

Research



Cite this article: Chou CT. 2018 Detection of persistent signals and its relation to coherent feed-forward loops. *R. Soc. open sci.* **5**: 181641. <http://dx.doi.org/10.1098/rsos.181641>

Received: 1 October 2018

Accepted: 8 October 2018

Subject Category:

Biochemistry and biophysics

Subject Areas:

systems biology

Keywords:

Coherent feed-forward loops, detection of persistent signals, detection theory, likelihood ratio, dynamical systems, time-scale separation

Author for correspondence:

Chun Tung Chou

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

Detection of persistent signals and its relation to coherent feed-forward loops

Chun Tung Chou

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

CTC, 0000-0003-4512-7155

Many studies have shown that cells use the temporal dynamics of signalling molecules to encode information. One particular class of temporal dynamics is persistent and transient signals, i.e. signals of long and short duration, respectively. It has been shown that the coherent type-1 feed-forward loop with an AND logic at the output (or C1-FFL for short) can be used to discriminate a persistent input signal from a transient one. This has been done by modelling the C1-FFL, and then using the model to show that persistent and transient input signals give, respectively, a non-zero and zero output. The aim of this paper is to make a connection between the statistical detection of persistent signals and the C1-FFL. We begin by first formulating a statistical detection problem of distinguishing persistent signals from transient ones. The solution of the detection problem is to compute the log-likelihood ratio of observing a persistent signal to a transient signal. We show that, if this log-likelihood ratio is positive, which happens when the signal is likely to be persistent, then it can be approximately computed by a C1-FFL. Although the capability of C1-FFL to discriminate persistent signals is known, this paper adds an information processing interpretation on how a C1-FFL works as a detector of persistent signals.

1. Introduction

By analysing the graph of the transcription networks of the bacterium *Escherichia coli* and the yeast *Saccharomyces cerevisiae*, the authors in [1–3] discovered that there were sub-graphs that appear much more frequently in these transcription networks than in randomly generated networks. These frequently occurring sub-graphs are called network motifs. A particular example of a network motif is the coherent type-1 feed-forward loop with an AND logic at the output (or C1-FFL for short). C1-FFL is the most abundant type of coherent feed-forward loop in the transcription networks of *E. coli* and *S. cerevisiae* [4].

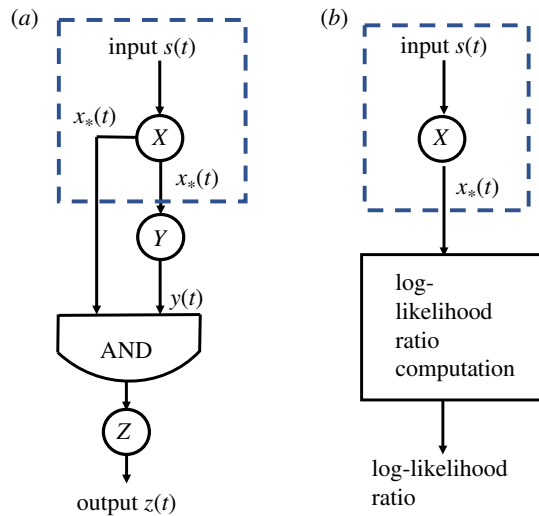


Figure 1. (a) The coherent type-1 feed-forward loop with AND logic. (b) The detection theory framework.

An example of C1-FFL in *E. coli* is the L-arabinose utilization system which activates the transcription of the *araBAD* operon when glucose is absent and L-arabinose is present [5]. By modelling the C1-FFL with ordinary differential equations (ODE), the authors in [2,4] show that the C1-FFL can act as a persistence detector to differentiate persistent input signals (i.e. signals of long duration) from transient signals (i.e. signals of short duration). The aim of this paper is to present a new perspective of the persistence detection property of C1-FFL from an information processing point of view.

In information processing, the problem of distinguishing signals which have some specific features from those which do not has been studied under the theory of statistical detection [6]. An approach to detection is to formulate a hypothesis testing problem where the alternative hypothesis (resp. null hypothesis) is that the observed signal does have (does not have) the specific features. The next step is to use the observed signal to compute the likelihood ratio to determine which hypothesis is more likely to hold. Since a C1-FFL can detect persistent signals, the question is whether the C1-FFL can be interpreted as a statistical detector. We show in this paper that the C1-FFL is related to a detection problem whose aim is to distinguish a long rectangular pulse (a prototype persistent signal) from a short rectangular pulse (a prototype transient signal). In particular, we show that, for persistent input signals, the output of the C1-FFL can be interpreted as the log-likelihood ratio of this detection problem. This result therefore provides an information processing interpretation of the computation being carried out by a C1-FFL.

2. Background

2.1. C1-FFL

The properties of coherent feed-forward loops have been studied in [2,4,7] using ODE models and in [5] experimentally. Here, we will focus on the property of C1-FFL with AND logic to detect persistent signals. We do that by using an idealized model of C1-FFL adapted from the text [7]. The model retains the important features of C1-FFL and is useful in understanding the derivation in this paper.

Figure 1a depicts the structure of the C1-FFL. One can consider both X and Y as transcription factors (TFs) which regulate the transcription of Z . The TF X is activated by the input signal $s(t)$ which acts as an inducer. We will denote the active form of X by X_* . Following [4], we assume that the activation of X (resp. the deactivation of X_*) is instantaneous when the input signal is present (absent). The active form X_* can be used to produce Y if its concentration exceeds a threshold K_{xy} . We use $[Y]$ to denote the concentration of Y . We write the reaction rate equation for Y as

$$\frac{d[Y]}{dt} = \beta_y \theta([X_*] > K_{xy}) - \alpha_y [Y], \quad (2.1)$$

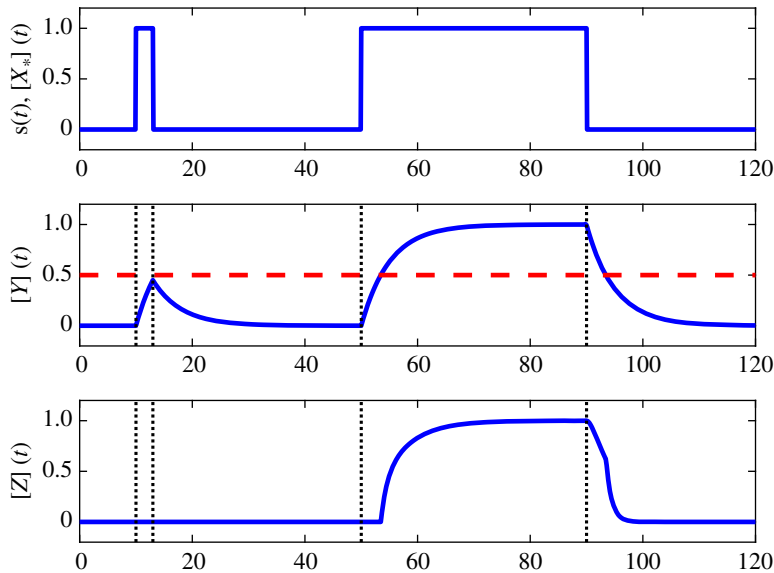


Figure 2. How C1-FFL detects persistent signals.

where β_y and α_y are reaction rate constants, and $\theta(c)$ is 1 if the Boolean condition c within the parentheses is true, or is otherwise 0.

The transcription of Z requires the concentration of X_* to be greater than K_{xz} and the concentration of Y to be greater than K_{yz} , which corresponds to the AND gate in figure 1a. The reaction rate equation for the output Z is:

$$\frac{d[Z]}{dt} = \beta_z \theta([X_*] > K_{xz}) \theta([Y] > K_{yz}) - \alpha_z [Z], \quad (2.2)$$

where β_z and α_z are reaction rate constants.

We now present a numerical example to demonstrate how the C1-FFL can be used to detect persistent signals. We assume the input signal $s(t)$ consists of a short pulse of duration 3 (the transient signal) followed by a long pulse of duration 40 (the persistent signal). We also assume that $s(t)$ has an amplitude of 1 when it is ON. The other parameter values are $\alpha_y = \beta_y = 0.2$, $K_{xy} = 0.6$, $\alpha_z = \beta_z = 1$, $K_{xz} = 0$ and $K_{yz} = 0.5$.

Since the activation of X or deactivation of X_* is instantaneous, we assume $[X_*](t) = s(t)$ for simplicity. The time profile of $s(t) = [X_*](t)$ is shown in the top plot in figure 2.

The middle plot of figure 2 shows $[Y](t)$. Since $[X_*](t) > K_{xy}$ when the input $s(t)$ is ON, the production of Y occurs during this period. When the pulse is short, the amount of Y being produced is limited and the maximum $[Y]$ is below K_{yz} , which is indicated by the dashed red horizontal line in the middle plot. Since the production of Z requires both $[X_*] > K_{xz}$ and $[Y] > K_{yz}$ (i.e. the AND gate) but the latter condition is not satisfied, therefore no Z is produced when the pulse is short. The bottom plot shows $[Z](t)$ is zero when a short pulse is applied. However, when the pulse is long, the concentration of $[Y]$ is given enough time to increase beyond the threshold K_{yz} and as a result we see the production of Z , as shown in the bottom plot. Note that when the pulse is long, the production of Z only starts after a delay; this is because the AND condition for the production of Z in equation (2.2) does not hold initially. This example shows that, for an ideal C1-FFL, a transient input will produce a zero output and a persistent input will give a non-zero output.

2.2. Detection theory

Detection theory is a branch of statistical signal processing. Its aim is to use the measured data to decide whether an event of interest has occurred. For example, detection theory is used in radar signal processing to determine whether a target is present or not. In the context of this paper, the events are whether the signal is transient or persistent. A detection problem is often formulated as a hypothesis testing problem, where each hypothesis corresponds to a possible event. Let us consider a detection problem with two hypotheses, denoted by \mathcal{H}_0 and \mathcal{H}_1 , which correspond to, respectively, the events

of transient and persistent signals. Our aim is to decide which hypothesis is more likely to hold. We define the log-likelihood ratio R :

$$R = \log \left(\frac{P[\text{measured data} | \mathcal{H}_1]}{P[\text{measured data} | \mathcal{H}_0]} \right), \quad (2.3)$$

where $P[\text{measured data} | \mathcal{H}_i]$ is the conditional probability that the measured data are generated by the signal specified in hypothesis \mathcal{H}_i . Note that we have chosen to use the log-likelihood ratio, rather than the likelihood ratio, because it will enable us to build a connection with C1-FFL later on. Intuitively, if the log-likelihood ratio R is positive, then the measured data are more likely to have been generated by a persistent signal or hypothesis \mathcal{H}_1 , and vice versa. Therefore, the key idea of detection theory is to use the measured data to compute the log-likelihood ratio and then use it to make a decision.

3. Connecting detection theory with C1-FFL

We will now present a big picture explanation of how we will connect detection theory with C1-FFL. The signal $x_*(t)$ in figure 1*a* is the output signal of Node X in the C1-FFL. We can view the C1-FFL as a two-stage signal processing engine. In the first stage, the input signal $s(t)$ is processed by Node X to obtain $x_*(t)$ and this is the part within the dashed box in figure 1*a*. In the second stage, the signal $x_*(t)$ is processed by the rest of the C1-FFL to produce the output signal $z(t)$. We will now make a connection to detection theory. Our plan is to apply detection theory to the dashed box in figure 1*a*. We consider $x_*(t)$ as the measured data and use them to determine whether the input signal is transient or persistent. Detection theory tells us that we should use $x_*(t)$ to compute the log-likelihood ratio. This means that we can consider the two-stage signal processing depicted in figure 1*b* where the input signal $s(t)$ generates $x_*(t)$ and the measured data $x_*(t)$ are used to calculate the log-likelihood ratio. If we can identify the log-likelihood ratio calculation in figure 1*b* with the processing by the part of C1-FFL outside of the dashed box, then we can identify the signal $z(t)$ with the log-likelihood ratio.

4. Detection of persistent signals

4.1. Defining the detection problem

We first define the problem for detecting a persistent signal using detection theory. Our first step is to specify the signalling pathway in Node X, which consists of three chemical species: signalling molecule S, molecular type X in an inactive form and its active form X_* . The activation and inactivation reactions are:



and



where k_+ and k_- are reaction rate constants. Let $x(t)$ and $x_*(t)$ denote, respectively, the *number* of X and X_* molecules at time t . Note that both $x(t)$ and $x_*(t)$ are piecewise constant because they are molecular counts. We assume that $x(t) + x_*(t)$ is a constant for all t and we denote this constant by M .

We assume that the input signal $s(t)$, which is the concentration of the signalling molecules S at time t , is a deterministic signal. We also assume that the signal $s(t)$ cannot be observed, so any characteristics of $s(t)$ can only be inferred.

We model the dynamics of the chemical reactions by using the chemical master equation [8]. This means that $x_*(t)$ is a realization of a continuous-time Markov chain. This also means that the same input signal $s(t)$ can result in different $x_*(t)$.

The measured datum at time t is $x_*(t)$. However, in the formulation of the detection problem, we will assume that at time t , the data available to the detection problem are $x_*(\tau)$ for all $\tau \in [0, t]$; in other words, the data are continuous in time and are the history of the counts of X_* up to time t inclusively. We will use $\mathcal{X}_*(t)$ to denote the continuous-time history of $x_*(t)$ up to time t inclusively. Note that even though we assume that the entire history $\mathcal{X}_*(t)$ is available for detection, we will see later on that the calculation of the log-likelihood ratio at time t does not require the storage of the past history.

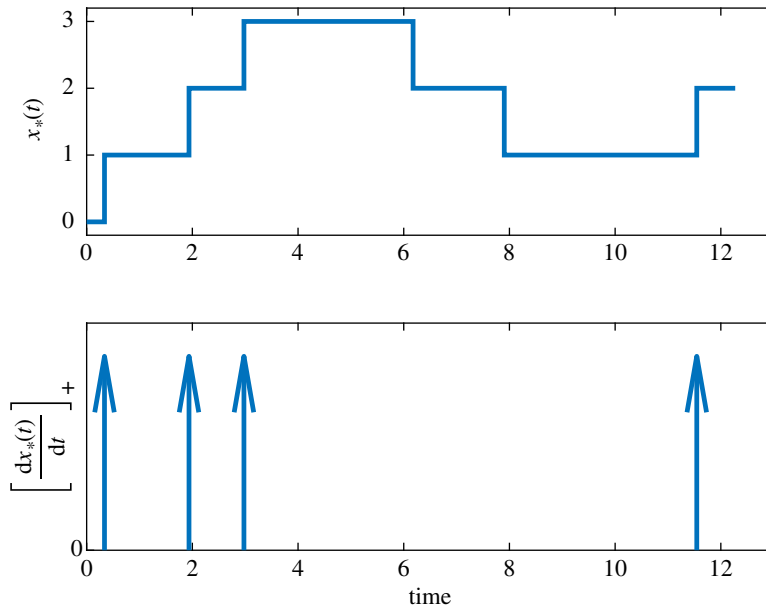


Figure 3. An example $x_*(t)$ and $[dx_*(t)/dt]_+$.

The last step in defining the detection problem is to specify the hypotheses \mathcal{H}_i ($i = 0, 1$). 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 the hypothesis \mathcal{H}_0 (resp. \mathcal{H}_1) is that the input signal $s(t)$ is the signal $c_0(t)$ (resp. $c_1(t)$) where $c_0(t)$ and $c_1(t)$ are two different deterministic signals. Intuitively, the aim of the detection problem is to use the history $\mathcal{X}_*(t)$ to decide which of the two signals, $c_0(t)$ or $c_1(t)$, is more likely to have produced the observed history.

4.2. Solution to the detection problem

Based on the definition of the detection problem, the log-likelihood ratio $L(t)$ at time t is given by:

$$L(t) = \log \left(\frac{P[\mathcal{X}_*(t) | \mathcal{H}_1]}{P[\mathcal{X}_*(t) | \mathcal{H}_0]} \right), \quad (4.2)$$

where $P[\mathcal{X}_*(t) | \mathcal{H}_i]$ is the conditional probability of observing the history $\mathcal{X}_*(t)$ given hypothesis \mathcal{H}_i . We show in appendix A 1 that $L(t)$ obeys the following ODE:

$$\begin{aligned} \frac{dL(t)}{dt} = & \left[\frac{dx_*(t)}{dt} \right]_+ \log \left(\frac{c_1(t)}{c_0(t)} \right) \\ & - k_+(M - x_*(t))(c_1(t) - c_0(t)), \end{aligned} \quad (4.3)$$

where $[w]_+ = \max(w, 0)$. We also assume that the two hypotheses are *a priori* equally likely, so $L(0) = 0$. Since $x_*(t)$ is a piecewise constant function counting the number of X_* molecules, its derivative is a sequence of Dirac deltas at the time instants that X is activated or X_* is deactivated. Note that the Dirac deltas corresponding to the activation of X carry a positive sign and the $[\]_+$ operator keeps only these. Figure 3 shows an example $x_*(t)$ and its corresponding $[dx_*(t)/dt]_+$. We remark that the derivation of equation (4.3) requires that both $c_0(t)$ and $c_1(t)$ are strictly positive for all t , otherwise (4.3) is not well defined.

Note that a special case of equation (4.3) with constant $c_i(t)$ and $M = 1$ appeared in [9]. An equation of the same form as equation (4.3) is used in [10] to understand how cells can distinguish between the presence and absence of a stimulus. A more general form of equation (4.3) which includes the diffusion of signalling molecules can be found in [11].

The importance of equation (4.3) is that, given the measured data $x_*(t)$, we can use it together with $c_i(t)$ to compute the log-likelihood ratio $L(t)$. We will use an example to illustrate how equation (4.3) can be used to distinguish between two signals of different durations. This example will also be used to illustrate what information is useful to distinguish such signals.

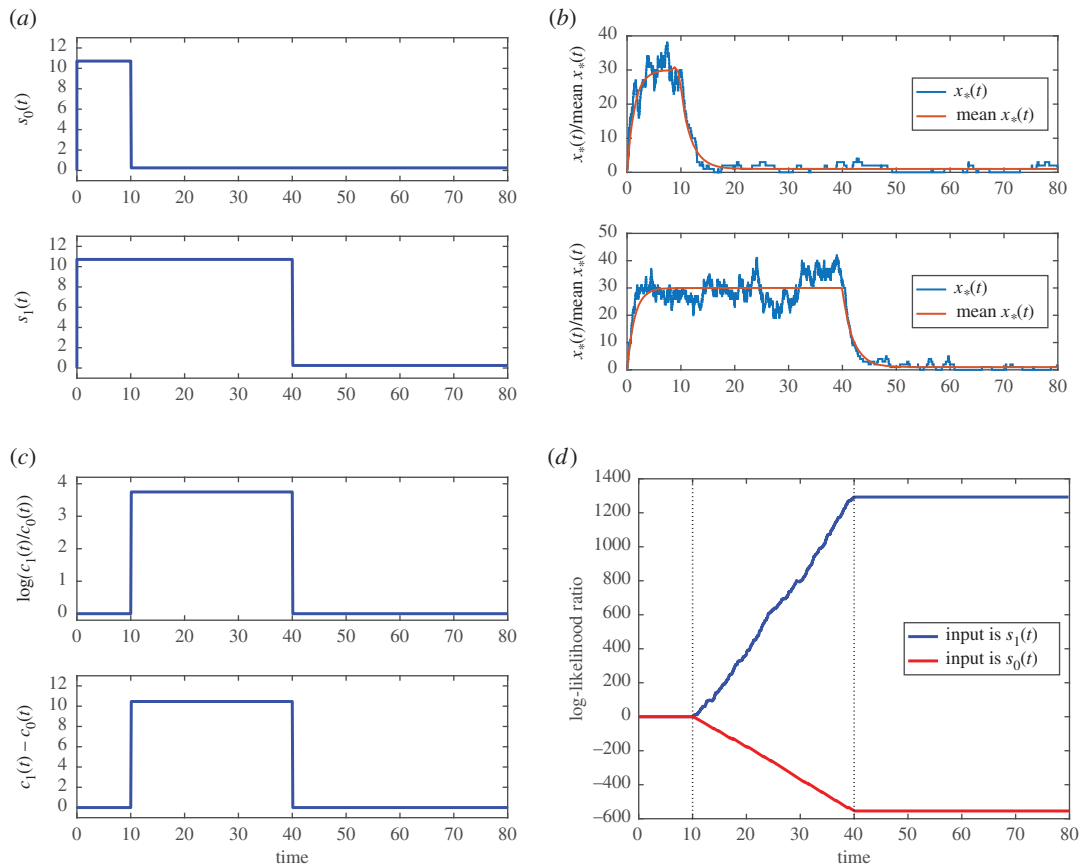


Figure 4. Example of distinguishing between a long and a short rectangular pulse. (a) Top graph: signal $s_0(t)$; bottom graph: signal $s_1(t)$; (b) top graph: $x_*(t)$ for input signal $s_0(t)$; bottom graph: $x_*(t)$ for input signal $s_1(t)$; (c) top graph: $\log(c_1(t)/c_0(t))$; bottom graph: $c_1(t) - c_0(t)$; and (d) log-likelihood ratio $L(t)$.

4.2.1. Example: using the log-likelihood ratio to distinguish between a long and a short pulse

In this example, we consider using equation (4.3) to distinguish between two possible input signals $s_0(t)$ and $s_1(t)$. Both $s_0(t)$ and $s_1(t)$ are rectangular pulses where $s_1(t)$ has a longer duration than $s_0(t)$. For simplicity, we assume that the reference signals $c_0(t) = s_0(t)$ and $c_1(t) = s_1(t)$.

In order to perform the numerical computation, we assume $k_+ = 0.02$, $k_- = 0.5$ and $M = 100$. The time profiles of $s_0(t)$ and $s_1(t)$ are shown in figure 4a. The durations of $s_0(t)$ and $s_1(t)$ are, respectively, 10 and 40 time units. The amplitude of the pulses when they are ON is 10.7 and it is 0.25 when they are OFF.

We use simulation to produce the measured data $x_*(t)$. We first use the input $s_0(t)$ together with the Stochastic Simulation Algorithm [12] to simulate the reactions (4.1). This produces the simulated $x_*(t)$ in the top plot of figure 4b. After that, we do the same with $s_1(t)$ as the input and this produces the simulated $x_*(t)$ in the bottom plot of figure 4b. It is important to point out that although we have plotted $s_0(t)$, $s_1(t)$ and the two time series of $x_*(t)$ in figures 4a,b using the same time interval, we are performing two separate numerical experiments: one with $s_0(t)$ as the input and the other with $s_1(t)$ as the input.

The log-likelihood ratio calculation in equation (4.3) uses the reference signals $c_0(t)$ and $c_1(t)$. We see from equation (4.3) that these two reference signals are used to form two weighting functions, $\log(c_1(t)/c_0(t))$ and $(c_1(t) - c_0(t))$. By using the assumed time profiles of $c_0(t)$ and $c_1(t)$, we can compute these two weighting functions and we have plotted them in figure 4c. It can be seen that both weighting functions are non-zero in the time interval $[10, 40)$ but are otherwise zero. This means that the computation of $L(t)$ is only using the measured data in the time interval $[10, 40)$ to determine whether the input signal is $c_0(t)$ or $c_1(t)$. This is because, outside of the time interval $[10, 40)$, the two data series $x_*(t)$ generated by $s_0(t)$ and $s_1(t)$ have the same statistical behaviour and therefore there is no information outside of $[10, 40)$ to say whether the input is long or short. Hence, a lesson we have learnt from this example is that the informative part of the data is when the long pulse is expected to be ON and the short pulse is expected to be OFF.

We first use the $x_*(t)$ generated by $s_0(t)$, together with the time profiles of $c_0(t)$ and $c_1(t)$, to compute the log-likelihood ratio $L(t)$ by numerically integrating equation (4.3). The resulting $L(t)$ is the red curve in figure 4d. Similarly, the blue curve in figure 4d shows the $L(t)$ corresponding to the input $s_1(t)$. We can see distinct behaviours in the two $L(t)$'s in the time intervals $[0, 10)$, $[10, 40)$ and $t \geq 40$. The behaviour in the time intervals $[0, 10)$ and $t \geq 40$ is simple to explain because $dL/dt = 0$ in these time intervals.

We next focus on the time interval $[10, 40)$. We first consider $s_1(t)$ as the input. In this time interval, a large $s_1(t)$ means the activation X continues to happen: see the bottom plot of figure 4b. The activation of X contributes to an increase in $L(t)$ due to the first term on the right-hand side (RHS) of equation (4.3). Although the second term of equation (4.3) contributes to a decrease in $L(t)$ via $(M - x_*(t))$, which is the number of inactive X, the contribution is comparatively small. Therefore, we see that the log-likelihood ratio $L(t)$, which is the blue curve in figure 4d, becomes more positive. Since a positive log-likelihood ratio means that the input signal is more likely to be similar to the reference signal $c_1(t)$, this is a correct detection. In a similar way, we can explain the behaviour of the red curve in figure 4d when $s_0(t)$ is applied.

A lesson that we can learn from the last paragraph is that, if our aim is to distinguish a persistent signal from a transient one accurately, then we want the persistent signal to produce a large positive $L(t)$. Since the positive contribution of $L(t)$ comes from the first term on the RHS of equation (4.3), we can get a large positive $L(t)$ by making sure that a persistent signal will produce many activations. This occurs when a persistent signal has a duration which is long compared with the time-scale of the activation and deactivation reactions (4.1)—we will make use of this condition later.

5. Connecting the log-likelihood calculation to C1-FFL

5.1. Choosing detection problem parameters to match the behaviour of C1-FFL

The detection problem defined in §4 is general and can be applied to any two chosen reference signals $c_0(t)$ and $c_1(t)$. In order to connect the detection problem in §4 to the fact that C1-FFL is a persistence detector, we will need to make specific choices for $c_0(t)$ and $c_1(t)$. In this paper, we will choose the reference signals $c_0(t)$ and $c_1(t)$ to be rectangular (or ON/OFF) pulses. Furthermore, we assume that when the reference signal is ON, its concentration level is a_1 , and when it is OFF, its concentration level is at the basal level a_0 with $a_1 > a_0 > 0$. The temporal profile of $c_i(t)$ (where $i = 0, 1$) is:

$$c_i(t) = \begin{cases} a_1 & \text{for } 0 \leq t < d_i \\ a_0 & \text{otherwise,} \end{cases} \quad (5.1)$$

where d_i is the duration of the pulse $c_i(t)$. In particular, we assume that the duration of $c_1(t)$ is longer than $c_0(t)$, i.e. $d_1 > d_0$. We can therefore identify $c_0(t)$ and $c_1(t)$ as the reference signals for, respectively, the transient and persistent signals.

We remark that there may be other choices of reference signals that can connect the detection problem in §4 to the one solved by C1-FFL; we will leave that for future work.

Remark 5.1. We would like to make a remark on the detection problem formulation. In this paper, we have chosen to formulate the detection problem by assuming that each hypothesis \mathcal{H}_0 and \mathcal{H}_1 consists of one reference signal. Such hypotheses, which consist of only one possibility per hypothesis, are known as simple hypotheses in the statistical hypothesis testing literature [6]. We know from [6] that if both hypotheses are simple, then the solution of the detection problem is to compute the likelihood ratio (2.3). In this paper, we have chosen to use simple hypotheses for \mathcal{H}_0 and \mathcal{H}_1 so as to make the problem trackable. In order to understand that, let us explore an alternative detection problem formulation.

An alternative formulation would be to assume that \mathcal{H}_0 (resp. \mathcal{H}_1) consists of all rectangular pulses with duration less than (greater than or equal to) a pre-defined threshold d_0 . In this case, both \mathcal{H}_0 and \mathcal{H}_1 are known as composite hypotheses. To the best of our knowledge, there are no standard solutions to the hypothesis testing problem with composite hypotheses at the moment. Although the text [6] presented two methods to deal with composite hypotheses, neither of them appears to be trackable because the Bayesian approach requires the evaluation of an integral and the generalized likelihood ratio test requires the solution to two optimization problems. Therefore, we have not considered them in this paper.

5.2. Computing an intermediate approximation

Our ultimate goal is to connect the computation of the log-likelihood ratio $L(t)$ in equation (4.3) to the computation carried out by C1-FFL. We will first derive an intermediate approximation for equation (4.3). In order to motivate why this intermediate approximation is necessary, one first needs to know that the C1-FFL realizes computation by using chemical reactions, and research from the molecular computation in synthetic biology has taught us that some computations are difficult to carry out by chemical reactions [13]. For equation (4.3), the difficulties are: (i) the log-likelihood ratio can take any real value but chemical concentration can only be non-negative; and (ii) it is difficult to calculate derivatives using chemical reactions. The aim of the intermediate approximation is to remove these difficulties. In addition, we want the computation to make use of $x_*(t)$ (number of active species X_*) instead of $M - x_*(t)$ (number of inactive species X) because signalling pathways typically use the active species to propagate information.

In order to analytically derive the intermediate approximation, we will need to assume that the input signal $s(t)$ has a certain form. Our derivation assumes that the input $s(t)$ is a rectangular pulse with the following temporal profile:

$$s(t) = \begin{cases} a & \text{for } 0 \leq t < d \\ a_0 & \text{otherwise,} \end{cases} \quad (5.2)$$

where d is the pulse duration, and a is the pulse amplitude when it is ON where $a > a_0$. Note that the parameters a and d are not fixed; and we will show that the intermediate approximation holds for a range of a and d .

In appendix A 2, we start from equation (4.3) and use a time-scale separation argument to derive the intermediate approximation $\hat{L}(t)$. The intermediate approximation $\hat{L}(t)$ has the following properties: if the input signal $s(t)$ is persistent, then $\hat{L}(t)$ approximates the log-likelihood ratio $L(t)$; if the input signal $s(t)$ is transient, then $\hat{L}(t)$ is zero. Note that the latter property is consistent with the behaviour of the ideal C1-FFL which gives a zero output for transient signals. The time evolution of $\hat{L}(t)$ is given by the following ODE:

$$\frac{d\hat{L}(t)}{dt} = x_*(t) \times \underbrace{\{k_- \pi(t) [\phi(s(t))]_+\}}_{=\eta(t)} \quad (5.3)$$

$$\text{where } \phi(u) = \log\left(\frac{a_1}{a_0}\right) - \frac{a_1 - a_0}{u}, \quad (5.4)$$

$$\pi(t) = \begin{cases} 1 & \text{for } d_0 \leq t < d_1 \\ 0 & \text{otherwise} \end{cases} \quad (5.5)$$

and

$$\hat{L}(0) = 0. \quad (5.6)$$

The behaviour of the intermediate approximation $\hat{L}(t)$ depends on the duration d of the input signal $s(t)$. Two important properties for $\hat{L}(t)$, which are discussed in further detail in appendix A 2, are:

- (1) If $d < d_0$, then $\hat{L}(t)$ is zero for all t .
- (2) If $d \geq d_0$ and if the duration $d - d_0$ is long compared to $1/k_+a + 1/k_-$, then $\hat{L}(t) \approx L(t)$ for $0 \leq t < \min\{d, d_1\}$ where $L(t)$ is given in equation (4.3).

We can consider those input signals $s(t)$ whose duration d is less than d_0 as transient signals. The first property says that these signals will give a zero $\hat{L}(t)$. Note that for the ideal C1-FFL considered in §2.1, a transient signal gives a zero output.

Those signals whose duration d is greater than or equal to d_0 are considered to be persistent. The second property concerns persistent signals with the property that the duration d and amplitude a have to be such that $d - d_0$ is long compared to $1/k_+a + 1/k_-$, which is the mean time between two consecutive activations of an X molecule. The physical effect of these signals is to produce a large number of activations and deactivations when the input signal $s(t)$ is ON. We argue in appendix A 2 that, if these conditions hold, then it is possible to use $\hat{L}(t)$ in (5.3) to approximate the log-likelihood $L(t)$ in the time interval $0 \leq t < \min\{d, d_1\}$.

We discussed in §4.2.1 that the detection of a persistent signal is best if there are many activations and deactivations when the persistent signal is ON. Fortunately, this is exactly the condition required for the second property to hold. Note that in the analysis of the ideal C1-FFL in [2,4,7] and in §2.1, both the

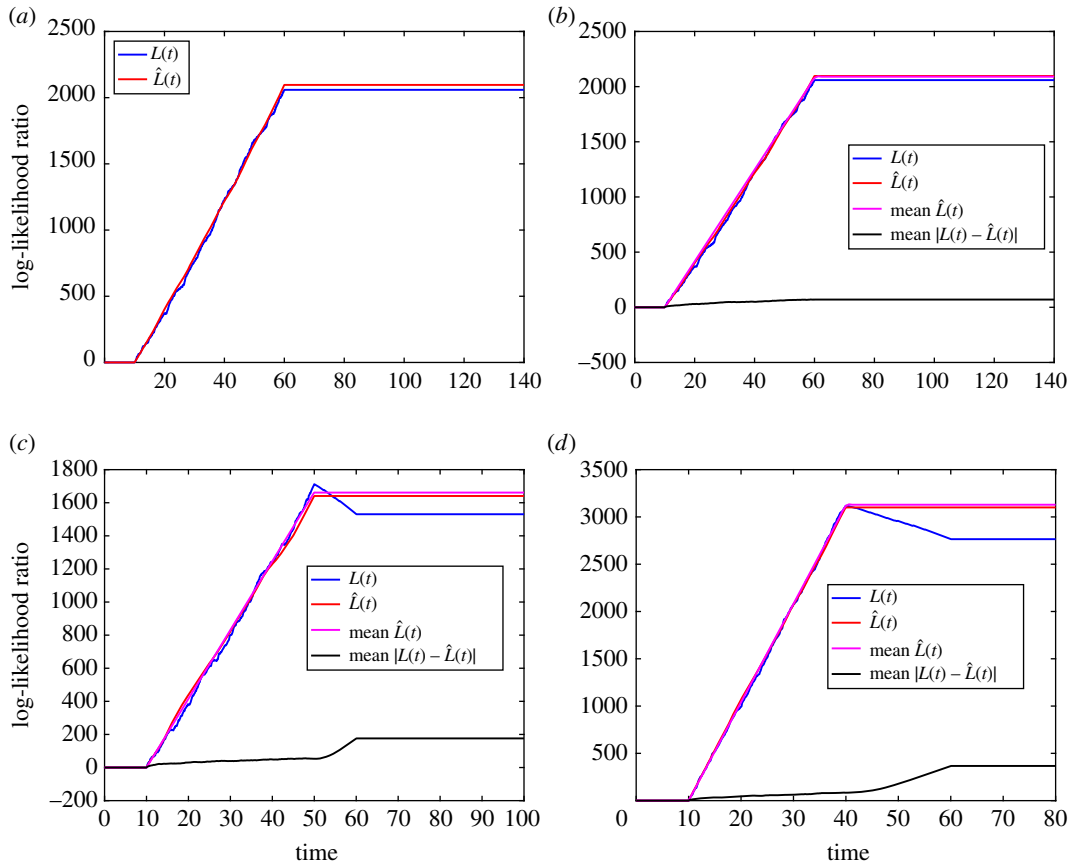


Figure 5. Numerical results for intermediate approximation.

activation and deactivation reactions (4.1) are assumed to be instantaneous, which can be viewed as k_+ and k_- being very large. This assumption can be justified from the fact that for C1-FFL, the molecule species S and X can be considered to be, respectively, an inducer and a transcription factor. It is known that the activation and deactivation dynamics of transcription factors are fast, see [7, Table 2.1]. Hence this assumption is not stringent and we will assume that reactions (4.1) are fast for the rest of this paper.

We remark that the second property does not cover all the persistent signals. For example, signals with a small amplitude a which do not produce a large enough number of activations and inactivations are not covered. These signals are persistent but are hard to detect.

At the beginning of this section, we mentioned some difficulties in realizing the computation of $L(t)$ in equation (4.3) using chemical reactions. We note that those difficulties are no longer present in the computation of $\hat{L}(t)$ using (5.3). In particular, we note that $\hat{L}(t)$ is always non-negative and can be interpreted as a log-likelihood ratio when the input is persistent.

5.2.1. Numerical illustration

We will now use a few numerical examples to illustrate that the intermediate approximation $\hat{L}(t)$ is approximately equal to the log-likelihood ratio $L(t)$ for persistent signals. For all these examples, we choose $k_+ = 0.02$, $k_- = 0.5$, $d_0 = 5$, $d_1 = 60$, $a_0 = 0.25$ and $a_1 = 10.7$.

For the first example, we choose $d = 70$ and $a = a_1$ for the input signal $s(t)$. We use the Stochastic Simulation Algorithm to obtain a realization of $x_*(t)$. We then use $x_*(t)$ to compute $L(t)$ and $\hat{L}(t)$. The results are shown in figure 5a. We can see that the approximation is good. We next generate 100 different realizations of $x_*(t)$ and use them to compute $L(t)$ and $\hat{L}(t)$. Figure 5b shows the mean of $|L(t) - \hat{L}(t)|$ over 100 realizations, as well as one realization of $L(t)$ and $\hat{L}(t)$. It can be seen that the approximation error is small. In figure 5b, we have also plotted the mean of $\hat{L}(t)$ obtained by solving the following system of ODEs:

$$\frac{d\bar{x}_*(t)}{dt} = k_+s(t)(M - \bar{x}_*(t)) - k_- \bar{x}_*(t) \quad (5.7)$$

and

$$\frac{d\bar{L}(t)}{dt} = \bar{x}_*(t) \times k_- \pi(t) [\phi(s(t))]_+, \quad (5.8)$$

where $\bar{x}_*(t)$ and $\bar{L}(t)$ are, respectively, the mean of $x_*(t)$ and $\hat{L}(t)$. It can be seen that a realization of $\hat{L}(t)$ is comparable to its mean.

We repeat the numerical experiment for $d = 40$ and $a = a_1$. Figure 5c shows a realization of $L(t)$, a realization of $\hat{L}(t)$, mean of $|L(t) - \hat{L}(t)|$ over 100 realizations, as well as the mean of $\hat{L}(t)$. We can see the approximation holds up till time $t = 40$, which is $\min\{d, d_1\}$. The purpose of this example is to illustrate why we need to include the condition $t \leq \min\{d, d_1\}$. This is because $L(t)$ and $\hat{L}(t)$ behave differently for $t > \min\{d, d_1\}$ if $d < d_1$. For $L(t)$, it falls after $t = 40$ because from this time onwards, the input signal $s(t)$ being used is small; this leads to a small number of activations and consequently a negative RHS for equation (4.3). However, for $\hat{L}(t)$, the RHS of equation (5.3) is zero because a small $s(t)$ makes $[\phi(s(t))]_+$ zero.

We have so far used $a = a_1$ and two different durations d . We now illustrate that the approximation holds for a different amplitude a . For the next numerical experiments, we keep $d = 40$ and use $a = 37.5$. The results are shown in figure 5d. We can see the approximation holds up till time $\min\{d, d_1\}$.

These examples demonstrate that, for persistent signals, the approximation $\hat{L}(t) \approx L(t)$ holds for different values of input duration d and amplitude a .

We also want to point out that the behaviour of $\hat{L}(t)$ for transient and persistent signals is consistent with that of the ideal C1-FFL discussed in §2.1. We have already pointed out that this is true for transient signals. For persistent signals, $\hat{L}(t)$ is zero initially and then followed by a non-zero output, i.e. there is a delay before $\hat{L}(t)$ becomes positive and this also holds for the ideal C1-FFL: see the bottom plot in figure 2. We will now map the intermediate approximation equation (5.3) to the reaction rate equations of a C1-FFL.

Remark 5.2. We want to remark that in the above formulation and numerical examples, the input signal $s(t)$ is allowed to differ from the two reference signals $c_0(t)$ and $c_1(t)$. Since the decision of the detection problem is based on the log-likelihood ratio in equation (4.2), we can interpret the detection problem as using the history $\mathcal{X}_*(t)$ (which is generated by $s(t)$) to decide which of the two signals, $c_0(t)$ or $c_1(t)$, is more likely to have produced the observed history. Furthermore, consider the case that $s(t)$ is parameterized by positive parameters a and d as in (5.2), then it can be shown that a small change in a or d will produce a small change in the mean of $L(t)$ and $\hat{L}(t)$.

5.3. Using C1-FFL to approximately compute $\hat{L}(t)$

The aim of this section is to show that the C1-FFL can be used to approximately compute the intermediate approximation $\hat{L}(t)$ in equation (5.3). Recall that the C1-FFL in figure 1a transforms the signal $x_*(t)$ into the output signal $z(t)$ using the the following components: Nodes Y and Z, and the AND logic. We will model these components using the following chemical reaction system:

$$\frac{dy(t)}{dt} = \frac{h_y x_*(t)^{n_y}}{\underbrace{K_y^{n_y} + x_*(t)^{n_y}}_{H_y(x_*(t))}} - d_y y(t) \quad (5.9a)$$

and

$$\frac{dz(t)}{dt} = x_*(t) \times \frac{h_z y(t)^{n_z}}{\underbrace{K_z^{n_z} + y(t)^{n_z}}_{H_z(y(t))}}, \quad (5.9b)$$

where h_y , n_y , K_y , etc. are coefficients of the Hill functions. We assume that the initial conditions are $y(0) = z(0) = 0$. Note that these two equations are comparable to the ideal C1-FFL model in §2.1. In particular, if we replace the θ -function in (2.1) by a Hill function, then it becomes (5.9a). Also, if we choose $K_{xz} = 0$ and $\alpha_z = 0$, and replace the θ -function in $z(t)$ by a Hill function in (2.2), then it becomes (5.9b).

By comparing the RHSs of the equations (5.3) and (5.9b), we see that the intermediate approximation $\hat{L}(t)$ and the output of the C1-FFL $z(t)$ can be made approximately equal if $k_- \pi(t) [\phi(s(t))]_+ (= \eta(t))$ in (5.3) and $H_z(y(t))$ in (5.9b) are approximately equal. We argue in appendix A 3 that it is possible to choose the parameters in (5.9) such that $\eta(t) \approx H_z(y(t))$ in the time interval $[0, \min\{d, d_1\})$. The argument consists of two parts, for the two time intervals $[0, d_0)$ and $[d_0, \min\{d, d_1\})$.

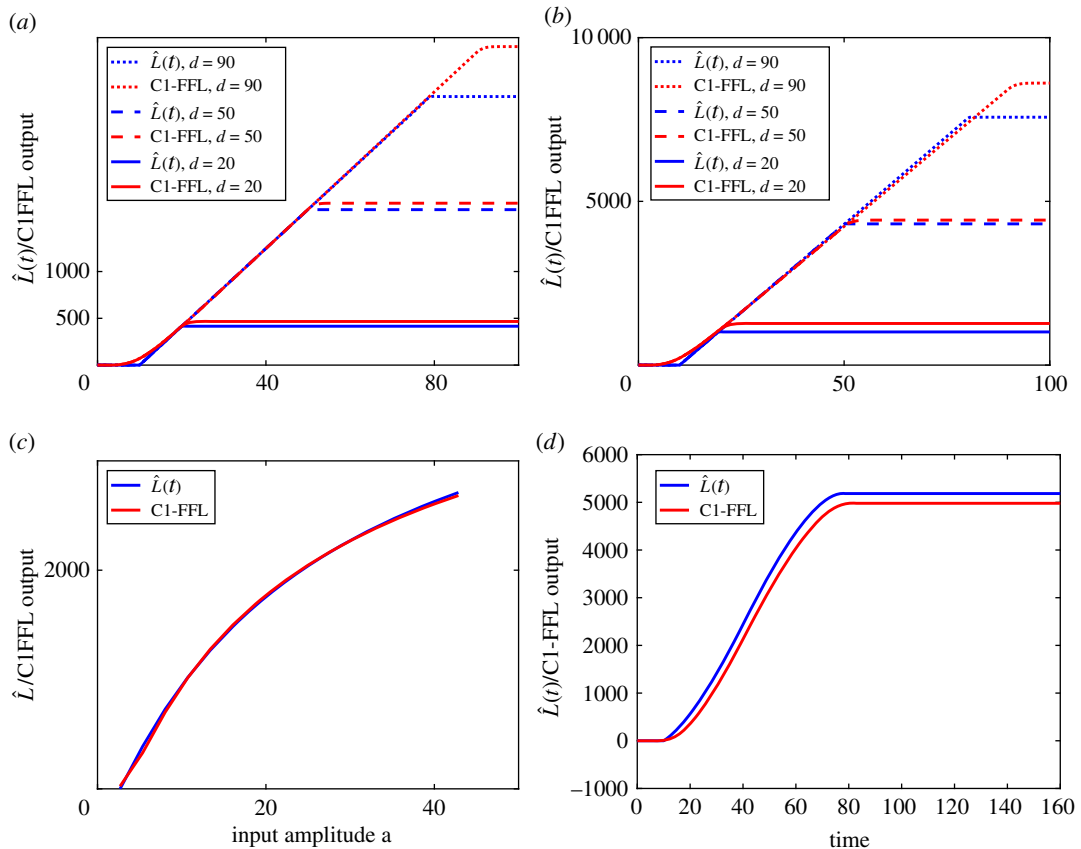


Figure 6. Numerical results on C1-FFL. (a) Comparing $\hat{L}(t)$ to C1-FFL output, $a = 5.4$; (b) comparing $\hat{L}(t)$ to C1-FFL output, $a = 40.2$; (c) comparing $\hat{L}(t)$ and C1-FFL output for different pulse input amplitudes; (d) comparing $\hat{L}(t)$ and C1-FFL output for a triangular pulse.

A major argument made in appendix A3 is to match $\eta(t)$ and $H_z(y(t))$ in the time interval $[d_0, \min\{d, d_1\})$ for persistent signals. We show in appendix A3 that this matching problem can be reduced to choosing the parameters in (5.9) so that the following two functions in a : $k_-[\phi(a)]_+$ and $H_z((1/d_y)H_y(Mk_+a/(k_+a + k_-)))$ are approximately equal for a large range of a where a , as defined in §5.2, is the amplitude of the input $s(t)$ when it is ON. We note in appendix A3 why these two functions in a can fit to each other.

5.4. Numerical examples

We now present numerical examples to show that C1-FFL can be used to compute $\hat{L}(t)$. We use the same k_+ , k_- , M , a_0 and a_1 values as in §5.2.1. We choose $d_0 = 10$ and $d_1 = 80$. We use parameter estimation to determine the parameters in equation (5.9) so that the C1-FFL output $z(t)$ matches $\hat{L}(t)$ for a range of a . The estimated parameters for the C1-FFL are: $h_y = 1.01$, $K_y = 8.04$, $n_y = 2.26$, $d_y = 0.24$, $h_z = 10.6$, $n_z = 5.84$ and $K_z = 5.43$. In this section, we will compare $\hat{L}(t)$ from (5.3) with $z(t)$ from (5.9) assuming the $x_*(t)$ in these two equations is given by $\bar{x}_*(t)$ in (5.7).

Figure 6a compares $\hat{L}(t)$ and $z(t)$ for input $s(t)$ with $a = 5.4$ and three different durations $d = 10, 30$ and 70 . When $d = 10$, the output of the C1-FFL is small. For $d = 30$ and 70 , the C1-FFL output matches well with $\hat{L}(t)$. To show that the match is also good for a different value of the input amplitude a , we show the results for $a = 40.2$ and $d = 20, 40$ and 90 in figure 6b. For the case of $d = 90$, we see that the match is good till $[0, d_1)$ because $d > d_1$.

We have demonstrated that $z(t)$ matches $\hat{L}(t)$ for two different values of a . We can show that the match is good for a large range of a . We fix the duration d to be 70 but vary the amplitude a from 2.7 to 85.7 . Figure 6c compares $\hat{L}(t)$ and $z(t)$ at $t = 70$. It can be seen that the C1-FFL approximation works for a wide range of a .

The previous examples show that we can match the C1-FFL output $z(t)$ to the intermediate approximation $\hat{L}(t)$ for pulse input $s(t)$ of different durations and amplitudes. We can also show that

the match extends to slowly-varying inputs. In this example, we assume $s(t)$ is a triangular pulse with $s(0) = 0$ and rises linearly to $s(40) = 42.8$ and then decreases linearly to $s(80) = 0$. Figure 6d shows the time responses $z(t)$ and $\hat{L}(t)$, and they are comparable.

Remark 5.3. We finish this section by making a number of remarks.

- Note that we have not included the degradation of Z in (5.9b) so that we can match it to $\hat{L}(t)$, which does not decay. It can be shown that if we add a degradation term $-\alpha z(t)$ to the RHS of (5.9b) and $-\alpha \hat{L}(t)$ to the RHS of equation (5.3), the resulting $z(t)$ will still be matched to $\hat{L}(t)$.
- Equation (5.9b) is not the most general form of C1-FFL. In the general form of equation (5.9b), which is presented in [4], the factor $x(t)$ is replaced by a Hill function of $x(t)$. We conjecture that it is possible to generalize the methodology in this paper to obtain the general case and we leave it as future work.
- The intermediate approximation $\hat{L}(t)$ is derived under the assumption that $s(t)$ is a rectangular pulse. Future work is needed to better understand the behaviour of the intermediate approximation $\hat{L}(t)$ when this assumption does not hold.
- Although we have shown that the approximate positive log-likelihood ratio in equation (5.3) can be computed by a C1-FFL, it is certainly not true that any C1-FFL can be used to realize equation (5.3). This can be seen from the fact that the C1-FFL model in (5.9) has seven parameters while the log-likelihood ratio calculation in equation (5.3) has only four parameters. A research question is whether any C1-FFL that can detect persistent signals has a corresponding log-likelihood ratio detector equation (5.3). We can answer this question by first characterizing the C1-FFL that can detect persistent signals and check whether such a correspondence exists. This is an open research problem to be addressed.
- We have so far assumed that $c_0(t)$ and $c_1(t)$ are strictly positive for all t by assuming that $a_0 > 0$. If $a_0 = 0$, then the log-likelihood ratio is no longer well defined because both (4.2) and (5.3) diverge. However, we can compute a shifted and scaled version of the log-likelihood ratio whose intermediate approximation for persistent signals is:

$$\frac{d\hat{L}(t)}{dt} = x_+(t) \times \pi(t).$$

It is still possible to use this intermediate approximation to detect persistent signals. This intermediate approximation can also be approximated by a C1-FFL. Details are omitted and will be studied in future work.

6. Conclusion and discussion

In this paper, we study the persistence detection property of C1-FFL from an information processing point of view. We formulate a detection problem on a chemical reaction cycle to understand how an input signal of a long duration can be distinguished from one of short duration. We solve this detection problem and derive an ODE which describes the time evolution of the log-likelihood ratio. An issue with this ODE is that it is difficult to realize it using chemical reactions. We then use time-scale separation to derive an ODE which can approximately compute the log-likelihood ratio when the input signal is persistent. We further show that this approximate ODE can be realized by a C1-FFL. It also provides an interpretation of the persistence detection property of C1-FFL as an approximate computation of the log-likelihood ratio.

The concept of the log-likelihood ratio (or a similar quantity) has been used to understand how cells make a decision in [9,10]. The paper [10] considers the problem of distinguishing between two environment states, which are the presence and absence of stimulus. It derives an ODE of the log-odds ratio and uses the ODE to deduce a biochemical network implementation in the form of a phosphorylation-dephosphorylation cycle. In this cycle, the fraction of phosphorylated substrate is the *posteriori* probability of the presence of stimuli. The paper [9] considers the problem of distinguishing between two different levels of concentration using the likelihood ratio. It also presents a molecular implementation that computes the likelihood ratio. This paper differs from [9,10] in one major way. We make a crucial approximation by considering only the positive log-likelihood ratio and ignoring the negative log-likelihood ratio. We are then able to connect the computation of the positive log-likelihood ratio with the computation carried out by a C1-FFL. This work therefore provides a connection between detection theory and C1-FFL using the positive log-likelihood ratio as the connecting point.

The computation of the positive log-likelihood ratio by C1-FFL, which is the key finding of this paper, is an example of using biochemical networks to perform the analog computation. There are a few other examples. The incoherent type-1 feed-forward loop, which is another network motif, is found to be able to compute fold change [14]. Allosteric protein is found to be able to compute logarithm approximately [15]. In addition, there is also work on using synthetic biochemical circuits to do analog computation [16,17].

In this paper, we use a methodology which is based on three key ingredients—statistical detection theory, time-scale separation and analog molecular computation—to derive a molecular network that can be used to discriminate persistent signals from transient ones. A possible application of the methodology of this paper in molecular biology is to derive the molecular networks that can decode temporal signals. According to the review paper on temporal signals in cell signalling [18], only some of the molecular networks for decoding temporal signals have been identified. In fact, the authors of [18] went further to state that ‘Identifying the mechanisms that decode dynamics remains one of the most challenging goals for the field.’ In [19], we used a methodology—which is similar to the one used in this paper and is based on the same three key ingredients — to derive a molecular network to decode concentration modulated signals. The derived molecular network was found to be consistent with the *Saccharomyces cerevisiae* DCS2 promotor data in [20], which were obtained from exciting the promotor by using various transcription factor dynamics, e.g. concentration modulation, duration modulation and others. Another possible application of the methodology of this paper is in synthetic biology. For example, in [21] we used a methodology—which is similar to the one used in this paper and in [19]—to derive a *de novo* molecular network for decoding concentration modulated signals. We remark that the molecular networks in [19,21] can be interpreted as an approximate log-likelihood detector of concentration modulated signals.

A recent report [22] considers the problem of determining the biochemical circuits that can be used to distinguish between a persistent and a transient signal. By searching over all biochemical circuits with a limited complexity, the authors find that there are five different circuits that can be used. One of these is C1-FFL. An open question is whether one can use the framework in this paper to deduce all circuits that can detect persistent signals. If this is possible, then it presents an alternative method to find the biochemical circuits that can realize a function.

Data accessibility. The source code for producing the results for this paper is available at Github, which is an open online code repository. The source code is at <https://github.com/ctchou-unsu/c1ffl-journal> and <https://doi.org/10.5061/dryad.20md774>.

Author’s contributions. C.T.C. designed the research, performed the mathematical analysis, wrote the simulation code and the paper.

Competing interests. The author declares no competing interests.

Funding. This work was not supported by any research grants.

Acknowledgements. The author wishes to thank Dr Guy-Bart Stan, Imperial College, for the suggestion to consider possible connections with motifs.

Appendix A. Proof and derivation

A.1. Proof of (4.3)

Recalling that $\mathcal{X}_*(t)$ is the history of $x_*(t)$ in the time interval $[0, t]$. In order to derive (4.3), we consider the history $\mathcal{X}_*(t + \Delta t)$ as a concatenation of $\mathcal{X}_*(t)$ and $x_*(t)$ in the time interval $(t, t + \Delta t]$. We assume that Δt is chosen small enough so that no more than one activation or deactivation reaction can take place in $(t, t + \Delta t]$. Given this assumption and right continuity of continuous-time Markov Chains, we can use $x_*(t + \Delta t)$ to denote the history of $x_*(t)$ in $(t, t + \Delta t]$.

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

$$P[\mathcal{X}_*(t + \Delta t) | \mathcal{H}_i] \tag{A 1}$$

$$= P[\mathcal{X}_*(t) \text{ AND } x_*(t + \Delta t) | \mathcal{H}_i] \tag{A 2}$$

$$= P[\mathcal{X}_*(t) | \mathcal{H}_i] P[x_*(t + \Delta t) | \mathcal{H}_i, \mathcal{X}_*(t)] \tag{A 3}$$

$$= P[\mathcal{X}_*(t) | \mathcal{H}_i] P[x_*(t + \Delta t) | \mathcal{H}_i, x_*(t)] \tag{A 4}$$

where we have expanded $\mathcal{X}_*(t + \Delta t)$ in equation (A 1) using concatenation and used Markov property to go from equation (A 3) to equation (A 4).

By using (A 4) in the definition of the log-likelihood ratio, we can show that:

$$L(t + \Delta t) = L(t) + \log \left(\frac{P[x_*(t + \Delta t) | \mathcal{H}_1, x_*(t)]}{P[x_*(t + \Delta t) | \mathcal{H}_0, x_*(t)]} \right) \quad (\text{A } 5)$$

The value of the expression $P[x_*(t + \Delta t) | \mathcal{H}_i, x_*(t)]$ depends on whether $x_*(t + \Delta t)$ is one greater than, one less than or equal to $x_*(t)$. These cases correspond, respectively, to the event that an X molecule had been activated, an X_* molecule had been deactivated, or there had been no change in the state of the molecules in the time interval $(t, t + \Delta t]$. Under the hypothesis \mathcal{H}_i , which means the input signal is assumed to be $c_i(t)$, the activation and deactivation rates are, respectively, $k_+(M - x_*(t))c_i(t)$ and $k_-x_*(t)$ when the number of X_* molecules is $x_*(t)$. We can therefore write the expression $P[x_*(t + \Delta t) | \mathcal{H}_i, x_*(t)]$ as:

$$\begin{aligned} P[x_*(t + \Delta t) | \mathcal{H}_i, x_*(t)] &= \delta_{x_*(t+\Delta t), x_*(t)+1} k_+(M - x_*(t)) c_i(t) \Delta t \\ &\quad + \delta_{x_*(t+\Delta t), x_*(t)-1} k_- x_*(t) \Delta t \\ &\quad + \delta_{x_*(t+\Delta t), x_*(t)} (1 - k_+(M - x_*(t)) c_i(t) \Delta t - k_- x_*(t) \Delta t), \end{aligned} \quad (\text{A } 6)$$

where $\delta_{a,b}$ is the Kronecker delta which is 1 when $a = b$.

Note that $P[x_*(t + \Delta t) | \mathcal{H}_i, x_*(t)]$ in (A 6) is a sum of three terms with multipliers $\delta_{x_*(t+\Delta t), x_*(t)+1}$, $\delta_{x_*(t+\Delta t), x_*(t)-1}$ and $\delta_{x_*(t+\Delta t), x_*(t)}$. Since these multipliers are mutually exclusive, we have:

$$\begin{aligned} &\log \left(\frac{P[x_*(t + \Delta t) | \mathcal{H}_1, x_*(t)]}{P[x_*(t + \Delta t) | \mathcal{H}_0, x_*(t)]} \right) \\ &= \delta_{x_*(t+\Delta t), x_*(t)+1} \log \left(\frac{k_+(M - x_*(t)) c_1(t) \Delta t}{k_+(M - x_*(t)) c_0(t) \Delta t} \right) \\ &\quad + \delta_{x_*(t+\Delta t), x_*(t)-1} \log \left(\frac{k_- x_*(t) \Delta t}{k_- x_*(t) \Delta t} \right) \\ &\quad + \delta_{x_*(t+\Delta t), x_*(t)} \log \left(\frac{1 - k_+(M - x_*(t)) c_1(t) \Delta t - k_- x_*(t) \Delta t}{1 - k_+(M - x_*(t)) c_0(t) \Delta t - k_- x_*(t) \Delta t} \right) \\ &\approx \delta_{x_*(t+\Delta t), x_*(t)+1} \log \left(\frac{c_1(t)}{c_0(t)} \right) \\ &\quad - \delta_{x_*(t+\Delta t), x_*(t)} k_+(M - x_*(t)) (c_1(t) - c_0(t)) \Delta t, \end{aligned} \quad (\text{A } 7)$$

where we have used the approximation $\log(1 + f\Delta t) \approx f\Delta t$ to obtain (A 7). Note also that the above derivation requires that both $c_0(t)$ and $c_1(t)$ must be strictly positive for all t .

By substituting equation (A 7) into equation (A 5), we have after some manipulations and after taking the limit $\Delta t \rightarrow 0$:

$$\begin{aligned} \frac{dL(t)}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{\delta_{x_*(t+\Delta t), x_*(t)+1}}{\Delta t} \log \left(\frac{c_1(t)}{c_0(t)} \right) \\ &\quad - \delta_{x_*(t+\Delta t), x_*(t)} k_+(M - x_*(t)) (c_1(t) - c_0(t)). \end{aligned} \quad (\text{A } 8)$$

In order to obtain equation (4.3), we use the following reasoning. First, the term $\lim_{\Delta t \rightarrow 0} (\delta_{x_*(t+\Delta t), x_*(t)+1} / \Delta t)$ is a Dirac delta at the time instant that an X molecule is activated. Second, the term $\delta_{x_*(t+\Delta t), x_*(t)}$ is only zero when the number of X_* molecule changes but the number of such changes is countable. In other words, $\delta_{x_*(t+\Delta t), x_*(t)} = 1$ with probability one. This allows us to drop $\delta_{x_*(t+\Delta t), x_*(t)}$. Hence equation (4.3).

A.2. Derivation of (5.3)

The aim of this appendix is to derive the intermediate approximation (5.3). We will split the derivations into two parts, depending on the length of the duration d relative to d_0 . We first consider the case where the input signal $s(t)$ has a duration longer than or equal to d_0 , which is also the more important case for the derivation because it deals with persistent signals.

Our aim is to find an approximation of the log-likelihood ratio $L(t)$ given in (4.3). Our strategy is to divide time into intervals such that, in each time interval, each of the time profiles of $c_0(t)$, $c_1(t)$ and $s(t)$ is a constant.

The first time interval is $[0, d_0)$ where $c_0(t) = c_1(t) = a_1$ and $s(t) = a$. Since $L(0) = 0$ and $dL(t)/dt = 0$ in this time interval, therefore $L(t) = 0$ in this time interval.

The next time interval to consider is $[d_0, \min\{d, d_1\})$ where $c_0(t) = a_0$, $c_1(t) = a_1$ and $s(t) = a$. For $t \in [d_0, \min\{d, d_1\})$, the log-likelihood ratio $L(t)$ in (4.3) can be written as $L(t) = L_1(t) + L_2(t)$ where

$$L_1(t) = \log\left(\frac{a_1}{a_0}\right) \underbrace{\int_{d_0}^t \left[\frac{dx_*(t)}{dt} \right]_+ dt}_{A(t)} \quad (\text{A } 9)$$

and

$$L_2(t) = -k_+(a_1 - a_0) \int_{d_0}^t (M - x_*(t)) dt. \quad (\text{A } 10)$$

We first consider finding an approximation of the integral $A(t)$ in (A 9), and the aim is to replace the positive derivative of $x_*(t)$ by some other arithmetic operations which can be computed by using chemical reactions. The integral $A(t)$ can be interpreted as the number of times that X is activated in the time interval $[d_0, t)$. For an X molecule, the time between two consecutive activations is a random variable with mean m and variance σ^2 where:

$$m = \frac{1}{k_+a} + \frac{1}{k_-} \quad (\text{A } 11)$$

and

$$\sigma^2 = \frac{1}{(k_+a)^2} + \frac{1}{k_-^2}. \quad (\text{A } 12)$$

This is because we can model the activation and deactivation of X by a two-state continuous-time Markov chain with transition rates k_+a and k_- .

We will now make a time-scale separation assumption by assuming that the duration $(t - d_0)$ is much bigger than m , i.e. $t - d_0 \gg 1/k_+a + 1/k_-$. This assumption can be met by having a sufficiently long duration d and large amplitude a . If this time-scale separation assumption holds, then there are many activations in the time interval $[d_0, t)$. In this case, using the renewal theorem [23] to approximate $A(t)$, we have:

$$\text{mean}(A(t)) \approx M \frac{t - d_0}{m} \quad (\text{A } 13)$$

and

$$\text{var}(A(t)) \approx M \frac{\sigma^2}{m^3} (t - d_0), \quad (\text{A } 14)$$

which implies that

$$\frac{\sqrt{\text{var}(A(t))}}{\text{mean}(A(t))} \approx \frac{\sigma}{m\sqrt{M}\sqrt{t - d_0}}. \quad (\text{A } 15)$$

This means we can approximate $A(t)$ by its mean and the error decreases with the reciprocal of the square root of the duration $t - d_0$. By using this approximation, we have:

$$L_1(t) \approx \log\left(\frac{a_1}{a_0}\right) \frac{M}{m} (t - d_0). \quad (\text{A } 16)$$

The time-scale separation assumption also implies that the ensemble average of $x_*(t)$ can be treated as a constant in the time interval $[d_0, t)$; we will denote this average by $x_{*,a}$ where

$$x_{*,a} = \frac{Mk_+a}{k_+a + k_-}. \quad (\text{A } 17)$$

This ensemble average is related to mean inter-activation time m in (A 11) by:

$$x_{*,a} = \frac{M}{mk_-}. \quad (\text{A } 18)$$

By using this relationship in (A 16), we have:

$$L_1(t) \approx k_- x_{*,a} \log\left(\frac{a_1}{a_0}\right)(t - d_0), \quad (\text{A } 19)$$

which means $L_1(t)$ can be computed from the ensemble average $x_{*,a}$. We will return to this expression shortly after studying the approximation of the integral in $L_2(t)$ in (A 10).

Since the Markov chain describing the reaction cycle of X and X_* is ergodic, the time average in (A 10) can be approximated by its ensemble average. By using the ensemble average $x_{*,a}$ in (A 17), we can show that:

$$L_2(t) \approx -k_- x_{*,a} \frac{(a_1 - a_0)}{a} (t - d_0). \quad (\text{A } 20)$$

Since $L(t) = L_1(t) + L_2(t)$, it follows from (A 19) and (A 20) that:

$$L(t) \approx k_- x_{*,a} \left(\log\left(\frac{a_1}{a_0}\right) - \frac{(a_1 - a_0)}{a} \right) (t - d_0), \quad (\text{A } 21)$$

in the time interval $[d_0, \min\{d, d_1\})$.

We can re-write the results that we have for the time interval $[0, \min\{d, d_1\})$ in differential form, as follows:

$$\frac{dL(t)}{dt} \approx k_- x_*(t) \left\{ \log\left(\frac{c_1(t)}{c_0(t)}\right) - \frac{(c_1(t) - c_0(t))}{s(t)} \right\}, \quad (\text{A } 22)$$

The derivation so far has shown that the ODEs (A 22) and (4.3) are approximately equal for t in $[0, \min\{d, d_1\})$. We will consider $t \geq \min\{d, d_1\}$. We need to split this into two cases: $d \geq d_1$ and $d < d_1$. For the first case, the time interval concerned is $t \geq d$. It can be verified that the RHSs of (A 22) and (4.3) are both zero for this time interval. Thus, if $d \geq d_1$, then $\hat{L} \approx L(t)$ for all t . We will consider the second case, where $d < d_1$, in the next paragraph.

If $d < d_1$, then the time interval $[d, d_1)$ is non-empty. In this interval, we have $s(t) = a_0$, $c_0(t) = a_0$ and $c_1(t) = a_1$, which means the term in curly brackets in (A 22) is equal to $\log(a_1/a_0) - (a_1 - a_0)/a_0$. Since $a_1 > a_0$, this term is negative. As a result, this may result in a negative $L(t)$. Although we learn from the research on a synthetic analog computation using chemical reactions [13] that it is possible to handle negative numbers using chemical reactions, the research also tells us that this is inherently a difficult process and the complexity is high. Therefore, we will use an approximation that does not result in a negative log-likelihood ratio. By adding the $[\]_+$ operator, where $[w]_+ = \max(w, 0)$, to the term in curly brackets in (A 22), we arrive at:

$$\frac{d\hat{L}(t)}{dt} \approx k_- x_*(t) \left\{ \left[\log\left(\frac{c_1(t)}{c_0(t)}\right) - \frac{(c_1(t) - c_0(t))}{s(t)} \right]_+ \right\}. \quad (\text{A } 23)$$

The addition of the $[\]_+$ operator does not affect what happens in the time interval $[0, \min\{d, d_1\})$. However, it means that the RHS of (A 23) does not equal the RHS of (4.3) in the time interval $[d, d_1)$; in fact, this is the only time interval when the RHSs of (A 23) and (4.3) are not approximately equal. This also means that, for input signals whose duration d is less than d_1 , the approximation $\hat{L}(t) \approx L(t)$ only holds in the time interval $[0, \min\{d, d_1\})$.

Our next step is to show that (A 23) can be written as (5.3). By using the form of $c_0(t)$ and $c_1(t)$, we can show that

$$\log\left(\frac{c_1(t)}{c_0(t)}\right) = \log\left(\frac{a_1}{a_0}\right) \pi(t) \quad (\text{A } 24)$$

and

$$c_1(t) - c_0(t) = (a_1 - a_0) \pi(t). \quad (\text{A } 25)$$

By substituting (A 24) and (A 25) into (A 23), we arrive at

$$\frac{d\hat{L}(t)}{dt} = x_*(t) \times \left\{ k_- \pi(t) \left[\log\left(\frac{a_1}{a_0}\right) - \frac{a_1 - a_0}{s(t)} \right]_+ \right\}, \quad (\text{A } 26)$$

which is the same as (5.3).

We conclude the derivation of $\hat{L}(t)$ by showing that $\hat{L}(t) = 0$ for all t for input signals $s(t)$ whose duration d is less than d_0 . This can be done by showing that the RHS of (5.3) is zero for all t . Since $\pi(t)$ is only non-zero in the time interval $[d_0, d_1)$, we only have to consider this time interval. In this time interval, we can show that $[\log(a_1/a_0) - (a_1 - a_0)/s(t)]_+$ is zero because $s(t) = a_0$.

A.3. Matching (5.3) to (5.9)

The aim of this appendix is to explain why it is possible to use the C1-FFL system in (5.9) to realize the intermediate approximation in (5.3). By comparing the RHSs of (5.3) and (5.9b), our aim is to show that, by using an appropriate choice of parameters in (5.9), $k_- \pi(t) [\phi(s(t))]_+ (= \eta(t))$ and $H_z(y(t))$ can be made to be approximately equal in the time interval $[0, \min\{d, d_1\})$. We will consider the time intervals $[0, d_0)$ and $[d_0, \min\{d, d_1\})$ separately.

We first consider the time interval $[d_0, \min\{d, d_1\})$. It is sufficient to consider only persistent input signals. Within this time interval, the persistent input $s(t)$ has an amplitude of a . Since we assume that the input $s(t)$ is long compared to the time-scale of the activation and deactivation reactions, therefore the mean of $x_*(t)$ is a plateau (see the bottom plot of figure 4b) whose height is $Mk_+a/(k_+a + k_-)$. Consequently, the time profiles of both $\eta(t)$ and $y(t)$ also contain a period of time that they plateau. The plateau in $\eta(t)$ contributes to the ramp-like increase in $\hat{L}(t)$ in figure 5.

This means that, if we want to match (5.3) and (5.9) in the time interval $[d_0, \min\{d, d_1\})$, we need to match the values of $\eta(t)$ and $H_z(y(t))$ at their plateau. Since the amplitude of the input $s(t)$ when it is ON is a , the heights of the plateau of $\eta(t)$ and $H_z(y(t))$ are, respectively, $k_- [\phi(a)]_+ (= f_1(a))$ and $H_z((1/d_y)H_y(Mk_+a/(k_+a + k_-))) (= f_2(a))$, and we want $f_1(a) \approx f_2(a)$ for as large a range of a as possible. Note that for all a such that $f_1(a) > 0$, both functions $f_1(a)$ and $f_2(a)$ are strictly increasing, and both $f_1(\infty)$ and $f_2(\infty)$ are constants. Therefore, we can choose the Hill function coefficients to fit $f_2(a)$ to $f_1(a)$. This argument takes care of the case when $s(t)$ is a persistent input which requires us to implement the function $\phi(\cdot)$ in (5.3) using the Hill functions in (5.9). We remark that we need to include the requirement $f_1(a) > 0$ in the above argument because $f_1(a)$ is not strictly increasing when $f_1(a) = 0$. This can be seen from the fact that there is a range of a such that $\phi(a) < 0$, which means that there is a range of a such that $f_1(a) = 0$, which in turn means that $f_1(a)$ is not monotonically increasing in this range.

We now consider the time interval $[0, d_0)$. In this time interval, $\hat{L}(t) = 0$ due to $\pi(t)$. This is a feature shared by the ideal C1-FFL model in [7]. The book [7] shows that this can be realized by choosing a big enough K_z in equation (5.9b) so that the production rate of $z(t)$ is small initially.

References

- Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. 2002 Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827. (doi:10.1126/science.298.5594.824)
- Shen-Orr SS, Milo R, Mangan S, Alon U. 2002 Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* **31**, 64–68. (doi:10.1038/ng881)
- Alon U. 2007 Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450–461. (doi:10.1038/nrg2102)
- Mangan S, Alon U. 2003 Structure and function of the feed-forward loop network motif. *Proc. Natl Acad. Sci. USA* **100**, 11 980–11 985. (doi:10.1073/pnas.2133841100)
- Mangan S, Zaslaver A, Alon U. 2003 The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *J. Mol. Biol.* **334**, 197–204. (doi:10.1016/j.jmb.2003.09.049)
- Kay SM. 1998 *Fundamentals of statistical signal processing, volume II: detection theory*. Englewood Cliffs, NJ: Prentice Hall.
- Alon U. 2006 *An introduction to systems biology: design principles of biological circuits*. London, UK: Chapman & Hall.
- Gardiner C. 2010 *Stochastic methods*. Berlin, Germany: Springer.
- Siggia ED, Vergassola M. 2013 Decisions on the fly in cellular sensory systems. *Proc. Natl Acad. Sci. USA* **110**, E3704–E3712. (doi:10.1073/pnas.1314081110)
- Kobayashi TJ, Kamimura A. 2011 Dynamics of intracellular information decoding. *Phys. Biol.* **8**, 055007. (doi:10.1088/1478-3975/8/5/055007)
- Chou CT. 2015 Maximum a-posteriori decoding for diffusion-based molecular communication using analog filters. *IEEE Trans. Nanotechnol.* **14**, 1054–1067. (doi:10.1109/TNANO.2015.2469301)
- Gillespie D. 1977 Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361. (doi:10.1021/j100540a008)
- Oishi K, Klavins E. 2011 Biomolecular implementation of linear I/O systems. *Systems Biology, IET* **5**, 252–260. (doi:10.1049/iet-syb.2010.0056)
- Goentoro L, Shoval O, Kirschner MW, Alon U. 2009 The incoherent feedforward loop can provide fold-change detection in gene regulation. *Mol. Cell* **36**, 894–899. (doi:10.1016/j.molcel.2009.11.018)
- Olsman N, Goentoro L. 2016 Allosteric proteins as logarithmic sensors. *Proc. Natl Acad. Sci. USA* **113**, E4423–E4430. (doi:10.1073/pnas.1601791113)
- Daniel R, Rubens JR, Sarpeshkar R, Lu TK. 2013 Synthetic analog computation in living cells. *Nature* **497**, 619–623. (doi:10.1038/nature12148)
- Chou CT. 2017 Chemical reaction networks for computing logarithm. *Synth. Biol.* **2**, 1–13. (doi:10.1093/synbio/ysx002)
- Purvis JE, Lahav G. 2013 Encoding and decoding cellular information through signaling dynamics. *Cell* **152**, 945–956. (doi:10.1016/j.cell.2013.02.005)
- Chou CT. 2018 Designing molecular circuits for approximate maximum a posteriori demodulation of concentration modulated signals. (<http://arxiv.org/abs/1808.01543>)
- Hansen AS, O'Shea EK. 2013 Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene

- expression. *Mol. Syst. Biol.* **9**, 1–14. (doi:10.1038/msb.2013.56)
21. Chou CT. 2018 Molecular circuit for approximate maximum a posteriori demodulation of concentration modulated signals. In *The 5th ACM Int. Conf. on Nanoscale Computing and Communication*. New York, NY: ACM.
22. Gerardin J, Lim WA. 2016 The design principles of biochemical timers: circuits that discriminate between transient and sustained stimulation. *bioRxiv*. pp. 1–51.
23. Grimmett GR, Stirzaker DR. 1997 *Probability and random processes*. Oxford, UK: Oxford University Press.