King's Research Portal

Link to publication record in King's Research Portal

CSK Realization for MC via Spatially Distributed Multicellular Consortia

Rinrada Jadsadaphongphaibool, Dadi Bi, and Yansha Deng

Abstract—**The design and engineering of molecular communication (MC) components capable of processing chemical concentration signals is the key to unleashing the potential of MC for interdisciplinary applications. By controlling the signaling pathway and molecule exchange between cell devices, synthetic biology provides the MC community with tools and techniques to achieve various signal processing functions. In this paper, we propose a design framework to realize any order concentration shift keying (CSK) systems based on simple and reusable single-input single-output cells. The design framework also exploits the distributed computation on multicellular consortia with spatial segregation, which has advantages in system scalability, low genetic manipulation, and signal orthogonality. We also create a small library of simple logic engineered cells and apply them to implement binary CSK (BCSK) and quadruple CSK (QCSK) systems to demonstrate the feasibility of our proposed design framework. The simplicity of our engineered cells allows for their reuse in other systems beyond CSK. Importantly, we establish a mathematical framework to theoretically characterize our proposed distributed multicellular systems. Specially, we divide a system into fundamental building blocks, from which we derive the impulse response of each block and the cascade of the impulse responses leads to the end-to-end response of the system. Simulation results obtained from the agent-based simulator BSim not only validate our CSK design framework but also demonstrate the accuracy of the proposed mathematical analysis.**

Index Terms—**Concentration shift keying, distributed computation, engineered cells, genetic circuits, molecular communication, spatial segregation, synthetic biology.**

I. INTRODUCTION

Molecular communication (MC) is a biologically-inspired communication paradigm that exchanges information via chemical signals, which facilitates various applications in the fields of healthcare for drug delivery, military for covert communication, environmental science and industry for pollution control, etc [2]. Over the past decade, the focus of MC has been on the use of communication engineering tools to model, evaluate, and optimize MC systems, such as molecule propagation modeling [3], [4], communication capacity [5], [6], and transmission and detection schemes [7]. Despite those advancements in communication theoretical studies, unified and coherent technologies to design and engineer MC components capable of processing chemical concentration signals are less explored, which hinders the development and implementation of interdisciplinary applications [8]–[10].

Like traditional communication systems, modulation-demodulation is one of the primary key signal processing functions for MC to encode information symbols into physical chemical signals. Among various modulation schemes, concentration shift keying (CSK) is the most popular modulation scheme in the MC literature, since it requires only one type of molecule, has low computation complexity, and can be extended to other modulation strategies [11]. In [12], we have realized several CSK modulation and demodulation functionalities via chemical circuits by designing and regulating a set of chemical reactions in microfluidic systems. In particular, it has shown that different concentration levels can be achieved by biological network motif-inspired chemical reactions and by adjusting the geometry of microfluidic systems [13]. The digital signal processing capability of chemical circuits was also demonstrated to be analogous to digital electronics [14], where any order CSK can be achieved. However, those chemical circuits are described by high-level languages and generic species, which requires an extra effort to find molecules to achieve these circuits in real biological scenarios.

The discipline of synthetic biology provides the engineering community with novel tools and techniques to process chemical concentration signals in cells for MC [10], [15]. Leveraging the activation and repression mechanisms of the gene expression process, engineered cells can imitate logic functions [16], [17], toggle switches [18], and a memory device [19]. This motivates several works on engineering MC transceivers in a single cell to perform the M-ary amplitude modulation-demodulation function [20] or the single parity-check coding-decoding function [21].

Although it is natural to engineer signal processing functions in a single cell from the system point of view, the confined volume of a cell can only accommodate a limited number of signal processing functions. It has shown that the maximum number of repressor-based logic gates per cell cannot exceed eight gates in both *Escherichia coli* (*E. Coli*) [22] and yeast [23]. A high number of logic gates inside a single cell can lead to a metabolic burden, reduced cell viability, and compromised reusability of the cell [22]. This substantially constrains the scalability and complexity of signal processing functions needed for practical applications. Luckily, these issues can be addressed by distributed computation in multicellular consortia, which divides complex signal processing functions into different cells with simple functions and wires them together via signaling molecules [24], [25]. In this way, the distributed computation can ultimately achieve improved

programmability and higher computational complexity through multi-cell cooperation [26].

Motivated by the aforementioned studies, the objective of this paper is to achieve CSK modulation and demodulation functions for MC via distributed multicellular consortia. Although we realized the binary CSK (BCSK) system using engineered cells in [1], it is unclear how this BCSK design can be extended to other higher orders of CSK schemes. Inspired by this, we propose a general design framework for CSK modulation and demodulation functions. The contributions of this paper are outlined as follows:

- We generalize and expand our conference version in [1] by proposing a general design framework to realize any order of CSK system for MC. Significantly, our design framework is based on single-input single-output cells along with spatial segregation in a microfluidic system, which not only provides signal orthogonality but also allows the simple engineered cells to be reused. Thus, we develop a small library of reusable engineered cells and utilize them to achieve the BCSK and quadruple CSK (QCSK) systems according to the general CSK design framework. Unlike the genetic circuits proposed in [20], [21], the gene expressions and molecules of our proposed system are designed and validated in the existing synthetic biology literature. Hence, our design principle can be used to develop other signal processing functions in MC using synthetic cells.

- We establish a novel mathematical framework to characterize our proposed BCSK and QCSK systems. This is accomplished by dividing the system into fundamental building blocks, deriving the impulse response of each building block, and finally deriving the end-to-end expression of the systems. Importantly, our mathematical framework can be applied to analyze other new and more complicated microfluidic-based multicellular circuits.

- We present the implementation of our proposed BCSK and QCSK systems in the BSim, which is an agent-based simulator that provides a different set of tools for an efficient simulation of cellular behavior in realistic environmental geometries. The simulation results obtained from BSim validate not only the functionalities of our proposed BCSK and QCSK systems but also our derived mathematical framework. We also evaluate the communication performance of the QCSK system in terms of bit error rate (BER) through BSim simulations. Our result shows that a well-tuned detection thresholding value and bit interval can reduce the effect of intersymbol interference (ISI).

The remainder of this paper is organized as follows. In Sec. II, we propose a general design framework to realize any order of CSK via distributed computation. In Sec. III, we apply the general CSK design framework to realize the BCSK and QCSK systems and establish a mathematical framework to characterize our proposed systems. Numerical results in Sec. IV validate the proposed CSK framework and their theoretical analyses. Finally, we conclude this paper in Sec. V.
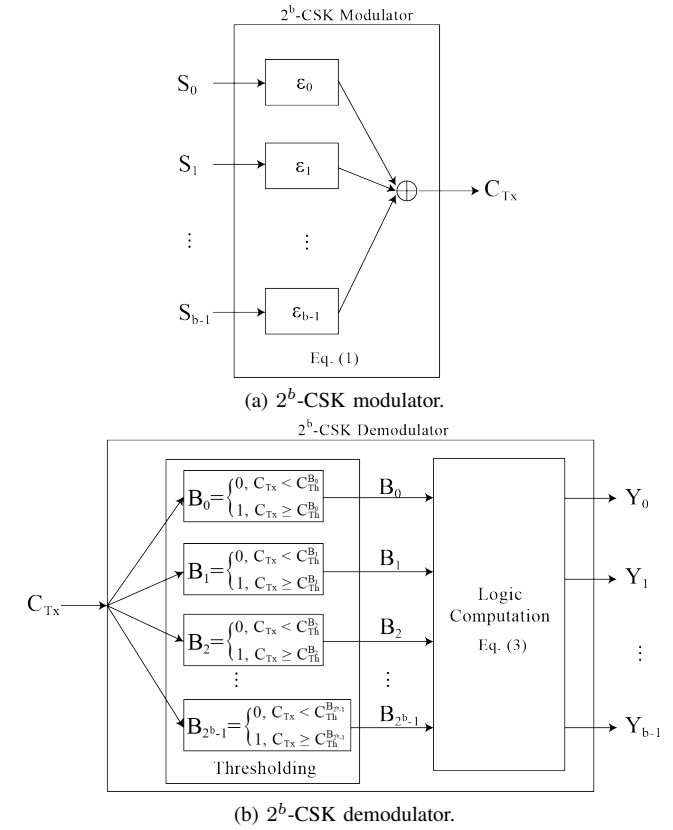

(a) 2^b-CSK modulator.

(b) 2^b-CSK demodulator.

Fig. 1. The design principle for realizing 2^b-CSK using genetic circuits.

II. DESIGN FRAMEWORK

In this section, we first propose a general design framework to realize any order of CSK using genetic circuits in Sec. II-A. The main idea of the design framework is to implement the 2^b-CSK function through a small library of reusable engineered cells, where the cells are distributed in different confined spaces and interconnected by signaling molecules to perform signal processing. With the complexity of the 2^b-CSK function reduced into simpler functions, the simple logic cell can be reused in other systems, e.g., the logic cell used in the BCSK system can be reused in the QCSK system, and the simpler function requires less editing of the gene to produce a desired output. Then, we present the fundamental processes of molecular signal processing inside an engineered cell in Sec. II-B and characterize the propagation environment of information molecules in Sec. II-C. The main notations of the proposed systems are summarized in Table I.

A. Design Principle

1) 2^b-CSK Modulation: The 2^b-CSK modulator (as shown in Fig. 1(a)) encodes b bits into 2^b different concentrations of a transmitted signal, which is equivalent to the amplitude shift keying (ASK) in conventional communications [27]. The key challenge of implementing 2^b-CSK modulation for MC is how to control the output concentration via the 2^b different input combinations. We address this by expressing the output concentration of a transmitted signal C_{Tx} through the base

TABLE I
TABLE OF NOTATIONS

Notation	Unit	Definition
b		Number of bits the system encodes/decodes.
C_{Tx}		Output concentration of a transmitted signal.
S_i, B_j, Y_i		Binary signal.
ε_i		Bit weighting of S_i.
$C_{\mathrm{Iin}}(t)$	nM	Intracellular concentration of input molecules.
η	μm s^{-1}	Exchange rate.
$C_{\mathrm{I}}(t)$	nM	General input concentration of the engineered cell block.
k_{d}	min^{-1}	Degradation rate of the molecules.
$C_{\mathrm{Oin}}(t)$	nM	Intracellular concentration of output molecules.
β	nM min^{-1}	Protein production rate.
θ	nM^{-1}	Data fitting parameter.
n		Hill coefficient.
ξ		Rate of the product molecules released from the gate cell.
$C_{\mathrm{R}}(t)$	nM	Concentration of repressor protein.
k_{f}		Forward reaction rate constant.
$f_{\mathrm{R}}(t)$		Rate of repressor production.
$C_{\mathrm{Th}}(t)$	nM	Concentration of threshold molecules.
$C_{\mathrm{O}}(t)$	nM	Concentration of molecules released from the cell.
$C(\mathbf{d}, t)$	nM	Molecular concentration inside microfluidic channel.
$\mathbf{d}$		Vector defining position in space.
D	μm^2 s^{-1}	Diffusion constant.
u	μm s^{-1}	Velocity.
$\mathbf{d}_0$		Initial position the molecules is released.
k_{a}		Absorption rate of a cell.
ψ		Exchange factor with the ratio $k_{\mathrm{a}} : \eta$.
$\mathcal{I}[\cdot]$	nM	Expected net change of the output concentration released by an ID cells on the surface of the cells during interval $[t, t + t_s]$.
$\mathcal{N}[\cdot]$	nM	Expected net change of the output concentration released by a NOT cells on the surface of the cells during interval $[t, t + t_s]$.
$\mathcal{T}[\cdot]$	nM	Expected net change of the output concentration released by a thresholding cells on the surface of the cells during interval $[t, t + t_s]$.
$\mathcal{P}[\cdot]$	nM	The concentration of molecules absorbed by a surface.
$C_{\mathrm{IP}}(t)$	nM	Normalized input concentration of the propagation block.
$\mathcal{S}_{\mathrm{a}}$		Absorption surface.
$\mathcal{S}_{\mathrm{e}}$		Emission surface.
$\mathcal{S}_{i,j}$		A surface on the y-z plane with $W_i \leq y \leq W_j$ and $0 \leq z \leq H$.

conversion method[1] [28], which is

$$C_{\mathrm{Tx}} = \sum_{i=0}^{b-1} \varepsilon_i S_i, \tag{1}$$

where S_i is the ith bit of the input $\{S_0, S_1, \cdots, S_{b-1}\}$ and ε_i is the bit weighting of S_i.

Eq. (1) can be translated into a biological implementation via b multicellular consortia located in different physically confined spaces [26]. In particular, the engineered cells in each consortium perform an identity (ID) function that produces a type of wiring molecules only in the presence of the input molecules, i.e., $S_i = 1$. The concentration of the wiring molecules of a consortium (i.e., the bit weighting

multiplication) can be controlled by adjusting the number of engineered cells, this is because the more cells, the higher the concentration [29]. Then, by mixing the wiring molecules released from multicellular consortia, we can implement the summation operator, thus achieving the 2^b-CSK modulation.

2) 2^b-CSK Demodulation: The 2^b-CSK demodulator (as shown in Fig. 1(b)) decodes 2^b different concentrations of transmitted signals to b bits. For this purpose, the demodulator is separated into a front-end and a back-end module. The front-end module consists of $2^b - 1$ biological thresholding units with different thresholding values to distinguish 2^b concentration levels into binary signals B_j ($0 \leq j \leq 2^b - 2$). In particular, the engineered cells in each thresholding unit produce a type of wiring molecules representing bit-1 only when the input concentration C_{Tx} exceeds a thresholding value. As the front-end module is only capable of decoding 2^b concentration

---

[1]Here, the base conversion in (1) converts binary values to decimal values. For instance, the binary value 11 can be converted to a decimal value $1 \times 2^1 + 1 \times 2^0 = 3$.

$$\begin{cases} Y_0 = (\overline{B_{2^b-2}} \cdots \overline{B_2}\ \overline{B_1}B_0) + (\overline{B_{2^b-2}} \cdots B_2B_1B_0) + \cdots + (B_{2^b-2} \cdots B_2B_1B_0) \\ Y_1 = (\overline{B_{2^b-2}} \cdots \overline{B_2}B_1B_0) + (\overline{B_{2^b-2}} \cdots B_2B_1B_0) + \cdots + (B_{2^b-2} \cdots B_2B_1B_0) \\ \quad\vdots \\ Y_{b-1} = (B_{2^b-2} \cdots B_2B_1B_0) \end{cases} \tag{2}$$

$$Y_i = \begin{cases} B_{2^{(b-1)}-1}, & i = b-1 \\ B_q + \displaystyle\sum_{l=1}^{2^{(b-i-1)}-1} \overline{(B_{q-(2l-1)(2^i)} + \overline{B_{q-2l(2^i)}})}, & 0 \le i < b-1 \end{cases} \tag{3}$$

TABLE II
CSK DEMODULATOR BACK-END TRUTH TABLE

Symbol Index	B_{2^b-2}	$\cdots$	B_2	B_1	B_0	Y_{b-1}	$\cdots$	Y_1	Y_0
0	0	$\cdots$	0	0	0	0	$\cdots$	0	0
1	0	$\cdots$	0	0	1	0	$\cdots$	0	1
2	0	$\cdots$	0	1	1	0	$\cdots$	1	0
3	0	$\cdots$	1	1	1	0	$\cdots$	1	1
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
2^b-1	1	$\cdots$	1	1	1	1	$\cdots$	1	1

levels to $2^b - 1$ binary signals, the back-end module is required to further process the signals into b bits. The back-end module extracts demodulated bit Y_i ($0 \le i \le b-1$) from B_j through logic computation, where the relationship between Y_i and B_j is presented in Table II. This table describes the output according to all possible input combinations, which can be transformed into Boolean equations in terms of sum-of-product (SOP) expression as (2), where $\overline{B_j}$ is the NOT operator of B_j and B_jB_i is the AND operation between B_j and B_i. To simplify the genetic engineering of the cells, we can rewrite (2) into an alternative expression excluding the AND logic based on the inverted logic formulation (ILF) [26]. According to De Morgan's Law [30], ILF is achieved by applying double negations. Thus, eq. (2) can be rewritten as (3), where $q = 2^{(b-1)} - 1 + \sum_{j=0}^{b-i-2} 2^{(b-(j+2))}$.

The back-end module, described in (3), can be translated into similar multicellular consortia as 2^b-CSK modulator. Apart from the ID function and summation operation, engineered cells in the consortium also perform a simple NOT function by inhibiting the production of wiring molecules in the presence of the input molecules. Hereby, we can achieve the back-end module and realize the 2^b-CSK demodulation by connecting the modules together.

B. Fundamental Processes of the Engineered Cells

From the discussion of our proposed design framework, the realization of 2^b-CSK modulation and demodulation requires three types of engineered cells: ID gate, NOT gate, and thresholding unit. As shown in Fig. 2, the molecule signal processing by all these three engineered cell types is based on three fundamental processes: 1) the *exchange* of input molecules from the extracellular to the intracellular environment, 2) the *signal processing* of output molecules through regulated gene expressions, and 3) the *release* of the output molecules

from the intracellular to the extracellular environment. In the following, we provide the details of these three processes along with their mathematical descriptions.
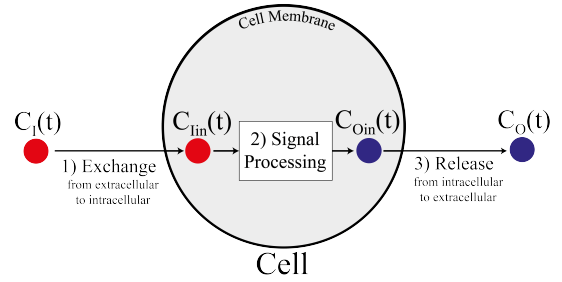

Fig. 2. Schematic representation of molecular processing by a cell.

1) Molecule exchange: When external molecules are absorbed on the cell membrane, the molecules are then transferred from the extracellular to the intracellular environment. The dynamics of molecule exchange can be expressed as [31, eq. (7)]

$$\frac{dC_{\mathrm{Iin}}(t)}{dt} = \eta C_{\mathrm{I}}(t) - k_d C_{\mathrm{Iin}}(t), \tag{4}$$

where $C_{\mathrm{Iin}}(t)$ is the concentration of input molecules inside a cell, η is the exchange rate, $C_{\mathrm{I}}(t)$ is the general input molecules and here it represents the molecules in the extracellular environment, and k_d is the degradation rate of the molecules.

2) Intracellular signal processing: Cells process signals through the gene regulatory network (GRN), which is a set of interacted gene expressions that control a specific cell function. The gene expression mainly consists of two steps: *transcription* and *translation* [19]. First, gene expression begins with transcription, where the RNA polymerase (RNAP) recognizes and binds to the promoter region found at the beginning of a gene to initiate the synthesis of a strand of mRNA. Once mRNA is produced, a ribosome uses tRNA to convert mRNA into protein via translation. The gene expression can be controlled by transcription factors (TFs), which can up-regulate (or down-regulate) the protein production by enhancing (or inhibiting) the binding of RNAP to the promoter site. By designing the GRN, we can realize the ID, NOT, and thresholding signal processing functions.

- **ID Gate**: As shown in Fig. 3, we engineer the ID gate through inducing the gene expression directly by the input molecules (i.e., acting as an activator). Thus,
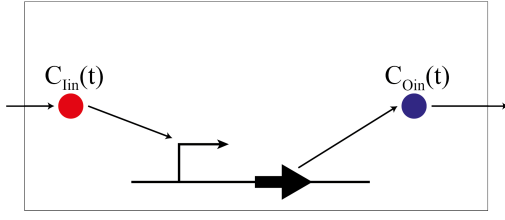
Fig. 3. The GRN of the ID gate. The promoter (bent arrow) indicates an area where transcription starts. The protein coding sequence (thick arrow) indicates the region of DNA directly read during translation to produce proteins.

the dynamics of output molecule concentration can be expressed as [32, eq. (7.8)]

$$\frac{dC_{\text{Oin}}(t)}{dt} = \beta \frac{C_{\text{Iin}}(t)^n}{1 + [\theta C_{\text{Iin}}(t)]^n} - (k_{\text{d}} + \xi)C_{\text{Oin}}(t), \quad (5)$$

where $C_{\text{Oin}}(t)$ is the concentration of the produced molecules inside a cell, β is the protein production rate, θ is a data fitting parameter, n is the Hill coefficient, and ξ is the rate of the product molecules released from the gate cell.
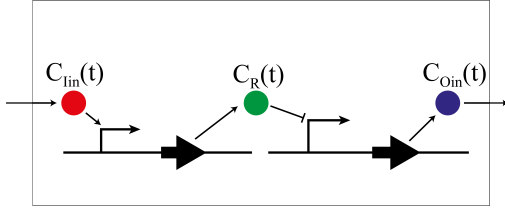

Fig. 4. The GRN of the NOT gate. The repression by the binding of TFs (indicated with a flat head arrow) prevents the expression of downstream proteins.

- **NOT Gate**: To implement the NOT function, we borrow an optimal configured and experimentally validated GRN of the engineered yeast cells from [19]. As shown in Fig. 4, the presence of the input molecules activates the production of a repressor protein, which can be modeled by

$$\frac{dC_{\text{R}}(t)}{dt} = \beta \frac{C_{\text{Iin}}(t)^n}{1 + [\theta C_{\text{Iin}}(t)]^n} - k_{\text{d}}C_{\text{R}}(t), \quad (6)$$

where $C_{\text{R}}(t)$ is the concentration of the repressor protein. Then, the produced repressor protein binds to the downstream gene and precludes the expression of the output molecules, which can be expressed as [32, eq. (7.10)]

$$\frac{dC_{\text{Oin}}(t)}{dt} = \beta \frac{1}{1 + [\theta C_{\text{R}}(t)]^n} - (k_{\text{d}} + \xi)C_{\text{Oin}}(t). \quad (7)$$
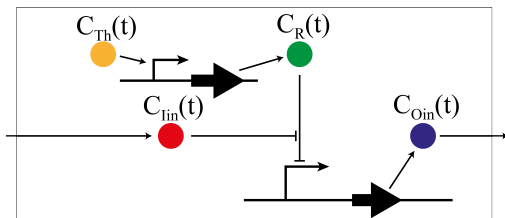

Fig. 5. The GRN of the thresholding unit.

- **Thresholding Unit**: As shown in Fig. 5, different from the GRN of the NOT gate, the input molecules react with

the repressor protein induced by the threshold molecules, thus controlling the amount of the output molecules. The reaction between the input molecules and the repressor can be expressed as

$$\text{I}_{\text{in}} + \text{R} \xrightarrow{k_{\text{f}}} \phi, \quad (8)$$

where k_{f} is the forward reaction rate constant and ϕ is a type of molecule irrelevant for signal processing. As such, the dynamics of input molecule concentration in (4) can be rewritten as

$$\frac{dC_{\text{Iin}}(t)}{dt} = \eta C_{\text{I}}(t) - k_{\text{f}}C_{\text{Iin}}(t)C_{\text{R}}(t) - k_{\text{d}}C_{\text{Iin}}(t), \quad (9)$$

and the dynamics of repressor protein concentration can be expressed as [33, eq. (22a)]

$$\frac{dC_{\text{R}}(t)}{dt} = f_{\text{R}}(t) - k_{\text{f}}C_{\text{Iin}}(t)C_{\text{R}}(t) - k_{\text{d}}C_{\text{R}}(t), \quad (10)$$

where $f_{\text{R}}(t) = \beta \frac{C_{\text{Th}}(t)^n}{1 + [\theta C_{\text{Th}}(t)]^n}$ is the rate of repressor production and $C_{\text{Th}}(t)$ is the concentration of the threshold molecules.[2] The thresholding value of the cell is the concentration of the repressor molecules in equilibrium without the consumption by the input molecules, i.e., $f_{\text{R}}(t) - k_{\text{d}}C_{\text{R}}(t) = 0$ in (10). Hence, the thresholding value can be tuned by controlling the production of repressor via $C_{\text{Th}}(t)$. Combining (7) with (10), we can express the signal processing of a thresholding unit.

3) Molecule release: With the production of the output molecules, they are further released from the intracellular to the extracellular. The dynamics of molecule release can be expressed as [31, eq. (7)]

$$\frac{dC_{\text{O}}(t)}{dt} = \xi C_{\text{Oin}}(t), \quad (11)$$

where $C_{\text{O}}(t)$ is the concentration of the released molecules from cells.

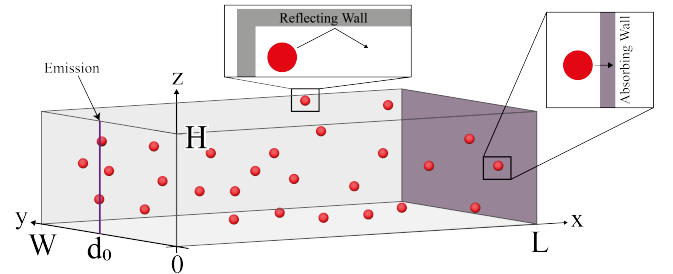### C. Propagation of Signaling Molecules

Fig. 6. The emission, propagation, and reception of signaling molecules.

Under the multicellular consortium setting, it is essential to establish communication between different consortia to achieve community level functions (e.g., 2^b-CSK modulation and demodulation). To provide wiring (or signaling) molecules

---

[2]While $C_{\text{Th}}(t)$ and $f_R(t)$ are constant for our proposed system, we aim to design a tunable thresholding cell that can be utilized in various scenarios. Therefore, we consider $C_{\text{Th}}(t)$ and $f_R(t)$ as time-dependent functions to accommodate this flexibility.

with guided transmission and increased propagation speed, we choose to incubate engineered cells in a microfluidic system. The integration of engineered cells with a microfluidic system not only provides an opportunity to realize spatial segregation by isolating different cell types at different locations [12], [26], but also avoids the cross-talk between cells by controlling the fluid direction.

For a 3D microfluidic channel, as shown in Fig. 6, the molecular concentration can be described by the advection-diffusion-reaction as [33, eq. (29)]

$$\frac{\partial C(\mathbf{d},t)}{\partial t} = D\nabla^2 C(\mathbf{d},t) - \nabla \cdot [\mathbf{u}(\mathbf{d})C(\mathbf{d},t)] - k_{\mathrm{d}}C(\mathbf{d},t),$$ (12)

where $\mathbf{d} = [x,y,z]$ is a vector defining the position in space, D is the diffusion constant, ∇^2 is the Laplace operator, ∇ is the Nabla operator, and $\mathbf{u}(\mathbf{d})$ is the local velocity. The flow inside the microfluidic channel is assumed to be laminar, which is a valid assumption in microfluidic settings [34]. For the tractability of the analysis, we further assume a uniform flow with constant velocity u along the y- and z-axis [35]. As a result, eq. (12) can be reduced to

$$\frac{\partial C(\mathbf{d},t)}{\partial t} = D\nabla^2 C(\mathbf{d},t) - u\frac{\partial C(\mathbf{d},t)}{\partial x} - k_{\mathrm{d}}C(\mathbf{d},t).$$ (13)

When molecules are uniformly released over $\mathbf{d}_0 = [0,y_0,z]$ into the microfluidic channel at time $t = 0$, the initial condition can be defined as

$$C(\mathbf{d}_0,t)\big|_{t=0} = \delta(x)\delta(y - y_0),$$ (14)

where $\delta(\cdot)$ is the Dirac delta function.

As the molecules propagate along the x-axis, the boundaries of the microfluidic channel are assumed to be reflective, and we can define their boundary conditions as [33, eq. (8)]

$$\frac{\partial C(\mathbf{d},t)}{\partial y}\bigg|_{y=0,W} = 0,$$ (15)

$$\frac{\partial C(\mathbf{d},t)}{\partial z}\bigg|_{z=0,H} = 0,$$ (16)

where W and H are the width and height of the channel, respectively.

The molecules propagated inside the microfluidic channel can be received by a column of well-aligned engineered cells located at $L+R$, where L is the distance between the releasing point and the cell and R is the radius of the cell. In this paper, the used signaling molecules (provided in Sec. III) are too large to cross the cell membrane, and their reception relies on the receptors distributed on the surface of the cell membrane [4], [10]. Therefore, each engineered cell can be modeled as an active receiver, and we assume the well-aligned cells create an absorbing wall along the y-z plane. For simplicity, we assume that the molecules will be absorbed at the left side of the cells (e.g., $x = L$), and the molecules produced by GRN will be released at the right side of the cells (e.g., $x = L + 2R$). Hence, another boundary condition can be defined as [4, eq. (5)]

$$D\frac{\partial C(\mathbf{d},t)}{\partial x}\bigg|_{x \to L} - k_{\mathrm{a}}C(\mathbf{d},t)\bigg|_{x \to L} = 0,$$ (17)
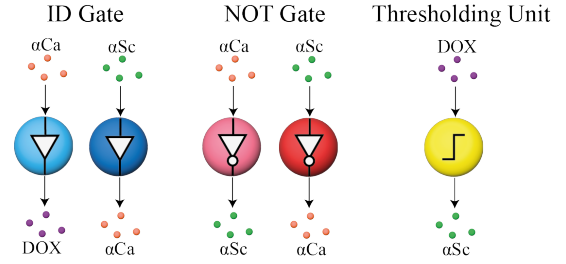

Fig. 7. The library of engineered cells used to achieve BCSK and QCSK.

where k_{a} is the absorption rate of one cell.

III. DESIGN AND ANALYSIS OF BCSK AND QCSK SYSTEMS

In this section, we apply the general 2^b-CSK design framework to realize the BCSK and QCSK systems to demonstrate the feasibility of the design principle. In order to implement the BCSK and QCSK systems according to the design principle presented in Sec. II-A, we develop a library of engineered yeast cells as shown in Fig. 7. This library consists of five types of cells that achieve three signal processing functions by responding and producing three signaling molecules: mating α-factor from *C.albicans* (αCa), mating α-factor from *Saccharomyces cerevisiae* (αSc), and doxycycline (DOX). In particular, the inputs and outputs of the engineered cells are:

- **ID Gate:** responds to αCa and αSc and produces DOX and αCa, respectively;
- **NOT Gate:** responds to αCa and αSc to express the LacI repressor (see Fig. 4) that prevents the expression of αSc and αCa, respectively;
- **Thresholding Unit:** responds to DOX, which reacts with TetR repressor (see Fig. 5) and inhibits the repression of αSc.

In the following, we first present how to use the cell library to achieve the BSCK and QCSK designs in Sec. III-A. We then analyze the building blocks of the BCSK and QCSK systems in Sec. III-B and apply them to examine the output responses of the systems in Sec. III-C.

A. Design of BCSK and QCSK Systems

1) BCSK Design: The BCSK system encodes one bit of information into two different concentration levels; the BCSK is also known as on-off keying (OOK) when zero molecules are released to send a bit-0. We present our proposed OOK design in Fig. 8. The OOK transmitter employs the ID cell (αCa-DOX) to respond to the input bit S_0, represented by molecules αCa, with a concentration $C_{S_0}(t)$. The reception of αCa molecules by the ID cells can lead to the expression and release of DOX into the microfluidic channel, generating the modulated signals denoted by $C_{\mathrm{Tx}}(t)$. Due to the ID logic of the engineered cells, no molecules will be released for bit-0 transmission, thus achieving the OOK modulation.

Once the DOX molecules arrive at the receiver, they can activate the thresholding cell (DOX-αSc) to produce αSc molecules and a binary signal B_0. For $b = 1$, eq. (3) can be expressed as $Y_0 = B_0$; therefore, the thresholding unit is sufficient to achieve OOK demodulation. The αSc molecules
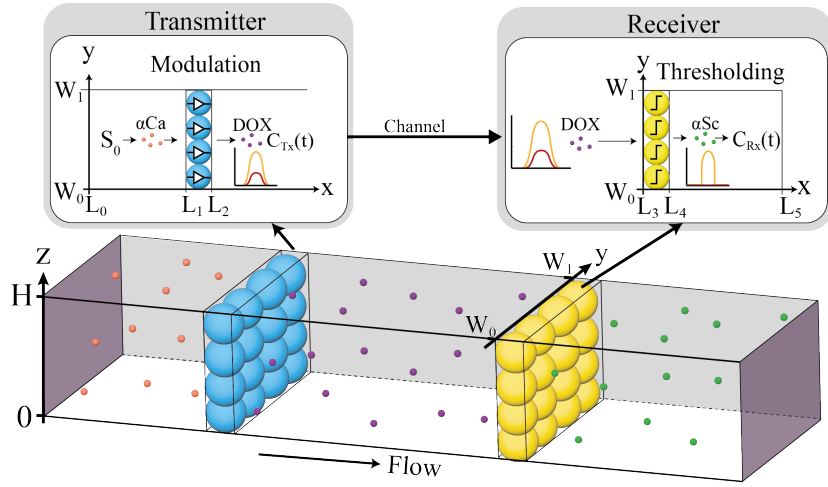
Fig. 8. The implementation of the BCSK system.

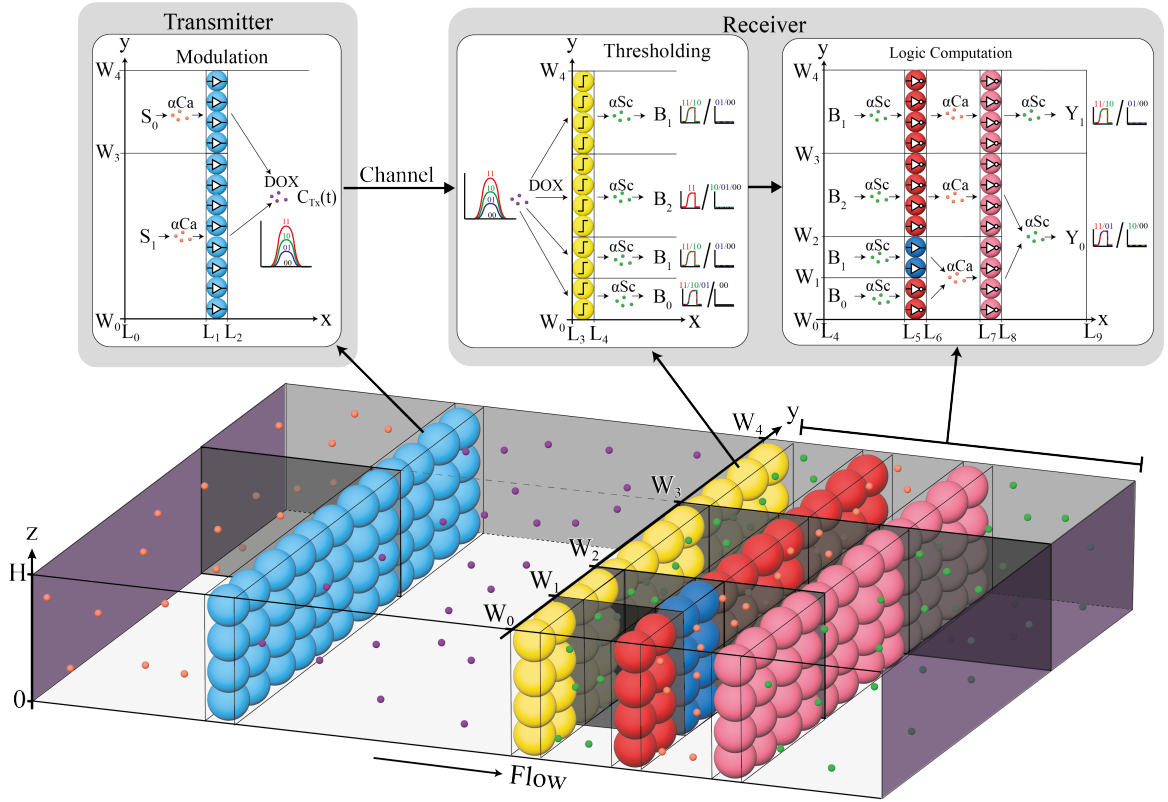Fig. 9. The implementation of the QCSK system.

then propagate to the detection wall, which generates the final output of Rx, denoted by $C_{\text{Rx}}(t)$. In practice, the concentration level can be visualized by fluorescent proteins or measured by electrochemical sensors experimentally.

2) QCSK Design: The QCSK system encodes two bits of information into four concentration levels. Fig. 9 presents our proposed QCSK design. For $b = 2$, eq. (1) can be rewritten as

$$C_{\text{Tx}} = \varepsilon_0 S_0 + \varepsilon_1 S_1. \tag{18}$$

The QCSK modulator encodes input bits, S_1 and S_0, by employing two populations of ID cells (αCa-DOX). The number of cells in the population responding to S_1 is twice of

that in the other population, i.e., $\varepsilon_1 = 2\varepsilon_0$ in (18). By doing so, bit S_1 leads to a higher concentration output than bit S_0 and the combination of S_1 and S_0 produces four concentration levels, thus achieving the QCSK modulation.

At the receiver side, DOX molecules activate three populations of thresholding cells (DOX-αSc) to distinguish the four concentration levels into B_0, B_1, and B_2. For $b = 2$, eq. (3) can be reduced to

$$Y_1 = B_1 = \overline{\overline{B_1}}, \tag{19}$$

$$Y_0 = B_2 + (\overline{B_1 + \overline{B_0}}) = \overline{\overline{B_2}} + (\overline{B_1 + \overline{B_0}}). \tag{20}$$

Although (19) reveals that the output of the thresholding unit B_1 can be directly interpreted as Y_1, we further apply a double NOT operation to B_1 to ensure that Y_1 has the same concentration level as Y_0. Specifically, the first NOT cells (αSc-αCa) take the thresholding output signal B_1, which inhibits the production of αCa molecules. The output of the first NOT cells propagates to the second population of NOT cells (αCa-αSc), where αCa inhibits the production of αSc molecules and output Y_1. For (20), three thresholding signals are processed and combined to output Y_0. The thresholding output signal B_2 experiences the same double NOT operation as the B_1 for Y_1. Simultaneously, the NOT cells (αSc-αCa) and ID cells (αSc-αCa) that respectively receive B_0 and B_1 signals from the thresholding cells, are employed to output αCa into the same propagation channel; the mixture of αCa molecules gives $B_1 + \overline{B_0}$. Then, the αCa molecules are processed by NOT cells (αCa-αSc) and the output is mixed with $\overline{\overline{B_2}}$ to obtain Y_0.

B. Circuit Analysis

To ease the analysis of the CSK design, we divide the end-to-end CSK system into four building blocks: ID block, NOT block, thresholding block, and propagation block. For the first three blocks, each block can contain one or more corresponding engineered cells. Recall that the engineered cell consists of three processes represented by ordinary differential equations (ODEs) in Sec. II-B; therefore, each cell can be described by a set of ODEs. As all three blocks experience the same molecule exchange, which is independent of any other processes, we first derive the impulse response of the molecule exchange in (4) as

$$h_1(t) = \eta e^{-k_d t}. \tag{21}$$

By solving the other processes of the engineered cell along with (21), we obtain the output of the first three blocks in the following. We then derive the output of the propagation block, which is modeled by a partial differential equation (PDE) and a set of constraints in Sec. II-C. Due to coupling the propagation building block and the engineered cell building block, η transforms into an exchange factor and (21) is rewritten as

$$h_1(t) = \psi e^{-k_d t}, \tag{22}$$

where ψ represent the exchange factor, which is determined as η / k_a.

1) ID Block: The ID block is governed by the ODEs in (4), (5), and (11). Relying on (22) and the Laplace transform, we solve this set of ODEs and obtain the accumulated concentration of the output molecules emitted by ID block as

$$C_O^{ID}(t) = \frac{\beta \xi [1 - e^{-(k_d + \xi)t}]}{\theta^n (k_d + \xi)} * \left\{ 1 - \frac{1}{1 + [\theta C_I(t) * h_1(t)]^n} \right\}, \tag{23}$$

where $*$ is the convolution operator. We also define operator $\mathcal{I}[\cdot]$ to characterize the expected net change of the output concentration released on the surface of the cells during the interval $[t, t+t_s]$ as

$$\mathcal{I}[C_I(t), I, O] = C_O^{ID}(t + t_s) - C_O^{ID}(t), \tag{24}$$

where I is the input molecule type, O is the output molecule type, and t_s is the time step. As different molecule types have different production rates, degradation rates, Hill function coefficients, and parameter fitting data, the variables I and O determine the values of β, θ, n, and k_d.

2) NOT Block: The NOT block is defined by the ODEs in (4), (6), (7), and (11). By solving (6) through Laplace transform, we can derive the concentration of the repressor molecules as

$$C_R^{NOT}(t) = \frac{\beta e^{-k_d t}}{\theta^n} * \left\{ 1 - \frac{1}{1 + [\theta C_I(t) * h_1(t)]^n} \right\}. \tag{25}$$

We then perform Laplace transform on (7) and (11) and arrive at the accumulated concentration of the output molecules emitted by NOT block as

$$C_O^{NOT}(t) = \frac{\xi \beta [1 - e^{-(k_d + \xi)t}]}{k_d + \xi} * \frac{1}{1 + [\theta C_R^{NOT}(t)]^n}. \tag{26}$$

We define operator $\mathcal{N}[\cdot]$ to characterize the net change of the output concentration and this can be expressed based on (26) as

$$\mathcal{N}[C_I(t), I, O] = C_O^{NOT}(t + t_s) - C_O^{NOT}(t). \tag{27}$$

3) Thresholding Block: The thresholding block is described by the ODEs in (4), (7), (10), and (11). Due to the non-linearity of (10) caused by the bimolecular reaction, we numerically solve the thresholding block by discretizing the processes into many time intervals with step t_s, such that $C_R^{Th}[m] = C_R^{Th}(mt_s)$, in the following theorem.

Theorem 1. *The concentration of the repressor at time mt_s can be approximated as* (28) *and shown in the top of next page, where* $C_{R_0}[m] = \{t_s f_R[m-1] + C_R^{Th}[m-1]\} e^{-k_d t_s}$ *and* $C_{I_0}[m] = \{t_s \eta C_I[m-1] + C_{Iin}^{Th}[m-1]\} e^{-k_d t_s}$ *are the concentrations of the repressor and input molecules at time* $(m-1)t_s$, *respectively.*

Proof: See APPENDIX A. ∎

The concentration of the input molecules is described by the ODE in (9), which has a similar form as (10); therefore, according to **Theorem 1**, the input concentration at mt_s can be derived as (29) and shown in the top of next page. Since the output production and the release processes for the thresholding block are the same as the NOT block, the output concentration emitted by the thresholding block can be expressed as

$$C_O^{Th}(t) = \frac{\xi \beta [1 - e^{-(k_d + \xi)t}]}{k_d + \xi} * \frac{1}{1 + [\theta C_R^{Th}(t)]^n}. \tag{30}$$

We then characterize the net change of the output concentration of the thresholding block as an operator $\mathcal{T}[\cdot]$, which can be expressed as

$$\mathcal{T}[C_I(t), C_{Th}(t), I, O] = C_O^{Th}[t + t_s] - C_O^{Th}(t). \tag{31}$$

4) Propagation Block: The propagation of signaling molecules is governed by the PDE in (13), with the initial condition in (14) and the boundary conditions in (15)-(17). For the propagation block, we are interested in the number of absorbed molecules after propagating through the channel, which can be derived in the following theorem.

$$C_{\text{R}}^{\text{Th}}[m] \approx \begin{cases} \dfrac{C_{\text{R}_0}[m](C_{\text{R}_0}[m] - C_{\text{I}_0}[m])}{C_{\text{R}_0}[m] - C_{\text{I}_0}[m]\exp[-k_{\text{f}}t_s(C_{\text{R}_0}[m] - C_{\text{I}_0}[m])]}, & C_{\text{R}_0}[m] \neq C_{\text{I}_0}[m] \\[4mm] \dfrac{C_{\text{R}_0}[m]}{1 + k_{\text{f}}t_s C_{\text{R}_0}[m]}, & C_{\text{R}_0}[m] = C_{\text{I}_0}[m] \end{cases} \tag{28}$$

$$C_{\text{Iin}}^{\text{Th}}[m] \approx \begin{cases} \dfrac{C_{\text{I}_0}[m](C_{\text{I}_0}[m] - C_{\text{R}_0}[m])}{C_{\text{I}_0}[m] - C_{\text{R}_0}[m]\exp[-k_{\text{f}}t_s(C_{\text{I}_0}[m] - C_{\text{R}_0}[m])]}, & C_{\text{I}_0}[m] \neq C_{\text{R}_0}[m] \\[4mm] \dfrac{C_{\text{I}_0}[m]}{1 + k_{\text{f}}t_s C_{\text{I}_0}[m]}, & C_{\text{I}_0}[m] = C_{\text{R}_0}[m] \end{cases} \tag{29}$$

$$h_2(t|L, \mathcal{S}_{\text{e}}, \mathcal{S}_{\text{a}}) = \frac{2H}{W} \sum_{l=1}^{\infty} \frac{(\lambda_l^2 + G_1^2)\cos(\lambda_l L)}{L(\lambda_l^2 + G_1^2) + G_1} e^{\frac{u(2L-ut)}{4D} - \lambda_l^2 Dt - k_{\text{d}}t} \Big[(y_{0_2} - y_{0_1})(y_2 - y_1)$$
$$+ 2\sum_{i=1}^{\infty} [\sin(\gamma_i y_{0_2}) - \sin(\gamma_i y_{0_1})][\sin(\gamma_i y_2) - \sin(\gamma_i y_1)]e^{-\gamma_i^2 Dt} \Big], \tag{33}$$

Theorem 2. *After a propagation of a distance L, the concentration of the molecules absorbed by a surface $\mathcal{S}_{\text{a}} = [y_1 \leq y \leq y_2, 0 \leq z \leq H]$ during $[t, t+t_s]$ is derived as*

$$\mathcal{P}[C_{\text{IP}}(t), L, \mathcal{S}_{\text{e}}, \mathcal{S}_{\text{a}}] = \frac{k_{\text{a}}t_s}{H(y_2 - y_1)} \Big[C_{\text{IP}}(t) * h_2(t|L, \mathcal{S}_{\text{e}}, \mathcal{S}_{\text{a}}) \Big], \tag{32}$$

where $C_{\text{IP}}(t) = \dfrac{\sum_{\text{d} \in \mathcal{S}_{\text{e}}} C(\mathbf{d}, t)}{H(y_{0_2} - y_{0_1})}$ is the normalized input concentration of the propagation block, $\mathcal{S}_{\text{e}} = [y_{0_1} \leq y \leq y_{0_2}, 0 \leq z \leq H]$ is the emission area along the y-z plane, and $h_2(t|L, \mathcal{S}_{\text{e}}, \mathcal{S}_{\text{a}})$ is derived as (33), where λ_l is the root of a transcendental equation $\lambda_l \tan(\lambda_l L) = G_1$, $\gamma_i = \frac{i\pi}{W}$, $\cos(\cdot)$ is a cosine function, and $G_1 = \frac{k_{\text{a}}}{D}$.

Proof: See APPENDIX B. ∎

C. Analysis of BCSK and QCSK Systems

Relying on the analyses of the building blocks, we analyze our proposed BCSK and QCSK systems in the following. Let $\mathcal{S}_{i,j}$ denotes a surface on the y-z plane with $W_i \leq y \leq W_j$ and $0 \leq z \leq H$.

1) BCSK Analysis: As shown in Fig. 8, the input molecules αCa propagate to a population of ID cells at the transmitter side, and the net concentration of the released molecules by ID cells (αCa-DOX) at $x = L_2$ during $[t, t+t_s]$ can be expressed as

$$C_{\text{Tx}}(t, t+t_s) = \mathcal{I}\Big[\mathcal{P}[C_{\text{S}_0}(t), L_1 - L_0, \mathcal{S}_{0,1}, \mathcal{S}_{0,1}], \alpha\text{Ca}, \text{DOX} \Big]. \tag{34}$$

The released DOX molecules propagate inside the channel to a population of thresholding cells, and the net concentration of the molecules absorbed by the thresholding cells at $x = L_3$ is derived as

$$C_{\text{Ch}}(t, t+t_s) = \mathcal{P}[C_{\text{Tx}}(t, t+t_s), L_3 - L_2, \mathcal{S}_{0,1}, \mathcal{S}_{0,1}]. \tag{35}$$

The thresholding cells (DOX-αSc) perform BCSK demodulation, where the net concentration of the molecules detected by

the detection wall at $x = L_5$ can be derived as

$$C_{\text{Rx}}(t, t+t_s) = \mathcal{P}\Big[\mathcal{T}[C_{\text{Ch}}(t, t+t_s), C_{\text{Th}}(t), \text{DOX}, \alpha\text{Sc}] , L_5 - L_4, \mathcal{S}_{0,1}, \mathcal{S}_{0,1} \Big]. \tag{36}$$

2) QCSK Analysis: For the QCSK transmitter in Fig. 9, the input bits S_1 and S_0 control the total concentration of DOX molecules generated at the transmitter. The net concentration of DOX molecules released by ID cells (αCa-DOX) at $x = L_2$ during $[t, t+t_s]$ with respect to input bit S_1 and S_0 can be expressed as

$$C_{\text{Tx}}^{S_1}(t, t+t_s) = \mathcal{I}\Big[\mathcal{P}[C_{\text{S}_1}(t), L_1 - L_0, \mathcal{S}_{3,4}, \mathcal{S}_{3,4}], \alpha\text{Ca}, \text{DOX} \Big], \tag{37}$$

and

$$C_{\text{Tx}}^{S_0}(t, t+t_s) = \mathcal{I}\Big[\mathcal{P}[C_{\text{S}_0}(t), L_1 - L_0, \mathcal{S}_{0,3}, \mathcal{S}_{0,3}], \alpha\text{Ca}, \text{DOX} \Big]. \tag{38}$$

Therefore, the total net concentration of DOX molecules generated by the QCSK transmitter can be expressed as

$$C_{\text{Tx}}(t, t+t_s) = \frac{W_3}{W_4} C_{\text{Tx}}^{S_1}(t, t+t_s) + \frac{W_4 - W_3}{W_4} C_{\text{Tx}}^{S_0}(t, t+t_s). \tag{39}$$

Then, the molecules propagate along a channel to populations of thresholding cells at the QCSK receiver and the net concentration of the modulated signals received by the thresholding cells located along $\mathcal{S}_{i,j}$ is expressed as

$$C_{\text{Ch}}[t, t+t_s|\mathcal{S}_{i,j}] = \mathcal{P}\Big[C_{\text{Tx}}^{S_1}(t, t+t_s), L_3 - L_2, \mathcal{S}_{3,4}, \mathcal{S}_{i,j} \Big] + \mathcal{P}\Big[C_{\text{Tx}}^{S_0}(t, t+t_s), L_3 - L_2, \mathcal{S}_{0,3}, \mathcal{S}_{i,j} \Big]. \tag{40}$$

Based on (31), the net concentration of αSc molecules generated by thresholding unit B_j can be derived as

$$C_{B_l}(t, t+t_s|\mathcal{S}_{i,j}) = \mathcal{T}\Big[C_{\text{Ch}}[t, t+t_s|\mathcal{S}_{i,j}], C_{\text{Th}}^{B_l}(t), \text{DOX}, \alpha\text{Sc} \Big]. \tag{41}$$

The αSc molecules generated from the thresholding cells propagate to the back-end module of the receiver, which is

$$C_{Y_1}(t, t+t_s) = \mathcal{P}\Bigg\{ \mathcal{N}\bigg[\mathcal{P}\Big\{ \underbrace{\mathcal{N}\big[\mathcal{P}[C_{B_1}(t, t+t_s|\mathcal{S}_{3,4}), L_5 - L_4, \mathcal{S}_{3,4}, \mathcal{S}_{3,4}], \alpha\mathrm{Sc}, \alpha\mathrm{Ca}\big], L_7 - L_6, \mathcal{S}_{3,4}, \mathcal{S}_{3,4}}_{\text{First population of NOT cells: } \overline{B_1}} \Big\}, \alpha\mathrm{Ca}, \alpha\mathrm{Sc} \bigg]$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{Second population of NOT cells: } \overline{\overline{B_1}}}$$

$$, L_9 - L_8, \mathcal{S}_{3,4}, \mathcal{S}_{3,4} \Bigg\} \tag{42}$$

$$C_1(t, t+t_s) = \mathcal{N}\bigg[\mathcal{P}\Big\{ \underbrace{\mathcal{N}\big[\mathcal{P}[C_{B_2}(t, t+t_s|\mathcal{S}_{2,3}), L_5 - L_4, \mathcal{S}_{2,3}, \mathcal{S}_{2,3}], \alpha\mathrm{Sc}, \alpha\mathrm{Ca}\big], L_7 - L_6, \mathcal{S}_{2,3}, \mathcal{S}_{2,3}}_{\text{First population of NOT cells: } \overline{B_2}} \Big\}, \alpha\mathrm{Ca}, \alpha\mathrm{Sc} \bigg] \tag{43}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{Second population of NOT cells: } \overline{\overline{B_2}}}$$

$$C_2(t, t+t_s) = \mathcal{N}\bigg[\mathcal{P}\Big\{ \underbrace{\mathcal{I}\big[\mathcal{P}[C_{B_1}(t, t+t_s|\mathcal{S}_{1,2}), L_5 - L_4, \mathcal{S}_{1,2}, \mathcal{S}_{1,2}], \alpha\mathrm{Sc}, \alpha\mathrm{Ca}\big], L_7 - L_6, \mathcal{S}_{1,2}, \mathcal{S}_{0,2}}_{\text{First population of ID cells: } B_1} \Big\}$$

$$+ \mathcal{P}\Big\{ \underbrace{\mathcal{N}\big[\mathcal{P}[C_{B_0}(t, t+t_s|\mathcal{S}_{0,1}), L_5 - L_4, \mathcal{S}_{0,1}, \mathcal{S}_{0,1}], \alpha\mathrm{Sc}, \alpha\mathrm{Ca}\big], L_7 - L_6, \mathcal{S}_{0,1}, \mathcal{S}_{0,2}}_{\text{First population of NOT cells: } \overline{B_0}} \Big\}, \alpha\mathrm{Ca}, \alpha\mathrm{Sc} \bigg] \tag{44}$$

$$C_{Y_0}(t, t+t_s) = \mathcal{P}[C_1(t, t+t_s), L_9 - L_8, \mathcal{S}_{2,3}, \mathcal{S}_{0,3}] + \mathcal{P}[C_2(t, t+t_s), L_9 - L_8, \mathcal{S}_{0,2}, \mathcal{S}_{0,3}] \tag{45}$$

responsible for processing the signal back into Y_1 and Y_0. From (19), the output of bit Y_1 is the signal processing of double NOT operation on B_1 signal. Thus, the net concentration of $\alpha\mathrm{Sc}$ molecules for bit Y_1 detected at $x = L_9$ by the detection wall can be expressed as (42). For output bit Y_0, it is the mix of $\overline{\overline{B_2}}$ and $\overline{B_1 + \overline{B_0}}$ as in (20). Similar to (42), the output of $\overline{\overline{B_2}}$ is generated through a double NOT operation at $x = L_8$, which can be expressed as (43). The output of $\overline{B_1 + \overline{B_0}}$ at $x = L_8$ can be expressed as a NOT operation of a summation of the concentration from B_1 and $\overline{B_0}$, which is (44). Therefore, the output of Y_0 detected at $x = L_9$ can be expressed as a summation of $\alpha\mathrm{Sc}$ molecules propagated from $x = L_8$ as (45).

IV. NUMERICAL RESULTS

In this section, we simulate our proposed BCSK and QCSK systems in BSim and Matlab to examine the effectiveness of our CSK design framework and validate our theoretical analysis. The simulation of molecule propagation in both simulators employs a particle-based method [4], [36]. This method involves storing the number of molecules and their respective locations. The displacement of molecules due to advection and diffusion in a 3D fluid environment in one time step is calculated as $\{\mathcal{N}(0, 2Dt_s) + ut_s, \mathcal{N}(0, 2Dt_s), \mathcal{N}(0, 2Dt_s)\}$. Additionally, molecules undergo degradation and are removed from the environment with a probability of $k_d t_s$ during the same time step. To account for the partially absorbing boundary conditions in (17), a collision with the engineered cell wall can result in either absorption with a probability $k_a\sqrt{\pi t_s/D}$ or reflection otherwise. The collided molecules are identified when the distance in the x-direction between each molecule and the transmitter exceeds the length L in each propagation channel. For the reflective boundary described in (15) and (16), molecules positioned outside the boundary $0 \le y \le W$ and $0 \le z \le H$ in the y- and z-directions, respectively, are reflected back into the environment.

In the simulation of the engineered cell, both simulators use an ODE solver to implement the system of ODEs. However, there is a fundamental difference between BSim and MATLAB. BSim is a computational agent-based simulator for modeling cell dynamics in a typical environment [37]. In this approach, each engineered cell is treated as an individual agent, which is simulated using an ODE solver. The number of agents simulated for BCSK and QCSK is 32 and 192, respectively. On the other hand, MATLAB considers all the engineered cells within a population as a single agent. In this MATLAB approach, the total concentration of information molecules colliding with the cell is inputted into the ODE solver for simulation, simplifying the process by collectively modeling the behavior of the entire cell population rather than representing each cell individually.

The simulation time interval t_s is set as 0.01s and the results are averaged over 2000 realizations. We set the molecules propagation parameters in (13) as $D = 89\mu\mathrm{m}^2$ s^{-1}, $u = 0.1\mu\mathrm{m}$ s^{-1}, and $k_a = 9\mu\mathrm{m}$ s^{-1}. We also set the parameters for engineered cell processes as $\xi = 20$s^{-1} and $k_f = 1$nM s^{-1}. It is noted that as the input molecules of an engineered cell are those absorbed on the cell membrane, we set the molecule exchange rate in (4) as $\eta = 1$. The other parameters are summarized in Tables III and IV. In all the figures, we also use "Ana." and "Sim." to abbreviate "Analytical" and "Simulation", respectively.

TABLE III
THE PARAMETERS OF THE ENGINEERED CELLS.

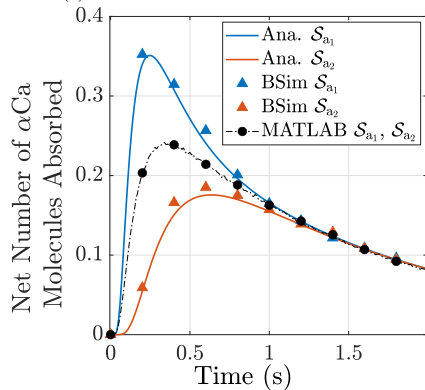| Molecule	β (nM min^{-1})	θ (nM^{-1})	n	k_{d} (min^{-1})	Reference
αCa	0.0369	0.26	0.9	0.05	[19]
αSc	0.162	0.167	1.2	0.023	[19]
DOX	0.162	0.167	1.2	0.023	-
R (Repressor)	0.615	1550	2	0.15	[19]

TABLE IV
THE DIMENSION OF THE CSK SYSTEMS.

Parameter	L_0	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	W_0	W_1	W_2	W_3	W_4	H
Value (μm)	0	1	4	39	42	43	46	47	50	55	0	2.5	5	10	15	3
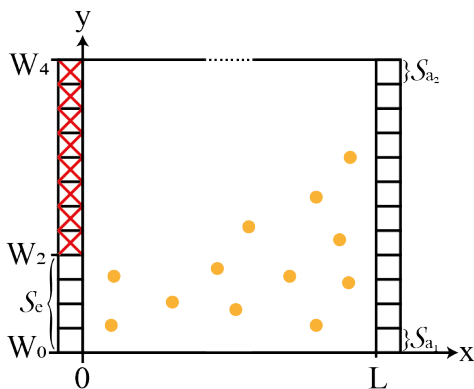
(a) The schematic of the simulation.


(b) The reception of molecules at $L = 10$.

Fig. 10. The comparison of BSim and MATLAB simulators. The molecules are released by the cells located along S_{e} with $W_0 \leq y \leq W_2$ and received by the cell at $x = L$ on $\mathcal{S}_{\mathrm{a}_1}$ and $\mathcal{S}_{\mathrm{a}_2}$ with $0 \leq y \leq 1.25$ and $13.75 \leq y \leq 15$, respectively. The cells at $x = 0$ along $W_2 \leq y \leq W_4$ (indicated as a box with a red cross) are not active and no molecules are released by those cells.

A. BSim vs. MATLAB

We first investigate the simulation accuracy of BSim and Matlab. Specifically, as illustrated in Fig. 10(a), only the cells located along S_{e} at $x = 0$ emit molecules into the channel, and two absorbing surfaces $\mathcal{S}_{\mathrm{a}_1}$ and $\mathcal{S}_{\mathrm{a}_2}$ are considered. The simulation results are averaged over 10,000 independent impulse emissions of 500 DOX molecules by the cells at time $t = 0$. Fig. 10(b) plots the simulation results by BSim and MATLAB and the corresponding analytical values. We observe that the simulation results obtained from BSim have a close match to the analytical values whereas MATLAB simulations
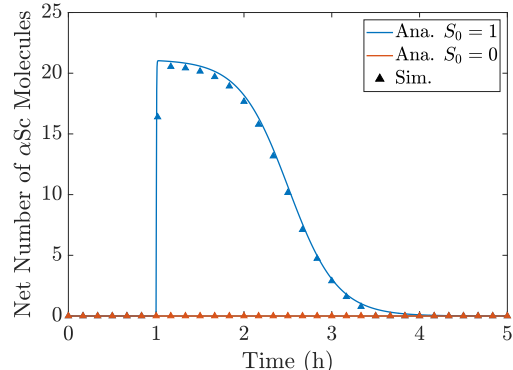
(a) BCSK transmitter.
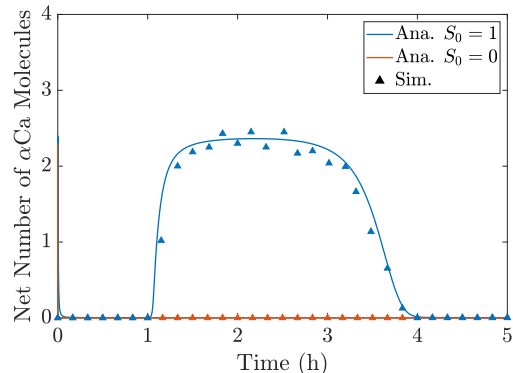
(b) BCSK receiver.

Fig. 11. The performance of our proposed BCSK system.

fail to capture the number of absorbed molecules. The reason for this difference is that the MATLAB simulates all the cells in the population located at $x = 0$ or $x = L$ as a single agent, therefore the number of absorbed molecules for $\mathcal{S}_{\mathrm{a}_1}$ and $\mathcal{S}_{\mathrm{a}_2}$ are the same, which ignores that the molecules released by S_{e} need more time to arrive at $\mathcal{S}_{\mathrm{a}_2}$. This signifies the importance of an agent-based simulator to capture the individual behavior of each cell. Based on this, only BSim is used for the rest of the simulations.

B. BCSK System

To evaluate our proposed BCSK transceiver design in Fig. 8, a rectangular input with the amplitude of 750 αCa molecules (50nM) and a duration of 10s are released to the BCSK transmitter at $t = 1$h to represent bit-1, i.e., $S_0 = 1$. We set the width of the channel as $W_1 = 5\mu$m. For the thresholding unit,
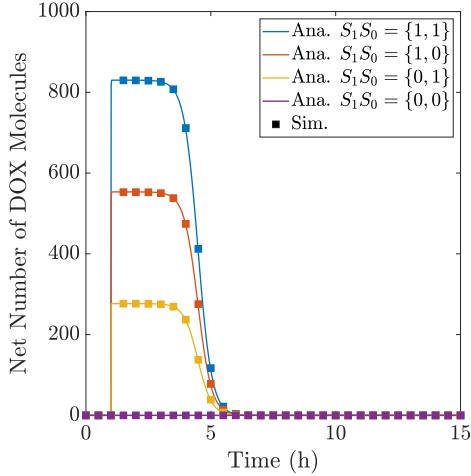
Fig. 12. The performance of our proposed QCSK transmitter.

we set the concentration of threshold molecules as 0.01nM. Fig. 11 plots the simulation results of our proposed BCSK system obtained from BSim and the analytical values obtained from (34) and (36). We observe that there is a close match between the simulation points and the analytical curves, which validates our theoretical analysis. We also observe that our proposed transmitter releases a certain number of molecules after the addition of 50nM and no molecules are released; at the same time, it also shows that our proposed receiver is able to release information molecules in response to the signals from the OOK transmitter, demonstrating the effectiveness of OOK modulation and demodulation functionalities.

C. QCSK System

To examine the performance of the QCSK design in Fig. 9, we consider all possible input combinations: $S_1 S_0 = \{1, 1\}$, $S_1 S_0 = \{1, 0\}$, $S_1 S_0 = \{0, 1\}$, and $S_1 S_0 = \{0, 0\}$. For bit-1 input, rectangular inputs with amplitude of 50nM of αCa molecules are released to the QCSK transmitter at $t = 1$h.[3] Due to the larger dimension of the QCSK system compared to BCSK, the information molecules take more time to traverse, leading to increased degradation during propagation. Moreover, multiple cell layers and an absorption-to-production ratio greater than one contribute to the additional consumption of information molecules during cell signal processing. Accordingly, we extend the duration of the rectangular inputs to 30 minutes to enhance the likelihood of molecules reaching the receiver. We also increase all the production rates by 40 folds, which can be achieved by amplifying circuits in practice [38]. For the thresholding unit, we set the thresholding parameters as $C_{\text{Th}}^{B_2}(t) = 0.7$nM, $C_{\text{Th}}^{B_1}(t) = 0.45$nM, and $C_{\text{Th}}^{B_0}(t) = 0.1$nM. Fig. 12 plots the simulation results of our proposed QCSK transmitter and the analytical results calculated by (39). We observe that the QCSK transmitter can modulate two bits into four different concentration levels, thus achieving QCSK modulation. Fig. 13 plots the simulation

[3]With respect to the geometry provided in Table IV, the amplitude is equivalent to 750 αCa molecules and 1500 αCa molecules for input bits S_0 and S_1, respectively.

(a) $S_1 S_0 = \{1, 1\}$

(b) $S_1 S_0 = \{1, 0\}$
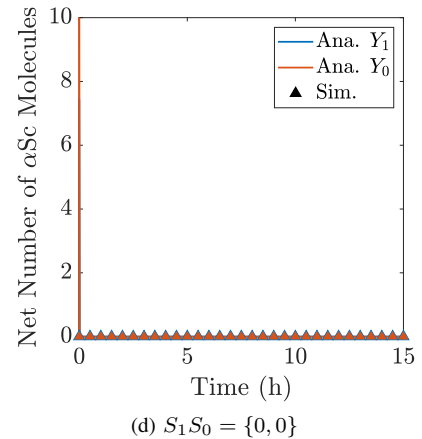

(c) $S_1 S_0 = \{0, 1\}$
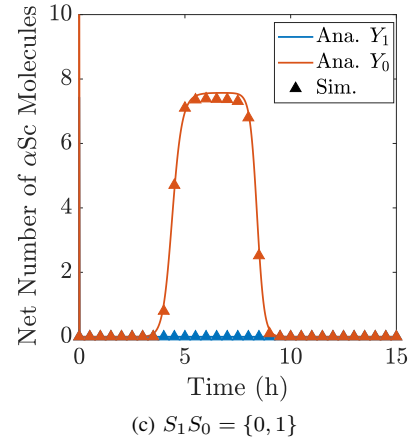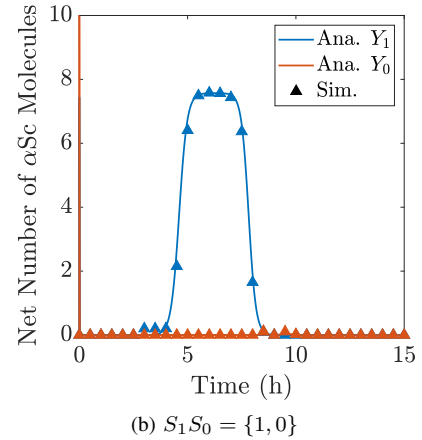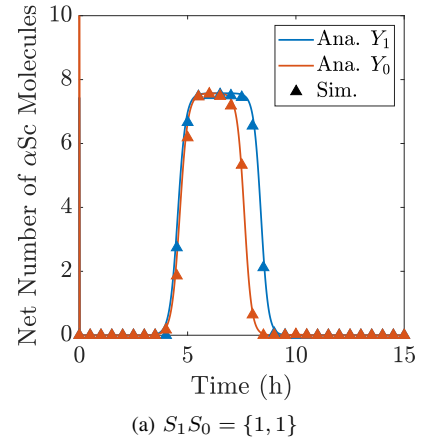

(d) $S_1 S_0 = \{0, 0\}$

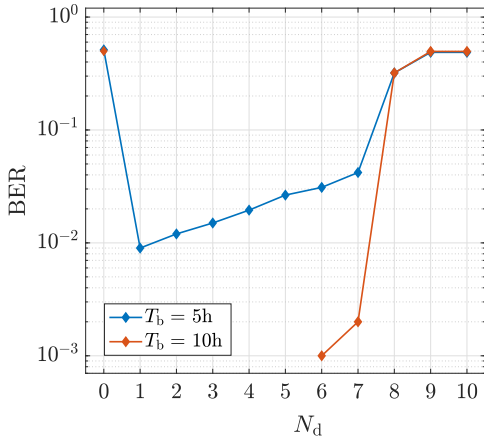Fig. 13. The performance of our proposed QCSK receiver.

Fig. 14. Investigation of the impact of detection parameter on the QCSK performance.

results of our proposed QCSK receiver and analytical results in (42) and (45). It is clear that the receiver is capable of distinguishing the concentration levels and demodulating the signals back to two bits. Furthermore, both figures present a close match between the analytical and simulation results, which further validates our theoretical analysis. Nevertheless, we can observe a disparity in the width of the signals Y_1 and Y_0 as depicted in Fig. 13(a). This discrepancy arises due to the variation in the time period of the thresholding bit outputs, wherein a higher thresholding concentration impedes the production of output molecules for a longer duration.

D. Communication Performance of QCSK Design

We also examine the communication performance of our proposed QCSK system. In particular, a 1000-bit sequence is randomly generated and transmitted under different bit intervals T_b and the simulation parameters for the QCSK system in Sec. IV-C are also used. According to the QCSK receiver performance in Fig. 13, the average response of bit-1 is roughly 8 molecules and appears at 5h after the start of the transmission; therefore, we sample the output bits 5h after each transmission. Fig. 14 plots the BER against the detection thresholding value N_d, which is the number of absorbed molecules required for the receiver to decide whether bit-1 is transmitted. For $T_b = 5$h, we first observe that when $N_d = 0$, the maximum BER for QCSK is obtained because all bits are detected as bit-1. The BER increases as N_d approaches the average number of detected molecules for bit-1. When N_d is greater than the average detected molecules, all bits are detected as bit-0 and this results in the maximum BER. Moreover, we find that $S_0 = 1$ is often detected as bit-0 when $S_1 S_0 = \{0, 1\}$ is transmitted after $S_1 S_0 = \{1, 1\}$. The thresholding output B_1 for Y_0 is bit-1 when $S_1 S_0 = \{1, 1\}$ while is bit-0 when $S_1 S_0 = \{0, 1\}$ (see Fig. 9). As $S_1 S_0 = \{1, 1\}$ leads to a higher concentration level than $S_1 S_0 = \{0, 1\}$ and the higher received concentration the longer duration of the thresholding output, the bit-1 of B_1 when $S_1 S_0 = \{1, 1\}$ can accumulate in the channel and interfere the bit-0 of B_1 when $S_1 S_0 = \{0, 1\}$, thus causing the wrong detection of S_0. This error can be mitigated by increasing T_b to ensure all molecules from the previous bit

undergo natural degradation before arriving $x = L_7$. From Fig. 14, we can observe that when $T_b = 10$h, the BER significantly decreases and no error is detected for $N_d < 6$.[4]

V. CONCLUSION

In this paper, we proposed a design principle to realize the concentration shift keying (CSK) system based on multicellular consortia and spatial segregation. According to our design principle, we developed a small library of engineered cells to realize binary CSK (BCSK) and quadruple CSK (QCSK) systems. Notably, the gene expressions and the involved molecules in our proposed system were specified in the synthetic biology domain. Then, we established an analytical framework that divides a system into four building blocks to theoretically characterized the proposed systems. We derived the impulse response of each building block, and the cascade of those impulse responses resulted in the analysis of the proposed system. Simulation results from BSim showed that our systems responded appropriately to input signals and closely matched our theoretical analysis. Additionally, we demonstrated that intersymbol interference (ISI) can be mitigated through a careful selection of bit interval and detection thresholding value. Our CSK design principle is inspired by successful *in vitro* implementations of 3-input majority and 4-input comparator using multicellular consortia with spatial segregation [26]. In addition, the gene regulatory networks in our library of engineered cells were verified through wet lab experiments [19], demonstrating the potential of our communication scheme in real-world biological settings. However, practical challenges in deploying this scheme in the real world remain, and our future work involves wet-lab validation and refining the mathematical framework by incorporating parabolic flow and stochastic analysis.

APPENDIX A
PROOF OF **THEOREM 1**

The dynamics of the repressor concentration in (10) can be divided into three processes using the operator splitting method [12] as

Production: $\qquad \dfrac{dC_R(t)}{dt} = f_R(t).$ (46a)

Degradation: $\qquad \dfrac{dC_R(t)}{dt} = -k_d C_R(t).$ (46b)

Bimolecular Reaction: $\qquad \dfrac{dC_R(t)}{dt} = -k_f C_{lin}(t) C_R(t).$ (46c)

We assume that the processes occur in the order of production, degradation, and bimolecular reaction[5], where the initial concentration of each process is the molecule concentration that resulted from previous processes. For the time interval $[(m-1)t_s, mt_s]$, the initial concentration for the degradation is a superposition of molecules produced by the gene expression

---

[4]The BER for $N_d < 6$ is presumably lower than 10^{-3}, thus a 1000-bit sequence is insufficient to detect the error accurately.

[5]When we consider the biomolecular reaction, the production and degradation occur simultaneously; however, the newly produced molecules do not experience the degradation instantly, but in the following time slot.

$$h_2(\mathbf{d}, t) = \frac{2}{W} \sum_{l=1}^{\infty} \frac{(\lambda_l^2 + G_1^2) \cos(\lambda_l x)}{L(\lambda_l^2 + G_1^2) + G_1} e^{\frac{u(2x - ut)}{4D} - \lambda_l^2 Dt - k_d t} \left[1 + 2 \sum_{i=1}^{\infty} \cos(\gamma_i y_0) \cos(\gamma_i y) e^{-\gamma_i^2 Dt} \right] \quad (53)$$

in (46a) and the residual repressor from the bimolecular reaction in the previous time interval $[(m-2)t_s, (m-1)t_s]$, which can be expressed as

$$t_s f_R[m-1] + C_R^{Th}[m-1], \quad (47)$$

where $f_R[m-1] = f_R\big((m-1)t_s\big)$. This portion of repressor experiences the natural degradation in (46b), resulting in the initial concentration for the bimolecular reaction, which can be expressed as

$$C_{R_0}[m] = \{t_s f_R[m-1] + C_R^{Th}[m-1]\} e^{-k_d t_s}. \quad (48)$$

Then, to find the repressor concentration at time mt_s, two cases of initial concentration need to be considered:

- **Case 1:** When $C_{R_0}[m] \neq C_{I_0}[m]$, chemical reaction in (8) occurs. Due to the one-to-one stoichiometric relation between reactants, eq. (46c) can be rewritten as

$$\frac{dC_R^{Th}(t)}{dt} = -k_f C_R^{Th}(t)\{C_R^{Th}(t) - (C_{R_0}[m] - C_{I_0}[m])\}. \quad (49)$$

By rearranging (49) and taking the integral during the time interval $[(m-1)t_s, mt_s]$, the concentration of the repressor at mt_s is derived as

$$C_R^{Th}[m] = \frac{C_{R_0}[m](C_{R_0}[m] - C_{I_0}[m])}{C_{R_0}[m] - C_{I_0}[m] \exp[-k_f t_s(C_{R_0}[m] - C_{I_0}[m])]}. \quad (50)$$

- **Case 2:** When $C_{R_0}[m] = C_{I_0}[m]$, eq. (49) can be rewritten as

$$\frac{dC_R^{Th}(t)}{dt} = -k_f \{C_R^{Th}(t)\}^2. \quad (51)$$

By applying the same step as in **Case 1** on (51), we arrive at

$$C_R^{Th}[m] = \frac{C_{R_0}[m]}{1 + k_f t_s C_{R_0}[m]}. \quad (52)$$

Combining those two cases with (48) as the initial condition, we arrive at (28).

APPENDIX B
PROOF OF **THEOREM 2**

We use the separation variable method [1] by assuming $C(\mathbf{d}, t) = X(x, t)Y(y)Z(z)$ to find the impulse response of (13). The impulse response $h_2(\mathbf{d}, t)$ of propagation block is derived as (53). The impulse response of the molecule propagation to the absorbing boundary $\mathcal{S}_a$ when the molecules are released by the emitting boundary $\mathcal{S}_e$ can be expressed as

$$h_2(t|L, \mathcal{S}_e, \mathcal{S}_a) = \int_{y_{0_1}}^{y_{0_2}} \int_0^H \int_{y_1}^{y_2} h_2(\mathbf{d} = [L, y, z], t) dy dz dy_0. \quad (54)$$

The above equation can lead to (33). Thus, the concentration of the molecules can be expressed as

$$C(t|L, \mathcal{S}_e, \mathcal{S}_a) = C_{IP}(t) * h_2(t|L, \mathcal{S}_e, \mathcal{S}_a). \quad (55)$$

To characterize the number of molecules absorbed on the absorbing wall, we define the change in the total number of particles due to the particle flux as [39, eq. (3.20)]

$$\frac{dN(t)}{dt} = \int_{\mathcal{S}} D\nabla C(\mathbf{d}, t) d\mathcal{S}, \quad (56)$$

where $\mathcal{S}$ is the surface of the given region with differential element $d\mathcal{S}$. The change in the total number of particles in (56) on the absorbing wall is equivalent to the flux at the boundary specified in (17); therefore, we can arrive at

$$N(t|L, \mathcal{S}_e, \mathcal{S}_a) = \int_0^t k_a C(\tau|L, \mathcal{S}_e, \mathcal{S}_a) d\tau. \quad (57)$$

Dividing (57) by the cross-section of the detection area $H(y_2 - y_1)$ to translate from the number of molecules to concentration and taking the net change, we can arrive at (32).

REFERENCES

[1] R. Jadsadaphongphaiboool, D. Bi, and Y. Deng, "Binary Concentration Shift Keying Realization for Molecular Communication via Genetic Circuits," in *Proc. ACM NanoCom*, Oct. 2022, pp. 1–6.

[2] N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo, "A comprehensive survey of recent advancements in molecular communication," *IEEE Commun. Surv. Tutor.*, vol. 18, no. 3, pp. 1887–1919, Feb. 2016.

[3] M. Pierobon and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 602–611, May 2010.

[4] Y. Deng, A. Noel, M. Elkashlan, A. Nallanathan, and K. C. Cheung, "Modeling and simulation of molecular communication systems with a reversible adsorption receiver," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 1, no. 4, pp. 347–362, Dec. 2015.

[5] S. Kadloor, R. S. Adve, and A. W. Eckford, "Molecular communication using brownian motion with drift," *IEEE Trans. Nanobiosci.*, vol. 11, no. 2, pp. 89–99, Jun. 2012.

[6] T. Nakano, Y. Okaie, and J.-Q. Liu, "Channel model and capacity analysis of molecular communication with brownian motion," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 797–800, Jun. 2012.

[7] D. Kilinc and O. B. Akan, "Receiver design for molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 705–714, Dec. 2013.

[8] M. Kuscu, E. Dinc, B. A. Bilgin, H. Ramezani, and O. B. Akan, "Transmitter and receiver architectures for molecular communications: A survey on physical design with modulation, coding, and detection techniques," *Proc. IEEE*, vol. 107, no. 7, pp. 1302–1341, Jul. 2019.

[9] C. A. Söldner, E. Socher, V. Jamali, W. Wicke, A. Ahmadzadeh, H.-G. Breitinger, A. Burkovski, K. Castiglione, R. Schober, and H. Sticht, "A survey of biological building blocks for synthetic molecular communication systems," *IEEE Commun. Surv. Tutor.*, vol. 22, no. 4, pp. 2765–2800, Jul. 2020.

[10] D. Bi, A. Almpanis, A. Noel, Y. Deng, and R. Schober, "A survey of molecular communication in cell biology: Establishing a new hierarchy for interdisciplinary applications," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 3, pp. 1494–1545, Mar. 2021.

[11] M. S. Kuran, H. B. Yilmaz, I. Demirkol, N. Farsad, and A. Goldsmith, "A Survey on Modulation Techniques in Molecular Communication via Diffusion," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 1, pp. 7–28, Jan. 2021.

[12] D. Bi and Y. Deng, "Microfluidic QCSK Transmitter and Receiver Design for Molecular Communication," *IEEE Trans. Commun.*, vol. 70, no. 9, pp. 5837–5852, Sep. 2022.

[13] D. Bi, Y. Deng, M. Pierobon, and A. Nallanathan, "Chemical reactions-based microfluidic transmitter and receiver design for molecular communication," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5590–5605, Sep. 2020.

[14] D. Bi and Y. Deng, "Digital signal processing for molecular communication via chemical-reaction-based microfluidic circuits," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 26–32, May 2021.

[15] A. S. Khalil and J. J. Collins, "Synthetic biology: applications come of age," *Nat. Rev. Genet.*, vol. 11, no. 5, pp. 367–379, May 2010.

[16] A. Tamsir, J. J. Tabor, and C. A. Voigt, "Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'," *Nature*, vol. 469, no. 7329, pp. 212–215, Dec. 2011.

[17] B. Wang, R. I. Kitney, N. Joly, and M. Buck, "Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology," *Nat. Commun.*, vol. 2, no. 1, pp. 1–9, Oct. 2011.

[18] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in Escherichia coli," *Nature*, vol. 403, no. 6767, pp. 339–342, Jan. 2000.

[19] A. Urrios, J. Macía, R. Manzoni, N. Conde, A. Bonforti, E. de Nadal, F. Posas, and R. Solé, "A Synthetic Multicellular Memory Device," *ACS Synth. Biol.*, vol. 5, no. 8, pp. 862–873, Aug. 2016.

[20] B. D. Unluturk, A. O. Bicen, and I. F. Akyildiz, "Genetically Engineered Bacteria-Based BioTransceivers for Molecular Communication," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1271–1281, Feb. 2015.

[21] A. Marcone, M. Pierobon, and M. Magarini, "Parity-check coding based on genetic circuits for engineered molecular communication between biological cells," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6221–6236, Dec. 2018.

[22] A. A. Nielsen, B. S. Der, J. Shin, P. Vaidyanathan, V. Paralanov, E. A. Strychalski, D. Ross, D. Densmore, and C. A. Voigt, "Genetic circuit design automation," *Sci.*, vol. 352, no. 6281, p. aac7341, Apr. 2016.

[23] M. W. Gander, J. D. Vrana, W. E. Voje, J. M. Carothers, and E. Klavins, "Digital logic circuits in yeast with crispr-dcas9 nor gates," *Nat. Commun.*, vol. 8, no. 1, pp. 1–11, May 2017.

[24] J. Macía, F. Posas, and R. V. Solé, "Distributed computation: the new wave of synthetic biology devices," *Trends Biotechnol.*, vol. 30, no. 6, pp. 342–349, Jun. 2012.

[25] M. A. Al-Radhawi, A. P. Tran, E. A. Ernst, T. Chen, C. A. Voigt, and E. D. Sontag, "Distributed Implementation of Boolean Functions by Transcriptional Synthetic Circuits," *ACS Synth. Biol.*, vol. 9, no. 8, pp. 2172–2187, Aug. 2020.

[26] J. Macía, R. Manzoni, N. Conde, A. Urrios, E. de Nadal, R. Solé, and F. Posas, "Implementation of Complex Biological Logic Circuits Using Spatially Distributed Multicellular Consortia," *PLoS Comput. Biol.*, vol. 12, no. 2, p. e1004685, Feb. 2016.

[27] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks," in *Proc. IEEE ICC*, Jun. 2011, pp. 1–5.

[28] H. Kettani, "On the conversion between number systems," *IEEE Trans. Circuits Syst. II*, vol. 53, no. 11, pp. 1255–1258, Nov. 2006.

[29] A. Urrios, E. Gonzalez-Flo, D. Canadell, E. De Nadal, J. Macía, and F. Posas, "Plug-and-Play Multicellular Circuits with Time-Dependent Dynamic Responses," *ACS Synth. Biol.*, vol. 7, no. 4, pp. 1095–1104, Apr. 2018.

[30] M. D. Ciletti and M. M. Mano, *Digital design: with an introduction to the Verilog HDL*. New Jersey, USA: Prentice-Hall, 2007.

[31] G. Fiore, A. Matyjaszkiewicz, F. Annunziata, C. Grierson, N. J. Savery, L. Marucci, and M. Di Bernardo, "In-silico analysis and implementation of a multicellular feedback control strategy in a synthetic bacterial consortium," *ACS Synth. Biol.*, vol. 6, no. 3, pp. 507–517, Dec. 2017.

[32] B. P. Ingalls, *Mathematical modeling in systems biology: an introduction*. Cambridge, Massachusetts, USA: MIT press, Jul. 2013.

[33] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel, and R. Schober, "Channel Modeling for Diffusive Molecular Communication—A Tutorial Review," *Proc. IEEE*, vol. 107, no. 7, pp. 1256–1301, Jul. 2019.

[34] C. S. Kumar, *Microfluidic devices in nanotechnology: Fundamental concepts*. Hoboken, New Jersey, USA: John Wiley & Sons, 2010.

[35] R. Mosayebi, A. Ahmadzadeh, W. Wicke, V. Jamali, R. Schober, and M. Nasiri-Kenari, "Early cancer detection in blood vessels using mobile nanosensors," *IEEE Trans. Nanobiosci.*, vol. 18, no. 2, pp. 103–116, Apr. 2019.

[36] A. Q. Cai, K. A. Landman, and B. D. Hughes, "Modelling Directional Guidance and Motility Regulation in Cell Migration," *Bull. Math. Biol. 2005 68:1*, vol. 68, no. 1, pp. 25–52, Mar. 2006.

[37] A. Matyjaszkiewicz, G. Fiore, F. Annunziata, C. S. Grierson, N. J. Savery, L. Marucci, and M. Di Bernardo, "BSim 2.0: An Advanced Agent-Based Cell Simulator," *ACS Synth. Biol.*, vol. 6, no. 10, pp. 1969–1972, Oct. 2017.

[38] S. Mogas-Díez, E. Gonzalez-Flo, and J. Macía, "2D printed multicellular devices performing digital and analogue computation," *Nat. Commun.*, vol. 12, no. 1, pp. 1–10, Mar. 2021.

[39] K. Schulten and I. Kosztin, *Lectures in Theoretical Biophysics*. Urbana, Illinois, USA: University of Illinois, Apr. 2000.