



HAL
open science

Stochastic Reaction and Diffusion Systems in Molecular Communications: Recent Results and Open Problems

Malcolm Egan, Bayram Cevdet Akdeniz, Bao Quoc Tang

► **To cite this version:**

Malcolm Egan, Bayram Cevdet Akdeniz, Bao Quoc Tang. Stochastic Reaction and Diffusion Systems in Molecular Communications: Recent Results and Open Problems. Digital Signal Processing, In press, pp.1-38. 10.1016/j.dsp.2021.103117 . hal-03233615

HAL Id: hal-03233615

<https://hal.science/hal-03233615>

Submitted on 25 May 2021

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

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stochastic Reaction and Diffusion Systems in Molecular Communications: Recent Results and Open Problems

Malcolm Egan, Bayram Cevdet Akdeniz and Bao Quoc Tang

Abstract

Chemical reactions and diffusion are two basic mechanisms governing the dynamics of molecules in a fluid. As such, they play a critical role in molecular communication for channel modeling, design of detection rules, implementation of molecular circuits for computation, and modeling interactions with external biochemical systems. For finite numbers of information-carrying molecules, stochastic models naturally arise with the simplest example given by the Wiener process, often known as Brownian motion. Nevertheless, the Wiener process fails to be accurate when external forces, friction, and chemical reactions are present. Recently, there have been several contributions that tailor molecular communication systems to these more challenging channel conditions. In this paper, we first overview a general family of stochastic models of reaction and diffusion systems, including both Langevin diffusion and the reaction-diffusion master equation. These models form a basis for the use of these models as molecular communication channels, from which modulation and detection schemes can be developed. We survey recent results on the design of these schemes, with a focus on a recently developed approach which is robust to a wide range of channel models, known as equilibrium signaling. We then turn to the implementation of these detection schemes and related parameter estimation problems via stochastic molecular circuits, based on stochastic chemical reaction networks. Finally, interactions between molecular communication systems and stochastic biological systems as well as open problems are discussed. Our overarching goal is to highlight how the consideration of general stochastic models of reaction and diffusion can be utilized in order to widen the application of molecular communications both within engineered systems, and also as motivation for advances in the mathematical characterization of these models.

Contents

1	Introduction	2
1.1	Main Contributions	3
1.2	Related Surveys	5
1.3	Organization	5
1.4	Notation	6
2	Langevin Diffusion	7
2.1	Dynamics of a Single Molecule	7

2.2	Multiple Non-Interacting Molecules	9
2.3	Simulation	11
3	Stochastic Reaction-Diffusion Models	12
3.1	The Chemical Master Equation and Chemical Reaction Networks	12
3.2	Reaction-Diffusion Master Equation	15
4	Modulation and Detection for Channels Governed by Stochastic Models	19
4.1	Detection in General Channels	20
4.2	Detection in RDME Channels	20
5	Inference via Stochastic Reaction-Diffusion Systems	23
5.1	Receiver Design	23
5.2	Parameter Estimation	24
6	Stochastic Biological Environments	25
7	Open Problems	26
7.1	Detection	26
7.2	Stochastic Molecular Circuits	27
7.3	External Biochemical Systems	27
7.4	Distributed Spectroscopy	28
8	Conclusions	29

1 **1. Introduction**

2 A basic feature of a wide range of biochemical systems is the presence of chemical signals, which can be
3 viewed as facilitating information exchange, or *communication*. A prototypical example is quorum sensing
4 [1], where bacteria cells coordinate their behavior based on local estimates of the quantity of proteins emitted
5 by each cell into a common fluid environment. Other examples include inter-cell communication via calcium
6 signaling [2] or in the development of the fruit fly *Drosophila melanogaster* via protein gradients [3].

7 In 2005, the notion of *molecular communications* for synthetic or artificial devices was introduced in [4]. In
8 contrast to chemical signals in natural biological systems, artificial molecular communication was developed
9 in order to support communication between nanoscale devices, which may act as sensors or actuators. The
10 potential applications of this idea in areas from healthcare to the environment has motivated a large body
11 of literature over the intervening 15 years, including the monograph [5].

12 It was quickly realized that a key challenge in molecular communications is uncertainty or randomness
13 in the motion of the molecules comprising the chemical signal. Often, the motion of information-carrying

14 molecules is driven by diffusion or convection within a fluid, and these molecules are larger than the other
15 molecules comprising the fluid; that is, they are colloidal. It is well known from statistical physics that the
16 motion of colloidal molecules depends on many factors including the composition of the fluid, the possibility
17 of chemical reactions, the presence of external forces, molecular degradation, the presence of obstacles, and
18 the structure of the fluid boundary [6].

19 One of the simplest stochastic models for molecular motion is the Wiener process, with non-interacting
20 information carrying molecules, which lies in the general family of Smoluchowski equations that in turn lie
21 in the family of Langevin diffusions [6, 7]. One advantage of this approach is that, under certain conditions,
22 one can obtain explicit solutions, which are desirable for the design of modulation and detection schemes.

23 Indeed, using the Wiener process model, there has been extensive work developing novel modulation and
24 detection schemes, as well as parameter estimation and channel coding (e.g., [8, 9, 10, 11]). Techniques from
25 information theory have also provided insights into the fundamental performance limits of these schemes
26 (e.g., [12, 13, 14]). Various boundary conditions have also been investigated in order to model reflection or
27 absorption, which can capture the effect of certain kinds of chemical reactions (e.g., [15]).

28 Nevertheless, the Wiener process does not capture all environments relevant for molecular communica-
29 tions, for example, the presence of external forces or general forms of chemical reactions between molecules in
30 the system. For such cases, more general stochastic reaction-diffusion systems provide a powerful framework
31 to develop implementations of fundamental signal processing algorithms and model biochemical systems.
32 Naturally, the additional complexity of these models makes it almost impossible to have any closed form
33 solutions. Instead, recent mathematical advances in the characterization of reaction-diffusion systems (e.g.,
34 [16]) play a key role in obtaining accurate and tractable approximations.

35 *1.1. Main Contributions*

36 This paper reviews recent contributions addressing these issues by introducing sophisticated stochastic
37 models for the motion of information-carrying molecules and also exploiting these models for computation.
38 We start by reviewing two key families of stochastic models: Langevin diffusion and the reaction-diffusion
39 master equation (RDME). The former model, in the form of the Smoluchowski equation, can be used to
40 describe motion of a single molecule or multiple non-interacting molecules. In particular, for the Smolu-
41 chowski equation, when the external force is induced by a bounded scalar potential, a closed form expression
42 is available for the equilibrium solution, which makes it straightforward to obtain the asymptotic statistics
43 of the system. We also briefly describe the Euler-Maruyama method [17] to obtain realizations of sample
44 paths (i.e., the realization of the trajectory for a single molecule) for the Smoluchowski equation. However,
45 in general situations, where chemical reactions are either dominant or non-negligible, the Langevin diffusion
46 model is not an accurate description for the dynamics of the system.

47 A first alternative is the chemical master equation (CME), a stochastic description of chemical reactions,
48 whose (large time) behavior has been developed using tools from chemical reaction network theory [18].

49 More precisely, when the chemical reaction network satisfies condition known as complex balance, then a
 50 stationary distribution can be determined explicitly using the complex balanced equilibrium. To capture
 51 *both* diffusion and chemical reactions, a second alternative called the *reaction-diffusion master equation*
 52 should be considered. The complexity of this model makes it more challenging to apply it to molecular
 53 communication. Nevertheless, there is a remarkable—but not, in general, completely rigorous—connection
 54 between this model and deterministic reaction-diffusion systems in the form of limit theorems [19, 20]. The
 55 qualitative and quantitative behavior of such deterministic reaction-diffusion systems have recently witnessed
 56 significant advances in the mathematics literature [16]. This opens up the opportunity to utilize these recent
 57 results in molecular communication, and, at the same time, requires new rigorous, analytical studies of the
 58 relationship between deterministic reaction-diffusion systems and their stochastic counterpart.

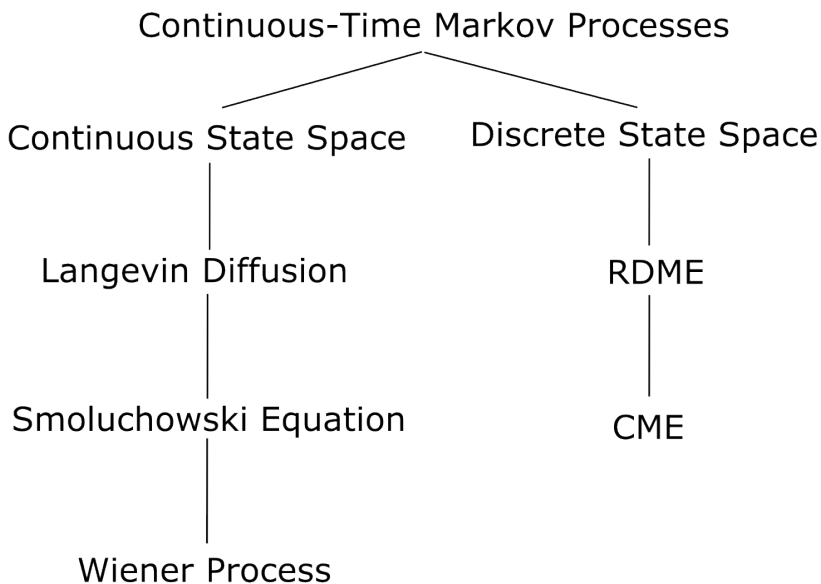


Figure 1: An overview of the stochastic models of reaction and diffusion considered in this paper. Models lower down in the hierarchy are special cases of models higher up.

59 After laying out the underlying theoretical models, which are summarized in Fig. 1 we then overview
 60 their application in three contexts relevant for signal processing in molecular communication:

- 61 (i) *Modulation and Detection*: In order to design reliable modulation and detection algorithms, it is
 62 necessary to characterize the statistics for the quantity of molecules observed by a receiver. In this
 63 case, stochastic models provide the basis for analysis of these statistics.
- 64 (ii) *Statistical Inference*: For low complexity synthetic devices, molecular circuits—where computation
 65 is performed via chemical reactions—are a highly desirable means of implementing algorithms for
 66 statistical inference (e.g., the detection schemes in (i) or parameter estimation). Stochastic models,
 67 particularly stochastic chemical reaction systems, form a useful basis for designing these molecular

68 circuits.

69 (iii) *Modeling Biochemical Environments*: In realistic molecular communication systems, information-
70 carrying molecules may interact with other biochemical processes in the environment. These processes
71 may themselves be governed by stochastic models, and therefore it is a key issue to understand how the
72 dynamics of the information-carrying molecules interact with the dynamics of such external processes.

73 We then discuss selected open problems when the channel, inference or the environment is modeled via
74 stochastic reaction-diffusion systems. We highlight both engineering challenges and mathematical problems
75 arising from system design.

76 1.2. Related Surveys

77 In recent years there have been a number of surveys on various aspects of molecular communication,
78 ranging from channel models to applications in biology, medicine, and nanonetworking. An incomplete list
79 of these surveys is summarized in Table 1.

Table 1: Summaries of Surveys on Molecular Communication

Topic	References
General Survey	[21]
Applications in Medicine	[22, 23, 24]
Applications in Biology	[11, 10, 25, 26]
Applications in Nanonetworking	[27, 28, 29]
Channel Models	[30]
Information Theoretic Limits	[14, 13]
Modulation and Equalization	[8, 9, 31]
Synaptic Communication	[32]

80 The vast majority of the existing surveys have focused on applications. A key novel aspect of this survey
81 is its focus on methodology. In particular, in the first part of this survey, we overview a general family
82 of stochastic models incorporating diffusion, external forces, and chemical reactions. This also contrasts
83 with the surveys in [30, 13], which either emphasize deterministic channel models or a restricted family of
84 Langevin diffusions. The other distinguishing aspect of this survey is in the second part, which overviews
85 applications of stochastic models of chemical reactions and diffusion to inference via biological circuits and
86 lab-on-a-chip devices, as well as the impact of external biochemical processes.

87 1.3. Organization

88 The paper is organized as follows. In Sec. 2, we overview stochastic models based on Langevin diffusion,
89 while Sec. 3 is devoted to chemical master equations and reaction-diffusion master equations. In Sec. 4,

we overview modulation and detection schemes applicable in general Smoluchowski and RDME models. In Sec. 5, we highlight methods for implementing statistical inference algorithms via stochastic models, which forms a basis for the construction of biological circuits. In Sec. 6, we overview recent work modeling the interaction between molecular communication systems and external biochemical processes. Sec. 7 outlines open problems and Sec. 8 concludes.

1.4. Notation

The notation in this paper is summarized in Table 2.

Table 2: Notation.

Variable	Definition	Variable	Definition
N	Number of voxels in the system.	$\mathbf{F}(\mathbf{x})$	external force
$\gamma(\mathbf{x})$	Drag parameter in Langevin diffusion.	m	mass of a molecule
$\boldsymbol{\sigma}(\mathbf{x})$	Covariance matrix of the random perturbation in Langevin diffusion.	$\mathbf{x}(t)$	location of a particle at time t
S_1, S_2	Chemical species.	$\mathbb{E}[\cdot]$	Expectation operator.
$\mathbf{M}_i(t)$	State vector of voxel i in time t .	∇f	Gradient of a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$.
κ_{ij}^l	Diffusive jump rate.	$\nabla \cdot f$	Divergence of a vector field $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$.
$a_i^l, l = 1, 2.$	Reaction rate constants.	∂_t	Partial derivative with respect to t .
$\nu_{l,k}$	Quantity of each species l produced or removed in reaction k .	∂_{xx}	Second partial derivative with respect to x .
$N_{\text{Rx}j}(t)$	Number of S_j molecules in the receiver at time t .	$\nabla^2 f(\mathbf{x})$ $= \nabla \cdot \nabla f(\mathbf{x})$	Laplacian of a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$.
T_s	Communication time interval.	V_{Vox}	volume of a voxel
Δ	Number of transmitted molecules for each bit 1 transmission.	S_n^m	Binary sequence with length n in which m bits are 1
$D_l(\mathbf{x})$	Spatially dependent diffusion coefficient for substance S_l	d	Spatial dimension of the system

97 **2. Langevin Diffusion**

98 A basic source of randomness in the motion of colloidal molecules in a fluid is diffusion. Before turning
 99 to more complex models relevant for molecular communications, we first review the dynamics of a single
 100 molecule in a domain $\Omega \subseteq \mathbb{R}^3$. The basic theory of diffusion for a single colloidal molecule was first developed
 101 by Einstein [33] in order to characterize Brownian motion and developed further by Smoluchowski [34] and
 102 Langevin [35].

103 *2.1. Dynamics of a Single Molecule*

Let $d \in \{1, 2, 3\}$ be the dimension of the system. The general form of the motion of a single molecule
 can be described by the differential equation, obtained via Newton's second law,

$$m \frac{d^2 \mathbf{x}}{dt^2} = -\gamma(\mathbf{x}) \frac{d\mathbf{x}}{dt} + \mathbf{F}(\mathbf{x}) + \boldsymbol{\sigma}(\mathbf{x}) \boldsymbol{\xi}(t), \quad (1)$$

104 where m is the mass of the molecule, $\mathbf{x}(t) \in \mathbb{R}^d$ is the particle's location at time t , $\gamma(\mathbf{x}) \frac{d\mathbf{x}}{dt} \in \mathbb{R}^d$ corresponds
 105 to viscous drag, $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^d$ is an external force (also known as convection), and $\boldsymbol{\sigma}(\mathbf{x}) \boldsymbol{\xi}(t) \in \mathbb{R}^d$ is a random
 106 perturbation due to collisions of the colloidal particle with the molecules comprising the fluid with $\boldsymbol{\sigma}(\mathbf{x}) \in$
 107 $\mathbb{R}^{d \times d}$. Due to the randomness, (1) forms a stochastic differential equation, with solution depending on the
 108 statistical structure of the random perturbation.

109 The term $\mathbf{F}(\mathbf{x})$ corresponds to external forces. For example, in [36], flow is induced by a pressure gradient,
 110 known as Poiseuille flow. A common scenario, arising when electrostatic or gravitational fields act on the
 111 molecule, is that $\mathbf{F}(\mathbf{x}) = \nabla V(\mathbf{x})$, where V is known as a scalar potential. A widely used example in molecular
 112 communications [13] is $V(\mathbf{x}) = v_0 + \mathbf{a} \cdot \mathbf{x}$, $v_0 \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^3$, where the force acting on the particle in each
 113 dimension is given by a_i , $i = 1, 2, 3$, commonly known as drift. Another common choice is the quadratic
 114 potential, which in one dimension is given by $V(x) = ax^2$, $a \in \mathbb{R}$ considered in [37], arising when electrostatic
 115 steering is present [38]. A further example occurs in the presence of acoustic tweezers [39].

116 The random perturbation $\boldsymbol{\xi}(t)$, is assumed to satisfy the following basic properties [6]:

117 (i) $\boldsymbol{\xi}(t)$ is independent of the location \mathbf{x} ;

118 (ii) $\mathbb{E}[\boldsymbol{\xi}(t)] = 0$;

119 (iii) $\mathbb{E}[\boldsymbol{\xi}(t)\boldsymbol{\xi}(t')^T] = \mathbf{I}\delta(t - t')$,

120 where \mathbf{I} is the identity matrix and $\delta(\cdot)$ is the Dirac delta function. On the other hand, $\boldsymbol{\sigma}(\mathbf{x}) \in \mathbb{R}^{d \times d}$ is a
 121 deterministic matrix, which can introduce location dependent variations in correlation and variance of the
 122 random fluctuations. Often, it is assumed that $\boldsymbol{\sigma}(\mathbf{x}) = \sigma(\mathbf{x})\mathbf{I}$, which means that the spatial components
 123 of the random perturbation are uncorrelated, potentially with location-dependent variance, often known as
 124 spatially inhomogeneous diffusion. In the vast majority of work on molecular communications and classical
 125 statistical mechanics, the term $\sigma(\mathbf{x}) = \sigma$ is independent of the location; however, this assumption can fail to
 126 hold in inhomogeneous fluid media [40].

Making sense of the differential equation in (1) requires a non-standard calculus, often in the form of the Itô calculus [7]. Roughly speaking, the starting point is to rewrite (1) as a stochastic integral equation. A rigorous definition of the random perturbation $\boldsymbol{\xi}(t)$ can be obtained via the Wiener process $\mathbf{W}(t)$ as [6]

$$\mathbf{W}(t) = \int_0^t \boldsymbol{\xi}(t') dt'. \quad (2)$$

In particular, the Wiener process is a continuous-time stochastic process, which can be viewed as a random function, satisfying

- (i) $\mathbf{W}(0) = 0$;
- (ii) $\mathbf{W}(t_2) - \mathbf{W}(t_1)$ is independent of $\mathbf{W}(t_4) - \mathbf{W}(t_3)$, with $[t_1, t_2] \cap [t_3, t_4] = \emptyset$;
- (iii) $\mathbf{W}(t_2) - \mathbf{W}(t_1)$ is multivariate Gaussian with covariance matrix $(t_2 - t_1)\mathbf{I}$;
- (iv) $\mathbf{W}(t)$ has continuous sample paths (i.e., realizations of $(\mathbf{W}(t), 0 \leq T)$ are continuous) almost surely.

For further properties of the Wiener process and its role in Langevin diffusion, we refer the reader to [6, 7].

An important special case of (1) arises when drag forces dominate inertial forces, which yields

$$\gamma(\mathbf{x}) \frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}) + \sigma(\mathbf{x})\boldsymbol{\xi}(t), \quad (3)$$

where $\boldsymbol{\sigma}(\mathbf{x}) = \sigma(\mathbf{x})\mathbf{I}$, known as the Smoluchowski equation. This situation often arises in molecular communications due to the fact that inertial forces are due to the mass of the particles, which are typically small. Associated with the Smoluchowski equation in (3) is a Fokker-Planck equation, which describes how the probability density function of the molecule's location evolves over time. For the Smoluchowski equation, the corresponding Fokker-Planck equation is given by [41]

$$\frac{\partial}{\partial t} p(\mathbf{x}, t | \mathbf{x}_0, t_0) = \nabla^2 (D(\mathbf{x}) p(\mathbf{x}, t | \mathbf{x}_0, t_0)) - \nabla \cdot \left(\frac{\mathbf{F}(\mathbf{x})}{\gamma(\mathbf{x})} p(\mathbf{x}, t | \mathbf{x}_0, t_0) \right), \quad (4)$$

where $p(\mathbf{x}, t | \mathbf{x}_0, t_0)$ is the probability density function for the location of the molecule at time t , given an initial location \mathbf{x}_0 at time t_0 , and $D(\mathbf{x}) = \frac{\sigma(\mathbf{x})^2}{2\gamma(\mathbf{x})^2}$ is the diffusion coefficient. Note that both the Smoluchowski equation in (3) and the Fokker-Planck equation in (4) are equivalent descriptions of the system dynamics. This formulation allows for the diffusion coefficient to be spatially inhomogeneous; i.e., the diffusion coefficient depends on the location $\mathbf{x} \in \mathbb{R}^d$.

In the special case where there are no external forces and the diffusion coefficient is constant, the Fokker-Planck equation for Brownian motion (commonly called Fick's second law) is obtained as [41]

$$\frac{\partial}{\partial t} p(\mathbf{x}, t | \mathbf{x}_0, t_0) = D \nabla^2 p(\mathbf{x}, t | \mathbf{x}_0, t_0), \quad (5)$$

where $D = \frac{\sigma^2}{2\gamma^2}$ is the spatially homogeneous diffusion coefficient, independent of the location \mathbf{x} , corresponding to constant coefficients σ, γ .

Solutions to the Fokker-Planck equations in (4) and (5) depend on boundary conditions. In the case of an infinite domain, conditions on the decay $p(\mathbf{x}, t | \mathbf{x}_0, t_0) \rightarrow 0$ as $\|\mathbf{x}\| \rightarrow \infty$ are imposed; for more details,

143 see [42]. In the case of finite domains (corresponding to finite volumes), which are common in the context
 144 of molecular communications, standard boundary conditions are reflective, reaction, Robin, or periodic.
 145 Each of the boundary conditions represents a different physical behavior of molecules at the boundary.
 146 We illustrate in Figure 2 such behavior for the simple case of the diffusion equation in one-dimension
 147 $\partial_t p(x, t) = D \partial_{xx} p(x, t)$ for $x \in \Omega := (0, 1)$, $D > 0$, subject to either (a): Homogeneous Dirichlet boundary
 148 condition $p(0, t) = p(1, t) = 0$, (b): Homogeneous Neumann boundary condition $p'(0, t) = p'(1, t) = 0$,
 149 (c): Periodic boundary condition $p(0, t) = p(1, t)$, or (d): Robin boundary condition $p'(1, t) = kp(1, t)$ and
 150 $-p'(0, t) = kp(0, t)$ for some $k > 0$.

For the general Fokker-Planck equation (4) in a bounded domain $\Omega \subset \mathbb{R}^d$ with boundary $\partial\Omega$, the boundary conditions in the case of reflective boundaries (or the Neumann boundary condition) are given by

$$\hat{\mathbf{n}}(\mathbf{x}) \cdot j(\mathbf{x}, t | \mathbf{x}_0, t_0) := \hat{\mathbf{n}}(\mathbf{x}) \cdot \left[\nabla(D(\mathbf{x})p(\mathbf{x}, t | \mathbf{x}_0, t_0)) - \frac{\mathbf{F}(\mathbf{x})}{\gamma(\mathbf{x})} p(\mathbf{x}, t | \mathbf{x}_0, t_0) \right] = 0, \quad \mathbf{x} \in \partial\Omega, \quad (6)$$

where $\hat{\mathbf{n}}(\mathbf{x})$ is the unit outward normal vector on $\partial\Omega$. The term $j(\mathbf{x}, t | \mathbf{x}_0, t_0)$ is known as the flux, and therefore (6) is interpreted as there is zero flux at the boundary. Another boundary condition is known as the homogeneous Dirichlet boundary condition, where

$$p(\mathbf{x}, t | \mathbf{x}_0, t_0) = 0, \quad \mathbf{x} \in \partial\Omega, \quad (7)$$

which corresponds to the scenario where every particle is consumed at the boundary $\partial\Omega$. It is also possible to combine the boundary conditions to yield the radiation boundary condition (or Robin boundary condition)

$$\hat{\mathbf{n}}(\mathbf{x}) \cdot j(\mathbf{x}, t | \mathbf{x}_0, t_0) = kp(\mathbf{x}, t | \mathbf{x}_0, t_0), \quad \mathbf{x} \in \partial\Omega, \quad (8)$$

where $k \in [0, \infty)$. The fourth boundary condition that is often imposed is known as the periodic boundary condition, in case Ω is a bounded interval in \mathbb{R} or a high dimensional torus \mathbb{T}^d . In one dimensional system defined on $\Omega = [0, b]$, it reads as

$$p(0, t | x_0, t_0) = p(b, t | x_0, t_0) \quad \forall t > 0. \quad (9)$$

151 In general, there do not exist closed-form expressions for $p(\mathbf{x}, t | \mathbf{x}_0, t_0)$ solving (4) under general boundary
 152 conditions. In special cases, such as for (5) subject to various boundary conditions, results can be found in
 153 [41, 43, 15, 44].

154 2.2. Multiple Non-Interacting Molecules

155 Recall the Smoluchowski model in (3). A basic question is what are the statistics for how many molecules
 156 are present in a given compact subdomain Ω' of $\Omega \subset \mathbb{R}^d$ at a given time? As will be discussed further in
 157 Section 4, this question is critical for receiver design in molecular communication. Suppose that all molecules
 158 are *non-interacting*; that is, their positions evolve independently. Physically, this means that the colloidal
 159 molecules do not collide with each other, or exert forces (e.g., electrostatic) on each other.

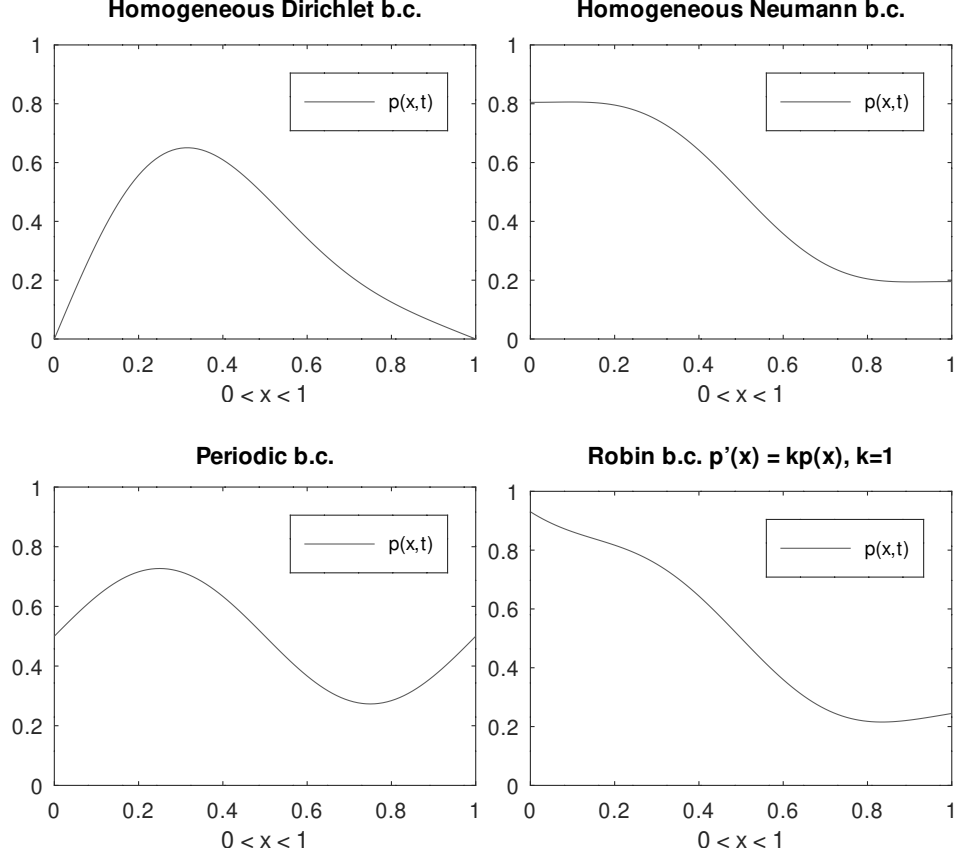


Figure 2: Solutions at time $t = 2$ for diffusion equation $\partial_t p(x, t) = D \partial_{xx} p(x, t)$, $x \in \Omega = (0, 1)$, $D = 0.01$ subject to different boundary conditions. For the homogeneous Dirichlet boundary condition, there are no molecules at the boundary for all time. For the homogeneous Neumann boundary condition, the flux of molecules at the boundary, in this case it is simply the slope of the solution state, is zero. For the periodic boundary condition, the number of molecules at the left and right boundary are always identical. Finally, for Robin boundary condition, the flux at the boundary is proportional to probability density of the molecule, i.e. the slope is steeper if the probability density is larger.

Let $X_i(t) \in \{0, 1\}$, $i = 1, \dots, W$ indicate whether or not molecule i lies in the subdomain Ω' at time t . The total number of molecules in Ω' at time t is then given by

$$X_{\text{tot}}(t) = \sum_{i=1}^W X_i(t), \quad (10)$$

where W is the total number of molecules in Ω . Since the $X_i(t)$ are independent by virtue of the absence of interactions, for W sufficiently large, the distribution of $X_{\text{tot}}(t)$ may be approximated via the central limit theorem [45]. In particular, a useful approximation for the distribution of $X_{\text{tot}}(t)$ is given by

$$\tilde{X}_{\text{tot}}(t) \sim \mathcal{N}(W p_{\Omega'}, W p_{\Omega'}), \quad (11)$$

160 where $\mathcal{N}(\mu, \sigma^2)$ corresponds to the Gaussian distribution, with mean μ and variance σ^2 . The term $p_{\Omega'}$
 161 denotes the probability a given colloidal molecule lies in Ω' at time t . In general computing $p_{\Omega'}$ requires
 162 the solution of the Smoluchowski equation in (3) or the Fokker-Planck equation in (4), which do not have a

163 closed-form expressions for general functional forms of the external forces and boundary conditions.

For t sufficiently large, the Smoluchowski equation can admit simple behavior. This arises when an *equilibrium solution* exists, which means that p_{Ω} converges to a quantity independent of the time t . Although the existence of an equilibrium solution must be postulated, as in the work of Einstein and Langevin, a sufficient condition in one-dimensional systems is the fluctuation-dissipation condition [42, 46], which is given by

$$\frac{\partial}{\partial x} D(x) = F(x)(\gamma(x)^{-1} - D(x)\beta), \quad x \in \Omega, \quad (12)$$

where $\beta = \frac{1}{k_{\text{B}}T}$, where k_{B} is Boltzmann's constant and T is the temperature of the system. For two and three-dimensional systems, see [42]. In the special case that the diffusion coefficient $D(x) = D$, $x \in \Omega$, then the fluctuation-dissipation condition corresponds to the Einstein relation; namely,

$$D = \frac{1}{\gamma\beta}. \quad (13)$$

Under the fluctuation-dissipation condition, the Fokker-Planck equation in (4) can be written as [42]

$$\frac{\partial}{\partial t} p(x, t|x_0, t_0) = \frac{\partial}{\partial x} \left(D(x) \frac{\partial}{\partial x} p(x, t|x_0, t_0) \right) - \frac{\partial}{\partial x} \beta F(x) D(x) p(x, t|x_0, t_0). \quad (14)$$

164 For a detailed derivation, we refer the reader to [42, Eq. 4.17].

If the external force is determined by a scalar potential function U , the force satisfies $F = -\frac{\partial U}{\partial x}$ and U is bounded in $[b_l, b_r]$ and infinite elsewhere, then the equilibrium solution to the Fokker-Planck equation in (4) is given by [41, Sec. 5.2]

$$p_{\infty}(x) = Z e^{-\beta U(x)}, \quad x \in \mathbb{R} \quad (15)$$

where $Z = \left(\int_{b_l}^{b_r} e^{-\beta U(x)} dx \right)^{-1}$ is the normalizing constant. For $d \in \{2, 3\}$ and a potential $U(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$, a similar result holds; namely,

$$p_{\infty}(\mathbf{x}) = \tilde{Z} e^{-\beta U(\mathbf{x})}, \quad \mathbf{x} \in \mathbb{R}^d, \quad (16)$$

165 where \tilde{Z} is the appropriate normalizing constant.

166 Note that scalar potentials are ubiquitous; for example, arising from electrostatic and gravitational fields.
 167 We also highlight that the equilibrium solution p_{∞} in (15) is remarkable in the sense that it only depends
 168 on β and the potential U . That is, given these parameters, it is straightforward to obtain the asymptotic
 169 statistics of the system. This is true even if the diffusion coefficient is spatially inhomogeneous. We note
 170 that for higher-dimensional systems, a similar equilibrium solution to (15) also holds [41].

171 2.3. Simulation

172 In order to obtain realizations of the sample paths for the Smoluchowski equation in (3), a standard
 173 approach is known as the Euler-Maruyama method [17]. Let $\mathbf{a}(t, \mathbf{x}) \equiv \frac{\mathbf{F}(\mathbf{x})}{\gamma(\mathbf{x})}$, $\mathbf{b}(t, \mathbf{x}) \equiv \frac{\sigma(\mathbf{x})}{\gamma(\mathbf{x})}$, and $\Delta \mathbf{W}_n$

174 be a realization drawn from $\mathcal{N}(0, \Delta t \mathbf{I})$. The Euler-Maruyama method is detailed in Algorithm 1, which is
 175 analogous to the Euler method for deterministic partial differential equations. Boundary conditions can be
 176 treated via the methods in [47].

Algorithm 1 The Euler–Maruyama Method

- 1: **INPUT:** $\mathbf{a}(t, \mathbf{x})$, $\mathbf{b}(t, \mathbf{x})$, number of iterations N .
 - 2: **Initialize:** $\mathbf{x}(0)$
 - 3: For $n = 1 : N$
 - 4: $t_n = t_{n-1} + \Delta t$
 - 5: $\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{a}(t_n, \mathbf{x}_n)\Delta t + \mathbf{b}(t_n, \mathbf{x}_n)\Delta \mathbf{W}_n$
-

177 **3. Stochastic Reaction-Diffusion Models**

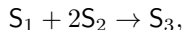
178 In the previous section, we overviewed Langevin models for diffusion and convection with a single
 179 molecule. In this section, we turn to the scenario where multiple molecules are present. In the case that no
 180 chemical reactions are present, this can be achieved via the Langevin models in combination with a central
 181 limit theorem or law of rare events. On the other hand, when chemical reactions are present, interactions
 182 between molecules must be considered. To this end, we first consider the case of well-mixed volumes where
 183 the chemical reactions can be exactly represented by the chemical reaction master equation. The relation of
 184 this stochastic description and deterministic reaction systems arising from chemical reaction network theory
 185 is also discussed. For systems with high heterogeneity, the occurrence of both reactions and diffusion are
 186 well captured by reaction-diffusion master equations (RDME). Note that the current setting of the RDME
 187 does not capture advection.

188 *3.1. The Chemical Master Equation and Chemical Reaction Networks*

Before introducing the RDME, we first overview the case where only reactions are present in a well-stirred container. Suppose that there are r_0 chemical reactions involving s_0 chemical species S_1, \dots, S_{s_0} . The network of chemical reactions is given by

$$\sum_{i=1}^{s_0} \nu_{ik} S_i \rightarrow \sum_{i=1}^{s_0} \nu'_{ik} S_i, \quad k = 1, \dots, r_0, \quad (17)$$

where $\nu_{ik}, \nu'_{ik} \in \mathbb{Z}_{\geq 0}$ are stoichiometric coefficients. For each $k \in \{1, \dots, r_0\}$, the k -th reaction in (17) means that a complex consisting of ν_{ik} molecules of S_i , $i = 1, \dots, s_0$, will produce, via reaction, a complex consisting of ν'_{ik} molecules of S_i , $i = 1, \dots, s_0$. We denote by $\boldsymbol{\nu}_k = [\nu_{1k}, \dots, \nu_{s_0k}]^T$ and $\boldsymbol{\nu}'_k = [\nu'_{1k}, \dots, \nu'_{s_0k}]^T$ the (column) vectors of stoichiometric coefficients. As an illustration, suppose that the chemical species S_1, S_2, S_3 are present and reaction k is of the form



189 then $\nu_{1k} = 1, \nu_{2k} = 2, \nu_{3k} = 0$ and $\nu'_{1k} = 0, \nu'_{2k} = 0, \nu'_{3k} = 1$. With this notation, the reaction systems (17)
 190 can be compactly expressed as

$$\boldsymbol{\nu}_k \rightarrow \boldsymbol{\nu}'_k, \quad k = 1, \dots, r_0. \quad (18)$$

At time t , the number of molecules for each species is described by a component of the random vector $\mathbf{X}(t) = (\mathbf{X}_1(t), \dots, \mathbf{X}_{s_0}(t)) \in \mathbb{Z}_{\geq 0}^{s_0}$. To account for the fact that reactions require individual molecules to be in close proximity, the rate that the k -th reaction occurs when the quantity of molecules for each species is given by \mathbf{x} is denoted by $\lambda_k(\mathbf{x})$, called the propensity. If the k -th reaction occurs at time t , the new state is then

$$\mathbf{X}(t) = \mathbf{X}(t-) + \boldsymbol{\nu}'_k - \boldsymbol{\nu}_k, \quad (19)$$

191 where $t-$ is an infinitesimal time before t and $\mathbf{X}(t-)$ is the state of the system just before the reaction
 192 occurs.

The number of times the k -th reaction occurs by time t is given by the counting process [48]

$$R_k(t) = Y_k \left(\int_0^t \lambda_k(\mathbf{X}(s)) ds \right), \quad (20)$$

where the Y_k , $k = 1, \dots, r_0$ are independent unit rate Poisson processes. As such, the state of the system at time t is given by

$$\mathbf{X}(t) = \mathbf{X}(0) + \sum_{k=1}^{r_0} R_k(t) (\boldsymbol{\nu}'_k - \boldsymbol{\nu}_k), \quad (21)$$

193 which describes the dynamics of a stochastic chemical reaction system. The associated Kolmogorov forward
 194 equation is called the chemical master equation and is given by

$$\frac{dP(\mathbf{x}, t)}{dt} = \sum_{k=1}^{r_0} (\lambda_k(\mathbf{x} - \boldsymbol{\nu}_k) P(\mathbf{x} - \boldsymbol{\nu}_k, t) - \lambda_k(\mathbf{x}) P(\mathbf{x}, t)), \quad (22)$$

195 where $P(\mathbf{x}, t)$ is the probability of the state being $\mathbf{x} \in \mathbb{Z}_{\geq 0}^{s_0}$ at time t , and $\lambda_k(\mathbf{x})$ is the propensity function for
 196 k -th reaction, which can be roughly interpreted as the probability per unit time that reaction k will occur
 197 given the system is in state \mathbf{x} [49].

198 As for the Smoluchowski equation, the evolution of the probability distribution function for the chemical
 199 master equation does not provide solutions in closed-form. It is therefore desirable to characterize equilibrium
 200 solutions, when they exist. It turns out that the conditions an equilibrium distribution exists is closely linked
 201 to the properties of an underlying deterministic chemical reaction system.

202 To illustrate the deterministic chemical reaction network framework (distinct from the above stochastic
 203 models) with a common example, we begin with the ubiquitous reversible¹ enzyme-activated chemical reac-
 204 tion network. This network consists of four chemical species: the enzyme E ; the substrate S ; the intermediate

¹To distinguish with irreversible enzyme reactions where there is no reverse reaction from the product back to the complex; i.e. $k_b = 0$ in (23)

205 complex \overline{ES} ; and the product P. The set of chemical species in this example is then $\mathcal{S}_E = \{E, S, \overline{ES}, P\}$. In
 206 this system, there are four chemical reactions



207 where $k_1, k_2, k_{\text{cat}}, k_b$ are positive reaction rate constants indicating the rate of the corresponding chemical
 208 reactions.

209 Denote by \mathcal{R}_E the set of all reactions in (23), and $\mathcal{K}_E = \{k_1, k_2, k_{\text{cat}}, k_b\}$ the set of reaction rate
 210 constants. The tuple $(\mathcal{S}_E, \mathcal{R}_E, \mathcal{K}_E)$ is called a chemical reaction network. In order to characterize this
 211 network, it is also necessary to consider suitable chemical kinetics (e.g., mass-action or Michaelis-Menten).
 212 Let $[E](t), [S](t), [\overline{ES}](t), [P](t)$ denote the concentration of each chemical species at time t . Under the standard
 213 assumption of mass-action kinetics [48], the concentrations of each species in the enzyme-activated system
 214 are governed by the following deterministic system of ordinary differential equations

$$\begin{aligned} \frac{d[E](t)}{dt} &= -k_1[E](t)[S](t) - k_b[E](t)[P](t) + k_2[\overline{ES}](t) + k_{\text{cat}}[\overline{ES}](t) \\ \frac{d[S](t)}{dt} &= -k_1[E](t)[S](t) + k_2[\overline{ES}](t) \\ \frac{d[\overline{ES}](t)}{dt} &= k_1[E](t)[S](t) + k_b[E](t)[P](t) - k_2[\overline{ES}](t) - k_{\text{cat}}[\overline{ES}](t) \\ \frac{d[P](t)}{dt} &= -k_b[E](t)[P](t) + k_{\text{cat}}[\overline{ES}](t), \end{aligned} \quad (24)$$

215 with non-negative initial conditions $[E](0) = E_0$, $[S](0) = S_0$, $[\overline{ES}](0) = \overline{ES}_0$, and $[P](0) = P_0$.

216 A general chemical reaction system, of which the example above is a special case, which is a chemical
 217 reaction network along with the corresponding reaction rate coefficients is defined as follows.

218 **Definition 3.1.** *A chemical reaction system is the tuple $(\mathcal{S}, \mathcal{R}, \mathcal{K})$ consisting of a set of chemical species*
 219 $\mathcal{S} = \{S_1, \dots, S_{s_0}\}$, *a set of reactions* $\mathcal{R} = \{\nu \rightarrow \nu', \nu, \nu' \in \mathbb{Z}_{\geq 0}^{s_0}\}$, *and the rate function* $\mathcal{K} : \mathcal{R} \rightarrow \mathbb{R}_+$.

For general chemical reaction systems, the dynamics can be described as follows. Let $\mathbf{z}(t) \in \mathbb{R}^{s_0}$ be
 the vector consisting of concentrations of each chemical species at time t . Under mass-action kinetics, the
 chemical reaction system is then governed by

$$\dot{\mathbf{z}}(t) = \sum_{\nu \rightarrow \nu' \in \mathcal{R}} k_{\nu \rightarrow \nu'} \mathbf{z}(t)^\nu (\nu' - \nu), \quad (25)$$

220 where $\mathbf{z}(t)^\nu = z_1(t)^{\nu_1} z_2(t)^{\nu_2} \dots z_{s_0}(t)^{\nu_{s_0}}$ and $k_{\nu \rightarrow \nu'}$ is the reaction rate constant associated to the reaction
 221 $\nu \rightarrow \nu' \in \mathcal{R}$, where the notation $\nu \rightarrow \nu'$ is understood as in (18).

222 A key property of chemical reaction networks is whether or not they admit a *deterministic steady-state*
 223 *concentration*; that is, for each species, the concentration converges to positive value as time $t \rightarrow \infty$. It is
 224 important to note that these steady-states are not guaranteed to be unique nor even to exist. In fact, a long-
 225 standing open problem in the theory of chemical reaction networks is conditions under which the deterministic
 226 steady-state exists and the positive equilibrium is the only attracting point, known as the global attractor
 227 conjecture [50]. One such condition requires the notion of complex balanced chemical reaction systems.

Definition 3.2. A deterministic chemical reaction system $(\mathcal{S}, \mathcal{R}, \mathcal{K})$ is said to be complex balanced if and only if there exists a point $\alpha \in \mathbb{R}_{>0}^{s_0}$ such that

$$\sum_{\nu \rightarrow \nu' \in \mathcal{R}} k_{\nu \rightarrow \nu'} \alpha^{\nu'} (\nu' - \nu) = \sum_{\nu'' \rightarrow \nu \in \mathcal{R}} k_{\nu'' \rightarrow \nu} \alpha^{\nu''} (\nu - \nu''). \quad (26)$$

228 Such a point α is called a positive complex balanced equilibrium.

229 One can show that the enzyme reaction system in (24) is complex balanced, since at equilibrium the reaction
230 rate of each forward reaction is equal to the reaction rate of the corresponding backward reaction.

Returning to stochastic chemical reaction systems (20), an analogous question is whether a stationary distribution exists; that is, whether there existing a limiting distribution π satisfying

$$\pi(\mathbf{m}) = \lim_{t \rightarrow \infty} \Pr(\mathbf{X}(t) = \mathbf{x} | \mathbf{X}(0) = \mathbf{x}_0). \quad (27)$$

231 For a wide range of stochastic chemical reaction systems, the stationary distribution can be characterized;
232 e.g., the following result in [51, Theorem 4.3].

Theorem 3.1. Let $\{\mathcal{S}, \mathcal{R}, \mathcal{K}\}$ be a weakly reversible² deterministic chemical reaction system with rate constants \mathcal{K} . Suppose that the deterministically modeled system is complex balanced with equilibrium concentration $\bar{\mathbf{c}} \in \mathbb{R}_{>0}^{s_0}$. Then, the stochastic system has a product form stationary distribution

$$\pi(\mathbf{x}) = M \frac{\bar{\mathbf{c}}^{\mathbf{x}}}{\mathbf{x}!}, \quad \mathbf{x} \in \mathbb{Z}_{\geq 0}^{s_0}, \quad (28)$$

233 where $M > 0$ is a normalizing constant, and $\mathbf{x}! = x_1! x_2! \cdots x_{s_0}!$.

234 For the case of two species reversible reaction $S_1 \rightleftharpoons S_2$, an application of Theorem 3.1 shows that the
235 distribution of the number of molecules from a given species is *binomially distributed* [51].

236 The time scale at which a stochastic chemical reaction system approaches an equilibrium state is highly
237 dependent on environmental conditions (e.g., temperature) which affect reaction propensities and the initial
238 quantity of each chemical species. One can obtain a reasonable estimate of the equilibration time from the
239 underlying deterministic chemical reaction network. For example, for certain enzyme reactions arising in
240 bacterial chemotaxis, the equilibration time is on the order of 0.2 s [52].

241 In the context of molecular communications, the chemical master equation has been used in [53] in order
242 to model ligand-receptor binding. As will be detailed in Sec. 5, stochastic chemical reaction systems play a
243 useful role in computation for signal processing.

244 3.2. Reaction-Diffusion Master Equation

245 Having defined stochastic chemical reaction systems, we now introduce, in addition, the presence of
246 diffusion. While stochastic reaction-diffusion systems can be defined for high order chemical reactions, for

²A weakly reversible chemical reaction system means that for any sequence of reactions from one chemical complex to another, there exists a reverse sequence of reactions to return to the original complex.

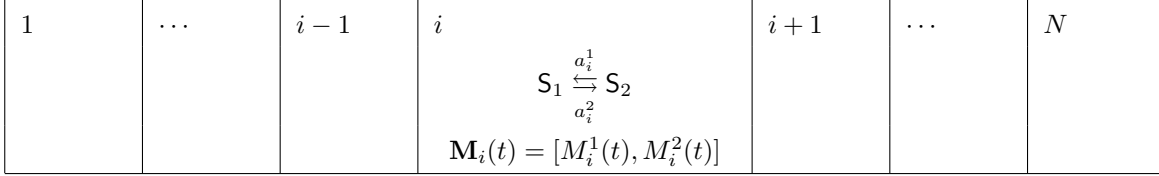


Figure 3: RDME model. The environment is divided into N voxels whose volumes are equal and V_{vox} . The molecules in i th voxel at time t are represented as $\mathbf{M}_i(t) = [M_i^1(t), M_i^2(t)]$. $\mathbf{M}(t) = [\mathbf{M}_1(t), \dots, \mathbf{M}_i(t), \dots, \mathbf{M}_N(t)]$ is governed by a reaction-diffusion master equation.

the purposes of exposition, we focus here on first-order reactions of the form



which have been recently considered in the context of molecular communications [54, 55, 56].

In order to capture the effect of small quantities of each chemical species in the system, we consider a stochastic model for the kinetics. To formally describe the scenario, we introduce the following notation. First, we divide the domain into N small voxels, each with a volume of V_{vox} as shown in Fig. 3. One main idea is that each voxel, well-mixing is assumed. Let $M_i^l(t)$, $l = 1, 2$, $i = 1, \dots, N$ denote the random variable for the number of molecules of species S_1 or S_2 in voxel i at time t . Denote $\mathbf{M}_i(t) = [M_i^1(t), M_i^2(t)]$ as the state vector in voxel i and the matrix consisting of all state vectors as $\mathbf{M}(t) = [\mathbf{M}_1(t), \dots, \mathbf{M}_N(t)]$. The probability that $\mathbf{M}(t)$ has value \mathbf{m} at time t conditioned on there being a quantity of molecules \mathbf{m}_0 in each voxel at time $t = 0$ is then denoted by

$$P(\mathbf{m}, t) = \Pr(\mathbf{M}(t) = \mathbf{m} | \mathbf{M}(0) = \mathbf{m}_0), \tag{30}$$

where $\mathbf{M}(0)$ is the initial quantity of molecules of each species in each voxel.

Remark 3.1. *If there is only a single voxel $i = 1$, then the system corresponds to a stochastic chemical reaction system with unimolecular reactions where $M_1^l(t)$ corresponds to $X_l(t)$ in (21).*

Since the reactions in (29) are unimolecular, it follows that in each reaction the number of molecules of the two species involved can only increase or decrease by one. Let $\mathbf{1}_i^l$ be the state where the number of molecules in all voxels is zero, except for species l in voxel i . That is, $\mathbf{M}(t) + \mathbf{1}_i^l$ means that the number of molecules of species l in voxel i is increased by one.

A popular model for stochastic kinetics of molecules is the reaction-diffusion master equation (RDME) [57], also utilized in the context of molecular communications in [56]. In this model, the diffusive jump rate—corresponding to the rate of the underlying Markov process and not the transition probability—is denoted by κ_{ij}^l for each individual molecule of the l -th species moving from voxel j into voxel i , with $\kappa_{ii}^l = 0$, $i = 1, \dots, N$. In particular, the probability per unit time that a molecule of S_l diffuses from voxel j to voxel i at time t is given by $\kappa_{ij}^l M_j^l(t)$.

262 In general, κ_{ij}^l depends on i, j and l ; that is, the probability of a molecule diffusing between two voxels
 263 is not spatially homogeneous (diffusive jump rates vary from voxel to voxel). The diffusive jump rate out of
 264 the boundary voxels is zero, which can be interpreted as a reflective boundary condition³. We remark that
 265 spatial inhomogeneity of diffusion does not enforce any boundary constraints (other than on the reflective
 266 boundary). Nevertheless, if the diffusive jump rates near the boundary of the receiver are small, the flux
 267 of probability for molecules of S_1 into the receiver will be low. This behavior can model the impact of a
 268 membrane surrounding a receiver in a molecular communication system, which is more difficult for molecules
 269 to pass through than in the rest of the fluid.

270 In the case of mass-action kinetics and first-order reactions, the probability per unit time that a molecule
 271 of S_l in voxel i reacts at time t is given by $a_i^l M_i^l(t)$ with rate constants a_i^l . In general, the reaction rate is
 272 dependent on the voxel index; i.e., we allow for voxels i to have different reaction propensities. For small scale
 273 systems typical in molecular communications, the homogeneity in key parameters such as temperature can be
 274 assumed, which implies few variations in the reaction propensities within individual voxels. The net change of
 275 each chemical species due to the reaction with substrate S_l is expressed via the vector $\boldsymbol{\nu}_l = [\nu_{l,1}, \nu_{l,2}] \in \mathbb{Z}_{\geq 0}^2$,
 276 where $\nu_{l,1}$ is the net gain of molecules of S_1 and $\nu_{l,2}$ is the net gain of molecules of S_2 . The term $\boldsymbol{\nu}_l \mathbf{1}_i$
 277 indicates that $\mathbf{M}(t)$ changes by $\boldsymbol{\nu}_l$ in the i -th voxel.

278 In the RDME model, the probability distribution $P(\mathbf{m}, t)$ evolves according to the system of differential
 279 equations given by

$$\begin{aligned} \frac{dP(\mathbf{m}, t)}{dt} = & \sum_{i=1}^N \sum_{j=1}^N \sum_{l=1}^2 (\kappa_{ij}^l (m_j^l + 1) P(\mathbf{m} + \mathbf{1}_j^l - \mathbf{1}_i^l, t) - \kappa_{ji}^l m_i^l P(\mathbf{m}, t)) \\ & + \sum_{i=1}^N \sum_{l=1}^2 (a_i^l (m_i^l + 1) P(\mathbf{m} - \boldsymbol{\nu}_l \mathbf{1}_i, t) - a_i^l m_i^l P(\mathbf{m}, t)), \end{aligned} \quad (31)$$

280 where the first triple sum corresponds to the net change in probability per unit time due to diffusion and the
 281 second double sum corresponds to the net increase in probability per unit time due to chemical reactions.
 282 We highlight that implicitly in (31), it is assumed that there is a finite, reflective boundary, no degradation
 283 of molecules, constant temperature, and no external forces (e.g. due to a linear potential inducing drift).
 284 For more details, we refer the reader to [57, 58]. A natural question is whether these assumptions can be
 285 relaxed. In [59], external forces were considered, with verification by numerical simulation. We postpone
 286 discussion of relaxing further the assumptions to Sec. 7.

287 The system of ordinary differential equations in (31) corresponds to the Kolmogorov forward equation
 288 for a continuous-time Markov chain, that is, the evolution of the system state is Markovian. In our setting,
 289 due to the reversible reactions and diffusive jump rates detailed above, the Markov chain corresponding to
 290 the RDME is irreducible and positive recurrent. Therefore, the stationary distribution for (31) exists.

291 In the context of molecular communications, the use of the RDME was initiated by [56] and developed in

³This is expressed mathematically as homogeneous Neumann boundary conditions.

several followup works [60, 61, 62, 63, 64, 65]. The RDME has also been considered recently in the context of equilibrium signaling [54, 55], which will be discussed further in Sec. 4.

For a sufficiently large number of molecules, the stationary distribution can in some cases be approximated via an underlying deterministic system. This approximation theory builds on results originally due to Kurtz for the chemical master equation [19] and related density dependent jump Markov processes. For instance, in the work [20], limit theorems approximating the solution of the RDME to that of an underlying deterministic reaction-diffusion system have been developed for the case of a single chemical species that may be produced or annihilated. For the RDME in (31), it may then be expected that for a sufficiently large number of molecules in the system, a good approximation for the concentrations of each species is given by the corresponding deterministic reaction diffusion system.

More precisely, denote by $\Omega \subset \mathbb{R}^d$, $d \geq 1$, a bounded domain comprising the communication channel, $u_l(\mathbf{x}, t)$, $l = 1, 2$ is the concentration of S_l at position $\mathbf{x} \in \Omega$ and time $t > 0$. By applying the law of mass action, the deterministic reaction-diffusion system modeling the reversible reaction (29) reads as

$$\begin{cases} \partial_t u_l - \nabla \cdot (D_l(\mathbf{x}) \nabla u_l) = \sum_{i=1}^2 a_{l,i} u_i, & \mathbf{x} \in \Omega, t > 0, l \in \{1, 2\}, \\ D_l(\mathbf{x}) \nabla u_l \cdot \hat{\mathbf{n}} = 0, & \mathbf{x} \in \partial\Omega, t > 0, l \in \{1, 2\}, \\ u_l(\mathbf{x}, 0) = u_{l0}(\mathbf{x}), & \mathbf{x} \in \Omega, l \in \{1, 2\}, \end{cases} \quad (32)$$

where the self-reaction rate $a_{l,l}$ is defined by $a_{l,l} = -a_{3-l,l}$, for an appropriate choice of reaction rates $a_{l,i}$ and diffusion coefficients $D_l(\mathbf{x})$. The vector-valued function $\hat{\mathbf{n}}(\mathbf{x})$ is the outer unit normal defined for $\mathbf{x} \in \partial\Omega$. The homogeneous Neumann boundary condition $D_l(\mathbf{x}) \nabla u_l \cdot \hat{\mathbf{n}} = 0$ corresponds to a reflective boundary condition, and the initial concentrations are assumed non-negative.

We remark that the unimolecular reaction (29) was chosen for the purposes of exposition. In general, one can consider a network of unimolecular reactions with arbitrary number of chemicals. The dynamical behaviour of the corresponding reaction-diffusion system has been completely characterized in [16].

For the model studied in [20], it is possible to rigorously establish central limit theorems [66, 67], which facilitates the use of Gaussian approximations for the number of molecules in a given subdomain of the system. For the RDME in (31), a Gaussian approximation has been justified empirically in [54].

For the purpose of molecular communications, in [68] a deterministic approximation has been used to approximate the expected number of molecules in a receiver in the presence of information molecule degrading enzymes in the fluid medium. In [69], the deterministic system has been used to model the expected quantity of hydrogen ions when there are reactions between hydrogen ions and acids or bases. Both of these deterministic models involve nonlinear reaction terms. A similar idea was introduced in [30, 70] exploiting a Poisson model with mean determined by an underlying deterministic reaction-diffusion system.

Simulation of the RDME in (31) can be performed via the Gillespie algorithm [71], which has been widely used within molecular communications. Further variations are discussed in [57, 72] and in [59] which

324 incorporates the presence of drift into the RDME.

325 4. Modulation and Detection for Channels Governed by Stochastic Models

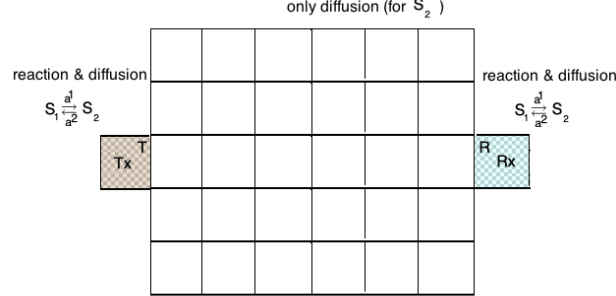


Figure 4: Communication Channel Model for equilibrium signaling. Transmitter (Tx), uses Δ new S_1 molecules to produce information molecules S_2 via $S_1 \xrightarrow{a_1} S_2$ and $S_2 \xrightarrow{a_2} S_1$. Once the information molecules S_2 are released from the transmitter to transmit a message, the receiver (RX) observes these molecules until the system reaches the equilibrium state.

326 A basic problem in molecular communications is how to encode messages to be transmitted in the state of
 327 a set of molecules, known as modulation. On the receiver side, the challenge is then to detect which message
 328 was sent given the observations by the receiving device. The problem of modulation and detection has been
 329 a central question since the inception of molecular communications and there is now a wide body of work,
 330 recently surveyed in [9]. In this section, we provide a brief review of the main approaches for modulation
 331 and detection, when molecular motion is governed by the RDME or Smoluchowski equation.

The general problem of maximum a posteriori detection can be written as

$$\hat{m} = \arg \max_{m \in \mathcal{M}} \Pr(N_{\text{Rx}} | m), \quad (33)$$

where \mathcal{M} is the finite set of possible messages and N_{Rx} is the number of molecules observed at the receiver. More concretely, suppose that the communication system is time slotted, with time slot duration T_s . Consider the n -th time slot. Due to the previous $n - 1$ transmissions, there are $N_{\text{Tx},l}(nT_s)$, $l = 1, 2$ molecules of each species S_l in the transmitter. At a time $nT_s + \delta$ shortly after the beginning of the time slot, the transmitter produces a quantity of S_1 depending on the bit to be transmitted. In particular, for binary signaling with $\mathcal{M} = \{0, 1\}$,

$$N_{\text{Tx},1}(nT_s + \delta) = \begin{cases} N_{\text{Tx},1}(nT_s) + \Delta & s_n = 1, \\ N_{\text{Tx},1}(nT_s) & s_n = 0, \end{cases} \quad (34)$$

332 for $\delta > 0$ a sufficiently small period of time, and $\Delta > 0$ corresponding to the number of molecules emitted
 333 with message $m = 1$ (not to be confused with the state \mathbf{m} for the RDME). Zero molecules are emitted when
 334 $m = 0$.

335 4.1. Detection in General Channels

336 The key challenge in this setting is to reliably detect the transmitted binary symbol; i.e., whether $m = 0$
337 or $m = 1$. A standard method, known as concentration shift keying (CSK), is often defined as the emission
338 of different quantities of molecules for each message; however, in practice, it is ubiquitous to restrict the
339 receiver to sample at a time T_s such that the quantity of molecules at the receiver is maximized in some
340 sense (e.g., the expected number of molecules at the receiver is maximized). For the simplest diffusion model
341 (5), the statistics for the number of molecules in the receiver, N_{RX} , can be computed in closed-form for an
342 infinite domain. However, a characterization of the statistics is more challenging when boundary conditions
343 or chemical reactions are present, although some results have been obtained in this setting (see e.g., [44]).
344 Extensions to non-binary signaling schemes have been studied in, for example, [55].

345 In the absence of chemical reactions, a number of approaches have been proposed; e.g., type based,
346 timing-based modulation, and spatial-based. Type-based signaling is one of the most widespread, where
347 multiple types of molecules can be used as information-carrying molecules. These benefits can be obtained
348 by representing each symbol with different types of molecules as proposed in [73] and known as Molecule
349 Shift Keying (MoSK). Alternatively, in Molecular Concentration Shift Keying (MCSK), it has been pro-
350 posed to use different types of molecules for two consecutive transmissions to reduce interference [74]. In
351 [75], different types of molecules are used as orthogonal channels and it is proposed to use multiple CSK
352 channels simultaneously which is called as D-MoSK in the literature. Timing based approaches are based
353 on representing the intended message by release time as done for binary signaling in [76] and multi level
354 signaling in [77]. Spatial-based approaches can also be developed by utilizing MIMO concepts [78] and index
355 modulation [79]. In addition to these, hybrid approaches have also been proposed such as [80] and [81].
356 In addition to these works and methods presented here, there are several other approaches available in the
357 literature and we refer the interested reader to the surveys [8] and [9] for more details.

358 4.2. Detection in RDME Channels

359 On the other hand, few approaches have been developed for molecular motion governed by the RDME.
360 A recently proposed approach [54, 55, 37], instead exploits the equilibrium behavior of the RDME in the
361 presence of a reflective boundary—leading to a scheme known as *equilibrium signaling*, which is also applicable
362 to a general family of Langevin diffusions [37]. In equilibrium signaling, T_s is chosen to be sufficiently large
363 such that the statistics for the number of molecules in the receiver can be approximated in closed-form via
364 (15) or a Gaussian approximation for the RDME (see [54]). An advantage of this approach is that the
365 stationary distribution is independent of most system parameters (e.g., diffusion coefficients, presence of
366 obstacles, and changes to the location of the transmitter or receiver). With the stationary distribution, it is
367 then feasible to derive near-optimal detection rules [54, 37].

368 To illustrate the idea behind equilibrium signaling, consider Fig. 4. Here, the transmitter seeks to send a
369 message to the receiver by releasing information molecules of species S_2 through the communication channel

370 by producing molecules via the first order chemical reactions $S_1 \xrightarrow{a^1} S_2$ and $S_2 \xrightarrow{a^2} S_1$. In equilibrium
 371 signaling, the receiver ideally observes the quantity of molecules of S_1, S_2 in its domain at the sampling time
 372 $T_s \rightarrow \infty$. In practice, as shown in [54], the required time t^* to reach the equilibrium state of the system with
 373 a tolerable error, can be much much smaller and such time can be derived in terms of channel parameters.
 374 Once, T_s is chosen as t^* , the transmitter communicates by releasing and the receiver observes the quantity
 375 of molecules in its domain.

376 One can view equilibrium signaling as lying in the family of CSK schemes, although—as noted earlier—it
 377 differs from how such schemes have been designed in practice. We also remark that equilibrium signaling
 378 can be used to increase the data rate for multi-level signaling [55]. The scheme is also highly robust to
 379 uncertainties in diffusion coefficients, the shape of the boundary, and locations of the transmitter and receiver
 380 [54]

To illustrate the robustness of equilibrium signaling to uncertainties in the channel, suppose that a
 sequence of bits, s_1, \dots, s_n , over a period of n sampling intervals is sent. We can represent this sequence
 as S_n^m to show that this sequence has m transmissions (each transmission corresponds to a bit 1). Further,
 let $N_{\text{Rx},1}(nT_s|S_n^m)$ and $N_{\text{Rx},2}(nT_s|S_n^m)$ denote the number of molecules of species S_1 and S_2 , respectively,
 observed by the receiver at the end of the n -th symbol period (i.e., at time $(n+1)T_s$), given the transmitted
 sequence S_n^m . Under these circumstances with sufficiently large Δ as shown in [54], the distribution of
 molecules in the receiver can be well approximated by

$$\begin{aligned} N_{\text{Rx},1}(nT_s|S_n^m) &\sim \mathcal{N}(m\mu_{r,1}, m\mu_{r,1}), \\ N_{\text{Rx},2}(nT_s|S_n^m) &\sim \mathcal{N}(m\mu_{r,2}, m\mu_{r,2}), \end{aligned} \quad (35)$$

where $\mu_{r,1}, \mu_{r,2} > 0$ are known constants, dependent on the volume of the enclosing container and not the
 specific geometry, and $\mathcal{N}(\mu, \sigma^2)$ denotes the Gaussian law with mean μ and variance σ^2 . In particular,

$$\mu_{r,1} = \frac{a^1}{a^2} \mu_{r,2}, \quad \mu_{r,2} = \frac{\Delta \frac{V_{\text{Rx}}}{V_{\text{tot}}}}{1 + \frac{a^1}{a^2} \frac{V_{\text{Tx}} + V_{\text{Rx}}}{V_{\text{tot}}}}, \quad (36)$$

where $V_{\text{tot}} = NV_{\text{Vox}}$ is the total volume of the system. Moreover,

$$\begin{aligned} N_{\text{Rx},1}(nT_s|S_n^m) + N_{\text{Rx},2}(nT_s|S_n^m) &\sim \\ \mathcal{N}(m(\mu_{r,1} + \mu_{r,2}), m(\mu_{r,1} + \mu_{r,2})). & \end{aligned} \quad (37)$$

381 Therefore, unlike other signaling schemes, in equilibrium signaling, there is no need to know exact geom-
 382 etry of the channel or the distance between the transmitter and the receiver. The only required information
 383 to model the signal is the volume of the channel and the volume of the receiver.

384 To show the performance of equilibrium signaling, we focus on the scenario of a bounded fluid domain
 385 and the presence of obstacles as in the three different channels presented in Fig. 5. In order to derive
 386 detection rules in these channels, it is necessary to evaluate the channel impulse responses, which cannot

387 be obtained analytically. Therefore, extensive simulations are typically required if conventional detection
 388 schemes are utilized. Furthermore, the detection methods will have different performance when the location
 389 of the transmitter or receiver, the fluid boundary, or obstacles varies. For example, if CSK is used for the
 390 channels given in Fig. 5, if the channel parameters (e.g., diffusion coefficients) are kept constant, we may
 391 expect different performance in the channels in Fig. 5(a) and 5(b) due to the differences in geometry. This is
 392 true even if the channels in Fig 5(b) and 5(c) have the same fluid boundary and obstacles, since the distance
 393 between the transmitter and the receiver are different. On the other hand, in equilibrium signaling, the
 394 channel response only depends on the volume of the channel and the receiver. As such, the performance in
 395 the three different channels in Fig. 5 is the same.

396 This phenomenon is illustrated in Fig 6. In this figure, three different communication schemes are
 397 evaluated: equilibrium signaling with two species S_1, S_2 ; equilibrium signaling with one species S_1 ; and CSK
 398 at the peak time of the observed signal (CSK_{opt}). Note that for equilibrium signaling with one or two species,
 399 all parameters are kept the same including the communication time slot which is equal to the approximate
 400 equilibrium time and can be approximated as in [54]. As expected and can also be observed in Fig 6, as
 401 the volume of the system and the receiver are kept constant, the performance of equilibrium signaling and
 402 equilibrium signaling with one species does not change for different scenarios.

403 On the other hand, using an optimum time slot for CSK communication can be obtained by selecting
 404 the time which maximizes the expected number of molecules as suggested in [82]. Therefore for all different
 405 scenarios— assuming that all channel parameters are known—it is possible to obtain an optimal detector
 406 for CSK. As can be seen in Fig 6, these curves differ for each scenario and equilibrium signaling can still
 407 outperform standard implementations of CSK.

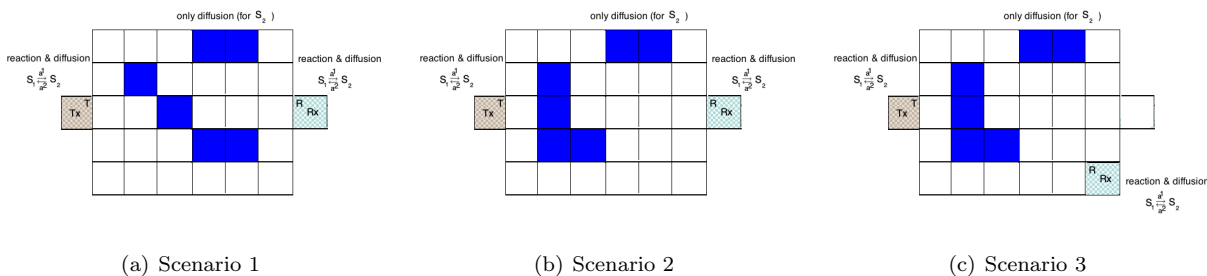


Figure 5: Channel with different geometries. Blue boxes indicate obstacles.

408 The main alternative to equilibrium signaling in the context of the RDME is known as reaction-shift
 409 keying (RSK) [64], where a continuous-time molecular signal is transmitted and detected via chemical reac-
 410 tions. In common with equilibrium signaling (with multiple chemical species) is the use of chemical reactions
 411 in the transmitter and receiver; however, like CSK and MoSK, RSK does not exploit an equilibrium state
 412 and requires computation of the expected quantity of molecules in the receiver at a given time. As this
 413 requires a full solution of the equation of motion, good estimates of transmitter-receiver distances and diffu-
 414 sion coefficients are implicitly required. This again contrasts with equilibrium signaling—which also has a

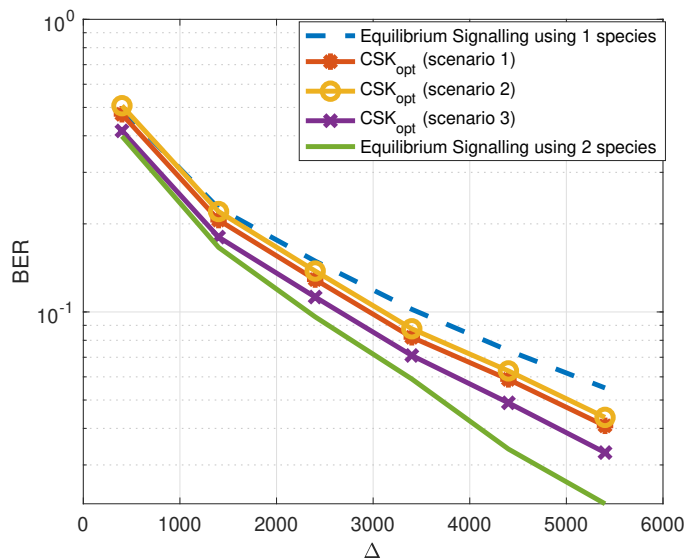


Figure 6: Performance comparison of CSK and equilibrium signaling for the three scenarios presented in Fig. 5. Note that for conventional CSK, the performances differ in different channels as expected while in equilibrium signaling since the volumes are kept constant, same BER values are observed.

415 dramatically lower complexity—where these estimates are not required. Nevertheless, RSK does not require
 416 the system to reach an approximate equilibrium state and has advantages for systems with short symbol
 417 periods.

418 5. Inference via Stochastic Reaction-Diffusion Systems

419 It is widely known that chemical reaction networks, particularly genetic circuits, form a powerful means
 420 of performing computation. This has recently been developed in the context of molecular communications
 421 via microfluidic circuits [11] and for encoding and decoding of channel codes [83]. It is less commonly
 422 observed, but equally true, that stochastic reaction-diffusion can be viewed as *stochastic molecular circuits*
 423 for statistical inference. In this section, we review the use of stochastic reaction-diffusion systems to solve
 424 statistical inference problems.

425 5.1. Receiver Design

426 In molecular communications, a key question is how the detection rules discussed in Section 4 can be
 427 implemented. Due to the small scale and limited resources available to receivers, particularly for nanoscale
 428 devices, computation via biochemical processes is highly desirable. The main work in this direction is due
 429 to Chou and his collaborators [60, 64, 61, 62, 65, 84], which exploits stochastic reaction-diffusion systems
 430 within the receiver.

A key example of this approach is in [63], which considers a receiver consisting of receptors that can exist in two states: an inactive state X ; and an active state X_* . An inactive receptor can be activated by a signaling molecule, where the system is governed by the reactions



431 where g_+ and g_- are the propensities (discussed in detail in Sec. 3) and S is a signaling molecule.

Detection is then based on the log posterior probability that symbol k has been sent given history $\mathcal{X}_*(t)$, denoted by $L_k(t) = \log P(k|\mathcal{X}_*(t))$. Chou then showed that $L_k(t)$ obeys

$$\frac{dL_k(t)}{dt} = \left[\frac{dx_*(t)}{dt} \right]_+ \log(\mathbb{E}[n_{\mathsf{R}}(t)|k, \mathcal{X}_*(t)]) - g_+(M - x_*(t))\mathbb{E}[n_{\mathsf{R}}(t)|k, \mathcal{X}_*(t)], \quad (39)$$

where $x_*(t)$ denotes the number of molecules X_* at time t and $n_{\mathsf{R}}(t)$ is the number of signaling molecules in the receiver. Computing the term $\mathbb{E}[n_{\mathsf{R}}(t)|k, \mathcal{X}_*(t)]$ is computationally challenging and therefore in [63], it is approximated via $\sigma_k(t) = \mathbb{E}[n_{\mathsf{R}}(t)|k]$. This leads to the demodulation filter

$$\frac{dZ_k(t)}{dt} = \left[\frac{dx_*(t)}{dt} \right]_+ \log(\sigma_k(t)) - g_+(M - x_*(t))\mathbb{E}[n_{\mathsf{R}}(t)|k, \mathcal{X}_*(t)], \quad (40)$$

432 with $Z_k(t)$ is initialized to the logarithm of the prior probability that symbol k is transmitted. Detection is
 433 then performed via $\hat{k} = \operatorname{argmax}_k Z_k(t)$. While (40) can be computed numerically, a key contribution in [63]
 434 shows that it can also be approximated via a molecular circuit based on the gene promotor circuit DCS2,
 435 which is found in the yeast *Saccharomyces cerevisiae*.

436 5.2. Parameter Estimation

One example of an inference problem relevant for molecular communications is to estimate the quantity of molecules of different types via non-specific receptors on the surface of the transmitter. In this setting, the number of each type of molecule produced by the receptors depends linearly on the number of molecules that bind with the receptors. In particular,

$$\mathbf{r} = \mathbf{O}\mathbf{l}, \quad (41)$$

437 where $\mathbf{r} \in \mathbb{Z}_{\geq 0}^{n_{\mathsf{R}}}$ is the number of each type of receptor molecule produced, which can be directly observed
 438 within the transmitting device. The vector $\mathbf{l} \in \mathbb{Z}_{\geq 0}^{n_{\mathsf{L}}}$ is the number of each type of species in the environment
 439 that bind to receptors on the transmitter. The matrix \mathbf{O} is an $n_{\mathsf{R}} \times n_{\mathsf{L}}$ observation matrix determining the
 440 production of receptor molecules given the number of molecules from the environment that bind with each
 441 receptor.

In order to estimate the quantity \mathbf{l} , a Bayesian approach can be adopted, to yield samples from $\Pr(\mathbf{l} = \mathbf{u}|\mathbf{r})$. Motivated by Theorem 3.1, consider the product-Poisson prior for the number of molecules in the environment that bind with the receptors, given by

$$\Pr(\mathbf{l} = \mathbf{u}) = \prod_{i=1}^{n_{\mathsf{L}}} e^{-\tilde{q}_i} \frac{\tilde{q}_i^{u_i}}{u_i!}, \quad (42)$$

442 where $\mathbf{l} \in \mathbb{Z}_{\geq 0}^{n_L}$, $\mathbf{r} \in \mathbb{Z}_{\geq 0}^{n_R}$, $\tilde{\mathbf{q}} = [\tilde{q}_1, \dots, \tilde{q}_{n_L}]^T$ are the positive rate parameters. In [85], it has been shown that
443 the distribution of the posterior $\Pr(\mathbf{l}|\mathbf{r})$ with the product-Poisson prior can be sampled from the quantities
444 of molecules of species X_1, \dots, X_n at a sufficiently large time t .

445 This approach has been adopted within molecular communications for signal processing in [86, 87].
446 Indeed, stochastic chemical reaction networks provide a powerful framework for computation. In particular,
447 chemical reaction networks have been shown to implement expectation-maximization [88] and Boltzmann
448 machines (binary stochastic neural networks) [89, 88]. This work focuses on the problem where it is desirable
449 to maximize the likelihood $\Pr(x|\theta)$, where x is data and θ is the parameter to be estimated. A key assumption
450 is that x is drawn from an exponential family and only observations $s = Ax$ are available; that is, the data x
451 is a latent variable and the observations s are obtained via a linear transformation A . We note that a similar
452 problem arises in the context of compressed sensing, which has been widely applied in various communication
453 systems (see e.g., [90]).

454 6. Stochastic Biological Environments

455 In molecular communication systems embedded in a biological environment, there may be interactions be-
456 tween information-carrying molecules and external biological organisms or competition for genetic resources
457 in the process of encoding or decoding information. This problem, addressed in [91, 92, 25, 93, 94, 95] and
458 often called *coexistence*, must account for the stochasticity due to small numbers of information carrying
459 molecules or, for example, limited ribosomes during DNA translation [94]. As such, stochastic reaction-
460 diffusion systems can play an important role.

461 The problem of coexistence is how to ensure that information-carrying molecules do not significantly
462 perturb an external biological system, where the external biological system is modeled via a stochastic
463 chemical reaction system, illustrated in Fig. 7 for two types of signaling molecules. In [25, 94], components
464 of a biological system are modeled via deterministic chemical reaction systems. There are two basic reasons
465 why the interactions between a molecular communication system and an external biological system may be
466 better modeled via stochastic models:

- 467 (i) the motion of the information-carrying molecules is governed via Langevin or RDME models;
- 468 (ii) when a biological system has relatively few molecules of a given species, the internal dynamics are
469 better captured via stochastic, typically chemical reaction network, models.

470 A general framework to establish information theoretic limits of molecular communication such that
471 perturbations of the biological system are limited was developed in [93]. At its heart, this approach placed
472 constraints on the statistics for the number of information-carrying molecules that bind with an external
473 biological system. As such, techniques from covert communications [96] can be adapted to derive scaling
474 laws for the number of messages that can be reliably communicated. An important conclusion is for a code

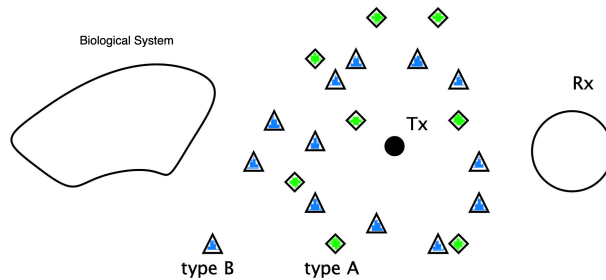


Figure 7: Molecular communication in the presence of a biological system [92].

475 blocklength n (corresponding to n symbols), the number of messages scales with \sqrt{n} when reliability and
 476 coexistence constraints are imposed. This is in contrast with the case where the biological system is not
 477 present, where via Shannon’s noisy channel coding theorem [97], the number of messages that can be reliably
 478 communicated scales with n . This suggests that coding in the presence of external biochemical systems may
 479 need to be adapted; namely, utilizing longer codes or transmission of less information.

480 From the perspective of practical modulation schemes, in [92] an approach based on reactive signaling
 481 [30] was proposed. In particular, the number and time of release of two signaling molecules that can react
 482 with each other was optimized subject to a constraint on the expected number of molecules that pass through
 483 an external biological system.

484 7. Open Problems

485 As detailed in the previous sections, there has been significant progress in developing molecular commu-
 486 nication systems in the presence of and exploiting stochastic models. Nevertheless, each scheme has some
 487 limits on its applicability. In this section, we highlight several open problems that we believe will enhance
 488 the range of applications that can be supported via molecular communications.

489 7.1. Detection

490 As noted in Sec. 4, a key challenge is to derive reliable detection rules for various modulation schemes
 491 in the presence of channels governed by stochastic models. While the simplest form of Langevin diffusion—
 492 namely standard Brownian motion—has been widely studied, there is very little work considering inhomo-
 493 geneous diffusion and external forces.

494 The main work in this direction is [37], based on equilibrium signaling; however, a key limitation is
 495 the requirement that information-carrying molecules do not degrade. Indeed, if molecules degrade, the
 496 equilibrium distribution for the number of molecules is concentrated around zero; that is, no information-
 497 carrying molecules remain in the system at the sampling time. A potential means of extending the work

498 in [37] to the case of molecule degradation is to instead study quasi-equilibrium distributions. While quasi-
499 equilibrium distributions have been considered in various chemical reaction systems [98], they have not yet
500 been studied in the context of molecular communications. A similar issue also arises for the equilibrium
501 signaling scheme in [54], in the context of the RDME model.

502 Another challenge related to equilibrium signaling is the choice of the sampling time. In order to select
503 the sampling time, it is necessary to determine when the system is “close” to equilibrium. In [54], this choice
504 was based on analysis of an approximate underlying deterministic system. In order to carry out this analysis
505 for general RDME models, a characterization of the optimal rate of convergence to equilibrium should be
506 investigated. While bounds on the equilibrium convergence rate are known for fairly general deterministic
507 reaction-diffusion systems [16], the optimal rate has not yet been obtained and remains a key open problem.

508 An important limitation of the vast majority of work on molecular communications is the assumption that
509 information-carrying molecules do not interact. Such an interaction can arise due to electrostatic forces; e.g.,
510 if the molecules are ionized. In order to derive reliable detection rules in this scenario it appears necessary
511 to draw techniques from the field of interacting particle systems [99].

512 Finally, when the fluid media is, for example, turbulent, Langevin diffusion is not an appropriate model.
513 Instead, anomalous diffusion models are more appropriate. While there is some work in this direction
514 [100, 101, 102, 103, 104, 105], the intractability of the solutions to the resulting generalized Fokker-Planck
515 equations introduces new difficulties for the design of detection rules.

516 *7.2. Stochastic Molecular Circuits*

517 At present, there has been very limited use of stochastic models to implement molecular circuits for the
518 purpose of signal processing. As transmitting and receiving devices typically require even limited capabilities
519 to perform coding and decoding, as well as parameter estimation, there are several opportunities for new
520 applications of stochastic models in these settings. On the one hand, there are extensions of the work in [63]
521 exploiting RDME models for detection. On the other hand, exploiting the capabilities of stochastic chemical
522 reaction systems, as in the work in [85, 88], appears to be a fruitful direction for future work.

523 A specific open problem is the development of molecular circuits for detection and channel estimation,
524 which can account for non-specific receptors. Indeed, nanoscale networks are expected to exploit devices
525 that are enclosed by a membrane and require a mechanism to transfer information-carrying molecules into
526 a receiving. Such receptors can bind to information-carrying molecules, and potentially also other chemical
527 species in the environment. The techniques developed in [85, 88] are ideally suited to coping with this
528 problem and an initial study has been carried out in [86].

529 *7.3. External Biochemical Systems*

530 Despite the possibility of information-carrying molecules interacting with external biochemical systems,
531 there has been very limited work investigating how these interactions can be mitigated and the impact on the

532 performance of the communication system. This is particularly true when the internal stochastic dynamics
533 of the biochemical systems are considered.

534 A starting point to begin characterizing the performance reduction can be found in [91], which exploited
535 a fundamental connection between the coexistence problem and covert communications to derive information
536 theoretic limits of communication. However, this work assumed that the biological system does not have
537 memory; that is, the biological system is “reset” after each symbol is transmitted. As realistic biological
538 systems are likely to have memory of their previous state (due to the Markovian nature of their dynamics),
539 addressing this issue seems important in order to understand the true limits of molecular communication
540 systems in the presence of external biochemical systems. A further extension would be to consider the “finite
541 blocklength regime” as considered in [106, 107].

542 Another open issue is the development of practical signaling strategies to minimize the impact of molecular
543 communication systems on their environment. One approach is to build on recent work studying the impact
544 of multiple receivers on reliability [108]. Another angle is to view the molecular communication link as a
545 perturbation of the reaction-diffusion system governing the external biological system. Here, variations on
546 perturbed global attractor theorems in chemical reaction network theory may play a useful role [109].

547 7.4. Distributed Spectroscopy

548 Many proposed engineering applications of molecular communications are in highly uncertain environ-
549 ments; e.g., *in vitro* applications in the human body. Nevertheless, in some settings it is possible to design
550 the communication channel in the sense that molecular motion can be controlled. A key example of such
551 a setting is in microfluidic systems, where the stochastic model governing motion can be adapted to the
552 engineering application. A basic question is then how to optimize the channel?

553 For deterministic models of microfluidic channels, there have been several works investigating how channel
554 parameters affect the dynamics [110, 111, 112, 113, 114, 115, 116]. For channels governed by stochastic mod-
555 els, how to tailor the channel to an engineering application has not been widely investigated. One exception
556 is the recent work in [117, 118], where channel design for the purpose of spectroscopy was investigated.

557 Indeed, a key challenge in molecular biology is to understand how environmental conditions affect the
558 dynamics of biochemical processes. To answer this questions in practice, it is necessary to measure the
559 concentration of chemical species produced by the biochemical process under observation—known as an assay.
560 However, it is not always straightforward to perform such measurements *in situ* or *in vivo* as the spectroscopy
561 measurement devices require special environmental conditions; for example, in force spectroscopy utilizing
562 atomic force microscopy, well-defined, flat and non-reactive substrates are required [119]. In this case, it
563 may desirable to extract samples from the process and perform spectroscopy elsewhere.

564 One approach to performing state estimation of a biochemical process in a distributed fashion has recently
565 been proposed in [117]. This approach is illustrated in Fig. 8, where a biochemical process under observation
566 contained in a chamber, connected to spectroscopy chamber—where the quantity of a given chemical species
567 can be directly observed—via a microfluidic channel.

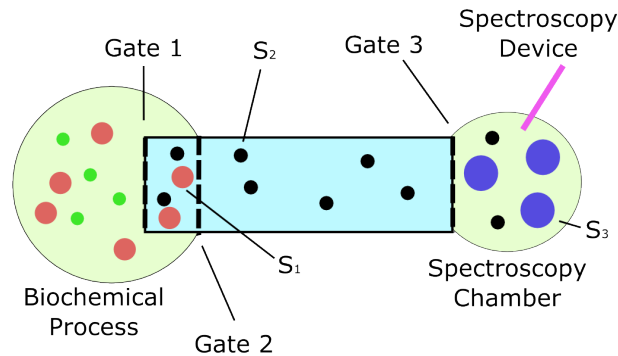


Figure 8: Distributed spectroscopy model in [117, 118].

568 A key open problem is how to design the joint sensing, communication and estimation system tailored
 569 to particular spectroscopy methods. Such design relies heavily on the behavior of a RDME analogous. In
 570 particular, instead of a molecular communication detection problem (as considered in Sec. 7.1), it is necessary
 571 to estimate an unknown quantity of a given chemical species. An even more challenging variant is to design
 572 such systems for *in vivo* applications, perhaps building on the Raman spectroscopy-based methods in [120].

573 8. Conclusions

574 Stochastic models for reaction and diffusion play a critical role in developing molecular communication
 575 systems. This is due to the need to characterize the statistics for the number of molecules that are observed
 576 by a receiver, to implement signal processing via molecular circuits, and to mitigate the impact on external
 577 biochemical processes. In this paper, we have surveyed general stochastic models and their applications in
 578 recent work. By outlining a number of open problems, we hope to encourage further developments to address
 579 outstanding issues of both an engineering and a mathematical nature.

580 References

- 581 [1] B. Miller, B. Bassler, Quorum sensing in bacteria, *Annual Reviews in Microbiology* 55 (1) (2001)
 582 165–199.
- 583 [2] M. Barros, P. Doan, K. Kandhavelu, B. Jennings, S. Balasubramaniam, Engineering calcum signaling
 584 of astrocytes for neural-molecular computing logic gates, *Scientific Reports* 11 (1) (2021) 1–10.
- 585 [3] M. Affolter, K. Basler, The Decapentaplegic morphogen gradient: from pattern formation to growth
 586 regulation, *Nature Reviews Genetics* 8 (9) (2007) 663–674.

- 587 [4] T. Nakano, T. Suda, M. Moore, R. Egashira, A. Enomoto, K. Arima, Molecular communication for
588 nanomachines using intercellular calcium signaling, in: *IEEE Conference on Nanotechnology*, 2005, pp.
589 478–481.
- 590 [5] T. Nakano, A. Eckford, T. Haraguchi, *Molecular Communication*, Cambridge University Press, 2013.
- 591 [6] W. Coffey, Y. Kalmykov, J. Waldron, *The Langevin Equation*, World Scientific, 2004.
- 592 [7] G. Pavliotis, *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and*
593 *Langevin Equations*, Springer, 2014.
- 594 [8] M. Kuscu, E. Dinc, B. A. Bilgin, H. Ramezani, O. B. Akan, Transmitter and receiver architectures
595 for molecular communications: A survey on physical design with modulation, coding, and detection
596 techniques, *Proceedings of the IEEE* 107 (7) (2019) 1302–1341.
- 597 [9] M. S. Kuran, H. B. Yilmaz, I. Demirkol, N. Farsad, A. Goldsmith, A survey on modulation techniques
598 in molecular communication via diffusion, *IEEE Communications Surveys & Tutorials* (2020).
- 599 [10] C. Söldner, et al., A survey of biological building blocks for synthetic molecular communication systems,
600 *IEEE Communications Surveys & Tutorials* 22 (4) (2020) 2765–2800.
- 601 [11] D. Bi, et al., A survey of molecular communication in cell biology: Establishing a new hierarchy for
602 interdisciplinary applications, *arXiv:2009.00090* (2020).
- 603 [12] K. Srinivas, A. Eckford, R. Adve, Molecular communication in fluid media: the additive inverse Gaus-
604 sian noise channel, *IEEE Transactions on Information Theory* 58 (7) (2012) 4678–4692.
- 605 [13] I. Akyildiz, M. Pierobon, S. Balasubramaniam, An information theoretic framework to analyze com-
606 munication systems based on statistical mechanics, *Proceedings of the IEEE* 107 (7) (2019) 1230–1255.
- 607 [14] C. Rose, I. Mian, M. Ozmen, Capacity bounds on point-to-point communication using molecules,
608 *Proceedings of the IEEE* 107 (7) (2019) 1342–1355.
- 609 [15] Y. Deng, et al., Modeling and simulation of molecular communication systems with a reversible ad-
610 sorption receiver, *IEEE Transactions on Molecular, Biological and Multi-Scale Communications* 1 (4)
611 (2015) 347–362.
- 612 [16] K. Fellner, W. Prager, B. Tang, The entropy method for reaction-diffusion systems without detailed
613 balance: first order chemical reaction networks, *Kinetic and Related Models* 10 (4) (2017) 1055–1087.
- 614 [17] P. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992.
- 615 [18] M. Feinberg, Chemical reaction network structure and the stability of complex isothermal reactors–I.
616 the deficiency zero and deficiency one theorems, *Chemical Engineering Science* 42 (10) (1987) 2229–
617 2268.

- 618 [19] T. Kurtz, Limit theorems for sequences of jump Markov processes, *J. Appl. Probab.* 8 (2) (1971)
619 344–356.
- 620 [20] L. Arnold, M. Theodesopulu, Deterministic limit of the stochastic model of chemical reactions with
621 diffusion, *Advances in Applied Probability* 12 (2) (1980) 367–379.
- 622 [21] N. Farsad, et al., A comprehensive survey of recent advancements in molecular communication, *IEEE*
623 *Communications Surveys & Tutorials* 18 (3) (2016) 1887–1919.
- 624 [22] U. Chude-Okonkwo, et al., Molecular communication and nanonetwork for targeted drug delivery: a
625 survey, *IEEE Communications Surveys & Tutorials* 19 (4) (2017) 3046–3096.
- 626 [23] Y. Chahibi, Molecular communication for drug delivery systems: a survey, *Nano Communication*
627 *Networks* 11 (2017) 90–102.
- 628 [24] L. Felicetti, et al., Applications of molecular communications to medicine: a survey, *Nano Communi-*
629 *cation Networks* 7 (2016) 27–45.
- 630 [25] C. McBride, R. Shah, D. Del Vecchio, The effect of loads in molecular communications, *Proceedings*
631 *of the IEEE* 107 (7) (2019) 1369–1386.
- 632 [26] E. Kim, et al., Redox is a global biodevice information processing modality, *Proceedings of the IEEE*
633 107 (7) (2019) 1402–1424.
- 634 [27] T. Nakano, et al., Methods and applications of mobile molecular communication, *Proceedings of the*
635 *IEEE* 107 (7) (2019) 1442–1456.
- 636 [28] K. Darchini, A. Alfa, Molecular communication via microtubules and physical contact in nanonetworks:
637 a survey, *Nano Communication Networks* 4 (2) (2013) 73–85.
- 638 [29] K. Yang, et al., A comprehensive survey on hybrid communication in context of molecular communi-
639 cation and terahertz communication for body-centric nanonetworks, *IEEE Transactions on Molecular,*
640 *Biological and Multi-Scale Communications* 6 (2) (2020) 107–133.
- 641 [30] V. Jamali, et al., Diffusive molecular communications with reactive molecules: Channel modeling
642 and signal design, *IEEE Transactions on Molecular, Biological and Multi-Scale Communications* 4 (3)
643 (2018) 171–188.
- 644 [31] M. C. Gursoy, M. Nasiri-Kenari, U. Mitra, Towards high data-rate diffusive molecular communications:
645 Performance enhancement strategies, *arXiv preprint arXiv:2101.02869* (2021).
- 646 [32] M. Veletić, I. Balasingham, Synaptic communication engineering for future cognitive brain–machine
647 interfaces, *Proceedings of the IEEE* 107 (7) (2019) 1425–1441.

- 648 [33] A. Einstein, On the movement of small particles suspended in stationary liquids required by the
649 molecular-kinetic theory of heat, *Annalen der Physik* 322 (8) (1905) 549–560.
- 650 [34] M. von Smoluchowski, Zur kinetischen theorie der Brownschen molekularbewegung und der suspensio-
651 nen, *Ann. Physik* 326 (21) (1906) 756–780.
- 652 [35] P. Langevin, On the theory of Brownian motion, *C. R. Acad. Sci. Paris* 146 (1908) 530–533.
- 653 [36] Y.-F. Lo, et al., Modeling molecular communications in tubes with Poiseuille flow and Robin boundary
654 condition, *IEEE Communications Letters* 23 (8) (2019) 1314–1318.
- 655 [37] M. Egan, B. Akdeniz, B. Tang, Equilibrium signaling in spatially inhomogeneous diffusion and external
656 forces, *IEEE Transactions on Molecular, Biological, and Multiscale Communications* (2021).
- 657 [38] R. Wade, R. Gabdouliline, S. Lüdemann, V. Lounnas, Electrostatic steering and ionic tethering in
658 enzyme-ligand binding: insights from simulations, *Proc. Nat. Acad. Sci.* 95 (1998) 5942–5949.
- 659 [39] S. Takatori, et al., Acoustic trapping of active matter, *Nature Communications* 7 (1) (2016) 1–7.
- 660 [40] A. Hughes, et al., An introduction to inhomogeneous liquids, density functional theory, and the wetting
661 transition, *American Journal of Physics* 82 (12) (2014) 1119–1129.
- 662 [41] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer, 1989.
- 663 [42] K. Schulten, I. Kosztin, *Lectures in Theoretical Biophysics*, University of Illinois, 2000.
- 664 [43] A. Ahmadzadeh, et al., Comprehensive reactive receiver modeling for diffusive molecular commu-
665 nication systems: Reversible binding, molecule degradation, and finite number of receptors, *IEEE*
666 *Transactions on NanoBioscience* 15 (7) (2016) 713–727.
- 667 [44] Y. Deng, et al., Analyzing large-scale multiuser molecular communication via 3-d stochastic geometry,
668 *IEEE Transactions on Molecular, Biological and Multi-Scale Communications* 3 (2) (2017) 118–133.
- 669 [45] P. Billingsley, *Probability and Measure*, John Wiley & Sons, 2008.
- 670 [46] P. Grassia, Dissipation, fluctuations, and conservation laws, *American Journal of Physics* 69 (2) (2001)
671 113–119.
- 672 [47] D. Lépine, Euler scheme for reflected stochastic differential equations, *Mathematics and Computers*
673 *in Simulation* 38 (1-3) (1995) 119–126.
- 674 [48] D. Anderson, T. Kurtz, *Stochastic Analysis of Biochemical Systems*, Springer International Publishing,
675 New York, 2015.
- 676 [49] D. Gillespie, The chemical Langevin equation, *The Journal of Chemical Physics* 113 (1) (2000) 297–306.

- 677 [50] P. Yu, G. Craciun, Mathematical analysis of chemical reaction systems, *Israel Journal of Chemistry*
678 58 (6-7) (2018) 733–741.
- 679 [51] D. Anderson, et al., Product-form stationary distributions for deficiency zero chemical reaction net-
680 works, *Bulletin of Mathematical Biology* 72 (8) (2010) 1947–1970.
- 681 [52] R. Silversmith, et al., Kinetic characterization of catalysis by the chemotaxis phosphatase CheZ, *Jour-
682 nal of Biological Chemistry* 283 (2) (2008) 756–765.
- 683 [53] M. Pierobon, I. Akyildiz, Noise analysis in ligand-binding reception for molecular communication in
684 nanonetworks, *IEEE Transactions on Signal Processing* 59 (9) (2011) 4168–4182.
- 685 [54] B. Akdeniz, M. Egan, B. Tang, Equilibrium signaling: Molecular communication robust to geometry
686 uncertainties, *IEEE Transactions on Communications* (2020).
- 687 [55] B. Akdeniz, M. Egan, Multi-level equilibrium signaling for molecular communication, in: *Proc. ACM
688 International Conference on Nanoscale Computing and Communications, 2020*, pp. 1–6.
- 689 [56] C. Chou, Extended master equation models for molecular communication networks, *IEEE Transactions
690 on NanoBioscience* 12 (2) (2013) 79–92.
- 691 [57] S. Isaacson, The reaction-diffusion master equation as an asymptotic approximation of diffusion to a
692 small target, *SIAM Journal on Applied Mathematics* 70 (1) (2009) 77–111.
- 693 [58] N. van Kampen, *Stochastic Processes in Physics and Chemistry*, Elsevier Science & Technology Books,
694 2007.
- 695 [59] A. Noel, D. Makrakis, Algorithm for mesoscopic advection-diffusion, *IEEE Transactions on NanoBio-
696 science* 17 (4) (2018) 543–554.
- 697 [60] C. Chou, Impact of receiver reaction mechanisms on the performance of molecular communication
698 networks, *IEEE Transactions on Nanotechnology* 14 (2) (2015) 304–317.
- 699 [61] C. Chou, A Markovian approach to the optimal demodulation of diffusion-based molecular communi-
700 cation networks, *IEEE Transactions on Communications* 63 (10) (2015) 3728–3743.
- 701 [62] C. Chou, Maximum a-posteriori decoding for diffusion-based molecular communication using analog
702 filters, *IEEE Transactions on Nanotechnology* 14 (6) (2015) 1054–1067.
- 703 [63] C. Chou, Designing molecular circuits for approximate maximum a posteriori demodulation of concen-
704 tration modulated signals, *IEEE Transactions on Communications* 67 (8) (2019) 5458–5473.
- 705 [64] H. Awan, C. Chou, Generalized solution for the demodulation of reaction shift keying signals in molec-
706 ular communication networks, *IEEE Transactions on Communications* 65 (2) (2016) 715–727.

- 707 [65] H. Awan, C. Chou, Improving the capacity of molecular communication using enzymatic reaction
708 cycles, *IEEE Transactions on NanoBioscience* 16 (8) (2017) 744–754.
- 709 [66] P. Kotelenez, Law of large numbers and central limit theorem for linear chemical reactions with diffu-
710 sion, *The Annals of Probability* (1986) 173–193.
- 711 [67] D. Blount, Comparison of stochastic and deterministic models of a linear chemical reaction with
712 diffusion, *The Annals of Probability* (1991) 1440–1462.
- 713 [68] A. Noel, C. Cheung, R. Schober, Improving receiver performance of diffusive molecular communication
714 with enzymes, *IEEE Transactions on NanoBioscience* 13 (1) (2014) 31–43.
- 715 [69] N. Farsad, A. Goldsmith, A molecular communication system using acids, bases and hydrogen ions, in:
716 *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*,
717 2016, pp. 1–6.
- 718 [70] M. Farahnak-Ghazani, et al., On medium chemical reaction in diffusion-based molecular communica-
719 tion: A two-way relaying example, *IEEE Transactions on Communications* 67 (2) (2018) 1117–1132.
- 720 [71] D. Gillespie, Stochastic simulation of chemical kinetics, *Annu. Rev. Phys. Chem.* 58 (2007) 35–55.
- 721 [72] J. Elf, A. Doncic, M. Ehrenberg, Mesoscopic reaction-diffusion in intracellular signaling, in: *Fluctua-*
722 *tions and noise in biological, biophysical, and biomedical systems*, Vol. 5110, International Society for
723 Optics and Photonics, 2003, pp. 114–124.
- 724 [73] M. S. Kuran, H. B. Yilmaz, T. Tugcu, I. F. Akyildiz, Modulation techniques for communication via
725 diffusion in nanonetworks, in: *2011 IEEE international conference on communications (ICC)*, IEEE,
726 2011, pp. 1–5.
- 727 [74] H. Arjmandi, A. Gohari, M. N. Kenari, F. Bateni, Diffusion-based nanonetworking: A new modulation
728 technique and performance analysis, *IEEE Communications Letters* 17 (4) (2013) 645–648.
- 729 [75] M. H. Kabir, S. R. Islam, K. S. Kwak, D-MoSK modulation in molecular communications, *IEEE*
730 *transactions on nanobioscience* 14 (6) (2015) 680–683.
- 731 [76] N. Garralda, I. Llatser, A. Cabellos-Aparicio, E. Alarcón, M. Pierobon, Diffusion-based physical chan-
732 nel identification in molecular nanonetworks, *Nano Communication Networks* 2 (4) (2011) 196–204.
- 733 [77] B. C. Akdeniz, A. E. Pusane, T. Tugcu, Position-based modulation in molecular communications,
734 *Nano communication networks* 16 (2018) 60–68.
- 735 [78] B. H. Koo, H. B. Yilmaz, C.-B. Chae, A. Eckford, Detection algorithms for molecular MIMO, in: *2015*
736 *IEEE International Conference on Communications (ICC)*, IEEE, 2015, pp. 1122–1127.

- 737 [79] M. C. Gursoy, E. Basar, A. E. Pusane, T. Tugcu, Index modulation for molecular communication via
738 diffusion systems, *IEEE Transactions on Communications* 67 (5) (2019) 3337–3350.
- 739 [80] M. C. Gursoy, E. Basar, A. E. Pusane, T. Tugcu, Pulse position-based spatial modulation for molecular
740 communications, *IEEE Communications Letters* 23 (4) (2019) 596–599.
- 741 [81] M. C. Gursoy, D. Seo, U. Mitra, Concentration and position-based hybrid modulation scheme for
742 molecular communications, in: *ICC 2020-2020 IEEE International Conference on Communications*
743 *(ICC)*, IEEE, 2020, pp. 1–6.
- 744 [82] I. Llatser, A. Cabellos-Aparicio, M. Pierobon, E. Alarcón, Detection techniques for diffusion-based
745 molecular communication, *IEEE Journal on Selected Areas in Communications* 31 (12) (2013) 726–
746 734.
- 747 [83] A. Marcone, et al., Parity-check coding based on genetic circuits for engineered molecular communi-
748 cation between biological cells, *IEEE Transactions on Communications* 66 (12) (2018) 6221–6236.
- 749 [84] M. Riaz, H. Awan, C. Chou, Using spatial partitioning to reduce the bit error rate of diffusion-based
750 molecular communications, *IEEE Transactions on Communications* 68 (4) (2020) 2204–2220.
- 751 [85] M. Virinchi, et al., A stochastic molecular scheme for an artificial cell to infer its environment from
752 partial observations, in: *Proc. International Conference on DNA-Based Computers*, 2017, pp. 82–97.
- 753 [86] M. Egan, T. Duong, M. Di Renzo, Biological circuits for detection in MoSK-based molecular commu-
754 nication, *IEEE Access* (2019).
- 755 [87] M. Egan, et al., Estimation and optimization for molecular communications with a coexistence con-
756 straint, in: *Proceedings of the Sixth Annual ACM International Conference on Nanoscale Computing*
757 *and Communication*, 2019, pp. 1–6.
- 758 [88] M. Virinchi, A. Behera, M. Gopalkrishnan, A reaction network scheme which implements the EM
759 algorithm, in: *Proc. International Conference on DNA Computing and Molecular Programming*, 2018,
760 pp. 189–207.
- 761 [89] W. Poole, et al., Chemical Boltzmann machines, in: *International Conference on DNA-Based Com-*
762 *puters*, 2017, pp. 210–231.
- 763 [90] M. Leinonen, M. Codreanu, G. Giannakis, Compressed sensing with applications in wireless networks,
764 *Foundations and Trends in Signal Processing* 13 (1-2) (2019).
- 765 [91] M. Egan, et al., Strategies for coexistence in molecular communication, *IEEE transactions on NanoBio-*
766 *science* 18 (1) (2018) 51–60.

- 767 [92] B. Akdeniz, M. Egan, A reactive signaling approach to ensure coexistence between molecular communi-
768 cation and external biochemical systems, *IEEE Transactions on Molecular, Biological and Multi-Scale*
769 *Communications* 5 (3) (2019) 247–250.
- 770 [93] M. Egan, et al., Coexistence in molecular communications, *Nano Communication Networks* 16 (2018)
771 37–44.
- 772 [94] A. Gyorgy, D. Del Vecchio, Limitations and trade-offs in gene expression due to competition for shared
773 cellular resources, in: *IEEE Conference on Decision and Control*, 2014, pp. 5431–5436.
- 774 [95] V. Loscri, A. Vegni, G. Fortino, On the interaction between a nanoparticulate system and the human
775 body in body area networks, *Micromachines* 6 (9) (2015) 1213–1235.
- 776 [96] M. Bloch, Covert communication over noisy channels: A resolvability perspective, *IEEE Transactions*
777 *on Information Theory* 62 (5) (2016) 2334–2354.
- 778 [97] C. Shannon, A mathematical theory of communication, *The Bell System Technical Journal* 27 (3)
779 (1948) 379–423.
- 780 [98] D. Anderson, et al., Stochastic analysis of biochemical reaction networks with absolute concentration
781 robustness, *Journal of The Royal Society Interface* 11 (93) (2014).
- 782 [99] T. Liggett, *Interacting Particle Systems*, Springer Science and Business Media, 2012.
- 783 [100] T. Mai, et al., Event detection in molecular communication networks with anomalous diffusion, *IEEE*
784 *Communications Letters* 21 (6) (2017) 1249–1252.
- 785 [101] D. Trinh, et al., Molecular communication with anomalous diffusion in stochastic nanonetworks, *IEEE*
786 *Transactions on Communications* 67 (12) (2019) 8378–8393.
- 787 [102] D. Trinh, et al., Molecular communication in H-diffusion, *IEEE Transactions on Communications*
788 (2020).
- 789 [103] T. Cao, et al., Anomalous diffusion in molecular communication, *IEEE Communications Letters* 19 (10)
790 (2015) 1674–1677.
- 791 [104] L. Chouhan, et al., Molecular communication in fractional diffusive channel, *IEEE Communications*
792 *Letters* 24 (10) (2020) 2172–2176.
- 793 [105] S. Huang, et al., Molecular communication via subdiffusion with a spherical absorbing receiver, *IEEE*
794 *Wireless Communications Letters* 9 (10) (2020) 1682–1686.
- 795 [106] Y. Polyanskiy, et al., Channel coding rate in the finite blocklength regime, *IEEE Transactions on*
796 *Information Theory* 56 (5) (2010) 2307–2359.

- 797 [107] H.-Y. Lin, et al., Weak flip codes and their optimality on the binary erasure channel, *IEEE Transactions*
798 *on Information Theory* 64 (7) (2018) 5191–5218.
- 799 [108] Y. Lu, et al., The effect of two receivers on broadcast molecular communication systems, *IEEE Trans-*
800 *actions on NanoBioscience* 15 (8) (2016) 891–900.
- 801 [109] M. Gopalkrishnan, A scheme for molecular computation of maximum likelihood estimators for log-
802 *linear models*, in: *Proc. International Conference on DNA-Based Computers*, 2016, pp. 3–18.
- 803 [110] E. De Leo, L. Donvito, L. Galluccio, A. Lombardo, G. Morabito, L. Zanolì, Communications and
804 *switching in microfluidic systems: pure hydrodynamic control for networking labs-on-a-chip*, *IEEE*
805 *Transactions on Communications* 61 (11) (2013) 4663–4677.
- 806 [111] M. Hamidović, U. Marta, G. Fink, R. Wille, A. Springer, W. Haselmayr, Information encoding in
807 *droplet-based microfluidic systems: First practical study*, in: *Proceedings of the Sixth Annual ACM*
808 *International Conference on Nanoscale Computing and Communication*, 2019, pp. 1–6.
- 809 [112] N. Farsad, A. Eckford, S. Hiyama, Y. Moritani, On-chip molecular communication: analysis and design,
810 *IEEE Transactions on NanoBioscience* 11 (3) (2012) 304–314.
- 811 [113] A. O. Bicen, J. J. Lehtomäki, I. F. Akyildiz, Shannon meets Fick on the microfluidic channel: diffusion
812 *limit to sum broadcast capacity for molecular communication*, *IEEE Transactions on NanoBioscience*
813 17 (1) (2018) 88–94.
- 814 [114] B. Krishnaswamy, et al., Time-elapse communication: bacterial communication on a microfluidic chip,
815 *IEEE Transactions on Communications* 61 (12) (2013) 5139–5151.
- 816 [115] L. Galluccio, et al., Capacity of a binary droplet-based microfluidic channel with memory and anticipa-
817 *tion for flow-induced molecular communications*, *IEEE Transactions on Communications* 66 (1) (2018)
818 194–208.
- 819 [116] A. Bicen, I. Akyildiz, System-theoretic analysis and least-squares design of microfluidic channels for
820 *flow-induced molecular communication*, *IEEE Transactions on Signal Processing* 61 (20) (2013) 5000–
821 5013.
- 822 [117] B. Akdeniz, M. Egan, A molecular communication scheme to estimate the state of biochemical processes
823 *on a lab-on-a-chip*, in: *Proc. ACM International Workshop on Nanoscale Computing, Communication,*
824 *and Applications*, 2020, pp. 8–13.
- 825 [118] B. Akdeniz, M. Egan, Molecular communication for equilibrium state estimation in biochemical pro-
826 *cesses on a lab-on-a-chip*, *IEEE Transactions on NanoBioscience* 20 (2) (2021) 193–201.

- 827 [119] L. Ferrari, J. Kaufmann, F. Winnefeld, J. Plank, Interaction of cement model systems with superplas-
828 ticizers investigated by atomic force microscopy, zeta potential, and adsorption measures, *Journal of*
829 *Colloid and Interface Science* 347 (2010) 15–24.
- 830 [120] H. Guo, Cooperative Raman spectroscopy for real-time in vivo nano-biosensing, *IEEE Transactions*
831 *on NanoBioscience* 16 (7) (2017) 571–584.