

Block-tridiagonal state-space realization of Chemical Master Equations: a tool to compute explicit solutions[☆]

A. Borri^a, F. Carravetta^a, G. Mavelli^a, P. Palumbo^a

^a*Istituto di Analisi dei Sistemi ed Informatica “A. Ruberti”,
Consiglio Nazionale delle Ricerche (IASI-CNR), Rome, Italy,
{[alessandro.borri](mailto:alessandro.borri@iasi.cnr.it),[francesco.carravetta](mailto:francesco.carravetta@iasi.cnr.it),
[gabriella.mavelli](mailto:gabriella.mavelli@iasi.cnr.it),[pasquale.palumbo](mailto:pasquale.palumbo@iasi.cnr.it)}@iasi.cnr.it.*

Abstract

Chemical Master Equations (CMEs) provide a comprehensive way to model the probabilistic behavior in biochemical networks. Despite their widespread diffusion in systems biology, the explicit computation of their solution is often avoided in favor of purely statistic Monte Carlo methods, due to the dramatically high dimension of the CME system.

In this work, we investigate some structural properties of CMEs and their solutions, focusing on the efficient computation of the stationary distribution. We introduce a generalized notion of one-step process, which results in a sparse dynamic matrix describing the collection of the scalar CMEs, showing a recursive block-tridiagonal structure as well. Further properties are inferred by means of a graph-theoretical interpretation of the reaction network. We exploit this structure by proposing different methods, including a dedicated LU decomposition, to compute the explicit solution.

Examples are included to illustrate the introduced concepts and to show the effectiveness of the proposed approach.

Keywords: Chemical Master Equation, Markov Processes, Systems Biology

[☆]This work has been partially supported by the Italian Ministry of Education and Research through the Flagship (PB05) InterOmics project.

1. Introduction

An important research topic in the field of Systems Biology is the detection of efficient methods for modeling complex cellular mechanisms. It has been recently highlighted the importance of the noise role in the dynamics of biological processes [1]. 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 modeling most of the molecular processes involved in cellular regulation, as well as in gene expression, see e.g. [2, 3]. These noisy behaviors are much more evident in cases when only few copies of the reacting species (DNA, RNA or proteins) are involved, and standard reaction rate equations, essentially dealing with the concentration dynamics in terms of Ordinary Differential Equations (ODE), fail to capture the inherent randomness of the phenomena. On the contrary, the Chemical Master Equation (CME) approach is known to be more appropriate in these cases, as it describes the biological process in terms of probability distribution of the underlying chemical population [4, 5, 6].

Such an attractive, stochastic approach, which allows to simulate and keep track of the reactions occurring in a single fixed volume, has recently become more and more appealing because of the biotechnology devices available nowadays, which are able to provide single-cell experimental data: see e.g. [7], where CME-based stochastic simulations have been used to validate a model of the Ras/cAMP/PKA signaling pathway, or the recent [8, 9, 10] where stochastic simulations have been used with the goal of *reverse engineering* from real data.

Except very basic cases, finding the exact solution of a CME (or just looking for the steady-state solution) is a *hard nut to crack*, even though the CME model results in a linear system. This is because the state vector collects the probabilities for all the possible combinations of copies of all the involved species, thus implying – even for a closed system of reacting species, i.e. when the CME is finite-dimensional – a dramatically high dimension of the state space. As a matter of fact, the computation of the matrix exponential or of the null space (in the stationary case) would require non-trivial numerical algorithms to be implemented. For these reasons, most efforts have been so far focused on implementing Monte Carlo methods (such as the Stochastic Simulation Algorithm (SSA), [5, 11], or the τ -leaping algorithms, [12, 13]) with the goal of approximating the exact solution. Indeed,

the performance of such algorithms is a tradeoff between the high number of Monte Carlo simulations required for approaching the exact solution and the time spent for running a single long-term simulation. It has to be stressed that, in some crucial cases, such a tradeoff could be not satisfactory. This is the case when some biological events happen rarely, thus requiring a very high number of Monte Carlo simulations in order to get a sufficiently precise statistics. Examples are usually taken from biological toggle switches such as, for instance, the ones related to the pyelonephritis-associated pili (Pap) epigenetic switch in *E. coli* [14] or the genetic toggle switch model of Gardner [15]. Therefore, a need exists to overcome purely statistical Monte Carlo methods and look for the solution of the original CME.

Important results on this field have been published in the last few years, aiming to highlight the properties characterizing the CMEs, and to provide implementable approximation schemes for their solutions. In this framework we may cite, among the others, the *Finite State Projection* method [16, 17, 18], based on a proper truncation of the set of the (possibly) infinite states of the CME system, and its improvements, such as [40], based on approximate Krylov-based methods to speed up the computation of the solution; the *Moment Closure* technique [19, 20, 21], aiming at estimating the statistical moments of the CMEs; *spectral decomposition* approaches providing the approximate solution in terms of a suitable class of basis functions [22, 23]; the *sparse grid approach* [41], specifically designed to deal with high dimensional problems arising in gene regulatory networks; the novel approach based on *reaction counts* [42, 43], in which the collection of the master equations is rewritten in terms of reactions instead of species, resulting in a simpler structure of the system; finally, model reduction techniques by *balanced truncation* [18, 46], exploited to efficiently approximate the probability distribution.

The present work investigates the structural properties of the CMEs and their usefulness in computing the solution, especially in the presence of a large number of states. The starting point of our investigation relies on the following observation: different state-space realizations (with different structural properties) can be achieved according to different ways of gathering probabilities, when building up the CME state vector. As a matter of fact, we will show that a very regular structure can be achieved for the aggregated system, and many nice features of the system can be inferred from it.

The mathematical setting that we here propose can be always adopted regardless of the number or kind of reactions/reagents. The method to build

up the CMEs exploits a recursive approach to aggregate the vector of probabilities and the dynamic matrix. The analysis of the dynamical properties is the novelty of the paper, as well as the characterization of the chemical master equations in a *graph-theoretical* fashion. Our results are suitably exploited to find efficiently the exact solution of the master equation, in a much faster way with respect to usual Monte Carlo SSA techniques; the method is validated in a couple of well-established biochemical examples, where a comparison with other solvers highlights the effectiveness of the approach.

The paper is organized as follows. In Section 2, we introduce the general setting of the (bio)chemical reaction framework. Section 3 briefly recaps the way to obtain the CMEs and introduces a novel structure to write them in a compact, useful form that reveals to show a recursive block-tridiagonal structure according to non-restrictive assumptions. Section 4 focuses on the characterization of the CME solutions, illustrating an intriguing graph-theoretical interpretation of the stationary distribution. Different computational approaches are proposed in Section 5 to achieve the steady-state solutions, most suitably exploiting the block-tridiagonal property coming out from the proposed state-space realization. In Section 6, we present simulation results in some biochemical applications. Section 7 offers concluding remarks. A preliminary version of this paper appeared in [24].

Before starting with the main topic, we introduce some notation extensively used throughout the paper. The symbols \mathbb{N} , \mathbb{N}_0 , \mathbb{R} and \mathbb{R}^+ denote the set of natural, nonnegative integer, real and positive real numbers, respectively. The cardinality of a set V is denoted by $|V|$. The transpose of a matrix A is written as A^T . The time-derivative $\frac{dp}{dt}$ is denoted by \dot{p} . Given a matrix A , the symbol $[A]_{ij}$ denotes the generic entry of matrix A , for any i, j .

A matrix is said to be block-tridiagonal if it has the following form (see e.g. [25], [26]):

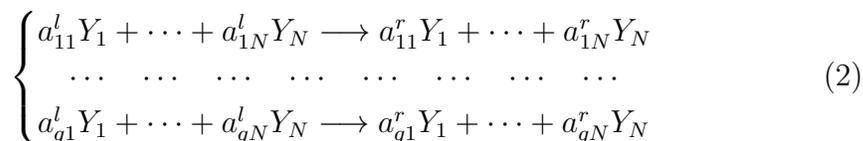
$$\Phi_\nu(\{A_i\}, \{B_i\}, \{C_i\}; i) = \begin{bmatrix} B_0 & C_1 & \emptyset & \cdots & \emptyset \\ A_0 & B_1 & C_2 & \emptyset & \emptyset \\ \emptyset & \ddots & \ddots & \ddots & \emptyset \\ \vdots & \ddots & A_{\nu-2} & B_{\nu-1} & C_\nu \\ \emptyset & \cdots & \emptyset & A_{\nu-1} & B_\nu \end{bmatrix} \quad (1)$$

where $\{A_i\} = \{A_i, i = 0, \dots, \nu - 1\}$, $\{B_i\} = \{B_i, i = 0, \dots, \nu\}$, $\{C_i\} = \{C_i, i = 1, \dots, \nu\}$ are sequences of suitably dimensioned square matrices,

and the \emptyset entries are zero matrices of proper dimensions. The operator $\Phi_\nu(\{A_i\}, \{B_i\}, \{C_i\}; i)$, where i is the *dummy index* of the sequences taken as inputs, is called *block-tridiagonal matrix builder*. If $\{A_i\}$, $\{B_i\}$, $\{C_i\}$ are sequences of scalar values and the \emptyset entries are scalar (zeros), the matrix in (1) is said to be *tridiagonal*.

2. State-space representation of Chemical Networks

Consider the following system (or network) of q (bio)chemical reactions in the general form [27]:



where Y_1, \dots, Y_N are the species involved. The number $\beta_{ij} = a_{ij}^r - a_{ij}^l$ is called *stoichiometric coefficient* of species j in reaction i , for all $j = 1, \dots, N$ and for all $i = 1, \dots, q$. In case the right-hand-side and the left-hand-side of a given reaction are equal, respectively, to the left-hand-side and the right-hand-side of a different reaction, we simply rephrase the two reactions as a unique formal *reversible* reaction by means of the symbol \rightleftharpoons .

Let us denote by $n(t)$ the state of the system at time t , with i -th component $n_i(t) \in \mathbb{N}_0$ being the number of copies of the i -th species at time t . The state function $n : [0, +\infty) \rightarrow \mathbb{N}_0^N$ is a realization of a discrete-valued continuous-time stochastic Markov process with initial conditions $n_i(0) = \bar{n}_i$, $i = 1, \dots, N$. We refer to a system (or to the underlying process) as *closed* if $n_i(t) \in \{0, 1, \dots, N_i\}$ for some fixed $N_i \in \mathbb{N}$, for all i . Due to mass balance constraints, vector $n(t)$ is usually redundant in closed systems, whose state can be univocally identified by means of a reduced set of the $n(t)$ components, which are called *independent species*. Thus, if N_C is the number of mass balance constraints, the reduced state vector $x(t)$ is $(N - N_C)$ -dimensional and can be built according to many different choices of the species involved. The following definition focuses on an important property of the above mentioned choices.

Definition 1 (Orthogonal species). *Consider a system of q chemical reactions as in (2). A set $\{Y_{r_1}, \dots, Y_{r_M}\}$ of independent species is a set of orthogonal species for the given system of reactions if for any reaction $i \in \{1, \dots, q\}$ there exists a unique index $j \in \{r_1, \dots, r_M\}$ such that $\beta_{ij} \neq 0$.*

In practice, given M independent species on a set of reactions, we say that these species are orthogonal if every chemical reaction changes just one component of the reduced state $x(t)$. Note that, although it is always possible to define a set of independent species out of a system of q reactions, it is not always possible to choose M independent *and* orthogonal species. Nevertheless, the orthogonality property can be recovered by defining state variables as appropriate linear combinations of the original species. This approach is conceptually similar to that based on the so-called “reaction counts” (see e.g. [42, 43]), where CMEs in terms of species are restated in terms of reactions by means of an appropriate redefinition of the state variables.

In the remainder of the work, a stochastic Markov process describing a network of reactions involving M independent orthogonal species will be denoted as an *orthogonal process*. An important class of orthogonal processes is the class of *one-step* processes [4], formally defined as the Markov processes associated to a set of M independent species such that the following assumptions are fulfilled:

- (H1) $\beta_{ij} \in \{-1, 0, 1\}$ for all i, j (unitary steps);
- (H2) the M species are orthogonal.

In the following, we will generalize the one-step setting by removing the orthogonality hypothesis (H2). Such a general case will be referred to as *generalized* (non-orthogonal) one-step process, or *unitary* process, because reactions are allowed to cause simultaneous changes of unitary amount in the state variables. A graphical example is given in Figure 1.

3. Structure of the Chemical Master Equation for generalized one-step processes

In the following, we denote by $x(t) \in \mathbb{N}_0^M$ the CME state vector, whose components are related to a choice of M independent species in the q reactions. As anticipated in the Introduction, different state-space realizations for the CME provide different structural properties of the underlying dynamical system, possibly suggesting a smart way to compute the solutions. To this end, we will briefly recap the route to achieve the CME, according to a new set of notations. We start from the definition of $p_{n_1, n_2, \dots, n_M}(t)$ as the joint probability of having n_i copies of the i -th species (for $i = 1, \dots, M$) at

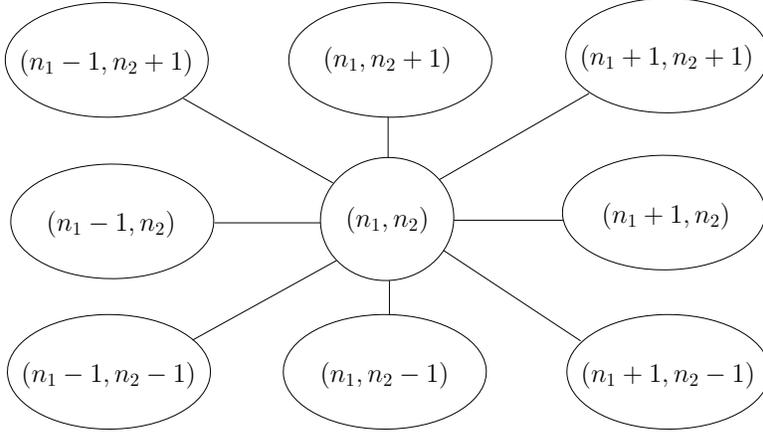


Figure 1: Example of bivariate generalized (non-orthogonal) one-step process. The pairs indicate the number of molecules of the two independent species; the links indicate non-zero transition probabilities per unit of time. Note that the diagonal links vanish in the classical one-step case.

time t :

$$p_{n_1, \dots, n_M}(t) \doteq P\left(x_1(t) = n_1, \dots, x_M(t) = n_M\right).$$

The transition from one state to another is ruled by the *propensities* or transition probabilities per unit of time, defined as:

$$g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \cdot P\left(x_i(t + \Delta t) = n_i + \alpha_i, i = 1, \dots, M \mid x_i(t) = n_i, i = 1, \dots, M\right),$$

in which P is the conditional probability for a step transition of the discrete amount α_i in each state variable $x_i(t)$, $i = 1, \dots, M$. According to standard hypotheses [4], it is assumed that such a probability does not depend explicitly on the time t and that only one reaction per time can occur.

It is worth noticing that not all the transitions are allowed for a given chemical network (2). As a matter of fact, the propensities are constrained to match the variations in the number of copies of all species in some reactions, namely:

$$g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} = 0 \quad \text{if } (\alpha_1, \dots, \alpha_M) \notin \{\beta_1, \dots, \beta_q\}. \quad (3)$$

with $\beta_i \doteq (\beta_{i1}, \dots, \beta_{iM})$ providing the aggregate vectors of the stoichiometric coefficients associated to reactions $i = 1, \dots, q$.

The following proposition formally defines generalized/orthogonal one-step processes by means of the network propensities previously introduced.

Proposition 2. *Consider a network of q chemical reactions involving M independent species, described by the stoichiometric coefficients β_{ij} , for $i = 1, \dots, q$ and $j = 1, \dots, M$. Then,*

(a) *Assumption (H1) (unitary or generalized one-step process) implies that for any state (n_1, \dots, n_M) of the process, one has:*

$$g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} = 0 \text{ if } \max_{j=1, \dots, M} |\alpha_j| > 1. \quad (4)$$

(b) *Assumptions (H1)+(H2) (orthogonal one-step process) imply that for any state (n_1, \dots, n_M) of the process, one has:*

$$g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} = 0 \text{ if } \sum_{j=1, \dots, M} |\alpha_j| > 1. \quad (5)$$

We give a simple example to clarify the previous result. Consider $M = 3$ and the transition probability per unit of time $g_{n_1, n_2, n_3}^{\alpha_1, \alpha_2, \alpha_3}$ with $(\alpha_1, \alpha_2, \alpha_3) = (1, -1, 0)$, namely the propensity of the transition to the state $(n_1 + 1, n_2 - 1, n_3)$. This transition is not allowed in orthogonal one-step processes, in fact $\sum_{j=1, \dots, 3} |\alpha_j| = 2 > 1$, implying $g_{n_1, n_2, n_3}^{\alpha_1, \alpha_2, \alpha_3} = 0$, according to Proposition 2(b). In the non-orthogonal case (Proposition 2(a)), instead, the transition is possible because $\max_{j=1, \dots, 3} |\alpha_j| = 1$, so the case $g_{n_1, n_2, n_3}^{\alpha_1, \alpha_2, \alpha_3} \neq 0$ is allowed.

The Chemical Master Equation is written for any general state (n_1, \dots, n_M) of the Markov process as the following probability balance equation [4]:

$$\begin{aligned} \dot{p}_{n_1, \dots, n_M}(t) &= \sum_{\alpha_1, \dots, \alpha_M} g_{n_1 + \alpha_1, \dots, n_M + \alpha_M}^{-\alpha_1, \dots, -\alpha_M} \cdot p_{n_1 + \alpha_1, \dots, n_M + \alpha_M}(t) \\ &\quad - \sum_{\alpha_1, \dots, \alpha_M} g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} \cdot p_{n_1, \dots, n_M}(t), \end{aligned} \quad (6)$$

where the first (second) summation takes into account the ingoing (outgoing) probability rates with respect to state (n_1, \dots, n_M) . In the case of closed systems, with each species i taking values in the set $\{0, 1, \dots, N_i\}$, the equations in (6) consist of a set of $\mathcal{M} = (N_1 + 1) \times \dots \times (N_M + 1)$ equations providing

the dynamics of the joint M -dimensional probability distribution. Note that the hypothesis of closed system is usually reasonable since, even in the case of open systems, one can truncate the system by constraining the population of the *unbounded species* below some reasonable levels, in order to limit the total number of states yet guaranteeing accurate results. This is justified because the CME approach is commonly used for reactions involving *few molecules* [4].

Hereafter we consider a proper way to collect the \mathcal{M} equations given by (6) in a compact form that satisfies a set of interesting properties. To this end, for any choice (n_1, \dots, n_{M-1}) of the $(N_1 + 1) \times \dots \times (N_{M-1} + 1)$ possible settings of the copies of the first $M - 1$ independent species, define the $(N_M + 1)$ -dimensional vector of probabilities:

$$\mathcal{P}_{n_1, \dots, n_{M-1}} \doteq \begin{bmatrix} p_{n_1, \dots, n_{M-1}, 0} \\ p_{n_1, \dots, n_{M-1}, 1} \\ \vdots \\ p_{n_1, \dots, n_{M-1}, N_M} \end{bmatrix} \in \mathbb{R}^{N_M+1}. \quad (7)$$

Then, the following vectors of probabilities can be recursively defined (for $1 \leq i \leq M - 2$)

$$\mathcal{P}_{n_1, \dots, n_i} \doteq \begin{bmatrix} \mathcal{P}_{n_1, \dots, n_i, 0} \\ \mathcal{P}_{n_1, \dots, n_i, 1} \\ \vdots \\ \mathcal{P}_{n_1, \dots, n_i, N_{i+1}} \end{bmatrix} \in \mathbb{R}^{(N_{i+1}+1) \times \dots \times (N_M+1)}, \quad (8)$$

up to the definition of vector \mathcal{P} , entailing all the probabilities involved by the CME:

$$\mathcal{P} \doteq \begin{bmatrix} \mathcal{P}_0 \\ \mathcal{P}_1 \\ \vdots \\ \mathcal{P}_{N_1} \end{bmatrix} \in \mathbb{R}^{\mathcal{M}}. \quad (9)$$

Since the right-hand side of Eq. (6) is a linear combination of the joint probabilities of the states of the Markov process, the equations for the joint probabilities of all states can be collected in the form of an autonomous linear system:

$$\dot{\mathcal{P}} = G\mathcal{P}, \quad G \in \mathbb{R}^{\mathcal{M} \times \mathcal{M}}. \quad (10)$$

It follows from (6) that G is a Metzler matrix (namely all the off-diagonal components are nonnegative), hence the system in (10) can be regarded as a positive linear dynamical system [28].

The remainder of the Section is devoted to characterize the block partition of matrix G , as naturally induced by the recursive definition of \mathcal{P} in (7)–(9). To this end, define the following scalar values for any state (n_1, \dots, n_M) and any vector of variations $(\alpha_1, \dots, \alpha_M)$:

$$G_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} = \begin{cases} g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} & \text{if } (\alpha_1, \dots, \alpha_M) \neq (0, \dots, 0), \\ - \sum_{\alpha'_1, \dots, \alpha'_M} g_{n_1, \dots, n_M}^{\alpha'_1, \dots, \alpha'_M} & \text{if } (\alpha_1, \dots, \alpha_M) = (0, \dots, 0). \end{cases} \quad (11)$$

Then, the following blocks can be recursively defined, backwards from (11):

$$G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i} = \begin{bmatrix} G_{n_1, \dots, n_i, 0}^{\alpha_1, \dots, \alpha_i, 0} & G_{n_1, \dots, n_i, 1}^{\alpha_1, \dots, \alpha_i, -1} & \cdots & G_{n_1, \dots, n_i, N_{i+1}}^{\alpha_1, \dots, \alpha_i, -N_{i+1}} \\ G_{n_1, \dots, n_i, 0}^{\alpha_1, \dots, \alpha_i, 1} & G_{n_1, \dots, n_i, 1}^{\alpha_1, \dots, \alpha_i, 0} & \cdots & G_{n_1, \dots, n_i, N_{i+1}}^{\alpha_1, \dots, \alpha_i, -N_{i+1}+1} \\ \vdots & \vdots & \ddots & \vdots \\ G_{n_1, \dots, n_i, 0}^{\alpha_1, \dots, \alpha_i, N_{i+1}} & G_{n_1, \dots, n_i, 1}^{\alpha_1, \dots, \alpha_i, N_{i+1}-1} & \cdots & G_{n_1, \dots, n_i, N_{i+1}}^{\alpha_1, \dots, \alpha_i, 0} \end{bmatrix}, \quad (12)$$

for $1 \leq i \leq M-1$, $0 \leq n_i \leq N_i$, $0 \leq |\alpha_i| \leq N_i$. Note that the blocks $G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i}$ are square matrices with dimension $(N_{i+1} + 1) \times \cdots \times (N_M + 1)$. Finally, matrix G can be built exploiting the set of blocks defined in (12) for the last ($i = 1$) backwards iteration:

$$G = \begin{bmatrix} G_0^0 & G_1^{-1} & G_2^{-2} & \cdots & G_{N_1}^{-N_1} \\ G_0^1 & G_1^0 & G_2^{-1} & \cdots & G_{N_1}^{-(N_1-1)} \\ G_0^2 & G_1^1 & G_2^0 & \cdots & G_{N_1}^{-(N_1-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ G_0^{N_1} & G_1^{N_1-1} & G_2^{N_1-2} & \cdots & G_{N_1}^0 \end{bmatrix}. \quad (13)$$

From the block decomposition of matrix G defined by eqs. (11)–(13), it follows that any subvector $\mathcal{P}_{n_1, \dots, n_i}$ in the right-hand side of Eq. (10) is multiplied by blocks of the type $G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i}$ for some choice of multi-index $(\alpha_1, \dots, \alpha_i)$. This happens for any choice of $i = 1, \dots, M$; therefore, for $i = M$, the scalar elements p_{n_1, \dots, n_M} are multiplied by scalars of the type $G_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M}$, hence contributing to the dynamics of $p_{n_1 + \alpha_1, \dots, n_M + \alpha_M}$ in (6). As a matter of fact, the case $(\alpha_1, \dots, \alpha_M) = (0, \dots, 0)$ defines the main diagonal of G which, accordingly to (11), is consistent with the CME in (6). The notation is illustrated in Figure 2 and an example is given in Figure 3, where it is shown the pattern of zeros of the 9×9 recursive block-tridiagonal matrix G obtainable with $M = 2$, $N_1 = N_2 = 2$.

The following theorem characterizes the block partitioning of matrix G in the case of generalized/orthogonal one-step processes, and generalizes the scalar one-step conditions given in (4)–(5) to all the blocks in G .

Theorem 3. *Consider a network of q chemical reactions involving M independent species, described by the stoichiometric coefficients β_{ij} , for $i = 1, \dots, q$ and $j = 1, \dots, M$, and assume the recursive block partitioning of G described in (11)–(13). Then:*

- (a) *Assumption (H1) (unitary or generalized one-step process) implies that, for any $i = 1, \dots, M$ and for any state (n_1, \dots, n_i) , one gets:*

$$G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i} = \emptyset \quad \text{if } \max_{j=1, \dots, i} |\alpha_j| > 1, \quad (14)$$

where \emptyset is a suitably dimensioned zero matrix.

- (b) *Assumptions (H1)+(H2) (orthogonal one-step process) imply that, for any $i = 1, \dots, M$, and for any state (n_1, \dots, n_i) , one gets:*

$$G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i} = \emptyset \quad \text{if } \sum_{j=1, \dots, i} |\alpha_j| > 1, \quad (15)$$

where \emptyset is a suitably dimensioned zero matrix.

Proof 4. Case (a). *We proceed by induction as follows. Base case ($i = M$). The condition in (14) rewrites*

$$g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} = 0 \quad \text{if } \max_{j=1, \dots, M} |\alpha_j| > 1, \quad (16)$$

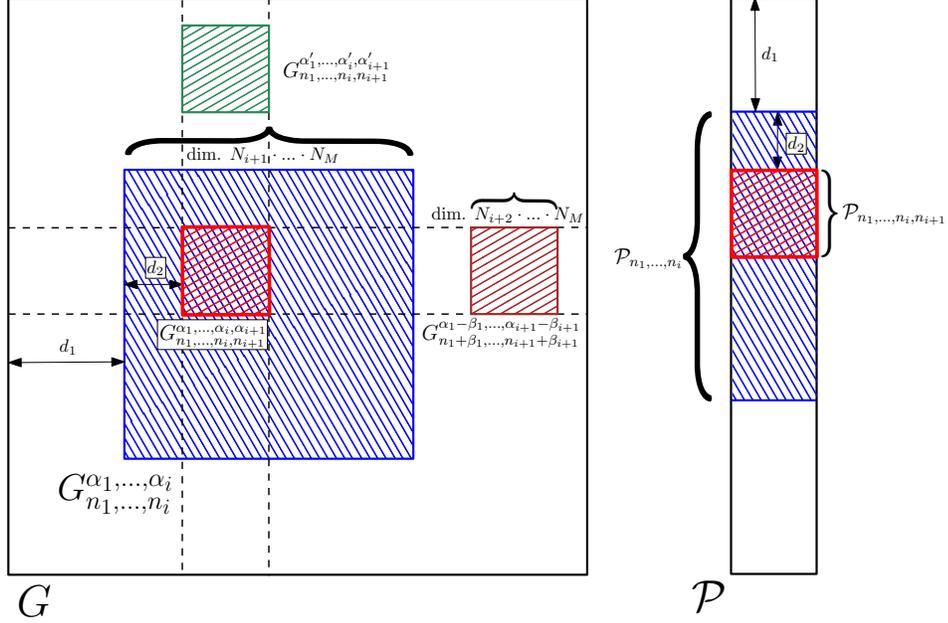


Figure 2: Recursive partitioning of the matrix G and of the probability vector \mathcal{P} . Note the same relative positions (expressed by the column shifts d_1 and d_2) of the submatrices $G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i}$ and $G_{n_1, \dots, n_i, n_{i+1}}^{\alpha_1, \dots, \alpha_i, \alpha_{i+1}}$ in G and of the subvectors $\mathcal{P}_{n_1, \dots, n_i}$ and $\mathcal{P}_{n_1, \dots, n_i, n_{i+1}}$ in \mathcal{P} . Moreover, at any level of the iteration, the blocks of G sharing the same set of columns (e.g. the red and the green blocks) are characterized by the same subscripts n_1, \dots, n_{i+1} ; instead the component-wise summations of subscripts and superscripts are constant for the blocks of G sharing the same set of rows (e.g. the red and the brown blocks).

where we plugged in the first case in Eq. (11). Since Assumption (H1) holds, the condition in (16) is directly implied by Proposition 2(a), concluding the proof of the base case.

Inductive step ($1 \leq i \leq M - 1$). We proceed backwards and we assume that the condition

$$G_{n_1, \dots, n_i, n_{i+1}}^{\alpha_1, \dots, \alpha_i, \alpha_{i+1}} = \emptyset \text{ if } \max_{j=1, \dots, i+1} |\alpha_j| > 1$$

		0			0	0	0	0
						0	0	0
0			0			0	0	0
		0			0			0
0			0			0		
0	0	0			0			0
0	0	0						
0	0	0	0			0		

Figure 3: Pattern of zeros of the 9×9 recursive block-tridiagonal matrix G obtainable with $M = 2$, $N_1 = N_2 = 2$.

holds, in order to show that the condition

$$G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i} = \emptyset \text{ if } \max_{j=1, \dots, i} |\alpha_j| > 1$$

holds too. Note that the condition $\max_{j=1, \dots, i} |\alpha_j| > 1$ implies that $\max_{j=1, \dots, i+1} |\alpha_j| \geq \max_{j=1, \dots, i} |\alpha_j| > 1$. Hence, by the inductive assumption, for any choice of n_{i+1} and α_{i+1} , $G_{n_1, \dots, n_i, n_{i+1}}^{\alpha_1, \dots, \alpha_i, \alpha_{i+1}} = \emptyset$, namely all the blocks in (12) are zero matrices. This implies that $G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i} = \emptyset$, which concludes the proof of case (a).

Case (b). The proof follows the same steps as case (a).

Remark 5 (Recursive block-tridiagonal structure of matrix G). Theorem 3 determines a recursive structure of matrix G , which we describe in the following. We distinguish two cases:

- **Generalized (non-orthogonal) one-step processes:** Theorem 3(a) implies that G is block-tridiagonal (see in Eq. (1)) and all the non-zero blocks of G are block-tridiagonal with the same structure of G . This can be seen by inspecting the block partitioning shown in (12–13) and noting that, at each level i of the iterative block partitioning, the blocks in the main diagonal are characterized by $\alpha_{i+1} = 0$, the first block diagonal below the main diagonal is characterized by $\alpha_{i+1} = 1$ and the first block diagonal above the main diagonal is characterized by $\alpha_{i+1} = -1$. Taking into account that a zero block at some iteration

implies that all its internal blocks are zero matrices, the recursive block-tridiagonal structure of G is directly implied by the condition in (14), which prevents the existence of non-zero blocks $G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i}$ with $|\alpha_j| > 1$ for some $j \leq i$.

- **Orthogonal one-step processes:** Theorem 3(b) determines further properties of G in the orthogonal case. In particular: G is still block-tridiagonal with all the non-zero off-diagonal blocks of G being diagonal matrices. This is implied by the fact that, for any off-diagonal block $G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i}$, one has $|\alpha_j| \geq 1$ for some $j \leq i$. As a consequence, all the elements $G_{n_1, \dots, n_i, n_{i+1}, \dots, n_M}^{\alpha_1, \dots, \alpha_i, \alpha_{i+1}, \dots, \alpha_M}$ with $(\alpha_{i+1}, \dots, \alpha_M) \neq (0, \dots, 0)$ (outside the main diagonal of the block) satisfy the condition $\sum_{j=1, \dots, M} |\alpha_j| > 1$ in (4) and are null.

It is worth noticing that the assumptions of generalized (non-orthogonal) one-step process are the mildest conditions preserving the recursive block-tridiagonal structure of G .

Finally note that, since matrix G is recursively block-tridiagonal, it can be expressed by means of the matrix builder in (1). For any recursion step $i = M - 1, \dots, 1$, the blocks in G can be expressed as:

$$G_{n_1, \dots, n_i}^{\alpha_1, \dots, \alpha_i} = \Phi_{N_{i+1}} \left(\{G_{n_1, \dots, n_i, n_{i+1}}^{\alpha_1, \dots, \alpha_i, 1}\}, \{G_{n_1, \dots, n_i, n_{i+1}}^{\alpha_1, \dots, \alpha_i, 0}\}, \{G_{n_1, \dots, n_i, n_{i+1}}^{\alpha_1, \dots, \alpha_i, -1}\}; n_{i+1} \right),$$

and the last step of the backward recursion provides the whole matrix:

$$G = \Phi_{N_1} \left(\{G_{n_1}^1\}, \{G_{n_1}^0\}, \{G_{n_1}^{-1}\}; n_1 \right).$$

Remark 6 (Multi-step processes). Assumption (H1) of non-orthogonal one-step process could be further extended to the more general case of non-orthogonal h -step process:

$$g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} = 0 \text{ if } \max_{j=1, \dots, M} |\alpha_j| > h.$$

This is the case, for instance, of networks of reactions generating bursts of (at most) h molecules [29]. As a consequence, Theorem 3 can be generalized with the aim of highlighting a recursive block- $(2h + 1)$ -diagonal structure of matrix G in generalized (non-orthogonal) h -step processes. A more formal characterization of multi-step processes is out of the scope of the present work.

4. Characterization of the CME solutions

In this Section, we show some interesting properties of matrix G that reveal to be useful to characterize the CME solutions, with a special focus on the steady-state distribution. From a theoretical point of view, the explicit solution of the CME in (10) at time t is:

$$\mathcal{P}(t) = e^{Gt}\mathcal{P}(0), \quad (17)$$

where the initial distribution $\mathcal{P}(0)$ is given at the initial time $t_0 = 0$.

Most of our attention will be focused on a very classical task when dealing with CMEs, namely the computation of the stationary distribution [4], providing fundamental information for understanding the equilibrium condition of a network of chemical reactions. A way to compute it, while avoiding the computation of the matrix exponential e^{Gt} , suitably exploits the stationary distribution properties; indeed, it is the probability vector that satisfies the following steady-state conditions:

$$\begin{cases} G\mathcal{P} &= \mathbf{0} \\ \mathbf{1}^T\mathcal{P} &= 1 \\ \mathcal{P} &\geq \mathbf{0} \end{cases} \quad (18)$$

where $\mathbf{0} = (0 \cdots 0)^T \in \mathbb{R}^M$, $\mathbf{1} = (1 \cdots 1)^T \in \mathbb{R}^M$ and the inequality is interpreted as component-wise. We point out that, instead of calculating the null space in (18), one can take advantage of some interesting properties of matrix G , coming from the *Algebraic Graph Theory* [30].

4.1. A graph-theoretical interpretation of biochemical networks

Let us briefly recall the concept of weighted directed graph (*digraph*) before getting to the details.

Definition 7. A weighted digraph is a triple (V, E, A) , where $V = \{v_k\}$ is a set of vertices (or nodes), $E \subseteq V \times V$ is a set of ordered pairs of vertices called edges (or links), and A is a weighted adjacency matrix such that, for any pair (i, j) , the entry $[A]_{ij}$ is strictly positive if (v_i, v_j) is an edge, whilst $[A]_{ij} = 0$ otherwise.

Note that the set of edges E can be derived from matrix A and can be therefore omitted in the previous definition. A very important matrix related to a weighted digraph is the Laplacian matrix.

Definition 8. [30] The Laplacian matrix L of the digraph (V, E, A) is defined as

$$L_{ij} = \begin{cases} \sum_k [A]_{ik} & i = j \\ -[A]_{ij} & \text{otherwise.} \end{cases}$$

We now introduce a formal graph-theoretical interpretation of a Markov process describing a network of chemical reactions.

Definition 9. The digraph associated to a continuous-time discrete-state stochastic Markov process is a weighted digraph (V, E, A) , where each vertex $v_k \in V$ is associated to a discrete state (n_1, \dots, n_M) of the process, and A is a matrix whose generic element $[A]_{ij}$ is the propensity $g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M}$ of reaching the state $v_j = (n_1 + \alpha_1, \dots, n_M + \alpha_M)$ from the state $v_i = (n_1, \dots, n_M)$. The set of edges E is uniquely defined by A and includes all the links (v_i, v_j) with non-zero probability per unit of time of reaching v_j from v_i .

The following Theorem shows that matrix G shares most of the properties of the graph Laplacian, according to a proper order of the nodes in V . This result, with its consequences, provides a novel characterization (to the best of the authors' knowledge) of the dynamic matrix G of a general CME in terms of well-known results of algebraic graph theory (see e.g. [30]).

Theorem 10. Let us assume that the set V of the nodes of the weighted digraph associated with the Markov process describing a chemical network is ordered according to the order of states induced by the recursive construction in Eqs. (7)–(9). Let L be the Laplacian of the graph. Then $G = -L^T$.

Proof 11. Consider any row i of the matrix G , corresponding to a state (n_1, \dots, n_M) of the Markov process. The master equation for such a state is given by (6), hence

$$[G]_{ii} = - \sum_{\alpha_1, \dots, \alpha_M} g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M} \quad (19)$$

and

$$[G]_{ij} = g_{n_1 + \alpha_1, \dots, n_M + \alpha_M}^{-\alpha_1, \dots, -\alpha_M},$$

for $i \neq j$, where the generic column j is referred to the node $(n_1 + \alpha_1, \dots, n_M + \alpha_M)$, for some $\alpha_1, \dots, \alpha_M$. Note that, from Definitions 8 and 9, $[G]_{ii} =$

$-[L]_{ii}$. Now consider the element $[L]_{ji} = -[A]_{ji}$ of the Laplacian which, from Definition 9, is (minus) the probability per unit of time of reaching state i , associated to (n_1, \dots, n_M) , from state j , associated to $(n_1 + \alpha_1, \dots, n_M + \alpha_M)$; hence by Definition 9 it is equal to $-g_{n_1 + \alpha_1, \dots, n_M + \alpha_M}^{-\alpha_1, \dots, -\alpha_M}$, in turn equal to $-[G]_{ij}$. This concludes the proof.

Known properties of G , which we rediscover here as a consequence of Theorem 10, are the following [30]:

- $\mathbf{1}^T G = \mathbf{0}^T$, following from (19) and from the fact that, for $j \neq i$, $[G]_{ji} = g_{n_1, \dots, n_M}^{\alpha_1, \dots, \alpha_M}$ for some $\alpha_1, \dots, \alpha_M$. Hence each column adds up to zero, implying that G is singular and admits a non-trivial null space;
- all eigenvalues of G have nonpositive real part, ensuring the convergence of the dynamics to the null space;
- $\mathbf{1}^T e^{Gt} = \mathbf{1}^T$ for all t , i.e e^{Gt} is a *column-stochastic* matrix. This ensures that $\mathcal{P}(t)$ is a probability vector (nonnegative entries which add up to 1) at any time t , provided that the initial condition $\mathcal{P}(0)$ is a probability vector.

4.2. Existence, uniqueness and properties of the stationary solution

We now rely upon the results stated in the previous subsection to deal with the solution of the Master Equation at the equilibrium.

The following proposition provides a necessary and sufficient condition for the existence of a 1-dimensional null space, i.e. a unique stationary distribution. The condition is that the digraph associated with the network of reactions has a globally reachable vertex¹, which is a milder assumption than strong connectivity; that is not a major assumption in networks of chemical reactions, where states can usually jump to adjacent states with non-zero probability per unit of time.

Proposition 12. *The stationary distribution of a discrete-state continuous-time Markov process is unique if and only if $\text{rank}(G) = \text{dim}(G) - 1$, namely if and only if the digraph associated with the Markov process has a globally reachable vertex. Under this assumption, 0 is a simple eigenvalue of G , with*

¹A globally reachable vertex v is a vertex of the digraph such that there exists a directed path from any node of the graph to v .

left eigenvector $\mathbf{1}^T$, and the other eigenvalues of G have negative real part. The stationary distribution is unique and is given by $\mathcal{P}_{ss} = \frac{u_0}{\mathbf{1}^T u_0}$, where u_0 is the right eigenvector corresponding to the eigenvalue 0. The second smallest eigenvalue of G (also called algebraic connectivity of the digraph associated with the Markov process) is related to the convergence speed to the stationary distribution.

The previous statement is a consequence of known properties of the Laplacian matrix [30], stating that $\text{rank}(L) = \text{dim}(L) - 1$ if and only if the digraph has a globally reachable vertex, and Theorem 10, implying that $\text{rank}(G) = \text{rank}(L)$.

Remark 13 (Duality with respect to consensus problems). *Theorem 10 and Proposition 12 highlight some interesting duality properties between the problem of finding the stationary solution of a CME and the so-called consensus problem, a very popular topic in the context of control of multi-agent systems [31]. In a common formulation of this problem, each node i in a network of agents, represented by a directed graph, is an integrator with dynamics $\dot{x}_i = u_i$. Starting from different local states x_i , the agents adopt an appropriate feedback rule u_i , just based on the state information of their own neighbors, in order to asymptotically reach a consensus (or agreement), namely a common value α of the individual state x_i for all the agents of the network. The closed-loop dynamics of the global state x , collecting all the individual states, can be restated as $\dot{x} = -Lx$, where L is the graph Laplacian. A consensus is reached if and only if $\text{rank}(L) = \text{dim}(L) - 1$, namely if and only if the digraph has a globally reachable vertex [30]. The right eigenvector $\mathbf{1}$ of L , associated to the 0 eigenvalue of L , corresponds to the left eigenvector $\mathbf{1}^T$ of G and is the consensus eigenvector, because the state converges to $\alpha\mathbf{1}$, namely all the local states asymptotically become equal. Note that, differently from the CME, the steady state vector is not uniquely determined, i.e. the consensus value α depends on the global initial state $x(0)$. The left normalized eigenvector w^T of L (with $\sum_i w_i = 1$), associated to the 0 eigenvalue of L , corresponds to right eigenvector (stationary distribution) \mathcal{P}_{ss} of G . The left eigenvector w^T plays a role in determining the agreement value of the network of agents as an appropriate convex combination of the local initial states: $\alpha = w^T x(0)$.*

5. Computational tools for the CME solution exploiting the block-tridiagonal realization

The computational burden in dealing with multi-dimensional CMEs is mainly dependent on the high dimension of G . This section is devoted to show the benefits in numerical computations provided by the block-tridiagonal realization of the system, induced by the state ordering considered in (7)–(9). Note that a random reordering of the state space destroys the recursive block-tridiagonal structure of the system, hence inhibiting, in general, the application of the methods illustrated hereafter. In the following, Assumption (H1) and the assumptions of Proposition 12 hold, namely we consider a generalized one-step process with a unique stationary distribution.

5.1. Tools to compute the stationary solution

The search for the stationary solution requires to solve the linear equation $G\mathcal{P} = \mathbf{0}$ in (18). The reader may refer to [44] as a reference for numerical tools involved in the solution of linear matrix equations for dense and sparse matrices. The proposed following approaches will properly account for the adopted block-tridiagonal realization.

Gaussian elimination method. Due to the sparsity of matrix G , the equilibrium problem $G\mathcal{P} = \mathbf{0}$ can be efficiently solved by means of the classical Gaussian elimination [32]. Moreover, because of the block-tridiagonal structure, matrix G is diagonally dominant and the diagonal pivoting is naturally employed. This ensures the numerical stability of the algorithm [44]. It is known that for dense matrices, the performance of Gaussian elimination is cubic with respect to the matrix dimension [48] but, in case of sparse matrices, one can benefit from the known position of the many zero entries. For instance, in the case of tridiagonal matrices, the Thomas algorithm [47] implements the Gaussian elimination with a performance that is linear with respect to the matrix dimension. The performance of the Gaussian elimination method will be tested in an example in Subsection 6.1.

Block LU decomposition method. The proposed block-tridiagonal structure to organize the CME plays an active role to lighten the computational burden when properly exploiting the LU decomposition. According to the Doolittle algorithm for the LU decomposition [33], matrix G can be factorized as $G = L \cdot U$, with L a unit (non-singular) lower triangular matrix and

U a (possibly singular) upper triangular matrix. Therefore, the computation of the stationary distribution provided by the solution of $G\mathcal{P}_{ss} = LU\mathcal{P}_{ss} = \mathbf{0}$ is reduced to the computation of the solution of the upper-triangular system $U\mathcal{P}_{ss} = \mathbf{0}$, by backward substitution.

Moreover, according to the block-tridiagonal structure and from Theorem 3, we know that matrix G in (13) assumes a simpler structure as:

$$G = \begin{bmatrix} G_0 & G_0^+ & \emptyset & \emptyset & \cdots & \cdots & \emptyset \\ G_1^- & G_1 & G_1^+ & \emptyset & \cdots & \cdots & \emptyset \\ \emptyset & G_2^- & G_2 & G_2^+ & \emptyset & \cdots & \emptyset \\ \vdots & \emptyset & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \emptyset \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & G_{N_1-1}^+ \\ \emptyset & \emptyset & \emptyset & \cdots & \emptyset & G_{N_1}^- & G_{N_1} \end{bmatrix}, \quad (20)$$

where we renamed $G_i^+ := G_{i+1}^{-1}$, $G_i^- := G_{i-1}^1$ and $G_i := G_i^0$ for all i , for an easier notation. Thus, without loss of generality, matrix G can be decomposed as $G = LU$, where L and U have the following structure:

$$L = \begin{bmatrix} L_0 & \emptyset & \emptyset & \cdots & \emptyset \\ L_1^- & L_1 & \emptyset & \cdots & \emptyset \\ \emptyset & L_2^- & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \emptyset \\ \emptyset & \cdots & \emptyset & L_{N_1}^- & L_{N_1} \end{bmatrix}, U = \begin{bmatrix} U_0 & U_0^+ & \emptyset & \cdots & \emptyset \\ \emptyset & U_1 & U_1^+ & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \emptyset \\ \vdots & \vdots & \ddots & \ddots & U_{N_1-1}^+ \\ \emptyset & \emptyset & \cdots & \emptyset & U_{N_1} \end{bmatrix}. \quad (21)$$

Algorithm 1 provides a scheme to compute the blocks in (21) (*Forward Elimination*) according to which the components of the steady-state distribution are achieved (*Backward Substitution*). Note that, taking advantage of

the block-bidiagonal structure of matrices L and U , the LU decomposition allows to manipulate blocks of reduced size with respect to G .

```

1: First Iteration:
    $[L_0, U_0] = \text{lu}(G_0);$ 
    $U_0^+ = L_0^{-1}G_0^+;$ 
2: Forward Elimination:
3: for ( $i = 1$  to  $N_1 - 1$ ) do
4:    $L_i^- = G_i^- U_{i-1}^{-1};$ 
    $[L_i, U_i] = \text{lu}(G_i - L_i^- U_{i-1}^+);$ 
    $U_i^+ = L_i^{-1}G_i^+;$ 
5: end for
6: Last Forward Elimination:
    $L_{N_1}^- = G_{N_1}^- U_{N_1-1}^{-1};$ 
    $[L_{N_1}, U_{N_1}] = \text{lu}(G_{N_1} - L_{N_1}^- U_{N_1-1}^+);$ 
7: Find  $\tilde{p}_{N_1} \geq \mathbf{0}$ ,  $\tilde{p}_{N_1} \neq \mathbf{0}$  s.t.  $U_{N_1} \tilde{p}_{N_1} = \mathbf{0}$ ;
8: Backward Substitution:
9: for ( $i = N_1 - 1$  to  $0$ ) do
10:   $\tilde{p}_i = -U_i^{-1} U_i^+ \tilde{p}_{i+1};$ 
11: end for
12:  $\tilde{p} = [\tilde{p}_0 \ \tilde{p}_1 \ \cdots \ \tilde{p}_{N_1}]'$ ;
13:  $\mathcal{P}_{ss} = \frac{\tilde{p}}{\mathbf{1}^T \tilde{p}};$ 

```

Algorithm 1: Computation of the stationary distribution via Block LU Decomposition for one-step processes. Differently from elsewhere in the paper (e.g. in Eq.(13)), the superscript '-1' here denotes the matrix inversion.

A few comments about Algorithm 1. We assume that the command *lu* denotes the Doolittle algorithm for the LU decomposition. Since G is singular and the stationary distribution is assumed to be unique, this necessarily results in matrix U_{N_1} being singular with 1-dimensional null space, hence \tilde{p}_{N_1} in line 7 exists and is uniquely determined up to the multiplication by a constant. Blocks L_i are invertible for all i and blocks U_i are invertible for $i < N_1$. The normalization in line 13 makes \mathcal{P}_{ss} unique. The algorithm will be exploited in an example in Section 6.1.

Reduced order system. A way to further speed-up the steady-state

computation is to properly exploit the mass-balance constraints of closed systems. In Eqs. (7)–(9), we proposed a recursive construction of \mathcal{P} including all the states (n_1, \dots, n_M) in the M -dimensional hyper-rectangular lattice $\{0, 1, \dots, N_1\} \times \dots \times \{0, 1, \dots, N_M\}$. Note that, in closed systems, all the states of the lattice not fulfilling the mass-balance constraints are not reachable (in the sense of Munsky et al. [17]) at any time from the set of *mass-balanced* states and have a zero steady-state probability. Therefore, one can erase *a priori* the components referred to those states from the vector \mathcal{P} , as well as the corresponding rows/columns in G , thus obtaining a reduced dynamics $\dot{\tilde{P}} = \tilde{G}\tilde{P}$ and a reduced equilibrium equation $\tilde{G}\tilde{P} = \mathbf{0}$. The reduced matrix \tilde{G} has dimension $\tilde{\mathcal{M}} \leq \mathcal{M}$, in general, and still possesses a recursive block-tridiagonal structure, and can be defined in terms of a more general matrix builder than the one in (1). In matrix \tilde{G} , in fact, keeping the same notation as in (1), the square blocks in $\{B_i\}$ are not equally dimensioned and the sequences $\{A_i\}$ and $\{C_i\}$ include, in general, rectangular blocks of appropriate dimensions. Matrix $\tilde{L} = -\tilde{G}^T$ is still the Laplacian of the subgraph induced by the set of mass-balanced states of the biochemical network. Hence \tilde{G} enjoys the same properties, with respect to such a subgraph, illustrated in Subsection 4.1 for G with respect to the original digraph, with the necessary modifications. In particular, the stationary distribution is unique if and only if the subgraph of the mass-balanced states contains a globally reachable node. In Subsection 6.2, we will show an example of the remarkable complexity reduction obtained by considering \tilde{G} instead of G in a practical case.

Clearly, the reduced system obtained by means of the mass balance constraints may also be applied to the Gaussian elimination algorithm and to the Block LU decomposition. In this last case, matrices $\{L_i\}$ and $\{U_i\}$ may well be rectangular, and the inverse matrices in Algorithm 1 may need to be reinterpreted as pseudo-inverses, which can lead to singularity issues.

5.2. Transient solution

The second part of the Section is devoted to addressing the computation of the explicit transient solution of the CME, formally given by (17). Unfortunately, the matrix exponential of a block tridiagonal matrix is not block-tridiagonal, in general, therefore the computation of the solution $\mathcal{P}(t)$ of the master equation does not take particular advantages of the one-step assumption. Furthermore, the large dimension of matrix G in general cases

makes the computation of e^{Gt} a hard task. Efficient methods to compute and approximate the matrix exponential can be found in [35, 39].

An alternative way to face the numerical problem could be to discretize the linear CME (17) according to a fixed step size Δt :

$$P_{k+1} = G_d P_k, \quad G_d = e^{G\Delta t}, \quad (22)$$

with $P_k = \mathcal{P}(k\Delta t)$ for any $k \in \mathbb{N}_0$. The exact discretization (22) does not introduce any approximation but, unfortunately, does not simplify the problem of computing the matrix exponential G_d . To this end, we can lighten the computational burden by truncating the Taylor expansion providing G_d . This way, for a given $h \in \mathbb{N}$, we build up the algorithm:

$$\tilde{P}_{k+1} = G_d^h \tilde{P}_k, \quad \tilde{P}_0 = \mathcal{P}(0), \quad G_d^h = \sum_{s=0}^h \frac{G^s (\Delta t)^s}{s!}, \quad (23)$$

providing \tilde{P}_k , a computationally affordable estimate of P_k , for any $k \in \mathbb{N}_0$.

Note that the matrix G_d^h is recursively block- $(2h + 1)$ -diagonal, so it is generally sparse. Furthermore, matrices G and G_d^h have the same eigenvectors with the eigenvalues given by

$$\Lambda(G_d^h) = \left\{ \sum_{s=0}^h \frac{\lambda_i^s (\Delta t)^s}{s!} : \lambda_i \in \Lambda(G) \right\},$$

where we denoted by $\Lambda(\cdot)$ the spectrum of a matrix. Note that, from Proposition 12, the right eigenvector u_0 , providing the unique stationary distribution as $\mathcal{P}_{ss} = \frac{u_0}{\mathbf{1}^T u_0}$, corresponds to the single eigenvalue 0 of G and to the single eigenvalue 1 of G_d^h , whatever order $h \in \mathbb{N}$ is chosen. As a matter of fact, if matrix G_d^h preserves the stability of the corresponding discrete-time system, the asymptotic solution coming out by iteratively running (23) definitely converges to the real steady-state solution, thus providing a further way to compute it. The problem is to design the step size Δt and the approximation order h in order to ensure G_d^h is Schur, i.e. has eigenvalues in the unitary circle.

The first order (Euler) approximation ($h = 1$), see e.g. [49], deserves a particular attention because the matrix G_d^1 is block-tridiagonal with same recursive structure of G , thus making the computation very efficient. The choice of the sampling time Δt in this case is addressed in the following result.

Proposition 14. Consider the first-order discrete-time approximate CME

$$\tilde{P}_{k+1} = G_d^1 \tilde{P}_k, \quad \tilde{P}_0 = \mathcal{P}(0), \quad (24)$$

where $G_d^1 = I + G\Delta t$, with I the \mathcal{M} -dimensional identity matrix. Assume that $\text{rank}(G) = \mathcal{M} - 1$ and

$$0 < \Delta t < \frac{1}{\max_{j=1, \dots, \mathcal{M}} |[G]_{jj}|}, \quad (25)$$

where the elements $[G]_{jj}$ are the diagonal entries of G . Then the discrete-time approximate CME in (24) converges (as $k \rightarrow +\infty$) to the exact stationary distribution \mathcal{P}_{ss} of the continuous-time CME in (10).

Proof 15. The proof is achieved by showing that (25) ensures that matrix G_d^1 is Schur. Indeed, by applying the Gershgorin circle theorem [34], the spectrum of G_d^1 , given by $\Lambda(G_d^1) = \{1 + \lambda_i \Delta t : \lambda_i \in \Lambda(G)\}$, can be bounded in the complex plane as follows:

$$\Lambda(G_d^1) \subset \bigcup_{j=1}^{\mathcal{M}} \mathbf{D}([G]_{jj}, \sum_{i \neq j} [G]_{ij}),$$

where $\mathbf{D}(a, b)$ denotes the closed disk centered at a with radius b . Since $G_d^1 = I + G\Delta t$, we can restate any element $[G]_{ij}$ in terms of the corresponding element $[G]_{ij}$ of G . So one gets:

$$\begin{aligned} \Lambda(G_d^1) &\subset \bigcup_{j=1}^{\mathcal{M}} \mathbf{D}(1 + [G]_{jj}\Delta t, \sum_{i \neq j} [G]_{ij}\Delta t) \\ &= \bigcup_{j=1}^{\mathcal{M}} \mathbf{D}(1 + [G]_{jj}\Delta t, -[G]_{jj}\Delta t), \end{aligned}$$

where we used the property that the columns of G add up to zero. Note that the disks are all centered on the real axis and the point $(1, 0)$ of the complex plane belongs to all the disks. Hence, each Gershgorin circle is tangent to the unit circle and can be made internally tangent provided that its radius $-[G]_{jj}\Delta t$ is smaller than 1. By imposing the joint conditions $-[G]_{jj}\Delta t < 1$ for all the circles, one gets:

$$\Delta t < \min_{j=1, \dots, \mathcal{M}} \frac{1}{|[G]_{jj}|} = \frac{1}{\max_{j=1, \dots, \mathcal{M}} |[G]_{jj}|}, \quad (26)$$

which concludes the proof.

Note that Proposition 14 provides an upper bound for Δt by avoiding the explicit computation of the spectrum of G_d^1 , which is computationally demanding. On the other hand, the choice of Δt imposed by (26) might be very conservative. As a matter of fact, it is common practice to try to increase its value until one gets a numerical blow-up.

Remark 16. *The tridiagonal form of (20) allows to efficiently compute the components of vector \tilde{P}_k . Indeed, analogously to partition (9), we may decompose*

$$\tilde{P}_k = [\tilde{P}_{k,0} \ \tilde{P}_{k,1} \ \cdots \ \tilde{P}_{k,N_1}]^T, \quad (27)$$

so that Eq. (24) becomes:

$$\tilde{P}_{k+1,j} = G_j^- \tilde{P}_{k,j-1} \Delta t + (I + G_j \Delta t) \tilde{P}_{k,j} + G_j^+ \tilde{P}_{k,j+1} \Delta t, \quad (28)$$

with $j = 0, \dots, N_1$, assuming $\tilde{P}_{k,0} = \tilde{P}_{k,N_1+1} = \mathbf{0}$ and $G_0^- = G_{N_1}^+ = \mathbf{0}$, and I the identity matrix with the same dimension of G_j . In Section 6.2, we will show the transient behavior of the Euler approximation in an example, in comparison with other CME solvers.

6. Simulation results

All the simulations in this section have been computed in the Matlab suite on an Apple MacBook Pro laptop with 2.5 GHz Intel Core i5 CPU and 16 GB RAM.

6.1. Orthogonal one-step process: miRNA-protein toggle switch

This first example is taken from [36], where a toggle switch involving a protein compound and a miRNA cluster is considered. According to the definitions introduced in Sections 2–3, the bivariate master equation modeling the biological framework under investigation provides an orthogonal one-step process with non-trivial transition probabilities:

$$\begin{cases} g_{n_1, n_2}^{1,0} &= \bar{\alpha} + \frac{k_1 n_1^2}{\Gamma_1 + n_1^2 + \Gamma_2 n_2} \\ g_{n_1, n_2}^{0,1} &= \beta + k_2 n_1 \\ g_{n_1, n_2}^{-1,0} &= \delta n_1 \\ g_{n_1, n_2}^{0,-1} &= \gamma n_2 \end{cases} \quad (29)$$

where species Y_1 and Y_2 represent the E2F-Myc complex and the miRNA cluster concentrations, respectively. The chosen parameters are $\bar{\alpha} = 1.68$, $\beta = 0.202$, $\delta = 0.2$, $\gamma = 0.2$, $k_1 = 90$, $k_2 = 0.05$, $\Gamma_1 = 10300$, $\Gamma_2 = 1006$ (see [36] and references therein for more details on meaning and measurement units of the model above). In [36], the stationary distribution for this process is not computed exactly, but a reduced one-dimensional model is studied, by exploiting the different time scale for the two reactions. In particular, n_2 is considered as the fast variable and the value of its steady state is computed by imposing $g_{n_1, n_2}^{0,1} = g_{n_1, n_2}^{0,-1}$ in (29), to obtain $n_2 = \frac{\beta + k_2 n_1}{\gamma}$. Such a value is substituted into the bivariate master equation, thus obtaining an approximate scalar CME, whose stationary distribution can be easily computed analytically. This example shows the possibility of poor agreement of the 1D approximation driven by different time scales with the original 2D system.

We applied the Gillespie Stochastic Simulation Algorithm (SSA) [5] to the described model. We repeated the stochastic simulation for the example above by means of $2 \cdot 10^4$ Monte Carlo runs of SSA, with a time horizon of 10^3 seconds and at most 10^7 observed reactions for each run, and we plotted the statistics of the occurrences of the steady states for n_1 . We compared them to the previously described 1D stationary distribution and to the solution of $G\mathcal{P} = \mathbf{0}$, obtained by means of the methods illustrated in Section 5. The matrix G was built by approximating the model with a closed system (as in [36]) with a sufficiently large number of copies ($N_1 = 300$, $N_2 = 80$), chosen so that the probability of generation of further molecules is negligible. While the Monte Carlo runs of the Gillespie Algorithm took several hours, the computation of the 2D stationary distribution by sparse Gauss Elimination (see Section 5.1) was computed in just 45 seconds. The computation time was further reduced in methods explicitly exploiting the block tridiagonal form of G . For instance, by employing the Block LU decomposition (Algorithm 1 in Section 5.1), it reduced to just 0.62 seconds, comparable to the most efficient (purely numerical) known methods to calculate the eigenvector corresponding to the smallest magnitude eigenvalue (computed in 0.4 seconds by using the Matlab sparse linear algebra function “eigs”). One can indeed speed up the statistical computation and reduce it to 24 minutes, by exploiting the *ergodic* properties of the process through a very long SSA run including over 10 million reactions, and by inferring the statistical distribution by computing the average recurrence time in each state of the process [37]. Although the time computation is much lower, such a method is still outperformed by the

aforementioned theoretical methods.

The plots in Figure 4 show the agreement of the statistical estimation (by SSA simulation) of the steady-state marginal distribution of species Y_1 with the one obtained from the $2D$ theoretical stationary distribution (which is the same for the two methods given in Sections 5.1), as well as the mismatch with respect to the $1D$ approximate steady-state distribution.

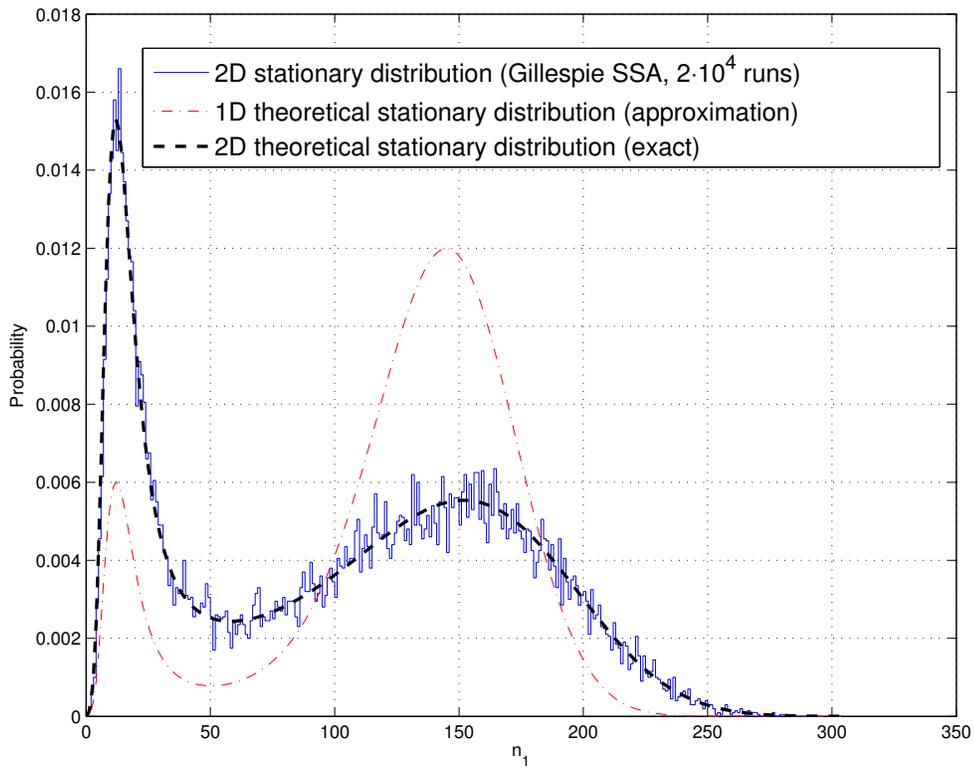
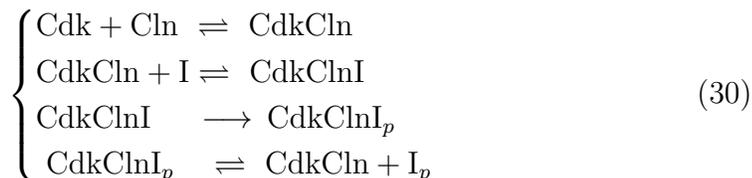


Figure 4: Comparison among the statistics of steady states provided by the Gillespie Algorithm (solid blue line), the $2D$ theoretical stationary distribution (dashed black line) and the $1D$ approximate stationary distribution (dash-dotted red line).

6.2. *Non-orthogonal one-step process: CDK activation*

Let us consider the following $q = 7$ biochemical reactions:



which represent a very general framework describing the activation of *cyclin-dependent kinases* (*Cdk*), a family of protein kinases playing a crucial role in regulating the cell cycle (see e.g. [38] and references therein for further details). In order to phosphorylate their target proteins, *Cdks* are required to bind to a proper cyclin, first reaction in (30); their action, however, is usually delayed by the binding of an inhibitor, second reaction in (30); the *Cdk* – *cyclin* complex eventually gets rid of the inhibitor after a (not reversible) phosphorylation, third reaction in (30), followed by the dissociation of the phosphorylated inhibitor, last reaction in (30).

We choose the $M = 4$ independent species as $Y_1 = \text{Cdk}$, $Y_2 = \text{I}$, $Y_3 = \text{CdkClnI}$ and $Y_4 = \text{I}_p$. According to the standard mass-action law, the transition probabilities can be considered to be proportional to the number of molecules of the reacting agents. As a result, a non-orthogonal one-step process can be defined, described by a 4D master equation, whose non-trivial transition probabilities are given by:

$$\left\{ \begin{array}{l} g_{n_1, n_2, n_3, n_4}^{-1, 0, 0, 0} = h_1 n_1 (b_2 - b_1 + n_1) \\ g_{n_1, n_2, n_3, n_4}^{1, 0, 0, 0} = h_2 (b_1 - b_3 - n_1 + n_2 + n_4) \\ g_{n_1, n_2, n_3, n_4}^{0, -1, 1, 0} = h_3 n_2 (b_1 - b_3 - n_1 + n_2 + n_4) \\ g_{n_1, n_2, n_3, n_4}^{0, 1, -1, 0} = h_4 n_3 \\ g_{n_1, n_2, n_3, n_4}^{0, 0, -1, 0} = h_5 n_3 \\ g_{n_1, n_2, n_3, n_4}^{0, 0, 0, 1} = h_6 (b_3 - n_2 - n_3 - n_4) \\ g_{n_1, n_2, n_3, n_4}^{0, 0, 0, -1} = h_7 n_4 (b_1 - b_3 - n_1 + n_2 + n_4) \end{array} \right.$$

where $b = (b_1, b_2, b_3)^T$ is the vector collecting the total mass of the three elementary species *Cdk*, *Cln* and *I*, being *I* and *I_p* two different forms of the same chemical player. We set $b = (30, 15, 30)^T$ and $h_j = 0.05$, for $j = 1, \dots, q$.

In the following, our goal is to find the transient solution $\mathcal{P}(t)$ of the CME in (10) for the 4D model described above, with initial condition $\mathcal{P}(0) = \bar{\mathcal{P}}$ defined by:

$$\bar{\mathcal{P}}_{n_1, n_2, n_3, n_4} = \begin{cases} 1 & \text{if } (n_1, n_2, n_3, n_4) = (15, 15, 15, 0), \\ 0 & \text{otherwise.} \end{cases}$$

Then, we use the transient distribution $\mathcal{P}(t)$ to compute the probability (as a function of time) that the number of molecules of the *active* kinase $CdkCln$ exceeds the number of molecules of the *inhibited* form $CdkClnI$. Indeed, the instant when a $CdkClnI$ gets rid of the inhibitor (formally when free $CdkCln$ exceeds $CdkClnI$ in deterministic methods) is usually considered as the onset of the Cdk functional activity [38]. For the simulation we consider a time horizon $[0, T]$, with $T = 300s$. The experiment can be repeated for higher absolute values of masses and different parameters, but the results and the plots are qualitatively similar.

Since $N_1 = N_2 = N_4 = 30$ and $N_3 = 15$, the dimension of matrix G is $\mathcal{M} = 31^3 \cdot 16 = 476,656$, but one can reduce this value by erasing from the vector of steady-state probabilities \mathcal{P} the components referred to states violating the mass constraints, as described in Section 5.2. The dimension of the reduced matrix \tilde{G} is $\tilde{\mathcal{M}} = 19,176$ (about 1/25 of the dimension \mathcal{M} of G).

Different solvers will be compared in achieving the CME solution. Ideally, computing the exact solution requires the computation of the matrix exponential of \tilde{G} . This is not an easy task, even for the reduced-order system, and the computation is performed in about 75 minutes. Alternatively, the exact iteration in (22) can be applied, and an accurate solution can be computed iteratively by means of Expokit [45], which implements the Krylov subspace method, see e.g. [35]. This approach required 26 seconds with a fixed time step of $\Delta t = 1s$. Note that the complexity of the method is strictly dependent on the number time points considered, so a lower computation time can be obtained at the expense of a more coarse sampling.

We then consider the first order (Euler) approximation of the matrix exponential described in Section 5.2, with time step $\Delta t = 0.04s$ computed by exploiting inequality (25) to ensure convergence to the exact stationary distribution \mathcal{P}_{ss} , in agreement with Proposition 14. Furthermore, the efficient iteration described in Eqs. (27)-(28) is exploited. The whole transient computation required 70 seconds.

As a term of comparison, we also consider a Krylov-based FSP Algorithm based on [40]. With respect to the original FSP [16], where the size of the state projection is tuned according to the final time T and on the approximation ε required, the Krylov-FSP algorithm allows to simulate the time horizon $[0, T]$ by approximating (via Krylov subspaces) the iteration in (22) and by using progressively increasing state projections, where the expansion is based on the concept of N -step reachability [16], which is conceptually analogous to the breadth-first search applied to the reaction graph. In our case, the whole network of reactions (including $\tilde{\mathcal{M}} = 19,176$ states) is visited in 64 expansion steps, while the Krylov-FSP projection at time T , with $\varepsilon = 0.01$, requires 52 expansion steps and includes 16850 states (-12% with respect to $\tilde{\mathcal{M}}$), so the computational saving in terms of spatial complexity is not remarkable. On the other hand, the time spent for the computation of the Krylov-based FSP solution resulted to be higher than the Euler approximation, mainly due to the computation of the successive approximation spaces, which required several minutes (although it may be possibly improved, depending on the actual implementation of the expansion algorithm and on the data structures involved therein).

The simulation results are shown in Figure 5. It is readily seen that the exact solution computed from (17) is almost indistinguishable from those obtained by means of the approximate methods: the one exploiting the Krylov subspace method [45], the Euler approximation in (27)-(28) and the one from the Krylov-based FSP [40]. The probability exceeds the median value 0.5 after about 114 seconds, when the event that the molecules of $CdkCln$ exceed those of $CdkClnI$ becomes more likely to happen than not. We let the reader note that, since reaction $CdkClnI \longrightarrow CdkClnI_p$ in (30) is not reversible, all the molecules of $CdkClnI$ eventually vanish. This is in agreement with Figure 5, where the probability goes to 1 as time increases.

We remark that the exact probabilistic information provided by our approach cannot be provided by any deterministic models based on the average or on the reagent concentrations, such as the macroscopic deterministic ODE computed from the first-order jump moment of the CME (see [4], Chapter 5, for further details). Indeed, since the obtained equation is not linear, it is known that the average dynamic behavior obtained by integrating numerically the nonlinear reaction-rate equation is an approximation of the result obtained solving exactly, by means of (17), the linear (higher-dimensional) CME. As shown in Figure 6, according to the numerical solution of the con-

centration equations, the expected number of *CdkCln* molecules exceeds the expected number of *CdkClnI* ones after about 138 seconds, with a percent relative error around 17%.

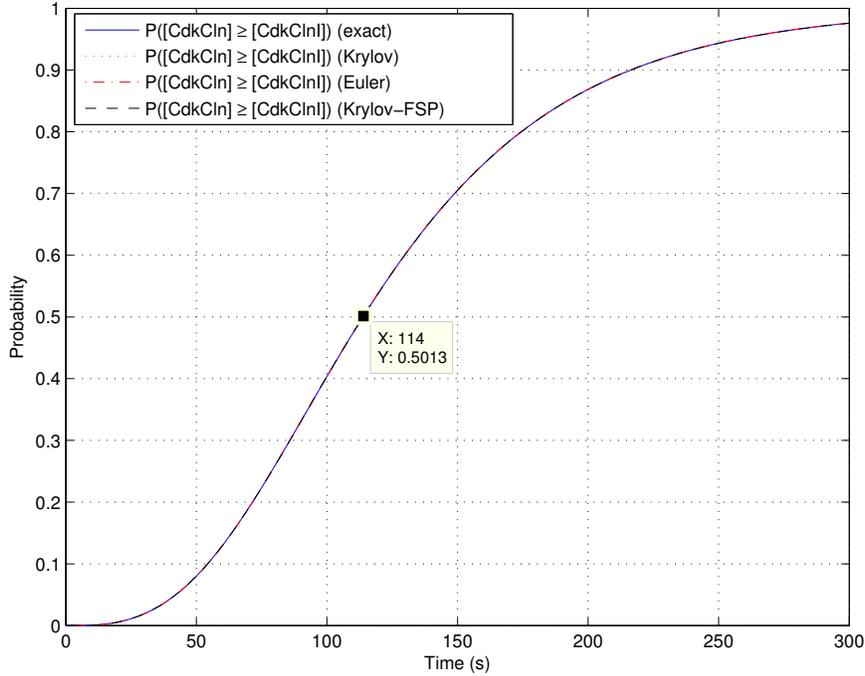


Figure 5: Plot of the probability (as a function of time) that the number of molecules of *CdkCln* exceeds the number of molecules of *CdkClnI* in the cyclin-dependent kinase (*Cdk*) reaction network, according to the exact solution (blue solid line) and to the approximate solutions: the one based on the Krylov subspace method (magenta dotted line), the Euler approximation (red dash-dotted line), and the one from the Krylov-based FSP (black dashed line). The trajectories are almost indistinguishable. After about 114 seconds, the event that the molecules of *CdkCln* exceed those of *CdkClnI* is more likely to happen than not.

7. Conclusions

In this work we presented some results on the dynamical properties and the efficient solution of the Chemical Master Equation, with a particular focus

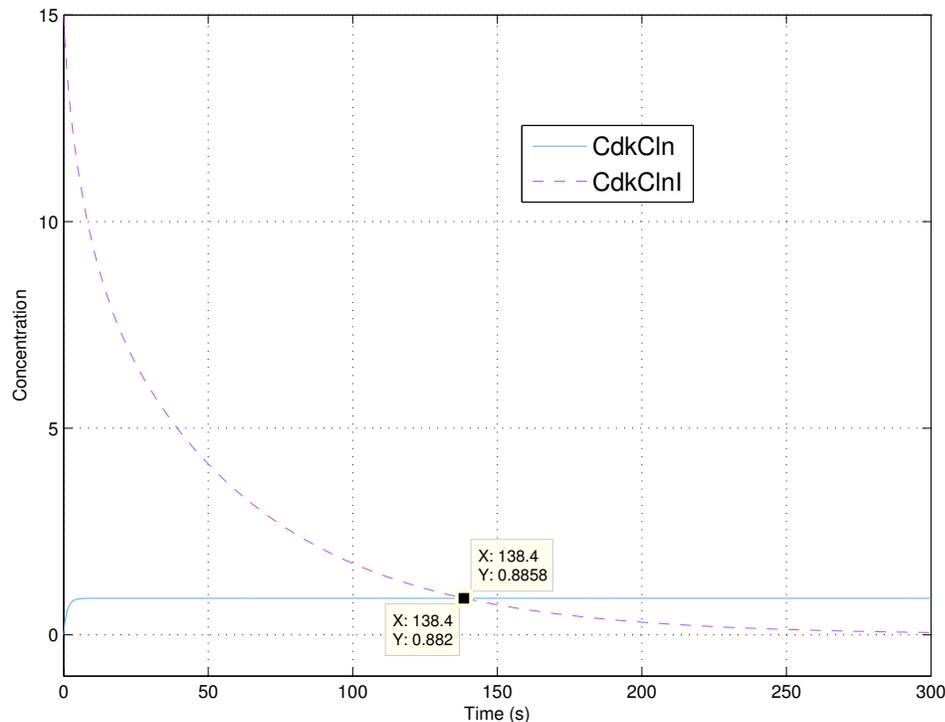


Figure 6: Plot of the solution of the reaction-rate deterministic equation obtained by means of the first jump moment [4] of the Cdk master equation. The state variables are here intended either as concentrations or as approximations of the expected populations. The expected number of $CdkCln$ molecules exceeds the expected number of $CdkClnI$ ones after about 138 seconds.

on the exact equilibrium distribution. The recursive structure of the dynamic matrix describing the stochastic evolution of the chemical population relies on a very common one-step assumption. Indeed, this is not a restriction because the property of orthogonal one-step process can always be recovered by means of a generalized definition of independent (non-redundant) species.

The computation of the solution does not exploit any further hypotheses, e.g. detailed balance property or reversibility of the Markov process, so it can be applied to very general cases. Moreover, many conceptual results of the present paper still hold in the infinite dimensional case, with the necessary

modifications, just requiring some additional technical effort.

The application of the method to real biochemical networks appears to be promising in that the approach is accurate and allows a cheap management of the computational resources when compared to other solvers and to the extensive use of stochastic simulation with the aim of approximating the probability distributions.

Acknowledgements. The authors wish to thank Gastone Castellani for introducing them to the topic of CMEs and Pino Pontrelli for the helpful discussions about numerical stability.

References

- [1] J. Mettetal and A. van Oudenaarden, Necessary noise, *Science*, 317, 463, 2007.
- [2] R. Bahar, C.H. Hartmann, K.A. Rodriguez, A.D. Denny, R.A. Busuttil, M.E. Doll, R.B. Calder, G.B. Chisholm, B.H. Pollock, C.A. Klein and J. Vijg, Increased cell-to-cell variation in gene expression in ageing mouse heart, *Nature*, 441, 1011–1014, 2006.
- [3] F.J. Bruggeman, N. Blüthgen and H.V. Westerhoff, *Noise management by molecular networks*, *PLoS Computational Biology*, 5(9), 1-11. 2009.
- [4] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North Holland, third edition, 2007.
- [5] D. T. Gillespie, Exact Stochastic Simulation of Coupled Chemical Reactions, *The Journal of Physical Chemistry* 81(25), 23402361, 1977.
- [6] D.T. Gillespie, A rigorous derivation of the chemical master equation, *Physica A*, 188, 404–425, 1992.
- [7] P. Cazzaniga, D. Pescini, D. Besozzi, G. Mauri, S. Colombo and E. Martegani, Modeling and stochastic simulation of the Ras/cAMP/PKA pathway in the yeast *Saccharomyces cerevisiae* evidences a key regulatory function for intracellular guanine nucleotides pools, *Journal of Biotechnology*, 133, 377-385, 2008.

- [8] G. Lillacci and M. Khammash, A distribution-matching method for parameter estimation and model selection in computational biology, *Int. J. Robust. Nonlinear Control*, 22, 1065-1081, 2012.
- [9] C. Zechner, J. Ruess, P. Krenn, S. Pelet, M. Peter, J. Lygeros and H. Koepl, Moment-based inference predicts bimodality in transient gene expression, *PNAS*, 109(21), 8340-8345, 2012.
- [10] H. Koepl, C. Zechner, A. Ganguly, S. Pelet and M. Peter, Accounting for extrinsic variability in the estimation of stochastic rate constants, *Int. J. Robust. Nonlinear Control*, 22, 1103-1119, 2012.
- [11] M.A. Gibson and J. Bruck, Efficient exact stochastic simulation of chemical systems with many species and many channels, *J. Phys. Chem.* 104, 18761889, 2000.
- [12] D.T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems, *J. Chem. Phys.* 115(4), 1716-1733, 2001.
- [13] Y. Cao, D.T. Gillespie and L.R. Petzold, Efficient step size selection for the tau-leaping simulation method, *J. Chem. Phys.* 124, 044109, 2006.
- [14] L.B. Blyn, B.A. Braaten, C.A. White-Ziegler, D.H. Rolfson and D.A. Low, Phase-variation of pyelonephritis-associated pili in *Escherichia coli*: evidence for transcriptional regulation, *EMBO J.*, 8(2), 613-620, 1989.
- [15] T. Gardner, C. Cantor and J. Collins, Construction of a toggle switch in *Escherichia coli*, *Nature* 403(6767), 339342, 2000.
- [16] B. Munsky and M. Khammash, The finite state projection algorithm for the solution of the chemical master equation, *J. Chem. Phys.*, 124, 044104, 2006.
- [17] B. Munsky and M. Khammash, The finite state projection approach for the analysis of stochastic noise in gene networks, *IEEE Trans. Autom. Control, Special Issue on Systems Biology*, 201-214, 2008.
- [18] B. Munsky and M. Khammash, Transient analysis of stochastic switches and trajectories with applications to gene regulatory networks, *IET Syst. Biol.*, Vol.2(5), 323-333, 2008.

- [19] C.S. Gillespie, Moment-closure approximations for mass-action models, *IET Syst. Biol.*, Vol.3(1), 52-58, 2009.
- [20] A. Singh and J. P. Hespanha, Approximate Moment Dynamics for Chemically Reacting Systems, *IEEE Trans. Automat. Contr.* 56(2), 414-418, 2011.
- [21] P. Smadbeck, Y.N. Kaznessis, A closure scheme for chemical master equations, *PNAS*, Vol.110(35), 14261-14265, 2013.
- [22] S. Engblom, Spectral approximation of solutions to the chemical master equation, *J. Comput. Appl. Math.*, 229, 208-221, 2009.
- [23] T. Jahnke, T. Udrescu, Solving chemical master equations by adaptive wavelet compression, *J. Comput. Phys.*, 229, 5724-5741, 2010.
- [24] A. Borri, F. Carravetta, G. Mavelli, P. Palumbo, Some Results on the Structural Properties and the Solution of the Chemical Master Equation, *Proceedings of the 2013 American Control Conference (ACC 2013)*, Washington, DC, USA, pp. 3777-3782, 2013.
- [25] F. Carravetta, Nearest-neighbour modelling of reciprocal chains, *Stochastics: An International Journal of Probability and Stochastic Processes.* 80 (6), 525-584, 2008.
- [26] F. Carravetta, 2D-Recursive Modelling of Homogeneous Discrete Gaussian Markov Fields, *IEEE Transactions on Automatic Control.* 56 (5), 1198-1203, 2011.
- [27] M. Ullah, O. Wolkenhauer, *Stochastic Approaches for Systems Biology*, Springer, 2011.
- [28] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*, Series on Pure and Applied Mathematics, Wiley-Interscience, New York, 2000.
- [29] K. R. Ghusinga, A. Singh, First-passage time calculations for a gene expression model, *Proceedings of the 2014 Conference on Decision and Control (CDC 2014)*, Los Angeles, CA, USA, pp. 3047-3052, 2014.

- [30] F. Bullo, J. Cortes and S. Martinez, Distributed Control of Robotic Networks, *Series in Applied Mathematics*, Princeton, 2009, ISBN 978-0-691-14195-4.
- [31] R. Olfati-Saber, J. Fax and R.M. Murray, Consensus and cooperation in multi-agent networked systems, *Proceedings of IEEE* 95(1), 215-233, 2007.
- [32] H. Anton, Elementary linear algebra, John Wiley & Sons, 2010.
- [33] W. H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, LU Decomposition and Its Applications. *Cambridge, England: Cambridge University Press*, pp. 34-42, 1992.
- [34] R. S. Varga, Gershgorin and His Circles, *Springer Series in Computational Mathematics*, vol. 36, 2004.
- [35] C. Moler, and C. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM review* 45.1 (2003): 3-49.
- [36] E. Giampieri, D. Remondini, L. de Oliveira, G. Castellani and P. Lió, Stochastic analysis of a miRNA-protein toggle switch, *Molecular Biosystems*, 7(10), 2796-2803, 2011.
- [37] P. Walters, An Introduction to Ergodic Theory, *Graduate Texts in Mathematics*, Springer Verlag, 1982.
- [38] L. Alberghina, G. Mavelli, G. Drovandi, P. Palumbo, S. Pessina, F. Tripodi, P. Coccetti, and M. Vanoni, Cell growth and cell cycle in *Saccharomyces cerevisiae*: basic regulatory design and protein-protein interaction network, *Biotechnology Advances*, 30 (1), 52-72, 2012.
- [39] W. J. Stewart, Introduction to the Numerical Solution of Markov Chains, Vol. 41, *Princeton University Press*, 1994.
- [40] K. Burrage, M. Hegland, S. Macnamara, and R. B. Sidje, A Krylov-based finite state projection algorithm for solving the chemical master equation arising in the discrete modelling of biological systems, *Proceedings of the Markov 150th Anniversary Conference*, pp. 21-37, 2006.

- [41] M. Hegland, C. Burden, L. Santoso, S. MacNamara and H. Booth, A solver for the stochastic master equation applied to gene regulatory networks, *J Comp. App. Math.*, 205(2), 708-724, 2007.
- [42] V. Sunkara, The Chemical Master Equation With Respect to Reaction Counts, *18th World IMACS/MODSIM Congress*, 2009.
- [43] V. Sunkara, Analysis and Numerics of the Chemical Master Equation, Ph.D. thesis, Australian National University, 2013.
- [44] I.S. Duff, A.M. Erisman and J.K. Reid, Direct Methods for Sparse Matrices, *Oxford University Press*, 1989.
- [45] R. B. Sidje, Expokit: a software package for computing matrix exponentials, *ACM Transactions on Mathematical Software (TOMS)*, 24.1, 130-156, 1998.
- [46] F. Lopez-Caamal, T.T. Marquez-Lago, Order Reduction of the Chemical Master Equation via Balanced Realisation, *PLoS ONE*, 9(8), 2014.
- [47] L.H. Thomas, Elliptic Problems in Linear Differential Equations over a Network, *Watson Sci. Comput. Lab Report*, Columbia University, New York, 1949.
- [48] V. S. Ryaben'kii, and S. V. Tsynkov, A theoretical introduction to numerical analysis, *CRC Press*, 2006.
- [49] S. D. Conte and C. W. De Boor, Elementary Numerical Analysis: An Algorithmic Approach (3rd ed.), *McGraw-Hill Higher Education*, 1980.