A Markovian Approach to the Optimal Demodulation of Diffusion-based Molecular Communication Networks

Chun Tung Chou

School of Computer Science and Engineering
University of New South Wales, Australia
ctchou@cse.unsw.edu.au

Technical Report
UNSW-CSE-TR-201503
April 2015

THE UNIVERSITY OF
NEW SOUTH WALES

School of Computer Science and Engineering
The University of New South Wales
Sydney 2052, Australia

Abstract

In a diffusion-based molecular communication network, transmitters and receivers communicate by using signalling molecules (or ligands) in a fluid medium. This paper assumes that the transmitter uses different chemical reactions to generate different emission patterns of signalling molecules to represent different transmission symbols, and the receiver consists of receptors. When the signalling molecules arrive at the receiver, they may react with the receptors to form ligand-receptor complexes. Our goal is to study the demodulation in this setup assuming that the transmitter and receiver are synchronised. We derive an optimal demodulator using the continuous history of the number of complexes at the receiver as the input to the demodulator. We do that by first deriving a communication model which includes the chemical reactions in the transmitter, diffusion in the transmission medium and the ligand-receptor process in the receiver. This model, which takes the form of a continuous-time Markov process, captures the noise in the receiver signal due to the stochastic nature of chemical reactions and diffusion. We then adopt a maximum a posterior framework and use Bayesian filtering to derive the optimal demodulator. We use numerical examples to illustrate the properties of this optimal demodulator.

Keywords: Molecular communication networks; modulation; demodulation; maximum a posteriori; optimal detection; stochastic models; Bayesian filtering; molecular receivers.

1 Introduction

Molecular communication is a promising approach to realise communications among nano-scale devices [1, 2, 3, 4]. There are many possible applications with these networks of nano-devices, for example, in-body sensor networks for health monitoring and therapy [5, 3]. This paper considers diffusion-based molecular communication networks.

In a diffusion-based molecular communication network, transmitters and receivers communicate by using signalling molecules or ligands. The transmitter uses different time-varying functions of concentration of signalling molecules (or emission patterns) to represent different transmission symbols. The signalling molecules diffuse freely in the medium. When signalling molecules reach the receiver, they react with chemical species in the receiver to produce output molecules. The counts of output molecules over time is the receiver output signal which the receiver uses to decode the transmitted symbols.

Two components in diffusion-based molecular communication system are modulation and demodulation. A number of different modulation schemes have been considered in the literature. For example, [6, 7] consider Concentration Shift Keying (CSK) where different concentrations of signalling molecules are used by the transmitter to represent different transmission symbols. Other modulation techniques that have been proposed include Molecule Shift Keying (MSK) [8, 9], Pulse Position Modulation (PPM) [10] and Frequency Shift Keying (FSK) [11]. This paper assumes that the transmitter uses different chemical reactions to generate the emission patterns of different transmission symbols. The motivation to use this type of modulation mechanism is that chemical reactions are a natural way to produce signalling molecules, e.g. the papers [12, 13] study a number of molecular circuits (which are sets of chemical reactions) that can produce oscillating signals, and the paper [14] discusses a number of signalling mechanisms in living cells.

We assume the receiver consists of receptors. When the signalling molecules (ligands) reach the receiver, they can react with the receptors to form ligand-receptor complexes (which are the output molecules in this paper). We consider the problem of using the continuous-time history of the number of complexes for demodulation assuming that the transmitter and receiver are synchronised. The ligand-receptor complex signal is a stochastic process with three sources of noise because the chemical reactions at the transmitter, the diffusion of signalling molecules and the ligand-receptor binding process are all stochastic. We derive a continuous-time Markov process (CTMP) which models the chemical reactions at the transmitter, the diffusion in the medium and the ligand-receptor binding process. By using this model and the theory of Bayesian filtering, we derive the maximum a posteriori (MAP) demodulator using the continuous-time history of the number of complexes as the input.

This paper makes two key contributions: (1) We propose to use a CTMP to model a molecular communication network with chemical reactions at the transmitter, a diffusive propagation propagation medium and receptors at the receiver. The CTMP captures all three sources of noise in the communication network. (2) We derive a closed-form expression for the MAP demodulation filter using the proposed CTMP. The closed-form expression gives insight into the important elements needed for optimal demodulation, these are the timings at which the receptor bindings occur, the number of unbound receptors and the mean concentration of signalling molecules around the receptors.

The rest of the paper is organised as follows. Section 2 discusses related work. Section 3 presents the system assumptions, as well as a mathematical model from the transmitter to the ligand-receptor complex signal based on CTMP. We derive the MAP demodulator in Section 4 and illustrate its numerical properties in Section 5. Finally, Section 6 concludes the paper.

2 Related work

There is a growing interest to understand molecular communication from the communication engineering point of view. For recent surveys of the field, see [1, 2, 3, 4]. We divide the discussion under these headings: transmitters, receivers, models and others.

Transmitters. A number of different types of transmission signals have been considered in the molecular communication literature. The papers [15, 7] assume that the transmitter releases the signalling molecules in a burst which can be modelled as either an impulse or a pulse with a finite duration. A recent work in [16] assumes that the transmitter releases the molecules according to a Poisson process. In this paper, we instead assume that the transmitter uses different sets of chemical reactions to generate different transmission symbols and we use CTMP to model these transmission symbols. Since a Poisson process can also be modelled by a CTMP, the transmission process in this paper is more general than that of [16]. Our CTMP model can also deal with an impulsive input by using an appropriate initial condition for the CTMP. The use of CTMP as an end-to-end model — which includes the transmitter, the medium and the receiver — does not appear to have been used before.

Receivers. Demodulation methods for diffusion-based molecular communication have been studied in [17, 18]. Both papers also use the MAP framework with discrete-time samples of the number of output molecules as the input to the demodulator. Instead, in this paper, we consider demodulation using continuous-time history of the number of complexes.

The demodulation from ligand-receptor signal has also been considered in [15]. The key difference is that [15] uses a linear approximation of the ligand-receptor process while we use a non-linear reaction rate.

The capacity of molecular communications based on ligand-receptor binding has been studied in [19, 20] assuming discrete samples of the number of complexes are available. A recent work [21] considers the capacity of such systems in the continuous-time limit. Instead of focusing on the capacity, our work focuses on demodulation.

Receiver design is an important topic in molecular communication and has been studied in many papers, some examples are [22, 17, 23, 18, 24]. These papers either use one sample or a number of discrete samples on the count of a specific molecule to compute the likelihood of observing a certain input symbols. This paper takes a different approach and uses continuous-time signals.

Another approach of receiver design for molecular communication is to derive molecular circuits that can be used for decoding. An attempt is made in [11] to design a molecular circuit that can decode frequency-modulated signals. However, the work does not take diffusion and reaction noise into consideration. A recent work in [25] analyses end-to-end molecular communication biological circuits from linear time-invariant system point of view. The work in [26] compares the information theoretic capacity of a number of different types of linear molecular circuits. This paper differs from the previous work in that it uses a non-linear ligand-receptor binding model.

The noise property of ligand-receptor for molecular communication has been characterised in [27]. The case for non-linear ligand-receptor binding does not appear to have an analytical solution and [27] derives an approximate characterisation using a linear reaction rate assuming that the number of signalling molecules around the receptor is large. This paper uses a non-linear ligand-receptor binding model and no approximation is used in solving the filtering problem.

Models. This paper uses the Reaction Diffusion Master Equation (RDME) [28] framework to model the reactions and diffusion in the molecular communication networks. RDME assumes that time is continuous while the diffusion medium is discretised into voxels. This results in a CTMP with finite number of (discrete) states. RDME has been used to model stochastic dynamics of cells in the biology literature [29]. An attraction of RDME is that it has the Markov property which means that one can leverage the rich theory behind Markov process.

The author of this paper has previously used an extension of the RDME model, called the RDME with exogenous input (RDMEX) model, to study molecular communication networks in [30, 31, 32]. The RDMEX assumes that the times at which the transmitter emits signalling molecules are deterministic. This results in a stochastic process which is piecewise Markov or

the Markov property only holds in between two consecutive emissions by the transmitter. In this paper, we assume the transmitter uses chemical reactions to generate the signalling molecules. Therefore, the emission timings are not deterministic but are governed by a stochastic process.

In this paper, we assume that the propagation medium is discretised in the voxels. An alternative modelling paradigm that has been used in a number of molecular communication network papers [7, 15, 16] is that the transmitter or receiver has a non-zero spatial dimension (commonly modelled by a sphere) while the propagation medium is assumed to be continuous. (Note that though [16] does not explicitly state the dimension of the receiver, one can infer from the fact that the receiver must have a non-zero dimension because it has a non-zero probability of receiving the signalling molecules.) We believe the technique in this paper can be adapted to this alternative modelling paradigm and we do not expect this alternative modelling paradigm will change the results in this paper; we will explain this in Section 4.4.

There is a rich literature in the modelling of biological systems discussing the difference between: (1) The particle approach which has a continuous state space because the state of a particle is its position; and (2) The mesoscopic approach (the approach in this paper) which discretises the medium into discrete voxels and consider the number of molecules in the voxels as the state. The first approach is more accurate but the computation burden can be high [33], while the second approach is accurate for appropriate discretisation [34, 28]. There are also hybrid approaches too. An overview of various modelling and simulation approaches can be found in [33].

Others: The results of this paper may also be of interest to biologists who are interested to understand how living cells can distinguish between different concentration levels. The result of this paper can be viewed as a generalisation of [35] which studies how cells can distinguish between two constant levels of ligand concentration.

3 End-to-end communication models

This paper considers diffusion-based molecular communication with one transmitter and one receiver in a fluid medium. Figure C.1 gives an overview of the setup considered in this paper. The transmitter uses different chemical reactions to generate the emission patterns of different transmission symbols. The transmitter acts as the source and emitter of signalling molecules. The signalling molecules diffuses in the fluid medium. The front-end of the receiver consists of a ligand-receptor binding process and the back-end consists of the demodulator with the number of complexes as its input.

In this section, we first describe the system assumptions in Section 3.1. We then present, in Section 3.2, an end-to-end model which includes the transmitter, the transmission medium and the ligand-receptor binding process in the receiver, see the dashed box in Figure C.1. The end-to-end model is a CTMP which includes chemical reactions in the transmitter, diffusion in the medium and the ligand-receptor binding process in the receiver.

3.1 Model assumptions

We assume that the medium (or space) is discretised into voxels while time is continuous. This modelling framework results in a RDME [28, 36, 37], which is a CTMP commonly used to model systems with both diffusion and reactions. In addition, we assume the communication uses only one type of signalling molecule (or ligand) denoted by S. We divide the description of our model into three parts: transmission medium, transmitter and receiver. We begin with the transmission medium. Table 3.1 summaries the frequently used notation and chemical symbols.

Transmission medium

We model the transmission medium as a three dimensional (3-D) space and partition the space into cubic *voxels* of volume W^3. Figure C.2 shows an example of a medium which has a dimension of 4 voxels along both the x and y-directions, and 1 voxel in the z-direction. (Note that Figure

Symbol	Meaning
W	Dimension of one side of a voxel
D	Diffusion constant
d	Diffusion rate between neighbouring voxels
N_v	Total number of voxels
M	Total number of receptors
$\tilde{\lambda}$	Reaction rate constant for the binding reaction
λ	$\lambda = \frac{\tilde{\lambda}}{W^3}$
μ	Reaction rate constant for the unbinding reaction
s	A transmission symbol
$b(t)$	Number of complexes at time t
$n_i(t)$	Number of signalling molecules in voxel i at time t
$N(t)$	Equation (3.4). A vector containing the number of signalling molecules in each voxel, the counts of intermediate chemical species in the transmitter and the cumulative count of the number of molecules that have left the system
$\sigma_s(t)$	The mean number of signalling molecules in the receiver voxel at time t if the transmitter sends symbol s
$U(t)$	The cumulative number of times the receptors have switched from the unbound to bound state at time t
E	An unbound receptor
S	A signalling molecule
C	A complex

Table 3.1: Notation and chemical symbols

C.2 should be viewed as a projection onto the $x-y$ plane.) In general, we assume the medium to have N_x, N_y and N_z voxels in the x, y and z directions where N_x, N_y and N_z are positive integers. In Figure C.2, $N_x = N_y = 4$ and $N_z = 1$. We also use $N_v = N_x N_y N_z$ to denote the total number of voxels.

We refer to a voxel by a triple (x, y, z) where x, y and z are integers or by a single index $\xi \in [1, N_v]$. Figure C.2 shows the triples for some of the voxels. The single index ξ is calculated from the triple (x, y, z) by using $\xi(x, y, z) = x + N_x(y-1) + N_x N_y(z-1)$. The single indices for voxels are shown in the top right-hand corner of the voxels in Figure C.2.

Diffusion is modelled by molecules moving from one voxel to a neighbouring voxel. For examples, in Figure C.2, molecules can diffuse from Voxel 1 to Voxels 2 or 5, from Voxel 2 to Voxels 1, 3 and 6, and so on. The diffusion of molecules between neighbouring voxels is indicated by the two-way arrows in Figure C.2.

We assume that the signalling molecule S is the only diffusible chemical species in our model and the diffusion coefficient for S is D. This means the signalling molecules diffuse from one voxel to a neighbouring voxel at a mean rate of d where $d = \frac{D}{W^2}$. In other words, within an infinitesimal time Δt, the probability that a signalling molecule diffuses to a neighbouring voxel is $d\,\Delta t$. Note that the expression for d can be derived from spatially discretising the diffusion equation into regular cubic voxels of volume W^3, see [36, p.341] or [38, Section 3].

The dashed lines in Figure C.2 indicate the boundary of our transmission medium. Many different boundary conditions are used by engineers and physicists to model what happens when a molecule reaches the boundary of a medium. Two typical boundary conditions are absorbing and reflecting boundaries [36]. An absorbing boundary means that a molecule can leave the transmission medium and once the molecule has left, it will not return to the medium. For example, in Figure C.2, we allow molecules to leave the medium via one surface of Voxel 16 as indicated by the one-way arrow. Mathematically, this is modelled by a rate of leaving the medium, similar to that of modelling the diffusion between the voxels. A reflecting boundary means that a molecule cannot leave the medium, i.e. a molecule hitting a reflecting boundary will stay in the voxel. Our model can capture these boundary conditions.

It has been shown in [34, 28] that in order for RDME to produce physically meaningful results, the voxel dimension W must be within a certain range. In this paper, we assume that W comes from a valid range. The choice of W is beyond the scope of the paper and the reader can refer to [34, 28] for further discussion.

For simplicity, we assume that the medium is homogeneous with a constant diffusion coefficient D. It is straightforward to extend the framework to cover inhomogeneous medium [32]. It is also possible to use non-cubic voxels, see [39, 40].

Transmitter

We assume the transmitter occupies one voxel. However, it is straightforward to generalise to the case where a transmitter occupies multiple voxels. We limit our consideration to one symbol interval without inter-symbol interference (ISI) at this moment. We will discuss the multiple symbol interval case with ISI in Section 4.5.

We assume that the transmitter can send K different symbols $s = 0, 1, .., K-1$ where each symbol s is characterised by an emission pattern $u_s(t)$. The role of emission pattern in molecular communication is the same as that of transmitted signal in electromagnetic communication. If a transmitter uses a deterministic emission pattern $u_s(t)$ to represent symbol s, it means the transmitter emits $u_s(t)$ signalling molecules into the transmitter voxel at time t. We use an example to illustrate the meaning of emission pattern. Consider an emission pattern $u_1(t)$ for Symbol 1 where $u_1(t) = \delta_{t,1.2} + \delta_{t,5.6} + 2\delta_{t,8.1}$ where $\delta_{i,j}$ denotes the Kronecker delta, which means $\delta_{i,j} = 1$ if and only if $i = j$. The emission pattern $u_1(t)$ means that, for Symbol 1, the transmitter emits one signalling molecule at times 1.2 and 5.6, two signalling molecules at time 8.1 and does not emit any molecules at any other times.

In this paper, we assume that the emission pattern for each symbol is produced by a set of chemical reactions located in the transmitter voxel. Given K symbols, the transmitter uses K different reactions to generate these symbols, see Figure C.1. As an example, a class of chemical reactions inside living cells [41] is

$$\text{RNA} \xrightarrow{\kappa} \text{RNA} + \text{A} \tag{3.1}$$

where ribonucleic acid (RNA, which is a molecule commonly found in living cells) produces the chemical species A. This class of chemical reactions can be modelled by a Poisson process where molecules of A are produced at a mean rate of κ [42]. Note that the emission patterns produced by chemical reactions are *not* deterministic, but *stochastic*. The mean emission pattern of this chemical reaction is $\mathbf{E}[u(t)] = \kappa$.

Following on from the above example, one can realise Amplitude Shift Keying (ASK) in molecular communication by using different chemical reactions that can produce signalling molecules at different mean rates. For example, if there are four different reactions that can produce signalling molecules at four different mean rates of κ_0, κ_1, κ_2 and κ_3, then one can use these four different reactions to produce 4 different symbols. Note that it is possible for the four chemical reactions to produce the same emission pattern (or realisation), though with different probabilities.

A standard result in physical chemistry shows that the dynamics of a set of chemical reactions can be modelled by a CTMP [43]. Therefore, we will model the transmitter by a CTMP. Note that, in this paper, we will not specify the sets of chemical reactions used by the transmitter except for simulation because the MAP demodulator does not explicitly depend on the sets of chemical reactions that the transmitter uses.

Receiver

We assume the receiver occupies one voxel and we use R to denote the index of the voxel at which the receiver is located. In Figure C.2, we assume the receiver is at Voxel 7 (light grey) and hence $R = 7$ for this example. In addition, we assume that the transmitter and receiver voxels are distinct.

We assume that the receiver has M non-interacting receptors and we use E as the chemical name for an unbound receptor. These receptors are fixed in space and do not diffuse, and they are only found in the receiver voxel. Furthermore, these receptors are assumed to be uniformly distributed in the receiver voxel.

The receptor E can bind to a signalling molecule S to form a ligand-receptor complex (or complex for short) C, which is a molecule formed by combining E and S. This is known as ligand-receptor binding in molecular biology literature [44]. The binding reaction can be written as the chemical equation:

$$\mathrm{S} + \mathrm{E} \xrightarrow{\tilde{\lambda}} \mathrm{C} \tag{3.2}$$

where $\tilde{\lambda}$ is the reaction rate constant. Since the receptors are only found in the receiver voxel, the binding reaction occurs in a volume of W^3, which is the volume of a voxel. This rate at which the complexes C is formed is given by the product of $\frac{\lambda}{W^3}$, the number of signalling molecules in the receiver voxel and the number of unbound receptors[1]. We define $\lambda = \frac{\tilde{\lambda}}{W^3}$ and will use λ in the CTMP. Note that this is equivalent to ligand-receptor binding model used in [27, Section V-B].

A ligand-receptor complex C can dissociate into an unbound receptor E and a signalling molecule S. This can be represented by the chemical equation

$$\mathrm{C} \xrightarrow{\mu} \mathrm{E} + \mathrm{S} \tag{3.3}$$

where μ is the reaction rate constant. The rate at which the complexes are dissociating is given by the product of μ and the number of complexes[2].

Since a receptor can either be in an unbound state E or in a complex C, we have the following conservation relation: the number of unbound receptors plus the number of complexes is equal to the total number of receptors M.

3.2 General end-to-end model

In order to derive the MAP demodulator, we need an end-to-end model which includes the transmitter, the medium and the ligand-binding process, see Figure C.1. Since chemical reactions (which includes the chemical reactions in the transmitter as well as the ligand-receptor binding process in the receiver) and diffusion can be modelled by CTMP, it is possible to use a CTMP as an end-to-end model. In this section we present a general end-to-end model that includes the transmitter, diffusion and the ligand-receptor process in the receiver. An excellent tutorial introduction to the modelling of chemical reactions and diffusion by using CTMP can be found in [38]. We have also included an example in C.

The aim of the end-to-end model is to determine the properties of the receiver signal from the transmitter signal. The receiver signal in our case is the number of complexes over time. Since the transmitter uses K symbols, the transmitter signal is generated by one of the K sets of chemical reactions. This means that we need K end-to-end models with a model for each of the K symbols or sets of chemical reactions. The principle behind building these K models are identical

[1] This footnote explains how $\frac{\lambda}{W^3}$ comes about. Consider a chemical reaction where reactants S and E react to form product C. We assume the reactions are taking place within a volume of W^3. Let c_{S}, c_{E} and c_{C} be, respectively, the concentration of S, E and C in the volume. The law of mass action says that $\frac{dc_{\mathrm{C}}}{dt} = \tilde{\lambda}\, c_{\mathrm{E}}\, c_{\mathrm{S}}$. In the case of the CTMP or RDME in this paper, we want to keep track of the number of molecules in a volume (the voxel) instead. Let n_{S}, n_{E} and n_{C} be, respectively, the number of S, E and C molecules in the volume. Since concentration and molecule counts are related by $c_{\mathrm{C}} W^3 = n_{\mathrm{C}}$ etc, we will, in a mathematically loose way, write $\frac{dn_{\mathrm{C}}}{dt} = \frac{\tilde{\lambda}}{W^3}\, n_{\mathrm{S}}\, n_{\mathrm{E}}$. Since n_{C} is a discrete quantity, the derivative $\frac{dn_{\mathrm{C}}}{dt}$ is not defined but we can interpret it as the production rate of molecules C. This explains how to convert the law of mass action, which is in terms of concentration, to the rate law used in RDME which is in terms molecular counts. This conversion is also discussed in [45, 38].

[2] The law of mass action for the dissociation reaction is $\frac{dc_{\mathrm{C}}}{dt} = -\mu\, c_{\mathrm{C}}$ where c_{C} is the concentration of the complexes. We can use the same argument in Footnote 1 to show that the dissociation rate of C is $\mu\, n_{\mathrm{C}}$ where n_{C} is the number of complexes. In particular, note that no scaling by volume W^3 of the voxel is required.

so without loss of generality, we will assume that the model here is for Symbol 0. We begin with a few definitions.

Let $n_i(t)$ (where $1 \leqslant i \leqslant N_v$) be the number of signalling molecules S in Voxel i at time t. In particular, since we have defined R to be the index of the receiver voxel, $n_R(t)$ is the number of signalling molecules in the receiver voxel. We assume the transmitter is a set of chemical reactions which uses H intermediate chemical species Q_1, Q_2, ... and Q_H and these intermediate species remain the transmitter voxel. Let $n_{Q_i}(t)$ be the number of chemical species Q_i in the transmitter voxel at time t. Molecules may also be degraded or leave the system forever if absorbing boundary condition is used. We use $n_A(t)$ to denote the cumulative number of molecules that have left the system. Note that since $n_i(t), n_{Q_i}(t)$ and $n_A(t)$ are molecular counts, they must belong to the set of non-negative integers $\mathbb{Z}_{\geqslant 0}$. We define the vector $N(t) \in \mathbb{Z}_{\geqslant 0}^{N_v+H+1}$ to be:

$$N(t) = \begin{bmatrix} n_1(t) & \ldots & n_{N_v}(t) & n_{Q_1}(t) & \ldots & n_{Q_H}(t) & n_A(t) \end{bmatrix}^T \tag{3.4}$$

where ^{T} denotes matrix transpose.

Let $b(t)$ denote the number of complexes or bound receptors at time t and $\mathbb{Z}_{[0,M]}$ denote the set of integers between 0 and M inclusively. We require that a valid $b(t)$ must be an element of $\mathbb{Z}_{[0,M]}$.

The state of the end-to-end model is the tuple $(N(t), b(t))$ where $b(t)$ is the number of complexes and $N(t)$ contains all the other molecular counts. Since a receptor can either be unbound or in a complex, the number of unbound receptors at time t is $M - b(t)$; therefore, the mathematical model only has to keep track of either the number of unbound receptors or the number of complexes, and we have chosen to keep track of the latter.

We will now specify the transition probabilities from state $(N(t), b(t))$ to state $(N(t+\Delta t), b(t+\Delta t))$. State transitions can be caused by and of these events: a chemical reaction in the transmitter, the diffusion of a signalling molecule from a voxel to neighbouring voxel, and the binding or unbinding of a receptor in the receiver. We know from the theory of CTMP that the probability of two events taking place in an infinitesimal duration of Δt is of the order of $o(\Delta t^2)$. Intuitively, this means only one event can occur within Δt. We can divide the transition probabilities from $(N(t), b(t))$ to $(N(t+\Delta t), b(t+\Delta t))$ into 2 groups depending on whether the number of complexes has changed or not in the time interval $(t, t+\Delta t)$. If the number of complexes has changed from time t to $t+\Delta t$, i.e. $b(t+\Delta t) \neq b(t)$, this means either a binding reaction (3.2) or a unbinding reaction (3.3) has occurred.

If a binding reaction (3.2) has occurred, then the number of signalling molecules in the receiver voxel is decreased by 1 and the number of complexes $b(t)$ is increased by 1. This reaction occurs at a mean rate of $\lambda\, n_R(t)\, (M - b(t))$. We use $\mathbb{1}_i$ to denote the standard basis vector with a '1' at the i-th position. We can write the state transition probability of the receptor binding reaction (3.2) as:

$$\mathbf{P}[N(t+\Delta t) = N(t) - \mathbb{1}_R, b(t+\Delta t) = b(t) + 1 | N(t), b(t)] = \lambda\, n_R(t)\, (M - b(t))\, \Delta t \tag{3.5}$$

Recalling that R is the index of the receiver voxel and $n_R(t)$ is the R-th element of $N(t)$ in (3.4), the expression $N(t+\Delta t) = N(t) - \mathbb{1}_R$ is equivalent to $n_R(t+\Delta t) = n_R(t) - 1$, which means the number of signalling molecules in the receiver voxel has decreased by 1. Similarly, the expression $b(t+\Delta t) = b(t) + 1$ says the number of complexes has increased by 1. The right-hand side (RHS) of (3.5) is the transition probability and is given by the product of mean reaction rate and Δt.

Similarly, the transition probability of the unbinding reaction is given by:

$$\mathbf{P}[N(t+\Delta t) = N(t) + \mathbb{1}_R, b(t+\Delta t) = b(t) - 1 | N(t), b(t)] = \mu\, b(t)\, \Delta t \tag{3.6}$$

where RHS of (3.6) is the transition probability.

We now specify the second group of transition probabilities with $b(t+\Delta t) = b(t)$. These transitions are caused by either a reaction in the transmitter or diffusion of signalling molecules

between neighbouring voxels. Let $\eta_i, \eta_j \in \mathbb{Z}_{\geqslant 0}^{N_v+H+1}$ be two valid $N(t)$ vectors; let also $\beta \in \mathbb{Z}_{[0,M]}$. For $\eta_i \neq \eta_j$, we write

$$\mathbf{P}[N(t+\Delta t) = \eta_i, b(t+\Delta t) = \beta | N(t) = \eta_j, b(t) = \beta] = d_{ij}\ \Delta t \tag{3.7}$$

where d_{ij} is the transition rate from state (η_j, β) to state (η_i, β). Since this transition is due to either a reaction in the transmitter or diffusion, d_{ij} is independent of the number of complexes β. Depending on the type of transition, the value of d_{ij} can depend on the reaction constants in the transmitter, diffusion rate and some states of η_j. For example, if the transition from η_j to η_i is caused by the diffusion of a signalling molecule from Voxel 1 to Voxel 2, we have $\eta_i = \eta_j - \mathbb{1}_1 + \mathbb{1}_2$ at a rate of $d\eta_{j,1}$ where $\eta_{j,1}$ is the first element in η_j or equivalently the number of signalling molecules in Voxel 1 in state η_j; so, for this example, $d_{ij} = d\eta_{j,1}$. The main advantage of using Equation (3.7) is that it allows us a cleaner abstraction to solve the Bayesian filtering problem when deriving the MAP demodulator. We also remark that we will not specify the exact expression of d_{ij} because d_{ij}'s do not appear explicitly in the demodulator.

Equations (3.5), (3.6) and (3.7) specify all the possible state transitions. The probability of no state transition is therefore:

$$\begin{aligned}&\mathbf{P}[N(t+\Delta t) = \eta_j, b(t+\Delta t) = b(t) | N(t) = \eta_j, b(t)] \\ &= 1 - d_{jj}\ \Delta t - \lambda\ n_R(t)\ (M - b(t))\ \Delta t - \mu\ b(t)\ \Delta t\end{aligned} \tag{3.8}$$

where

$$d_{jj} = \sum_{j \neq i} d_{ij} \tag{3.9}$$

We have now specified all the state transition probabilities for Symbol 0. If a different symbol is used, the value of H, the dimension of $N(t)$ and the d_{ij} parameters can change. However, the state transition probabilities still can be summarised by Equations of the form (3.5), (3.6) and (3.7). In any case, the derivation of the MAP demodulator only requires us to work with one symbol at a time. Hence, we will use Equations (3.5)-(3.9) for any transmission symbol.

3.3 Discussion

Note that the CTMP includes all the three sources of noise in our system, due to chemical reactions in the transmitter, random diffusive movements in the medium and the ligand-receptor binding process at the receiver. Some of the these noise components have also been characterised in earlier literature. For example, [47] discusses sampling noise at the transmitter and counting noise at the receiver. One can also study these noises using the derived CTMP and let us take sampling noise as an example. For the moment, let us isolate the transmitter voxel from the propagation medium, i.e. we do not allow the signalling molecules to leave or enter the transmitter voxel. In this case, the number of signalling molecules in the transmitter voxel is due entirely to chemical reactions and we use $n_{\text{isolated}}(t)$ to denote the number of signalling molecules in the isolated transmitter voxel at time t. Now, let us consider the transmitter voxel again but we allow signalling molecules to diffuse in and out of the voxel; we use $n_{\text{connected}}(t)$ to denote the number of signalling molecules in the transmitter voxel in this case. It is natural to consider $n_{\text{connected}}(t)$ as the transmitter signal because this is the number of signalling molecules in the transmitter voxel. However, $n_{\text{connected}}(t)$ can be different from $n_{\text{isolated}}(t)$ because signalling molecules can diffuse in and out of the transmitter voxel, which is the cause of sampling noise.

Equations (3.5) to (3.8) hold for any valid state $(N(t), b(t))$. If we collect all the transition probability equations for all valid states, then we can form the infinitesimal generator of the CTMP. For a given initial probability distribution of the initial state $(N(0), b(0))$, one can in principle solve the first order ordinary differential equation (ODE) associated with the infinitesimal generator to compute the probability of the number of complexes $b(t)$, or the property of the receiver signal. However, in practice, this ODE is of a very high dimension and it is an active area of research to

derive algorithms to solve this ODE efficiently and accurately [48]. We remark that this ODE is commonly known as the reaction-diffusion Master equation [38, 37, 36] because it describes the dynamics of chemical reactions and diffusion. For this paper, it suffices for us to use the equations of the form in (3.5) to (3.8), therefore, we will not present the Master equation.

4 The MAP demodulator

This section aims to derive the optimal demodulator using the CTMP derived in the previous section. We assume the input to the demodulator is the continuous-time signal $b(t)$ which is the number of complexes at time t. There are a number of reasons why we choose to work with the continuous-time signal $b(t)$, rather than its sampled version. First, the signal $b(t)$ may **not** be strictly band limited in the frequency domain. Second, our results show that the optimal demodulator needs to know the time instances at which a receptor is switching from the unbound to bound state. This timing information, which is essentially an impulse, is unfortunately lost by sampling $b(t)$. Third, the solution of the proposed decoding problem can be used to benchmark molecular circuit [26] based decoders. Since molecular circuits use chemical reactions for computation, they are fundamentally analogue circuits. Fourth, there is an increasing interest in the circuit design community to design low-power analogue signal processing circuits [49].

An intermediate step to derive the MAP demodulator is to solve a continuous-time filtering problem. In a filtering problem, one uses all the observations available up till time t to predict the system state at time t. For our case, the observations are the number of complexes $b(t)$ in the continuous interval $[0, t]$. (This means the number of observations is infinite because we are considering the continuous-time signal $b(t)$ in a non-zero time interval $[0, t]$.) We use $\mathcal{B}(t) = \{b(\tau); 0 \leqslant \tau \leqslant t\}$ to denote the section of the signal $b(t)$ in the time interval $[0, t]$. (Note that the $\{\}$ is not a set notation. Here $\mathcal{B}(t)$ is a realisation of the number of complexes in $[0, t]$ of the CTMP. This notation is used in some non-linear filtering literature, e.g. [50].) We can think of $\mathcal{B}(t)$ as the history of the number of complexes up till time t. The demodulation problem is to use the history $\mathcal{B}(t)$ to determine which symbol the transmitter has sent.

In Sections 4.1 to 4.3, we consider only one symbol interval and do not consider ISI. We consider the ISI case in Section 4.5.

4.1 The MAP framework

We adopt a MAP framework for detection. Let $\mathbf{P}[s|\mathcal{B}(t)]$ denote the posteriori probability that symbol s has been sent given the history $\mathcal{B}(t)$. If the demodulation decision is to be done at time t, then the demodulator decides that symbol $\hat{s}$ has been sent if

$$\hat{s} = \arg\max_{s=0,\ldots,K-1} \mathbf{P}[s|\mathcal{B}(t)] \tag{4.1}$$

Instead of working with $\mathbf{P}[s|\mathcal{B}(t)]$, we will work with its logarithm. Let

$$L_s(t) = \log(\mathbf{P}[s|\mathcal{B}(t)]) \tag{4.2}$$

The first step is to determine $L_s(t + \Delta t)$ from $L_s(t)$. Given $\mathcal{B}(t)$ is the section of $b(t)$ in the time interval $[0, t]$, one can consider $\mathcal{B}(t + \Delta t)$ as the concatenation of $\mathcal{B}(t)$ and the section of $b(t)$ in the time interval $(t, t + \Delta t]$. Over an infinitesimal Δt, we can consider the signal $b(t)$ is a constant in the time interval $(t, t + \Delta t]$; we therefore abuse the notation and use $b(t + \Delta t)$ to denote section of $b(t)$ in the time interval $[0, t]$. By using Bayes' rule, it can be shown that

$$L_s(t + \Delta t) = L_s(t) + \log(\mathbf{P}[b(t + \Delta t)|s, \mathcal{B}(t)]) - \log(\mathbf{P}[b(t + \Delta t)|\mathcal{B}(t)]) \tag{4.3}$$

where $\mathbf{P}[b(t + \Delta t)|s, \mathcal{B}(t)]$ is the probability that there are $b(t + \Delta t)$ complexes given that the transmitter has sent the symbol s and the previous history $\mathcal{B}(t)$. The last term on the RHS of (4.3), i.e. $\mathbf{P}[b(t + \Delta t)|\mathcal{B}(t)]$, is independent of the transmission symbol so we do not need it for the purpose of detection. We will focus on determining $\mathbf{P}[b(t + \Delta t)|s, \mathcal{B}(t)]$.

4.2 Computing $\mathbf{P}[b(t+\Delta t)|s,\mathcal{B}(t)]$

The problem of determining the probability $\mathbf{P}[b(t+\Delta t)|s,\mathcal{B}(t)]$ is essentially a Bayesian filtering or hidden Markov model problem. Recall that the complete state of the system is $(N(t),b(t))$ and the receiver can only observe $b(t)$, therefore the task of the receiver is to use the history $\mathcal{B}(t)$ and the system model to do prediction. Standard method can be used to derive $\mathbf{P}[b(t+\Delta t)|s,\mathcal{B}(t)]$ but the derivation is long, especially because of the diffusion terms; the derivation can be found in Appendix A. The result is

$$\begin{aligned}&\mathbf{P}[b(t+\Delta t)|s,\mathcal{B}(t)]\\ =&\delta_{b(t+\Delta t),b(t)+1}\ \lambda(M-b(t))\ \Delta t\ \mathbf{E}[n_R(t)|s,\mathcal{B}(t)]+\delta_{b(t+\Delta t),b(t)-1}\ \mu b(t)\ \Delta t+\\ &\delta_{b(t+\Delta t),b(t)}\ (1-\lambda(M-b(t))\mathbf{E}[n_R(t)|s,\mathcal{B}(t)]\ \Delta t-\mu b(t)\ \Delta t)\end{aligned}\tag{4.4}$$

Note that only one of the three terms on the RHS of Equation (4.4) is non-zero depending on whether the observed $b(t+\Delta t)$ is one more, one less or equal to that of $b(t)$; or, in other words, whether the number of complexes has increased by one, decreased by one or stayed the same. The term $\mathbf{E}[n_R(t)|s,\mathcal{B}(t)]$ is the expected number of signalling molecules in the receiver voxel given the history and the symbol s. The meaning of this term is that the receiver uses the history to predict what the expected number of signalling molecules in the receiver voxel is. Note that only the chemical kinetic parameters λ and μ of the receptor appear explicitly in Equation (4.4). Other parameters, such as the set of chemical reaction that generate Symbol s and the diffusion coefficient, do not appear explicitly in Equation (4.4) but influence the system behaviour via the term $\mathbf{E}[n_R(t)|s,\mathcal{B}(t)]$.

4.3 The demodulation filter

By substituting Equation (4.4) into Equation (4.3) and let Δt go to zero, we show in Appendix B that

$$\frac{dL_s(t)}{dt}=\frac{dU(t)}{dt}\log(\mathbf{E}[n_R(t)|s,\mathcal{B}(t)])-\lambda(M-b(t))\mathbf{E}[n_R(t)|s,\mathcal{B}(t)]+\tilde{L}(t)\tag{4.5}$$

with $L_s(0)$ initialised to the logarithm of the prior probability that Symbol s is sent. The term $U(t)$ is the cumulative number of times that the receptors have turned from the unbound to bound state. The meaning of $U(t)$ is illustrated in Figure C.3 assuming that there are two receptors. The top two pictures in Figure C.3 show the state transitions for the two receptors. The third picture shows the function $U(t)$ which is increased by one every time a receptor switches from the unbound to bound state. The bottom picture shows $\frac{dU(t)}{dt}$ which is the derivative of $U(t)$. Note that $\frac{dU(t)}{dt}$ consists of a train of impulses (or Dirac deltas) where the timings of the impulses are the times at which a receptor binding event occurs. Loosely speaking, one may also view $\frac{dU(t)}{dt}$ as $\max(\frac{db(t)}{dt},0)$.

The function $\tilde{L}(t)$, which is the last term on the RHS of (4.5), contains all the terms that are independent of Symbol s. Since $L_s(t)$ does not appear on the RHS of (4.5), this means that $\tilde{L}(t)$ adds the same contribution to all $L_s(t)$ for all $s=0,...,K-1$. We can therefore ignore $\tilde{L}(t)$ for the purpose of demodulation.

The term $\mathbf{E}[n_R(t)|s,\mathcal{B}(t)]$ in Equation (4.5) is the prediction of the mean number of signalling molecules in the receiver voxel using the history of receptor state. This is a filtering problem which requires extensive computation. Instead, we assume that the receiver has prior knowledge that if Symbol s is transmitted, then the mean number of signalling molecules in the receiver voxel is $\sigma_s(t)$ and the receiver uses $\sigma_s(t)$ for demodulation. We can view $\sigma_s(t)$ as internal models that the demodulator uses. The use of internal models is fairly common in signal processing and communication, e.g. a matched filter correlates the measured data with an expected response.

After making the modifications described in the last two paragraphs, we are now ready to describe the demodulator. Using $b(t)$ as the input, the demodulator runs the following K continuous-

time filters in parallel:

$$\frac{dZ_s(t)}{dt} = \frac{dU(t)}{dt}\log(\sigma_s(t)) - \lambda(M - b(t))\sigma_s(t) \tag{4.6}$$

where $Z_s(0)$ is initialised to the logarithm of the prior probability that the transmitter sends Symbol s. If the demodulator makes the decision at time t, then the demodulator decides that Symbol $\hat{s}$ has been transmitted if

$$\hat{s} = \arg\max_{s=0,\ldots,K-1} Z_s(t) \tag{4.7}$$

The demodulator structure is illustrated in Figure C.4. By comparing Equations (4.5) and (4.6), it can be shown that $L_{s_1}(t) - L_{s_2}(t) = Z_{s_1}(t) - Z_{s_2}(t)$ for any two symbols s_1 and s_2. An interpretation of the demodulation filter output $Z_s(t)$ is that $\exp(Z_s(t))$ is proportional to the posteriori probability $\mathbf{P}[s|\mathcal{B}(t)]$.

We see from Equation (4.6) that the calculation of the demodulator output requires a number of pieces of information. For the calculation of the first term on the RHS of Equation (4.6), it needs to know the time instances at which the receptor bindings occur and this timing is used to determine a contribution from the weighting function $\log(\sigma_s(t))$. The second term on the RHS of Equation (4.6) requires the number of unbound receptors at time t as well as the weighting function $\sigma_s(t)$.

In order to understand Equation (4.6), we consider the situation where Symbol 1 generates a lot more signalling molecules than Symbol 0 such that it results in more signalling molecules in the receiver voxel, or $\sigma_1(t) > \sigma_0(t)$ for all t. If the transmitter sends Symbol 1, then more signalling molecules are expected to reach the receiver voxel. The consequence is that there are more receptor binding events and the number of unbound receptors $(M - b(t))$ is smaller. Therefore, in Equation (4.6), we expect a big positive contribution from the first term on the RHS and a small negative contribution from the second term. The net effect is a big $Z_1(t)$. On the other hand, if the transmitter sends Symbol 0, the number of receptor binding events is smaller and $(M - b(t))$ is big. This results in a smaller $Z_0(t)$. Therefore, $Z_1(t)$ is likely to be bigger than $Z_0(t)$, which means correct detection.

4.4 Discussions

Implementation issues

The implementation of the demodulator is an open research problem. The demodulation filter (4.6) is an analogue filter and it requires the internal model $\sigma_i(t)$. The design of analogue circuits using molecules and chemical reactions for calculations is an active research area, see [51] for a recent overview. The demodulation filter requires logarithm, multiplication, subtraction, integration and counting the number of times the receptors have switched from the unbound to bound state. An analogue molecular circuit for calculating logarithm is presented in [52]. There are also circuits that can perform multiplication and subtraction [51]. Living cells are known to use integration [53]. It may be possible to implement the counting using a chemical reaction [35]. It may be possible to approximate the internal models by using some lower order chemical reactions. This discussion shows that some components to implement the demodulator exist but the exact implementation remains an open problem.

In order to bypass the difficulty of computing $\mathbf{E}[n_R(t)|s, \mathcal{B}(t)]$, we have proposed to use internal models $\sigma_s(t)$. An open research problem is to study sub-optimal estimation of $\mathbf{E}[n_R(t)|s, \mathcal{B}(t)]$.

Note that [51] shows that analogue computation is more efficient than digital computation if high precision is not required. An interesting problem is to study the impact of low precision analogue calculations on the demodulation performance.

Continuous transmission medium versus voxels

We mention in Section 2 that some papers in molecular communications assume a continuous medium (rather than discretising the medium into voxels) and a non-zero receiver size. If a

continuous medium is assumed, the state of a signalling molecule in the transmission medium is its position. Let $p(\vec{x}, t)$ be the probability that a molecule is at position $\vec{x}$ at time t. The time evolution of $p(\vec{x}, t)$ can again be modelled by a CTMP, which is continuous in both time and space. The time evolution of $p(\vec{x}, t)$ is described by the differential Chapman-Kolmogorov Equation (CKE) [36]. It is possible to derive an alternative CTMP by replacing the diffusion of signalling molecules in Equation (3.7) by differential CKE as well as by adding equations to describe how the signalling molecules enter or leave the receiver voxel. We can then apply Bayesian filtering to this alternative CTMP. We expect this alternative CTMP will give the same demodulator filter because in our derivation based on discrete voxels, the MAP demodulator depends only on the number of the signalling molecules in the receiver voxel and the diffusion parameters do not appear explicitly in the demodulation filter. Further support of this argument is given in the derivation in Appendix B where we show that the d_{ij} parameters in (3.7) are cancelled out in deriving the Bayesian filter.

4.5 ISI

It is in principle possible to use the demodulation filters (4.6) to deal with the case with ISI. We use T_x to denote one symbol duration and we assume that ISI only lasts for a finite amount of time. Specifically, consider a symbol sent in $[kT_x, (k+1)T_x]$, we assume that the effect of this transmission can be neglected after the time $(k+n)T_x$. This can be realised by appropriately choosing the transmitter and receiver parameters, T_x and n.

In order to make the explanation here a bit more concrete, we assume that the transmitter uses $K = 2$ symbols and $n = 3$. Over a duration of $n = 3$ symbols, the possible sequences sent by the transmitter are 000, 001, 010, 011, ... , 111. Let $\sigma_{0,0,0}(t)$ denote the mean number of signalling molecules at the receiver voxel if the sequence 000 is sent. We can similarly define $\sigma_{0,0,1}(t), ..., \sigma_{1,1,1}(t)$. Consider the transmission of three consecutive symbols s_{k-2}, s_{k-1} and s_k. Assuming that we have an estimation of the first two symbols $\hat{s}_{k-2}$ and $\hat{s}_{k-1}$, then the decoding of s_k can be done by using the demodulation filter (4.6) by replacing $\sigma_s(t)$ by $\sigma_{\hat{s}_{k-2},\hat{s}_{k-1},s}$. For example, if $s_{k-2} = 1$ and $s_{k-1} = 0$, then one can decode what s_k is by using the demodulator filters $\sigma_{1,0,1}(t)$ and $\sigma_{1,0,1}(t)$. Although the decision feedback based method can solve the ISI problem, the number of internal models increases exponentially with the memory length parameter n.

The reason why we need to consider all 2^n possible transmission sequences is that the ligand-receptor binding process has a non-linear reaction rate. A method to reduce the number of internal models is to design the system so that $\sigma_{0,0,0}(t)$ etc. can be decomposed into a sum. Let $\sigma_s(t)$ $(s = 0, 1)$ be the mean number of signalling molecules at the receiver voxel if the symbol s is sent for one symbol duration and in the absence of ISI. If

$$\sigma_{s_1,s_2,s_3}(t) \approx \sigma_{s_1}(t - 2T_x) + \sigma_{s_2}(t - T_x) + \sigma_{s_3}(t) \tag{4.8}$$

holds for all s_1, s_2 and s_3, then one can again make use of decision feedback to decode the ISI signal. However, this time, only K internal models are needed. Equation (4.8) can be made to hold approximately if the number of receptors is large. This can be explained as follows. First of all, if ligand-receptor binding is absent, this means there is only free diffusion then Equations (4.8) holds because the mean number of signalling molecules obeys the diffusion equation which is linear. This means that we need to create an environment that "looks like" free diffusion even when ligand-receptor binding is present. This can be realised if the number of signalling molecules that are bound to the receptors is small compared to those that are free. A method to achieve this is to increase the number of receptors. We will demonstrate this with a numerical example in Section 5. However, it is still an open problem to solve the ISI in the general case.

5 Properties of the demodulator

The aim of this section is to study the properties of the MAP demodulator numerically. We begin with the methodology.

5.1 Methodology

We consider a medium of $2\mu\text{m} \times 2\mu\text{m} \times 1\ \mu\text{m}$. We assume a voxel size of $(\frac{1}{3}\mu\text{m})^3$ (i.e. $W = \frac{1}{3}\ \mu\text{m}$), creating an array of $6 \times 6 \times 3$ voxels. The transmitter and receiver are located at (0.5,0.8,0.5) and (1.5,0.8,0.5) (in μm) in the medium. The voxel co-ordinates are (2,3,2) and (5,3,2) respectively.

We assume the diffusion coefficient D of the medium is 1 $\mu\text{m}^2\text{s}^{-1}$. The receptor parameters are $\tilde{\lambda} = 0.005\ \mu\text{m}^3\ \text{s}^{-1}$, $\lambda = \frac{\tilde{\lambda}}{W^3}$, and $\mu = 1\ \text{s}^{-1}$. These values are similar to those used in [45] and [27] [1]. We assume an absorbing boundary for the medium and the signalling molecules escape from a boundary voxel surface at a rate of $\frac{d}{50}$. The above parameter values will be used for all the numerical experiments.

For each experiment, the transmitter uses either $K = 2$ or $K = 3$ symbols. Each symbol is generated by a different sets of chemical reactions. Different experiments may use different sets of chemical reactions and will be described later. The number of receptors also varies between the experiments.

We use the Stochastic Simulation Algorithm (SSA) [54] to obtain realisations of $b(t)$ which is the number of complexes over time. SSA is a standard algorithm in chemistry to simulate diffusion and reactions; it is essentially an algorithm to simulate a CTMP.

In order to use Equation (4.6), we require the mean number of signalling molecules $\sigma_s(t)$ in the receiver voxel when Symbol s is sent. Unfortunately, it is not possible to analytically compute $\sigma_s(t)$ from the CTMP because of moment closure problem which arises when the transition rate is a non-linear function of the state [55]. We therefore resort to simulation to estimate $\sigma_s(t)$. Each time when we need an $\sigma_s(t)$, we run SSA simulation 500 times and average the results to obtain $\sigma_s(t)$. Note that these simulations are different from those that we use to generate $b(t)$ for the performance study. In other words, the simulations for estimating $\sigma_s(t)$ and for performance study are completely independent.

Once $b(t)$ and $\sigma_s(t)$ are obtained, we use numerical integration to calculate $Z_s(t)$ using Equation (4.6). We assume that all symbols appear with equal probability, so we initialise $Z_s(0) = 0$ for all s.

5.2 Properties of the demodulator output

For this experiment, we use $K = 2$ symbols and $M = 50$ receptors. Both Symbols 0 and 1 use a reaction of the form:

$$\text{RNA} \xrightarrow{\kappa} \text{RNA} + \text{S} \tag{5.1}$$

where the κ's for Symbols 0 and 1 causes, respectively, 40 and 80 molecules to be generated per second on average by the transmitter. The simulation time is about 3 seconds.

Figure C.5 shows the demodulation filter outputs $Z_0(t)$ and $Z_1(t)$ if the transmitter sends a Symbol 0. It can be seen that $Z_0(t) > Z_1(t)$ most of the time after $t = 1.2$, which means the detection is likely to be correct after this time. The sawtooth like appearance of $Z_0(t)$ and $Z_1(t)$ is due to the fact that every time when a receptor is bound, there is a jump in the filter output according to Equation (4.6). Figure C.6 shows the filter outputs $Z_0(t)$ and $Z_1(t)$ if the transmitter sends a Symbol 1; the behaviour is similar.

Figure C.7 shows the *mean* filter outputs $Z_0(t)$ and $Z_1(t)$ if the transmitter sends a Symbol 0. The mean is computed over 200 realisations of $b(t)$. It can be seen that the mean filter output of $Z_0(t)$ is greater than that of $Z_1(t)$. Similarly, if Symbol 1 is sent, then we expect of the mean of $Z_1(t)$ to be bigger. The figure is not shown for brevity.

Figure C.8 shows the mean symbol error rates (SERs) for Symbols 0 and 1 if the detection is done at time t. The SER for Symbol 1 is high initially but as more information is processed over time, the SER drops to a low value. This experiment shows that it is possible to use the analogue

[1] The paper [27] considers ligand-receptor binding in the chemical master equation setting. In our notation, the parameter values in [27] are $D = 100\ \mu\text{m}^2\text{s}^{-1}$, $\tilde{\lambda} = 0.2\ \mu\text{m}^3\ \text{s}^{-1}$ and $\mu = 10\ \text{s}^{-1}$. These parameters are 10–100 times faster than ours and can be considered as a time-scaling. Note that [27] uses k_+ and k_- instead of, respectively, $\tilde{\lambda}$ and μ.

demodulation filter (4.6) to compute a quantity that allows us to distinguish between two emission patterns at the receiver.

5.3 Impact of number of receptors

We continue with the setting of 5.2 but we vary the number of receptors between 1 and 20. We assume the demodulator makes the decision at $t = 2.5$ and calculate the mean SER for both symbols at $t = 2.5$. Figure C.9 plots the SERs versus the number of receptors. It can be seen that the SER drops with increasing number of receptors.

We have used $K = 2$ symbols so far. We retain the current Symbols 0 and 1, and add a Symbol 2 which is also of the form of Reaction (5.1) but its mean rate of production of signalling molecules is 3 times that of Symbol 0. The number of receptors M used are: 1, 10, 20, ..., 150. We compute the average SER at $t = 2.5$ assuming each symbol is transmitted with equal probability. We plot the logarithm of the average SER against $\log(M)$ in Figure C.10. It can be seen that the SER drops with increasing number of receptors M. The plot in Figure C.10 suggests that, when the number of receptors M is large, the relationship between logarithm of SER and $\log(M)$ is linear. We perform a least-squares fit for M between 50 and 150. The fitted straight line is shown in Figure C.10 and it has a slope of -1.13. A possible explanation is that, because the receptors are non-interacting, each receptor provides an independent observation. The empirical evidence suggests that the average SER scales according to $\frac{1}{M}$ asymptotically.

5.4 Distinguishability of different chemical reactions

Equation (4.6) suggests that if the transmitter uses two sets of reactions which have almost the same mean number of signalling molecules in the receiver voxel, then it may be difficult to distinguish between these two symbols. In this study, Symbol 0 is generated by Reaction (5.1) with a rate of κ while Symbol 1 is generated by:

$$[\mathrm{RNA}]_{\mathrm{ON}} \longleftrightarrow [\mathrm{RNA}]_{\mathrm{OFF}} \tag{5.2}$$

$$[\mathrm{RNA}]_{\mathrm{ON}} \xrightarrow{2\kappa} [\mathrm{RNA}]_{\mathrm{ON}} + \mathrm{S} \tag{5.3}$$

where we assume that RNA can be in an ON or OFF state, and signalling molecules S are only produced when the RNA is in the ON-state. We assume that the there is an equal probability for the RNA to be in the two states and the reaction rate constant for the production of signalling molecule S from $[\mathrm{RNA}]_{\mathrm{ON}}$ is 2κ. This means that the mean rate of production of signalling molecules S by Symbols 0 and 1 are the same. This gives rise to very similar $\sigma_0(t)$ and $\sigma_1(t)$. Figure C.11 shows the demodulation filter outputs $Z_0(t)$ and $Z_1(t)$ for one simulation. It can be seen that the two outputs are almost indistinguishable. Consequently, the SER is pretty high. This shows that symbols generated by reactions which have similar mean number of signalling molecules at the receiver voxel can be hard to distinguish.

5.5 ISI

The aim of this experiment is to study the performance of the demodulator in the presence of ISI. We use the decision feedback method described in Section 4.5 together with the approximation decomposition in (4.8). We vary the number of receptors from 25 to 150. We use two different memory lengths ℓ of 4 and 5. If the memory length is n, we express the mean number of output molecules at the current symbol interval as a sum of ℓ terms, i.e. a generalisation of (4.8) to ℓ terms. We carry out the simulation for a duration of 20 symbol lengths and compute the average SER over 20 symbols. Figure C.12 shows the average SER versus the number of receptors. It can be seen that an increasing number of receptors can also be used to deal with ISI.

6 Conclusions and future work

This paper studies a diffusion-based molecular communication network that uses different sets of chemical reactions to represent different transmission symbols. We focus on the demodulation problem. We assume the receiver uses a ligand-receptor binding process and uses the continuous history of the number of ligand-receptor complexes over time as the input signal to the demodulator. We derive the maximum a posteriori demodulator by solving a Bayesian filtering problem.

A Proof of Equation (4.4)

Let s denote the transmitted symbol, our aim is to determine $\mathbf{P}[b(t+\Delta t)|s, \mathcal{B}(t)]$ in terms of the quantity at time t. Recalling that $(N(t), b(t))$ is the state of the CTMP and since only $\mathcal{B}(t)$ is observed, the problem of predicting $b(t+\Delta t)$ from $\mathcal{B}(t)$ is a Bayesian filtering or hidden Markov model problem. The first step is to condition on the state of the system, as follows:

$$\mathbf{P}[b(t+\Delta t)|s, \mathcal{B}(t)] \tag{A.1}$$
$$= \sum_i \mathbf{P}[N(t+\Delta t) = \eta_i, b(t+\Delta t)|s, \mathcal{B}(t)] \tag{A.2}$$
$$= \sum_i \sum_j \mathbf{P}[N(t+\Delta t) = \eta_i, b(t+\Delta t)|s, N(t) = \eta_j, \mathcal{B}(t)]\mathbf{P}[N(t) = \eta_j|s, \mathcal{B}(t)] \tag{A.3}$$
$$= \sum_i \sum_j \mathbf{P}[N(t+\Delta t) = \eta_i, b(t+\Delta t)|s, N(t) = \eta_j, b(t)]\mathbf{P}[N(t) = \eta_j|s, \mathcal{B}(t)] \tag{A.4}$$

where we have used the Markov property $\mathbf{P}[N(t+\Delta t) = \eta_i, b(t+\Delta t)|s, N(t) = \eta_j, \mathcal{B}(t)] = \mathbf{P}[N(t+\Delta t) = \eta_i, b(t+\Delta t)|s, N(t) = \eta_j, b(t)]$ to arrive at Equation (A.4).

We now focus on the term $\mathbf{P}[N(t+\Delta t) = \eta_i, b(t+\Delta t)|s, N(t) = \eta_j, b(t)]$ in Equation (A.4). This term is the state transition probability. Using the CTMP in Section 3, we have

$$\mathbf{P}[N(t+\Delta t) = \eta_i, b(t+\Delta t)|s, N(t) = \eta_j, b(t)] \tag{A.5}$$
$$= \delta(b(t+\Delta t) = b(t)+1)P_1 + \delta(b(t+\Delta t) = b(t)-1)P_2 + \delta(b(t+\Delta t) = b(t))P_3$$

where

$$P_1 = \delta(\eta_i = \eta_j - \mathbb{1}_R)\lambda\eta_{j,R}(M - b(t))\ \Delta t \tag{A.6}$$
$$P_2 = \delta(\eta_i = \eta_j + \mathbb{1}_R)\mu b(t)\ \Delta t \tag{A.7}$$
$$P_3 = \eta_{i\to j} \tag{A.8}$$

where $\eta_{j,R}$ is the R-th element of η_j, i.e. there are $\eta_{j,R}$ signalling molecules in the receiver voxel, and

$$\eta_{i\to j} = \begin{cases} d_{ij}\ \Delta t & \text{if } i \neq j \\ 1 - (\lambda\eta_{j,R}(M - b(t)) - \mu b(t) - d_{jj})\ \Delta t & \text{if } i = j \end{cases} \tag{A.9}$$

where

$$d_{jj} = \sum_{i\neq j} d_{ij} \tag{A.10}$$

By substituting Equation (A.6) into Equation (A.4), we have

$$\mathbf{P}[b(t+\Delta t)|s, \mathcal{B}(t)] = \delta(b(t+\Delta t) = b(t)+1)Q_1 + \delta(b(t+\Delta t) = b(t)-1)Q_2 + \delta(b(t+\Delta t) = b(t))Q_3 \tag{A.11}$$

where

$$Q_\ell = \sum_i \sum_j P_\ell \mathbf{P}[N(t) = \eta_j|s, \mathcal{B}(t)] \tag{A.12}$$

We will now determine Q_1, Q_2 and Q_3.

For Q_1, we have

$$\begin{aligned}
Q_1 &= \sum_i \sum_j \delta(\eta_i = \eta_j - \mathbb{1}_R)\lambda\eta_{j,R}(M - b(t))\ \Delta t\ \mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] \\
&= \lambda(M - b(t))\ \Delta t\ \sum_i \sum_j \delta(\eta_i = \eta_j - \mathbb{1}_R)\eta_{j,R}\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] \\
&= \lambda(M - b(t))\ \Delta t \sum_{j\ s.t.\ \eta_{i,R} \geqslant 1} \eta_{j,R}\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] && \text{(A.13)} \\
&= \lambda(M - b(t))\ \Delta t\ \sum_j \eta_{j,R}\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] && \text{(A.14)} \\
&= \lambda(M - b(t))\ \Delta t\ \mathbf{E}[n_R(t) | s, \mathcal{B}(t)] && \text{(A.15)}
\end{aligned}$$

Note that in Equation (A.13), the sum is over all states η_i with at least one signalling molecule in the receiver voxel, i.e. $\eta_{j,R} \geqslant 1$. Since the summand in Equation (A.13) is zero if $\eta_{j,R} = 0$, we get the same result if we are to sum over all possible states, that is why Equation (A.14) holds.

For Q_2, we have

$$\begin{aligned}
Q_2 &= \sum_i \sum_j \delta(\eta_i = \eta_j + \mathbb{1}_R)\mu b(t)\ \Delta t\ \mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] \\
&= \mu b(t)\ \Delta t \sum_i \sum_j \delta(\eta_i = \eta_j + \mathbb{1}_R)\ \mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] && \text{(A.16)} \\
&= \mu b(t)\ \Delta t \sum_j \mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] && \text{(A.17)} \\
&= \mu b(t)\ \Delta t && \text{(A.18)}
\end{aligned}$$

Note that Equation (A.17) follows from Equation (A.16) because for every η_j, there is a unique η_i such that $\eta_i = \eta_j + \mathbb{1}_R$ holds.

For Q_3, we have

$$\begin{aligned}
Q_3 =& \sum_i \sum_j \eta_{i \to j}\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] \\
=& \sum_i \sum_{j \neq i} (d_{ij}\ \Delta t)\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] + \\
& \sum_j (1 - \lambda\eta_{j,R}(M - b(t))\ \Delta t - \mu b(t)\ \Delta t - d_{jj}\ \Delta t)\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] \\
=& \sum_j (1 - \lambda\eta_{j,R}(M - b(t))\ \Delta t - \mu b(t)\ \Delta t)\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] + \\
& (\sum_i \sum_{j \neq i} d_{ij}\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] - \sum_j d_{jj}\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)])\ \Delta t \\
=& (1 - \lambda(M - b(t))\mathbf{E}[n_R(t) | s, \mathcal{B}(t)]\ \Delta t - \mu b(t)\ \Delta t) + \\
& \underbrace{(\sum_i \sum_{j \neq i} d_{ij}\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)] - \sum_j \sum_{i \neq j} d_{ij}\mathbf{P}[N(t) = \eta_j | s, \mathcal{B}(t)])}_{=0}\ \Delta t \\
=& (1 - \lambda(M - b(t))\mathbf{E}[n_R(t) | s, \mathcal{B}(t)]\ \Delta t - \mu b(t)\ \Delta t) && \text{(A.19)}
\end{aligned}$$

Having obtained Q_1, Q_2 and Q_3, we arrive at:

$$\begin{aligned}
\mathbf{P}[b(t + \Delta t) | s, \mathcal{B}(t)] =& \delta(b(t + \Delta t) = b(t) + 1)\lambda(M - b(t))\ \Delta t\ \mathbf{E}[n_R(t) | s, \mathcal{B}(t)] + \\
& \delta(b(t + \Delta t) = b(t) - 1)\mu b(t)\ \Delta t + \\
& \delta(b(t + \Delta t) = b(t))(1 - \lambda(M - b(t))\mathbf{E}[n_R(t) | s, \mathcal{B}(t)]\ \Delta t - \mu b(t)\ \Delta t) && \text{(A.20)}
\end{aligned}$$

Note that Equation (A.20) is the same as Equation (4.4) in the main text.

B Proof of Equation (4.5)

From Equation (4.3), we have:

$$\frac{dL_s(t)}{dt} = \lim_{\Delta t \to 0} \frac{\log(\mathbf{P}[b(t+\Delta t)|s, \mathcal{B}(t)])}{\Delta t} - \lim_{\Delta t \to 0} \frac{\log(\mathbf{P}[b(t+\Delta t)|\mathcal{B}(t)])}{\Delta t} \tag{B.1}$$

Note that the second term on the RHS is independent of transmission symbol s, we will focus on the first term.

Note that $\mathbf{P}[b(t+\Delta t)|s, \mathcal{B}(t)]$, which is given in Equation (A.20), is a sum three terms with multipliers $\delta(b(t+\Delta t) = b(t)+1)$, $\delta(b(t+\Delta t) = b(t)-1)$ and $\delta(b(t+\Delta t) = b(t))$. Since these multipliers are mutually exclusive, we have:

$$\begin{aligned}
\log\left(\mathbf{P}[b(t+\Delta t)|s, \mathcal{B}(t)]\right) =& \delta(b(t+\Delta t) = b(t)+1) \log\left(\lambda(M - b(t))\ \Delta t\ \mathbf{E}[n_R(t)|s, \mathcal{B}(t)]\right) + \\
& \delta(b(t+\Delta t) = b(t)-1) \log\left(\mu b(t)\ \Delta t\right) + \\
& \delta(b(t+\Delta t) = b(t)) \log\left((1 - \lambda(M - b(t))\mathbf{E}[n_R(t)|1, \mathcal{B}(t)]\ \Delta t - \mu b(t)\ \Delta t\right)) \\
\approx& \delta(b(t+\Delta t) = b(t)+1) \log\left(\mathbf{E}[n_R(t)|s, \mathcal{B}(t)]\right) - \\
& \delta(b(t+\Delta t) = b(t))\lambda(M - b(t))\mathbf{E}[n_R(t)|s, \mathcal{B}(t)]\ \Delta t + \\
& \tilde{P}(t)
\end{aligned} \tag{B.2}$$

where we have used the approximation $\log(1+\alpha x) \approx \alpha x$ and have collected all terms that do not depend on s in $\tilde{P}(t)$.

By substituting Equation (B.2) into Equation (B.1), and taking limit $\Delta t \to 0$, we have

$$\begin{aligned}
\frac{dL_s(t)}{dt} =& \lim_{\Delta t \to 0} \frac{\delta(b(t+\Delta t) = b(t)+1)}{\Delta t} \log\left(\mathbf{E}[n_R(t)|s, \mathcal{B}(t)]\right) - \\
& \delta(b(t+\Delta t) = b(t))\lambda(M - b(t))\left(\mathbf{E}[n_R(t)|s, \mathcal{B}(t)]\right]) + \tilde{L}(t) \qquad \text{(B.3)} \\
=& \frac{dU(t)}{dt} \log\left(\mathbf{E}[n_R(t)|s, \mathcal{B}(t)]\right) - \lambda(M - b(t))\left(\mathbf{E}[n_R(t)|s, \mathcal{B}(t)]\right) + \tilde{L}(t) \qquad \text{(B.4)}
\end{aligned}$$

where all terms that are independent of s have been collected in $\tilde{L}(t)$. Note that $\tilde{L}(t)$ contains some terms that diverges but this is not an issue because for demodulation it is their relative difference $L_{s_1}(t) - L_{s_2}(t)$ (for any two symbols s_1 and s_2) that matters.

Note also that we have used the following reasonings to arrive at Equation (B.4) from Equation (B.3):

1. The term $\lim_{\Delta t \to 0} \frac{\delta(b(t+\Delta t)=b(t)+1)}{\Delta t}$ is an impulse whenever a receptor changes from the unbound to the bound state. This is precisely $\frac{dU(t)}{dt}$.
2. The term $\delta(b(t+\Delta t) = b(t))$ is only zero when the number of bound receptors changes and the number of such changes is finite. In other words, $\delta(b(t+\Delta t) = b(t)) = 1$ with probability one. This allows us to drop $\delta(b(t+\Delta t) = b(t))$.

Finally, note that Equation (B.4) is the same as Equation (4.5).

C An example end-to-end model

For this example, we assume the transmission medium consists of 3 voxels as illustrated in Figure C.13. The transmitter and receiver are assumed to be located in, respectively, Voxels 1 and 3. We assume reflecting boundary condition which means the signalling molecules cannot leave the medium.

This Appendix presents an example end-to-end model. For this end-to-end model, the transmitter is assumed to send Symbol 0 which means it uses the set of chemical reactions corresponding to this symbol. We therefore view a transmitter as a set of chemical reactions located within the

transmitter voxel. It is still an open problem what chemical reactions are good for communication performance. The example being used here is not meant to promote the use of a particular set of chemical reactions but our purpose is to show how a set of chemical reactions can be modelled by a CTMP.

For this example, we assume that the production of the signalling molecules S requires two intermediate chemical species F and G, which are produced by RNA_1 and RNA_2. There are four reactions and they are assumed to take place within the transmitter voxel only. The four chemical reactions are:

$$\mathrm{RNA}_1 \xrightarrow{k_1} \mathrm{RNA}_1 + \mathrm{F} \tag{C.1}$$

$$\mathrm{RNA}_2 \xrightarrow{k_2} \mathrm{RNA}_2 + \mathrm{G} \tag{C.2}$$

$$\mathrm{F} \xrightarrow{k_3} \mathrm{S} \tag{C.3}$$

$$\mathrm{S} + \mathrm{G} \xrightarrow{k_4} \phi \tag{C.4}$$

Reaction (C.1) says that the molecules of F are produced at a mean rate of k_1. Similarly, according to Reaction (C.2), G is produced at a mean rate of k_2. Reaction (C.3) says that F is converted to S at a mean rate equals to k_3 times the number of F molecules in the transmitter voxel. Reaction (C.4) says that S and G can react to produce a molecule ϕ that we are not interested to keep track of in the mathematical model. If an S (or a G) molecule takes part in Reaction (C.4), we can consider this S (G) molecule has left the system permanently after the reaction. The rate of Reaction (C.4) is k_4 times the number of G molecules and the number of signalling molecules S in the transmitter voxel.

We assume that the chemical species RNA_1, RNA_2, F and G are found in the transmitter voxel only, and they cannot leave the transmitter voxel. This means that we do not need to consider the diffusion of these chemical species. The only diffusible chemical species in the entire system is the signalling molecule S. We also assume that there is only one of each RNA_1 and RNA_2 and their counts remain constant.

In order to define the state of the system, we make the following definitions: $n_i(t)$ is the number of signalling molecules in Voxel i at time t, $n_F(t)$ and $n_G(t)$ are respectively the number of F and G molecules at time t, $n_A(t)$ is the cumulative number of molecules that have left the system at time t and $b(t)$ is the number of complexes (or bound receptors) at time t. Since a receptor can either be unbound or in a complex, the number of unbound receptors at time t is $M - b(t)$; therefore, the mathematical model only has to keep track of either the number of unbound receptors or the number of complexes, and we have chosen to keep track of the latter. The state of the system is completely specified by these seven molecular counts: $n_1(t), n_2(t), n_3(t), n_F(t), n_G(t), n_A(t)$ and $b(t)$. All the molecular counts should be non-negative integers (i.e. belonging to the set $\mathbb{Z}_{\geqslant 0}$) and a further restriction is that $0 \leqslant b(t) \leqslant M$ or we write $b(t) \in \mathbb{Z}_{[0,M]}$.

We define the vector $N(t)$ as

$$N(t) = \begin{bmatrix} n_1(t) & n_2(t) & n_3(t) & n_F(t) & n_G(t) & n_A(t) \end{bmatrix}^T \tag{C.5}$$

where the superscript ^{T} is used to denote matrix transpose.

Based on the definition of $N(t)$, the state of the system is the tuple $(N(t), b(t))$ and a valid state must be an element of the set $\mathcal{S} = \mathbb{Z}^6_{\geqslant 0} \times \mathbb{Z}_{[0,M]}$. The state of the system changes when a reaction or diffusion event occurs. Our modelling assumptions mean that reactions can only take place in the transmitter or the receiver voxels. The reactions in the transmitter voxel are (C.1)−(C.4). The reactions taking place in the receiver voxel are (3.2) and (3.3). The only diffusible chemical species in this system is the signalling molecule S. Within an infinitesimal time Δt, **at most one** diffusion or reaction event can occur. Therefore, the dynamics of the system can be specified by the transition probability from state $(N(t), b(t))$ to $(N(t + \Delta t), b(t + \Delta t))$. We will now specify these transition probabilities and we begin with the transmitter.

Four possible reaction events (C.1)–(C.4) can take place in the transmitter voxel. An occurrence of Reaction (C.1) increases the number of F molecules in the transmitter voxel by 1 and

this occurs at a mean rate of k_1. By defining $\mathbb{1}_i$ to be the standard basis vector with a '1' at the i-th position, we can write the state transition probability due to Reaction (C.1) as:

$$\mathbf{P}[N(t+\Delta t) = N(t) + \mathbb{1}_4, b(t+\Delta t) = b(t) | N(t), b(t)] = k_1 \, \Delta t \tag{C.6}$$

Note that we have used $\mathbb{1}_4$ because $n_F(t)$ is increased by 1 if Reaction (C.1) occurs and $n_F(t)$ is the fourth element of $N(t)$ in the definition of $N(t)$ in (C.5). The right-hand side (RHS) of Equation (C.6) is the transition probability that Reaction (C.1) occurs in $(t, t+\Delta t)$, which is given by the reaction rate k_1 times Δt.

We can write the transition probabilities due to Reactions (C.2)−(C.4) as:

$$\mathbf{P}[N(t+\Delta t) = N(t) + \mathbb{1}_5, b(t+\Delta t) = b(t) | N(t), b(t)] = k_2 \, \Delta t \tag{C.7}$$

$$\mathbf{P}[N(t+\Delta t) = N(t) - \mathbb{1}_4 + \mathbb{1}_1, b(t+\Delta t) = b(t) | N(t), b(t)] = k_3 \, n_F(t) \, \Delta t \tag{C.8}$$

$$\mathbf{P}[N(t+\Delta t) = N(t) - \mathbb{1}_5 - \mathbb{1}_1 + 2\mathbb{1}_6, b(t+\Delta t) = b(t) | N(t), b(t)] = k_4 \, n_G(t) \, n_1(t) \Delta t \tag{C.9}$$

The rationale behind Equation (C.7) is similar to that of (C.6). Equation (C.8) models Reaction (C.3). If Reaction (C.3) occurs, an F molecule is converted to an S molecule, so the number of F molecules $n_F(t)$ (which is the fourth element of $N(t)$) is decreased by 1 and the number of signalling molecule in the transmitter voxel $n_1(t)$ (which is the first element of $N(t)$) is increased by 1; this change in the number of molecules as a result of Reaction (C.3) can be written as $N(t+\Delta t) = N(t) - \mathbb{1}_4 + \mathbb{1}_1$ in (C.8). Equation (C.9) models Reaction (C.4). When Reaction (C.4) occurs, a G and an S molecule in the transmitter are consumed, hence $-\mathbb{1}_5 - \mathbb{1}_1$ in (C.9). We are not interested to keep track of the molecules as a result of this reaction, we consider these two molecules have left the system permanently and add '2' to $n_A(t)$ which is at the sixth position of $N(t)$. The letter 'A' here comes from 'absorbing' because once a molecule is added to $n_A(t)$, it will not leave. Note that the RHSs of (C.6)–(C.9) show the transition probabilities and they are of the form of the transition rate times Δt.

The state of the system can also be changed by signalling molecules diffusing from one voxel to another. For this example, there are four possible diffusion events, which take place when a signalling molecule diffuses from a voxel to its neighbouring voxel. The four diffusion events are: from Voxel 1 to Voxel 2, from Voxel 2 to Voxel 1, from Voxel 2 to Voxel 3, and from Voxel 3 to Voxel 2. The transition probabilities of these four events are:

$$\mathbf{P}[N(t+\Delta t) = N(t) - \mathbb{1}_1 + \mathbb{1}_2, b(t+\Delta t) = b(t) | N(t), b(t)] = d \, n_1(t) \, \Delta t \tag{C.10}$$

$$\mathbf{P}[N(t+\Delta t) = N(t) + \mathbb{1}_1 - \mathbb{1}_2, b(t+\Delta t) = b(t) | N(t), b(t)] = d \, n_2(t) \, \Delta t \tag{C.11}$$

$$\mathbf{P}[N(t+\Delta t) = N(t) - \mathbb{1}_2 + \mathbb{1}_3, b(t+\Delta t) = b(t) | N(t), b(t)] = d \, n_2(t) \, \Delta t \tag{C.12}$$

$$\mathbf{P}[N(t+\Delta t) = N(t) + \mathbb{1}_2 - \mathbb{1}_3, b(t+\Delta t) = b(t) | N(t), b(t)] = d \, n_3(t) \, \Delta t \tag{C.13}$$

Equation (C.10) is the probability that a signalling molecules diffuses from Voxel 1 to Voxel 2. The occurrence of this event means the number of signalling molecules in Voxel 1 ($= n_1(t)$, which is the first element of $N(t)$) is decreased by 1 while the number of signalling molecules in Voxel 2 ($= n_2(t)$, which is the second element of $N(t)$) is increased by 1. The probability of this occurring is $d \, \Delta t$. The explanation for the other three transition probabilities are similar.

The last category of state transitions occurs when a receptor is bound or unbound according to chemical reactions (3.2) and (3.3). The state transition probabilities are:

$$\mathbf{P}[N(t+\Delta t) = N(t) - \mathbb{1}_3, b(t+\Delta t) = b(t) + 1 | N(t), b(t)] = \lambda \, n_3(t) \, (M - b(t)) \, \Delta t \tag{C.14}$$

$$\mathbf{P}[N(t+\Delta t) = N(t) + \mathbb{1}_3, b(t+\Delta t) = b(t) - 1 | N(t), b(t)] = \mu \, b(t) \, \Delta t \tag{C.15}$$

Equation (C.14) is the transition probability for receptor binding or the formation of new complex. This event occurs when a signalling molecule in the receiver voxel reacts with a unbound receptor to form a complex. As a result of this reaction, the number of signalling molecules in the receiver voxel (which is Voxel 3 in this example) is decreased by 1 and the number of complexes $b(t)$ is increased by 1. The rate of this event is proportional to the product of the number of signalling

molecules in the receiver voxel $n_3(t)$ and the number of unbound receptors $(M - b(t))$. Equation (C.15) is the transition probability for a receptor to unbind. The unbinding reaction causes the number of signalling molecules in the receiver voxel $n_3(t)$ to increase by 1 while the number of complexes $b(t)$ to decrease by 1. The rate of this reaction is proportional to number of complexes $b(t)$.

Equations (C.6) to (C.15) give the transition probabilities of the possible events that can occur when the state of the system is $(N(t), b(t))$. It is possible that no transitions occurs in the time interval $(t, t + \Delta t)$, the probability of this occurring is given by the complementary to that of an event occurring, that is, one minus the sum of the RHSs of Equations (C.6) to (C.15). This completes the model.

Bibliography

[1] I. Akyildiz, F. Brunetti, and C. Blázquez, "Nanonetworks: A new communication paradigm," *Computer Networks*, vol. 52, pp. 2260–2279, 2008.

[2] S. Hiyama and Y. Moritani, "Molecular communication: Harnessing biochemical materials to engineer biomimetic communication systems," *Nano Communication Networks*, vol. 1, no. 1, pp. 20–30, May 2010.

[3] T. Nakano, M. J. Moore, F. Wei, A. V. Vasilakos, and J. Shuai, "Molecular Communication and Networking: Opportunities and Challenges," *IEEE Transactions on Nanobioscience*, vol. 11, no. 2, pp. 135–148, 2012.

[4] T. Nakano, T. Suda, Y. Okaie, M. J. Moore, and A. V. Vasilakos, "Molecular Communication Among Biological Nanomachines: A Layered Architecture and Research Issues," *NanoBioscience, IEEE Transactions on*, vol. 13, no. 3, pp. 169–197, 2014.

[5] B. Atakan, O. Akan, and S. Balasubramaniam, "Body area nanonetworks with molecular communications in nanomedicine," *Communications Magazine, IEEE*, vol. 50, no. 1, pp. 28–34, 2012.

[6] B. Atakan and O. B. Akan, "Deterministic capacity of information flow in molecular nanonetworks," *Nano Communication Networks*, vol. 1, no. 1, pp. 31–42, May 2010.

[7] M. Mahfuz, D. Makrakis, and H. Mouftah, "On the characterization of binary concentration-encoded molecular communication in nanonetworks," *Nano Communication Networks*, vol. 1, pp. 289–300, 2010.

[8] H. ShahMohammadian, G. G. Messier, and S. Magierowski, "Optimum receiver for molecule shift keying modulation in diffusion-based molecular communication channels," *Nano Communication Networks*, vol. 3, no. 3, pp. 183–195, Sep. 2012.

[9] M. Kuran, H. Yilmaz, T. Tugcu, and I. Akyildiz, "Modulation Techniques for Communication via Diffusion in Nanonetworks," in *Communications (ICC), 2011 IEEE International Conference on*, 2011, pp. 1–5.

[10] S. Kadloor, R. S. Adve, and A. W. Eckford, "Molecular Communication Using Brownian Motion With Drift," *NanoBioscience, IEEE Transactions on*, vol. 11, no. 2, pp. 89–99, 2012.

[11] C. T. Chou, "Molecular circuits for decoding frequency coded signals in nano-communication networks," *Nano Communication Networks*, vol. 3, no. 1, pp. 46–56, 2012.

[12] Q. Ni, A. Ganesan, N.-N. Aye-Han, X. Gao, M. D. Allen, A. Levchenko, and J. Zhang, "Signaling diversity of PKA achieved via a Ca2+-cAMP-PKA oscillatory circuit," *Nature Chemical Biology*, vol. 7, no. 1, pp. 34–40, Nov. 2010.

[13] A. Jovic, B. Howell, M. Cote, S. M. Wade, K. Mehta, A. Miyawaki, R. R. Neubig, J. J. Linderman, and S. Takayama, "Phase-Locked Signals Elucidate Circuit Architecture of an Oscillatory Pathway," *PLoS Computational Biology*, vol. 6, no. 12, p. e1001040, Dec. 2010.

[14] F. Tostevin and P. ten Wolde, "Mutual information in time-varying biochemical systems," *Physical Review E*, vol. 81, no. 6, Jun. 2010.

[15] H. ShahMohammadian, G. G. Messier, and S. Magierowski, "Nano-machine molecular communication over a moving propagation medium," *Nano Communication Networks*, vol. 4, no. 3, pp. 142–153, Sep. 2013.

[16] R. Mosayebi, H. Arjmandi, A. Gohari, M. Nasiri-Kenari, and U. Mitra, "Receivers for Diffusion-Based Molecular Communication: Exploiting Memory and Sampling Rate," *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 12, pp. 2368–2380, 2014.

[17] A. Noel, K. C. Cheung, and R. Schober, "Optimal Receiver Design for Diffusive Molecular Communication With Flow and Additive Noise," *IEEE Transactions on Nanobioscience*, vol. 13, no. 3, pp. 350–362, Sep. 2014.

[18] M. U. Mahfuz, D. Makrakis, and H. T. Mouftah, "Strength-based optimum signal detection in concentration-encoded pulse-transmitted OOK molecular communication with stochastic ligand-receptor binding," *Simulation Modelling Practice and Theory*, vol. 42, pp. 189–209, 2014.

[19] A. Einolghozati, M. Sardari, and F. Fekri, "Capacity of diffusion-based molecular communication with ligand receptors." *ITW*, pp. 85–89, 2011.

[20] A. Einolghozati, M. Sardari, A. Beirami, and F. Fekri, "Capacity of discrete molecular diffusion channels." *ISIT*, pp. 723–727, 2011.

[21] P. J. Thomas and A. W. Eckford, "Capacity of a Simple Intercellular Signal Transduction Channel," *arXiv.org*, Nov. 2014.

[22] A. Noel, K. C. Cheung, and R. Schober, "Improving Receiver Performance of Diffusive Molecular Communication With Enzymes," *NanoBioscience, IEEE Transactions on*, vol. 13, no. 1, pp. 31–43, 2014.

[23] D. Kilinc and O. B. Akan, "Receiver Design for Molecular Communication," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 12, pp. 705–714, Dec. 2013.

[24] L.-S. Meng, P.-C. Yeh, K.-C. Chen, and I. F. Akyildiz, "On Receiver Design for Diffusion-Based Molecular Communication," *IEEE Transactions on Signal Processing*, vol. 62, no. 22, pp. 6032–6044, Nov. 2014.

[25] M. Pierobon, "A Molecular Communication System Model Based on Biological Circuits," in *ACM The First Annual International Conference on Nanoscale Computing and Communication*. New York, New York, USA: ACM Press, 2014, pp. 1–8.

[26] C. T. Chou, "Molecular communication networks with general molecular circuit receivers," in *ACM The First Annual International Conference on Nanoscale Computing and Communication*. New York, New York, USA: ACM Press, 2014, pp. 1–9.

[27] M. Pierobon and I. F. Akyildiz, "Noise Analysis in Ligand-Binding Reception for Molecular Communication in Nanonetworks," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4168–4182, 2011.

[28] D. T. Gillespie, A. Hellander, and L. R. Petzold, "Perspective: Stochastic algorithms for chemical kinetics," *The Journal of Chemical Physics*, vol. 138, no. 17, pp. 170 901–170 901–14, 2013.

[29] M. J. Lawson, B. Drawert, M. Khammash, L. Petzold, and T.-M. Yi, "Spatial Stochastic Dynamics Enable Robust Cell Polarization," *PLoS Computational Biology*, vol. 9, no. 7, p. e1003139, Jul. 2013.

[30] C. T. Chou, "Extended master equation models for molecular communication networks," *IEEE Transactions on Nanobioscience*, vol. 12, no. 2, pp. 79–92, 2013.

[31] ——, "Noise properties of linear molecular communication networks," *Nano Communication Networks*, vol. 4, pp. 87–97, 2013.

[32] ——, "Impact of receiver reaction mechanisms on the performance of molecular communication networks," *IEEE Transactions on Nanotechnology*, 2015.

[33] K. Burrage, P. Burrage, A. Leier, T. Marquez-Lago, and D. Nicolau, Jr, "Stochastic simulation for spatial modelling of dynamic processes in a living cell," *Design and Analysis of Biomolecular Circuits*, p. 43, 2011.

[34] S. Hellander, A. Hellander, and L. Petzold, "Reaction-diffusion master equation in the microscopic limit," *Physical Review E*, vol. 85, no. 4, p. 042901, 2012.

[35] E. D. Siggia and M. Vergassola, "Decisions on the fly in cellular sensory systems." *Proceedings of the National Academy of Sciences*, vol. 110, no. 39, pp. E3704–12, Sep. 2013.

[36] C. Gardiner, *Stochastic methods.* Springer, 2010.

[37] N. van Kampen, Ed., *Stochastic Processes in Physics and Chemistry*, 3rd ed. North-Holland, 2007.

[38] R. Erban, J. Chapman, and P. Maini, "A practical guide to stochastic simulations of reaction-diffusion processes," *arXiv preprint arXiv:0704.1908*, 2007.

[39] S. Isaacson and C. Peskin, "Incorporating diffusion in complex geometries into stochastic chemical kinetics simulations," *SIAM Journal on Scientific Computing*, vol. 28, no. 1, pp. 47–74, 2007.

[40] S. Engblom, L. Ferm, A. Hellander, and P. Lötstedt, "Simulation of Stochastic Reaction-Diffusion Processes on Unstructured Meshes," *SIAM Journal on Scientific Computing*, vol. 31, no. 3, Feb. 2009.

[41] G. Tkačik and A. M. Walczak, "Information transmission in genetic regulatory networks: a review," *Journal of Physics: Condensed Matter*, vol. 23, no. 15, p. 153102, Apr. 2011.

[42] V. Shahrezaei and P. S. Swain, "Analytical distributions for stochastic gene expression," *Proceedings of the National Academy of Sciences*, vol. 105, no. 45, pp. 17 256–17 261, Nov. 2008.

[43] D. Gillespie, "A rigorous derivation of the chemical master equation," *Physica A: Statistical Mechanics and its Applications*, vol. 188, pp. 404–425, 1992.

[44] B. Alberts *et al.*, *Molecular Biology of the Cell*, 5th ed. Garland Science, 2007.

[45] R. Erban and S. J. Chapman, "Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions," *Physical Biology*, Mar. 2009.

[46] C. T. Chou, "A markovian approach to the optimal demodulation of diffusion-based molecular communication networks," arXiv, Tech. Rep., 2015, arXiv:XXX.

[47] M. Pierobon and I. F. Akyildiz, "Diffusion-based Noise Analysis for Molecular Communication in Nanonetworks," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2532–2547, 2011.

[48] B. Munsky and M. Khammash, "The finite state projection algorithm for the solution of the chemical master equation," *The Journal of Chemical Physics*, vol. 124, p. 044104, 2006.

[49] D. Y. Barsakcioglu, Y. Liu, P. Bhunjun, J. Navajas, A. Eftekhar, A. Jackson, R. Quian Quiroga, and T. G. Constandinou, "An Analogue Front-End Model for Developing Neural Spike Sorting Systems," *Biomedical Circuits and Systems, IEEE Transactions on*, vol. 8, no. 2, pp. 216–227, 2014.

[50] S. Haykin, P. Yee, and E. Derbez, "Optimum nonlinear filtering," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2774–2786, 1997.

[51] R. Sarpeshkar, "Analog synthetic biology," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 372, no. 2012, p. 20130110, Mar. 2014.

[52] R. Daniel, J. R. Rubens, R. Sarpeshkar, and T. K. Lu, "Synthetic analog computation in living cells," *Nature*, vol. 497, no. 7451, pp. 619–623, May 2013.

[53] T. Yi, Y. Huang, M. Simon, and J. Doyle, "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 9, p. 4649, 2000.

[54] D. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *The journal of physical chemistry*, 1977.

[55] P. Smadbeck and Y. N. Kaznessis, "A closure scheme for chemical master equations," *Proceedings of the National Academy of Sciences*, vol. 110, no. 35, pp. 14 261–14 265, Aug. 2013.

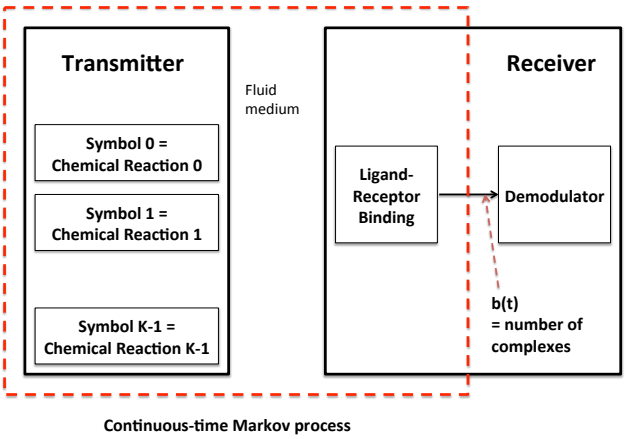

Figure C.1: An overview of the system considered in this paper.

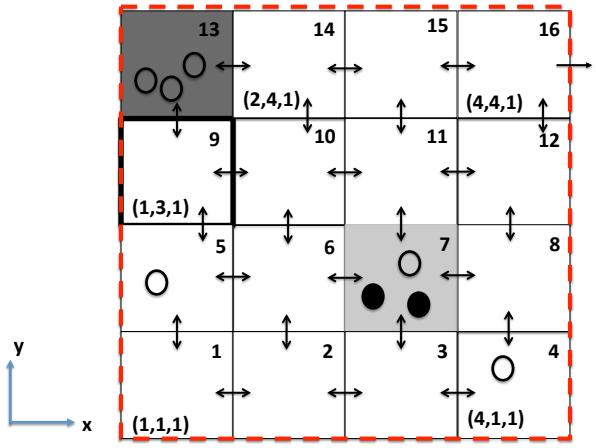

Figure C.2: A model of molecular communication network. The volume is divided into voxels. The indices of the voxels are given in the top right hand corner. Unfilled circles are signalling molecules. Filled circles are receptors. The dashed lines show the boundary of the medium.

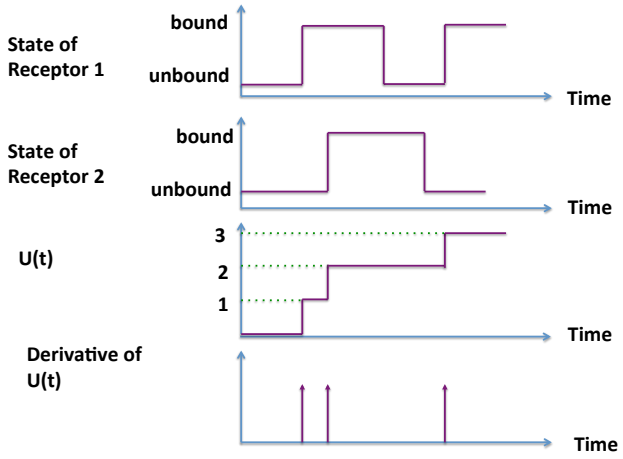

Figure C.3: This figure explains the meaning of the function $U(t)$, which is the total number of unbound-to-bound transitions for all receptors.

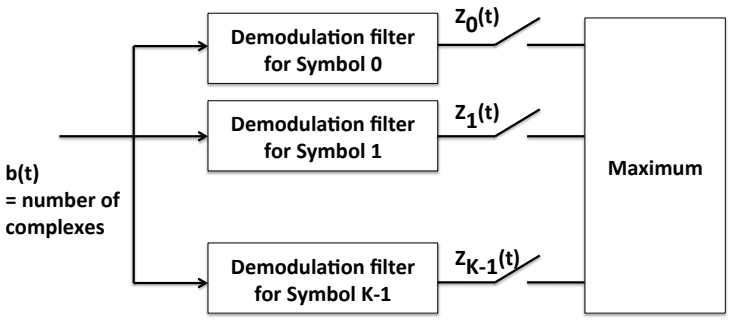

Figure C.4: The demodulator structure.

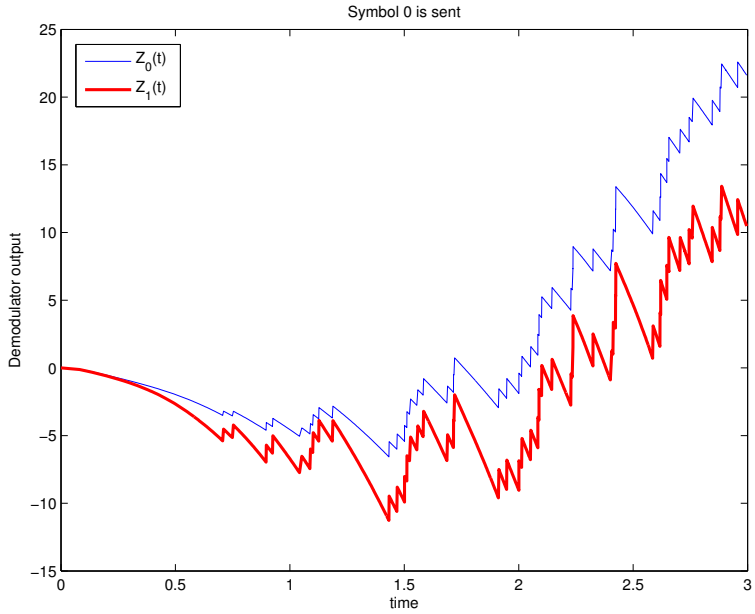

Figure C.5: The output of the modulators $Z_0(t)$ (thin line) and $Z_1(t)$ (thick line) for Symbol 0.

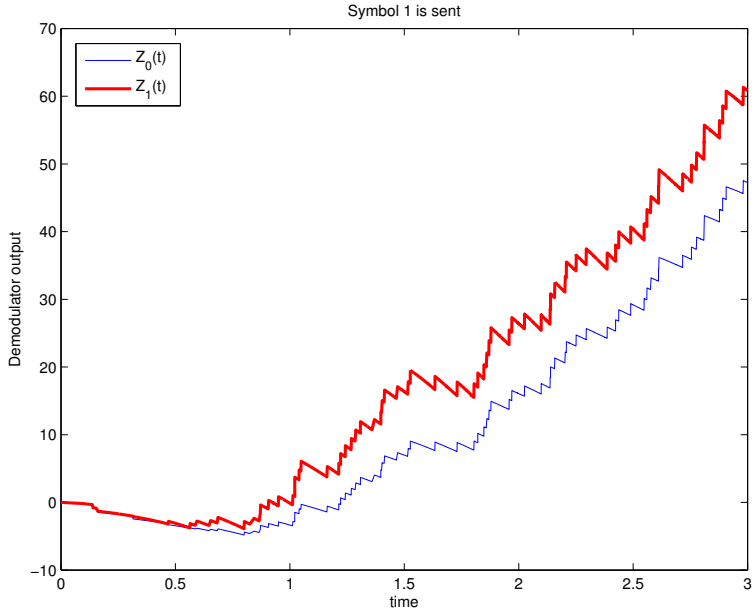

Figure C.6: The output of the modulators $Z_0(t)$ (thin line) and $Z_1(t)$ (thick line) for Symbol 1.

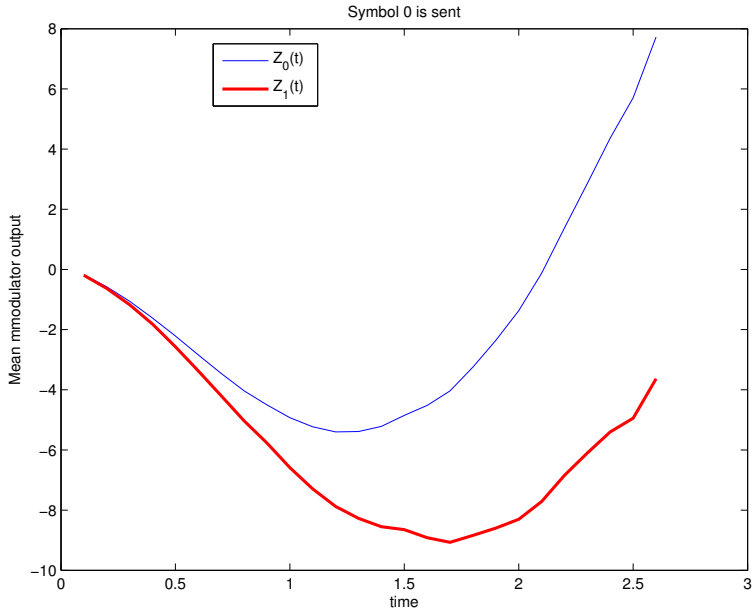

Figure C.7: The mean output of the modulators $Z_0(t)$ (thin line) and $Z_1(t)$ (thick line) for Symbol 0.

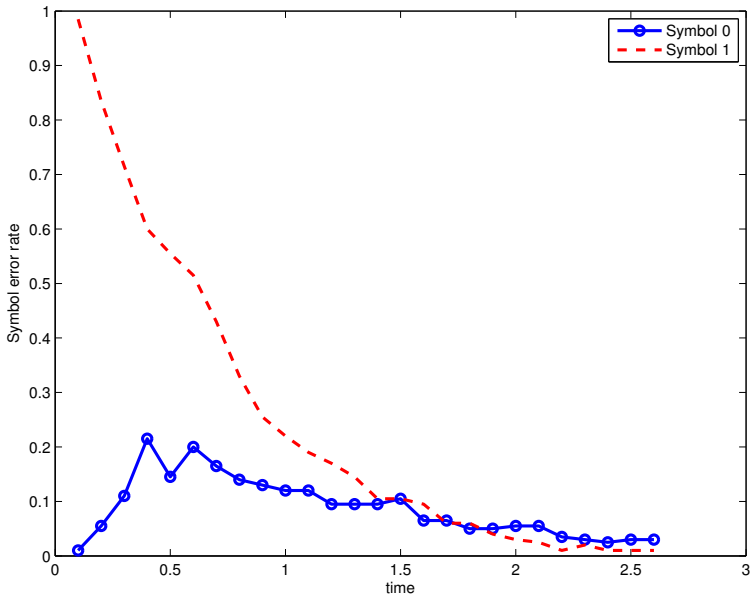

Figure C.8: The SER for Symbols 0 and 1.

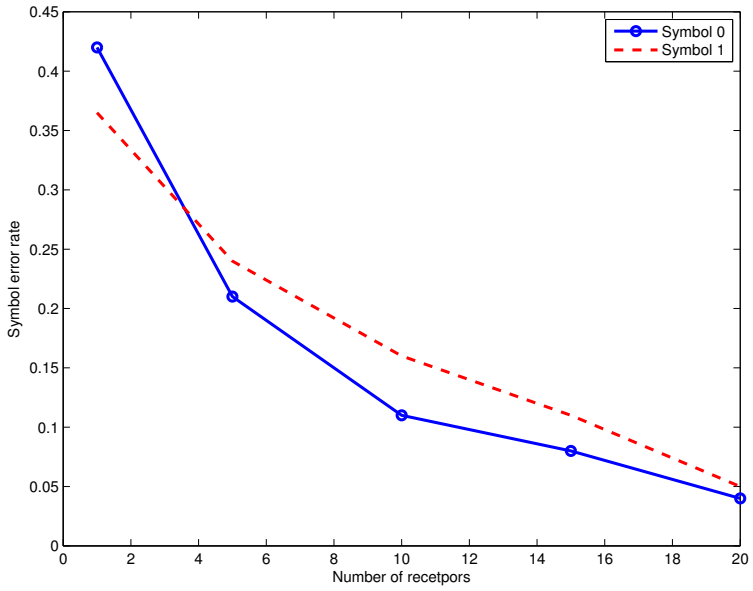

Figure C.9: The SER for Symbols 0 and 1 for varying number of receptors.

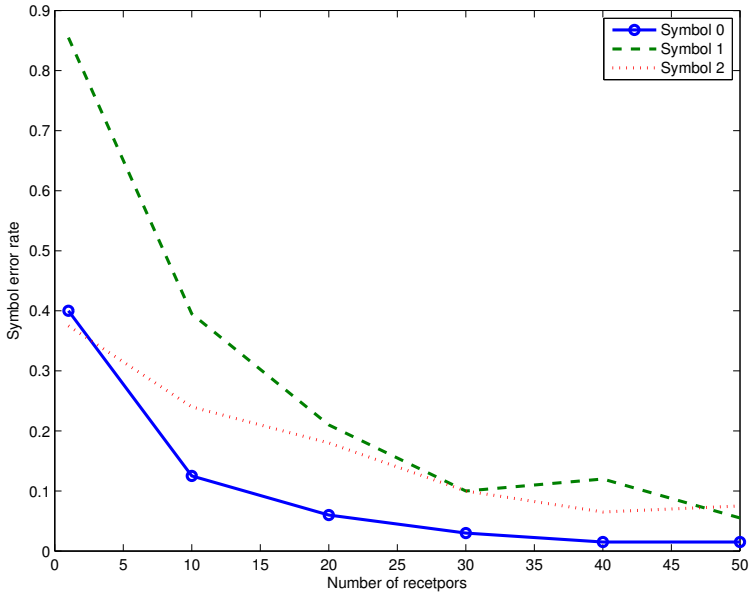

Figure C.10: The SER for Symbols 0, 1 and 2 for varying number of receptors.

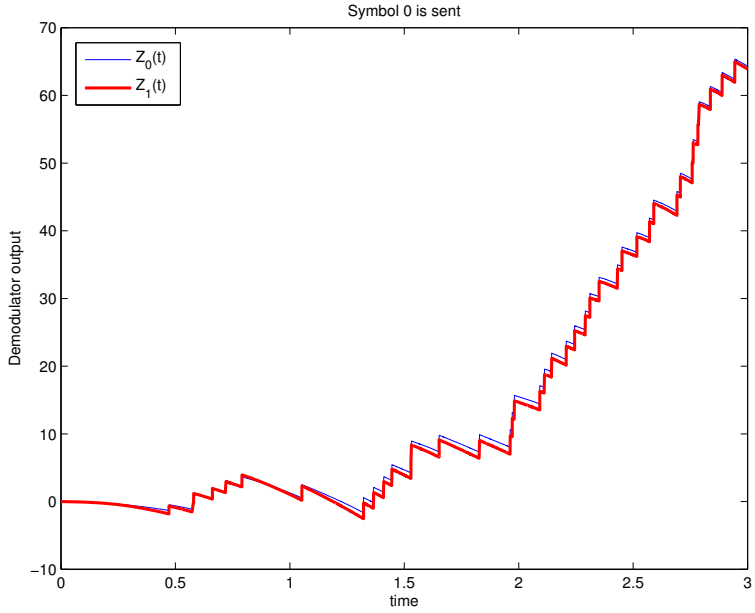

Figure C.11: The output of the modulators $Z_0(t)$ (thin line) and $Z_1(t)$ (thick line) for Symbol 0. The mean number of signalling molecules at the receiver voxel for both symbols is similar.

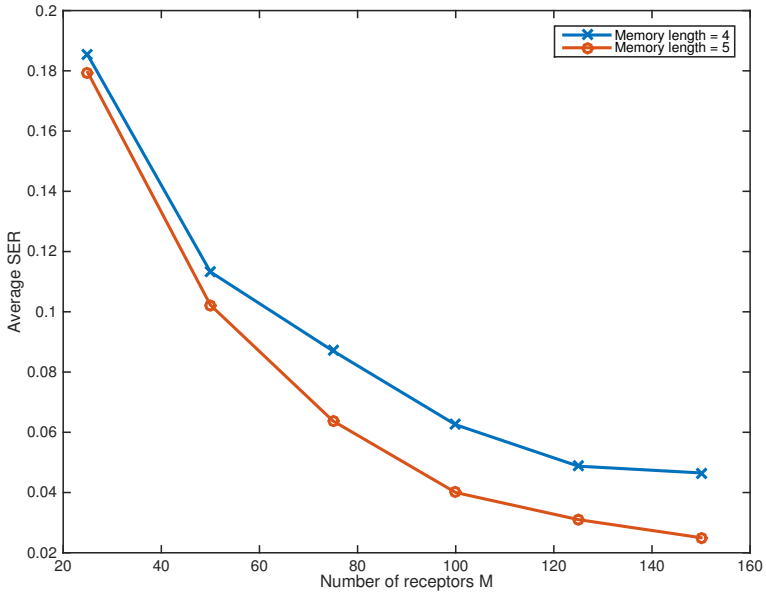

Figure C.12: The average SER versus number of receptors. The ISI case.

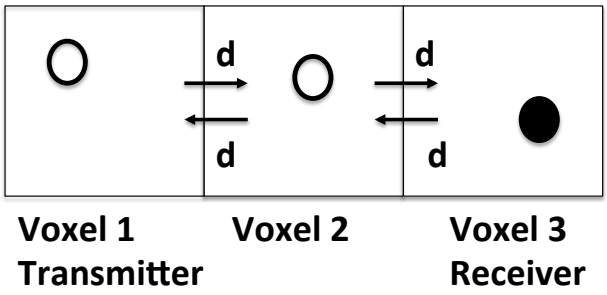

Figure C.13: An example system consisting of 3 voxels. The indices of the voxels are as indicated. Unfilled circles are signalling molecules. Filled circles are receptors.