (equations (3.5) and (3.14)), we can define the analogous scale factor as

$$S := \frac{\int_\Omega \langle u(x) \rangle \mathrm{d}x}{\lambda p / \gamma}. \tag{3.19}$$

This integral can be computed from the numerical PDE solution. Then, we posit that n approximately follows a Poisson-beta distribution with parameters $\tilde{\lambda} = S\lambda/\gamma$, $\tilde{\alpha} = S\alpha/\gamma$, and $\tilde{\beta} = S\beta/\gamma$. The resulting predicted distribution is shown in comparison against the counts from simulations in figure 7. This prediction using the numerical PDE solution for the scaling factor shows remarkable

agreement with simulations. Although numerically solving a PDE on cellular geometries may be computationally costly, we emphasize that it seems far less costly than any current alternative approach for characterizing the distributions of both counts and positions. For instance, a generating function approach would seemingly require a PDE solution for each discretized pixel.

4. Inference on rates with heterogeneous cells

In this last section, we outline a possible avenue to use the findings in this work towards inference of model parameters with data. Although we have thus far achieved an understanding of the forward predictions of the model, the Cox process observations provide distinct technical challenges in their inference. The likelihood for a single observation (cell) of the Cox process (equation (2.1)) with parameters θ is [76]:

$$\mathcal{L}(\theta; X) := p_\theta(X = x_1, \dots, x_n) = \int_u \left[\frac{1}{n!} e^{-\int_\Omega u(x)dx} \prod_{i=1}^n u(x_i) \right] du(x). \tag{4.1}$$

In other words, evaluation of the likelihood requires (infinite dimensional) integration over all possible realizations of the solution to the SPDE (2.2). While there is extensive literature on sophisticated numerical approaches for performing inference with this Cox process likelihood [76], we instead focus on the possibility of a simpler approximation that uses our earlier discussed findings.

As a motivating detour, we momentarily consider the likelihood of a simple Poisson process with deterministic intensity $u(x)$ and underlying parameters θ. In this case, the likelihood becomes the term inside the integral of the Cox process likelihood:

$$L_{\text{Poiss}}(\theta; X) = \frac{1}{n!} e^{-\int_\Omega u(x)dx} \prod_{i=1}^n u(x_i) \tag{4.2}$$

The observation we hope to emphasize can be seen by rearranging this likelihood and denoting $\langle n \rangle = \int_\Omega u(x)dx$ and $p(x_i) = u(x_i)/\langle n \rangle$. Then, the likelihood (equation (4.2)) can be written as

$$L_{\text{Poiss}}(\theta; X) = \frac{\langle n \rangle^n e^{-\langle n \rangle}}{n!} \prod_{i=1}^n p(x_i)$$

$$= f_{\text{Poiss}}(n; \langle n \rangle) \prod_{i=1}^n p(x_i). \tag{4.3}$$

In other words, for the Poisson spatial point process, the likelihood can be decomposed into the contributions of the counts and positions. Motivated by this observation and the findings presented thus far, we approximate the full Cox process likelihood by

$$\mathcal{L}(\theta; X) = \int_u \left[\frac{1}{n!} e^{-\int_\Omega u(x)dx} \prod_{i=1}^n u(x_i) \right] du(x) \approx p_\theta(n) \prod_{i=1}^n p_\theta(x_i), \tag{4.4}$$

where the spatial distribution is computed from the expected positions:

$$p_\theta(x) = \frac{\langle u(x) \rangle}{\langle n \rangle} = \frac{\langle u(x) \rangle}{\int_\Omega \langle u(x) \rangle dx}. \tag{4.5}$$

Aside from the motivating example with deterministic intensity $u(x)$, this approximation does not seem to be exact. Instead, equation (4.4) can be interpreted as a 'mean-field' approximation that does not fully account for correlations between counts and the spatial positions. Importantly, however, the quantities in this approximation are all tractable using the previously discussed results of this work. The mean behaviours used in equation (4.5) can be computed exactly. The count distributions $p_\theta(n)$ were found to be well approximated by a Poisson-beta distribution with parameters computed from $\langle n \rangle$. Therefore, the evaluation of this approximate likelihood reduces to straightforward analytical or numerical solutions to a single PDE for the mean intensity. In most scenarios, we can imagine, this is far less costly than any Monte Carlo sampling technique for evaluating the true Cox process likelihood.

The obvious question that remains is whether this approximation to the likelihood is sufficiently accurate to do reliable inference. We address this question with a demonstration of inference on synthetically generated data, presented in figure 8. To mimic the realistic challenge of cell-to-cell heterogeneity, we generate $M = 500$ synthetic observations from the one-dimensional model, where each cell i has a randomly generated z_i and R_i, but all kinetic parameters are fixed across the cells. We again assume steady-state conditions, so not all kinetic parameters are identifiable. We assume that the diffusion coefficient D may be measured through other means (for instance, live-cell tracking [50]), but the remaining parameters are unknown $\theta = [\lambda, \gamma, \alpha, \beta, \kappa]$. With this synthetic dataset, we use the Cox process likelihood approximation (equation (4.4)) for maximum likelihood inference using `MATLAB`'s `fminsearch`. To simultaneously diagnose the identifiability of these parameters, we also compute the profile likelihood for each parameter, defined by

$$PL(\theta_i) := \max_{\theta_j : j \neq i} \mathcal{L}(\theta). \tag{4.6}$$

That is, the parameter of interest is fixed to a specified value (determined by the functional input to the profile likelihood) and the remaining parameters are optimized. The profile likelihood is a standard diagnostic for identifiability of parameters [77]. The
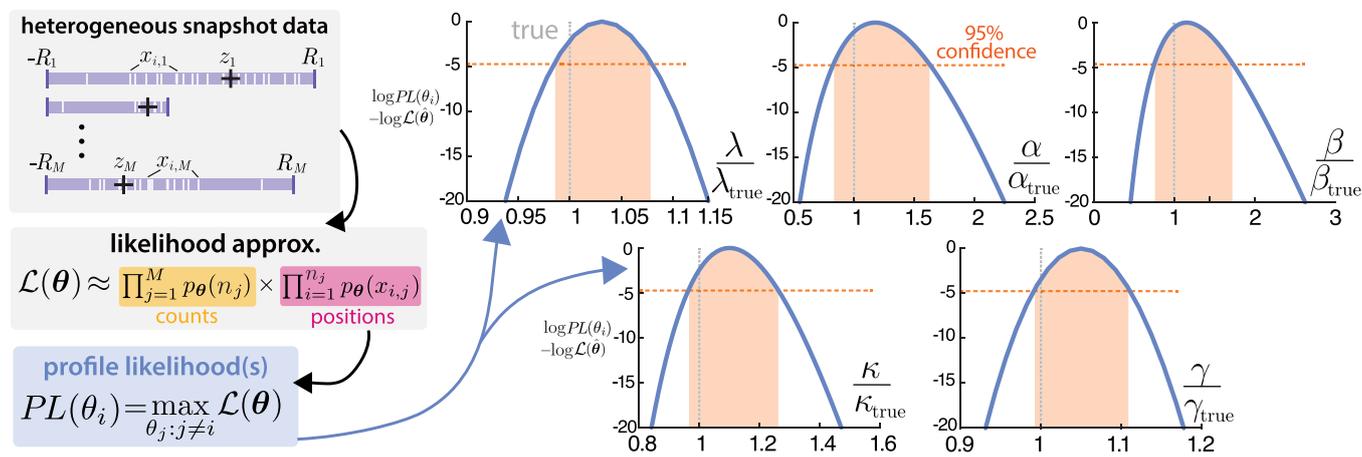
Figure 8. Demonstration of inference and identifiability on a heterogeneous dataset. A synthetic dataset of $M = 500$ images, each with randomly chosen R and z was generated. This is then used in an approximation of the full likelihood (equation (4.4)) from which profile likelihoods are computed. The profile likelihoods are plotted as $\chi(\theta_i) := \log PL(\theta_i) - \log \mathcal{L}(\hat{\theta})$ to facilitate comparison with the threshold for an asymptotic confidence region for simultaneous inference of the parameters. Since all profile likelihoods decrease sufficiently fast over the windows of interest, structural and practical identifiability can be concluded.

output of each profile can be compared with the maximizing value of the likelihood over some window of interest, often corresponding to a confidence region. If the profile likelihood takes on sufficiently distinct values over this window, the parameter can be interpreted as both structurally and practically identifiable [78].

In figure 8, we see that the maximizing values of the profile likelihoods closely agree with the true simulated parameters. Moreover, we consider a window corresponding to the asymptotic 95% confidence region for the recovery of all parameters with a threshold of $\chi^2_{0.95,4}$, accounting for the four-degrees of freedom in the optimization of each profile likelihood [78]. Over this confidence region, all five parameters appear to be structurally and practically identifiable. We remind the reader of the non-spatial variant, a Poisson-beta distribution with only three identifiable parameters. It seems worth noting that all true parameters fall at the lower end of the confidence region, which may be indicative of biased estimates arising from this approximate inference procedure. We believe this proof-of-concept demonstration highlights how the findings of this work can be used for inference, but optimization and dissection of this inference are left for the future.

5. Discussion

We have formulated and investigated a toy model of nuclear mRNA that explicitly incorporates nuclear spatial diffusion and telegraph transcriptional dynamics. The most fundamental finding is that observations of the model form a Cox process, a spatial point process with intensity corresponding to the solution of a stochastically switching PDE. This spatial telegraph PDE lends itself to some analytical tractability in one spatial dimension. The mean and variance of both spatial distributions and counts were computed and verified against stochastic simulation. The model predictions about the role of spatial processes qualitatively agree with experimental findings, most notably that slowing down nuclear export increases the dispersion of nuclear mRNA counts [25]. However, these basic calculations were unwieldy even for the simplest of the models, making quantitative comparison challenging. The main upside of our work is the observation that a Poisson-beta distribution well approximates the full count distribution and parameters of this distribution can be straightforwardly computed. This distributional approximation was shown to generalize broadly, including Robin boundaries that model nuclear export and generically shaped two-dimensional cells with spatially heterogeneous diffusion. The computational tractability of the count distribution empowers the ability to perform approximate inference. We show that with heterogeneous snapshots of cells with distinct sizes and gene sites, kinetic parameters are identifiable from the spatial distributions and counts. Altogether, our work paves fundamental theoretical progress towards connecting imaging data (for instance, from smFISH) of spatial distributions of nuclear mRNA to infer the spatiotemporal gene expression dynamics underlying them.

Our findings should not be viewed as in tension with work that models gene expression dynamics non-spatially, for instance, those that treat mRNA export as a unimolecular reaction between homogeneous nuclear and cytoplasmic compartments [36,40]. Instead, our work sheds light on why these non-spatial models have found such success in their ability to fit observed RNA counts. Even with explicit spatial diffusion, nuclear export, and spatial heterogeneities, we found that a Poisson-beta distribution, the same prediction as the non-spatial process, well-describes the count distribution across parameters. The observation that explicit diffusion until export in the nucleus can be modelled as a single-step reaction has been noted before [25], but we emphasize that care must be taken in interpreting the parameters. For instance, a decay term must be interpreted as a mix of degradation and spatial export, and our findings show how these terms are differently affected by spatial geometric factors. It seems challenging, perhaps impossible, to disentangle these effects through fitting to non-spatial models. By contrast, we have shown that fitting the distributional counts with spatial information empowers a new resolution of detail to interrogate the dynamics.

On the biological relevance of the work, the model is a caricature of reality and should be treated as such. We have chosen various parameters throughout only to highlight qualitative features of the model. Although the nucleus is crowded and heterogeneous, approximately diffusive motion has been observed for mRNA [50] with a diffusion coefficient on the order of $D \approx 0.03$ [23] to $D \approx 0.12\ \mu\mathrm{ms}^{-1}$ [79]. The Kuramoto length $\ell = \sqrt{D/\gamma}$ has been previously estimated to be $\mu\mathrm{m}$ for nuclear mRNA. The accordance of this length with nuclear diameters [29] suggests informational value in the spatial patterns. Moreover, these relative scales give a diffusive timescale of $R^2/D \approx 1$ h. This is in agreement with the observation that export and degradation are on approximately the same time scales [25]. However, nuclear export is not instantaneous, nor occurs at every location on the nuclear surface [24,27]. This suggests finite κ seems most appropriate, but we were unable to identify an approximate numerical value. It is unlikely that the full spatial distribution of diffusion coefficients can be identified as in figure 7. However, we believe this demonstrates the possibility of imaging nuclear condensates (nucleoli, speckles) that are known to mediate spatial organization of mRNA [23,80] and using this spatial information in the modelling and inference process. There is evidence nuclear mRNA degradation is far more elaborate than considered here [81], but an effective rate serves as a starting point for approximating these complexities. The transcriptionally active gene site was assumed to be fixed in location based on extensive experimental [42,82] and computational [83] evidence of chromosomal territories. This confinement leads to an effective diffusivity several orders of magnitude lower than that of mRNA. In the model, the point source could be replaced by a more realistic steady-state distribution reflecting this DNA diffusion with no apparent technical difficulties.

The telegraph process for transcriptional dynamics was chosen here as a minimally complex example of transcriptional dynamics that produces super-Poissonian dispersion of molecular counts [84]. This dispersion is intimately linked with the correlation between the particles induced by the production process and causes the particle-wise machinery of our previous work [49] to fail. We anticipate this approach lends itself better to generalizing into other complexities of interest. In this vein, there are several avenues of promising future direction. We have considered only a single mRNA population, but one could imagine extending the framework to account for multiple species available from imaging such as multiple genes [31,85] or both nuclear and cytoplasmic RNA [31,40]. The telegraph transcriptional process also lends itself to generalization to more realistic multi-state production processes [86–90], feedback (with delay) [11,91], or time-dependent cell-cycle dynamics [13,63,92]. Even more broadly, we hope the machinery of this work can be used to study other spatial regulations of gene expression [93,94] that occur with distinct subcellular localization such as RNA splicing [29,95].

On the theory and computational side, the finding of a Poisson-beta distribution for the counts was heuristically motivated. Owing to the lack of proof, it remains unclear whether the distribution of counts genuinely follows this distribution or is merely well approximated by it. Over a broad set of parameters, the approximation did not exhibit any discernible pattern of failure. We believed that the Poisson-beta approximation corresponds to a truncation of the Green's function series. However, the leading term of the series solution (equation (3.7)) does not seem to follow the form (equation (3.10)) conjectured from scaling arguments. Nonetheless, we believe this line of investigation is worthwhile to share and warrants further investigation in the future owing to its computational tractability. Alternative computational approaches that seem likely to be fruitful include simulation-based inference [96,97] and graph neural networks for stochastic reaction-diffusion that can generalize to arbitrary geometries [98].

Appendix A. Poisson representation of the model

In this section, we provide further details on the claim in the main text that the observations of the discrete particle process follow a Cox process (equation (2.1)) with intensity governed by the SPDE (2.2). The idea of spatial Poisson representations dates back to Gardiner [56,57]. More accessible recent work relating mechanistic models and point processes can be found in [48,100] and the geophysics literature [64]. We believe the explicit computation of this model with dichotomous noise has not been considered.

For simplicity that can be easily generalized, consider a one-dimensional domain discretized into windows of size Δx, each with counts $n_1, \ldots, n_M$. Denote the promoter state corresponding to $\Lambda(t)$ in the main text as $j(t) = \{0, 1\}$. In state $j = 0$, transcription occurs at rate $\tilde{\lambda}_{\mathrm{off}} = \lambda_{\mathrm{off}}(\Delta x)^{-1}$ and in state $j = 1$, at rate $\tilde{\lambda} = \lambda(\Delta x)^{-1}$. The gene site is fixed to be in the spatial location indexed k. In the discretized model, diffusion corresponds to a unimolecular reaction at rate $\mathcal{D} = D/(\Delta x^2)$.

Denote $p_0(\boldsymbol{n}, t) = p(\boldsymbol{n}, 0, t)$ and $p_1(\boldsymbol{n}, t) = p(\boldsymbol{n}, 1, t)$, respectively. The stochastic vector of counts $\boldsymbol{n}$ and promoter state j evolves through the Chapman–Kolmogorov equations:

$$\frac{\partial p_j}{\partial t}(\boldsymbol{n}, t) = \sum_{i=1}^{n} \mathcal{D}(n_{i+1} + 1)p_j(\boldsymbol{n} + \boldsymbol{r}_{i+1}^i) + \mathcal{D}(n_{i-1} + 1)p_j(\boldsymbol{n} + \boldsymbol{r}_{i-1}^i) - 2\mathcal{D}n_i p_j(\boldsymbol{n}, t) + \gamma(n_i + 1)p_j(\boldsymbol{n} + \boldsymbol{r}_i^+, t) - \gamma n_i p_j(\boldsymbol{n}, t)$$

$$+ \tilde{\lambda}_{\mathrm{off}}p_j(\boldsymbol{n} + \boldsymbol{r}_i^-, t)\delta_{j0}\delta_{ik} - \tilde{\lambda}_{\mathrm{off}}p_j(\boldsymbol{n}, t)\delta_{j0}\delta_{ik} + \tilde{\lambda}p_j(\boldsymbol{n} + \boldsymbol{r}_i^-, t)\delta_{j1}\delta_{ik} - \tilde{\lambda}p_j(\boldsymbol{n}, t)\delta_{j1}\delta_{ik}. \tag{A 1}$$

We have abbreviated the perturbation vectors $\boldsymbol{r}_i^{\pm} = [0, \ldots, \pm 1, 0, \ldots]$ and $\boldsymbol{r}_i^{i+1} = [0, \ldots, -1, +1, 0, \ldots]$. The Poisson representation takes the ansatz:

$$p_j(\boldsymbol{n}, t) = \int_{u_i} \prod_{i=1}^{M} \frac{e^{-u_i} u_i^{n_i}}{n_i!} f_j(\boldsymbol{u}, t) \mathrm{d}u_i, \tag{A 2}$$

which gives the evolution of the intensities as

$$\frac{\partial f_j}{\partial t} = -\sum_{i=1}^{M} \frac{\partial}{\partial u_i} \left[\left(\mathcal{D}u_{i+1} + \mathcal{D}u_{i-1} - 2\mathcal{D}u_i - \gamma u_i + \tilde{\lambda}_{\mathrm{off}} \delta_{j0} \delta_{ik} + \tilde{\lambda} \delta_{j1} \delta_{ik} \right) f_j(\boldsymbol{u}, t) \right] + \sum_{q=0,1} A_{jq} f_q(\boldsymbol{u}, t), \tag{A 3}$$

where A is the 2×2 generator matrix for the promoter switching process. This is the Chapman–Kolmogorov equation for the piecewise deterministic Markov process (PDMP) system:

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = \mathcal{D}u_{i+1} + \mathcal{D}u_{i-1} - 2\mathcal{D}u_i - \gamma u_i + \tilde{\lambda}_{\mathrm{off}} \delta_{ik}$$

$$\beta \uparrow\downarrow \alpha$$

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = \mathcal{D}u_{i+1} + \mathcal{D}u_{i-1} - 2\mathcal{D}u_i - \gamma u_i + \tilde{\lambda}\delta_{ik}. \tag{A 4}$$

In the limit as $\Delta x \to 0$ with $u(x, t) = u_i(t)/\Delta x$, the switching PDE (2.2) is recovered. We note the structural contrast of this result compared with related works [48,64] that arrive at Gaussian SPDEs from biomolecular reactions or white-noise limits.

Appendix B. Stochastic simulation details

For all stochastic simulations, the promoter state switching, birth of molecules, and their decay are handled by a standard Gillespie simulation [43] until some time T. To ensure steady state, T is chosen to be $T = 100/\min\{\alpha, \beta, \gamma, \lambda, R^2/D\}$. For all particles living at the end of the simulation, their birth times are recorded. These candidate molecules are then simulated spatially, starting at their birth time until either T or they exit the domain. The final positions of all remaining molecules that have survived the Gillespie and spatial simulation steps are outputs of the simulation. In simulations with finite κ, the boundary is assumed to be a partially reflected diffusion [72]. If a particle Euler steps out of the domain, with probability $P\Delta t$ it is absorbed. Based on that work, $P = \kappa\sqrt{\pi/D}$. If the particle is not absorbed, it is then reflected. All finite κ simulations throughout the article are in one spatial dimension, where reflection is straightforward. In simulations with heterogeneous $D(x)$, the Euler step is interpreted in an Itô sense , $x(t + \Delta t) = x(t) + \xi\sqrt{2D(x)\Delta t}$, $\xi \sim \mathcal{N}(0, 1)$. [75].

(a) Parameter values used in figures

Figure 2: $z = 0$, $R = 3$, $\alpha = 0.5$, $\beta = 0.1$, $\lambda = 100$, $\lambda_{\mathrm{off}} = 0$, $\gamma = 1.5$, $D = 1$. $N_{\mathrm{sims}} = 5000$.
Figure 3: $z = 0$, $R = 3$, $\alpha = 0.5$, $\beta = 0.1$, $\lambda = 100$, $\lambda_{\mathrm{off}} = 0$, $\gamma = 1.5$, $D = 1$. $N_{\mathrm{sims}} = 5000$.
Figure 4: $z = 0$, $R = 3$, $\alpha = 0.5$, $\beta = 0.1$, $\lambda = 100$, $\lambda_{\mathrm{off}} = 0$, $\gamma = 1.5$, $D = 1$. $N_{\mathrm{sims}} = 5000$.
Figure 5: $z = 0$, $R = 3$, $\alpha = 0.5$, $\beta = 0.1$, $\lambda = 100$, $\lambda_{\mathrm{off}} = 0$, $\gamma = 1.5$, $D = 5$. $N_{\mathrm{sims}} = 5000$.
Figure 6: $z = 0$, $R = 1$, $\alpha = 0.5$, $\beta = 0.1$, $\lambda = 100$, $\lambda_{\mathrm{off}} = 0$, $\gamma = 0.75$, $D = 1$, panels b–d: $\kappa = 10^3$.
Figure 7: $z = [0, -0.5]$, $\langle R \rangle = 1$, $\alpha = 0.75$, $\beta = 0.25$, $\lambda = 100$, $\lambda_{\mathrm{off}} = 0$, $\gamma = 1$, $D(x, y) = D_- + (D_+ - D_-)(\cos(3\pi x) + \cos(3\pi y) + 2)/4$., $D_+ = 0.1$, $D_- = 10^{-3}$, $N_{\mathrm{sims}} = 5000$.
Figure 8: domains are randomly generated by $R_i \sim \Gamma(4, 1/4)$, $z_i \sim \beta(4/3, 4/3) \times 2R - R$, so $\langle R \rangle = 1$, $\mathrm{var}(R) = 0.5$, $\langle z \rangle = 0$, $\mathrm{var}(z) \approx 0.34$. The remaining parameters are defined the same as the main text and taken to be: $\alpha = 1$, $\beta = 0.25$, $\lambda = 250$, $\lambda_{\mathrm{off}} = 0$, $\gamma = 1.5$, $D = 1$, $\kappa = 5$, $M_{\mathrm{data}} = 500$.

Appendix C. Series solution for variance of population

In this section, we show the derivation of the variance result in §3.2. The calculations closely follow [67] on the stochastic cable equation. As noted in equation (3.8), the SPDE (2.2) can formally be solved by integrating the stochastic point source noise against the Green's function $G(x, t)$ that satisfies

$$\partial_t G = D\partial_{xx}G - \gamma G,$$
$$G(-R, t) = 0, \ G(R, t) = 0, \tag{C 1}$$
$$G(x, 0) = \delta(x - z).$$

This Green's function is [69]:

$$G(x, t) = \sum_{m=1}^{\infty} g_m(t)\phi_m(x)\phi_m(z), \tag{C 2}$$

with

$$g_m(t) := e^{-t(\gamma + D\pi^2 m/(2R)^2)} := e^{-\mu_m t}$$

$$\phi_m(x) := \sqrt{1/R}\,\sin(m\pi(x+R)/2R). \tag{C3}$$

By equation (3.6), computation of the variance of n requires knowledge of the second moment of the underlying stochastic intensity. That is:

$$C(x,y,t) = \langle u(x,t)u(y,t) \rangle$$

$$= \left\langle \left[\int_0^t G(x,t-s_1)\Lambda(s_1)\mathrm{d}s_1 \right]\left[\int_0^t G(x,t-s_2)\Lambda(s_2)\mathrm{d}s_2 \right] \right\rangle \mathrm{d}s_2\mathrm{d}s_1$$

$$= \int_0^t \int_0^t G(x,t-s_1)G(y,t-s_2)\langle \Lambda(s_1)\Lambda(s_2) \rangle \mathrm{d}s_2\mathrm{d}s_1. \tag{C4}$$

The autocorrelation of the dichotomous process is classical [62], and satisfies

$$\langle \Lambda(s_1)\Lambda(s_2) \rangle = \frac{\lambda^2 \alpha\beta}{(\alpha+\beta)^2}e^{-|s_1-s_2|(\alpha+\beta)} := \kappa e^{-|s_1-s_2|/\omega}. \tag{C5}$$

Ultimately, this gives us

$$C(x,y,t) = \kappa \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \phi_n(x)\phi_n(z)\phi_m(y)\phi_m(z) \int_0^t \int_0^t g_n(t-s_1)g_m(t-s_2)e^{-|s_1-s_2|/\tau}\mathrm{d}s_2\mathrm{d}s_1. \tag{C6}$$

We defer evaluating this, noting that ultimately we want to spatially integrate $C(x,y,t)$, and that

$$\int_{-R}^{R}\int_{-R}^{R}\phi_n(x)\phi_m(y)\phi_n(z)\phi_m(z)\mathrm{d}x\mathrm{d}y = \frac{4\left((-1)^m-1\right)\left((-1)^n-1\right)\phi_m(z)\phi_n(z)}{\pi^2 mn}. \tag{C7}$$

With some help from `Mathematica`, we can also integrate

$$\tilde{g}(t) := \int_0^t \int_0^t g_m(t-s_1)g_m(t-s_2)e^{-|s_1-s_2|\omega}\mathrm{d}s_2\mathrm{d}s_1. \tag{C8}$$

The expression is unwieldy, but after taking $\lim_{t\to\infty}$, we get

$$\tilde{g} = \lim_{t\to\infty}\tilde{g}(t) = \frac{\mu_m + \mu_n + 2\omega}{(\mu_m+\omega)(\mu_n+\omega)(\mu_m+\mu_n)}. \tag{C9}$$

The temporal terms (equation (C 9)) and spatial terms (equation (C 7)) can be combined into equation (C 6) to nearly give the infinite sum in the main text. The last ingredient is the observation that only odd terms of equation (C 7) are non-zero, so we re-index $m \to 2m-1$, $n \to 2n-1$, and equation (3.7) is recovered. Figures with predictions from the series truncate after 1000 terms.

References

1. Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002 Stochastic gene expression in a single cell. *Science* **297**, 1183–1186. (doi:10.1126/science.1070919)

2. Blake WJ, KÆrn M, Cantor CR, Collins JJ. 2003 Noise in eukaryotic gene expression. *Nature* **422**, 633–637. (doi:10.1038/nature01546)

3. Raj A, van Oudenaarden A. 2008 Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226. (doi:10.1016/j.cell.2008.09.050)

4. Munsky B, Neuert G, van Oudenaarden A. 2012 Using gene expression noise to understand gene regulation. *Science* **336**, 183–187. (doi:10.1126/science.1216379)

5. Sanchez A, Golding I. 2013 Genetic determinants and cellular constraints in noisy gene expression. *Science* **342**, 1188–1193. (doi:10.1126/science.1242975)

6. McAdams HH, Arkin A. 1997 Stochastic mechanisms in gene expression. *Proc. Natl Acad. Sci. USA* **94**, 814–819. (doi:10.1073/pnas.94.3.814)

7. Smiley MW, Proulx SR. 2010 Gene expression dynamics in randomly varying environments. *J. Math. Biol.* **61**, 231–251. (doi:10.1007/s00285-009-0298-z)

8. Paulsson J. 2005 Models of stochastic gene expression. *Phys. Life Rev.* **2**, 157–175. (doi:10.1016/j.plrev.2005.03.003)

9. Friedman N, Cai L, Xie XS. 2006 Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Phys. Rev. Lett.* **97**, 168302. (doi:10.1103/PhysRevLett.97.168302)

10. Shahrezaei V, Swain PS. 2008 Analytical distributions for stochastic gene expression. *Proc. Natl Acad. Sci. USA* **105**, 17256–17261. (doi:10.1073/pnas.0803850105)

11. Kumar N, Platini T, Kulkarni RV. 2014 Exact distributions for stochastic gene expression models with bursting and feedback. *Phys. Rev. Lett.* **113**, 268105. (doi:10.1103/PhysRevLett.113.268105)

12. Kumar N, Singh A, Kulkarni RV. 2015 Transcriptional bursting in gene expression: analytical results for general stochastic models. *PLoS Comput. Biol.* **11**, e1004292. (doi:10.1371/journal.pcbi.1004292)

13. Cao Z, Grima R. 2020 Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells. *Proc. Natl Acad. Sci. USA* **117**, 4682–4692. (doi:10.1073/pnas.1910888117)

14. Kim JK, Marioni JC. 2013 Inferring the kinetics of stochastic gene expression from single-cell RNA-sequencing data. *Genome Biol.* **14**, R7. (doi:10.1186/gb-2013-14-1-r7)

15. Zechner C, Unger M, Pelet S, Peter M, Koeppl H. 2014 Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings. *Nat. Methods* **11**, 197–202. (doi:10.1038/nmeth.2794)

16. Munsky B, Fox Z, Neuert G. 2015 Integrating single-molecule experiments and discrete stochastic models to understand heterogeneous gene transcription dynamics. *Methods* **85**, 12–21. (doi:10.1016/j.ymeth.2015.06.009)

17. Gorin G, Vastola JJ, Pachter L. 2023 Studying stochastic systems biology of the cell with single-cell genomics data. *Cell Syst.* **14**, 822–843. (doi:10.1016/j.cels.2023.08.004)

18. Kilic Z, Schweiger M, Moyer C, Shepherd D, Pressé S. 2023 Gene expression model inference from snapshot RNA data using Bayesian non-parametrics. *Nat. Comput. Sci.* **3**, 174–183. (doi:10.1038/s43588-022-00392-0)

19. Weidemann DE, Holehouse J, Singh A, Grima R, Hauf S. 2023 The minimal intrinsic stochasticity of constitutively expressed eukaryotic genes is sub-Poissonian. *Sci. Adv.* **9**, eadh5138. (doi:10.1126/sciadv.adh5138)

20. Grima R, Esmenjaud PM. 2024 Quantifying and correcting bias in transcriptional parameter inference from single-cell data. *Biophys. J.* **123**, 4–30. (doi:10.1016/j.bpj.2023.10.021)

21. Gorin G, Yoshida S, Pachter L. 2023 Assessing Markovian and delay models for single-nucleus RNA sequencing. *Bull. Math. Biol.* **85**, 114. (doi:10.1007/s11538-023-01213-9)

22. Berry S, Pelkmans L. 2022 Mechanisms of cellular mRNA transcript homeostasis. *Trends Cell Biol.* **32**, 655–668. (doi:10.1016/j.tcb.2022.05.003)

23. Vargas DY, Raj A, Marras SAE, Kramer FR, Tyagi S. 2005 Mechanism of mRNA transport in the nucleus. *Proc. Natl Acad. Sci. USA* **102**, 17008–17013. (doi:10.1073/pnas.0505580102)

24. Mor A, Suliman S, Ben-Yishay R, Yunger S, Brody Y, Shav-Tal Y. 2010 Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. *Nat. Cell Biol.* **12**, 543–552. (doi:10.1038/ncb2056)

25. Hansen MMK, Desai RV, Simpson ML, Weinberger LS. 2018 Cytoplasmic amplification of transcriptional noise generates substantial cell-to-cell variability. *Cell Syst.* **7**, 384–397. (doi:10.1016/j.cels.2018.08.002)

26. Xia C, Fan J, Emanuel G, Hao J, Zhuang X. 2019 Spatial transcriptome profiling by merfish reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proc. Natl Acad. Sci. USA* **116**, 19490–19499. (doi:10.1073/pnas.1912459116)

27. Lefaudeux D, Sen S, Jiang K, Hoffmann A. 2022 Kinetics of mRNA nuclear export regulate innate immune response gene expression. *Nat. Commun.* **13**, 7197. (doi:10.1038/s41467-022-34635-5)

28. Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. 2008 Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877–879. (doi:10.1038/nmeth.1253)

29. Ding F, Elowitz MB. 2019 Constitutive splicing and economies of scale in gene expression. *Nat. Struct. Mol. Biol.* **26**, 424–432. (doi:10.1038/s41594-019-0226-x)

30. Eng CHL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, Yun J. 2019 Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature* **568**, 235–239. (doi:10.1038/s41586-019-1049-y)

31. Li G, Neuert G. 2019 Multiplex RNA single molecule FISH of inducible mRNAs in single yeast cells. *Sci. Data* **6**, 94. (doi:10.1038/s41597-019-0106-6)

32. Gómez-Schiavon M, Chen LF, West AE, Buchler NE. 2017 BayFish: Bayesian inference of transcription dynamics from population snapshots of single-molecule RNA FISH in single cells. *Genome Biol.* **18**, 164. (doi:10.1186/s13059-017-1297-9)

33. Samacoits A *et al.* 2018 A computational framework to study sub-cellular RNA localization. *Nat. Commun.* **9**, 4584. (doi:10.1038/s41467-018-06868-w)

34. Imbert A, Ouyang W, Safieddine A, Coleno E, Zimmer C, Bertrand E, Walter T, Mueller F. 2022 FISH-Quant v2: a scalable and modular tool for smFISH image analysis. *RNA* **28**, 786–795. (doi:10.1261/rna.079073.121)

35. Wyart M, Botstein D, Wingreen NS. 2010 Evaluating gene expression dynamics using pairwise RNA FISH data. *PLoS Comput. Biol.* **6**, e1000979. (doi:10.1371/journal.pcbi.1000979)

36. Singh A, Bokes P. 2012 Consequences of mRNA transport on stochastic variability in protein levels. *Biophys. J.* **103**, 1087–1096. (doi:10.1016/j.bpj.2012.07.015)

37. Sturrock M, Li S, Shahrezaei V. 2017 The influence of nuclear compartmentalisation on stochastic dynamics of self-repressing gene expression. *J. Theor. Biol.* **424**, 55–72. (doi:10.1016/j.jtbi.2017.05.003)

38. Lin YT, Buchler NE. 2019 Exact and efficient hybrid Monte Carlo algorithm for accelerated bayesian inference of gene expression models from snapshots of single-cell transcripts. *J. Chem. Phys.* **151**, 024106. (doi:10.1063/1.5110503)

39. Filatova T, Popović N, Grima R. 2022 Modulation of nuclear and cytoplasmic mRNA fluctuations by time-dependent stimuli: analytical distributions. *Math. Biosci.* **347**, 108828. (doi:10.1016/j.mbs.2022.108828)

40. Wang Y, Szavits-Nossan J, Cao Z, Grima R. 2024 Joint distribution of nuclear and cytoplasmic mRNA levels in stochastic models of gene expression: analytical results and parameter inference. *Syst. Biol.* (doi:10.1101/2024.04.29.591679)

41. Balakrishnan S, Raju SR, Barua A, Pradeep RP, Ananthasuresh GK. 2021 Two nondimensional parameters for characterizing the nuclear morphology. *Biophys. J.* **120**, 4698–4709. (doi:10.1016/j.bpj.2021.09.035)

42. Yunger S, Rosenfeld L, Garini Y, Shav-Tal Y. 2010 Single-allele analysis of transcription kinetics in living mammalian cells. *Nat. Methods* **7**, 631–633. (doi:10.1038/nmeth.1482)

43. Smith S, Grima R. 2019 Spatial stochastic intracellular kinetics: a review of modelling approaches. *Bull. Math. Biol.* **81**, 2960–3009. (doi:10.1007/s11538-018-0443-1)

44. Gorin G, Vastola JJ, Fang M, Pachter L. 2022 Interpretable and tractable models of transcriptional noise for the rational design of single-molecule quantification experiments. *Nat. Commun.* **13**, 7620. (doi:10.1038/s41467-022-34857-7)

45. Vo HD, Forero-Quintero LS, Aguilera LU, Munsky B. 2023 Analysis and design of single-cell experiments to harvest fluctuation information while rejecting measurement noise. *Front. Cell Dev. Biol.* **11**, 1133994. (doi:10.3389/fcell.2023.1133994)

46. Jiang Q, Fu X, Yan S, Li R, Du W, Cao Z, Qian F, Grima R. 2021 Neural network aided approximation and parameter inference of non-Markovian models of gene expression. *Nat. Commun.* **12**, 2618. (doi:10.1038/s41467-021-22919-1)

47. Gorin G, Carilli M, Chari T, Pachter L. 2024 Spectral neural approximations for models of transcriptional dynamics. *Biophys. J.* **123**, 2892–2901. (doi:10.1016/j.bpj.2024.04.034)

48. Schnoerr D, Grima R, Sanguinetti G. 2016 Cox process representation and inference for stochastic reaction–diffusion processes. *Nat. Commun.* **7**, 11729. (doi:10.1038/ncomms11729)

49. Miles CE, McKinley SA, Ding F, Lehoucq RB. 2024 Inferring stochastic rates from heterogeneous snapshots of particle positions. *Bull. Math. Biol.* **86**, 74. (doi:10.1007/s11538-024-01301-4)

50. Tadakuma H, Ishihama Y, Shibuya T, Tani T, Funatsu T. 2006 Imaging of single mRNA molecules moving within a living cell nucleus. *Biochem. Biophys. Res. Commun.* **344**, 772–779. (doi:10.1016/j.bbrc.2006.03.202)

51. Tunnacliffe E, Chubb JR. 2020 What is a transcriptional burst? *Trends Genet.* **36**, 288–297. (doi:10.1016/j.tig.2020.01.003)

52. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011 Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**, 472–474. (doi:10.1126/science.1198817)

53. Dar RD, Razooky BS, Singh A, Trimeloni TV, McCollum JM, Cox CD, Simpson ML, Weinberger LS. 2012 Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc. Natl Acad. Sci. USA* **109**, 17454–17459. (doi:10.1073/pnas.1213530109)

54. Munsky B, Li G, Fox ZR, Shepherd DP, Neuert G. 2018 Distribution shapes govern the discovery of predictive models for gene regulation. *Proc. Natl Acad. Sci. USA* **115**, 7533–7538. (doi:10.1073/pnas.1804060115)

55. Luo S, Zhang Z, Wang Z, Yang X, Chen X, Zhou T, Zhang J. 2023 Inferring transcriptional bursting kinetics from single-cell snapshot data using a generalized telegraph model. *R. Soc. Open Sci.* **10**, 221057. (doi:10.1098/rsos.221057)

56. Gardiner CW, Chaturvedi S. 1977 The Poisson representation. I. A new technique for chemical master equations. *J. Stat. Phys.* **17**, 429–468. (doi:10.1007/BF01014349)

57. Chaturvedi S, Gardiner CW. 1978 The Poisson representation. II Two-time correlation functions. *J. Stat. Phys.* **18**, 501–522. (doi:10.1007/BF01014520)

58. Stinchcombe AR, Peskin CS, Tranchina D. 2012 Population density approach for discrete mRNA distributions in generalized switching models for stochastic gene expression. *Phys. Rev. E* **85**, 061919. (doi:10.1103/PhysRevE.85.061919)

59. Iyer-Biswas S, Jayaprakash C. 2014 Mixed poisson distributions in exact solutions of stochastic autoregulation models. *Phys. Rev. E* **90**, 052712. (doi:10.1103/PhysRevE.90.052712)

60. Sugár I, Simon I. 2014 Self-regulating genes: exact steady state solution by using Poisson representation. *Open Phys.* **12**, 615–627. (doi:10.2478/s11534-014-0497-0)

61. Wang X, Li Y, Jia C. 2023 Poisson representation: a bridge between discrete and continuous models of stochastic gene regulatory networks. *J. R. Soc. Interface* **20**, 20230467. (doi:10.1098/rsif.2023.0467)

62. Fuliński A. 1994 Non-Markovian noise. *Phys. Rev. E* **50**, 2668–2681. (doi:10.1103/PhysRevE.50.2668)

63. Jia C, Li Y. 2023 Analytical time-dependent distributions for gene expression models with complex promoter switching mechanisms. *SIAM J. Appl. Math.* **83**, 1572–1602. (doi:10.1137/22M147219X)

64. Ancey C, Bohorquez P, Heyman J. 2015 Stochastic interpretation of the advection-diffusion equation and its relevance to bed load transport. *J. Geophys. Res.* **120**, 2529–2551. (doi:10.1002/2014JF003421)

65. Lawley SD, Mattingly JC, Reed MC. 2015 Stochastic switching in infinite dimensions with applications to random parabolic PDE. *SIAM J. Math. Anal.* **47**, 3035–3063. (doi:10.1137/140976716)

66. Heyman J, Ma HB, Mettra F, Ancey C. 2014 Spatial correlations in bed load transport: evidence, importance, and modeling. *J. Geophys. Res.* **119**, 1751–1767. (doi:10.1002/2013JF003003)

67. Tuckwell HC, Wan FYM, Wong YS. 1984 The interspike interval of a cable model neuron with white noise input. *Biol. Cybern.* **49**, 155–167. (doi:10.1007/BF00334461)

68. Tuckwell HC, Wan FYM, Rospars JP. 2002 A spatial stochastic neuronal model with Ornstein–Uhlenbeck input current. *Biol. Cybern.* **86**, 137–145. (doi:10.1007/s004220100283)

69. Stakgold I. 2000 Boundary Value Problems of Mathematical Physics: Volume 1. In *Classics in applied mathematics no. 29*. Philadelphia, PA: Society for Industrial and Applied Mathematics. (doi:10.1137/1.9780898719888)

70. Lindsay AE, Bernoff AJ, Ward MJ. 2017 First passage statistics for the capture of a Brownian particle by a structured spherical target with multiple surface traps. *Multiscale Model. Simul.* **15**, 74–109. (doi:10.1137/16M1077659)

71. Leech V, Hazel JW, Gatlin JC, Lindsay AE, Manhart A. 2022 Mathematical modeling accurately predicts the dynamics and scaling of nuclear growth in discrete cytoplasmic volumes. *J. Theor. Biol.* **533**, 110936. (doi:10.1016/j.jtbi.2021.110936)

72. Singer A, Schuss Z, Osipov A, Holcman D. 2008 Partially reflected diffusion. *SIAM J. Appl. Math.* **68**, 844–868. (doi:10.1137/060663258)

73. Klindziuk A, Kolomeisky AB. 2021 Understanding the molecular mechanisms of transcriptional bursting. *Phys. Chem. Chem. Phys.* **23**, 21399–21406. (doi:10.1039/D1CP03665C)

74. Garner RM, Molines AT, Theriot JA, Chang F. 2024 Incorporating the diffusivity gradient term in Brownian dynamics simulations of diffusion. *Biophys. J.* **123**, 2641–2643. (doi:10.1016/j.bpj.2024.07.032)

75. Pacheco-Pozo A, Balcerek M, Wyłomanska A, Burnecki K, Sokolov IM, Krapf D. 2024 Langevin equation in heterogeneous landscapes: how to choose the interpretation. *Phys. Rev. Lett.* **133**, 067102. (doi:10.1103/PhysRevLett.133.067102)

76. Moller J, Waagepetersen RP. 2003 *Statistical inference and simulation for spatial point processes*. New York, NY, USA: CRC Press.

77. Simpson MJ, Maclaren OJ. 2023 Profile-wise analysis: a profile likelihood-based workflow for identifiability analysis, estimation, and prediction with mechanistic mathematical models. *PLoS Comput. Biol.* **19**, e1011515. (doi:10.1371/journal.pcbi.1011515)

78. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J. 2009 Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **25**, 1923–1929. (doi:10.1093/bioinformatics/btp358)

79. Coulon A, Ferguson ML, de Turris V, Palangat M, Chow CC, Larson DR. 2014 Kinetic competition during the transcription cycle results in stochastic RNA processing. eLife **3**, e03939. (doi:10.7554/eLife.03939)

80. Roussel MR, Tang T. 2012 Simulation of mRNA diffusion in the nuclear environment. *IET Syst. Biol.* **6**, 125–133. (doi:10.1049/iet-syb.2011.0032)

81. Bresson S, Tollervey D. 2018 Surveillance-ready transcription: nuclear RNA decay as a default fate. *Open Biol.* **8**, 170270. (doi:10.1098/rsob.170270)

82. Bornfleth H, Edelmann P, Zink D, Cremer T, Cremer C. 1999 Quantitative motion analysis of subchromosomal foci in living cells using four-dimensional microscopy. *Biophys. J.* **77**, 2871–2886. (doi:10.1016/S0006-3495(99)77119-5)

83. Bohrer CH, Larson DR. 2023 Synthetic analysis of chromatin tracing and live-cell imaging indicates pervasive spatial coupling between genes. *eLife* **12**, e81861. (doi:10.7554/eLife.81861)

84. Jiao F, Li J, Liu T, Zhu Y, Che W, Bleris L, Jia C. 2024 What can we learn when fitting a simple telegraph model to a complex gene expression model? *PLoS Comput. Biol.* **20**, e1012118. (doi:10.1371/journal.pcbi.1012118)

85. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. 2015 Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090. (doi:10.1126/science.aaa6090)

86. Dattani J, Barahona M. 2017 Stochastic models of gene transcription with upstream drives: exact solution and sample path characterization. *J. R. Soc. Interface* **14**, 20160833. (doi:10.1098/rsif.2016.0833)

87. Herbach U. 2019 Stochastic gene expression with a multistate promoter: breaking down exact distributions. *SIAM J. Appl. Math.* **79**, 1007–1029. (doi:10.1137/18M1181006)

88. Lippitt W, Sethuraman S, Tang X. 2022 Stationarity and inference in multistate promoter models of stochastic gene expression via stick-breaking measures. *SIAM J. Appl. Math.* **82**, 1953–1986. (doi:10.1137/21M1440876)

89. Wang Z, Luo S, Zhang Z, Zhou T, Zhang J. 2023 4D nucleome equation predicts gene expression controlled by long-range enhancer-promoter interaction. *PLoS Comput. Biol.* **19**, e1011722. (doi:10.1371/journal.pcbi.1011722)

90. Ma M, Szavits-Nossan J, Singh A, Grima R. 2024 Analysis of a detailed multi-stage model of stochastic gene expression using queueing theory and model reduction. *Math. Biosci.* **373**, 109204. (doi:10.1016/j.mbs.2024.109204)

91. Karamched BR, Miles CE. 2023 Stochastic switching of delayed feedback suppresses oscillations in genetic regulatory systems. *J. R. Soc. Interface* **20**, 20230059. (doi:10.1098/rsif.2023.0059)

92. Jia C, Grima R. 2023 Coupling gene expression dynamics to cell size dynamics and cell cycle events: exact and approximate solutions of the extended telegraph model. *iScience* **26**, 105746. (doi:10.1016/j.isci.2022.105746)

93. Zon JS van, Morelli MJ, Tănase-Nicola S, ten Wolde PR. 2006 Diffusion of transcription factors can drastically enhance the noise in gene expression. *Biophys. J.* **91**, 4350–4367. (doi:10.1529/biophysj.106.086157)

94. Sturrock M, Hellander A, Matzavinos A, Chaplain MAJ. 2013 Spatial stochastic modelling of the Hes1 gene regulatory network: intrinsic noise can explain heterogeneity in embryonic stem cell differentiation. *J. R. Soc. Interface* **10**, 20120988. (doi:10.1098/rsif.2012.0988)

95. Coté A *et al*. 2024 Post-transcriptional splicing can occur in a slow-moving zone around the gene. *eLife* **12**, RP91357. (doi:10.7554/eLife.91357)

96. Daigle BJ, Soltani M, Petzold LR, Singh A. 2015 Inferring single-cell gene expression mechanisms using stochastic simulation. *Bioinformatics* **31**, 1428–1435. (doi:10.1093/bioinformatics/btv007)

97. Cao Z, Chen R, Xu L, Zhou X, Fu X, Zhong W, Grima R. 2024 Efficient and scalable prediction of stochastic reaction–diffusion processes using graph neural networks. *Math. Biosci.* **375**, 109248. (doi:10.1016/j.mbs.2024.109248)

98. Lenive O, W Kirk PD, H Stumpf MP. 2016 Inferring extrinsic noise from single-cell gene expression data using approximate Bayesian computation. *BMC Syst. Biol.* **10**, 81. (doi:10.1186/s12918-016-0324-x)

99. Miles C. 2025 chris-miles/BurstDiffusion1D. *Zenodo*. (doi:10.5281/zenodo.14599827)

100. Roques L, Allard D, Soubeyrand S. 2022 Spatial statistics and stochastic partial differential equations: a mechanistic viewpoint. *Spat. Stat.* **50**, 100591. (doi:10.1016/j.spasta.2022.100591)