Noise-induced transitions in gene circuits: A perturbative approach for slow noise

Gerardo Aquino[1,*] and Andrea Rocco [2,3,†]

[1]*Department of Computing, Goldsmiths, University of London, SE14 6AD London, United Kingdom*
[2]*Department of Physics, University of Surrey, GU2 7XH Guildford, United Kingdom*
[3]*Department of Microbial Sciences, University of Surrey, GU2 7XH, Guildford, United Kingdom*

We consider a generic class of gene circuits affected by nonlinear extrinsic noise. To address this nonlinearity we introduce a general perturbative methodology based on assuming timescale separation between noise and gene dynamics, with fluctuations exhibiting a large but finite correlation time. We apply this methodology to the case of the toggle switch, and by considering biologically relevant log-normal fluctuations, we find that the system exhibits noise-induced transitions. The system becomes bimodal in regions of the parameter space where it would be deterministically monostable. We show that by including higher order corrections our methodology allows one to obtain correct predictions for the occurrence of transitions even for not so large correlation time of the fluctuations, thereby overcoming limitations of previous theoretical approaches. Interestingly we find that at intermediate noise intensities the noise-induced transition in the toggle switch affects one of the genes involved, but not the other one.

I. INTRODUCTION

Extrinsic noise has been investigated thoroughly over the last decades in a variety of physical and chemical systems [1]. It has emerged as an active dynamical player, which can produce so-called noise-induced transitions to bimodal dynamics in systems otherwise deterministically monostable.

In chemical and physical systems, the search and the possible emergence of such transitions relies on the multiplicative nature of the corresponding stochastic dynamics, which is mathematically well characterized when the noise appears linearly in the system's dynamical equations, and it is Gaussian and white. In this case the Stratonovich prescription to interpret the corresponding stochastic integral leads to the emergence of the so-called Stratonovich drift in the Langevin and Fokker-Planck dynamics, which is responsible for dramatic changes in the stability properties of stationary states, and possibly in their number [2].

In biological systems, however, none of the assumptions on the noise being linear, Gaussian, and white is usually fulfilled. In gene regulatory networks, for instance, the noise is often non-Gaussian [3,4], is usually characterized by correlation times that can easily exceed cell cycles times [5], and it appears mostly in a nonlinear fashion in the dynamics determining gene expression.

*Gerardo.Aquino@apha.gov.uk
†A.Rocco@surrey.ac.uk

While it can be very hard to treat exactly the corresponding stochastic dynamics under these conditions, a geometric construction based on noise filtering has been introduced [6,7]. In essence, this approach relies on the transformation properties of probability distributions of stochastic variables related by a given input/output transformation. Conservation of probability allows one to express the output probability distribution in terms of the input probability distribution, which results in nontrivial dynamics when the relation between the stochastic variables is nonlinear.

This approach has been rederived dynamically in [8] by extending to nonlinear noise the so-called switching-curve approximation introduced in [9,10]. The dynamical derivation highlights the crucial role played by the timescales of the system. The approximation works well when the input noise is much slower (in the limit, infinitely slower) than the dynamics of the system of interest. In a gene expression process, this translates in extrinsic noise being much slower than protein degradation and synthesis [8].

In this paper we take up this issue again, and develop a perturbative framework that allows one to go beyond the approximation defined in [8] by extending that approach further to large but finite correlation times of the noise. We show that the inclusion of higher order perturbative terms can modify dramatically the analysis of the emergence of noise-induced transitions, by removing them when present at lower orders, and therefore providing a qualitative and quantitative more accurate description of the stochastic dynamics.

We apply our formalism to the case of the toggle switch [11], and show that when log-normal noise is considered, the system undergoes a transition in a region of the phase space where the deterministic system would be monostable. We also find that tuning the noise intensity may give rise to a situation in which the marginalized probabilities of the genes

involved show different qualitative features, with overlapping unimodal and bimodal behaviors. We name this feature "partial bimodality" and discuss its origin in the context of noise propagation across the network.

The paper is organized as follows. In Sec. II we present the genetic circuits considered here. In Sec. III we introduce our perturbative approach, and develop it up to the second perturbative order for the generic circuit introduced in Sec. II (with the third perturbative order described for the toggle switch in the Appendix). In Sec. IV we introduce the toggle switch as an application of the methodology, and assess the perturbative framework by comparing the contributions of different orders to direct stochastic simulations. In Sec. V we draw conclusions on the methodology introduced, and make some final remarks.

II. GENETIC CIRCUITS

We consider a generic circuit of interacting genes. We make the assumption that the half-life of mRNAs is much shorter than that of the corresponding proteins, so that mRNA dynamics can be adiabatically eliminated, and gene expression can be described in a single step of combined transcription and translation [12]. We then describe the system by decomposing the rate equations for each synthesized protein into a production term and a degradation term as follows:

$$\frac{d\mathbf{x}}{dt} = \mathbf{g}(\mathbf{x}, R) - \mathbf{K}\mathbf{x} \,. \qquad (1)$$

Here, $\mathbf{x}$ is an n-dimensional vector whose components are the x_i species involved, while $\mathbf{K}$ is a diagonal matrix whose diagonal elements are the degradation rates $k_i = 1/\tau_{x_i}$ of each species x_i, with τ_{x_i} the respective half-lifes. The function $\mathbf{g}(\mathbf{x}, R)$ is a $\mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ Hill-like function of the concentrations of the n species $\mathbf{x}$, and R is a factor that exerts control on the dynamics of $\mathbf{x}$. The factor R can be thought of as an external control parameter (for instance a binding constant), or any other species at high enough concentrations so that it is unaffected by the dynamics of $\mathbf{x}$ (for instance an external signaling molecule, or a transcription factor). For the sake of simplicity, here we limit ourselves to a single factor R, but the procedure can be further generalized to multiple factors.

III. PERTURBATIVE CALCULATION

Let us now consider fluctuations acting on R. We describe them by adding a generic Langevin-type dynamics to Eq. (1),

$$\frac{d\mathbf{x}}{dt} = \mathbf{g}(\mathbf{x}, R) - \mathbf{K}\mathbf{x} \,, \qquad (2)$$

$$\frac{dR}{dt} = \frac{\mu(R)}{\tau} + \sqrt{\frac{D}{\tau}}\nu(R)\xi(t) \,, \qquad (3)$$

with τ the timescale of the process R. In Eq. (3) the functions $\mu(R)$ and $\nu(R)$ are kept generic for now. Their specification allows us to reproduce different noise distributions. Also, the noise $\xi(t)$ is assumed Gaussian and white, with correlator $\langle \xi(t)\xi(t')\rangle = 2\delta(t - t')$.

By rescaling time as $t \to t/\tau_M$, with $\tau_M = 1/k_M$ the largest degradation time of the species involved (we assume that $\tau_{x_i} \lesssim \tau_M$ for $x_i \neq x_M$) we obtain

$$\frac{d\mathbf{x}}{dt} = \varepsilon\boldsymbol{\gamma}(\mathbf{x}, R) - \bar{\mathbf{K}}\mathbf{x} \,, \qquad (4)$$

$$\frac{dR}{dt} = \varepsilon\mu(R) + \sqrt{\varepsilon D}\nu(R)\xi(t) \,, \qquad (5)$$

where $\varepsilon = \tau_M/\tau$, $\boldsymbol{\gamma} = \tau\mathbf{g}$, and $\bar{\mathbf{K}} = \mathbf{K}/k_M$ with $\bar{\mathbf{K}}$ still diagonal and $\bar{\mathbf{K}}_{ii} = \kappa_{ii} = k_{x_i}/k_M = O(1)$. In order for the dynamics of R to be slower than that of all the species x_i, we consider the condition $\tau_M \ll \tau$, implying that ε is a small parameter, with $\varepsilon \ll 1$.

This naturally leads to seeking a solution of the system (4) and (5) in terms of a perturbative expansion in orders of ε. This perturbative approach is inspired by the one presented in [8], with the important difference that the latter was based on assuming the limit $\tau \to \infty$, while here we consider instead the limit $\tau_M \to 0$ (i.e., $\varepsilon \to 0$) with τ large, but fixed and finite. This different perspective keeps the function $\boldsymbol{\gamma} = \tau\mathbf{g}$ in (4) fixed in the limit $\varepsilon \to 0$, and allows us to obtain a systematic perturbative expansion in ε. As a result, the approach presented here allows us to go beyond the lowest order calculation of [8], accounting for the noise filtering approach of [6,7], and to obtain all perturbative corrections analytically at any order in ε.

Let us define the Fokker-Planck operator L_{FP} for the R process as

$$L_{\text{FP}}(R) = \frac{\partial}{\partial R}\left[-(\mu + D\nu\nu') + D\frac{\partial}{\partial R}(\nu^2)\right], \qquad (6)$$

where we have assumed the Stratonovich interpretation for the multiplicative noise equation (5).

The total Fokker-Planck equation for the time-dependent probability density $W_t(\mathbf{x}, R)$ is therefore

$$\frac{\partial}{\partial t}W_t(\mathbf{x}, R) = [-\nabla \cdot (\varepsilon\boldsymbol{\gamma}(\mathbf{x}, R) - \bar{\mathbf{K}}\mathbf{x}) + \varepsilon L_{\text{FP}}(R)]W_t(\mathbf{x}, R) \,. \qquad (7)$$

We now aim to compute the stationary solution of Eq. (7), namely the time-independent probability density $W(\mathbf{x}, R)$ satisfying

$$[\nabla \cdot (\varepsilon\boldsymbol{\gamma}(\mathbf{x}, R) - \bar{\mathbf{K}}\mathbf{x}) - \varepsilon L_{\text{FP}}(R)]W(\mathbf{x}, R) = 0 \,, \qquad (8)$$

obtained by imposing $\frac{\partial}{\partial t}W_t(\mathbf{x}, R) = 0$ in (7). We seek a solution of the type

$$W(\mathbf{x}, R) = \bar{W}(R)\psi(\mathbf{x}, R), \qquad (9)$$

where $\bar{W}$ is chosen so that $L_{\text{FP}}\bar{W}(R) = 0$, since the dynamics of the variable R is independent of x, y (but not viceversa). In general it can be shown [2] that $\bar{W}(R)$ can be written as

$$\bar{W}(R) = \frac{\mathcal{N}}{\nu(R)}\exp\left(\frac{1}{D}\int^R dR'\frac{\mu(R')}{\nu^2(R')}\right), \qquad (10)$$

and it must be as well

$$\int \mathbf{dx}W(\mathbf{x}, R) = \bar{W}(R) \Rightarrow \int \mathbf{dx}\psi(\mathbf{x}, R) = 1 \,. \qquad (11)$$

A. Zeroth order solution

We begin by assuming the following perturbative expansion for $W(\mathbf{x}, R)$ of generic order m in powers of ε

$$W(\mathbf{x}, R) = \sum_{j=0}^{m} \frac{\varepsilon^j}{j!} W^{(j)}(\mathbf{x}, R) + O(\varepsilon^{m+1}) . \quad (12)$$

By replacing this expansion in the lefthand side of (8), and by keeping only the zeroth order terms in ε, we obtain

$$(\nabla \cdot \mathbf{x}) W^{(0)}(\mathbf{x}, R) = 0 , \quad (13)$$

which immediately leads to the zeroth order solution

$$W^{(0)}(\mathbf{x}, R) = \delta^n(\mathbf{x}) \bar{W}(R) = \bar{W}(R) \prod_{i=1}^{n} \delta(x_i) . \quad (14)$$

B. First order correction

Having evaluated $W^{(0)}$ we can again replace the expansion (12) in the lefthand side of (8), and keep this time all terms up to first order in ε. We obtain

$$-\nabla \cdot (\boldsymbol{\gamma}(\mathbf{x}, R) W^{(0)} - \bar{\mathbf{K}} \mathbf{x} W^{(1)}) + L_{\mathrm{FP}}(R) W^{(0)} = 0 ,$$

which, using the property $x \frac{d^n \delta(x)}{dx} = -n \frac{d^{n-1} \delta(x)}{dx}$ and noticing that $L_{\mathrm{FP}}(R) W^{(0)} = 0$, leads to the solution

$$W^{(1)}(\mathbf{x}, R) = -\bar{W}(R) \sum_i \left(\frac{\delta'(x_i)}{\kappa_{ii}} \gamma_i(\mathbf{x}, R) \prod_{j \neq i} \delta(x_j) \right) . \quad (15)$$

Therefore, up to the first order in ε we can write

$$W(\mathbf{x}, R) = W^{(0)}(\mathbf{x}, R) + \varepsilon W^{(1)}(\mathbf{x}, R)$$

$$= \bar{W}(R) \left[\delta^n(\mathbf{x}) - \varepsilon \sum_i \left(\frac{\delta'(x_i)}{\kappa_{ii}} \gamma_i(\mathbf{x}, R) \prod_{j \neq i} \delta(x_j) \right) \right]$$

$$\simeq \bar{W}(R) \, \delta^n \left(\mathbf{x} - \frac{\varepsilon}{\bar{\mathbf{K}}} \boldsymbol{\gamma}(\mathbf{x}, R) \right) , \quad (16)$$

where the last equality is meant in the sense of distributions, and is also valid up to the first order in ε:

$$\int d\mathbf{x} \, \delta^n \left(\mathbf{x} - \varepsilon \frac{1}{\bar{\mathbf{K}}} \boldsymbol{\gamma}(\mathbf{x}, R) \right) h(\mathbf{x})$$

$$= \int d\mathbf{x} \left[\delta^n(\mathbf{x}) - \varepsilon \sum_i \left(\frac{\delta'(x_i)}{\kappa_{ii}} \gamma_i(\mathbf{x}, R) \prod_{j \neq i} \delta(x_j) \right) \right] h(\mathbf{x})$$

$$+ O(\varepsilon^2) \quad (17)$$

for any test function $h(\mathbf{x})$. We notice here that this first order result coincides with the result of the nonlinear noise filtering approach proposed in [6] and derived dynamically in [8]. In fact, the delta function in Eq. (16) defines the (approximate) steady state solution, determined by the condition

$$\mathbf{z}(\mathbf{x}, R) = \mathbf{x} - \varepsilon \frac{1}{\bar{\mathbf{K}}} \boldsymbol{\gamma}(\mathbf{x}, R) = 0. \quad (18)$$

This solution coincides with the steady state solution of Eqs. (2) and (3) in the limit $\tau \to \infty$ of infinitely slow noise

(i.e., constant R on the time scale of the dynamics of $\mathbf{x}$) and corresponds trivially to a change of variable $R \to \mathbf{x}$.

For instance, in the one-dimensional case, when $\gamma(x, R) = \gamma(R)$ [i.e., $g(x, R) = g(R)$] depends monotonically only on R, one can use the one-dimensional version of Eq. (16), and after integrating over R via

$$p(x) = \int dR \, W(x, R) = \int dR \, \bar{W}(R) \delta \left(x - \frac{\varepsilon}{\kappa} \gamma(x, R) \right)$$

$$= \int dR \, \bar{W}(R) \delta \left(x - \frac{\varepsilon}{\kappa} \gamma(R) \right), \quad (19)$$

obtain the transformation of the probability density function

$$p(x) = \bar{W}(r(x)) |dr(x)/dx|. \quad (20)$$

In Eq. (20), $r(x)$ descends from the one-dimensional version of (18), i.e., $x = \frac{\varepsilon}{\kappa} \gamma(R)$, and therefore $r(x) = R = \gamma^{-1}(\kappa x / \varepsilon) = g^{-1}(kx)$, with $\kappa = k/k_M$ being the only value of the one-dimensional matrix $\bar{\mathbf{K}}$.

This solution corresponds to the nonlinear filtering approach introduced in [6,7], where the Hill-like g function acts as transfer function of the regulator R, assumed as an infinitely slow fluctuating variable, and produces a corresponding distribution of stationary protein concentration x.

However, this perturbative approach can be extended to describe noise with finite τ beyond the first order, providing nontrivial corrections to the resulting dynamics as it is shown in the next subsection, as well as to generic n-dimensional systems by using the implicit functions theorem.

In the multidimensional case, in fact, when $\mathbf{z}(\mathbf{x}, R)$ is a continuously differentiable function $\mathbb{R}^{n+1} \to \mathbb{R}^n$, the implicit functions theorem tells us that a function $r(\mathbf{x}) = R$ exists which is locally invertible so that $\mathbf{z}(r^{-1}(R), R) = 0$. It follows that the probability density for $\mathbf{x}$ can be determined from $\bar{W}(R)$ via a change of variable as

$$p(\mathbf{x}) = \bar{W}(r(\mathbf{x})) |J|. \quad (21)$$

Here, $|J|$ is the Jacobian determinant of the transformation $(R = r(\mathbf{x}), x_2 \cdots x_n) \to (x_1 \cdots x_n)$ which can be obtained, without knowing the explicit form of r, from the function $\mathbf{z}$ in Eq. (18) using the implicit function theorem in n dimensions. Equation (21) generalizes Eq. (20) to the case of n species, with $r(\mathbf{x})$ now depending on all species $x_1 \cdots x_n$ and defined only implicitly via Eq. (18).

C. Second order correction

In order to evaluate the second order correction we first need to evaluate $L_{\mathrm{FP}}(R) W^{(1)}(\mathbf{x}, R)$. From Eq. (15) and the property $L_{\mathrm{FP}}(R) \bar{W}(R) = 0$, it follows that

$$- L_{\mathrm{FP}}(R) W^{(1)}(\mathbf{x}, R)$$

$$= \sum_i \frac{\delta'(x_i)}{\kappa_{ii}} \prod_{k \neq i} \delta(x_k) L_{\mathrm{FP}}(R) [\gamma_i(\mathbf{x}, R) \bar{W}(R)]$$

$$= \sum_i \frac{\delta'(x_i)}{\kappa_{ii}} G_i(\mathbf{x}, R) \prod_{k \neq i} \delta(x_k), \quad (22)$$

where $G_i(\mathbf{x}, R) = L_{FP}(R)[\gamma_i(\mathbf{x}, R)\bar{W}(R)]$ which, using Eq. (6), leads to

$$G_i(\mathbf{x}, R) = -(\mu(R) - 3D\nu(R)\nu'(R))\bar{W}(R)\partial_R\gamma_i(\mathbf{x}, R)$$
$$+ D\nu(R)^2(2\bar{W}'(R)\partial_R\gamma_i(\mathbf{x}, R) + \bar{W}(R)\partial_R^2\gamma_i(\mathbf{x}, R). \tag{23}$$

Having evaluated $W^{(1)}(\mathbf{x}, R)$ we can now repeat the same procedure as done for the first order, i.e., replace expansion (12) in the lefthand side of (8) and this time keep all terms up to second order in ε. We obtain

$$-\nabla \cdot \left(\boldsymbol{\gamma}(\mathbf{x}, R)W^{(1)} - \bar{\mathbf{K}}\mathbf{x}\frac{W^{(2)}}{2}\right) + L_{FP}(R)W^{(1)}(\mathbf{x}, R) = 0, \tag{24}$$

which, using (22), can be rewritten as

$$\nabla \cdot \left(\boldsymbol{\gamma}(\mathbf{x}, R)W^{(1)} - \bar{\mathbf{K}}\mathbf{x}\frac{W^{(2)}}{2}\right)$$
$$+ \sum_i \frac{\delta'(x_i)}{k_{ii}}G_i(\mathbf{x}, R)\prod_{k\neq i}\delta(x_k) = 0. \tag{25}$$

From the structure of this equation and from (16) we deduce that the solution for $W^{(2)}(\mathbf{x}, R)$ must have the following form:

$$W^{(2)}(\mathbf{x}, R) = \bar{W}(R)\sum_{i,q}[(A_{iq}\delta'(x_i) + B_{iq}\delta''(x_i))\delta(x_q)$$
$$+ C_{iq}\delta'(x_i)\delta'(x_q)]\prod_{k\neq i,q}\delta(x_k). \tag{26}$$

The coefficients A_{iq}, B_{iq}, C_{iq} can be obtained by inserting (26) in (25) and result in

$$A_{iq} = \left(-\frac{2}{k_{ii}^2}\bar{G}_i(\mathbf{x}, R) + \tilde{A}_i(\mathbf{x}, R)\right)\delta_{iq}, \tag{27a}$$

with $\quad \frac{\tilde{A}_i(\mathbf{x}, R)}{x_i} - \partial_{x_i}\tilde{A}_i(\mathbf{x}, R) = -\frac{2}{\kappa_{ii}^2}\partial_{x_i}\bar{G}_i(\mathbf{x}, R)$,

$$B_{iq} = \left(\frac{\gamma_i(\mathbf{x}, R)}{\kappa_{ii}}\right)^2\delta_{iq}, \tag{27b}$$

$$C_{iq} = \frac{\gamma_i(\mathbf{x}, R)\gamma_q(\mathbf{x}, R)}{\kappa_{qq}\kappa_{ii}}(1 - \delta_{iq}), \tag{27c}$$

with with $\bar{G}_i = G_i/\bar{W}(R)$. Therefore, we make the following ansatz for the stationary probability density $W(\mathbf{x}, R)$:

$$W(\mathbf{x}, R) = \bar{W}(R)\delta^n\left(\mathbf{x} - \frac{\varepsilon}{\bar{\mathbf{K}}}\boldsymbol{\gamma}(\mathbf{x}, R) - \frac{\varepsilon^2}{2}\mathbf{Q}(\mathbf{x}, R)\right) + O(\varepsilon^3) \tag{28}$$

with

$$\mathbf{Q}(\mathbf{x}, R) = \left(-\frac{2}{\bar{\mathbf{K}^2}}\bar{\mathbf{G}}(\mathbf{x}, R) + \tilde{\mathbf{A}}(\mathbf{x}, R)\right). \tag{29}$$

In fact, the solution given by the ansatz (28) coincides up to second order in ε with the solution derived with the perturbative expansion (12) to the same order, i.e., including Eq. (26) with coefficients (27).

The delta function in Eq. (28) is similar to the one in (16) and therefore, as it was seen for the first order in (18), it
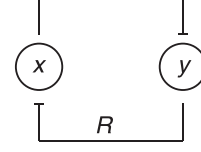

FIG. 1. The toggle switch. The expression level of gene x is downregulated by gene y upon binding with the repressor R. We make the assumption that when active, both genes synthesize the proteins x and y in one single step of combined transcription and translation.

defines the (approximate) steady state solution of Eqs. (2) and (3). This is determined by the condition

$$\mathbf{z}(\mathbf{x}, R) = \mathbf{x} - \frac{\varepsilon}{\bar{\mathbf{K}}}\boldsymbol{\gamma}(\mathbf{x}, R) - \frac{\varepsilon^2}{2}\mathbf{Q}(\mathbf{x}, R) = 0, \tag{30}$$

which in the case that $\nabla \frac{1}{\bar{\mathbf{K}}}\boldsymbol{\gamma}(\mathbf{x}, R) = $ const, in turn implies that $\nabla \frac{1}{\bar{\mathbf{K}}}\mathbf{G}(\mathbf{x}, R) = 0$, and therefore, (neglecting null contributions in distributional sense) $\tilde{\mathbf{A}}(\mathbf{x}, R) = 0$ simplifies to

$$\mathbf{z}(\mathbf{x}, R) = \mathbf{x} - \frac{\varepsilon}{\bar{\mathbf{K}}}\boldsymbol{\gamma}(\mathbf{x}, R) - \varepsilon^2\frac{1}{\bar{\mathbf{K}}^2}\bar{\mathbf{G}}(\mathbf{x}, R) = 0. \tag{31}$$

Note that both $\mathbf{Q}(\mathbf{x}, R)$ and $\bar{\mathbf{G}}(\mathbf{x}, R)$ depend linearly on $\boldsymbol{\gamma} = \tau\mathbf{g}$. Therefore, when multiplied by their expansion factor $\varepsilon^2 = \frac{\tau_M^2}{\tau^2}$, they disappear in the limit $\tau \to \infty$, leaving only the first order correction in Eqs. (28), (30), or (31). This can be verified to apply to all the higher order corrections in the same limit. Therefore Eq. (18), obtained as a first order correction within our approach, is actually an exact solution in such limits, thereby corresponding to the noise filtering approach of [6,7].

Under the assumption that $\mathbf{z}(\mathbf{x}, R)$ is a continuously differentiable function $\mathbb{R}^{n+1} \to \mathbb{R}^n$, the implicit functions theorem can be applied again. As a result, the probability density for $\mathbf{x}$ can be determined from $\bar{W}(R)$ via a change of variable as shown in Eq. (21), where the Jacobian can now be obtained from the function $\mathbf{z}$ in Eqs. (30) or (31).

In the following section, we show how this approach is applied with a concrete example.

IV. THE TOGGLE SWITCH

As an application of our methodology we consider the toggle switch, a positive feedback loop of two genes mutually repressing each other (Fig. 1) [11]. The dynamic equations describing each of the synthesized proteins have the form (1), namely,

$$\frac{dx}{dt} = g(y, R) - k_1 x = \frac{g_1}{1 + \rho_1(Ry)^{\beta_1}} - k_1 x, \tag{32}$$

$$\frac{dy}{dt} = f(x) - k_2 y = \frac{g_2}{1 + \rho_2 x^{\beta_2}} - k_2 y. \tag{33}$$

Here, we have assumed that protein synthesis occurs in one single stage of merged transcription and translation, which, as discussed, is valid when the half-life of mRNA species is much shorter than the half-life of protein species [12]. Further to standard degradation dynamics for x and y proteins, characterized by degradation rates k_1 and k_2 respectively, the

structure of the equations is based on two Hill-like functions, which account for each of the two synthesized proteins to act as a repressor on the other gene. Here parameters $\rho_{1,2}$ are the association constants for the binding of x (y, respectively) to the promoter of the target gene, $g_{1,2}$ are the maximal expression rates of each gene, and $\beta_{1,2}$ are cooperativity parameters representing homodimerization of the relevant protein.

The bifurcation diagram of this system has been analyzed in detail in [11]. The system exhibits deterministic multistable behavior for $\beta_1, \beta_2 > 1$, and it is otherwise monostable. To probe the effect of noise in a situation of deterministic monostability, we therefore set in the following $\beta_1 = \beta_2 = 1$. In this way any transition to bimodal dynamics will only be due to the effect of noise, accounting thereby for the possible emergence of a noise-induced transition in the system.

To explore this further, we have made the assumption that regulation of x happens upon formation of a heterodimer of y with the external (repressive) transcription factor R. Hence, fluctuations in R will correspond to insertion of nonlinear extrinsic noise in the system, and will be amenable to be studied with the methodology developed here.

A. Perturbative calculation

As for the more general case, let us add Eq. (3) to Eqs. (32) and (33) for the toggle switch and let us assume $\tau_x = 1/k_1 > \tau_y = 1/k_2$. We then redefine time as $t \to t/\tau_x$; we obtain

$$\frac{dx}{dt} = \varepsilon \gamma(y, R) - x = \varepsilon \frac{\gamma_1}{1 + \rho_1 R y} - x , \qquad (34)$$

$$\frac{dy}{dt} = \varepsilon f(x) - \kappa y = \varepsilon \frac{\gamma_2}{1 + \rho_2 x} - \kappa y , \qquad (35)$$

$$\frac{dR}{dt} = \varepsilon \mu(R) + \sqrt{\varepsilon D} \nu(R) \xi(t) , \qquad (36)$$

where $\varepsilon = \tau_x/\tau$, $\gamma_{1,2} = g_{1,2}\tau$, and $\kappa = k_2/k_1 \simeq 1$. For simplicity, here we assume $\kappa = 1$.

We also assume that the timescale τ of the dynamics of R satisfies $\tau \gg \tau_x, \tau_y$ and therefore ε is a small parameter. In this way we are assuming that the dynamics of R is slower than that of x and y.

With L_{FP} defined as in Eq. (6), the Fokker-Planck (7) becomes

$$\frac{\partial}{\partial t} W_t(x, y, R) = \left[-\frac{\partial}{\partial x}(\varepsilon \gamma(y, R) - x) - \frac{\partial}{\partial y}(\varepsilon f(x) - y) \right.$$
$$\left. + \varepsilon L_{\mathrm{FP}}(R) \right] W_t(x, y, R), \qquad (37)$$

whose stationary solution satisfies

$$\left[\frac{\partial}{\partial x}(x - \varepsilon \gamma(y, R)) + \frac{\partial}{\partial y}(y - \varepsilon f(x)) + \varepsilon L_{\mathrm{FP}}(R) \right] W(x, y, R)$$
$$= 0. \qquad (38)$$

B. Zeroth order solution

Assuming the perturbative expansion (12), Eq. (13) becomes

$$\left(\frac{\partial}{\partial x} x + \frac{\partial}{\partial y} y \right) W^{(0)}(x, y, R) = 0, \qquad (39)$$

which immediately leads to the zeroth order solution:

$$W^{(0)}(x, y, R) = \delta(x)\delta(y)\bar{W}(R). \qquad (40)$$

C. First order correction

We now proceed with the calculation of first order corrections $W^{(1)}(x, y, R)$. Upon substitution of (12) in (38) and by collecting terms up to first order in ε, we obtain

$$-\frac{\partial}{\partial x}(\gamma(y, R)W^{(0)} - xW^{(1)}) - \frac{\partial}{\partial y}(f(x)W^{(0)} - yW^{(1)})$$
$$+ L_{\mathrm{FP}}(R)W^{(0)} = 0, \qquad (41)$$

which leads to the solution

$$W^{(1)}(x, y, R) = -\bar{W}(R)\big(\delta'(x)\delta(y)\gamma(y, R) + \delta'(y)\delta(x)f(x)\big) \qquad (42)$$

as could be derived directly from Eq. (15). Following our general procedure, we can write to the first order in ε:

$$W(x, y, R) = W^{(0)} + \varepsilon W^{(1)}$$
$$= \bar{W}(R)[(\delta(x) - \varepsilon \gamma(y, R)\delta'(x))\delta(y)$$
$$- \delta'(y)\delta(x)f(x)]$$
$$= \bar{W}(R)\delta(x - \varepsilon \gamma(y, R))\delta(y - \varepsilon f(x)), \qquad (43)$$

analogous to Eq. (16). In fact, if we let

$$\tilde{\delta}(x, y) = (\delta(x) - \varepsilon \gamma(y, R)\delta'(x))\delta(y) - \varepsilon \delta'(y)\delta(x)f(x), \qquad (44)$$

then from the distributional point of view,

$$\iint dx \, dy \, \tilde{\delta}(x, y)h(x, y)$$
$$= h(0, 0) + \varepsilon \left(\gamma(y, R)\frac{\partial}{\partial x} h(0, 0) + f(x)\frac{\partial}{\partial y} h(0, 0) \right)$$
$$= h(\varepsilon \gamma(y, R), \varepsilon f(x)) + O(\varepsilon^2), \qquad (45)$$

and since

$$h(\varepsilon \gamma(y, R), \varepsilon f(x))$$
$$= \iint dx \, dy \, \delta(x - \varepsilon \gamma(y, R))\delta(y - \varepsilon f(x))h(x, y) \qquad (46)$$

for any $h(x, y)$, the last equality in (43) follows. This result coincides at this order with the noise filtering approach proposed in [6] with $p(x)$ stemming from $\bar{W}(R)$ via the change of variables $R \to x$ induced by the (analytically solvable in this case) condition $x - \varepsilon \gamma(\varepsilon f(x), R) = 0$.

D. Second order correction

We can now carry out the evaluation of the second order correction by computing $L_{\mathrm{FP}}(R)W^{(1)}$. Equation (22) becomes

$$L_{\mathrm{FP}}(R)W^{(1)}(x, y, R) = -\delta'(x)\delta(y)L_{\mathrm{FP}}(R)(\gamma(y, R)\bar{W}(R))$$
$$= -\delta'(x)\delta(y)G(y, R), \qquad (47)$$

where

$$
\begin{aligned}
G(y, R) &= L_{\mathrm{FP}}(R)(\gamma(y, R)\bar{W}(R)) \\
&= -(\mu(R) - 3D\nu(R)\nu'(R))\bar{W}(R)\partial_R\gamma(y, R) \\
&\quad + D\nu(R)^2\big(2\bar{W}'(R)\partial_R\gamma(y, R) + \bar{W}(R)\partial_R^2\gamma(y, R)\big).
\end{aligned}
\tag{48}
$$

Equation (24) now becomes

$$
\begin{aligned}
&-\frac{\partial}{\partial x}\Big(\gamma(y, R)W^{(1)} - \frac{x}{2}W^{(2)}\Big) - \frac{\partial}{\partial y}\Big(f(x)W^{(1)} - \frac{y}{2}W^{(2)}\Big) \\
&\quad + L_{\mathrm{FP}}(R)W^{(1)} = 0,
\end{aligned}
\tag{49}
$$

which, using (47), can be rewritten as

$$
\begin{aligned}
&-\frac{\partial}{\partial x}\Big(\gamma(y, R)W^{(1)} - \frac{x}{2}W^{(2)}\Big) - \frac{\partial}{\partial y}\Big(f(x)W^{(1)} - \frac{y}{2}W^{(2)}\Big) \\
&\quad + \delta'(x)\delta(y)G(y, R) = 0.
\end{aligned}
\tag{50}
$$

The form of the solution $W^{(2)}(x, y, R)$ is analogous to (26), namely given by

$$
\begin{aligned}
W^{(2)}(x, y, R) &= \bar{W}(R)(A_{10}\delta'(x)\delta(y) + A_{11}\delta'(x)\delta'(y) \\
&\quad + A_{20}\delta''(x)\delta(y) + A_{02}\delta(x)\delta''(y)).
\end{aligned}
\tag{51}
$$

Inserting this expression in (50), one obtains the following conditions on the coefficients A_{ij}

$$
\begin{aligned}
A_{10} &= -2\bar{G}(y, R), & A_{20} &= (\gamma(y, R))^2, \\
A_{11} &= 2\gamma(y, R)f(x), & A_{02} &= (f(x))^2
\end{aligned}
$$

with $\bar{G} = G/\bar{W}$, as could be derived as well directly from Eq. (27) with the condition $\tilde{\mathbf{A}} = 0$. The ansatz for the steady state solution for the probability density $W(x, y, R)$ that coincides up to the second order in ε with the perturbative evaluation shown above is therefore

$$
\begin{aligned}
W(x, y, R) &= \bar{W}(R)\delta(x - \varepsilon\gamma(y, R) - \varepsilon^2\bar{G}(y, R)) \\
&\quad \times \delta(y - \varepsilon f(x)) + O(\varepsilon^3).
\end{aligned}
\tag{52}
$$

E. Log-normal Noise

The power of the approach that we propose here also relies on the fact that we can assume different types of noise, as long as these can be represented in terms of the Langevin dynamics (3). By following [8], we focus on log-normal noise, which in contrast to Gaussian noise preserves the positivity of parameter values on which it acts. It also describes well extrinsic fluctuations [5,13] and universal behaviors in bacteria and yeast [14,15].

The stationary log-normal distribution

$$
\bar{w}(R) = \frac{1}{R\sqrt{2\pi}}\exp\left[-\frac{1}{2D}\left(\log\frac{R}{\bar{R}} + \frac{D}{2}\right)^2\right]
\tag{53}
$$

can be derived [8] by defining the stochastic process

$$
R(t) = \bar{R}e^{\eta(t)}e^{-D/2},
\tag{54}
$$

where $\eta(t)$ is the standard Ornstein-Uhlenbeck noise,

$$
\frac{d\eta}{dt} = -\frac{\eta}{\tau} + \sqrt{\frac{D}{\tau}}\xi(t),
\tag{55}
$$

with $\xi(t)$ zero average Gaussian white noise, and $\bar{R} = \langle R(t)\rangle$, since $\langle e^{\eta(t)}\rangle = e^{D/2}$. Hence, the stochastic process (54) obeys

$$
\frac{dR}{dt} = -\frac{R}{\tau}\log\left(\frac{R}{\bar{R}} + \frac{D}{2}\right) + \sqrt{\frac{D}{\tau}}R\xi(t),
\tag{56}
$$

which corresponds to (3) with the choice $\mu(R) = -R\log(R/\bar{R} + D/2)$ and $\nu(R) = R$. Equation (10) then allows the derivation of (53).

F. Comparison with numerical simulations

The joint probability density $p(x, y)$ for the protein distributions can be obtained by marginalizing (52) with respect to R:

$$
p(x, y) = \int dR\, W(x, y, R).
\tag{57}
$$

The probability density $p_x(x)$ for the distribution of protein x can then be obtained by marginalizing (57) with respect to y. In this case it is simpler to first integrate in y after replacing (52) in (57); we obtain

$$
\begin{aligned}
p_x(x) &= \int dR\bar{W}(R)\delta(x - \varepsilon\gamma(\varepsilon f(x), R) - \varepsilon^2\bar{G}(\varepsilon f(x), R)) \\
&= \int dR\bar{W}(R)\delta(z(x, R)),
\end{aligned}
\tag{58}
$$

and further integration with respect to R amounts to a change of variable $R = r(x)$ where the function $r(x)$ is defined implicitly by the equation

$$
z(x, R) = x - \varepsilon\gamma(\varepsilon f(x), R) - \varepsilon^2\bar{G}(\varepsilon f(x), R) = 0.
\tag{59}
$$

We notice that in this particular case we can also easily develop the third order correction. This is obtained following the same procedure used for the second order (see Appendix for details) leading to the following modified expression for the steady state solution for the probability density $W(x, y, R)$:

$$
\begin{aligned}
W(x, y, R) &= \bar{W}(R)\delta(x - \varepsilon\gamma(y, R) - \varepsilon^2\bar{G}(y, R) + \\
&\quad - \varepsilon^3\bar{G}_3(x, y, R))\delta(y - \varepsilon f(x)) + O(\varepsilon^4).
\end{aligned}
\tag{60}
$$

Here, we have used (see Appendix):

$$
\bar{G}_3(x, y, R) = \bar{G}_M(y, R) + \frac{\bar{G}_2(y, R)}{2x},
\tag{61}
$$

and therefore, by integrating over y we obtain

$$
p_x(x) = \int dR\bar{W}(R)\delta(z(x, R))
\tag{62}
$$

with

$$
\begin{aligned}
z(x, R) &= x - \varepsilon\gamma(\varepsilon f(x), R) - \varepsilon^2\bar{G}(\varepsilon f(x), R) + \\
&\quad - \varepsilon^3\bar{G}_3(x, \varepsilon f(x), R),
\end{aligned}
\tag{63}
$$

which replaces (59).

According to the implicit functions theorem, from the condition $z(x, R) = 0$, it follows that a function $r(x) = R$ exists such that

$$
p_x(x) = \bar{W}(r(x))\left|\frac{dr}{dx}\right|,
\tag{64}
$$

where dr/dx can be computed by applying the implicit functions theorem either to (59) for the calculation up to the second order, or to (63) for the third order. Hence,

$$\frac{dr}{dx} = -\frac{\partial_x z(x, R)}{\partial_R z(x, R)}, \qquad (65)$$

which allows us to explicitly evaluate $p_x(x)$ for the toggle switch.

Finally, in order to derive

$$p_y(y) = \int dx \int dR\, W(x, y, R), \qquad (66)$$

it suffices to notice that at steady state we have exactly $y = \varepsilon f(x)$, and therefore, $p_y(y)$ in this case can be derived directly from $p_x(x)$ via a change of variable as

$$p_y(y) = \frac{1}{\varepsilon} p_x\left(f^{-1}\left(\frac{y}{\varepsilon}\right)\right)\left|\frac{df^{-1}}{dy}\right|$$

$$= \frac{p_x(f^{-1}(y/\varepsilon))}{\varepsilon |f'(x = f^{-1}(y/\varepsilon)|}. \qquad (67)$$

Equations (64) and (67) so constructed can then be compared at any given perturbative order with direct stochastic simulations, in order to highlight the predictive power of the theoretical approach developed here.

We perform direct stochastic simulations of fluctuations in R by using the standard approach to generate an Ornstein-Uhlenbeck process, converting it via Eq. (54) into log-normal noise and then by performing ordinary integration of Eqs. (32) and (33). The probability densities are obtained by averaging a single trajectory over long times. Parameters are set so as to reproduce biologically relevant conditions. In all simulations and in the analytical solution, we fix the maximal gene expression rates at the values $g_1 = g_2 = g_0 = 10^{-2} s^{-1}$, which is representative of typical gene expression in bacteria [16], and the degradation rates $k_1 = k_2 = 10^{-4} s^{-1}$, which are representative of proteins' half-life in the range of two hours. We also fix the dissociation constants of transcription factors to DNA $\rho_1 = 10 nM^{-1}$ and $\rho_2 = 1 nM^{-1}$, and $\bar{R} = 0.87 nM$ [8].

We then focus on two different values of ε in order to explore the contribution of different perturbative orders. Since we assume $k = 10^{-4}$ s^{-1}, choosing $\tau = 10^7$ s corresponds to assume $\varepsilon = 0.001$, while the choice $\tau = 10^5$ s will reflect a larger value of ε, namely $\varepsilon = 0.1$. It is this second regime, when the timescale separation between dynamics of x and of the R fluctuations is less pronounced, that we predict that higher order terms in the perturbative expansion will become relevant.

In Fig. 2 we show the comparison between perturbative expansion and stochastic simulations for $\varepsilon = 0.001$. The theoretical analysis predicts bimodality for the concentration x, and this is confirmed excellently by the numerical result. In this case the first order solution via Eq. (43), second order solution via Eq. (59), and third order solution via Eq. (63) practically coincide.

In Fig. 3 (top panel) we show a particular case where using the first order solution in Eq. (43) leads erroneously to conclude that bimodality emerges, while including second and
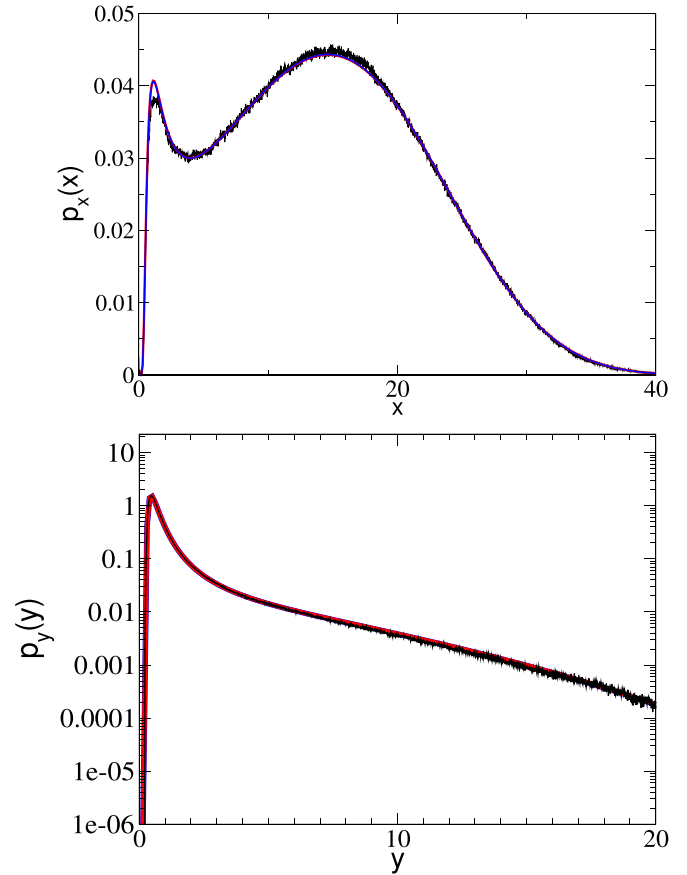

FIG. 2. Probability distributions $p_x(x)$ (top panel) and $p_y(y)$ (bottom panel) as given by numerical simulation (black line) compared to theoretical solution (64) for p_x and (67) for p_y, respectively, at the first order (red line), second order via Eq. (59) (blue line), and third order via Eq. (63) (blue continuous line). These solutions are indistinguishable. Same colors and line styles are used in both panels. Here $\alpha = 10$, $\bar{R} = 0.87 nM$, $g_0 = 0.01$, $\tau = 10^7$, $\varepsilon = 0.001$, $D = 0.012$.

third order correction gives a quantitatively and qualitatively correct description of the properties of the density function, correctly predicting a unimodal distribution for both x and y protein concentrations.

G. Noise-induced transitions: Partial bimodality

The analysis and simulations shown in Fig. 2 highlight the emergence of a particular phenomenon, which we name "partial bimodality," whereby bimodal behavior emerges for the protein concentration x, but not for y. This is clearly unexpected in standard toggle switches affected by intrinsic noise only, as in this case the bimodality would instead simultaneously affect both genes involved.

In order to obtain further insight into these dynamics and clarify the nature of the noise-induced transition identified here, we perform simulations and solve the corresponding first order dynamics for three different values of noise intensities.

In Fig. 4 we show simulations and analytical predictions for the toggle switch in a regime where at low noise intensity ($D = 0.01$) the dynamics are unimodal. In this case
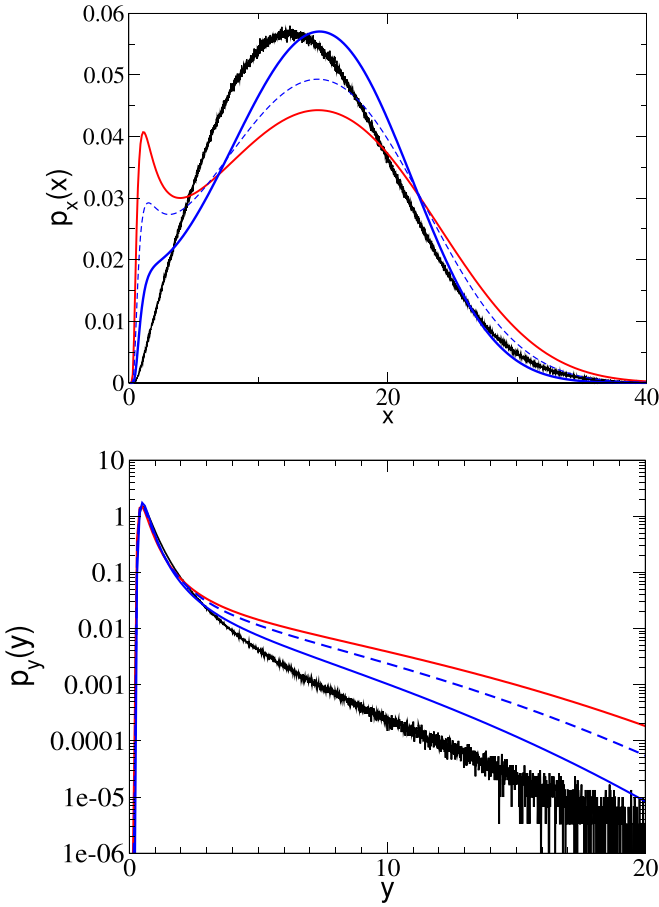
FIG. 3. Probability distributions $p_x(x)$(top panel) and $p_y(y)$ (bottom panel) as given by numerical simulation (black line) compared to theoretical solution (64) for p_x and (67) for p_y, respectively, at the first order (red lines), second order via Eq. (59) (blue dashed lines), and third order via Eq. (63) (blue continuous lines). Same colors and line styles are used in both panels. Including the third order correctly shows only one mode for $p_x(x)$ (blue continuous line, top panel). Here $\alpha = 10$, $\bar{R} = 0.87nM$, $g_0 = 0.01$, $\tau = 10^5$, $\varepsilon = 0.1$, $D = 0.012$.

we observe that increasing the noise intensity up to $D = 10$, bimodal behavior emerges for both genes. However, interestingly, we also observe that for intermediate values of noise, $D \simeq 0.1$, a regime exists where bimodality emerges only for one gene and not the other.

For this case we show two exemplary trajectories for variables x and y in the upper panel of Fig. 5. While the variable x makes complete transitions between a high value around the second mode ($x = 32.5$) and low values close to the first mode ($x = 0.24$), the variable y remains more consistently around $y = 0.2$, as testified by the higher peak of its distribution at the origin as compared to that of p_x (see inset in the middle of Fig. 4). In correspondence of transitions of the x variable to the low state, excursions to slightly larger values of y are permitted, but their amplitude is not large enough to exit the basin of attraction of the low y state. In the bottom panel of Fig. 5 instead, the case where bimodality occurs for both species is shown. These trajectories correspond to the bimodal distributions in the bottom panel of Fig. 4 with modes
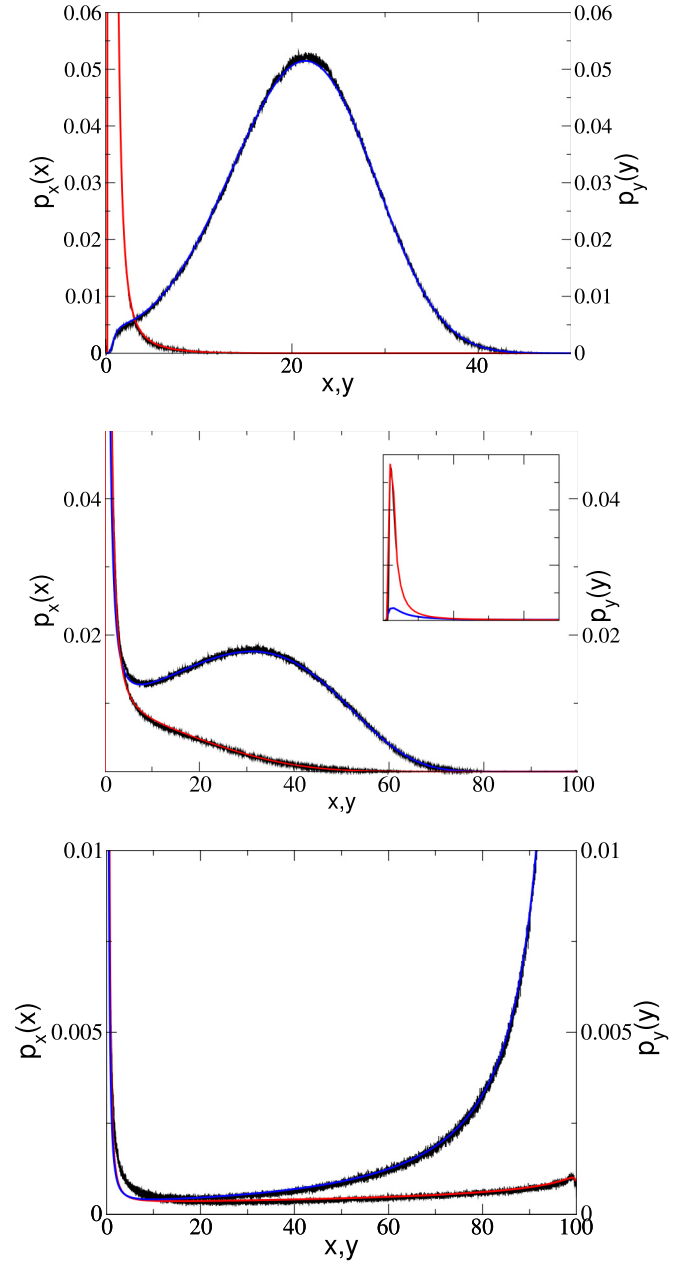

FIG. 4. Noise-induced transitions and partial bimodality in the toggle switch. The distributions $p_x(x)$ and $p_y(y)$ as given by numerical simulations (black lines) compared to analytical solutions (43) (red line and blue line for $p_y(y)$ and $p_x(x)$, respectively). Here $\alpha = 10$, $\bar{R} = 0.8nM$, $g_0 = 0.01$, $\tau = 10^7$, $\varepsilon = 0.001$, from top to bottom $D = 0.01, 0.1, 10$.

in $x, y = 0$ and in $x = 100, y = 99$. Since p_x has a higher peak in $x = 100$ and p_y has a higher peak in $y = 0$, these species spend most of their time around these values with sporadic transitions to values around $x = 0$ and $y = 99$, respectively. The amplitude of the fluctuations of the y variable in this case allows the variable y to exit the basin of attraction of its high state.

We attribute the different noise amplitudes experienced by the variables x and y to a noise propagation effect, whereby the fluctuations of the external parameter R are perceived
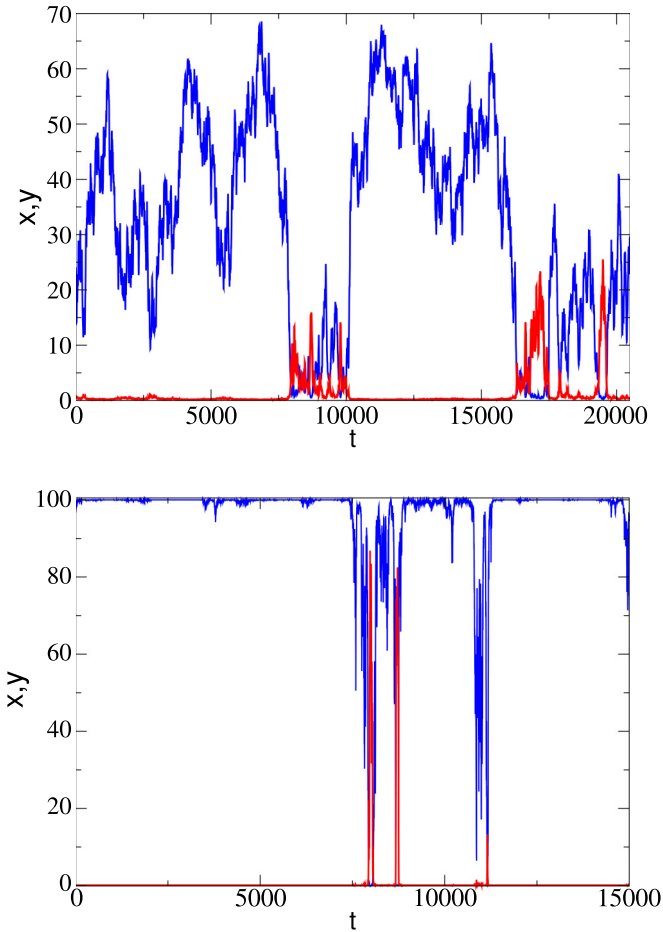
FIG. 5. Exemplary trajectories for the variable x (blue line) and y (red line) for parameters $\alpha = 10$, $\bar{R} = 0.8$nM, $g_0 = 0.01$, $\tau = 10^7$, $\varepsilon = 0.001$, and $D = 0.1$ (upper panel, corresponding to distributions with partial bimodality in middle panel in Fig. 4) and $D = 10$ (lower panel, corresponding to bimodal distributions in bottom panel in Fig. 4).

differently by the two genes. In fact, the variance of the distribution of x is consistently larger of the variance of the distribution of y, for all input noises on R (parametrized by D), as shown in Table I.

While x is under the direct influence of fluctuations of R, gene y feels those fluctuations, propagated through the repressive action from x, at a lower level. This means that while gene x can be above threshold for transitions to bimodal behavior, gene y may still remain below the threshold, and therefore may still exhibit unimodal behavior. This mechanism is in agreement with the result of some extensive studies performed

TABLE I. Variances of x and y distributions corresponding to the probability distributions shown in Fig. 4. Parameter values are set as in Fig. 4.

D	0.01	0.1	10
σ_x	56.62	325.31	590.67
σ_y	0.88	58.57	191.17

in [17], which demonstrated that positive feedback loops limit the range of fluctuations of the target gene.

V. CONCLUSIONS

In this paper we have introduced an efficient methodology to deal with nonlinear extrinsic noise when the correlation time of the fluctuations is large. The derivation proposed here is based on a timescale expansion of the master equation and reproduces to the first order of the phenomenological approach based on the nonlinear noise filtering introduced in [6]. Importantly, our perturbative approach allows one to go beyond the dynamical derivation presented in [8] and obtain correct predictions even in a regime when the timescales characterizing fluctuations and gene dynamics are comparable, with the lowest perturbative orders failing to capture the qualitative dynamical features of the system.

We also find that the toggle switch is capable of a partial bimodality, in that while one of the two participating genes shows full bimodal behavior, the other gene still preserves unimodality. This feature is intermediate in terms of noise intensities to the two extreme cases of no bimodality for any of the two genes, to full bimodality for both of them. As a followup to this observation, it will be interesting to analyze the effect of simultaneous intrinsic and extrinsic noise on the system. We expect that the interplay between intrinsic and extrinsic noise would provoke even richer dynamics, in which the balance between noise intensities will be crucial in determining the structure of the phase space.

The mechanism for noise-induced transitions described here and the methodology introduced are promising and worthy of further analysis and experimental validation. They will further our understanding of the fundamental and active role played by noise in biological systems.

All details of the mathematical models developed and the data deriving from their simulation are available in the article.

APPENDIX: EVALUATION OF THIRD ORDER PERTURBATIVE CORRECTION FOR THE TOGGLE SWITCH

We can evaluate the third order correction in the perturbative expansion just by carrying on the procedure introduced in the main text to include the third order. We start with $W(x,R) = W^{(0)} + \varepsilon W^{(1)} + \varepsilon^2 W^{(2)}/2 + \varepsilon^3 W^{(3)}/6$ and this leads to

$$-\frac{\partial}{\partial x}\left(\gamma(y,R)W^{(2)} - \frac{xW^{(3)}}{3}\right)$$
$$-\frac{\partial}{\partial y}\left(f(x)W^{(2)} - \frac{yW^{(3)}}{3}\right) + L_{\text{FP}}(R)W^{(2)} = 0, \quad \text{(A1)}$$

where we know $W^{(2)}(x, y, R)$ from previous calculations. Therefore, we can evaluate $L_{\text{FP}}(R)W^{(2)}$ and we obtain

$$
\begin{aligned}
L_{\text{FP}}(R)&W^{(2)}(x, y, R) \\
&= G_2(y, R)\delta''(x)\delta(y) \\
&\quad + 2f(x)G(y, R)\delta'(x)\delta'(y) - 2G_M(y, R)\delta'(x)\delta(y), \quad \text{(A2)}
\end{aligned}
$$

with

$$
\begin{aligned}
G_2(y, R) &= L_{\text{FP}}(\gamma^2(y, R)\bar{W}(R)) \\
&= 2(3D\nu(R)\nu'(R) - \mu(R))\bar{W}(R)\gamma(y, R)\partial_R\gamma(y, R) \\
&\quad + D\nu(R)^2\big[4\bar{W}'(R)\gamma(y, R)\partial_R\gamma(y, R) \\
&\quad + 2\bar{W}(R)\big(\gamma(y, R)\partial_R^2\gamma(y, R) + (\partial_R\gamma(y, R))^2\big)\big],
\end{aligned}
$$
$$
\text{(A3)}
$$

where it can be shown that

$$
\begin{aligned}
G_M(y, R) &= L_{\text{FP}}(R)(G(y, R)) \\
&= \alpha(R)G(y, R) + \sigma(R)\partial_R G(y, R) + \gamma(R)\partial_R^2 G(y, R),
\end{aligned}
$$
$$
\text{(A4)}
$$

with

$$
\alpha(R) = -(\mu'(R) - D\nu'(R)^2 - D\nu(R)\nu''(R)), \quad \text{(A5)}
$$

$$
\sigma(R) = -(\mu(R) - 3D\nu(R)\nu'(R)), \quad \text{(A6)}
$$

$$
\gamma(R) = D\nu^2(R). \quad \text{(A7)}
$$

From Eq. (A1) we can therefore expect $W_3(x, R)$ of the form

$$
\begin{aligned}
W_3(x, y, R) &= \bar{W}(R)(A_{10}\delta'(x)\delta(y) + A_{11}\delta'(x)\delta'(y) \\
&\quad + A_{12}\delta'(x)\delta''(y) + A_{20}\delta''(x)\delta(y) + A_{21}\delta''(x)\delta'(y) + \\
&\quad + A_{30}\delta'''(x)\delta(y) + A_{03}\delta(x)\delta'''(y)).
\end{aligned}
$$
$$
\text{(A8)}
$$

Replacing Eq. (A8) in (A1) and solving for A_{ij} leads to

$$
A_{10} = -6\bar{G}_M(y, R) - (3/x)\bar{G}_2(y, R), \quad \text{(A9)}
$$

$$
A_{11} = 6\bar{G}(y, R)f(x), \quad \text{(A10)}
$$

$$
A_{12} = 3\gamma(y, R)f^2(x), \quad \text{(A11)}
$$

$$
A_{20} = 6\bar{G}(y, R)\gamma(y, R), \quad \text{(A12)}
$$

$$
A_{21} = -3\gamma^2(y, R)f(x), \quad \text{(A13)}
$$

$$
A_{30} = -\gamma^3(y, R), \quad \text{(A14)}
$$

$$
A_{03} = -f^3(x). \quad \text{(A15)}
$$

[1] W. Horsthemke and R. Lefever, *Noise-Induced Transitions*, (Springer 1984).

[2] C. W. Gardiner, *Handbook of stochastic methods*, Springer (1985).

[3] L. Cai, N. Friedman, and X. S. Xie, Nature (London) **440**, 358 (2006).

[4] M. Bengtsson, A. Stahlberg, P. Rorsman and Mikael Kubista, Genome Res. **15**, 1388 (2005).

[5] N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz, Science **307**, 1962 (2005).

[6] A. Ochab-Marcinek and M. Tabaka, Proc. Natl. Acad. Sci. USA **107**, 22096 (2010).

[7] A. Ochab-Marcinek, J. Jedrak, M. Tabaka, Phys. Chem. Chem. Phys. **19**, 22580 (2017).

[8] G. Aquino and A. Rocco, Math. Biosci. Eng. **17**, 6993 (2020).

[9] L. Arnold, W. Horsthemke, and R. Lefever, Z. Phys. B **29**, 367 (1978).

[10] L. E. Reichl and W. C. Schieve, *Instabilities, bifurcation, and Fluctuations in Chemical Systems*, University of Texas Press, Austin, 1982.

[11] T. S. Gardner, C. R. Cantor, and J. J. Collins, Nature (London) **403**, 339 (2000).

[12] N. Rosenfeld, M. B. Elowitz, and U. Alon, J. Mol. Biol. **323**, 785 (2002).

[13] V. Shahrezaei, J. F. Ollivier, and P. S. Swain, Mol. Sys. Bio. **4**, 196 (2008).

[14] H. Salman, N. Brenner, C.K. Tung, N. Elyahu, E. Stolovicki, L. Moore, A. Libchaber, and E. Braun, Phys. Rev. Lett. **108**, 238105 (2012).

[15] N. Brenner, C. M. Newman, D. Osmanovic, Y. Rabin, H. Salman, and D. L. Stein, Phys. Rev. E **92**, 042713 (2015).

[16] M. Thattai and A. van Oudenaarden, Proc. Natl. Acad. Sci. USA **98**, 8614 (2001).

[17] G. Hornung and N. Barkai, PLOS Comp. Biol. **4**, e8 (2008).