Strategies for Coexistence in Molecular Communication
Malcolm Egan, Valeria Loscri, Trung Duong, Marco Renzo

To cite this version:
Malcolm Egan, Valeria Loscri, Trung Duong, Marco Renzo. Strategies for Coexistence in Molecular Communication. IEEE Transactions on NanoBioscience, Institute of Electrical and Electronics Engineers, 2019. <hal-01928205>

HAL Id: hal-01928205
https://hal.archives-ouvertes.fr/hal-01928205
Submitted on 30 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Strategies for Coexistence in Molecular Communication

Malcolm Egan, Valeria Loscri, Trung Q. Duong and Marco Di Renzo

Abstract—Some of the most ambitious applications of molecular communications are expected to lie in nanomedicine and advanced manufacturing. In these domains, the molecular communication system is surrounded by a range of biochemical processes, some of which may be sensitive to chemical species used for communication. Under these conditions, the biological system and the molecular communication system impact each other. As such, the problem of coexistence arises, where both the reliability of the molecular communication system and the function of the biological system must be ensured. In this paper, we study this problem with a focus on interactions with biological systems equipped with chemosensing mechanisms, which arises in a large class of biological systems. We motivate the problem by considering chemosensing mechanisms arising in bacteria chemotaxis, a ubiquitous and well-understood class of biological systems. We then propose strategies for a molecular communication system to minimize disruption of biological system equipped with a chemosensing mechanism. This is achieved by exploiting tools from the theory of chemical reaction networks. To investigate the capabilities of our strategies, we obtain fundamental information theoretic limits by establishing a new connection with the problem of covert communications.

Index Terms—Molecular Communications, Coexistence, Chemical Reaction Networks, Covert Communications

I. INTRODUCTION

Molecular communication is well known to play an important role in a range of biological systems, ranging from bacteria colonies [1] to the development of the embryos of the fruit fly Drosophila melanogaster [2], and also in cell-cell communication in humans. Following the seminal paper of Hiyama et al. [3], the last decade has seen the proposal and initial development of artificial molecular communications. Motivated in large part by the possibility of coordinating nanoscale networks in domains including nanomedicine [4], [5] and high efficiency manufacturing, artificial molecular communication aims to develop mechanisms to transfer information via the propagation of molecules. One version of this vision has been called the internet of nanobio things [6], which follows the ideas underlying the internet of things into the nanoscale regime.

The vision of nanoscale networking and the internet of nanobio things remains riddled with many challenges. One fundamental issue is how molecules should propagate from a transmitter to a receiver. Diffusion-based propagation, flow assisted propagation, active transport and bacterial assisted propagation have all been considered [7]. Recent surveys of these approaches, their models and challenges are available in [8], [9].

Based on these channel models, there have also been many studies of receiver structures from a communication point of view; e.g., algorithms for detection of transmitted symbols [10]–[13]. On the transmitter side, coding strategies accounting for strict energy and complexity constraints have been introduced in [14]. Intersymbol interference reduction strategies have been proposed in [15] and synchronization in [16]. Simulation tools have also been developed in [17], [18].

Despite the progress designing communication strategies, the large majority of existing work has focused on the case of an isolated molecular communication system. That is, the molecular communication system is the only biochemical system in the environment. As a consequence, little is known about how a molecular communication system interacts with its environment.

Exceptions examining interactions between a molecular communication system and external biochemical systems have focused on the problem of drug delivery. That is, how an external biological system degrades the quality of the communication channel. For example, in [19] the impact of immune responses on the reliability of drug delivery was studied. A recent survey of drug delivery methods based on molecular communication is available in [20].

In this paper, we adopt an alternative perspective where our focus is instead on the impact of molecular communications on the function of a nearby biochemical system. As biological systems are governed by complex networks of chemical reactions or other biochemical processes, it may be unavoidable that chemical species used to transfer information interact with the biological system. This situation is particularly likely when—as has been widely proposed [21]—the molecular communication system adopts a bio-inspired approach. For example, a Ca,
$^{2+}$-based mechanism may alter balances of electrostatic forces and interfere with the structure of the lipid bilayer, compromising the membrane of a cell [22]. As such, it is critical to understand the impact of the molecular communication system on the function of the biological system.

In our earlier work on coexistence in molecular communications [23], we introduced tools from the theory of chemical reaction networks to understand how molecular communication systems impact the dynamics of chemical systems. In particular, we established how steady states of a class of simple chemical systems are perturbed by the presence of information molecules.
In this paper, we apply chemical reaction networks to study how a molecular communication link impacts more realistic models of biological systems. Here, we consider biological systems that exploit chemosensing mechanisms [24]. Chemosensing provides a means for a cell to observe the presence of chemical species in its environment and adjust internal parameters based on observed concentrations. As such, chemosensing is exploited in a wide range of biological systems.

A well-studied setting of chemosensing arises in the bacteria *Escherichia coli*, which is used for chemotaxis [25]. This mechanism provides a means for each bacterium to move or “tumble” towards attractants such as glucose or away from repellants such as phenol. The first part of this paper describes this mechanism and known features of its underlying chemical kinetics. We then provide a general model of chemosensing systems that we later use to understand the impact of the presence of molecular communications.

Having established a concrete model of a biological chemosensing system, we discuss general strategies for how a molecular communication system can coexist with biological systems. The challenge is to minimize the impact of information molecules (or other chemical species aiding communication) on the concentration of chemical species governing biological processes. Our strategies are based on a new analogy with cognitive radio developed in the context of wireless communications [26].

One such strategy is the underlay strategy, where the molecular communication system chooses a transmission strategy to ensure that changes to the concentration of chemical species inside a biological system do not exceed a given level. To characterize fundamental limits of the underlay strategy, we establishing a novel connection to the problem of covert communications [27].

A. Overview of Contributions

The main goal of this paper is to establish a framework to model molecular communications in the presence of an external biological system, in which both systems coexist. In this work, we account for the fact that most interactions between the molecular communication system and the biological system will be through a chemosensing mechanism.

Our main contributions are summarized as follows:

1) We formalize chemosensing mechanisms in biological systems and their interaction with a molecular communication system. This formalization is based on the theory of chemical reaction networks, which provides methods to establish stability of the biological system and hence conditions under which biological function is preserved. Unlike earlier work on coexistence between molecular communication systems and external biological systems [23], this work accounts for a wider range of biological systems and utilizes a more realistic model.

2) We propose the framework of cognitive molecular communications, which provides a basis to design systems that do not disrupt the function of nearby biological systems. Our framework is based on an analogy with cognitive radio and we discuss different means of implementing the underlying ideas within the context of molecular communications.

3) We study the underlay strategy in detail and show that it leads to a problem closely related to covert communications or communications with a low probability of detection. Under certain conditions on the behavior of the biological system, we show that known fundamental limits for the covert communications problem can be directly applied. As such, we obtain scaling laws for the number of messages that can be sent as the number of channel uses tends to infinity. In particular, the square-root law holds, which implies that the number of bits of information scales with the square-root of the number of channel uses.

B. Organization of the Paper

In Section II, we overview chemosensing in biological systems and how they can be modeled via chemical reaction networks. In Section III, we develop our framework of cognitive molecular communications. In Section IV, we study the underlay strategy and establish a new connection with cognitive radio and we discuss different means of implementing the underlying ideas within the context of molecular communications.

In Section VI, we discuss limitations of the models and the impact of the structure of the chemosensing mechanism. Finally, in Section VII we conclude and outline avenues of future work.

II. CHEMOSENSING IN BIOLOGICAL SYSTEMS

A. Motivating Example

On the surface of the membrane of a cell, proteins—held in place by lipids—provide a means for the cell to interact with its environment. These proteins bind with ligands in the environment, which triggers a sequence of biochemical processes whereby a chemical complex used for signaling is produced. The signaling complex then forms the basis for regulation of internal parameters of the cell.

Perhaps the most well-understood chemosensing mechanism arises in bacteria chemotaxis [25], which is used by bacteria including *Escherichia coli* to move through a fluid. In particular, each bacterium is equipped with a rotary motor. The probability of clockwise motor rotation is governed by a biochemical regulator system based on inputs from a chemosensing mechanism. In the presence of no attractants (e.g., glucose) or repellants (e.g., phenol), the bacterium follows a random walk. However, when attractants or repellants are present, the movement of bacteria is biased towards or away from the attractant or repellent, respectively.

The biochemical mechanism is described as follows [28]. Upon activation by ligand binding of attractants or repellants, the ligand-occupied form of a given binding protein on the surface of the membrane docks to a transmembrane receptor, initiating a signaling event. This leads to up or down regulation of a histidine kinase, which is a class of enzyme. The histidine kinase provides a mechanism to transfer a phosphate group...
from ATP to an aspartate in the active site of a response regulator protein, CheY. An aspartate is an amino acid of a protein (i.e., CheY), and once the phosphate group is transferred to the aspartate, it modifies the protein to its phosphorylated form, phospho-CheY. At this point phospho-CheY diffuses due to thermal fluctuations or a concentration gradient to the rotary motor where it docks and increases the probability of clockwise motor rotation.

A key feature of bacteria chemotaxis is that the steady state level of phospho-CheY serves as a signal that controls the rotor motion. It is modulated by two opposing reactions: creation of phospho-CheY by the receptor-kinase complex; and destruction of phospho-CheY by hydrolysis of its acyl phosphate.

Due to the fact bacteria chemotaxis relies heavily on steady state levels of certain chemical species, the function of the biological system is tightly linked to their steady state concentrations. Therefore it is critical to understand the behavior of steady state concentrations of chemical species arising in biological chemosensing mechanisms. This is achieved by exploiting tools from the theory of chemical reaction networks, which we recall in the following section.

B. Chemical Reaction Network Preliminaries

Chemical reaction networks are mathematical models governing the time evolution of chemical concentrations [29]. To illustrate these concepts, we begin with the ubiquitous class of enzyme-activated chemical reaction networks. These networks consist of four chemical species: the enzyme $E$; the reactant $S$; the complex $ES$; and the product $P$. The set of chemical species in this example is then $S_E = \{E, S, ES, P\}$. In this system there are three reactions:

$$
\begin{align*}
E + S & \xrightarrow{k_1} ES \\
ES & \xrightarrow{k_2} E + S \\
ES & \xrightarrow{k_3} E + P,
\end{align*}
$$

where $k_1, k_2, k_3, \text{ and } k_{\text{cat}}$ are called reaction rate coefficients and control the rate of the chemical reactions.

A convenient way of representing each of these reactions is as a map\(^1\) from $\mathbb{N}^{S_E}$ to $\mathbb{N}^{S_E}$. For example, the first reaction is written as $(1, 1, 0, 0) \rightarrow (0, 0, 1, 0)$. In this way, we can define a set of reactions $R_E = \{y_i \rightarrow y'_i, i = 1, 2, 3\}$, where $y_i \in \mathbb{N}^{S_E}$ is the vector of reactants in reaction $i$ and $y'_i \in \mathbb{N}^{S_E}$ is the vector of products.

The pair of chemical species and chemical reactions $(S_E, R_E)$ is called a chemical reaction network. In order to characterize this network, we also need to consider the kinetics. Let $[E](t), [S](t), [ES](t), [P](t)$ denote the concentration of each chemical species at time $t$. Under the standard assumption of mass-action kinetics, the concentrations of each species in the enzyme-activated system are governed by the following system of ODEs

$$
\begin{align*}
\frac{d[E](t)}{dt} &= -k_1 [E](t)[S](t) + k_{-1}[ES](t) \\
\frac{d[ES](t)}{dt} &= k_1 [E](t)[S](t) - k_{-1}[ES](t) - k_{\text{cat}}[ES](t) \\
\frac{d[P](t)}{dt} &= k_{\text{cat}}[ES](t),
\end{align*}
$$

with initial conditions $[E](0) = E_0$, $[S](0) = S_0$, $[ES](0) = ES_0$, and $[P](0) = P_0$. The biochemical system can then be written as the tuple $(S_E, R_E, k_E)$, where $k_E : \{1, 2, 3\} \rightarrow \{k_1, k_2, k_{\text{cat}}\}$.

We now consider general chemical reaction systems.

**Definition 1.** A chemical reaction system is the tuple $(S, R, k)$ consisting of a set of chemical species $S$, a set of reactions $R = \{y_i \rightarrow y'_i, i = 1, 2, \ldots\}$, and the rate function $k$.

Let $x(t) \in \mathbb{R}^S$ be the vector consisting of concentrations of each chemical species at time $t$. Under mass-action kinetics, the chemical reaction system is then governed by

$$
\dot{x}(t) = \sum_{y \rightarrow y' \in R} k_{y \rightarrow y'} x(t)^y (y' - y),
$$

where $x(t)^y = x_1(t)^{y_1} x_2(t)^{y_2} \cdots$. A key property of chemical reaction networks is their steady state concentration; that is the concentration of each species when $\dot{x}(t) = 0$. In the context of biological systems, the steady state is often tightly linked to function. However, it is important to note that these steady states are not guaranteed to be unique nor even to exist. In fact, a long-standing open problem in the theory of chemical reaction networks is conditions under which the steady state exists and is unique, known as the global attractor theorem [30].

In this paper, we only assume that there exists at least one steady state of the chemical reaction network, which may depend on initial conditions. As such, we denote the steady state as $\Phi_{x(0)}$, where $x(0)$ denotes the initial concentration levels.

C. Chemosensing Chemical Reaction Network Models

Chemical reaction networks provide a basis to formalize the behavior of biological systems exploiting chemosensing. For our purposes, there are three classes of chemical species involved: the species to which receptors are sensitive, denoted by the set $S$; the species produced by the chemosensing mechanism, denoted by the set $C$; and species produced by the adaptation mechanism of interest for a particular biological process, denoted by the set $A$. The behavior of such a biological system is assumed to be governed by the ordinary differential equations in (3).

To illustrate this model in the context of chemosensing, consider the example of bacteria chemotaxis detailed in Section II. The chemosensing-regulation system is called the

---

\(^1\)The notation $\mathbb{N}^{S_E}$ represents the set of functions from $S_E \rightarrow \mathbb{N}$, which map the chemical complexes in $S_E$ to multiplicities. For example, in the reaction $2A + C \rightarrow B$, the multiplicity of $A$ is two.
EnvZ-OmpR system and is described by the following chemical reactions.

\[
X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} XT \overset{k_2}{\rightarrow} X_p
\]

\[
X_p + Y \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} X_pY \overset{k_4}{\rightarrow} X + Y_p
\]

\[
XT + Y_p \underset{k_{-5}}{\overset{k_5}{\rightleftharpoons}} XTY_p \overset{k_6}{\rightarrow} XT + Y,
\]

where \( X \) is the sensor kinase, \( X_p \) is its phosphorylated form, \( Y \) is the response regulator CheY, and \( Y_p \) is phospho-CheY. Therefore, using the notation defined above, \( C = \{X, X_p, XT\} \) and \( A = \{Y, Y_p, XTY_p\} \). The species \( T \) is linked to levels of ATP in the bacterium. Although not explicitly stated in the system (4), the concentration of \( X \) is determined by the concentration of species in \( S \) to which the receptors are sensitive.

As noted earlier, the steady state concentrations are closely linked to function and in the context of chemotaxis, determine the rate that the rotor turns. To this end, denote the vector of concentrations at time \( t \) by \( x(t) \in \mathbb{R}^C \cup A \), which can be obtained by solving the system of ordinary differential equations induced by the reactions in (4). The EnvZ-OmpR system then admits a steady state concentration vector \( \Phi_{x(0)} \), which depends on the initial conditions \( x(0) \) [31]. These initial conditions are determined by the quantity of attractants or repellants in the environment. That is, in order to understand the behavior of each bacterium, it is necessary to characterize the impact of the environment.

More generally, we are concerned with the impact of the environment on the chemosensing mechanism of a biological system. As observed in the bacteria chemotaxis example, the function of a biological system is determined by a small number of chemical species. In particular, the key species in bacteria chemotaxis is phospho-CheY, which controls the rate of the rotor. For general chemosensing mechanisms, we are therefore interested in how the concentration of species in the set \( S \) affects the steady state behavior of these key chemical species in \( A \) via the chemosensing species in the set \( C \). The fundamental question is then: how do we ensure that chemical species in the molecular communication system do not significantly disturb the steady state behavior of the key species via the chemosensing mechanism? Put simply, how can the molecular communication system and the biological system coexist?

III. General Strategies for Coexistence

In this section, we propose three strategies to minimize the impact of molecular communication systems on biological systems, where interactions are mediated via a chemosensing mechanism. The three strategies differ in their impact, complexity and level of cooperation between the communication and biological systems. Each strategy is developed by analogy with a problem in wireless communication systems, known as cognitive radio [26]. For this reason, we call our framework cognitive molecular communications.

To begin, we briefly review the key ideas behind cognitive radio. In cognitive radio, there are primary and secondary wireless communication networks. The primary network has priority and the goal is to design the secondary network’s access protocol to ensure that the reliability of the primary network is not degraded. Cognitive radio strategies are known as [32]: underlay, where the secondary network transmits independently of the primary network under an interference constraint; overlay, where the primary and secondary networks cooperate by sharing their transmission strategy to remove interference; and interweaving, where the secondary network senses the activity of the primary network and only transmits when the primary network is not active.

In cognitive radio there are a number of technical difficulties due to the nature of electromagnetic propagation in wireless communications. However, for our purposes only the basic principles are important. In particular, we seek to exploit a new analogy between the primary wireless network and the biological system, as well as the secondary wireless network and the molecular communication system.

While the molecular communication system and the biochemical system exchange mass due to reactions between chemical species in each system, they also exchange information, either through dedicated reaction pathways or indirectly through observations made locally in the molecular communication system. The basis for the analogy is that a similar situation occurs in cognitive radio, except instead of the exchange of matter, the interaction is through electromagnetic interference.

The first strategy is underlay, where cooperation is not possible but the molecular communication system has knowledge of its impact on the biochemical system. The impact of the molecular communication system is determined by any changes to the kinetics of the biochemical system. More precisely, this impact can be formalized as changes to initial conditions or parameters of the differential equations describing the kinetics of the biochemical system. The underlay strategy is necessary when the information molecules used for the molecular communication link either overlap or react with the species in the biochemical system. In the case where the biochemical system is complex—i.e., containing a very large number of species—this situation is likely to occur and therefore it is important to design the molecular communication link to ensure the concentration of species in the biochemical system are not significantly perturbed.

In the context of biological systems exploiting chemosensing, the underlay strategy requires that the molecular communication systems does not introduce chemical species that result in large changes to the steady state of key chemical species in the biological system. An advantage of this approach is that it does not require high complexity sensing or cooperation mechanisms at the cost of knowledge of the behavior of the biological system.

The second strategy is overlay. In the context of cognitive radio, the overlay strategy is applicable when the secondary transmitting device has knowledge of the codebook and messages of transmissions in the primary network. That is, a high level of side information is available to the secondary network. The analogous setting in molecular communications is when the molecular communication link and the biochemical system are jointly designed.
Overlay communication can be viewed as a cooperative transmission strategy present in biological systems such as bacteria colonies. In Vibrio fischeri or Vibrio harveyi, for example, a dedicated communication link is established between bacteria, which is supported by the production of autoinducer molecules during the DNA transcription process [1]. By encoding molecular communication messages in this way, communication is intimately linked to the function of each bacterium. Crucially, communication does not reduce the ability of the bacteria to find nutrient sources or reproduce as could be the case if the underlay strategy was employed.

The third strategy is interweaving. In cognitive radio, this strategy ensures that the secondary network does not interfere with the primary network by detecting spectrum holes. That is, no side information assumptions are required at the cost of additional signal processing to detect whether there is a primary network transmission on a given band at a given time.

In molecular communications, the analogous situation is when the information molecules either do not react with any species in the biochemical system or, more generally, the steady state concentrations are unaffected by any interactions. In this case the molecular communication link can be viewed as an isolated chemical system, able to coexist with the biochemical system. This is an assumption widely used in existing communication and information theoretic studies, e.g., [33]. Nevertheless, it may be challenging to implement unless the kinetics of the biochemical system are sufficiently slow. Conditions where the interweaving strategy can be applied were obtained in our earlier work in [23].

We remark that an implementation of the interweaving approach would require a sensing strategy to detect which chemical species interact with the biological system. This is likely to be more challenging than in common cognitive radio models due to the fact that the behavior of the biological system is governed by differential equations.

Each of the three cognitive communication strategies have advantages and drawbacks. At present, developing systematic methods to select the appropriate approach remains an open problem. In order to address this challenge, a formalization is required. To begin a study of these strategies, in the following section we focus on the underlay approach by exploiting the chemical reaction network models introduced in Section II.

IV. THE UNDERLAY APPROACH

A. System Model

A key question in order to implement underlay methods is how to characterize the impact of a molecular communication system on a biological system exploiting chemosensing. Based on the discussion in Section II, we model the biological system via the chemical reaction network \((S \cup C \cup A, R, k)\), defined in Section II-C. Motivated by the bacteria chemotaxis model, we assume that the chemical reaction network admits a steady state for initial conditions \(w(0)\). Moreover, only one key species directly impacts the function of the chemotaxis model; namely, phospho-CheY. We denote the steady state concentration of this key species by \(\Phi_w(0)\).

In the example of bacteria chemotaxis, there are ligands in the environment that bind with receptor proteins, initiating a response by the chemosensing mechanism. At each time interval \(t\), we model the initial concentrations of the environmental ligands as random variables. This implies that the initial concentrations observed by the chemosensing mechanism form a discrete random vector, which we denote by \(W(0)\) with elements supported on \([0, W_{\text{max}}]\) and distribution \(P_{W(0)}\).

We assume that the biological system reaches its steady state \(\Phi_W(0)\), or equilibrates, within a time period \(T\). Moreover, after a further time period \(T'\), the concentration of all species in \(C \cup A\) return to fixed values, independent of \(W(0)\). These assumptions provide a basis to characterize fundamental limits of the system, and we will return to their validity in Section VI.

We now introduce the molecular communication system, which transmits information by modulating the number of molecules that are emitted, known as concentration shift keying (CSK). We assume that the communication system is time-slotted with the transmission each symbol separated by a time exceeding \(T + T'\); that is, the system operates with time slots of duration greater than \(T + T'\).

The molecular communication system seeks to transmit a message \(m \in M = \{1, 2, \ldots, M\}\) over \(n\) time slots. To do this, the system is equipped with an encoder \(E : M \rightarrow X^n\), where \(X\) is a finite set corresponding to the set of symbols. Each symbol is the number of information molecules emitted by the molecular communication system. The chemical species corresponding to each information molecule is assumed to be a single species in \(S\). Moreover, the set \(X\) is assumed to contain an innocent symbol, denoted by \(x_0\), corresponding to the case where the transmitter does not send any molecules.

At time slot \(i\), we model the channel between the transmitter and the receiver by the transition probability \(P_{Y_i|X_i}\). \(X_i\) is the number of molecules emitted in time slot \(i\) and \(Y_i\) is the number of information molecules observed at the receiver, with a quantity lying in the finite set \(Y\). In general, the transition probability \(P_{Y_i|X_i}\) may depend on \(i\) due to the presence of inter-symbol interference. In this work, we assume that this is not the case (i.e., all information molecules are absorbed by the end of the time slot). As such, the channel is memoryless and determined by the transition probability \(P_{Y_i|X_i}\), independent of the time slot. As a consequence, communication occurs over the discrete memoryless channel \((X, P_{Y_i|X_i})\).

The receiver in the molecular communication system is equipped with a decoder \(D : Y^n \rightarrow M\). After observations from \(n\) time slots, the decoder forms an estimate of transmitter message \(m \in M\), denoted by \(\hat{m} \in M\). For codewords of length \(n\), the decoder has an average probability of error given by

\[
P_{\text{err}}^{(n)} = \frac{1}{M} \Pr(\hat{m} \neq m|m\text{ sent}).
\] (5)

Some information molecules are also observed by the biological system. Since the information molecule species lies in the set \(S\), these molecules interact with the biological system’s chemosensing mechanism. The quantity of molecules in the set \(S\), denoted by \(Z_i\) with values in the discrete set \(Z\), observed by the biological system in time slot \(i\) is assumed to be governed by the transition probability \(P_{Z_i|X_i}\). As the biological system is sensitive to all molecules in \(S\), it follows that \(Z_i\) is determined not only by \(X_i\) but also by
the distribution of the other molecules in $S$. We assume that the distribution of molecules in $S$ that are not information molecules is fixed. As a consequence, the channel between the transmitter and the biological system is also a discrete memoryless channel, with a transition probability denoted by $P_{Z|X}$. We denote this channel by $(\mathcal{X}, P_{Z|X}, \mathcal{Z})$.

We also note that in each time slot $i$, the quantity of molecules that bind with receptors, $Z_i$, form an element of the concentration vector $W(0)$. Moreover, since all other chemical species are assumed to have fixed initial concentrations, the steady state of the biological system only depends on $Z_i$. As such, in the following we denote the steady state concentration in time slot $i$ by $\Phi(Z_i)$. We also introduce the notation $\Phi(Z) = [\Phi(Z_1), \ldots, \Phi(Z_n)]^T$ for the vector of steady state concentrations of the key chemical species determining biological function; e.g., phospho-CheY in the bacteria chemotaxis model.

B. Communication Constraints

Our goal is to design a transmission strategy such that the molecular communication system is reliable and the function of the biological system is not disrupted; that is, the two systems coexist. The reliability requirement is modeled by requiring

\[
\lim_{n \to \infty} P_{\text{err}}(n) = 0. 
\]

(6)

Due to the importance of steady state concentrations on biological function, we formalize the coexistence constraint as follows. Recall that $\Phi(Z_i)$ is the steady state concentration of the key chemical species determining the function of the biological system and $Z_i$ is the quantity of molecules observed by the chemosensing system in time slot $i$. A natural approach is therefore to ensure that for some $\delta > 0$,

\[
|\mathbb{E}[P_{Z_i} \Phi(Z_i)] - \mathbb{E}[P_{Z_i|X_i=x_0} \Phi(Z_i)]| < \delta, \quad i = 1, 2, \ldots, n, 
\]

(7)

where $P_{Z_i}$ is the distribution of the quantity of molecules that bind with receptors within time slot $i$. Similarly, the distribution $P_{Z_i|X_i=x_0}$ corresponds to the the quantity of molecules that bind with receptors when there is no communication. Observe that the coexistence constraint requires that the expected steady state concentration within the biological system maintain a certain amount of the molecules.

Working directly with the constraint in (7) is presently not tractable and therefore it is necessary to consider bounds. Let $Z_{\text{max}} = \max \{ \Phi(z) : z \in \mathcal{Z} \}$ and define the $n$-dimensional vector $x_0 = [x_0, \ldots, x_0]^T$. Then,

\[
|\mathbb{E}[P_{Z_i} \Phi(Z_i)] - \mathbb{E}[P_{Z_i|X_i=x_0} \Phi(Z_i)]| 
\leq Z_{\text{max}} \|P_{Z_i} - P_{Z_i|X_i=x_0}\|_1 
\leq Z_{\text{max}} \sqrt{2D(P_{Z_i}||P_{Z_i|X_i=x_0})} 
\leq Z_{\text{max}} \sqrt{2D(P_{Z_i}||P_{Z_i|X_i=x_0})}, 
\]

(8)

where $D(\cdot||\cdot)$ is the Kullback-Leibler divergence and the bounds follow from Pinsker’s inequality and the chain rule for Kullback-Leibler divergence [34]. An appropriate coexistence constraint is therefore that

\[
\lim_{n \to \infty} D(P_{Z_i}||P_{Z_i|X=x_0}) = 0. 
\]

(9)

A key insight is that the constraint in (9) bears a strong similarity to the problem of covert communications or communication with a low probability of detection [27]. Next, we formalize and exploit this non-trivial connection to obtain fundamental limits on the rate of reliable molecular communication which can coexist with a biological system.

C. Connections with Covert Communications

The covert communication problem involves two legitimate users who attempt to transmit over a discrete memoryless channel $(\mathcal{X}, P_{Y|X}, \mathcal{Y})$ without being detected by a warden, who observes the signals through another discrete memoryless channel $(\mathcal{X}, P_{Z|X}, \mathcal{Z})$. The key assumptions in this problem are [35]:

- There exists an innocent symbol $x_0 \in \mathcal{X}$, corresponding to the input of the channel when no communication takes place. The distributions on the outputs of the two channels in this case are

\[
P_0 = P_{Y|X=x_0}, \quad Q_0 = P_{Z|X=x_0}. 
\]

(10)

- There exists another symbol $x_1 \in \mathcal{X}$ with $x_1 \neq x_0$, inducing output distributions

\[
P_1 = P_{Y|X=x_1}, \quad Q_1 = P_{Z|X=x_1}. 
\]

(11)

Note that it is possible to also consider input alphabets containing more than two symbols [35]; however, for the purposes of establishing a connection between the covert communication and overlay molecular communication problems it is sufficient to only consider the two symbol case.

Let $Q_1 \ll Q_0$ and $Q_1 \neq Q_0$, which excludes the situations where either the warden would always detect a transmission with non-vanishing probability or would never detect it.

- $P_1 \ll P_0$, which guarantees that the receiver does not have an unfair advantage over the warden.

In the problem of covert communications, the transmitter aims to transmit a message $W$ uniformly distributed on $[1, M]$ by encoding it into a codeword $X = (X_1, \ldots, X_n)$ of $n$ symbols with the help of a secret key $S$ uniformly distributed on $[1, K]$. In particular, at the beginning of a block of $n$ channel uses, the transmitter sets a switch $T$:

- if $T = 1$, the input is connected to the channel;
- otherwise the innocent symbol $x_0$ is sent $n$ times.

Upon observing a noisy version $Y = (Y_1, \ldots, Y_n)$ of $X$ and knowing $S$, the objective is for the receiver to form reliable estimates $\hat{T}$ and $\hat{W}$ of $T$ and $W$. For a codeword of length $n$, the error is defined as

\[
P_{\text{err}}(n) = \mathbb{E}_S[\Pr(W \neq \hat{W}|S, T = 1)] + \Pr(\hat{T} \neq 0|T = 0). 
\]

(12)

Let $Q, P$ be two probability measures. Then, $Q \ll P$ denotes that $Q$ is absolutely continuous with respect to $P$. 

2
The goal in covert communications is then to establish scalings of \( \log M \) and \( \log K \) with \( n \) for which

\[
\lim_{n \to \infty} P_{err}(n) = 0 \quad (13)
\]

and

\[
\lim_{n \to \infty} D(\hat{Q}^n || Q_0^\otimes n) = 0,
\]

where \( Q_0^\otimes n = \prod_{i=1}^{n} Q_0 \).

It is now straightforward to identify an equivalence between the covert communication problem and our formalization of the overlay molecular communication strategy. First observe that, for all \( n \),

\[
D(\hat{Q}^n || Q_0^\otimes n) = D(P_Z || P_{Z|X=x_0}),
\]

where \( Z = [Z_1, \ldots, Z_n]^T \). Moreover, the probability of error in (13) is precisely the probability of error in (5) under the switching communication strategy.

In both the overlay molecular communication and the covert communications problems, the channels between the transmitter and intended receiver as well as the transmitter and the biological system or warden are discrete and memoryless. As such, the formal descriptions of both communication problems are equivalent.

### D. On the Secret Key

Although under the switching communication strategy, there is an equivalence between the two problems, a natural question is the role of the secret key in molecular communications. Due to the fact that the biological system is a passive observer, does the secret key still play such an important role?

The answer to this question is that to guarantee the divergence constraint, a secret key is required for the switching strategy, independent of chemical species observed directly or indirectly by the biological system. However, due to the fact that the biological system is not actively attempting to gain knowledge of the secret key—as in the case of covert communications—the problem of constructing it is much simpler. For example, it may be obtained by observing the concentration of a chemical species observable by both the transmitter and the receiver that does not interact with the chemosensing mechanism used by the biological system.

### E. Fundamental Limits of the Underlay Strategy

By the equivalence of the molecular communication problem with underlay and covert communication, any fundamental limit obtained for either problem holds for the other. In particular, the following theorem holds.

**Theorem 1.** Consider a molecular communication system using the underlay strategy with \( X = \{x_0, x_1\} \) where \( x_0 \) is the innocent symbol corresponding to no transmission. Further, assume that

\[
\begin{align*}
P_{Y|X=x_1} &\ll P_{Y|X=x_0}, & P_{Z|X=x_1} &\ll P_{Z|X=x_0}, \\
& & P_{Z|X=x_1} &\neq P_{Z|X=x_0}
\end{align*}
\]

(16)

and denote the chi-square distance by \( \chi_2(\cdot || \cdot) \). Then for any \( \xi \in (0, 1) \), there exists a communication strategy such that

\[
\lim_{n \to \infty} D(P_Z || P_{Z|X=x_0}) = 0, \quad \lim_{n \to \infty} P_{err} = 0,
\]

\[
\lim_{n \to \infty} \sqrt{n}D(P_Z || P_{Z|X=x_0}) = 0,
\]

\[
\lim_{n \to \infty} \frac{\log M}{\sqrt{n}} = 0
\]

\[
\lim_{n \to \infty} \frac{\log K}{\sqrt{n}} = 0
\]

\[
\frac{2}{\chi_2(P_{Z|X=x_1} || P_{Z|X=x_0})^2} D(P_{Y|X=x_1} || P_{Y|X=x_0}),
\]

\[
(1 - \xi) \sqrt{\frac{2}{\chi_2(P_{Z|X=x_1} || P_{Z|X=x_0})^2}} D(P_{Y|X=x_1} || P_{Y|X=x_0})
\]

\[
(1 - \xi) D(P_{Y|X=x_1} || P_{Y|X=x_0}) + (1 + \xi) D(P_{Z|X=x_1} || P_{Z|X=x_0})^+. \quad (17)
\]

**Proof.** Apply the equivalence between the overlay molecular communication and covert communications problems in Section IV-C, then use [35, Corollary 2].

Observe that the rate obeys the square-root scaling law, which implies that for a secret key scaling with \( \sqrt{n} \) the number of messages that can be reliably sent satisfying the constraint in (7) also scales with \( \sqrt{n} \). We investigate the impact of model parameters on the scaling law in the following section.

### V. Numerical Results

In this section, we provide numerical results to illustrate features of our model. In particular, we demonstrate the key differences between the model introduced in this paper and our earlier model studied in [23]. We also study the behavior of the scaling law in Theorem 1 to gain further insights into fundamental limits of communication in the proposed model.

#### A. Illustration of the Model

In our earlier work in [23], we studied conditions under which the steady-state concentrations of chemical species are not perturbed by an external molecular communication system. These conditions were on the type of chemical species used for communication. To illustrate this setup, consider the following example in Section IV-C in [23].

Suppose that the chemical reaction system \( (S_B, R_B, k_B) \) models a biological system where \( S_B = \{X_1, X_2, X_3\} \) and \( R_B \) consists of the reactions

\[
X_1 + X_3 \underset{k_B}{\overset{k_B}{\rightleftharpoons}} 2X_2
\]

\[
2X_2 \underset{k_B}{\overset{k_B}{\rightleftharpoons}} X_1 + X_3
\]

Further, suppose that the effect of the molecular communication link that emits an information molecule \( I \) is to introduce the additional reactions

\[
X_1 + I \underset{k_B}{\overset{k_B}{\rightleftharpoons}} X_4
\]

\[
X_4 \overset{k_B}{\rightleftharpoons} X_1 + I
\]

\[
I \overset{k_B}{\rightleftharpoons} \emptyset.
\]

The last reaction models the absorption of information molecules by the receiver.
As a consequence of [23, Theorem 1], the steady-state concentration levels of the biological system do not change. This is illustrated in Fig. 1, which plots the time evolution for the concentration of $X_1$ in the presence and absence of information molecules. Note that the quantity of emitted information molecules does not matter as long as all information molecules are absorbed as the time $t \to \infty$.

![Graph showing concentration of species A](image1)

Fig. 1. Plot of the time evolution for the concentration of species $X_1$ in (18). The parameters are: $k_1 = 1$, $k_2 = 1$, $k_3 = 1$, $k_4 = 1$, $k_5 = 0.1$, $[X_1](0) = 1 M$, $[X_2](0) = 2 M$, $[X_3](0) = 3 M$, $[X_4](0) = 0 M$, and $[I](0) = 0.05 M$.

In this present work, we study the more general scenario detailed in Section II-C, where the quantity of emitted information molecules can change the steady-state concentrations of molecules within the biological system. This occurs due to the presence of a chemosensing mechanism and is illustrated in Fig. 2.

In particular, Fig. 2 shows the behavior of a biological system modeled by the reaction network with mass-action kinetics

$$A + B \overset{\alpha}{\to} 2B$$

$$B \overset{\beta}{\to} A,$$

which corresponds to a simplified model of the system in (4) introduced in [31]. In particular, $B$ corresponds to the inactive form of a protein which is affected by the presence of information molecules and $A$ is the active form. The active form $A$ is assumed to directly impact the function of the biological system.

Fig. 2 shows the impact of changing the initial concentration of the species $B$. Observe that unlike the model illustrated in Fig. 1, the steady-state concentrations are affected by the molecular communication system. This motivates the need for communication strategies that limit the perturbation of the steady-state, as obtained in Theorem 1.

![Graph showing concentration of species A](image2)

Fig. 2. Plot of the time evolution for the concentration of species $A$ in (19). The parameters are $[A](0) = 0.1 M$, $\alpha = 0.5$ and $\beta = 0.2$.

parameters affect the term

$$T(P_{Z|X}, P_{Y|X}) = \sqrt{\frac{2}{\chi^2(P_{Z|X=x_1}|P_{Z|X=x_0})}D(P_{Y|X=x_1}|P_{Y|X=x_0}).} \quad (20)$$

Suppose that the number of molecules observed by the receiver, $Y$, and the biological system, $Z$, in the absence of a transmission is $Y \sim \text{Pois}(\lambda_{Y_0})$ and $Z \sim \text{Pois}(\lambda_{Z_0})$, respectively. Note that the random variable $Z$ can be interpreted as the number of molecules observed by the biological system in its natural environment.

During a transmission, the communication system emits either 0 or $L$ molecules. A transmitted molecule reaches the receiver with probability $p_C$. Molecules that are not absorbed by the receiver are assumed to be observed by the biological system. These molecules in turn affect the steady state concentrations of the biological system.

Fig. 3 plots the scaling law in (20) for varying success probability $p_C$. Observe that for low $p_C$, an increasing expected number of environmental molecules observed by the biological system has a limited impact. This is due to the fact that very few molecules reach the receiver and hence it is difficult to transmit information, irrespective of $\lambda_Z$.

On the other hand for large $p_C$, the scaling law rapidly increases when the expected number of environmental molecules observed by the biological system increases. Intuitively, this arises due to the fact that it is more difficult for the biological system to discriminate between the case in which there are information molecules present and the case when there are not.

VI. DISCUSSION

A. Model Assumptions

Our analysis of the underlay strategy in Section IV relied on two key assumptions:

1) The biological system reaches its steady state $\Phi_{W(0)}$, or equilibrates, within a time period $T$.
2) After a further time period $T'$, the concentration of all species in $C \cup A$ return to a fixed value, independent of $W(0)$. The main reason to introduce these assumptions was to obtain memoryless channels for the receiver and the biological system. The consequence of relaxing either assumption will be to introduce memory into the model.

We believe that that the first assumption is reasonable as long as transmissions of symbols by the transmitter are separated sufficiently in time. On the other hand, the second assumption is a property of the biological system and may not be satisfied in every case.

Observe that if the second assumption is relaxed, the steady state concentrations of chemical species in the biological system (i.e., in $C$ and $A$) will form new initial conditions for the next symbol. As such, the channel will have a dependence between outputs at different time periods. It is an interesting avenue of future work to establish scaling laws for the communication rate and the secret key in order for the constraint in (7) to be satisfied.

B. On the Interweaving Strategy

An important feature of some biological systems, including the EnvZ-OmpR system used in bacteria chemotaxis is absolute robust adaptation. This is a fundamental property arising from the structure of the chemical reaction network, which means that once the initial concentration of species in $C$ produced by the chemoreceptor ($X$ in the EnvZ-OmpR system) reaches a given level, the steady state value of the key chemical species $Y_p$ is independent of $X$. This means that the response of bacteria is highly sensitive even to small quantities of attractants or repellants.

We note that sufficient conditions for absolute concentration robustness are known. These conditions depend on the graph structure induced by the chemical reaction network. See [31] for evidence the conditions are present in a range of biological systems.

Clearly if a biological system has absolute concentration robustness, then it is not sensitive to the presence of a molecular communication system. Formally, this means that $P_Z|X=x_1 = P_Z|X=x_0$. In this case, the molecular communication system will never impact the biological system. As such, the molecular communication link can operate at an arbitrarily high rate without disrupting the function of the biological system. The problem of identifying whether or not a biological system has absolute concentration robustness can be viewed within the framework of the interweaving strategy developed in Section III. In particular, to apply the interweaving strategy it is necessary to develop methods for the molecular communication system to detect the presence of absolute concentration robustness.

Another situation where the interweaving strategy is applicable arises when the concentration of all chemical species is not sensitive to the molecular communication system. Such a scenario was identified in our earlier work in [23], where we established conditions for all species in the biological system preserve their steady state concentrations in the presence of molecular communications. An interesting open issue is to establish wider classes of biological systems for which the interweaving strategy can be applied.

VII. CONCLUSIONS

At present, the vast majority of work on artificial nanoscale molecular communication systems has focused on scenarios where a coexisting biological system is not present. In contrast, this paper has established a model of the interactions between molecular communication systems and external biological systems via tools from chemical reaction networks. In particular, we studied the setting where the biological system observes its environment through a chemosensing mechanism.

In order to develop strategies for a molecular communication system to operate without disrupting the function of the external biological system, we proposed three cognitive molecular communication strategies. These strategies bear strong analogies with those in cognitive radio in wireless communication networks.

To explore the performance the applicability of the cognitive molecular communication strategies, we focused on the underlay strategy. For a model motivated by biological systems exploiting chemosensing, we established an equivalence with the problem of covert communications. Through this equivalence, scaling laws for the number of messages that could be sent as the number of channel uses tended to infinity.

This work has provided initial models for the interactions between molecular communication systems and external biological systems, as well as strategies to ensure the function of the biological system is not disrupted. As such, there are a number of future research directions.

The first direction is to explore other models for the kinetics of chemical systems. One such approach is via the chemical master equation, which models the concentrations of each chemical species as a stochastic process rather than through deterministic chemical reaction networks. It is known that features such as absolute concentration robustness are not guaranteed within these models, even when these features are present in the underlying deterministic model. As such, it is
likely that the impact of a molecular communication system may be even more pronounced.

The second direction is to relax the assumptions detailed in Section VI on the system model introduced to study the underlying strategy. This will require the development of information theoretic techniques to cope with molecular communication channels with memory and the divergence constraints.

The third direction is to explore the reliability of communication using the overlay and interweaving cognitive molecular communication strategies. Although there is limited work on interweaving [23] at present, only a simple model was considered. More sophisticated models and fundamental limits of communication remain open issues.

REFERENCES


Valeria Loscri (A09-M09-SM17) is a permanent researcher of the FUN Team at Inria Lille-Nord Europe since Oct. 2013. From Dec. 2006 to Sept. 2013, she was Research Fellow in the TITAN Lab of the University of Calabria, Italy. She received her MSc and PhD degrees in Computer Science in 2003 and 2007, respectively, from the University of Calabria. Her research interests focus on heterogeneous communication technologies and cooperation of heterogeneous devices. She is in the editorial board of Elsevier ComNet, JNCA, IEEE Trans. on Nanobioscience. Since 2015 she is member of the Committee for Technological Development at Inria Lille-Nord Europe. Since 2016, she is Scientific European Responsible for Inria Lille-Nord Europe. She received her Habilitation à diriger des recherches (HDR) in 2018 from Université Lille 1.

Marco Di Renzo (S’05-A’07-M’09-SM’14) was born in L’Aquila, Italy, in 1978. He received the Laurea (cum laude) and Ph.D. degrees in electrical engineering from the University of L’Aquila, Italy, in 2003 and 2007, respectively, and the Habilitation à Diriger des Recherches (Doctor of Science) degree from University Paris-Sud, France, in 2013. Since 2010, he has been a Chargé de Recherche CNRS (CNRS Associate Professor) in the Laboratory of Signals and Systems (L2S) of Paris-Saclay University - CNRS, CentraleSupelec, Univ Paris Sud, Paris, France. He serves as the Associate Editor-in-Chief of IEEE Communications Letters, and as an Editor of IEEE Transactions on Communications, and IEEE Transactions on Wireless Communications. He is a Distinguished Lecturer of the IEEE Vehicular Technology Society and IEEE Communications Society, and a Senior Member of the IEEE. He is a recipient of several awards, including the 2013 IEEE-COMSOC Best Young Researcher Award for Europe, Middle East and Africa, the 2013 NoE-NEWCOM# Best Paper Award, the 2014-2015 Royal Academy of Engineering Distinguished Visiting Fellowship, the 2015 IEEE Jack Neubauer Memorial Best Systems Paper Award, the 2015-2018 CNRS Award for Excellence in Research and Ph.D. Supervision, the 2016 MSCA Global Fellowship (declined), the 2017 SEE-IEEE Alain Glavieux Award “for outstanding results in developing several mathematical abstractions (for mobile network modeling), innovating ideas, and demonstrating their usefulness in future wireless communication systems”, the 2018 IEEE COMSOC Best Young Professional Award in Academia “for outstanding contributions to the academia in terms of innovative Communications Research and IEEE ComSoc Services, especially in the area of emerging physical-layer technologies”, and 7 conference Best Paper Awards (2012 and 2014 IEEE CAMAD, 2013 IEEE VTC-Fall, 2014 IEEE ATC, 2015 IEEE ComManTel, 2017 IEEE SigTelCom, 2018 INISCOM).

Trung Q. Duong (S’05-M’12-SM’13) received his Ph.D. degree in Telecommunications Systems from Blekinge Institute of Technology (BTH), Sweden in 2012. Currently, he is with Queen’s University Belfast (UK), where he was a Lecturer (Assistant Professor) from 2013 to 2017, and a Reader (Associate Professor) from 2018. His current research interests include Internet of Things (IoT), wireless communications, molecular communications, and signal processing. He is the author or co-author of over 300 technical papers published in scientific journals (181 articles) and presented at international conferences (130 papers).

Dr. Duong currently serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, IET COMMUNICATIONS, and a Lead Senior Editor for IEEE COMMUNICATIONS LETTERS. He was awarded the Best Paper Award at the IEEE Vehicular Technology Conference (VTC-Spring) in 2013, IEEE International Conference on Communications (ICC) 2014, IEEE Global Communications Conference (GLOBECOM) 2016, and IEEE Digital Signal Processing Conference (DSP) 2017. He is the recipient of prestigious Royal Academy of Engineering Research Fellowship (2016-2021) and has won a prestigious Newton Prize 2017.