



HAL
open science

Sensitivity estimation for stochastic models of biochemical reaction networks in the presence of extrinsic variability

Jakob Ruess, Heinz Koepl, Christoph Zechner

► **To cite this version:**

Jakob Ruess, Heinz Koepl, Christoph Zechner. Sensitivity estimation for stochastic models of biochemical reaction networks in the presence of extrinsic variability. *Journal of Chemical Physics*, 2017, 146 (12), pp.124122. 10.1063/1.4978940 . hal-01547484

HAL Id: hal-01547484

<https://hal.science/hal-01547484>

Submitted on 26 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sensitivity Estimation for Stochastic Models of Biochemical Reaction Networks in the Presence of Extrinsic Variability

Jakob Ruess,^{1,2, a)} Heinz Koepl,³ and Christoph Zechner^{4, b)}

¹⁾*Inria Saclay - Ile-de-France, 91120 Palaiseau, France*

²⁾*Institut Pasteur, 75724 Paris Cedex 15, France*

³⁾*TU Darmstadt, Department of Electrical Engineering and Information Technology, 64283 Darmstadt, Germany*

⁴⁾*Max Planck Institute of Molecular Cell Biology and Genetics and Center for Systems Biology Dresden, 01307 Dresden, Germany*

Determining the sensitivity of certain system states or outputs to variations in parameters facilitates our understanding of the inner working of that system and is an essential design tool for the de-novo construction of robust systems. In cell biology, the output of interest is often the response of a certain reaction network to some input (e.g. stressors or nutrients) and one aims to quantify the sensitivity of this response in the presence of parameter heterogeneity. We argue that for such applications, parametric sensitivities in their standard form do not paint a complete picture of a system's robustness since one assumes that all cells in the population have the same parameters and are perturbed in the same way. Here, we consider stochastic reaction networks in which the parameters are randomly distributed over the population and propose a new sensitivity index that captures the robustness of system outputs upon changes in the characteristics of the parameter distribution, rather than the parameters themselves. Subsequently, we make use of Girsanov's likelihood ratio method to construct a Monte Carlo estimator of this sensitivity index. However, it turns out that this estimator has an exceedingly large variance. To overcome this problem, we propose a novel estimation algorithm that makes use of a marginalization of the path distribution of stochastic reaction networks and leads to Rao-Blackwellized estimators with reduced variance.

I. INTRODUCTION.

Many models of biochemical reaction networks have been found to be robust in the sense that their outputs are relatively insensitive to local changes in most of the model parameters¹. Contrary to this, results in synthetic biology have shown that building molecular circuits that robustly perform a certain function in a population of cells is a major challenge². Possibly the main reason for this discrepancy is that cell populations are heterogeneous. Even genetically identical cells may show large variations in key quantities such as the number of ribosomes that are present in the cell. On the other hand, on the modeling side, the traditional approach has been to write down one model (e.g. as a system of differential equations or as a stochastic process) and to assume that this model is representative of every cell in the population. Accordingly, it is not very surprising that local parameter sensitivities computed from such models cannot provide a full picture of robustness in heterogeneous populations. Alternatively, one could determine global parameter sensitivities and search for cases where the system is robust over the entire space of possible parameters, or at least over some reasonably large region in parameter space. However, global sensitivity analysis such as Sobol's decomposition³ or Morris screening⁴ can

be computationally very demanding for complex models, and one may argue that the existence of globally robust systems does not seem very likely, except maybe in some special cases⁵.

Recently, it has been proposed to incorporate heterogeneity in models by allowing the rate constants of chemical reactions to be random variables with some distribution over the population^{6,7}. It has been shown in several studies⁸⁻¹⁰ that this allows one to capture the experimentally observed variability between cells, but the analysis of the resulting models may be challenging. The goal of this paper is to make a first step towards establishing sensitivity and robustness analysis for models that include both intrinsic noise coming from the randomness of reaction events as well as heterogeneity of the population. Specifically, we focus on stochastic reaction networks described by continuous-time Markov chain (CTMC) models with random reaction rate parameters. We propose a Monte Carlo algorithm that can be used to estimate the sensitivities of arbitrary functionals on the path space of the stochastic process with respect to hyperparameters of the model that describe the population distribution of the random reaction rate parameters.

For classical CTMC models without randomness in model parameters, sensitivity estimation has been an active research topic in the past years¹¹⁻¹⁶. Here, we start by providing an extension of the sensitivity estimation method proposed by Plyasunov and Arkin¹⁷ to CTMC models where the model parameters are random. We investigate the performance of this method on simple toy models and show that the provided estimates have extremely large variances. Motivated by these results, we propose a novel method that is specifically geared to

^{a)}Contributed equally to this work; Electronic mail: jakob.ruess@inria.fr

^{b)}Contributed equally to this work; Electronic mail: zechner@mpi-cbg.de

wards models with random parameters. In particular, we combine the Girsanov method with the recently introduced marginal process framework for stochastic reaction networks in random environments¹⁸ and show that this allows one to construct unbiased Rao-Blackwellized estimators^{19,20} that have a much lower variance.

II. STOCHASTIC REACTION NETWORKS IN RANDOM ENVIRONMENTS AND THEIR SENSITIVITIES.

A. Mathematical modeling.

We describe the time-evolution of a stochastic reaction network by a continuous-time Markov chain (CTMC) X consisting of M molecular species that interact with each other through N reaction channels. Each reaction i is associated with a stoichiometric change vector $\nu_i \in \mathbb{R}^M$ and reaction propensity $h_i(x, z_i) = z_i g_i(x)$ with z_i a kinetic rate constant and $g_i(x)$ a polynomial determined by the law of mass-action. The stochastic state $X(t)$ of the reaction network can then be described by an integral equation of the form²¹

$$X(t) = X(0) + \sum_{i=1}^K Y_i \left(z_i \int_0^t g_i(X(s)) ds \right) \nu_i \quad (1)$$

with Y_i as a time-transformed unit Poisson process corresponding to the occurrences of reaction i up to time t .

In practice, some of the reaction propensities of X may depend on so-called extrinsic factors that are not explicitly included in the model. Variations of such factors will cause an additional source of heterogeneity – commonly known as extrinsic variability. For instance, if protein translation is modeled by a one-step reaction, the rate of this reaction will be different in each cell depending on the number of available ribosomes, the ATP abundance and so forth. To account for extrinsic variability, we resort to a mixed-effects modeling approach, meaning that we allow some of the kinetic parameters to be randomly distributed across cells. For simplicity, we assume that these parameters fluctuate slowly compared to the timescale of X , such that we can treat them as static random variables. We denote these random variables by $(Z_1, \dots, Z_N) \mid (A = a)$ and their joint probability distribution by $p(\cdot \mid a)$. The goal of this paper will be to calculate sensitivities with respect to the hyperparameters a . The distribution $p(\cdot \mid a)$ determines how the extrinsic variables $Z = (Z_1, \dots, Z_N) \in \mathcal{Z}$ are distributed across cells and a contains the hyperparameters of interest, which may, but do not have to be, a full parametrization of the distribution. For instance, if there is only one extrinsic variable with a log-normal distribution, a could comprise location and scale parameter of that distribution. For the sake of simplicity, throughout this manuscript we will use for a the collection of means μ_i and variances σ_i^2 of Z_i , i.e. $a = (\mu_1, \dots, \mu_N, \sigma_1^2, \dots, \sigma_N^2)$.

Overall, a cell j from a heterogeneous population can be described by a randomly drawn parameterization $z^j \sim p(\cdot \mid a)$ together with a conditional Markov chain $X \mid (Z = z^j)$ satisfying

$$X^j(t) = X^j(0) + \sum_{i=1}^K Y_i \left(z_i^j \int_0^t g_i(X^j(s)) ds \right) \nu_i. \quad (2)$$

B. Sensitivity Estimation

Local sensitivity analysis provides a common and intuitive approach to assess a network's susceptibility to parameter variations. A sensitivity index for some parameter z_i can be defined as

$$\frac{\partial}{\partial z_i} y(z) = \frac{\partial}{\partial z_i} \mathbb{E}[f(X)], \quad (3)$$

with $X \in \mathcal{X}$ as a complete sample path of $X(t)$ on an interval $[0, T]$ and f as some functional of that path. If we had $f(X) := X(T)$, for instance, then $y(z)$ would correspond to the mean of $X(t)$ at time T .

Sensitivity indices like the one from (3) can capture how a certain property of a population (defined by y) changes upon small perturbations in the parameter z_i . Implicitly, this assumes that all cells in a population are equipped with the *same* parameter z_i and that these parameters are perturbed in the same way. This appears to be a suitable proxy for robustness in cases where z_i is a parameter that applies to all cells in a population, such as the concentration of an external signal. However, this does not hold true if we want to predict how a network behaves when it is exposed to extrinsic variability. In that case, we need a way to characterize the behavior of the network upon changes in the characteristics of the distribution of Z rather than its particular realization z . Under our model from (2), this would correspond to computing the sensitivities with respect to the hyperparameters μ_i and σ_i^2 , i.e.,

$$S_i = \left(\frac{\partial}{\partial \mu_i} \xi(a), \frac{\partial}{\partial \sigma_i^2} \xi(a) \right) \quad (4)$$

with $\xi(a) = \mathbb{E}[y(Z)] = \mathbb{E}[\mathbb{E}[f(X) \mid Z]]$ for all $i = 1, \dots, N$. This two-dimensional index allows us to quantify the sensitivity of $\xi(a)$ upon *shifting* and *stretching* the distribution over Z , respectively (Fig. 1a). We want to stress that $\partial/\partial \mu_i \xi(a)$ and $\partial/\partial \sigma_i^2 \xi(a)$ are different and correspondingly, may lead to opposing conclusions about a network's robustness. For instance, a circuit that is found robust under changes in μ_i , may show high sensitivity with respect to changes in σ_i^2 and vice versa.

Note that eq. (4) can be understood as a generalization of the conventional sensitivity index (3) to random parameters. Informally, this can be seen by letting $\sigma_i^2 \rightarrow 0$, in which case (4) approaches (3), i.e., $\partial/\partial \mu_i \xi(a) \rightarrow \partial/\partial z_i y(z)$.

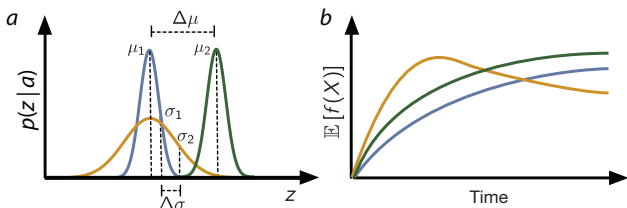


FIG. 1. Quantifying robustness with respect to extrinsic variations. (a) Conventional parameter sensitivities are based on the implicit assumption that each cell undergoes the same parameter change. For small $\sigma_1 \rightarrow 0$, this corresponds to shifting the mean of the probability distribution over extrinsic factors by $\Delta\mu$. This is in contrast to the second scenario that corresponds to a transition $\sigma_1 \rightarrow \sigma_2$, i.e., the variability changes from very small to large by $\Delta\sigma$. (b) The two types of perturbations $\Delta\mu$ and $\Delta\sigma$ can have opposing effects and therefore lead to contradicting conclusions about a network's robustness.

C. The Girsanov method for random parameters

Many reaction networks of interest are complex and involve nonlinearities such that (4) becomes analytically intractable. We therefore seek efficient simulation algorithms that allow us to approximate (4) as an n -sample Monte Carlo estimate. One of these simulation algorithms is known as the Girsanov likelihood ratio (GLR) method¹⁷. For our model, a Girsanov estimator can be obtained by noting that

$$\begin{aligned} \frac{\partial}{\partial a_k} \xi(a) &= \int \int f(x) p(x|z) \frac{\partial}{\partial a_k} p(z|a) dz dx \\ &= \int \int f(x) p(x|z) \left(\frac{\partial}{\partial a_k} \ln p(z|a) \right) p(z|a) dz dx \\ &= \mathbb{E}[f(X)W_k(Z)] \end{aligned} \quad (5)$$

with $W_k(z) = \partial/\partial a_k \ln p(z|a)$. Correspondingly, we can compute an n -sample sensitivity estimate as

$$S_k(a) = \frac{1}{n} \sum_{i=1}^n f(x^i) W_k(z^i) \quad (6)$$

with (x^i, z^i) as i.i.d. samples drawn from the joint distribution $p(x, z) = p(x|z)p(z|a)$. Unfortunately, it turns out that estimators of the form (6) often have a very large variance such that a reasonable accuracy can be achieved only if n is very large. In the following, we develop an improved estimator that is guaranteed to achieve a lower variance than the original one from (6).

D. A Rao-Blackwellized sensitivity estimator

One reason for the exceedingly high variance of (6) is that the Monte Carlo averaging is performed over the augmented sampling space $\mathcal{X} \times \mathcal{Z}$. To obtain variance

reduction, we will construct in the following a marginalized estimator that requires samples only from the path space \mathcal{X} and not \mathcal{Z} . To that end, we rewrite (5) such that

$$\begin{aligned} \frac{\partial}{\partial a_k} \xi(a) &= \int f(x) \frac{\partial}{\partial a_k} \left(\int p(x|z) p(z|a) dz \right) dx \\ &= \int f(x) \left(\frac{\partial}{\partial a_k} \ln p(x|a) \right) p(x|a) dx \quad (7) \\ &= \mathbb{E}[f(X)\widehat{W}_k(X)] \end{aligned}$$

with $\widehat{W}_k(x) = \partial/\partial a_k \ln p(x|a)$. The corresponding Monte Carlo estimator reads

$$\widehat{S}_k(a) = \frac{1}{n} \sum_{i=1}^n f(x^i) \widehat{W}_k(x^i), \quad (8)$$

with x^i as i.i.d. samples drawn from the marginal path distribution $p(x|a)$. Intuitively, one would argue that the marginalized estimator $\widehat{S}_k(a)$ achieves an improved performance since the dimension of the sampling space is reduced. Indeed, variance reduction can be proven for this estimator leading to the following theorem.

Theorem 1. *Let $S_k(a)$ and $\widehat{S}_k(a)$ be sensitivity estimators defined by (6) and (8), respectively. Then, for any finite n we have that*

$$\text{Var}[\widehat{S}_k(a)] \leq \text{Var}[S_k(a)]. \quad (9)$$

Proof. See Appendix A. \square

What remains to be addressed is how one can calculate the weights $\widehat{W}_k(x)$ for a given realization x . To this end, we build on the recently developed marginal process framework^{8,18,22}, which allows one to integrate a CTMC (2) with respect to a random (and possibly time-varying) parametrization Z . The resulting marginal process can be shown to be non-Markovian with propensities depending on the history of that process. Most importantly, it is the process that admits the marginal path distribution $p(x|a)$ that appears in (7) and in the weights $\widehat{W}_k(x)$. However, the exact analytical construction of that process and its path distribution is possible only under certain distributional assumptions about $p(z|a)$. Several examples of such assumptions are provided in^{18,22} but in the present work, we restrict ourselves to the case where Z consists of independent and Gamma-distributed random variables, i.e. $p(z|a) = \prod_{i=1}^N \mathcal{G}(z_i | \alpha_i, \beta_i)$. The shape and inverse scale parameters α_i and β_i can be expressed in terms of the means and variances of z_i , i.e., $\alpha_i = \mu_i^2/\sigma_i^2$ and $\beta_i = \mu_i/\sigma_i^2$. Under these assumptions, one can show²² that the likelihood of observing a path x

for some a becomes

$$p(x | a) = \int p(x | z)p(z | a)dz \\ \propto \prod_{i=1}^N \frac{\Gamma\left(R_i + \frac{\mu_i^2}{\sigma_i^2}\right)}{\Gamma\left(\frac{\mu_i^2}{\sigma_i^2}\right)} \left(G_i + \frac{\mu_i}{\sigma_i}\right)^{-\frac{\mu_i^2}{\sigma_i^2} - R_i} \left(\frac{\mu_i}{\sigma_i}\right)^{\frac{\mu_i^2}{\sigma_i^2}}, \quad (10)$$

with R_i as the number of occurrences of reaction i in x and $G_i = \int_0^T g_i(x(s))ds$. Differentiating the logarithm of (10) with respect to μ_i and σ_i^2 yields the desired weights $\widehat{W}_k(x)$ (see Appendix B).

As indicated earlier, the Monte Carlo estimate $\widehat{S}_k(a)$ is obtained by averaging over sample paths from the marginal distribution $p(x | a)$. Such paths can be simulated indirectly by first drawing a random parameter $z^i \sim p(\cdot | a)$ and subsequently drawing $x^i \sim p(\cdot | Z = z^i)$ using the stochastic simulation algorithm (SSA)²³. However, it is also possible to draw samples x^i directly from $p(x | a)$ by employing the marginal simulation algorithm (MSA)¹⁸. In cases where Z is expensive to simulate (e.g., if it is a stochastic process itself) this can yield substantially reduced simulation times compared to SSA. Under the assumption of static random parameters Z that we employed in this manuscript, however, the MSA is unlikely to yield a noticeable improvement. A possible implementation of the Rao-Blackwellized sensitivity estimation algorithm is given below.

Algorithm 1: Rao-Blackwellized Sensitivity Estimation

1. **for** $j = 1, \dots, n$
2. Simulate $z^j \sim \prod_{i=1}^N \mathcal{G}(\cdot | \mu_i^2/\sigma_i^2, \mu_i/\sigma_i^2)$
3. Simulate $x^j \sim p(\cdot | z^j)$ using SSA
4. Set weights $\widehat{W}_k(x^j) \leftarrow \frac{\partial}{\partial a_k} \ln p(x^j | a)$ using (10)
5. **end**
6. Compute estimate $\widehat{S}_k(a) \leftarrow \frac{1}{n} \sum_{j=1}^n f(x^j) \widehat{W}_k(x^j)$

We would like to specifically point out that this implementation can be used to estimate the sensitivities to multiple parameters in parallel without requiring additional simulation runs. In particular, all that is needed is the computationally negligible calculation of the weights corresponding to the parameters of interest. It follows that the computational cost of the algorithm does not scale with the number of parameters and can be used for complex reaction networks, as long as standard stochastic simulation is feasible.

III. RESULTS

A. Sensitivity estimation for a simple model of gene expression.

To validate the proposed estimators and to test how large the variance reduction obtained by the Rao-

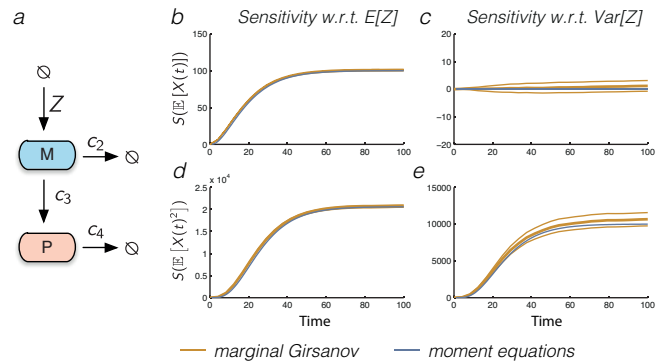


FIG. 2. Validation of the sensitivity estimation. (a) Two stage model of gene expression where we chose $M(0) = P(0) = 0$, $\mathbb{E}[Z] = 1$, $\text{Var}(Z) = 0.5$, $c_2 = 0.1$, $c_3 = 1$ and $c_4 = 0.1$. (b-e) Estimated (brown) and true (blue) sensitivities of the mean of $P(t)$ (b,c) and the second uncentered moment of $P(t)$ (d,e) with respect to the mean of Z (b,d) and the variance of Z (c,e). The estimates (thick lines) were computed from $n = 2e5$ simulated trajectories x^1, \dots, x^n and their $\pm\sigma$ -confidence region was determined via bootstrapping (thin lines).

Blackwellization is, we consider the simple two stage model of gene expression, consisting only of messenger RNA M and protein P , as depicted in Figure 2a. We assume that the production rate of M can vary between the cells, and accordingly, model this rate with a random variable Z . We used our algorithm to estimate the sensitivities of the first two moments of P with respect to the mean of Z (Figure 2b and 2d) and the variance of Z (Figure 2c and 2e). The results agree very well with the true values of these sensitivities, which can be obtained from the exact system of moment equations for this simple example²⁴.

In addition to this, we performed a comparison of the variance at stationarity of the Girsanov estimator (given in eq. 6) and the Rao-Blackwellized estimator (given in eq. 8) for the sensitivities of the mean amount of mRNA $\mathbb{E}[M(t)]$. Figure 3 shows that the Rao-Blackwellization can reduce the variance of the estimator by several orders of magnitude and is especially dramatic for low heterogeneity, i.e. when the coefficient of variation η_Z of Z is small. The reason for this is that the quality of the Girsanov estimator deteriorates when η_Z becomes small.

B. Sensitivity estimation for a bistable switch.

We next employ our algorithm to characterize the robustness of a stochastic bistable switch with respect to extrinsic variability. In particular, we consider a variant of the Schloegl system²⁵ consisting of four mass-actions reactions with corresponding parameters c_1, \dots, c_4 . Extrinsic variability is assumed to enter the dynamics through parameter $c_1 = Z$ (Figure 4a). We are now interested in quantifying the probability of finding a cell in the

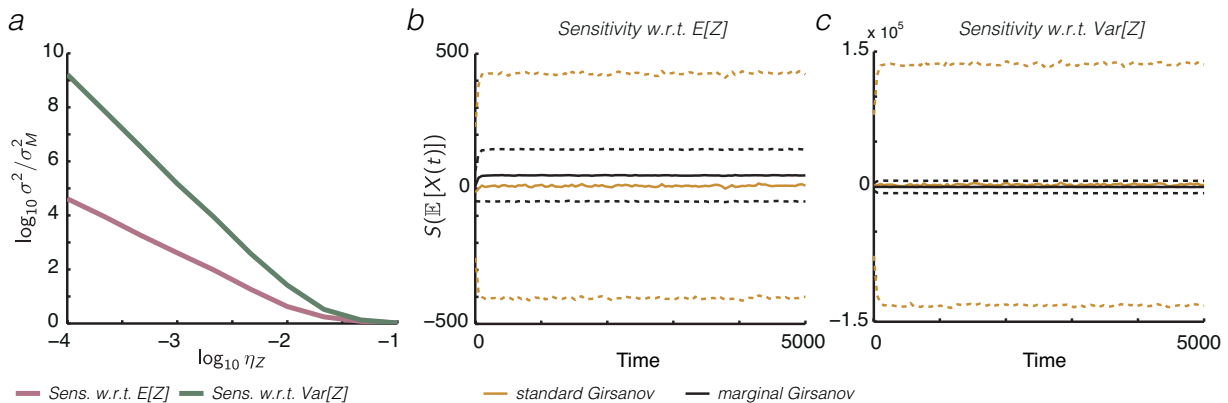


FIG. 3. **Variance reduction through Rao-Blackwellization.** We calculated the sensitivities of the expected amount of mRNA $\mathbb{E}[M(t)]$ of the model in Figure 2a with respect to mean and variance of Z for different coefficients of variation η_Z of Z . All estimates were computed 300 times using $n = 1000$ i.i.d. samples each to determine the mean and $\pm\sigma$ confidence regions. (a) Logarithmic variance reduction (ratio of the variances of the two estimators) as a function of the logarithm of η_Z for the sensitivity of the mean of $M(t)$ at stationarity ($T = 5000$) with respect to the mean (purple) and the variance (green) of Z . For $\eta_Z = 10^{-4}$ the variance reduction is more than 9 orders of magnitude. (b,c) Sensitivities of the mean of $M(t)$ with respect to mean (b) and variance (c) of Z . The estimates and their confidence regions were calculated with the Girsanov estimator (brown) and the Rao-Blackwellized estimator (black). Mean values and $\pm\sigma$ confidence regions are shown as solid and dashed lines, respectively.

more abundant of the two modes, that is, $P(X(t) > \varepsilon)$ with ε as an arbitrary threshold. For instance, if we had tuned a synthetic switch to maintain a constant percentage of ON-cells, we may ask how this number changes in the presence of extrinsic noise. In contrast to the previous example, this choice of output functional does not allow for any analytical or efficient numerical solution through the use of moment equations.

Figure 4b shows the values of $P(X(t) > \varepsilon)$ as a function of ε for a homogeneous population ($\text{Var}[Z] = 0$) and a heterogeneous population with $CV(Z) = 0.1$. The difference between these two curves, i.e. the change in the output when extrinsic noise changes from zero to substantial ($CV(Z) = 0.1$), is maximal around $\varepsilon = 35$ as can be seen in Figure 4c (blue curve). To test whether this can be correctly predicted from the sensitivities, we use our algorithm to estimate $\partial/\partial\mu P(X(t) > \varepsilon)$ and $\partial/\partial\sigma^2 P(X(t) > \varepsilon)$ with $\mu = \mathbb{E}[Z]$ and $\sigma^2 = \text{Var}[Z]$ for $\varepsilon \in [0, 60]$ using $n = 2e5$ sampled trajectories. The results (Figure 4c) show that the sensitivity to the extrinsic variance σ^2 is a good indicator of the actual change of the output with respect to increasing extrinsic noise levels. Contrary to this, the sensitivity to the extrinsic mean μ is not a good indicator in our setting and would only be appropriate if Z changes in all cells in the same way, for instance due to some global perturbation that affects the entire cell population simultaneously.

C. Hog1 induced gene expression in yeast.

A biological system that has received widespread attention in the last years is the transcriptional response to osmotic stress in budding yeast^{26–30}. Upon hyperos-

motomic shock, yeast cells activate a signaling cascade that leads to the activation and translocation to the nucleus of the MAPK (mitogen-activated protein kinase) Hog1. In the nucleus, active Hog1 leads to the initiation of a transcriptional stress response in which roughly 300 genes are transiently upregulated until the stress has been counterbalanced through an appropriate increase of cell internal osmolarity³¹. In several studies, one of these stress genes (STL1) has been labeled with a fluorescent marker and its dynamics after the addition of salt (NaCl) to the cell media have been observed^{10,32}. An intriguing feature of this system is that for intermediate stress levels, single cell measurements show that some cells initiate their transcriptional response program whereas other cells do not³³. These observations have motivated researchers to investigate whether this observed variability stems from extrinsic or intrinsic noise sources. It is therefore of interest to study how the fraction of responding cells changes with increasing levels of extrinsic noise. To address this question, we made use of a stochastic model of the system (see Figure 5a and Appendix C), akin to previously published models^{7,34}, but allowing all reaction rates to be influenced by extrinsic noise. We chose a threshold protein level $\epsilon = 40$, sufficiently above any levels that could be reached through leakage in gene expression, and defined a cell as responding to the stress if the amount of protein reached values larger than the chosen threshold at any time t in an interval $[0, T]$ corresponding to the period of exposure to the stress. Subsequently, we used our sensitivity estimation scheme to determine the sensitivities of the fraction of responding cells with respect to changes in means and variances of the random reaction rates. That is, we calculated the sensitivity indexes S_i , $i = 1, \dots, 8$ defined in (4) for the 8 reaction rate

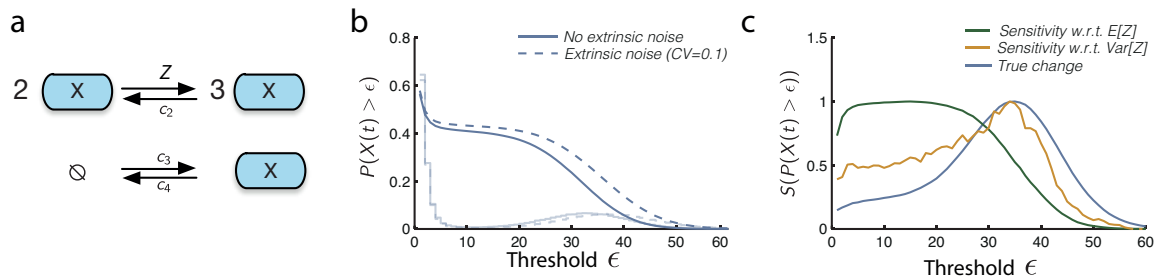


FIG. 4. **Sensitivity estimation of a bistable switch.** (a) Model of the Schloegl system. The rate $c_1 = Z$ is assumed to be influenced by extrinsic noise and modeled as a gamma distributed random variable Z with $\mathbb{E}[Z] = 1.05$. The remaining parameters are chosen as $c_2 = 0.075$, $c_3 = 3.5$ and $c_4 = 3$ (b) Value of $P(X(t) > \epsilon)$ as a function of ϵ for a homogeneous population (solid blue) and a heterogeneous population with $CV(Z) = 0.1$ (dashed blue) at time $t = 50$. For reference, also the underlying bimodal probability distributions (light blue) are shown. (c) The change in $P(X(t) > \epsilon)$ when extrinsic noise changes from zero to substantial ($CV=0.1$) (blue) is compared to the sensitivity of $P(X(t) > \epsilon)$ with respect to the mean $\mu = \mathbb{E}[Z]$ (green) and the variance $\sigma^2 = \text{Var}[Z]$ (brown) of the extrinsic variable. These sensitivities were computed for a small CV of Z (i.e., $CV(Z) = 0.01$) corresponding to very little or no extrinsic noise. All estimates were computed using $n = 2e5$ sample paths.

constants of the model using the path functional

$$f(x) := \begin{cases} 1, & \text{if } \exists t \in [0, T] : x(t) > \epsilon, \\ 0, & \text{otherwise.} \end{cases}$$

To allow for quantitative comparisons of different parameters, we chose to display sensitivities to the logarithm of the parameters. Figure 5c shows the estimated sensitivities to the log-means $\frac{\partial}{\partial \ln \mu_i} \mathbb{E}[f(X)]$ (left) and log-variances $\frac{\partial}{\partial \ln \sigma_i^2} \mathbb{E}[f(X)]$ (right). The fraction of responding cells is most sensitive to the mean of the chromatin remodeling reaction and overall relatively insensitive to parameter variances. This is in good agreement with previously published results where it was found that the bimodality of the response of cell populations is caused by intrinsic noise, and in particular, by a slow step in the activation of the gene^{7,33}. To conclude, we would like to point out that most of the sensitivities $\frac{\partial}{\partial \ln \mu_i} \mathbb{E}[f(X)]$ have a different sign from the sensitivities $\frac{\partial}{\partial \ln \sigma_i^2} \mathbb{E}[f(X)]$. This shows once more that the two types of perturbations can have opposing effects on the network's output (as claimed in Figure 1).

IV. DISCUSSION

Understanding the robust behavior of biochemical processes in living cells is among the prime interests in the field of systems- and synthetic biology. Local sensitivity analysis is one of the most effective mathematical tools that we have available to predict how a biological circuit is affected by noise and perturbations. Methods for computing sensitivities are well established for deterministic models described by ordinary differential equations. For models including intrinsic molecular noise, however, this is not the case and the calculation of sensitivities for these models remains a major challenge and an ac-

tive research topic^{35,36}. The goal of this paper was to go beyond existing methods and develop a sensitivity estimation scheme for models that include both intrinsic and extrinsic sources of variability by allowing the reaction rate constants to randomly vary between cells. We developed a novel sensitivity index that allows the quantification of a circuit's robustness with respect to changes in characteristics of the distribution over a parameter of interest as opposed to the parameter itself. In particular, we have shown, how sensitivities can be computed with respect to the mean and variance of this distribution. We remark that the presented method can be extended to arbitrary characteristics of the distribution in a straightforward fashion, for instance, to compute sensitivities with respect to higher-order moments or normalized measures of variation (e.g., the coefficient of variation).

However, we found that a naive extension of established Monte Carlo estimators may lead to an unreasonably large variance for these types of models. For instance, the confidence regions in Figure 3c show that with a naive implementation, the sensitivity with respect to the extrinsic variance, which is in truth zero, may be estimated as anything ranging between -10^5 and 10^5 despite the fact that a fairly large number ($n = 1000$) of sampled system trajectories was used in the estimation. Other more sophisticated sensitivity estimation techniques, e.g. based on finite differences¹², regularized pathwise differentiation³⁷ or the auxiliary path algorithm¹³ might lead to improved results but such approaches have not been tested or even developed for stochastic models that include extrinsic variability. Motivated by this, we made use of the recently introduced marginal process framework for stochastic reaction networks in random environments¹⁸ to propose Rao-Blackwellized estimators that are guaranteed to have lower variances than naively constructed estimators that do not profit from an analytical marginalization. Importantly, our numerical results (Figure 3a) showed that the

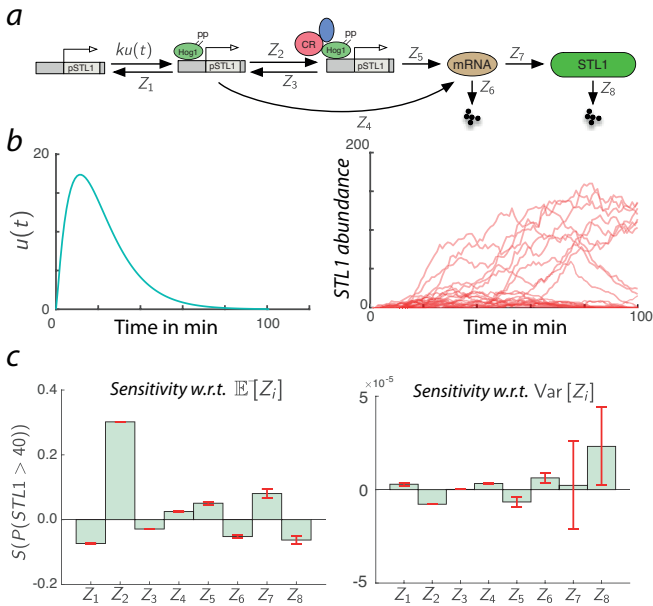


FIG. 5. **Hog1 induced gene expression in yeast.** (a) Stochastic model of the system. Hog1 binding to the STL1 promoter depends on the nuclear Hog1 concentration and happens at the time varying rate $ku(t)$ (see panel (b) and Appendix C). When Hog1 is bound, messenger RNA can be produced, but only at a low rate Z_4 . Efficient mRNA production (at rate Z_5) requires an additional chromatin remodeling step, which takes place at a rate Z_2 . Subsequently, mRNA is translated (at rate Z_7) to yield the stress response protein STL1. When the stress is counterbalanced, the nuclear Hog1 signal $u(t)$ returns to zero, the promoter returns to its inactive state (rates Z_1 and Z_3), and mRNA and protein are degraded (rates Z_6 and Z_8). (b) Left: Nuclear Hog1 dynamics acting as input to the gene expression system. Right: Representative STL1 trajectories obtained by stochastic simulation. (c) Estimated sensitivities (and $\pm\sigma$ confidence regions) of the fraction of responding cells to the logarithm of means and variances of all the reaction rates. The estimators were evaluated using the Rao-Blackwellized estimator with $n = 5e6$ sample paths. Simulations were performed with Gamma distributed small extrinsic noise ($CV(Z) = 0.01$ for all rates). Means of Z_1, \dots, Z_8 are provided in Appendix C.

resulting variance reduction may be dramatic, i.e. up to several orders of magnitude. We expect that these results can serve as the basis for establishing a theory, and the required methodology, for sensitivity analysis of stochastic models of biochemical reaction networks with random parameters.

ACKNOWLEDGMENTS

The authors would like to thank Ankit Gupta and Mustafa Khammash for their valuable feedback. J.R. acknowledges support from the Agence Nationale de la Recherche (ANR) under grant no. ANR-16-CE33-0018 (MEMIP). H.K. acknowledges support from the LOEWE

research priority program CompuGene.

Appendix A: Proof of Theorem 1

Assuming i.i.d. samples, the variances $S_k(a)$ and $\hat{S}_k(a)$ are given by

$$\text{Var}[S_k(a)] = \frac{1}{n} \text{Var}[f(X)W_k(Z)] \quad (\text{A1})$$

$$\text{and } \text{Var}[\hat{S}_k(a)] = \frac{1}{n} \text{Var}[f(X)\widehat{W}_k(X)], \quad (\text{A2})$$

respectively. Therefore, we need to show that $\text{Var}[f(X)\widehat{W}_k(X)] \leq \text{Var}[f(X)W_k(Z)]$. We start by rewriting the modified weights $\widehat{W}_k(a)$ from eq. (7) as

$$\begin{aligned} \widehat{W}_k(X) &= \frac{\partial}{\partial a_k} \ln p(x|a) \\ &= \frac{1}{p(x|a)} \int p(x|z) \frac{\partial}{\partial a_k} p(z|a) dz. \end{aligned}$$

Applying Bayes' rule, we further obtain

$$\begin{aligned} \widehat{W}_k(X) &= \frac{1}{p(x|a)} \int \frac{p(z|x,a)p(x|a)}{p(z|a)} \frac{\partial}{\partial a_k} p(z|a) dz \\ &= \int \underbrace{\frac{1}{p(z|a)} \frac{\partial}{\partial a_k} p(z|a)}_{= \frac{\partial}{\partial a_k} \ln p(z|a)} p(z|x,a) dz \\ &= \mathbb{E}[W_k(Z) | X]. \end{aligned}$$

Since the weights $W_k(Z)$ of the original estimator $S_k(a)$ are replaced by their expectation conditional on the path X , we know by the Rao-Blackwell theorem^{19,20} that $\hat{S}_k(a)$ can only have a lower or equal variance than $S_k(a)$. \square

Intuitively, this can also be seen by the law of total variance

$$\begin{aligned} \text{Var}[S_k(a)] &= \frac{1}{n} \text{Var}[f(X)W_k(Z)] = \\ &= \underbrace{\frac{1}{n} \mathbb{E}[\text{Var}[f(X)W_k(Z) | X]]}_{\geq 0} \\ &+ \underbrace{\frac{1}{n} \text{Var}[f(X)\mathbb{E}[W_k(Z) | X]]}_{\text{Var}[\hat{S}_k(a)]}. \end{aligned}$$

The first term on the r.h.s is always greater than zero unless Z can be predicted from X (e.g., when X conditional on Z is deterministic). The second term corresponds to the variance of the Rao-Blackwellized estimator, which is consequently smaller than the total variance $\text{Var}[S_k(a)]$.

Appendix B: Weights for independent Gamma distribution

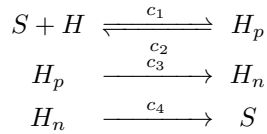
To facilitate the implementation of our estimation procedure, we provide here the weights needed in line 4 of *Algorithm 1* for independent and Gamma distributed extrinsic factors Z_1, \dots, Z_N , i.e. $p(z | a) = \prod_{i=1}^N \mathcal{G}(z_i | \alpha_i, \beta_i)$. The weights for estimating the sensitivities with respect to means μ_i and variances σ_i^2 , $i = 1, \dots, N$ can then be obtained by differentiating the logarithm of Eq.(10). This leads to

$$\begin{aligned} \frac{\partial}{\partial \mu_i} \ln p(x | a) &= \frac{G_i \mu_i - R_i}{\mu_i + G_i \sigma_i^2} + \frac{2\mu_i}{\sigma_i^2} \left\{ \psi \left(R_i + \frac{\mu_i^2}{\sigma_i^2} \right) \right. \\ &\quad \left. - \psi \left(\frac{\mu_i^2}{\sigma_i^2} \right) - \ln \left(G_i + \frac{\mu_i}{\sigma_i^2} \right) + \ln \left(\frac{\mu_i}{\sigma_i^2} \right) \right\}, \\ \frac{\partial}{\partial \sigma_i^2} \ln p(x | a) &= -\frac{\mu_i}{\sigma_i^2} \cdot \frac{G_i \mu_i - R_i}{\mu_i + G_i \sigma_i^2} + \frac{\mu_i^2}{\sigma_i^4} \left\{ \psi \left(\frac{\mu_i^2}{\sigma_i^2} \right) \right. \\ &\quad \left. - \psi \left(R_i + \frac{\mu_i^2}{\sigma_i^2} \right) + \ln \left(G_i + \frac{\mu_i}{\sigma_i^2} \right) - \ln \left(\frac{\mu_i}{\sigma_i^2} \right) \right\}, \end{aligned}$$

where ψ is the digamma function.

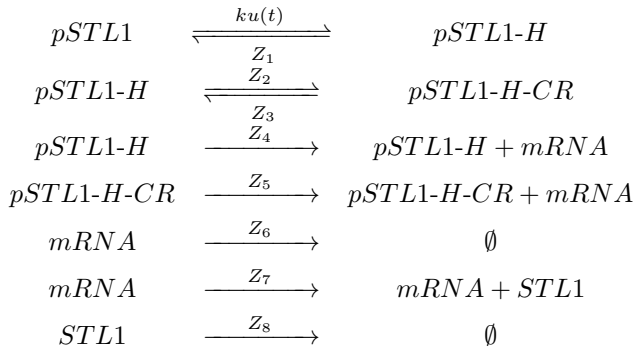
Appendix C: Model of Hog1 induced transcriptional activation

To generate the transient nuclear Hog1 signal displayed in Figure 5, we used a deterministic model consisting of the following pseudo-reactions:



with parameters $c_i = 0.1$, $i = 1, \dots, 4$, and $S(0) = 50$, $H(0) = 50$, $H_p(0) = H_n(0) = 0$.

The stochastic reaction network describing gene activation is defined by the following reactions:



where the means μ_1, \dots, μ_8 of Z_1, \dots, Z_8 have been chosen as $\mu_1 = 0.5$, $\mu_2 = \mu_3 = 0.01$, $\mu_4 = 0.1$, $\mu_5 = 1$, $\mu_6 = 0.1$, $\mu_7 = 1$, and $\mu_8 = 0.1$. Furthermore, we chose

$k = 0.5$ and we assumed that the gene is initially in the off state and that no mRNA and protein molecules are present.

- ¹R. Gutenkunst, J. Waterfall, F. Casey, K. Brown, C. Myers, and J. Sethna, *PLoS Computational Biology* **3**, e189 (2007).
- ²S. Cardinale and A. P. Arkin, *Biotechnology Journal* **7**, 856 (2012).
- ³I. M. Sobol', *Matematicheskoe Modelirovanie* **2**, 112 (1990).
- ⁴M. D. Morris, *Technometrics* **33**, 161 (1991).
- ⁵T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle, *Proceedings of the National Academy of Sciences of the USA* **97**, 4649 (2000).
- ⁶H. Koepl, C. Zechner, A. Ganguly, S. Pelet, and M. Peter, *Int J Robust Nonlin* **22**, 1103 (2012).
- ⁷C. Zechner, J. Ruess, P. Krenn, S. Pelet, M. Peter, J. Lygeros, and H. Koepl, *Proceedings of the National Academy of Sciences of the USA* **109**, 8340 (2012).
- ⁸C. Zechner, M. Unger, S. Pelet, M. Peter, and H. Koepl, *Nature Methods* **11**, 197 (2014).
- ⁹J. Ruess, F. Parise, A. Miliadis-Argeitis, M. Khammash, and J. Lygeros, *Proceedings of the National Academy of Sciences of the USA* **112**, 8148 (2015).
- ¹⁰A. Llamasi, A. Gonzalez-Vargas, C. Versari, E. Cinquemani, G. Ferrari-Trecate, P. Hersen, and G. Batt, *PLoS Computational Biology* **12**, e1004706 (2016).
- ¹¹R. Gunawan, Y. Cao, L. Petzold, and F. Doyle, *Biophysical Journal* **88**, 2530 (2005).
- ¹²D. Anderson, *SIAM Journal on Numerical Analysis* **50**, 2237 (2012).
- ¹³A. Gupta and M. Khammash, *Journal of the Royal Society Interface* **11**, 20140979 (2014).
- ¹⁴A. Miliadis-Argeitis, J. Lygeros, and M. Khammash, *The Journal of Chemical Physics* **141**, 024104 (2014).
- ¹⁵E. Skubak Wolf and D. Anderson, *The Journal of Chemical Physics* **142**, 034103 (2015).
- ¹⁶A. Tsourtis, Y. Pantazis, M. Katsoulakis, and V. Harmandaris, *The Journal of Chemical Physics* **143**, 014116 (2015).
- ¹⁷S. Plyasunov and A. Arkin, *Journal of Computational Physics* **221**, 724 (2006).
- ¹⁸C. Zechner and H. Koepl, *PLoS Computational Biology* **10**, e1003942 (2014).
- ¹⁹C. Rao, *Bulletin of the Calcutta Mathematical Society* **37**, 81 (1947).
- ²⁰D. Blackwell, *The Annals of Mathematical Statistics* **18**, 105 (1947).
- ²¹D. Anderson and T. Kurtz, in *Design and Analysis of Biomolecular Circuits* (Springer, 2011) pp. 3–42.
- ²²C. Zechner, S. Deb, and H. Koepl, 12th European Control Conference (ECC), Zurich, Switzerland, 4269 (2013).
- ²³D. Gillespie, *Journal of computational physics* **22**, 403 (1976).
- ²⁴J. Ruess, A. Miliadis-Argeitis, and J. Lygeros, *Journal of the Royal Society Interface* **10**, 20130588 (2013).
- ²⁵F. Schlögl, *Zeitschrift für Physik* **253**, 147 (1972).
- ²⁶S. Hohmann, *Microbiol Mol Biol Rev* **66**, 300 (2002).
- ²⁷A. Capaldi, T. Kaplan, Y. Liu, N. Habib, A. Regev, N. Friedman, and E. O'Shea, *Nat Genet* **40**, 1300 (2008).
- ²⁸J. Macia, S. Regot, T. Peeters, N. Conde, R. Sol, and F. Posas, *Sci Signal* **2**, ra13 (2009).
- ²⁹D. Muzzey, C. Gomez-Urbe, J. Mettetal, and A. van Oudenaarden, *Cell* **138**, 160 (2009).
- ³⁰Z. Zi, W. Liebermeister, and E. Klipp, *PloS One* **5**, e9522 (2010).
- ³¹A. Gasch, P. Spellman, C. Kao, O. Carmel-Harel, M. Eisen, G. Storz, D. Botstein, and P. Brown, *Mol Biol Cell* **11**, 4241 (2000).
- ³²J. Uhlenhof, A. Miermont, T. Delaveau, G. Charvin, F. Fages, S. Bottani, G. Batt, and P. Hersen, *Proc Natl Acad Sci USA* **109**, 1427114276 (2012).
- ³³S. Pelet, F. Rudolf, M. Nadal-Ribelles, E. de Nadal, F. Posas, and M. Peter, *Science* **332**, 732 (2011).
- ³⁴A. Gonzalez, J. Uhlenhof, E. Schaul, E. Cinquemani, G. Batt,

- and G. Ferrari-Trecate, Proceedings of the 12th European Control Conference (2013).
- ³⁵G. Arampatzis, M. Katsoulakis, and L. Rey-Bellet, The Journal of Chemical Physics **144**, 104107 (2016).
- ³⁶A. Hashemi, M. Nunez, P. Plechac, and D. Vlachos, The Journal of Chemical Physics **144**, 074104 (2016).
- ³⁷P. Sheppard, M. Rathinam, and M. Khammash, The Journal of Chemical Physics **136**, 034115 (2012).