



**HAL**  
open science

# Stochastic reaction networks with input processes: Analysis and application to gene expression inference

Eugenio Cinquemani

► **To cite this version:**

Eugenio Cinquemani. Stochastic reaction networks with input processes: Analysis and application to gene expression inference. *Automatica*, 2019, 101, pp.150-156. 10.1016/j.automatica.2018.11.047 . hal-01925923

**HAL Id: hal-01925923**

**<https://inria.hal.science/hal-01925923v1>**

Submitted on 27 Nov 2018

**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.

# Stochastic reaction networks with input processes: Analysis and application to gene expression inference <sup>\*</sup>

Eugenio Cinquemani <sup>a</sup>

<sup>a</sup>*Univ. Grenoble Alpes, Inria, 38000 Grenoble, France*

---

## Abstract

Stochastic reaction network modelling is widely utilized to describe the probabilistic dynamics of biochemical systems in general, and gene interaction networks in particular. The statistical analysis of the response of these systems to perturbation inputs is typically dependent on specific perturbation models. Motivated by reporter gene systems, widely utilized in biology to monitor gene activity in individual cells, we address the analysis of reaction networks with state-affine rates in presence of an input process. We develop a generalization of the so-called moment equations that precisely accounts for the first- and second-order moments of arbitrary inputs without the need for a model of the input process, as well as spectral relationships between the network input and state. We then apply these results to develop a method for the reconstruction of the autocovariance function of gene activity from reporter gene population-snapshot data, a crucial step toward the investigation of gene regulation, and demonstrate its performance on a simulated case study.

*Key words:* Moment Equations; Gene networks; Spectral analysis; Chemical Master Equation; Systems Biology

---

## 1 Introduction

Stochastic Reaction Network (SRN) modelling is widespread [37]. In cellular biology, fundamental SRN modelling and analysis tools such as the Chemical Master Equation (CME, [11]) or the Moment Equations (ME, [15]) are widely utilized to investigate, in particular, gene expression dynamics and interactions in single cells [36,17,29]. Most often in this context, network dynamics depend on perturbation processes that are partially or completely unknown. Analysis of SRNs in presence of perturbations is thus an important challenge *per se* [16,21,39]. A common technique to monitor gene expression dynamics are reporter gene systems, DNA constructs enabling quantification of the activity of a gene of interest *via* the monitoring of a visible (*e.g.* fluorescent) reporter protein [13,41]. Gene activity acts as a perturbation input on the reporter dynamics, and is the object of biological interest. The challenge is to learn its properties from the available readouts.

In this paper, we address statistical analysis of SRNs with input processes, and its application to the recon-

struction of input process statistics in reporter gene systems. As a first contribution, for the fundamental class of SRN with state-affine reaction rates, we generalize the standard MEs to account for the presence of virtually arbitrary input processes, and establish the relationship between input and state power spectra. Using our generalized MEs in conjunction with tools from regularized estimation, our second contribution is an original algorithm for the inference of first- and second-order gene activity statistics (in particular, its autocovariance function) from population-snapshot data, that is, time-profiles of reporter abundance statistics easily obtained *e.g.* by flow-cytometry experiments [13,40].

Our contribution to the analysis of SRNs with stochastic input processes, vastly extending the preliminary results in [6], is novel to the best of our knowledge. Contrary to existing work where mean-field, linear-noise, or other approximations of the CME are utilized (see *e.g.* [27,34]), our analysis is a general, exact consequence of the CME. A similar approach is taken in [20], but in absence of stochastic inputs, and occasionally for specific case studies (see *e.g.* [38]). On the other hand, the key novelty of our statistics inference method is that no model for the input process is required. Current competing techniques (*e.g.* [40,19,18,13,24,5]) reduce the problem to estimating unknown parameters of a given model for the input dynamics, with much loss

---

<sup>\*</sup> This paper was not presented at any IFAC meeting. Corresponding author Tel. +33 4 76 61 53 09.

*Email address:* eugenio.cinquemani@inria.fr (Eugenio Cinquemani).

of generality, or rely on observations of individual cell trajectories that require far more complex experiments (see *e.g.* [35,41,22]). Exceptions are few and typically based on approximations [26,14,31]. Our method can be seen as a generalization of existing deconvolution methods [42,9,33] from the case of single input profiles to that of input process statistics, providing a new tool for the investigation of unknown gene regulatory mechanisms. In a broader perspective, this work leverages approaches from stochastic processes, estimation and control theory to develop novel results on the modelling and analysis of biochemical processes, and methods for their reconstruction from experimental data. Due to the nature of SRN models, our results contribute to the theory on Continuous-Time Markov Chains (CTMCs, [37,25]) in particular, and Stochastic Hybrid Systems [15,23] in general, with further potential applications for these systems. The application treated here also reiterates the interest of control-theoretic approaches in the ever-increasing fields of systems and synthetic biology.

The paper is organized as follows. Sec. 2 discusses SRNs with state-affine rates. Our original analysis of SRNs with input processes is developed in Sec. 2.1. Sec. 3 is dedicated to reporter gene systems, with our original method for the reconstruction of gene activation statistics developed in Sec. 3.1. Performance of the method is evaluated in simulation on a realistic case study in Sec. 4. Perspectives of the work are discussed in Sec. 5.

*Notation:*  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{R}$  and  $\mathbb{R}_+$  denote natural, integer, real and nonnegative real numbers, respectively. For a set  $T \subset \mathbb{R}$ ,  $\mathbb{1}_T(\cdot)$  is the indicator function of  $T$ , and  $\mathbb{1}(\cdot)$  is the unit step function  $\mathbb{1}_{[0,+\infty)}(\cdot)$ .  $\|\cdot\|$  denotes Euclidean norm. For two random vectors  $X$  and  $Y$ ,  $\mathbb{E}[X]$  denotes expectation of  $X$ ,  $\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^T]$  (superscript “ $T$ ” denoting transposition) and  $\text{Var}(X) = \text{Cov}(X, X)$ .  $\mathbb{P}[\cdot]$  denotes probability.

## 2 Stochastic reaction networks with inputs: Generalized Moment Equations

A reaction network is a family of  $n \in \mathbb{N}$  chemical species  $\mathcal{S}_1, \dots, \mathcal{S}_n$  and  $m \in \mathbb{N}$  reactions  $\mathcal{R}_1, \dots, \mathcal{R}_m$  that may occur among them in a given reaction volume. The stoichiometry matrix,  $S = [S_1, \dots, S_m] \in \mathbb{Z}^{n \times m}$ , is defined such that the  $i$ th row of  $S_j$  is the net change in the number of molecules of  $\mathcal{S}_i$  when reaction  $\mathcal{R}_j$  occurs, with  $i = 1, \dots, n$  and  $j = 1, \dots, m$ . Let  $X_i(t)$  be the number of molecules of  $\mathcal{S}_i$  at time  $t$ , and  $X(t) = [X_1(t) \cdots X_n(t)]^T$ . In a well-stirred volume, the dynamics of  $X(t)$  is governed by the transition probabilities

$$\mathbb{P}[X(t+dt) = x + S_j | X(t) = x] = w_j(x, t)dt + o(dt),$$

where  $w_j(x, t) \geq 0$  is the so-called rate function of reaction  $\mathcal{R}_j$ , while  $o(dt)$  denotes terms that vanish faster

than  $dt$  when  $dt \rightarrow 0$  [37]. In this limit,  $w_j$  has the interpretation of probability that  $\mathcal{R}_j$  occurs per unit time  $dt$ . It depends on  $x$ , the current abundance of the different species, as dictated by the mass-action laws [11], and possibly on time due to *e.g.* environmental perturbations. Hence  $X(t)$  is a random state vector, and it follows the laws of a CTMC. We concentrate on reaction rates  $w_j(x, t)$  of the form

$$W_j^T x + f_j(t), \quad j = 1, \dots, m, \quad (1)$$

where  $W_j \in \mathbb{R}_+^n$  is a constant column vector and  $f_j : \mathbb{R} \rightarrow \mathbb{R}_+$  is a piecewise continuous function. This form is peculiar of networks with zeroth- or first-order reactions, and constitutes the basis for approximate description of more complex reaction dynamics [36] (see also comments in Sec. 5). Let  $W = [W_1, \dots, W_m]^T$  and  $f = [f_1, \dots, f_m]^T$ . For a given initial condition  $X(0) = x_0$ , let us write  $X^{f, x_0}$  to emphasize the dependence of the state process  $X$  on  $f(\cdot)$  and  $x_0$ . The evolution of the mean  $\mu^{f, x_0}(t) = \mathbb{E}[X^{f, x_0}(t)]$  and covariance matrix  $\Sigma^{f, x_0}(t) = \text{Var}(X^{f, x_0}(t))$  are given by the well-known moment equations (see *e.g.* [15])

$$\dot{\mu}^{f, x_0}(t) = SW\mu^{f, x_0}(t) + Sf(t), \quad (2)$$

$$\dot{\Sigma}^{f, x_0}(t) = SW\Sigma^{f, x_0}(t) + \Sigma^{f, x_0}(t)W^T S^T + Q^{f, x_0}(t), \quad (3)$$

with  $Q^{f, x_0}(t) = S \text{diag}(W\mu^{f, x_0}(t) + f(t))S^T$ ,  $\mu^{f, x_0}(0) = x_0$ ,  $\Sigma^{f, x_0}(0) = 0$ . Moreover, the autocovariance function  $\rho^{f, x_0}(z, t) = \text{Cov}(X(z), X(t))$  obeys [20]

$$\frac{\partial}{\partial z} \rho^{f, x_0}(z, t) = SW\rho^{f, x_0}(z, t), \quad (4)$$

with  $z \geq t$  and  $\rho^{f, x_0}(t, t) = \Sigma^{f, x_0}(t)$ . As the notation suggests, all these statistics depend on  $f$  and  $x_0$ . For the later developments, we need to consider the closed-form solution of Eq. (2)–(4). Define  $\ell(t) = \exp(SWt)\mathbb{1}(t)$ . For any vector or matrix function  $g$ , also define  $\mathcal{L}_t g = \int d\tau \ell(t-\tau)g(\tau)$ ,  $\mathcal{L}_t^* g = \int d\zeta g(\zeta)\ell(t-\zeta)^T$  and  $\mathcal{L}_t^\circ g = \int d\tau \ell(t-\tau)g(\tau)\ell(t-\tau)^T$ , where all integrals extend over  $\mathbb{R}_+$ . For a function  $g(\tau, \zeta)$  depending on two arguments  $\tau$  and  $\zeta$ , let  $\mathcal{L}_t$  and  $\mathcal{L}_t^*$  operate integration respectively in the first argument ( $\tau$ ) and second argument ( $\zeta$ ). We have the following result (see proof in Appendix A).

**Proposition 1** For any  $t \geq 0$  and  $z \geq t$ ,  $\mu^{f, x_0}(t) = \ell(t)x_0 + \mathcal{L}_t(Sf)$  and  $\rho^{f, x_0}(z, t) = \ell(z-t)\mathcal{L}_t^\circ(Q^{f, x_0})$ .

### 2.1 Generalized moment equations

Consider a network with affine rates (1). We now consider the case where  $f$  is one possible outcome of a stochastic process  $F = [F_1, \dots, F_m]^T$ . We make a standing assumption.

**Assumption 1** For all  $t$  and  $j = 1, \dots, m$  it holds that

$$\mathbb{P}[X(t+dt) = X(t) + S_j | X(t), F(\cdot)] = [W_j^T X(t) + F_j(t)]dt + o(dt). \quad (5)$$

In other words we stipulate that, when  $f$  is an outcome of the random process  $F$ , Eq. (1) are the conditional reaction rates of the network. This is a nontrivial fact. Indeed, according to (5), the evolution of  $X$  at time  $t$  given  $X(t)$  and the whole history of  $F$  depends on  $F(t)$  but is statistically independent of the future of  $F$ . This is analogous to the notion of stochastic causality used to characterize absence of feedback from a stochastic dynamical process ( $X$ ) into its own input ( $F$ ) [12,3]. For this class of networks, we wish to characterize the first- and second-order statistics of  $X$  as a function of the statistics of  $F$ . In agreement with the earlier definition of  $f$ , we require that  $\mathbb{E}[F(t)] \geq 0$  elementwise for all  $t$ . Denote

$$\begin{aligned} \mu_F(t) &= \mathbb{E}[F(t)], & \rho_F(z, t) &= \text{Cov}(F(z), F(t)), \\ \mu(t) &= \mathbb{E}[X(t)], & \rho(z, t) &= \text{Cov}(X(z), X(t)), \end{aligned}$$

as well as  $\Sigma(t) = \rho(t, t)$  and  $\xi_F(t) = \text{Cov}(X(0), F(t))$ .

**Proposition 2** Assume that first- and second-order moments of  $F$  are uniformly bounded. For any  $t \geq 0$  and  $z \geq t$ , it holds that  $\mu(t) = \ell(t)\mu(0) + \mathcal{L}_t(S\mu_F)$  and

$$\begin{aligned} \rho(z, t) &= \ell(z)\Sigma(0)\ell(t)^T + \mathcal{L}_z \mathcal{L}_t^*(S\rho_F S^T) + \\ &\ell(z)\mathcal{L}_t^*(\xi_F S^T) + \mathcal{L}_z(S\xi_F^T)\ell(t)^T + \ell(z-t)\mathcal{L}_t^\circ Q, \end{aligned}$$

with  $Q(t) = S\text{diag}(W\mu(t) + \mu_F(t))S^T$ . It follows that

$$\dot{\mu}(t) = SW\mu(t) + S\mu_F(t), \quad (6)$$

$$\dot{\Sigma}(t) = SW\Sigma(t) + \Sigma(t)W^T S^T + Q(t) + \quad (7)$$

$$V_\xi(t, t) + V_\xi^T(t, t) + V_\rho(t, t) + V_\rho^T(t, t),$$

$$\frac{\partial}{\partial z} \rho(z, t) = SW\rho(z, t) + V_\xi(z, t) + V_\rho(z, t), \quad (8)$$

$$V_\xi(z, t) = S\xi_F(z)^T \ell(t)^T, \quad V_\rho(z, t) = \mathcal{L}_t^*(S\rho_F(z, \cdot)S^T).$$

The proof of this (see appendix) relies on Assumption 1 to treat (2)–(4) as the equations for the conditional moments of  $X$  given  $F = f$ , and to infer the moments of  $X$  by marginalization. Comparing (6)–(8) with (2)–(4), one sees that the stochasticity of  $F$  contributes the input terms  $V_\xi$  and  $V_\rho$ . Moreover, first- and second-order statistics of  $X$  are related with those of  $F$  via linear transformations. In the next section we will exploit this fact to develop effective reconstruction methods for the statistics of  $F$  from measured statistics of  $X$  for a case study of high interest in microbiology. To conclude this section, consider the case where  $F$  has stationary mean

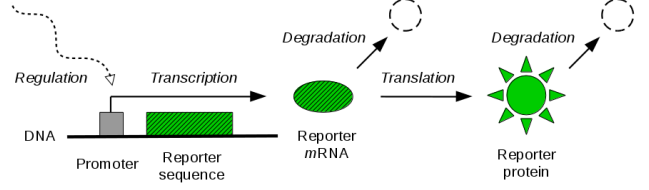


Fig. 1. Reporter gene system. The coding sequence of a reporter protein is placed under the control of the promoter of a gene of interest. Upon gene expression, reporter mRNA molecules are transcribed from the gene and further translated into visible (quantifiable) protein molecules. Both mRNA and protein molecules are subject to degradation.

$\bar{\mu}_F = \mu_F(\cdot)$  and autocovariance  $\bar{\rho}_F(\delta) = \rho_F(\cdot + \delta, \cdot)$ , with  $\delta \in \mathbb{R}_+$ . Let  $\ell_-(\cdot) = \ell(-\cdot)$  and “ $\ast$ ” denote convolution.

**Proposition 3** If all eigenvalues of  $SW$  have negative real part, for  $t \rightarrow +\infty$ ,  $\mu(t) \rightarrow \bar{\mu} = -(SW)^{-1}S\bar{\mu}_F$  and  $\rho(t + \delta, t) \rightarrow \bar{\rho}(\delta) = [\ell \ast (S\bar{\rho}_F S^T) \ast \ell_-^T](\delta) + \ell(\delta)\bar{\Sigma}^\circ$ , where, for  $\bar{Q} = S\text{diag}(W\bar{\mu} + \bar{\mu}_F)S^T$ ,  $\bar{\Sigma}^\circ$  is the unique solution of  $0 = SW\bar{\Sigma}^\circ + \bar{\Sigma}^\circ W^T S^T + \bar{Q}$ .

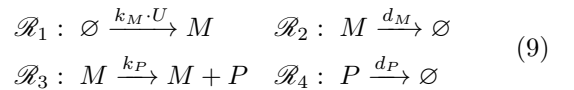
This result, proven in the appendix, allows one to generalize existing results on the power spectrum of  $X$  [20] to the presence of a random input process. Let  $R = \mathcal{F}\bar{\rho}$  and  $R_F = \mathcal{F}\bar{\rho}_F$  be the power spectral densities of  $X$  and  $F$ , where  $\mathcal{F}$  denotes Fourier transform [28]. From the expression of  $\bar{\rho}$  in Prop. 3, using well-known properties of Fourier transforms,  $R(i\omega)$  is found to be

$$L(i\omega)SR_F(i\omega)S^T L(-i\omega)^T + L(i\omega)\bar{\Sigma}^\circ + \bar{\Sigma}^\circ L(-i\omega)^T,$$

where  $\omega \in \mathbb{R}$ ,  $i$  is the imaginary unit, and  $L = \mathcal{F}\ell$ . The randomness of  $F$  contributes the leftmost term, which is equivalent to the output spectrum of a linear filter with frequency response  $L(i\omega)S$  and input  $F$  [28]. The remaining terms depend on  $\bar{\mu}_F$  via  $\bar{\Sigma}^\circ$ . They account for the deterministic component of  $F$  and correspond to the results in [20].

### 3 Reporter systems: Inference of gene activity statistics from population-snapshot data

In experimental biology, reporter systems are genetic constructs engineered in cells to monitor the activity of a gene of interest by the experimental quantification of a fluorescent protein that is synthesized when the gene is expressed. The system is illustrated in Fig. 1. Gene expression kinetics may be described by the reactions



[10,17], where  $M$  and  $P$  denote mRNA and protein species, respectively, while  $k_M$ ,  $d_M$ ,  $k_P$  and  $d_P$  are positive rate parameters.  $\mathcal{R}_1$  (transcription) represents synthesis (appearance) of one new mRNA molecule.  $\mathcal{R}_3$

(translation) describes the synthesis of one new molecule of protein  $P$  (from an existing  $m$ RNA molecule).  $\mathcal{R}_2$  and  $\mathcal{R}_4$  describe the degradation (disappearance) of one  $m$ RNA and one protein molecule, respectively. Finally,  $U$  is a binary process representing promoter activity, such that, at time  $t$ ,  $U(t) = 1$  if the gene can be transcribed ( $\mathcal{R}_1$  is enabled), and  $U(t) = 0$  otherwise [29]. In the context of this paper,  $P$  is the fluorescent protein.

In individual cells, reactions are best described as random events. System (9) is then a stochastic reaction network that can be modelled in accordance with Sec. 2. Let  $X_1(t)$  and  $X_2(t)$  denote the abundance of  $M$  and of  $P$  at time  $t$ , in the same order. Define  $F(t) = [k_M U(t), 0, 0, 0]^T$ . From the laws of mass-action [11], in agreement with standard models in the literature [29,10,17], the rates of reactions (9) are in the affine form of Sec. 2, that is, they are equal to  $WX(t) + F(t)$ . In turn, matrix  $W$  and the stoichiometry of the system are given by

$$S = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \quad W = \begin{bmatrix} 0 & d_M & k_P & 0 \\ 0 & 0 & 0 & d_P \end{bmatrix}^T.$$

For this system, we postulate that Assumption 1 holds. This is biologically reasonable since by definition reporter genes are foreign to the host cell, therefore products  $M$  and  $P$  are not expected to participate in the regulation of the cell genes.

### 3.1 Promoter statistics reconstruction method

We now develop a method to reconstruct second-order statistics of the promoter activity  $U$  from mean and variance measurements of the reporter protein abundance  $Y = X_2$ . This type of data stems from flow-cytometry experiments, where the empirical statistics of  $Y$  at different times  $t_k$ , with  $k = 0, \dots, M-1$ , are computed from samples of many (*e.g.*  $10^4$  or more) cells from a given cellular population. To develop our method, we apply the results of Sec. 2 assuming that the gene expression model described above is exact. Since no hypothesis is made on the mechanisms governing  $U$ , and the modelling of the transcription and translation steps is widely accepted, reconstruction errors due to modelling inaccuracy are expected to be negligible for a variety of real systems.

Following [40], by the law of large numbers, mean and variance measurements obey

$$\tilde{\mu}_Y(t_k) = \mu_Y^*(t_k) + e_k^\mu, \quad \tilde{\sigma}_Y^2(t_k) = \sigma_Y^{*2}(t_k) + e_k^\sigma, \quad (10)$$

where  $\mu_Y^*(t_k)$  and  $\sigma_Y^*(t_k)$  denote the true mean and variance of  $Y$  at time  $t_k$ , while the errors ( $e_k^\mu, e_k^\sigma$ ) are approximately Gaussian, zero-mean and uncorrelated across  $k$ , with covariance matrix determined from the data.

We consider the case where  $U(t)$  is wide-sense stationary, that is, with constant mean  $\bar{\mu}_U^*$  and autocovariance function  $\bar{\rho}_U^*(\delta)$  only depending on time lag  $\delta \in \mathbb{R}$ , and  $X(0) = 0$ . In practice, these conditions should be warranted by suitable experiment design. Expressing the statistics of  $F$  in terms of those of  $U$ , Eq. (6)–(7) become

$$\dot{\mu}(t) = SW\mu(t) + \begin{bmatrix} k_M \bar{\mu}_U & 0 \end{bmatrix}^T, \quad (11)$$

$$\dot{\Sigma}(t) = SW\Sigma(t) + \Sigma(t)W^T S^T + Q(t) + \Lambda_{\bar{\rho}_U}(t), \quad (12)$$

with  $\mu(0) = 0$  and  $\Sigma(0) = 0$ , where  $\Lambda_{\bar{\rho}_U}(t) = V_\rho(t, t) + V_\rho^T(t, t)$  with  $\rho(t+\delta, t) = \text{diag}(k_M^2 \bar{\rho}_U(\delta), 0)$  (for the definition of  $V_\rho(t, t)$  and  $Q(t)$  see Prop. 2). The true mean and variance in (10) then obey  $\mu_Y^*(t_k) = C\mu^*(t_k)$  and  $\sigma_Y^{*2}(t_k) = C^T \Sigma^*(t_k) C$ , where  $\mu^*$  and  $\Sigma^*$  are the solution of (11)–(12) corresponding to the true but unknown statistics  $\bar{\mu}_U = \bar{\mu}_U^*$  and  $\bar{\rho}_U = \bar{\rho}_U^*$ , and  $C = [0, 1]^T$ . Eq. (11)–(12) depend on  $\theta = (k_M, d_M, k_P, d_P)$ . These are design parameters of the reporter system, hence we assume they are known. We may then formulate the reconstruction problem as follows.

**Problem 1** *Given measurements (10), estimate the unknown quantities  $\bar{\mu}_U^*$  and  $\bar{\rho}_U^*(\cdot)$  using model (11)–(12) with all parameters known.*

Mean  $\bar{\mu}_U^*$  can be immediately found by fitting the data-points  $\tilde{\mu}_Y$  with the solution of (11) as a function of the only unknown  $\bar{\mu}_U$ . Therefore we assume  $\bar{\mu}_U^*$  known, call  $\mu^*(\cdot)$  the corresponding solution of (11), and  $Q^*(t)$  the expression of  $Q(t)$  determined by  $\bar{\mu}_U^*$  and  $\mu^*(\cdot)$ . We thus concentrate on the following problem reformulation.

**Problem 2** *Given data  $\tilde{\sigma}_Y^2(\cdot)$ , estimate  $\bar{\rho}_U^*(\cdot)$  using (12) with all parameters and  $Q(\cdot) = Q^*(\cdot)$  known.*

Let  $\hat{\Sigma}(t|\bar{\rho}_U, Q)$  be the solution of (12) at time  $t$  for generic  $Q$  and  $\bar{\rho}_U$ . Since  $\Lambda_{\bar{\rho}_U}(t)$  is a linear functional of  $\bar{\rho}_U$ ,  $\hat{\Sigma}(t|\bar{\rho}_U, Q)$  is linear in  $\bar{\rho}_U$  and  $Q$ . Problem 2 is thus a linear inverse one. However, data sampling and the infinite dimension of  $\bar{\rho}_U(\cdot)$  make it ill-posed [2]. The solution we propose is based on Tikhonov regularization [8]. An estimate  $\hat{\rho}_U$  of  $\bar{\rho}_U^*$  is defined as a solution of

$$\inf_{\bar{\rho}_U \in \mathcal{C}} \sum_{k=0}^{M-1} \alpha_k^2 \left( \tilde{\sigma}_Y^2(t_k) - C^T \hat{\Sigma}(t_k|\bar{\rho}_U, Q^*) C \right)^2 + \gamma \mathcal{Q}(\bar{\rho}_U) \quad (13)$$

where the  $\alpha_k^2 = \text{var}(e_k^\sigma)^{-1}$  are fixed from the data as in [40] (see discussion of Eq. (10)), and for some regularization parameter  $\gamma \geq 0$ , term  $\gamma \mathcal{Q}(\bar{\rho}_U)$  penalizes irregular solutions.  $\mathcal{C}$  is the convex cone of positive semi-definite functions that autocovariance  $\bar{\rho}_U$  must belong to [28]. In view of the linearity of  $\hat{\Sigma}(t|\bar{\rho}_U, Q^*)$ , this is a convex problem as long as  $\mathcal{Q}$  is convex. The challenge we address

here is to turn this formal solution into a computationally tractable yet effective procedure. To do this, we consider approximate solutions  $\hat{\rho}_U$  of the form  $\mathcal{R} \cdot \hat{c}$ , where, for some  $N \in \mathbb{N}$ ,  $\hat{c} \in \mathbb{R}^N$ ,  $\mathcal{R}(\cdot) = [r_1(\cdot), \dots, r_N(\cdot)]$  and the  $r_\ell(\delta)$ , with  $\ell = 1, \dots, N$ , are suitably chosen functions, symmetric around  $\delta = 0$  in view of the symmetry of  $\bar{\rho}_U$ . Define  $\mathcal{V}(\cdot) = [v_1(\cdot), \dots, v_N(\cdot)]$  where, for every  $\ell$ ,  $v_\ell(\cdot) = C^T \hat{\Sigma}(\cdot | r_\ell, 0) C$ , and  $v_0(\cdot) = C^T \hat{\Sigma}(\cdot | 0, Q^*) C$ . Using equality  $C^T \hat{\Sigma}(\cdot | \mathcal{R}c, Q^*) C = v_0(\cdot) + \mathcal{V}(\cdot)c$ , Eq. (13) becomes

$$\inf_{c \in \mathbb{R}^N} \sum_{k=0}^{M-1} \alpha_k^2 (\tilde{\sigma}_Y^2(t_k) - v_0(t_k) - \mathcal{V}(t_k)c)^2 + \gamma \mathcal{Q}(\mathcal{R}c) \quad (14)$$

$$\text{s.t. } \mathcal{R}(\cdot)c \in \mathcal{C}. \quad (15)$$

Among the possible choices of  $\mathcal{R}$  that allow for a dense approximation of  $\mathcal{C}$  (as  $N \rightarrow \infty$ ), we propose one that makes computation of  $\mathcal{V}$  easy and at the same time leads to an explicit formulation of constraint (15). First notice that  $\mathcal{V}(t_0), \dots, \mathcal{V}(t_{M-1})$  are independent of  $\mathcal{R}(\delta)$  for  $\delta \notin T = [0, t_{M-1}]$ . In view of this, let  $\tau_0, \dots, \tau_N$  be a uniform partition of  $T$  ( $0 = \tau_0 < \tau_1 < \dots < \tau_{N-1} < \tau_N = t_{M-1}$ ). For  $\ell = 1, \dots, N$ , define the functions  $r_\ell$  over  $(-t_{M-1}, t_{M-1})$  as  $r_\ell(\delta) = \mathbb{1}_{[\tau_{\ell-1}, \tau_\ell]}(|\delta|)$ . In this case, for every  $\ell$ , the explicit expression of  $\Lambda_{r_\ell}(t)$  is easily computed, and thus  $v_\ell(t_k) = C^T \hat{\Sigma}(t_k | r_\ell, 0) C$  can be evaluated by straightforward numerical integration of (12) (with  $\Lambda_{r_\ell}$  in place of  $\Lambda_{\bar{\rho}_U}$ ) at increasing time points  $k = 0, \dots, M-1$ . Moreover, since  $\mathcal{R}(\cdot)c$  is piecewise constant with values given by the entries of  $c$ , the infinite-dimensional constraint (15) can be replaced by its finite-dimensional version

$$\mathcal{I}(c) \in \mathcal{C}_N, \quad (16)$$

where  $\mathcal{I}(c)$  denotes the symmetric Toeplitz matrix with first column equal to  $c$ , and  $\mathcal{C}_N$  is the convex cone of positive semi-definite matrices of order  $N$  [28]. Finally,  $\mathcal{Q}(\mathcal{R}c)$  can be simply defined as  $\|\Delta^2 c\|$ , with  $\Delta^2 c$  the vector of finite second-order differences of  $c$ , a natural adaptation of the standard second-order roughness penalty  $\int (\ddot{r}(\delta))^2 d\delta$  [8,30] to the piecewise-constant function  $\mathcal{R}c$ . In summary, estimate  $\hat{\rho}_U = \mathcal{R}\hat{c}$  of  $\bar{\rho}_U^*$  is obtained from the solution  $\hat{c}$  of the quadratic problem (14) with linear constraints (16), which is straightforward to compute by numerical optimization [4].

Of course, the solution depends on the regularization factor  $\gamma$ , the larger the  $\gamma$  the smoother the solution at the price of reduced fit to the data. Yet its choice can be automated in several ways. In our case, since the measurement error statistics  $\alpha_k$  are known, we can rely on the discrepancy principle [1]. Given the fitting error  $\hat{e}_k^\gamma = \tilde{\sigma}_Y^2 - v_0(t_k) - \mathcal{V}(t_k)\hat{c}^\gamma$  (here superscript “ $\gamma$ ” emphasizes dependency of the solution on  $\gamma$ )

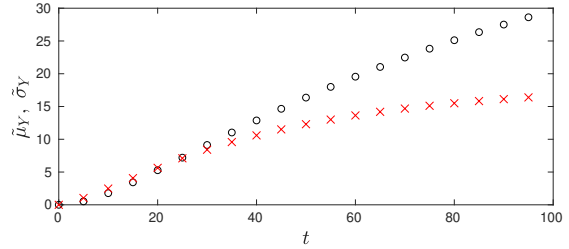


Fig. 2. Example of fluorescent reporter snapshot data  $\tilde{\mu}_Y$  (black circles) and  $\tilde{\sigma}_Y$  (red crosses).

this approach selects  $\gamma$  as the largest value such that  $\|[\alpha_0 \hat{e}_0^\gamma, \dots, \alpha_{M-1} \hat{e}_{M-1}^\gamma]\| \leq \sqrt{M}$ . Intuitively, this corresponds to just the right amount of regularization such that fitting error is within the range of measurement error. In practice, this is implemented by a search over a logarithmic grid of candidate values of  $\gamma$  spanning different orders of magnitude. Note that this procedure only relies on the available experimental measurements.

## 4 Numerical results

We now demonstrate the autocovariance reconstruction method of Sec. 3.1 on a simulated example of the gene expression model (9). We consider a case where the binary process  $U$  has a fixed switch-off rate  $\lambda_-$  and a stochastic switch-on rate  $B \cdot \lambda_+$ , where  $B$  is a latent binary process with constant switch-on and switch-off rates  $\beta_+$  and  $\beta_-$ , respectively. Biologically,  $B$  models the random binding and unbinding of a transcription factor that is necessary for the activation of gene expression ( $U$  can switch-on only when  $B = 1$ ). For simulating fluorescent reporter snapshot data, we use this model with parameter values of biological relevance [17],  $(\lambda_+, \lambda_-) = (0.1, 0.05)$  [ $\text{min}^{-1}$ ] and  $(\beta_+, \beta_-) = (0.1351, 0.1)$  [ $\text{min}^{-1}$ ]. These values are also such that  $U$  spends on average an equal time in the on and off states (that is,  $\bar{\mu}_U = 0.5$  by construction). For reconstruction instead, we consider that the above model for the dynamics of  $U$  is not known and aim at reconstructing the autocovariance function of  $U$ . This is simply not possible with currently existing methods, which would require a known model for the reactions governing the dynamics of  $U$  in order to estimate its parameters (in this case,  $\lambda_+, \lambda_-$  and  $\beta_+, \beta_-$ ). Conversely, for both simulation and reconstruction, we assume that parameters  $\theta$  are known (recall from Sec. 3.1 that these parameters do not enter the dynamics of  $U$  and are design parameters of the reporter gene system). Inspired by [17], we use the realistic parameter values  $\theta = (0.5, 0.1, 0.2, 0.01)$  [ $\text{min}^{-1}$ ].

In accordance with a realistic experimental scenario, we assume that fluorescent reporter snapshot measurements  $\tilde{\mu}_Y(t_k)$  and  $\tilde{\sigma}_Y^2(t_k)$  are collected at  $M = 20$  time instants  $t_k = 5 \cdot k$  [min], with  $k = 0, \dots, M-1$ , from empirical statistics over independent samples of  $10^5$  cells.

Here we generate this data via repeated random simulations of the reaction network using StochKit [32]. Fig. 2 shows one simulated data set. The reconstruction method of Sec. 3.1 with automated choice of the regularization factor  $\gamma$  from the set  $\{10^p, p = 1, \dots, 10\}$  is implemented in Matlab, using CVX [7] for the optimization of (14) under constraint (16). The method is implemented with the partition of the measurement horizon  $[0, 95]$  defined by  $\tau_\ell = \ell \cdot T_s - T_s/2$ , with  $\ell = 1, \dots, N-1$ , where  $T_s = 95/(N-1)$  and  $N = 150$ . Notice that  $N \gg M$  makes regularization crucial. To assess statistical performance in the reconstruction of  $\bar{\rho}_U^*$ , estimation is performed over 1000 randomly generated datasets, yielding 1000 different estimates that are used to compute empirical mean and standard deviation of  $\hat{\rho}_U(\delta)$ , with  $\delta \in T$ . For comparison, estimation is additionally performed with regularization parameter fixed to  $\gamma = 10^6$ , a convenient choice determined in advance by trial and error, and estimation performance statistics are also drawn in this case.

Results are illustrated in Fig. 3 and compared with  $\bar{\rho}_U^*$ , the true autocovariance function of  $U$ . (The computation of the latter is easy since the joint process  $(B, U)$  is a finite-state Markov chain, but is not reported due to space limitations.) For both the automated and the fixed choice of  $\gamma$ , the true  $\bar{\rho}_U^*$  is duly contained in the confidence band defined by  $\text{mean}(\hat{\rho}_U(\delta)) \pm 2 \cdot \text{std}(\hat{\rho}_U(\delta))$ ,  $\delta \in T$ . Confidence bands are similar in the two cases, remarkably narrow for small lags and nicely bounded for  $\delta$  roughly up to 75. In addition, despite regularization, estimates show very limited bias, reconfirming excellent estimation performance overall. Not surprisingly, uncertainty diverges at larger lags, comparable with the duration of the experiment. Mathematically, this follows from the strict stability of matrix  $SW$  (eigenvalues are  $-d_M < 0$  and  $-d_P < 0$ ), implying that (12) is a dynamical transformation with exponentially decaying memory;  $SW$  also enters  $\Lambda_{\bar{\rho}_U}$  in a way that cancels out the tail of  $\bar{\rho}_U$ , thus preventing its reconstruction from the data. Intuitively, all this reflects the fact that, for large  $\delta$ , the correlation between  $U(t-\delta)$  and  $U(t)$  is not detected by  $Y(t)$ , since fluctuations of  $U(t-\delta)$  are “forgotten” at a rate determined by the reporter mRNA and protein degradation rates  $d_M$  and  $d_P$ . The increased bias above  $\delta = 75$ , where reconstruction is most uncertain, is arguably an artifact of the finite number of repetitions of the estimation experiment. Compared with the fixed choice of  $\gamma$ , the automated choice of regularization only marginally increases bias and variance of the estimates around  $\delta = 0$  (yet variance is smaller for large  $\delta$ ). For the case into study, because  $U$  is a Bernoulli process, the convex constraint  $\bar{\mu}_U^*(1 - \bar{\mu}_U^*) = \bar{\rho}_U(0)$ , relating the mean (assumed known) with the variance  $\text{var}(U) = \bar{\rho}_U(0)$  of the process, could also be added into the optimization problem to ameliorate estimation around  $\delta = 0$ . However, results not reported here show that this deteriorates estimation performance for  $\delta$  away from 0.

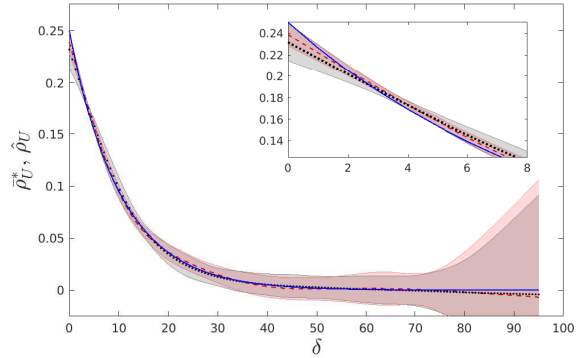


Fig. 3. Statistics of estimator  $\hat{\rho}_U$ . Blue line: True autocovariance  $\bar{\rho}_U^*(\cdot)$ ; Red dashed line and shaded region:  $\text{mean}(\hat{\rho}_U(\cdot)) \pm 2 \cdot \text{std}(\hat{\rho}_U(\cdot))$  as computed from 1000 simulated estimation experiments, using the fixed regularization factor  $\gamma = 10^6$  (bottom-cropped for better scaling). Black dotted line and dark shaded region: Same but with automated choice of regularization factor. Inset: Magnification of the estimator statistics around  $\delta = 0$ .

## 5 Conclusions

In this paper we have presented generalized MEs for the analysis of SRNs with state-affine rates and input processes, and derived the relationship between input and state spectra, under the assumption of absence of feedback. An important direction of research is the investigation of analogous generalized MEs for nonlinear rate models. Several applications can be considered, among which the testing of reaction network causality based on joint measurements of input and output statistics.

We have also developed and demonstrated performance of a method for the reconstruction of gene activity statistics. Though applied to the case of a binary process  $U$ , the method equally applies to arbitrary input processes. From an experimental viewpoint, application to real biological data is the next step. From a methodological standpoint, the method is amenable of much generalization, notably reconstruction of nonstationary input process statistics from output autocovariance data, and application to other reaction networks. Analysis of reconstruction error in presence of model mismatch is another research direction of interest. Applications of spectral analysis to gene reporter systems are also pursued.

Extensions and applications of analysis and reconstruction methods in the broader field of CTMCs is an intriguing research perspective. At the same time, further exploitation of control-theoretic methods in the analysis of microbiological data carries great potential.

## Acknowledgements

This work was funded in part by the French national research agency (ANR) via project MEMIP (ANR-16-CE33-0018), and by the Inria IPL CoSy. The author thanks Hidde de Jong for suggestions on the manuscript.



## References

- [1] F. Bauer and M.A. Lukas. Comparing parameter choice methods for regularization of ill-posed problems. *Math. Comput. Simulat.*, 81:1795–1841, 2011.
- [2] M. Bertero. Linear inverse and ill-posed problems. *Adv. Electron. El. Phys.*, 75:1 – 120, 1989.
- [3] C.G. Bowsher, M. Voliotis, and P.S. Swain. The fidelity of dynamic signaling by noisy biomolecular networks. *PLoS Comput. Biol.*, 9(3):e1002965, 2013.
- [4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, USA, 2004.
- [5] E. Cinquemani. Reconstruction of promoter activity statistics from reporter protein population snapshot data. In *Proc. of the 54th IEEE CDC*, pages 1471–1476, Dec 2015.
- [6] E. Cinquemani. On observability and reconstruction of promoter activity statistics from reporter protein mean and variance profiles. In *Proc. of the 5th HSB workshop, Grenoble, France (LNCS 9957)*, pages 147–163, Cham, 2016. Springer.
- [7] CVX Research, Inc. CVX: Matlab software for disciplined convex programming. <http://cvxr.com/cvx>, 2012.
- [8] G. De Nicolao, G. Sparacino, and C. Cobelli. Nonparametric input estimation in physiological systems: Problems, methods, and case studies. *Automatica*, 33(5):851 – 870, 1997.
- [9] B. Finkenstädt, E.A. Heron, M.I. Komorowski, K. Edwards, S. Tang, C.V. Harper, J.R.E. Davis, M.R.H. White, A.J. Millar, and D.A. Rand. Reconstruction of transcriptional dynamics from gene reporter data using differential equations. *Bioinformatics*, 24(24):2901–2907, 2008.
- [10] N. Friedman, L. Cai, and X. S. Xie. Linking stochastic dynamics to population distribution: An analytical framework of gene expression. *Phys. Rev. Lett.*, 97:168302, 2006.
- [11] D.T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188:404–425, 1992.
- [12] C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969.
- [13] J. Hasenauer, S. Waldherr, M. Doszczak, N. Radde, P. Scheurich, and F. Allgower. Identification of models of heterogeneous cell populations from population snapshot data. *BMC Bioinformatics*, 12(1):125, 2011.
- [14] J. Hasenauer, V. Wolf, A. Kazeroonian, and F. J. Theis. Method of conditional moments (MCM) for the Chemical Master Equation. *J. Math. Biol.*, 69(3):687–735, 2014.
- [15] J.P. Hespanha. Modelling and analysis of stochastic hybrid systems. *IEE Proceedings – Control Theory and Applications*, 153(5):520–535, Sept 2006.
- [16] A. Hilfinger, T.M. Norman, G. Vinnicombe, and J. Paulsson. Constraints on fluctuations in sparsely characterized biological systems. *Phys. Rev. Lett.*, 116(058101), 2016.
- [17] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins. Stochasticity in gene expression: From theories to phenotypes. *Nat. Rev. Gen.*, 6:451–464, 2005.
- [18] M. Komorowski, B. Finkenstädt, C. Harper, and D. Rand. Bayesian inference of biochemical kinetic parameters using the linear noise approximation. *BMC Bioinformatics*, 10(1):343, 2009.
- [19] M. Komorowski, B. Finkenstädt, and D. Rand. Using a single fluorescent reporter gene to infer half-life of extrinsic noise and other parameters of gene expression. *Biophys. J.*, 98(12):2759–2769, 2010.
- [20] I. Lestas, J. Paulsson, N. E. Ross, and G. Vinnicombe. Noise in gene regulatory networks. *IEEE Trans. Autom. Control.*, 53(Special Issue):189–200, 2008.
- [21] I. Lestas, G. Vinnicombe, and J. Paulsson. Fundamental limits on the suppression of molecular fluctuations. *Nature*, 467(7312):174–178, 2010.
- [22] A. Llamasi, A. M. Gonzalez-Vargas, C. Versari, E. Cinquemani, G. Ferrari-Trecate, P. Hersen, and G. Batt. What population reveals about individual cell identity: Single-cell parameter estimation of models of gene expression in yeast. *PLoS Comput. Biol.*, 12:e1004706, 2016.
- [23] J. Lygeros and M. Prandini. Stochastic hybrid systems: A powerful framework for complex, large scale applications. *Eur. J. Control*, 16:583–594, 2010.
- [24] G. Neuert, B. Munsy, R.Z. Tan, L. Teytelman, M. Khammash, and A. van Oudenaarden. Systematic identification of signal-activated stochastic gene regulation. *Science*, 339(6119):584–587, 2013.
- [25] J.R. Norris. *Markov Chains*. Number 2 in Cambridge series on statistical and probabilistic mathematics. Cambridge University Press, New York, USA, 1997.
- [26] A. Ocone, L. Haghverdi, N.S. Mueller, and F.J. Theis. Reconstructing gene regulatory dynamics from high-dimensional single-cell snapshot data. *Bioinformatics*, 31(12):i89–i96, 2015.
- [27] E.M. Ozbudak, M. Thattai, I. Kurtser, A.D. Grossman, and A. van Oudenaarden. Regulation of noise in the expression of a single gene. *Nat. Genet.*, 31:69–73, 2002.
- [28] A. Papoulis. *Probability, random variables, and stochastic processes*. McGraw-Hill series in electrical engineering. McGraw-Hill, New York, 1991.
- [29] J. Paulsson. Models of stochastic gene expression. *Phys. Life Rev.*, 2(2):157 – 175, 2005.
- [30] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, 2006.
- [31] J. Ruess, A. Miliadis-Argeitis, S. Summers, and J. Lygeros. Moment estimation for chemically reacting systems by extended kalman filtering. *J. Chem. Phys.*, 135(16):165102, 2011.
- [32] K. R. Sanft, S. Wu, M. Roh, J. Fu, R. K. Lim, and L. R. Petzold. Stochkit2: Software for discrete stochastic simulation of biochemical systems with events. *Bioinformatics*, 27(17):2457–2458, 2011.
- [33] M. Schelker, A. Raue, J. Timmer, and C. Kreutz. Comprehensive estimation of input signals and dynamics in biochemical reaction networks. *Bioinformatics*, 28(18):i529–i534, 2012.
- [34] M.L. Simpson, C.D. Cox, and G.S. Saylor. Frequency domain Langevin analysis of stochasticity in gene transcriptional regulation. *J. Theor. Biol.*, 229:383–394, 2004.
- [35] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef. Mammalian genes are transcribed with widely different bursting kinetics. *Science*, 332:472–474, 2011.
- [36] M. Thattai and A. van Oudenaarden. Intrinsic noise in gene regulatory networks. *PNAS*, 98(15):8614–8619, 2001.
- [37] N.G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library, 1992.
- [38] X. Wang, B. Errede, and T.C. Elston. Mathematical analysis and quantification of fluorescent proteins as transcriptional reporters. *Biophys. J.*, 94(6):2017–2026, 2008.



- [39] E. Yeung, J.L. Beck, and R.M. Murray. Modeling environmental disturbances with the Chemical Master Equation. In *Proc. of the 52nd IEEE CDC*, pages 1384–1391, 2013.
- [40] C. Zechner, J. Ruess, P. Krenn, S. Pelet, M. Peter, J. Lygeros, and H. Koeppl. Moment-based inference predicts bimodality in transient gene expression. *PNAS*, 109(21):8340–8345, 2012.
- [41] C. Zechner, M. Unger, S. Pelet, M. Peter, and H. Koeppl. Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings. *Nat. Methods*, 11:197–202, 2014.
- [42] V. Zulkower, M. Page, D. Ropers, J. Geiselmann, and H. de Jong. Robust reconstruction of gene expression profiles from reporter gene data using linear inversion. *Bioinformatics*, 31(12):i71–i79, 2015.

## A Proofs

**Proposition 1.** The explicit solution of (2) is well-known since (2) is a linear vector ordinary differential equation. The expression of  $\rho^{f,x_0}(t, t) = \Sigma^{f,x_0}(t)$  follows from the general solution of matrix Lyapunov equations of the form  $\dot{\Sigma}(t) = A\Sigma(t) + \Sigma(t)A^T + Q(t)$ , which is  $\Sigma(t) = \exp(At)\Sigma(0)\exp(At)^T + \int_0^t \exp(A(t-\tau)Q(\tau)\exp(A(t-\tau))^T d\tau$ . In the case of Eq. (3),  $A = SW$  and  $Q(t) = Q^{f,x_0}(t)$ , while  $\Sigma(0) = \Sigma^{f,x_0}(0) = 0$ . Thus  $\rho^{f,x_0}(t, t) = \Sigma^{f,x_0}(t) = \mathcal{L}_t^\circ(Q^{f,x_0})$ . Using the latter as initial condition at  $z = t$  of Eq. (4), the expression of  $\rho^{f,x_0}(z, t)$  for  $z \geq t$  follows.

**Proposition 2.** Under Assumption 1, the Chemical Master Equation for the evolution of  $\mathbb{P}[X(t)|F(\cdot) = f(\cdot)]$  over  $t$  takes the usual form [11], with rates given by (1). Hence, for any initial state  $X(0) = x_0$ , the conditional system moments that follow from this equation [15] are as in Prop. 1. We now use this fact to infer the moments of  $X$  by marginalization. To do so, we will repeatedly exchange integration with expectation [28]. First of all,  $\mu(t) = \mathbb{E}[\mu^{F,X_0}(t)] = \mathbb{E}[\ell(t)X(0) + \mathcal{L}_t(SF)] = \ell(t)\mu_0 + \mathcal{L}_t(S\mu_F)$ . Now define  $M(z, t) \triangleq \mathbb{E}[X(z)X(t)^T]$ . One has

$$\begin{aligned} M(z, t) &= \mathbb{E}[\rho^{F,X_0}(z, t)] + \mathbb{E}[\mu^{F,X_0}(z)\mu^{F,X_0}(t)^T] \\ &= \mathbb{E}[\ell(z-t)\Sigma^{F,X_0}(t)] + \mathbb{E}[(\ell(z-t)\mu^{F,X_0}(t) + \mathcal{L}_{t,z}(SF))\mu^{F,X_0}(t)^T] \\ &= \ell(z-t)\mathbb{E}[\Sigma^{F,X_0}(t) + \mu^{F,X_0}(t)\mu^{F,X_0}(t)^T] + \mathbb{E}[\mathcal{L}_{t,z}(SF)\mu^{F,X_0}(t)^T] \\ &= \ell(z-t)M(t, t) + \mathcal{L}_{t,z}(S\mathbb{E}[F\mu^{F,X_0}(t)^T]), \end{aligned}$$

$$\begin{aligned} M(t, t) &= \mathbb{E}[\Sigma^{F,X_0}(t)] + \mathbb{E}[\mu^{F,X_0}(t)\mu^{F,X_0}(t)^T] \\ &= \mathbb{E}[\mathcal{L}_t^\circ(S\text{diag}(W\mu^{F,X_0} + F)S^T)] + \mathbb{E}[(\ell(t)X(0) + \mathcal{L}_t(SF))(\dots)^T] \\ &= \mathcal{L}_t^\circ Q + \ell(t)M(0)\ell(t)^T + \mathbb{E}[\mathcal{L}_t(SF)\mathcal{L}_t^*(F^T S^T)] + \ell(t)\mathbb{E}[X(0) \times \mathcal{L}_t^*(F^T S^T)] + \mathbb{E}[\mathcal{L}_t(SF)X(0)^T]\ell(t)^T. \end{aligned}$$

Then, using  $M(0, 0) - \mu_0\mu_0^T = \Sigma_0$ , we have

$$\begin{aligned} \Sigma(t) &= M(t, t) - \mu(t)\mu(t)^T \\ &= \mathcal{L}_t^\circ Q + \ell(t)\Sigma_0\ell(t)^T + (\mathbb{E}[\mathcal{L}_t(SF)\mathcal{L}_t^*(F^T S^T)] - \mathcal{L}_t(S\mu_F)\mathcal{L}_t^*(\mu_F^T S^T)) + \ell(t)(\mathbb{E}[X(0)\mathcal{L}_t^*(F^T S^T)] - \mu_0\mathcal{L}_t^*(\mu_F^T S^T)) + (\mathbb{E}[\mathcal{L}_t(SF)X(0)^T] - \mathcal{L}_t(S\mu_F)\mu_0^T)\ell(t)^T \\ &= \mathcal{L}_t^\circ(Q) + \ell(t)\Sigma_0\ell(t)^T + \mathcal{L}_t\mathcal{L}_t^*(S\rho_F S^T) + \ell(t)\mathcal{L}_t^*(\xi_F S^T) + \mathcal{L}_t(S\xi_F^T)\ell(t)^T \end{aligned}$$

where we also used the fact that, for the relevant  $g$  and  $h$ ,  $\mathbb{E}[(\mathcal{L}g)(\mathcal{L}^*h)] = \mathbb{E}[\mathcal{L}\mathcal{L}^*(gh)] = \mathcal{L}\mathcal{L}^*(\mathbb{E}[gh])$ . Next,

$$\begin{aligned} \rho(z, t) &= M(z, t) - \mu(z)\mu(t)^T \\ &= \ell(z-t)M(t, t) + \mathcal{L}_{t,z}(S\mathbb{E}[F\mu^{F,X_0}(t)^T]) - (\ell(z-t)\mu(t) + \mathcal{L}_{t,z}(S\mu_F))\mu(t)^T \\ &= \ell(z-t)(M(t, t) - \mu(t)\mu(t)^T) + \mathcal{L}_{t,z}(S\mathbb{E}[F\mu^{F,X_0}(t)^T] - S \cdot \mu_F\mu(t)^T) \\ &= \ell(z-t)\Sigma(t) + \mathcal{L}_{t,z}(SCov(F, \mu^{F,X_0}(t))). \end{aligned}$$

Since, for the generic  $g$ ,  $(\mathcal{L}g)^T = \mathcal{L}^*(g^T)$ , the rightmost term of the last line can be expanded as  $\mathcal{L}_{t,z}(SCov(F, \ell(t)X(0) + \mathcal{L}_t(SF))) = \mathcal{L}_{t,z}(S\xi_F^T)\ell(t)^T + \mathcal{L}_{t,z}\mathcal{L}_t^*(S\rho_F S^T)$ . Finally, inserting the expression of  $\Sigma(t)$ , using the equality  $\ell(z-t)\ell(t) = \ell(z)$  and regrouping terms,

$$\begin{aligned} \rho(z, t) &= \ell(z-t)\mathcal{L}_t^\circ Q + \ell(z)\Sigma(0)\ell(t)^T + (\ell(z-t)\mathcal{L}_t\mathcal{L}_t^*(S\rho_F S^T) + \mathcal{L}_{t,z}\mathcal{L}_t^*(S\rho_F S^T)) + (\ell(z-t)\mathcal{L}_t(S\xi_F^T)\ell(t)^T + \mathcal{L}_{t,z}(S\xi_F^T)\ell(t)^T) + \ell(z)\mathcal{L}_t^*(\xi_F S^T) \\ &= \ell(z-t)\mathcal{L}_t^\circ Q + \ell(z)\Sigma(0)\ell(t)^T + \mathcal{L}_z\mathcal{L}_t^*(S\rho_F S^T) + \mathcal{L}_z(S\xi_F^T)\ell(t)^T + \ell(z)\mathcal{L}_t^*(\xi_F S^T). \end{aligned}$$

From the integral expressions of  $\mu(t)$  and  $\rho(z, t)$  just proven, Eq. (6)–(8) follow by taking derivatives.

**Proposition 3.** Given the strict stability of  $SW$ , the expression of  $\bar{\mu}$  is promptly found as the solution of (6) for  $\mu_F(\cdot) = \bar{\mu}_F$  and  $\dot{\mu} = 0$ . From the expression

of  $\rho(z, t) = \rho(t + \delta, t)$  in Prop. 2, since  $\ell(t)$  vanishes for  $t \rightarrow +\infty$  and  $\mathcal{L}_z(S\xi_F^T)$  is bounded because  $\xi_F$  is,  $\lim_{t \rightarrow +\infty} \rho(t + \delta, t)$  is given by the limit of

$$\mathcal{L}_t^* \mathcal{L}_{t+\delta}(S\rho_F S^T) + \ell(\delta) \mathcal{L}_t^\circ (S\text{diag}(W\mu + \mu_F) S^T).$$

For  $\rho_F(t + \delta, t) = \bar{\rho}_F(\delta)$ , the first term can be written as

$$\int_{-\infty}^t d\zeta' \int_{-\infty}^{t+\delta} d\tau' \ell(\tau') S \bar{\rho}_F(\delta + \zeta' - \tau') S^T \ell(\zeta')^T,$$

which converges to  $[\ell * (S\bar{\rho}_F S^T) * \ell_-^T](\delta)$  in the limit. Next, for  $\mu_F(\cdot) = \bar{\mu}_F$ , factor  $\mathcal{L}_t^\circ(\dots)$  of the second term is by definition the solution at  $t$  of

$$\begin{aligned} \dot{\mu} &= SW\mu + S\bar{\mu}_F, \\ \dot{\Sigma}^\circ &= SW\Sigma^\circ + \Sigma^\circ W^T S^T + S\text{diag}(W\mu + \bar{\mu}_F) S^T, \end{aligned}$$

with  $\Sigma^\circ = 0$ . For  $t \rightarrow +\infty$ ,  $\mu \rightarrow \bar{\mu}$  and  $\Sigma^0 \rightarrow \bar{\Sigma}^0$ , where  $\bar{\Sigma}^0$  must satisfy  $0 = SW\bar{\Sigma}^0 + \bar{\Sigma}^0 W^T S^T + \bar{Q}$ , as stated.