Cubification of Nonlinear Stochastic Differential Equations and Approximate Moments Calculation of the Langevin Equation

Alessandro Borri, Francesco Carravetta, and Pasquale Palumbo

***Abstract*— For the class of Ito-type nonlinear Stochastic Differential Equations (SDE), where the drift and the diffusion are $\sigma\pi$-functions ($\sigma\pi$-SDE), we prove that the (infinite) set of all moments of the solution satisfies a system of infinite ordinary differential equations (ODEs), which is always linear. The result is proven by showing first that a $\sigma\pi$-SDE can be *cubified*, i.e. reduced to a system of SDE of larger (but still finite) dimension in general, where drifts and diffusions are at most third-degree polynomial functions. Our motivation for deriving a moment equation in closed form comes from systems biology, where second-order moments are exploited to quantify the stochastic variability around the steady-state average amount of the molecular players involved in a bio-chemical reaction framework. Indeed, the proposed methodology allows to write the moment equations in the presence of non-polynomial nonlinearities, when exploiting the Chemical Langevin Equations (which are SDE) as a model abstraction. An example is given, associated to a protein-gene production model, where non-polynomial nonlinearities are known to occur.**

I. INTRODUCTION

The problem of describing through an ODE the moments evolution of the distribution associated to a nonlinear SDE, is up to now solved for a very restricted class of systems only (essentially: for polynomial systems). Of course, 'equations of moments' can be found in the literature, even as far as *conditional* moments (with respect to some observed process) are concerned, but they are not given in general as a *closed* system of SDEs. In regard to this, an equation – which is called *the general equation of optimal nonlinear filtering* (OFE) – can be found in the §8 of [1] (in particular: eq. (8.10)). Such equation is the result of showing that, for a very general class of partially observable processes (those given by a sum of an increasing process plus a martingale) an any-order conditional-moment can be written as an Ito process, but not as the strong solution of a SDE, which is shown to be possible only for the class of *conditionally linear* systems (CLS i.e. nonlinear SDEs which are linear for fixed values of the observed variables). In the latter case, it can be shown that the OFE for the i-th order conditional moment, depends on the $i+1$-th order one, and thus the OFE is equivalent to an *infinite dimensional* SDE. A further step is showing that for CLS the process distribution is conditionally Gaussian, which implies that any moment of such a distribution can be written as a function of the first two moments, and this allows to obtain a *closed* system of SDEs for the first two conditional moments[1]. For systems with non-stochastic coefficients, taking the expectation of the above equations leads directly to a closed system of ODEs for the *unconditional* moments, but in this way the class of systems reduces to just the linear ones.

In this paper we are concerned with the sub-problem of calculating unconditional moments for nonlinear SDEs, which is by itself an important problem in many areas of science and engineering where the quantity of interest is the probability distribution of a random process. We will show that a wide class of nonlinear SDE – namely those whose drift and/or diffusion is expressed through ratios of *$\sigma\pi$-functions*[2] – can be *exactly cubified* (i.e. polynomialized as a third degree polynomial). This will allow us to write the exact moment equation for the same class of SDE, which results to be a (deterministic) *linear system*, even though *infinite dimensional*, and in particular, a kind of infinite-dimensional system given by an infinite set of ODE's.

Our seeking of a general moment equations for nonlinear SDE's is motivated, in particular, by the applications in systems biology, where the steady-state of the second order moments of the chemical players involved in biochemical reaction networks allow to quantify the stochastic variability around the steady-state average solutions and have been recently investigated in the synthetic biology framework with the aim of quantifying noise reduction in presence of feedbacks [6], [7], [14]. When approaching the problem by means of Chemical Master Equations, it is known how to write the moment equations when involved nonlinearities are polynomials [15]. It is known as well that, even when nonlinearities are polynomials, the solution cannot be written in closed form, and approximations are required to achieve a solution; among this framework, a widely adopted possibility is based on moment closure techniques [10]. Coherently with our theory developed for SDE's, here we adopt the Chemical Langevin Equations (CLE) to model a chemical reaction network. It is known that CLE dramatically reduce the complexity of the underlining Chemical Master Equations, at the expense of the approximation that the amount of the species under investigation is modeled continuously rather than in the proper discrete fashion as for CME [9], [12]. Unfortunately, the use of CLE does not solve the problem to write the moment equations for non-polynomial systems. It worths noticing that non-polynomial nonlinearities are not so seldom

All authors are with the Istituto di Analisi dei Sistemi ed Informatica 'A. Ruberti', Consiglio Nazionale delle Ricerche, Roma 00185, Italy. PP is with SYSBIO.IT Center for Systems Biology, Italy. {alessandro.borri, francesco.carravetta, pasquale.palumbo}@iasi.cnr.it

[1] Which is a *conditionally Gaussian* Kalman Filter.

[2] A quite large class of nonlinear functions which has been studied in [2], in regard to the problem of the *exact quadratization* for nonlinear deterministic control systems.

to find: for instance, sigmoidal functions (like logistic, or more complex Hill functions) are widely exploited to model saturations naturally arising in biological frameworks [4]. Instead, it can be shown that such nonlinearities can be modeled as $\sigma\pi$-functions, thus providing the natural ground to apply our proposed methodology. Even with the proposed approach, moment equations can be exactly written, but they are not in closed form, thus we require as well some moment closure technique to obtain a solution.

The paper is organized into two sections. In §II the class of $\sigma\pi$-systems is defined, and the first main result is presented (exact cubification) showing that any $\sigma\pi$-system can be transformed into an equivalent *cubic* SDE, evolving, in general, in a larger dimensional state space. Then the general (linear, and infinite-dimensional) moments equation is written down (which is the second main result of the paper). In §II we present an example of application: a very common and simple model taken from system biology, a typical network motif in gene transcription networks, for which we show how to jointly apply our moments equation and the moment closure method in order to compute the first two moments of the associated distribution.

II. Cubification of $\sigma\pi$-SDE's.

A. Notation

Throughout the paper we adopt the following convention: If a *scalar* quantity has been 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 the omission of the rightmost index shall denote the *column* vector which collects the related entries, the omission of the two rightmost indices will denote 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 $\xi \in \mathbb{R}^{\nu_1\cdots\nu_p}$). and so on. We also sometimes denote by a single index, say k, a double index (i,j), by writing $\xi_k = \xi_{i,j}$, where k runs as follows:

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

We denote by c the linear map (a matrix) swapping the first two indices of its argument, i.e.[3]:

$$(c\xi)_{i_2,i_1,i_3,\dots,i_p} = \xi_{i_1,\dots,i_p}. \tag{1}$$

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

$$c_{i_2}\xi = (c\xi)_{i_2}. \tag{2}$$

If $\xi \in \mathbb{R}^\alpha$ is a random vector, $\mathbf{E}\{\xi\}$ shall denote the expectation. The symbol I_α will denote the identity in $\mathbb{R}^\alpha$.

[3]Note that c is a *commutation matrix*, for the Kronecker (tensor) product between vectors, i.e. for any couple x, y of vectors (even of different dimension), c is the matrix such that $x \otimes y = c(y \otimes x)$.

B. Problem Setting

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 $[0,T] \subset \mathbb{R}$ as the unique strong solution of an Ito-type SDE

$$dx(t) = f(t,x(t))dt + g(t,x(t))dW(t), \tag{3}$$

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 $d\mathbf{E}\{WW^T\} = I_d dt$. Let us consider equation (3) written component-wise, and omit the time-dependencies hereinafter:

$$dx_i = f_i(x)dt + \sum_{s=1}^{d} g_{i,s}(x)dW_s. \tag{4}$$

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:

$$f_i(x) = \sum_{l=1}^{\nu_i} \alpha_{i,l} X_{i,l}; \quad X_{i,l} = \prod_{j=1}^{n} x_j^{p^l_{i,j}}, \tag{5}$$

$$g_{i,s}(x) = \sum_{l=1}^{\nu^*_{i,s}} \alpha^*_{i,s,l} X^*_{i,s,l}; \quad X^*_{i,s,l} = \prod_{j=1}^{n} x_j^{p^{*,s,l}_{i,j}}, \tag{6}$$

where $p^l_{i,j}$ and $p^{*,s,l}_{i,j}$ are *real* exponents, and $\alpha_{i,l}, \alpha^*_{i,s,l}$ real parameters, time-varying in general. A stochastic system (4) with drift and diffusion given by $\sigma\pi$-functions, as in (5), (6), is said a *$\sigma\pi$-stochastic system*. We call the functions $X_{i,l}$ given in (5) *drift-monomials*, and the functions $X^*_{i,s,l}$ given in (6) *diffusion-monomials*. By substituting (5), (6) into (4), and introducing the compound double index $k = (s,l) = (1,1),\dots(1,\nu^*_{i,1}),\dots(d,\nu^*_{i,d})$, we obtain:

$$dx_i = \sum_{l=1}^{\nu_i} \alpha_{i,l} X_{i,l} dt + \sum_{k=(1,1)}^{(d,\nu^*_{i,d})} \alpha^*_{i,k} X^*_{i,k} dW^*_k. \tag{7}$$

$$W^*_k = W^*_{s,l} = W_s, \quad \forall l = 1,\dots,\nu^*_{i,s}. \tag{8}$$

Let us define $\nu^*_i = \nu^*_{i,1} + \dots + \nu^*_{i,d}$, which is the number of values that the double index k takes on, and let ι an enumeration of these values, i.e. a function (invertible) such that, for any k there is an $l = 1,\dots,\mu_i = \nu_i + \nu^*_i$ such that $k = \iota(l)$. Equation (7) can be written in a compact form by introducing the *formal coefficients* $v_{i,l}$ defined as:

$$v_{i,l} = \alpha_{i,l} dt, \quad \text{for} \quad l = 1,\dots,\nu_i, \tag{9}$$

$$v_{i,l} = \alpha^*_{i,\iota(l-\nu_i)} dW^*_{\iota(l-\nu_i)}, \quad l = \nu_i+1,\dots,\mu_i, \tag{10}$$

and extending the definition of $\alpha_{i,l}, X_{i,l}$ as

$$\alpha_{i,l} = \alpha^*_{i,\iota(l-\nu_i)}, \;\; X_{i,l} = X^*_{i,\iota(l-\nu_i)}, \;\; \text{for} \;\; l = \nu_i+1,\dots,\mu_i, \tag{11}$$

from which eq. (7) turns into the short form:

$$dx_i = \sum_{l=1}^{\mu_i} v_{i,l} X_{i,l} = v_i^T X_i. \tag{12}$$

Note that X_i^* is a subvector of X_i – indeed, the vector including the last $\mu_i - \nu_i$ entries of X_i – and thus we have:

$$X_i^* = \epsilon_i X_i, \tag{13}$$

for a suitable $0-1$ matrix ϵ_i. Moreover, $X \in \mathbb{R}^{\mathrm{r}}$, $X^* \in \mathbb{R}^{\mathrm{r}^*}$, with $r = \mu_1 + \ldots + \mu_n$, $r^* = \nu_1^* + \ldots, \nu_n^*$, and we have

$$X^* = \epsilon X, \qquad \text{with} \quad \epsilon = \mathrm{diag}\{\epsilon_1, \ldots, \epsilon_n\}. \tag{14}$$

The short form (12) 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 in fact obtained by directly applying the convention on the indices to the coefficients α, and α^*. Thus f_i and $g_{i,s}$ in (5), (6) rewrites

$$f_i(x) = \alpha_i^T X_i, \tag{15}$$
$$g_{i,s}(x) = {\alpha^*}_{i,s}^T X_{i,s}^*, \tag{16}$$

and (4) becomes:

$$dx_i = \alpha_i^T X_i dt + \sum_{s=1}^{d} {\alpha^*}_{i,s}^T X_{i,s}^* dW_s. \tag{17}$$

In the following we use both the representations (12) and (17). The reader is to be aware that $X_{i,s}^*$ in (17) is a vector, while $X_{i,k}^*$ is a scalar, and indeed the latter is a quantity with three indices (remind that k in (7) is a double index) and $X_{i,s}^*$ is the vector aggregating the $X_{i,k}^*$ $(= X_{i,s,l}^*)$, by saturating the index l, accordingly with the convention on indices.

C. Cubification of $\sigma\pi$ SDE's.

Recall that the Ito formula, for a general (scalar) function $\phi(x)$, and an Ito process x satisfying (4), writes as follows:

$$d\phi(x) = \sum_{j=1}^{n} \frac{\partial \phi}{\partial x_j} dx_j + \frac{1}{2} \sum_{j,j',s}^{n,n,d} g_{j,s}(x) g_{j',s}(x) \frac{\partial^2 \phi}{\partial x_j \partial x_{j'}} dt. \tag{18}$$

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. (12). Let us define the functions $Z_{i,l}(x)$, $(l = \ldots, \mu_i)$:*

$$Z_{i,l} = X_{i,l} x_i^{-1}, \quad l = 1, \ldots, \nu_i \tag{19}$$
$$Z_{i,l} = Z^*_{i,\iota(l-\nu_i)}; \quad l = \nu_i + 1 + \ldots, \mu_i, \tag{20}$$
$$Z^*_{i,k} = X^*_{i,k} x_i^{-1}, \ \ k \quad (= (s,l)) = (1,1), \ldots, (d, \nu^*_{i,d}) \tag{21}$$

and the exponents $\pi_{i,j}^l$ as ($\delta_{i,j}$ Kronecker symbol):

$$\pi_{i,j}^l = p_{i,j}^l - \delta_{i,j} \qquad l = 1, \ldots, \nu_i \tag{22}$$
$$\pi_{i,j}^l = \pi_{i,j}^{*,\iota(l-\nu_i)}; \quad l = \nu_i + 1 + \ldots, \mu_i, \tag{23}$$
$$\pi_{i,j}^{*,k} = p_{i,j}^{*,k} - \delta_{i,j} \quad k \ \ (= (s,l)) \ \ = (1,1), \ldots, (d, \nu^*_{i,d}) \tag{24}$$
$$\tilde{\pi}_{i,j,j'}^l = \pi_{i,j}^l (\pi_{i,j'}^l - \delta_{j,j'}) \tag{25}$$

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

$$dx_i = (\alpha_i^T Z_i) x_i dt + \sum_{s=1}^{d} ({\alpha^*}_{i,s}^T Z_{i,s}^*) x_i dW_s, \tag{26}$$

$$\begin{aligned} dZ_{i,l} &= \sum_{j=1}^{n} \pi_{i,j}^l (\alpha_j^T Z_j) Z_{i,l} dt \\ &+ \frac{1}{2} \sum_{j,j',s}^{n,n,d} \tilde{\pi}_{i,j,j'}^l ({\alpha^*}_{j,s}^T Z_{j,s}^*)({\alpha^*}_{j',s}^T Z_{j',s}^*) Z_{i,l} dt \\ &+ \sum_{j,s}^{n,d} \pi_{i,j}^l ({\alpha^*}_{j,s}^T Z_{j,s}^*) Z_{i,l} dW_s, \end{aligned} \tag{27}$$

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

Proof. (*Omitted*).

Similarly as in [2], we name the SDE (27) the *stochastic driver*, and the *bilinear* SDE (26), *the stochastic final stage*, associated to the SDE (4). We name the system of SDE's constituted by the stochastic driver and final stage *an exact cubification* of the SDE (4). Note that (27) is an autonomous SDE (in Z), since Z^* just collects a part of the entries of Z, as one can read out from (19), (21).

D. Vector form

Let us consider the vectors $\alpha_i, \alpha_{i,s}^*$ in the final stage equation (26) and define

$$\beta_{i,l,j,l'} = \pi_{i,j}^l \alpha_{j,l'}. \tag{28}$$
$$\gamma_{s,i,l,j,m} = \pi_{i,j}^l (c\alpha^*)_{s,j,m} \tag{29}$$
$$h_{s,i,l,j,m,j',m'} = \tilde{\pi}_{i,j,j'}^l (c\alpha^*)_{s,j,m} (c\alpha^*)_{s,j',m'}. \tag{30}$$

Moreover let us build up the matrices

$$A = \mathrm{diag}\{\alpha_1^T \ldots, \alpha_n^T\}, \tag{31}$$
$$A_s^* = \mathrm{diag}\{(c\alpha^*)_{s,1}^T \ldots, (c\alpha^*)_{s,n}^T\}, \tag{32}$$
$$\mathbf{A} = \mathrm{diag}\{A_1^T, \ldots, A_n^T\}, \tag{33}$$
$$\mathbf{A}_s^* = \mathrm{diag}\{(A_s^*)_1^T, \ldots, (A_s^*)_n^T\}, \tag{34}$$
$$\mathbf{F} = \mathrm{diag}\{\beta_{1,1}^T, \ldots, \beta_{n,\mu_n}^T\}, \tag{35}$$
$$\mathbf{H}_s^* = \mathrm{diag}\{h_{s,1,1}^T, \ldots, h_{s,n,\nu_n}^T\}, \tag{36}$$
$$\mathbf{G}_s^* = \mathrm{diag}\{\gamma_{s,1,1}^T, \ldots, \gamma_{s,n,\mu_n}^T\}, \tag{37}$$

where M_i^T denotes the i-th row of $M = A, A_s^*$.

Theorem 2 *(Exact cubification in vector form): let us define*

$$\mathbf{B}_s = \mathbf{A}_s^* (I_n \otimes c_s)(I_n \otimes \epsilon), \tag{38}$$
$$\mathbf{H} = \frac{1}{2} \sum_{s=1}^{d} \mathbf{H}_s^* (I_r \otimes c_s^{[2]})(I_r \otimes \epsilon^{[2]}), \tag{39}$$
$$\mathbf{G}_s = \mathbf{G}_s^* (I_r \otimes c_s)(I_r \otimes \epsilon), \tag{40}$$

where ϵ is the matrix defined through (13), (14). Then the final stage and driver are given, in vector form, by:

$$dx = \mathbf{A}(x \otimes Z)dt + \sum_{s=1}^{d} \mathbf{B}_s(x \otimes Z)dW_s, \qquad (41)$$

$$dZ = \left(\mathbf{F}Z^{[2]} + \mathbf{H}Z^{[3]}\right)dt + \sum_{s}^{d} \mathbf{G}_s Z^{[2]} dW_s. \qquad (42)$$

by defining the aggregate process $\mathbf{Z} \in \mathbb{R}^{\mathrm{q}}$, $q = n + r$:

$$\mathbf{Z}^T = [x^T, Z^T], \qquad (43)$$

we can write the exact cubification (41), (42), as a single vector SDE:

$$d\mathbf{Z} = (\Phi\mathbf{Z}^{[2]} + \Psi\mathbf{Z}^{[3]})dt + \sum_{s=1}^{d} \Gamma_s \mathbf{Z}^{[2]} dW_s. \qquad (44)$$

where Φ, Ψ, Γ_s are easily calculated matrices.

Proof. (*Omitted*)

E. Equation of the moments

Theorem 3 *(Equation of the moments). The set of all moments $\mathbf{E}\{\mathbf{Z}^{[m]}(t)\}$, for $m \in \mathbb{N}\backslash\{0\}$ of the random process $\mathbf{Z}(t)$ defined in (43), is given by the following infinite set of ODEs:*

$$\frac{d}{dt}\mathbf{E}\{\mathbf{Z}^{[m]}\} = \Phi^m \mathbf{E}\{\mathbf{Z}^{[m+1]}\} + \Psi^m \mathbf{E}\{\mathbf{Z}^{[m+2]}\}, \qquad (45)$$

where

$$\Phi^m = U_q^m(\Phi \otimes I_{q^{m-1}}), \qquad (46)$$

$$\Psi^m = U_q^m(\Psi \otimes I_{q^{m-1}}) + \sum_{s,s'} O_q^m(\Gamma_s \otimes \Gamma_{s'} \otimes I_{q^{m-2}}) \qquad (47)$$

$$\Gamma_s^m = U_q^m(\Gamma_s \otimes I_{q^{m-1}}). \qquad (48)$$

where $U_\cdot^\cdot, O_\cdot^\cdot$ are suitably defined matrices whose formula can be found in Lemma 5.1 of [3].

Proof. (*Omitted*)

III. Application to a basic network motif in Systems Biology

Consider the case of gene Z encoding for a protein which is a transcription factor for itself, which means Z autoregulates its own level of transcription. This is a typical network motif in gene transcription networks, i.e. a subnetwork pattern emerging in a high level of occurrences with respect to the occurrences of other patterns in random networks (see [4] for more details). We will consider the case of negative autoregulation and, according to a common simplifying assumption [4], [13], [11], will confuse the transcript (i.e. mRNA) with the corresponding protein. The copy number of transcript will be denoted by z and the following reactions of degradation (of 1 molecule) and production (of b molecules) will be considered:

$r_0)$ $z \mapsto z - 1$ $\quad a_0(z) = k_0 z,$

$r_{1b})$ $z \mapsto z + b$ $\quad a_{1b}(z) = k_1 f(z) p_b,$

where:

- $a_0 = k_0 z$ is the linear propensity for degradation (with $k_0 \in \mathbb{R}^+$), which is a common assumption in the literature (see, e.g., [6], [7], [8]);
- $f(z)$ is an inhibitory Hill function typically exploited to model negative regulation in gene transcription networks [4], [13], [11]:

$$f(z) = \frac{\theta^\nu}{z^\nu + \theta^\nu}, \qquad (49)$$

  with θ (*repression threshold*) denoting the copy number of z according to which the propensity is reduced by one half, and with the Hill coefficient ν (*promoter sensitivity*) denoting the sensitivity of the promoter to changes in z around θ;
- $\{p_b\}_{b=1}^{B_{\max}}$ is the probability mass function ($\sum_{b=1}^{B_{\max}} p_b = 1$) of the discrete random variable B denoting the burst size associated to the production reaction. We assume $B_{\max} \leq +\infty$, so possibly including the case of arbitrarily large bursts. We denote by $\langle B \rangle = \sum_{b=1}^{B_{\max}} b \cdot p_b$ the mean value of the burst size.

By defining the stoichiometric matrix $N = [-1\ 1\ 2\ \cdots]$ and aggregating the propensities in the reaction rates vector $V(z) = [a_0(z)\ a_{11}(z)\ a_{12}(z)\ \cdots]^T$, a popular stochastic formulation of a reaction network is given by the Chemical Langevin Equation (CLE, see e.g. [9], [12]), which is a stochastic differential equation describing the evolution in time of the copy number z_t, as follows:

$$dz_t = NV(z_t)dt + N\sqrt{(diag(V(z_t)))}dW_t, \qquad (50)$$

with $\{W_t\}$ being a vector of mutually independent standard Wiener processes $\{W_{bt}\}$, for $b = 0, 1, \ldots, B_{\max}$. CLE dramatically reduces the complexity of the underlying Chemical Master Equation, at the expense of the approximation that the amount of the species under investigation is continuous rather than discrete.

Neglecting the diffusion part of the CLE and exploiting the approximation $\langle V(z_t) \rangle = V(\langle z_t \rangle)$, which holds only for linear propensities (where $\langle \cdot \rangle$ denotes the expectation operator $\mathbf{E}\{\cdot\}$), the well known deterministic Reaction Rate Equations (RRE) are obtained:

$$\begin{aligned} \dot{\langle z_t \rangle} &= N \cdot V(\langle z_t \rangle) = -k_0 \langle z_t \rangle + k_1 f(\langle z_t \rangle) \sum_{b=1}^{B_{\max}} b \cdot p_b \\ &= -k_0 \langle z_t \rangle + k_1 \langle B \rangle f(\langle z_t \rangle). \end{aligned} \qquad (51)$$

Let us focus our attention on the moment equations for z_t, according to the SDE mathematical model coming from the CLE. In particular, we are interested in the second order moments of the involved chemical players: their steady-state computation allows 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 feedbacks [6], [7], [14]. In general nonlinear cases, to properly write the dynamics of moments equations it is required to constraint nonlinear terms (propensities in our

case) to polynomial functions [15]. However, even in case of polynomial systems, moment equations can be written but they do not provide a closed form expression, because the right-hand side of lower order moment equations depends on higher-order moments. There can be found in the literature different approaches aiming at providing an approximate solution, at least for the second-order moments at steady-state. One of this is Van Kampen's Linear Noise Approximation [9]. Recent approaches are based on moment closure techniques [10] enabling the computation of higher-order moments by means of nonlinear functions of the lower-order ones. In general, moment closure techniques can be applied only if the propensities are polynomial functions of the state, accordingly to what previously stated.

In the following, we will show a practical application of the SDE cubification techniques illustrated in the previous section, where, at first, the CLE in (50) is turned into a $\sigma\pi$-SDE by means of a proper coordinate change, and then the moment equation (45) is applied at the equilibrium. We focus on the first part of the procedure, which requires analytical computation, while numerical results on the actual computation of (45) (depending on the number of moments considered) and comparisons with existing approximate methods will be object of future work. Note that, according to [10], the accuracy of the moment closure scheme applied to the $\sigma\pi$-SDE can be arbitrarily increased by enlarging the number of moments considered in (45).

By expanding the CLE in (50), one obtains:

$$\begin{aligned} dz_t = &\left(-k_0 z_t + k_1 \langle B \rangle f(z_t)\right) dt - \sqrt{k_0 z_t} dW_{0t} \\ &+\sqrt{k_1 f(z_t)} \sum_{b=1}^{B_{\max}} b\sqrt{p_b} dW_{bt}. \end{aligned} \tag{52}$$

Note that the equation (52) is not in the form (7), due to the presence of the non-polynomial nonlinearity $f(z)$. In order to turn the system (52) into a $\sigma\pi$-stochastic system, we define an additional state variable ω_t as follows:

$$\omega_t := g(z_t) = z_t^\nu + \theta^\nu \tag{53}$$

such that $f(z_t) = \theta^\nu \omega_t^{-1}$. Hence the evolution of z_t in (52) can be rewritten as:

$$\begin{aligned} dz_t = &\left(-k_0 z_t + k_1 \langle B \rangle \theta^\nu \omega_t^{-1}\right) dt \\ &-\sqrt{k_0 z_t} dW_{0t} + \sqrt{k_1 \theta^\nu \omega_t^{-1}} \sum_{b=1}^{B_{\max}} b\sqrt{p_b} dW_{bt} \\ = &\left(-k_0 z_t + k_1 \langle B \rangle \theta^\nu \omega_t^{-1}\right) dt \\ &-\sqrt{k_0} z_t^{\frac{1}{2}} dW_{0t} + \sum_{b=1}^{B_{\max}} b\sqrt{k_1 p_b \theta^\nu} \omega_t^{-\frac{1}{2}} dW_{bt} \end{aligned} \tag{54}$$

while the evolution of ω_t is given by:

$$\begin{aligned} d\omega_t = &\tfrac{dg(z_t)}{dz_t} dz_t = \nu z_t^{\nu-1} \left(\left(-k_0 z_t + k_1 \langle B \rangle f(z_t)\right) dt\right) \\ &+\nu z_t^{\nu-1} \left(-\sqrt{k_0 z_t} dW_{0t} + \sqrt{k_1 f(z_t)} \sum_{b=1}^{B_{\max}} b\sqrt{p_b} dW_{bt}\right) \\ = &\left(-k_0 \nu z_t^\nu + k_1 \nu \langle B \rangle \theta^\nu z_t^{\nu-1} \omega_t^{-1}\right) dt \\ &+\left(-\nu\sqrt{k_0} z_t^{\nu-\frac{1}{2}} dW_{0t} + \sum_{b=1}^{B_{\max}} \nu b \sqrt{k_1 p_b \theta^\nu} z_t^{\nu-1} \omega_t^{-\frac{1}{2}} dW_{bt}\right) \end{aligned} \tag{55}$$

By defining the augmented state $x(t) = [z_t \quad \omega_t]^T$ and the noise vector $W(t) = [W_{0t} \quad W_{1t} \quad \cdots \quad W_{B_{\max} t}]^T$, and omitting the time-dependencies hereinfter, it is readily seen that the system is now $\sigma\pi$, namely can be rewritten in the form (5)–(8), where:

- $n = 2$ (state dimension)
- $d = B_{\max} + 1$ (dimension of the Wiener process)
- $\nu_i = 2$ for $i = 1, 2$ (number of drift monomials for each state equation i)
- $\nu^*_{i,s} = 1$, for $i = 1, 2$, for $s = 1, \ldots, d$ (number of diffusion monomials for each state equation i and for each noise component s)
- $\alpha_{1,1} = -k_0 \qquad \alpha_{1,2} = k_1 \langle B \rangle \theta^\nu$
  $\alpha_{2,1} = -k_0 \nu \qquad \alpha_{2,2} = k_1 \nu \langle B \rangle \theta^\nu$
  (drift coefficients)
- $X_{1,1} = x_1, \qquad X_{1,2} = x_2^{-1}$
  $X_{2,1} = x_1^\nu, \qquad X_{2,2} = x_1^{\nu-1} x_2^{-1}$
  (drift monomials)
- $p^1_{1,1} = 1 \quad p^1_{1,2} = 0 \quad p^2_{1,1} = 0 \quad p^2_{1,2} = -1$
  $p^1_{2,1} = \nu \quad p^1_{2,2} = 0 \quad p^2_{2,1} = \nu - 1 \quad p^2_{2,2} = -1$
  (drift exponents)
- $\alpha^*_{1,1,1} = -\sqrt{k_0} \qquad \alpha^*_{1,s,1} = (s-1)\sqrt{k_1 p_{s-1} \theta^\nu}$ for $s = 2, ..., d$
  $\alpha^*_{2,1,1} = -\nu\sqrt{k_0} \qquad \alpha^*_{2,s,1} = \nu(s-1)\sqrt{k_1 p_{s-1} \theta^\nu}$ for $s = 2, ..., d$
  (diffusion coefficients)
- $X^*_{1,1,1} = x_1^{\frac{1}{2}} \qquad X^*_{1,s,1} = x_2^{-\frac{1}{2}}$ for $s = 2, ..., d$
  $X^*_{2,1,1} = x_1^{\nu-\frac{1}{2}} \qquad X^*_{2,s,1} = x_1^{\nu-1} x_2^{-\frac{1}{2}}$ for $s = 2, ..., d$
  (diffusion monomials)
- $p^{*,1,1}_{1,1} = \frac{1}{2} \qquad p^{*,1,1}_{1,2} = 0 \quad p^{*,s,1}_{1,1} = 0 \qquad p^{*,s,1}_{1,2} = -\frac{1}{2}$ for $s = 2, ..., d$
  $p^{*,1,1}_{2,1} = \nu - \frac{1}{2} \quad p^{*,1,1}_{2,2} = 0 \quad p^{*,s,1}_{2,1} = \nu - 1 \quad p^{*,s,1}_{2,2} = -\frac{1}{2}$ for $s = 2, ..., d$
  (diffusion exponents)

Note that, since $d = B_{\max} + 1$ and $\nu^*_{i,s} = 1$, for $i = 1, 2$, for $s = 1, \ldots, d$, one gets:

- $\nu^*_i = \sum_{s=1}^d \nu^*_{i,s} \quad = d = B_{\max} + 1 \qquad$ for $i = 1, 2$,
- $\mu_i = \nu_i + \nu^*_i = d + 2 = B_{\max} + 3 \qquad$ for $i = 1, 2$.

From (8) one can equivalently define, for all i, the compound double index $k = (s, l)$, with $s = 1, ..., d$, $l = \nu^*_{i,s} = 1$, from which we can extend the definition of $\alpha_{i,l}$ and $X_{i,l}$, in agreement with (11). Such quantities are included in the vectors α_i and X_i:

- $\alpha_i = [\alpha_{i,1} \quad \alpha_{i,2} \quad \alpha^*_{i,1,1} \cdots \alpha^*_{i,d,1}]^T \in \mathbb{R}^{\mu_i}$, for $i = 1, 2$,
- $X_i = [X_{i,1} \; X_{i,2} \; X^*_{i,1,1} \cdots X^*_{i,d,1}]^T \in \mathbb{R}^{\mu_i}$, for $i = 1, 2$,

which are useful, in turn, to define the vector Z in (19)–(21) as follows:

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

with $r = \mu_1 + \mu_2 = 2(d+2) = 2(B_{\max}+3)$. Finally, the aggregate process $\mathbf{Z}$ in (43) is defined as:

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

with $q = n + r = 2 + 2(d+2) = 2B_{\max} + 8$.

The previous definitions allow to compute all the matrices required to explicitly write the moment equation (45). Since we are interested in the second-order moments, we consider the following system of $\left(q+q^2\right)$ differential equations:

$$\begin{cases} \frac{d\left\langle \mathbf{Z}^{[1]} \right\rangle}{dt} = \Phi^1 \left\langle \mathbf{Z}^{[2]} \right\rangle + \Psi^1 \left\langle \mathbf{Z}^{[3]} \right\rangle, \\ \frac{d\left\langle \mathbf{Z}^{[2]} \right\rangle}{dt} = \Phi^2 \left\langle \mathbf{Z}^{[3]} \right\rangle + \Psi^2 \left\langle \mathbf{Z}^{[4]} \right\rangle, \end{cases} \tag{56}$$

where we used the shortcut $\left\langle \mathbf{Z}^{[i]} \right\rangle = \mathbf{E}\{\mathbf{Z}^{[i]}\}$. Since the aforementioned system depends, in general, on the $\left(q^3+q^4\right)$ third-order and fourth-order moments, it is possible to apply the Separable Derivative-Matching (SDM) moment closure technique in [10] to formally define two nonlinear approximating functions $\bar{\varphi}_3 : \mathbb{R}^q \times \mathbb{R}^{q^2} \to \mathbb{R}^{q^3}$ and $\bar{\varphi}_4 : \mathbb{R}^q \times \mathbb{R}^{q^2} \to \mathbb{R}^{q^4}$ (depending just on the lower-order moments but matching the time derivative of the higher-order moments at some initial time), such that we can set:

$$\begin{array}{l} \left\langle \mathbf{Z}^{[3]} \right\rangle \simeq \bar{\varphi}_3 \left(\left\langle \mathbf{Z}^{[1]} \right\rangle, \left\langle \mathbf{Z}^{[2]} \right\rangle \right), \\ \left\langle \mathbf{Z}^{[4]} \right\rangle \simeq \bar{\varphi}_4 \left(\left\langle \mathbf{Z}^{[1]} \right\rangle, \left\langle \mathbf{Z}^{[2]} \right\rangle \right), \end{array} \tag{57}$$

implying that the system in (56) can be rewritten and solved with respect to the $\left(q+q^2\right)$ first-order and second-order moments:

$$\begin{cases} \frac{d\left\langle \mathbf{Z}^{[1]} \right\rangle}{dt} = \Phi^1 \left\langle \mathbf{Z}^{[2]} \right\rangle + \Psi^1 \bar{\varphi}_3 \left(\left\langle \mathbf{Z}^{[1]} \right\rangle, \left\langle \mathbf{Z}^{[2]} \right\rangle \right), \\ \frac{d\left\langle \mathbf{Z}^{[2]} \right\rangle}{dt} = \Phi^2 \bar{\varphi}_3 \left(\left\langle \mathbf{Z}^{[1]} \right\rangle, \left\langle \mathbf{Z}^{[2]} \right\rangle \right) + \Psi^2 \bar{\varphi}_4 \left(\left\langle \mathbf{Z}^{[1]} \right\rangle, \left\langle \mathbf{Z}^{[2]} \right\rangle \right), \end{cases} \tag{58}$$

from which the dynamics of the second-order moment of the CLE in (50) can be obtained by extracting the first component of $\left\langle \mathbf{Z}^{[2]} \right\rangle$, which is equal to $\left\langle z_t^2 \right\rangle$.

At the steady state, eq. (58) turns into a system of $\left(q+q^2\right)$ algebraic equations

$$\begin{cases} \Phi^1 \left\langle \mathbf{Z}_{ss}^{[2]} \right\rangle + \Psi^1 \bar{\varphi}_3 \left(\left\langle \mathbf{Z}_{ss}^{1]} \right\rangle, \left\langle \mathbf{Z}_{ss}^{[2]} \right\rangle \right) = \mathbf{0}_q, \\ \Phi^2 \bar{\varphi}_3 \left(\left\langle \mathbf{Z}_{ss}^{1]} \right\rangle, \left\langle \mathbf{Z}_{ss}^{[2]} \right\rangle \right) + \Psi^2 \bar{\varphi}_4 \left(\left\langle \mathbf{Z}_{ss}^{1]} \right\rangle, \left\langle \mathbf{Z}_{ss}^{[2]} \right\rangle \right) = \mathbf{0}_{q^2}, \end{cases} \tag{59}$$

where the subscripts ss indicate the equilibrium moments and $\mathbf{0}_q$ and $\mathbf{0}_{q^2}$ denote the q-dimensional and q^2-dimensional zero vectors, respectively. Similarly to the transient equation (58), the approximate steady-state second-order moment of the CLE in (50) is obtained by extracting the first component of $\left\langle \mathbf{Z}_{ss}^{[2]} \right\rangle$, which is equal to $\left\langle z_{ss}^2 \right\rangle$.

IV. Conclusion

Theorems 1 and 3 include the two main results of the paper: Theorem 1 states that every $\sigma\pi$-SDE can be cubified into a larger state-space, which amounts to say that there exists a *cubic*-SDE, given component-wise by eqs. (26)-(27), which generates *the same* process x, solution of the $\sigma\pi$-SDE (7) through the bilinear eq. (26) only. Theorem 3 uses an aggregate of eqs (26)-(27), built up in Theorem 2 (cubification in vector form), and a vector version of the Ito formula (**??**) (which is taken from the paper [3]), and builds up a (infinite) sequence of SDE's for the Kronecker powers of the state (x, Z) of eq. (44). Then, taking the expectations, the general (infinite-dimensional) moments equation (45) is obtained. An application in the systems biology framework is proposed, aiming at calculating the second-order moments for the Langevin equation associated to a very common reaction network motif, for which moment equations cannot be written according to the reference literature, because of the non-polynomial nonlinearities involved.

V. Acknowledgements

The financial support of the MIUR grant SysBioNet - Italian Roadmap for ESFRI Research Infrastructures - to SYSBIO Centre of Systems Biology is gratefully acknowledged

References

[1] R.S. Liptser, A.N. Shiryayev *Statistics of Random Processes – General Theory*, Springer-Verlag, 1977.

[2] F. Carravetta: Global Exact Quadratization of Continuous-Time Nonlinear Control Systems. *SIAM J. Control Optim.* 53, 235-261, 2015.

[3] F. Carravetta, A. Germani, M.K. Shuakayev : A new Suboptimal Approach to the Filtering Problem for Bilinear Stochastic Differential Systems, *SIAM J. Control Optim.*, 38, 1171-1203, 2000.

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

[5] I. Golding, J. Paulsson, S. Zawilski, E. Cox, Real-time kinetics of gene activity in individual bacteria, *Cell* 123, 1025-1036, 2005.

[6] Oyarzun, D.A, Lugagne, J.-B., Stan, G.-B.V.: Noise propagation in synthetic gene circuits for metabolic control, ACS Synthetic Biology, 2014.

[7] A. Singh, J.P. Hespanha, Optimal feedback strength for noise suppression in autoregulatory gene networks, *Biophysical Journal* 96, 4013-4023, 2009.

[8] M. Soltani, C. Vargas, A. Singh, Conditional moment closure schemes for studying stochastic dynamics of genetic circuits. *IEEE Transactions on Biomedical Circuits and Systems*, 9(4), 518–526, 2015.

[9] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North Holland, third edition, 2007.

[10] A. Singh, J.P. Hespanha, Approximate moment dynamics for chemically reacting systems, *IEEE Transactions on Automatic Control.* 56, 414-418, 2011.

[11] P. Mendes, W. Sha, K. Ye, Artificial gene networks for objective comparison of analysis algorithms, *Bioinformatics* 19, 122–129, 2003.

[12] R. Khanin, D. J. Higham, Chemical Master Equation and Langevin regimes for a gene transcription model *Theoretical Computer Science* 408, 31-40, 2008.

[13] N. Soranzo, G. Bianconi, C. Altafini, Comparing association network algorithms for reverse engineering of large scale gene regulatory networks: synthetic vs real data, *Bioinformatics* 23, 1640–1647, 2007.

[14] A. Borri, P. Palumbo and A. Singh, Metabolic noise reduction for enzymatic reactions: The role of a negative feedback, *Proceedings of the 54th IEEE Conference on Decision and Control (CDC)*, Osaka, pp. 2537-2542, 2015.

[15] J.P. Hespanha, A. Singh, Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems, *Int. J. of Robust and Nonlinear Control* 15, 669-689, 2005.