

Extending the Artificial Chemistry to Design Networking Algorithms with Controllable Dynamics

Massimo Monti, Manolis Sifalakis, Christian F. Tschudin

Technical Report CS-2012-003

University of Basel

July, 2012

Abstract

Predicting and controlling the dynamics of communication protocols is as important as designing their main functional purpose. However, describing networking protocols' dynamics and analyzing their stability are difficult tasks when the classical finite-state-machine approach is used.

We propose to design networking algorithms which exhibit predictable dynamics. More specifically in a bottom-up manner, we design algorithms whose dynamics are driven by reaction laws on microscopic level and that, on macroscopic level, exhibit predictable behaviors emerging from the underlying models.

Our approach to design algorithms for networking (i) focuses on dynamics, and (ii) is not overspecialized on particular applications but rather represents a systematic design approach.

Keywords: Distributed Algorithms, Non-Work-Conserving-Schemes, Chemical Networking Protocols, Flow Control.



Introduction

Predicting and controlling the dynamics of communication protocols are difficult but important challenges in communication systems engineering. While in the early days of computer networks, the researchers' attention was on merely functional aspects, in the last three decades, the focus has shifted towards mastering also the protocol dynamics.

Common techniques for analyzing protocols and related algorithms rely first on building an abstract model and then reducing it to an analytically tractable system. For example, fluid models focus on the average flow-level properties by abstracting away the details of the underlying stochastic behavior [1, 2, 3]. Such abstractions however, bear significant limitations: they have to be extracted manually from a given source code, and they often over-simplify the behavior of the real implementation.

In this paper, we extend the design methodology of Chemical Networking Protocols (CNPs) proposed in [4] and apply it to design various network algorithmic operations. This methodology focuses on dynamics and analyzability since the early design-stages. By using these Pseudo-Chemically-Driven Algorithms (PCDAs), we design mechanisms whose dynamics are driven by *reaction* laws on microscopic level, and that, on macroscopic level, exhibit predictable behaviors emerging from the underlying functional models. The result is that the average trajectory of these algorithms can be described by a set of Ordinary Differential Equations (ODEs) and analyzed at steady-states as well as at transient-phases.

We propose PCDAs as predictable non-work-conserving schemes for scheduling events in communication networks at different layers, e.g. regulation of packets' transmission as a flow control mechanism at ISO-OSI transport layer, or regulation of media access at link layer.

We study PCDAs' steady-state behavior and their transient behavior for arbitrary initial conditions. This analysis bases on fluid models, is performed in the frequency domain, and exploits tools that are well established in fields such as signal processing or control theory. With the gained insights, we are able to simplify the design phase of these special algorithms by identifying and optimizing their key parameters that dictate their dynamic behavior.

The remainder of this paper is organized as follows: Sec.1 introduces Pseudo-Chemically Driven Algorithms (PCDAs) and Sec.1.1 clarifies this new algorithm-design paradigm through a simple example: the PCD rate-controller. Sect. 2 describes the analysis method, Sec.2.5 clarifies the analysis method by applying it to the rate-controller, and Sec.2.6 compares the analysis results with measurements done in a simulation environment. Sect. 3 focuses on algorithms' microkinetics. Finally, Sec.4 discusses the features of our approach.

Space and Scope of our Contribution

In network protocol design, the idea of exploiting behaviors and systems of our surrounding nature is not new: the literature reports many trials (successful or not) that exploit biological models (e.g. swarm-based systems, firstly introduced in [5]), physical fields (e.g. EM, temperature, and other generic artificial force fields such as potentials in [6]) and bio-chemical models (e.g. reaction diffusion for epidemic routing in [7]).¹ However, the majority of research takes care of the spatial distribution/diffusion of information (e.g. routing) or exploits models of other fields as a solution

¹We cite one work for each example only but many others could be mentioned.

only for optimization problems. Instead, we propose a new paradigm for the complete design of algorithms which (i) focuses on dynamics, (ii) does not relate to specific applications but rather represents a general design paradigm, and (iii) does not care directly about the spatial-aspect of the distribution of contents (information) but rather cares about timing and event scheduling.

This work extends and broadens the scope of some of the ideas that Meyer has proposed in [4]: Chemical Networking Protocols (CNPs) – the use of a chemical metaphor and reaction model in the design of self-healing protocols for computer networks whose dynamics are induced by artificial chemical reaction networks. Following the CNP approach, we let algorithms be driven by underlying reaction models. The direct coupling between packet quantity and packet rate, which is enforced by reaction laws (e.g. Law of Mass Action–LoMA), permits to automatically build a deterministic ODE model that describes the algorithm’s dynamics. Actually, CNPs try to exactly reproduce chemical reactions’ dynamics by implementing the “real-time next reaction algorithm”, a variant of an algorithm originally proposed in [8] that simulates real chemical reactions. In our work, we extend CNPs by removing this constraint and deviating from what happens in real chemistry. By allowing also non-exponentially distributed inter-reaction times, which define the model’s kinetics at the microscopic level, PCDAs exhibit controllability, enhanced functionality and analyzability. Our contribution lies also in the type and the scope of analysis: we care about steady-states analysis and ODE description (in the time domain) only as a first step towards the study of algorithms’ transient behavior. We analyze the algorithm with a signal-processing approach in the frequency-domain, gaining in this way insights on the algorithm’s response to arbitrary perturbations and on the stability of the algorithm response.

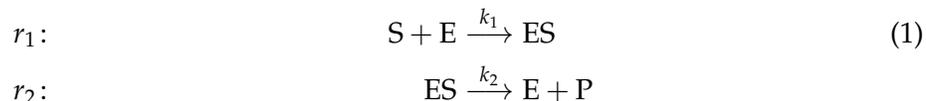
Our work is close conceptually to what proposed in [9]: design algorithms through a control-theoretic approach by basing on an analytical model of network transients. As suggested in [9], we avoid to use queueing network models (cumbersome or impossible) but rather we base on a deterministic description of the average of algorithm’s trajectory. Again as in [9], we use a state-space description of the algorithm and we make observations on the algorithm stability for small perturbations around fixed stable points. However, unlike what has been proposed in [9], (i) PCDAs are actually driven by the underlying, completely-predictable reaction model and thus the fluid model related to the algorithm dynamics comes “for free” in form of a set of Ordinary Differential Equations (ODEs). (ii) The fluid model we study does not try to describe the algorithm (also making use of assumptions) but rather represents the algorithm itself. (iii) Our analysis is in continuous-domain rather than in discrete one.

1 Pseudo-Chemically Driven Algorithms (PCDAs)

Traditionally, algorithm execution is handled by a state machine that, upon reception of a packet, synchronously changes its internal state. Additionally, algorithms for computer networks are usually designed based on empirically gained experiences and most of the time, their actual dynamics turn out to be unpredictable and not fully understood. Here, following in Meyer’s footsteps [4], we propose networking algorithms (albeit not for protocols only) whose dynamics are driven by Pseudo-Chemical (PC) reaction models: Network events, e.g. packet arrivals or generations, cause the production of virtual molecules in the PC model, which represents the control plane of the algorithm’s dynamics. Molecules react with other molecules and the obtained web of reactions can be used to perform a computation. When several computer-network nodes (e.g. hosts) interact with each other and implement PCDAs, the overall system can be seen as

unique PC reaction network that performs a *distributed* computation.

Each network node contains a multiset of finite sets of molecular species $\mathcal{S} = \{s_1, \dots, s_{|\mathcal{S}|}\}$. In addition, a set of reaction rules $\mathcal{R} = \{r_1, \dots, r_{|\mathcal{R}|}\}$ expresses which molecules can react and which molecules are generated during this process. Such a set of reactions, which constitute an algorithm, can be represented by reaction equations, e.g.



For example, reaction r_1 consumes, if present, two molecules of species S and E from the local multiset and generates an ES-species molecule.

Formally, a reaction is given by the equation

$$r \in \mathcal{R}_i: \quad \sum_{s \in \mathcal{S}_i} \chi_{sr} s \xrightarrow{k_r} \sum_{s \in \mathcal{S}_i} \zeta_{sr} s.$$

The coefficient k_r defines the reaction speed; the non-negative integers χ_{sr} and ζ_{sr} are the stoichiometric *reactant* and *product* coefficients that denote the number of s -species molecules *consumed* and *produced* by reaction r , respectively.

A node that receives a molecule places it into its local multiset. Reaction rules that may consume (=process) the new molecules are however not executed immediately but rather scheduled for a later, random time. Similarly to what artificial chemistry proposes, such an algorithm \mathcal{A} forces a reaction law on macroscopic level, e.g. the Law of Mass Action – LoMA.² In chemistry, the LoMA [10, 11] states that the reaction rate is proportional to the concentration (quantity) c of all reactant molecules: In our case, e.g. referring to reaction r_1 , S and E-species molecules are consumed at *average* rate $v_1 = k_1 c_S c_E$.

1.1 PCDA example – A simple rate-controller

As an example of what introduced so far, we propose a simple PCDA that can be used as an algorithm (i) to decouple the generation (or arrival) rate of packets from the transmission (sending) rate in a certain host and (ii) to guarantee that the transmission rate always respects a certain prefixed threshold. By deviating from traditional work-conserving schemes, the host equipped with this rate limiting algorithm waits before sending the newly generated packet. The sending event is defined by the Pseudo-Chemical (PC) model dynamics. Figure 1 shows a host that includes a source module, modeling the process of packet generation (or arrival process), a queue where packets pending for the transmission are stored, and a server that, upon the reception of a triggering signal from the PC model, serves the aforementioned queue and sends the packet.

For each generated packet, an S-molecule is created in the PC model. S and E-molecules are consumed to create an ES-molecule through r_1 -reaction, and r_2 -reaction consumes an ES-molecule to create again an E-molecule and a P-molecule. In biochemistry, this model is well known since it regulates enzyme-substrate interactions, and its kinetics are known as *Michaelis-Menten* kinetics [12, 13].

The reaction network’s topology in fig.1 is characterized by a conserved moiety: a molecular subgroup, e.g. E-ES, whose mass is constant during the evolution of the reaction network, e.g.

²The algorithm \mathcal{A} is the one for implementing the artificial chemistry. It is not the networking algorithm itself.

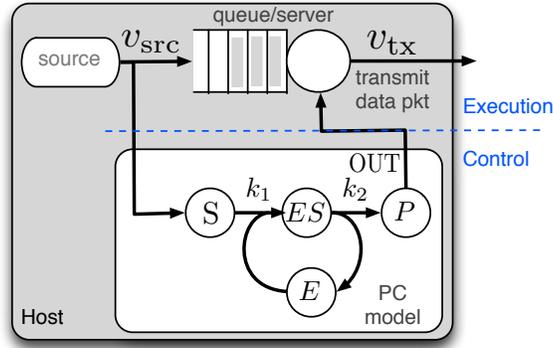


Figure 1: Host elements. A PCDA for rate limiting that regulates the host’s packet transmission, guaranteeing the respect of a prefixed, adjustable threshold of the host transmission rate.

$c_E + c_{ES} = e_0$. The moiety represents the key mechanism in order to implement a rate-limiting action. Indeed, according to the LoMA, any reaction that consumes a species part of the moiety experiences a limitation in the rate, e.g. the rate v_2 of r_2 -reaction has $e_0 k_2$ as a maximum value when all molecules in the closed loop are of ES-species. The described phenomenon is caught by the Michaelis-Menten equation:

$$v_2 = k_2 c_{ES} = e_0 k_2 \frac{c_S}{(k_2/k_1) + c_S} \quad (2)$$

By looking at (2), we can state that even if S-species increases without limits ($c_S \rightarrow \infty$), the rate of r_2 -reaction assumes a finite value: $v_2 = e_0 k_2$.

2 Analyzing PCDA’s dynamics

This section exemplifies how Metabolic Control Analysis (MCA) [14, 15] and frequency response analysis can be applied to PCDA’s in order to obtain the control over the parameters that affect the dynamic behavior of the system. First, we represent the algorithm’s dynamic state through a system of ODEs (Sec.2.1); second, we show how to linearly approximate eventual, non-linear systems around their steady state (Sec.2.2); third, we express the reaction network as a Linear Time Invariant (LTI) system (Sec.2.3); and fourth, we analyze the system’s response in the frequency domain (Sec.2.4). Section 2.5 clarifies the analysis approach by studying the transient state of the simple rate-controller introduced previously and Sec.2.6 compares, in the time-domain, the analysis results with empirical measurements done in simulations.

2.1 ODE-description

Following the description of an artificial chemistry in [16], Meyer in [4] describes a networking protocol as the tuple of “species \mathcal{S} , reactions \mathcal{R} , and an algorithm \mathcal{A} ” that simply relates packet quantities and rates. The algorithm that Meyer proposed, which implements the artificial chemistry (i.e. schedules and execute virtual reactions), is the “real-time next reaction algorithm”, a variant of an algorithm originally proposed in [17] to simulate real chemical reactions. However, as we only

wish to “model chemically” a desired behavior rather than reproduce artificially an exact chemical behavior, we see no reason for abiding to the constraint of the “real-time next reaction algorithm”. In fact, we add an extra degree of freedom at the microscopic level, enabling in this way significant algorithmic-design flexibility (further details are in Sec.3.). On macroscopic level, PCDAs are still characterized by the relationship between packet quantities and rates and this allows us to describe PCDAs’ behaviors through an automatically-derived fluid model: the system of ODEs

$$\dot{\mathbf{c}}(t) = \mathbf{\Psi} \cdot \mathbf{v}(\mathbf{c}(t)) \quad (3)$$

where $\mathbf{\Psi} = [\psi]_{sr}$ is the stoichiometric matrix that defines the reaction network topology ($[\psi]_{sr} = \xi_{sr} - \chi_{sr}$), and where the vector $\mathbf{v} = (v_1, \dots, v_{|\mathcal{R}|})^T$ combines all reaction rates.

2.2 Linearization of the ODEs

Often PC models that characterize PCDAs are non linear: the variation of a molecular concentration c is not proportional to one molecular concentration only (unimolecular reaction) but rather it depends on the multiplication of two (bimolecular reaction) or more molecular concentrations. In order to avoid methods that directly deal with the systems’ non-linearities such as [18, 19], we propose a sensitivity analysis similar to the one proposed in [20] and [21] for studying models in biology. We can rewrite the ODE system (3) by separating external perturbations \mathbf{p} (i.e. input signals of our analysis) from the state variables \mathbf{c} (i.e. concentrations):

$$\dot{\mathbf{c}}(t) = \mathbf{\Psi} \cdot \mathbf{v}(\mathbf{c}(t), \mathbf{p}(t)) \quad (4)$$

For the following perturbation analysis, we focus on fluctuations around steady states only, $\mathbf{c}^*, \mathbf{p}^*$.³ The new vectors $\mathbf{x}(t) = \mathbf{c}(t) - \mathbf{c}^*$ and $\mathbf{u}(t) = \mathbf{p}(t) - \mathbf{p}^*$ denote deviations from nominal state and input signal values, respectively. We reduced the system to a set of linear ODEs,

$$\dot{\mathbf{x}}(t) = \mathbf{\Psi} \cdot \left. \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{x}(t) + \mathbf{\Psi} \cdot \left. \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{u}(t) \quad (5)$$

which provides a good approximation around the steady state for small concentration and input signal variations [15].⁴ The matrixes $\left. \frac{\partial \mathbf{v}(\mathbf{c}^*, \mathbf{p}^*)}{\partial \mathbf{c}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)}$ and $\left. \frac{\partial \mathbf{v}(\mathbf{c}^*, \mathbf{p}^*)}{\partial \mathbf{p}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)}$ can be seen as *elasticity coefficients* and provide a measure for how strongly a reaction in isolation is changed by infinitesimal perturbations either to concentrations c or inputs p .

When PCDAs do not base on non-linear models, the linearization of PC models is obviously not required: eq. (5) turns out to be independent from steady states values $\mathbf{c}^*, \mathbf{p}^*$ and thus its validity is not restricted around steady states only.

2.3 LTI System Representation

As a next step, by following [15], we represent (5) as a Linear Time Invariant (LTI) system:

$$\dot{\mathbf{x}}(t) = \mathbf{A} \cdot \mathbf{x}(t) + \mathbf{B} \cdot \mathbf{u}(t) \quad (6a)$$

$$\mathbf{y}(t) = \mathbf{C} \cdot \mathbf{x}(t) + \mathbf{D} \cdot \mathbf{u}(t) \quad (6b)$$

³Generally, steady states are referred as the molecular concentrations at equilibrium, when the value of concentrations is steady. They can be obtained by solving (4) for $\dot{\mathbf{c}}(t) \equiv \mathbf{0}$.

⁴The validity of (5) around the fixed point $(\mathbf{c}^*, \mathbf{p}^*)$ only is a direct consequence of having linearized the system. However, we remind that we make use of a linearization since this reduces the computational complexity of the analysis and makes possible the use of classical control theory.

where \mathbf{y} represents the output vector of the quantities of our interest. Equation (6b) describes the system behavior which is dependent on the system state and the input whereas (6a) represents the constraint on how the state evolution of the system is subjected to the system state and the input. In a sense the uniqueness of the characterization of a system though (6b) is guaranteed by imposing the satisfaction of (6a). The Jacobian state matrix \mathbf{A} defines how variations of the concentrations affects their future changes; the input matrix \mathbf{B} indicates how external perturbations cause fluctuations of the system state. Both are evaluated at the fixed point as follows:

$$\mathbf{A} = \Psi \cdot \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \Big|_{(\mathbf{c}^*, \mathbf{p}^*)} \quad \mathbf{B} = \Psi \cdot \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \Big|_{(\mathbf{c}^*, \mathbf{p}^*)} \quad (7)$$

The output matrix \mathbf{C} and the feedforward matrix \mathbf{D} must be chosen according to the response we want to analyze, i.e. \mathbf{y} -vector definition. Often, we are interested in two typical outputs: (a) species concentrations, or (b) reaction rates. (a) To analyze the fluctuations of concentrations with respect to perturbations of the input, we define

$$\mathbf{C} = \mathbf{I} \quad \mathbf{D} = \mathbf{0} \quad (8a)$$

yielding $\mathbf{y}(t) = \mathbf{x}(t)$ (Obtained by using (8a) in (6b)). (b) To look at the fluctuations of the reaction rates with respect to a perturbation of the input, we define

$$\mathbf{C} = \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \Big|_{(\mathbf{c}^*, \mathbf{p}^*)} \quad \mathbf{D} = \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \Big|_{(\mathbf{c}^*, \mathbf{p}^*)} \quad (8b)$$

yielding $\mathbf{y}(t) = \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \Big|_{(\mathbf{c}^*, \mathbf{p}^*)} \mathbf{x}(t) + \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \Big|_{(\mathbf{c}^*, \mathbf{p}^*)} \mathbf{u}(t)$, which represents the deviation from the nominal rates.

Again, when PCDAs do not base on non-linear models, the description of PCDA as a LTI systems does not depend on steady states values (i.e. the mathematical definitions given in this section do not depend on $\mathbf{c}^*, \mathbf{p}^*$) and thus its validity is not restricted around steady states only.

2.4 Signal Analysis in the Frequency Domain

We can now use the Laplace transform $\mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t) dt$ to transpose the LTI system (6) from the time- to the frequency-domain:

$$s \cdot \mathbf{x}(s) - \mathbf{x}_0 = \mathbf{A} \cdot \mathbf{x}(s) + \mathbf{B} \cdot \mathbf{u}(s) \quad (9a)$$

$$\mathbf{y}(s) = \mathbf{C} \cdot \mathbf{x}(s) + \mathbf{D} \cdot \mathbf{u}(s) \quad (9b)$$

By assuming that the initial conditions are set to steady state (i.e., there are no initial perturbation: $\mathbf{x}_0 = \mathbf{0}$), the *Transfer Function* (TF) of the LTI system results in

$$\mathbf{H}(s) = \frac{\mathbf{y}(s)}{\mathbf{u}(s)} = \mathbf{C} (s \cdot \mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D} \quad (10)$$

The TF matrix \mathbf{H} describes the transient behavior of all steady-state output deviations \mathbf{y} with respect to perturbations on all inputs \mathbf{u} .

2.5 Analyzing the simple rate-controller

This section shows how to apply the introduced methodology by analyzing the transients of the simple rate-controller whose PC model is depicted in fig.1.

Like any other PCDA, the dynamics of the algorithm in fig.1 are described by a system of coupled ODEs that can be easily extracted from the PC model:

$$\dot{c}_S = v_{\text{src}} - k_1 c_S c_E \quad (11a)$$

$$\dot{c}_{\text{ES}} = k_1 c_S c_E - k_2 c_{\text{ES}} \quad (11b)$$

$$\dot{c}_E = k_2 c_{\text{ES}} - k_1 c_S c_E \quad (11c)$$

Let us focus for example on (11a): because of our design decision to use the LoMA as a macroscopic reaction law, the reaction rates are proportional to reactant concentrations and thus (i) outflows from S-species (in this case one outflow only) are proportional (through the reaction coefficient) to S-species' concentration c_S times the concentration of all other involved reactants (i.e. c_E only), (ii) outflows have negative sign, and (iii) inflows (i.e. v_{src}) have positive sign. The set of equations in (11) have a direct analogy with the inflow-outflow balance that fluid models exhibit.

The next step is to find the steady states. We start by solving (11) with respect to species concentrations, when the left-hand side is equal to 0:

$$c_S^* = \frac{v_{\text{src}}}{k_2 e_0 - v_{\text{src}}} \frac{k_2}{k_1} \quad (12)$$

$$c_E^* = \frac{k_2 e_0 - v_{\text{src}}}{k_2}$$

$$c_{\text{ES}}^* = \frac{v_{\text{src}}}{k_2}$$

Since there exists a conserved moiety E-ES (i.e. a closed loop where the sum of the concentrations of involved species is constant: $c_{\text{ES}} + c_E = \text{const.} = e_0$), if we impose that molecular concentrations must be positive quantities, we limit the validity of (12) to the following region:

$$v_{\text{src}} < k_2 e_0 \quad (13)$$

The obtained result suggests that, if the generation rate of the host is sufficiently low, see (13), the host does not experience any rate-limitation and can transmit with a rate v_{tx} that equals its generation rate: $v_{\text{tx}} = v_{\text{OUT}} = k_2 c_{\text{ES}}^* = v_{\text{src}}$. On the contrary, when reaction $r_1 : S + E \xrightarrow{k_1} \text{ES}$ does not drain S-species with high-enough rates ((13) is not respected), S-species' concentration increases without limit and the mass conserved in the moiety lies mainly in ES-species only, i.e. $c_{\text{ES}} = e_0, c_E = 0$.

The next step is the definition of the rate vector as $\mathbf{v} = [k_1 c_S c_E \quad k_2 c_{\text{ES}} \quad v_{\text{src}}]$ and of the stoichiometric matrix (reaction network topology) as

$$\Psi = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

whose rows represent the species $\mathbf{s} = [S \text{ ES } E]$ and columns represent the reactions $\mathbf{r} = [r_1 \quad r_2 \quad r_{\text{src}}]$. For example, the first column expresses the fact that reaction r_1 consumes an S-molecule and an

E-molecule and produces an ES-molecule. Each structural conservation, e.g. the moiety E-ES, corresponds to linearly dependent rows in the stoichiometric matrix. As suggested in [22], we avoid redundant terms and proceed taking into account the reduced stoichiometric matrix Ψ_R that can be obtained removing the row related to E-species from Ψ , i.e. the last row:

$$\Psi_R = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

As it is suggested in [23], the relationship between Ψ and Ψ_R is formalized by the link matrix \mathbf{L} ($\Psi = \Psi_R \mathbf{L}$), which makes explicit how dependent species' concentrations, $\mathbf{c}_d = [c_E]$, can be derived from independent species' concentrations, $\mathbf{c}_i = [c_S \ c_{ES}]$:

$$\mathbf{L} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix}$$

The link matrix is composed by the identity matrix whose size equals the number of independent species ($\text{rank}(\Psi)$), and an additional row vector describing the relationship between dependent and independent species. In this case there is one moiety only and the additional row vector expresses that the dependent species E can be derived by subtracting ES-species' concentration c_{ES} from the initial amount that characterizes the moiety (e_0). Namely, species can be represented in terms of independent species only as $\mathbf{c} = \mathbf{L}\mathbf{c}_i + \mathbf{T}$ where \mathbf{T} is an appropriate constant vector – i.e. referring

to our rate-controller example we have $[c_S \ c_{ES} \ c_E] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} c_S \\ c_{ES} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ e_0 \end{bmatrix}$

In order to represent the PC module shown in fig.1 as a LTI system, we must evaluate the behavior of the model around steady states (nominal states) and input signal values. Namely, we calculate the jacobians of the rate vector with respect to state vector \mathbf{c} and perturbation vector \mathbf{p} respectively. The state vector \mathbf{c} includes independent species' concentrations \mathbf{c}_i and dependent species' concentrations \mathbf{c}_d , and, assuming we want to study the PCD Rate controller's response to variations on packet generation rate, we fix $\mathbf{p} = [v_{\text{src}}]$:

$$\frac{\partial \mathbf{v}}{\partial \mathbf{c}} = \begin{bmatrix} k_1 c_E & 0 & k_1 c_S \\ 0 & k_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\frac{\partial \mathbf{v}}{\partial \mathbf{p}} = [0 \ 0 \ 1]^T$$

We know that the state matrix \mathbf{A} defines how a perturbation on the concentrations affects their future changes and the input matrix \mathbf{B} indicates how external perturbations cause fluctuations of the system state. By using the reduced form of the stoichiometric matrix Ψ_R , the definitions of \mathbf{A}

and \mathbf{B} differs from the one in (7) as follows:⁵

$$\mathbf{A} = \mathbf{\Psi}_R \cdot \left. \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{L}$$

$$\mathbf{B} = \mathbf{\Psi}_R \cdot \left. \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)}$$

The output matrix \mathbf{C} and the feedforward matrix \mathbf{D} must be chosen according to the response we want to analyze. Since we look at the fluctuations of the reaction rates with respect to a perturbation of the input (i.e. $\mathbf{y}(t) = \left. \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{L} \cdot \mathbf{x}(t) + \left. \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{u}(t)$ – the output vector reports variations on reaction rates), we define

$$\mathbf{C} = \left. \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{L} \quad \mathbf{D} = \left. \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)}$$

Again, the output matrix \mathbf{C} includes the relationship between dependent and independent species, which is formalized by the link matrix \mathbf{L} .

We can now calculate, by using (10), the TF $\mathbf{H}(s)$ in the Laplace domain, which describes the transient behavior of output deviations \mathbf{y} from the nominal state with respect to perturbations on the chosen input $\mathbf{u} = [v_{\text{src}}]$. Since the triggering rate for packet transmissions is the rate of r_2 -reaction $v_{\text{OUT}} = v_2 (= k_2 c_{\text{ES}})$, we can focus on the last row of the TF matrix, the row vector that describes how v_2 deviates from the steady-state:

$$H(s) = \begin{cases} \frac{k_1(k_2 e_0 - v_{\text{src}})}{s^2 + s \left(k_2 + k_1 \left(\frac{k_2 e_0 - v_{\text{src}}}{k_2} + \frac{k_2 v_{\text{src}}}{k_1(k_2 e_0 - v_{\text{src}})} \right) \right) + k_1(k_2 e_0 - v_{\text{src}})} & (v_{\text{src}} < k_2 e_0 = R) \\ 0 & (v_{\text{src}} \geq k_2 e_0 = R, c_S \rightarrow \infty) \end{cases} \quad (14)$$

In this seemingly complex result, we can observe several interesting components: If the generation rate respects the prefixed threshold R , (i) the rate-controller dampens all sharp variations on the generation rate in a way to reduce their effect on the transmission rate. This is represented by the low pass filtering-behavior described by the top equation in (14). (ii) The cut-off frequency of the implemented low pass filter is adjustable through k_2 coefficient. Indeed, one of the two TF's poles matches k_2 coefficient. (iii) The module lets the host adapt and send with a rate that matches the new generation rate. This is represented by the fact that the low-pass filter has a unity Continuous Wave (CW) gain. (iv) If the generation rate exceeds the threshold R and thus the host is already transmitting at the maximum allowed rate, the module prevents that variations on the generation rate affect the host's transmission rate. Indeed, in this case ($v_{\text{src}} \geq k_2 e_0 = R$), the S-concentration increases without limits and no variations on v_{src} have effect on v_{tx} : the host has a large amount of packets to transmit (infinite number of molecules), which sends at the maximum allowed rate $k_2 e_0 = R$. (v) The exact shape of TFs depends on the input. Indeed, we observe v_{src} -dependency in (14). (vi) The adaptation of the algorithm to bounded variations of the packet generation rate will always produce a limited variation on the transmission rate since the TF always exhibits negative real-part poles – Bounded-Input Bounded-Output (BIBO) stability of the linearized model.

⁵This definition is a direct consequence of describing the state-evolution as

$$\dot{\mathbf{x}}(t) = \mathbf{\Psi}_R \cdot \left. \frac{\partial \mathbf{v}}{\partial \mathbf{c}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{L} \cdot \mathbf{x}(t) + \mathbf{\Psi}_R \cdot \left. \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{u}(t)$$

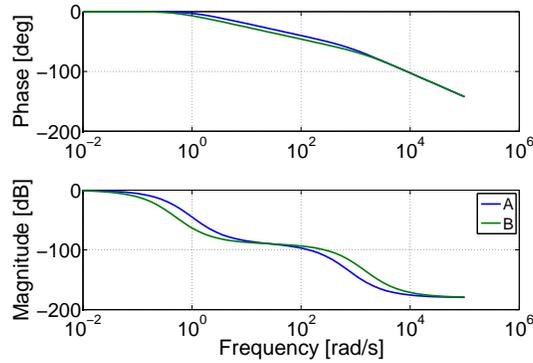


Figure 2: Transfer Function (TF): if the rate-limit mechanism is not functioning, the PCD controller depicted in fig.1 behaves as a Low-Pass filter with respect to variations on the generation rate. Model’s settings were (a) $e_0 = 1000$, $k_1 = 1$, and $k_2 = 1$ and (b) $e_0 = 500$, $k_1 = 1$, and $k_2 = 0.5$.

All other rows of the TF matrix represent the effect of variations of the generation rate v_{src} on the rate of other reactions composing the model.

2.6 On time-domain dynamics of the simple rate-controller

The TFs allow us to predict the time-domain behavior of the algorithm in response to arbitrary perturbations on the inputs (i.e. packet generation rate). In this section, we look at (i) the step response, which represents a sort of stability (convergence) measurement of the algorithm, and (ii) the impulse response, which instead relates to the sensibility of the algorithm. Namely, the step response reveals how the transmission rate changes in response to a sudden increase of packet generation rates. Whereas, the impulse response reveals the variation of the transmission rate in response to the injection of a certain amount of packets (also one only) into the queue. Additionally, to evaluate the quality of the analysis results, we compare the analytically predicted response of the simple rate-controller to measurements obtained in simulations. (We have implemented a host as depicted in fig.1 in OMNeT++ [24].)

Figure 3(a) shows the plots of (A) the normalized average over 0.5s-period of the generation rate when the generation process is modeled as of Poisson, (B) the normalized instantaneous transmission rate, (C) the predicted transmission rate as a step response to the TF in (14), and (D) the normalized instantaneous transmission rate obtained when the generation process is deterministic instead of Poisson. In the simulation, the step-perturbation on v_{src} was a sudden increase at time $t = 10$ s of the host’s generation rate from 250 pkt/s to 500 pkt/s.

From this comparison we can observe that the transient analysis predicts the actual behavior of a PC algorithm. There are some deviations between B and C curves due to the fact that the generation rate is highly stochastic. By considering the same generation rate but modeling this process as deterministic, the deviations between the analytical prediction and a realization of the actual algorithm’s behavior are undetectable (C and D curves almost overlap). In case of the generation process being of Poisson, the oscillations in the transmission rate and thus the deviations between real behavior and predicted one can be smoothed by reducing the cutoff frequency of the implemented low pass filter: in fig.3(b) the normalized instantaneous transmission rate (B-curve) is closer to the normalized predicted rate (C-curve) compared to what shown in fig.3(a).

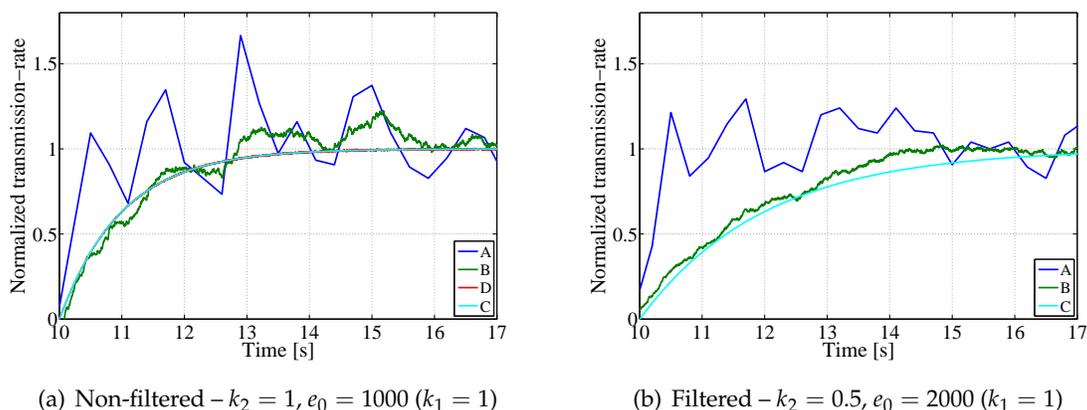


Figure 3: Comparison between the transient analysis results and the normalized transmission rate measured in a simulation run: The initial generation rate was $v_{src} = 250\text{pkt/s}$ and the applied perturbations brought it to 500pkt/s at time $t = 10\text{s}$. A-curve is the normalized 0.5s-averaged generation rate, B-curve is the normalized instantaneous transmission rate, C-curve is the predicted transmission rate, and D-curve is the normalized instantaneous transmission rate obtained when the generation process is deterministic instead of Poisson.

Figure 4 shows the impulse response: in a host equipped with the PCD controller depicted in fig.1 with parameters $e_0 = 1000$, $k_1 = 1$, and $k_2 = 1$ and a generation rate of 500pkt/s , if 50 packets are injected into the queue, after a transient increase, the host's transmission rate stabilizes back to the original value.

3 On PCDAs' microkinetics

The importance of microkinetics in studying and designing PCDAs might be hidden: Even though we propose to design and study algorithms (i.e. PCDAs) by observing from the macroscopic level (e.g. average algorithm trajectory and LoMA), events in the algorithms are dictated by the microkinetics of the underlying model. For instance, we can characterize macroscopically the dynamics of the simple rate-controller in fig.1 with the Michaelis-Menten equation in (2). However, each transmission executed by the rate-controller module is triggered by the single execution of r_2 -reaction (namely to the single production of P-molecules).

Unlike algorithms in [4, 8, 17], our approach does not try to mimic actual chemical reaction kinetics on microscopic level but rather it forces algorithms to have law-based behavior on macroscopic level (e.g. LoMA-based). The difference between CNPs, which stick to the original artificial chemistry, and PCDAs lies in the way how the inter-reaction time (i.e. the time between two consecutive reactions) is modeled: In CNPs, this random time is exponentially distributed with a mean that is inversely proportional to the concentration c of all reactant molecules (matching the LoMA principle). Whereas, in PCDAs the inter-reaction time is not strictly related to real-chemical phenomena and thus can be modeled with different probability distributions, e.g. the reaction rate can be a normally distributed variable.

The exponential distribution may limit the performance of algorithms: (i) The variance is fixed to the square of the mean value which is dictated by the applied reaction law, e.g. LoMA. This

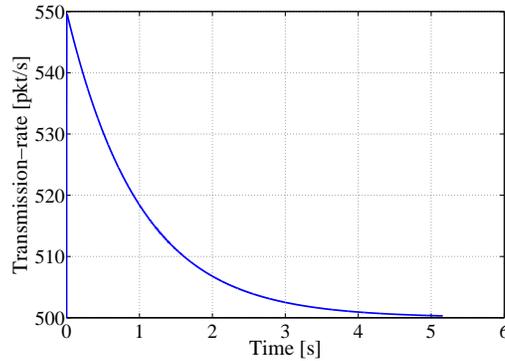


Figure 4: Impulse response: in a host equipped with PCD controller depicted in fig.1 with parameters $e_0 = 1000$, $k_1 = 1$, and $k_2 = 1$ and a generation rate of 500 pkt/s, the effect on transmission rate of the injection of 50 packets in its queue.

leads to have a poor predictability of models that work with low concentrations. That is, the more packets an algorithm deals with, the more closely its actual behavior will approach the predicted average. (ii) Skewness, ex.kurtosis, and mode, suggest that the inter-reaction time often (with very high probability) assumes very low values and sporadically (with very low probability) assumes high values and on average only, it matches the value that the law (e.g. LoMA) predicts. In the simple rate-controller in fig.1, by using an exponentially distributed inter-reaction time, the packet transmissions are often instantaneous as in a work-conserving scheme (as soon as packets are generated, they are sent) and rarely, transmissions are characterized by very long delays. Thus, the transmission rate, although it matches the LoMA-predicted value on average, comes out to be bursty. This phenomenon is shown in fig.5 that reports the rates of the PC rate-controller in case of exponentially distributed inter-reaction times. Specifically in fig.5(a), the instantaneous sending rate presents very high peaks and only on average, fig.5(b), it is more constant around the mean generation rate value. Also when the rate limiting mechanism is acting, by using an exponentially distributed inter-reaction time, the instantaneous transmission rate is very spiky (fig.6(a)) and only on average it fluctuates around the prefixed threshold (fig.6(b)).

PCDAs could be characterized by deterministic inter-reaction time. However, it is known that relaxing the determinism of algorithms can (i) increase their robustness – e.g. gossip-style protocols, (ii) can help to break symmetry – e.g. leader election in anonymous networks, and (iii) can help in resolving access conflict – e.g. CSMA family. For this reason, we may want to keep some sort of controllable randomness in algorithms’ dynamics.

A candidate distribution may be the normal distribution, which (i) is very tractable analytically, and (ii) arises as the outcome of the central limit theorem (the sum of a large number of random variables is distributed approximately normally). Differently from what we have previously argued about exponentially distributed variables, the probability of seeing a normally-distributed variable that is far (i.e. more than a few standard deviations) from its mean decreases extremely rapidly. In addition, the Gaussian distribution is not fully described by the mean only but also by the variance, which, from a designing point of view, represents the parameter to trim the level of algorithm randomness. The fact that Gaussian distributed variables can assume negative values (not reasonable for modeling time) can be solved through the use of the truncated, re-normalized

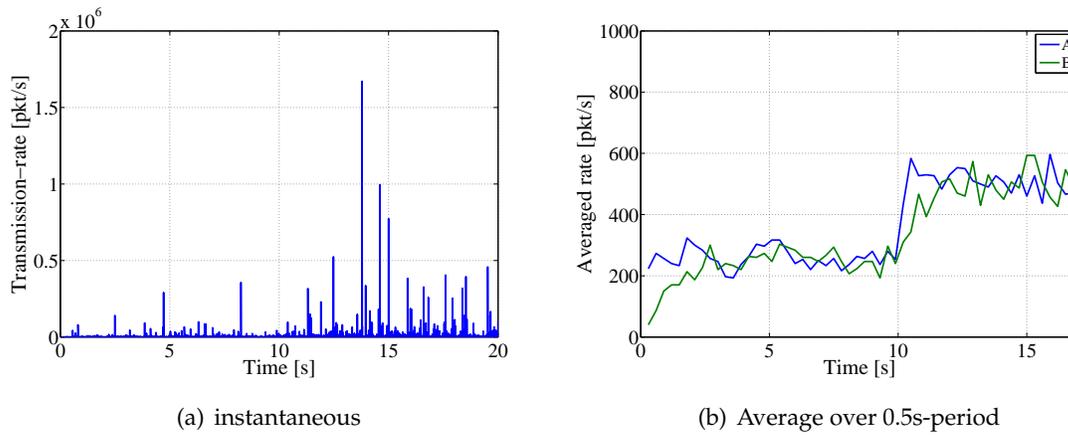


Figure 5: Host's transmission rate when the inter-reaction time is exponentially distributed: Instantaneous transmission rate in fig.5(a), 0.5s-average of the generation rate (A-curve) and of the transmission rate (B-curve) in fig.5(b). Model settings were $e_0 = 1000$, $k_1 = 1$, and $k_2 = 1$ and at time 10s, the generation rate changed from 250 pkt/s to 500 pkt/s.

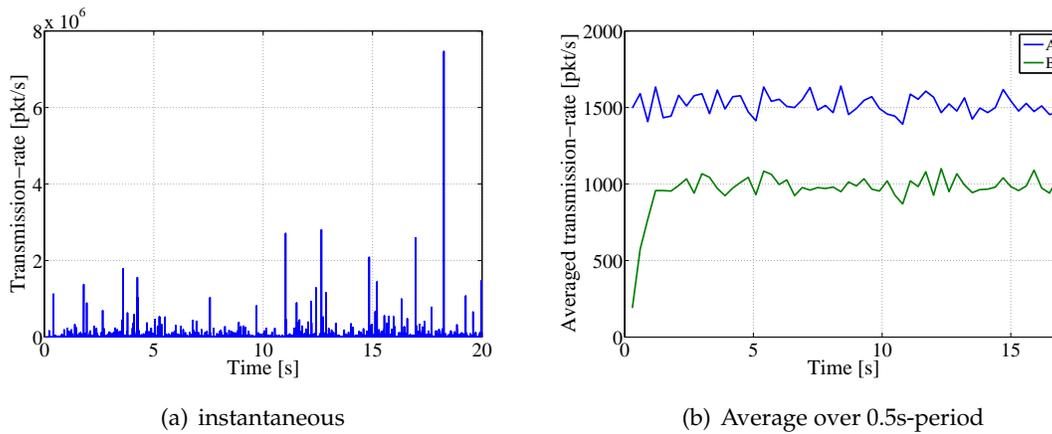


Figure 6: Host's transmission rate when the inter-reaction time is exponentially distributed: Instantaneous transmission rate in fig.6(a), 0.5s-average of the generation rate (A-curve) and of the transmission rate (B-curve) in fig.6(b). Model settings were $e_0 = 1000$, $k_1 = 1$, and $k_2 = 1$ and at time 10s, the generation rate was 1500 pkt/s.

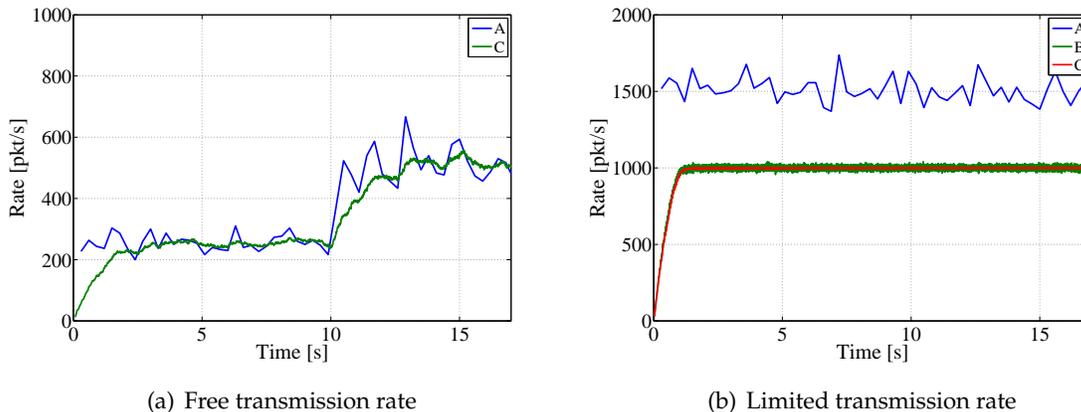


Figure 7: Host’s transmission rate when the inter-reaction time is Gaussian distributed and the related standard deviation is adjustable. A-curve represents the 0.5s-average of the generation rate; B-curve represents the instantaneous transmission rate when the standard deviation of is set to $0.01 \cdot \text{mean}$; C-curve represents the instantaneous transmission rate when the standard deviation of is set to $0.001 \cdot \text{mean}$. Model settings were $e_0 = 1000$, $k_1 = 1$, and $k_2 = 1$ and, for fig.7(a), at time 10s the generation rate changed from 250 pkt/s to 500 pkt/s, whereas for fig.7(b) the generation rate was 1500 pkt/s on average.

version, and in any case by taking appropriate values of mean and variance, the probability to have negative times is negligible. Figure 7 shows the transmission rate of the PCD rate-controller in case of Gaussian-distributed inter-reaction rates: the instantaneous transmission rate (C-curve) can be smooth and very close to the average generation rate value (A-curve). Additionally, the level of the algorithmic randomness can be trimmed through the variance (B-curve *versus* C-curve).⁶

The use of an exponential distribution may be required for specific case when designers need algorithms whose dynamics must be related to a Poisson process, e.g. in order to *exactly describe* the algorithm’s behavior with queuing and Markov theory on microscopic level. Note that the extraordinary benefit of PCDA that are characterized by an exponential distribution (and thus of CNPs proposed in [4]) lies in the fact that we do not model algorithms’ behavior as a tractable stochastic process, i.e. Poisson process, but rather the tractable process itself constitutes the desired algorithms’ behavior.

3.1 On the microscopic analysis of PCDA’s dynamics

In this section, we explore the possibility to microscopically describe PCDA: We provide an analytical point of view which is closer to the traditional approach used in queuing theory, we highlight the benefits of using chemically driven algorithms, and we make clear the limitation of this approach. Finally, we formalize how the average, macroscopic behavior of PCDA is related to their microscopic, stochastic trajectory.

The macroscopic analysis of PCDA that we have explained in previous sections does neither describe networking algorithms in terms of state probability evolutions nor in terms of stationary probability distributions. However, with the design choice of having exponentially distributed mi-

⁶The stochasticity characterizing the algorithm appears as fluctuation on the transmission rate around the average value with a controllable variance. The stochasticity can improve the algorithm’s performance.

crokinetics, the state evolution of PCDAs can be directly associated with a time-continuous, discrete-space Markov jump process. From this point of view, the state of a PCDA represents a stochastic process which is described by the vector of random variables $\mathbf{N}(t) = [N_{s_1}(t), N_{s_2}(t), \dots, N_{|S|}(t)]$ whose elements are the concentrations of the species at a certain time t (note that each species adds a dimension to the problem). At this microscopic level, the state probability $P(\mathbf{n}, t) = P[\mathbf{N}(t) = \mathbf{n} | \mathbf{N}(t_0) = \mathbf{n}_0]$ represents the probability that at time t , the process is in the state $\mathbf{n} \in \mathcal{X}$ (where \mathcal{X} is the state space). We can further describe the state transition of a PCDA through the Chemical Master Equation (CME):⁷

$$\frac{dP(\mathbf{n}, t)}{dt} = \sum_{r \in \mathcal{R}} \left(a_r(\mathbf{n} - \psi_r) P(\mathbf{n} - \psi_r, t) - a_r(\mathbf{n}) P(\mathbf{n}, t) \right) \quad (15)$$

Using the description in [27], the CME states that the change of the probability of being in the state \mathbf{n} at time t is equal to the probability of arriving at \mathbf{n} by executing the reaction $r \in \mathcal{R}$, subtracted by the probability of leaving \mathbf{n} by executing r . With such a prospective, a reaction reflects the transition from a state to another and the transition probability does relate to the reaction propensity.

In case the system state-space is finite (and by making lighter the notation with $p_i(t) = P(n_i, t) = P[N(t) = n_i | N(t_0) = n_0]$), the CME can be simplified to:

$$\frac{d\mathbf{p}(t)}{dt} = \mathbf{p}(t)\mathbf{Q} \quad (16)$$

where \mathbf{Q} is the transition rate matrix. \mathbf{Q} is a stochastic matrix (row-sums are null) which is square with size that matches the size of the state-space, $|\mathcal{X}|$, and whose $q_{i,j}$ -th off-diagonal element is obtained by summing up the propensity functions evaluated at state n_i of those reactions that lead from state n_i to state n_j , whereas the diagonal element $q_{i,i}$ is obtained by imposing the stochasticity condition on the related row, i.e. $\sum_{j \neq i} q_{i,j} + q_{i,i} = 0$.

From the simplified version of the CME in (16), by integration we can obtain the *state probability evolution* $\mathbf{p}(t) = \mathbf{p}(0)e^{\mathbf{Q}t}$, and by imposing the condition of no state-transitions (i.e. left hand-side of (16) equalling null), we obtain the state probability distribution of the system at equilibrium, i.e. the *stationary probability distribution*.

The microscopic description of a PCDA does not base on simplifying assumptions that make the model analyzable (as it is in many studies of queueing systems, where the memory-less assumption and thus the speculation of exponential distributed times and of Poisson processes allow the use of Markov's theory). In case of PCDAs that have exponentially distributed inter-reaction times (a mere design decision rather than a simplifying assumption), the microscopic description of the algorithm's behavior in terms of state probability evolutions and stationary probability distributions comes for free and actually formalizes the intrinsic, wanted stochasticity affecting the algorithm's dynamics.

The CME cannot be solved for many chemical reactions [28] and definitely it suffers from the curse of dimensionality: each species adds one dimension to the problem leading to an exponential growth of the computational complexity required by the analysis. Usually, we need reaction models with many species and that do include complex interactions (even non-linear ones) in order to

⁷This claim derives directly from the fact that PCDAs are driven in this case exactly by an artificial chemistry. We can thus make use of the theory and all already-proven theorems and laws that characterize chemical reaction systems in terms of stochastic processes and probability distributions [25, 26].

design algorithms to compute interesting networking tasks. For this reason, we discard the analysis at the microscopic level and do focus on the deterministic description of the algorithms' trajectory.

In case we make use of exponentially distributed inter-reaction times, we can base on chemistry theory and describe the average (deterministic) trajectory of a PCDA basing on the expectation operator: $c_s(t) = E(N_s(t))$. By deriving from the CME in (16), according to [29] we can write the variation of the concentration of a species s as

$$\frac{dc_s(t)}{dt} = \sum_{r \in \mathcal{R}} \psi_{s,r} \cdot E(a_r(\mathbf{N}(t))) \quad (17)$$

For unimolecular reactions, the propensity function is linear and thus we can derive directly the ODE form

$$\frac{dc_s(t)}{dt} = \sum_{r \in \mathcal{R}} \psi_{s,r} \cdot a_r(\mathbf{c}(t)) \quad (18)$$

For multi-molecular reaction we can still use (18) but in this case the equation represents an approximation only of the average trajectory of the molecular species where the higher order moments are ignored (for further details please refer to [29]).

We can similarly argue about the average trajectory of PCDAs in case we use non-exponentially distributed inter-reaction times. This can be easily understood if we look at how the underlying reaction model implements what is described by (17): the variation of a molecular concentration c_s is dictated by the execution (one per time) of those reactions, e.g. $r \in \mathcal{R}$, that involve s -species (namely, a production of one or more s -molecules in case s -species is a product or a consumption of one or more s -molecules in case it is a reactant). This relationship is formalized by the stoichiometric coefficient $\psi_{s,r}$ that appears in (17). The time at which the next r -reaction is executed is calculated as an independent random variable with a mean inversely proportional to the propensity function a_r , which according to LoMA is defined as

$$a_r = k_r \cdot \prod_{i=1}^{|\mathcal{S}|} c_i^{\alpha_i}$$

where c_i is the concentration of the i -th species, k_r is a "speed" coefficient for reaction r , $|\mathcal{S}|$ is the set of all molecule species, and α_i is the number of molecules from species i that are consumed by reaction r .

Simulations confirm, in most cases, the usability of the ODE approximation to analyze and predict the behavior of PCDAs that use different distributions of inter-reaction times (also deterministic times). However, we may experience a problem derived from the use of a non-exponential distribution time and due to how PCDAs actually implement an ODE model (discrete space – concentrations can assume values belonging to natural numbers only). To explain the problem we look at the simple example where we have a single species A whose molecules are injected with reaction $r_1 : \emptyset \xrightarrow{k_1} A$ and consumed by two reactions $r_2 : A \xrightarrow{k_2} \emptyset$ and $r_3 : A \xrightarrow{k_3} \emptyset$. Assuming $k_2 \neq k_3$ and according to the ODE at the macroscopic level, we have different outflow rates v_2 and v_3 : $c_A^* = k_1/(k_2 + k_3)$ and thus $v_2 = (k_2 k_1)/(k_2 + k_3)$ and $v_3 = (k_3 k_1)/(k_2 + k_3)$. Assuming that $c_A^{(0)} = 0$ initially, the egress reactions r_2 and r_3 are inactive. The ingress reaction r_1 is scheduled for time $t = 1/k_1$ s. At that time, one A-molecule is generated and r_1 is rescheduled for $1/k_1$ s later. At the same time, both egress reactions become active and are scheduled according to the

LoMA. Let us further assume that $k_1 < k_3$ and $k_3 > k_2$ and thus that r_1 -reaction is not fast enough to fill the A-species and r_3 -reaction occurs earlier (with a deterministic inter-reaction time or with a “non-enough-random” Gaussian inter-reaction time) than r_2 -reaction. At time $1/k_1 + 1/k_3$ s, the only A-molecule is consumed by r_3 -reaction and both egress reactions become inert again, because there is no molecule left. At each $1/(n \cdot k_1)$ instant (with $n \in \mathcal{N}$) this cyclic trend starts again and an A-molecule is produced with r_1 -reaction and then later consumed by r_3 -reaction. In other words, for low concentrations of A-species, by using a deterministic inter-reaction time (or a “non-enough-random” Gaussian inter-reaction time), the slower reaction never occurs because the faster reaction deterministically (or with non-enough stochasticity) fires always before the slower one. The aforementioned phenomenon affects models that include (i) more competing, linear, egress reactions, which have different reaction coefficients, that drain a single species with a sufficiently low concentration, (ii) more competing, non-linear, egress reactions that, even if the reaction coefficients are the same, do depend on reactant species with different concentrations and one of these concentrations is sufficiently low.

3.2 Analyzing the algorithm’s stochasticity on macroscopic level

So far, we have avoided to treat statistically the time-evolution of the algorithms’ behavior on microscopic level, thus gaining in analysis’ simplicity. However we show here that, by following the signal processing approach, we can describe the randomness that actually characterizes PCDA’s dynamics on macroscopic level.

Approaches have been proposed to describe the dynamics of chemical reaction systems by separating them to a systemic, deterministic trajectory and a stochastic noise. The first describes the average behavior of the chemical system over the possible realizations of the stochastic process that represents the system state-evolution. We have already seen that this deterministic trajectory derives from the macroscopic description of the system through the ODE model. The second catches the stochastic fluctuations observed in the algorithm’s real-behavior around the average trajectory. The derivation of this noise refers to a mesoscopic description of the chemical model: Chemical Langevin Equation (CLE) [30], Linear Noise Approximation (LNA) [31], and Two Moment Approximation (2MA) [32]. All these three descriptions take for granted that inter-reaction times are independent, exponentially-distributed variables and their accuracy relates to the quantity of molecules. All these three descriptions approximate the stochastic fluctuations to a *Gaussian* noise.

The description of the dynamics of PCDA’s in these terms would preserve our ability to look at algorithms’ dynamics on macroscopic level but also to further take into account the algorithms’ stochasticity, i.e. to provide a probabilistic description of the fluctuations observed in the algorithm’s real-behavior around the predicted, average trajectory obtained by the transient analysis. However, CLE, LNA, and 2MA are not trivial to be extracted and solved and they might present problems of accuracy especially for high-order reaction systems and for systems counting a low amount of molecules [33].

The use of a Gaussian distribution to model the algorithm’s microkinetics, which interdicts the use of Markov theory for studying the model’s microscopic kinetics, makes the characterization of the stochasticity in terms of noise straightforward in some special cases: For reactions where reactants’ concentrations remain constant, we can easily extend the ODE-based prediction of the deterministic, average trajectory of rates with an analysis of a superimposed white Gaussian noise. Indeed, since we let the inter-reaction rate values be *independent Gaussian distributed variables* with a mean dictated by the reaction law (e.g. LoMA) and a preset variance, we can split the instantaneous-

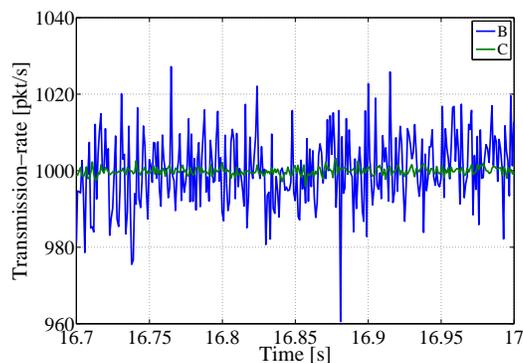


Figure 8: 0.3s-zoom on instantaneous transmission rates in fig.7(b), for different values of standard deviation.

transmission-rate signal into (i) the deterministic signal that is steady at the value announced by the reaction law (e.g. LoMA) and (ii) a *superimposed white Gaussian noise* with null mean value and set variance. We can further suppose that at steady states, the instantaneous transmission rate of the simple rate-controller (B-curve in fig.8) can again be split in a deterministic signal and a white Gaussian noise.

Suppose a generation rate that on average is constant, e.g. 1000 pkt/s, or that is sufficiently high to not respect the prefixed rate limit, e.g. 1500 pkt/s when the threshold is fixed to 1000 pkt/s. In both these situations, if the PC-model in the rate-controller makes use of Gaussian distributed inter-reaction rates, the instantaneous sending rate will fluctuate by a controllable amount around the constant value, e.g. of 1000 pkt/s, similarly to what is shown in fig.8. Fluctuations around the steady value of 1000 pkt/s can be approximated to a time-discrete wide-sense-stationary random process, where first and second moments do not vary with respect to time. Additionally, we assume the ergodicity of the process: for infinite realizations, the time and space (ensemble) averages coincide. Under these assumptions, we can estimate the Probability Density Function (PDF) and the autocorrelation of the measured noise (fluctuations) affecting the transmission rate of the rate-control algorithm during a simulation run: Figure 9(a) shows the estimated probability density function of such an additive noise which, as we have supposed, resulted to be *Gaussian distributed* with estimated mean $\mu = -1.16 \cdot 10^{-11}$ (approximately null) and variance $\sigma^2 = 1.0038$ (in the PC model, we set unitary variance of the inter-reaction time). Figure 9(b) shows the estimated autocorrelation of the noise sequence. These estimations refer to a “noise sequence” of 44996 samples when the instantaneous sending rate fluctuated around 1000 pkt/s.

4 Discussion

We propose a novel approach for designing and analyzing algorithms that schedule system events in a non-work-conserving manner. Our work extends Chemical Networking Protocols (CNPs) design and methodology that was proposed in [4]: At the microscopic level, Pseudo-Chemically Driven Algorithms (PCDAs) can deviate from actual chemical reaction laws and probability distributions of the inter-reaction time. The benefit of using such a novel algorithmic design paradigm is threefold: (A) The algorithm’s dynamics as well as the global system dynamics are analyzable and predictable. (B) The algorithm’s dynamics can be shaped as desired in the design stage. (C) The analysis can

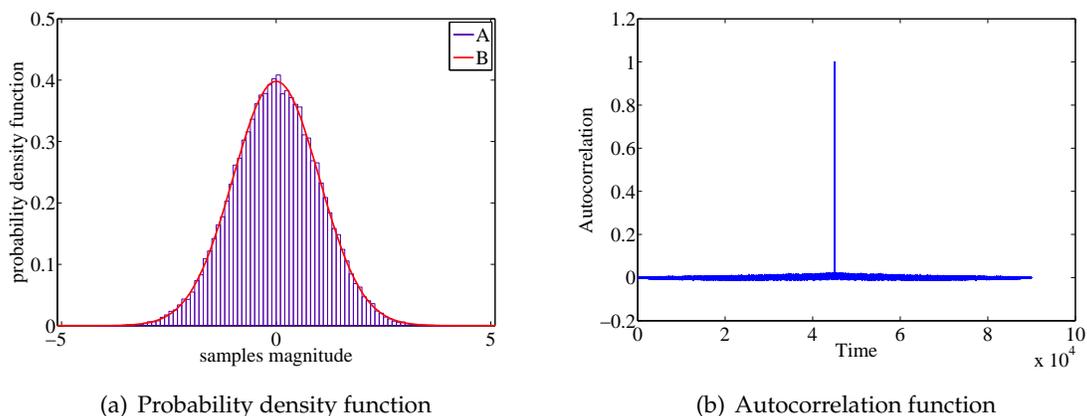


Figure 9: Noise characterization. In fig.9(a) the estimation of the density probability function (A-curve) and the fitted Gaussian function with $\mu = -1.16 \cdot 10^{-11}$ and $\sigma^2 = 1.0038$ (B-curve).

base on well-established control theory tools and this can help to transpose the ground knowledge of system engineering into the algorithms' design stage.

4.1 Analyzability

The analysis of PCDAs does not regard in detail the model kinetics on microscopic level but rather evaluates the emerged model behavior on macroscopic level. Moreover, the enabled flexibility that allows the use of new reaction laws that do not model the time with an exponential distribution interdicts the use of Markov theory for studying the model's microscopic kinetics. However, we have pointed out that this detailed analysis is cumbersome or often impossible, and therefore its practical use is limited to the study of a system's behavior rather than in designing a networking algorithm.

The use of other distributions still guarantees the algorithm's analyzability and may improve the predictability of the algorithm's macroscopic behavior. For example in most cases, modeling the inter-reaction rate with a normal distribution enables us to predict the trajectory of algorithms based on ODEs. The phenomenon described previously in section 3.1, which affects models that use non-exponential distributed times and which includes competing egress reactions characterized by different rates, can be bypassed by following different directions: (i) we can simply make a design decision and let the critical reactions always be characterized by exponentially distributed times; (ii) we can modify the implementation of the core-scheduler of PCDAs by monitoring the concentrations on which the critical reactions depend and, when concentrations get low, by forcing the inter-reaction time of the concerned critical reaction to be exponentially distributed.

The analysis of the PCDAs is divided into steady-state analysis and transient-state analysis. The first gives us already an understanding of the algorithm's behavior when steady conditions are reached, e.g. in case of the PC rate-controller, for constant loads and after a sufficient time, the transmission rate will stabilize at a certain predictable amount. The second gives us insights on the general dynamics of the algorithm, e.g. we can exactly predict the averaged transmission rate of a host equipped with the PC rate-controller and predict its response to arbitrary perturbations on the generation rate.

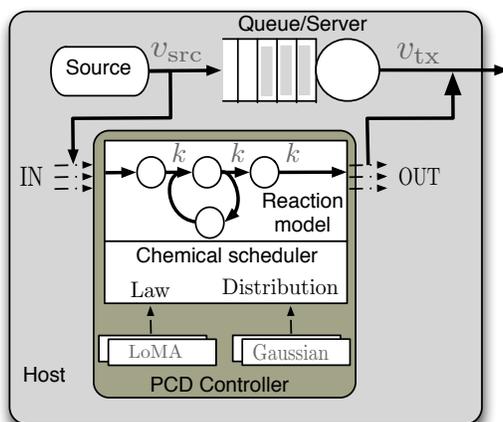


Figure 10: Pseudo Chemically Driven (PCD) Controller

4.2 Design flexibility

By not adhering to the governing rules of real chemistry but simply developing a methodology of algorithm design, PCDAs can use new reaction laws and different inter-reaction time distributions, which broaden the design freedom for implementing networking mechanism and protocols (please refer to fig.10).

As an example of this flexibility, we can use the Michaelis-Menten reaction equation, which governs the global behavior of the PCD rate-controller, as a stand-alone reaction law. Every algorithm that has to implement any sort of rate control over the scheduling of events (e.g. transmissions, memory accesses, media accesses) could then be driven by the Michaelis-Menten law. That is, we could replace the rate-controller's PC model in fig.1 with a molecule whose outflow rate is defined by the Michaelis-Menten equation in (2).

As yet another example of the benefit derived from the gained flexibility in designing, we can imagine the introduction of a law that has a complementary functionality with respect to LoMA: Since LoMA regulates the rate according to the concentrations of reactants, we can see LoMA as a reaction law that bases on a sort of repulsive force. On the contrary, we envision the possible need of an attraction-based law, which would regulate the rate proportionally to the concentrations of products.⁸ Basing on first simulations, we believe that such a law could make easier and more compact some "chemical motifs" proposed in [4] (e.g. the "difference motif" proposed by Meyer can be replaced with a simpler, linear model that makes use of the attraction-based law.)

4.3 Control theory and signal processing to analyze and design algorithms

The strength of PCDAs lies in the possibility to apply well-known tools of mechanical and electrical engineering into the analysis and the design of networking algorithms. Designers can start from algorithms' functional aspects but still focus on algorithms' dynamics and engineer the desired functional flow models. With the help of control theory and signal processing approach, PCDAs can be seen as a combination of simple, reusable building blocks whose dynamics and functionality

⁸The attraction-based law should still depend on the reactants' concentration in order to guarantee a sufficient amount of reactant molecules that are needed by the concerned reaction to occur.

can be described by the Transfer Function (TF). For example, we can see the rate-controller's PC model in fig.1 as a black box where the relationship between output and input rate is formalized by the TF in (14). In this sense, PCDA's design follows Kelly's quote "Complexity must be grown from simple systems": As the complexity increases, the algorithm still consists of analyzable blocks (e.g. low pass filters, rate-controller) that will be interconnected through simple composition rules (e.g. parallel and series).

5 Conclusion and future work

In Pseudo-Chemically Driven Algorithms (PCDAs), there is a separation between the scheduling of actions and their execution. The first is implemented through an extended artificial chemistry whose dynamics can be engineered on flow-level without the need of regulating directly the events on microscopic level; the second does not take care of dynamics but rather acts promptly to the triggering signal of the chemical model.

Here, we proposed to extend the artificial chemistry for networking with further laws (i.e. not LoMA only) and to apply different distributions to the inter-reaction time. We have shown that networking mechanisms benefit of this design flexibility and that the analysis is still accurate. In addition, we plan to formalize the derivation of the ODE approximation of pseudo-chemical driven algorithms and to extend and generalize the knowledge about the intrinsic stochasticity that these algorithms may exhibit.

Future works should aim at designing PCDA's for further specific applications, at identifying simple design patterns that can be easily interconnected to compute complex tasks, and at identifying new reaction laws that enable new algorithmic functionalities still guaranteeing the description of the algorithmic macroscopic behavior through ODEs. We envision that PCDA's could cover not higher layers only of communication systems (e.g. application or transport layers) but rather should include solutions for the physical layer.

Acknowledgment

The present work benefited from the input of Thomas Meyer who provided valuable comments and ideas to the writing of the research summarized here. The author would like also to thank the Swiss National Science Foundation for the partial support of this research through grant #132525.

References

- [1] V. Misra, W. Gong, and D. Towsley, "A fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED," *Proc. SIGCOMM*, vol. 30, pp. 151–160, 2000.
- [2] S. Shakkottai and R. Srikant, "How good are deterministic fluid models of internet congestion control?" in *INFOCOM*, vol. 2, 2002, pp. 497–505.
- [3] S. H. Low, "A duality model of TCP and queue management algorithms," *IEEE/ACM Trans. Net.*, vol. 11, no. 4, pp. 525–536, 2003.

- [4] T. Meyer, "On chemical and self-healing networking protocols," PhD Thesis, University of Basel, Switzerland, 2010.
- [5] S. Ruud, E. H. Owen, L. Janet, and J. M. R. Leon, "Ant-based load balancing in telecommunications networks," *Adaptive Behavior*, vol. 5, no. 2, pp. 169–207, 1996.
- [6] B. Anindya, L. Alvin, and R. Sharad, "Routing using potentials: A dynamic traffic-aware routing algorithm," in *SIGCOMM'03; Karlsruhe; Germany*, August 2003, pp. 37–48.
- [7] D. Klein, J. Hespanha, and U. Madhow, "A reaction-diffusion model for epidemic routing in sparsely connected manets," in *INFOCOM, 2010 Proceedings IEEE*, 14–19 March 2010, pp. 1 – 9.
- [8] R. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.*, vol. 81, no. 25, pp. 2340–2361, 1977.
- [9] S. Keshav, "A control-theoretic approach to flow control," in *ACM SIGCOMM '91*, September 1991, pp. 3–15.
- [10] H. I. Abrash, "Studies concerning affinity," *J. Chem. Edu.*, vol. 63, pp. 1044–1047, 1986.
- [11] P. Waage and C. M. Guldberg, "Studies concerning affinity," *Forhandlinger: Videnskabs - Selskabet i Christiania*, vol. 35, 1864.
- [12] L. Michaelis and M. L. Menten, *Biochem. Z.*, vol. 49, no. 352, pp. 333–369, 1913.
- [13] K. A. Johnson and R. S. Goody, "The original michaelis constant: Translation of the 1913 michaelis–menten paper," *Biochemistry*, vol. 50, no. 39, pp. 8264–8269, 2011.
- [14] C. Reder, "Metabolic control theory: A structural approach," *J. Theor. Biol.*, vol. 135, no. 2, pp. 175–201, 1988.
- [15] B. Ingalls, "A frequency domain approach to sensitivity analysis of biochemical networks," *J. Phys. Chem. B*, vol. 108, no. 3, pp. 1143–1152, 2004.
- [16] P. Dittrich, J. Ziegler, and W. Banzhaf, "Artificial chemistries - a review," *Artificial Life*, vol. 7, no. 3, pp. 225–275, 2001.
- [17] M. A. Gibson and J. Bruck, "Efficient exact stochastic simulation of chemical systems with many species and many channels," *J. Phys. Chem. A*, vol. 104, no. 9, pp. 1876–1889, 2000.
- [18] A. P. Kenneth R. Harris, "Studies on the analysis of nonlinear processes via functional expansions–1. solution of nonlinear ODEs," *Journal of Chemical Engineering Science*, vol. 52, no. 18, pp. 3183 – 3195, 1997.
- [19] J. C. P. Jones and S. A. Billings, "Describing volterra series, and analysis of non-linear system in the frequency domain," *Int. J. Control*, vol. 53, no. 1, pp. 871–887, 1991.
- [20] R. Heinrich, S. M. Rapoport, and T. A. Rapoport, "Metabolic regulation and mathematical models," *Progress in Biophysics and Molecular Biology*, vol. 32, no. 1, pp. 1–82, 1977.
- [21] H. Kacser and J. A. Burns, "The control of flux," *Symposia of the Society for Experimental Biology*, vol. 27, pp. 65–104, 1973.

- [22] C. V. Rao, H. M. Sauro, and A. P. Arkin, "Putting the "control" in metabolic control analysis," in *Proc. DYCOPS 7*, 2004.
- [23] H. M. Sauro and B. Ingalls, "Conservation analysis in biochemical networks: computational issues for software writers," *Biophysical Chemistry*, vol. 109, no. 1, pp. 1 – 15, 2004.
- [24] A. Varga and R. Hornig, "An overview of the OMNeT++ simulation environment," in *Proc. SIMUTOOLS*, 2008, pp. 1–10.
- [25] D. A. McQuarrie, "Stochastic approach to chemical kinetics," *Journal of Applied Probability*, vol. 4, no. 3, pp. 413–478, 1967.
- [26] D. T. Gillespie, "A rigorous derivation of the chemical master equation," *Physica A: Statistical Mechanics and its Applications*, vol. 188, no. 1–3, pp. 404–425, 1992.
- [27] V. Sunkara, "The chemical master equation with respect to reaction counts," in *Proc. 18th World IMACS/MODSIM Congress*, 2009.
- [28] T. Jahnke and W. Huisinga, "Solving the chemical master equation for monomolecular reaction systems analytically," *Journal of Mathematical Biology*, vol. 54, pp. 1–26, 2007.
- [29] N. G. V. Kampen, "Stochastic processes in physics and chemistry," 2007.
- [30] D. T. Gillespie, "The chemical Langevin equation," *J. Chem. Phys.*, vol. 113, no. 1, 2000.
- [31] N. G. van Kampen, "The expansion of the master equation," *Advances in Chemical Physics*, vol. 34, 1976.
- [32] M. Ullah and O. Wolkenhauer, "Investigating the two-moment characterisation of subcellular biochemical networks," *Journal of Theoretical Biology*, vol. 260, no. 3, pp. 340 – 352, 2009.
- [33] L. Ferm, P. Lötstedt, and A. Hellander, "A hierarchy of approximations of the master equation scaled by a size parameter," *J. Sci. Comput*, vol. 34, no. 2, pp. 127–151, 2008.