Cubification of $\sigma\pi$-SDE
and Exact Moment Equations

Alessandro Borri[1], Francesco Carravetta[1,*], Pasquale Palumbo[1]

*Istituto di Analisi dei Sistemi ed Informatica, Consiglio Nazionale delle Ricerche,
Via dei Taurini, 19, 00185 Roma, Italy*

Abstract

For an Itô-like Stochastic Differential Equation (SDE) system, with drift and diffusion that are *formal* polynomials of the independent variables, we show that all moments satisfy an infinite, countable, set of *linear* ordinary differential equations. This result is achieved by means of the *exact cubification* of the SDE, which consists in a set of *deterministic* transformations of the state variables, giving place to a new SDE with further finitely many state variables. Exact cubification can be considered as an extension to the 'stochastic case' of the *exact quadratization* of deterministic nonlinear systems, available in the literature. An example is finally shown, taken from systems biology, in which, for a basic chemical reaction network the exact moment equation is written down, and an approximate solution is calculated through a *moment closure* method.

Keywords: Stochastic Differential Equations, Moments Equations, Moments Closure, Exact Cubification, Exact Quadratization, $\sigma\pi$-functions.

1. Introduction

The problem of describing through an Ordinary Differential Equations (ODE) system the moments evolution of the distribution associated to a nonlinear Stochastic Differential Equations (SDE) system is, up to now, solved for a very restricted class of SDE only (essentially, for SDE with drift and

---
*Corresponding Author
  Email addresses: `alessandro.borri@iasi.cnr.it` (Alessandro Borri),
`francesco.carravetta@iasi.cnr.it` (Francesco Carravetta),
`pasquale.palumbo@iasi.cnr.it` (Pasquale Palumbo)

diffusion nonlinearities provided by *standard* polynomials, i.e. polynomials with non-negative integer exponents). Within the filtering theory framework, moments equations use to refer to *conditional* moments (with respect to some observed process), therefore they are stochastic processes instead of deterministic quantities, and they are not given in general as a *closed* system of SDE. In this regard it is shown in Liptser-Shiryayev (1977)[Ch. 8, eq. (8.10)] that for a very general class of partially observable processes (those given by the sum of an increasing process plus a martingale) an any-order conditional moment can be written as an Itô process, but not as the strong solution of a SDE, since this point is verified only for the smaller class of Conditionally Linear Systems (CLS), i.e. nonlinear SDEs which are linear for fixed values of the observed variable. In the latter case, it can be shown that the SDE for the i-rd order conditional moment depends on the $(i + 1)$-rd order one, eventually leading the moments equations to an *infinite dimensional* SDE. In case the process distribution of the CLS is conditionally Gaussian, which implies that any moment of such a distribution can be written as a function of the first two moments, this allows to obtain a *closed* system of SDEs for the first two conditional moments. For systems with non-stochastic coefficients, taking the expectation of the aforementioned conditional moments equations provides the moments equations as an ODE system, with closed ODE arising only for original linear SDE.

This paper investigates how to achieve the ODE system of the (unconditioned) moments of the probability distribution associated to a general class of nonlinear SDE extending that of standard polynomial nonlinearities for which the problem is known to be solved, though not in closed form. Motivation stems from the many applicative fields of stochastic modelling, where relevant knowledge from the system is usually inferred by means of first and second order moments. Among these fields, systems biology has shown an increasing interest in stochastic modelling and computation. It has been nowadays established the importance of the noise role in biological processes Mettetal et al. (2007). Random fluctuations, provided by a wide set of concurring factors including, for instance, thermal noise or asynchronous occurrence of synthesis and degradation events, need to be considered when modelling most of the molecular processes involved in cellular regulation, as well as in gene expression, see e.g. Bahar et al. (2006); Bruggerman et al. (2009); Kiviet et al. (2014). These processes are usually described, up to a discrete molecular scale, by the *Chemical Master Equation* (CME) or, up to a continuous macroscopic scale, by a SDE called *Chemical Langevin Equa-*

tion (CLE) Van Kampen (2007). Within this framework, the computation of moment equations could be fruitfully exploited to infer information on how noise propagates according to different network wirings or noise sources, with applications in metabolic as well as in enzymatic reactions or gene expression networks (see, e.g., Sontag et al. (2015); Oyarzun et al. (2014); Borri et al (2016b)). Provided that nonlinearities are standard polynomials, moment equations can be computed (though, usually, they are not in closed form) according to Hespanha et al. (2005); Singh et al. (2011) for both CME and CLE, actually for any generic stochastic hybrid system, i.e., a SDE (or an ODE) that allows discrete resets according to a continuous time Markov chain modeling discrete stochastic events. In those papers, suitable approximation schemes are also provided to close the moment equations.

In case of SDE with different nonlinearities than those of standard polynomials, to achieve the exact moment equations in explicit form is an open problem. The main contribution of this paper is to solve such a problem for a wide class of nonlinearities that non-trivially extends the one of standard polynomials and includes, for instance, *formal* polynomials, i.e., polynomials with generic real exponents. According to Carravetta (2015), formal polynomials will be also said $\sigma\pi$-functions. More in details, the exact moment equations in explicit form are achieved by means of a preliminary result, which can be regarded as the *stochastic version* of the *exact quadratization* of nonlinear deterministic systems described in Carravetta (2015), where it was proven that for a very large class of nonlinearities, there always exists a change of variables that transforms a finite set of nonlinear ODE into another (larger) set of ODE, in the old and new variables, with only *quadratic* nonlinearities. As we will see in the following, this result cannot be extended to SDE of Itô type, but it can be proven that the same set of change of variables used in the deterministic case gives place, in the stochastic case, to a *cubic* (instead of quadratic) SDE in an augmented set of differential stochastic unknowns. Similarly to the deterministic case, we call *exact cubification* the transformation of a $\sigma\pi$-SDE into a cubic one. We are not going to prove this result for the same class of nonlinearities considered in Carravetta (2015), and limit ourselves to the more restricted (but, nonetheless still wide) class of $\sigma\pi$-functions, representing indeed the *basic* class of functions for which exact quadratization was proven. However, the range of applicability of the exact cubification can be extended to the larger (than $\sigma\pi$) class of analytic ICF (Integral Closed Form) functions, according to the same procedure described in Carravetta (2015) that substantially consists in a preliminary set

of relations transforming the original ICF system into a $\sigma\pi$ one. In a preliminary version of the present theory Borri et al (2016), an example from the systems biology framework was presented, modelling a basic, though meaningful, case of gene regulation system. In that case, the SDE was in an ICF form, thus requiring a preliminary transformation into a $\sigma\pi$ system before to apply cubification. Here, we present a different case from chemical reaction networks with reaction rates modelled according to a generalised mass action law, a framework exploited when dealing with ultrasensitivity responses in systems biology (cf. Ferrell et al. (2014); Ferrell et al. (2014b,c)).

The paper is organised as follows. After §II, where main notations are introduced following the same lines of Carravetta (2015), including SDE in the $\sigma\pi$ formal definition, in §III the first main result is presented (exact cubification) showing that any $\sigma\pi$-SDE can be transformed into an equivalent *cubic* SDE, evolving, in general, in a larger dimensional state space. The compact vector form of exact cubification is reported in §IV. In §V the general (linear, infinite-dimensional) moments equation is achieved (which is the second main result). The applicative example from the systems biology framework is reported in §VI.

2. Notation

2.1. Indices convention

Throughout the paper we adopt the following *indices convention*. If a *scalar* quantity is defined through a multi-indexed symbol, for instance $\xi_{i_1,\dots,i_p} \in \mathbb{R}$, $(i_j = 1,\dots,\nu_j \in \mathbb{N})$, then omitting the rightmost index denotes a *column* vector, as follows:

$$\xi_{i_1,\dots,i_{p-1}} = [\xi_{i_1,\dots,i_{p-1},1},\dots,\xi_{i_1,\dots,i_{p-1},\nu_p}]^T \in \mathbb{R}^{\nu_p}, \tag{1}$$

omitting the two rightmost indices denotes the stack of the vectors $\xi_{i_1,\dots,i_{p-1}} \in \mathbb{R}^{\nu_p}$, for $i_{p-1} = 1,\dots,\nu_{p-1}$, and so on (up to the vector $\xi \in \mathbb{R}^{\nu_1\cdots\nu_p}$). We also sometimes denote by a single index, say $\mathbf{k}$, a double index (i,j), by writing $\xi_{\mathbf{k}} = \xi_{i,j}$, where $\mathbf{k}$ runs as follows:

$$\mathbf{k} = (1,1),\dots,(1,n_j),(2,1),\dots,(n_i,n_j).$$

Double indices are denoted as a bold latin character for the sake of clarity.

We denote by c the linear map (a matrix) swapping the first two indices of its argument, i.e.:

$$(c\xi)_{i_2,i_1,i_3,\ldots,i_p} = \xi_{i_1,\ldots,i_p}. \tag{2}$$

Moreover, we define c_{i_2} as the matrix extracting the i_2-rd sub-vector of $c\xi$ (or, in other words, the vector obtained by stacking ξ_{i_1,i_2} with respect to i_1):

$$c_{i_2}\xi = (c\xi)_{i_2}. \tag{3}$$

Notice that c_{i_2} (resp: c) is a $(\bar{\nu}_2 \times \nu)$-dimensioned (resp: $(\nu \times \nu)$-dimensioned) matrix, with

$$\nu = \nu_1 \cdot \nu_2 \ldots \cdot \nu_p; \quad \bar{\nu}_2 = \nu/\nu_2. \tag{4}$$

The calculation of the matrix c is given in Appendix A, Theorem 5. If $\xi \in \mathbb{R}^\alpha$ is a random vector, $\mathbf{E}\{\xi\}$ shall denote its expectation. The symbol I_α denotes the identity in $\mathbb{R}^\alpha$. The symbol $\delta_{i,j}$ is the discrete Kronecker function, defined as $\delta_{i,j} = 1$ for $i = j$ and $\delta_{i,j} = 0$ elsewhere.

2.2. $\sigma\pi$-SDE and their short forms

We consider a stochastic system in $\mathbb{R}^n$ whose state variable $x(t) \in \mathbb{R}^n$ is supposed to be a diffusion process well defined in some time interval $[0, T] \subset \mathbb{R}$ (where T can be $+\infty$ as well) as the unique strong solution – for a given random variable $\bar{x}$ assigned as initial condition: $x(0) = \bar{x}$ – of an Ito-type SDE:

$$dx(t) = f(t, x(t))dt + g(t, x(t))dW(t), \tag{5}$$

where f (resp. g) is a suitable vector function in $\mathbb{R}^n$ (resp. a suitable matrix function in $\mathbb{R}^{n \times d}$), and $W(t) \in \mathbb{R}^d$ is the standard Wiener process with incremental covariance $d\mathbf{E}\{WW^T\} = I_d dt$. [1] Let us consider equation (5) written component-wise, and omit the time-dependencies hereinafter:

$$dx_i = f_i(x)dt + \sum_{s=1}^{d} g_{i,s}(x)dW_s. \tag{6}$$

---

[1]Under mild hypotheses, which are essentially the same as for ordinary differential equations (i.e. that all systems functions are locally Lipschitz in their domain of definition) we can of course guarantee existence and unicity for a strong-solution over some interval $[0, T]$. The class of functions (formal polynomials) we are going to introduce, fulfills these requirements. However, note that one cannot assure in general that $T = +\infty$ with probability one, which, on the other hand, is not assured in the deterministic case either.

We consider the case of f_i and $g_{i,s}$ being $\sigma\pi$-*functions*, that is to say *formal* polynomials in the variable x, accordingly to the following formulas:

$$f_i(x) = \sum_{l=1}^{\nu_i} \alpha_{i,l} X_{i,l}; \quad X_{i,l} = \prod_{j=1}^{n} x_j^{p_{i,j}^l}, \tag{7}$$

$$g_{i,s}(x) = \sum_{\lambda=1}^{\nu_{i,s}^*} \alpha_{i,s,\lambda}^* X_{i,s,\lambda}^*; \quad X_{i,s,\lambda}^* = \prod_{j=1}^{n} x_j^{p_{i,j}^{*,s,\lambda}}, \tag{8}$$

where $p_{i,j}^l$ and $p_{i,j}^{*,s,\lambda}$ are *real* exponents, and $\alpha_{i,l}, \alpha_{i,s,\lambda}^*$ real parameters, time-varying in general. A stochastic system (6) with drift and diffusion given by $\sigma\pi$-functions, as in (7), (8), is said a $\sigma\pi$-*stochastic system*. We call the functions $X_{i,l}$ given in (7) *drift-monomials*, and the functions $X_{i,s,\lambda}^*$ given in (8) *diffusion-monomials*. By substituting (7), (8) into (6), and introducing the compound double index $\mathbf{k} = (s, \lambda) = (1, 1), \ldots (1, \nu_{i,1}^*), \ldots (d, \nu_{i,d}^*)$, we obtain:

$$dx_i = \sum_{l=1}^{\nu_i} \alpha_{i,l} X_{i,l} dt + \sum_{\mathbf{k}=(1,1)}^{(d,\nu_{i,d}^*)} \alpha_{i,\mathbf{k}}^* X_{i,\mathbf{k}}^* dW_{\mathbf{k}}^*, \tag{9}$$

$$W_{\mathbf{k}}^* = W_{s,\lambda}^* = W_s, \quad \forall \lambda = 1, \ldots, \nu_{i,s}^*. \tag{10}$$

Let us define

$$\nu_i^* = \nu_{i,1}^* + \ldots + \nu_{i,d}^*, \tag{11}$$

which is the number of values that the double index $\mathbf{k}$ takes on, and let ι an enumeration of these values, i.e. an invertible function ι such that $\mathbf{k} = \iota(l^*)$, for $l^* = 1, \ldots, \nu_i^*$. Equation (9) can be written in a compact form by introducing the *formal coefficients* $v_{i,l}$ defined as:

$$v_{i,l} = \alpha_{i,l} dt, \quad \text{for} \qquad l = 1, \ldots, \nu_i, \tag{12}$$

$$v_{i,l} = \alpha_{i,\iota(l-\nu_i)}^* dW_{\iota(l-\nu_i)}^*, \quad l = \nu_i + 1, \ldots, \mu_i, \tag{13}$$

where $\mu_i = \nu_i + \nu_i^*$, and extending the definition of $\alpha_{i,l}, X_{i,l}$ for $l = \nu_i + 1, \ldots, \mu_i$ as follows

$$\alpha_{i,l} = \alpha_{i,\iota(l-\nu_i)}^*, \quad X_{i,l} = X_{i,\iota(l-\nu_i)}^*. \tag{14}$$

Then eq. (9) turns into the short form:

$$dx_i = \sum_{l=1}^{\mu_i} v_{i,l} X_{i,l} = \bar{v}_i^T \bar{X}_i, \tag{15}$$

where the barred symbols $\bar{v}_i$ (resp: $\bar{X}_i$) are used to indicate that the indices convention has been applied with respect to the *extended* definitions of $v_{i,l}$ (resp: $X_{i,l}$) i.e. for the index l running up to μ_i. We will use the standard notation v_i (resp: X_i) when the indices convention is applied to the original $v_{i,l}$ (resp: $X_{i,l}$), i.e. the index l runs up to ν_i only.

Note that X_i^* is a sub-vector of $\bar{X}_i$, and in fact:

$$\bar{X}_i = [(X_i)^T, (X_i^*)^T]^T, \tag{16}$$

and thus we have:

$$X_i^* = \epsilon_i \bar{X}_i, \tag{17}$$

for a suitable $\{0,1\}$-valued matrix ϵ_i. Moreover, $X \in \mathbb{R}^{r-r*}$, $X^* \in \mathbb{R}^{r^*}$, with

$$r = \mu_1 + \ldots + \mu_n, \tag{18}$$
$$r^* = \nu_1^* + \ldots, \nu_n^*, \tag{19}$$

and we have

$$X^* = \epsilon \bar{X}, \qquad \text{with} \quad \epsilon = \text{diag}\{\epsilon_1, \ldots, \epsilon_n\}. \tag{20}$$

The short form (15) is a SDE where the drift and diffusion are hidden into the formal coefficients $v_{i,l}$. Another short form, which does not hide the drift and the diffusion, can be obtained by directly applying the convention on the indices to the coefficients α and α^*. Thus f_i and $g_{i,s}$ in (7), (8) rewrite

$$f_i(x) = \alpha_i^T X_i, \tag{21}$$
$$g_{i,s}(x) = \alpha_{i,s}^{*T} X_{i,s}^*, \tag{22}$$

and (6) becomes:

$$dx_i = \alpha_i^T X_i dt + \sum_{s=1}^d \alpha_{i,s}^{*T} X_{i,s}^* dW_s. \tag{23}$$

In the following we use both the representations (15) and (23). The reader has to be aware that $X_{i,s}^*$ in (23) is a vector, while $X_{i,\mathbf{k}}^*$ is a scalar, and indeed the latter is a quantity with three indices (remind that $\mathbf{k}$ in (9) is a double index) and $X_{i,s}^*$ is the vector which aggregates the $X_{i,\mathbf{k}}^*$ $(= X_{i,s,\lambda}^*)$ by saturating the index λ, accordingly with the convention on indices.

3. Exact cubification of $\sigma\pi$-SDE.

Recall that the Itô formula, for a general (measurable and scalar) function $\phi(x)$, and an Itô process x satisfying (6), writes as follows (cf. Liptser-Shiryayev (1977)):

$$d\phi(x) = \sum_{j=1}^{n} \frac{\partial \phi}{\partial x_j} dx_j + \frac{1}{2} \sum_{j,j',s}^{n,n,d} g_{j,s}(x) g_{j',s}(x) \frac{\partial^2 \phi}{\partial x_j \partial x_{j'}} dt. \tag{24}$$

The main theoretical result of the paper is stated in the following Theorem.

Theorem 1. *Let us consider the Ito-type $\sigma\pi$ stochastic system described (in short form) by eq. (15). Let us define, for $i = 1, \ldots, n$ and $l = 1, \ldots, \mu_i$ the functions $Z_{i,l}(x)$:*

$$Z_{i,l} = X_{i,l} x_i^{-1}, \quad l = 1, \ldots, \nu_i, \tag{25}$$

$$Z_{i,l} = Z_{i,\iota(l-\nu_i)}^*, \quad l = \nu_i + 1 + \ldots, \mu_i, \tag{26}$$

where, for $\mathbf{k} = (1,1), \ldots, (d, \nu_{i,d}^*)$, the functions $Z_{i,\iota(l-\nu_i)}^*$ are defined as:

$$Z_{i,\mathbf{k}}^* = X_{i,\mathbf{k}}^* x_i^{-1}. \tag{27}$$

Moreover let the exponents $\pi_{i,j}^l$ be defined as:

$$\pi_{i,j}^l = p_{i,j}^l - \delta_{i,j}, \quad l = 1, \ldots, \nu_i, \tag{28}$$

$$\pi_{i,j}^l = \pi_{i,j}^{*,\iota(l-\nu_i)}, \quad l = \nu_i + 1 + \ldots, \mu_i, \tag{29}$$

$$\pi_{i,j}^{*,\mathbf{k}} = p_{i,j}^{*,\mathbf{k}} - \delta_{i,j}, \quad \mathbf{k} = (1,1), \ldots, (d, \nu_{i,d}^*), \tag{30}$$

$$\tilde{\pi}_{i,j,j'}^l = \pi_{i,j}^l (\pi_{i,j'}^l - \delta_{j,j'}). \tag{31}$$

Then, the processes x_i and $Z_{i,l}$ satisfy the (Ito-type) SDE:

$$dx_i = (\alpha_i^T Z_i) x_i dt + \sum_{s=1}^{d} (\alpha_{i,s}^{*T} Z_{i,s}^*) x_i dW_s, \tag{32}$$

$$dZ_{i,l} = \sum_{j=1}^{n} \pi_{i,j}^l (\alpha_j^T Z_j) Z_{i,l} dt + \frac{1}{2} \sum_{j,j',s}^{n,n,d} \tilde{\pi}_{i,j,j'}^l (\alpha_{j,s}^{*T} Z_{j,s}^*)(\alpha_{j',s}^{*T} Z_{j',s}^*) Z_{i,l} dt$$

$$+ \sum_{j,s}^{n,d} \pi_{i,j}^l (\alpha_{j,s}^{*T} Z_{j,s}^*) Z_{i,l} dW_s, \tag{33}$$

for $i = 1, \ldots, n$, and $l = 1, \ldots, \mu_i$.

8

Proof. Equation (32) comes directly from the definition of $Z_{i,l}$ given in (25). By Itô formula (24), with $\phi(x) = Z_{i,l}(x)$, using (15), we have

$$dZ_{i,l} = \sum_{j=1}^{n} \frac{\partial Z_{i,l}}{\partial x_j} (\bar{v}_j^T \bar{X}_j) + \chi, \tag{34}$$

where χ, through (22), is given by:

$$\chi = \frac{1}{2} \sum_{j,j',s}^{n,n,d} (\alpha^{*T}_{j,s} X^*_{j,s})(\alpha^{*T}_{j',s} X^*_{j',s}) \frac{\partial^2 Z_{i,l}}{\partial x_j \partial x_{j'}} dt. \tag{35}$$

By (25) and the definition of $X_{i,l}$ given in (7) one has

$$\begin{aligned}
\frac{\partial Z_{i,l}}{\partial x_j} &= \frac{\partial}{\partial x_j} X_{i,l} x_i^{-1} = \frac{\partial}{\partial x_j} \prod_{s=1}^{n} x_s^{p^l_{i,s}} x_i^{-1} \\
&= p^l_{i,j} X_{i,l} x_j^{-1} x_i^{-1} - X_{i,l} x_i^{-2} \delta_{i,j} = \pi^l_{i,j} X_{i,l} x_j^{-1} x_i^{-1} \\
&= \pi^l_{i,j} Z_{i,l} x_j^{-1}
\end{aligned}$$

which multiplied by $\bar{v}_j^T \bar{X}_j$, again by (25) yields

$$\frac{\partial Z_{i,l}}{\partial x_j} (\bar{v}_j^T \bar{X}_j) = \pi^l_{i,j} Z_{i,l} (\bar{v}_j^T \bar{Z}_j), \tag{36}$$

where

$$\bar{Z}_j = \bar{X}_j x_j^{-1}. \tag{37}$$

Moreover

$$\begin{aligned}
\frac{\partial^2 Z_{i,l}}{\partial x_j \partial x_{j'}} &= \frac{\partial}{\partial x_{j'}} \frac{\partial Z_{i,l}}{\partial x_j} = \frac{\partial}{\partial x_{j'}} \pi^l_{i,j} Z_{i,l} x_j^{-1} = \pi^l_{i,j} \frac{\partial}{\partial x_{j'}} \left(Z_{i,l} x_j^{-1} \right) \\
&= \pi^l_{i,j} \pi^l_{i,j'} Z_{i,l} x_j^{-1} x_{j'}^{-1} - \pi^l_{i,j} Z_{i,l} x_j^{-2} \delta_{j,j'} = \pi^l_{i,j} Z_{i,l} x_j^{-1} x_{j'}^{-1} (\pi^l_{i,j'} - x_j^{-1} x_{j'} \delta_{j,j'}). \tag{38}
\end{aligned}$$

Observe that $\pi^l_{i,j'} - x_j^{-1} x_{j'} \delta_{j,j'} = \pi^l_{i,j'} - \delta_{j,j'}$, therefore by (31), eq. (38) rewrites

$$\frac{\partial^2 Z_{i,l}}{\partial x_j \partial x_{j'}} = \tilde{\pi}^l_{i,j,j'} Z_{i,l} x_j^{-1} x_{j'}^{-1} \tag{39}$$

which, replaced into (35) yields

$$\chi = \frac{1}{2} \sum_{j,j',s}^{n,n,d} \tilde{\pi}^l_{i,j,j'} (\alpha^{*T}_{j,s} Z^*_{j,s})(\alpha^{*T}_{j',s} Z^*_{j',s}) Z_{i,l} dt \tag{40}$$

Now, from (15), (23), and recalling (37), we have

$$\bar{v}_j^T \bar{Z}_j = \alpha_j^T Z_j dt + \sum_{s=1}^{d} \alpha^{*T}_{j,s} Z^*_{j,s} dW_s,$$

which has to be used into (36). The result is

$$\frac{\partial Z_{i,l}}{\partial x_j}(\bar{v}_j^T \bar{X}_j) = \pi^l_{i,j}(\alpha_j^T Z_j) Z_{i,l} dt + \sum_{s=1}^{d} \pi^l_{i,j}(\alpha^{*T}_{j,s} Z^*_{j,s}) Z_{i,l} dW_s, \qquad (41)$$

which, used into (34), with (40) replaced as well into (34), finally yields (33).
●

Similarly as in Carravetta (2015), we name the SDE (33) the *stochastic driver*, and the *bilinear* SDE (32), *the stochastic final stage*, associated to the SDE (6). We name the SDE systems constituted by the stochastic driver and final stage *an exact cubification* of the original SDE in (6).

4. Cubification in vector form

By using Kronecker products (see Appendix A), we can gather equations (32) and (33) in order to obtain a couple of vector equations. Note that (33) is an autonomous SDE, in the vector variable $\bar{Z}$ aggregating (37), since Z^* just collects a part of the entries of $\bar{Z}$, as one can read out from (25), (27).

Let us consider the vectors $\alpha_i, \alpha^*_{i,s}$ in the final stage equation (32), and the coefficients $\pi, \tilde{\pi}$ defined in (28)–(31). Define

$$\beta_{i,l,j,l'} = \pi^l_{i,j}\alpha_{j,l'}, \qquad (42)$$

$$\gamma_{s,i,l,j,m} = \pi^l_{i,j}(c\alpha^*)_{s,j,m}, \qquad (43)$$

$$h_{s,i,l,j,m,j',m'} = \tilde{\pi}^l_{i,j,j}(c\alpha^*)_{s,j,m}(c\alpha^*)_{s,j',m'}, \qquad (44)$$

where, looking at (8), and on account of (2), we infer that the indices s, j, m range from $(1,1,1)$ to $(d,n,\nu^*_{j,s})$, and thus c is a square matrix having dimension $q = n \cdot d \cdot r^*$, with r^* given by (19). The matrix c can be calculated by applying Theorem 5 to the present case ($\nu_1 = n$, $\nu_2 = d$, $\nu_3 = r^*$) the result is:

$$c = C_{n,d}^T \otimes I_{r^*}. \qquad (45)$$

where $C_{n,d}^T$ are commutation matrices defined in Appendix (Theorem 5). Note that c has q rows, which can be partitioned into d subsets of rows, each of cardinality $n \cdot r^*$. Then, accordingly with the notation stated in (2), in the following we denote by c_s the $(n \cdot \nu^* \times q)$-dimensional matrix obtained from c by extracting the s-th subset of rows. That said, let us build up the matrices

$$A = \mathrm{diag}\{\alpha_1^T, \ldots, \alpha_n^T\}, \tag{46}$$

$$A_s^* = \mathrm{diag}\{(c\alpha^*)_{s,1}^T, \ldots, (c\alpha^*)_{s,n}^T\}, \tag{47}$$

$$\mathbf{A} = \mathrm{diag}\{A_1^T, \ldots, A_n^T\}, \tag{48}$$

$$\mathbf{A}_s^* = \mathrm{diag}\{(A_s^*)_1^T, \ldots, (A_s^*)_n^T\}, \tag{49}$$

$$\mathbf{F} = \mathrm{diag}\{\beta_{1,1}^T, \ldots, \beta_{n,\mu_n}^T\}, \tag{50}$$

$$\mathbf{H}_s^* = \mathrm{diag}\{h_{s,1,1}^T, \ldots, h_{s,n,\mu_n}^T\}, \tag{51}$$

$$\mathbf{G}_s^* = \mathrm{diag}\{\gamma_{s,1,1}^T, \ldots, \gamma_{s,n,\mu_n}^T\}, \tag{52}$$

where A_i (resp: $(A_s^*)_i$) denotes the i-th column of the matrix A (resp: A_s^*). Moreover, let us define

$$\mathbf{B}_s = \mathbf{A}_s^*(I_n \otimes c_s)(I_n \otimes \epsilon), \tag{53}$$

$$\mathbf{H} = \frac{1}{2}\sum_{s=1}^{d}\mathbf{H}_s^*(I_r \otimes c_s^{[2]})(I_r \otimes \epsilon^{[2]}), \tag{54}$$

$$\mathbf{G}_s = \mathbf{G}_s^*(I_r \otimes c_s)(I_r \otimes \epsilon), \tag{55}$$

where ϵ is the matrix defined through (17), (20). We can prove the following Theorem.

Theorem 2. *(Cubification in vector form). The final stage and driver are given, in vector form, by:*

$$dx = \mathbf{A}(x \otimes Z)dt + \sum_{s=1}^{d}\mathbf{B}_s(x \otimes Z)dW_s, \tag{56}$$

$$dZ = \left(\mathbf{F}Z^{[2]} + \mathbf{H}Z^{[3]}\right)dt + \sum_{s}\mathbf{G}_s Z^{[2]}dW_s. \tag{57}$$

By defining the aggregate process $\mathbf{Z} \in \mathbb{R}^{n+r}$:

$$\mathbf{Z}^T = [x^T, Z^T], \tag{58}$$

11

we can write the exact cubification (56), (57) according to the following compact SDE:

$$dZ = (\Phi Z^{[2]} + \Psi Z^{[3]})dt + \sum_{s=1}^{d} \Gamma_s Z^{[2]} dW_s. \tag{59}$$

with

$$\Phi = \begin{bmatrix} 0 & \mathbf{A} & 0 & 0 \\ 0 & \cdots & 0 & \mathbf{F} \end{bmatrix} \in \mathbb{R}^{(n+r)\times(n+r)^2}; \tag{60}$$

$$\Psi = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & \mathbf{H} \end{bmatrix} \in \mathbb{R}^{(n+r)\times(n+r)^3}; \tag{61}$$

$$\Gamma_s = \begin{bmatrix} 0 & \mathbf{B}_s & 0 & 0 \\ 0 & \cdots & 0 & \mathbf{G}_s \end{bmatrix} \in \mathbb{R}^{(n+r)\times(n+r)^2}. \tag{62}$$

Proof. The following preliminary results are required. Given two vectors $v, w \in \mathbb{R}^p$ we define $v \circ w \in \mathbb{R}^p$ as $(v \circ w)_i = v_i w_i$. Let M be a $\nu_1 \times \nu_2$ matrix, and $x \in \mathbb{R}^{\nu_2}$, $y \in \mathbb{R}^{\nu_1}$. Then, there obviously exists a matrix $\mathbf{M}$, $(\nu_1 \times \nu_1 \nu_2)$-dimensioned, such that:

$$(Mx) \circ y = \mathbf{M}(y \otimes x). \tag{63}$$

As a matter of fact, such a matrix $\mathbf{M}$ can be easily derived by noting that the i-th entry $(i = 1, \ldots, \nu_1)$ of the vector $(Mx) \circ y$ is $\sum_j M_{i,j} y_i x_j$, and $y_i x_j$ is the (i, j)-th entry of the $\nu_1 \nu_2$ dimensioned vector $y \otimes x$. Thus, defining $\mathbf{M} = \text{diag}\{M_1^T, \ldots, M_n^T\}$ we get (63).

By suitably exploiting these results, since $(\alpha_{i,s}^{*T} Z_{i,s}^*)x_i = ((c\alpha^*)_{s,i}^T (cZ^*)_{s,i})x_i$, by aggregating in i for eq. (32), we obtain

$$dx = (AZ) \circ x dt + \sum_{s=1}^{d} (A_s^*(cZ^*)_s) \circ x dW_s$$

$$= \mathbf{A}(x \otimes Z)dt + \sum_{s=1}^{d} \mathbf{A}_s^*(x \otimes (cZ^*)_s))dW_s. \tag{64}$$

Let us consider the three terms on the right-hand side of (33), we have

$$\pi_{i,j}^l \alpha_{j,l'} Z_{j,l'} = \beta_{i,l,j,l'} Z_{j,l'}$$

$$\tilde{\pi}_{i,j,j'}^l (\alpha^*_{j,s,m} Z^*_{j,s,m})(\alpha^*_{j',s,m'} Z^*_{j',s,m'})$$

$$= \tilde{\pi}_{i,j,j'}^l (c\alpha^*)_{s,j,m} (c\alpha^*)_{s,j',m'} (cZ^*)_{s,j,m} (cZ^*)_{s,j',m'}$$

$$= h_{s,i,l,j,m,j',m'} (cZ^*)_{s,j,m} (cZ^*)_{s,j',m'}$$

$$\pi_{i,j}^l \alpha^*_{j,s,m} Z^*_{j,s,m} = \pi_{i,j}^l (c\alpha^*)_{s,j,m} (cZ^*)_{s,j,m} = \gamma_{s,i,l,j,m} (cZ^*)_{s,j,m},$$

12

thus, taking the sums:

$$\sum_{j,l'} \pi^l_{i,j} \alpha_{j,l'} Z_{j,l'} = \beta^T_{i,l} Z, \tag{65}$$

$$\sum_{j,j',m,m'} \tilde{\pi}^l_{i,j,j'} (\alpha^*_{j,s,m} Z^*_{j,s,m})(\alpha^*_{j',s,m'} Z^*_{j',s,m'}) = h^T_{s,i,l}(cZ^*)^{[2]}_s, \tag{66}$$

$$\sum_{j,m} \pi^l_{i,j} \alpha^*_{j,s,m} Z^*_{j,s,m} = \gamma^T_{s,i,l}(cZ^*)_s. \tag{67}$$

Notice that the same matrix c swaps the first two indices in both α^* and Z^*. Substituting in (33), we have

$$dZ_{i,l} = (\beta^T_{i,l} Z) Z_{i,l} dt + \frac{1}{2} \sum_s^d h^T_{s,i,l}(cZ^*)^{[2]}_s Z_{i,l} dt + \sum_s^d \gamma^T_{s,i,l}(cZ^*)_s Z_{i,l} dW_s. \tag{68}$$

Let F, H^*_s, G^*_s be the matrices whose (i,l)-th rows are $\beta^T_{i,l}, h^T_{s,i,l}, \gamma^T_{s,i,l}$, respectively. By aggregating with respect to the double index (i,l) eq. (68), we obtain

$$dZ = \Big((FZ) \circ Z + \frac{1}{2} \sum_s^d (H^*_s(cZ^*)^{[2]}_s) \circ Z \Big) dt + \sum_s^d (G^*_s(cZ^*)_s) \circ Z dW_s$$

$$= \Big(\mathbf{F}Z^2 + \frac{1}{2} \sum_s^d (\mathbf{H}^*_s(Z \otimes (cZ^*)^{[2]}_s)) \Big) dt + \sum_s^d (\mathbf{G}^*_s(Z \otimes (cZ^*)_s) dW_s. \tag{69}$$

By suitably exploiting the Knonecker product properties (see, e.g. Carravetta et al. (1996) and references therein), we have

$$x \otimes (cZ^*)_s = (I_n \cdot x) \otimes (c_s Z^*) = (I_n \otimes c_s)(x \otimes Z^*)$$
$$= (I_n \otimes c_s)(I_n \cdot x \otimes \epsilon Z) = (I_n \otimes c_s)(I_n \otimes \epsilon)(x \otimes Z),$$

which, with Z playing the role of x, and recalling that $Z \in \mathbb{R}^r$ turns into:

$$Z \otimes (cZ^*)_s = (I_r \otimes c_s)(I_r \otimes \epsilon)Z^{[2]}.$$

Moreover,

$$Z \otimes (cZ^*)^{[2]}_s = (I_r \cdot Z) \otimes (c_s Z^*)^{[2]}$$
$$= (I_r \cdot Z) \otimes c^{[2]}_s Z^{*[2]} = (I_r \otimes c^{[2]}_s)(Z \otimes Z^{*[2]})$$
$$= (I_r \otimes c^{[2]}_s)(I_r \cdot Z \otimes \epsilon^{[2]} Z^{[2]}) = (I_r \otimes c^{[2]}_s)(I_r \otimes \epsilon^{[2]})Z^{[3]}.$$

The identities above, used in (64), (69), yield (56), (57). Finally, by exploiting (58), we have

$$
\mathbf{Z}^{[2]} = \begin{bmatrix} x^{[2]} \\ x \otimes Z \\ Z \otimes x \\ Z^{[2]} \end{bmatrix}, \qquad
\mathbf{Z}^{[3]} = \begin{bmatrix} x^{[3]} \\ x^{[2]} \otimes Z \\ x \otimes Z \otimes x \\ x \otimes Z^{[2]} \\ Z \otimes x^{[2]} \\ Z \otimes x \otimes Z \\ Z^{[2]} \otimes x \\ Z^{[3]} \end{bmatrix}, \tag{70}
$$

from which we readily derive the matrices (60)–(62).   ●

5. Equations of the Moments

Consider the exact cubification form described by (59) and suppose that all moments of $\mathbf{Z}$ exist and are finite, namely:

$$
\mathbf{m}^h(t) = \mathbf{E}\{\mathbf{Z}^{[h]}(t)\} < +\infty, \quad t \in [0, +\infty), \quad h = 1, 2, \dots
$$

Note that all moments of h-th order of $x(t)$ are included in the entries of $\mathbf{m}^h(t)$. In order to write the equations of all $\mathbf{m}^h$, $h = 1, 2 \dots$, we need some additional tools that can be found in Carravetta et al. (2000), and we report here for the ease of the reader.

For any $\mathcal{C}^\infty$ matrix function $F : \mathbb{R}^n \mapsto \mathbb{R}^{m \times p}$, we introduce the differential operator $(d/dx)\otimes$ defined as

$$
\frac{d}{dx} \otimes F(x) = \begin{bmatrix} \dfrac{\partial F(x)}{\partial x_1} & \cdots & \dfrac{\partial F(x)}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times np}. \tag{71}
$$

In case of a vectorial map $F : \mathbb{R}^n \mapsto \mathbb{R}^m$, $(d/dx)\otimes$ provides the Jacobian matrix. Higher-order derivatives of F can be represented by the following recursive formalism

$$
\frac{d^{[i]}}{dx^{[i]}} \otimes F(x) = \frac{d}{dx} \otimes \left[\frac{d^{[i-1]}}{dx^{[i-1]}} \otimes F(x) \right] \in \mathbb{R}^{m \times n^i p}. \tag{72}
$$

In case of polynomial transformations (denoted according to the Kronecker formalism, see the Appendix), the following formulas involved in the first and second order derivatives will be exploited.

14

Lemma 1. *(Lemma 4.1 in Carravetta et al. (2000)) For any integer $h \geq 1$ and $x \in \mathbb{R}^n$, it results*

$$\frac{d}{dx} \otimes x^{[h]} = U_n^h(I_n \otimes x^{[h-1]}) \tag{73}$$

and for any $h > 1$:

$$\frac{d^{[2]}}{dx^{[2]}} \otimes x^{[h]} = O_n^h(I_{n^2} \otimes x^{[h-2]}), \tag{74}$$

where

$$U_n^h = \left(\sum_{\tau=0}^{h-1} C_{n,n^{h-1-\tau}}^T \otimes I_{n^\tau} \right), \tag{75}$$

$$O_n^h = \sum_{\tau=0}^{h-1} \sum_{s=0}^{h-2} (C_{n,n^{h-1-\tau}}^T \otimes I_{n^\tau})(I_n \otimes C_{n,n^{h-2-s}}^T \otimes I_n). \tag{76}$$

and matrices $C_{u,v}^T$, $u, v \in \mathbb{N}$, are commutation matrices (see Appendix, Theorem 5).

Theorem 3. *The set of all moments $\mathbf{m}^h$, for $h \in \mathbb{N} \setminus \{0\}$ of the random process $\mathbf{Z} \in \mathbb{R}^q$, with $q = n + r$, defined by the SDE (59), is given by the following infinite set of ODE:*

$$\dot{\mathbf{m}}^h = \Phi_h \mathbf{m}^{h+1} + \Psi_h \mathbf{m}^{h+2}, \qquad h = 1, 2, \dots \tag{77}$$

where

$$\Phi_h = U_q^h(\Phi \otimes I_{q^{h-1}}), \tag{78}$$

$$\Psi_h = U_q^h(\Psi \otimes I_{q^{h-1}}) + \sum_{s=1}^{d} O_q^h(\Gamma_s^{[2]} \otimes I_{q^{h-2}}), \tag{79}$$

with Φ, Ψ, Γ_s being the matrices defined in (60)-(61)-(62), and U_q^h, O_q^h being the matrices defined in (75), (76).

Proof. In order to get a SDE for the processes $\mathbf{Z}^{[h]}$, we need the Itô formula in the Kronecker formalism (see Theorem 5.2 in Carravetta et al. (2000)):

$$d\mathbf{Z}^{[h]} = \left(\frac{d}{d\mathbf{Z}} \otimes \mathbf{Z}^{[h]} \right) d\mathbf{Z} + \frac{1}{2} \left(\frac{d^{[2]}}{d\mathbf{Z}^{[2]}} \otimes \mathbf{Z}^{[h]} \right) \sum_{s=1}^{d} \Gamma_s^{[2]} \mathbf{Z}^{[4]} dt. \tag{80}$$

15

By properly exploiting Lemma 1 we have

$$\frac{d}{d\mathbf{Z}} \otimes \mathbf{Z}^{[h]} = U_q^h(I_q \otimes \mathbf{Z}^{[h-1]}), \qquad \frac{d^{[2]}}{d\mathbf{Z}^{[2]}} \otimes \mathbf{Z}^{[h]} = O_q^h(I_{q^2} \otimes \mathbf{Z}^{[h-2]}),$$

so that one has

$$\left(\frac{d}{d\mathbf{Z}} \otimes \mathbf{Z}^{[h]}\right) d\mathbf{Z} = U_q^h(I_q \otimes \mathbf{Z}^{[h-1]})\left(\Phi\mathbf{Z}^{[2]} + \Psi\mathbf{Z}^{[3]}\right) dt$$
$$+ \sum_{s=1}^d U_q^h(I_q \otimes \mathbf{Z}^{[h-1]})\Gamma_s\mathbf{Z}^{[2]}dW_s, \tag{81}$$

and

$$\left(\frac{d^{[2]}}{d\mathbf{Z}^{[2]}} \otimes \mathbf{Z}^{[h]}\right)\sum_{s=1}^d \Gamma_s^{[2]}\mathbf{Z}^{[4]}dt = \sum_s O_q^h(I_{q^2} \otimes \mathbf{Z}^{[h-2]})\Gamma_s^{[2]}\mathbf{Z}^{[4]}dt. \tag{82}$$

By means of the usual tensor product calculations we obtain (by using the formulas in the Appendix):

$$(I_q \otimes \mathbf{Z}^{[h-1]})\Phi\mathbf{Z}^{[2]} = (I_q \otimes \mathbf{Z}^{[h-1]})\left((\Phi\mathbf{Z}^{[2]}) \otimes 1\right) = (\Phi\mathbf{Z}^{[2]}) \otimes \mathbf{Z}^{[h-1]}$$
$$= (\Phi \otimes I_{q^{h-1}})\mathbf{Z}^{[h+1]},$$

and, similarly:

$$(I_q \otimes \mathbf{Z}^{[h-1]})\Psi\mathbf{Z}^{[3]} = (\Psi \otimes I_{q^{h-1}})\mathbf{Z}^{[h+2]},$$
$$(I_q \otimes \mathbf{Z}^{[h-1]})\Gamma_s\mathbf{Z}^{[2]} = (\Gamma_s \otimes I_{q^{h-1}})\mathbf{Z}^{[h+1]},$$
$$(I_{q^2} \otimes \mathbf{Z}^{[h-2]})\Gamma_s^{[2]}\mathbf{Z}^{[4]} = (\Gamma_s^{[2]} \otimes I_{q^{h-2}})\mathbf{Z}^{[h+2]}$$

which, used in (81) and (82), turn equation (80) into:

$$d\mathbf{Z}^{[h]} = \left(\Phi_h\mathbf{Z}^{[h+1]} + \Psi_h\mathbf{Z}^{[h+2]}\right)dt + \sum_s^d \Gamma_s^h\mathbf{Z}^{[h+1]}dW_s, \tag{83}$$

with Φ_h, Ψ_h given by (78)–(79), and

$$\Gamma_s^h = U_q^h(\Gamma_s \otimes I_{q^{h-1}}). \tag{84}$$

Taking expectations in (83) we finally derive (77). ●

16

6. A systems biology application

The applicative example is taken from a basic chemical reaction network, aiming at describing ultra-sensitivity response, a phenomenon widely investigated in systems biology, the challenging research line conceived as a mixture of life science investigation, rigorous mathematical modeling and engineering methodologies. In this framework, stochastic modeling revealed to be an unavoidable means to properly investigate the inherent stochasticity of the many and diverse noise sources affecting the molecular and/or cellular processes under investigation Mettetal et al. (2007); Bahar et al. (2006); Bruggerman et al. (2009); Kiviet et al. (2014). To this end, to achieve the whole probability distribution of all the species under investigation could be not computationally affordable (or even unreliable); therefore, first- and second-order moments are usually sought as the proper compromise between the obtainable information and the computational burden. Indeed, the second-order moments allow to quantify the stochastic variability around the steady-state average solution and have been recently investigated in the synthetic biology framework with the aim of quantifying noise reduction in presence of feedback Oyarzun et al. (2014); Sontag et al. (2015); Borri et al (2016b).

Ultra-sensitivity response has been recently investigated in detail, seeming to be a potentially important biochemical property and a commonplace in cell signaling Ferrell et al. (2014); Ferrell et al. (2014b,c). The example under investigation takes inspiration from signaling cascades like, e.g., the three-step Mitogen-Activated Protein Kinase (MAPK) cascade: each step can be thought of as an input/output module where the input is the active kinase that catalyzes the activation of a downstream kinase. The following example considers one such module with W^* being the active kinase playing the role of the input that catalyzes the activation of another kinase $Y \rightleftharpoons Y^\star$ (asterisks will denote activated species). The ratio $Y^\star/(Y+Y^\star)$ is the output of the module, with $Y^\star$ being the input of another downstream module. In Ferrell et al. (2014c), it is shown that the upstream signal can be better transmitted without degradation through the cascade of subsequent activations if the steady-state response of each module is described by a steep Hill function of the type

$$\frac{Y^\star}{Y + Y^\star} = \frac{(W^\star)^m}{(W^\star)^m + \theta^m} \tag{85}$$

where m is the Hill coefficient providing a measure of the steepness of the Hill

function (i.e. a measure of the sensitivity of the cascade module response). The higher is m, the more accurate is the signal transmission through the cascade. Thinking at the entries of 85 as steady-state concentrations, the Hill coefficient can be estimated from experimental data, therefore it may well result as a real number. In terms of a reverse engineering problem, when looking for the chemical reaction network providing the experimental steady-state results, it readily comes out that these can be obtained according to the following bidirectional biochemical reaction

$$Y + mW^\star \underset{h_{\text{off}}}{\overset{h_{\text{on}}}{\rightleftharpoons}} Y^\star \tag{86}$$

where W, Y (as well as their activated counterparts $W^\star$, $Y^\star$) are the species involved and $h_{\text{on}}, h_{\text{off}} \in \mathbb{R}^+$ are the on/off reaction coefficients. If m is treated as the usual *stoichiometric coefficient*, according to mass-action law, we can write the forward/backward reaction rates

$$v^+ = h_{\text{on}} Y (W^\star)^m \qquad v^- = h_{\text{off}} Y^\star \tag{87}$$

so that, at steady state, when $v^+ = v^-$, by properly exploiting the mass constraint $Y + Y^\star = Y_{TOT}$, we find again the input/output relationship (85).

Standard mass action law usually sets stoichiometric coefficients as integer numbers. On the other hand, the ultrasensitivity is a typical case where generalized mass-action laws would be preferred (see also Vlad et al (2009)), so that it is meaningful to consider the reactions (86) with m being any positive real number.

According to the discrete stochastic modeling framework, define $\bar{x} = [w^\star \ y^\star]^T$ as the vector associated to the chemical reaction network, which uniquely determines the state of the system, thanks to the mass constraints. Symbols $w, w^\star, y, y^\star$ denote the copy numbers of species $W, W^\star, Y, Y^\star$, respectively, and $y_T = y + y^\star$ refers to the overall (and constant) amount of species Y. Then, the following reactions of degradation/production can be considered:

$$\text{Reaction 1} : \begin{cases} w^\star \mapsto w^\star - m \\ y^\star \mapsto y^\star + 1 \end{cases} \quad \begin{aligned} & N_1 = [-m \quad 1]^T \\ & a_1(\bar{x}) = h_{\text{on}}(w^\star)^m (y_T - y^\star), \end{aligned}$$

$$\text{Reaction 2} : \begin{cases} w^\star \mapsto w^\star + m \\ y^\star \mapsto y^\star - 1, \end{cases} \quad \begin{aligned} & N_2 = [m \quad -1]^T \\ & a_2(\bar{x}) = h_{\text{off}} y^\star, \end{aligned}$$

where N_i is the *stoichiometric vector* of reaction $i = 1, 2$ (i.e. reaction i resets $\bar{x}$ to $\bar{x} + N_i$) and $a_i(\bar{x})$ is its *propensity* (or probability for unit time). In the stochastic approach, reaction propensities and stoichiometry completely define a continuous-time Markov chain, whose dynamics (in terms of probability distribution) is represented in a discrete and noisy way by the chemical master equations (CMEs) Van Kampen (2007). In the general case, however, the CME is solvable neither in explicit nor in numerical form, since its complexity grows exponentially with the number of species involved. An alternative formulation, which dramatically reduces the complexity of the underlying CME, consists in switching from discrete to continuous states by means of the chemical Langevin equation (CLE, see e.g. Van Kampen (2007); Khanin et al. (2008)), whose computational cost is just linear with respect to the number of species. The CLE is a stochastic differential equation describing the evolution in time of the copy number $\bar{x}(t)$ (approximated as continuous rather than discrete), as follows:

$$d\bar{x}(t) = \sum_{i=1}^{2} N_i a_i(\bar{x}(t))dt + \sum_{i=1}^{2} N_i \sqrt{a_i(\bar{x}(t))}dW_i(t), \tag{88}$$

where $W_i(t)$ are mutually independent standard Wiener processes. By expanding the CLE in (88), which is in the general form (5), by defining the state variables $x_1 := w^\star$, $x_2 := y^\star$ and omitting the time dependencies, one obtains the following equations:

$$dx_1 = \left(-mh_{\text{on}}x_1^m(y_T - x_2) + mh_{\text{off}}x_2\right)dt - m\sqrt{h_{\text{on}}x_1^m(y_T - x_2)}dW_1$$
$$+ m\sqrt{h_{\text{off}}x_2}dW_2,$$
$$dx_2 = \left(h_{\text{on}}x_1^m(y_T - x_2) - h_{\text{off}}x_2\right)dt + \sqrt{h_{\text{on}}x_1^m(y_T - x_2)}dW_1 - \sqrt{h_{\text{off}}x_2}dW_2,$$
$$\tag{89}$$

which do not define a $\sigma\pi$-stochastic system in the form (6)–(8). In order to turn the system (89) into a $\sigma\pi$-stochastic system, we define an additional state variable $x_3 := y_T - x_2$, so that the evolutions can be rewritten as:

$$dx_1 = \left(-mh_{\text{on}}x_1^m x_3 + mh_{\text{off}}x_2\right)dt - m\sqrt{h_{\text{on}}x_1^m x_3}dW_1 + m\sqrt{h_{\text{off}}x_2}dW_2,$$
$$dx_2 = \left(h_{\text{on}}x_1^m x_3 - h_{\text{off}}x_2\right)dt + \sqrt{h_{\text{on}}x_1^m x_3}dW_1 - \sqrt{h_{\text{off}}x_2}dW_2,$$
$$dx_3 = -dx_2 = \left(-h_{\text{on}}x_1^m x_3 + h_{\text{off}}x_2\right)dt - \sqrt{h_{\text{on}}x_1^m x_3}dW_1 + \sqrt{h_{\text{off}}x_2}dW_2.$$
$$\tag{90}$$

19

We stress that the coefficient m being any real number prevents writing finite-order moments in a closed way, unless the proposed cubification methodology is applied. By defining the augmented state $x = [x_1 \quad x_2 \quad x_3]^T$ and the noise vector $W = [W_1 \quad W_2]^T$, it is readily seen that the system (90) is $\sigma\pi$, namely it can be rewritten in the form (6)–(8), with the following settings (zero drift/diffusion exponents are omitted):

- $n = 3$ (state dimension)

- $d = 2$ (dimension of the Wiener process)

- $\nu_i = 2$, for $i = 1, 2, 3$ (number of drift monomials for each state equation i)

- $\nu^*_{i,s} = 1$, for $i = 1, 2, 3$, for $s = 1, 2$ (number of diffusion monomials for each state equation i and for each noise component s)

- $\alpha_{1,1} = -mh_{\text{on}}, \qquad \alpha_{1,2} = mh_{\text{off}},$
  $\alpha_{2,1} = -\alpha_{3,1} = h_{\text{on}}, \qquad \alpha_{2,2} = -\alpha_{3,2} = -h_{\text{off}}$ (drift coefficients)

- $X_{i,1} = x_1^m x_3, \qquad X_{i,2} = x_2, \qquad i = 1, 2, 3$ (drift monomials)

- $p_{i,1}^1 = m$, $p_{i,3}^1 = p_{i,2}^2 = 1$, $p_{i,1}^1 = p_{i,1}^2 = p_{i,3}^2 = 0 \quad i = 1, 2, 3$ (drift exponents)

- $\alpha^*_{1,1,1} = -m\sqrt{h_{\text{on}}}, \qquad \alpha^*_{1,2,1} = m\sqrt{h_{\text{off}}}, \qquad \alpha^*_{2,1,1} = -\alpha^*_{3,1,1} = \sqrt{h_{\text{on}}},$
  $\alpha^*_{2,2,1} = -\alpha^*_{3,2,1} = -\sqrt{h_{\text{off}}}$ (diffusion coefficients)

- $X^*_{1,1,1} = X^*_{2,1,1} = X^*_{3,1,1} = x_1^{\frac{m}{2}} x_3^{\frac{1}{2}},$
  $X^*_{1,2,1} = X^*_{2,2,1} = X^*_{3,2,1} = x_2^{\frac{1}{2}}$ (diffusion monomials)

- $p^{*,1,1}_{1,1} = p^{*,1,1}_{2,1} = p^{*,1,1}_{3,1} = \frac{m}{2}, \qquad p^{*,1,1}_{i,2} = p^{*,2,1}_{i,3} = p^{*,2,1}_{i,1} = 0,$
  $p^{*,1,1}_{1,3} = p^{*,1,1}_{2,3} = p^{*,1,1}_{3,3} = p^{*,2,1}_{1,2} = p^{*,2,1}_{2,2} = p^{*,2,1}_{3,2} = \frac{1}{2}$ (diffusion exponents)

Note that, since $d = 2$ and $\nu^*_{i,s} = 1$, for $i = 1, 2, 3$, then one gets:

$$\nu^*_i = \sum_{s=1}^{d} \nu^*_{i,s} = d = 2 \qquad i = 1, 2, 3,$$
$$\mu_i = \nu_i + \nu^*_i = 2 + d = 4 \qquad i = 1, 2, 3. \tag{91}$$

From (10) one can equivalently define, for all i, the compound double index $k = (s, l)$, with $s = 1, 2$, $l = \nu_{i,s}^* = 1$, from which we can extend the definition of $\alpha_{i,l}$ and $X_{i,l}$, in agreement with (14). Such quantities are included in the vectors α_i and X_i, for all i:

$$\begin{aligned}
\alpha_i &= [\alpha_{i,1} \quad \alpha_{i,2} \quad \alpha_{i,1,1}^* \quad \alpha_{i,2,1}^*]^T \in \mathbb{R}^{\mu_i}, \\
X_i &= [X_{i,1} \quad X_{i,2} \quad X_{i,1,1}^* \quad X_{i,2,1}^*]^T \in \mathbb{R}^{\mu_i},
\end{aligned} \tag{92}$$

which are useful, in turn, to define the vector Z in (25)–(27) as follows:

$$Z = [X_1^T x_1^{-1} \quad X_2^T x_2^{-1} \quad X_3^T x_3^{-1}]^T \in \mathbb{R}^r,$$

with $r = \mu_1 + \mu_2 + \mu_3 = 12$. Finally, the aggregate process $\mathbf{Z}$ in (58) is defined as:

$$\mathbf{Z} = [x^T \quad Z^T]^T \in \mathbb{R}^q,$$

with $q = n + r = 15$. The settings above allow to compute all the matrices required to explicitly write the moment equation (77). Since we are interested in the second-order moments, we consider the following system of $q + q^2 = 240$ differential equations:

$$\begin{aligned}
\dot{\mathbf{m}}^1 &= \Phi_1 \mathbf{m}^2 + \Psi_1 \mathbf{m}^3, \\
\dot{\mathbf{m}}^2 &= \Phi_2 \mathbf{m}^3 + \Psi_2 \mathbf{m}^4.
\end{aligned} \tag{93}$$

An approximate solution of (93) can be obtained by using a moment closure procedure, as e.g. in Hespanha et al. (2005); Singh et al. (2011). We do not drift along into this topic, which goes beyond the scope of the paper: we just point out that numerical simulation results would strongly depend on the adopted moment closure approximation scheme.

7. Conclusion

It is well known in the literature how to write exact moment equations for nonlinear SDE with drift and diffusion terms provided by standard polynomials. Here we extend such a result to the case of drift and diffusion terms provided by formal polynomials, i.e., polynomials with real exponents. The result is achieved by means of a suitable embedding of the original nonlinear system into a larger state space where nonlinearities are at most third-order standard polynomials with respect to the new state variables (*cubification*).

Similarly to the case of standard-polynomial nonlinearities, the achieved exact moment equations cannot be straightforwardly solved, since they require a suitable moment closure technique. A motivating example taken from systems biology is reported to show a nontrivial framework where the present theory could be applied.

Appendix A. Tensor algebras

Here follows, for the ease of the readers, a list of the main definitions and few properties of Kronecker products, up to an extent just sufficient for the present paper. For more details we address the reader to Carravetta et al. (2000), and Rodgers (1980). The interested reader, familiar with algebraic concepts, is referred to Dummit et al. (2004) for more insight about the interplay with tensor algebra.

For $u \in \mathbb{R}^n$, $v \in \mathbb{R}^m$, we define their *Kronecker product* $u \otimes v$ as follows:

$$u \otimes v = [u_1 v^T, \ldots u_n v^T]^T \in \mathbb{R}^{nm}, \tag{A.1}$$

Of course, definition (A.1) can be applied even if u, v are matrices, by identifying $u \in \mathbb{R}^{n \times m}$ to an element of $\mathbb{R}^{nm}$. The Knonecker product is associative and distributive with respect to sum of vectors, however it is non-commutative. Moreover, we recall the following properties. For A, B, C, D suitably defined matrices, and u, v vectors, provided that both sides are well defined, the following identities hold (see Carravetta et al. (2000)–Rodgers (1980)):

$$(A \cdot C) \otimes (B \cdot D) = \quad (A \otimes B) \cdot (C \otimes D), \tag{A.2}$$
$$(A \otimes B)^T = \quad A^T \otimes B^T, \tag{A.3}$$
$$\mathrm{st}(A \cdot B \cdot C) = \quad (C^T \otimes A) \cdot \mathrm{st}(B), \tag{A.4}$$
$$u \otimes v = \quad \mathrm{st}(v \cdot u^T), \tag{A.5}$$
$$\mathrm{tr}(A \otimes B) = \quad \mathrm{tr}(A) \cdot \mathrm{tr}(B), \tag{A.6}$$

where $\mathrm{tr}(M)$ denotes the trace of a square matrix M, and $\mathrm{st}(M)$ is the stack of any matrix M (i.e. the vector piling up the columns of M). Although the Kronecker product is not commutative, the following property holds (see Carravetta et al. (2000) for a proof).

Theorem 4. *For any given pair of matrices $A \in \mathbb{R}^{r \times s}$, $B \in \mathbb{R}^{n \times m}$, we have*

$$B \otimes A = C_{r,n}^T (A \otimes B) C_{s,m} , \tag{A.7}$$

where the commutation matrix $C_{u,v}$ is the $(u \cdot v) \times (u \cdot v)$ matrix such that its (h, l) entry is given by:

$$\{C_{u,v}\}_{h,l} = \delta_{l,(|h-1|_v)u+\left(\left[\frac{h-1}{v}\right]+1\right)} \qquad (A.8)$$

where $[\cdot]$ and $|\cdot|_s$ denote integer part and s-modulo, respectively.

Observe that $C_{1,1} = 1$, hence in the vector case when $a \in \mathbb{R}^r$ and $b \in \mathbb{R}^n$, (A.7) becomes

$$b \otimes a = C_{r,n}^T(a \otimes b). \qquad (A.9)$$

We can now find the expression of the matrix c defined in (2)

Theorem 5. *The matrix c, defined in (2), is given by*

$$c = C_{\nu_1,\nu_2}^T \otimes I_{\nu_3 \cdots \nu_p}, \qquad (A.10)$$

Proof. Let $u^j \in \mathbb{R}^{\nu_j}$, $j = 1, \ldots, p$, a set of p vectors, and $h = \{h_1, \ldots, h_p\}$ a permutation of the ordered set $\{1, \ldots, p\}$. Denote by $\mathcal{H}$ the set of all of these permutations. Define the following sets of symbols, namely S_1, S_2:

$$S_1 = \{\xi_{i_{h_1},\ldots,i_{h_p}} : \quad h \in \mathcal{H}\}, \qquad S_2 = \{u_{i_{h_1}}^{h_1} \cdots u_{i_{h_p}}^{h_p} : \quad h \in \mathcal{H}\},$$

Then, we have the natural bijection between S_1 and S_2:

$$\xi_{i_{h_1},\ldots,i_{h_p}} \leftrightarrow u_{i_{h_1}}^{h_1} \cdots u_{i_{h_p}}^{h_p}. \qquad (A.11)$$

By (A.9), and using formula (A.2), we have

$$u^2 \otimes u^1 \otimes u^3 \otimes \ldots \otimes u^p = \left[C_{\nu_1,\nu_2}^T(u^1 \otimes u^2)\right] \otimes u^3 \otimes \ldots \otimes u^p$$
$$= \left(C_{\nu_1,\nu_2}^T \otimes I_{\nu_3 \cdots \nu_p}\right)(u^1 \otimes u^2 \otimes u^3 \otimes \ldots \otimes u^p) \qquad (A.12)$$

Moreover, by reason of the bijection (A.11), looking at (2) we realize that (component-wise):

$$c\xi \leftrightarrow u^2 \otimes u^1 \otimes u^3 \otimes \ldots \otimes u^p \qquad \xi \leftrightarrow u^1 \otimes u^2 \otimes \ldots \otimes u^p$$

through which, identity (A.12) implies $c\xi = \left(C_{\nu_1,\nu_2}^T \otimes I_{\nu_3 \cdots \nu_p}\right)\xi$. Thus, (A.10) ensues.

23

References

R.S. Liptser, and A.N. Shiryayev, *Statistics of Random Processes – General Theory*, (Springer-Verlag, 1977).

F. Carravetta, SIAM J. Cont & Opt. **53**, 235 (2015).

U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman and Hall/CRC, 2006).

E. Altszyler, A.C. Ventura, A. Colman-Lerner, A. Chernomoretz, PLoS ONE, **12** (2017).

R. Bahar, C.H. Hartmann, K.A. Rodriguez, A.D. Denny, R.A. Busuttil, M.E. Dollè, R.B. Calder, G.B. Chisholm, B.H. Pollock, C.A. Klein and J. Vijg, Nature, **441**, 1011 (2006).

A. Borri, F. Carravetta, and P. Palumbo, Proc. IEEE Conf. Dec. & Control, Las Vegas, USA, 4540 (2016).

A. Borri, P. Palumbo, A. Singh, IET Syst. Biol, 1 (2016).

F.J. Bruggeman, N. Blüthgen and H.V. Westerhoff, PLoS Computational Biology, **5**, 1 (2009).

F. Carravetta, A. Germani and M. Raimondi, SIAM J. Cont. & Opt, **34**, 1666 (1996).

F. Carravetta, A. Germani, M.K. Shuakayev, SIAM J. Cont & Opt, **38**, 1171 (2000).

D. S. Dummit, R. M. Foote *Abstract Algebra*, (John Wiley & Sons, Third Edition, 2004).

J.E. Ferrell, Jr, and S.H. Ha, Trends Biochem. Sci., **39** 496 (2014).

J.E. Ferrell, Jr, and S.H. Ha, Trends Biochem. Sci., **39**, 556 (2014).

J.E. Ferrell, Jr, and S.H. Ha, Trends Biochem. Sci., **39**, 612 (2014).

J.P. Hespanha and A. Singh, Int. J. of Robust and Nonlinear Control, **15**, 669 (2005).

R. Khanin and D. J. Higham, Theoretical Computer Science, **408**, 31 (2008).

D.J. Kiviet, P. Nghe, N. Walker, S. Boulineau, V. Sunderlikova, S.J. Tans, Nature, **514**, 376 (2014).

P. Mendes, W. Sha and K. Ye, Bioinformatics, **19**, 122 (2003).

J. Mettetal and A. van Oudenaarden, Science, **317**, 463 (2007).

Oyarzun, D.A, Lugagne, J.-B., Stan, G.-B.V., ACS Synth. Biol., 116 (2014).

G.S. Rodgers, *Matrix Derivatives*, (Marcel Dekker, 1980).

A. Singh and J.P. Hespanha, IEEE Transactions on Automatic Control. **56**, 414 (2011).

M. Soltani, C. Vargas and A. Singh, IEEE Transactions on Biomedical Circuits and Systems, **9**, 518 (2015).

E.D. Sontag and A. Singh, IEEE Life Sicence Letters, **1**, 26 (2015).

N. Soranzo, G. Bianconi and C. Altafini, Bioinformatics, **23** 1640 (2007).

N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*, (North Holland, third edition, 2007).

M.O. Vlad, A.D. Corlan, F. Morán, R. Spang, P. Oefner and J. Ross, Proc. Natl. Acad. Sci. USA, **106** 6465 (2009).

J.N. Weiss, The FASEB Journal, **11**, 835 (1997).