Analyzing the Dynamics of Chemical Networking Protocols with a Signal Processing Approach

Massimo Monti, Thomas Meyer,
Marco Luise and Christian Tschudin

Technical Report CS-2011-002
University of Basel
July 5th, 2011

Abstract

One important problem in communication protocol engineering is to predict and analyze the dynamic behavior of protocols. Chemical Networking Protocols (CNPs) – a recently proposed design method inspired by chemical reactions – offers the opportunity to transpose the good and well-understood analyzability of chemical reactions to networking protocols. In this paper, we provide a novel method to analyze the dynamics of "chemical protocols". We suggest to study their average flow properties by exploiting tools from several well-established fields, such as model linearization proposed in Metabolic Control Analysis, a state-space description classically used in control theory, and the system's characterization in the frequency domain, which is central to signal theory. With our contribution, CNP designers can easily determine the key parameters of their protocols and understand how to calibrate them in order to obtain the desired behavior. We demonstrate the feasibility of our method by applying it to an actual chemical congestion control protocol and highlight formerly unknown protocol features.

Keywords: dynamic analysis, chemical networking protocols, frequency response, non-linear reaction networks, congestion control algorithm.

Universität Basel — Departement Mathematik und Informatik — Computer Networks Group
Bernoullistr. 16 | CH–4056 Basel | Switzerland Tel/Fax: +41 61 267 05 55/59 http://cn.cs.unibas.ch/

1 Introduction

Predicting and controlling the dynamics of communication protocols are difficult but important challenges in communication systems engineering. Understanding the protocol dynamics is becoming difficult in the complex and large scale networks of today such as the Internet. While in the early days of computer networks, the researchers' attention was on merely functional aspects, in the last three decades, their focus shifted towards mastering the dynamics. This is reflected in the literature, which has been enriched by numerous complex studies about dynamic properties of packet streams [19, 2].

The first step towards understanding the dynamics of protocols is to build an abstract model and to reduce it to an analytically tractable system. Fluid models, for example, focus on the average flow-level properties by abstracting away from the underlying detailed stochastic behavior [25, 28, 20]. Such abstractions however bear some problems: they have to be extracted manually from the already existing source code, and they often over-simplify the behavior of the real protocol.

By drawing an analogy with chemistry, Chemical Networking Protocols (CNPs) [23] offer a new way to describe protocol dynamics through a fluid model that is accurate and easy to obtain. Because the dynamics of chemical protocols is very close to the dynamics of chemical reactions, engineers can exploit well-established tools from chemistry in order to design and analyze the behavior of their protocols. In CNPs, each packet is represented by a virtual molecule that reacts with other packets according to the stochastic laws of chemical kinetics. As a result, the dynamic evolution of CNPs can be exactly described by the Chemical Master Equation (CME) [22]. Because the CME suffers from a combinatorial state explosion, finding a solution is cumbersome and often even impossible [14].

A better approach is to reduce the model to a set of non-linear Ordinary Differential Equations (ODEs). Gillespie proved that this model accurately predicts the trajectory of the system for most of the cases [8]. This fluid model has two advantages over the detailed stochastic analysis: First, it reduces the complexity of the analysis, and second, it can be generated directly and automatically from the topology of the corresponding reaction network.

This paper proposes an extension to the study of CNP dynamics. Beyond previous analysis approaches, where only the steady-state behavior has been looked at, or where numerical ODE solutions have been given, we aim at studying the transient behavior of CNPs for arbitrary initial conditions. We therefore analyze the fluid model of CNPs in the frequency domain by exploiting tools that are well established in fields such as signal processing or control theory. In order to analyze generic non-linear reaction models, we propose to evaluate the system responses in terms of small external perturbations around the equilibrium. With a linearization around the steady state, we avoid the exceptional complexity of methods that are based on the multi-dimensional Laplace transform [15].

In particular, this paper shows how the Metabolic Control Analysis (MCA) (first proposed in [11], introduced in [27], and then further extended in [12]) can be transposed from biology to the analysis of CNPs. With our analysis, we also simplify the design phase of CNPs by identifying the protocols' key parameters and showing how they can be calibrated for obtaining the optimal performance.

The remainder of this paper is organized as follows: Sect. 2 reviews Chemical Networking Protocols (CNPs) and demonstrates with a simple example how this paradigm establishes a relationship between networking protocols and chemistry. Section 3 introduces an actual chemical

protocol that implements a simple congestion control algorithm. Section 4 describes the tools that are well established in other fields and that we apply to study CNP dynamics, whereas in Sect. 5, we show an application of our method through the analysis of the chemical congestion control algorithm. Finally, Sect. 6 discusses the features and limitations of our approach.

2 Chemical Networking Protocols (CNPs)

We start this paper by providing a compact introduction to chemical protocols. This should help the reader to understand our contribution in this novel research area. We first introduce the chemical packet metaphor followed by the concept of artificial chemistry and its use in computer networks. We then summarize already existing approaches to analyze the dynamics of CNPs and identify their problems.

2.1 Modeling Communication in a Chemical Way

Traditionally, protocol execution is handled by a state machine that, upon reception of a packet, synchronously changes its internal state. Here, we introduce a "molecule metaphor" where each packet is treated as a virtual molecule. Molecules react with other molecules in a reaction vessel (node). A reaction may produce other molecules being delivered to the application or being sent over the network. In such a chemical perspective, we obtain a web of reactions that together perform a distributed computation, called Chemical Networking Protocol (CNP).

Each network node contains a multiset of a finite set of molecule species $\mathcal{S} = \{s_1, \ldots, s_{|\mathcal{S}|}\}$ (=packet types). In addition, each node defines a set of reaction rules $\mathcal{R} = \{r_1, \ldots, r_{|\mathcal{R}|}\}$, expressing which reactant molecules can collide and which molecules are generated during this process. Such a set of reactions, which constitute a protocol, is typically represented by chemical reaction equations, e.g.,

$$
\begin{aligned}
r_1&: && \mathrm{W}_i + \mathrm{D}_i \xrightarrow{k_1} \mathrm{W}_i + \mathrm{D}_j \\
r_2&: && \mathrm{W}_i + \mathrm{L}_i \xrightarrow{k_2} \mathrm{L}_i \\
r_3&: && \mathrm{L}_i \xrightarrow{k_3} \emptyset \\
r_4&: && \emptyset \xrightarrow{v_{\mathrm{inc}}} \mathrm{W} \\
r_5&: && \emptyset \xrightarrow{v_{\mathrm{loss1}}} \mathrm{L}
\end{aligned}
\tag{1}
$$

For example, reaction r_1 in node i consumes, if present, two molecules W and D from the local multiset, regenerates W and sends molecule D to neighbor node j. Reactions may also destroy molecules (r_3) or generate new molecules (r_4, r_5). The reaction network above is also depicted in the gray box of Fig. 2(b).

A node that receives a molecule places it into its local multiset. Reaction rules that may consume (=process) the new molecule are however not executed immediately. They are rather scheduled for a later time based on an exponential time distribution. Such an algorithm $\mathcal{A}$ (e.g., [7, 6]) mimics chemical reaction kinetics on the microscopic level and forces the Law of Mass Action (LoMA) on the macroscopic level. The LoMA [1, 30] states that the reaction rate is proportional to the concentration (quantity) c of all reactant molecules: In our reaction r_1, molecules W_1 and D_1 react at average rate $v_1 = k_1 c_{\mathrm{W}_1} c_{\mathrm{D}_1}$, which is equal to the rate at which node i sends D-molecules to node j.

2.2 A Formal Artificial Chemistry for Networking

A CNP can be formally specified as a distributed version [23, 22] of an artificial chemistry [5, 4] by the quadruple $\mathcal{ADC} = (\mathcal{G}, \mathcal{S}, \mathcal{R}, \mathcal{A})$. The graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ represents the computer network, where $\mathcal{V}$ is the set of nodes (chemical reaction vessels) and $\mathcal{E}$ is the set of undirected network links connecting the nodes.

The set of reaction rules $\mathcal{R}$ is formally expressed by the chemical reaction equation

$$r \in \mathcal{R}_i : \quad \sum_{s \in \mathcal{S}_i} \chi_{sr} s \xrightarrow{k_r} \sum_{s \in \mathcal{S}_i} \xi_{sr} s. \tag{2}$$

The reaction coefficient k_r defines the reaction speed, the integers χ_{sr} and ξ_{sr} are the stoichiometric *reactant* and *product* coefficients that denote the number of s-molecules *consumed* and *produced* by reaction r, respectively. Note that all reactants must reside in the same node in order to react, while the products may also be produced in one of the direct neighbors of node i.

The algorithm $\mathcal{A}$ defines the dynamical behavior of the artificial chemistry, namely *which* reaction is executed *when*. A possible implementation of $\mathcal{A}$ is the "real-time next reaction algorithm" [22], a variant of an algorithm originally proposed in [6] to simulate real chemical reactions. The algorithm computes the probability that reaction $r \in \mathcal{R}_i$ occurs in the next infinitesimal period $[t, t + dt)$. This probability is given by the propensity function v_r, which depends on the concentration vector $\mathbf{c} = (c_1, \ldots, c_{|\mathcal{S}|})^T$:

$$r \in \mathcal{R} : \quad v_r(\mathbf{c}) = k_r \prod_{s \in \mathcal{S}} x_s^{\chi_{sr}} \tag{3}$$

2.3 Approximation of the Chemical Model Behavior

Real chemical reactions exhibit the same reaction probabilities, and it is well known that the stochastic evolution of chemical reactions can be described by the Chemical Master Equation (CME) [21]. In theory, CNP designers could use the CME to study the exact behavior of protocols at the microscopic level. However, the CME suffers from the known curse of dimensionality: Each species adds one dimension to the problem, leading to a combinatorial growth of the computational complexity. For a model of N species, each being represented by a maximum of M instances, the CME consists of one equation per state, resulting in M^N coupled differential equations. Even though the CME is solvable for simple systems such as linear first order reaction networks [14], it cannot be solved for more complex systems.

At the macroscopic level, chemical kinetics can be approximated by deterministic Differential Rate Equations (DREs) [17]. The computational work required to solve them is considerably smaller. The average trajectory of a reaction network can now be evaluated by solving a system of ODEs, containing only one ODE per species. This deterministic approximation is valid only if we are dealing with large species concentrations; intuitively, the higher the number of molecules per species is, the lower is the effect of random fluctuations. This approximation is also supported by the empirically observed "Law of Mass Action" [1, 30], which states that in a free and homogeneous medium the reaction rate is proportional to the concentration of the involved reactants.

In the CNP setting, this means that packets of the same type are building up a "pressure" for the reactions processing them. We are able to express the dynamics of CNPs as time derivative of the species concentration vector $\mathbf{c}$ expressed as

$$\dot{\mathbf{c}} = \mathbf{\Psi} \cdot \mathbf{v}(\mathbf{c}). \tag{4}$$

In this expression, $\mathbf{\Psi} = [\psi]_{sr}$ is the stoichiometric matrix whose elements represent the difference between the stoichiometric product and reactant coefficients: $\psi_{sr} = \xi_{sr} - \chi_{sr}$. The propensity vector $\mathbf{v} = (v_1, \ldots, v_{|\mathcal{R}|})^T$ simply lists the reaction rates.

Even though we have reduced the complexity of the mathematical model, there are still problems in obtaining useful information about the protocols' behavior: We often deal with multi-molecular reactions, resulting in non-linear ODEs that are hard to solve. We could solve the ODEs numerically, but only for one particular scenario at a time (e.g., one external stimulus, one set of initial conditions). We rather aim at studying the dynamics in a more general way. But before we provide a solution to this problem, we present a simple example that illustrates the concepts introduced so far.

2.4 A Simple Example

The following example demonstrates how the CNP paradigm establishes a relationship between a networking entity (a packet queue) and chemistry, through the unimolecular reaction depicted in Fig. 1. A molecular species S can be seen as a packet queue that temporarily stores molecules

(a) Abstract Chemical Model: Unimolecular (Linear) Reaction

(b) Corresponding FIFO Queue

Figure 1: Species/Queue analogy: Packets (=molecules) arrive at queue (=molecular species) S with a rate of v_{in}. The service rate of the queue is proportional to (i) its concentration (=fill level) c_{S} and (ii) the reaction coefficient k: $v_{\text{out}} = kc_{\text{S}}$.

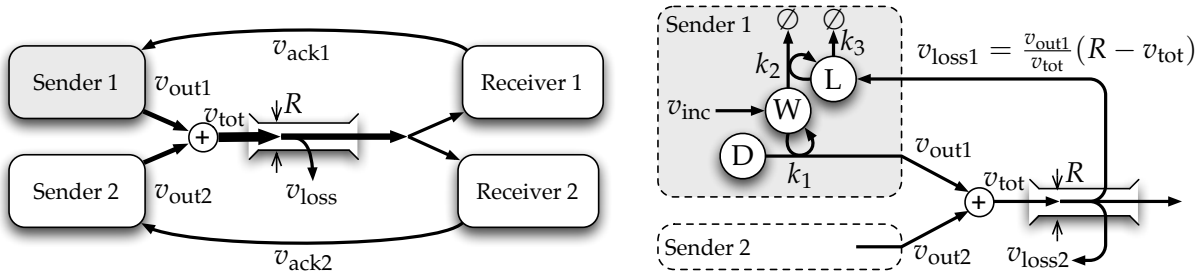
(=packets) until an egress reaction consumes them with a rate of $v_{\text{out}} = kc_{\text{S}}$, according to the LoMA. The term c_{S} denotes the concentration of the S-species, i.e., the fill level of the queue. The egress reaction is equivalent to a non-work-conserving server, which services the queue with a rate proportional to its fill level c_{S} and to the coefficient of the reaction k. A zero-order reaction models the arrival process with a rate of v_{in} packets per second. The behavior of the queue's reaction network in Fig. 1(a) can be given in the chemical ODE notation (4) as

$$\dot{c}_{\text{S}} = v_{\text{in}} - kc_{\text{S}}. \tag{5}$$

Such a chemical queue cannot be used to limit the transmission rate to the link capacity as work-conserving queues usually do. Instead, the queue exhibits low-pass characteristics, meaning that it is able to effectively dampen packet bursts. We aim to provide methods to find out what kind of filter or behavior a chemical protocol implements – not only for this simple queue, but for an arbitrary "chemical queueing network".

3 A Chemical Congestion Control Algorithm

This section introduces an actual chemical protocol that is more complex than the simple queuing example discussed before – a Chemical Congestion Control Algorithm (C_3A). The role of a



(a) Channel sharing communication model

(b) The chemical reaction network of C_3A

Figure 2: A general model of two packet streams sharing the same channel (a), and the chemical reaction network of C_3A that implements the chemical congestion control algorithm in sender 1 (b).

congestion control algorithm is to provide efficient and fair utilization of the network's capacity for multiple competing streams running over the same capacity-limited channel. Without loosing too much generality, Fig. 2(a) shows the simplest scenario where only two streams are sharing a channel with capacity R. Sender 1 is equipped with an implementation of C_3A, originally proposed in [22] and [24]. Sender 1 contains the reaction network shown in Fig. 2(b). These virtual chemical reactions take care to immediately throttle the transmission rate if a lost packet is detected via a sequence number gap in the stream of returned acknowledgments. The protocol has a behavior similar to TCP Renos additive increase / multiplicative decrease mechanism [13], but is implemented by just coupling the "right" chemical reactions as follows:

The application layer (not shown) is responsible to take the next data packet from the transmit queue and put it into the reaction vessel as a D-molecule as soon as a data packet has been sent. In chemical terms, data species D is buffered and contains a constant amount of instances (e.g., $c_D = 1\,\text{pkt} = \text{const.}$). The concentration c_W of the window-species W controls the transmission rate of sender 1 according to the LoMA: $v_{out1} = k_1 c_D c_W$. The concentration of W-molecules increases continuously at constant rate v_{inc} as long as no packet loss is detected. This inflow of W-molecules mimics the additive (linear) increase mechanism of TCP-Reno.

Depending on the packet rate v_{out2} of sender 2, which competes for the link capacity, some data packets may be lost. The total link loss rate is $v_{loss} = \max(v_{tot} - R, 0)$. If we assume statistical multiplexing, this total loss rate affects the two streams from sender 1 and 2 *pro rata*, resulting in a packet loss rate for sender 1 of $v_{loss1} = v_{loss}(v_{out1}/v_{tot})$.

This loss rate triggers the selective feedback mechanism [3] of C_3A. In particular, each lost packet generates an L-molecule. In proportion to the concentration of L, c_L, a fast reaction (coefficient k_2) quickly removes W-molecules, which causes an immediate reduction of the transmission rate v_{out1}. The duration of this decrease is controlled by the lifetime of L-molecules, which decay via the reaction with coefficient k_3. This feedback mechanism leads to an exponential decrease of the transmission rate, which is comparable to the multiplicative decrease of TCP-Reno.

Since this algorithm implements congestion control with coupled chemical reactions only, it is straight forward to automatically derive a mathematical description of the system's behavior. The dynamics of the chemical model in Fig. 2(b) can be described using the ODE approximation (4),

leading to one equation for each molecule species:

$$\dot{c}_{\mathrm{W}} = \overbrace{v_{\mathrm{inc}}}^{\text{lin. increase}} - \overbrace{k_2 c_{\mathrm{W}} c_{\mathrm{L}}}^{\text{exp. decrease}} \tag{6a}$$

$$\dot{c}_{\mathrm{L}} = \underbrace{(k_1 c_{\mathrm{W}} c_{\mathrm{D}} + v_{\mathrm{out2}} - R)\frac{k_1 c_{\mathrm{W}} c_{\mathrm{D}}}{k_1 c_{\mathrm{W}} c_{\mathrm{D}} + v_{\mathrm{out2}}}}_{\textit{pro rata}\text{ packet loss}} - \underbrace{k_3 c_{\mathrm{L}}}_{\text{decay of L}} \tag{6b}$$

Note that (6) is a system of coupled non-linear differential equations that is not trivial to solve directly. In addition, from the mere solution of (6), it is difficult for protocol designers to know *which* control parameters define the protocol behavior and *how* these parameters should be calibrated in order to let the protocol behave in the desired way. Another challenge that we are facing is the problem to predict the protocol dynamics in response to several different input stimuli: e.g., how does sender 1 adapt its transmission rate in response to a burst of a competing stream, or in response to different packet loss probabilities? These challenges can be addressed with a signal processing and control theory approach.

4 Signal Processing and Control Theory Applied to CNPs

This section provides the foundation of our approach to analyze CNP dynamics. It summarizes the tools that are available in other fields and that we have transposed to CNPs. But first, we give some motivation for their application to the analysis of protocol dynamics.

4.1 Motivation and Method Overview

Our aim is to make the study of CNP dynamics as simple as possible while preserving sufficient accuracy. For this reason, we avoid the complexity of the CME and use coupled ODEs instead. That is, we focus on the average trajectory of the stochastic process. Note that our analysis deals with deterministic continuous-valued signals rather than stochastic time-continuous discrete-valued signals. This approximation implicitly assumes that molecular concentrations are sufficiently high.

Direct solutions of coupled ODEs have two limitations: they are non-trivial to obtain for systems with non-linearities (often, CNPs require such non-linear reaction models) and they describe system responses to one specific input stimulus only. Both limitations can be overcome by first linearizing the model around its steady state and by analyzing it in the frequency domain.

We linearize reaction networks by using a method from biology – the Metabolic Control Analysis (MCA). MCA was first proposed in [11] and [16] and then introduced in [27] with the purpose of understanding the sensitivity of biological processes. Such a linearization reduces the analytical complexity that we would encounter otherwise (e.g., when using methods that directly account for the systems' non-linearities [18, 15]).

We are now dealing with a linearized model and are allowed to apply the Laplace Transform to determine the transfer function of the system in the frequency domain [12]. With this transformation – from functions with a real dependent variable (i.e., the time) into functions with a complex dependent variable – we convert linear coupled ODEs into algebraic expressions that are easier to manipulate. Furthermore, the frequency transform of the system is valid for arbitrary stimuli. This allows us to characterize the system's response in a more general way than in the time domain.

As a next step, we make use of "control theory" tools, in particular the state-space description of Linear Time Invariant (LTI) models, as proposed in [12]. In this way, the behavior of CNPs can be predicted not only in response to constant perturbations, but also in response to other varying internal and external perturbation channels. External perturbations include for example tunable protocol parameters, inflow and outflow rates, and protocol configurations, whereas internal perturbations cover reaction coefficients, molecule concentrations, and internal reaction rates, to name a few.

In summary, our analysis procedure makes use of the LTI state-space model description. For this purpose, the system must first be linearized around its fixed points (Sect. 4.2). Second, the inputs and outputs of the system must be identified (Sect. 4.3). Third, standard methods from control theory can be applied to calculate the frequency response of the system (Sect. 4.4). Each element of the frequency transform matrix represents how a perturbation on one input parameter affects one of the observable outputs.

4.2 Linearization of the ODEs

We first make the ODE description of a CNP (4) dependent on a vector of external perturbations $\mathbf{p}$, which represent the input signals of our analysis:

$$\dot{\mathbf{c}}(t) = \mathbf{\Psi} \cdot \mathbf{v}\left(\mathbf{c}(t), \mathbf{p}(t)\right) \tag{7}$$

Again, $\mathbf{c}$ is the vector of approximated concentrations per species, $\mathbf{\Psi}$ the stoichiometric matrix of the reaction network, and $\mathbf{v}$ is the propensity (reaction rate) vector.

Then, we perform a local perturbation analysis that avoids to deal with the non-linearities of (7) by focusing only on the fluctuations around the steady states of the system, $\mathbf{c}^*$, $\mathbf{p}^*$:[1]

$$\mathbf{x}(t) = \mathbf{c}(t) - \mathbf{c}^* \tag{8a}$$

$$\mathbf{u}(t) = \mathbf{p}(t) - \mathbf{p}^* \tag{8b}$$

The vectors $\mathbf{x}$ and $\mathbf{u}$ represent deviations from the nominal state and from nominal parameter values, respectively. To this end, we can approximate the behavior of a CNP with the linear equation

$$\dot{\mathbf{x}}(t) = \mathbf{\Psi} \cdot \left.\frac{\partial \mathbf{v}}{\partial \mathbf{c}}\right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{x}(t) + \mathbf{\Psi} \cdot \left.\frac{\partial \mathbf{v}}{\partial \mathbf{p}}\right|_{(\mathbf{c}^*, \mathbf{p}^*)} \cdot \mathbf{u}(t) \tag{9}$$

This linear ODE is a good approximation around the steady state. It describes the system's response to small variations of all parameters $\mathbf{p}$, i.e., to all kinds of internal and external perturbations [12].

Note that (7) and (9) may contain redundant equations. This happens if the chemical reaction network contains loops, i.e., when there are conservation constraints [26]. In this case, MCA usually simplifies the stoichiometric matrix $\mathbf{\Psi}$ by using Gauss Jordan elimination with partial pivoting, resulting in a reduced row Echelon form of $\mathbf{\Psi}$. As this "simplification" does not provide an improvement in the analysis procedure, we ignore this step in this paper.

[1] Steady states are calculated by studying (4) at equilibrium, i.e., by solving $\dot{\mathbf{c}}(c, p) = \mathbf{0}$.

4.3 State-Space Representation and Definition of Input and Output Signals

After we linearized the ODEs around the steady state, we express the system in "control theory" terms and refer to the traditional state-space representation of LTI systems, as proposed in [26]:

$$\dot{\mathbf{x}}(t) = \mathbf{A} \cdot \mathbf{x}(t) + \mathbf{B} \cdot \mathbf{u}(t) \tag{10a}$$

$$\mathbf{y}(t) = \mathbf{C} \cdot \mathbf{x}(t) + \mathbf{D} \cdot \mathbf{u}(t) \tag{10b}$$

The state matrix $\mathbf{A}$ defines how a perturbation of the concentrations (state variables) affect their future change. This matrix is equal to the Jacobian matrix evaluated at the fixed point:

$$\mathbf{A} = \boldsymbol{\Psi} \cdot \left.\frac{\partial \mathbf{v}}{\partial \mathbf{c}}\right|_{(\mathbf{c}^*, \mathbf{p}^*)} \tag{11}$$

The input matrix $\mathbf{B}$ indicates how external perturbations (i.e., perturbations of protocol input configurations) affect fluctuations of the state; this is also evaluated at the fixed point:

$$\mathbf{B} = \boldsymbol{\Psi} \cdot \left.\frac{\partial \mathbf{v}}{\partial \mathbf{p}}\right|_{(\mathbf{c}^*, \mathbf{p}^*)} \tag{12}$$

The output matrix $\mathbf{C}$ and the feedforward matrix $\mathbf{D}$ define what output we are interested in. They must be appropriately chosen according to the response we want to analyze. Often we are interested in two typical output configurations: (*i*) molecular concentrations, and (*ii*) reaction/transmission rates. (*i*) In order to analyze the fluctuations of molecular concentrations with respect to perturbations of the input parameters, we define

$$\mathbf{C} = \mathbf{I} \tag{13a}$$

$$\mathbf{D} = \mathbf{0} \tag{13b}$$

yielding $\mathbf{y}(t) = \mathbf{c}(t)$. (*ii*) In order to look at the fluctuations of the reaction rates with respect to a perturbation of the input parameters, we define

$$\mathbf{C} = \left.\frac{\partial \mathbf{v}}{\partial \mathbf{c}}\right|_{(\mathbf{c}^*, \mathbf{p}^*)} \tag{14a}$$

$$\mathbf{D} = \left.\frac{\partial \mathbf{v}}{\partial \mathbf{p}}\right|_{(\mathbf{c}^*, \mathbf{p}^*)} \tag{14b}$$

4.4 Signal Analysis in the Frequency Domain

Since our new model is linear, we can make use of the Laplace transform $\mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t)\, dt$ to transpose the time-domain interpretation (10) into the following frequency-domain interpretation:

$$s \cdot \mathbf{x}(s) - \mathbf{x}_0 = \mathbf{A} \cdot \mathbf{x}(s) + \mathbf{B} \cdot \mathbf{u}(s) \tag{15a}$$

$$\mathbf{y}(s) = \mathbf{C} \cdot \mathbf{x}(s) + \mathbf{D} \cdot \mathbf{u}(s) \tag{15b}$$

The vector $\mathbf{x}(s)$ is the Laplace transform of the deviation from the nominal state, $\mathbf{y}(s)$ is the Laplace transform of the output, $\mathbf{u}(s)$ is the Laplace transform of deviation from the nominal parameters, and $\mathbf{x}_0$ is the initial condition vector. The state vector results in

$$\mathbf{x}(s) = (s \cdot \mathbf{I} - \mathbf{A})^{-1} \mathbf{x}_0 + (s \cdot \mathbf{I} - \mathbf{A})^{-1} \mathbf{B} \cdot \mathbf{u}(s) \tag{16}$$

and the output vector is

$$\mathbf{y}(s) = \mathbf{C} (s \cdot \mathbf{I} - \mathbf{A})^{-1} \mathbf{x}_0 + \Big(\mathbf{C} (s \cdot \mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D} \Big) \mathbf{u}(s) \tag{17}$$

From (17), and by assuming no initial perturbation ($\mathbf{x}_0 = \mathbf{0}$), we are able to derive the *Transfer Function* (TF) of the LTI system:

$$\mathbf{H}(s) = \frac{\mathbf{y}(s)}{\mathbf{u}(s)} = \mathbf{C} (s \cdot \mathbf{I} - \mathbf{A})^{-1} \mathbf{B} + \mathbf{D} \tag{18}$$

The TF matrix $\mathbf{H}$ gives us the transient behavior of all steady-state output deviations $\mathbf{y}$ w.r.t. perturbations on all inputs $\mathbf{u}$.

5 Analyzing C_3A through a "Signal Processing Approach"

This section provides a tangible example of our approach to analyze the dynamics of CNPs. It reports specific results for the Chemical Congestion Control Algorithm (C_3A) that we introduced in Sect. 3.

First, we identify inputs and outputs of the analysis and the initial system setting (Sect. 5.1). In Sect. 5.2, we then provide the transfer function that characterizes the C_3A. Next, we look at its step-response in Sect. 5.3, where we also identify the key-quantities related to its time-behavior. In Sect. 5.4, we review the findings made so far. Then, in Sect. 5.5, we complement our analysis by comparing it to simulation results and discuss it in Sect. 5.6.

5.1 Defining Inputs and Outputs of the Analysis

The aim of our analysis is to detect the key parameters of C_3A. We study the ability of sender 1 to adapt its transmission rate v_{out1} dynamically to the channel capacity R and to the traffic generated by competing senders with a cumulative transmission rate of v_{out2} (see Fig. 2).

For these requirements we define our system as follows: The internal state of the system is given by the concentration of the two volatile species: $(c_{\text{W}}, c_{\text{L}})$. The vector of input signals contains only one element – the perturbation of v_{out2}: $\mathbf{u} = (v_{\text{out2}})$. The observable output signal of our model shall be the transmission rate of sender 1, $v_{\text{out1}} = k_1 c_{\text{W}}$.[2] Note however, that this transmission rate is neither a concentration of a species nor a reaction rate that appears in our ODE model (6), and consequently, it cannot be derived directly with one of the two methods (13) or (14). But we can use a little trick here: Since the transmission rate is proportional to the concentration of W, we use (13) to obtain the response to perturbations of c_{W} in the frequency domain and multiply it by the reaction coefficient k_1 to obtain v_{out1}.

[2] We set $c_{\text{D}} = 1 = \text{const.}$ because D is buffered.

We assume that initially the system is in steady state and that sender 2 is silent. Consequently, sender 1 should be able to fully exploit the channel capacity R as no other users are sharing the channel. We then perturb the transmission rate v_{out2} of sender 2 around this steady state and study sender one's response.

5.2 Obtaining the Transfer Function

After we have defined the input and output signals, we determine the Transfer Function (TF) of the linearized system according to Sect. 4. The obtained TF is quite complex. However, the influence of some terms is marginal for sufficiently large capacity values:

$$R \gg \sqrt{\frac{4k_1k_3v_{\text{inc}}}{k_2}} \tag{19}$$

In this region, the TF can be simplified to

$$H(s) = -\frac{Rk_2}{s^2 + sk_3 + Rk_2} \tag{20}$$

From this TF, we conclude that C_3A acts as a second order low pass filter to a perturbation of sender two's transmission rate, because (20) has no finite zeros (roots of the TF's numerator) and two poles (roots of the TF's denominator).

We have to distinguish three different ways a second order system behaves: The system is under-dampened if the two poles are complex conjugate, it is said to be over-dampened if they are real, and the system is critically dampened for the special case where the real poles coincide.

For C_3A the nature of the poles only depends on the channel capacity R and the reaction coefficients k_2 and k_3. This indicates that the behavior of C_3A can be controlled by these parameters. We further conclude that C_3A is bounded-input/bounded-output stable, because its poles always have negative real part for positive values of R, k_2 and k_3.

Figure 3 shows the TF for three parameter sets: *A* and *B* are under-dampened, whereas *C* is over-dampened.

5.3 Step Response

Other than the transfer function, the step response gives us a more intuitive perception of how C_3A behaves in face of a sudden competition. Figure 4 depicts the response of sender one's transmission rate when sender 2 starts to transmit on the same channel with a rate of half the channel capacity: $v_{\text{out2}}(t \to 0^-) = 0\,\text{pkt/s}$, $v_{\text{out2}}(t \to 0^+) = R/2$; the amplitude (y-axis) is normalized to the channel capacity R. In order to reveal the properties of C_3A in the time domain, we first derive the *natural frequency* $\omega_n = \sqrt{Rk_2}$ and the *damping ratio* $\zeta = k_3/(2\sqrt{Rk_2})$ from (20). From these two quantities and according to the nature of the poles, we obtain the settling time, the oscillation frequency, and the overshoot.

Under-dampened The response of the under-dampened C_3A (i.e., for $R > k_3^2/(4k_2)$) is an exponentially decaying sinusoid that approaches the final transmission rate (see scenarios *A* and *B* in

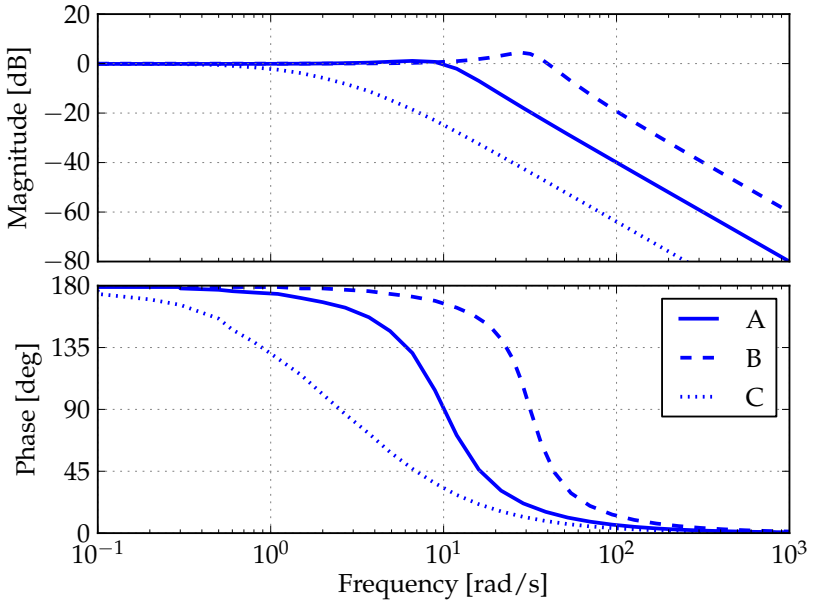


Figure 3: Bode diagram showing the response of sender one's output rate v_{out1}; sender 2 is initially silent; the injection rate is fixed to $v_{\text{inc}} = 10^3$ pkt/s; the parameters are chosen as follows:
A: $k_1 = 0.079\,/(\text{pkt}\,\text{s})$, $k_2 = 0.1000\,/(\text{pkt}\,\text{s})$, $k_3 = 10\,/\text{s}$,
B: $k_1 = 0.079\,/(\text{pkt}\,\text{s})$, $k_2 = 1.0000\,/(\text{pkt}\,\text{s})$, $k_3 = 20\,/\text{s}$, and
C: $k_1 = 0.030\,/(\text{pkt}\,\text{s})$, $k_2 = 0.0064\,/(\text{pkt}\,\text{s})$, $k_3 = 6\,/\text{s}$

Fig. 4). In this case, the *settling time* (i.e., the time that sender 1 takes to adapt its transmission rate) is given by

$$t_{\text{set}} \simeq \frac{4.6}{\zeta\omega_n} = \frac{9.2}{k_3} \tag{21}$$

Another important quantity of under-dampened systems is the *oscillation period* given by

$$t_{\text{osc}} = \frac{2\pi}{\omega_n\sqrt{1-\zeta^2}} = \frac{4\pi}{\sqrt{4Rk_2 - k_3^2}} \tag{22}$$

Finally, the *overshoot* quantifies the maximum amplitude of the response, measured from the desired response of the system:

$$\text{overshoot}\,[\%] = 100 \cdot e^{\frac{-\pi\zeta}{\sqrt{1-\zeta^2}}} = 100 \cdot e^{-\frac{\pi k_3}{\sqrt{4Rk_2 - k_3^2}}} \tag{23}$$

The A-curve in Fig. 4 exhibits a small overshoot (16 % of its final value) and a settling time of about 0.92 s. The settling time for scenario B is shorter, but its oscillations are more evident (overshoot of 35 % and an oscillation period of 0.46 s).

Over-dampened If C_3A is over-dampened (i.e., for $R < k_3^2/(4k_2)$), its response is a decaying exponential (see scenario *C* in Fig. 4). In this mode, the equations for the settling time is slightly

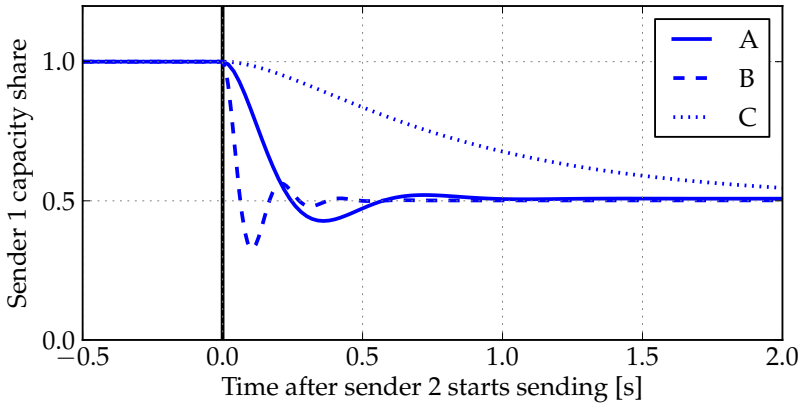


Figure 4: Step response of the normalized output rate v_{out1}/R of sender 1 to a sudden start of sender 2: $v_{\text{out2}}(t \to 0^+\,\text{s}) = R/2$; the injection rate is fixed to $v_{\text{inc}} = 10^3\,\text{pkt/s}$; the parameters are chosen as follows (same as in Fig. 3):
A: $k_1 = 0.079\,/\,(\text{pkt}\,\text{s})$, $k_2 = 0.1000\,/\,(\text{pkt}\,\text{s})$, $k_3 = 10\,/\text{s}$,
B: $k_1 = 0.079\,/\,(\text{pkt}\,\text{s})$, $k_2 = 1.0000\,/\,(\text{pkt}\,\text{s})$, $k_3 = 20\,/\text{s}$, and
C: $k_1 = 0.030\,/\,(\text{pkt}\,\text{s})$, $k_2 = 0.0064\,/\,(\text{pkt}\,\text{s})$, $k_3 = 6\,/\text{s}$.

different, because there are no oscillations:

$$t_{\text{set}} \simeq \frac{5}{(\zeta - \sqrt{\zeta^2 - 1})\omega_n} = \frac{10}{k_3 - \sqrt{k_3^2 - 4Rk_2}} \tag{24}$$

The C-curve in Fig. 4 is characterized by a damping ratio of 1.19 and approaches the final steady value in 3.6 s.

5.4 Derived Guidelines for a C_3A Implementation

Our analysis of the abstract reaction network of C_3A reveals several important insights about its behavior. In order to build a real protocol implementation of C_3A, we have to determine "good" parameters for the reaction coefficients k_1, k_2, and k_3, and the injection rate v_{inc}. Thereby, our main objective is to obtain a short settling time and a small overshoot.

Interestingly, the behavior of C_3A only depends on the channel capacity R and the two parameters k_2 and k_3, but neither on k_1 nor on v_{inc}, if we assume that the channel capacity R is large enough according to (19).[3] The transmission coefficient k_1 and the injection rate v_{inc} define how aggressive the packets are sent and how aggressive this rate increases, respectively. But since the negative feedback from lost packets is a linear function of the transmission rate, the rate-throttling behavior is as aggressive as the transmission strategy. That is, the effects of k_1 and v_{inc} approximately cancel each other out; the two parameters only marginally influence the behavior of our abstract C_3A model.

The coefficients k_2 and k_3 define how C_3A reacts to lost packets. According to (21), the settling time shrinks for larger values of k_3. Intuitively, this makes sense as k_3 specifies how fast information from lost packets is destroyed. On the other hand, the parameter k_2 defines how aggressively the transmission rate is reduced for each lost packet. As (23) suggests, the smaller k_2 is w.r.t. k_3,

[3] The parameters k_1 and v_{inc} however co-decide for which capacity values the simplification holds.

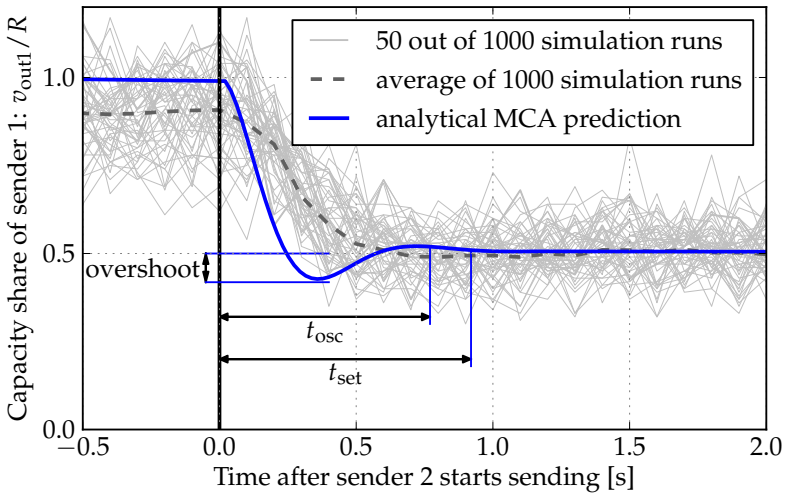


Figure 5: Comparison between the analytically predicted response and 1000 OMNeT++ simulation runs of the chemical congestion control algorithm. Parameter set A: $k_1 = 0.079\,/(\text{pkt}\,\text{s})$, $k_2 = 0.1\,/(\text{pkt}\,\text{s})$, $k_3 = 10\,/\text{s}$, $v_{inc} = R = 1000\,\text{pkt}/\text{s}$.

the more the system integrates the feedback over time and the more overshoot it produces in its response.

Thus, as a general recipe, we would like k_2 and k_3 to be large in order to obtain a quick and accurate response. We also want k_1 and v_{inc} to be large so that C_3A quickly probes for the channel capacity, even though this is not reflected by the transfer function. But at the same time we have to keep k_1, k_3, and v_{inc} small compared to k_2 such that our assumption in (19) still holds.

5.5 Simulation Results

In order to complement our analysis, we compare the analytically predicted response of C_3A to empirical results obtained from OMNeT++ simulations [29, 10]. We implemented C_3A with parameter set A (see Fig. 3) that shows a quick adaptation with moderate overshoot, and we simulated the same sudden competitive burst from $v_{out2} = 0\,\text{pkt}/\text{s}$ to $R/2$.

Figure 5 compares the step response of v_{out1} as predicted by our analysis (continuous line) to the average over thousand simulation runs. Additionally, the light gray lines show the response of fifty randomly-picked simulation runs. The real C_3A protocol implements a stochastic process, whose realization is a specific simulation run with seemingly chaotic behavior. However, according to chemical kinetics [7], the average over many runs is deterministic and should reflect the ODE approximation, which was the basis for our analytical approach. Our analysis predicts a settling time of $t_{set} \sim 0.92\,\text{s}$, an overshoot of approximately 16.3, %, and dampened oscillation with a period of $t_{osc} \sim 0.73\,\text{s}$. The simulation results are close to these numbers, but a deviation is clearly recognizable.

5.6 The Influence of Stochasticity in the C_3A Implementation

There are several reasons for the discrepancies between the predicted response, the average of simulation runs, and a certain particular simulation run.

The law of large numbers The reason why a single simulation run deviates from the average of all runs roots in the law of large numbers. The packet scheduler of C_3A and similar CNPs is a stochastic algorithm that mimics real chemical reactions [22]. The more packets a CNP deals with, the more closely a simulation run (realization of the stochastic process) will approach its average. This well-known behavior of the law of large numbers was derived for real chemical reactions in [8] and [9]. According to these findings, a CNP behaves more deterministically, the higher the fill levels of its packet queues are, i.e., the higher the molecule concentration are.

Differences in the state-space representation The fact that even the average over many simulation runs deviate from the predicted behavior can be explained by the difference in the state-space representation. The concentration values in the ODEs are real-valued numbers, whereas in the real implementation of C_3A, concentrations represent packet quantities – the state space is discrete. While the continuous ODE model changes the output rate with infinite resolution and smoothly stabilizes and optimizes it, a real implementation may produce oscillations as we argue in the following paragraphs.

Example of the discrepancy between analytical prediction and simulation Fig. 5 shows small differences between our analytical prediction and the real implementation. For example, when sender 2 is silent, our analysis predicts that sender 1 is able to send with a transmission rate equal to the channel capacity, while the C_3A implementation is not able to exploit the full capacity.

We can either configure the C_3A implementation with high or with low molecule concentrations. When we parametrize the protocol such that the resulting concentrations are high, the protocol is able to achieve a stable transmission rate as predicted, but only at the cost of high packet loss. This clearly is undesirable. If, on the other hand, the concentrations are kept low, the packet loss rate can be reduced by letting the protocol quickly and immediately react to a single lost packet. In this mode, the concentration of the L-molecule is either zero or one. A feedback based on such a discrete signal leads to (intended) oscillations where the protocol continuously alternates its transmission behavior between linear increase (in absence of packet loss) and exponential decrease (in presence of packet loss). The exact timing of these two phases is different from run to run, due to the stochasticity of the scheduling algorithm. For the same reason, the transmission rate is stochastic, which leads to small bursts that trigger too early packet losses. This is why the full capacity cannot be exploited.

6 Discussion

Our analysis approach to CNPs bases on a (deterministic, continuous-valued) ODE description of abstract chemical reaction networks proposed in [8]. On one hand, this approximation enables us to have a general prediction of the average behavior of protocols, but on the other hand, we miss some particular features of the stochastic process. Specifically, the more packets a CNP deals with, the more closely the analysis will predict every possible realization of the protocol behavior.

Despite the discrepancies we expect, there are several arguments that support our mean-field approach: First, CNPs usually operate with high molecule concentrations, which lowers possible differences between empirical results and our prediction. Second, in an early stage, designers are often more interested in understanding the average trend of their protocols rather than the

stochastic details. Third, by taking a macroscopic view on the protocol dynamics, we are able to quickly detect key parameters of a CNP and understand how these parameters should be calibrated in order to obtain specific protocol features. This is hardly possible for the detailed stochastic treatment of complex CNPs.

Linearizing the system around the steady state is another simplification we suggested in order to study complex systems. This additional simplification is feasible in many cases, for example, when we are interested in how protocols evolve starting from stable initial conditions.

The strength of our analysis lies in the possibility to automatically generate a generic matrix description of an arbitrary CNP. We just have to determine the input perturbation vector, and obtain a transfer function matrix, whose elements describe all different aspects of the protocol's dynamics we are interested in.

Finally, we have to note that the analysis proposed in this paper does not consider propagation delays. In previous simulations of C_3A [22, 24], we also studied the influence of non-zero round-trip times. We proposed in [22, Sect. 12.1.2] to model such delays as additional first-order reactions within the chemical model. The approach proposed in this paper can then be applied directly to the extended model.

7 Conclusion

In this paper, we proposed an approach to analyze and understand dynamics of Chemical Networking Protocols (CNPs). The analysis exploits tools from several well-established fields: a model linearization proposed in Metabolic Control Analysis (MCA), the state-space description of Linear Time Independent (LTI) models classically used in control theory, and the characterization of the system in the frequency domain, central to signal theory. This analysis leads to general observations about the average behavior of CNPs.

Our approach starts with an automatically generated deterministic description of the protocol as chemical reaction network, the behavior of which is approximated by ODEs. The larger the packet buffers of CNPs are, the more accurate our analysis is. Future work could be oriented towards a rigorous treatment of the stochastic aspects of CNP dynamics, i.e., the exact microscopic description of CNPs, valid also when dealing with a very low number of packets. We envision to still use our deterministic approach in order to determine the systemic trajectory, but to extend it by an analysis of the superimposed stochastic noise. The latter would describe the deviation between the deterministic model and the real protocol behavior. This would enable protocol designers to define the margins in which the protocol should operate despite stochastic fluctuations.

In the future, we would also like to improve the analysis presented in this paper by incorporating delays into the chemical model. We believe that we can still use the approach presented here by extending special reactions that model these delays.

References

[1] Henry I. Abrash. Studies concerning affinity. *J. Chem. Edu.*, 63:1044–1047, 1986.

[2] Christos G. Cassandras. *Discrete Event Systems*. Richard D. Irwin, Inc., and Asken Associates, Inc., Burr Ridge, IL, USA, 1993.

[3] Dah-Ming Chiu and Raj Jain. Analysis of the increase and decrease algorithms for congestion avoidance in computer networks. *J. Comp. Net. ISDN Sys.*, 17(1):1–14, 1989.

[4] P. Dittrich, J. Ziegler, and W. Banzhaf. Artificial chemistries - a review. *Artificial Life*, 7(3):225–275, 2001.

[5] Peter Dittrich. *On Artificial Chemistries*. Doctoral thesis, University of Dortmund, Jan 2001.

[6] Michael A. Gibson and Jehoshua Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A*, 104(9):1876–1889, 2000.

[7] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361, 1977.

[8] Daniel T. Gillespie. The chemical Langevin equation. *J. Chem. Phys.*, 113(1), 2000.

[9] Daniel T. Gillespie. The chemical Langevin and Fokker-Planck equations for the reversible isomerization reaction. *Journal of Physical Chemistry A*, 106(20):5063–5071, 2002.

[10] Pongor György. OMNeT++: Objective modular network testbed. In *Proc. International Workshop on Modeling, Analysis, and Simulation on Computer and Telecommunication Systems (MASCOTS '93)*, pages 323–326, 1993.

[11] Reinhard Heinrich, S. M. Rapoport, and T. A. Rapoport. Metabolic regulation and mathematical models. *Progress in Biophysics and Molecular Biology*, 32(1):1–82, 1977.

[12] Brian Ingalls. A frequency domain approach to sensitivity analysis of biochemical networks. *J. Phys. Chem. B*, 108(3):1143–1152, 2004.

[13] V. Jacobson. Congestion avoidance and control. In *SIGCOMM'88 Symposium Proc. Comm. Arch. Proto.*, pages 314–329, 1988.

[14] Tobias Jahnke and Wilhelm Huisinga. Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of Mathematical Biology*, 54:1–26, 2007.

[15] J. C. P Jones and S. A. Billings. Describing volterra series, and analysis of non-linear system in the frequency domain. *Int. J. Control*, 53(1):871–887, 1991.

[16] H. Kacser and J. A. Burns. The control of flux. *Symposia of the Society for Experimental Biology*, 27:65–104, 1973.

[17] N. G. Van Kampen. Stochastic processes in physics and chemistry. 2007.

[18] Ahmet Palazoglu Kenneth R. Harris. Studies on the analysis of nonlinear processes via functional expansions–1. solution of nonlinear ODEs. *Journal of Chemical Engineering Science*, 52(18):3183 – 3195, 1997.

[19] Jean-Yves Le Boudec and Patrick Thiran. *Network Calculus: A Theory of Deterministic Queuing Systems for the Internet*, volume 2050 of *Lecture Notes in Computer Science*. Springer, Berlin / Heidelberg, 2001.

[20] Steven H. Low. A duality model of TCP and queue management algorithms. *IEEE/ACM Trans. Net.*, 11(4):525–536, 2003.

[21] Donald Allan McQuarrie. Stochastic approach to chemical kinetics. *Journal of Applied Probability*, 4(3):413–478, 1967.

[22] Thomas Meyer. *On Chemical and Self-Healing Networking Protocols*. Phd thesis, University of Basel, Switzerland, 2010.

[23] Thomas Meyer and Christian Tschudin. Chemical networking protocols. In *Proc. Hotnets VIII*, Oct 2009.

[24] Thomas Meyer and Christian Tschudin. Flow management in packet networks through interacting queues and law-of-mass-action scheduling. Technical Report CS-2011-001, University of Basel, Jan 2011.

[25] Vishal Misra, WeiBo Gong, and Don Towsley. A fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED. *Proc. SIGCOMM*, 30:151–160, 2000.

[26] Christopher V. Rao, Herbert M. Sauro, and Adam P. Arkin. Putting the "control" in metabolic control analysis. In *Proc. DYCOPS 7*, 2004.

[27] Christine Reder. Metabolic control theory: A structural approach. *J. Theor. Biol.*, 135(2):175–201, 1988.

[28] Sanjay Shakkottai and R. Srikant. How good are deterministic fluid models of internet congestion control? In *Proc. INFOCOM*, volume 2, pages 497–505.

[29] András Varga and Rudolf Hornig. An overview of the OMNeT++ simulation environment. In *Proc. SIMUTOOLS*, pages 1–10, 2008.

[30] Peter Waage and Cato M. Guldberg. Studies concerning affinity. *Forhandlinger: Videnskabs - Selskabet i Christiania*, 35, 1864.