

Article

What Is the Trait d'Union between Retroactivity and Molecular Communication Performance Limits?

Francesca Ratti ^{1,2,*} , Maurizio Magarini ²  and Domitilla Del Vecchio ¹ 

¹ Mechanical Engineering Department, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; ddv@mit.edu

² Department of Information, Electronics, and Bioengineering, Politecnico di Milano, 20133 Milan, Italy; maurizio.magarini@polimi.it

* Correspondence: fratti@mit.edu or francesca.ratti@polimi.it

Abstract: Information exchange is a critical process in all communication systems, including biological ones. Retroactivity describes the load that downstream modules apply to their upstream systems in biological circuits. The motivation behind this work is that of integrating retroactivity, a concept proper of biochemical circuits, with the metrics defined in Information Theory and Digital Communications. This paper focuses on studying the impact of retroactivity on different biological signaling system models, which present analogies with well-known telecommunication systems. The mathematical analysis is performed both in the high and low molecular counts regime, by mean of the Chemical Master Equation and the Linear Noise Approximation, respectively. The main goal of this work is to provide analytical tools to maximize the reliable information exchange across different biomolecular circuit models. Results highlight how, in general, retroactivity harms communication performance. This negative effect can be mitigated by adding to the signaling circuit an independent upstream system that connects with the same pool of downstream circuits.

Keywords: molecular communication; mutual information; systems biology; synthetic biology; communication systems; retroactivity



Citation: Ratti, F.; Magarini, M.; Del Vecchio, D. What Is the Trait d'Union between Retroactivity and Molecular Communication Performance Limits? *Molecules* **2022**, *27*, 3130. <https://doi.org/10.3390/molecules27103130>

Academic Editors: Ergin Dinc, Bige Deniz Unluturk, Michael Barros, Sasitharan Balasubramaniam, Kerstin Lenk and Murat Kuscu

Received: 15 March 2022

Accepted: 11 May 2022

Published: 13 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Molecular Communication (MC) is a field of research that has gained relevance in recent years. This emerging discipline is directly inspired by natural communications between cells in biology [1,2]. Characterizing living cells from an information and communication theoretical perspective is one of the keys to understand the fundamentals of MC system engineering [3]. In parallel with the experimental applications of MC [4–6], the development of mathematical models for many of the phenomena occurring in an MC system has gained notice lately. There exist several types of MC channels, e.g., single-cell, multiple-cell, and cell-population [7]. From a modeling perspective, great attention has been paid to the methods of propagation of molecules in the extracellular environment [8–12]. A step forward has been that of studying the communication performance in biochemical circuits [13,14]. In fact, an open research problem in MC is that of finding optimal ways to transfer information in biochemical circuits [15]. Several works have investigated this problem recently. For example, in [16] the authors analyze the channel capacity of a biological system considering both diffusion-based channel and ligand-based receiver. Moreover, in [17] the authors provide a closed-form expression for the information capacity of an MC system with a noisy channel. In [18], a different approach is taken, where the authors use enzymatic reaction cycles to improve the upper bound on the mutual information (MI) for a diffusion-based MC system. The work in [19] presents an analysis of the channel capacity in diffusive MC by considering intersymbol interference from all the previous time slots and the channel transmission probability in each time slot.

Furthermore, there exist several works in the literature investigating parallelisms between well-known telecommunication and MC models and evaluating information exchange performance in different biological scenarios. For example, in [20] the authors focus on the interference that is generated in the case of a broadcast channel, i.e., when the same transmitter (e.g., a cell) sends the same message simultaneously to multiple receivers (e.g., several cells). In [21], the authors present and develop models for the molecular multiple-access, broadcast, and relay channels in a MC system and perform a numerical analysis on their capacity expressions. The authors of [22] study the capacity of a multiple-access channel that is affected by the parameters of the diffusive channel and ligand-receptor binding mechanisms. The work in [23] presents a training-based channel impulse response estimation for diffusive *multiple-input multiple-output (MIMO)* channels. In [24], a MIMO design for MC is proposed, where multiple molecular emitters are used at the transmitter and multiple molecular detectors are used at the receiver. Various diversity techniques for MIMO transmissions based on molecular diffusion are proposed in [25] to improve the communication performance in nanonetworks in the presence of multi-user interference.

The common thread among the aforementioned works is the focus on the communication in the extracellular environment, where propagation of molecules plays a fundamental role in the communication performance. In this paper, we concentrate on biomolecular circuits whose communication is not prominently affected by propagation, and, in particular, we focus on isolating the impact of *retroactivity* on the communication performance. Retroactivity is the effect that downstream systems receiving a signal apply to upstream ones sending the signal. It can be seen as an extension of the concept of impedance of electrical circuits to biomolecular systems because the additional binding/unbinding reaction which constitutes the downstream system competes with the biochemical interactions constituting the upstream system. Thus, it may disrupt the operation of the upstream system [26].

Since its introduction in the literature [26], retroactivity has been studied in the contexts of control theory and systems biology, with the aim of characterizing and efficiently designing modular biomolecular circuits [27–35]. The back-propagated signal generated by the interconnected components plays a role not only in the design of biomolecular circuits but also in the communication performance of the upstream system.

There exist a few works in the literature considering the effects of retroactivity on the information exchange in molecular circuits. The author of [36] presents preliminary results on the impact of retroactivity on the communication performance of a diffusive MC system, composed of one transmitter and one receiver. A review of retroactivity in different signaling systems and genetic circuits can be found in [37].

The main contribution of this paper consists of the analytical investigation of the impact of retroactivity on the information exchange for different MC system models. This is achieved by making parallelisms with some well-known telecommunication ones. The ultimate objective is to lay foundations of how to efficiently maximize the reliable information exchange when considering different signaling circuits. In the following, we illustrate the results we achieved in this direction. Preliminary outcomes were presented in [38].

The paper is organized as follows. Section 2 provides a high-level summary of the work, by introducing its motivations and by presenting a visual scheme resuming the main steps. In Section 3, we introduce the theoretical concepts on which our work lays its foundations, and we contextualize them by explaining the preliminary assumptions we made. Then, in Section 4, we present the biochemical system models on which we perform the communication performance evaluation, by making parallelisms with well-known telecommunication models. In Section 5, we present and discuss the analytical results. Last, in Section 6 we conclude the paper.

2. Overview

The fields of MC and of systems and synthetic biology have as a final goal that of understanding biological phenomena, and that of “engineering” them to improve upon their nature when possible or necessary, e.g., to combat diseases [39,40], or to enhance agricultural processes [41,42]. We asked ourselves if the way a biological system is built can affect its communication performance, and, in particular, *if and how retroactivity affects MC communication performance*. In fact, to the best of our knowledge, this question has not been answered in the literature yet. Awareness regarding the effect of retroactivity on communication performance is relevant both in the context of understanding and maximizing the information flow [43,44] and in that of designing biological circuits [45,46]. The goal of this paper is that of filling this gap.

Figure 1 visually summarizes the main steps we have taken. We consider five biochemical systems, which enable us to quantify the effect of retroactivity in a variety of scenarios. All of them present similarities with most of the common digital communication models. These will be detailed in Section 4.

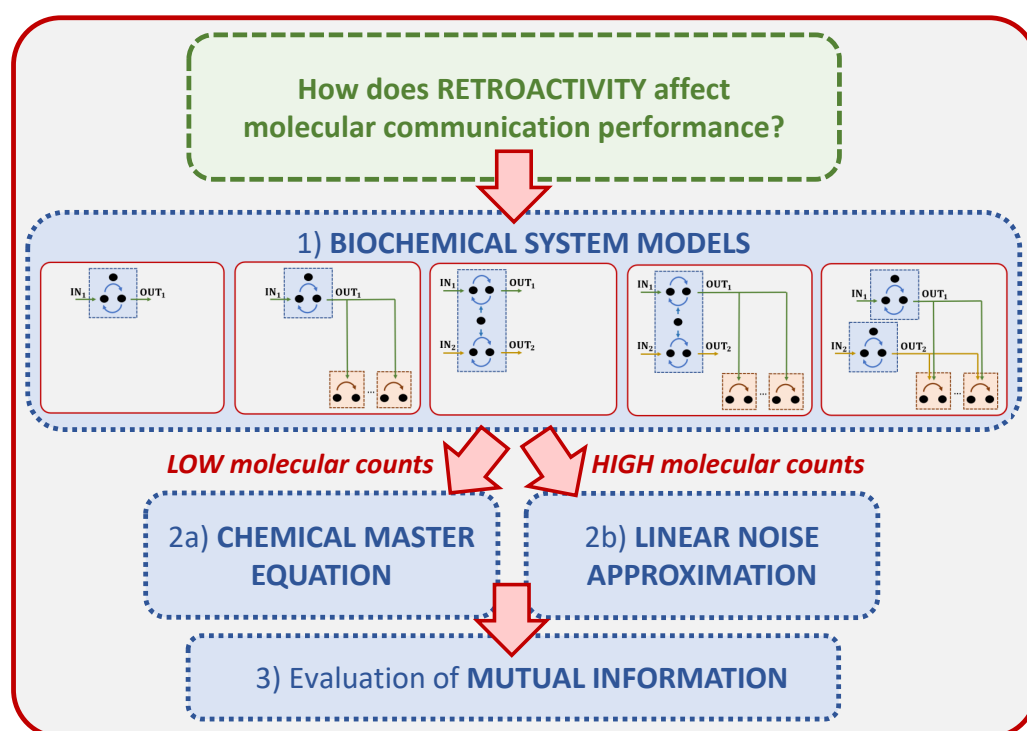


Figure 1. Overview of the structure of the paper. The biochemical system models are detailed later in Figure 3. The Chemical Master Equation (CME) is a system of differential equations that describes the rate of change of the probability of the system to be in any give state at time t [47]. The Linear Noise Approximation (LNA) of the CME is a linear time-varying stochastic differential equation that allows a stochastic characterization of the evolution of a chemical reaction network, still maintaining scalability comparable to that of the deterministic models [48].

We model the stochastic behavior of each biochemical system both in the low molecular counts regime via the Chemical Master Equation (CME) [47], and in the high molecular counts regime via the Linear Noise Approximation (LNA) [48]. Although the CME can describe systems in high molecular count regimes, its analysis in these cases is computationally demanding. Therefore, supposing that the biochemical reaction rates are time-invariant, we opted for using the LNA in the high molecular count regime. In this way, we can compute a tractable analysis for the simpler models, and to make some speculations on the more complex ones. Both the CME and the LNA are solved at *steady state*. In fact, we consider the symbol completely received when the system reaches steady state, so that we avoid the

effects of the transient behavior of the biomolecular circuit on the communication exchange performance, and we are able to isolate the contribution of the retroactivity. Please note that when dealing with non-equilibrium systems, steady state solutions may not be achieved. This would imply the necessity to compute the solution of the CME and of the LNA with respect to time, thus the evaluation of the MI in time. In this work we suppose the steady state solution of the considered systems to exist. The CME and LNA allow determining the probability mass functions (pmf) and the probability density functions (pdf), respectively, necessary for the evaluation of the MI between the input and the output of the system.

3. Preliminaries

The core of this work lays its foundations on concepts coming from different fields, i.e., digital communication, information and control theory, and biology. In this section, we introduce the main theoretical background of this study.

3.1. Retroactivity

Each chemical reaction that, starting from a compound, produces another molecule can be regarded as an input/output system. Such an input/output system can, in turn, be regarded as a module that, when connected with others, forms a more complex system. A fundamental issue that arises when interconnecting different components is how the process of transmitting a signal to a “downstream” module affects the dynamic state of the “upstream” subsystem sending the signal. In fact, on interconnection, a signal that goes from the downstream to the upstream system is generated. This phenomenon is called *retroactivity*, formally defined as the back action from the downstream system to the upstream one [26].

Retroactivity conceptually differentiates from feedback. In fact, *while it is not possible to transmit the output of the upstream system to the downstream one without retroactivity, on the contrary, feedback from the downstream system can be eliminated even when the upstream system is transmitting the signal*. As previously mentioned, while the impact of retroactivity on the design of biomolecular circuit has already been investigated, and solutions to mitigate this effect have been presented, the effects of retroactivity on the communication performance have not been analyzed in depth in the literature.

In our paper, we consider different models of signaling circuits. These are molecular circuits that, given external stimuli (inputs), through a series of chemical reactions, transform them to signals (outputs) that can control how cells respond to their environment. Then, we analyze to what extent retroactivity from downstream systems affects the communication performance between the input and the output of the upstream system.

Figure 2 shows a general scheme of signal exchange between upstream and downstream systems. The blue arrows represent the signal passing through the input of the upstream system to the downstream system. The mauve arrows represent the effect of retroactivity, i.e., the signal that goes from the downstream to the upstream system. We will perform the analysis for both the low and high molecular count regimes. In the figure, the upstream system describes the communication models representing both scenarios, which are detailed in the next sections.

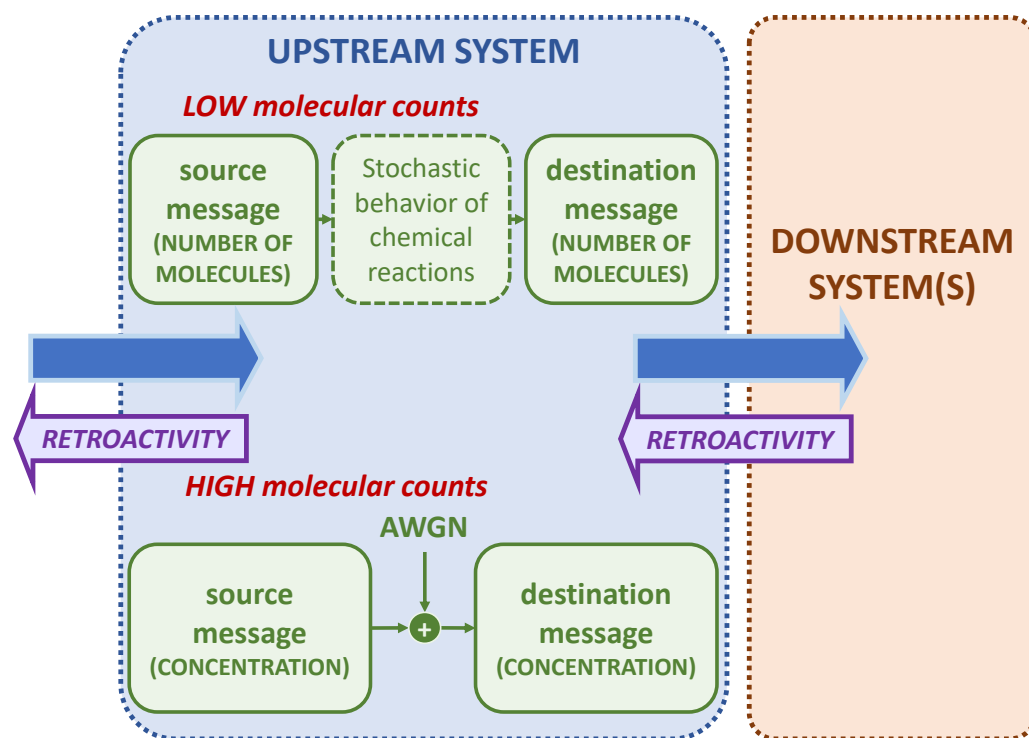


Figure 2. Upstream signaling system in the case of low and high molecular counts. The interconnection between the upstream and the downstream system generates the signal going from right to left in the figure (mauve arrow), that is the effect of retroactivity. The blue arrow from left to right represents the signal transmitted throughout the system.

3.2. Stochastic Models for Biochemical Systems

The evaluation of the MI as difference of entropies [49] requires the knowledge of the marginal pmfs (pdfs in the continuous case) $P(X)$, $P(Y)$, and the conditional pmfs $P(Y | X)$, $P(X | Y)$. Thus, a stochastic modeling to approximate the behavior of the considered systems at steady state is needed. Several modelings have been proposed in the literature [48,50–53]. As anticipated in Section 2, in this paper we model the behavior of the system in the low molecular count regime via the CME [47,51], while we use the LNA [48] under the assumption of high molecular counts.

3.2.1. The Chemical Master Equation

The CME describes the rate of change of the probability of a molecular microstate in a chemical reaction system [47]. For a system with N chemical species, a microstate represents the number of molecules present for each species in the system at a given time. Thus, it requires the enumeration of all the possible microstates of the system (i.e., each possible microscopic configuration). In this work, as previously mentioned, we consider steady state, i.e., we set to zero the rate of change of the probability in the CME.

The motivations behind the choice of the CME to model the behavior of the system are two. First, we consider the accuracy of the results worth the complexity of the calculations. Furthermore, the CME does not take into account any stochastic effects other than the ones due to intrinsic noise in the chemical reactions. In this way, we can selectively quantify the effect of retroactivity on information exchange for different biomolecular system models. The major drawback of the CME is the computational complexity in computing the probability of being in each microstate at time t , which increases exponentially as the number of molecules and reactions in the system increases. For this reason, we use the CME only in the low molecular counts regime, and we rely on the LNA for studying the behavior of systems with a higher number of molecules.

3.2.2. The Linear Noise Approximation

The LNA lays its foundations on the central limit theorem from probability theory [54], which states that, when independent random variables are added, their properly normalized sum tends toward a normal distribution, even if the original variables themselves are not normally distributed. This is why the LNA is valid only in the case of high molecular counts. Thus, when applying the LNA to model the stochastic behavior of a biochemical circuit, the resulting distributions of the species concentrations of the system in time are Gaussian. When using the LNA, the channel noise is assumed Additive White Gaussian (AWGN), with variance $\sigma_{noise}^2 = 1$, hence simplifying the communication model. The mean of the distribution is obtained by the reaction rate equations of the system [47]. In our case, since we work at steady state, these differential equations are solved for the equilibrium. The steady state covariance matrix of the system is computed by solving the Lyapunov Equation [47]. As it can be noticed in the upstream system of Figure 2, in the case of the LNA model, the unit of measure of the biochemical species is concentration rather than number of molecules as in the CME. Thus, the number of molecules in the system is to be divided by the reaction volume.

4. System Model

In this section, we present the chemical reaction models analyzed in the paper. The chemical reactions are divided into subgroups representing the upstream and downstream systems. The addition of complex and realistic phenomena, e.g., particle movement dynamics, intersymbol interference, and memory is out of the scope of this work. Retroactivity is an effect present in all biocommunication systems, and the goal is to analyze and isolate its impact on MI. The construction of a realistic biological scenario implies the introduction of many factors in the evaluation of the MI, making unnecessarily difficult the isolation of the impact of retroactivity in this context.

More specifically, we consider the five signaling system models shown in Figure 3. Figure 3a represents an isolated signaling system, which is not connected to any downstream target. In digital communication, this could be viewed as a *Single-Input Single-Output (SISO)* system, where the input is I_1 , and the consequent output is Z_1 . Then, in Figure 3b, the same isolated system is connected through Z_1 to N chemical circuits. Thus, in Figure 3b, the system of Figure 3a becomes an upstream system directly linked to N downstream ones. This model could be viewed as a *Broadcast channel (BC)*, where the output of the upstream system Z_1 is the same message sent to different receivers, i.e., the downstream systems. The third signaling system (Figure 3c) models the same upstream system in the presence of a second circuit that shares the same enzyme E . We model the second circuit as a replica of the first one. This recalls the *MIMO* setup typical of the telecommunication systems, where the multiple inputs are I_1 and I_2 and the outputs are Z_1 and Z_2 , respectively. The fourth considered signaling system is illustrated in Figure 3d. It is a MIMO model in which the output of the first upstream system Z_1 is connected to N downstream targets. The last model considers two isolated SISO upstream systems that can connect to the same N downstream systems. Please note that only one of the two upstream systems can connect with a specific downstream system at a time. This means that when both upstream systems are sending information, the number of downstream systems available to receive the message Z_1 is $Q \leq N$. This recalls the *Multiple-Access Channel (MAC)* of digital communication, where the two upstream systems are the different transmitters that do not coordinate, and the downstream systems are the receivers that communicate with both the transmitters.

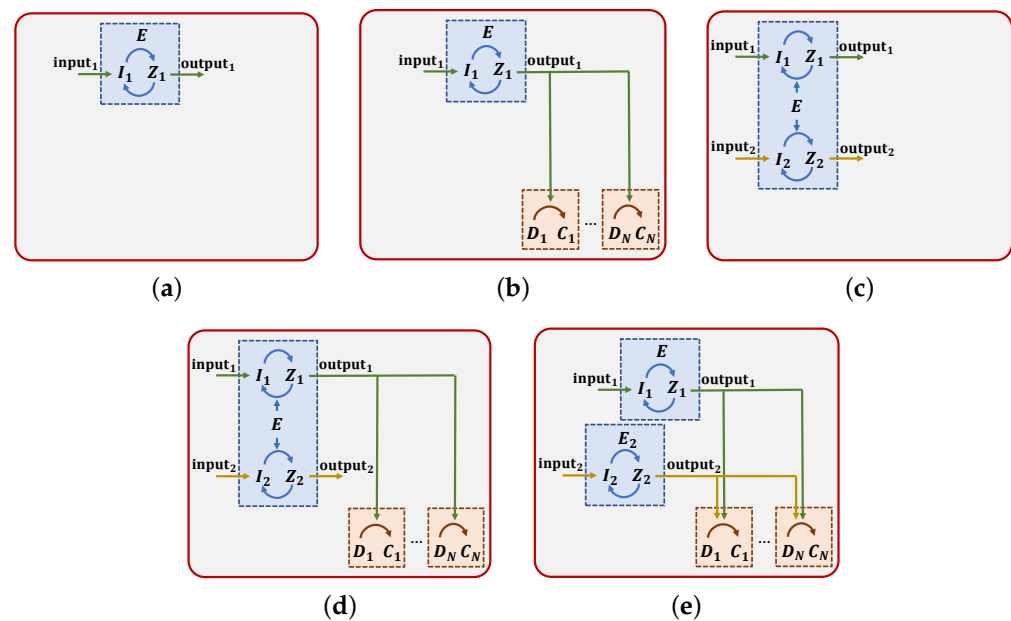
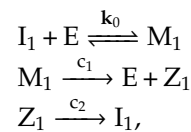


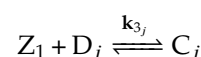
Figure 3. Diagrams of the signaling systems analyzed in the paper: (a) isolated SISO, (b) SISO with N downstream targets, (c) isolated MIMO, (d) MIMO with N downstream targets, (e) two isolated SISO with MAC.

The two-step enzymatic reaction representing the upstream system can be described as



with conservation laws $E_{\text{tot}} = E + M_1$, and $I_{\text{tot}_1} = I_1 + M_1 + Z_1$, where I_1 is the protein (input message) that binds to an enzyme E to form the complex M_1 , that in turn is transformed in the output protein Z_1 [47]. The coefficients c_1 and c_2 are the catalytic rates of the unidirectional chemical reactions, and $k_0 = k_0^-/k_0^+$ is the so-called dissociation constant of the reversible binding reaction, k_0^+ and k_0^- being the association and dissociation rate constants [47]. The last reaction is represented by a one-step rather than a two-step enzymatic reaction to reduce the complexity of the calculations in the following. Furthermore, note that this last reaction makes the system cyclic. This is a major discrepancy with respect to traditional communication models, which are unidirectional. The secondary upstream circuit in the last three models is composed of an analogous two-step enzymatic reaction, with input I_2 and output Z_2 , with conservation law $I_{\text{tot}_2} = I_2 + M_2 + Z_2$. Thus, the conservation law that preserves the total amount of enzyme present in the system becomes $E_{\text{tot}} = E + M_1 + M_2$ for the models in Figure 3c,d. The model in Figure 3e is characterized by two separate conservation laws that preserve the quantity of the enzymes E and E_2 , that are, respectively, $E_{\text{tot}_1} = E + M_1$ and $E_{\text{tot}_2} = E_2 + M_2$.

Each j th downstream system is composed of one reversible reaction



with conservation law $D_{\text{tot}_j} = D_j + C_j$. Here, the output of the upstream system Z_1 (and also Z_2 in Figure 3e) binds with the DNA D to form the complex C . Please note that for the model in Figure 3e, the conservation law becomes $D_{\text{tot}_j} = D_j + C_{j1} + C_{j2}$, where C_{j1} is the complex formed by the binding of Z_1 with D and C_{j2} the one coming from the binding of Z_2 with D . The coefficient k_{3j} is the dissociation constant of the reversible binding reactions and is equal to k_{3j}^-/k_{3j}^+ , where k_{3j}^+ and k_{3j}^- are the association and dissociation rate constants,

respectively. The presence of the downstream systems modifies the conservation laws of $I_{\text{tot}1}$ and $I_{\text{tot}2}$. These become $I_{\text{tot}1} = I_1 + M_1 + Z_1 + \sum_{j=1}^Q C_j$ and $I_{\text{tot}2} = I_2 + M_2 + Z_2 + \sum_{j=1}^{N-Q} C_j$, where Q is the number of upstream systems connected to the first upstream system, that in the second and in the fourth model is equal to N , while in the last one we have $0 \leq Q \leq N$. Thus, supposing no M_1 , M_2 , Z_1 , Z_2 , C_j present in the system before t_0 , we can write $I_{\text{tot}1}(t_0) = I_1(t_0)$ and $I_{\text{tot}2}(t_0) = I_2(t_0)$.

5. Results: Communication Performance Evaluation

In this section, we present the results obtained supposing a low and a high molecular count regime. Please note that due to the complexity of the calculations, not for all the case studies we determine an analytical formula of the MI. Nevertheless, we discuss whether retroactivity causes a reduction of the communication performance in all the considered scenarios.

5.1. Low Molecular Counts

For all the signaling system models, we make the same hypotheses on the values of E_{tot} , $D_{\text{tot}j}$ and on the number of input symbols n_{I_1} , n_{I_2} , and their corresponding $I_1(t_0)$, $I_2(t_0)$ to compute the steady state solution of the CME. We choose $n_{I_1} = n_{I_2} = 2$. This means that both the upstream systems have two available input symbols, one composed by 0 molecules of I ($I(t_0) = 0$, no transmission) while the second composed by one molecule of I ($I(t_0) = 1$). This corresponds to a Concentration On-Off Keying modulation [55]. This choice allows the explicit enumeration of all the microstates of the system, thus the analytical investigation of the role of retroactivity in the information exchange between $I_1(t_0)$ and $Z_1(t_s)$, where t_s stands for steady state time. Accordingly, we also set $E_{\text{tot}}, D_{\text{tot}j} = 1$.

By solving the CME to obtain the necessary pmfs and by substituting X with $I_1(t_0)$ and Y with $Z_1(t_s)$, it is possible to show (see the derivations in the Appendix A) that the formula of the MI we obtain as difference of entropies H in the five considered cases is the same, corresponding to

$$\begin{aligned} I(I_1(t_0), Z_1(t_s)) &= H(I_1(t_0)) - H(I_1(t_0) | Z_1(t_s)) \\ &= \log_e \left(P_{0_1}^{(-P_{0_1})} P_{1_1}^{(-P_{1_1})} \right) - \log_e \left(\left(\left(\frac{P_{0_1}}{P_{0_1} + AP_{1_1}} \right)^{(-P_{0_1})} \right) \cdot \left(\left(\frac{AP_{1_1}}{P_{0_1} + AP_{1_1}} \right)^{(-AP_{1_1})} \right) \right) \end{aligned} \quad (1)$$

nat/symbol, where P_{0_1} and P_{1_1} are the probabilities of the two possible input symbols $I_1(t_0) = 0$ and $I_1(t_0) = 1$. The term A is a constant depending on the coefficients of the chemical reactions of the upstream and downstream systems, and on the probabilities P_{0_2} and P_{1_2} of the input symbols of the second upstream system in the two MIMO models.

Isolated SISO. In the first case, when the model is composed by a single isolated SISO signaling system (Figure 3a), we obtain

$$A = A_0 = \frac{(1 + \mathbf{k}_0)c_2}{c_1 + (1 + \mathbf{k}_0)c_2}, \quad (2)$$

where $\mathbf{k}_0 = (k_0^- \Omega) / k_0^+$, Ω being the volume in which the reactions take place, as detailed in the Appendix A. This shows that the probability of the output $Z_1(t_s)$ is affected by the presence of the reaction $Z_1 \xrightarrow{c_2} I_1$ that transforms the output protein Z_1 in the input I_1 , not allowing the complete consumption of I_1 at steady state.

SISO with N downstream targets. The term A in the second system (Figure 3b) captures the impact of retroactivity on the information exchange due to N downstream systems. In this case, the formula of A is affected by the coefficients of the reactions connecting the upstream and the downstream systems. We obtain

$$A = A_N = 1 - \frac{c_1}{(1 + \mathbf{k}_0)c_2 + \left(1 + \sum_{j=1}^N (1/k_{3_j}) \right) c_1}, \quad (3)$$

where $\mathbf{k}_{3_j} = (\mathbf{k}_{3_j}^- \Omega) / \mathbf{k}_{3_j}^+$. Please note that if we remove the downstream systems effect $\sum_{j=1}^N (1/\mathbf{k}_{3_j})$, we return to A_0 . This system captures the impact of retroactivity on the information exchange performance in a generic biomolecular signaling BC. If we assume that all the N downstream systems have the same dissociation constant \mathbf{k}_3 , we can rewrite $A_N = 1 - \frac{c_1}{(1+\mathbf{k}_0)c_2 + (1+N/\mathbf{k}_3)c_1}$. We observe that, for $N \rightarrow \infty$, A_N is equal to 1. By substituting $A = 1$ in (1), we obtain that $I(I_1(t_0), Z_1(t_s)) = 0$, i.e., $I_1(t_0)$, $Z_1(t_s)$ are completely independent. This is a coherent result if we note that, for infinite downstream systems, the probability of the number of free molecules of $Z_1(t_s)$ being different from 0 tends to 0, independently of the value of $I_1(t_0)$.

We also note that, for $c_2 = 0$, i.e., if the upstream system is not cyclic, A_0 becomes 0, leading to $I(I_1(t_0), Z_1(t_s)) = \log_e \left(P_{0_1}^{(-P_{0_1})} P_{1_1}^{(-P_{1_1})} \right)$. This corresponds to having the MI equal to the entropy of $I_1(t_0)$. Thus, in this case $I_1(t_0)$ and $Z_1(t_s)$ are fully dependent. On the contrary, removing the cycle in the upstream system results in $A_N = 1 - \frac{c_1}{(1+\sum_{j=1}^N (1/\mathbf{k}_{3_j}))c_1}$. This is because the load given by the connected downstream systems still has an impact on the information exchange between I_1 and Z_1 by subtracting free Z_1 (i.e., the output message) from the system environment. If $c_2 \rightarrow \infty$ we obtain $A = A_0 = A_N = 1$, thus the MI becomes 0, meaning that the input and the output are completely independent. In fact, for $c_2 \rightarrow \infty$, the produced Z_1 is immediately transformed into I_1 , so that the number of free molecules $Z_1(t_s)$ is equal to 0 for all values of $I_1(t_0)$. In general, the higher the A , the lower the communication exchange performance measured via (1).

Isolated MIMO. The formula of A for the third model (Figure 3c) is influenced by the presence of the second upstream system

$$A = A_{0, \text{MIMO}} = A_0 P_{0_2} + B P_{1_2}, \quad (4)$$

where A_0 comes from (2), and B is a constant depending on the chemical reaction rates as follows. Supposing all equal catalytic rates $c_1 = c_2 = c$ and supposing binding/unbinding rates $\mathbf{k}_0^+, \mathbf{k}_0^- \gg c$, we can write B as

$$B = \frac{3(\mathbf{k}_0^+)^2 + (\mathbf{k}_0^-)^2 + 4\mathbf{k}_0^- \mathbf{k}_0^+ + 3\mathbf{k}_0^- c + 6\mathbf{k}_0^+ c}{5(\mathbf{k}_0^+)^2 + (\mathbf{k}_0^-)^2 + 5\mathbf{k}_0^- \mathbf{k}_0^+ + 3\mathbf{k}_0^- c + 8\mathbf{k}_0^+ c}, \quad (5)$$

where $\mathbf{k}_0^+ = \mathbf{k}_0^+ / \Omega$ (see Appendix A). If $P_{1_2} = 0$ and by consequence $P_{0_2} = 1$, we obtain $A_{0, \text{MIMO}} = A_0$, thus we return to the single isolated SISO model. We demonstrated via Wolfram Mathematica[®] software that $A_0 P_{0_2} + B P_{1_2} \geq A_0$ is always true, for every value of P_{0_2} and P_{1_2} , with the constraint that their sum should always be 1 and they cannot assume negative values. Thus, since $A_{0, \text{MIMO}}$ is always greater than A_0 , the MI in (1) is always lower in the case of isolated MIMO rather than in the case of isolated SISO. As in digital communication, also in MC, the communication performance of a MIMO system is worse than the ones of a SISO system given the same boundary conditions.

MIMO with N downstream targets. Likewise, the formula of A for the fourth system (Figure 3d) becomes

$$A = A_{N, \text{MIMO}} = A_N P_{0_2} + G P_{1_2} \quad (6)$$

where A_N is (3) and G depends on the rates of the chemical reactions. The analytical formula of G cannot be easily expressed in closed form, although it is worth noticing that it depends not only on the reaction rates of the upstream, but also on the binding/unbinding rates of the downstream systems. This is a key difference with respect to B (5). Under some assumptions, i.e., all equal (un)binding rates $\mathbf{k}^+, \mathbf{k}^-$, all equal catalytic rates c , and $\mathbf{k}^+, \mathbf{k}^- \gg c$, we obtain via Wolfram Mathematica[®] software $G > B$ and $A_N > A_0$. From these, it follows $A_{N, \text{MIMO}} > A_{0, \text{MIMO}}$ (4), i.e., when the output Z_1 of the MIMO upstream system is connected with N downstream systems, the MI lowers.

In this case, when $P_{1_2} = 0$, we return to the second model, i.e., the SISO upstream system and BC. Notwithstanding this, we would like to understand if $A_{N, \text{MIMO}}$ is always greater than A_N , i.e., if the presence of a second upstream system lowers the MI. To do that, we set the inequality $A_{N, \text{MIMO}} \geq A_N$, and we solve it with respect to G . Specifically

$$\begin{aligned} A_N P_{0_2} + G P_{1_2} &\geq A_N \\ \Rightarrow A_N (P_{0_2} - 1) - G A_N (P_{0_2} - 1) &\geq 0 \\ \Rightarrow G &\geq A_N. \end{aligned} \quad (7)$$

From (7), we observe that it would be theoretically possible that for some values of G (i.e., for $G < A_N$), $I(I_1(t_0), Z_1(t_s))$ of the system in Figure 3d becomes higher than the one in Figure 3b. This would imply that a MIMO upstream system binding with the same number N of downstream as the single SISO upstream can mitigate the impact of retroactivity on $I(I_1(t_0), Z_1(t_s))$.

Two isolated SISO with MAC. The expression of A in the model of Figure 3e is noteworthy. In fact, although there are two upstream systems in the environment, they are completely independent, meaning that the presence of the second does not have any influence on the communication performance of the first one. Since the two systems share the binding with the same N downstream systems, A becomes

$$A = A_{Q, \text{MAC}} = 1 - \frac{c_1}{(1 + \mathbf{k}_0)c_2 + \left(1 + \sum_{j=1}^Q (1/k_{3_j})\right)c_1}. \quad (8)$$

The number of downstream systems binding with the second upstream is always complementary to Q , and it is equal to $N - Q$. Supposing that the two upstream systems emit the same number of molecules, i.e., Z_1, Z_2 , then on average $Q = N/2$. If $Q = 0$, then (8) becomes (2), while in the opposite case, when $Q = N$, then $A_{Q, \text{MAC}} = A_N$ (3). From this, we note that the fifth model represents an improvement in relation to the communication performance with respect to the model of Figure 3b. In fact, since the second upstream system binds with some of the available downstream systems, the retroactivity observed at the first upstream system is lowered.

Study of A term via Z-channel capacity and capacity bounds. Please note that regardless of the specific signaling system, each of these cases can be traced back to a Z-channel. In fact, the joint pmf $P(I_1(t_0), Z_1(t_s))$ is for all the five models

$$P(I_1(t_0), Z_1(t_s)) = \begin{bmatrix} P_{0_1} & 0 \\ A P_{1_1} & (1 - A) P_{1_1} \end{bmatrix}, \quad (9)$$

where the rows represent the 0 and 1 possible values of $I_1(t_0)$ and the columns that of $Z_1(t_s)$ (see Appendix A for the exact derivation). From (9), it is clear that if $I_1(t_0) = 0$, then $Z_1(t_s)$ is certainly equal to 0, while if $I_1(t_0) = 1$, then $Z_1(t_s)$ is 1 with a certain probability $(1 - A)$ and 0 with probability A .

For this reason, we validate our MI Formula (1), by plotting the Z-channel capacity [56], where the conditional probability $P(X = 1 | Y = 0)$ in our case is to be substituted with A , and by varying the values of A and of P_{0_1} . Please note that since both A and P_{0_1} are probabilities, the range of valid values that they can assume is included between 0 and 1. Furthermore, we plot on the same graph the MI we obtained (1). By doing this, as it can be observed from the graph, we are also finding a lower bound on the channel capacity for each (A, P_{0_1}) and, for each A , what is the value of P_{0_1} that maximize it, thus the value of the channel capacity. We plot also the upper bound on the channel capacity starting from the definition of MI as Kullback–Leibler divergence as described in [57] in 3D space. In this way, we note that the minimum of the upper bound, the Z-channel capacity, and the maximum of the lower bound for each value of A coincide (red line). Figure 4a is the visualization of what has just been described. Thanks to this 3D plotting, we are able to

isolate the channel capacity value for each value of A (Figure 4b), that, as for each value of the MI, decreases as A increases, confirming our theoretical results.

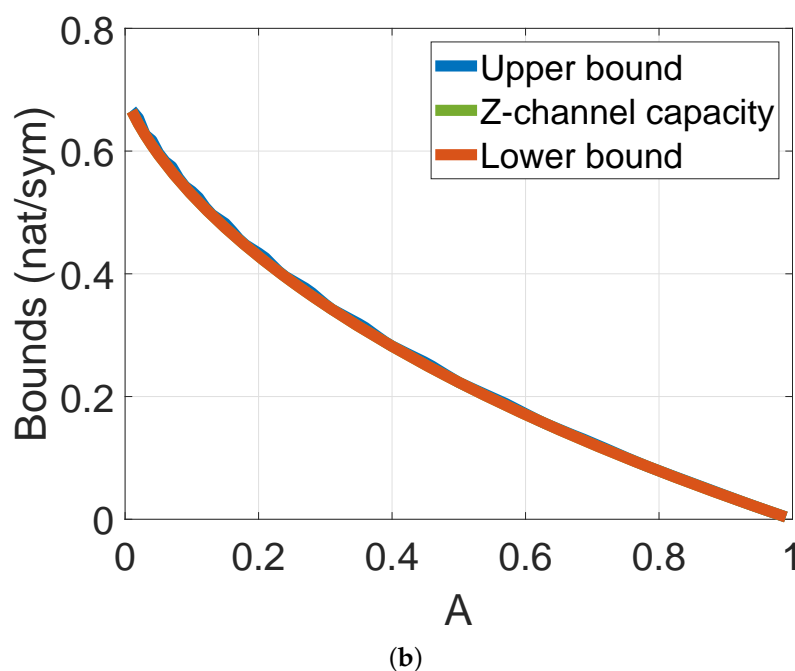
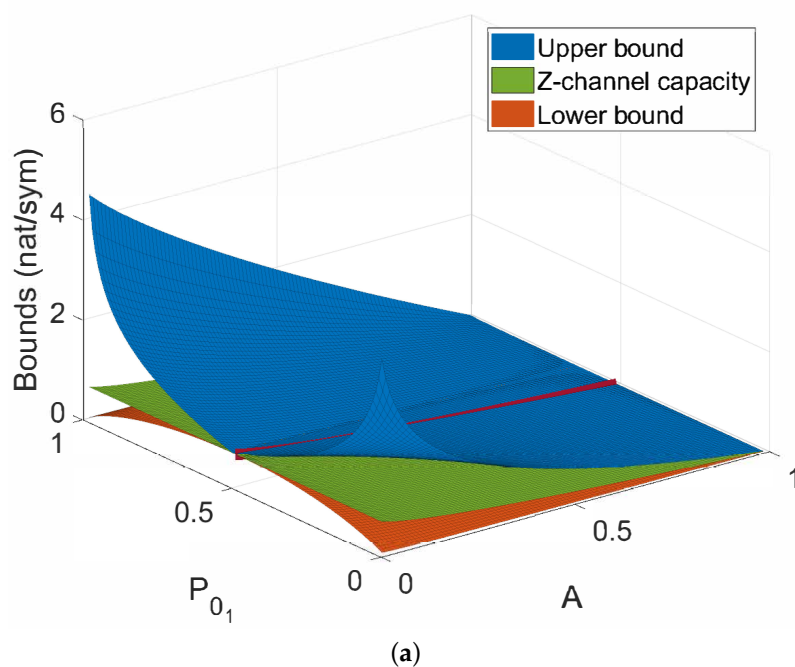
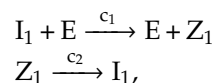


Figure 4. (a) Upper, lower bounds on the channel capacity, and Z-channel capacity for the four signaling systems, by varying P_{01} and A . (b) Channel capacity bounds and Z-Channel capacity with respect to the valid range of A values.

5.2. High Molecular Counts

As mentioned in Section 3.2, the LNA describes the system through an ordinary differential equation and a linear (time-varying) stochastic differential Equation [48]. To

mitigate the analytical complexity, when only a single upstream system is present, we reduce it to a one-step cyclic enzymatic reaction



with conservation law $I_{\text{tot}1} = I_1 + Z_1 + \sum_{j=1}^N C_j$. Thus, in the case of isolated SISO system, the differential equation characterizing the variation of the mean of Z_1 with time can be written as

$$\frac{dZ_1}{dt} = c_1 E (I_{\text{tot}1} - Z_1) - c_2 Z_1, \quad (10)$$

where $\sum_{j=1}^N C_j$ is equal to 0 (no downstream systems). Then, the mean of the Gaussian distribution characterizing $Z_1(t_s)$ can be found easily by setting (10) equal to 0 and by solving the resulting equation with respect to Z_1

$$\mu_0 = \frac{c_1 E I_{\text{tot}1}}{c_2 + c_1 E}. \quad (11)$$

The variance σ_0^2 of the same distribution is obtained by solving the Lyapunov equation

$$\begin{aligned} 2(-c_1 E - c_2)\sigma_0^2 &= -\frac{1}{\Omega} \left(c_1 E \left(I_{\text{tot}1} - \frac{c_1 E I_{\text{tot}1}}{c_2 + c_1 E} \right) + c_2 \frac{c_1 E I_{\text{tot}1}}{c_2 + c_1 E} \right) \\ \Rightarrow \sigma_0^2 &= \frac{c_1 E I_{\text{tot}1} c_2}{\Omega (c_1 E + c_2)^2}, \end{aligned} \quad (12)$$

where Ω is the volume of the environment (i.e., usually a cell). Then, remembering that the LNA supposes $\sigma_{\text{noise}}^2 = 1$, it is straightforward deriving the channel capacity by substituting (12) in its well-known formula [49].

Supposing equal dissociation constant k_3 for all the downstream systems, their characterizing differential equation is

$$\frac{dC_j}{dt} = k_3^+ Z_1 D_j - k_3^- C_j = k_3^+ Z_1 (D_{\text{tot}j} - C_j) - k_3^- C_j, \quad (13)$$

where Z_1 is to be substituted with Z_2 when the second upstream system connects with the j th downstream one. Thus, the steady state value of C_j when connected to the first upstream system is

$$C_j = \frac{Z_1 D_{\text{tot}j}}{k_3 + Z_1}. \quad (14)$$

Then, for the second model (SISO with N downstream systems, Figure 3b) we obtain

$$\frac{dZ_1}{dt} = c_1 E \left(I_{\text{tot}1} - Z_1 - \sum_{j=1}^N C_j \right) + k_3^- \sum_{j=1}^N C_j - c_2 Z_1 - k_3^+ Z_1 \sum_{j=1}^N (D_{\text{tot}j} - C_j). \quad (15)$$

To obtain the mean value of $Z(t_s)$ distribution, we should substitute (14) in (15). Then, supposing that the N downstream systems have equal conservation law $D_{\text{tot}} = D + C$, we obtain $\mu_N = \mu_0$. This means that in the high molecular count regime the mean of the $Z(t_s)$ distribution is not impacted by the presence of the downstream systems. Due to the higher computational complexity, we do not provide the formula of the variance σ_N^2 in this case. However, to determine it one would have to solve the Lyapunov equation for this system with respect to σ_N^2 .

A two-step enzymatic reaction for the upstream systems is intrinsic of MIMO models. Thus, we reuse the chemical reactions modeling of the low molecular counts regime. For this reason, the system of differential equations characterizing the third model (Figure 3c) is

$$\begin{cases} \frac{dZ_1}{dt} = c_1M_1 - c_2Z_1 \\ \frac{dM_1}{dt} = k_0^+EI_1 - (k_0^- + c_1)M_1 \\ \frac{dZ_2}{dt} = c_1M_2 - c_2Z_2 \\ \frac{dM_2}{dt} = k_0^+EI_2 - (k_0^- + c_1)M_2. \end{cases} \quad (16)$$

From the conservation laws, $I_1 = I_{\text{tot}1} - M_1 - Z_1$, $I_2 = I_{\text{tot}2} - M_2 - Z_2$, and $E = E_{\text{tot}} - M_1 - M_2$. Similarly, the system in Figure 3d can be written as

$$\begin{cases} \frac{dZ_1}{dt} = c_1M_1 + k_3^- \sum_{j=1}^N C_j - Z_1 \left(c_2 + k_3^+ \sum_{j=1}^N (D_{\text{tot}j} - C_j) \right) \\ \frac{dM_1}{dt} = k_0^+ (E_{\text{tot}} - M_1 - M_2) \left(I_{\text{tot}1} - M_1 - Z_1 - \sum_{j=1}^N C_j \right) - (k_0^- + c_1)M_1 \\ \frac{dZ_2}{dt} = c_1M_2 - c_2Z_2 \\ \frac{dM_2}{dt} = k_0^+ (E_{\text{tot}} - M_1 - M_2) (I_{\text{tot}2} - M_2 - Z_2) - (k_0^- + c_1)M_2, \end{cases} \quad (17)$$

where C_j is to be substituted with (14) in the first two differential equations.

Both (16) and (17), after being set equal to 0, are to be solved with respect to Z_1 to obtain the mean value of the Gaussian-distributed $Z(t_s)$. We do not report here these results and the Lyapunov equation that would give the variance in these cases, due to their analytical complexity.

The differential equation characterizing Z_1 of the fifth system (Figure 3e) is analogous to (15), considering the number of downstream systems $0 \leq Q \leq N$ and remembering $D_{\text{tot}} = D + C_{j_1} + C_{j_2}$. The same analytical complexity applies for the calculation of the variance.

6. Conclusions

In this paper, we evaluate the effect of retroactivity on communication performance of different signaling systems. More specifically, we considered five different biomolecular circuits with analogies to some well-known digital communication system models. For each of them, we evaluated the analytical formula of the MI between the input at time 0 (i.e., the beginning of the transmission of a symbol) and the output evaluated at steady state. The steady state assumption implies the absence of intersymbol interference and the mitigation of the transient behavior of the system.

Our results show that retroactivity has a negative impact on the communication performance, by reducing the MI as the number of downstream systems increases. We note also how the effect of retroactivity can be mitigated by the presence of a second independent upstream system, that connects with the same pool of downstream systems. The same behavior occurs for all the five systems both in the high and in the low molecular count regime.

The analytical computation of the Gaussian distributed output in the high molecular count regime for the remaining cases, as well as a biological interpretation of the A constant in the low molecular count regime are left for future work.

Nevertheless, we can make some hypotheses on the analytical results in the high molecular counts. For the SISO with N downstream targets model, given that in the low molecular count regime the MI of this system is lower than the one in the isolated SISO case, and by remembering the AWGN channel capacity formula [49], we can speculate that $\sigma_N^2 = \sigma_0^2 - K$, where K is a positive constant. If true, this would imply a lower channel capacity value than in the isolated SISO scenario. Regarding the two isolated SISO with MAC model, since the number of downstream systems connected to Z_1 is less or equal than

the one of SISO with N downstream targets case, we speculate that in this case the variance $\sigma_{Q, \text{MAC}}^2$ would be greater than σ_N^2 .

Future work will develop in multiple directions. First of all we are interested in linking the effect of retroactivity on cell behavior, in the same line of [58]. Another line of research is that of comparing strategies to quantify the MI and information transmission and observe if the quantification of the impact of retroactivity holds. Furthermore, it would be worth generalizing our method to evaluate the impact of retroactivity in the case of more complex systems, as for example in the case of a set of non linear differential equations.

Author Contributions: Conceptualization, F.R., D.D.V. and M.M.; methodology, F.R. and D.D.V.; formal analysis, F.R.; investigation, F.R. and D.D.V.; writing—original draft preparation, F.R. and D.D.V.; writing—review and editing, F.R., D.D.V. and M.M.; supervision, D.D.V. and M.M.; project administration, D.D.V. and M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank Theodore Wu Grunberg and Yili Qian for their valuable comments. F.R. acknowledges the support of the “Progetto Roberto Rocca” Doctoral Fellowship.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

This appendix is devoted to the derivation of the expression of the MI (1). In all the five considered scenarios, the first step to perform is that of recalling the marginal distribution of the input

$$P(I_1(t_0)) = \begin{bmatrix} P_{0_1} \\ P_{1_1} \end{bmatrix}, \quad (\text{A1})$$

and that of evaluating $P(I_1(t_0) | Z_1(t_s))$. To do this, it is necessary to solve the CME. The first step consists of enumerating all the possible microstates of the biomolecular system. Clearly, for the same number of molecules, as the number of species present in the system increases, also the number of possible microstates becomes higher. In our work, we consider $n_{I_1} = n_{I_2} = 2$, thus $I_{\text{tot}_1} \leq 1$ and $I_{\text{tot}_2} \leq 1$. Furthermore, we set $E_{\text{tot}} = D_{\text{tot}_j} = 1$. The following Tables A1 and A2 resume all the possible microstates q_i for each considered system when $I_{\text{tot}_1} = 0$, $I_{\text{tot}_2} = 1$ and $I_{\text{tot}_1} = 1$, $I_{\text{tot}_2} = 1$, respectively. Please note that we consider $I_{\text{tot}_2} = 1$ because we are interested in the value of the MI in the presence of the molecule emitted by the second upstream system. For simplicity, we suppose to have only one downstream system, i.e., $j = 1$.

Table A1. Microstates of the systems when $I_{\text{tot}_1} = 0$, $I_{\text{tot}_2} = 1$.

System Model	State	M_1	I_1	Z_1	C_1	M_2	I_2	Z_2	C_2	E	E_2	D
Isolated SISO	q_1	0	0	0	0	0	0	0	0	1	0	0
SISO + downst.	q_1	0	0	0	0	0	0	0	0	1	0	1
Isolated MIMO	q_1	0	0	0	0	0	0	1	0	1	0	0
	q_2	0	0	0	0	0	1	0	0	1	0	0
	q_3	0	0	0	0	1	0	0	0	0	0	0
MIMO with 1 downstream target	q_1	0	0	0	0	0	0	1	0	1	0	1
	q_2	0	0	0	0	0	1	0	0	1	0	1
	q_3	0	0	0	0	1	0	0	0	0	0	1
Two isolated SISO with MAC	q_1	0	0	0	0	0	0	0	1	0	1	0
	q_2	0	0	0	0	0	0	1	0	0	1	1
	q_3	0	0	0	0	0	1	0	0	0	1	1
	q_4	0	0	0	0	1	0	0	0	0	0	1

Table A2. Microstates of the systems when $I_{\text{tot}1} = 1$, $I_{\text{tot}2} = 1$.

System Model	State	M_1	I_1	Z_1	C_1	M_2	I_2	Z_2	C_2	E	E_2	D
Isolated SISO	q_1	0	0	1	0	0	0	0	0	1	0	0
	q_2	0	1	0	0	0	0	0	0	1	0	0
	q_3	1	0	0	0	0	0	0	0	0	0	0
SISO with 1 downstream target	q_1	0	0	0	1	0	0	0	0	1	0	0
	q_2	0	0	1	0	0	0	0	0	1	0	1
	q_3	0	1	0	0	0	0	0	0	1	0	1
	q_4	1	0	0	0	0	0	0	0	0	0	1
Isolated MIMO	q_1	0	0	1	0	0	0	1	0	1	0	0
	q_2	0	0	1	0	0	1	0	0	1	0	0
	q_3	0	0	1	0	1	0	0	0	0	0	0
	q_4	0	1	0	0	0	0	1	0	1	0	0
	q_5	0	1	0	0	0	1	0	0	1	0	0
	q_6	0	1	0	0	1	0	0	0	0	0	0
	q_7	1	0	0	0	0	1	0	0	0	0	0
	q_8	1	0	0	0	1	0	0	0	0	0	0
MIMO with 1 downstream target	q_1	0	0	0	1	0	0	1	0	1	0	0
	q_2	0	0	0	1	0	1	0	0	1	0	0
	q_3	0	0	0	1	1	0	0	0	0	0	0
	q_4	0	0	1	0	0	0	1	0	1	0	1
	q_5	0	0	1	0	0	1	0	0	1	0	1
	q_6	0	0	1	0	1	0	0	0	0	0	1
	q_7	0	1	0	0	0	0	1	0	1	0	1
	q_8	0	1	0	0	0	1	0	0	1	0	1
	q_9	0	1	0	0	1	0	0	0	0	0	1
	q_{10}	1	0	0	0	0	0	1	0	0	0	1
	q_{11}	1	0	0	0	0	1	0	0	0	0	1
Two isolated SISO with MAC	q_1	0	0	0	1	0	0	1	0	1	1	0
	q_2	0	0	0	1	0	1	0	0	1	1	0
	q_3	0	0	0	1	1	0	0	0	1	0	0
	q_4	0	0	1	0	0	0	0	1	1	1	0
	q_5	0	0	1	0	0	0	1	0	1	1	1
	q_6	0	0	1	0	0	1	0	0	1	1	1
	q_7	0	0	1	0	1	0	0	0	1	0	1
	q_8	0	1	0	0	0	0	0	1	1	1	0
	q_9	0	1	0	0	0	0	1	0	1	1	1
	q_{10}	0	1	0	0	0	1	0	0	1	1	1
	q_{11}	0	1	0	0	1	0	0	0	1	0	1
	q_{12}	1	0	0	0	0	0	0	1	0	1	0
	q_{13}	1	0	0	0	0	0	1	0	0	1	1
	q_{14}	1	0	0	0	0	1	0	0	0	1	1
	q_{15}	1	0	0	0	1	0	0	0	0	0	1

From now on, we report the mathematical steps only for the isolated SISO system, analogous results are obtained for the remaining models. Both in the case of $I_{\text{tot}1} = 0$ and of $I_{\text{tot}1} = 1$, the set of chemical reactions representing the biomolecular system can be rewritten as a set of transitions between the enumerated microstates, as depicted in Figure A1 for the isolated SISO system in the case of $I_{\text{tot}1} = 1$. Clearly, since the case with $I_{\text{tot}1} = 0$ is characterized only by one microstate, the system remains in q_1 with probability equal to 1 in each instant of time t , i.e., $P(q_1, t) = 1$. When $I_{\text{tot}1} = 1$, and assuming that all reactions

take place in a volume Ω , it is possible to use the propensity functions, i.e., the probability that a given reaction occurs in a sufficiently small amount of time, to obtain

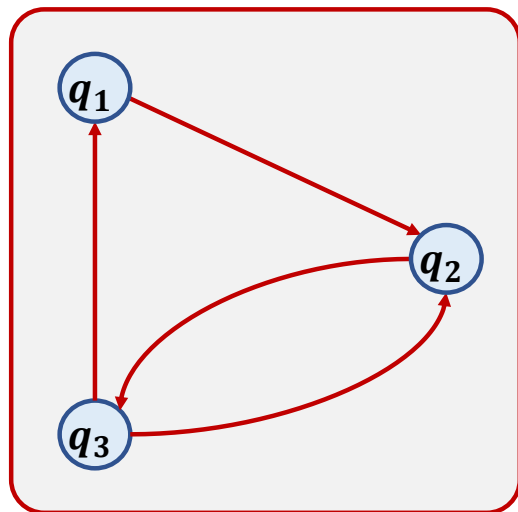
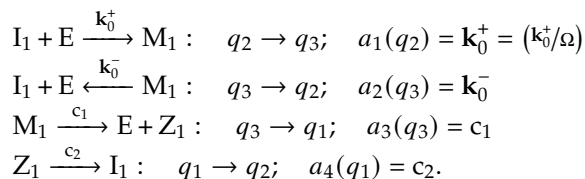


Figure A1. Transitions between microstates for the isolated SISO system model.

Then, the CME can be written down using the propensity functions for each reaction. For the isolated SISO system it becomes

$$\frac{d}{dt} \begin{bmatrix} P(q_1, t) \\ P(q_2, t) \\ P(q_3, t) \end{bmatrix} = \begin{bmatrix} -c_2 & 0 & c_1 \\ c_2 & -\mathbf{k}_0^+ & \mathbf{k}_0^- \\ 0 & \mathbf{k}_0^+ & -\mathbf{k}_0^- - c_1 \end{bmatrix} \begin{bmatrix} P(q_1, t) \\ P(q_2, t) \\ P(q_3, t) \end{bmatrix}. \tag{A2}$$

In turn, the steady state solution for the probabilities $P(q_i, t_s)$ can be solved by setting $\dot{P} = 0$, which yields

$$\begin{aligned}
 P(q_1, t_s) &= \frac{\mathbf{k}_0^+ c_1}{\mathbf{k}_0^+ c_1 + c_2 (\mathbf{k}_0^- + \mathbf{k}_0^+ + c_1)} \\
 P(q_2, t_s) &= \frac{c_2 (\mathbf{k}_0^- + c_1)}{\mathbf{k}_0^+ c_1 + c_2 (\mathbf{k}_0^- + \mathbf{k}_0^+ + c_1)} \\
 P(q_3, t_s) &= \frac{\mathbf{k}_0^+ c_2}{\mathbf{k}_0^+ c_1 + c_2 (\mathbf{k}_0^- + \mathbf{k}_0^+ + c_1)}.
 \end{aligned} \tag{A3}$$

Note that these probabilities can be seen as the joint pmf $P(I_1(t_s), Z_1(t_s))$, where

$$\begin{aligned}
 P(q_1, t_s) &= P(I_1(t_s) = 0, Z_1(t_s) = 1), \\
 P(q_2, t_s) &= P(I_1(t_s) = 1, Z_1(t_s) = 0), \\
 P(q_3, t_s) &= P(I_1(t_s) = 0, Z_1(t_s) = 0).
 \end{aligned} \tag{A4}$$

Thus, if we multiply $P(q_1, t) = 1$ by P_{0_1} and (A4) by P_{1_1} , we easily obtain the joint probability $P(I_1(t_s), Z_1(t_s), I_1(t_0))$. By marginalizing over $I_1(t_s)$, we finally obtain

$$P(I_1(t_0), Z_1(t_s)) = \begin{bmatrix} P_{0_1} & 0 \\ \frac{c_2(k_0^- + k_0^+ + c_1)}{k_0^+ c_1 + c_2(k_0^- + k_0^+ + c_1)} P_{1_1} & \left(1 - \frac{c_2(k_0^- + k_0^+ + c_1)}{k_0^+ c_1 + c_2(k_0^- + k_0^+ + c_1)}\right) P_{1_1} \end{bmatrix}, \quad (\text{A5})$$

where the rows represent the possible values assumed by $I_1(t_0)$, i.e., 0 and 1, and the columns the ones assumed by $Z_1(t_s)$.

By remembering that the dissociation constant is defined as $k_0 = k_0^-/k_0^+$, we introduce $k_0 = k_0^-/k_0^+$. Furthermore, we hypothesize $c_1, c_2 \ll k_0^-, k_0^+$. Then, we rewrite (A5) as

$$P(I_1(t_0), Z_1(t_s)) = \begin{bmatrix} P_{0_1} & 0 \\ \frac{(1+k_0)c_2}{c_1+(1+k_0)c_2} P_{1_1} & \left(1 - \frac{(1+k_0)c_2}{c_1+(1+k_0)c_2}\right) P_{1_1} \end{bmatrix}, \quad (\text{A6})$$

where $\frac{(1+k_0)c_2}{c_1+(1+k_0)c_2} = A$.

Having $P(I_1(t_0), Z_1(t_s))$, it is straightforward to obtain $P(Z_1(t_s))$ by marginalizing over $I_1(t_0)$, and then

$$P(I_1(t_0) | Z_1(t_s)) = \frac{P(I_1(t_0), Z_1(t_s))}{P(Z_1(t_s))} = \begin{bmatrix} \frac{P_{0_1}}{P_{0_1} + AP_{1_1}} & 0 \\ \frac{AP_{1_1}}{P_{0_1} + AP_{1_1}} & 1 \end{bmatrix}. \quad (\text{A7})$$

The entropies are now easily evaluated

$$H(I_1(t_0)) = - \sum_{j=0}^1 P(I_{1_j}(t_0)) \log(P(I_{1_j}(t_0))) = \log_e \left(P_{0_1}^{(-P_{0_1})} P_{1_1}^{(-P_{1_1})} \right), \quad (\text{A8})$$

and

$$\begin{aligned} H(I_1(t_0) | Z_1(t_s)) &= - \sum_{i=0}^1 \sum_{j=0}^1 P(Z_{1_i}(t_s)) P(I_{1_j}(t_0) | Z_{1_i}(t_s)) \log_e \left(P(I_{1_j}(t_0) | Z_{1_i}(t_s)) \right) \\ &= \log_e \left(\left(\left(\frac{P_{0_1}}{P_{0_1} + AP_{1_1}} \right)^{(-P_{0_1})} \right) \cdot \left(\left(\frac{AP_{1_1}}{P_{0_1} + AP_{1_1}} \right)^{(-AP_{1_1})} \right) \right). \end{aligned} \quad (\text{A9})$$

In turn, the MI is

$$\begin{aligned} I(I_1(t_0), Z_1(t_s)) &= H(I_1(t_0)) - H(I_1(t_0) | Z_1(t_s)) \\ &= \log_e \left(P_{0_1}^{(-P_{0_1})} P_{1_1}^{(-P_{1_1})} \right) - \log_e \left(\left(\left(\frac{P_{0_1}}{P_{0_1} + AP_{1_1}} \right)^{(-P_{0_1})} \right) \cdot \left(\left(\frac{AP_{1_1}}{P_{0_1} + AP_{1_1}} \right)^{(-AP_{1_1})} \right) \right) \end{aligned} \quad (\text{A10})$$

nat/symbol.

References

1. Akyildiz, I.F.; Brunetti, F.; Blázquez, C. Nanonetworks: A new communication paradigm. *Comput. Netw.* **2008**, *52*, 2260–2279. [[CrossRef](#)]
2. Akan, O.B.; Ramezani, H.; Khan, T.; Abbasi, N.A.; Kuscü, M. Fundamentals of molecular information and communication science. *Proc. IEEE* **2017**, *105*, 306–318. [[CrossRef](#)]
3. Forney, G.D.; Ungerboeck, G. Modulation and coding for linear Gaussian channels. *IEEE Trans. Inf. Theory* **1998**, *44*, 2384–2415. [[CrossRef](#)]
4. Tuccitto, N.; Li-Destri, G.; Messina, G.M.L.; Marletta, G. Reactive messengers for digital molecular communication with variable transmitter–receiver distance. *Phys. Chem. Chem. Phys.* **2018**, *20*, 30312–30320. [[CrossRef](#)]

5. Giannoukos, S.; McGuinness, D.T.; Marshall, A.; Smith, J.; Taylor, S. A chemical alphabet for macromolecular communications. *Anal. Chem.* **2018**, *90*, 7739–7746. [[CrossRef](#)]
6. Llopis-Lorente, A.; Díez, P.; Sánchez, A.; Marcos, M.D.; Sancenón, F.; Martínez-Ruiz, P.; Villalonga, R.; Martínez-Mañez, R. Interactive models of communication at the nanoscale using nanoparticles that talk to one another. *Nat. Commun.* **2017**, *8*, 15511. [[CrossRef](#)]
7. Wada, T.; Hironaka, K.I.; Kuroda, S. Cell-to-cell variability serves as information not noise. *Curr. Opin. Syst. Biol.* **2021**, *27*, 100339. [[CrossRef](#)]
8. Gresho, P.M. Advection-diffusion. In *Von Karman Inst. of Fluid Dynamics Computational Fluid Dynamics*; Brussels CEST: Brussels, Belgium, 1985.
9. Chang, R. *Physical Chemistry for the Biosciences*; University Science Books: Sausalito, CA, USA, 2005.
10. Hundsdorfer, W.; Verwer, J.G. *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013; Volume 33.
11. Llatser, I.; Cabellos-Aparicio, A.; Pierobon, M.; Alarcón, E. Detection techniques for diffusion-based molecular communication. *IEEE J. Sel. Areas Commun.* **2013**, *31*, 726–734. [[CrossRef](#)]
12. Mosayebi, R.; Arjmandi, H.; Gohari, A.; Nasiri-Kenari, M.; Mitra, U. Receivers for diffusion-based molecular communication: Exploiting memory and sampling rate. *IEEE J. Sel. Areas Commun.* **2014**, *32*, 2368–2380. [[CrossRef](#)]
13. Ruhi, N.A.; Bogdan, P. Multiscale modeling of biological communication. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; pp. 1140–1145.
14. Abbaszadeh, M.; Huang, Y.; Thomas, P.J.; Wen, M.; Ji, F.; Guo, W. Kolmogorov Turbulence and Information Dissipation in Molecular Communication. *IEEE Trans. Mol. Biol. Multi-Scale Commun.* **2021**, *7*, 262–270. [[CrossRef](#)]
15. Koo, B.H.; Lee, C.; Pusane, A.E.; Tugcu, T.; Chae, C.B. MIMO Operations in Molecular Communications: Theory, Prototypes, and Open Challenges. *IEEE Commun. Mag.* **2021**, *59*, 98–104. [[CrossRef](#)]
16. Liu, Q.; Yang, K. Channel capacity analysis of a diffusion-based molecular communication system with ligand receptors. *Int. J. Commun. Syst.* **2015**, *28*, 1508–1520. [[CrossRef](#)]
17. Pierobon, M.; Akyildiz, I.F. Capacity of a diffusion-based molecular communication system with channel memory and molecular noise. *IEEE Trans. Inf. Theory* **2013**, *59*, 942–954. [[CrossRef](#)]
18. Awan, H.; Chou, C.T. Improving the capacity of molecular communication using enzymatic reaction cycles. *IEEE Trans. Nanobiosci.* **2017**, *16*, 744–754. [[CrossRef](#)]
19. Cheng, Z.; Zhu, Y.; Chi, K.; Li, Y.; Xia, M. Capacity analysis for diffusive molecular communication with ISI channel. *Nano Commun. Netw.* **2017**, *13*, 43–50. [[CrossRef](#)]
20. Lu, Y.; Higgins, M.D.; Noel, A.; Leeson, M.S.; Chen, Y. The effect of two receivers on broadcast molecular communication systems. *IEEE Trans. Nanobiosci.* **2016**, *15*, 891–900. [[CrossRef](#)]
21. Atakan, B.; Akan, O.B. On molecular multiple-access, broadcast, and relay channels in nanonetworks. In Proceedings of the 3rd International Conference on Bio-Inspired Models of Network, Information and Computing Systems, Hyogo, Japan, 25–28 November 2008; pp. 1–8.
22. Liu, Q.; Yang, K. Multiple-access channel capacity of diffusion and ligand-based molecular communication. In Proceedings of the 16th ACM International Conference on Modeling, Analysis & Simulation of Wireless and Mobile Systems, Barcelona, Spain, 3–8 November 2013; pp. 151–158.
23. Rouzgar, S.M.; Spagnolini, U. Channel estimation for diffusive MIMO molecular communications. In Proceedings of the 2017 European Conference on Networks and Communications (EuCNC), Oulu, Finland, 12–15 June 2017; pp. 1–5.
24. Koo, B.H.; Lee, C.; Yilmaz, H.B.; Farsad, N.; Eckford, A.; Chae, C.B. Molecular MIMO: From theory to prototype. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 600–614. [[CrossRef](#)]
25. Meng, L.S.; Yeh, P.C.; Chen, K.C.; Akyildiz, I.F. MIMO communications based on molecular diffusion. In Proceedings of the 2012 IEEE Global Communications Conference (GLOBECOM), Anaheim, CA, USA, 3–7 December 2012; pp. 5380–5385.
26. Del Vecchio, D.; Ninfa, A.J.; Sontag, E.D. Modular cell biology: Retroactivity and insulation. *Mol. Syst. Biol.* **2008**, *4*, 161. [[CrossRef](#)]
27. Saez-Rodriguez, J.; Gayer, S.; Ginkel, M.; Gilles, E.D. Automatic decomposition of kinetic models of signaling networks minimizing the retroactivity among modules. *Bioinformatics* **2008**, *24*, i213–i219. [[CrossRef](#)]
28. Del Vecchio, D.; Sontag, E.D. Engineering principles in bio-molecular systems: From retroactivity to modularity. In Proceedings of the 2009 European Control Conference (ECC), Budapest, Hungary, 23–26 August 2009; pp. 658–664.
29. Jayanthi, S.; Del Vecchio, D. Retroactivity attenuation in bio-molecular systems based on timescale separation. *IEEE Trans. Autom. Control* **2010**, *56*, 748–761. [[CrossRef](#)]
30. Ossareh, H.R.; Ventura, A.C.; Merajver, S.D.; Del Vecchio, D. Long signaling cascades tend to attenuate retroactivity. *Biophys. J.* **2011**, *100*, 1617–1626. [[CrossRef](#)] [[PubMed](#)]
31. Anderson, J.; Chang, Y.C.; Papachristodoulou, A. Model decomposition and reduction tools for large-scale networks in systems biology. *Automatica* **2011**, *47*, 1165–1174. [[CrossRef](#)]
32. Del Vecchio, D. A control theoretic framework for modular analysis and design of biomolecular networks. *Annu. Rev. Control* **2013**, *37*, 333–345. [[CrossRef](#)]

33. Sivakumar, H.; Hespanha, J.P. Towards modularity in biological networks while avoiding retroactivity. In Proceedings of the 2013 American Control Conference, Washington, DC, USA, 17–19 June 2013; pp. 4550–4556.
34. Pantoja-Hernández, L.; Martínez-García, J.C. Retroactivity in the context of modularly structured biomolecular systems. *Front. Bioeng. Biotechnol.* **2015**, *3*, 85. [[CrossRef](#)]
35. Bhaskaran, S.; Nair, A.S. Effect of retroactivity in the dynamics of mixed feedback loop. In Proceedings of the 2016 International Conference on Emerging Technological Trends (ICETT), Kollam, India, 21–22 October 2016; pp. 1–6.
36. Awan, H. Effect of retroactivity on the performance of molecular communication networks. In Proceedings of the 2016 Australian Communications Theory Workshop (AusCTW), Melbourne, Australia, 20–22 January 2016; pp. 35–40.
37. McBride, C.; Shah, R.; Del Vecchio, D. The effect of loads in molecular communications. *Proc. IEEE* **2019**, *107*, 1369–1386. [[CrossRef](#)]
38. Ratti, F.; Magarini, M.; Del Vecchio, D. The impact of retroactivity on information exchange in molecular communications. In Proceedings of the 7th ACM International Conference on Nanoscale Computing and Communication, Virtual Conference, 23–25 September 2020; pp. 1–2.
39. Chahibi, Y. Molecular communication for drug delivery systems: A survey. *Nano Commun. Netw.* **2017**, *11*, 90–102. [[CrossRef](#)]
40. Veletić, M.; Barros, M.T.; Balasingham, I.; Balasubramaniam, S. A molecular communication model of exosome-mediated brain drug delivery. In Proceedings of the 6th Annual ACM International Conference on Nanoscale Computing and Communication, Dublin, Ireland, 25–27 September 2019; pp. 1–7.
41. Van Bueren, E.L.; Backes, G.; De Vriend, H.; Østergård, H. The role of molecular markers and marker assisted selection in breeding for organic agriculture. *Euphytica* **2010**, *175*, 51–64. [[CrossRef](#)]
42. Farsad, N.; Pan, D.; Goldsmith, A. A novel experimental platform for in-vessel multi-chemical molecular communications. In Proceedings of the GLOBECOM 2017–2017 IEEE Global Communications Conference, Singapore, 4–8 December 2017; pp. 1–6.
43. Kim, Y.A.; Przytycki, J.H.; Wuchty, S.; Przytycka, T.M. Modeling information flow in biological networks. *Phys. Biol.* **2011**, *8*, 035012. [[CrossRef](#)]
44. Kadloor, S.; Adve, R.S.; Eckford, A.W. Molecular communication using Brownian motion with drift. *IEEE Trans. Nanobiosci.* **2012**, *11*, 89–99. [[CrossRef](#)]
45. Alon, U. *An introduction to Systems Biology: Design Principles of Biological Circuits*; CRC Press: Boca Raton, FL, USA, 2019.
46. Grunberg, T.W.; Del Vecchio, D. Modular analysis and design of biological circuits. *Curr. Opin. Biotechnol.* **2020**, *63*, 41–47. [[CrossRef](#)]
47. Del Vecchio, D.; Murray, R.M. *Biomolecular Feedback Systems*; Princeton University Press: Princeton, NJ, USA, 2015.
48. Van Kampen, N.G. *Stochastic Processes in Physics and Chemistry*; Elsevier: Amsterdam, The Netherlands, 1992; Volume 1.
49. Cover, T.M. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 1999.
50. Hougen, O.A.; Watson, K.M. *Chemical Process Principles*; John Wiley & Sons: Hoboken, NJ, USA, 1943; Volume 1.
51. Gillespie, D.T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **1977**, *81*, 2340–2361. [[CrossRef](#)]
52. Risken, H. Fokker-Planck equation. In *The Fokker-Planck Equation*; Springer: Berlin/Heidelberg, Germany, 1996; pp. 63–95.
53. Gillespie, D.T. The chemical Langevin equation. *J. Chem. Phys.* **2000**, *113*, 297–306. [[CrossRef](#)]
54. Montgomery, D.C.; Runger, G.C. *Applied Statistics and Probability for Engineers*; Wiley: Hoboken, NJ, USA, 2014.
55. Kuran, M.S.; Yilmaz, H.B.; Tugcu, T.; Akyildiz, I.F. Modulation techniques for communication via diffusion in nanonetworks. In Proceedings of the 2011 IEEE International Conference on Communications (ICC), Kyoto, Japan, 5–9 June 2011; pp. 1–5.
56. MacKay, D.J. *Information Theory, Inference and Learning Algorithms*; Cambridge University Press: Cambridge, MA, USA, 2003.
57. Csiszar, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*; Cambridge University Press: Cambridge, MA, USA, 2011.
58. Koorehdavoudi, H.; Bogdan, P.; Wei, G.; Marculescu, R.; Zhuang, J.; Carlsen, R.W.; Sitti, M. Multi-fractal characterization of bacterial swimming dynamics: A case study on real and simulated *serratia marcescens*. *Proc. R. Soc. Math. Phys. Eng. Sci.* **2017**, *473*, 20170154. [[CrossRef](#)] [[PubMed](#)]