THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Extending the linear-noise approximation to biochemical systems influenced by intrinsic noise and slow lognormally distributed extrinsic noise

OPEN ACCESS

Extending the linear-noise approximation to biochemical systems influenced by intrinsic noise and slow lognormally distributed extrinsic noise

Emma M. Keizer,[1, *] Björn Bastian,[2, *] Robert W. Smith,[1] Ramon Grima,[3, †] and Christian Fleck[1, 4, ‡]

[1]*Laboratory of Systems & Synthetic Biology,*
Wageningen UR, Wageningen, The Netherlands
[2]*Institute for Ion Physics and Applied Physics,*
University of Innsbruck, Innsbruck, Austria
[3]*School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom*
[4]*Current address: Department of Biosystems Science*
and Engineering, ETH Zurich, Basel, Switzerland

(Dated: April 10, 2019)

Abstract

It is well known that the kinetics of an intracellular biochemical network is stochastic. This is due to intrinsic noise arising from the random timing of biochemical reactions in the network as well as due to extrinsic noise stemming from the interaction of unknown molecular components with the network and from the cell's changing environment. While there are many methods to study the effect of intrinsic noise on the system dynamics, few exist to study the influence of both types of noise. Here we show how one can extend the conventional linear-noise approximation to allow for the rapid evaluation of the molecule numbers statistics of a biochemical network influenced by intrinsic noise and by slow lognormally distributed extrinsic noise. The theory is applied to simple models of gene regulatory networks and its validity confirmed by comparison with exact stochastic simulations. In particular, we consider three important biological examples. Firstly, we investigate how extrinsic noise modifies the dependence of the variance of the molecule number fluctuations on the rate constants. Secondly, we show how the mutual information between input and output of a network motif is affected by extrinsic noise. And thirdly, we study the robustness of the ubiquitously found feed-forward loop motifs when subject to extrinsic noise.

Keywords: extrinsic noise, LNA, nonlinear effects, synthetic network design

I. INTRODUCTION

Studying the dynamics of biological systems is essential in understanding the design principles underlying biochemical and synthetic networks. However, this task is also challenging given the complexity of interactions between the system's components and its environment. At the molecular level, biological processes possess a certain degree of randomness as chemical reactions are probabilistic events [1]. This stochasticity or *noise* in biological networks has widely varying functional roles, and can be both advantageous and detrimental to cells. Positive effects include phenotypic diversity and the ability to quickly adapt to changing environmental conditions, thus increasing the probability of survival [2, 3]. In contrast, noise can also restrict the ability of a cell to resolve input signals of different strengths and hence

---

* These authors contributed equally

† Correspondence: Ramon.Grima@ed.ac.uk

‡ Correspondence: cfleck@ethz.ch

reduces the information that can be accessed about the external environment [4]. For these reasons, the network's topology either exploits or attenuates noise.

Due to the high complexity of the intracellular machinery, one often can only study a set of reactions between a certain number of observable molecular components. We call this subset of reactions and components *a biochemical system of interest*. The rest of the intracellular and extracellular reactions, species and environmental cues not accounted for, we call *the dynamic environment*. Noise in the biochemical system's dynamics can then stem from either itself or from the dynamic environment. That originating from itself, i.e. from the inherent discreteness of the molecules participating in the biochemical system is called *intrinsic noise* [5]; this type of noise increases with decreasing average molecule numbers and hence is particularly relevant to intracellular dynamics due to the low copy number of genes, messenger RNA (mRNA) and some proteins in a single cell [6]. The noise stemming from the dynamic environment is termed *extrinsic noise* and this affects the biochemical system of interest via modulation of its rate constants. Several studies have shown that a consideration of both types of noise is crucial to understanding gene expression and biochemical dynamics in both prokaryotic and eukaryotic systems [6–8].

The well-mixed stochastic description of point-particle biochemical systems is given by the chemical master equation (CME) [5]. However, this equation rarely can be solved exactly for gene regulatory networks with feedback interactions and when it can, it invariably is for the case of zero extrinsic noise [9, 10]. There are two distinct ways to proceed: (i) exact stochastic simulation or (ii) approximate analytic techniques, which we discuss in this order next.

The stochastic simulation algorithm (SSA), as formulated by Gillespie [1], provides trajectories which are consistent with the CME provided the rate constants are time-independent, i.e, it can only describe intrinsic noise since extrinsic noise manifests as noise in the rate constants. The SSA is noteworthy because it is exact in the limit of an infinite number of samples. Modifications to the SSA that take into account both types of noise have been devised by Shahrezaei *et al.* [11], Anderson [12] and by Voliotis *et al.* [13]. The latter is the most computationally efficient algorithm of the three and also the only one not suffering from numerical integration error. However, the disadvantage of these methods is that a large number of simulations may be required to obtain statistically significant results [3].

Alternatively, various approximation schemes have been derived to obtain analytical ex-

pressions for statistical moments and the marginal distributions of molecular numbers (for a recent review see [14]). Scott *et al.* [15] and Toni *et al.* [16] develop approximate methods based on the linear-noise approximation [5] which allow the calculation of moments for the case of small intrinsic noise together with extrinsic noise originating from rate constants with a static (time-independent) normal distribution. Roberts *et al.* [17] develop a different type of approximation based on the WKB (Wentzel-Kramers-Brillouin) method and the assumption of extrinsic noise originating from rate constants with a static negative binomial distribution. The advantage of the linear-noise approximation methods is the ease with which they can be calculated for systems with a large number of interacting components since the method amounts to solving a Lyapunov equation that can be computed very efficiently [14] while the major disadvantage is that the fluctuating rate parameters can become negative due to the assumption of a normal distribution. In contrast, the WKB method is difficult to extend to more than one variable, however the fluctuating rates are positive.

In this work, we present a novel method based on the linear-noise approximation that is applicable to systems with intrinsic noise together with extrinsic noise originating from rate constants with a static lognormal distribution. It is assumed that the timescale of the extrinsic noise is much longer than that of intrinsic noise, a biologically realistic scenario [2] (correlation times for extrinsic fluctuations in *Escherichia coli* are of the order of 40 minutes which corresponds to the cell cycle period [18, 19], while intrinsic processes typically happen on the order of a minute or shorter timescales). Our method is computationally efficient (due to the use of the linear-noise approximation) while maintaining physical realism by enforcing positive fluctuating rate constants. It hence overcomes the disadvantages of the aforementioned existing frameworks (see previous paragraph). It is also the case that the lognormal distribution appears to be ubiquitous in cell biology [20–22] and hence it is the obvious choice to characterise the generally non-Gaussian distribution of positive fluctuating rates.

The paper is divided as follows. In Section II we develop the theory and derive general expressions for the mean, variances and power spectra of fluctuating molecule numbers in a general biochemical system subjected to lognormal extrinsic noise and intrinsic noise. In Section III, the theory is applied to study how extrinsic noise affects: (1) the second moments of protein numbers in a three-stage model of gene expression and an auto-regulatory genetic feedback loop, (2) the information transfer through a simple biochemical system and (3) the

robustness of feed-forward motifs.

II. THEORY

To correctly model the effects of extrinsic noise, the variables describing extrinsic fluctuations are introduced at the level of the CME. This renders the theory intractable, besides in simplest cases. To overcome this obstacle we propose an asymptotic expansion method relying on three main steps. In Section II A, we follow van Kampen's system size expansion [23] which is truncated at first order to obtain the well-known linear-noise approximation (LNA) [24] as a function of the time-dependent extrinsic variables. Subsequently, we assume timescale separation between the fast intrinsic fluctuations and slowly changing extrinsic variables (Section II B). This allows us, as a final step, to employ a small noise expansion of the extrinsic stochastic variables and obtain closed-form expressions for the means, variances and power spectra of the biochemical system components in Section II C.

A. LNA with extrinsic variables

We consider a chemical network with system volume Ω and N molecular species with copy numbers $X_i(t)$ and concentrations $x_i(t) = X_i(t)/\Omega$, where $i \in \{1, \ldots, N\}$. The chemical reactions are described by R reaction channels with rate functions $\tilde{f}_j(\mathbf{x}, \boldsymbol{\eta}, \Omega)$ with $j \in \{1, \ldots, R\}$ and stoichiometric matrix $\mathbf{S} \in \mathbb{Z}^{N \times R}$. External fluctuations in the rates, that are not included in the microscopic description, are described by slowly changing stochastic variables $\eta_k(t)$, $k \in \{1, \ldots, M\}$, where M equals the number of fluctuating parameters $\bar{c}_k$,

$$\bar{c}_k(t) = c_k \nu_k(t) = c_k \left(1 + \eta_k(t)\right), \tag{1}$$

such that $\langle \bar{c}_k(t) \rangle = c_k$. Lognormal coloured noise [25] $\nu_k(t)$ ensures positive rate constants which avoids spurious production or degradation. Furthermore, lognormal rather than normal distributions have been measured for gene expression rates [11, 18]. The fluctuations around the mean values c_k are then proportional to $\eta_k(t)$. As shown in Appendix A, the lognormal variables $\nu_k(t) = \exp\left(\mu_k(t) - \frac{1}{2}\epsilon_k^2\right)$ with mean 1 may by constructed from an Ornstein-Uhlenbeck process for the normal variables $\mu_k(t)$ with standard deviations ϵ_k.

The probability that the network is in a state with copy numbers $\mathbf{X}$ and stochastic variables $\boldsymbol{\eta}$ at time t is given by $P(\mathbf{X}, \boldsymbol{\eta}, t)$, and $\Omega \tilde{f}_j(\mathbf{x}, \boldsymbol{\eta}, \Omega)\, dt$ is the probability that a

reaction of type j occurs in time $\mathrm{d}t$. The system's state after the reaction is defined by copy numbers $X_i + S_{ij}$ and the value of $\boldsymbol{\eta}$ at that time.

The chemical master equation describing the microscopic system dynamics is then given by

$$\frac{\mathrm{d}P(\mathbf{X}, \boldsymbol{\eta}, t)}{\mathrm{d}t} = \Omega \sum_{j=1}^{R} \left(\prod_{i=1}^{N} E_i^{-S_{ij}} - 1 \right) \tilde{f}_j(\mathbf{x}, \boldsymbol{\eta}, \Omega) P(\mathbf{X}, \boldsymbol{\eta}, t), \tag{2}$$

where E_i is a step operator that is defined by the action $E_i^n g(X_1, ..., X_i, ..., X_N) = g(X_1, ..., X_i + n, ..., X_N)$, the product of which takes into account all system states that can evolve to the state given by $\mathbf{X}$ and $\boldsymbol{\eta}$. The probability flux away from state $(\mathbf{X}, \boldsymbol{\eta})$ due to reaction j is given by $-\Omega \tilde{f}_j(\mathbf{x}, \boldsymbol{\eta}, \Omega) P(\mathbf{X}, \boldsymbol{\eta}, t)$.

If the system volume is sufficiently large, the fluctuations of the concentrations $\mathbf{x}$ due to a couple of reactions on short timescales are relatively small. Thus, the rate functions change more or less continuously on a larger timescale. We can therefore define a macroscopic limit $\Omega \to \infty$ with the concentrations, transition rates and deterministic reaction rate equations given by:

$$\boldsymbol{\phi} = \lim_{\Omega \to \infty} \mathbf{x}, \tag{3a}$$

$$\mathbf{f}(\boldsymbol{\phi}, \boldsymbol{\eta}) = \lim_{\Omega \to \infty} \tilde{\mathbf{f}}(\mathbf{x}, \boldsymbol{\eta}, \Omega), \tag{3b}$$

$$\frac{\mathrm{d}\boldsymbol{\phi}}{\mathrm{d}t} = \mathbf{g}(\boldsymbol{\phi}, \boldsymbol{\eta}) = \mathbf{S}\mathbf{f}(\boldsymbol{\phi}, \boldsymbol{\eta}). \tag{3c}$$

Following the system size expansion by van Kampen [24] we relate the microscopic and macroscopic vectors via

$$\mathbf{x} = \boldsymbol{\phi} + \Omega^{-\frac{1}{2}} \boldsymbol{\xi}, \tag{4}$$

where a new variable $\boldsymbol{\xi}$ denotes the microscopic fluctuations around the macroscopic concentrations $\boldsymbol{\phi}$. It obeys the stochastic differential equation [25]

$$\mathrm{d}\boldsymbol{\xi}(t) = \mathbf{A}(\boldsymbol{\eta}(t)) \boldsymbol{\xi}(t) \, \mathrm{d}t + \mathbf{B}(\boldsymbol{\eta}(t)) \, \mathrm{d}\mathbf{W}(t) \tag{5}$$

with the Wiener process increments $\mathrm{d}\mathbf{W}(t)$, Jacobian matrix $\mathbf{A}(\boldsymbol{\eta})$ and diffusion matrix $\mathbf{B}(\boldsymbol{\eta})$.

B. Timescale separation between dynamics of intrinsic and extrinsic fluctuations

1. Integrating a stationary solution

As Eq. (9) generally cannot be solved analytically, we assume that the supremum of the timescales of intrinsic noise as given by the absolute value of the inverse of the eigenvalues of the Jacobian $\mathbf{A}$ is much less than the timescale of extrinsic noise. This allows us to split the time axis into intervals $[t_n, t_n + \triangle t]$, on which the extrinsic variables $\boldsymbol{\eta}$ are treated as constants $\boldsymbol{\eta}^n \equiv \boldsymbol{\eta}(t_n)$. For well defined stationary solutions we require the existence of a unique macroscopic stationary solution $\boldsymbol{\phi}^s(\boldsymbol{\eta}^n)$ of Eq. (3c),

$$\mathbf{g}(\boldsymbol{\phi}^s(\boldsymbol{\eta}^n), \boldsymbol{\eta}^n) = 0, \tag{6}$$

and that the Jacobian matrix has only eigenvalues with negative real part (*i.e.* a stable monotonic system). The stationary Jacobian and diffusion matrices are

$$\mathbf{A}(\boldsymbol{\eta}^n) = \frac{\partial \mathbf{g}}{\partial \boldsymbol{\phi}}(\boldsymbol{\phi}^s(\boldsymbol{\eta}^n), \boldsymbol{\eta}^n), \tag{7}$$

$$\mathbf{B}(\boldsymbol{\eta}^n)\mathbf{B}(\boldsymbol{\eta}^n)^T = \mathbf{S}\,\mathrm{diag}\left(\mathbf{f}(\phi^s(\boldsymbol{\eta}^n), \boldsymbol{\eta}^n)\right)\mathbf{S}^T. \tag{8}$$

Aiming at a stationary solution $\mathbf{x}(t)$ that makes it possible to obtain expressions for the mean, variance and power spectrum, we further follow the well known *linear-noise approximation* [24]: we linearise the rate equations (3c) and use Eqs. (4) and (5) to obtain the linear stochastic differential equation for $t \in [t_n, t_n + \triangle t]$. The stochastic differential equation describing the fluctuations in molecule numbers can be derived by using Eq. (4) (with ϕ replaced by $\phi^s(\eta^n)$) together with Eq. (5) (with $\eta(t)$ replaced by η^n) to obtain:

$$\mathrm{d}\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\eta}^n)(\mathbf{x}(t) - \boldsymbol{\phi}^s(\boldsymbol{\eta}^n))\,\mathrm{d}t + \frac{1}{\sqrt{\Omega}}\mathbf{B}(\boldsymbol{\eta}^n)\,\mathrm{d}\mathbf{W}(t). \tag{9}$$

with the stationary solution [25]:

$$\mathbf{x}^s(t) = \boldsymbol{\phi}^s(\boldsymbol{\eta}^n) + \frac{1}{\sqrt{\Omega}}\boldsymbol{\xi}(\boldsymbol{\eta}^n, t), \tag{10a}$$

$$\boldsymbol{\xi}(\boldsymbol{\eta}^n, t) = \int_{-\infty}^t e^{\mathbf{A}(\boldsymbol{\eta}^n)(t-t')}\mathbf{B}(\boldsymbol{\eta}^n)\,\mathrm{d}\mathbf{W}(t'). \tag{10b}$$

2. Calculating mean concentrations and variances

To evaluate averages and variances of the stationary concentrations $\mathbf{x}^s$ we denote averaging over intrinsic fluctuations $\boldsymbol{\xi}$ by $\langle\ \rangle_i$ and over the extrinsic variables $\boldsymbol{\eta}$ by $\langle\ \rangle_e$. The mean

concentrations then simplify to

$$\langle\langle \mathbf{x}^s \rangle_i \rangle_e = \langle \boldsymbol{\phi}^s(\boldsymbol{\eta}) \rangle_e \tag{11}$$

where $\langle \boldsymbol{\xi} \rangle_i$ evaluates to zero on each of the time intervals [24]. In the same way, the covariance matrix of $\mathbf{x}^s$ in the timescale separation approximation can be written as

$$
\begin{aligned}
\mathbf{V}(\mathbf{x}^s) &= \left\langle (\mathbf{x}^s - \langle \mathbf{x}^s \rangle)(\mathbf{x}^s - \langle \mathbf{x}^s \rangle)^T \right\rangle \\
&= \mathbf{V}(\boldsymbol{\phi}^s) + \frac{1}{\Omega}\mathbf{V}(\boldsymbol{\xi}).
\end{aligned}
\tag{12}
$$

where $\langle\ \rangle$ abbreviates $\langle\langle\ \rangle_i\rangle_e$ and we define the covariance matrices of $\boldsymbol{\phi}^s$ and $\boldsymbol{\xi}$ as

$$\mathbf{V}(\boldsymbol{\phi}^s) = \left\langle \boldsymbol{\phi}^s \boldsymbol{\phi}^{sT} \right\rangle_e - \langle \boldsymbol{\phi}^s \rangle_e \left\langle \boldsymbol{\phi}^{sT} \right\rangle_e, \tag{13}$$

$$\mathbf{V}(\boldsymbol{\xi}) = \left\langle \langle \boldsymbol{\xi} \boldsymbol{\xi}^T \rangle_i \right\rangle_e. \tag{14}$$

One can obtain $\mathbf{V}(\boldsymbol{\xi})$ algebraically. For the sake of brevity we anticipate the result using Eq. (20) and (22)

$$\mathbf{V}(\boldsymbol{\xi}) = \left\langle \langle \boldsymbol{\xi} \boldsymbol{\xi}^T \rangle_i \right\rangle_e = \langle \mathbf{G}_i(t,t) \rangle_e = \langle \mathbf{C}(\boldsymbol{\eta},\boldsymbol{\eta}) \rangle_e \tag{15}$$

were the Lyapunov matrix $\mathbf{C}$ is evaluated at equal times and we calculate the time correlation function $\mathbf{G}(t_1, t_2)$ as an intermediate step to calculate the power spectrum in what follows.

3. *Calculating power spectra via Fourier transformation*

The spectrum matrix of a stationary stochastic process is connected to the time correlation function by the *Wiener-Khinchin theorem* via a Fourier transformation if the time correlation is sufficiently smooth [24, 25],

$$\mathbf{P}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\omega\tau} \mathbf{G}(\tau, 0) \, \mathrm{d}\tau. \tag{16}$$

We first introduce the time correlation function of a stationary solution $\mathbf{x}^s$ (Eq. 10) as

$$\mathbf{G}(t_1, t_2) = \left\langle \left(\mathbf{x}^s(t_1) - \langle \mathbf{x}^s \rangle\right)\left(\mathbf{x}^s(t_2) - \langle \mathbf{x}^s \rangle\right)^T \right\rangle \tag{17}$$

$$= \mathbf{G}_e(t_1, t_2) + \frac{1}{\Omega} \langle \mathbf{G}_i(t_1, t_2) \rangle_e. \tag{18}$$

The second equality holds under timescale separation conditions and represents the sum of the time correlation function of the macroscopic stationary state

$$\mathbf{G}_e(t_1, t_2) = \left\langle \boldsymbol{\phi}^s(\boldsymbol{\eta}^1)\boldsymbol{\phi}^s(\boldsymbol{\eta}^2)^T \right\rangle_e - \langle \boldsymbol{\phi}^s \rangle_e \left\langle \boldsymbol{\phi}^{sT} \right\rangle_e \tag{19}$$

8

and the time correlation function of the intrinsic noise subject to slow extrinsic fluctuations

$$\mathbf{G}_i(t_1, t_2) = \left\langle \boldsymbol{\xi}\left(\boldsymbol{\eta}^1\right) \boldsymbol{\xi}\left(\boldsymbol{\eta}^2\right)^T \right\rangle_i. \tag{20}$$

To evaluate Eq. (20) we follow the calculation of the variance in stationary solution by Gardiner [25] (having defined $\mathbf{A}$ as the Jacobian it has opposite sign to the matrix in the reference) and generalise it by evaluating the results at different times to obtain

$$\mathbf{G}_i(t_1, t_2) = e^{\mathbf{A}(\boldsymbol{\eta}^1)(t_1 - \min(t_1, t_2))} \ \mathbf{C}\left(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2\right) e^{\mathbf{A}(\boldsymbol{\eta}^2)^T (t_2 - \min(t_1, t_2))} \tag{21}$$

where the $\mathbf{C}$ matrix is defined by the Lyapunov equation

$$\mathbf{A}(\boldsymbol{\eta}^1)\, \mathbf{C}\left(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2\right) + \mathbf{C}\left(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2\right) \mathbf{A}(\boldsymbol{\eta}^2)^T$$
$$= -\mathbf{B}\left(\boldsymbol{\eta}^1\right) \mathbf{B}\left(\boldsymbol{\eta}^2\right)^T. \tag{22}$$

Exploiting timescale separation, we split the spectrum matrix Eq. (16) into two terms according to Eq. (18),

$$\mathbf{P}(\omega) = \mathbf{P}_e(\omega) + \frac{1}{\Omega} \left\langle \mathbf{P}_i(\omega) \right\rangle_e. \tag{23}$$

With Eq. (21) and $\boldsymbol{\eta}^n \equiv \boldsymbol{\eta}(t_n)$ with $t_1 = \tau$ and $t_2 = 0$ we can express $\left\langle \mathbf{P}_i(\omega) \right\rangle_e$ in terms of the matrices

$$\mathbf{R}(\omega) = \left\langle \int_0^\infty e^{-\left(-\mathbf{A}(\boldsymbol{\eta}^1) + i\omega\right)\tau} \mathbf{C}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)\, \mathrm{d}\tau \right\rangle_e, \tag{24a}$$

$$\mathbf{R}(\omega)^{*T} = \left\langle \int_0^\infty \mathbf{C}(\boldsymbol{\eta}^2, \boldsymbol{\eta}^1) e^{\left(\mathbf{A}(\boldsymbol{\eta}(\tau))^T + i\omega\right)\tau}\, \mathrm{d}\tau \right\rangle_e. \tag{24b}$$

Assuming stationarity of $\boldsymbol{\eta}(t)$ it can be shown that

$$\mathbf{R}(\omega) + \mathbf{R}(\omega)^{*T} = \int_{-\infty}^{+\infty} e^{-i\omega\tau} \mathbf{G}_i(\tau, 0)\, \mathrm{d}\tau. \tag{25}$$

Further, we split the Taylor expansion of the Jacobian in two,

$$\mathbf{A}(\boldsymbol{\eta}) = \mathbf{A}^0 + \left(\mathbf{A}(\boldsymbol{\eta}) - \mathbf{A}^0\right), \qquad\qquad \mathbf{A}^0 \equiv \mathbf{A}(\mathbf{0}), \tag{26}$$

9

and summarise the quantities of interest for Eq. (23):

$$\mathbf{P}_e(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\omega\tau} \Big(\big\langle \boldsymbol{\phi}^s(\boldsymbol{\eta}^1)\boldsymbol{\phi}^s(\boldsymbol{\eta}^2)^T \big\rangle_e \tag{27}$$

$$- \big\langle \boldsymbol{\phi}^s \big\rangle_e \big\langle \boldsymbol{\phi}^{sT} \big\rangle_e \Big) \, d\tau,$$

$$\mathbf{R}(\omega) = \int_0^\infty e^{-\left(-\mathbf{A}^0 + i\omega\right)\tau} \times \tag{28}$$

$$\left\langle e^{\left(\mathbf{A}(\boldsymbol{\eta}^1) - \mathbf{A}^0\right)\tau} \mathbf{C}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \right\rangle_e \, d\tau,$$

$$\langle \mathbf{P}_i(\omega) \rangle_e = \frac{1}{2\pi} \left(\mathbf{R}(\omega) + \mathbf{R}(\omega)^{*T} \right). \tag{29}$$

C.   Small noise expansion

In the third and last step, we expand equations (11), (13), (15), (27) and (28) in Taylor series in the M noise variables $\boldsymbol{\eta}$. To calculate expected values, also of the time dependent integrands in (27) and (28), we consequently need the n-point time correlation functions of the extrinsic noise variables. To this end, we need the following results derived in Appendix B.

For smooth functions $y_k(\mu_k)$ of normal stochastic variables μ_k with $(m_k)^{\text{th}}$ derivatives $y_k^{(m_k)}(\mu_k)$, the derived n-point time correlation function reads:

$$\langle y_1 \ldots y_n \rangle = \sum_{|\mathbf{d}^n|=0}^\infty \left(\prod_{k=1}^n \left\langle y_k^{(m_k)}(\mu_k) \right\rangle \prod_{\substack{i,j=1 \\ i<j}}^n \frac{\Delta_{ij}^{d_{ij}}}{d_{ij}!} \right) \tag{30}$$

which generalises a previous result for $n = 2$ [26]. Each value of $|\mathbf{d}^n| = \sum d_{ij}$ involves summation over all index tuples $\mathbf{d}^n = (d_{12}, d_{13}, d_{23}, \ldots, d_{(n-1)\,n})$. The 2-point time correlation functions of the normal stochastic variables are denoted by $\Delta_{ij} = \langle \mu_i \mu_j \rangle$. The number m_k of variables μ_k in the product of Δ_{ij} functions defines the order of the derivatives, and is given in terms of the $\mathbf{d}^n$ tuple by $m_k = \sum_{i=1}^{k-1} d_{ik} + \sum_{j=k+1}^n d_{kj}$. We point out the simplicity of this result for lognormal variables with mean $\langle \nu_k \rangle = 1$,

$$\langle \nu_1 \ldots \nu_n \rangle = \exp\left(\sum_{1 \le i < j \le n} \Delta_{ij} \right). \tag{31}$$

In the small noise expansion we use the shifted $\eta_k = \nu_k - 1$ with the time correlation functions (see Appendix B)

$$\langle \eta_1 \ldots \eta_n \rangle = \sum_{u=\lfloor \frac{n+1}{2} \rfloor}^\infty \sideset{}{'}\sum_{|\mathbf{d}^n|=u} \prod_{\substack{i,j=1 \\ i<j}}^n \frac{\Delta_{ij}^{d_{ij}}}{d_{ij}!} \tag{32}$$

10

where the prime denotes the restriction of the sum to terms where for each $j \in \{1, \ldots, n\}$ there is an $i < j$ or $k > j$ with $d_{ij} \neq 0$ or $d_{jk} \neq 0$. The floor function $\lfloor \frac{n+1}{2} \rfloor$ gives the smallest u for which this can be satisfied. Selected terms of the final results that we present in the following are evaluated in Appendix G to further clarify the complex notation.

1. Mean

We define the Taylor expansion of the mean concentrations in Eq. (11) with the unusual but beneficial notation

$$\langle \mathbf{x}^s(\boldsymbol{\eta}) \rangle = \langle \boldsymbol{\phi}^s(\boldsymbol{\eta}) \rangle = \sum_{n=0}^{\infty} \sum_{\#\mathbf{r}=n} \phi^s_{(\mathbf{r})} \langle \eta_{(\mathbf{r})} \rangle \tag{33}$$

where the multi-indices

$$\mathbf{r} = (r_1, \ldots, r_n), \quad \#\mathbf{r} = n, \quad r_i \in \{1, \ldots, M\} \tag{34a}$$

are sorted by their length $\#\mathbf{r}$ and we denote the Taylor coefficients and products of the noise variables as

$$\phi^s_{(\mathbf{r})} \equiv \frac{1}{n!} \frac{\partial}{\partial \eta_{r_1}} \cdots \frac{\partial}{\partial \eta_{r_n}} \phi^s(\boldsymbol{\eta}) \big|_{\boldsymbol{\eta}=\mathbf{0}}, \tag{34b}$$

$$\eta_{(\mathbf{r})} \equiv \eta_{r_1} \ldots \eta_{r_n}. \tag{34c}$$

The sum over $\#\mathbf{r} = n$ involves all $\mathbf{r}$ tuples of length n. By inserting Eq. (32) into Eq. (33) we obtain the final result for the mean concentrations in terms of the covariances of the independent normal stochastic variables μ_i with standard deviations ϵ_i, that is, their 2-point correlation functions evaluated at equal times

$$\Gamma_{ij} \equiv \Delta_{ij}(0) = \delta_{ij} \epsilon_i^2 \tag{35}$$

with the Kronecker delta (see details in Appendix C):

$$\langle \mathbf{x}^s \rangle = \sum_{u=0}^{\infty} \sum_{n=0}^{2u} \sum_{\#\mathbf{r}=n} \phi^s_{(\mathbf{r})} \sideset{}{'}\sum_{|\mathbf{d}^n|=u} \prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{1}{d_{ij}!} \left(\Gamma_{r_i r_j} \right)^{d_{ij}}. \tag{36}$$

The prime at the sum denotes the same restriction as in Eq. (32): the sum is running over all tuples $\mathbf{d}^n = (d_{12}, d_{13}, d_{23}, \ldots, d_{(n-1)\,n})$ with $|\mathbf{d}^n| = \sum d_{ij} = u$ that obey the condition that for each $j \in \{1, \ldots, n\}$ there is a $i < j$ or $k > j$ such that $d_{ij} \neq 0$ or $d_{jk} \neq 0$.

11

2. Variance

With Eq. (15) we write the matrix $V(\boldsymbol{\xi}) = \langle \mathbf{C}(\boldsymbol{\eta}, \boldsymbol{\eta}) \rangle$ in analogy to $\langle \mathbf{x}^s \rangle = \langle \boldsymbol{\phi}^s \rangle$. Because $\mathbf{C}$ is evaluated at equal times here, we may expand it according to (34b) with Taylor coefficients $\mathbf{C}_{(\mathbf{r})}$. From Eq. (36) we then obtain

$$\mathbf{V}(\boldsymbol{\xi}) = \sum_{u=0}^{\infty} \sum_{n=0}^{2u} \sum_{\#\mathbf{r}=n} \mathbf{C}_{(\mathbf{r})} {\sum_{|\mathbf{d}^n|=u}}' \prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{1}{d_{ij}!} \left(\Gamma_{r_i r_j} \right)^{d_{ij}}. \tag{37}$$

The derivation of the series for $\mathbf{V}(\boldsymbol{\phi}^s)$ is conducted similarly. For the expansion of $\boldsymbol{\phi}^s \boldsymbol{\phi}^{s\,T}$ in Eq. (13), an ordinary multi-index $\mathbf{q} = (q_1, q_2)$, $|\mathbf{q}| = q_1 + q_2$ with integer $q_1, q_2 \geq 1$ is used to define the lengths of two multi-indices $\mathbf{r}^1$ and $\mathbf{r}^2$ (see Eq. 34) and finally

$$\mathbf{V}(\boldsymbol{\phi}^s) = \sum_{u=1}^{\infty} \sum_{n=2}^{2u} \sum_{|\mathbf{q}|=n} \sum_{\#\mathbf{r}^1=q_1} \sum_{\#\mathbf{r}^2=q_2} \boldsymbol{\phi}^s_{(\mathbf{r}^1)} \boldsymbol{\phi}^{s\,T}_{(\mathbf{r}^2)} {\sum_{|\mathbf{d}^n|=u}}'' \prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{1}{d_{ij}!} \left(\Gamma_{f_i^2 f_j^2} \right)^{d_{ij}}. \tag{38}$$

The doubly primed sum (see Appendix D) denotes the restriction of the summation by one additional condition: there is at least one $d_{ij} \neq 0$ for which $i \leq q_1$ and $j > q_1$ since all other terms cancel in the subtraction in Eq. (13). For later use, we define the index functions $f_i^k(\mathbf{r}^1, \ldots, \mathbf{r}^k)$ for a generalised set of multi-indices $\mathbf{r}^1, \ldots, \mathbf{r}^k$ with $\#\mathbf{r}^i = q_i$ and $\mathbf{q} = (q_1, \ldots, q_k)$,

$$f_i^k = \begin{cases} r_i^1 & \text{if } 1 \leq i \leq q_1 \\ r_{(i-q_1)}^2 & \text{if } q_1 < i \leq q_1 + q_2 \\ \vdots & \vdots \\ r_{(i-|\mathbf{q}|+q_k)}^k & \text{if } q_1 + \cdots + q_{k-1} < i \leq |\mathbf{q}| \end{cases} \tag{39}$$

that we have used in Eq. (38) to refer to $\mathbf{r}^1$ and $\mathbf{r}^2$.

3. Power spectrum

In Eq. (27) for the spectrum matrix $\mathbf{P}_e(\omega)$ the integral from $-\infty$ to 0 is the complex conjugate of the integral from 0 to ∞. Thus the small noise expansion proceeds in complete analogy to $\mathbf{V}(\boldsymbol{\phi}^s)$ in Eq. (13) with the result in Eq. (38) when the product of covariances

12

$\Gamma_{f_i^2 f_j^2}$ is replaced by the Fourier transform plus its complex conjugate

$$\frac{1}{2\pi} \int_0^\infty e^{-i\omega\tau} \prod_{\substack{i,j=1 \\ i<j}}^n \frac{1}{d_{ij}!} \left(\Delta_{f_i^2 f_j^2}(t_i - t_j) \right)^{d_{ij}} \mathrm{d}\tau \tag{40}$$

where the 2-point time correlation functions are

$$\Delta_{ij}(t_i - t_j) \equiv \langle \mu_i(t_i)\mu_j(t_j) \rangle = \Gamma_{ij} e^{-K_i|t_i - t_j|} \tag{41}$$

with inverse correlation times K_i. We finally obtain

$$\mathbf{P}_e(\omega) = \sum_{u=1}^\infty \sum_{n=2}^{2u} \sum_{|\mathbf{q}|=n} \sum_{\#\mathbf{r}^1=q_1} \sum_{\#\mathbf{r}^2=q_2} \boldsymbol{\phi}^s_{(\mathbf{r}^1)} \boldsymbol{\phi}^{s}_{(\mathbf{r}^2)}{}^T \times$$

$$\sum_{|\mathbf{d}^n|=u}'' \frac{\pi^{-1} \Theta(\mathbf{d}^n, \mathbf{r}^1)}{\omega^2 + \Theta^2(\mathbf{d}^n, \mathbf{r}^1)} \prod_{\substack{i,j=1 \\ i<j}}^n \frac{1}{d_{ij}!} \left(\Gamma_{f_i^2 f_j^2} \right)^{d_{ij}} \tag{42}$$

with $\Theta(\mathbf{d}^n, \mathbf{r}^1)$ derived from integral (40) in Appendix E,

$$\Theta(\mathbf{d}^n, \mathbf{r}^1) = \sum_{i=1}^{\#\mathbf{r}^1} \sum_{j=\#\mathbf{r}^1+1}^n d_{ij} K_{r_i^1} . \tag{43}$$

To calculate the second spectrum matrix under influence of extrinsic fluctuations, $\langle \mathbf{P}_i(\omega) \rangle$ in Eq. (29), the $\mathbf{R}(\omega)$ matrix in Eq. (28) requires expanding. First, we expand the Jacobian matrix $\mathbf{A}(\boldsymbol{\eta}^1)$ at time t_1 in the same way as $\boldsymbol{\phi}^s(\boldsymbol{\eta})$ in Eq. (33) with Taylor coefficients $\mathbf{A}_{(\mathbf{r})}$ using the multi-index $\mathbf{r}$ from Eq. (34). The Lyapunov matrix $\mathbf{C}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$ (Eq. 22) is evaluated with respect to two arguments corresponding to different times. Thus, the Taylor expansion requires a second multi-index

$$\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_a), \qquad \#\boldsymbol{\sigma} = a, \qquad \sigma_i \in \{1, 2\}, \tag{44}$$

such that index σ_i specifies the argument with respect to which the derivative is taken for the Taylor coefficients and to which the components η_i belong,

$$\mathbf{C}_{(\mathbf{r}, \boldsymbol{\sigma})} \equiv \frac{1}{a!} \frac{\partial}{\partial \eta_{r_1}^{\sigma_1}} \cdots \frac{\partial}{\partial \eta_{r_a}^{\sigma_a}} \mathbf{C}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)\big|_{\boldsymbol{\eta}^1 = \boldsymbol{\eta}^2 = \mathbf{0}}, \tag{45}$$

$$\eta_{(\mathbf{r}, \boldsymbol{\sigma})} \equiv \eta_{r_1}^{\sigma_1} \cdots \eta_{r_a}^{\sigma_a}. \tag{46}$$

The Taylor expansion of $\exp\left((\mathbf{A}(\boldsymbol{\eta}^1) - \mathbf{A}^0))\tau\right)$ in Eq. (28) in c'th order involves c different Taylor coefficients of $\mathbf{A}$. To distinguish them, we use sets of multi-indices $\mathbf{r}^1, \ldots, \mathbf{r}^c$ with an

ordinary multi-index

$$\mathbf{q} = (q_1, \ldots, q_c), \qquad\qquad |\mathbf{q}| = q_1 + \cdots + q_c \qquad (47)$$

and $\#\mathbf{r}^i = q_i \geq 1$ for all $i = (1, \ldots, c)$. The result of the expansion and integration of $\mathbf{R}(\omega)$ in Appendix F reads

$$\mathbf{R}(\omega) = \sum_{u=0}^{\infty} \sum_{n=0}^{2u} \sum_{a=0}^{n} \sum_{|\mathbf{q}|=n-a} \sum_{\#\mathbf{r}^1=q_1} \cdots \sum_{\#\mathbf{r}^c=q_c} \sum_{\#\mathbf{r}^{c+1}=a} \sum_{\#\boldsymbol{\sigma}=a}$$
$$\times \sum_{|\mathbf{d}^n|=u}' \frac{1}{\left(-\mathbf{A}^0 + \theta(\mathbf{d}^n, |\mathbf{q}|, \mathbf{r}^{c+1}, \boldsymbol{\sigma}) + i\omega\right)^{c+1}}$$
$$\times \mathbf{A}_{(\mathbf{r}^1)} \ldots \mathbf{A}_{(\mathbf{r}^c)} \mathbf{C}_{(\mathbf{r}^{c+1}, \boldsymbol{\sigma})} \prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{1}{d_{ij}!} \left(\Gamma_{f_i^{c+1} f_j^{c+1}}\right)^{d_{ij}} \qquad (48)$$

where the sum $\sum_{|\mathbf{q}|=n-a}$ is carried out over all possible c, $(q_1, \ldots, q_c)$, $q_i \geq 1$ with $q_1 + \cdots + q_c + a = n$ and

$$\theta(\mathbf{d}^n, |\mathbf{q}|, \mathbf{r}^{c+1}, \boldsymbol{\sigma}) = \sum_{\substack{i,j=1 \\ i<j}}^{n} d_{ij} \, K_{r_{j-|\mathbf{q}|}^{c+1}} \, \beta_{ij}(|\mathbf{q}|), \qquad (49)$$

$$\beta_{ij}(x) = \begin{cases} 1 & \text{if } i \leq x < j \text{ and } \sigma_{(j-x)} \neq 1, \\ 1 & \text{if } x < i < j \text{ and } \sigma_{(j-x)} \neq \sigma_{(i-x)}, \\ 0 & \text{else.} \end{cases} \qquad (50)$$

Finally, we add the complex conjugate to $\mathbf{R}(\omega)$ and divide by 2π to obtain the spectrum matrix $\langle \mathbf{P}_i(\omega) \rangle$ as described by Eq. (29).

4. Automated sum evaluation

While the derivation of the presented results involves a double expansion and a rather complicated notation, its strength lies in the closed-form expressions for the mean, variance and power spectrum of a biochemical system under time scale separation conditions. Stochastic modelling of the underlying chemical master equation with extrinsic fluctuations requires an extension of the Gillespie algorithm [11, 27] and it is usually difficult or inefficient to achieve accuracy at different timescales. A fast automated evaluation of the closed-form

expressions allows for a systematic approach to analyse the effect of extrinsic fluctuations as a function of different parameters.

The automated sum evaluation has been implemented with the SymPy library for symbolic mathematics [28] in the `ext_noise_expansion` program that can be obtained from `github.com` [29]. Simple systems can be partially or fully evaluated symbolically while larger systems may require numerical parameter values. The limiting step is the calculation of the stationary state ϕ^s in Eq. (6) as an analytical function of the extrinsic variables. For this task, an external solver adapted to the system of interest may be used. All sums are formally evaluated before inserting the Taylor coefficients and subsequent term simplification. Taylor coefficients for ϕ^s, $\mathbf{A}$ and $\mathbf{B}$ are directly calculated using memoisation to avoid multiple evaluations. Taylor coefficients of $\mathbf{C}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$ from the Lyapunov equation Eq. (22) are obtained by expansion and recursive coefficient comparison (this leads to Lyapunov equations for each coefficient that can be constructed explicitly for any order [30]). More details are presented in Appendix G.

D. Stochastic simulations

Stochastic simulations were performed using the Extrande algorithm by Voliotis *et al.* [13] implemented in `C++11`. Each stochastic simulation data point in Figures 1 and 2 was obtained from 100 trajectories of 10^8 seconds each.

E. Optimisation routine

All optimisations were performed using the L-BFGS-B algorithm (`fmin_l_bfgs_b()`) provided in the SciPy library [31]. Parameters were generated from a uniform distribution with bounds $[0.0001, 20]$.
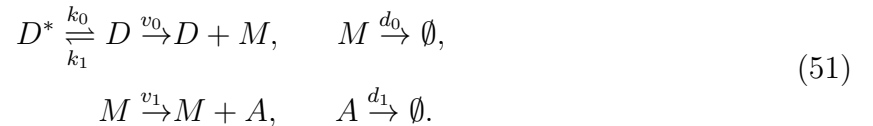
III. RESULTS

Here, we will present several applications of the theory to study features of noisy biochemical networks. We first introduce a linear model of gene expression to establish the (limits of) validity of our method and quantify the effects of extrinsic noise in different

system parameters on intrinsic and extrinsic cell-to-cell variability. We go on to study the potential of negative feedback control to suppress gene expression noise in the presence of extrinsic noise. In a third example, we apply the notion of mutual information to a simple biochemical system to analyse how extrinsic fluctuations affect a network's ability to relay information. Finally, we present an efficient multi-objective optimisation scheme that combines our analytical framework with deterministic dynamics to obtain optimal network topologies and parameters for feed-forward loops.

A. Nonlinear effects of extrinsic fluctuations on a linear model of gene expression

We start by verifying the validity of our method on the well-known three-stage model of gene expression [7] by introducing coloured noise in various system parameters and studying its effect on the mean and variance of protein numbers. In this model, the promoter can switch between the active and inactive state through e.g. the binding of transcription factors, changes in the chromatin structure, or through the binding of RNA polymerase [7, 32]. Only when the promoter is in the active state, mRNA can be transcribed and subsequently translated into proteins. Although each step represents several biochemical reactions, the processes of promoter configuration switching, transcription, translation, as well as degradation of mRNA and protein molecules are assumed to obey first-order kinetics [33]. The gene expression model, shown in Figure 1(a), includes the dynamics of active promoter sites D, inactive promoter sites D^*, mRNA molecules M, and protein molecules A (note that this notation is unrelated to any terms of the same name in Section II), and is described by the following reactions:

$$D^* \underset{k_1}{\overset{k_0}{\rightleftharpoons}} D \overset{v_0}{\rightarrow} D + M, \qquad M \overset{d_0}{\rightarrow} \emptyset,$$
$$M \overset{v_1}{\rightarrow} M + A, \qquad A \overset{d_1}{\rightarrow} \emptyset. \tag{51}$$

We will now show how to write down the LNA as in Eq. (9) for extrinsic noise in one of the reaction rates of this system. To introduce extrinsic fluctuations η into the rate function, we multiply a rate constant, e.g. d_1, by a lognormally distributed stochastic variable $\nu = 1 + \eta$

$$\bar{d}_1(t) = d_1 \nu(t) = d_1 (1 + \eta(t)). \tag{52}$$

16

The lognormally distributed variable can be constructed from a normal stochastic variable μ with mean zero, variance ϵ^2, and correlation time $\tau = K^{-1}$ defined by the Ornstein-Uhlenbeck process

$$\mathrm{d}\mu(t) = -K\mu(t) + \sqrt{2K}\epsilon \, \mathrm{d}W(t). \tag{53}$$

The shifted lognormal stochastic variable η is then given by

$$\eta(t) = \exp\left(\mu(t) - \frac{1}{2}\epsilon^2\right) - 1, \tag{54}$$

and we define the magnitude of the extrinsic noise as the coefficient of variation

$$CV = \sqrt{\exp(\epsilon^2) - 1}, \tag{55}$$

such that the average of the lognormal variable $\nu(t) = 1$ and its standard deviation is equal to Eq. (55). To be able to write down the LNA in the way of Eq. (9), we specify the Jacobian matrix $\mathbf{A}$ and the diffusion matrix $\mathbf{BB}^T$. The deterministic rate equations for the macroscopic concentrations $\boldsymbol{\phi}$ of the molecular species (where $\phi_1 = D$, $\phi_2 = M$, and $\phi_3 = A$) are

$$\frac{\mathrm{d}\boldsymbol{\phi}}{\mathrm{d}t} = \begin{pmatrix} -k_1\phi_1 + k_0(1-\phi_1) \\ v_0\phi_1 - d_0\phi_2 \\ v_1\phi_2 - d_1(1+\eta)\phi_3 \end{pmatrix}. \tag{56}$$

The dynamics of D^* can be eliminated as the total promoter concentration is conserved $D + D^* = 1/\Omega$ (assuming there is only one copy of the promoter). One of the conditions for the LNA is that it is valid in the limit of large system size Ω. However, as the current system is linear (i.e. it contains no bimolecular reactions), the LNA will give the exact expressions for the mean and variance independent of the value of Ω [34]. For this reason, we have chosen $\Omega = 1$ here.

The Jacobian as defined in Eq. (7) is then given by

$$\mathbf{A} = \begin{pmatrix} -(k_0+k_1) & 0 & 0 \\ v_0 & -d_0 & 0 \\ 0 & v_1 & -d_1(1+\eta) \end{pmatrix}, \tag{57}$$

and the diffusion matrix as defined in Eq. (8) for the system is given by

$$\mathbf{BB}^T = \begin{pmatrix} k_1\phi_1^s + k_0(1-\phi_1^s) & 0 & 0 \\ 0 & v_0\phi_1^s + d_0\phi_2^s & 0 \\ 0 & 0 & v_1\phi_2^s + d_1(1+\eta)\phi_3^s \end{pmatrix}, \tag{58}$$

where ϕ_i^s are the steady-state concentrations of the active promoter, mRNA, and protein molecules, respectively. The procedure described in Eqs. (52)-(58) is the same for extrinsic

noise in any other system parameter. In summary, for the three-stage model of gene expression the SDE studied in Eq. (9) is defined by the steady state concentrations $\boldsymbol{\phi^s}$, the Jacobian as given in Eq. (57), the diffusion matrix as given in Eq. (58), and the extrinsic noise $\nu(t)$ in d_1 as defined by Eq. (52) and Eq. (53).

The reaction rates used to obtain the results in Figure 1 are shown in Table I and are representative for gene expression in mammalian cells as determined by Schwanhäusser *et al.* and Suter *et al.* [35, 36]. Schwanhäusser *et al.* experimentally determined transcription and translation rates for over $5\,000$ genes in mammalian cells [35]. We selected the mode of these parameter distributions as parameter values for our linear model of gene expression. To ensure relatively fast intrinsic timescales, we chose a protein degradation and mRNA degradation rate associated with a gene that was classified as having both unstable protein and mRNA (see Figure 5 in [35]), while also enforcing that the protein lifetime is much longer than the mRNA lifetime. Promoter activation and deactivation rates were chosen as the average over various mammalian genes as measured by Suter *et al.* [36].

In the three-stage gene expression model, the three intrinsic timescales are $1/(k_0 + k_1)$, $1/d_0$, and $1/d_1$, corresponding to the lifetimes of the molecular species D, M, and A respectively. The timescale of the extrinsic noise process is $\tau = K^{-1}$ as defined in Eq. (53). The ratio of the extrinsic and intrinsic timescales is then

$$\lambda = \tau / \max(\tau_{\text{int}}), \tag{59}$$

where $\max(\tau_{\text{int}})$ is the longest intrinsic timescale in the model, here $1/d_1 = 15\,625$ s. To see for which values of λ our method gives good results, we simulate the model in Figure 1 with the mRNA degradation rate d_0 subject to extrinsic noise with various values of τ using the Extrande algorithm. Figure 1(b) shows that for longer extrinsic correlation times τ (corresponding to larger values of λ), the stochastic simulations (denoted "SSA" in Figure 1) approach the analytically calculated mean number of protein A, given by $\langle \phi_3^s \rangle$ (see Eq. (36)). For both extrinsic noise magnitudes $CV = 0.1$ and $CV = 0.25$, timescale separation conditions are satisfied for $\lambda \approx 10$ and we are able to accurately predict the mean number of proteins. This corresponds to an extrinsic timescale of $\tau = 10^5$ s, which is of the same order of magnitude as the period of the mammalian cell cycle (approximately 27.5 h [35]).

Noise in different parameters can affect the system in different ways. Figure 1(c) shows that the mean protein number can decrease (k_0), increase (d_0, d_1, k_1), or remain constant
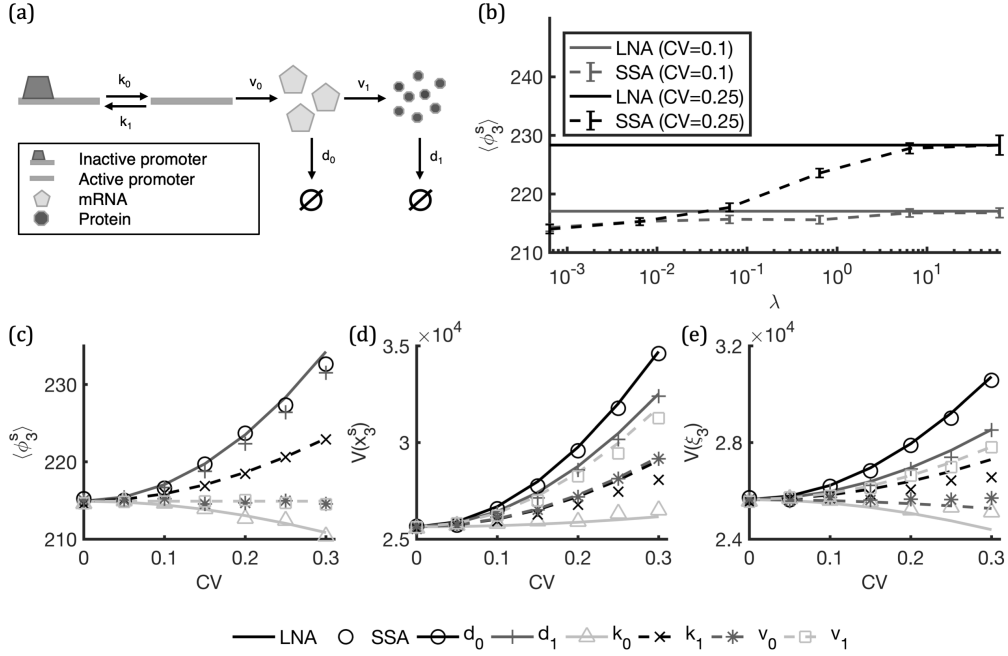
FIG. 1. (a) Gene expression model in which the promoter can switch between the active and inactive state. (b) Validity of the timescale separation condition depends on the ratio of extrinsic and intrinsic timescales λ and the magnitude of the extrinsic noise denoted by the CV (here in parameter d_0). The mean number of protein A at steady state, denoted by $\langle \phi_3^s \rangle$, with grey lines corresponding to extrinsic noise with $CV = 0.1$, and black lines corresponding to extrinsic noise with $CV = 0.25$. Solid lines are analytical predictions (denoted by "LNA"), dashed lines are stochastic simulation results (denoted by "SSA") from the Extrande algorithm [13]. 95% confidence intervals for the mean were calculated following Brančík *et al.* in [37]. (c) Effect of extrinsic noise in different parameters on the mean number of protein A. Analytical predictions are denoted by lines (black: d_0, dark grey: d_1, light grey: k_0, dashed black: k_1, dashed dark grey: v_0, and dashed light grey: v_1), stochastic simulation results by markers (circles: d_0, plus signs: d_1, triangle: k_0, crosses: k_1, asterisks: v_0, and squares: v_1). Note that the analytical predictions for extrinsic noise in parameters d_0 and d_1 overlap, as well for v_0 and v_1. We computed Eq. (36) up to sixth order in ϵ ($u \leq 3$). (d) Effect of extrinsic noise in different parameters on the total variance of protein A, given by $V(x_3^s)$ (see Eq. 12). We computed Eqs. (38), (37) up to second order in ϵ ($u \leq 1$). Line colours and marker styles correspond to extrinsic noise in the same parameters as in (c). (e) Effect of extrinsic noise in different parameters on the intrinsic variance of protein A, given by $V(\xi_3)$ (see Eq. 12). We computed Eq. (37) up to second order in ϵ ($u \leq 1$). We used the dual reporter technique [6] to compute estimates of the intrinsic variance from stochastic simulations. Line colours and marker styles correspond to extrinsic noise in the same parameters as in (c).

19

TABLE I. Parameter values used for the gene expression model.

parameter	value (s^{-1})
k_0	0.00085
k_1	0.0017
v_0	0.00028
v_1	0.028
d_0	0.00019
d_1	0.000064

(v_0, v_1) in response to increasing extrinsic noise magnitude (CV). The total variance in protein A, denoted by $V(x_3^s)$ (see Eq. (12)), always increases when a system is affected by extrinsic noise (Figure 1(d)). Extrinsic noise in a parameter that controls the lifetime of molecular components, such as the mRNA or protein degradation rate, might cause intrinsic and extrinsic timescales to mix and thus have non-trivial effects. This can be seen, for example, when extrinsic fluctuations are applied to the promoter activation rate k_0, causing a decrease in the intrinsic variance, denoted by $V(\xi_3)$ (Figure 1(e)). The potential of extrinsic fluctuations to decrease the intrinsic noise in gene expression was already noted by Shahrezaei *et al.* [11].

B. Regulated gene expression

Feedback control is often proposed as a possible strategy to reduce gene expression noise, and therefore it is considered to have great potential for the design of robust synthetic gene regulatory networks [16]. Several studies have confirmed that such strategies are indeed capable of reducing the variability in protein concentrations as well as influencing the number of modes of the protein number distribution [16, 38–43]. To study the potential of negative feedback to reduce protein noise in a system subject to extrinsic noise, we adapt the genetic feedback loop model proposed by Grima *et al.* [9]. In this model, the mRNA dynamics are omitted, which is a valid assumption in the absence of translational bursting [44, 45]. The genetic feedback model contains a negative feedback loop where the gene product can bind to the promoter, thus preventing protein production (Figure 2(a)). This system consists of
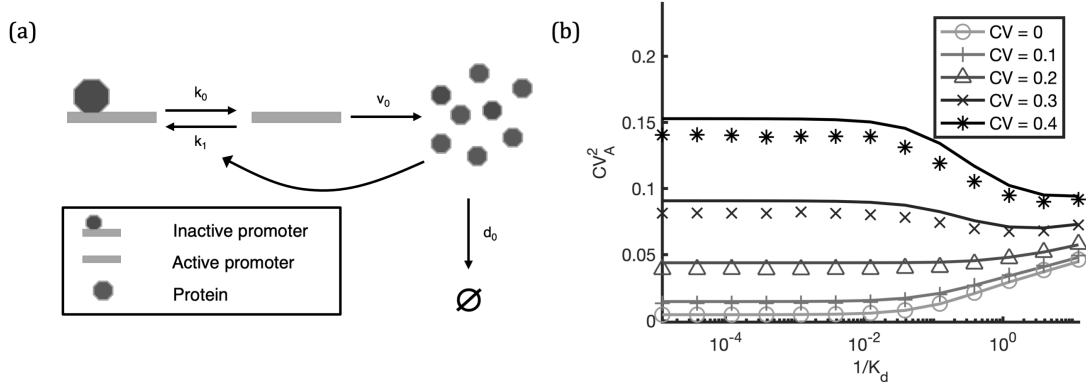
FIG. 2. (a) Model of autoregulatory gene expression. The feedback strength $1/K_d = k_1/(k_0\Omega)$ is determined by the binding affinity of the protein to the promoter. (b) Analytical predictions (solid lines) and stochastic simulation results (markers) for the CV_A^2 of the number of molecules of protein A as a function of feedback strength $1/K_d$ for extrinsic noise in protein production rate v_0 ranging in magnitude from no extrinsic noise ($CV = 0$, light grey lines and circles) to strong extrinsic noise ($CV = 0.4$, black lines and asterisks). We computed the mean up to sixth order in ϵ ($u \leq 3$ in Eq. (36)) and the variances up to second order in ϵ ($u \leq 1$ in Eqs. (38), (37)).

the following reactions

$$D \xrightarrow{v_0} D + A \underset{k_0}{\overset{k_1}{\rightleftharpoons}} D^*, \qquad A \xrightarrow{d_0} \emptyset. \tag{60}$$
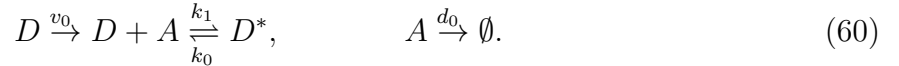
Here, D denotes the unbound promoter, D^* the bound, inactive promoter, and A the protein. The values for the parameters k_0, v_0, and d_0 (Table II) were chosen so that they are consistent with those in the linear gene expression model, as follows. Since the mRNA dynamics are omitted, d_0 now refers to the protein degradation rate. To ensure the same steady-state species concentrations as in the unregulated gene expression model, we choose the protein production rate as

$$v_0 = \phi_2^s v_1, \tag{61}$$

where ϕ_2^s refers to the macroscopic steady-state concentration of mRNA in the linear gene expression model, respectively, and v_1 is equal to the parameter value of the translation rate of the linear gene expression model as stated in Table I. Note that for a promoter which can be in two states, D and D^*, the average concentration of D (D^*) normalised by the total promoter concentration, is the same as the probability that the promoter is in state D (D^*). The probability that the binding reaction occurs in a small time interval is proportional

21

TABLE II. Parameter values used for the autoregulatory gene expression model.

parameter	value (s^{-1})
k_0	0.00085
k_1/Ω	$6.7 \times 10^{-11} - 6.7 \times 10^{-5}$
v_0	0.014
d_0	0.000064

to k_1/Ω, where k_1 is the protein-DNA binding rate constant and Ω is the nuclear volume (which occupies $\sim 8\%$ of the cellular volume of approximately 2 picolitres [46, 47]). We vary the binding rate k_1 of protein A to the promoter over a range of biologically relevant specificities [48] and define the feedback strength as the inverse of the non-dimensional dissociation constant K_d

$$\frac{1}{K_d} = \frac{k_1}{k_0\Omega}. \tag{62}$$

We note that for very small values of $1/K_d$, the system is weakly non-linear since protein binding becomes a rare event compared to promoter activation.

In a negative feedback system with no extrinsic noise, it is well known that the noise increases with $1/K_d$ [49–51]. This is confirmed in Figure 2(b), where the light grey lines and circles ($CV = 0$) show that the noise in the protein population, quantified by its coefficient of variation squared CV_A^2, increases as the negative feedback strength is amplified. We consider the case of extrinsic noise affecting the protein production rate v_0 and show in Figure 2(b) how this changes the noise in the number of proteins A. Slow fluctuations in the protein production rate could for example be the result of changes in the cell's mitochondrial or ribosomal content over time [52]. While extrinsic noise in the protein production rate v_0 does increase protein noise, as the extrinsic noise magnitude increases beyond $CV \sim 0.3$ a higher level of negative feedback decreases CV_A^2. This finding is supported by experimental data that shows that negative autoregulation mechanisms are able to negate the effects of slow extrinsic noise [19]. In the absence of extrinsic noise, the variance of protein numbers has only an intrinsic component. The magnitude of intrinsic noise is influenced the average expression level of molecular species and the response time of the system, which is the time it takes for any initial perturbation to decay and the system to return to its equilibrium.

In one instance, protein expression levels decrease monotonically with increasing negative feedback strength, and smaller molecular numbers are associated with higher intrinsic noise. In the second, negative feedback is known to speed up the response time of a system, leading to the attenuation of protein noise [53]. These effects can be seen in Figure 2(b) for $CV = 0$, $CV = 0.1$, and $CV = 0.2$, showing that smaller values of $1/K_d$ result in a less noisy system given the system is subject to no or low extrinsic noise. This implies that the decrease of intrinsic noise due to a faster response time cannot compensate for the increase in noise resulting from smaller protein levels. The extrinsic contribution to the protein noise is also a function of the response time of the system [51], which causes the extrinsic component of the protein noise to decrease monotonically with increasing feedback strength. If the CV of the extrinsic noise source is increased above 0.3 (Figure 2(b), dark grey line and crosses, and black line and asterisks), the reduction in extrinsic noise is larger than the increase in intrinsic noise up until $1/K_d \approx 10^{-1}$. With increasing negative feedback strength, the reduction of extrinsic noise is finally negated by an increase in intrinsic noise due to decreasing protein levels, and CV_A^2 will increase again.

C. Signal transduction and extrinsic noise

Cells are embedded in highly fluctuating environments. It is vital for biological systems that they can sense external stimuli and process this information in order to adapt to their environment accordingly. Information theoretic approaches have, for example, been applied to biological systems to address the question of how well a network subject to biochemical noise is able to transmit information that arrives at cell receptors into the intracellular environment.

In order to analyse the effects of extrinsic noise on the signal transduction process, we consider the simple two-stage gene expression model shown in Figure 3(a), which contains mRNA molecules M and protein molecules A, the concentrations of which fluctuate over time

$$\emptyset \xrightarrow{v_0} M \xrightarrow{d_0} \emptyset, \qquad M \xrightarrow{v_1} M + A, \qquad A \xrightarrow{d_1} \emptyset. \tag{63}$$

We again choose parameters for this motif such that they are consistent with those in the linear gene expression model. The values for v_1, d_0 and d_1 are the same as in Table I, while

the mRNA production rate v_0 is calculated as

$$v_0 = \phi_1^s v_0^*, \tag{64}$$

where ϕ_1^s corresponds to the macroscopic steady-state concentration of the active promoter in the linear gene expression model and the asterisk refers to parameters as stated in Table I.

We are interested in how the rate of information transfer from mRNA to protein is affected by an extrinsic noise source that perturbs the translation process (extrinsic fluctuations in v_1). Fluctuations in the rate of translation arise because this process is dependent on the number of free ribosomes in the cytoplasm, which changes stochastically over time. To quantify how well networks can transmit information in noisy environments, we can make use of the mutual information rate (MIR) as a metric. Calculation of the mutual information of trajectories is a challenging task [54, 55], but for linear Gaussian processes an expression involving the power spectra of continuous-time input signal $s(t)$ and output signal $x(t)$ can be derived, where $s(t)$ and $x(t)$ are the deviations of the copy number of species S and X from their respective steady-state values:

$$\text{MIR}(s(t), x(t)) = -\frac{1}{4\pi} \int_{-\infty}^{\infty} \ln\left[1 - \frac{|P_{sx}(\omega)|^2}{P_{ss}(\omega)P_{xx}(\omega)}\right] d\omega. \tag{65}$$

Here, $P_{sx}(\omega)$ is the cross-power spectrum of $s(t)$ and $x(t)$ (the Fourier transform of the cross-correlation function of $s(t)$ and $x(t)$), and $P_{ss}(\omega)$ and $P_{xx}(\omega)$ are the power spectra of $s(t)$ and $x(t)$, respectively. Note that for non-Gaussian and/or non-linear systems this expression provides a lower bound for the channel capacity [56]. The protein content of a cell contains information about the mRNA concentration of the cell at an earlier time point (since more mRNA will lead to a higher amount of protein). We consider the case where the mRNA concentration in a cell is translated to a protein concentration, but this process is disrupted by an external process affecting the translation rate v_1. The MIR then quantifies how much information about the input signal (i.e. the deviations from the mRNA steady state) can still be found back in the output signal (i.e. the deviations of the protein concentration from its steady state value). In [54], Tostevin *et al.* have derived the analytical expression for the MIR of the input and output trajectories of the motif in Figure 3(a) in the absence of extrinsic noise. We extend this result by considering an external process that affects the translation process, causing the parameter v_1 to fluctuate over time. By substituting the analytical expressions for the (cross-)power spectra resulting from Eq. (23) into Eq. (65), we
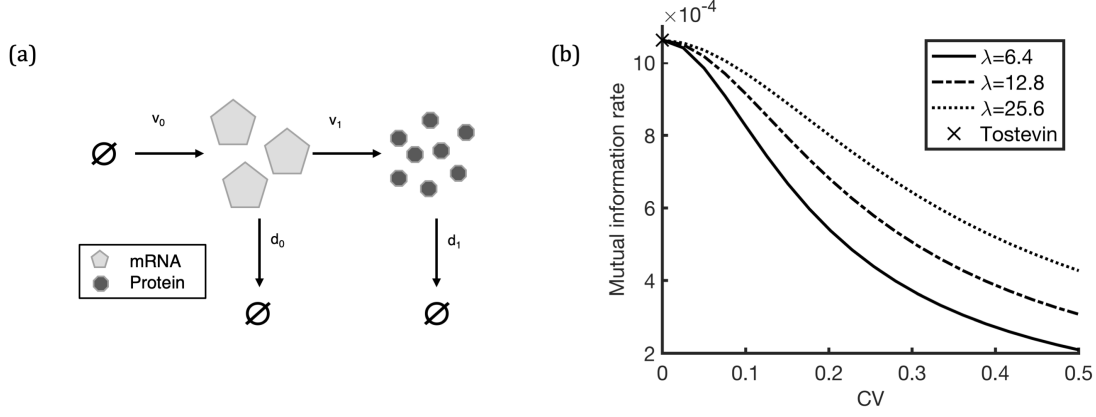
FIG. 3. (a) Two-stage model of gene expression, in which information about the concentration of mRNA (input) is transduced to the concentration of protein (output). (b) Mutual information rate between mRNA (input) and protein (output) variables as a function of extrinsic noise magnitude CV for extrinsic noise in parameter v_1 with different correlation times, corresponding to a ratio of timescale separation λ between intrinsic and extrinsic processes of 6.4 (solid line), 12.8 (dash-dot line), and 25.6 (dotted line). The black cross denotes the result by Tostevin *et al.* (motif III) [54] in the absence of extrinsic noise. In the expression for the MIR, Eq. (65), the analytical solution of the power spectrum was calculated up to second order in ϵ ($u \leq 1$ in Eqs. (42), (48)).

TABLE III. Parameter values used for the information processing motif.

parameter	value (s^{-1})
v_0	0.000093
v_1	0.028
d_0	0.00019
d_1	0.000064

are able to quantify how the accuracy of information transmission from mRNA to protein concentration is affected by this disrupting process.

Figure 3(b) shows that in the absence of extrinsic noise, our approximate solution agrees with the solution by Tostevin *et al.* (marked by the black cross) [54]. The presence of extrinsic noise decreases the fidelity of information processing, and the extent of this effect depends on both the extrinsic noise magnitude CV and the ratio of timescale separation λ, as defined in Eq. (59), between the intrinsic and extrinsic processes. Extrinsic noise sources

with long extrinsic correlation times (large λ) are less disruptive than those with shorter correlation times (smaller λ). This implies that slowly fluctuating environmental variables have a smaller negative effect on the MIR than external variables that fluctuate quickly. This result is intuitive, as in the extreme case where faster fluctuating extrinsic processes happen on roughly the same timescale as the intrinsic processes, it may prove harder to distinguish between signal and noise and information might be lost. In the limit of infinitely slowly fluctuating external variables, the value of the external variable is constant with respect to the intrinsic timescale, and will have no effect on the signal transduction process regardless of the noise magnitude.

D. Robustness of feed-forward loop motifs

Feed-forward loops (FFLs) are capable of responding in a precise, robust manner to external signals [57]. These motifs are defined by a gene X that regulates a second gene Y. Both X and Y then regulate a target gene Z and are very common, appearing in many organisms, including *E. coli*, *B. subtilis*, *S. cerevisiae*, *D. melanogaster*, *C. elegans*, and humans [58]. There are multiple types of FFLs (see Figure 4(a)) since regulation can take place either through activation or repression, and much effort has been devoted to extract the general features of each one. However, this is not straightforward because both the transient and equilibrium behaviour is characteristic of a particular system [59]. For these reasons, constructing an optimal system that fulfils certain design requirements can be a considerable computational task. Here, we aim to present an efficient optimisation scheme to generate optimal parameters for a FFL to ensure it responds in a precise manner to input signals but remains robust to noise. To do this, we devise two objective functions that quantify both the dynamic as well as the stochastic behaviour of the system. We aim to generate parameter sets for the network such that the system responds to a switching-on of the input signal X by a negative pulse in the concentration of Z, before going back to its original steady state (black line, Figure 4(b)). Moreover, the variation around this steady state concentration of Z for a noisy input signal X should be minimal. Taking these requirements into account, the form of the objective function is

$$S = w_1 c_{ODE} + w_2 c_{LNA}, \qquad (66)$$

26

where w_1 and w_2 are the weights of the respective objective functions c_{ODE} and c_{LNA}, with $w_1 + w_2 = 1$ and

$$c_{ODE} = \sum_{i=1}^{5} q_i s_i, \tag{67}$$

$$c_{LNA} = CV_Z^2 = \frac{V(x_3^s)}{\langle x_3^s \rangle^2}, \tag{68}$$

where x_3^s is the steady-state concentration of protein Z and

$$s_1 = (\phi_1(t_f) - \phi_1(t_f - \Delta t))^2, \tag{69}$$

$$s_2 = (\phi_2(t_f) - \phi_2(t_f - \Delta t))^2, \tag{70}$$

$$s_3 = (\phi_3(t_f) - \phi_3(t_f - \Delta t))^2, \tag{71}$$

$$s_4 = (\phi_3(t_f) - \phi_3(t_0))^2, \tag{72}$$

$$s_5 = \frac{\min(\phi_3)}{\langle \phi_3 \rangle}, \tag{73}$$

with ϕ_i the macroscopic concentrations of the proteins X, Y, and Z respectively, $q_i = \frac{1}{5}$, $i = 1, \ldots, 5$ the subweights of each ODE (ordinary differential equation) objective s_i, $t_0 = 0$ the initial time point, $t_f = 1000$ the final time point of the simulation, and $\Delta t = 5$. To obtain c_{ODE} we simulate the FFLs using an ODE solver, where Eqs. (69)–(71) ensure each of the three system components reaches steady state, Eq. (72) ensures that Z reaches pre-input concentration, and Eq. (73) aims to produce a significant drop in the concentration of Z upon a change in input X (Figure 4(b)). The score c_{LNA} is obtained from our analytical solution (Eq. (11) and Eq. (12)).

We used the most general model of the FFL from Macía *et al.* in [59], that is able to describe all eight FFL topologies. The macroscopic rate equations for this model are given by

$$\frac{d\boldsymbol{\phi}}{dt} = \begin{pmatrix} \alpha_0(1+\eta) - d_0\phi_1 \\ \alpha_1\left(\frac{1+\beta_0 K_1\phi_1}{1+K_1\phi_1}\right) - d_1\phi_2 \\ \alpha_2\left(\frac{1+\beta_1 K_2\phi_1+\beta_2 K_3\phi_2+\beta_3 K_2 K_3\phi_1\phi_2}{1+K_2\phi_1+K_3\phi_2+K_2 K_3\phi_1\phi_2}\right) - d_2\phi_3 \end{pmatrix}. \tag{74}$$

In this model, α_i describes the basal production of the proteins X, Y, and Z, and d_i denotes the degradation rate of a species. The type of regulatory interaction between the regulator gene and the gene it targets is described by parameters β_i, where values $\beta_i < 1$ correspond

27

to an inhibitory interaction as the production rate decreases proportionally to the basal level, whereas $\beta_i > 1$ describes activation [59]. K_i describes the binding equilibrium of the regulator with the gene it targets. Extrinsic noise enters the model in the production rate of X, α_0, with a magnitude of $CV = 0.5$. We fix the ratio of basal production/degradation ($\alpha_i/d_i = 100$) of all three molecular species to ensure that all species are sufficiently abundant for the LNA to hold. Initial conditions are $X(t_0) = 0$, $Y(t_0) = Z(t_0) = 100$. We generate 2×10^6 random parameter sets for the 10 remaining system parameters and optimise for the objective score S for the cases $\{w_1 = 1, w_2 = 0\}$ and $\{w_1 = 0.5, w_2 = 0.5\}$ (for details on the optimisation procedure, see Section II E). The 0.01% top scoring parameter sets are then selected for either set of objective scores, and their corresponding topology is determined. The results in Table IV show that only three of eight possible topologies, shown in the red boxes in Figure 4(a), are present in the results when only the ODE objective function is considered ($w_1 = 1$), and that cFFL4 is the most prevalent motif. If both the ODE and LNA objectives are given equal importance ($w_1 = w_2 = 0.5$), then the cFFL3 and iFFL1 topologies are also present among the top-scoring parameter sets (solid line boxes, Figure 4(a)). However, since these do not occur when only considering the ODE criteria, this implies that they are not very suitable to give the desired dynamic behaviour. In addition, Macía et al. also find that the iFFL1 motif is not capable of producing the desired negative pulse [59].

Since the cFFL4 topology is the most prevalent, this motif has the highest probability to produce the desired system behaviour. For this reason, we perform an optimisation within the local parameter space of the cFFL4 motif for 2×10^4 randomly generated parameter sets and select the $1\,500$ best-scoring sets. Figure 4(c) shows how the parameter space changes with the addition of the c_{LNA} objective. If only the robustness of the network to noise is considered ($w_1 = 0$, light grey circles), no specific parameter values for the parameters K_1 and d_2 are preferred. When only the dynamic behaviour of the FFL is prioritised ($w_1 = 1$, medium grey squares), high values of K_1 and d_2 more likely result in the desired dynamics. Interestingly, when both the c_{ODE} and c_{LNA} objectives are given equal weight ($w_1 = 0.5$, black downward-pointing triangles) the parameter space is further constrained compared to optimising for a single objective. Thus, for robust FFLs the degradation rate of Z and the binding equilibrium of protein X to the gene that produces Y needs to be tuned.
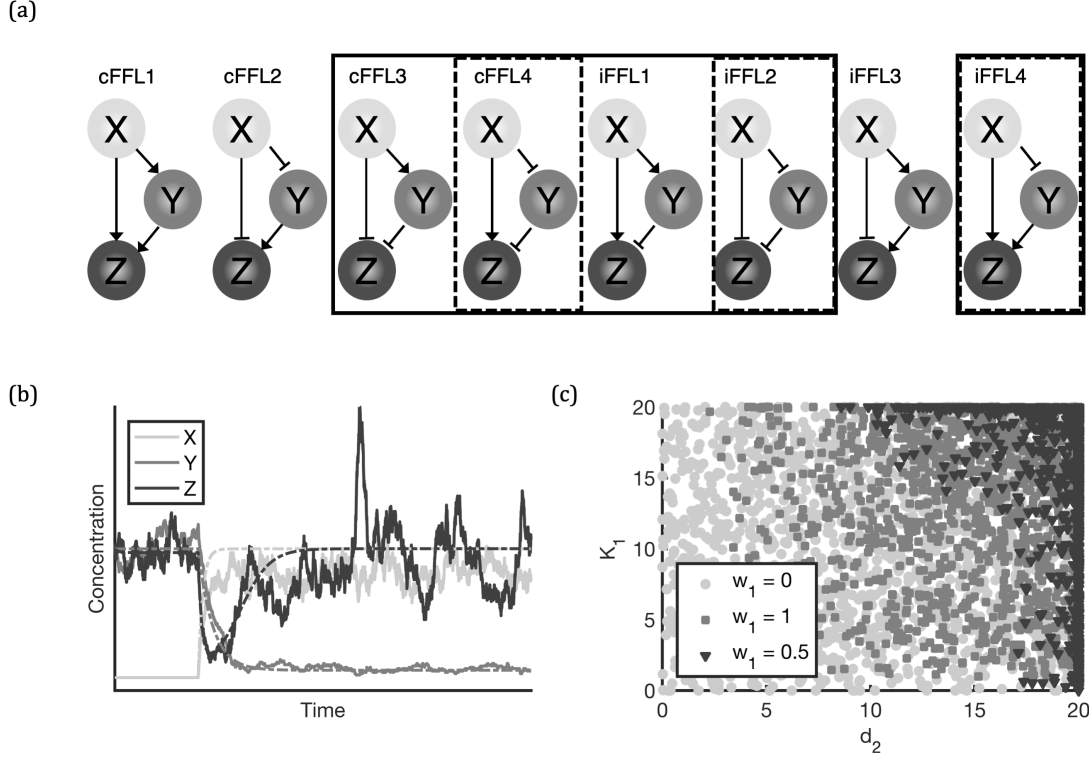
28

FIG. 4. (a) All eight feed-forward loop topologies, where the arrow indicates the type of regulation (activation or repression). The motifs enclosed by a dashed line denote the subset of optimal FFL topologies for the case of $\{w_1 = 1, w_2 = 0\}$ in Eq. (66), whereas the motifs enclosed by a line denote the subset of optimal topologies when $\{w_1 = 0.5, w_2 = 0.5\}$ (See Table IV). (b) Sample result of the FFL dynamic behaviour of our optimisation scheme. After the production of protein X is induced, this produces a drop in the concentration of protein Z. The concentration of Z then recovers to its original steady-state level. Solid lines denote example stochastic simulations of the concentration of proteins X (light grey), Y (medium grey), and Z (black) over time, while dashed lines correspond to the respective deterministic dynamics. (c) How the parameter space of parameters K_1 and d_2 is constrained by considering an additional objective function for the system's stochastic behaviour. $w_1 = 0$ (light grey circles) refers to the case where only the dynamic behaviour is prioritised, $w_1 = 1$ (medium grey squares) refers to the case where only the stochastic behaviour is prioritised, and $w_1 = 0.5$ (black downward-pointing triangles) refers to the case where both objectives are given equal importance.

TABLE IV. Occurrence of FFL topologies of top (0.01%) scoring parameter sets out of 2×10^6 parameter sets.

topology	$w_1 = 1$	$w_1 = 0.5$
iFFL2	1%	2%
cFFL4	75%	39.5%
iFFL4	24%	1%
cFFL3	0%	14.5%
iFFL1	0%	43%

IV. DISCUSSION & CONCLUSION

In this work, we have derived an analytical framework to quantify the contribution of coloured extrinsic noise to fluctuations in gene expression. We have shown that when the conditions underlying the theory are satisfied, we are able to accurately describe the mean, variance and power spectrum of molecule number fluctuations subject to both intrinsic and extrinsic noise sources. Using several examples, we show that the theory is relevant in a wide range of applications, and can be used to distil the principles underlying fundamental system behaviour, noise sources, and information processing in biochemical networks.

Our framework relies on three main approximations: the linear-noise approximation, the separation between timescales of the intrinsic and extrinsic fluctuations, and the small extrinsic noise expansion. First, the LNA will give an accurate approximation of the CME when the molecular species populations are sufficiently large, when the nonlinearity in the reaction rates is sufficiently weak or else for special classes of biochemical systems ([34]; see discussion later). Specifically, the LNA is valid whenever the next term in the system-size expansion gives a very small correction to the moments of the LNA [60]. Second, our theory requires the extrinsic fluctuations to be slow with respect to the system's intrinsic dynamics. As most intrinsic processes happen on the timescale of seconds or sub-seconds and extrinsic fluctuations have a typical correlation time corresponding to the cell cycle period (many minutes), we expect that our timescale separation assumption is reasonable. The assumption of small extrinsic noise might appear at the first sight very limiting, however, in practice we find that the approximation yields sufficiently accurate results for noise magnitudes

at least as large as $CV = 0.3$ for the mean protein number (Figure 1(b)) and variances (Figure 1(d)) of the linear model and $CV = 0.2$ for the means and variances of the regulated gene expression model (Figure 2(b)).

A more general issue with studying stochastic systems is computational speed. To obtain statistics of biochemical systems subject to both intrinsic and extrinsic noise with reasonable confidence levels, one needs to simulate many trajectories of the system for a considerable time. The advantage of obtaining closed-form expressions for these statistics is that parameter values simply need to be substituted, and there is no need to re-evaluate the system. For example, it takes approximately 3 hours to generate one data point (100 trajectories of 10^8 seconds each) for the model in Figure 1(a), whereas evaluating the analytical expressions for the mean and variance of each molecular component takes less than a second on a typical desktop computer. As pointed out in Section II C 4, the limiting step in the automated sum evaluation is the calculation of the stationary state ϕ^s. This task becomes increasingly computationally intensive for more complex systems. For this reason, the limit of complexity that we can study is determined by the computational power available. For specific cases, a more specialised external solver could be employed to accelerate this task.

Due to this speed-up, we have been able to use the analytical framework to perform a computationally efficient multi-objective optimisation of FFLs. With this optimisation routine, we are able to explore both network topology and parameter space to generate systems with optimal dynamic and stochastic features, which is generally infeasible for non-trivial systems using simulation-based approaches. Our analysis shows that even in simple networks such as FFLs, there exists a complex relationship between system structure and function. With this optimisation scheme we are able to quickly generate recommendations for an optimal network topology and parameter ranges. Compared to optimisation using stochastic simulation algorithms, this optimisation scheme gives an improvement in computational time of several orders of magnitude. The results of the optimisation scheme suggest that not every FFL motif is capable of producing a specific dynamic response, and that not all FFL types have the same extrinsic noise tolerance. Although there does not appear to be a trade-off between these two objectives, choosing optimal networks and their parameters such that they fulfil both requirements can be a substantial task given the high dimensionality of the problem. In this case, combining deterministic dynamics with stochastic analysis of equilibrium behaviour is an efficient and effective approach.

31

Since analytical expressions for both intrinsic and extrinsic contributions to variability can be obtained, the proposed method allows a systematic analysis of how changing the properties of extrinsic fluctuations affects intrinsic variability and total noise in gene expression models. Similarly, these expressions can provide predictions on which network parameters are susceptible to perturbations and contribute to high variability and can be used as a tool for stochastic sensitivity analysis. Such a method can be of interest for synthetic biology applications as it could provide universal design principles for network construction that exploits (suppresses) the positive (negative) effects of cellular stochasticity.

The framework developed here rests on the validity of the linear-noise approximation first and foremostly. This limits the current approach to analysis of nonlinear biochemical systems with large numbers of molecules in all species or else to those systems with arbitrary number of molecules but weakly nonlinear reaction rates. However, we note that the linear-noise approximation has been, over the past decade, extended to estimate the first and second moments of the molecule number distributions of nonlinear biochemical systems in which one or more molecular species is present in low copy numbers [60–63]. The corrections to the LNA power spectrum of fluctuations due to low molecule numbers have also been systematically studied [64, 65]. Hence by starting from these frameworks and repeating the same analysis as we performed here, i.e., applying the assumption of timescale separation between intrinsic and extrinsic noise and subsequently assuming small extrinsic noise, would likely result in a new theory which overcomes the major limitations of the present approach.

In conclusion, we have proposed a fast, systematic analytical framework to assess the effects of coloured environmental noise on biochemical systems. We have shown that the mathematical framework provides accurate predictions of system characteristics for a wide range of biological networks. Given the speed and flexibility of our approach, the research community can now further access the sources of variability in gene expression data. This will lead to a better understanding of how biological systems exploit or suppress environmental signals. There is, thus, the potential to uncover new design principles to aid the construction of new, robust *in vivo* synthetic circuits.

ACKNOWLEDGMENTS

Appendix A: Construction of the noise variables

To introduce extrinsic fluctuations to a rate constant c_k, we multiply it with a lognormal variable $\bar{\nu}_k(t)$,

$$\bar{c}_k(t) = c_k \bar{\nu}_k(t) = c_k \left(1 + \bar{\eta}_k(t)\right). \tag{A1}$$

We require $\langle \bar{c}_k(t) \rangle = c_k$ and also define a shifted stochastic variable $\bar{\eta}_k(t)$ that will be needed in the small noise expansion. The lognormal variables $\bar{\nu}_k(t)$ can be constructed from normally distributed variables $\bar{\mu}_k(t)$ with variances $\langle \bar{\mu}_k \bar{\mu}_k \rangle = \epsilon_k^2$ and inverse correlation times K_k. The latter may be described by the Ornstein-Uhlenbeck process [25]

$$\mathrm{d}\bar{\mu}_k(t) = -K_k \bar{\mu}_k(t) \, \mathrm{d}t + \sqrt{2K_k} \epsilon_k \, \mathrm{d}W(t). \tag{A2}$$

We define the lognormally distributed stochastic variable by

$$\bar{\nu}_k(t) = \exp\left(\bar{\mu}_k(t) - \frac{1}{2}\epsilon_k^2\right) \tag{A3}$$

and use Wick's theorem [66] to calculate its mean (only even powers in $\bar{\mu}_k^c$ do not vanish):

$$\langle \bar{\nu}_k \rangle = \exp\left(-\frac{1}{2}\epsilon_k^2\right) \sum_{c=0}^{\infty} \frac{1}{(2c)!} \langle \bar{\mu}_k^{2c} \rangle = \exp\left(-\frac{1}{2}\epsilon_k^2\right) \sum_{c=0}^{\infty} \frac{1}{2^c c!} \underbrace{\langle \bar{\mu}_k \bar{\mu}_k \rangle^c}_{\epsilon_k^2} = 1 \tag{A4}$$

in agreement with the requirement $\langle \bar{c}_k(t) \rangle = c_k$. The shifted lognormal stochastic variable is then

$$\bar{\eta}_k(t) = \exp\left(\bar{\mu}_k(t) - \frac{1}{2}\epsilon_k^2\right) - 1. \tag{A5}$$

Appendix B: Arbitrary time correlation functions of lognormal stochastic variables

Given a set of independent normally distributed stochastic variables $\{\bar{\mu}_k\}$, a tuple of indices $(r_1, \ldots, r_n)$ and a tuple of times $(t_1, \ldots, t_n)$ we define $\mu_k \equiv \bar{\mu}_{r_k}(t_k)$, and the two-point time correlation functions Δ_{ij} for $1 \le i < j \le n$,

$$\Delta_{ij} \equiv \langle \mu_i \mu_j \rangle \tag{B1}$$

for all $i < j$. We consider a smooth function and its mth derivatives

$$y_k \equiv V(\mu_k) = \sum_{c=0}^{\infty} \frac{a_c}{c!} \mu_k^c, \qquad y_k^{(m)}(\mu_k) \equiv \frac{\mathrm{d}^m V(\mu_k)}{\mathrm{d}\mu_k^m} = \sum_{c=m}^{\infty} \frac{1}{(c-m)!} a_c \mu_k^{c-m} \tag{B2}$$

to derive the generalisation of the $n = 2$ result by Malakhov [26] for the n-point time correlation function

$$\langle y_1 \ldots y_n \rangle = \sum_{c_1=0}^{\infty} \cdots \sum_{c_n=0}^{\infty} \frac{a_{c_1} \ldots a_{c_n}}{c_1! \ldots c_n!} \langle \mu_1^{c_1} \ldots \mu_n^{c_n} \rangle . \tag{B3}$$

According to Wick's theorem [66], the correlation functions $\langle \mu_1^{c_1} \ldots \mu_n^{c_n} \rangle$ decompose into sums of partitions into two-point correlation functions and they are zero for odd n. We apply the theorem partially to isolate the two-point correlation functions $\Delta_{ij} = \langle \mu_i \mu_j \rangle$ with $i < j$,

$$\langle y_1 \ldots y_n \rangle = \sum_{l_1=0}^{\infty} \cdots \sum_{l_n=0}^{\infty} \sum_{u=0}^{\infty} \sum_{|\mathbf{d}^n|=u} \left(\frac{c_1! \ldots c_n!}{l_1! \ldots l_n! \prod_{i<j} d_{ij}!} \frac{a_{c_1} \ldots a_{c_n}}{c_1! \ldots c_n!} \langle \mu_1^{l_1} \rangle \ldots \langle \mu_n^{l_n} \rangle \prod_{\substack{i,j=1 \\ i<j}}^{n} \Delta_{ij}^{d_{ij}} \right) \tag{B4}$$

where the inner sum is taken over all tuples $\mathbf{d}^n = (d_{12}, d_{13}, d_{23}, \ldots, d_{(n-1)n})$ with $|\mathbf{d}^n| = \sum d_{ij} = u$. There are $c_k!/l_k!$ possibilities to assign $m_k = c_k - l_k$ from a total of c_k factors μ_k to the u Δ_{ij} pairs. However, to obtain the number of *different* partitions into the pairs, the product of the former must be divided by the product of $d_{ij}!$ permutations of d_{ij} identical factors Δ_{ij}. We notice that $l_k = c_k - m_k$ to recognise the derivatives $y_k^{(m_k)}$ from Eq. (B2) so finally

$$\langle y_1 \ldots y_n \rangle = \sum_{u=0}^{\infty} \sum_{|\mathbf{d}^n|=u} \left(\prod_{k=1}^{n} \langle y_k^{(m_k)}(\mu_k) \rangle \prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{\Delta_{ij}^{d_{ij}}}{d_{ij}!} \right) . \tag{B5}$$

The orders m_k of the derivatives are given by the number of indices k that appear in the product of two-point correlation functions, that is $m_k = \sum_{i=1}^{k-1} d_{ik} + \sum_{j=k+1}^{n} d_{kj}$.

For $n = 2$ with $k = m_1 = m_2 = d_{12}$ and $B_\mu[\tau] = \Delta_{12}(\tau) = \langle \mu(0)\mu(\tau) \rangle$ we recover the result by Malakhov [26],

$$B_y[\tau] = \langle y(0)y(\tau) \rangle - \langle y \rangle^2 = \sum_{k=1}^{\infty} \frac{1}{k!} \langle y^{(k)}(\mu) \rangle^2 B_\mu^k[\tau]. \tag{B6}$$

a. Evaluation for normal stochastic variables With normally distributed $y_k = \mu_k$ (mean 0) the term $\langle y_k^{(m_k)}(\mu_k) \rangle$ is 1 for $m_k = 1$ and 0 else. Consequently, $m_k = 1$ for $k \in \{1, \ldots, n\}$, that is each index must occur exactly once in the product of two-point correlation functions in Eq. (B5) so also all d_{ij} are 1 and Wick's theorem is recovered.

b.   Evaluation for lognormal stochastic variables   Lognormally distributed $y_k = \nu_k$ with mean 1 is invariant under differentiation with respect to μ_k (see Eq. A3) so the term $\langle y_k^{(m_k)}(\mu_k) \rangle$ becomes identically 1 for all m_k. This leads to significant simplification of Eq. (B5) and we use the multinomial theorem to obtain

$$\langle \nu_1 \dots \nu_n \rangle = \sum_{u=0}^{\infty} \sum_{|\mathbf{d}^n|=u} \left(\prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{\Delta_{ij}^{d_{ij}}}{d_{ij}!} \right) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{\substack{i,j=1 \\ i<j}}^{n} \Delta_{ij} \right)^k = \exp\left(\sum_{\substack{i,j=1 \\ i<j}}^{n} \Delta_{ij} \right). \qquad \text{(B7)}$$

c.   Evaluation for shifted lognormal stochastic variables   The mean of the stochastic variables η_k in Eq. (A5) is 0 and all derivatives with respect to μ_k are identical to ν_k in Eq. (A3) with mean 1. Therefore, the product of $\langle y_k^{(m_k)}(\mu_k) \rangle$ terms in Eq. (B5) vanishes if $m_k = 0$ for any $k \in \{1, \dots, n\}$ and is 1 otherwise. The final result is

$$\langle \eta_1 \dots \eta_n \rangle = \sum_{u=0}^{\infty} \sideset{}{'}\sum_{|\mathbf{d}^n|=u} \left(\prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{\Delta_{ij}^{d_{ij}}}{d_{ij}!} \right) \qquad \text{(B8)}$$

where the prime denotes the condition that for each $j \in \{1, \dots, n\}$ there is a $i < j$ or $k > j$ such that $d_{ij} \neq 0$ or $d_{jk} \neq 0$ (consequently $m_j \neq 0$). For example, with $n = 2$ we obtain $\langle \eta_1 \eta_2 \rangle = \exp(\Delta_{12}) - 1$. Evaluation for $r_1 = r_2 = k$ at equal times gives the variance $\langle \bar{\eta}_k^2 \rangle = \exp(\epsilon_k^2) - 1$ for $\bar{\eta}_k$ in Eq. (A5) where the variance of μ_k is $\epsilon_k^2 = \langle \mu_k \mu_k \rangle$.

Appendix C: Small noise expansion for the mean

The calculation of the mean concentrations $\langle \mathbf{x}^s \rangle$ in Eq. (33) is mediated by the multi-index notation defined by Eq. (34) so the correlation functions in the small noise expansion read $\langle \eta_{r_1} \dots \eta_{r_n} \rangle$ and can be evaluated by means of Eq. (32):

$$\langle \mathbf{x}^s \rangle = \sum_{n=0}^{\infty} \sum_{\#\mathbf{r}=n} \phi_{(\mathbf{r})}^s \sum_{u=\lfloor \frac{n+1}{2} \rfloor}^{\infty} \sideset{}{'}\sum_{|\mathbf{d}^n|=u} \prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{1}{d_{ij}!} \left(\Gamma_{r_i r_j} \right)^{d_{ij}}. \qquad \text{(C1)}$$

Γ_{ij} denotes the covariances $\langle \mu_i(t) \mu_j(t) \rangle$ of the normal stochastic variables. We rearrange the order of summation with u as the principal summation index, then n runs from 0 to $2u$, and obtain the final result in Eq. (36).

Appendix D: Small noise expansion for the covariance matrix

The small noise expansion of $\mathbf{V}(\boldsymbol{\xi})$ in analogy to $\langle \mathbf{x}^s \rangle$ is detailed in the main text, see Eq. (37). For $\mathbf{V}(\boldsymbol{\phi}^s)$ we use the multi-index $\mathbf{q} = (q_1, q_2)$, $|\mathbf{q}| = q_1 + q_2$, and the Taylor series of $\boldsymbol{\phi}^s$ in Eq. (33) to expand Eq. (13):

$$\mathbf{V}(\boldsymbol{\phi}^s) = \sum_{n=0}^{\infty} \sum_{|\mathbf{q}|=n} \sum_{\#\mathbf{r}^1=q_1} \sum_{\#\mathbf{r}^2=q_2} \boldsymbol{\phi}^s_{(\mathbf{r}^1)} \boldsymbol{\phi}^{s\,T}_{(\mathbf{r}^2)} \left(\left\langle \eta_{(\mathbf{r}^1)} \eta_{(\mathbf{r}^2)} \right\rangle - \left\langle \eta_{(\mathbf{r}^1)} \right\rangle \left\langle \eta_{(\mathbf{r}^2)} \right\rangle \right). \tag{D1}$$

To evaluate the correlation functions we use Eq. (32) and the index functions f_{ij}^2 from Eq. (39) to obtain

$$\left\langle \eta_{(\mathbf{r}^1)} \eta_{(\mathbf{r}^2)} \right\rangle = \sum_{u=\lfloor \frac{n+1}{2} \rfloor}^{\infty} {\sum_{|\mathbf{d}^n|=u}}' \prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{1}{d_{ij}!} \left(\Gamma_{f_i^2 f_j^2} \right)^{d_{ij}}, \tag{D2}$$

$$\left\langle \eta_{(\mathbf{r}^1)} \right\rangle \left\langle \eta_{(\mathbf{r}^2)} \right\rangle = \sum_{v=\lfloor \frac{n_1+1}{2} \rfloor}^{\infty} {\sum_{|\mathbf{d}^{n_1}|=v}}' \prod_{\substack{i,j=1 \\ i<j}}^{n_1} \frac{1}{d_{ij}!} \left(\Gamma_{r_i^1 r_j^1} \right)^{d_{ij}} \times \sum_{w=\lfloor \frac{n_2+1}{2} \rfloor}^{\infty} {\sum_{|\mathbf{d}^{n_2}|=w}}' \prod_{\substack{i,j=1 \\ i<j}}^{n_2} \frac{1}{d_{ij}!} \left(\Gamma_{r_i^2 r_j^2} \right)^{d_{ij}}. \tag{D3}$$

Each term in Eq. (D3) appears as well in Eq. (D2) when $u = v + w$. The other way round, every term in (D2) that contains only two-point correlation functions that occur in one of the two factors in Eq. (D3) cancels in the difference

$$\left\langle \eta_{(\mathbf{r}^1)} \eta_{(\mathbf{r}^2)} \right\rangle - \left\langle \eta_{(\mathbf{r}^1)} \right\rangle \left\langle \eta_{(\mathbf{r}^2)} \right\rangle = \sum_{u=\lfloor \frac{n+1}{2} \rfloor}^{\infty} {\sum_{|\mathbf{d}^n|=u}}'' \prod_{\substack{i,j=1 \\ i<j}}^{n} \frac{1}{d_{ij}!} \left(\Gamma_{f_i^2 f_j^2} \right)^{d_{ij}}. \tag{D4}$$

The new restriction indicated by the second prime in the sum, there is at least one $d_{ij} \neq 0$ for which $i \leq n_1$ and $j > n_1$, asserts that only those terms from Eq. (D2) are kept that do not cancel with the corresponding term of the sum in Eq. (D3). We substitute this result into Eq. (D1) and change the order of summation. For $n < 2$ this restriction cannot be fulfilled so the sum in the result Eq. (38) in the main text starts with $u = 1$ and $n = 2$.

Appendix E: First integral for the spectrum matrix

To compute Eq. (40) for the spectrum matrix $\mathbf{P}_e(\omega)$ we first evaluate the two-point time correlation functions $\Delta_{ij}(t_i - t_j)$ of the normal stochastic variables μ_i according to Eq. (41),

$$
\frac{1}{2\pi} \int_0^\infty e^{-i\omega\tau} \prod_{\substack{i,j=1 \\ i<j}}^n \frac{1}{d_{ij}!} \left(\Delta_{f_i^2 f_j^2}(t_i - t_j) \right)^{d_{ij}} \mathrm{d}\tau = \frac{1}{2\pi} \int_0^\infty e^{-i\omega\tau} \prod_{\substack{i,j=1 \\ i<j}}^n \frac{1}{d_{ij}!} \left(\Gamma_{f_i^2 f_j^2} e^{-K_{f_i^2}|t_i - t_j|} \right)^{d_{ij}} \mathrm{d}\tau
$$

$$
= \frac{1}{2\pi} \int_0^\infty e^{-\left(i\omega + \sum_{i<j=1}^n d_{ij} K_{f_i^2} |t_i - t_j| \tau^{-1} \right)\tau} \mathrm{d}\tau \prod_{\substack{i,j=1 \\ i<j}}^n \frac{1}{d_{ij}!} \left(\Gamma_{f_i^2 f_j^2} \right)^{d_{ij}}. \quad \text{(E1)}
$$

So far we have not treated the times t_i, t_j precisely and need to follow them back to equation Eq. (27) where $t_1 = \tau$ corresponds to $\boldsymbol{\eta}^1$ and $t_2 = 0$ to $\boldsymbol{\eta}^2$. The information needed for their correct evaluation is traceable in the definition of the index function f_i^2 in Eq. (39) that maps to components of $\boldsymbol{\eta}^1$ if $i \leq \#\mathbf{r}^1$ ($t_i = \tau$) and to $\boldsymbol{\eta}^2$ else ($t_i = 0$). The difference $|t_i - t_j|\tau^{-1}$ becomes either 0 or 1 and the sum in the exponent reduces to pairs i, j that obey $1 \leq i \leq \#\mathbf{r}^1 < j \leq n$. With finite and positive inverse correlation times K_i, the exponent has a negative real part so the integral with the solution

$$
\frac{1}{2\pi} \left(i\omega + \sum_{i=1}^{\#\mathbf{r}^1} \sum_{j=\#\mathbf{r}^1+1}^n d_{ij} K_{f_i^2} \right)^{-1} \quad \text{(E2)}
$$

exists. Finally, we evaluate the index function $f_i^2 = r_i^1$ for $i \leq \#\mathbf{r}^1$, denote the double sum as $\Theta(\mathbf{d}^n, \mathbf{r}^1)$ in Eq. (43) and add the complex conjugate of the whole expression to obtain $\mathbf{P}_e(\omega)$ in Eq. (42).

Appendix F: Second integral for the spectrum matrix

To calculate the spectrum matrix $\langle \mathbf{P}_i(\omega) \rangle$ in Eq. (29), we expand Eq. (28) with the Taylor coefficients for $\mathbf{A}(\boldsymbol{\eta}_1)$, using Eq. (34), and $\mathbf{C}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$, using Eq. (44) and (45),

$$
\mathbf{R}(\omega) = \int_0^\infty e^{-\left(-\mathbf{A}^0 + i\omega \right)\tau} \left\langle \sum_{c=0}^\infty \frac{\tau^c}{c!} \left(\sum_{q=1}^\infty \sum_{\#\mathbf{r}=q} \mathbf{A}_{(\mathbf{r})} \eta_{(\mathbf{r})} \right)^c \left(\sum_{a=0}^\infty \sum_{\#\mathbf{r}=a} \sum_{\#\boldsymbol{\sigma}=a} \mathbf{C}_{(\mathbf{r},\boldsymbol{\sigma})} \eta_{(\mathbf{r},\boldsymbol{\sigma})} \right) \right\rangle \mathrm{d}\tau.
$$

$$
\text{(F1)}
$$

The term in the average is a sum of terms of the form

$$
\frac{\tau^c}{c!} \left(\sum_{\#\mathbf{r}^1=q_1} \mathbf{A}_{(\mathbf{r}^1)} \eta_{(\mathbf{r}^1)} \right) \cdots \left(\sum_{\#\mathbf{r}^c=q_c} \mathbf{A}_{(\mathbf{r}^c)} \eta_{(\mathbf{r}^c)} \right) \left(\sum_{\#\mathbf{r}^{c+1}=a} \sum_{\#\boldsymbol{\sigma}=a} \mathbf{C}_{(\mathbf{r}^{c+1},\boldsymbol{\sigma})} \eta_{(\mathbf{r}^{c+1},\boldsymbol{\sigma})} \right). \quad \text{(F2)}
$$

38

With the multi-index $\mathbf{q} = (q_1, \dots, q_c)$ from Eq. (47) we change to $n = |\mathbf{q}| + a$ (the order in $\boldsymbol{\eta}$) as principal sum index,

$$\mathbf{R}(\omega) = \sum_{n=0}^{\infty} \sum_{a=0}^{n} \sum_{|\mathbf{q}|=n-a} \sum_{\#\mathbf{r}^1=q_1} \cdots \sum_{\#\mathbf{r}^c=q_c} \sum_{\#\boldsymbol{\sigma}=a}$$
$$\times \int_0^{\infty} e^{-(-\mathbf{A}^0+i\omega)\tau} \frac{\tau^c}{c!} \Big(\mathbf{A}_{(\mathbf{r}^1)} \dots \mathbf{A}_{(\mathbf{r}^c)} \mathbf{C}_{(\mathbf{r}^{c+1},\boldsymbol{\sigma})} \big\langle \eta_{(\mathbf{r}^1)} \dots \eta_{(\mathbf{r}^c)} \eta_{(\mathbf{r}^{c+1},\boldsymbol{\sigma})} \big\rangle \Big) \, \mathrm{d}\tau.$$

The sum $\sum_{|\mathbf{q}|=n-a}$ is carried out over all possible c, $(q_1, \dots, q_c)$ with $q_1 + \cdots + q_c + a = n$. The correlation functions are calculated according to Eq. (32). After changing the order of summation, a comparison to $\mathbf{R}(\omega)$ in Eq. (48) gives

$$\left(-\mathbf{A}^0 + \theta(\mathbf{d}^c, |\mathbf{q}|, \mathbf{r}^{c+1}, \boldsymbol{\sigma}) + i\omega\right)^{-(c+1)} = \frac{1}{c!} \int_0^{\infty} e^{-\left(-\mathbf{A}^0+i\omega+\sum_{i<j=1}^{n} d_{ij} K_{f_i^{c+1}} |t_i-t_j|\tau^{-1}\right)\tau} \tau^c \, \mathrm{d}\tau$$
$$= \left(-\mathbf{A}^0 + i\omega + \sum_{\substack{i,j=1\\i<j}}^{n} d_{ij} K_{f_i^{c+1}} |t_i - t_j|\tau^{-1}\right)^{-(c+1)} \quad \text{(F3)}$$

where the sum in the exponent stems from $\Delta_{ij}(t_i - t_j)$ in Eq. (41) and $\int_0^{\infty} e^{-a\tau} \tau^c \, \mathrm{d}\tau = \left(-\frac{\mathrm{d}}{\mathrm{d}a}\right)^c \int_0^{\infty} e^{-a\tau} \, \mathrm{d}\tau = \frac{c!}{a^{c+1}}$ was used in the last equality. We identify the θ function as the sum in the last term that we evaluate following the arguments in section E. The difference $|t_i - t_j|\tau^{-1}$ is 0 or 1 and non zero if and only if f_j^{c+1} in Eq. (39) maps to a component of r^{c+1}, that is $j > |\mathbf{q}|$, and the index function f_i^{c+1} either maps to a component of $\mathbf{r}^i$, $i \in \{1, \dots, c\}$ (corresponding to t_1 in Eq. 28) nd $\sigma_{i-|\mathbf{q}|} = 2$ (corresponding to t_2) or also $i > |\mathbf{q}|$ and $\sigma_{i-|\mathbf{q}|} \neq \sigma_{j-|\mathbf{q}|}$ (so $t_1 \neq t_2$). This result is formalised in Eq. (49) and (50) in the main text. For simplification of the notation in Eq. (49), $K_{f_i^{c+1}}$ is evaluated to $K_{r_{j-|\mathbf{q}|}^{c+1}}$ which is allowed due to the $\Gamma_{f_i^{c+1} f_j^{c+1}}$ in Eq. (48) that is proportional to $\delta_{f_i^{c+1} f_j^{c+1}}$.

Appendix G: Evaluation of the small noise expansion

While the closed-form expressions obtained from the small noise expansion are well suited for automated evaluation, the notation is rather complicated and will be unfamiliar to most readers. To facilitate reading of the sums to the reader we evaluate the first terms in more detail here. The simplest of the sums is Eq. (36) for the mean concentrations

$$\langle \mathbf{x}^s \rangle = \sum_{u=0}^{\infty} \sum_{n=0}^{2u} \sum_{\#\mathbf{r}=n} \phi_{(\mathbf{r})}^s \sum_{|\mathbf{d}^n|=u}' \prod_{\substack{i,j=1\\i<j}}^{n} \frac{1}{d_{ij}!} \left(\Gamma_{r_i r_j}\right)^{d_{ij}}$$

39

for which we evaluate all terms for $u = 0$ and $u = 1$:

$$u = 0 \quad n = 0 \quad \mathbf{r} = () \quad \phi^s_{(\mathbf{r})} = \phi^s_{()} = \phi^s(\mathbf{0})$$

$$u = 1 \quad n = 0 \quad \mathbf{r} = () \quad 0$$

$$n = 1 \quad \mathbf{r} = (i) \quad 0$$

$$n = 2 \quad \mathbf{r} = (i,j) \quad \phi^s_{(i,j)}\Gamma_{ij} = \tfrac{1}{2}\tfrac{\partial}{\partial\eta_i}\tfrac{\partial}{\partial\eta_j}\,\phi^s(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\,\delta_{ij}\epsilon_i^2 = \tfrac{1}{2}\tfrac{\partial^2}{\partial\eta_i^2}\,\phi^s(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\,\epsilon_i^2$$

where the last term is a sum over the index i that enumerates the extrinsic noise variables η_i (Einstein notation). The $\Gamma_{r_i r_j}$ symbol has been evaluated according to Eq. (35). We obtain the first order result in the variances ϵ_i^2 of the stochastic variables $\mu_i(t)$ from which we constructed the lognormal variables (Eq. A5)

$$\langle \mathbf{x}^s \rangle = \phi^s(\mathbf{0}) + \frac{1}{2}\frac{\partial^2}{\partial\eta_i^2}\,\phi^s(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\,\epsilon_i^2 + \mathcal{O}(\epsilon_i^4). \tag{G1}$$

The contribution $\mathbf{V}(\boldsymbol{\xi})$ in Eq. (37) to the total variance is formally equivalent and we obtain

$$\mathbf{V}(\boldsymbol{\xi}) = \mathbf{C}(\mathbf{0}) + \frac{1}{2}\frac{\partial^2}{\partial\eta_i^2}\,\mathbf{C}(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\,\epsilon_i^2 + \mathcal{O}(\epsilon_i^4). \tag{G2}$$

As opposed to Eq. (45), $\frac{\partial}{\partial\eta_i}\,\mathbf{C}(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}$ is a Taylor coefficient of $\mathbf{C}(\boldsymbol{\eta}) \equiv \mathbf{C}(\boldsymbol{\eta},\boldsymbol{\eta})$ from Eq. (22) at equal times.

For the purely extrinsic contribution to the variance in Eq. (38)

$$\mathbf{V}(\boldsymbol{\phi}^s) = \sum_{u=1}^{\infty}\sum_{n=2}^{2u}\sum_{|\mathbf{q}|=n}\sum_{\#\mathbf{r}^1=q_1}\sum_{\#\mathbf{r}^2=q_2}\phi^s_{(\mathbf{r}^1)}\phi^{s\,T}_{(\mathbf{r}^2)}\sideset{}{''}\sum_{|\mathbf{d}^n|=u}\prod_{\substack{i,j=1\\i<j}}^{n}\frac{1}{d_{ij}!}\left(\Gamma_{f_i^2 f_j^2}\right)^{d_{ij}}$$

the sum starts with $u = 1$ and with the double prime only terms with derivatives of both ϕ^s and $\phi^{s\,T}$ contribute,

$$u = 1 \quad \mathbf{q} = (2,0) \quad \mathbf{r}^1 = (i,j) \quad \mathbf{r}^2 = () \quad 0$$

$$\mathbf{q} = (1,1) \quad \mathbf{r}^1 = (i) \quad \mathbf{r}^2 = (j) \quad \phi^s_{(i)}\phi^{s\,T}_{(j)}\Gamma_{ij} = \left(\tfrac{\partial}{\partial\eta_i}\,\phi^s(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\right)\left(\tfrac{\partial}{\partial\eta_j}\,\phi^s(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\right)^T\delta_{ij}\epsilon_i^2$$

$$\mathbf{q} = (0,2) \quad \mathbf{r}^1 = () \quad \mathbf{r}^2 = (i,j) \quad 0 \ .$$

The $\Gamma_{f_i^2 f_j^2}$ symbol has been evaluated with Eq. (35) and (39). The first order result in the variances ϵ_i^2 is

$$\mathbf{V}(\boldsymbol{\phi}^s) = \left(\frac{\partial}{\partial\eta_i}\,\phi^s(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\right)\left(\frac{\partial}{\partial\eta_i}\,\phi^s(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\right)^T\epsilon_i^2 + \mathcal{O}(\epsilon_i^4). \tag{G3}$$

The total variance is $\mathbf{V}(\mathbf{x}^s) = \mathbf{V}(\boldsymbol{\phi}^s) + \frac{1}{\Omega}\mathbf{V}(\boldsymbol{\xi})$ according to Eq. (12).

In comparison to $\mathbf{V}(\boldsymbol{\phi}^s)$, the purely extrinsic contribution to the spectrum matrix $\mathbf{P}_e(\omega)$ in Eq. (42) needs evaluation of the additional factor $\frac{\pi^{-1}\Theta}{\omega^2+\Theta^2}$ with $\mathbf{d}^2 = (d_{12}) = (1)$, $\mathbf{r}^1 = (i)$ and $n = 2$. Eq. (43) then gives $\Theta(\mathbf{d}^2, \mathbf{r}^1) = K_i$ and

$$\mathbf{P}_e(\omega) = \left(\frac{\partial}{\partial\eta_i}\,\phi^s(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\right)\left(\frac{\partial}{\partial\eta_i}\,\phi^s(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}}\right)^T\frac{K_i\epsilon_i^2}{\pi\,(\omega^2 + K_i^2)} + \mathcal{O}(\epsilon_i^4). \tag{G4}$$

For the second spectrum matrix $\mathbf{P}_e(\omega)$ we need to evaluate Eq. (48),

$$\mathbf{R}(\omega) = \sum_{u=0}^{\infty}\sum_{n=0}^{2u}\sum_{a=0}^{n}\sum_{|\mathbf{q}|=n-a}\sum_{\#\mathbf{r}^1=q_1}\cdots\sum_{\#\mathbf{r}^c=q_c}\sum_{\#\mathbf{r}^{c+1}=a}\sum_{\#\boldsymbol{\sigma}=a}{\sum_{|\mathbf{d}^n|=u}}'$$

$$\times \frac{1}{\left(-\mathbf{A}^0 + \theta(\mathbf{d}^n, |\mathbf{q}|, \mathbf{r}^{c+1}, \boldsymbol{\sigma}) + i\omega\right)^{c+1}}\,\mathbf{A}_{(\mathbf{r}^1)}\dots\mathbf{A}_{(\mathbf{r}^c)}\mathbf{C}_{(\mathbf{r}^{c+1},\boldsymbol{\sigma})}\prod_{\substack{i,j=1\\i<j}}^{n}\frac{1}{d_{ij}!}\left(\Gamma_{f_i^{c+1}f_j^{c+1}}\right)^{d_{ij}}.$$

To elucidate the full complexity of the sum, we here evaluate some terms for $u = 2$. We use an abbreviated notation for derivatives as defined by:

$$\mathbf{A}_{ij}'' \equiv \frac{1}{2!}\frac{\partial}{\partial\eta_i}\frac{\partial}{\partial\eta_j}\,\mathbf{A}(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\mathbf{0}} \qquad \text{and} \qquad \mathbf{C}_i^{\sigma} \equiv \frac{\partial}{\partial\eta_i^{\sigma}}\,\mathbf{C}(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)|_{\boldsymbol{\eta}^1=\boldsymbol{\eta}^2=\mathbf{0}}\,. \tag{G5}$$

For $u = 2$, $n = 4$, $a = 1$ and $c = 2$ we obtain the terms

$$\sum_{\substack{|\mathbf{q}|=3\\\#\mathbf{q}=2}}\sum_{\#\mathbf{r}^1=q_1}\sum_{\#\mathbf{r}^2=q_2}\sum_{r_1^3}\sum_{\sigma_1}{\sum_{|\mathbf{d}^4|=2}}'\frac{1}{\left(-\mathbf{A}^0 + \theta(\mathbf{d}^4, 3, \mathbf{r}^3, \boldsymbol{\sigma}) + i\omega\right)^3}\,\mathbf{A}_{(\mathbf{r}^1)}\mathbf{A}_{(\mathbf{r}^2)}\mathbf{C}_{(\mathbf{r}^3,\boldsymbol{\sigma})}\prod_{\substack{i,j=1\\i<j}}^{4}\frac{1}{d_{ij}!}\left(\Gamma_{f_i^3f_j^3}\right)^{d_{ij}}$$

$$= \left(\frac{\delta_{ij}\delta_{kl}\epsilon_i^2\epsilon_k^2}{\left(-\mathbf{A}^0 + K_{kl}\delta_{2\sigma} + i\omega\right)^3} + \frac{\delta_{ik}\delta_{jl}\epsilon_i^2\epsilon_j^2}{\left(-\mathbf{A}^0 + K_{jl}\delta_{2\sigma} + i\omega\right)^3} + \frac{\delta_{il}\delta_{jk}\epsilon_i^2\epsilon_l^2}{\left(-\mathbf{A}^0 + K_{il}\delta_{2\sigma} + i\omega\right)^3}\right)\left(\mathbf{A}_i'\mathbf{A}_{jk}'' + \mathbf{A}_{ij}''\mathbf{A}_k'\right)\mathbf{C}_l^{\sigma}. \tag{G6}$$

The sum over $|\mathbf{q}| = 3$ was evaluated by writing $c = 2$ symbols $\mathbf{A}$ with all possibilities to assign at least one of a total of three indices to each of them. The Einstein notation for the extrinsic noise components $i, j, k, l \in \{1, \dots, M\}$ of $\boldsymbol{\eta}$ and times t_{σ}, $\sigma \in \{1, 2\}$ accounts for all other sums except the primed sum over $|\mathbf{d}^4| = 2$. The latter involves 3 tuples $(d_{12}, d_{13}, d_{23}, d_{14}, d_{24}, d_{34})$, namely all components zero but $d_{12} = d_{34} = 1$, $d_{13} = d_{24} = 1$ or $d_{14} = d_{23} = 1$. The covariances $\Gamma_{f_i^3f_j^3}$ and the θ function evaluate according to Eq. (35), (39) and (49).

With $n = 3$ instead, $\mathbf{q}$ needs to be $(1, 1)$ and we sum over $|\mathbf{d}^3| = 2$ and obtain the three

terms

$$\left(\frac{\delta_{ij}\delta_{ik}\epsilon_i^4}{\left(-\mathbf{A}^0 + K_{ik}\delta_{2\sigma} + i\omega\right)^3} + \frac{\delta_{ij}\delta_{jk}\epsilon_i^4}{\left(-\mathbf{A}^0 + K_{jk}\delta_{2\sigma} + i\omega\right)^3} + \frac{\delta_{ik}\delta_{jk}\epsilon_i^4}{\left(-\mathbf{A}^0 + (K_{ik}+K_{jk})\delta_{2\sigma} + i\omega\right)^3}\right)\mathbf{A}_i'\mathbf{A}_j'\mathbf{C}_k^\sigma.$$

(G7)

Finally, we evaluate the terms for $u = n = 2$ and $a = 0$ for which $|\mathbf{d}^2| = 2$ only allows $d_{12} = 2$ and with $a = 0$ no inverse correlation times K_i are involved but $\mathbf{q}$ may be both (2) with $c = 1$ or $(1, 1)$ with $c = 2$ which gives

$$\frac{\frac{1}{2!}\delta_{ij}^2\epsilon_i^4}{\left(-\mathbf{A}^0 + i\omega\right)^2}\mathbf{A}_{ij}''\mathbf{C}_0 + \frac{\frac{1}{2!}\delta_{ij}^2\epsilon_i^4}{\left(-\mathbf{A}^0 + i\omega\right)^3}\mathbf{A}_i'\mathbf{A}_j'\mathbf{C}_0 = \frac{\frac{1}{2}\epsilon_i^4}{\left(-\mathbf{A}^0 + i\omega\right)^2}\left(\mathbf{A}_{ii}'' + \frac{1}{\left(-\mathbf{A}^0 + i\omega\right)}\mathbf{A}_i'\mathbf{A}_i'\right)\mathbf{C}_0.$$

(G8)

We note here that the matrix multiplication is non commutative so the order of terms is important. In general, using Einstein notation for fixed u and n, one writes down all possible terms $\mathbf{A}\dots\mathbf{AC}$ with n indices ($\mathbf{C}_0$ without index is allowed but not so for the Taylor coefficients of $\mathbf{A}$) and then for each term evaluates the remaining sum over $|\mathbf{d}^n| = u$ in order to derive the factors containing the θ function (here c is the number of $\mathbf{A}$-symbols) and covariances Γ_{ij}.

The spectrum matrix $\langle\mathbf{P}_i(\omega)\rangle = \frac{1}{2\pi}\left(\mathbf{R}(\omega) + \mathbf{R}(\omega)^{*T}\right)$ in zeroth order is the power spectrum of the Ornstein-Uhlenbeck process for the concentrations $\mathbf{x}$ in the absence of extrinsic noise. The Ornstein-Uhlenbeck result is obtained from Eq. (9) by setting $\boldsymbol{\eta}$ to zero. Explicitly, its power spectrum is given by the well known result [25]

$$\mathbf{R}(\omega) = \left(-\mathbf{A}^0 + i\omega\right)^{-1}\mathbf{C}_0 + \mathcal{O}\left(\epsilon_i^2\right)$$

(G9)

$$\Rightarrow \langle\mathbf{P}_i(\omega)\rangle = \frac{1}{2\pi}\left(-\mathbf{A}^0 + i\omega\right)^{-1}\mathbf{B}\mathbf{B}^T\left(-\mathbf{A}^{0T} - i\omega\right)^{-1} + \mathcal{O}\left(\epsilon_i^2\right)$$

(G10)

where we have used the Lyapunov equation (22) for the last equality. According to Eq. (23), the total spectrum matrix is $\mathbf{P}(\omega) = \mathbf{P}_e(\omega) + \frac{1}{\Omega}\langle\mathbf{P}_i(\omega)\rangle$.

---

[1] D. Gillespie, Exact stochastic simulation of coupled chemical reactions, Journal of Physical Chemistry **81**, 2340 (1977).

[2] J. M. Raser and E. K. O'shea, Noise in gene expression: origins, consequences, and control, Science **309**, 2010 (2005).

[3] C. Rao, D. Wolf, and A. Arkin, Control, exploitation and tolerance of intracellular noise, Nature **420**, 231 (2002).

[4] R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko, Information transduction capacity of noisy biochemical signaling networks, Science **334**, 354 (2011).

[5] N. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, 1997).

[6] M. Elowitz, A. Levine, E. Siggia, and P. Swain, Stochastic gene expression in a single cell, Science **297**, 1183 (2002).

[7] J. Raser and E. O'Shea, Control of stochasticity in eukaryotic gene expression, Science **304**, 1811 (2004).

[8] I. S. Araújo, J. M. Pietsch, E. M. Keizer, B. Greese, R. Balkunde, C. Fleck, and M. Hülskamp, Stochastic gene expression in arabidopsis thaliana, Nature communications **8**, 2132 (2017).

[9] R. Grima, D. Schmidt, and T. Newman, Steady-state fluctuations of a genetic feedback loop: An exact solution, Journal of Chemical Physics **137**, 035104 (2012).

[10] N. Kumar, T. Platini, and R. V. Kulkarni, Exact distributions for stochastic gene expression models with bursting and feedback, Physical review letters **113**, 268105 (2014).

[11] V. Shahrezaei, J. Ollivier, and P. Swain, Colored extrinsic fluctuations and stochastic gene expression, Molecular Systems Biology **4**, 196 (2008).

[12] D. F. Anderson, A modified next reaction method for simulating chemical systems with time dependent propensities and delays, The Journal of chemical physics **127**, 214107 (2007).

[13] M. Voliotis, P. Thomas, R. Grima, and C. G. Bowsher, Stochastic simulation of biomolecular networks in dynamic environments, PLoS computational biology **12**, e1004923 (2016).

[14] D. Schnoerr, G. Sanguinetti, and R. Grima, Approximation and inference methods for stochastic biochemical kinetics?a tutorial review, Journal of Physics A: Mathematical and Theoretical **50**, 093001 (2017).

[15] M. Scott, B. Ingalls, and M. Kaern, Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks, Chaos: An Interdisciplinary Journal of Nonlinear Science **16**, 026107 (2006).

[16] T. Toni and B. Tidor, Combined model of intrinsic and extrinsic variability for computational network design with application to synthetic biology, PLoS Computational Biology **9**, e1002960 (2013).

[17] E. Roberts, S. Be'er, C. Bohrer, R. Sharma, and M. Assaf, Dynamics of simple gene-network

motifs subject to extrinsic fluctuations, Physical Review E **92**, 062717 (2015).

[18] N. Rosenfeld, J. Young, P. Swain, and M. Elowitz, Gene regulation at the single-cell level, Science **307**, 1962 (2005).

[19] Z. Hensel, H. Feng, B. Han, C. Hatem, J. Wang, and J. Xiao, Stochastic expression dynamics of a transcription factor revealed by single-molecule noise analysis, Nature Structural and Molecular Biology **19**, 797 (2012).

[20] C. Furusawa, T. Suzuki, A. Kashiwagi, T. Yomo, and K. Kaneko, Ubiquity of log-normal distributions in intra-cellular reaction dynamics, Biophysics **1**, 25 (2005).

[21] M. Bengtsson, A. Ståhlberg, P. Rorsman, and M. Kubista, Gene expression profiling in single cells from the pancreatic islets of langerhans reveals lognormal distribution of mrna levels, Genome research **15**, 1388 (2005).

[22] P. Lu, C. Vogel, R. Wang, X. Yao, and E. M. Marcotte, Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation, Nature biotechnology **25**, 117 (2007).

[23] N. G. van Kampen, A power series expansion of the master equation, Canadian Journal of Physics **39**, 551 (1961).

[24] N. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).

[25] C. Gardiner, *Handbook of stochastic methods for physics, chemistry and natural sciences*, 2nd ed. (Springer-Verlag, Berlin; Heidelberg [etc.], 1990).

[26] A. Malakhov, *Cumulant Analysis of Non-Gaussian Random Processes and their Transformations [in Russian]* (Sovetskoe Radio, Moscow, 1978).

[27] D. Gillespie, Stochastic simulation of chemical kinetics, Annual Review of Physical Chemistry **58**, 35 (2007).

[28] A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, T. Rathnayake, S. Vig, B. E. Granger, R. P. Muller, F. Bonazzi, H. Gupta, S. Vats, F. Johansson, F. Pedregosa, M. J. Curry, A. R. Terrel, v. Roučka, A. Saboo, I. Fernando, S. Kulal, R. Cimrman, and A. Scopatz, Sympy: symbolic computing in python, PeerJ Computer Science **3**, e103 (2017).

[29] B. Bastian, ext_noise_expansion – Linear noise approximation with slow extrinsic fluctuations, `https://github.com/basbjo/ext_noise_expansion` (2013).

[30] B. Bastian, *Spatial and temporal patterns in biology*, Master's thesis, University of Freiburg, Germany (2013).

[31] E. Jones, T. Oliphant, P. Peterson, *et al.*, SciPy: Open source scientific tools for Python (2001–), [Online; accessed April 10, 2019].

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

[33] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins, Stochasticity in gene expression: from theories to phenotypes, Nature Reviews Genetics **6**, 451 (2005).

[34] R. Grima, Linear-noise approximation and the chemical master equation agree up to second-order moments for a class of chemical systems, Physical Review E **92**, 042124 (2015).

[35] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, Global quantification of mammalian gene expression control, Nature **473**, 337 (2011).

[36] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef, Mammalian genes are transcribed with widely different bursting kinetics, Science **332**, 472 (2011).

[37] L. Brančík and E. Kolářová, Simulation of multiconductor transmission lines with random parameters via stochastic differential equations approach, Simulation **92**, 521 (2016).

[38] Y. Dublanche, K. Michalodimitrakis, N. Kümmerer, M. Foglierini, and L. Serrano, Noise in transcription negative feedback loops: simulation and experimental analysis., Molecular Systems Biology **2** (2006).

[39] J. Zhang, Z. Yuan, and T. Zhou, Physical limits of feedback noise-suppression in biological networks, Physical Biology **6**, 046009 (2009).

[40] I. Lestas, G. Vinnicombe, and J. Paulsson, Fundamental limits on the suppression of molecular fluctuations, Nature **467**, 174 (2010).

[41] A. Grönlund, P. Lötstedt, and J. Elf, Transcription factor binding kinetics constrain noise suppression via negative feedback, Nature communications **4**, 1864 (2013).

[42] D. A. Oyarzún, J.-B. Lugagne, and G.-B. V. Stan, Noise propagation in synthetic gene circuits for metabolic control, ACS Synthetic Biology **4**, 116 (2014).

[43] Z. Cao and R. Grima, Linear mapping approximation of gene regulatory networks with stochastic dynamics, Nature communications **9**, 3305 (2018).

[44] V. Shahrezaei and P. S. Swain, Analytical distributions for stochastic gene expression, Pro-

ceedings of the National Academy of Sciences **105**, 17256 (2008).

[45] J. Paulsson, Models of stochastic gene expression, Physics of life reviews **2**, 157 (2005).

[46] F. Lang, M. Ritter, E. Wöll, H. Weiss, D. Häussinger, J. Hoflacher, K. Maly, and H. Grunicke, Altered cell volume regulation in ras oncogene expressing nih fibroblasts, Pflügers Archiv **420**, 424 (1992).

[47] M. D. Huber and L. Gerace, The size-wise nucleus: nuclear volume control in eukaryotes, The Journal of cell biology **179**, 583 (2007).

[48] A. Bar-Even, E. Noor, Y. Savir, W. Liebermeister, D. Davidi, D. S. Tawfik, and R. Milo, The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters, Biochemistry **50**, 4402 (2011).

[49] T. Marquez-Lago and J. Stelling, Counter-intuitive stochastic behaviour of simple gene circuits with negative feedback, Biophysical Journal **98**, 1742 (2010).

[50] M. Voliotis and C. G. Bowsher, The magnitude and colour of noise in genetic negative feedback systems, Nucleic acids research **40**, 7084 (2012).

[51] A. Singh and J. P. Hespanha, Optimal feedback strength for noise suppression in autoregulatory gene networks, Biophysical journal **96**, 4013 (2009).

[52] I. G. Johnston, B. Gaal, R. P. das Neves, T. Enver, F. J. Iborra, and N. S. Jones, Mitochondrial variability as a source of extrinsic cellular noise, PLoS computational biology **8**, e1002416 (2012).

[53] N. Rosenfeld, M. B. Elowitz, and U. Alon, Negative autoregulation speeds the response times of transcription networks, Journal of molecular biology **323**, 785 (2002).

[54] F. Tostevin and P. R. Ten Wolde, Mutual information between input and output trajectories of biochemical networks, Physical Review Letters **102**, 218101 (2009).

[55] D. Bernardi and B. Lindner, A frequency-resolved mutual information rate and its application to neural systems, Journal of Neurophysiology **113**, 1342 (2015).

[56] F. Gabbiani, Coding of time-varying signals in spike trains of linear and half-wave rectifying neurons, Network: Computation in Neural Systems (1996).

[57] S. Mangan and U. Alon, Structure and function of the feed-forward loop network motif, Proceedings of the National Academy of Sciences **100**, 11980 (2003).

[58] E. Dekel, S. Mangan, and U. Alon, Environmental selection of the feed-forward loop circuit in gene-regulation networks, Physical biology **2**, 81 (2005).

[59] J. Macía, S. Widder, and R. Solé, Specialized or flexible feed-forward loop motifs: a question of topology, BMC systems biology **3**, 84 (2009).

[60] R. Grima, An effective rate equation approach to reaction kinetics in small volumes: Theory and application to biochemical reactions in nonequilibrium steady-state conditions, The Journal of Chemical Physics **133**, 035101 (2010).

[61] F. Fröhlich, P. Thomas, A. Kazeroonian, F. J. Theis, R. Grima, and J. Hasenauer, Inference for stochastic chemical kinetics using moment equations and system size expansion, PLoS computational biology **12**, e1005030 (2016).

[62] R. Grima, P. Thomas, and A. V. Straube, How accurate are the nonlinear chemical fokker-planck and chemical langevin equations?, The Journal of chemical physics **135**, 084103 (2011).

[63] R. Grima, A study of the accuracy of moment-closure approximations for stochastic chemical kinetics, The Journal of chemical physics **136**, 04B616 (2012).

[64] P. Thomas, A. V. Straube, J. Timmer, C. Fleck, and R. Grima, Signatures of nonlinearity in single cell noise-induced oscillations, Journal of Theoretical Biology (2013).

[65] P. Thomas, C. Fleck, R. Grima, and N. Popović, System size expansion using feynman rules and diagrams, Journal of Physics A: Mathematical and Theoretical **47**, 455007 (2014).

[66] M. Le Bellac, *Quantum and Statistical Field Theory* (Clarendon Press, Oxford, 1991).