Using Biochemical Circuits to Approximately Compute Log-Likelihood Ratio for Detecting Persistent Signals

CHUN TUNG CHOU, (Member, IEEE)
School of Computer Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia

e-mail: c.t.chou@unsw.edu.au

ABSTRACT Given that biochemical circuits can process information by using analog computation, a question is: What can biochemical circuits compute? This paper considers the problem of using biochemical circuits to distinguish persistent signals from transient ones. We define a statistical detection problem over a reaction pathway consisting of three species: an inducer, a transcription factor (TF) and a gene promoter, where the inducer can activate the TF and an active TF can bind to the gene promoter. We model the pathway using the chemical master equation so the counts of bound promoters over time is a stochastic signal. We consider the problem of using the continuous-time stochastic signal of the counts of bound promoters to infer whether the inducer signal is persistent or not. We use statistical detection theory to derive the solution to this detection problem, which is to compute the log-likelihood ratio of observing a persistent signal to a transient one. We then show, using time-scale separation and other assumptions, that this log-likelihood ratio can be approximately computed by using the continuous-time signals of the number of active TF molecules and the number of bound promoters when the input is persistent. Finally, we show that the coherent feedforward gene circuits can be used to approximately compute this log-likelihood ratio when the inducer signal is persistent.

INDEX TERMS Statistical signal processing, signal detection, molecular computing, analog computation, biochemical circuits.

I. INTRODUCTION

The fact that living organisms use biomolecular circuits for information processing has provided much inspiration for engineers to design new biomolecular circuits. Firstly, engineers have been inspired to use engineering theory to design synthetic circuits. E.g., control theory has been used to design circuits for homeostatic control [1], concentration regulation [2] and to counter the effect of loading [3]; signal processing theory has been used to design molecular circuits that can be used for communications [4]–[6] and filtering [7]. Secondly, one can view the information processing carried out by biomolecular circuits as analog computation [8], [9]. This view point has helped engineers to design molecular circuits that can perform logarithmic sensing [10], parity check [11] and integral control [12]. This paper considers the problem of persistence detection, i.e. the problem of deciding whether a particular chemical species has been present in

sufficient quantity for a long enough time. In the natural world, the bacteria Escherichia coli (E. coli) perform persistence detection [13]. In the synthetic world, cell-based therapy [14] can make use of persistence detection on biomarkers. In this paper, we will solve the persistence detection problem by using statistical detection theory [15] and show that a gene circuit (which is specific type of biochemical circuit) can be used to approximately compute the solution of this detection problem.

A gene circuit that can perform persistence detection in E. coli is the Coherent Type-1 Feedforward Loop with an AND logic at the output [16] or the C1-FFL for short. The C1-FFL is a network motif and is a frequently found circuit in both E. coli and Saccharomyces cerevisiae (yeast) [16], [17]. This means that the C1-FFL carries out important functions in cells. The authors in [16] showed that the C1-FFL can act as a persistence detector. They did this by modelling the gene expression in the C1-FFL by using ordinary differential equations (ODEs) and show that a persistent (resp. transient) input to the C1-FFL will result in a high (zero) output.

The paper [16] took a deterministic approach to understand persistence detection. Given that the biochemical environment is stochastic, it is therefore necessary to understand how cells can infer information on the environment from a stochastic point of view [18]. This paper considers a reaction pathway consisting of three chemical species: an inducer, a transcription factor (TF) and a gene promoter. In this reaction pathway, the inducer can activate the TF and the activated TF can bind with the gene promoter. In order to model a stochastic biochemical environment, we model the reaction pathway using the chemical master equation [19]. We consider a detection problem whose aim is to infer whether the inducer signal is persistent or not by using the signal of the number of bound promoters over time. According to detection theory, the solution to this detection problem is to compute a log-likelihood ratio and we derive an ODE which describes the evolution of this log-likelihood ratio over time. In order to connect this ODE to the C1-FFL, we use time-scale separation and other assumptions to derive an intermediate approximation which is an ODE that can approximately compute the log-likelihood ratio for persistent signals. We then show that this intermediate approximation can be realised by using a C1-FFL. The key contribution of this paper is to show that it is possible to find the parameters of a C1-FFL so that its output is approximately equal to the log-likelihood ratio of statistical detection problem when the inducer signal is persistent. More specifically, the aim of this statistical detection problem is to detect whether the inducer signal is persistent in an inducer-TF-gene pathway. In addition, the methodology in this paper can be useful for designing synthetic molecular circuits for performing other signal processing tasks.

This paper makes advances compared to our previous work [20]. In comparison to [20], this paper makes two different assumptions: (i) This paper considers an inducer-TF-gene pathway but [20] considered only an inducer-TF pathway; (ii) This paper assumes that the inducer signal is stochastic while [20] assumed that the inducer signal is deterministic. These two different assumptions mean that new methodologies are needed to show that the C1-FFL can be used to approximately compute the log-likelihood ratio of a persistence detection problem. First, we need to show how the log-likelihood ratio can be computed exactly (Sec. III-B) in an inducer-TF-gene pathway. Second, we need to derive a new method to approximately compute the log-likelihood ratio (Sec. IV). In particular, this paper needs to approximate the solution to a Bayesian filtering problem [21] but this is not required in [20] as it considered a deterministic inducer signal. Third, we need to derive a method to show how the parameters of the approximate log-likelihood ratio computation can be mapped to the C1-FFL parameters (Sec. V). These three aspects are the new elements of this paper in comparison to [20]. Furthermore, in comparison with our earlier conference paper [22], this paper explains why a C1-FFL can be used to approximately compute the approximate log-likelihood ratio (Sec. V) and provide full derivation
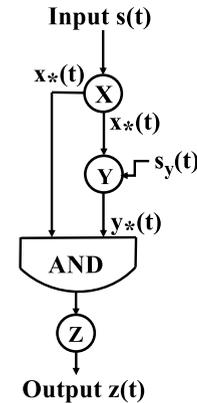

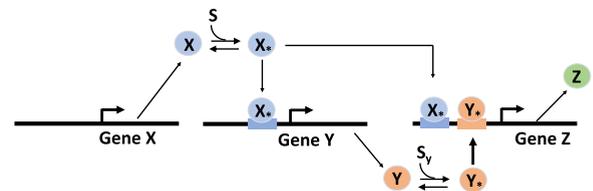**FIGURE 1.** Network representation of C1-FFL.

FIGURE 2. Representation of C1-FFL using inducers, transcription factors and genes.

on the computation of exact and approximate log-likelihood ratios (Appendices A and B).

The rest of this paper is organised as follows. Sec. II presents background information on the C1-FFL. We then define the detection problem and present its solution in Sec. III. After that in Sec. IV, we present a method to approximately compute the log-likelihood ratio and use this approximation in Sec. V to show that the C1-FFL can be used to approximately compute the log-likelihood ratio when it is positive. Finally, Sec. VI presents a discussion and concludes the paper.

II. BACKGROUND ON C1-FFL

The C1-FFL can be depicted as a network where each link is associated with a signal and each node transforms the input signal(s) into an output signal. Fig. 1 shows the network of the C1-FFL. The input signal is $s(t)$ and output signal is $z(t)$. Both $x_*(t)$ and $y_*(t)$ are intermediate signals, and $s_y(t)$ is an external signal.

The C1-FFL in Fig. 1 is an abstraction of the molecular interactions which are depicted in Fig. 2. In the figure, both S and S_y are inducers. Both X and Y are TFs, which are expressed by their corresponding gene. The inducer S (resp. S_y) turns the inactive form X (Y) into its active form X_* (Y_*). The activation of gene Z requires the binding of both X_* and Y_* to the promoter of Z, i.e. the AND gate in Fig. 1.

Note that there is a one-to-one correspondence between the chemical species in Fig. 2 with their corresponding time signals in Fig. 1, e.g. $x_*(t)$ is the concentration of X_* at time t and so on. In this paper, we will assume that the inducer S_y is always present and its concentration is always above the threshold needed to activate Y. Furthermore, we assume

the activation of Y by S_y is fast, this allows us to write $y_*(t) = y(t)$ and we will use $y(t)$ for $y_*(t)$ from now on. By using Hill function to model the gene expression, [13] presents an ODE model for the C1-FFL, as follows:

$$\frac{dx_*(t)}{dt} = k_+(M - x_*(t))s(t) - k_-x_*(t) \tag{1a}$$

$$\frac{dy(t)}{dt} = \underbrace{\frac{h_{xy}x_*(t)^{n_{xy}}}{K_{xy}^{n_{xy}} + x_*(t)^{n_{xy}}}}_{H_{xy}(x_*(t))} - d_y y(t) \tag{1b}$$

$$\frac{dz(t)}{dt} = \underbrace{\frac{h_{xz}x_*(t)^{n_{xz}}}{K_{xz}^{n_{xz}} + x_*(t)^{n_{xz}}}}_{H_{xz}(x_*(t))} \times \underbrace{\frac{h_{yz}y(t)^{n_{yz}}}{K_{yz}^{n_{yz}} + y(t)^{n_{yz}}}}_{H_{yz}(y(t))} - d_z z(t) \tag{1c}$$

where k_+, k_-, d_y and d_z are reaction rate constants; h_{xy}, n_{xy}, K_{xy}, h_{xz}, n_{xz}, K_{xz}, h_{yz}, n_{yz} and K_{yz} are coefficients for Hill functions $H_{xy}(x_*(t))$, $H_{xz}(x_*(t))$ and $H_{yz}(y(t))$. Lastly, $x(t) + x_*(t)$ is the constant M. The multiplication of H_{xz} and H_{yz} on the right-hand side (RHS) of (1c) implements the AND gate in Fig. 1. With suitably chosen parameter values, the C1-FFL in (1) acts as a persistence detector in the sense that if the input signal $s(t)$ is a persistent (resp. transient), then the output $z(t)$ has a high (low) value.

III. STATISTICAL DETECTION ON A REACTION PATHWAY

Our aim is to consider a statistical detection problem to determine whether the input signal is persistent or not. However, in this section, we will consider a more general detection problem because it can readily be solved and we will specialise it to persistence detection in Sec. IV. This section is divided into two parts. We define the detection problem in Sec. III-A and present its solution in Sec. III-B.

Convention: In this paper, we use upper case letters to denote a chemical species, e.g. S, X_* etc. For each chemical species, there are two corresponding continuous-time signals based on its concentration and molecular counts. E.g. for the chemical species X_*, we denote its concentration over time as $x_*(t)$ (note: lower case x) and its molecular counts over time is $X_*(t)$ (note: upper case X).

A. DETECTION PROBLEM

In order that we can connect the detection problem to the C1-FFL later on, we will define the detection problem using a reaction pathway which is a subset of the C1-FFL species and reactions in Fig. 2. We have depicted the reaction pathway used in the detection problem in Fig. 3. The reaction pathway consists of five chemical species: S, inactive X and its corresponding active form X_*, as well as inactive $\widetilde{Z}$ and the complex $\widetilde{Z}_{X_*}$ which is formed by the binding of X_* to $\widetilde{Z}$. These five species take part in the following four chemical reactions:

$$S + X \xrightarrow{k_+} S + X_* \tag{2a}$$
$$X_* \xrightarrow{k_-} X \tag{2b}$$
$$X_* + \widetilde{Z} \xrightarrow{g_+} \widetilde{Z}_{X_*} \tag{2c}$$
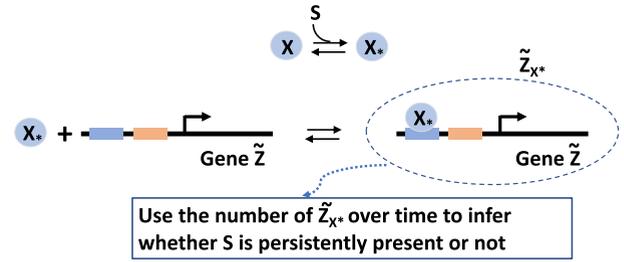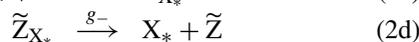$$\widetilde{Z}_{X_*} \xrightarrow{g_-} X_* + \widetilde{Z} \tag{2d}$$

FIGURE 3. The reaction pathway for the detection problem.

where k_+, k_-, g_+ and g_- are reaction propensity constants. For the time being, we will make the simplifying assumption that the volume scaling needed to convert between propensity and reaction rate constants is 1. This simplification allows us to equate propensity constants with reaction rate constants. We will explain how non-unit volume can be dealt with in Remark 3. With this assumption, note that k_+ and k_- in (2a) and (2b) are equal to those in (1a).

In terms of molecular biology, S is an inducer and X is a TF. In Reaction (2a), the species S activates X to produce X_*. Reaction (2b) is a deactivation reaction. The reactions (2a) and (2b) are depicted in both Figs. 2 and 3.

The species $\widetilde{Z}$ is a gene. In fact, $\widetilde{Z}$ in Fig. 3 is the same as Z in Fig. 2. Note that Fig. 2 follows the standard convention in molecular biology where a gene and the protein that it expresses are given the same symbol Z. However, in this paper, we need different symbols for the gene and the protein that the gene expresses so that we can clearly distinguish their corresponding time signals. Therefore, we have chosen to use $\widetilde{Z}$ to denote the gene and use Z to denote the protein expressed by $\widetilde{Z}$. In Reaction (2c), an active X_* binds with the promoter of $\widetilde{Z}$ to produce the complex $\widetilde{Z}_{X_*}$. Lastly, Reaction (2d) is an unbinding reaction.

Let $S(t)$, $X(t)$, $X_*(t)$, $\widetilde{Z}(t)$ and $\widetilde{Z}_{X_*}(t)$ denote, respectively, the number of S, X, X_*, $\widetilde{Z}$ and $\widetilde{Z}_{X_*}$ molecules at time t. Note these signals are piecewise constant because they are molecular counts. We assume that $X(t) + X_*(t)$ (resp. $\widetilde{Z}(t) + \widetilde{Z}_{X_*}(t)$) is a constant for all t and we denote this constant by M (N).

We will refer to $S(t)$ as the input signal. The goal of the detection is to use the signal $\widetilde{Z}_{X_*}(t)$ to determine whether the input is persistent or not. We assume that the signal $S(t)$ is generated by some species and chemical reactions upstream of (2). We will model these upstream reactions and (2) using the chemical master equation [19] which is a specific type of continuous-time Markov chain. We further assume that the upstream species do not react with X, X_*, $\widetilde{Z}$ and $\widetilde{Z}_{X_*}$. Intuitively, this means we can predict the behaviour of X, X_*, $\widetilde{Z}$ and $\widetilde{Z}_{X_*}$ from that of S.

We have now defined the reaction pathway and its model. The next task is to specify the measured data and the hypotheses for the detection problem. The measured datum at time t is $\widetilde{Z}_{X_*}(t)$. However, in the formulation of the detection problem, we will assume that at time t, the data available to the detection problem are $\widetilde{Z}_{X_*}(\tau)$ for all $\tau \in [0, t]$; in other words, the data are continuous in time and are the history of
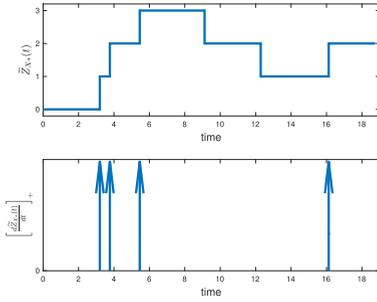
FIGURE 4. Illustrating $\widetilde{Z}_{X_*}(t)$ and $\left[\frac{d\widetilde{Z}_{X_*}(t)}{dt}\right]_+$.

the counts of $\widetilde{Z}_{X_*}$ up to time t inclusively. We will use $\widetilde{\mathcal{Z}}_{X_*}(t)$ to denote the continuous-time history of $\widetilde{Z}_{X_*}(t)$ up to time t inclusively.

We now specify the hypotheses $\mathcal{H}_i$ ($i = 0, 1$) for the detection problem. Later on, we will identify $\mathcal{H}_0$ and $\mathcal{H}_1$ with, respectively, transient and persistent signals. However, at this stage, we want to solve the detection problem in a general way. We assume that $\mathcal{H}_0$ and $\mathcal{H}_1$ are two distinct subsets of the set of all possible $S(t)$. Intuitively, the aim of the detection problem is to decide which signal class $\mathcal{H}_0$ or $\mathcal{H}_1$ is more likely to have produced the observed history.

We remark that in the definition of the detection problem, the input signal $S(t)$ is not directly observable. Since S reacts with the molecules in the reaction pathway in Fig. 3, the downstream signal $\widetilde{Z}_{X_*}(t)$ contains information on $S(t)$. The aim of the detection problem is to infer the information on $S(t)$ from this downstream signal. Given that we model the chemical system with the chemical master equation, both signals $S(t)$ and $\widetilde{Z}_{X_*}(t)$ are noisy.

B. SOLUTION TO THE DETECTION PROBLEM

The aim of the detection problem is to decide which hypothesis $\mathcal{H}_i$ ($i = 0, 1$) is more likely to have generated the observed history $\widetilde{\mathcal{Z}}_{X_*}(t)$. We assume that the two hypotheses are equally likely, therefore the log-ratio of posteriori probabilities $\frac{P[\mathcal{H}_1|\widetilde{\mathcal{Z}}_{X_*}(t)]}{P[\mathcal{H}_0|\widetilde{\mathcal{Z}}_{X_*}(t)]}$ is equal to the log-likelihood ratio $L(t)$:

$$L(t) = \log\left(\frac{P[\widetilde{\mathcal{Z}}_{X_*}(t)|\mathcal{H}_1]}{P[\widetilde{\mathcal{Z}}_{X_*}(t)|\mathcal{H}_0]}\right) \tag{3}$$

where $P[\widetilde{\mathcal{Z}}_{X_*}(t)|\mathcal{H}_i]$ is the conditional probability of observing the history $\widetilde{\mathcal{Z}}_{X_*}(t)$ given hypothesis $\mathcal{H}_i$.

In Appendix A, we show that the time evolution of $L(t)$ is given by the following ODE:

$$\frac{dL(t)}{dt} = \left[\frac{d\widetilde{Z}_{X_*}(t)}{dt}\right]_+ \log\left(\frac{J_1(t_-)}{J_0(t_-)}\right)$$
$$- g_+(N - \widetilde{Z}_{X_*}(t))(J_1(t) - J_0(t)) \tag{4}$$

$$J_i(t) = E[X_*(t)|\widetilde{\mathcal{Z}}_{X_*}(t), \mathcal{H}_i] \tag{5}$$

where $[w]_+ = \max(w, 0)$, $E[\ ]$ denotes the expectation and $E[X_*(t)|\widetilde{\mathcal{Z}}_{X_*}(t), \mathcal{H}_i]$ is the conditional expectation of $X_*(t)$ given the history and $\mathcal{H}_i$. Note that in deriving (4), we assume that the hypotheses have been properly chosen so

that $J_i(t) > 0$ (for $t > 0$ and $i = 0, 1$) which in turn implies that $\log\left(\frac{J_1(t)}{J_0(t)}\right)$ is well defined.

Since $\widetilde{Z}_{X_*}(t)$ is a piecewise constant function counting the number of $\widetilde{Z}_{X_*}$ molecules, its derivative is a sequence of Dirac deltas at the time instants that $\widetilde{Z}_{X_*}$ forms or unbinds. Note that the Dirac deltas corresponding to the formation of $\widetilde{Z}_{X_*}$ carries a positive sign and the $[\ ]_+$ operator keeps only these. Fig. 4 shows an example $\widetilde{Z}_{X_*}(t)$ and its corresponding $\left[\frac{d\widetilde{Z}_{X_*}(t)}{dt}\right]_+$.

At the time instant t that $\widetilde{Z}_{X_*}$ is formed or unbind, the expectation $J_i(t)$ also has a jump in value at time t. The term $J_i(t_-)$ in (4) refers to the value of $J_i(t)$ just before the jump. Lastly, we assume that the two hypotheses are a priori equally likely, so $L(0) = 0$.

We next present a numerical example to illustrate the properties of (4) and to explain what information is important for persistence detection. This example will also provide some intuition on how we will approximately compute the log-likelihood in Sec. IV.

1) NUMERICAL EXAMPLE

The aim of this example is to show that we can use the log-likelihood ratio computed by (4) to distinguish a persistent signal from a transient one. In order to conduct the numerical study, we will assume that the inducer S is produced and degraded by the following chemical reactions:

$$S_{pre} \xrightarrow{f_+} S_{pre} + S \tag{6a}$$

$$S \xrightarrow{f_-} \phi \tag{6b}$$

where S_{pre} is a precursor that produces S, and f_+ and f_- are reaction rate constants. The reason why we choose to use these reactions is that they allow us to use the time profile $S_{pre}(t)$, which is the number of precursor molecules a time t, to control the amplitude and duration of $S(t)$. The chemical master equation is used to model the reactions (6) and (2).

We assume that both hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ include the knowledge of the reactions (6) and (2) because this knowledge is required for computing $E[X_*(t)|\widetilde{\mathcal{Z}}_{X_*}(t), \mathcal{H}_i]$ in (5). In order to define the hypotheses, we first define reference signals $R_0(t)$ and $R_1(t)$ for $S_{pre}(t)$. Both reference signals are deterministic ON-OFF pulses and their time profile has the form:

$$R_i(t) = \begin{cases} S_{pre}^{ON,ref} & \text{for } 0 \le t < d_i \\ S_{pre}^{OFF,ref} & \text{for } t \ge d_i \end{cases} \tag{7}$$

where d_i is the ON-duration for $R_i(t)$; and, $S_{pre}^{ON,ref}$ and $S_{pre}^{OFF,ref}$, are, respectively, the ON and OFF amplitudes. We assume that $S_{pre}^{ON,ref} \gg S_{pre}^{OFF,ref}$ so that one may consider $S_{pre}^{OFF,ref}$ as a reference value for basal concentration. Furthermore, we assume that the duration of $R_1(t)$ is longer than $R_0(t)$, i.e. $d_1 > d_0$. We define $\mathcal{H}_0$ (resp. $\mathcal{H}_1$) as the set of all $S(t)$ that are generated by using $R_0(t)$ ($R_1(t)$) as the time profile for $S_{pre}(t)$. Since $d_1 > d_0$, we will identify
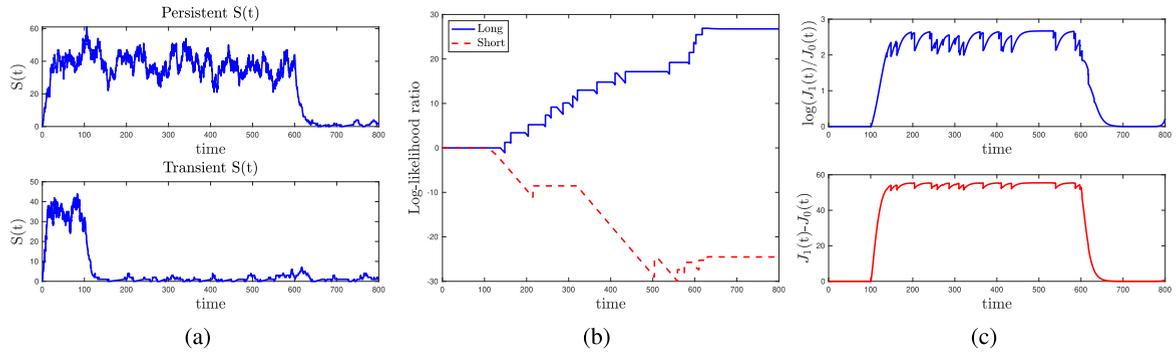
FIGURE 5. (a) Sample persistent $S(t)$ (top plot) and transient $S(t)$ (bottom plot). (b) The log-likelihood ratio for a long signal (or persistent signal) and a short signal. (c) The top and bottom plots show $\log\left(\frac{J_1(t)}{J_0(t)}\right)$ and $J_1(t) - J_0(t)$ in (4).

$\mathcal{H}_0$ (resp. $\mathcal{H}_1$) as, respectively, the set of transient (persistent) signals.

For this example, the kinetic parameters for the reaction pathway (2) are: $k_+ = 0.02$, $k_- = 0.5$, $g_+ = 0.002$ and $g_- = 0.05$. The total number of TFs M is 100 and the number of genes N is 1. Furthermore, for reactions (6), $f_+ = 0.37$ and $f_- = 0.1$. For the reference signals, we choose $S_{pre}^{ON,ref}$ and $S_{pre}^{OFF,ref}$ to be, respectively, 100 and 3 molecules; and $d_1 = 600$ and $d_0 = 100$.

We will now use the above set-up to detect a persistent signal. For simplicity, we assume that the actual $S_{pre}(t)$ is the reference signal $R_1(t)$ which results in a persistent input $S(t)$; see the top plot of Fig. 5a for a sample of persistent $S(t)$. We first use the chosen $S_{pre}(t)$ and the Stochastic Simulation Algorithm (SSA) [23] to produce a realisation of $\widetilde{Z}_{X_*}(t)$. The simulation assumes that there are zero S, X_* and $\widetilde{Z}_{X_*}$ molecules initially. We then use optimal Bayesian filtering [21] to numerically compute $E[X_*(t)|\widetilde{Z}_{X_*}(t), \mathcal{H}_i]$ for both $i = 0, 1$. The numerical solution to this optimal Bayesian filtering problem requires us to solve a master equation given the hypothesis. In order to solve this master equation exactly, we have purposely chosen the problem parameter values so that the number of S and X molecules are limited. After solving the Bayesian filtering problem, we numerically integrate (4) to obtain the log-likelihood ratio $L(t)$, which is plotted as the solid blue line in Fig. 5b. We see that the log-likelihood ratio is zero for $t \leq 100$, ramps up in the time interval $[100, 600]$ and plateaus after $t \geq 600$. The log-likelihood ratio reaches a positive value at the end, which means correct detection because it says the $\widetilde{Z}_{X_*}(t)$ signal is more likely to have been generated by a persistent signal.

Next, we use a transient input. We assume that the actual $S_{pre}(t)$ is the reference signal $R_0(t)$; see the bottom plot of Fig. 5a for a sample of transient $S(t)$. We perform the same steps as before, namely SSA simulation, optimal Bayesian filtering and numerical integration to obtain the log-likelihood ratio for this transient input. The resulting log-likelihood ratio $L(t)$ is plotted as red dashed lines in Fig. 5b. This $L(t)$ becomes negative which means correct detection.

Fig. 5c shows the weighting factors $\log\left(\frac{J_1(t)}{J_0(t)}\right)$ and $J_1(t) - J_0(t)$ in (4) for the case when the $S(t)$ is persistent.

(The curves are similar when $S(t)$ is transient.) It shows that these two weighting factors are mostly positive in the time interval $[100,600]$ but are zero outside. This means the contribution to the log-likelihood ratio comes from the signal within $[100, 600]$. This can also be seen from Fig. 5b where the log-likelihood ratio does not change outside of $[100,600]$ but increases (resp. decreases) for persistent (transient) signal within $[100,600]$. This makes intuitive sense because the persistent input is different from the transient input within this time interval, so the signal in this time interval is useful for discriminating persistent signals from transient ones.

If a persistent signal can give a large positive log-likelihood ratio, then the probability of correctly detecting the persistent signal is higher. For this example, the positive contribution to log-likelihood ratio comes from the first term on the RHS of (4) because the weighting factors are non-negative, see Fig. 5c. In fact, each time when a X_* binds to a $\widetilde{Z}$ in the time interval $[100,600]$, it creates a positive jump in the magnitude of the log-likelihood ratio, which can be seen in Fig. 5b. This means that a persistent signal becomes easier to detect if X_* binds to $\widetilde{Z}$ many times when the signal is ON. This can be achieved if the ON duration of the persistent signal has a longer time-scale compared to those of the binding and unbinding reactions of $\widetilde{Z}_{X_*}$ (i.e. reactions (2c) and (2d)) so that these reactions occur many times when the input is ON.

Although this example shows that the solution (4) can distinguish a persistent signal from a transient one by computing the log-likelihood ratio, the solution is hard to implement in a biochemical environment because Bayesian filtering requires extensive computation and a model of the pathway. In the next section, we will explore how we can compute the log-likelihood ratio approximately without using Bayesian filtering.

IV. COMPUTING THE LOG-LIKELIHOOD RATIO APPROXIMATELY

Our ultimate goal is to show that the computation of the log-likelihood ratio $L(t)$ in (4) can be carried out by a C1-FFL in (1), i.e. there exists a set of parameters for the C1-FFL such that $z(t)$ in (1) is approximately equal to $L(t)$ in (4). It is not obvious from the expression of (4) that this can be done. The aim of this section is to derive an ODE, which will be referred

to as the **intermediate approximation**, such that the output of this ODE is approximately equal to $L(t)$ when the input is persistent. We will then use this intermediate approximation in Sec. V to relate to the C1-FFL.

A. ASSUMPTIONS

The detection problem and its solution in Sec. III are general in the sense that they apply to any reaction pathways of the form (2) and hypotheses $\mathcal{H}_i$. In order to connect the detection problem to the C1-FFL model in (1), we will need to make specific assumptions to derive the intermediate approximation. We will specify these assumptions in this subsection.

We make the following two assumptions on the reaction pathway (2):

- The time-scale of the inducer-TF reactions (2a) and (2b) is faster than that of the TF-gene promoter reactions (2c) and (2d). This translates to small g_+ and g_-.
- The number of TF molecules M is much higher than the number of genes N.

We believe these are realistic assumptions. First, according to [24, Table 2.2], for E. coli, the time-scale for equilibrium binding of small molecules to protein is of the order of 1 ms and the time-scale for TF binding to gene promoter is of the order of 1s. Second, the copy number of most genes is either 1 or 2.

In order to make the derivation of the intermediate approximation analytically tractable, we assume that hypotheses $\mathcal{H}_i$ are defined in the same way as in the numerical example in Sec. III-B1. This is so that we can approximate some signal, e.g. the weighting functions in Fig. 5c, by a piecewise constant function.

For analytical tractability, we further assume that the actual $S_{pre}(t)$ is a deterministic ON-OFF pulse of the form:

$$S_{pre}(t) = \begin{cases} \dfrac{f_-}{f_+}\alpha & \text{for } 0 \leq t < d_i \\ S_{pre}^{OFF,ref} & \text{for } t \geq d_i \end{cases} \tag{8}$$

where d is the pulse duration and the ON-amplitude $\frac{f_-}{f_+}\alpha$ will result in a steady state $\mathrm{E}[S(t)]$ of amplitude α when the pulse is ON.

We assume that the input signal $S(t)$ is produced by using $S_{pre}(t)$ with the reactions (6), and the reaction rate constants f_+ and f_- have been chosen such that the signal $S(t)$ is slowly time varying compared to the time-scale of the inducer-TF reactions (2a) and (2b). This makes intuitive sense because a persistent or long $S(t)$ needs to be "measured" by faster reactions.

We assume that if a persistent input is applied to the reaction pathway (2), the pathway is almost at steady state by d_0 where d_0 is the duration of the reference signal $R_0(t)$ in (7). This requirement can be met if the duration d_0 is long enough. It may be instructive to recall from the discussion in the numerical example in Sec. III-B1 that there is a time interval which is informative for persistence detection. For the assumptions in this section, the informative time interval can

be shown to be $[d_0, \min(d, d_1)]$. Intuitively, this assumption allow us to use the steady state statistics in the time interval $[d_0, \min(d, d_1)]$ to replace $\left[\frac{d\widetilde{Z}_{X_*}(t)}{dt}\right]_+$ and $g_+(N - \widetilde{Z}_{X_*}(t))$ in (4) by, respectively, $g_-\widetilde{Z}_{X_*}(t)$ and $g_-\frac{\widetilde{Z}_{X_*}(t)}{X_*(t)}$.

B. INTERMEDIATE APPROXIMATION

An ideal persistence detector has the properties that a transient input will result in a zero output and a persistent input will result in a positive output [24]. The C1-FFL, when acting as a persistence detector, can be considered to be an approximation of this ideal behaviour [24]. However, it is not possible to map the log-likelihood ratio detector in (4) to the C1-FFL because the log-likelihood ratio becomes negative for transient signals but the concentration in the C1-FFL can only be non-negative. We will use the intermediate approximation as a bridge to connect (4) to the C1-FFL. The intermediate approximation has two key properties. First, if the input is transient, then the output of the intermediate approximation is zero. Second, if the input is persistent, then the output of the intermediate approximation is approximately equal to the log-likelihood ratio given by (4). Another purpose of the intermediate approximation is to replace the complex computation in (4), e.g. derivative and optimal Bayesian filtering, by simpler computation that can be implemented by chemical reactions.

The derivation of the intermediate approximation is given in Appendix B, using the assumptions stated in Sec. IV-A. The derivation shows that the time evolution of the intermediate approximation $\hat{L}(t)$ is given by the following ODE:

$$\frac{d\hat{L}(t)}{dt} = \widetilde{Z}_{X_*}(t)\, g_-\, \pi(t)\, [\phi(X_*(t))]_+ \tag{9}$$

where

$$\phi(X_*(t)) = \log\left(\frac{X_1}{X_0}\right) - \frac{X_1 - X_0}{X_*(t)}, \tag{10}$$

$$X_i = \frac{Mk_+a_i}{k_+a_i + k_-} \text{ for } i = 0, 1 \tag{11}$$

$$a_0 = \frac{f_+}{f_-}S_{pre}^{OFF,ref} \tag{12}$$

$$a_1 = \frac{f_+}{f_-}S_{pre}^{ON,ref} \tag{13}$$

$$\pi(t) = \begin{cases} 1 & \text{for } d_0 \leq t < d_1 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

$$\hat{L}(0) = 0 \tag{15}$$

Furthermore, it can be shown that time evolution of $\mathrm{E}[\hat{L}(t)]$ obeys the following ODE:

$$\frac{d\mathrm{E}[\hat{L}(t)]}{dt} = \mathrm{E}[\widetilde{Z}_{X_*}(t)]\, g_-\, \pi(t)\, [\phi(\mathrm{E}[X_*(t)])]_+ \tag{16}$$

The behaviour of the intermediate approximation $\hat{L}(t)$ depends on two parameters of the input signal $S(t)$: its mean ON-amplitude α and duration d (see (8)). Three important properties for $\hat{L}(t)$ are:

1) If $d < d_0$, then for all t, we have $\hat{L}(t)$ and $E[\hat{L}(t)]$ are zero or small. This is due to $\pi(t)$, which is zero outside of $[d_0, d_1)$, and the fact that $X_*(t)$ is likely to be small for $t \geq d_0$.

2) If the amplitude α is lower than a threshold, then for all t, we have $\hat{L}(t)$ is zero or small and $E[\hat{L}(t)]$ is zero. We will explain this for $E[\hat{L}(t)]$. Since $E[X_*(t)]$ is an increasing function of α, this means a small α will give a small $E[X_*(t)]$. If $E[X_*(t)]$ is less than $\frac{X_1 - X_0}{\log\left(\frac{X_1}{X_0}\right)}$ for all t, then $[\phi(E[X_*(t)])]_+$ on the RHS of (16) is zero and this implies $E[\hat{L}(t)]$ is zero for all t. The explanation for $\hat{L}(t)$ is similar.

3) If d is longer than d_0 and α is sufficiently large, then for $0 \leq t < \min\{d, d_1\}$ we have $\hat{L}(t) \approx L(t)$ where $L(t)$ is given in (4).

The first two properties are concerned with transient signals, which are those input signals whose duration is no longer than d_0 or whose amplitude α is small. The intermediate approximation says that transient signals give a small $\hat{L}(t)$. On the other hand, persistence signals have a duration longer than d_0 and have a sufficiently large amplitude α. For persistent signals, the intermediate approximation $\hat{L}(t)$ is approximately equal to the log-likelihood ratio $L(t)$ in the time interval $0 \leq t < \min\{d, d_1\}$. Note in particular that this approximation holds for a range of d and α values. From now on, we will choose d_1 to be ∞ so that $\hat{L}(t) \approx L(t)$ holds for $0 \leq t < d$, i.e. when the persistent signal is ON. Note that an infinite d_1 means $\pi(t)$ in (14) becomes a step function which changes from 0 to 1 at time d_0.

1) NUMERICAL EXAMPLES

The numerical examples in this section use the following kinetic parameters for the reaction pathway (2): $k_+ = 0.02$, $k_- = 0.5$, $g_+ = 0.002$ and $g_- = 0.0125$. These parameters have been chosen such that the time-scale of the inducer-TF reactions are faster than those of the TF-gene promoter. The number of TF M is 100 and the number of genes N is 1, which means $M \gg N$. The values of f_+, f_-, $S_{pre}^{ON,ref}$ and $S_{pre}^{OFF,ref}$ are the same as those in Sec. III-B1. The duration parameters for the reference signals are: $d_0 = 100$, $d_1 = \infty$. The signal duration $d = 3000$. All the above parameters are fixed. We will vary the values of amplitude α.

For the first numerical experiment, we use $\alpha = 37$, which means $S_{pre}(t)$ has an amplitude of $S_{pre}^{ON,ref}$ when it is ON. This excitation means there is a mean probability of 0.6 that X is active. We use SSA simulation to generate 100 realisations of $X_*(t)$ and $\tilde{Z}_{X_*}(t)$, and use them to compute the true log-likelihood ratio $L(t)$ (which requires only $\tilde{Z}_{X_*}(t)$) and the intermediate approximation $\hat{L}(t)$ (which requires both $X_*(t)$ and $\tilde{Z}_{X_*}(t)$). Fig. 6a compares one realisation of $L(t)$ (blue line) and $\hat{L}(t)$ (red line) for a persistent input. It can be seen that the intermediate approximation is fairly accurate.

We then use all 100 realisations to compute the mean of $L(t)$, the mean of $\hat{L}(t)$, and the root-mean-square (RMS) error of $L(t) - \hat{L}(t)$. The results are plotted in Fig. 6b.
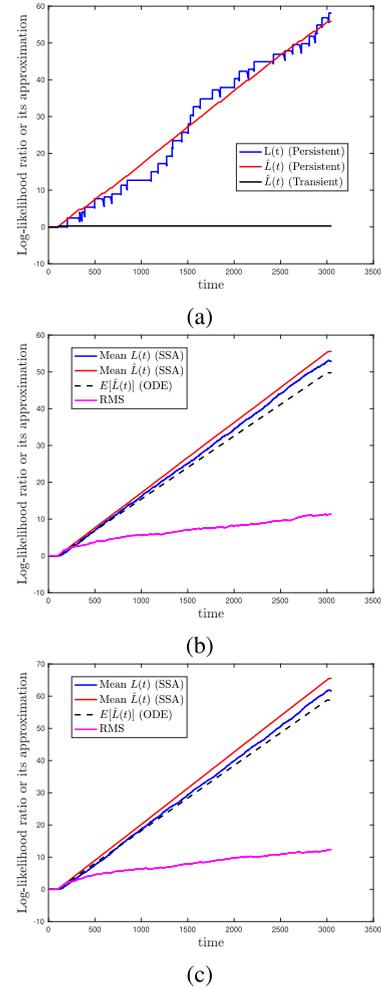
(a)

(b)

(c)

FIGURE 6. Numerical results for Sec. IV-B1. (a) Comparing one realisation of $L(t)$ and $\hat{L}(t)$ for a persistent signal. Also shows a realisation of $\hat{L}(t)$ for a transient signal. (b) Comparing the mean of $L(t)$ from SSA, mean of $\hat{L}(t)$ from SSA, mean of $\hat{L}(t)$ by (16). Also shows RMS error between $L(t)$ and $\hat{L}(t)$ from SSA. (c) Same type of comparison as (b) but for different values of α and d.

This shows that $\hat{L}(t)$ is a good approximation of $L(t)$ for many realisations. Next, we check the accuracy of using (16) to compute $E[\hat{L}(t)]$. The black dashed lines in Fig. 6b plot $E[\hat{L}(t)]$ computed from (16), which is almost the same as the mean obtained via SSA. This shows that (16) is an accurate method to compute $E[\hat{L}(t)]$.

The discussion so far focused on persistent signals. The black line in Fig. 6a shows the intermediate approximation for a transient input with $d = 100$. It can be seen that the output is small. This agrees with our earlier prediction.

For a given set of hypotheses, the intermediate approximation holds for a range of α and d. We now change α to 99.9 (which means $S_{pre}(t)$ has an amplitude of $1.33 S_{pre}^{ON,ref}$ when it is ON). By using 100 rounds of SSA simulations, we compute the means of both $L(t)$ and $\hat{L}(t)$, as well as the RMS error between them. The results are in Fig. 6c and they show that the intermediate approximation is accurate for different values of α. Fig. 6c also shows $E[\hat{L}(t)]$ computed from (16) and we can see that the accuracy is good.
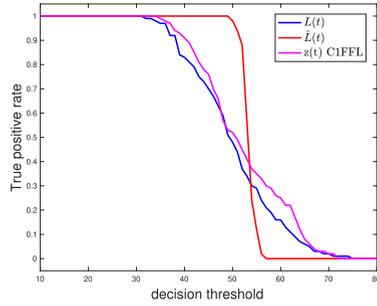
FIGURE 7. The impact of the decision threshold on the TP rates.

FIGURE 8. Relating the computation of the mean approximate log-likelihood ratio $E[\hat{L}(t)]$ in (16) to the C1-FFL.

Remark 1: The reader may wonder why we do not define the hypotheses of the detection problem as: $\mathcal{H}_0$ (resp. $\mathcal{H}_1$) means the duration of the input signal is shorter (longer) than a given threshold. The reason is that these are composite hypotheses and the solution to the resulting detection problem is much harder, see [20, Remark 5.1] for a more in-depth discussion. We also want to point out that, even with our simpler formulation, the resulting detector gives a small output for any signal whose duration is shorter than d_0.

2) TRUE POSITIVE (TP) AND FALSE POSITIVE (FP) RATES
In this section, we want to study the performance of using the approximate log-likelihood ratio $\hat{L}(t)$ for decision making. Since our interest is to detect persistent signals, we can do that by testing whether the log-likelihood ratio is greater than or equal to a positive threshold ρ. We use a numerical example to illustrate the impact of the positive decision threshold ρ on the TP and FP rates. For the TP rate, we assume the input is a long pulse of duration $d = 3000$ and generate 100 realisations of $X_*(t)$ and $\widetilde{Z}_{X_*}(t)$. We use the data to compute 100 values of $L(d)$ and $\hat{L}(d)$ where $L(d)$ and $\hat{L}(d)$ are, respectively, the log-likelihood ratio (4) and approximate log-likelihood ratio (9) at time $t = d$. We then compare these values against the threshold to obtain the TP rates. We vary the threshold between 10 and 80. The blue and red lines in Fig. 7 show the TP rates for $L(t)$ and $\hat{L}(t)$ respectively. It shows that there is a wider range of thresholds that can be used to obtain a high TP rate for $\hat{L}(t)$. This is because $\hat{L}(t)$ has a lower variance in comparison. We will discuss the magenta curve in Fig. 7 in Sec. V-B.

For the FP rate, we generate 100 realisations using a short pulse of duration $d = 100(= d_0)$ and use the data to compute $L(d)$ and $\hat{L}(d)$. For all the 100 realisations, the $L(d)$'s are negative and $\hat{L}(t) \approx 0$. Therefore for all the thresholds in $10 - 80$, the FP rates are zero for both $\hat{L}(t)$ and $L(t)$. This shows that we are able to find thresholds that give a large TF while keeping FP low.

The above numerical experiment on computing the FP rate also shows that it is not a problem to "round" the negative log-likelihood ratio in $L(d)$ to a near zero approximation $\hat{L}(d)$. This is because, if the test is to check whether the log-likelihood ratio is above a sufficiently positive threshold, then a negative $L(d)$ or an almost zero $\hat{L}(d)$ will lead to the same decision.
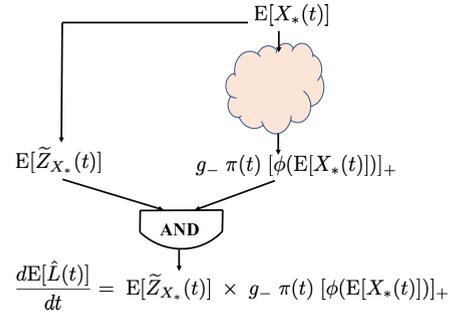
V. USING THE C1-FFL TO APPROXIMATELY COMPUTE LOG-LIKELIHOOD RATIO
We have shown in the previous section that the mean of the intermediate approximation $E[\hat{L}(t)]$ is an accurate approximation of the mean log-likelihood ratio $E[L(t)]$ when the input is persistent. The aim of this section is to show that we can use the C1-FFL in (1) to approximately compute $E[\hat{L}(t)]$ in (16).

A. RELATING $E[\hat{L}(t)]$ TO THE C1-FFL
For the time being, we will assume d_z in (1c) is zero and show that it is possible to find C1-FFL parameters such that $z(t)$ in (1) is approximately equal to $E[\hat{L}(t)]$ in (16). We will explain in Remark 2 how a non-zero d_z can be handled.

We depict the calculations of (16) in Fig. 8. We split the computation on the RHS of (16) as the product of $E[\widetilde{Z}_{X_*}(t)]$ and $g_- \pi(t) [\phi(E[X_*(t)])]_+$ where the multiplication operation is depicted as an and gate in the figure. Note that the two-branch structure of the computation in Fig. 8 has a direct resemblance with that of the C1-FFL in Fig. 2. We first consider the computation of the two branches separately.

We first consider the computation of $E[\widetilde{Z}_{X_*}(t)]$ from $E[X_*(t)]$, which is the branch on the left in Fig. 8. By using the volume scaling assumption (Sec. III), we equate molecular count $\widetilde{Z}_{X_*}(t)$ with concentration $\widetilde{z}_{X_*}(t)$, and similarly, $X_*(t)$ with $x_*(t)$. We propose to compute $\widetilde{z}_{X_*}(t)$ from $x_*(t)$ by using:

$$\widetilde{z}_{X_*}(t) = \text{computed by } \frac{Ng_+ x_*(t)}{g_+ x_*(t) + g_-} \qquad (17)$$

Note that the above approximation is obtained from assuming that the TF-gene reactions (2c) and (2d) are close to steady state. By using the parameters in the numerical example in Section IV-B1, we have plotted the two sides of (17) in Fig. 9a when the input is persistent with a duration d of 200. Note that there are two transients of $\widetilde{z}_{X_*}(t)$ in the figure, one after time 0 when the input turns ON and the other at time d when the input turns OFF. We can see that, other than these two transients, the two sides of (17) are almost equal. We will show later on these two transients have little effect on the accuracy of the overall computation. Note that the RHS of (17) has the form of a Hill function and we can identify it with $H_{xz}(x_*(t))$ in (1c).

We next consider the computation of $g_- \pi(t)[\phi(E[X_*(t)])]_+$ from $E[X_*(t)]$, which is depicted by a cloud in Fig. 8. We first argue that, for most of the admissible choices of d_0, there must be a time delay element in the cloud. This implies that there must be some chemical reactions in the cloud in order to create this time delay. To understand which d_0 is admissible, we recall that we assume in Section IV-A that the pathway (2) is almost at steady state by the time d_0. As an illustration of this assumption, consider Fig. 9a which shows the time profile of $\widetilde{z}_{X_*}(t)$, the vertical dashed line shows the time (which we will denote as t_{ss}) by which $\widetilde{z}_{X_*}(t)$ is sufficiently close to steady state, the assumption means an admissible d_0 must be greater than or equal to t_{ss}.

In our technical report [25, Appendix D], we show that there must be a delay in the cloud if d_0 is strictly greater than t_{ss}. Intuitively, a delay is needed because $\pi(t)$ is zero in $[0, d_0]$ but X_* almost reaches steady state before d_0. Given that there must be a delay element in the cloud in Fig. 8, we can achieve that by inserting a transcription node (e.g. Node Y in Fig. 2) in the cloud. With this insertion, we can identify the branch on the right in Fig. 8 with the indirect branch in the C1-FFL in Fig. 2.

The next step is to show that we can find Hill functions $H_{xy}()$ and $H_{yz}()$ in (1) which will enable us to compute $g_- \pi(t)[\phi(x_*(t))]_+$. Since the inducer-TF pathway is fast, we can approximate $x_*(t)$ by its steady state value $x_{*,ss}$ for $t < d$. We can use (1b) to show that

$$y(t) = \frac{H_{xy}(x_{*,ss})}{d_y}(1 - \exp(-d_y t)) \qquad (18)$$

Our aim is to achieve the approximation:

$$H_{yz}(y(t)) \approx g_- \pi(t)[\phi(x_{*,ss})]_+. \qquad (19)$$

We will turn this into two requirements. First, in order to imitate $\pi(t)$, we require $H_{yz}(y(t))$ to have a sharp transition from a low value to a high value around d_0. Note that this requirement implies that $y(t)$ should not have reached steady state by d_0 and can be achieved by suitable choice of d_y. Second, when $H_{yz}(y(t))$ reaches steady state, which happens some time after time d_0, its amplitude is determined by $g_-[\phi(x_*)]_+$. We can consider these two requirements separately because the first requirement is concerned with a time event around time d_0 while the second requirement is concerned with the steady state amplitude which is reached at a time later than d_0.

We first consider how $H_{yz}(y(t))$ can be used to realised a transition around time d_0. Let us recall that $H_{yz}(y(t))$ has the form $\frac{h_{yz} y(t)^{n_{yz}}}{K_{yz}^{n_{yz}} + y(t)^{n_{yz}}}$. Given that $y(t)$ is an increasing function of t, if we choose K_{yz} to be $y(d_0)$, then we have $y(t) < K_{yz}$ for $t < d_0$ and $y(t) > K_{yz}$ for $t > d_0$. Hence, if n_{yz} is sufficiently large, then $H_{yz}(y(t))$ will have a sharp transition around d_0.

The argument in the last paragraph works for a particular choice of $x_{*,ss}$, which corresponds to a particular choice of input signal amplitude. We will now explain how $H_{yz}(y(t))$ can be used to realised a transition at time d_0 for a range of input amplitudes. Since $H_{xy}(x_{*,ss})$ is a Hill function, then for sufficiently large $x_{*,ss}$, the value of $H_{xy}(x_{*,ss})$ does not change much because $H_{xy}(x_{*,ss})$ saturates. As a result, there is a range of $x_{*,ss}$ such that $y(t)$ does not change a lot. This means if we choose K_{yz} to be the $y(d_0)$ corresponding to a particular $x_{*,ss}$ which saturates $H_{xy}(x_{*,ss})$, we can a achieve sharp transition around time d_0 for a range $x_{*,ss}$.

We will now give an illustration on the first requirement. Fig. 9b plots the two sides of (19) for a particular input amplitude. We see that we can use H_{yz} to approximate $g_- \pi(t) [\phi(x_{*,ss})]_+$ well except near the falling edge of $g_- \pi(t) [\phi(x_{*,ss})]_+$. We can now consider the two branches in Fig. 8 together. We now compare the whole RHS of (16), which is $E[\widetilde{Z}_{X_*}(t)]$ times $g_- \pi(t) [\phi(E[X_*(t)])]_+$, and the RHS of (1c), which is $H_{xz}(x_*(t))$ times $H_{yz}(y(t))$. Fig. 9c compares these two expressions. An interesting observation is that the mismatch in some time intervals in Fig. 9a and 9b are cancelled out when the multiplication is made.

We now consider the second requirement which requires the steady state value of $H_{yz}(y(t))$ to be $g_-[\phi(x_*)]_+$. According to (18), the steady state amplitude of $y(t)$ is $\frac{H_{xy}(x_{*,ss})}{d_y}$, therefore the requirement is equivalent to finding H_{xy} and H_{yz} such that:

$$H_{yz}\left(\frac{H_{xy}(x_{*,ss})}{d_y}\right) \approx g_-[\phi(x_{*,ss})]_+ \qquad (20)$$

holds for a large range of $x_{*,ss}$. Let $\chi = \frac{1}{x_{*,ss}}$. By using the expressions of $H_{xy}()$ and $H_{yz}()$ in (1) as well as $\phi()$ in (10), we can rewrite (20) in terms of χ as:

$$\frac{h_{yz}}{\eta(1 + K_{xy}^{n_{xy}} \chi^{n_{xy}})^{n_{yz}} + 1} \approx g_- \left[\log\left(\frac{X_1}{X_0}\right) - (X_1 - X_0)\chi\right]_+ \qquad (21)$$

where $\eta = \left(\frac{K_{yz} d_y}{h_{xy}}\right)^{n_{yz}}$. If we consider the expression on the RHS of (21) as a function of χ, it has a linearly decreasing segment followed by a constant segment at zero. Recall that we have chosen n_{yz} to be large earlier, and if we furthermore choose n_{xy} is to be 1, then the LHS may be approximated by three segments: a constant for small χ, linearly decreasing for intermediate χ and zero for large χ. This shows that it is possible to find Hill functions H_{xy} and H_{yz} such that the second requirement holds. We will demonstrate this with a numerical example.

B. NUMERICAL EXAMPLE

This numerical example uses the first set of parameter values as in Sec. IV-B1. Our aim is to show that we can choose a set of C1-FFL parameters so that $z(t) \approx \hat{L}(t)$ for a range of input amplitude α. We do this by fixing $H_{xz}()$ as $\frac{Ng + x_*(t)}{g + x_*(t) + g_-}$, and by fitting the C1-FFL parameters in $H_{xy}()$ and $H_{yz}()$ in (1) by nonlinear optimisation. The data for fitting is obtained from varying α from 3.125 to 62.5. For each α, we compute $E[\hat{L}(t)]$ using (16) and use them to fit the parameters of the C1-FFL. The fitted values of the C1-FFL parameters are: $k_{xy} = 9.66$,
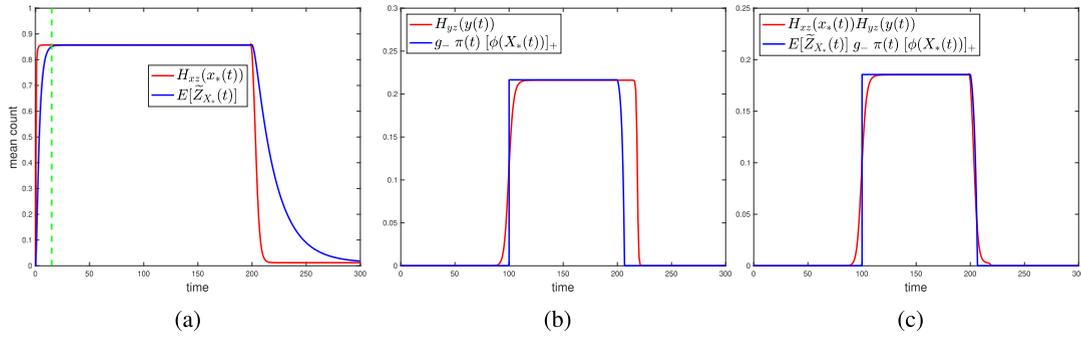
FIGURE 9. (a) Comparing $\widetilde{z}_{X_*}(t)$ and $\frac{Ng_+x_*(t)}{g_+x_*(t)+g_-}$. (b) Comparing $g_-\,\pi(t)\,[\phi(x_*)]_+$ and $H_{yz}(y(t))$. (c) Comparing $\widetilde{z}_{X_*}(t)g_-\,\pi(t)\,[\phi(x_*)]_+$ and $H_{XZ}(x_*(t))H_{yz}(y(t))$.
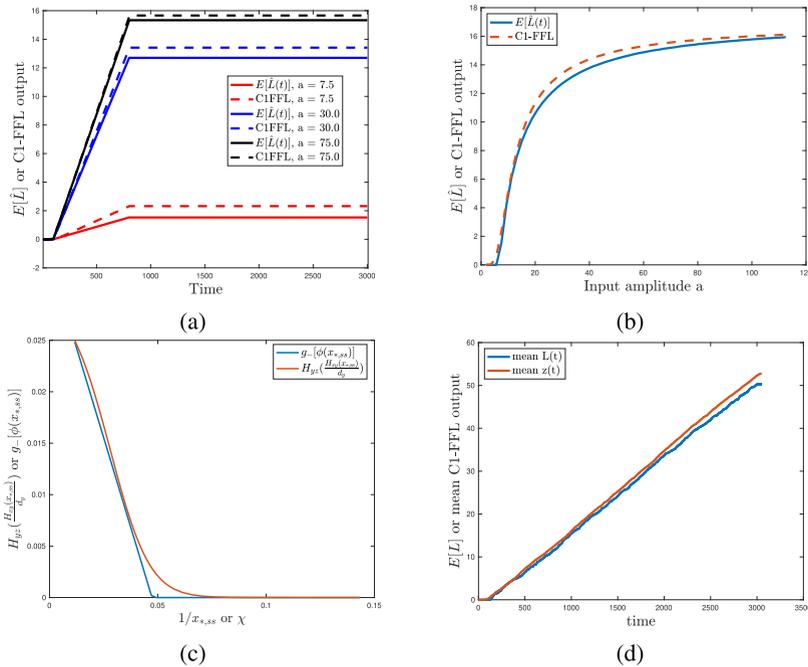

FIGURE 10. Numerical results for Sec. V. (a) Comparing the C1-FFL output against $E[\hat{L}(t)]$ for three different values of α. (b) Compare $z(t)$ (C1-FFL) and $\hat{L}(t)$ at time $t = 800$ for $a \in [12.5, 100]$. (c) Comparing $g_-\,[\phi(x_{*,ss})]_+$ and $H_{yz}(\frac{H_{xy}(x_{*,ss})}{d_y})$. Note the horizontal axis is $\chi = \frac{1}{x_{*,ss}}$ (d) Comparing the mean C1-FFL output from tau-leaping against mean $L(t)$ from SSA.

$n_{xy} = 1.65$, $K_{xy} = 4.36$, $d_y = 0.037$, $h_{yz} = 0.135$, $n_{yz} = 50.0$ and $K_{yz} = 255.4$.

Fig. 10a compares $E[L(t)]$ and the C1-FFL output $z(t)$ for three different values of α: 7.5, 30 and 75. It can be seen that they match very well. Next, we compare the value of $z(t)$ and $\hat{L}(t)$ at time $t = 800$ for $\alpha \in [1.9, 112]$. Fig. 10b shows that the match is good for a large range of α.

Although we use optimisation to obtain the parameters of the Hill functions $H_{xy}()$ and $H_{yz}()$. We find that their values are compatible to our intuitive argument in Sec. V-A. First, we say that K_{yz} should be chosen close to the value of $y(d_0)$ for a range of input amplitudes. For input amplitudes of 5, 15 and 480 (which correspond to a mean probability of 0.4, 0.6 and 0.8 that the promoter $\widetilde{Z}_{X_*}$ is bound), the values of $y(d_0)$

are respectively, 229.2, 247.3 and 252.9. These values are pretty close to $K_{yz} = 255.4$ obtained from using optimization. Second, the steady state value of $y(t)$ is around 259 for a range of input amplitudes and $y(d_0)$ has not yet reached the steady state. Third, we argue that $H_{yz}(\frac{H_{xy}(x_{*,ss})}{d_y})$ and $g_-\,[\phi(x_{*,ss})]_+$ should be approximately equal, i.e. (20) or (21). Fig. 10c plots the two sides of (21) and shows that they match well except near the turning point. Fourth, the optimised value of n_{xy} is 1.65 which is similar to the predicted value of 1.

We next study the behaviour of the fitted C1-FFL when $s(t)$ in (1) is the stochastic persistent input used in Sec. IV-B1. We apply the tau-leaping simulation algorithm in [26] to simulate the stochastic behaviour of the C1-FFL from its ODE model. We perform 100 rounds of simulation to obtain

the mean of the C1-FFL output $z(t)$. We compare the mean of $z(t)$ against the mean log-likelihood computed in Sec. IV-B1 in Fig. 10d and we can see that they match fairly well. Next, we use $z(t)$ to study the TP (true positive) and FP (false positive) rates in the same way as in Sec. IV-B2. Both persistent and transient inputs are used in this study, and 100 simulation runs are performed for each type of input. The magenta line in Fig. 7 (Page 129003) shows how the TP rates vary with the decision threshold. It can be seen that the TP curve for the C1-FFL is in between those for $L(t)$ and $\hat{L}(t)$. We find that we can again find a range of positive thresholds such that the TP rate is high and FP rate is zero.

Remark 2: We have shown that on the condition that d_z in (1c) is zero, then we can find parameters of the C1-FFL such that $z(t)$ in (1) is approximately equal to $E[\hat{L}(t)]$ in (16). Let us now assume d_z is non-zero and we add the term $-d_z E[\hat{L}(t)]$ to the RHS of (16), then we can use the same C1-FFL parameters as before to make $z(t)$ in (1) to be approximately equal the new $E[\hat{L}(t)]$.

Remark 3: The derivations in Secs.III-V assume that the reactions take place in a volume V of 1. If the volume is not 1, (16) remains the same. Note that both g_- and the term inside $[\ ]_-$ in (16) are independent of V. If we divide both sides of (16) by V, we have $\frac{d(E[\hat{L}(t)]/V)}{dt} = E[\tilde{z}_{X_*}(t)] \, g_- \, \pi(t) \, [\ldots]_-$. Thus, the argument in Sec. V-A is still valid if we interpret the output of the C1-FFL as log-likelihood ratio per volume.

Remark 4: Our interpretation of the C1-FFL as a statistical persistence detector shows an interesting signal processing architecture involving parallel processing. The C1-FFL has two arms. The short arm produces the signals $X_*(t)$ and $\tilde{Z}_{X_*}(t)$ which contain information on whether the inducer signal is persistent or not. The longer arm then makes use of the signals $X_*(t)$ and $\tilde{Z}_{X_*}(t)$ to approximately compute, in a parallel manner, the approximate log-likelihood ratio. It is also interesting to see that the gene $\tilde{Z}$, with its two promoter sites, has the dual roles of generating the signal to be processed as well as processing the signal.

C. DISCUSSION

Biochemical circuits are known to be impacted by both intrinsic and extrinsic noise [27]. An open problem is to design a C1-FFL, which is based on the mass-reaction kinetics, for persistence detection using the probabilistic framework studied in this paper. In particular, we envisage using this design problem to study the impact intrinsic noise (which comes from reaction kinetics) and extrinsic noise (which comes from input signal properties such as how the signal is generated, amplitude, duration etc.) on the detection and false alarm rates.

Another open problem is to study how our work can be used in conjunction with the sequential probability ratio test (SPRT) [28]. The SPRT uses the log-likelihood ratio $L(t)$, such as that defined in (4) which uses data up till time t, with two thresholds θ_1 and θ_0 where $\theta_1 > \theta_0$. At any point in time, if $L(t) > \theta_1$ or $L(t) < \theta_0$ is true, then a decision is reached (deciding for, respectively, $\mathcal{H}_1$ or $\mathcal{H}_1$); otherwise

a decision has not been reached and data collection will continue. It is suggested in [29] that the SPRT is useful for biological detection because theoretically SPRT has a shorter decision time. We can use our approximate log-like ratio $\hat{L}(t)$ in the spirit of SPRT. In our case, only one threshold $\theta_1(> 0)$ will be used. At any point in time, if $\hat{L}(t) > \theta_1$ then a decision is reached for $\mathcal{H}_1$; otherwise a decision has not been reached. (Note the similarity to the study in Sec. IV-B2.) This set up will be useful for practical applications where an action is needed if $\mathcal{H}_1$ is true but otherwise no actions are required. We envisage that this set up will have a shorter decision time as in SPRT. It will be interesting to study how the design parameters mentioned in the last paragraph will influence the decision time and accuracy. This study is important for practical implementation.

VI. CONCLUSION

This paper shows that the C1-FFL gene circuits can be used to approximately compute the log-likelihood ratio for statistical detection of persistent signals. In order to derive this result, we use new methods to show how log-likelihood ratio can be approximately computed and to show how this approximation can be mapped to the parameters of a C1-FFL. An open problem is to study the impact of intrinsic and extrinsic noise on the performance of the C1-FFL based statistical detector. Another open problem is to consider using the approximate log-likelihood ratio computation together with sequential probability ratio test to realise fast persistent detection. Finally, note that the methodology in this paper can be used to derive biochemical circuits that can be used to solve other statistical signal processing problems.

APPENDIX A
DERIVING THE ODE FOR THE EXACT LOG-LIKELIHOOD RATIO

Recalling that $\tilde{\mathcal{Z}}_{X_*}(t)$ is the history of $\tilde{Z}_{X_*}(t)$ in the time interval $[0, t]$. In order to derive (4), we consider the history $\tilde{\mathcal{Z}}_{X_*}(t + \Delta t)$ as a concatenation of $\tilde{\mathcal{Z}}_{X_*}(t)$ and $\tilde{Z}_{X_*}(t)$ in the time interval $(t, t + \Delta t)$. We assume that Δt is chosen small enough so that no more than one reaction can take place in $(t, t + \Delta t)$. Given this assumption and right continuity of continuous-time Markov Chains, we can use $\tilde{Z}_{X_*}(t + \Delta t)$ to denote the history of $\tilde{Z}_{X_*}(t)$ in $(t, t + \Delta t)$.

Consider the likelihood of observing the history $\tilde{\mathcal{Z}}_{X_*}(t + \Delta t)$ given hypothesis $\mathcal{H}_i$:

$$P[\tilde{\mathcal{Z}}_{X_*}(t + \Delta t) | \mathcal{H}_i] \tag{22}$$
$$= P[\tilde{\mathcal{Z}}_{X_*}(t) \text{ and } \tilde{Z}_{X_*}(t + \Delta t) | \mathcal{H}_i] \tag{23}$$
$$= P[\tilde{\mathcal{Z}}_{X_*}(t) | \mathcal{H}_i] \, P[\tilde{Z}_{X_*}(t + \Delta t) | \mathcal{H}_i, \tilde{\mathcal{Z}}_{X_*}(t)] \tag{24}$$

where we have expanded $\tilde{\mathcal{Z}}_{X_*}(t + \Delta t)$ in (22) using concatenation.

By using (24) in the definition of log-likelihood ratio, we can show that:

$$L(t + \Delta t) = L(t) + \log\left(\frac{P[\tilde{Z}_{X_*}(t + \Delta t) | \mathcal{H}_1, \tilde{\mathcal{Z}}_{X_*}(t)]}{P[\tilde{Z}_{X_*}(t + \Delta t) | \mathcal{H}_0, \tilde{\mathcal{Z}}_{X_*}(t)]}\right) \tag{25}$$

The condition probability $P[\widetilde{Z}_{X_*}(t + \Delta t)|\mathcal{H}_i, \widetilde{\mathcal{Z}}_{X_*}(t)]$ is the prediction of the number of $\widetilde{Z}_{X_*}$ molecules at time $t + \Delta t$ based on its history up till time t. This conditional probability can be obtained by solving a Bayesian filtering problem over a continuous-time Markov chain which describes the dynamics of the chemical reactions in (2) and those that produce S [21]. We considered how this conditional probability could be evaluated in our earlier work [30]. The key result in [30] says that $P[\widetilde{Z}_{X_*}(t + \Delta t)|\mathcal{H}_i, \widetilde{\mathcal{Z}}_{X_*}(t)]$ can be expressed in terms of the predicted rate of the chemical reactions that $\widetilde{Z}_{X_*}$ are involved in. By using [21], [30], we have:

$$P[\widetilde{Z}_{X_*}(t + \Delta t)|\mathcal{H}_i, \widetilde{\mathcal{Z}}_{X_*}(t)]$$
$$= \delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)+1} \, g_+(N - \widetilde{Z}_{X_*}(t)) \, J_i(t_-) \, \Delta t$$
$$+ \delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)-1} \, g_- \widetilde{Z}_{X_*}(t) \, \Delta t$$
$$+ \delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)}$$
$$\times (1 - g_+(N - \widetilde{Z}_{X_*}(t))J_i(t) \, \Delta t - g_- \widetilde{Z}_{X_*}(t) \, \Delta t) \quad (26)$$

where $\delta_{a,b}$ is the Kronecker delta which is 1 when a equals to b and zero otherwise, and $J_i(t) = E[X_*(t)|\mathcal{H}_i, \widetilde{\mathcal{Z}}_{X_*}(t)]$ is the expected number of X_* molecules at time t given Hypothesis i and the history $\widetilde{\mathcal{Z}}_{X_*}(t)$.

Note that $P[\widetilde{Z}_{X_*}(t+\Delta t)|\mathcal{H}_i, \widetilde{\mathcal{Z}}_{X_*}(t)]$ in (26) is a sum of three terms with multipliers $\delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)+1}$, $\delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)-1}$ and $\delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)}$. Since these multipliers are mutually exclusive, we have:

$$\log \left(\frac{P[\widetilde{Z}_{X_*}(t + \Delta t)|\mathcal{H}_1, \widetilde{\mathcal{Z}}_{X_*}(t)]}{P[\widetilde{Z}_{X_*}(t + \Delta t)|\mathcal{H}_0, \widetilde{\mathcal{Z}}_{X_*}(t)]} \right)$$
$$= \delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)+1} \log \left(\frac{g_+(N - \widetilde{Z}_{X_*}(t)) J_1(t_-) \Delta t}{g_+(N - \widetilde{Z}_{X_*}(t)) J_0(t_-) \Delta t} \right)$$
$$+ \delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)-1} \log \left(\frac{g_- \widetilde{Z}_{X_*}(t) \Delta t}{g_- \widetilde{Z}_{X_*}(t) \Delta t} \right)$$
$$+ \delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)}$$
$$\times \log \left(\frac{1 - g_+(N - \widetilde{Z}_{X_*}(t))J_1(t) \Delta t - g_- \widetilde{Z}_{X_*}(t) \Delta t}{1 - g_+(N - \widetilde{Z}_{X_*}(t))J_0(t) \Delta t - g_- \widetilde{Z}_{X_*}(t) \Delta t} \right)$$
$$\approx \delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)+1} \log \left(\frac{J_1(t_-)}{J_0(t_-)} \right)$$
$$- \delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)} g_+(N - \widetilde{Z}_{X_*}(t)) (J_1(t) - J_0(t)) \, \Delta t \quad (27)$$

where we have used the approximation $\log(1 + f \, \Delta t) \approx f \, \Delta t$ and have ignored terms of order $(\Delta t)^2$ or higher to obtain (27). Note also that the above derivation assumes that $\frac{J_1(t)}{J_0(t)}$ is strictly positive so its logarithm is well defined; this can be achieved by proper choice of the hypotheses.

By substituting (27) into (25), we have after some manipulations and after taking the limit $\Delta t \to 0$:

$$\frac{dL(t)}{dt} = \lim_{\Delta t \to 0} \frac{\delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)+1}}{\Delta t} \log \left(\frac{J_1(t_-)}{J_0(t_-)} \right)$$
$$- \delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)} g_+(N - \widetilde{Z}_{X_*}(t)) (J_1(t) - J_0(t)) \quad (28)$$
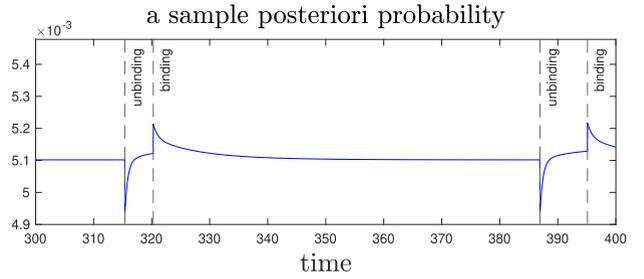
FIGURE 11. Typical temporal behaviour of a $\hat{p}_t(s, x_*)$.

In order to obtain (4), we use the following reasonings. First, the term $\lim_{\Delta t \to 0} \frac{\delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)+1}}{\Delta t}$ is a Dirac delta at the time instant that an X_* molecule binds with $\widetilde{Z}$ to form a $\widetilde{Z}_{X_*}$. Since the binding instants are also the times at which $\widetilde{Z}_{X_*}(t)$ jumps by $+1$, we can identify this term with $\left[\frac{d\widetilde{Z}_{X_*}(t)}{dt} \right]_+$ where $[w]_+ = \max(w, 0)$. Second, the term $\delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)}$ is only zero when the number of $\widetilde{Z}_{X_*}$ molecule changes but the number of such changes is countable. In other words, $\delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)} = 1$ with probability one. This allows us to drop $\delta_{\widetilde{Z}_{X_*}(t+\Delta t),\widetilde{Z}_{X_*}(t)}$. Hence (4). We remark that a more general framework of deriving log-likelihood ratio and log-posteriori probability in the reaction-diffusion master equation framework can be found in [30]–[32].

APPENDIX B
DERIVATION OF THE INTERMEDIATE APPROXIMATION

The aim of this section is to show that the log-likelihood ratio computation in (4) can be approximated by the intermediate approximation in (9) for persistent signals in the time interval $[0, \min(d, d_1)]$. The derivation is based on the assumptions stated in Sec. IV-A. We have divided the derivation into three steps.

(Step 1) The aim of this step is to approximate $E[X_*(t)|\widetilde{\mathcal{Z}}_{X_*}(t), \mathcal{H}_i]$ in (4). The computation of this expectation requires the solution of a Bayesian filtering problem which is computationally intensive. We will argue that if g_+ and g_- are small, then we can use the approximation:

$$E[X_*(t)|\widetilde{\mathcal{Z}}_{X_*}(t), \mathcal{H}_i] \approx \begin{cases} X_1 & \text{for } 0 \le t < d_i \\ X_0 & \text{otherwise.} \end{cases} \quad (29)$$

where $i = 0, 1$, and X_0 and X_1 are defined in (11). In other words, we approximate $E[X_*(t)|\widetilde{\mathcal{Z}}_{X_*}(t), \mathcal{H}_i]$ by a rectangular pulse.

In order to argue for (29), we start by deriving the solution of the Bayesian filtering problem of using the history $\widetilde{\mathcal{Z}}_{X_*}(t)$ to determine the posteriori probability distribution $\hat{p}_t(s, x_*) \triangleq P[S(t) = x, X_*(t) = x_* | \widetilde{\mathcal{Z}}_{X_*}(t), \mathcal{H}_i]$. The typical temporal behaviour of $\hat{p}_t(s, x_*)$ is depicted in Fig. 11. We can see that at the time instants where $\widetilde{Z}$ binds or $\widetilde{Z}_{X_*}$ unbinds, the probability $\hat{p}_t(s, x_*)$ has a discrete jump; at other times, i.e. between two consecutive jumps, the probability $\hat{p}_t(s, x_*)$ varies continuously. By using [21, Eq. (21)], we can show that the time evolution of the continuously varying part of $\hat{p}_t(s, x_*)$
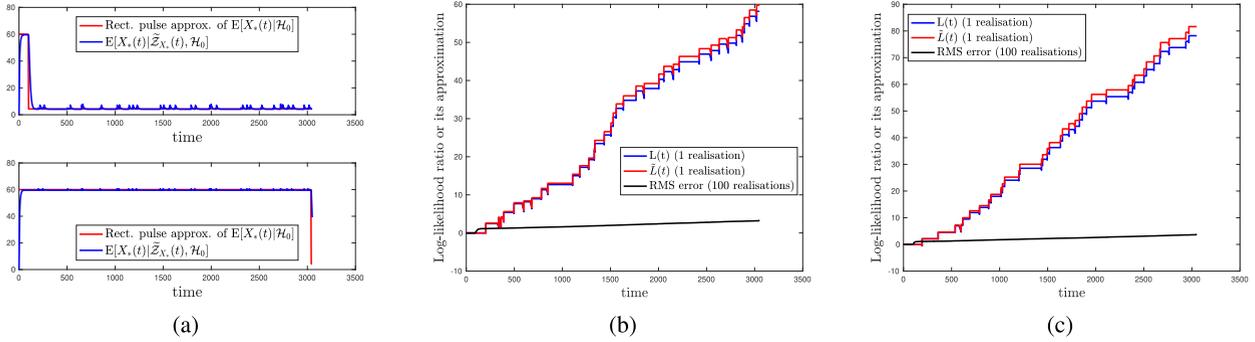
FIGURE 12. (a) Demonstrating the accuracy of (29). Top plot: $\mathcal{H}_0$; Bottom plot: $\mathcal{H}_1$. (b) and (c) Comparing exact log-likelihood ratio $L(t)$ (4) and the approximation $\tilde{L}(t)$ (33). For (b), $S_{ref}^{ON,ref} = 101$, and for (c), $S_{ref}^{ON,ref} = 270$.

is governed by the following ODE:

$$\frac{d\hat{p}_t(s, x_*)}{dt} = [\mathcal{L}_i \hat{p}_t](s, x_*)$$
$$- g_+(x_* - \mathrm{E}[X_*(t)|\widetilde{\mathcal{Z}}_{X_*}(t), \mathcal{H}_i]) (N - \widetilde{\mathcal{Z}}_{X_*}(t)) \quad (30)$$

where

$$[\mathcal{L}_i \hat{p}_t](s, x_*)$$
$$= f_+ \, R_i(t) \, \hat{p}(s-1, x_*) - f_+ \, R_i(t) \, \hat{p}_t(s, x_*)$$
$$+ f_- \, (s+1) \, \hat{p}(s+1, x_*) - f_- \, s \, \hat{p}_t(s, x_*)$$
$$+ k_+ \, s \, (M - x_* + 1) \, \hat{p}(s, x_* - 1)$$
$$- k_+ \, s \, (M - x_*) \, \hat{p}_t(s, x_*)$$
$$+ k_- \, (x_* + 1) \, \hat{p}(s, x_* + 1) - k_- \, x_* \, \hat{p}_t(s, x_*) \quad (31)$$

If g_+ is small, then the last term in (30) can be viewed as a perturbation to the ODE:

$$\frac{dp_t(s, x_*)}{dt} = [\mathcal{L}_i p_t](s, x_*) \quad (32)$$

This ODE is in fact the chemical master equation which governs the reactions (6), (2) with $g_+ = g_- = 0$. For small g_+, we have $\hat{p}_t(s, x_*) \approx p_t(s, x_*)$. In addition, if both g_+ and g_- are small, then the time between two jumps in $\hat{p}_t(s, x_*)$ is long; this further implies that $\hat{p}_t(s, x_*)$ is mostly around steady state, e.g. see Fig. 11 in the time interval [320,386]. Overall, this means that we can approximate $\mathrm{E}[X_*(t)|\widetilde{\mathcal{Z}}_{X_*}(t), \mathcal{H}_i]$ by using the steady state mean number of X_* molecules assuming $g_+ = g_- = 0$. When $g_+ = g_- = 0$, the steady state mean number of X_* molecules is approximately given by the expression in (11). Hence (29).

After using the approximation in Step 1, (4) becomes:

$$\frac{d\tilde{L}(t)}{dt} \approx \left[\frac{d\widetilde{\mathcal{Z}}_{X_*}(t)}{dt} \right]_+ \log\left(\frac{X_1}{X_0} \right) \pi(t)$$
$$- g_+(N - \widetilde{\mathcal{Z}}_{X_*}(t))(X_1 - X_0) \pi(t) \quad (33)$$

where $\pi(t)$ is defined in (14).

We now demonstrate the accuracy of (29) and (33) using numerical examples. We use the same parameter values as in Sec. IV-B1. Fig. 12a plots the two sides of (29) for the two hypotheses. It can seen that, other than the transients,

the approximation is fairly accurate. Next, we use Fig. 12b to demonstrate the accuracy of (33). By using one realisation of $\widetilde{\mathcal{Z}}_{X_*}(t)$, we calculate the exact log-likelihood ratio $L(t)$ (4) and the approximation $\tilde{L}(t)$ (33). The figure shows that the approximation is accurate. In addition, the figure also shows the RMS error between $L(t) - \tilde{L}(t)$ for 100 realisations of $\widetilde{\mathcal{Z}}_{X_*}(t)$. It can be seem that the RMS error is small. In order to show that the approximation holds for different parameter settings, we use a different value $S_{ref}^{ON,ref}$ and plot the result in Fig. 12c.

(Step 2) The aim of this step is to replace $\left[\frac{d\widetilde{\mathcal{Z}}_{X_*}(t)}{dt} \right]_+$ and $\widetilde{\mathcal{Z}}_{X_*}(t)$ in (33) by alternative expressions. Since $\pi(t)$ in (33) is zero for $t < d_0$, we only have to consider input signals whose duration $d \geq d_0$. Recall that we assume in Sec. IV-A that the pathway (2) reaches steady state by d_0 if the input has a duration of at least d_0. This means that the probability distributions of X_* and $\widetilde{\mathcal{Z}}_{X_*}$ are in steady state in the time interval $[d_0, d]$

Note that the RHS of (33) is the sum of two terms that do not depend on $L(t)$. We can therefore consider the contribution of each term to $L(t)$ separately.

First, we consider the contribution of the first term on the RHS of (33) to $L(t)$, which we will call $L_1(t)$:

$$\frac{dL_1(t)}{dt} = \left[\frac{d\widetilde{\mathcal{Z}}_{X_*}(t)}{dt} \right]_+ \log\left(\frac{X_1}{X_0} \right) \pi(t) \quad (34)$$

By using a method similar to that in [5], which is based on the renewal theorem [33], we can show that (34) can be approximated by:

$$\frac{dL_1(t)}{dt} \approx g_- \, \widetilde{\mathcal{Z}}_{X_*}(t) \log\left(\frac{X_1}{X_0} \right) \pi(t) \quad (35)$$

Next, we consider the contribution of the second term on the RHS of (33) to $L(t)$, which we will call $L_2(t)$:

$$\frac{dL_2(t)}{dt} = g_+(N - \widetilde{\mathcal{Z}}_{X_*}(t))(X_1 - X_0) \pi(t) \quad (36)$$

By integrating the above equation, we have $L_2(t) = 0$ for $t \in [0, d_0)$; for $t \in [d_0, \min(d, d_1)]$, we have:

$$L_2(t) = (X_1 - X_0) g_+ \int_{d_0}^t (N - \widetilde{\mathcal{Z}}_{X_*}(\tau)) \, d\tau \quad (37)$$

Since the reaction pathway is in steady state in the time interval $[d_0, \min(d, d_1)]$, we can replace the time average in (37) by its ensemble average. In this part, we will overload the symbol $\widetilde{Z}_{X_*}$ to use it to refer to the random variable of the number of $\widetilde{Z}_{X_*}$ molecules at steady state. This should not cause any confusion because the meaning should be clear from the context. In addition, we will overload the symbol X_* in the same way. With this overloading, the mean number of X_* and $\widetilde{Z}_{X_*}$ molecules at steady state are denoted by $E[X_*]$ and $E[\widetilde{Z}_{X_*}]$ respectively. We can now rewrite (37) as:

$$L_2(t) \approx (X_1 - X_0)\, g_+(N - E[\widetilde{Z}_{X_*}])\, (t - d_0) \quad (38)$$

In order to be able to connect to the C1-FFL, we will need to replace the expression $(N - E[\widetilde{Z}_{X_*}])$ by a different expression. The derivation of this replacement expression requires a few auxiliary results.

(Auxiliary Result 1) By considering the global balance of the steady state of the reaction pathway (2), we have:

$$g_+ E[(N - \widetilde{Z}_{X_*})X_*] = g_- E[\widetilde{Z}_{X_*}] \quad (39)$$

(Auxiliary Result 2) If the amplitude of the input is sufficiently high, then the number of X_* molecules can be approximately modelled as a binomial distribution with M trials and a success probability of $\frac{k_+\alpha}{k_+\alpha+k_-}$.

Consider a binomial distribution $B(Q; m, f)$ with parameters m (number of trials) and f (success probability), then for sufficiently large m and f, we have

$$\frac{1}{E[Q]} \approx E[I(\frac{1}{Q})] \quad (40)$$

where

$$I(\frac{1}{q}) = \begin{cases} 0 & \text{for } q = 0 \\ \dfrac{1}{q} & \text{for } q \geq 1 \end{cases} \quad (41)$$

This result essentially says that the mean of the reciprocal of a binomial random variable (with $\frac{1}{0}$ excluded) is approximately equal to the reciprocal of the mean of the binomial random variable. If $f = 1$ and $m \geq 1$, the binomial distribution has a single outcome with a non-zero probability so (40) is exact. Intuitively, if a probability has a single modal distribution with a narrow spread, then (40) holds approximately. For $f = 0.1$, the relative error of using (40) is 3.21% for $m = 300$ and drops to 1.87% for $m = 500$. In general, the approximation is better for large m and f.

(Auxiliary Result 3) Since the inducer-TF reactions are faster than the TF-gene reactions and $M \gg N$, we can show that:

$$E[(N - \widetilde{Z}_{X_*})X_*] \approx E[(N - \widetilde{Z}_{X_*})]\, E[X_*] \quad (42)$$

We will argue that the above approximation holds by using time average to compute $E[(N - \widetilde{Z}_{X_*})X_*]$. Let $t_0, t_1, \ldots$ be a sequence of time instants at which $\widetilde{Z}_{X_*}$ changes its value. Since the continuous-time Markov chain associated with the

chemical system is ergodic, we have:

$$E[(N - \widetilde{Z}_{X_*})X_*] = \sum_{i=0}^{\infty}(N - \widetilde{Z}_{X_*}(t_i)) \int_{t_i}^{t_{i+1}} X_*(t)\, dt \quad (43)$$

Since the TF-gene reactions are slow in comparison, $\widetilde{Z}_{X_*}$ is a slow species while X_* is a fast species. This means the time interval $[t_i, t_{i+1})$ (during which the count of $\widetilde{Z}_{X_*}$ is a constant) is likely to be long compared to the time-scale of the fast species X_*. This allows us to approximate the integral on the RHS of (43) by $E[X_*](t_{i+1} - t_i)$. Hence (42). Note that the above argument is identical to the one used in [34] to derive the slow-scale tau-leaping simulation algorithm.

(Auxiliary Result 4) By using the same argument as in Auxiliary Result 3, we can show that:

$$E[(N - \widetilde{Z}_{X_*})I(\frac{1}{X_*})] \approx E[(N - \widetilde{Z}_{X_*})]\, E[I(\frac{1}{X_*})] \quad (44)$$

We will now use the above auxiliary results and (38) to derive the replacement expression. By using Auxiliary Results 1 and 3, we have:

$$g_+ E[(N - \widetilde{Z}_{X_*})] \approx g_- E[\widetilde{Z}_{X_*}]\frac{1}{E[X_*]} \quad (45)$$

We then apply Auxiliary Result 2 to the RHS of (45) to obtain:

$$g_+ E[(N - \widetilde{Z}_{X_*})] \approx g_- E[\widetilde{Z}_{X_*}]E[I(\frac{1}{X_*})] \quad (46)$$

By applying Auxiliary Result 4 to the RHS of (46), we have:

$$g_+ E[(N - \widetilde{Z}_{X_*})] \approx g_- E[\widetilde{Z}_{X_*}I(\frac{1}{X_*})] \quad (47)$$

By substituting (47) into (38), we have:

$$L_2(t) \approx (X_1 - X_0)\, g_- E[\widetilde{Z}_{X_*}I(\frac{1}{X_*})](t - d_0) \quad (48)$$

By turning the above equation into the differential form, we have:

$$\frac{dL_2(t)}{dt} \approx (X_1 - X_0)g_- \widetilde{Z}_{X_*}(t)I(\frac{1}{X_*(t)})\, \pi(t) \quad (49)$$

Next, by combining (35) and (49), we have:

$$\frac{dL(t)}{dt} \approx g_- \widetilde{Z}_{X_*}(t)\, \pi(t) \left\{ \log\left(\frac{X_1}{X_0}\right) - (X_1 - X_0)I(\frac{1}{X_*(t)}) \right\} \quad (50)$$

(Step 3) Since a set of chemical reactions can be modelled by a set of ODEs, we want to turn the ODE in (50) into a form that can be implemented by a set of chemical reactions. However, (50) cannot be directly implemented by chemical reactions because log-likelihood ratio can take both positive and negative values but chemical concentration is always non-negative. Although [35] has derived a chemical computation system that can have both positive and negative numbers, it requires double the number of species and reactions. As in our previous work [5], [20], we choose to compute only

the log-likelihood ratio when it is positive. We do that by applying $[\ ]_+$ to the RHS of (50); we have:

$$\frac{dL(t)}{dt} \approx g_- \widetilde{Z}_{X_*}(t)\, \pi(t)$$
$$\times \left[\log\left(\frac{X_1}{X_0}\right) - (X_1 - X_0)I(\frac{1}{X_*(t)}) \right]_+ \quad (51)$$

We now replace $I(\frac{1}{X_*(t)})$ in (51) by $\frac{1}{X_*(t)}$ to obtain:

$$\frac{dL(t)}{dt} \approx g_- \widetilde{Z}_{X_*}(t)\, \pi(t)$$
$$\times \left[\log\left(\frac{X_1}{X_0}\right) - (X_1 - X_0)\frac{1}{X_*(t)} \right]_+ \quad (52)$$

The removal of $I(\ )$ will not make much difference because the probability of having $X_*(t)$ equals to 0 is small when the input signal is persistent. Note that (52) is the same as (9). This completes the derivation for (9).

In order to derive (16), we start from (52) and take expectation on both sides. If the amplitude α is sufficiently high, then there is a high probability that $X_*(t)$ is large. This means we can take the expectation operator to the inside of the $[\ ]_+$ operator. After that we apply Auxiliary Results 2, 3 and 4 to obtain (16).

REFERENCES

[1] J. J. Y. Teo, R. Weiss, and R. Sarpeshkar, "An artificial tissue homeostasis circuit designed via analog circuit techniques," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 3, pp. 540–553, Jun. 2019.

[2] L. Pasotti, M. Bellato, N. Politi, M. Casanova, S. Zucca, M. G. C. De Angelis, and P. Magni, "A synthetic close-loop controller circuit for the regulation of an extracellular molecule by engineered bacteria," *IEEE Trans. Biomed. Circuits Syst.*, vol. 31, no. 1, pp. 248–258, Feb. 2018.

[3] C. McBride, R. Shah, and D. D. Vecchio, "The effect of loads in molecular communications," *Proc. IEEE*, vol. 107, no. 7, pp. 1369–1386, May 2019.

[4] M. Kuscu and O. B. Akan, "Channel sensing in molecular communications with single type of ligand receptors," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6868–6884, Oct. 2019.

[5] C. T. Chou, "Designing molecular circuits for approximate maximum a posteriori demodulation of concentration modulated signals," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5458–5473, Aug. 2019.

[6] H. Awan and C. T. Chou, "Improving the capacity of molecular communication using enzymatic reaction cycles," *IEEE Trans. Nanobiosci.*, vol. 16, no. 8, pp. 744–754, Dec. 2018.

[7] C. Zechner, G. Seelig, M. Rullan, and M. Khammash, "Molecular circuits for dynamic noise filtering," *Proc. Nat. Acad. Sci. USA*, vol. 113, pp. 4729–4734, Apr. 2016.

[8] R. Sarpeshkar, "Analog synthetic biology," *Philos. Trans. Roy. Soc. A, Math. Phys. Eng. Sci.*, vol. 372, no. 2012, Mar. 2014, Art. no. 20130110.

[9] J. J. Y. Teo, S. S. Woo, and R. Sarpeshkar, "Synthetic biology: A unifying view and review using analog circuits," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 4, pp. 453–474, Aug. 2015.

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

[11] 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.

[12] C. Briat, A. Gupta, and M. Khammash, "Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks," *Cell Syst.*, vol. 2, pp. 15–26, Jan. 2016.

[13] S. Mangan, A. Zaslaver, and U. Alon, "The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks," *J. Mol. Biol.*, vol. 334, no. 2, pp. 197–204, Nov. 2003.

[14] M. Mimee, A. C. Tucker, C. A. Voigt, and T. K. Lu, "Programming a human commensal bacterium, bacteroides thetaiotaomicron, to sense and respond to stimuli in the murine gut microbiota," *Cell Syst.*, vol. 1, no. 1, pp. 62–71, Jul. 2015.

[15] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 2. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[16] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon, "Network motifs in the transcriptional regulation network of escherichia coli," *Nature Genet.*, vol. 31, no. 1, pp. 64–68, Apr. 2002.

[17] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network motifs: Simple building blocks of complex networks," *Science*, vol. 298, no. 5594, pp. 824–827, Oct. 2002.

[18] E. Libby, T. J. Perkins, and P. S. Swain, "Noisy information processing through transcriptional regulation," *Proc. Nat. Acad. Sci. USA*, vol. 104, no. 17, pp. 7151–7156, Apr. 2007.

[19] C. Gardiner, *Stochastic Methods*. Berlin, Germany: Springer, 2010.

[20] C. T. Chou, "Detection of persistent signals and its relation to coherent feed-forward loops," *Roy. Soc. Open Sci.*, vol. 5, no. 11, Nov. 2018, Art. no. 181641.

[21] L. Bronstein and H. Koeppl, "Marginal process framework: A model reduction tool for Markov jump processes," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 97, no. 6, pp. 1–14, Jun. 2018.

[22] C. T. Chou, "Using detection theory and molecular computation to understand signal processing in living cells," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 1822–1826.

[23] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.*, vol. 81, no. 25, pp. 2340–2361, Dec. 1977.

[24] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*. London, U.K.: Chapman & Hall, 2006.

[25] C. T. Chou, "Using biochemical circuits to approximately compute log-likelihood ratio for detecting persistent signals," 2019, *arXiv:1907.09841*. [Online]. Available: http://arxiv.org/abs/1907.09841

[26] T. Tian and K. Burrage, "Stochastic models for regulatory networks of the genetic toggle switch," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 22, pp. 8372–8377, May 2006.

[27] P. S. Swain, M. B. Elowitz, and E. D. Siggia, "Intrinsic and extrinsic contributions to stochasticity in gene expression," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 20, pp. 12795–12800, Oct. 2002.

[28] A. Wald, "Sequential tests of statistical hypotheses," *Ann. Math. Statist.*, vol. 16, no. 2, pp. 117–186, 1945.

[29] E. D. Siggia and M. Vergassola, "Decisions on the fly in cellular sensory systems," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 39, pp. E3704–E3712, Sep. 2013.

[30] H. Awan and C. T. Chou, "Generalized solution for the demodulation of reaction shift keying signals in molecular communication networks," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 715–727, Feb. 2017.

[31] C. T. Chou, "Maximum a-posteriori decoding for diffusion-based molecular communication using analog filters," *IEEE Trans. Nanotechnol.*, vol. 14, no. 6, pp. 1054–1067, Nov. 2015.

[32] C. T. Chou, "A Markovian approach to the optimal demodulation of diffusion-based molecular communication networks," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3728–3743, Oct. 2015.

[33] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. Oxford, U.K.: Oxford Univ. Press, 1997.

[34] Y. Cao, D. T. Gillespie, and L. R. Petzold, "The slow-scale stochastic simulation algorithm," *J. Chem. Phys.*, vol. 122, no. 1, Jan. 2005, Art. no. 014116.

[35] K. Oishi and E. Klavins, "Biomolecular implementation of linear I/O systems," *IET Syst. Biol.*, vol. 5, no. 4, pp. 252–260, Jul. 2011.

CHUN TUNG CHOU (Member, IEEE) received the B.A. degree in engineering science from the University of Oxford, U.K., and the Ph.D. degree in control engineering from the University of Cambridge, U.K. He is currently an Associate Professor with the School of Computer Science and Engineering, University of New South Wales, Australia. His research interests include molecular communication, molecular computing, and pervasive computing. He was on the Editorial Board of IEEE Transactions on Molecular, Biological, and Multi-Scale Communications, IEEE Wireless Communications Letters, and *Nano Communication Networks*.