



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Steady-state fluctuations of a genetic feedback loop with fluctuating rate parameters using the unified colored noise approximation

Citation for published version:

Holehouse, J, Gupta, A & Grima, R 2020, 'Steady-state fluctuations of a genetic feedback loop with fluctuating rate parameters using the unified colored noise approximation', *Journal of Physics A: Mathematical and Theoretical*, vol. 53, 405601. <https://doi.org/10.1088/1751-8121/aba4d0>

Digital Object Identifier (DOI):

[10.1088/1751-8121/aba4d0](https://doi.org/10.1088/1751-8121/aba4d0)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Journal of Physics A: Mathematical and Theoretical

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Steady-state fluctuations of a genetic feedback loop with fluctuating rate parameters using the unified colored noise approximation

James Holehouse¹, Abhishek Gupta¹ and Ramon Grima² 

School of Biological Sciences, University of Edinburgh, United Kingdom

E-mail: Ramon.Grima@ed.ac.uk

Received 4 April 2020, revised 19 June 2020

Accepted for publication 10 July 2020

Published 1 September 2020



CrossMark

Abstract

A common model of stochastic auto-regulatory gene expression describes promoter switching via cooperative protein binding, effective protein production in the active state and dilution of proteins. Here we consider an extension of this model whereby colored noise with a short correlation time is added to the reaction rate parameters—we show that when the size and timescale of the noise is appropriately chosen it accounts for fast reactions that are not explicitly modeled, e.g., in models with no mRNA description, fluctuations in the protein production rate can account for rapid multiple stages of nuclear mRNA processing which precede translation in eukaryotes. We show how the unified colored noise approximation can be used to derive expressions for the protein number distribution that is in good agreement with stochastic simulations. We find that even when the noise in the rate parameters is small, the protein distributions predicted by our model can be significantly different than models assuming constant reaction rates.

Keywords: Fokker–Planck equation, gene regulatory networks, stochastic gene expression, stochastic simulations

(Some figures may appear in colour only in the online journal)

¹These authors contributed equally.

²Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

Proteins perform a large range of cellular functions and hence it is of great interest to understand how the genes that produce them operate. Autoregulation is a mechanism to regulate gene expression whereby proteins expressed by a certain gene can subsequently bind to the same gene and cause an increase or a decrease in its expression (positive and negative feedback, respectively) [1]. Autoregulation is common; for example in *Escherichia coli* it is estimated that 40% of all transcription factors are self-regulated [2, 3].

For at least two decades, it has been known that gene expression is inherently stochastic [4, 5], and as such the modeling of gene regulatory networks must account for this stochasticity. Following van Kampen [6], given a system of interest, noise can be seen as originating from two different sources: (i) noise that is inherent to the system itself and cannot be turned off, also called internal or intrinsic noise; (ii) noise coming from a source outside the system of interest, known as external or extrinsic noise. This noise can be switched off via a coupling parameter. If we specify a system of interest that is described by a set of reactions with constant rate parameters, then it follows that any fluctuations in the molecule numbers must be due to the inherent randomness in the time at which the reactions fire, and hence the noise is intrinsic. In contrast, if we add fluctuations to the rate parameters to account for external processes, then it follows that this noise is extrinsic. For example, if one models mRNA transcription from a gene by a first-order reaction with a constant rate parameter then one is only modeling intrinsic noise. However, if one adds noise to the transcription rate to account for fluctuating numbers of polymerases and transcription factors that are not explicitly described in the system, then one is modeling both intrinsic and extrinsic noise sources. Note that these definitions of intrinsic and extrinsic noise are generally different from, and not to be confused with, the definitions proposed using dual-reporter methods in [7].

The division of noise into these two categories is of course artificial but it is useful from a conceptual and modeling point of view. The simulation of stochastic biochemical processes is most commonly done using the stochastic simulation algorithm (SSA) [8] which assumes that the rate parameter of a reaction will not change in the interval between two successive reaction events, i.e., it models intrinsic noise only. While this may be the case in many situations, it is not generally true. This is because whenever we have an effective reaction that lumps together a large number of intermediate reactions (a multi-stage reaction process), we are making the inherent assumption that these intermediate reactions occur very fast and hence naturally the effective rate parameter is fluctuating on a fast timescale.

Taking into account these fluctuations is however not a simple feat. The chemical master equation (CME, [6, 9]) describing the Markov process simulated by the SSA has been solved exactly or approximately to obtain the protein number distribution in steady-state for a wide variety of models of autoregulation [10–22], provided the rate parameters are assumed to be constant. There are however a number of studies that have analyzed stochastic models with fluctuating rate parameters. The importance of studying these models stems from the fact that they potentially offer a compromise between model precision (how well can the model capture the complexity of the underlying biochemical dynamics by means of fluctuating parameters) and analytical tractability (how easy it is to solve the stochastic model). Modifications of the linear noise approximation (LNA) (a type of Fokker–Planck approximation of the CME) incorporating noise in the rate parameters have proved popular to approximate moments for systems subject to small magnitudes of noise with certain properties: (i) for time-independent Gaussian colored noise [23, 24] and (ii) more realistic lognormally distributed noise [25]. Wentzel–Kramers–Brillouin methods have also been utilised for cases where the correlation time of the colored noise is tending either to zero or to infinity [26]. These methods

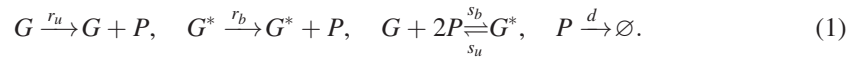
provide probability distributions for systems where the noise on the rate parameters is drawn from a negative binomial distribution, however their analysis does not easily translate to finding good approximations for steady state probability distributions where the correlation time of colored noise is neither small or large.

The focus of the present article is threefold: (i) to provide a general method by which one can obtain analytical expressions for the steady-state protein distributions of auto-regulatory gene circuits with fluctuating rate parameters, through the use of the *unified colored noise approximation* (UCNA) [27], (ii) to use this method to investigate the effects that extrinsic noise of different magnitude and timescales has on auto-regulatory gene expression and (iii) to show how the colored noise formalism can be used to describe complex models of autoregulation that involve multi-stage protein production and multi-stage protein degradation. We note that the UCNA was previously utilised in a gene expression context [28] for linear reaction networks that are deterministically monostable and in which there is no feedback mechanism. Our analysis goes further, exploring the addition of colored noise to a non-linear reaction network which expresses deterministic bistability, whilst also incorporating intrinsic fluctuations from the core gene expression processes.

The structure of our paper is as follows. In section 2 we introduce the cooperative auto-regulatory reaction scheme that we will study in this article. We also show that for non-fluctuating rate parameters, the analytical protein distribution given by the chemical Fokker–Planck equation (FPE) provides an excellent approximation of the protein distribution solution of the CME, in the limit of fast gene switching. In section 3 we add colored noise to each reaction in the auto-regulatory reaction scheme (assuming fast gene switching) and use the UCNA to derive the protein number distribution solution of the chemical FPE. The solution is shown to be in good agreement with a SSA modified to account for extrinsic noise on the rate parameters. We also use the solution to investigate the effect that extrinsic noise has on the number of modes of the protein distribution and clarify the limits of the UCNA derivation, including the three main conditions which cause it to breakdown. In section 4 we extend the analysis to the limit of slow gene switching by introducing a conditional version of the UCNA. In section 5 we show two examples of how one can successfully model complex auto-regulatory systems by means of simpler ones with colored noise on the reaction rate parameters, here done for multi-stage protein production and multi-stage degradation. We conclude in section 6 with a discussion of our results and further problems to be addressed on this topic.

2. Approximate solution for autoregulation with non-fluctuating rate parameters

We consider the reaction scheme for a genetic non-bursty cooperative feedback loop, where for simplicity we neglect the presence of mRNA:



The reactions $G \xrightarrow{r_u} G + P$ and $G^* \xrightarrow{r_b} G^* + P$ model the production of protein P in each gene state, $G + 2P \xrightleftharpoons[s_u]{s_b} G^*$ models the binding and unbinding of the gene to the proteins (with cooperativity 2), and $P \xrightarrow{d} \emptyset$ models the dilution/degradation of proteins inside the cell. For simplicity we assume that there is only one gene copy present in the system and it can be in one of two states, G or G^* , at any one time (some models in the literature consider more states than two

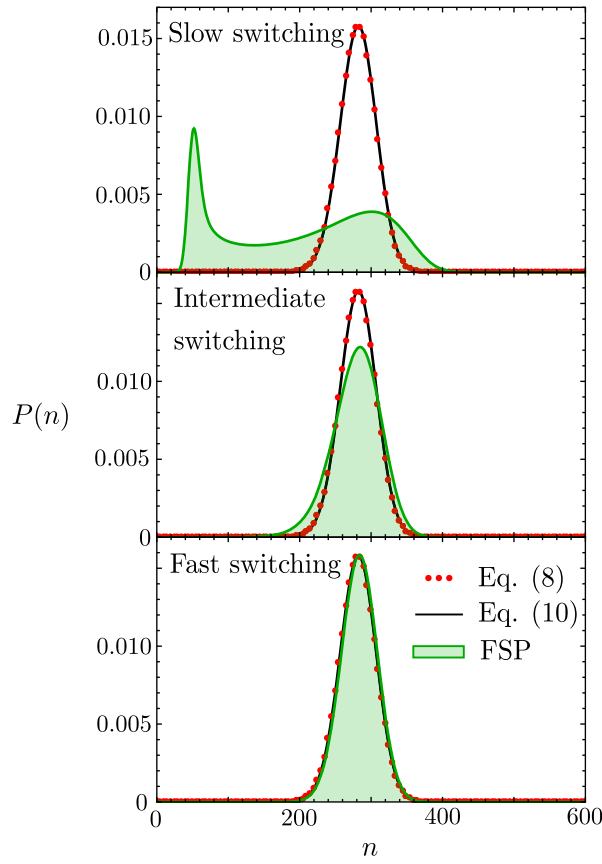


Figure 1. Comparison of the heuristic reduced master equation solution from equation (8) (red dots), the FPE solution from equation (10) (black line) and the solution of the full cooperative network using FSP (green shaded region). Shared parameters in each plot are $r_u = 50$, $r_b = 400$, $\Omega = 200$ and $d = 1$. The FSP gives the exact solution for a truncated state space chosen such that the neglected probability mass is negligible. The top plot shows distributions for the case $s_u = s_b = 5$, where clearly the heuristic master equation and FPE solutions are a poor approximation of the FSP. The middle plot shows distributions for the case $s_u = s_b = 50$ where we can observe a convergence of the heuristic master equation and FPE solutions toward the FSP solution. The bottom plot shows excellent agreement of the FSP with the heuristic master equation and FPE solutions for fast switching where $s_u = s_b = 5 \times 10^3$.

[29–31]). Note also that the reaction modeling protein binding to the gene is to be understood as an effective reaction in cases where the protein binds to enhancer regions rather than directly to the promoter [32]. Before considering the addition of colored noise to the reaction rate parameters above, we first consider the solution with constant rate parameters to provide a reference point for approximations made in section 3, and to clarify the approximation of a CME by a one variable chemical FPE.

The CME for the reaction scheme in equation (1) does not have a known exact solution, even at steady-state for constant reaction rate parameters, and so approximations are necessary. Note that in what follows, we will use the terminology ‘reaction rate parameters’ and ‘rates’ interchangeably. We first consider the limit of fast gene switching—i.e., the frequency of gene

activation and inactivation events is much larger than the frequency of any other reaction in the system. Later in section 4 we will discuss approximations for the slow switching limit. Where $[g^*]$ and $[g]$ are the deterministic mean number of bound and unbound gene respectively and $[n]$ is the mean protein number, the rate equations for the reaction scheme in equation (1) are:

$$\frac{d[g]}{dt} = s_u[g^*] - \frac{s_b}{\Omega^2}[g][n]^2, \quad (2)$$

$$\frac{d[n]}{dt} = 2s_u[g^*] - \frac{2s_b}{\Omega^2}[g][n]^2 + r_u[g] + r_b[g^*] - d[n], \quad (3)$$

where $[g] + [g^*] = 1$. For clarity we state the units of each rate parameter: r_u , r_b , d and s_u have units of s^{-1} , and s_b has units of $\text{volume}^2 s^{-1}$. This ensures a matching of the units with the left-hand side of equations (2) and (3), which has units of s^{-1} . In the fast switching limit, the gene rapidly equilibrates to quasi-steady state conditions, i.e., $d[g]/dt \approx d[g^*]/dt \approx 0$ and hence the deterministic rate equation for mean protein number reduces to a much simpler form:

$$\frac{d[n]}{dt} = \frac{Lr_u + r_b([n]/\Omega)^2}{L + ([n]/\Omega)^2} - d[n], \quad (4)$$

where $L = s_u/s_b$. Note that the reaction scheme here described exhibits deterministic bistability over some regions of the parameter space. This equation is consistent with a birth–death process where proteins are produced via a reaction with a rate that is dependent on the number of proteins and are destroyed by a first-order reaction [33]. The CME for this reduced process is given by:

$$\frac{dP_a(n, t)}{dt} = T^+(n-1)P_a(n-1, t) + T^-(n+1)P_a(n+1, t) - (T^+(n) + T^-(n))P_a(n, t), \quad (5)$$

where $P_a(n, t)$ is the probability that at a time t there are n proteins in the system; $T^+(n)$ and $T^-(n)$ are the propensities of protein production and degradation respectively. The subscript a denotes that this is the probability for the reduced system, an *approximate* solution to the master equation of the full system. $T^+(n)dt$ is the probability, given n proteins are in the system, that a protein production reaction occurs, increasing the protein number of the system by 1, in the time interval $[t, t + dt)$. Similarly, $T^-(n)dt$ is the probability, given n proteins are in the system, that a protein degradation event occurs, decreasing the protein number by 1, in the time interval $[t, t + dt)$. These propensities are given by:

$$T^+(n) = \frac{r_u L + r_b(n/\Omega)^2}{L + (n/\Omega)^2}, \quad (6)$$

$$T^-(n) = d n. \quad (7)$$

These propensities are deduced directly from the form of the effective rate equation in equation (4). Essentially, the probability for a particular reaction per unit time is taken to be the same as the reaction rate in the effective deterministic rate equation with $[n]$ replaced by n . *We emphasise that while this appears to be a heuristic rule with no apparent fundamental microscopic basis, it has been shown that the reduced master equation based on it provides an accurate approximation to the SSA of the full reaction system in fast gene switching conditions provided the low protein number states are rarely visited [15, 33].*

The exact steady state solution of the one variable master equation given by equation (5) can be found using standard methods [34]:

$$P_a(n) = P_a(0) \prod_{z=1}^n \frac{T^+(z-1)}{T^-(z)}, \quad (8)$$

where $P_a(0)$ is the steady state probability evaluated at $n = 0$ (acting effectively here as a normalisation constant). We can further approximate the reduced master equation in equation (5) by the Stratonovich form of the FPE [35]:

$$\frac{\partial P(n,t)}{\partial t} = -\frac{\partial}{\partial n} [(a_1(n) + a_2'(n)a_2(n))P(n,t)] + \frac{1}{2} \frac{\partial^2}{\partial n^2} [a_2(n)P(n,t)], \quad (9)$$

where $a_1(n)$ and $a_2(n)$ are the first two jump moments, given by $a_1(n) = T^+(n) - T^-(n)$ and $a_2(n) = T^+(n) + T^-(n)$ respectively, and $P(n)$ denotes the FPE solution (a notation used throughout the paper). The purpose of this further approximation by means of an FPE will be made clear in section 3.1. Equation (9) has a steady state solution of the form [6]:

$$P(n) = \frac{N}{\sqrt{T^+(n) + T^-(n)}} \exp\left(2 \int^n \frac{T^+(z) - T^-(z)}{T^+(z) + T^-(z)} dz\right), \quad (10)$$

where N is a normalisation constant. Although the integral in the exponent of equation (10) can be solved exactly with propensities of the form of equations (6) and (7) since it is the integral of the ratio of two cubic polynomials, the solution is too complicated to be detailed here. The approximations made by the FPE approximation are that (i) fluctuations in the protein number are small and (ii) we are in the fast switching regime between the gene states. Figure 1 compares the FPE solution equation (10) with the solution of the heuristic CME in equation (8) and the solution of the full CME of the reaction scheme in equation (1) using the finite space projection method (FSP) [36]. Note that provided the state space is truncated large enough, the FPE solution matches the solution of the heuristic CME almost exactly. Clearly, when gene switching is fast (bottom plot of figure 1) all three solutions agree with each other. However, when gene switching is not fast (top and middle plots on figure 1) both the reduced CME and FPE solutions are a poor approximation of the true distribution from FSP.

3. Accounting for fluctuating rates using the UCNA

Fluctuating rate parameters can be used to include a description of processes not explicitly taken into account in the formulation of a model. In figure 2 we illustrate this idea. In this section, we add fluctuations to the rate parameters of the FPE description derived earlier and use the UCNA to obtain a new effective FPE that is valid when the timescale of the noise on the rates is either very small or very large. We remind the reader that in this section we consider gene switching to be fast, and we consider the case of slow switching in section 4.

3.1. Fluctuating degradation rate

We begin by considering the case of a fluctuating degradation rate. These fluctuations could for example stem from details of the degradation machinery that are not explicitly described in the model, e.g. multi-step degradation mediated by enzymatic reactions.

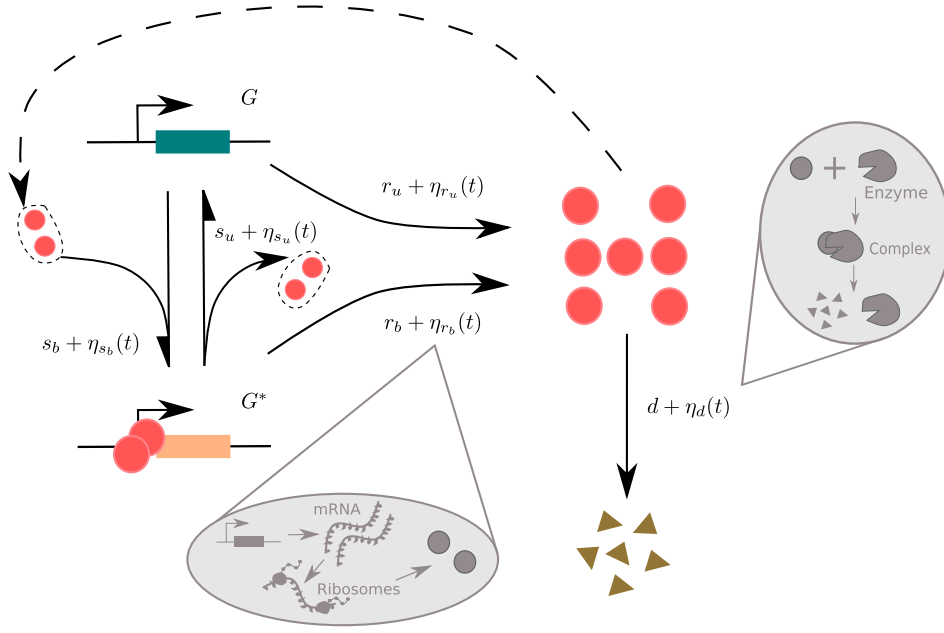


Figure 2. Illustration of the cooperative auto-regulatory reaction scheme, with colored noise included on each individual reaction. For the case of non-fluctuating rates explored in section 2 the noise terms, η_i , on the rate parameter can be set to zero. Where colored noise is included in section 3 these noise terms are not set to zero. The addition of noise onto rate parameters can be thought of as accounting for processes that are not explicitly included in the gene expression model. Here we show two examples, where colored noise on the rate parameters of the reduced model can be used to account for mRNA number fluctuations during protein translation, or the degradation of proteins via an enzyme catalytic mechanism.

The equivalent Langevin equation to the chemical FPE from equation (9) using the propensities from equations (6) and (7) is given by [6, 37]:

$$\frac{dn}{dt} = \frac{r_u L + r_b (n/\Omega)^2}{L + (n/\Omega)^2} - dn + \sqrt{\frac{r_u L + r_b (n/\Omega)^2}{L + (n/\Omega)^2} + dn} \cdot \Gamma(t), \quad (11)$$

where $\Gamma(t)$ is Gaussian white noise with zero mean and correlator $\langle \Gamma(t)\Gamma(t') \rangle = \delta(t - t')$. Now we introduce a fluctuating degradation rate by setting $d = d_0(1 + \eta(t))$, where $\eta(t)$ is Gaussian colored noise with a mean of zero and correlator $\langle \eta(t)\eta(t') \rangle = (D/\tau)\exp(-|t - t'|/\tau)$ [27, 38]. Here, τ is the *correlation time* of the colored noise, D/τ is the *noise strength* (the variance of fluctuations) and d_0 is the mean degradation rate. Since D/τ is the noise strength, i.e., D scales the noise strength at constant τ , we occasionally refer to D itself as the *noise strength* (where τ is a fixed parameter). In the limit of $\tau \rightarrow 0$ colored noise becomes white noise since $\lim_{\tau \rightarrow 0} \langle \eta(t)\eta(t') \rangle = D\delta(t - t')$. Note that $\eta(t)$ must satisfy $|\eta(t)| \ll 1$ such that d is a positive quantity (and hence admits physical interpretation as a rate parameter). The inclusion of colored noise can be approximated by the following two component system [27]:

$$\frac{dn}{dt} = \frac{r_u L + r_b (n/\Omega)^2}{L + (n/\Omega)^2} - d_0(1 + \eta)n + \sqrt{\frac{r_u L + r_b (n/\Omega)^2}{L + (n/\Omega)^2} + d_0 n} \cdot \Gamma(t), \quad (12)$$

$$\frac{d\eta}{dt} = -\frac{1}{\tau}\eta + \frac{1}{\tau}\theta(t), \quad (13)$$

where $\theta(t)$ is Gaussian white noise with zero mean and correlator $\langle\theta(t)\theta(t')\rangle = 2D\delta(t-t')$, and the time dependence on the protein number $n(t)$ and noise $\eta(t)$ is suppressed for notational convenience. Note that in the argument of the square root above we have replaced $\eta(t)$ by its mean of zero; this constitutes a mean-field type of approximation, and is useful such that one can solve equations (12) and (13) analytically—however, where the noise is small this is generally a very good approximation. Note that we also use this mean-field assumption in sections 3.2 and 3.3. For transparency, we rewrite equations (12) and (13) as:

$$\frac{dn}{dt} = h(n) + g_1(n)\eta + g_2(n)\Gamma(t), \quad (14)$$

$$\frac{d\eta}{dt} = -\frac{1}{\tau}\eta + \frac{1}{\tau}\theta(t), \quad (15)$$

with

$$h(n) = \frac{r_u L + r_b (n/\Omega)^2}{L + (n/\Omega)^2} - d_0 n, \quad (16)$$

$$g_1(n) = -d_0 n, \quad (17)$$

$$g_2(n) = \sqrt{\frac{r_u L + r_b (n/\Omega)^2}{L + (n/\Omega)^2} + d_0} n. \quad (18)$$

In order to approximately solve equations (14) and (15) we next utilize the UCNA to obtain reduced Langevin equations when the noise η is either very fast or very slow. For completeness, we present a non-rigorous but intuitive proof of the UCNA along the lines found in [27] which essentially consists of a direct adiabatic elimination on the stochastic differential equations (SDEs) in equations (14) and (15). For a more rigorous derivation of a UCNA-like FPE we advise reader to read the seminal work of Fox, who introduced a functional calculus approach to the study of colored noise SDEs [39–42]. A review of the differing UCNA-like derivations can be found in [43].

It has been discussed in [27, 43] that the adiabatic elimination we employ below is exact for $\tau \rightarrow 0$ (white noise) or $\tau \rightarrow \infty$ (highly correlated noise) but that it should give a useful approximation for intermediate values of τ . We note that the theory provided by Roberts *et al* [26] does not provide such a result as they consider separately the cases of $\tau \rightarrow 0$ and $\tau \rightarrow \infty$. For the biological applications we consider in section 5 the limit of $\tau \rightarrow \infty$ is not of interest, and we will later focus on the limit of τ small, although the derivation shown here holds for large τ too. First, where we use overdots to represent derivatives with respect to time t , one should proceed in rearranging equation (14) for η :

$$\eta(n, \dot{n}) = \frac{1}{g_1(n)}(\dot{n} - h(n) - g_2(n)\Gamma(t)). \quad (19)$$

In what follows we will utilise a mean-field approximation (denoted by the subscript *mf*) to approximately calculate the time derivative of $\eta(n, \dot{n})$. We start by defining the mean-field approximation of $\eta(n, \dot{n})$ as:

$$\eta_{mf}(n_{mf}, \dot{n}_{mf}) = \frac{1}{g_1(n_{mf})}(\dot{n}_{mf} - h(n_{mf})). \quad (20)$$

Taking the time derivative with respect to non-dimensional time $\hat{t} = t/\tau$ (denoted by the overdot) we obtain:

$$\dot{\eta}_{mf} = \frac{1}{g_1(n_{mf})} \left(\frac{h(n_{mf})g'_1(n_{mf})}{g_1(n_{mf})} - h'(n_{mf}) \right) \dot{n}_{mf} + \frac{\tau^{-1}}{g_1(n_{mf})} \left(\ddot{n}_{mf} - \frac{g'_1(n_{mf})}{g_1(n_{mf})} \dot{n}_{mf}^2 \right), \quad (21)$$

where the prime on each function of n_{mf} denotes the derivative with respect to n_{mf} . In the limit of $\tau \rightarrow 0$, the second term on the right-hand side of equation (21) goes to infinity and hence the only way to keep the time derivative finite is to impose the condition:

$$\ddot{n}_{mf} - \frac{g'_1(n_{mf})}{g_1(n_{mf})} \dot{n}_{mf}^2 = 0. \quad (22)$$

This then implies that in this limit we have:

$$\dot{\eta}_{mf} \approx \frac{1}{g_1(n_{mf})} \left(\frac{h(n_{mf})g'_1(n_{mf})}{g_1(n_{mf})} - h'(n_{mf}) \right) \dot{n}_{mf}. \quad (23)$$

Note that taking the limit of $\tau \rightarrow \infty$ gives the same result and hence the approximation equation (23) is valid in both the limit of small and large τ . This can be shown to be self-consistently true; taking the time-derivative of equation (14) alongside a mean-field approximation we get,

$$\ddot{n}_{mf} = (h'(n_{mf}) + g'_1(n_{mf})\eta_{mf}) \dot{n}_{mf} + g_1(n_{mf})\dot{\eta}_{mf}. \quad (24)$$

Assuming equations (20) and (23) to be true one then recovers

$$\ddot{n}_{mf} - \frac{g'_1(n_{mf})}{g_1(n_{mf})} \dot{n}_{mf}^2 = 0, \quad (25)$$

which means that if equation (23) holds true then so does equation (22) (and vice versa).

In equation (15) we can now substitute η from (19) and $\dot{\eta}_{mf}$ for $\dot{\eta}$ from equation (23) giving us the UCNA for the system with colored noise on the degradation rate, which is *exact* in the limits $\tau \rightarrow 0$ or $\tau \rightarrow \infty$:

$$\dot{n} \approx \frac{h(n)}{C(n, \tau)} + \frac{1}{C(n, \tau)} (g_1(n)\theta(t) + g_2(n)\Gamma(t)), \quad (26)$$

where

$$C(n, \tau) = 1 + \tau \left(\frac{g'_1(n)h(n)}{g_1(n)} - h'(n) \right). \quad (27)$$

Note that we have dropped off the *mf* subscript for clarity. Finally, in order to get a simplified Langevin equation, we modify equation (26) such that we only have one effective Gaussian white noise term. We begin by proposing:

$$g(n)\tilde{\Gamma}(t) = g_1(n)\theta(t) + g_2(n)\Gamma(t), \quad (28)$$

where $\tilde{\Gamma}(t)$ is Gaussian white noise with mean zero and correlator $\langle \tilde{\Gamma}(t)\tilde{\Gamma}(t') \rangle = 2\delta(t-t')$, and then use relations between the correlators to find our unknown $g(n)$. Note that we assume zero correlation between $\Gamma(t)$ and $\theta(t)$, i.e. $\langle \Gamma(t)\theta(t') \rangle = \langle \Gamma(t')\theta(t) \rangle = 0$. Explicitly, utilising the correlators, we find:

$$g(n)^2 \langle \tilde{\Gamma}(t)\tilde{\Gamma}(t') \rangle = g_1(n)^2 \langle \theta(t)\theta(t') \rangle + g_2(n)^2 \langle \Gamma(t)\Gamma(t') \rangle, \quad (29)$$

which gives us

$$g(n) = \sqrt{Dg_1(n)^2 + \frac{1}{2}g_2(n)^2}. \quad (30)$$

Hence, our final reduced Langevin equation is given by:

$$\dot{n} = \frac{h(n)}{C(n, \tau)} + \frac{g(n)}{C(n, \tau)} \tilde{\Gamma}(t), \quad (31)$$

which corresponds to the result in [38]. Note that equations (26) and (31) are *identical*. Here we pause to make a couple of comments on $C(n, \tau)$, which can be interpreted as a renormalisation of the Langevin equation in equation (14) to account for the addition of colored noise to the rate parameters. In fact, when $\tau = 0$, equation (31) recovers the correct Langevin equation for a process with white noise on the rate parameters. One should also note the independence of $C(n, \tau)$ from the strength of the noise D ; the renormalisation with respect to the addition of colored noise on the degradation rate is not specific to the size of the noise, it simply accounts for the finite correlation time.

The FPE corresponding to this SDE should be chosen in the Stratonovich form, following from [41, 44, 45], as this is the physical implementation of an SDE with colored noise having a non-zero correlation time τ . This FPE is:

$$\frac{\partial P(n, t)}{\partial t} = -\frac{\partial}{\partial n} \left[\left(\tilde{h}(n) + \tilde{g}(n)\tilde{g}'(n) \right) P(n, t) \right] + \frac{\partial^2}{\partial n^2} \left[\tilde{g}(n)^2 P(n, t) \right], \quad (32)$$

where $\tilde{h}(n) = h(n)/C(n, \tau)$ and $\tilde{g}(n) = g(n)/C(n, \tau)$. Following equations (9) and (10) in section 2 and [6], the steady state solution to this equation is then given by:

$$P(n) = \frac{N}{\tilde{g}(n)^2} \exp \left(\int^n \frac{\tilde{h}(z) + \tilde{g}(z)\tilde{g}'(z)}{\tilde{g}(z)^2} dz \right) = \frac{N}{\tilde{g}(n)} \exp \left(\int^n \frac{\tilde{h}(z)}{\tilde{g}(z)^2} dz \right), \quad (33)$$

where N is the normalisation constant, chosen over the domain $n \in [0, \infty)$.

Having made various approximations to arrive at equation (33) we now pause and summarise the approximations made thus far, clarifying the conditions under which we expect equation (33) to produce meaningful distributions. We started by considering equation (11) which is the chemical Langevin equation describing protein dynamics and which was derived from the CME describing the reaction scheme in equation (1) under the approximations of large protein numbers and fast promoter switching. Subsequently we added colored noise to the degradation rate in equation (11) and made a mean-field approximation (valid for small fluctuations about the mean degradation rate) to obtain the coupled Langevin equations equation (12) and (13). These equations were then reduced to a single effective Langevin equation equation (31) by the UCNA under the assumption that the correlation time of colored noise is very small or very large. Finally the solution of this Langevin equation is given by equation (33). Hence summarizing, we expect the latter solution to be an accurate stochastic description of the protein fluctuations in reaction scheme (1) with a fluctuating degradation rate provided the protein numbers are large, promoter switching is fast and the correlation time of fluctuations in the degradation parameter is either very small or very large.

To test the accuracy of the distributions for colored noise provided by the UCNA in equation (33), we compare the UCNA solution to a distribution produced from a modified SSA that explicitly accounts for the colored noise on the degradation rate. This modification

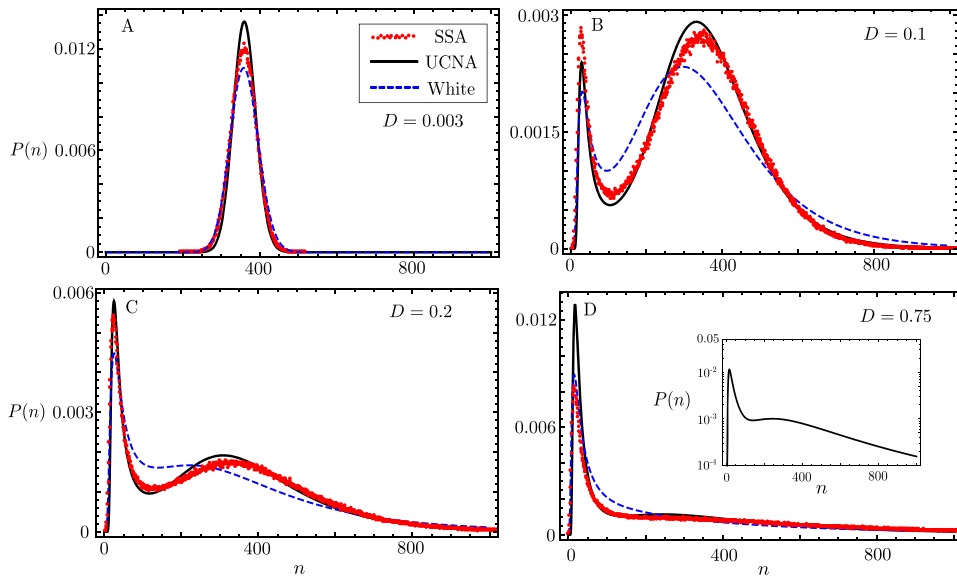


Figure 3. Comparison of the UCNA (black line) from equation (33) and white extrinsic noise (UCNA with $\tau = 0$, dashed blue line) with stochastic simulations using the modified SSA (red points) of the cooperative reaction scheme in equation (1), where the colored noise is added to the degradation rate. Aside from variation in the strength of noise D (shown on each plot), the shared parameters are $r_u = 24$, $r_b = 464$, $s_b = s_u = 1000$, $d_0 = 1$, $\Omega = 200$ and $\tau = 1$. Parameters s_b and s_u are chosen to be large compared to other system parameters such that the frequency of gene activation and inactivation events is much larger than the frequency of other reaction events, i.e. the fast gene switching assumption. Note that for this choice of rate parameters, the rate equations are bistable with equilibrium points at $n = 47.4, 360.4$. The criterion $\sqrt{D/\tau} < 1$ is required to ensure positivity of the degradation rate. As the extrinsic noise is increased, the mass of the distribution shifts from the mode at 360.4 to the mode at 47.4. The inset of (D) shows the same distribution but with the y-axis on a log scale, emphasising the exponential tail of the distribution for large n . SSA data in each case comes from a single steady state trajectory of 9×10^6 s.

is given in full detail in appendix A. Essentially, the dilution/degradation reaction $P \rightarrow \emptyset$ is replaced by three new reactions alongside the introduction of a ghost species Y , these being (i) $\emptyset \rightleftharpoons Y$ and (ii) $P + Y \rightarrow Y$. The rates of these new reactions are then chosen to ensure the magnitude of *effective* external noise on the degradation reaction, due to fluctuations in molecule numbers of the ghost species, match the colored noise SDE given in equation (13).

Figure 3 shows steady state probability distributions produced by the UCNA for various values of D for a deterministically bistable set of parameters. The UCNA correctly captures the shift of the probability mass from the equilibrium point of higher molecule number (referred to as the *upper mode*) to the lower equilibrium point (referred to as the *lower mode*) as D is increased. Importantly, this shows that when gene switching is assumed to be fast, colored noise can induce bimodality—one should keep this in mind for when we look at slow gene switching in section 4. Readers should also note that the parameter choices have been selected such that the Fokker–Planck approximation is good, notably that the system size is large, i.e., $\Omega \gg 1$, and the mean number of proteins in the system is also large. In all cases $\sqrt{D/\tau} < 1$

so that the degradation rate remains positive. The behavior seen as D increases in figure 3 can be explained as follows. When D is small (figure 3(A)) the colored noise η in equation (15) is also small compared to the mean number of molecules in the system, and the noise cannot force the system out of the upper mode. As D gets larger (figures 3(B) and (C)) the fluctuations η at the upper mode also become larger, allowing the system to explore the lower mode. When the system is found in the lower mode the pre-factor of the colored noise in equation (14), $g_1(n) \propto n$, is lesser in magnitude, and the fluctuations in η are much smaller than when the system inhabits the upper mode hence the increased probability mass at the lower mode. That the system is less noisy at the lower mode means that the system struggles more to get a fluctuation large enough to propel it into the upper mode. These properties of the system as D increases can be further seen through (i) the increase in probability mass found at the lower mode as D increases throughout all of figures 3(A)–(D), and (ii) the increased probability mass found in the tail of the distribution for large n (figure 3(D)); while the tail is very slowly decaying it is still exponential and hence the distribution is not heavy-tailed (see the inset of figure 3(D)). This ability to induce bimodality through a more detailed description of the details of the degradation process is important in the context of *cellular decision-making*. It is hence possible for regions of the reaction rate parameter space previously thought unable to induce multiple phenotypic states to do so with an increasing influence of more complex degradation mechanisms. Note that for the majority of cases in figure 3, the UCNA provides a much better approximation than the white noise approximation, hence one cannot simply assume that since the correlation time τ is relatively small that it can be approximated as zero.

Figure 4 shows how the UCNA responds to increasing correlation time τ while the noise strength, D/τ , remains fixed. For all cases where τ is small, the UCNA performs very well. As τ increases however the UCNA starts to predict ever increasing negative probabilities for some values of n . Notably though, figure 4(B) shows that even where significant negative probability is predicted at large τ , the UCNA still manages to capture the rest of the distribution. This negativity of $C(n, \tau)$ is commented on in both [27, 41]. The former deals with this negativity by taking the absolute magnitude of the pre-factor of the exponential in equation (33), while the latter comments that the proof of their UCNA-like FPE is only formally valid where $C(n, \tau) > 0, \forall n$. Here we choose not to take the magnitude of the pre-factor in equation (33), since although this leads to a positive probability for all n it is nonetheless a poor approximation; but we take careful note of the comment made by Fox in [41], as this indicates where the UCNA will perform well. The intuition behind the argument of Fox can be stated as: if for some n , $C(n, \tau) < 0$ there must be a transitory value of n for which $C(n, \tau) = 0$, at this point the equation (31) becomes physically ill-defined and our solution is invalid.

Finally, we observe that the parameter values chosen for both plots in figure 4 correspond to deterministically monostable systems. The bimodality that is observed in figure 4 is hence *noise induced bimodality*. The mode that appears for small τ corresponds to the deterministic equilibrium point, whereas the noise induced mode does not correspond to an equilibrium point of the deterministic system. We notice that the ability to exhibit a noise induced mode as τ becomes large is especially true for monostable parameter sets which are in close proximity to bistable parameter sets in the parameter space. This can be explained by occasional jumps between the monostable and bistable regimes due to sufficiently large fluctuations in the degradation rate. Hence a measure of the distance here is the difference in the magnitude of d_0 needed such that the system is deterministically bistable divided by the noise strength, defined as $\Delta d_0 = |d_0 - d_c|/(D/\tau)$, where d_c is the closest value of the mean degradation rate to d_0 expressing bistability. For example, the parameter set chosen in figure 4(A), although

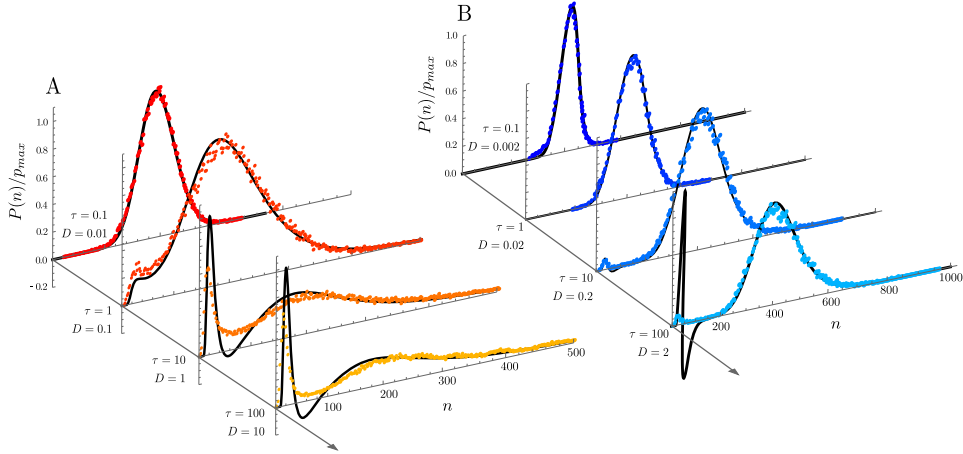


Figure 4. Comparison of the UCNA (black line) against the modified SSA (colored dots) as the correlation time τ is increased at constant noise size D/τ . Note that the y-axis shows $P(n)/p_{\max}$, where $P(n)$ is defined in equation (33) and p_{\max} is equal to the maximum value of $P(n)$. (A) Shows the performance of increasing τ for a system with parameters $r_u = 20$, $r_b = 250$, $d_0 = 1$, $s_u = 3 \times 10^2$, $s_b = 10^3$ and $\Omega = 100$. Deterministically this system is monostable with an equilibrium point at $n = 194.7$, however as τ is increased a shift toward a lower mode is observed. When τ is sufficiently large, the UCNA predicts a negative probability. (B) Shows similar to (A) but with parameters $r_u = 25$, $r_b = 480$, $d_0 = 1$, $s_u = 8 \times 10^2$, $s_b = 10^3$ and $\Omega = 200$. This too is a deterministically monostable system with equilibrium point $n = 406.0$. As τ increases, the breakdown of the UCNA is more apparent than for (A) with the prediction of negative probability for small n more drastic. Both (A) and (B) show that unless τ is large, while D/τ is small, the UCNA provides a very good approximation, even where the colored noise induces bimodality in deterministically monomodal systems. SSA data in each case comes from a single steady state trajectory of 9×10^5 s.

monostable, is very close to a parameter set that exhibits deterministic bistability ($\Delta d_0 = 2.12$). On the other hand, the parameter set of figure 4(B) is far from the bistable parameter regime ($\Delta d_0 = 57.5$)—and hence the bimodality shown is very limited as τ becomes large. The reason for this noise induced bimodality then can be seen by the ability of a system, through fluctuations in the rate parameters, to access parameter regimes which in fact do exhibit deterministic bistability. Importantly, even when it seems bimodality is not induced (e.g., figures 3(A) or 4(B)), using the extremal equation of $P(n)$ from [38], i.e., $\tilde{h}(n) = \tilde{g}(n)\tilde{g}'(n)$, one can show that the UCNA still predicts the presence of two modes. This explanation of the induced bimodality in cooperative autoregulation is further supported by the lack of noise induced bimodality when colored noise is included on the degradation rate of the FPE describing non-cooperative autoregulation; here the UCNA's extremal equation only ever predicts the existence of one mode for the probability distribution.

3.2. Fluctuating effective protein production rates

We now extend the analysis from section 3.1 to the effective protein production rates. Colored noise on the effective production rates can be used to implicitly model multi-step protein production, including multiple stages of mRNA processing before translation (see figure 2). We add colored noise onto the effective protein production rates via, $r_u = r_u^{(0)}(1 + \eta_1(t))$ and $r_b = r_b^{(0)}(1 + \eta_2(t))$, which upon substituting in the Langevin equation describing the feedback

loop equation (11) we obtain the following set of SDEs:

$$\frac{dn}{dt} = \frac{r_u^{(0)}L + r_b^{(0)}(n/\Omega)^2}{L + (n/\Omega)^2} - dn + \frac{r_u^{(0)}L\eta_1 + r_b^{(0)}(n/\Omega)^2\eta_2}{L + (n/\Omega)^2} + \sqrt{\frac{r_u^{(0)}L + r_b^{(0)}(n/\Omega)^2}{L + (n/\Omega)^2}} + dn \cdot \Gamma(t), \tag{34}$$

$$\frac{d\eta_1}{dt} = -\frac{1}{\tau}\eta_1 + \frac{1}{\tau}\theta_1(t), \tag{35}$$

$$\frac{d\eta_2}{dt} = -\frac{1}{\tau}\eta_2 + \frac{1}{\tau}\theta_2(t), \tag{36}$$

where $\theta_1(t)$ and $\theta_2(t)$ are Gaussian white noise terms with zero mean and correlators $\langle \theta_1(t)\theta_1(t') \rangle = 2D_1\delta(t - t')$ and $\langle \theta_2(t)\theta_2(t') \rangle = 2D_2\delta(t - t')$ respectively. Note that here we have used a mean-field approximation for the terms under the square root, as was done in section 3.1. In a similar style to equation (28) we now propose a new noise term $\tilde{\eta}(t)$, which couples $\eta_1(t)$ and $\eta_2(t)$, satisfying:

$$F(n)\tilde{\eta}(t) = f_1(n)\eta_1(t) + f_2(n)\eta_2(t), \tag{37}$$

where $f_1(n) = r_u^{(0)}L/(L + (n/\Omega)^2)$, $f_2(n) = r_b^{(0)}(n/\Omega)^2/(L + (n/\Omega)^2)$ and $\tilde{\eta}(t)$ is colored noise with zero mean and correlator $\langle \tilde{\eta}(t)\tilde{\eta}(t') \rangle = e^{-|t-t'|/\tau}/\tau$, satisfying the following equation:

$$\frac{d\tilde{\eta}}{dt} = -\frac{1}{\tau}\tilde{\eta} + \frac{1}{\tau}\theta(t), \tag{38}$$

where $\theta(t)$ is Gaussian white noise with correlator $\langle \theta(t)\theta(t') \rangle = 2\delta(t - t')$. The correlators for $\eta_1(t)$ and $\eta_2(t)$ are $\langle \eta_1(t)\eta_1(t') \rangle = D_1 e^{-|t-t'|/\tau}/\tau$ and $\langle \eta_2(t)\eta_2(t') \rangle = D_2 e^{-|t-t'|/\tau}/\tau$, where we have assumed that the colored noise on both production rates has the same correlation time but a differing magnitude of noise strength. Using the properties of the correlators of η_1 , η_2 and $\tilde{\eta}$ we then find:

$$F(n) = \sqrt{f_1(n)^2D_1 + f_2(n)^2D_2}. \tag{39}$$

Sharing the notation adopted in section 3.1, we define the following:

$$h(n) = \frac{r_u^{(0)}L + r_b^{(0)}(n/\Omega)^2}{L + (n/\Omega)^2} - dn, \tag{40}$$

$$g_2(n) = \sqrt{\frac{r_u^{(0)}L + r_b^{(0)}(n/\Omega)^2}{L + (n/\Omega)^2}} + dn. \tag{41}$$

This gives us the following SDE which is coupled to equation (38):

$$\frac{dn}{dt} = h(n) + F(n)\tilde{\eta} + g_2(n)\Gamma(t). \tag{42}$$

Then, following the same UCNA procedure as in equations (19)–(26), we obtain the following approximate Langevin equation:

$$\dot{n} \approx \frac{h(n)}{C(n, \tau)} + \frac{1}{C(n, \tau)}(F(n)\theta(t) + g_2(n)\Gamma(t)), \quad (43)$$

where

$$C(n, \tau) = 1 + \tau \left(\frac{F'(n)h(n)}{F(n)} - h'(n) \right). \quad (44)$$

In this case it is interesting to note that unlike the case of a fluctuating degradation rate, here $C(n, \tau)$ does depend on both the correlation time τ and the strength of the colored noise D_1, D_2 (unless $D_1 = D_2$ in which case there is only dependence on τ). To simplify equation (43) further, we again propose:

$$g(n)\tilde{\Gamma}(t) = F(n)\theta(t) + g_2(n)\Gamma(t), \quad (45)$$

where $\tilde{\Gamma}(t)\tilde{\Gamma}(t') = 2\delta(t - t')$, and find using the correlators that $g(n) = \sqrt{F(n)^2 + g_2(n)^2}/2$. This leads to the final approximate SDE:

$$\dot{n} = \frac{h(n)}{C(n, \tau)} + \frac{g(n)}{C(n, \tau)}\tilde{\Gamma}(t), \quad (46)$$

which is identical in notation to equation (31) but where $h(n), C(n, \tau)$ and $g(n)$ are all defined in this section. The equivalent FPE for this SDE is then:

$$\frac{\partial P(n, t)}{\partial t} = -\frac{\partial}{\partial n} \left[\left(\tilde{h}(n) + \tilde{g}(n)\tilde{g}'(n) \right) P(n, t) \right] + \frac{\partial^2}{\partial n^2} \left[\tilde{g}(n)^2 P(n, t) \right]. \quad (47)$$

Again our solution for the probability distribution will then be:

$$P(n) = \frac{N}{\tilde{g}(n)} \exp \left(\int^n \frac{\tilde{h}(z)}{\tilde{g}(z)^2} dz \right), \quad (48)$$

with $\tilde{h}(n) = h(n)/C(n, \tau)$ and $\tilde{g}(n) = g(n)/C(n, \tau)$.

We now describe the modified SSA that takes into account extrinsic noise on the effective protein production rates. This modification replaces the protein production reaction in each gene state, i.e., $G_k \rightarrow G_k + P$ where G_k represents either G or G^* , by three new reactions alongside the introduction of a ghost species Y_k for each gene state. These new reactions are $\emptyset \xrightleftharpoons[r_2]{r_1} Y_k$ and $G_k + Y_k \xrightarrow{r_3} G_k + Y_k + P$. Utilising the LNA (assuming Y_k to be abundant), as was done for colored noise on the degradation rate in appendix A, one finds these rates to be $r_1 = 1/(D_k\Omega), r_2 = 1/\tau$ and $r_3 = r_k^{(0)}D_k\Omega/\tau$, which ensure matching to the colored noise SDE given in equation (34), where $r_k^{(0)}$ represents $r_u^{(0)}$ or $r_b^{(0)}$ in G and G^* respectively.

Figure 5 shows a good performance of the UCNA when compared to the modified SSA described above. This performance is shown for each differing qualitative behavior expressed by cooperative bimodality, i.e., (i) monostable positive feedback, (ii) bistable positive feedback, and (iii) monostable negative feedback. In all three plots shown the UCNA matches the modified SSA well, and clearly performs better than if one were to approximate the colored noise with white noise (i.e., $\tau = 0$).

We find the same qualitative behavior of the creation and eventual destruction of bimodality (see figures 6(A)(i)–(iii)) as the noise strengths, D_1 and D_2 , become large for the colored noise on the protein production rates as was found in figure 3 for colored noise on the degradation rate. Note that for the chosen parameter set in figure 6(A) that the white noise approximation

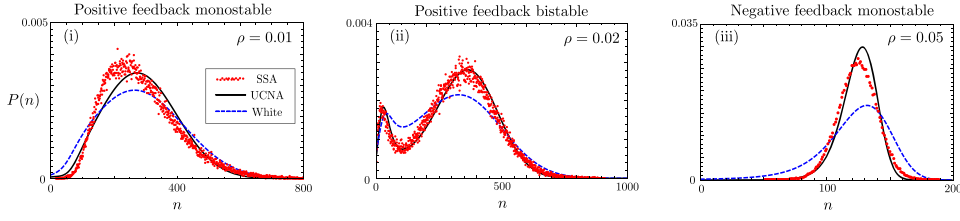


Figure 5. This figure shows the agreement of the UCNA on the protein production rates to the modified SSA (detailed in section 3.2), also compared to the case of white extrinsic noise ($\tau = 0$). The plots show agreement of the UCNA over the three main qualitative regimes of cooperative autoregulation at large molecule number, showing respectively: (i) monostable positive feedback with parameters $r_u^{(0)} = 150$, $r_b^{(0)} = 300$, $s_u = s_b = 10^3$, $d = 1$, $D_1 = 0.25$, $D_2 = 0.25$, $\tau = 0.5$, $\Omega = 100$; (ii) bistable positive feedback with parameters $r_u^{(0)} = 24$, $r_b^{(0)} = 468$, $s_u = s_b = 10^3$, $d = 1$, $D_1 = 0.75$, $D_2 = 0.1$, $\tau = 1$, $\Omega = 200$; (iii) monostable negative feedback with parameters $r_u^{(0)} = 470$, $r_b^{(0)} = 20$, $s_u = s_b = 10^3$, $d = 1$, $D_1 = 0.1$, $D_2 = 0.1$, $\tau = 1$, $\Omega = 70$. In the top right hand corner of each plot is the value of ρ for the distribution, defined and discussed in section 3.4.3, here showing that for good UCNA performance ρ should be small to satisfy condition 3. SSA data in each case comes from a single steady state trajectory of 9×10^5 s.

performs generally very well compared to the UCNA. For $\tau \leq 1$, the white extrinsic noise approximation can typically perform quite well compared to the modified SSA, but note that this is not always the case (e.g., see figure 3).

3.3. Fluctuating binding/unbinding rates

Finally, we apply the UCNA to the case of colored noise added to the binding and unbinding rates of the protein to the gene. This could be utilised to implicitly model the effect of multiple gene states in the transition of G to G^* , as has been experimentally and theoretically investigated [46–48], accounting for DNA looping via distal enhancers or chromatin conformational states. For convenience we define $s_b = s_b^{(0)}(1 + \eta_1(t))$, $s_u = s_u^{(0)}(1 + \eta_2(t))$ and

$$L_\eta = L_0 \left(\frac{1 + \eta_1(t)}{1 + \eta_2(t)} \right), \quad \text{with } L_0 = \frac{s_u^{(0)}}{s_b^{(0)}}. \quad (49)$$

Substituting equation (49) in the Langevin equation describing the feedback loop equation (11) (and making a mean-field approximation for the terms under the square root) we obtain the following set of SDEs:

$$\frac{dn}{dt} = \frac{L_\eta r_u + r_b(n/\Omega)^2}{L_\eta + (n/\Omega)^2} + \sqrt{\frac{r_u L_0 + r_b(n/\Omega)^2}{L_0 + (n/\Omega)^2}} \, dn \cdot \Gamma(t), \quad (50)$$

$$\frac{d\eta_1}{dt} = -\frac{1}{\tau} \eta_1 + \frac{1}{\tau} \theta_1(t), \quad (51)$$

$$\frac{d\eta_2}{dt} = -\frac{1}{\tau} \eta_2 + \frac{1}{\tau} \theta_2(t), \quad (52)$$

where $\theta_1(t)$ and $\theta_2(t)$ are Gaussian white noise terms with zero mean and correlators $\langle \theta_1(t)\theta_1(t') \rangle = 2D_1\delta(t-t')$ and $\langle \theta_2(t)\theta_2(t') \rangle = 2D_2\delta(t-t')$ respectively. In order to proceed using the UCNA we must linearise the drift term in equation (50) with respect to η_1 and η_2

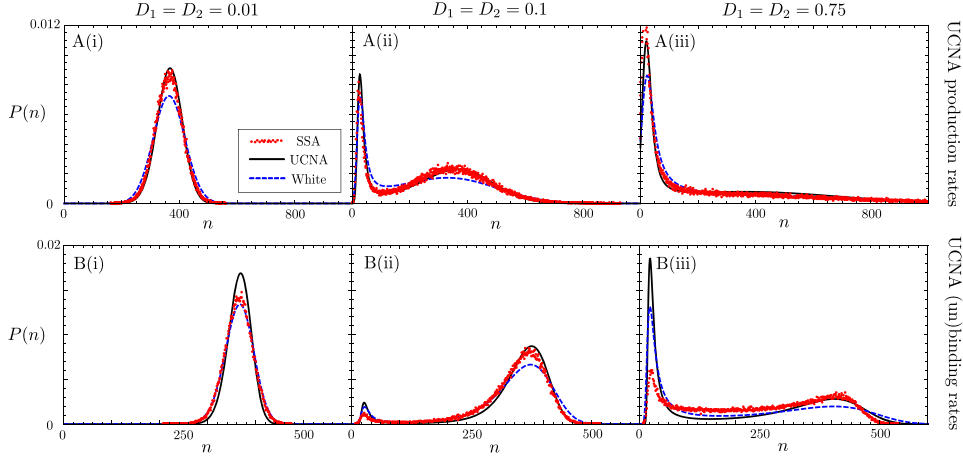


Figure 6. Plots showing the creation and eventual destruction of bimodality in the probability distributions for colored noise on the (A) protein production rates, (B) binding/unbinding rates (denoted on the figure by (un)binding rates), analogously to what was observed in figure 3 for colored noise on the degradation rate. For (A) it is clear that the UCNA performs well where the noise strength is both small in (A)(i) and large in (A)(iii). For (B) we see that the low (B)(i) and intermediate (B)(ii) noise cases are well predicted by the UCNA and white noise approximation, however where the noise becomes large (B)(iii) the UCNA breaks down, whereas the white noise approximation still performs well compared to the modified SSA prediction. Other than the noise strengths given on the figure, the parameters for both (A) and (B) are $r_u^{(0)} = 24$, $r_b^{(0)} = 468$, $s_u = s_b = 10^3$, $d = 1$, $\tau = 1$, $\Omega = 200$ (i.e., the same parameters used in figure 3 which express deterministic bistability). SSA data in each case comes from a single steady state trajectory of 9×10^5 s.

through the small noise approximation $\eta_1, \eta_2 \ll 1$:

$$\frac{dn}{dt} \approx \frac{r_u L_0 + r_b (n/\Omega)^2}{L_0 + (n/\Omega)^2} - dn + \left(\frac{L_0 n^2 \Omega^2 (r_u - r_b)}{(L_0 \Omega^2 + n^2)^2} \right) (\eta_1 - \eta_2) + \sqrt{\frac{r_u L_0 + r_b (n/\Omega)^2}{L_0 + (n/\Omega)^2} + dn} \cdot \Gamma(t). \quad (53)$$

For convenience we now define:

$$h(n) = \frac{r_u L_0 + r_b (n/\Omega)^2}{L_0 + (n/\Omega)^2} - dn, \quad (54)$$

$$g_1(n) = \frac{L_0 n^2 \Omega^2 (r_u - r_b)}{(L_0 \Omega^2 + n^2)^2}, \quad (55)$$

$$g_2(n) = \sqrt{\frac{r_u L_0 + r_b (n/\Omega)^2}{L_0 + (n/\Omega)^2} + dn}, \quad (56)$$

$$F(n) = g_1(n) \sqrt{D_1 + D_2}, \quad (57)$$

$$g(n) = \sqrt{F(n)^2 + g_2(n)^2/2}. \quad (58)$$

In terms of these new functions equation (53) becomes,

$$\frac{dn}{dt} = h(n) + g_1(n)(\eta_1 - \eta_2) + g_2(n)\Gamma(t). \tag{59}$$

Following section 3.2 we then arrive at the UCNA for colored noise on the binding rates where $\eta_1, \eta_2 \ll 1$:

$$\frac{dn}{dt} = \frac{h(n)}{C(n, \tau)} + \frac{1}{C(n, \tau)} (F(n)\theta(t) + g_2(n)\Gamma(t)). \tag{60}$$

Then, using the properties of the correlators of $\theta(t)$ and $\Gamma(t)$ we arrive at:

$$\frac{dn}{dt} = \frac{h(n)}{C(n, \tau)} + \frac{g(n)}{C(n, \tau)} \tilde{\Gamma}(t), \tag{61}$$

where,

$$C(n, \tau) = 1 + \tau \left(\frac{g'_1(n)h(n)}{g_1(n)} - h'(n) \right), \tag{62}$$

and $\tilde{\Gamma}(t)$ is Gaussian white noise with mean zero and correlator $\langle \tilde{\Gamma}(t)\tilde{\Gamma}(t') \rangle = 2\delta(t - t')$. Here, as for the UCNA applied to the degradation rate, $C(n, \tau)$ is again independent of the strengths of the colored noise terms. This UCNA, as we shall see, should be a good approximation where both D_1 and D_2 are small—by ‘small’ we explicitly mean that D_1 and D_2 should be smaller than noise strengths used on the UCNA for protein production rates or the degradation rate. The solution to equation (61) is given by:

$$P(n) = \frac{N}{\tilde{g}(n)} \exp \left(\int^n \frac{\tilde{h}(z)}{\tilde{g}(z)^2} dz \right), \tag{63}$$

with $\tilde{h}(n) = h(n)/C(n, \tau)$ and $\tilde{g}(n) = g(n)/C(n, \tau)$.

Now we evaluate the performance of the UCNA on the binding and unbinding rates, and compare it with the modified SSA. In this case the modified SSA replaces the binding and unbinding reactions, $G + 2P \xrightleftharpoons[s_u]{s_b} G^*$, by the following: $\emptyset \xrightleftharpoons[r_2]{r_1} Y_1$, $G + Y_1 + 2P \xrightarrow{r_3} Y_1 + G^*$, $\emptyset \xrightleftharpoons[r_5]{r_4} Y_2$, and $G^* + Y_2 \xrightarrow{r_6} G + Y_2 + 2P$, where Y_1 and Y_2 are ghost species. The rates of these reactions are determined via the LNA (assuming the ghost species to be numerous) and are $r_1 = 1/(D_1\Omega)$, $r_2 = 1/\tau$, $r_3 = s_b^{(0)}D_1\Omega/\tau$, $r_4 = 1/(D_2\Omega)$, $r_5 = 1/\tau$ and $r_6 = s_u^{(0)}D_2\Omega/\tau$.

In figure 6(B) we test the UCNA on the binding and unbinding rates compared to the modified SSA described above. Clearly, the same qualitative behavior of the creation and destruction of bimodality, as noise strength is increased, is observed, as was also observed for colored noise on the degradation rate (figure 3) and protein production rates (figure 6(A)). The resultant expression of bimodality however, is clearly different than for these cases. Notably, this UCNA does ascribe to an additional limitation compared to the UCNA of degradation or production rates; a limitation due to the further small noise approximation made in equation (53). This limitation is seen in figure 6(B)(iii), showing that the UCNA applied to the binding and unbinding rates is much more sensitive to increased noise strength than the other UCNA applications. One also observes that the white noise approximation in figure 6(B) performs almost as well as the UCNA (figures 6(B)(i)–(ii)) or in some cases

better than the UCNA (figure 6(B)(iii)); hence, the white noise approximation may be a safer approximation than the UCNA for colored noise applied to the binding and unbinding rates.

3.4. Breakdown conditions of the UCNA

Having now applied the UCNA to approximate distributions for colored noise on the (i) degradation rate, (ii) protein production rates and (iii) the binding/unbinding rates, we now assess the conditions which cause the UCNA to breakdown. The application of the UCNA to colored noise on the protein production rates presents a somewhat more complex problem than the application of the UCNA to colored noise on the degradation rate or the binding/unbinding rates; hence, we more easily see that there are *three main conditions for the breakdown of the UCNA*—conditions beside the large system size or large molecule number requirement needed to approximate the discrete master equation by a one variable FPE, or even the need for τ to be chosen small or large enough such that equations (22) and (23) are approximately satisfied. Below we detail these three conditions, in each case explaining why the disagreement occurs. Note that although the analysis of breakdown conditions below is done for the UCNA on the protein production rates, the same arguments hold for the other applications of the UCNA previously presented.

3.4.1. Condition 1. The first of these conditions concerns the positivity condition required on $C(n, \tau)$, that is $C(n, \tau) > 0 \forall n$. We refer to this as *condition 1*. Since we have already discussed this condition in a previous section we will not repeat the discussion here, and refer the reader to section 3.1. In figure 7(A)(i) we see a disagreement between the UCNA and the modified SSA for a parameter set that exhibits bimodality, and in figure 7(A)(ii) it is verified that this since $C(n, \tau) < 0$ where $n \approx 100$. Note however, that if $C(n, \tau)$ becomes negative outside of the region containing most of the probability mass that the UCNA can still provide a good approximation to the true modified SSA solution.

3.4.2. Condition 2. The second condition observed for the breakdown of the UCNA concerns the violation of the *characteristic ‘length’ scale* (the length here being a distance measure in the n space), which we now discuss. In appendix B we show in more detail why the arguments we present below hold. Based on the noise intensity of the noise term arising from the colored noise in equation (43), we can introduce the characteristic length scale L over which fluctuations in the colored noise term are damped:

$$L(n, \tau) = \frac{F(n)}{C(n, \tau)}, \tag{64}$$

noting that the requirement of condition 1 means that this length is always positive. Our approximate one variable FPE in equation (47) will then be valid under the condition that the drift term varies slowly with respect to L (following appendix B), meaning that one needs to satisfy

$$L \left| \partial_n \left(\tilde{h}(n) + \tilde{g}(n)\tilde{g}'(n) \right) \right| \ll \left| \tilde{h}(n) + \tilde{g}(n)\tilde{g}'(n) \right| \tag{65}$$

in order for the UCNA to hold. More succinctly, this condition is:

$$C(n, \tau) \gg \kappa(n, \tau), \tag{66}$$

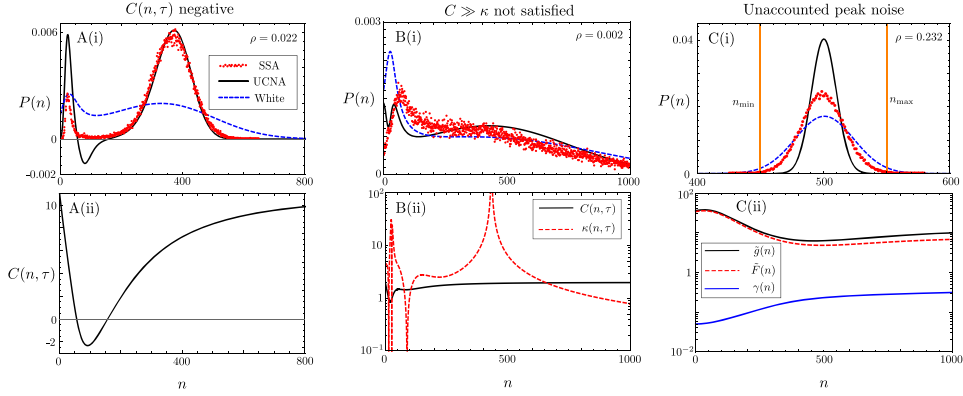


Figure 7. This figure shows the disagreement of the UCNA on the protein production rates to the ground truth modified SSA predictions (detailed in section 3.2), also compared to the case of white extrinsic noise ($\tau = 0$). Each disagreement corresponds to a single breakdown condition of the UCNA being violated. The legend in (A)(i) applies to (A)(i), (B)(i) and (C)(i). Plots in (A) show the breakdown of the UCNA due to condition 1. (A)(i) shows the prediction of negative probability due to the negativity of $C(n, \tau)$ in (A)(ii) around the same value of n . Parameters for (A) are $r_u^{(0)} = 20$, $r_b^{(0)} = 470$, $s_u = s_b = 10^3$, $d = 1$, $D_1 = 1$, $D_2 = 0.1$, $\tau = 10$, $\Omega = 200$. Plots in (B) show the breakdown of the UCNA due to condition 2. (B)(ii) shows that $\kappa(n, \tau) > C(n, \tau)$ over a large range of n , corresponding to the poor UCNA prediction seen in (B)(i) over this entire region. Parameters for (B) are $r_u^{(0)} = 50$, $r_b^{(0)} = 450$, $s_u = s_b = 10^3$, $d = 1$, $D_1 = 1$, $D_2 = 1$, $\tau = 1$, $\Omega = 100$. Plots in (C) show the breakdown of the UCNA due to condition 3. (C)(ii) shows a relatively large value of $\gamma(n)$ over most of the defined region \mathcal{D} , and also shows the pre-factors of the total UCNA noise $\tilde{g}(n)$ and that arising only from the colored noise $\tilde{F}(n) = F(n)/C(n, \tau)$. Vertical orange lines in (C)(i) indicate the limits of the region \mathcal{D} . The value ρ in the top right-hand corner of (C)(i) can be compared to the smaller values of ρ seen for other parameter sets in (A)(i) and (B)(i), indicating that the breakdown observed is truly associated to condition 3. The parameters for (C) are $r_u^{(0)} = 2300$, $r_b^{(0)} = 120$, $s_u = s_b = 10^4$, $d = 1$, $D_1 = 0.002$, $D_2 = 0.04$, $\tau = 2$, $\Omega = 230$. SSA data in each case comes from a single steady state trajectory of 9×10^5 s.

where we henceforth define the function

$$\kappa(n, \tau) = F(n) \left| \frac{\partial_n \left(\tilde{h}(n) + \tilde{g}(n)\tilde{g}'(n) \right)}{\tilde{h}(n) + \tilde{g}(n)\tilde{g}'(n)} \right|. \quad (67)$$

We refer to equation (66) as *condition 2*. In figure 7(B) we explore this breakdown for a parameter set that breaks condition 2 over a large region of the parameter space, between $0 < n < 650$. Clearly the UCNA provides a poor approximation in this regime; note however that, similar to condition 1, if condition 2 is violated (i) outside of the domain where most of the probability mass is contained, or (ii) over a small region of the domain containing most of the probability mass, then the UCNA can still provide a good approximation.

3.4.3. Condition 3. The final condition resulting in the breakdown of the UCNA concerns the underestimation of noise. We refer to this as *unaccounted peak noise*, and this forms our final breakdown condition, *condition 3*. The explanation behind condition 3 is that the UCNA in general will always underestimate the Poisson noise for a particular value of n , arising from

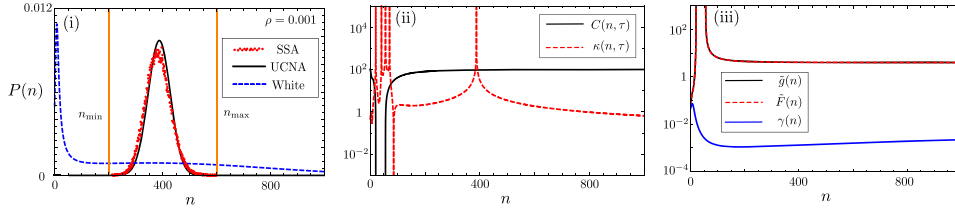


Figure 8. Plots showing a good performance of the UCNA for $\tau = 100$. (i) Shows the probability distributions from the modified SSA, UCNA and white noise approximation. The vertical orange lines show the limits of the region \mathcal{D} in this case. (ii) Shows that in the region \mathcal{D} that condition 1 is satisfied since $C(n \in \mathcal{D}) > 1$, and condition 2 is satisfied since $C(n \in \mathcal{D}, \tau) \gg \kappa(n \in \mathcal{D}, \tau)$. (iii) Shows that condition 3 is satisfied in \mathcal{D} , i.e., $\gamma(n) \approx 0$, since $\tilde{g}(n) \approx \tilde{F}(n)$, which is corroborated by the small value of ρ shown in (i). Other parameters here are $r_u^{(0)} = 10$, $r_b^{(0)} = 400$, $s_u = s_b = 10^3$, $d = 1$, $D_1 = 0.5$, $D_2 = 1$, $\Omega = 70$. SSA data in each case comes from a single steady state trajectory of 9×10^5 s.

the necessary neglect of Poisson noise terms in the derivation of the UCNA: (i) neglect of the noise terms under the square root of the Poisson noise pre-factor in equation (34) (a form of mean-field approximation), and (ii) neglect of Poisson noise term $g_2(n)\Gamma(t)$ and its time derivative from the $\dot{\eta}$ term in equations (19)–(21) via the use of another mean-field approximation. However, the error on the UCNA caused by condition 3 will be small when colored noise dominates the Poisson noise. To investigate the degree to which colored noise is dominant, identifying $F(n)/C(n, \tau)$ from equation (42) and $g(n)/C(n, \tau)$ from equation (46), we define

$$\gamma(n) = \left| \frac{g(n)/C(n, \tau) - F(n)/C(n, \tau)}{g(n)/C(n, \tau)} \right| = \left| \frac{g(n) - F(n)}{g(n)} \right| \quad (68)$$

where, for some n , if $\gamma(n) \approx 1$ then Poisson noise dominates, else if $\gamma(n) \approx 0$ then colored noise dominates. Intermediate values of $\gamma(n)$ mean that both Poisson and colored noise is apparent in the system. To investigate whether noise is underestimated generally over the region containing most of the probability, defined as $\mathcal{D} = [n_{\min}, n_{\max}]$, we further define

$$\rho = \frac{1}{|\mathcal{D}|} \int_{n_{\min}}^{n_{\max}} \gamma(n) dn. \quad (69)$$

Here, if $\rho \approx 1$ then Poisson noise dominates over the entire region \mathcal{D} , else if $\rho \approx 0$ then colored noise dominates over the entire region \mathcal{D} . Figure 7(C) explores this disagreement, where figure 7(C)(i) shows the clear underestimation of noise in the UCNA distribution when compared to the modified SSA distribution. Sample values of n_{\min} and n_{\max} are also shown on figure 7(C)(i). Figure 7(C)(ii) shows how the total UCNA noise $\tilde{g}(n)$ varies with respect to the contribution of colored noise $\tilde{F}(n) = F(n)/C(n, \tau)$. Also shown on figure 7(C)(ii) is the variation of $\gamma(n)$. Values of ρ are shown in the top right-hand corner for all probability distributions in figure 5; unlike the other distributions shown in figures 5 and 7, in figure 7(C)(i) $\rho \neq 0$ does not hold, clarifying that the reason for the UCNA’s disagreement for this parameter set is due to condition 3.

3.4.4. Large τ UCNA distributions. Having successfully identified the three main conditions causing the breakdown of the UCNA, we are now able to determine where the UCNA will perform well, *even in the large τ limit*. In figure 8 we explore an example of the UCNA performing

exceptionally well for $\tau = 10^2$ (see figure 8(i)). Clearly the UCNA does not violate any of the three conditions here: (1) $C(n, \tau)$ is not negative in \mathcal{D} (see figure 8(ii)); (2) $C(n, \tau) \gg \kappa(n, \tau)$ in \mathcal{D} (again, see figure 8(ii)); (3) $\gamma(n)$ is small for all n in \mathcal{D} (see figure 8(iii)) as evidenced by the small value of $\rho = 0.001$. As expected, the prediction of white noise on the protein production rates is very poor in the regime where $\tau \gg 1$.

4. Slow gene switching: the conditional UCNA

In the previous sections we have focused on fast gene switching, whereby a Hill function can then be used to approximate the production of proteins from two different gene states, shown in the reaction scheme of equation (1). We now consider the case where the switching rates s_u and s_b are very small; small enough that the system has two dominant modes of behavior, one pertaining to each gene state. The approach followed here is very similar to the *conditional LNA* studied in [49], but instead of approximating the distribution conditional on each gene state as a Gaussian we instead utilise the UCNA in each gene state. We shall refer to this method as the *conditional UCNA* (cUCNA). We begin by stating the law of total probability for the marginal distribution of proteins that we are interested in approximating:

$$P(n, t) = \sum_{\underline{G}} P(G_i, t) P(n|G_i, t). \tag{70}$$

Here \underline{G} is the set of possible gene state (in our case $\underline{G} = \{G, G^*\}$), $P(G_i, t)$ is the marginal distribution of being in gene state G_i at a time t and $P(n|G_i, t)$ is the conditional probability of having n proteins at a time t given that the system is in state G_i . Our task now is to find suitable approximations for $P(G_i, t)$ and $P(n|G_i, t)$ that allow us then to construct an approximation of the full steady state distribution in equation (70). In our case we have two different gene states, G and G^* , and hence we can construct the reaction schemes conditional on each gene state. The reaction scheme conditional on gene state G is (i) $G \xrightarrow{r_u} G + P, P \xrightarrow{d} \emptyset$, and the reaction scheme conditional on gene state G^* is (ii) $G^* \xrightarrow{r_b} G^* + P, P \xrightarrow{d} \emptyset$. This then allows us to approximately find the steady state mean number of proteins conditional on each gene state when s_u and s_b are very small (where the subscript a denotes approximate): $\langle n|G \rangle_a = r_u/d$ and $\langle n|G^* \rangle_a = r_b/d$. We can use these conditional means to find the marginal probabilities of being in a specific gene state at steady state. Note that in this calculation we will ignore the influence of noise on the rate parameters; the inherent assumption is that extrinsic noise does not much influence the probability of being in each gene state. First we write an approximate master equation for the transitions between differing gene states:

$$\frac{d}{dt} P(G, t) \approx s_u P(G^*, t) - \frac{s_b \langle n|G \rangle_a^2}{\Omega^2} P(G, t). \tag{71}$$

We can then solve the above equation at steady state (denoted by the subscript s) by utilising conservation of probability, $P_s(G) = 1 - P_s(G^*)$, giving:

$$P_s(G^*) = \left(1 + \frac{s_u}{s_b} \left(\frac{d \cdot \Omega}{r_u} \right)^2 \right)^{-1}, \tag{72}$$

$$P_s(G) = \left(1 + \frac{s_b}{s_u} \left(\frac{r_u}{d \cdot \Omega} \right)^2 \right)^{-1}. \tag{73}$$

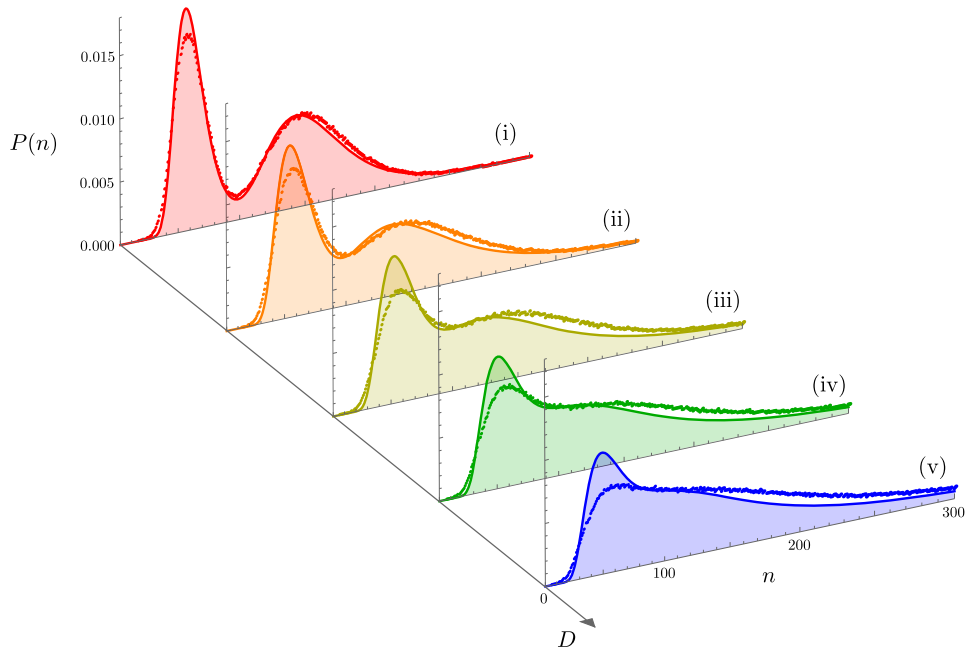


Figure 9. Comparison of the cUCNA with the modified SSA for increasing values of the colored noise strength for both gene states, D . It is seen that the cUCNA (solid lines) is a good approximation to the true distribution (dots, simulated using the modified SSA described in appendix A), especially for small values of D . The noise strengths for each plot are (i) $D = 0.1$, (ii) $D = 0.2$, (iii) $D = 0.3$, (iv) $D = 0.4$, (v) $D = 0.5$, and the shared parameters are $r_u = 30$, $r_b = 75$, $s_b = 0.01$, $s_u = 0.001$, $d_0 = 0.5$ and $\tau = 1$. Clearly as D gets larger the bimodality exhibited by the slow switching between the gene states is destroyed by the extrinsic noise added to the degradation rate.

Since now we have the $P_s(G_i)$ needed for equation (70) we need to find the $P_s(n|G_i)$ terms. Here we show how to calculate these terms for noise on the degradation rate, although this can be easily extended to the case where we have noise on the protein production rates. In each gene state, the system we are concerned to study is $G_i \xrightarrow{r_i} G_i + P$, $P \xrightarrow{d_i} \emptyset$, where G_i , d_i and r_i represent either gene state G or G^* , the corresponding gene state dependent decay rate, and production rate r_u or r_b respectively. Adding colored noise to the degradation rate $d_i = d_0(1 + \eta_i)$, where d_i is the degradation rate in gene state G_i given colored noise η_i , we then have the following set of SDEs in each gene state (here we have applied the mean-field approximation to the terms in the square root):

$$\frac{dn}{dt} = r_i - d_0 n - (d_0 n)\eta_i + \sqrt{r_i + d_0 n} \cdot \Gamma(t), \tag{74}$$

$$\frac{d\eta_i}{dt} = -\frac{1}{\tau_i}\eta_i + \frac{1}{\tau_i}\theta_i(t), \tag{75}$$

where $\Gamma(t)$ and $\theta_i(t)$ are Gaussian white noise terms, each with zero mean and correlators $\langle \Gamma(t)\Gamma(t') \rangle = \delta(t - t')$ and $\langle \theta_i(t)\theta_i(t') \rangle = 2D_i\delta(t - t')$ respectively. Processing the usual steps of the UCNA method, detailed explicitly in section 3, we find the approximate steady state

probability for each gene state:

$$P_s(n|G_i) \approx N \exp(u(n, r_i)) n^{2r_i\tau_i-1} (r_i + d_0 n(2d_0 D_i n + 1))^{-\frac{1}{2} - \frac{1}{2d_0 D_i} - r_i\tau_i} (n + r_i\tau_i), \quad (76)$$

where N is a normalisation constant and we have defined,

$$u(n, r_i) = \frac{(D_i(4 - 6d_0\tau_i)r_i + 1) \tan^{-1}\left(\frac{4d_0 D_i n + 1}{\sqrt{8D_i r_i - 1}}\right)}{d_0 D_i \sqrt{8D_i r_i - 1}}. \quad (77)$$

Hence, using equations (72), (73) and (76) we can now approximate equation (70) as:

$$P(n) \approx P_s(G)P_s(n|G) + P_s(G^*)P_s(n|G^*). \quad (78)$$

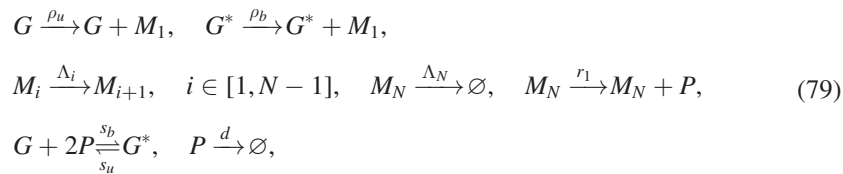
Figure 9 compares the cUCNA with the modified SSA—which is the same as the modified SSA found in section 3.1. Figure 9(i) shows that for small switching rates, the cUCNA can correctly capture the bimodality exhibited where the colored noise on the degradation rate is small. As the noise on the degradation rate gets larger the cUCNA still provides a decent approximation to the true distributions; it is also clear that the bimodality of the protein distribution is destroyed as the size of this noise increases. One can contrast this to the cases observed in figures 3 and 6 which showed that where the gene switching rates are fast, increased colored noise strength can in fact induce bimodality. *In summary, we find that extrinsic noise on the degradation rate of a slow switching auto-regulatory system generally destroys bimodality, but for fast switching it is common to observe the opposite phenomenon.*

5. Applications

In this section we explicitly show, by means of two examples, how one can use the colored noise formulation that was introduced earlier to describe intricate molecular details of cooperative autoregulation. We first show this for multi-stage protein production with fast gene switching, and then for multi-stage protein degradation with slow gene switching.

5.1. Multi-stage protein production

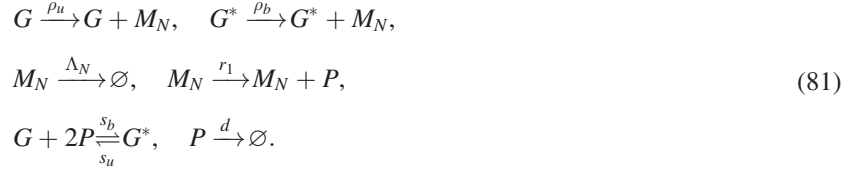
The first example of using colored noise as a form of model reduction is that of mapping multi-stage protein production onto a simpler system, where colored noise accounts for processes not explicitly considered in the simpler model. Consider multi-stage protein production on the cooperative auto-regulatory feedback loop:



where it is assumed the system contains only one gene copy, either in state G or in state G^* . The simpler model that we will then map this system onto the cooperative auto-regulatory feedback loop:



where $r_u = r_u^{(0)}(1 + \eta_1(t))$ and $r_b = r_b^{(0)}(1 + \eta_2(t))$, and assigning the properties of colored noises $\eta_1(t)$ and $\eta_2(t)$ such that equation (79) can be mapped onto equation (80) is the task we have assigned ourselves. One can think of the different M_i for $i < N$ as the various stages of nascent mRNA, before it is eventually fully transcribed in stage M_N (mature mRNA) where it can then begin translation [50, 51]. Utilising the slow scale LNA [19] one can show that if $\Lambda_i \gg \max\{\Lambda_N, \rho_u, \rho_b\}$ for $i \in [1, N-1]$ then the nascent mRNA M_1, \dots, M_{N-1} are fast species, and the reaction system in equation (79) is consistent with the following reaction scheme describing fluctuations in the slow species G, G^*, M_N and P :



We now apply the van Kampen ansatz to the number of mature mRNA, M_N . In gene state G this gives us $n_1(t) = \Omega\phi_1 + \Omega^{1/2}\epsilon_1(t)$, and in gene state G^* this gives us $n_2(t) = \Omega\phi_2 + \Omega^{1/2}\epsilon_2(t)$, where $\phi_1 = \rho_u/(\Lambda_N\Omega)$ and $\phi_2 = \rho_b/(\Lambda_N\Omega)$ are the steady state solutions to the rate equation describing the mature mRNA in the gene states G and G^* respectively, and $\epsilon_1(t)$ and $\epsilon_2(t)$ describe small fluctuations about these means. Note the occurrence of $1/\Omega$ in ϕ_1 and ϕ_2 follows since the concentration of a single gene in a volume Ω is $1/\Omega$. Using these ansatzes allows us to construct the effective protein production rates in gene states G and G^* respectively:

$$r_u = r_1 n_1(t) = \frac{r_1 \rho_u}{\Lambda_N} \left(1 + \Omega^{1/2} \frac{\Lambda_N}{\rho_u} \epsilon_1(t) \right), \quad (82)$$

$$r_b = r_1 n_2(t) = \frac{r_1 \rho_b}{\Lambda_N} \left(1 + \Omega^{1/2} \frac{\Lambda_N}{\rho_b} \epsilon_2(t) \right). \quad (83)$$

One can then see that $r_u^{(0)} = r_1 \rho_u / \Lambda_N$, $r_b^{(0)} = r_1 \rho_b / \Lambda_N$ and that the noise terms have the form:

$$\eta_1(t) = \Omega^{1/2} \frac{\Lambda_N}{\rho_u} \epsilon_1(t), \quad (84)$$

$$\eta_2(t) = \Omega^{1/2} \frac{\Lambda_N}{\rho_b} \epsilon_2(t). \quad (85)$$

In order to fully specify $\eta_1(t)$ and $\eta_2(t)$ we need to find the correlators $\langle \eta_1(t) \eta_1(t') \rangle$ and $\langle \eta_2(t) \eta_2(t') \rangle$, which can be done by application of the LNA [6]. Note that since we are already restricted to the large system size, large molecule number regime following the FPE approximation to the CME (discussed in section 3.1), we can apply the LNA without further restricting the validity of the final solution. The same can also be said for the use of the LNA in section 5.2. Applying the LNA to $n_1(t)$ and $n_2(t)$, whose fluctuations are fully specified by the reactions $G \xrightarrow{\rho_u} G + M_N$, $G^* \xrightarrow{\rho_b} G^* + M_N$ and $M_N \xrightarrow{\Lambda_N} \emptyset$, gives us the two following one variable FPEs:

$$\frac{\partial \Pi(\epsilon_1, t)}{\partial t} = \Lambda_N \frac{\partial}{\partial \epsilon_1} (\epsilon_1 \Pi(\epsilon_1, t)) + \frac{1}{2} \left(\frac{2\rho_u}{\Omega} \right) \frac{\partial^2 \Pi(\epsilon_1, t)}{\partial \epsilon_1^2}, \quad (86)$$

$$\frac{\partial \Pi(\epsilon_2, t)}{\partial t} = \Lambda_N \frac{\partial}{\partial \epsilon_2} (\epsilon_2 \Pi(\epsilon_2, t)) + \frac{1}{2} \left(\frac{2\rho_b}{\Omega} \right) \frac{\partial^2 \Pi(\epsilon_2, t)}{\partial \epsilon_2^2}, \quad (87)$$

where $\Pi(\epsilon_i, t)$ is the probability of having a fluctuation of size ϵ_i at a time t . These FPEs, combined with equation (84) and (85), admit equivalent Langevin equations for $\eta_1(t)$ and $\eta_2(t)$, given by:

$$\frac{d\eta_1(t)}{dt} = \Lambda_N \left(-\eta_1(t) + \sqrt{\frac{2}{\rho_u}} \beta_1(t) \right), \quad (88)$$

$$\frac{d\eta_2(t)}{dt} = \Lambda_N \left(-\eta_2(t) + \sqrt{\frac{2}{\rho_b}} \beta_2(t) \right), \quad (89)$$

where $\beta_1(t)$ and $\beta_2(t)$ are independent Gaussian white noises with zero mean and correlator $\langle \beta_1(t) \beta_1(t') \rangle = \langle \beta_2(t) \beta_2(t') \rangle = \delta(t - t')$. From here one can find the correlators of $\eta_1(t)$ and $\eta_2(t)$:

$$\langle \eta_1(t) \eta_1(t') \rangle = \frac{\Lambda_N}{\rho_u} \exp(-\Lambda_N |t - t'|), \quad (90)$$

$$\langle \eta_2(t) \eta_2(t') \rangle = \frac{\Lambda_N}{\rho_b} \exp(-\Lambda_N |t - t'|). \quad (91)$$

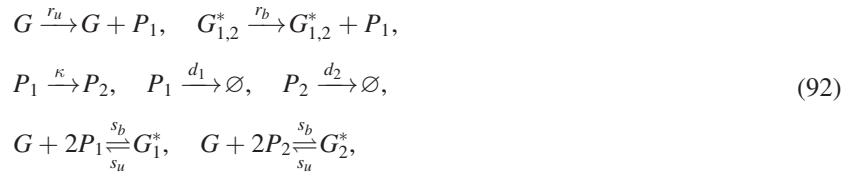
Comparing to the results of section 3.2 it is clear that $\eta_1(t)$ and $\eta_2(t)$ satisfy the definition of colored noise, with noise strengths $D_1 = 1/\rho_u$, $D_2 = 1/\rho_b$ and shared correlation time $\tau = 1/\Lambda_N$. This completes the mapping between the full complex system in equation (79) and our reduced process in equation (80). We can hence utilise our solution for the probability distribution with colored noise on the effective protein production rates in equation (48). Note that the colored noise in this case can model transcriptional bursting, namely the production of proteins in bursts due to rapid translation from short lived mRNA [52, 53]; bursty expression has been previously modeled in the literature by an effective first-order reaction with constant rate parameter but with the special property that when the reaction fires, the number of proteins produced is sampled from a geometric distribution [33, 54].

In figure 10(A) we show how effective the UCNA can be in approximating the protein distribution from the full system described in equation (79), where we have for simplicity assumed that there are three mRNA states: M_1 , M_2 and M_3 (i.e., $N = 3$). Figure 10(A)(i) shows the approximation for a parameter set exhibiting bimodality: the red points represent the *standard* SSA of the full system in equation (79); the black line represents the distribution predicted from the UCNA (i.e., using equation (48) with $D_1 = 1/\rho_u$, $D_2 = 1/\rho_b$ and $\tau = 1/\Lambda_N$); the blue dotted line represents the distribution if one put white noise of the same magnitude on the protein production rates (i.e., the UCNA at $\tau = 0$); and the orange line with circles shows the distribution if one was to neglect noise on the reaction rates entirely (i.e., $r_u = r_u^{(0)}$ and $r_b = r_b^{(0)}$). Clearly, in figure 10(A)(i) the UCNA is the only distribution that fits the SSA prediction, showing both the effectiveness of our model reduction as well as the need to properly account for the correlation time of colored noise in model reduction. This makes sense, since one would expect processes occurring in the full system to be correlated over short times, i.e., *that noise events in close temporal proximity are not independent*, and one cannot simply neglect these effects. Figure 10(A)(ii) instead shows the various approximations for a monomodal parameter set. In this case, white noise is a poor approximation, and it is clear that one cannot neglect the finite correlation time. However, it is interesting to note that properly accounting for the correlation time using the UCNA returns the same distribution as if one had not added noise to the production rates at all—this is due to the small magnitudes of D_1/τ and D_2/τ respectively.

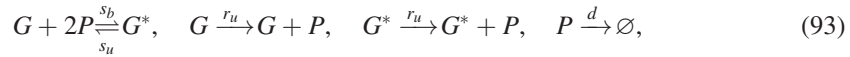
These two examples shows that where correlation time is finite, it is imperative that one models it correctly.

5.2. Multi-stage protein degradation

Proteins in cells are often degraded via multi-step processes. For example, a major degradation pathway in eukaryotic cells is the ubiquitin-proteasome degradation pathway [55], and more recent experiments have shown that a subset of proteins in the mammalian proteome have age-dependent degradation rates [56, 57]. From figure 2 in [56] we consider a system with two different stages of protein with differing degradation rates combined with the cooperative auto-regulatory feedback loop:



where $G_{1,2}^*$ indicates *either* the state G_1^* or G_2^* . This reaction system models age dependent protein states, since the protein P_1 is always produced from the gene, and eventually undergoes a transition to protein state P_2 , where P_1 and P_2 have differing degradation rates. We will show how to map this system to the reduced system:



where the total number of P is given as the sum of the number of P_1 and P_2 , i.e., $n = n_1 + n_2$, G^* is simply the sum of G_1^* and G_2^* , and $d = d_0(1 + \eta(t))$, where $\eta(t)$ is colored noise. Our task is to find the properties of the noise $\eta(t)$ such that one can map the full system in equation (92) onto the reduced system in equation (93). Note that although we here look at two different stages of protein, the analysis presented below can be easily extended for several different stages of protein degradation.

One finds that the effective degradation rate of the sum of protein number P_1 and P_2 is:

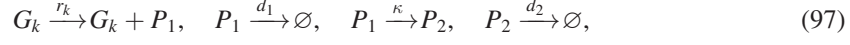
$$d = \frac{n_1 d_1 + n_2 d_2}{n_1 + n_2}. \tag{94}$$

In the following analysis we consider gene switching to be slow, which allows us to apply the cUCNA from section 4. We first consider the probability of being in each gene state at steady state $P_s(G_k)$, where G_k represents either gene state G or G^* . Note that we assume both protein stages P_1 and P_2 can bind and unbind to the gene at the same respective rates, and note that $\langle n|G \rangle = \langle n_1|G \rangle + \langle n_2|G \rangle$. Following the analysis from section 4 in equations (72) and (73) we find:

$$P_s(G^*) = \left(1 + \frac{s_u \Omega^2}{s_b} \left(\frac{d_2(\kappa + d_1)}{r_u(\kappa + d_2)} \right)^2 \right)^{-1}, \tag{95}$$

$$P_s(G) = \left(1 + \frac{s_b}{s_u \Omega^2} \left(\frac{r_u(\kappa + d_2)}{d_2(\kappa + d_1)} \right)^2 \right)^{-1}. \tag{96}$$

We can now proceed to find the probability distribution conditional on each gene state $P_s(n_1, n_2)$. In gene state G_k the conditional reaction system is:



where the protein is always produced in stage P_1 , and $r_k \equiv r_u$ in gene state G , and $r_k \equiv r_b$ in gene state G^* . Now we employ the van Kampen ansatz [6] on n_1 and n_2 in gene state G_k , i.e. $n_1^{(k)}(t) = \Omega \phi_1^{*(k)} + \Omega^{1/2} \epsilon_1^{(k)}(t)$ and $n_2^{(k)}(t) = \Omega \phi_2^{*(k)} + \Omega^{1/2} \epsilon_2^{(k)}(t)$, where $\phi_1^{*(k)}$ and $\phi_2^{*(k)}$ are the deterministic steady state mean concentrations of P_1 and P_2 in gene state G_k respectively, and $\epsilon_1^{(k)}(t)$ and $\epsilon_2^{(k)}(t)$ are fluctuations about these mean values. In the following we drop the superscript (k) notation for aesthetic reasons, although one should keep in mind that the process below must be individually conducted on each gene state. The purpose of using the van Kampen ansatz can be seen upon its substitution into equation (94) which for a large system size, Ω , gives:

$$d = \frac{d_1 \phi_1^* + d_2 \phi_2^*}{\phi_1^* + \phi_2^*} \left(1 + \Omega^{-1/2} \left(\epsilon_1(t) \left(\frac{d_1}{d_1 \phi_1^* + d_2 \phi_2^*} - \frac{1}{\phi_1^* + \phi_2^*} \right) + \epsilon_2(t) \left(\frac{d_2}{d_1 \phi_1^* + d_2 \phi_2^*} - \frac{1}{\phi_1^* + \phi_2^*} \right) \right) \right) + \mathcal{O}(\Omega^{-1}). \quad (98)$$

By comparing to the effective degradation from the reduced model in gene state G_k , $d = d_0(1 + \eta_k(t))$, one can see that to match the two models we must have

$$d_0 = (d_1 \phi_1^* + d_2 \phi_2^*) / (\phi_1^* + \phi_2^*), \quad (99)$$

and

$$\eta_k(t) = \Omega^{-1/2} (\epsilon_1(t) y_1(d_1, d_2, \phi_1^*, \phi_2^*) + \epsilon_2(t) y_2(d_1, d_2, \phi_1^*, \phi_2^*)), \quad (100)$$

where we have defined the functions,

$$y_1(d_1, d_2, \phi_1^*, \phi_2^*) = \frac{d_1}{d_1 \phi_1^* + d_2 \phi_2^*} - \frac{1}{\phi_1^* + \phi_2^*}, \quad (101)$$

$$y_2(d_1, d_2, \phi_1^*, \phi_2^*) = \frac{d_2}{d_1 \phi_1^* + d_2 \phi_2^*} - \frac{1}{\phi_1^* + \phi_2^*}. \quad (102)$$

If the correlators $\langle \epsilon_1(0) \epsilon_1(t) \rangle$, $\langle \epsilon_2(0) \epsilon_2(t) \rangle$, $\langle \epsilon_1(0) \epsilon_2(t) \rangle$ and $\langle \epsilon_2(0) \epsilon_1(t) \rangle$ are known, then one can also find the correlator of $\eta_k(t)$, i.e., $\langle \eta_k(0) \eta_k(t) \rangle$, given by:

$$\langle \eta_k(0) \eta_k(t) \rangle = \frac{1}{\Omega} (y_1^2 \langle \epsilon_1(0) \epsilon_1(t) \rangle + y_2^2 \langle \epsilon_2(0) \epsilon_2(t) \rangle + y_1 y_2 (\langle \epsilon_1(0) \epsilon_2(t) \rangle + \langle \epsilon_2(0) \epsilon_1(t) \rangle)). \quad (103)$$

Note in equation (100) that if $d_1 = d_2 = d$, then the magnitude of $\eta_k(t)$ is zero for all t since the system $P_1 \xrightarrow{\kappa} P_2, \quad P_1 \xrightarrow{d} \emptyset, \quad P_2 \xrightarrow{d} \emptyset$ is equivalent to $P \xrightarrow{d} \emptyset$ where one is only interested in the total number of proteins.

To proceed in finding $\eta_k(t)$ in equation (100), we first need to find the steady state concentrations ϕ_1^* and ϕ_2^* from the deterministic rate equations. These are,

$$\frac{d\phi_1}{dt} = \frac{r_k}{\Omega} - (\kappa + d_1) \phi_1, \quad (104)$$

$$\frac{d\phi_2}{dt} = \kappa \phi_1 - d_2 \phi_2, \quad (105)$$

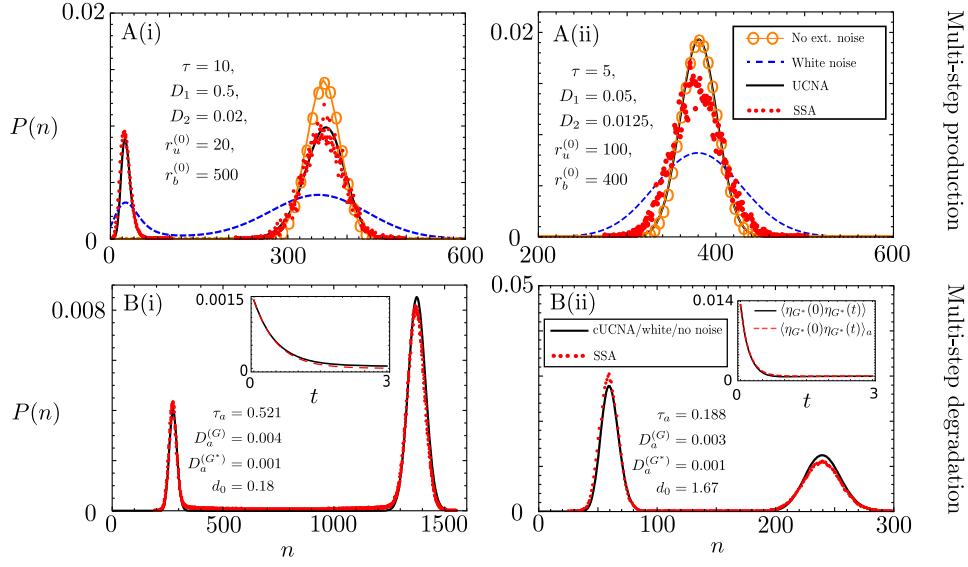


Figure 10. (A) Shows distributions of the *standard* SSA of the reaction scheme in equation (79) for multistage protein production in three intermediate species M_1 , M_2 and M_3 , against (1) the UCNA, (2) white noise (i.e., $\tau = 0$) and (3) no extrinsic colored noise (i.e., the solution given by equation (10)). Each plot shows the colored noise parameters used for the UCNA solution from equation (48), which are determined from the full multi-stage protein production process in equation (79). Note the legend in (A)(ii) applies only to distributions in (A)(i)–(ii). Parameter values for the standard SSA in (A)(i) are $\rho_u = 2$, $\rho_b = 50$, $\Lambda_1 = 1000$, $\Lambda_2 = 1000$, $\Lambda_3 = 0.1$, $r_1 = 1$, $s_u = s_b = 1000$, $d = 1$ and $\Omega = 230$. Parameter values for the SSA in (A)(ii) are $\rho_u = 20$, $\rho_b = 80$, $\Lambda_1 = 1000$, $\Lambda_2 = 1000$, $\Lambda_3 = 0.2$, $r_1 = 1$, $s_u = s_b = 1000$, $d = 1$ and $\Omega = 100$. (B) Shows distributions of the *standard* SSA, with protein decay following the multi-step process of equation (92), against the cUCNA. The colored noise parameters used for the cUCNA solution of equation (118) are shown on each plot, with these values being determined from the full model using equation (99), (116), (117). The insets show a comparison of the approximate correlator (see equation (115)) and the full double exponential correlator (see equation (114)) for gene state G^* . Note the legend in (B)(ii) applies only to distributions in (B)(i)–(ii), and the legend on the inset of (B)(ii) applies also to the inset in (B)(i). Parameter values for the SSA in (B)(i) are $r_u = 50$, $r_b = 250$, $s_b = 2.5 \times 10^{-3}$, $s_u = 10^{-3}$, $\Omega = 200$, $d_1 = 1$, $k = 1$ and $d_2 = 0.1$. Parameter values for the SSA in (B)(ii) are $r_u = 100$, $r_b = 400$, $s_b = 10^{-3}$, $s_u = 10^{-4}$, $\Omega = 200$, $d_1 = 1$, $k = 1$ and $d_2 = 5$. SSA data for (A)(i) and (A)(ii) come from a single steady state trajectory of length 10^8 s and 9×10^5 s respectively. Note that (A)(i) presents a very long relaxation to the steady state due to the systems inertia in staying in one of the two modes of the distribution. SSA data for (B)(i) and (B)(ii) come from a single steady state trajectory of 9×10^6 s.

where again the $1/\Omega$ in equation (104) follows since the concentration of a single gene in a volume Ω is $1/\Omega$. Enforcing the steady state condition allows us to find ϕ_1^* and ϕ_2^* ,

$$\phi_1^* = \frac{r_k}{\Omega(\kappa + d_1)}, \quad \phi_2^* = \frac{\kappa r_k}{d_2(\kappa + d_1)}.$$

Note that the linear dependence of ϕ_1^* and ϕ_2^* on r_k means that the effective degradation rate d_0 from equation (99) is independent of the gene state. Assuming that both P_1 and P_2 are numerous, we now proceed to the LNA [6, 58] of the system in equation (97), which will allow us to find the correlators $\langle \epsilon_1(0)\epsilon_1(t) \rangle$, $\langle \epsilon_2(0)\epsilon_2(t) \rangle$, $\langle \epsilon_1(0)\epsilon_2(t) \rangle$ and $\langle \epsilon_2(0)\epsilon_1(t) \rangle$. Where \mathbf{S} is the stoichiometric matrix, $\underline{\phi} = (\phi_1, \phi_2)$ and $\underline{f}(\underline{\phi})$ is the macroscopic rate vector one can computationally find the required matrices needed to perform the LNA: (i) the steady state Jacobian matrix $A_{ij} = d(\mathbf{S} \cdot \underline{f}(\underline{\phi}))_j / d\phi_i|_{\underline{\phi}=\underline{\phi}^*}$, and (ii) the steady state diffusion matrix $(\mathbf{B} \cdot \mathbf{B}^T)_{ij} = \mathbf{S} \cdot \text{Diag}(f(\underline{\phi})) \cdot \mathbf{S}^T|_{\underline{\phi}=\underline{\phi}^*}$. The Jacobian matrix then allows us to find the time evolution of both $\langle \epsilon_1(t) \rangle$ and $\langle \epsilon_2(t) \rangle$ since $\partial_t \langle \underline{\epsilon} \rangle = \mathbf{A} \cdot \langle \underline{\epsilon} \rangle$, where $\langle \underline{\epsilon} \rangle = (\langle \epsilon_1(t) \rangle, \langle \epsilon_2(t) \rangle)$. Solving these coupled first order ODEs gives us:

$$\langle \epsilon_1(t) \rangle = \langle \epsilon_1(0) \rangle e^{-(d_1 + \kappa)t}, \quad (106)$$

$$\langle \epsilon_2(t) \rangle = \frac{-\kappa \langle \epsilon_1(0) \rangle e^{-(d_1 + \kappa)t} + (\kappa \langle \epsilon_1(0) \rangle + (d_1 - d_2 + \kappa) \langle \epsilon_2(0) \rangle) e^{-d_2 t}}{d_1 - d_2 + \kappa}, \quad (107)$$

where $-d_2$ and $-(d_1 + \kappa)$ are eigenvalues of \mathbf{A} . Clearly, in the limit $t \rightarrow \infty$ the fluctuations about the steady state concentrations ϕ_1^* and ϕ_2^* , $\langle \epsilon_1(t) \rangle$ and $\langle \epsilon_2(t) \rangle$, tend to zero as required. The final step of the LNA then requires us to find the covariance matrix \mathbf{C} at steady state, which has the steady state variances $\langle \epsilon_1^2 \rangle$ and $\langle \epsilon_2^2 \rangle$ as diagonal components and covariance $\langle \epsilon_1 \epsilon_2 \rangle = \langle \epsilon_2 \epsilon_1 \rangle$ in the off-diagonal components. \mathbf{C} is then given by the Lyapunov equation [58]:

$$\mathbf{A} \cdot \mathbf{C} + \mathbf{C} \cdot \mathbf{A}^T + \mathbf{B} \cdot \mathbf{B}^T = 0, \quad (108)$$

whose solution is given by:

$$\mathbf{C} = \begin{pmatrix} \frac{r_k}{(d_1 + \kappa)\Omega} & 0 \\ 0 & \frac{\kappa r_k}{d_2(d_1 + \kappa)\Omega} \end{pmatrix}. \quad (109)$$

From van Kampen [6] p 259 we assert that for some fluctuation ϵ_i , $\langle \epsilon_i(0)\epsilon_j(t) \rangle = \langle \epsilon_i(0)\langle \epsilon_j(t) \rangle \rangle$, and that at $t = 0$ we have $\underline{\phi} = \underline{\phi}^*$ so that $\langle \epsilon_i(0)\epsilon_j(0) \rangle = \langle \epsilon_i \epsilon_j \rangle$. For example, for $\langle \epsilon_1(0)\epsilon_1(t) \rangle$ we have, using $\langle \epsilon_1(t) \rangle$ from equation (106) and $\langle \epsilon_1^2 \rangle$ from equation (109), $\langle \epsilon_1(0)\epsilon_1(t) \rangle = \langle \epsilon_1(0)\langle \epsilon_1(t) \rangle \rangle = \langle \epsilon_1^2 \rangle e^{-(d_1 + \kappa)t}$. Explicitly, one can then calculate all the correlators, which are given by:

$$\langle \epsilon_1(0)\epsilon_1(t) \rangle = \frac{r_k}{(d_1 + \kappa)\Omega} e^{-(d_1 + \kappa)t}, \quad (110)$$

$$\langle \epsilon_2(0)\epsilon_2(t) \rangle = \frac{\kappa r_k}{d_2(d_1 + \kappa)\Omega} e^{-d_2 t}, \quad (111)$$

$$\langle \epsilon_2(0)\epsilon_1(t) \rangle = 0, \quad (112)$$

$$\langle \epsilon_1(0)\epsilon_2(t) \rangle = \frac{\kappa r_k (e^{-d_2 t} - e^{-(d_1 + \kappa)t})}{(d_1 + \kappa)(d_1 - d_2 + \kappa)\Omega}. \quad (113)$$

Now that these correlators have been determined, we can substitute them into equation (103) giving us the following for the correlator of $\eta_k(t)$:

$$\langle \eta_k(0)\eta_k(t) \rangle = \frac{(d_1 - d_2)^2 \kappa (\kappa (d_1 + \kappa) e^{-(d_1 + \kappa)t} + (d_1 - d_2) d_2 e^{-d_2 t})}{(d_1 + \kappa)(d_1 - d_2 + \kappa)(d_2 + \kappa)^2 r_k}, \quad (114)$$

noting the only dependence on the gene state G_k comes from the pre-factor $1/r_k$. Comparing this equation to the colored noise seen in equation (12) in section 3.1 we see however that we have two exponentials in the correlator. This sum of exponentials in equation (114) can be approximated by a single exponential through a small t expansion. This gives us:

$$\langle \eta_k(0)\eta_k(t) \rangle \approx \langle \eta_k(0)\eta_k(t) \rangle_a = \frac{D_a^{(k)}}{\tau_a} e^{-t/\tau_a}, \quad (115)$$

where $D_a^{(k)}$ and τ_a are the approximate noise strength and correlation time given by:

$$D_a^{(k)} = \frac{(d_1 - d_2)^2 \kappa}{(d_1 + \kappa)(d_1 \kappa + d_2(d_2 + \kappa) + \kappa^2) r_k}, \quad (116)$$

$$\tau_a = \frac{d_2 + \kappa}{d_1 \kappa + d_2(d_2 + \kappa) + \kappa^2}. \quad (117)$$

Clearly, the small t expansion allows us to roughly interpret the noise $\eta_k(t)$, present in gene state G_k , as colored noise with strength D_a/τ_a and correlation time τ_a . Note that even when both exponentials equally contribute to the correlator in equation (114), this is generally a very good approximation for few protein stages. Knowing $D_a^{(k)}$ and τ_a for $\eta_k(t)$ we can now substitute them into equations (76) and (77) in section 4, then using equations (95) and (96) we find

$$P(n) \approx P_s(G)P_s(n|G) + P_s(G^*)P_s(n|G^*). \quad (118)$$

Figure 10(B) shows two different cases of the cUCNA predicting distributions for multi-stage degradation and slow gene switching: in (B)(i) for the case of $d_1 > d_2$ (true for around 80% of proteins in [56]); in (B)(ii) for the case of $d_2 > d_1$ (true for around 20% of proteins in [56]). On the main plots red dots show the standard SSA prediction of the full reaction scheme in equation (92), and the black lines show the cUCNA from equation (118), which in both cases is almost indistinguishable from the white noise (cUCNA with $\tau = 0$) and no external colored noise predictions (discussed further in the following paragraph). The insets show the correlators in gene state G^* , where the red dashed line represents $\langle \eta_{G^*}(0)\eta_{G^*}(t) \rangle_a$ and the black line shows $\langle \eta_{G^*}(0)\eta_{G^*}(t) \rangle$. Note that the correlators for gene state G are not shown because they show very similar to what is seen for state G^* . Even given the complex model reduction from two protein species to one effective protein species the cUCNA performs very well in predicting distributions from the standard SSA of the full system in equation (92). Note that as one considers more protein stages with differing degradation rates it becomes more different to fit the correlator to a single exponential, which presents a limitation of this method for more protein stages.

However, we find that since our analysis is restricted to the large protein number regime, and the noise strength D_a is inversely proportional to the production rate r_k which is typically large, that D_a is typically very small in both gene states. This means that the cUCNA probability distribution is almost identical to probability distributions that assume white noise (cUCNA with $\tau = 0$) or even no colored noise. However, what the analysis in this section provides is the *quantitative reasons why one could necessarily neglect the contribution of colored noise in model reduction from the full system in equation (92) to the simpler system in equation (93)*.

6. Conclusion

In this paper we have explored the addition of colored noise onto the reaction rates for a cooperative auto-regulatory circuit. Starting from a reduced chemical Fokker–Planck description, we used the UCNA to derive approximate expressions for the probability distribution of protein numbers in the limits of fast and slow promoter switching. The approximation is valid provided the colored noise on the reaction rates is small and the correlation time is short. By means of stochastic simulations, we verified the accuracy of the approximate distributions; we also verified the predictions of the UCNA, namely that under fast promoter switching conditions the addition of colored noise can induce bimodality whereas under slow promoter switching conditions, noise can destroy bimodality.

We also have shown how complex models of gene expression can be mapped onto simpler models with noisy rates. In particular we have shown that: (i) an auto-regulatory feedback loop with multi-stage protein production, including different stages of mRNA processing, can be mapped onto an auto-regulatory feedback loop with a single protein production reaction step having colored noise added to its reaction rate. (ii) A feedback loop with multi-stage protein degradation can be mapped onto a feedback loop with a single protein degradation reaction with a fluctuating rate. We have also verified that in many instances, one cannot simply approximate colored noise with white noise, or else neglect it entirely, since this does not match behavior seen from the full underlying models of multi-stage protein production or degradation.

While here we focused on a self-regulatory example, the UCNA and its conditional variant (cUCNA) provide an easily extendable analysis to model more complex gene regulatory networks with fluctuating parameters such as those with cross-regulation [59–61]. Our analysis is the first to our knowledge, to analytically find steady state probability distributions where colored noise is added to a non-linear reaction (the protein-gene binding reaction) in a gene regulatory context; a previous study applied the UCNA to study the effects of extrinsic noise in genetic circuits composed of purely linear reactions [28]. Given that our calculations show that the protein distributions for auto-regulatory circuits with extrinsic noise on reaction parameters can be dramatically different than models assuming constant reaction rates, an interesting future research direction would be to develop UCNA based methods that can directly infer the properties of colored noise on reaction rates from protein expression data.

Acknowledgments

This work was supported by a BBSRC EASTBIO PhD studentship for JH, a Darwin Trust scholarship for AG and a Leverhulme Trust Grant (RPG-2018-423) for RG.

Appendix A. Stochastic simulations of autoregulation with extrinsic noise

In this paper extrinsic noise is accounted for in the SSA through the introduction of a new ghost species Y and some new ghost reactions. For example, consider the case where we want to model a fluctuating degradation rate $d = d_0(1 + \eta(t))$, where $\langle \eta(t) \rangle = 0$ and $\langle \eta(t)\eta(t') \rangle = (D/\tau) \exp(-|t - t'|/\tau)$. We will then replace the degradation reaction, $P \xrightarrow{d} \emptyset$, by the following set of reactions:



We now show why this equivalence exists. The *propensity* for the degradation reaction in equation (119) is $(n_p n_Y d_0 D \Omega / \tau) / \Omega$, meaning that the *effective degradation rate* is $d = (n_Y d_0 D \Omega / \tau) / \Omega$. Assuming there are large numbers of Y , we apply the van Kampen ansatz that fluctuations in Y occur around its deterministic steady state mean [6]:

$$\frac{n_Y}{\Omega} = \frac{\tau}{D\Omega} + \Omega^{-1/2} \epsilon(t). \quad (120)$$

Then, employing the system size expansion, and enforcing the LNA, we obtain a linear FPE for the probability of having a fluctuation of size $\epsilon(t)$ at a time t , denoted $\Pi(\epsilon, t)$ [6, 58]:

$$\frac{\partial \Pi(\epsilon, t)}{\partial t} = \frac{1}{\tau} \frac{\partial}{\partial \epsilon} (\epsilon \Pi(\epsilon, t)) + \frac{1}{2} \frac{2}{D\Omega} \frac{\partial^2 \Pi(\epsilon, t)}{\partial \epsilon^2}. \quad (121)$$

This FPE then admits an equivalent Langevin equation given by:

$$\frac{d\epsilon(t)}{dt} = -\frac{1}{\tau} \epsilon(t) + \sqrt{\frac{2}{D\Omega}} \beta(t), \quad (122)$$

where $\beta(t)$ is Gaussian white noise with zero mean and correlator $\langle \beta(t) \beta(t') \rangle = \delta(t - t')$. Hence, from equation (120) it follows that d goes as:

$$d = d_0 D \frac{\Omega}{\tau} \frac{n_Y}{\Omega} = d_0 (1 + \eta(t)), \quad (123)$$

$$\frac{d\eta(t)}{dt} = -\frac{1}{\tau} \eta(t) + \frac{\sqrt{2D}}{\tau} \beta(t), \quad (124)$$

where $\eta(t) = \Omega^{1/2} D \epsilon(t) / \tau$. Equations (123) and (124) are consistent with the definition of colored noise described at the beginning of this section. This modified SSA requires that where τ and D are both individually large, that $\tau \gg D$ such that slow switching is not enforced between differing numbers of the ghost species.

Appendix B. Detailed explanation of condition 2

In order to explain the origin of condition 2—a condition on the length scale of colored noise fluctuation compared to the rate of variation of the drift term in equation (47)—we will first consider a more intuitive example. Consider a Brownian particle subject to a force $F(x)$, whose state is specified by both its position x , as well as its velocity v . The set of SDEs governing the state of this particle is then [6]:

$$\frac{dx}{dt} = v, \quad (125)$$

$$\frac{dv}{dt} = \frac{F(x)}{m} - \gamma v + \frac{k_b T \gamma}{m} \Gamma(t), \quad (126)$$

where m is the mass of the particle, γ is the damping coefficient of the frictional force surrounding the particle (frictional force is $-\gamma m v$), $k_b T$ is the thermal energy of the particle, and $\Gamma(t)$ is Gaussian white noise with zero mean and correlator $\langle \Gamma(t) \Gamma(t') \rangle = \delta(t - t')$. The equivalent multivariate FPE for this set of SDEs is [34]:

$$\frac{\partial P(x, v; t)}{\partial t} = \gamma \left[\frac{\partial(vP)}{\partial v} + \frac{k_b T}{m} \frac{\partial^2 P}{\partial v^2} \right] - v \frac{\partial P}{\partial x} - \frac{F(x)}{m} \frac{\partial P}{\partial v}. \quad (127)$$

Now following van Kampen p 216–218 [6], one can utilise singular perturbation theory assuming that the damping coefficient γ is small (although the same procedure could be done for γ large) in order to reduce the above FPE in two variables to an FPE in the position variable x alone. The result after having done this procedure is:

$$\frac{\partial P(x; t)}{\partial t} = -\frac{\partial}{\partial x} \left(\frac{F(x)}{m\gamma} P \right) + \frac{k_b T}{m\gamma} \frac{\partial^2 P}{\partial x^2}. \tag{128}$$

Aside from the requirement that γ must be small, there is another condition required of equation (128) such that it reasonably approximates equation (127). This condition arises *physically* since we realise that if we are to approximate equation (127) by equation (128), then the drift term $F(x)/m\gamma$ must be approximately constant over the distance that the velocity is damped. One finds that the associated ‘length scale’ L over which the velocity is damped is simply the pre-factor of diffusion term in equation (128), i.e., $L = \frac{k_b T}{m\gamma}$ [62]. Enforcing the requirement that $F(x)$ is slowly varying over this length scale we find the inequality $L|F'(x)| \ll |F(x)|$, explicitly:

$$\frac{m\gamma}{k_b T} \gg \left| \frac{F'(x)}{F(x)} \right|, \tag{129}$$

which must be satisfied for our one variable FPE to be a good approximation.

We now return to our colored noise problem and recall the set of SDEs that define our system:

$$\frac{dn}{dt} = h(n) + F(n)\eta + g_2(n)\Gamma(t), \tag{130}$$

$$\frac{d\eta}{dt} = -\frac{1}{\tau}\eta + \frac{1}{\tau}\theta(t), \tag{131}$$

where all functions of n and t are defined in section 3.2. This set of SDEs has an equivalent bi-variate (Stratonovich) FPE given by:

$$\begin{aligned} \frac{\partial P(n, \eta; t)}{\partial t} = & -\frac{\partial}{\partial n} \left(\left(h(n) + F(n)\eta - \frac{1}{2}g_2(n)g_2'(n) \right) P \right) + \frac{1}{\tau} \frac{\partial}{\partial \eta} (\eta P) \\ & + \frac{1}{2} \frac{\partial^2}{\partial n^2} (g_2(n)^2 P) + \frac{1}{\tau} \frac{\partial^2}{\partial \eta^2} P. \end{aligned} \tag{132}$$

We now recall equation (47), i.e., our UCNA approximated one variable FPE, where η was adiabatically eliminated:

$$\frac{\partial P(n; t)}{\partial t} = -\frac{\partial}{\partial n} \left[\left(\tilde{h}(n) + \tilde{g}(n)\tilde{g}'(n) \right) P(n, t) \right] + \frac{\partial^2}{\partial n^2} [\tilde{g}(n)^2 P(n, t)]. \tag{133}$$

Analogously to the case of the Brownian particle, this one variable FPE can only be approximately correct where the variation of the drift term with respect to the length scale of colored noise fluctuations is small. From equation (43), we identify our length scale as the pre-factor of the noise term whose origin is the adiabatic elimination of η , i.e.,

$$L = \frac{F(n)}{C(n, \tau)}. \tag{134}$$

Hence, $C(n, \tau)$ must satisfy the following length scale condition

$$C(n, \tau) \gg F(n) \left| \frac{\partial_n \left(\tilde{h}(n) + \tilde{g}(n)\tilde{g}'(n) \right)}{\tilde{h}(n) + \tilde{g}(n)\tilde{g}'(n)} \right| \quad (135)$$

if equation (133) is to be a good approximation of equation (132), as seen in equation (66) from the main text. Note that one can also make the argument that the diffusion term $\tilde{g}(n)^2$ should also slowly vary with respect to L . However, we generally find that this is satisfied if equation (135) is satisfied, and hence we do not include this as an additional condition on the validity of the UCNA.

ORCID iDs

Ramon Grima  <https://orcid.org/0000-0002-1266-8169>

References

- [1] Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K and Walter P 2015 *Molecular Biology of the Cell* 6th edn (New York: Garland Science)
- [2] Shen-Orr S S, Milo R, Mangan S and Alon U 2002 Network motifs in the transcriptional regulation network of escherichia coli *Nat. Genet.* **31** 64
- [3] Rosenfeld N, Elowitz M B and Alon U 2002 Negative autoregulation speeds the response times of transcription networks *J. Mol. Biol.* **323** 785–93
- [4] Ko M S H 1991 A stochastic model for gene induction *J. Theor. Biol.* **153** 181–94
- [5] Elowitz M B, Levine A J, Siggia E D and Swain P S 2002 Stochastic gene expression in a single cell *Science* **297** 1183–6
- [6] van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* vol 1 (Amsterdam: Elsevier)
- [7] Swain P S, Elowitz M B and Siggia E D 2002 Intrinsic and extrinsic contributions to stochasticity in gene expression *Proc. Natl Acad. Sci.* **99** 12795–800
- [8] Gillespie D T 1977 Exact stochastic simulation of coupled chemical reactions *J. Phys. Chem.* **81** 2340–61
- [9] Gardiner C W *et al* 1985 *Handbook of Stochastic Methods* vol 3 (Berlin: Springer)
- [10] Holehouse J, Cao Z and Grima R 2020 Stochastic modeling of auto-regulatory genetic feedback loops: a review and comparative study *Biophys. J.* **118** 1517–25
- [11] Grima R, Schmidt D R and Newman T J 2012 Steady-state fluctuations of a genetic feedback loop: an exact solution *J. Phys. Chem.* **137** 035104
- [12] Kumar N, Platini T and Kulkarni R V 2014 Exact distributions for stochastic gene expression models with bursting and feedback *Phys. Rev. Lett.* **113** 268105
- [13] Liu P, Yuan Z, Wang H and Zhou T 2016 Decomposition and tunability of expression noise in the presence of coupled feedbacks *Chaos* **26** 043108
- [14] Jia C, Yin G G, Zhang M Q *et al* 2019 Single-cell stochastic gene expression kinetics with coupled positive-plus-negative feedback *Phys. Rev. E* **100** 052406
- [15] Jia C and Grima R 2020 Small protein number effects in stochastic models of autoregulated bursty gene expression *J. Phys. Chem.* **152** 084115
- [16] Cao Z and Grima R 2018 Linear mapping approximation of gene regulatory networks with stochastic dynamics *Nat. Commun.* **9** 3305
- [17] Kurasov P, Lück A, Mugnolo D and Wolf V 2018 Stochastic hybrid models of gene regulatory networks—a pde approach *Math. Biosci.* **305** 170–7
- [18] Andreychenko A, Bortolussi L, Grima R, Thomas P and Wolf V 2017 Distribution approximations for the chemical master equation: comparison of the method of moments and the system size expansion *Modeling Cellular Systems* (Berlin: Springer) pp 39–66

- [19] Thomas P, Straube A V and Grima R 2012 The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions *BMC Syst. Biol.* **6** 39
- [20] Friedman N, Cai L and Xie X S 2006 Linking stochastic dynamics to population distribution: an analytical framework of gene expression *Phys. Rev. Lett.* **97** 168302
- [21] Ochab-Marcinek A and Tabaka M 2015 Transcriptional leakage versus noise: a simple mechanism of conversion between binary and graded response in autoregulated genes *Phys. Rev. E* **91** 012704
- [22] Jędrak J and Ochab-Marcinek A 2016 Influence of gene copy number on self-regulated gene expression *J. Theor. Biol.* **408** 222–36
- [23] Scott M, Ingalls B and Kærn M 2006 Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks *Chaos* **16** 026107
- [24] Toni T and Tidor B 2013 Combined model of intrinsic and extrinsic variability for computational network design with application to synthetic biology *PLoS Comput. Biol.* **9** e1002960
- [25] Keizer E M, Bastian B, Smith R W, Grima R and Fleck C 2019 Extending the linear-noise approximation to biochemical systems influenced by intrinsic noise and slow lognormally distributed extrinsic noise *Phys. Rev. E* **99** 052417
- [26] Roberts E, Be'er S, Bohrer C, Sharma R and Assaf M 2015 Dynamics of simple gene-network motifs subject to extrinsic fluctuations *Phys. Rev. E* **92** 062717
- [27] Jung P and Hänggi P 1987 Dynamical systems: a unified colored-noise approximation *Phys. Rev. A* **35** 4464
- [28] Shahrezaei V, Ollivier J F and Swain P S 2008 Colored extrinsic fluctuations and stochastic gene expression *Mol. Syst. Biol.* **4** 196
- [29] da Costa Pereira G, Forger M, Ramos A F, Radulescu O and Hornos J E M 2013 Multimodality and flexibility of stochastic gene expression *Bull. Math. Biol.* **75** 2600–30
- [30] Dattani J and Barahona M 2017 Stochastic models of gene transcription with upstream drives: exact solution and sample path characterization *J. R. Soc., Interface* **14** 20160833
- [31] Herbach U 2019 Stochastic gene expression with a multistate promoter: Breaking down exact distributions *SIAM J. Appl. Math.* **79** 1007–29
- [32] Li C, Cesbron F, Oehler M, Brunner M and Höfer T 2018 Frequency modulation of transcriptional bursting enables sensitive and rapid gene regulation *Cell Syst.* **6** 409–23
- [33] Holehouse J and Grima R 2019 Revisiting the reduction of stochastic models of genetic feedback loops with fast promoter switching *Biophys. J.* **117** 1311–30
- [34] Gardiner C 2009 *Stochastic Methods* vol 4 (Berlin: Springer)
- [35] Risken H 1996 Fokker–Planck equation *The Fokker–Planck Equation* (Berlin: Springer) pp 63–95
- [36] Minsky B and Khammash M 2006 The finite state projection algorithm for the solution of the chemical master equation *J. Phys. Chem.* **124** 044104
- [37] Gillespie D T 2000 The chemical langevin equation *J. Chem. Phys.* **113** 297–306
- [38] Li C, Da-Jin W and Sheng-Zhi K 1995 Bistable kinetic model driven by correlated noises: unified colored-noise approximation *Phys. Rev. E* **52** 3228
- [39] Fox R F 1986 Functional-calculus approach to stochastic differential equations *Phys. Rev. A* **33** 467
- [40] Fox R F 1986 Uniform convergence to an effective Fokker–Planck equation for weakly colored noise *Phys. Rev. A* **34** 4525
- [41] Fox R F 1987 Stochastic calculus in physics *J. Stat. Phys.* **46** 1145–57
- [42] Fox R F and Roy R 1987 Steady-state analysis of strongly colored multiplicative noise in a dye laser *Phys. Rev. A* **35** 1838
- [43] Grigolini P, Lugiato L A, Mannella R, McClintock P V E, Merri M and Pernigo M 1988 Fokker-planck description of stochastic processes with colored noise *Phys. Rev. A* **38** 1966
- [44] Wong E and Zakai M 1965 On the convergence of ordinary integrals to stochastic integrals *Ann. Math. Stat.* **36** 1560–4
- [45] Pesce G, McDaniel A, Hottovy S, Wehr J and Volpe G 2013 Stratonovich-to-itô transition in noisy systems with multiplicative feedback *Nat. Commun.* **4** 2733
- [46] Nicolas D, Phillips N E and Naef F 2017 What shapes eukaryotic transcriptional bursting? *Mol. BioSyst.* **13** 1280–90
- [47] Cao Z, Filatova T, Oyarzún D A and Grima R 2019 Multi-scale bursting in stochastic gene expression (bioRxiv:717199)

- [48] Zhang J and Zhou T 2019 Stationary moments, distribution conjugation and phenotypic regions in stochastic gene transcription *Mathematical Biosciences and Engineering: MBE* **16** 6134–66
- [49] Thomas P, Popović N and Grima R 2014 Phenotypic switching in gene regulatory networks *Proc. Natl Acad. Sci.* **111** 6994–9
- [50] Choubey S, Kondev J and Sanchez A 2015 Deciphering transcriptional dynamics *in vivo* by counting nascent rna molecules *PLoS Comput. Biol.* **11** e1004345
- [51] Filatova T, Popovic N and Grima R 2020 Statistics of nascent and mature rna fluctuations in a stochastic model of transcriptional initiation, elongation, pausing and termination (<https://doi.org/10.1101/2020.05.13.092650>)
- [52] Chong S, Chen C, Ge H and Xie X S 2014 Mechanism of transcriptional bursting in bacteria *Cell* **158** 314–26
- [53] Halpern K B, Tanami S, Landen S, Chapal M, Szlak L, Hutzler A, Nizhberg A and Itzkovitz S 2015 Bursty gene expression in the intact mammalian liver *Molecular Cell* **58** 147–56
- [54] Mackey M C, Tyrán-Kaminska M and Yvinec R 2013 Dynamic behavior of stochastic gene expression models in the presence of bursting *SIAM J. Appl. Math.* **73** 1830–52
- [55] Cooper G M 2000 *The Cell: A Molecular Approach* 2nd edn (Sunderland, MA: Sinauer Associates)
- [56] McShane E *et al* 2016 Kinetic analysis of protein stability reveals age-dependent degradation *Cell* **167** 803–15
- [57] Pedraza J M and Paulsson J 2008 Effects of molecular memory and bursting on fluctuations in gene expression *Science* **319** 339–43
- [58] Elf J and Ehrenberg M 2003 Fast evaluation of fluctuations in biochemical networks with the linear noise approximation *Genome Res.* **13** 2475–84
- [59] Walczak A M, Sasai M and Wolynes P G 2005 Self-consistent proteomic field theory of stochastic gene switches *Biophys. J.* **88** 828–50
- [60] Pájaro M, Alonso A A, Otero-Muras I and Vázquez C 2017 Stochastic modeling and numerical simulation of gene regulatory networks with protein bursting *J. Theor. Biol.* **421** 51–70
- [61] Herbach U, Bonnaïffoux A, Espinasse T and Gandrillon O 2017 Inferring gene regulatory networks from single-cell data: a mechanistic approach *BMC Syst. Biol.* **11** 105
- [62] Hänggi P and Jung P 1995 *Colored Noise in Dynamical Systems (Advances in Chemical Physics vol 89)* (Hoboken, NJ: JohnWiley & Sons) pp 239–326