Symbol-by-Symbol Maximum Likelihood Detection for Cooperative Molecular Communication

Yuting Fang, Student Member, IEEE, Adam Noel, Member, IEEE, Nan Yang, Senior Member, IEEE, Andrew W. Eckford, Senior Member, IEEE, and Rodney A. Kennedy, Fellow, IEEE

Abstract—In this paper, symbol-by-symbol maximum likelihood (ML) detection is proposed for a cooperative diffusion-based molecular communication (MC) system. In this system, the transmitter (TX) sends a common information symbol to multiple receivers (RXs) and a fusion center (FC) chooses the TX symbol that is more likely, given the likelihood of its observations from all RXs. The transmission of a sequence of binary symbols and the resultant intersymbol interference are considered in the cooperative MC system. Three ML detection variants are proposed according to different RX behaviors and different knowledge at the FC. The system error probabilities for two ML detector variants are derived, one of which is in closed form. The optimal molecule allocation among RXs to minimize the system error probability of one variant is determined by solving a joint optimization problem. Also for this variant, the equal distribution of molecules among two symmetric RXs is analytically shown to achieve the local minimal error probability. Numerical and simulation results show that the ML detection variants provide lower bounds on the error performance of simpler, non-ML cooperative variants and demonstrate that these simpler cooperative variants have error performance comparable to ML detectors.

Index Terms—Molecular communication, multi-receiver cooperation, symbol-by-symbol maximum likelihood detection, error performance, optimization

I. INTRODUCTION

Molecular communication (MC) has been heralded as one of the most promising paradigms to implement communication in bio-inspired nanonetworks, due to the potential benefits of bio-compatibility and low energy consumption [2]. In MC, the information transmission between devices is realized through the exchange of molecules. Since no source of external energy is required for free diffusion, it is the simplest molecular propagation mechanism. One of the primary challenges posed by diffusion-based MC is that its reliability rapidly decreases when the transmitter (TX)-receiver (RX) distance increases. A naturally-inspired approach, which also makes use of the envisioned collaboration between nanomachines, is allowing multiple RXs to share information for cooperative detection. Often, cells or organisms share common information to the FC and the FC chooses the TX symbol that is more likely, given the likelihood of its observations from a TX. Also, in [18, 19] the FC makes a single decision about the presence of an abnormality, such that there is only one information symbol and no symbol-by-symbol detection.

In this paper, we present symbol-by-symbol ML detection for a cooperative diffusion-based MC system, based on [8–10], which consists of one TX, K RXs, and an FC. The significance of this paper is that our results provide lower bounds on the error performance that can be achieved by the detectors considered in [8–10]. We consider relatively simple RXs with an energy detector or a signal amplifier. The computations required at the RXs can be implemented at the molecular level [20]. We keep the relatively high complexity required for ML detection at the FC. This is because the FC could have a direct interface with the macroscopic world and easier access to computational resources. In our proposed system, the transmission of each information symbol from the TX to the FC via the RXs is completed in two phases, as shown in Fig. 1. In the first phase, the TX sends a symbol that is observed by all RXs. In the second phase, the RXs send their detected information to the FC and the FC chooses the TX symbol that is more likely, given the likelihood of its observations from all RXs.

Since binary symbols are the easiest to transmit and detect [21], and we assume that the TX needs to send multiple
Multiple DF Single ML TX are more realistic in biological environments. ML detection in RXs outperforms AF [22] and multi-type outperforms single-type detection at the FC (SA-ML). We design the detectors according to different variants of ML detection at the FC (SD-ML), and 3) DF with single-molecule-type and ML detection at the FC (MD-ML), 2) DF with single-molecule-type and ML detection at the FC (SD-ML), and SA-ML are represented by red, blue, and green arrows, respectively.

Fig. 1. An example of a cooperative MC system with 2 RXs. The transmission from the TX to the RXs is represented by black dashed arrows. “D” and “A” denote the RXs making decisions and amplifying observations, respectively, and $A_k$ denotes the type of released molecule. The transmission from the RXs to the FC in MD-ML, SD-ML, and SA-ML are represented by red, blue, and green arrows, respectively.

| Acronym | Relaying at RXs | Molecule Type used in RXs | Behavior at FC | Complexity Comparison |
|---------|----------------|----------------------------|----------------|------------------------|
| MD-ML  | DF             | Multiple                   | ML Detection   | MD-ML > SD-ML > SA-ML  |
| SD-ML  | DF             | Single                     | ML Detection   |                        |
| SA-ML  | AF             | Single                     | ML Detection   |                        |

In this section, we present the system model (i.e., physical environment and general behaviors of devices) for the cooperative MC system and some preliminary results that are needed.

The current symbol interval requires knowing the previously-transmitted symbols by the TX (and by all RXs for DF). For convenience, we refer to the FC-estimated previous symbols as local history and the perfect knowledge of the previous symbols as genie-aided history. Furthermore, the memory required at the FC may be implemented by synthesizing a memory unit into the FC based on [23].

Our major contributions are summarized as follows:

1) We present novel symbol-by-symbol ML detection designs for the cooperative MC system with all detector variants, i.e., SD-ML, MD-ML, and SA-ML. For practicality, we consider the FC chooses the current symbol using its local history and design the methods for the FC to obtain the local history. We also derive the likelihood of observations for all detectors.

2) We derive analytical expressions for the system error probability for SD-ML and SA-ML using the genie-aided history. The assumption of genie-aided history leads to tractable error performance analysis. The analytical error probabilities for SD-ML with $K = 1$ and SA-ML are given in closed form. The error performance of MD-ML is mathematically intractable.

3) We determine the optimal molecule allocation among RXs to minimize the system error probability of SD-ML. To achieve this, we formulate and solve a joint optimization problem in terms of molecule allocation and a constant threshold. In this problem, the objective function is the closed-form approximation of error probability of SD-ML since there is no closed-form expression for the error probability of SD-ML. We also analytically prove that the equal distribution of molecules among two symmetric RXs achieves the local minimal error probability of SD-ML.

4) We validate the accuracy of our analytical expressions of error probability via a particle-based simulation method where we track the motions of molecules over time due to diffusion. Using simulation and numerical results, we also demonstrate the FC’s effectiveness in estimating the previously-transmitted symbols and confirm the effectiveness of our optimization method.

In contrast to our preliminary work in [1], which only presents ML detection design of SD-ML in a symmetric topology, and did not derive the system error probability, this paper presents additional detector variants, i.e., MD-ML and SA-ML, relaxes the constraint of symmetric topology, and derives and optimizes the system error probability.

Notations: We use the following notations: $Pr(\cdot)$ denotes probability, $\lceil x \rceil$ denotes the greatest integer that is less than or equal to $x$, $\lfloor x \rfloor$ denotes the smallest integer that is greater than or equal to $x$, and $\lceil \cdot \rceil$ denotes the nearest integer. $\log(\cdot)$ is the natural logarithm, $\operatorname{erf}(\cdot)$ is the error function, and $\exp(\cdot)$ is the exponential function. $| \cdot |$ is the cardinality of a set.
in Section III. We will describe specific behaviors of the RXs and the FC for the ML detector variants in Section III.

A. System Model

We consider a cooperative MC system in unbounded three-dimensional space. An example of the system is illustrated in Fig. 1. We assume that all RXs and the FC are passive spherical observers. Accordingly, we denote \( V_{\text{RX}_k} \) and \( t_{\text{RX}_k} \) as the volume and radius of the \( k \)th RX, \( \text{RX}_k \), respectively, where \( k \in \{1, 2, \ldots, K\} \). We also denote \( V_{\text{FC}} \) and \( t_{\text{FC}} \) as the volume and radius of the FC, respectively. We use the terms “sample” and “observation” interchangeably to refer to the number of molecules observed by a RX or the FC at some time \( t \) and assume each observation is independent of each other. The symbol interval from the TX to the FC is given by \( T = t_{\text{trans}} + t_{\text{report}} \), where \( t_{\text{trans}} \) is the transmission interval from the TX to the RXs and \( t_{\text{report}} \) is the report interval time from the RXs to the FC.

In the following, we describe the timing schedules and general behaviors of the TX, the RXs, and the FC. An example of the timing schedule for the system is shown in Fig. 2. The timing schedules of the devices could be implemented by introducing oscillators to control the timing of releasing molecules [24]. Also, various methods can be adopted to achieve time synchronization among nanomachines, e.g., [25, 26].

TX: At the beginning of the \( j \)th symbol interval, i.e., \((j - 1)T\), the TX transmits \( W_{\text{TX}}[j] \). The TX transmits \( W_{\text{TX}}[j] \) to the RXs over the diffusive channel via type \( A_0 \) molecules which diffuse independently. The TX uses ON/OFF keying [21] to convey information, i.e., the TX releases \( S_0 \) molecules of type \( A_0 \) to convey information symbol “1” with probability \( P_{\text{TX}[j]} = 1 = P_1 \), but no molecules to convey information symbol “0”. The TX then keeps silent until the start of the \((j + 1)\)th symbol interval. We denote \( L \) as the number of symbols transmitted by the TX. We define \( W_{\text{TX}} = \{W_{\text{TX}}[1], \ldots, W_{\text{TX}}[L]\} \) as an \( L \)-length subsequence of the symbols transmitted by the TX, where \( l \leq L \). Throughout the paper, \( W \) is a single symbol and \( W \) is a vector of symbols. We do not consider channel codes for this system since the required encoder and the decoder may not be practical for MC systems [2, 27].

RX: Each RX\(_k\) observes type \( A_0 \) molecules over the TX – RX\(_k\) link and takes \( M_{\text{RX}_k} \) samples in each symbol interval at the same times. The time of the \( m \)th sample by each RX in the \( j \)th symbol interval is given by \( t_{\text{RX}_k}(j, m) = (j - 1)T + m \Delta t_{\text{RX}_k} \), where \( \Delta t_{\text{RX}_k} \) is the time step between two successive samples by each RX, \( m \in \{1, 2, \ldots, M_{\text{RX}_k}\} \). The RXs operate in half-duplex mode, such that they do not receive the information and report their decisions at the same time. This is because half-duplex mode is more appropriate in a biological environment since it requires lower computational complexity than full-duplex mode. At the time \( (j - 1)T + t_{\text{trans}} \), each RX transmits molecules via a diffusion-based channel to the FC. For MD-ML and SD-ML, each RX detects with a relatively simple energy detector [13]. We denote \( \hat{W}_{\text{RX}_k}[j] \) as RX\(_k\)’s binary decision on the \( j \)th transmitted symbol. Based on the energy detector, RX\(_k\) makes decision \( \hat{W}_{\text{RX}_k}[j] = 1 \) if \( s_k[j] \geq \xi_{\text{RX}_k} \), otherwise \( \hat{W}_{\text{RX}_k}[j] = 0 \), where \( s_k[j] \) is the value of the realization of \( S_{\text{ob}}[j] \) and \( \xi_{\text{RX}_k} \) is the constant detection threshold at RX\(_k\), independent of \( W_{\text{TX}} \). We define \( W_{\text{RX}_k} = \{W_{\text{RX}_k}[1], \ldots, W_{\text{RX}_k}[L]\} \) as an \( L \)-length subsequence of RX\(_k\)’s binary decisions.

FC: The FC takes the \( m \)th sample in the \( j \)th symbol interval at \( t_{\text{FC}}(j, m) = (j - 1)T + m \Delta t_{\text{FC}} + \tilde{m} \Delta t_{\text{FC}} \), where \( \Delta t_{\text{FC}} \) is the time step between two successive samples by the FC and \( \tilde{m} \in \{1, 2, \ldots, M_{\text{FC}}\} \). We denote \( \hat{W}_{\text{FC}}[j] \) as the FC’s decision on the \( j \)th symbol transmitted by the TX. We define \( W_{\text{FC}} = \{\hat{W}_{\text{FC}}[1], \ldots, \hat{W}_{\text{FC}}[L]\} \) as an \( L \)-length subsequence of the FC’s decisions on the symbols transmitted by the TX. We denote \( W_{\text{FC}}[j] \) as the FC’s estimated binary decision of RX\(_k\) on the \( j \)th transmitted symbol. We define \( W_{\text{FC}_k} = \{W_{\text{FC}_k}[1], \ldots, W_{\text{FC}_k}[L]\} \) as the FC’s estimate of the first \( L \) binary decisions by RX\(_k\).

B. Preliminaries

In this subsection, we establish some preliminary results for a TX – RX\(_k\) link and a RX\(_k\) – FC link. We first evaluate the probability \( P_{\text{ob}}(W_{\text{RX}_k}) \) of observing a given type \( A_0 \) molecule, emitted from the TX at \( t = 0 \), inside \( V_{\text{RX}_k} \) at time \( t \). Based on [28, Eq. (27)], we write \( P_{\text{ob}}(W_{\text{RX}_k}) \) as

\[
P_{\text{ob}}(W_{\text{RX}_k}) = \frac{1}{2} \left[ \text{erf} (\tau_1) + \text{erf} (\tau_2) \right] - \frac{\sqrt{D_0 t}}{4D_0^2 \sqrt{\pi}} \left[ \exp \left( -\tau_1^2 \right) - \exp \left( -\tau_2^2 \right) \right],
\]

where \( \tau_1 = \frac{r_{\text{RX}_k} + d_{\text{RX}_k}}{2\sqrt{D_0 t}}, \tau_2 = \frac{r_{\text{RX}_k} - d_{\text{RX}_k}}{2\sqrt{D_0 t}}, D_0 \) is the diffusion coefficient of type \( A_0 \) molecules in \( \text{m}^2/\text{s} \), \( d_{\text{RX}_k} \) is the distance between the TX and RX\(_k\) in \( m \). We denote the sum of \( M_{\text{RX}_k} \).
TABLE II
ILLUSTRATION OF THE FC’S LOCAL HISTORY

| Interval | The FC’s decisions | The FC’s local history |
|----------|--------------------|------------------------|
| 1        | \(W_{\text{FC}}[1]\) and \(W_{\text{FC}}[1]\) | No History |
| 2        | \(W_{\text{FC}}[2]\) and \(W_{\text{FC}}[2]\) | \(W_{\text{FC}}[1]\) and \(W_{\text{FC}}[1]\) |
| ...      | ...                | ...                    |
| L        | \(W_{\text{FC}}[L]\) and \(W_{\text{FC}}[L]\) | \(W_{\text{FC}}[L-1], \ldots, W_{\text{FC}}[1]\) and \(W_{\text{FC}}[L-1], \ldots, W_{\text{FC}}[1]\) |

samples by RX\(_k\) in the \(j\)th symbol interval by \(S_{\text{ob}}^{\text{RX}}[j]\). As discussed in [29, 30], \(S_{\text{ob}}^{\text{RX}}[j]\) can be accurately approximated by a Poisson random variable (RV). The mean of \(S_{\text{ob}}^{\text{RX}}[j]\) is then given by

\[
S_{\text{ob}}^{\text{RX}}[j] = \sum_{i=1}^{j} S_{\text{RX}}[i] \max_{m=1} \sum_{i=1}^{j} P_{\text{RX}}^{(\text{RX, RX})}(j-i) T + m \Delta t_{\text{RX}}. \tag{2}
\]

We denote \(P_{\text{ob, k}}^{(\text{RX, RX})}(t)\) as the probability of observing a given \(A_k\) molecule, emitted from the center of RX\(_k\) at \(t = 0\), inside RX\(_k\) at time \(t\). We obtain \(P_{\text{ob, k}}^{(\text{RX, RX})}(t)\) by replacing \(r_{\text{RX}}, d_{\text{RX}},\) and \(D_0\) with \(r_{\text{RX}}, d_{\text{RX}},\) and \(D_k\), respectively, where \(D_k\) is the diffusion coefficient of type \(A_k\) molecules in \(\text{m}^2/\text{s}\) and \(d_{\text{RX}}\) is the distance between RX\(_k\) and the FC in \(\text{m}\).

III. ML DETECTION DESIGN AND DERIVATION

In this section, we design and derive three symbol-by-symbol ML detectors. Throughout this section, the FC uses its local history to choose the current symbol, i.e., the FC evaluates the likelihood of the observations \(W_{\text{FC}}^{j-1}\) and \(W_{\text{FC}}^{j-1}\) (\(W_{\text{FC}}^{j-1}\) is not needed for SA-ML) in the \(j\)th symbol interval, as shown in Table II, where \(k \in \{1, 2, \ldots, K\}\). Using the local history at the FC, we formulate the general decision rule of ML detection in the \(j\)th interval as

\[
\hat{W}_{\text{FC}}[j] = \arg\max_{W_{\text{RX}}[j] \in \{0, 1\}} \mathcal{L}[j|W_{\text{RX}}[j], \hat{W}_{\text{FC}}^{j-1}], \tag{3}
\]

or

\[
\hat{W}_{\text{FC}}[j] = \arg\max_{W_{\text{RX}}[j] \in \{0, 1\}} \mathcal{L}[j|W_{\text{RX}}[j], \hat{W}_{\text{FC}}^{j-1}, \hat{W}_{\text{FC}}^{j-1}], \tag{4}
\]

where we define \(\mathcal{L}[j|\cdot]\) \(\triangleq\) \(\Pr(\text{FC’s observations in } j\text{th interval} | \cdot)\). Eq. (3) applies to SA-ML and (4) applies to SD-ML and MD-ML. For simplicity, we also write the likelihoods in (3) and (4) as \(\mathcal{L}[j]\). In the following, we present the specific behaviors of the RXs and the FC of each ML detector, derive the corresponding \(\mathcal{L}[j]\), and compare the complexities of the detectors.

A. MD-ML

Each RX\(_k\) in MD-ML transmits type \(A_k\) molecules, which can be independently detected by the FC, to report \(W_{\text{RX}}[j]\) to the FC. Similar to the TX, each RX uses ON/OFF keying to report its decision to the FC and the RX releases \(S_k\) molecules of type \(A_k\) to convey information symbol “1”. The FC receives type \(A_k\) molecules over the RX\(_k\) – FC link and takes \(M_{\text{FC}}\) samples of each of the \(K\) types of molecules transmitted by all RXs in every reporting interval. The FC adds \(M_{\text{FC}}\) observations for each RX\(_k\) – FC link in the \(j\)th symbol interval. We denote \(S_{\text{ob}}^{\text{FC}}[j]\) as the total number of \(A_k\) molecules observed within \(V_{\text{FC}}\) in the \(j\)th symbol interval, due to both current and previous emissions of molecules by RX\(_k\). The TX and RX\(_k\) use the same modulation method and the TX – RX\(_k\) and RX\(_k\) – FC links are both diffusion-based. Therefore, like \(S_{\text{ob}}^{\text{RX}}[j], S_{\text{ob}}^{\text{FC}}[j]\) can also be accurately approximated as a Poisson RV. We denote \(\hat{S}_{\text{ob}}^{\text{FC}}[j]\) as the mean of \(S_{\text{ob}}^{\text{FC}}[j]\). Values of realizations of \(S_{\text{ob}}^{\text{FC}}[j]\) are labeled \(\hat{s}_{\text{k}}[j]\). We assume that the \(K\) RX\(_k\) – FC links are independent, so the FC has \(K\) independent sums \(\hat{s}_{\text{k}}[j]\) from the \(K\) RX\(_k\) – FC links. The FC chooses the symbol \(\hat{W}_{\text{FC}}[j]\) that is more likely, given the joint likelihood of the \(K\) sums \(\hat{s}_{\text{k}}[j]\) in the \(j\)th interval. We obtain \(\mathcal{L}[j]\) by

\[
\mathcal{L}[j] = \prod_{k=1}^{K} \left[ \Pr \left( W_{\text{FC}}[j] = 1 | W_{\text{RX}}[j], W_{\text{FC}}^{j-1} \right) \times \Pr \left( S_{\text{ob}}^{\text{FC}}[j] = \hat{s}_{\text{k}}[j] | W_{\text{RX}}[j], W_{\text{FC}}^{j-1} \right) + \Pr \left( W_{\text{FC}}[j] = 0 | W_{\text{RX}}[j], W_{\text{FC}}^{j-1} \right) \times \Pr \left( S_{\text{ob}}^{\text{FC}}[j] = \hat{s}_{\text{k}}[j] | W_{\text{RX}}[j], W_{\text{FC}}^{j-1} \right) \right]. \tag{5}
\]

For the evaluation of the likelihood in all future intervals, i.e., \(\mathcal{L}[j+1], \ldots, \mathcal{L}[L]\), the FC also chooses the symbol \(\hat{W}_{\text{FC}}[j]\) in the \(j\)th interval given the likelihood of the sum \(\hat{s}_{\text{k}}[j]\) from the RX\(_k\) – FC link in the \(j\)th interval. By doing so, \(\hat{W}_{\text{FC}}[j]\) is obtained by

\[
\hat{W}_{\text{FC}}[j] = \arg\max_{\hat{s}_{\text{k}}[j] \in \{0, 1\}} \mathcal{L}[j|W_{\text{RX}}[j], \hat{W}_{\text{FC}}^{j-1}], \tag{6}
\]

Eqs. (5) and (6) can be evaluated by applying the conditional cumulative distribution function (CDF) of the Poisson RV \(S_{\text{ob}}^{\text{RX}}[j]\) and the conditional PMF of the Poisson RV \(S_{\text{ob}}^{\text{FC}}[j]\). The conditional means \(\hat{S}_{\text{ob}}^{\text{FC}}[j]\) given \(W_{\text{RX}}[j]\) are obtained by replacing \(S_0, W_{\text{RX}}[j], P_{\text{ob}}^{(\text{RX, RX})}, M_{\text{RX}}, m, \) and \(\Delta t_{\text{RX}}\) in (2) with \(S_k, W_{\text{RX}}[j], P_{\text{ob}}^{(\text{RX, RX})}, M_{\text{FC}}, m, \) and \(\Delta t_{\text{FC}}\), respectively.

B. SD-ML

The behavior of each RX\(_k\) in SD-ML is the same as that in MD-ML, except we assume that each RX\(_k\) transmits type \(A_k\) molecules to report \(W_{\text{RX}}[j]\) to the FC. This is because it may not be realistic for each RX to release a unique type of molecule. For simplicity, the number of released type \(A_k\) molecules for each RX\(_k\) in SD-ML is also denoted by \(S_k\). The FC receives type \(A_k\) molecules over all \(K\) RX\(_k\) – FC links and takes \(M_{\text{FC}}\) samples of type \(A_k\) molecules in each symbol interval. The FC adds \(M_{\text{FC}}\) observations for all RX\(_k\) – FC links in the \(j\)th symbol interval. We denote \(S_{\text{ob}}^{\text{FC}}[j]\) as the total number of \(A_k\) molecules observed within \(V_{\text{FC}}\) in the \(j\)th symbol interval, due to both current and previous emissions of molecules by all RXs. We note that \(S_{\text{ob}}^{\text{FC}}[j] = \sum_{k=1}^{K} S_{\text{ob}}^{\text{FC}}[j]\) is also a Poisson RV whose mean is given by \(S_{\text{ob}}^{\text{FC}}[j] = \sum_{k=1}^{K} S_{\text{ob}}^{\text{FC}}[j]\). Values of realizations of \(S_{\text{ob}}^{\text{FC}}[j]\) are labeled \(\hat{s}_{[j]\). The FC chooses the symbol \(\hat{W}_{\text{FC}}[j]
that is more likely, given the likelihood of \( \hat{s}[j] \) in the \( j \)th interval. To facilitate the evaluation of \( \mathcal{L}[j] \) for SD-ML, we define \( \mathcal{W}_{\text{rc}} \) as \( \{ \hat{W}_{\text{rc}}[t], \ldots, \hat{W}_{\text{rc,k}}[t] \} \). Using the notation \( \mathcal{W} \), we derive \( \mathcal{L}[j] \) as

\[
\mathcal{L}[j] = \sum_{h=1}^{2^K} \left[ \Pr \left( \hat{W}_{\text{rc}}[h, j] = \hat{s}[j] | \mathcal{W}_{\text{rc}}[j-1] \right) \times \Pr \left( S_{\text{ob}}^{\text{EC}}[j] = \hat{s}[j] | \mathcal{W}_{\text{rc}}[j-1], \mathcal{W}_{\text{rc}}^\text{EC} \right) \right],
\]

where \( \hat{W}_{\text{rc}}[h, j] \) is the \( h \)th realization of the vector \( \{ \hat{W}_{\text{rc}}[j], \ldots, \hat{W}_{\text{rc,k}}[j] \} \), \( h \in \{1, 2, \ldots, 2^K\} \). For (7), we need to consider each \( \hat{W}_{\text{rc}}[j] \) and the corresponding probability leading to \( \hat{s}[j] \). For the evaluation of the likelihood in all future intervals, the FC chooses \( \mathcal{W}_{\text{rc}}^\text{EC} \) that gives the maximum likelihood of \( \hat{s}[j] \). By doing so, \( \mathcal{L}_{\text{EC}}[j] \) is obtained by

\[
\mathcal{L}_{\text{EC}}[j] = \arg\max_{\mathcal{W}_{\text{rc}}^\text{EC}} \Pr \left( S_{\text{ob}}^{\text{EC}}[j] = \hat{s}[j] | \mathcal{W}_{\text{rc}}[j-1], \mathcal{W}_{\text{rc}}^\text{EC} \right).
\]

We now derive the conditional mean of \( S_{\text{ob}}^{\text{EC}}[j] \) given \( \mathcal{W}_{\text{rc}}[j] \) and \( \mathcal{W}_{\text{rc}}[j-1], \ldots, \mathcal{W}_{\text{rc}}^\text{EC} \). To do so, we evaluate \( \mathcal{S}_{\text{ob}}^{\text{EC}}[j] \) as

\[
\mathcal{S}_{\text{ob}}^{\text{EC}}[j] = \sum_{k=1}^{M_{\text{ec}}} \sum_{i=1}^{\hat{m}} \mathcal{S}_{\text{ec},k} \left( \hat{W}_{\text{ec}}[j], \mathcal{S}_{\text{ob}}^{\text{EC}}[j] \right) \left( \hat{m} \Delta_{\text{ec}} \right)
\]

\[
+ \sum_{i=1}^{\hat{m}} \hat{W}_{\text{ec}}[i] \mathcal{P}_{\text{ob},k} \left( \mathcal{S}_{\text{ec},k} \right) \left( \left( j - i \right) T + \hat{m} \Delta_{\text{ec}} \right).
\]

D. Comparison of Complexity

We summarize the complexity comparison in Table I. MDML requires higher complexity than SD-ML. This is because each RX releases a unique type of molecule in MD-ML, whereas in SD-ML the RXs release a single type of molecule. SD-ML requires higher complexity than SA-ML. This is because the RXs need to decode the TX’s symbols and the FC needs to estimate the RXs’ decisions in SD-ML, but in SA-ML the RXs only need to amplify the received signal and the FC does not need to estimate the RXs’ decisions.

IV. ERROR PERFORMANCE ANALYSIS

In this section, we derive the error probability of SD-ML and SA-ML using the genie-aided history, which leads to tractable expressions. Also, the error probability with genie-aided history provides a lower bound on that with local history. We denote \( Q_{\text{ec}}[j] \) as the error probability of the system in the \( j \)th symbol interval for a TX sequence \( W_{\text{ec}}[j] \). The closed-form expressions of \( Q_{\text{ec}}[j] \) for SD-ML with \( K = 1 \) and SA-ML are mathematically tractable.

To derive \( Q_{\text{ec}}[j] \), we first derive equivalent decision rules with lower-complexity than (3) and (4) for SD-ML and SA-ML in Theorems 1 and 2, respectively. The decision rules when not all previously-transmitted symbols are “0” cannot be directly applied to the case where all previously-transmitted symbols are “0”. Based on these theorems, the general forms of these lower-complexity decision rules are that the FC compares the observation with adaptive thresholds when not all previously-transmitted symbols are “0” and the FC compares the observation with 0 when all previously-transmitted symbols are “0”. Notably, these adaptive thresholds adapt to different ISI in different symbol intervals.

A. SD-ML

We now derive \( Q_{\text{ec}}[j] \) for the SD-ML variant. To this end, we first define \( \hat{N}_{\text{ec}}[j] \) as the expected ISI at the FC in the \( j \)th symbol interval due to the emissions of molecules from the TX and the previous intervals by RXs. The TX – RXk and RXk – FC links are both diffusion-based. Therefore, \( S_{\text{ob},k}^\text{EC}[j] \) can be accurately approximated as a Poisson RV. We denote \( S_{\text{ob},k}^\text{EC}[j] \) as the mean of \( S_{\text{ob},k}^\text{EC}[j] \). The FC adds \( M_{\text{ec}} \) observations for all RXk – FC links in the \( j \)th symbol interval and this sum is denoted by the RV \( S_{\text{ob},k}^\text{EC}[j] \). We note that \( S_{\text{ob},k}^\text{EC}[j] = \sum_{k=1}^{K} S_{\text{ob},k}^{\text{EC}}[j] \) is also a Poisson RV whose mean is given by \( \sum_{k=1}^{K} S_{\text{ob},k}^\text{EC}[j] \). Values of realizations of \( S_{\text{ob},k}^\text{EC}[j] \) are labeled \( s_{\text{ec}}[j] \). The FC chooses the symbol \( W_{\text{ec}}[j] \) that is more likely given the likelihood of \( s_{\text{ec}}[j] \) in the \( j \)th interval and \( \mathcal{L}[j] \) is given in (10) at the top of the following page. where \( S_{\text{ob}}^\text{EC}[i] \) and \( \mathcal{s}_{\text{ec}}[i] \), \( i \in \{1, \ldots, j\} \) and \( k \in \{1, \ldots, K\} \), are defined in Section II-A.

Theoretically, any number of molecules between 0 and \( S_{0} \) can be observed at each RX. Thus, there is a large number of realizations for each Poisson RV \( S_{\text{ob}}^\text{EC}[i] \) in (10), which makes the complete evaluation of (10) cumbersome. To simplify the evaluation of (10), we consider finitely many random realizations of each Poisson RV \( S_{\text{ob}}^\text{EC}[i] \). For example, we generate 5000 random realizations of each \( S_{\text{ob}}^\text{EC}[i] \) for a given \( W_{\text{ec}}[j] \), which is sufficient to ensure the accuracy of (10). It is shown that (10) can be evaluated by applying the conditional PMF of the Poisson RV \( S_{\text{ob}}^\text{EC}[i] \). We obtain the conditional mean of \( S_{\text{ob}}^\text{EC}[i] \) by replacing \( S_{\text{ob}}[j], P_{\text{ob}}^\text{EC}[j] \), and \( M_{\text{ec}} \), \( m \), and \( \Delta t_{\text{ec}} \) in (2) with \( S_{\text{ob}}[j], P_{\text{ob}}^\text{EC}[j] \), \( M_{\text{ec}} \), \( m \), and \( \Delta t_{\text{ec}} \), respectively. Based on \( S_{\text{ob}}^\text{EC}[i] = \sum_{k=1}^{K} S_{\text{ob},k}^\text{EC}[j] \), we can then obtain the conditional mean of \( S_{\text{ob}}^\text{EC}[j] \).

\[ Q_{\text{ec}}[j] = \sum_{i=1}^{\hat{m}} \hat{W}_{\text{ec}}[i] P_{\text{ob},k} \left( \mathcal{S}_{\text{ec},k} \right) \left( \left( j - i \right) T + \hat{m} \Delta_{\text{ec}} \right). \]
\[ \mathcal{L}[j] = \sum_{s_0, s_1, \ldots, s_K = 0}^{S_0, S_1, \ldots, S_K} \prod_{i=1}^{K} \Pr\left( s_{ob,i}^k = s_i^k \right) \sum_{s_0} S_0^j \sum_{s_1} S_1^j \sum_{s_K} S_K^j \times \Pr\left( s_{ob,k}^j = s_k^j, \ldots, s_{ob,k}^j = s_K^j, s_{ob,k}^j = s_K^j \right) \]

where the conditional CDF of the Poisson RV \( S_{ob,k}^{sd,j} \) can be evaluated by
\[ \Pr\left( s_{ob,k}^{sd,j} < \xi_{ad,ob}^{sd,j} | W_{tx}^j, \hat{\lambda}_{0}^j \right) = \frac{\sum_{h=1}^{2^K} \left[ \Pr\left( \hat{W}_{tx}^{rk,h} | W_{tx}^j, W_{tx}^{j-1} \right) \exp\left( -\hat{\lambda}_p^j - \hat{\lambda}_{0,ob}^j \right) \right] \eta^{\hat{\lambda}_p^j + \hat{\lambda}_{0,ob}^j \eta} }{\eta^{\hat{\lambda}_p^j + \hat{\lambda}_{0,ob}^j \eta} } , \]

where \( \hat{\lambda}_p^j \) can be evaluated by (11) via the approximated \( W_{tx}^{j-1} \), \( k \in \{1, 2, \ldots, K\} \). The approximated \( W_{tx}^{j-1} \) can be obtained using the biased coin toss method introduced in [31]. Specifically, we model the 4th decision at RX, \( W_{tx}^{j-1} \), as \( W_{tx}^{j-1} = |\lambda - W_{tx}^j| \), where \( i \in \{1, 2, \ldots, j-1\} \) and \( \lambda \in \{0, 1\} \) is the outcome of the coin toss with \( \Pr(\lambda = 1) = \Pr(\hat{W}_{tx}^{rk,h} = 0 | W_{tx}^j = 1) \) if \( W_{tx}^j = 1 \) and \( \Pr(\lambda = 1) = \Pr(\hat{W}_{tx}^{rk,h} = 1 | W_{tx}^j = 0) \) if \( W_{tx}^j = 0 \). When \( \hat{\lambda}_p^j = 0 \), we evaluate the conditional \( Q_{\text{EC}} \) as
\[ Q_{\text{EC}} \left[ j | \hat{\lambda}_p^j = 0 \right] = P_1 \Pr\left( S_{ob,k}^{sd,j} = 0 | W_{tx}^j = 1, \hat{\lambda}_p^j = 0 \right) + (1 - P_1) \Pr\left( S_{ob,k}^{sd,j} > 0 | W_{tx}^j = 0, \hat{\lambda}_p^j = 0 \right) , \]
performance analysis, we consider only one random realization of $S^{\text{RXs}}_{i}$ with the mean $S^{\text{RXs}}_{i}$ for the given previous symbols transmitted by the TX, $W^{i-1}_{\text{TX}}$. We define $\hat{\lambda}_i^{[j]}$ as the expected ISI at the FC in the $j$th symbol interval due to $W^{i-1}_{\text{TX}}$. We define $\check{\lambda}_i^{[j]}$ as the number of the signal molecules at the FC in the $j$th symbol interval due to $W^{i}_{\text{TX}}$. By modeling the realization of $S^{\text{RXs}}_{i}$ as its mean $S^{\text{RXs}}_{\text{av}}[i]$, we write $\hat{\lambda}_i^{[j]}$ and $\check{\lambda}_i^{[j]}$ as

$$
\hat{\lambda}_i^{[j]} = \sum_{k=1}^{K} \left( \sum_{i=1}^{j} \alpha_{k} S^{\text{RXs}}_{ob}[i] \sum_{m=1}^{M_{\text{FC}}} P^{(\text{RXs,R})}_{ob,k} ((j-i) T + \bar{m} \Delta t_{\text{FC}}) \right)
+ \sum_{i=1}^{j} \alpha_{k} S^{0}_{ob} \sum_{m=1}^{M_{\text{FC}}} P^{(\text{TX,R})}_{ob} ((j-i) T + m \Delta t_{\text{FC}}) \times \sum_{m=1}^{M_{\text{FC}}} P^{(\text{RXs,R})}_{ob,k} (\bar{m} \Delta t_{\text{FC}}),
$$

and

$$
\check{\lambda}_i^{[j]} = \sum_{k=1}^{K} \alpha_{k} S^{0}_{ob} \sum_{m=1}^{M_{\text{FC}}} P^{(\text{TX,R})}_{ob} (m \Delta t_{\text{FC}}) \sum_{m=1}^{M_{\text{FC}}} P^{(\text{RXs,R})}_{ob,k} (\bar{m} \Delta t_{\text{FC}}),
$$

respectively. $\hat{\lambda}_i^{[j]}$ in (20) consists of two components. The first summation over $i$ is the expected ISI at the FC in the $j$th symbol interval due to the molecules released by the RXs but without the amplification of the RXs' ISI from the TX. The second summation over $i$ accounts for the amplification of the ISI in the $j$th symbol interval at all RXs due to $W^{i-1}_{\text{TX}}$. We note that the conditional mean of $S^{\text{RXs}}_{ob}[i]$ is $\lambda_i^{[j]} + \check{\lambda}_i^{[j]}$ when $W^{i}_{\text{TX}}[j] = 1$, and the conditional mean of $S^{\text{RXs}}_{ob}[i]$ is $\hat{\lambda}_i^{[j]}$ when $W^{i}_{\text{TX}}[j] = 0$. If not all previous symbols transmitted by the TX are “0”, i.e., $W^{i}_{\text{TX}}[j] \neq 0$, then we have $\hat{\lambda}_i^{[j]} > 0$. If all previous symbols transmitted by the TX are “0”, i.e., $W^{i-1}_{\text{TX}} = 0$, then we have $\hat{\lambda}_i^{[j]} = 0$. For the sake of brevity, for SA-ML, we define $\mathcal{L}_1[j|W^{i}_{\text{TX}}[j] = 1, W^{i-1}_{\text{TX}}] \triangleq \mathcal{L}_1^{[j]}$ and $\mathcal{L}_0[j|W^{i}_{\text{TX}}[j] = 0, W^{i-1}_{\text{TX}}] \triangleq \mathcal{L}_0^{[j]}$. Applying the conditional PMF of the Poisson RV $S^{\text{RXs}}_{ob}[j]$ to (10), we derive $\mathcal{L}_1^{[j]}$ and $\mathcal{L}_0^{[j]}$ as

$$
\mathcal{L}_1^{[j]} = \frac{\bar{s}_j^{[j]} \exp \left( -\left( \bar{\lambda}_i^{[j]} + \check{\lambda}_i^{[j]} \right) \right)}{\bar{s}_j^{[j]}!},
$$

and

$$
\mathcal{L}_0^{[j]} = \frac{\bar{s}_j^{[j]} \exp \left( -\bar{\lambda}_i^{[j]} \right)}{\bar{s}_j^{[j]}!},
$$

respectively. Based on (22) and (23), we rewrite the general decision rule of SA-ML in (3) as a lower-complexity decision rule in the following theorem.

**Theorem 2:** When $\hat{\lambda}_i^{[j]} > 0$, the decision rule of SA-ML is

$$
\hat{W}_{\text{ML}}[j] = \begin{cases} 
1, & \text{if } \bar{s}_j^{[j]} \geq \xi_{\text{ML}}^{\text{RXs}}[j], \\
0, & \text{otherwise},
\end{cases}
$$

where $\xi_{\text{ML}}^{\text{RXs}}[j] = \left[ \bar{\lambda}_i^{[j]}/\log(\bar{\lambda}_i^{[j]} + \check{\lambda}_i^{[j]}) \right]$. When $\hat{\lambda}_i^{[j]} = 0$, the decision rule is

$$
\hat{W}_{\text{ML}}[j] = \begin{cases} 
1, & \text{if } \bar{s}_j^{[j]} > 0, \\
0, & \text{otherwise}.
\end{cases}
$$

**Proof:** Applying (22) and (23) to (3), we rewrite the decision rule for SA-ML as

$$
\left( \bar{\lambda}_i^{[j]} + \check{\lambda}_i^{[j]} \right) \bar{s}_j^{[j]} \exp \left( -\left( \bar{\lambda}_i^{[j]} + \check{\lambda}_i^{[j]} \right) \right) \\
\hat{W}_{\text{ML}}[j]=1 \quad \hat{W}_{\text{ML}}[j]=0 \quad \bar{s}_j^{[j]} \exp \left( -\bar{\lambda}_i^{[j]} \right).
$$

We then discuss the case when $W^{i-1}_{\text{TX}} = 0$ and $W^{i-1}_{\text{TX}} \neq 0$. When $W^{i}_{\text{TX}} \neq 0$, then we have $\hat{\lambda}_i^{[j]} > 0$ and we rewrite (26) as

$$
\left( \bar{\lambda}_i^{[j]} / \check{\lambda}_i^{[j]} + 1 \right) \bar{s}_j^{[j]} \hat{W}_{\text{ML}}[j]=1 \quad \hat{W}_{\text{ML}}[j]=0 \quad \exp \left( \check{\lambda}_i^{[j]} \right).
$$

We rearrange (27) and obtain (24). We next discuss the case $W^{i}_{\text{TX}} = 0$, which leads to $\hat{\lambda}_i^{[j]} = 0$. If $\lambda_i^{[j]} = 0$ and $\bar{s}_j^{[j]} = 0$, we write (26) as

$$
\exp \left( -\bar{\lambda}_i^{[j]} \right) \hat{W}_{\text{ML}}[j]=1 \quad \hat{W}_{\text{ML}}[j]=0 \quad 1,
$$

where the decision at the FC is always $\hat{W}_{\text{ML}}[j] = 0$ since $\bar{s}_j^{[j]}$ always holds. If $\lambda_i^{[j]} = 0$ and any $\bar{s}_j^{[j]} > 0$, we write (26) as

$$
\left( \bar{\lambda}_i^{[j]} / \check{\lambda}_i^{[j]} \right) \bar{s}_j^{[j]} \hat{W}_{\text{ML}}[j]=1 \quad \hat{W}_{\text{ML}}[j]=0 \quad 0,
$$

where the decision at the FC is always $\hat{W}_{\text{ML}}[j] = 1$ since $\bar{s}_j^{[j]}$ always holds. Thus, we obtain the decision rule in (25).

Based on Theorem 2, when $W^{i-1}_{\text{TX}} \neq 0$, we evaluate $Q_{\text{ML}}[j]$ for SA-ML as

$$
Q_{\text{ML}}[j] = (1 - P_1) \Pr \left( S^{\text{RXs}}_{ob}[j] \geq \xi_{\text{ML}}^{\text{RXs}}[j] \middle| W^{i}_{\text{TX}}[j] = 0, W^{i-1}_{\text{TX}} \right)
+ P_1 \Pr \left( S^{\text{RXs}}_{ob}[j] < \xi_{\text{ML}}^{\text{RXs}}[j] \middle| W^{i}_{\text{TX}}[j] = 1, W^{i-1}_{\text{TX}} \right),
$$

where $\Pr \left( S^{\text{RXs}}_{ob}[j] < \xi_{\text{ML}}^{\text{RXs}}[j] \middle| W^{i}_{\text{TX}}[j] \right)$ can be evaluated by replacing $\hat{W}_{\text{ML}}[j]$ and $S^{\text{RXs}}_{ob}[j] = \bar{s}_j^{[j]}$ in (10) with $W^{i}_{\text{TX}}[j]$ and $S^{\text{RXs}}_{ob}[j] < \xi_{\text{ML}}^{\text{RXs}}[j]$, respectively. Similar to the evaluation of (10), we consider finitely many random realizations of $S^{\text{RXs}}_{ob}[i]$ in (30). When $W^{i-1}_{\text{TX}} = 0$, $Q_{\text{ML}}[j]$ for SA-ML can be obtained by replacing $\geq$, $<$, and $\xi_{\text{ML}}^{\text{RXs}}[j]$ with $>$, $=,$ and $0$ in (30), respectively.

**V. ERROR PERFORMANCE OPTIMIZATION**

In this section, we determine the optimal molecule distribution among RXs that minimizes the error probability of SD-ML using the genie-aided history, inspired by the fact the quantity of any type of molecule is usually constrained in practical biological environments. We also analytically prove that the equal allocation of molecules among two symmetric RXs achieves the local minimal error probability of SD-ML.
To this end, we first formulate the optimization problem as follows:

$$\min S \quad Q_{rc}[j] \quad \text{in (19)}$$

s.t.  
$$S_1 + S_2 + \cdots + S_K - N = 0,$$
$$S_k \geq 0,$$  \hspace{1cm} (31)

where $S = \{ S_1, S_2, \ldots, S_K \}, k \in \{1, 2, \ldots, K\}$, and $N$ is the total number of molecules released by $K$ RXs for symbol “1”. Combining (16) and (19), we note that $\xi_{rc,ad,SD}^{j}$ is required to evaluate $Q_{rc}[j]$. Based on Theorem 1, the adaptive threshold $\xi_{rc,ad,SD}^{j}$ is obtained by numerically solving $L^{\pi}_{\xi} = L_{\xi}^{ad,SD}$ in terms of $\tilde{s}[j]$, while the closed-form expression for $\xi_{rc,ad,SD}^{j}$ is mathematically intractable. Therefore, there is no closed-form expression for $Q_{rc}[j]$, which makes it very hard to optimize $Q_{rc}[j]$ in (19). To tackle this challenge, we find a closed-form approximation for $Q_{rc}[j]$ in (19) by considering a constant threshold $\xi$ in (16). By doing so, we find the approximation of $Q_{rc}[j]$ as

$$Q_{rc}^\xi[j] = P_1 \sum_{h=1}^{2^K} \left[ \Pr \left( \tilde{W}_{TX}^{\xi,j,h} | W_{TX}[j] = 1, W_{TX}^{-1} \right) \right]$$

$$+ (1 - P_1) \sum_{h=1}^{2^K} \left[ \Pr \left( \tilde{W}_{TX}^{\xi,j,h} | W_{TX}[j] = 0, W_{TX}^{-1} \right) \right] \times (1 - \Lambda), \hspace{1cm} (32)$$

where $Q_{rc}^\xi[j]$ is the approximation of $Q_{rc}[j]$, $\Lambda$ is given by

$$\Lambda = \sum_{\eta=0}^{\xi-1} \exp \left( -\tilde{\lambda}_p^\xi[j] - \tilde{\lambda}_v^\xi,j,h [\eta] \right) \left( \frac{\tilde{\lambda}_p^\xi[j] + \tilde{\lambda}_v^\xi,j,h [\eta]}{\eta} \right) \right)^n, \hspace{1cm} (33)$$

and $\xi$ is a constant. In (33), $\tilde{\lambda}_p^\xi[j]$ and $\tilde{\lambda}_v^\xi,j,h [\eta]$ are the functions of $S$ based on (11) and (12).

Lemma 1: The approximation of $Q_{rc}[j]$ by $Q_{rc}^\xi[j]$ is tight when $\xi = \xi_{rc,ad,SD}^{j}$.

Proof: We note that the likelihood of the occurrence that all previous symbols transmitted by all RXs are “0” is very small. Thus, we approximate $\Pr \left( \tilde{\lambda}_p^\xi[j] = 0 | W_{TX}^{-1} \right) \approx 0$ and $\Pr \left( \tilde{\lambda}_p^\xi[j] > 0 | W_{TX}^{-1} \right) \approx 1$. Using these approximations in (19), we obtain $Q_{rc}[j] \approx Q_{rc}^\xi[j]$ for $\xi = \xi_{rc,ad,SD}^{j}$. We then note that $Q_{rc}[j]_{\xi = \xi_{rc,ad,SD}^{j}} = Q_{rc}^\xi[j]_{\xi = \xi_{rc,ad,SD}^{j}}$. Thus, $Q_{rc}[j]$ is accurately approximated by $Q_{rc}^\xi[j]$ when $\xi = \xi_{rc,ad,SD}^{j}$.

Lemma 2: Since the adaptive threshold $\xi_{rc,ad,SD}^{j}$ adapts to different ISI for different symbol intervals, $\xi_{rc,ad,SD}^{j}$ is the optimal $\xi$ that minimizes $Q_{rc}^\xi[j]$ if $P_1 = \frac{1}{2}$, i.e., $\xi_{rc,ad,SD}^{j} = \arg\min Q_{rc}^\xi[j]$.

$\xi$

Proof: Please see Appendix B.

Based on Lemma 1 and Lemma 2, the approximation of $Q_{rc}[j]$ by $Q_{rc}^\xi[j]$ is tight when $\xi = \xi_{rc,ad,SD}^{j}$ and $\xi_{rc,ad,SD}^{j}$ is the optimal $\xi$ which minimizes $Q_{rc}^\xi[j]$. Therefore, the optimal $S$ that minimizes $Q_{rc}[j]$ in (19) can be obtained by finding the jointly optimal $S$ and $\xi$ to minimize $Q_{rc}^\xi[j]$ in (32), i.e., the approximate solution to the problem (31) can be obtained by solving the optimization problem given by:

$$\min S, \xi \quad Q_{rc}^\xi[j]$$

s.t.  
$$S_1 + S_2 + \cdots + S_K - N = 0,$$
$$S_k \geq 0.$$  \hspace{1cm} (34)

To solve (34), we examine its convexity. The convexity of an optimization problem can be proven by showing that its objective function and constraints are convex with respect to the optimization variables. Since the constraints in (34) are affine, they are convex. The convexity of the objective function, i.e., $Q_{rc}^\xi[j]$, can be proven by showing that its Hessian is positive semidefinite with respect to its optimization variables. For the convexity of $Q_{rc}^\xi[j]$, we have the following proposition:

Proposition 1: The Hessian of $Q_{rc}^\xi[j]$ is not positive semidefinite with respect to $S$ and $\xi$.

Proof: Please see Appendix C.

Based on Proposition 1, the multi-dimensional optimization problem (34) is not a convex optimization problem. To overcome this challenge, we use GlobalSearch in MATLAB to repeatedly run a local solver with the sequential quadratic programming (SQP) algorithm until convergence is achieved (i.e., the global minimum is found) to solve the problem (34). Our numerical results in Section VI confirm the effectiveness of this optimization method.

To obtain additional analytical insights in molecule distribution, we discuss the optimal distribution of the number of molecules in a symmetric topology. Intuitively, we expect that an equal distribution of molecules among symmetric RXs is the optimal allocation to minimize the error probability. To confirm this conjecture, we first find that the equal distribution locally minimizes $Q_{rc}^\xi[j]$ under certain conditions. We derive such conditions in the following Lemma:

Lemma 3: In the symmetric topology with $K = 2$, if $T(\xi) > 0$, $Q_{rc}^\xi[j]$ achieves a local minimum when $S_1 = \frac{N}{2}$; otherwise, it achieves a local maximum, where $T(\xi)$ is given by

$$T(\xi) = (\alpha(P_1 - 1) + \beta P_1) (2 + N(\nu + 2\sigma) - 2\xi), \hspace{1cm} (35)$$

where

$$\sigma_1 = \sigma_2 = \sigma, \quad \nu_1 = \nu_2 = \nu,$$

$$\alpha(1, 0) = \alpha(0, 1) = \alpha, \quad \text{and} \quad \beta(1, 0) = \beta(0, 1) = \beta, \hspace{1cm} (36)$$

$$\sigma_k = \sum_{i=1}^{j-1} \tilde{W}_{RX} [i] \sum_{m=1}^{M_{rc}} P_{ob,k}^{(RX, rc)} ((j - i) T + \tilde{m} \Delta t_{rc}), \hspace{1cm} (37)$$

$$\nu_k = \sum_{m=1}^{M_{rc}} P_{ob,k}^{(RX, rc)} (\tilde{m} \Delta t_{rc}), \hspace{1cm} (38)$$

$$\alpha(a_1, a_2) = \Pr \left( \tilde{W}_{RX} [j] = a_1 | W_{TX}[j] = 0, W_{TX}^{-1} \right) \times \Pr \left( \tilde{W}_{RX} [j] = a_2 | W_{TX}[j] = 0, W_{TX}^{-1} \right), \hspace{1cm} (39)$$
and

\[ \beta(a_1, a_2) = \Pr\left( W_{RX_{j}}[j] = a_1 | W_{TX}[j] = 1, W_{TX}^{j-1} \right) \]
\[ \times \Pr\left( \hat{W}_{RX_{j}}[j] = a_2 | W_{TX}[j] = 1, W_{TX}^{j-1} \right). \]

(40)

**Proof:** Please see Appendix D.

Using Lemma 1, Lemma 2, and Lemma 3, we find that the equal distribution of molecules always achieves the local minimal error probability for SD-ML in a two-RX system, as stated in the following theorem:

**Theorem 3:** In the symmetric topology with two RXs, \( Q_{RC}[j] \) achieves a local minimal value when \( S_1 = \frac{\sigma_1^2}{2} \) if \( P_1 = \frac{1}{2} \).

**Proof:** Please see Appendix E.

## VI. NUMERICAL RESULTS AND SIMULATIONS

In this section, we present numerical and simulation results to examine the error performance of the ML detectors. We simulate using a particle-based method considered in [32], where we track the precise locations of all individual molecules. Unless otherwise noted, we consider the environmental parameters in Table III.

Throughout this section, we keep the TX and the FC fixed at \((0\mu m, 0\mu m, 0\mu m)\) and \((2\mu m, 0\mu m, 0\mu m)\), respectively. To clearly demonstrate the impact of the number of samples and the number of RXs on the error probability of the system, we consider a symmetric topology in Section VI-A. To clearly show the impact of asymmetric RX location on the error probability of the system and the corresponding optimal molecule distribution, we consider an asymmetric topology in Section VI-B.

We assume that the TX releases \(10^6\) molecules for symbol “1”. We also assume the total number of molecules released by all RXs for symbol “1” is fixed at 2000 throughout this section to ensure the fairness of error performance comparison for different \( K \). For MD-ML and SD-ML, in Figs. 3-5, each RX releases \( S_0 = \left[ 2000/K \right] \) molecules to report a decision of “1”. For SA-ML, in Figs. 3-5, each RX uses an amplification factor to ensure that the average number of molecules released by all RXs for transmission of one symbol is 1000 for the fair comparison among SA-ML, SD-ML, and MD-ML. \( Q_{RC} \) is obtained by averaging \( Q_{RC}[j] \) over all symbol intervals and 50000 random-generated realizations of \( W_{TX}^{j-1} \), and then the value of \( Q_{RC}^* \) is the minimum \( Q_{RC} \) found by numerically optimizing the corresponding constant decision thresholds via exhaustive search. To decrease the complexity of exhaustive search, we consider the same decision threshold at all RXs such that \( \hat{\xi}_{RX_i} = \xi_{RF}, \forall k \).

In Figs. 3-5, for each ML detection variant, we plot the error probability with the local history and genie-aided history. We observe that the error performance using the local history has a very small degradation from that using the genie-aided history. This demonstrates the effectiveness of our proposed method to estimate the previous symbols. We also observe that the simulations have very strong agreement with the analytical results, thereby validating our analytical results. In Figs. 3-5, we observe that the error performance degradation with the local history compared to the genie-aided history for SA-ML is more noticeable than that for SD-ML and MD-ML. This is because in SD-ML and MD-ML, the FC directly estimates previous RX symbols from the RX-FC links. However, for SA-ML, the FC does not directly estimate the previous RX emissions from the RX-FC links and the error in the estimation of previous TX symbols propagates to the estimated previous RX emissions.

### A. Symmetric Topology

We consider at most 6 RXs in this subsection and the specific locations of RXs are: \((2\mu m, \pm 0.6\mu m, 0)\) and \((2\mu m, \pm 0.3\mu m, \pm 0.5196\mu m)\), where the RXs are placed on a circle perpendicular to the line passing from the TX, the FC, and the center of the circle.

In order to provide trade-offs between the performance versus the information available, we compare the error performance of the ML detectors with the majority rule [8, 9] and SD-Constant [10]. Notably, we also propose a new variant for comparison, namely, SA-Constant. In SA-Constant, the behavior of each RX is the same as that in SA-ML, but the FC makes a decision \( \hat{W}_{RC}[j] \) by comparing \( \hat{s}[j] \) with a constant threshold \( \hat{\xi}_{RF} \), independent of \( W_{TX}^{j-1} \). It can be shown that \( Q_{RC}[j] \) for SA-Constant with any realization of \( W_{TX}^{j-1} \) can be obtained by replacing \( \xi_{RF}^{sd,sa} [j] \) with the threshold \( \hat{\xi}_{RF} \) in (30). We summarize all variants considered in this subsection in Table IV. For these variants, we consider the same parameters as the ML detectors for the fairness of our comparisons.

### TABLE III

| Parameter                  | Symbol     | Value                     |
|----------------------------|------------|---------------------------|
| Volume of each RX          | V_{RX}     | \( \frac{4}{3} \pi \times 0.2 \text{ \mu m}^3 \) |
| Radius of FC               | r_{FC}     | 0.2 \text{ \mu m}         |
| Time step at RXs           | \Delta t_{RX} | 100 \text{ \mu s}     |
| Time step at FC            | \Delta t_{FC} | 50 \text{ \mu s}        |
| Number of samples by RXs   | M_{RX}     | 5                         |
| Number of samples by FC    | M_{FC}     | 10                        |
| Transmission time interval  | t_{trans}  | 1 \text{ ms}             |
| Report time interval       | t_{report} | 0.3 \text{ ms}           |
| Bit interval time          | T          | 1.3 \text{ ms}           |
| Diffusion coefficient      | D_0 = D_k  | 5 \times 10^{-6} \text{ m}^2/\text{s} |
| Length of symbol sequence  | L          | 20                        |
| Probability of binary 1    | P_1        | 0.5                       |

### TABLE IV

| Variants          | Relaxing at RXs | Molecule Type Used in RXs | Behavior at FC |
|-------------------|-----------------|---------------------------|----------------|
| Majority Rule [8,9]| DF              | Multiple                  | Constant Threshold |
| MD-ML             | DF              | Multiple                  | ML Detection    |
| SD-Constant [10]   | DF              | Single                    | Constant Threshold |
| SD-ML             | DF              | Single                    | ML Detection    |
| SA-Constant        | AF              | Single                    | Constant Threshold |
| SA-ML             | AF              | Single                    | ML Detection    |
In Fig. 3, we plot the optimal average global error probability $Q_{ec}^*$ of different variants versus the number of RXs $K$. The analytical error performance of SD-Constant and the majority rule is presented in [10] and [9], respectively.

In Fig. 4(a), we observe that SD-Constant and SA-ML using the local history achieve similar error performance. In Fig. 4(b), we observe that the majority rule has similar error performance with SD-ML and the majority rule outperforms SA-ML using the local history. These observations demonstrate the good performance of the majority rule, relative to SD-ML and SA-ML. Importantly, we observe that MD-ML outperforms SD-ML and SA-ML using the local history.

As $K$ increases, even though the total number of molecules is constrained. The same observation of error performance improvement may be observed in a channel with additive signal dependent noise if our results can be well approximated by the Gaussian signal dependent noise model [33]. The system error performance does not always improve as $K$ increases. This is because if we keep increasing $K$, the number of released molecules for each RX decreases, which leads to the RX-FC link becoming unreliable. The system error performance would improve as the volume of the FC increases for the fixed $K$, since the FC can observe more molecules, but the volume of microorganisms cannot be easily altered.

In Fig. 5, we consider a three-RX system and plot the optimal average error probability $Q_{ec}^*$ of different variants versus the distance $d_{RX_k}$ between the TX and RX$_3$. RX1 and RX2 are fixed at $(2\mu m, 0, 0.6\mu m)$ and $(2\mu m, 0, -0.6\mu m)$, respectively. The locations of RX$_3$ are (1) $(2\mu m, 0.48\mu m, 0)$, (2) $(1.6\mu m, 0.48\mu m, 0)$, (3) $(1.2\mu m, 0.36\mu m, 0)$, (4) $(0.8\mu m, 0.24\mu m, 0)$, (5) $(0.4\mu m, 0.12\mu m, 0)$.

B. Asymmetric Topology

In Fig. 5, we consider a three-RX system and plot the optimal average error probability of different variants versus the distance between the TX and RX$_3$. We keep the positions of RX$_1$ and RX$_2$ fixed and move RX$_3$ along the line segment.
between the symmetric position and the TX, as indicated in the caption. We observe for our three variants that the error performance first improves and then decreases as RX3 moves toward the TX. This is because both the TX-RX3 link and the RX2-FC link contribute to the error performance of the system. When \( d_{RX3} \) is relatively large, the system error performance is dominated by the TX-RX3 link and this link becomes more reliable as \( d_{RX3} \) decreases. For \( d_{RX2} \) is relatively small, the system error performance is dominated by the RX2-FC link, which becomes weaker when \( d_{RX2} \) decreases. We also observe that MD-ML outperforms SD-ML and SD-ML outperforms SA-ML, which is consistent with our observations in Fig. 4(b).

In the following figures, we present results to assess the accuracy of our proposed optimization method in Section V. We denote the solution to problem (34) by \( S^\dagger = \{S_1\dagger, S_2\dagger, \ldots, S_K\dagger\} \). We denote the optimal solution via exhaustive search by \( S^* = \{S_1^*, S_2^*, \ldots, S_K^*\} \).

In Fig. 6, we consider a two-RX system and plot the error probability of SD-ML versus the number of molecules released by RX1, for different locations of RX1, where we keep RX2 fixed at \((2\mu m, 0.6\mu m, 0\mu m)\) and move RX1 along the line segment between the symmetric position and the TX, as indicated in the caption. The x-axis coordinate of \( \diamond \) is the solution \( S_1^\dagger \) to problem (34) and the corresponding y-axis coordinate is the \( Q_{EC}[j] \) achieved at \( S^\dagger \). We observe that \( S_1^\dagger \) and \( Q_{EC}[j]|_{S=S^\dagger} \) are almost identical to \( S_1^* \) and \( Q_{EC}[j]|_{S=S^*} \), respectively, which confirms the validity of Lemma 1 and Lemma 2, the effectiveness of problem (34), and the accuracy of our method to solve problem (34). In Fig. 6(a), we observe that \( S_1 = 1000 \) achieves the minimal \( Q_{EC}[j] \), which verifies Theorem 3. Interestingly, we observe that from Figs. 6(a)–(d), when we move RX1 towards the TX, the optimal molecule allocation for RX1 first increases and then decreases. This is because, when RX1 approaches to the TX, the TX–RX1–FC link becomes more reliable, so increasing the number of molecules for RX1 optimizes the whole system; and when RX1 is very close to the TX, the TX–RX1–FC link becomes less reliable due to a weak RX1–FC link. In particular, in Fig. 6(c), the optimal solution is to allocate all molecules to RX1. This is because when RX1 is at \((1\mu m, 0.3\mu m, 0\mu m)\), RX1 is very close to the optimal relay location, i.e., the midpoint between the TX and the FC, thus the TX–RX1–FC link is much more reliable than the TX–RX2–FC link and allocating all molecules to RX1 optimizes the whole system.

In Fig. 7, we consider a three-RX system and plot the error probability of SD-ML versus the number of molecules released by RX1, RX2, and RX3. The locations of the three RXs are generated randomly, as indicated in the caption. The x-axis, y-axis, and z-axis coordinates of \( \spadesuit \) are the solutions to problem (34). The corresponding 4th coordinate (i.e., color bar) is the \( Q_{EC}[j] \) achieved at \( S^\dagger \). We observe that \( S^\dagger \) and \( Q_{EC}[j]|_{S=S^\dagger} \) are almost identical to \( S^* \) and \( Q_{EC}[j]|_{S=S^*} \), respectively, which again verifies Lemma 1, Lemma 2, and the effectiveness of problem (34).

VII. CONCLUSIONS

Combined with our initial work in [1], we presented for the first time symbol-by-symbol ML detection for the cooperative diffusion-based MC system with multiple communication phases. We considered the transmission of a sequence of binary symbols and accounted for the resultant ISI in the design and analysis of the cooperative MC system. We presented three ML detectors, i.e., MD-ML, SD-ML, and SA-ML. For practicality, the FC chooses the current symbol using its own local history. For tractability, we derived the system error probabilities for SD-ML and SA-ML using the genie-aided history. We formulated and solved a multi-dimensional optimization problem to find the optimal molecule allocation among RXs that minimizes the system error probability of SD-ML. We analytically proved that the equal distribution of molecules among two symmetric RXs obtains the local minimal error probability of SD-ML. Using numerical and simulation results, we corroborated the accuracy of these analytical expressions and the effectiveness of the formulated optimization problem. Our results revealed trade-offs between
the performance, knowledge of previous symbols, the types of molecule available, relaying modes, and computational complexity.

**APPENDIX A**

**PROOF OF THEOREM 1**

We first prove the decision rule for SD-ML when \( \hat{x}_p[j] > 0 \). To this end, based on (4), we first rewrite the general decision rule for SD-ML as \( W_{rc}[j] = 1 \) if \( \frac{\mathcal{L}_0^{sd}[j]}{\mathcal{L}_0^{sd}[j]} \geq 1 \), otherwise \( \hat{W}_{rc}[j] = 0 \). Thus, if \( \frac{\mathcal{L}_0^{sd}[j]}{\mathcal{L}_0^{sd}[j]} \) is a monotonically increasing function with respect to \( \hat{s}[j] \), then we can obtain the decision rule (14). We can prove that \( \frac{\mathcal{L}_0^{sd}[j]}{\mathcal{L}_0^{sd}[j]} \) is a monotonically increasing function with respect to \( \hat{s}[j] \) by proving that \( \frac{\mathcal{L}_0^{sd}[j]}{\mathcal{L}_0^{sd}[j]} \geq 1 \). Based on (13), we first rewrite \( \mathcal{L}_0^{sd}[j] \) and \( \mathcal{L}_0^{sd}[j] \) as

\[
\mathcal{L}^{sd}_1[j] = \sum_{h_1=1}^{2^K} \left[ \text{Pr}(h_1|1) \exp \left( -\hat{x}_p[j] - \hat{x}_{b,\text{tot}}[j] \right) \right] \left( \hat{s}[j] \right)^{-1},
\]

and

\[
\mathcal{L}^{sd}_0[j] = \sum_{h_0=1}^{2^K} \left[ \text{Pr}(h_0|0) \exp \left( -\hat{x}_p[j] - \hat{x}_{b,\text{tot}}[j] \right) \right] \left( \hat{s}[j] \right)^{-1},
\]

respectively, where \( \text{Pr}(h|b) = \text{Pr}(W_{rx}^{ks}|W_{tx}[j] = b, W_{rx}^{ws-1}) \).

Based on (41) and (42), we find the first derivative of \( \frac{\mathcal{L}_0^{sd}[j]}{\mathcal{L}_0^{sd}[j]} \) with respect to \( \hat{s}[j] \) as

\[
\left( \frac{\mathcal{L}_1^{sd}[j]}{\mathcal{L}_0^{sd}[j]} \right)' = \sum_{h_1=1}^{2^K} \sum_{h_0=1}^{2^K} \left[ \text{Pr}(h_1|1) \text{Pr}(h_0|0) \Pi(h_1, h_0) \right] \left( \mathcal{L}_0^{sd}[j] \left( \hat{s}[j] \right) \right)^{-2},
\]

where

\[
\Pi(h_1, h_0) = \exp \left( -2\hat{x}_p[j] - \hat{x}_{b,\text{tot}}[j] - \hat{x}_{b,\text{tot}}[j] \right) \left( \hat{s}[j] \right)^{-1} \times \log \left( \hat{x}_p[j] + \hat{x}_{b,\text{tot}}[j] \right) \left( \hat{x}_p[j] + \hat{x}_{b,\text{tot}}[j] \right).
\]

We observe that in (43), all terms are positive except for the \( \log(.) \) term. Since \( \log(x) > 0 \) when \( x > 1 \), we separate (43) into two parts: \( \log(.) > 0 \) and \( \log(.) < 0 \). By doing so, we rewrite (43) as the sum of A and B, i.e.,

\[
A = \sum_{h_1=1}^{2^K} \sum_{h_0=1}^{2^K} \left[ \text{Pr}(h_1|1) \text{Pr}(h_0|0) \Pi(h_1, h_0) \right] \left( \mathcal{L}_0^{sd}[j] \left( \hat{s}[j] \right) \right)^{-2}
\]

and

\[
B = \sum_{h_1=1}^{2^K} \sum_{h_0=1}^{2^K} \left[ \text{Pr}(h_1|1) \text{Pr}(h_0|0) \Pi(h_1, h_0) \right] \left( \mathcal{L}_0^{sd}[j] \left( \hat{s}[j] \right) \right)^{-2}.
\]

We further rearrange the summation orders and exchange \( h_1 \) and \( h_0 \) in (46) to rewrite B as

\[
B = \sum_{h_1=1}^{2^K} \sum_{h_0=1}^{2^K} \left[ \text{Pr}(h_0|0) \Pi(h_1, h_0) \right] \frac{\mathcal{L}_0^{sd}[j] \left( \hat{s}[j] \right)}{\left( \mathcal{L}_0^{sd}[j] \left( \hat{s}[j] \right) \right)^2}.
\]

Combining (45) and (47) and applying \( \Pi(h_1, h_0) = -\Pi(h_0, h_1) \), we have

\[
\left( \frac{\mathcal{L}_1^{sd}[j]}{\mathcal{L}_0^{sd}[j]} \right)' = \sum_{h_1=1}^{2^K} \sum_{h_0=1}^{2^K} \left[ \text{Pr}(h_1, h_0) \right] \left( \mathcal{L}_0^{sd}[j] \left( \hat{s}[j] \right) \right)^{-2},
\]

where \( \text{Pr}(h_1, h_0) = \text{Pr}(h_1|1) \text{Pr}(h_0|0) - \text{Pr}(h_0|1) \text{Pr}(h_1|0) \).

We find that (48) holds when \( \text{Pr}(h_1, h_0) > 0 \) is valid, i.e., where \( \hat{x}_{b,\text{tot}}[j] > \hat{x}_{b,\text{tot}}[j] \). We note that \( \hat{x}_{b,\text{tot}}[j] > \hat{x}_{b,\text{tot}}[j] \) leads to \( W_{rx}^{ks} \| x \|_1 \geq \| W_{rx}^{ks} \|_{\eta_1} \), where \( \| x \|_1 \) is the 1-norm of the vector \( x \). When \( \| W_{rx}^{ks} \|_{\eta_1} > \| W_{rx}^{ks} \|_{\eta_1} \) holds, we have \( \text{Pr}(h_1, h_0) = \text{Pr}(h_1|1) \text{Pr}(h_0|0) > \text{Pr}(h_1|0) \), which leads to \( \vartheta(h_1, h_0) > 0 \). Thus, \( \text{Pr}(h_1, h_0) > 0 \) holds if \( \hat{x}_{b,\text{tot}}[j] > \hat{x}_{b,\text{tot}}[j] \). This proves that \( \left( \frac{\mathcal{L}_1^{sd}[j]}{\mathcal{L}_0^{sd}[j]} \right)' > 0 \) and thus proves the decision rule for SD-ML when \( \hat{x}_p[j] > 0 \).

We finally prove the decision rule when \( \hat{x}_p[j] = 0 \). We recall that \( \hat{x}_p[j] = 0 \) means all previous RX symbols are “0”. It probably occurs when all previous TX symbols are “0” (i.e., no ISI at RXs) if the error probability of the first phase is small. Hence, there is no likelihood that “1” is detected at RXs when “0” is transmitted by the TX, which leads to \( \text{Pr}(W_{rx}^{ks}|W_{tx}[j] = 0, W_{rx}^{ws-1}) \approx 1 \) and \( \text{Pr}(W_{rx}^{ks} \neq 0|W_{tx}[j] = 0, W_{rx}^{ws-1}) \approx 0 \). Using these approximations and \( \hat{x}_p[j] = 0 \), we approximate \( \mathcal{L}_0^{sd}[j] \approx \exp(0)(0)^{\hat{s}[j]} \). When \( \hat{x}_p[j] = 0 \) and \( \hat{s}[j] > 0 \), \( \mathcal{L}_0^{sd}[j] > 0 \), thus the decision at the FC is always \( \hat{W}_{rc}[j] = 1 \) since \( \mathcal{L}_0^{sd}[j] > 0 \).

**APPENDIX B**

**PROOF OF LEMMA 2**

We take the first derivative of (32) with respect to \( \xi \). However, \( Q_t^{rc}[j] \) is a discrete function with respect to \( \xi \), which makes \( Q_t^{rc}[j] \) not differentiable in terms of \( \xi \). To tackle this challenge, we approximate the sum in (33) with an integral with respect to \( \eta \), i.e.,

\[
A \approx \int_{\eta_0}^{\xi} \exp \left( -\hat{x}_p[j] - \hat{x}_{b,\text{tot}}[j] \right) \left( \hat{x}_p[j] + \hat{x}_{b,\text{tot}}[j] \right)^{\eta} \eta \exp \left( -\hat{x}_p[j] - \hat{x}_{b,\text{tot}}[j] \right) d\eta.
\]

Using the continuous approximation of \( A \) in (49) and \( \partial \int_{\eta_0}^{\xi} f(t)dt/\partial x = f(x) \), we take the first derivative of (32) with respect to \( \xi \) as \( \partial Q_t^{rc}[j]/\partial \xi = R_{\psi_1}(\xi) - (1 - R_{\psi_2}(\xi)) \psi_b(\xi) \), where \( \psi_b(\xi) \) is defined by

\[
\psi_b(\xi) = \sum_{b=0}^{2^K} \left[ \text{Pr}(W_{rx}^{ks}|W_{tx}[j] = b, W_{rx}^{ws-1}) \right] \left( \xi \right)^{-1} \times \exp \left( -\hat{x}_p[j] - \hat{x}_{b,\text{tot}}[j] \right) \left( \hat{x}_p[j] + \hat{x}_{b,\text{tot}}[j] \right) \xi.
\]
Comparing (50) with (13), we find that \( \psi_b(s[j]) = L^0_{\text{opt}}[j] \).

We recall that \( \xi_{\text{rc}}[j] \) is the solution to \( L^0_{\text{opt}}[j] = L^0[j] \) in terms of \( s[j] \). Hence, \( \xi_{\text{rc}}[j] \) is the solution to \( P_1 \psi_1(\xi) - (1 - P_1) \psi_2(\xi) = \partial Q^T_{\text{rc}}[j]/\partial \xi = 0 \) if \( P_1 = \frac{1}{2} \). Therefore, \( \xi_{\text{rc}}[j] \) is the optimal \( \xi \) which minimizes (32).

**APPENDIX C
Proof of Proposition 1**

The problem (34) has \( K + 1 \) optimization variables and the evaluation of its Hessian requires very high computational complexity. To decrease the complexity, we first consider the simplest case with \( K = 2 \) and investigate the Hessian of \( Q^T_{\text{rc}}[j] \) with respect to \( S_1 \) for a fixed \( \xi \). To this end, we take the first derivative of \( Q^T_{\text{rc}}[j] \) with respect to \( S_1 \). In (32), \( \Lambda \) is a discrete function in terms of \( S_1 \), which makes the derivative cumbersome. If we approximate \( \Lambda \) using (49), there is no closed-form for the first derivative of (49) with respect to \( S_1 \). To overcome this challenge, we approximate \( \Lambda \) by another continuous approximation, i.e., the continuous regularized incomplete Gamma function. By doing so, we have

\[
\Lambda \approx \frac{\Gamma\left(\lfloor \xi \rfloor, \lambda x, 0 \right) + \lambda x, \lfloor \xi \rfloor \Theta S_1}{\Gamma\left(\lfloor \xi \rfloor\right)},
\]

where \( \Gamma(\gamma, \delta) \) is the incomplete Gamma function and the Gamma function \( \Gamma(\gamma) \) is a special case of \( \Gamma(\gamma, \delta) \) with \( \delta = 0 \). Applying this approximation to (32), we obtain the continuous approximation of \( Q^T_{\text{rc}}[j] \). Using \( \partial \Gamma(\gamma, \delta)/\partial \delta = -\exp(\gamma - \delta)\delta^{-1} \), we take the first derivative of \( Q^T_{\text{rc}}[j] \) as

\[
\frac{\partial Q^T_{\text{rc}}[j]}{\partial S_1} \approx \frac{1}{\Gamma\left(\lfloor \xi \rfloor\right)} \left( \sum_{a_1=0}^{\lfloor \xi \rfloor} \sum_{a_2=0}^{\lfloor \xi \rfloor} \left( \left( (1 - P_1) \alpha(a_1, a_2) - P_1 \beta(a_1, a_2) \right) \exp\left( -\Xi(a_1, a_2) \right) \times \left( \Xi(a_1, a_2) \right)^{-1+\lfloor \xi \rfloor} \Omega(a_1, a_2) \right) \right),
\]

where \( \Xi(a_1, a_2) = \sigma_1 s_1 + \sigma_2 (N - s_1) + \sigma_3 s_1 + \sigma_4 (N - s_1) \) and \( \Omega(a_1, a_2) = \sigma_5 - \sigma_2 + \sigma_3 l_{\text{th}} - \sigma_4 s_{\text{th}} \), where \( \sigma_k \) and \( \nu_k \) are given in (37) and (38), respectively. We then find the second derivative of \( Q^T_{\text{rc}}[j] \) with respect to \( S_1 \) as

\[
\frac{\partial^2 Q^T_{\text{rc}}[j]}{\partial S_1^2} = -\frac{1}{\Gamma\left(\lfloor \xi \rfloor\right)} \left( \sum_{a_1=0}^{\lfloor \xi \rfloor} \sum_{a_2=0}^{\lfloor \xi \rfloor} \exp\left( -\Xi(a_1, a_2) \right) \times \left( \Xi(a_1, a_2) \right)^{-2+\lfloor \xi \rfloor} \left( \Omega(a_1, a_2) \right)^2 \varpi \right),
\]

where

\[
\varpi = \left( (1 - P_1) \alpha(a_1, a_2) - P_1 \beta(a_1, a_2) \right) \left( 1 - \lfloor \xi \rfloor + \Xi(a_1, a_2) \right).
\]

In (53), all terms are nonnegative except for \( \varpi \). Thus, if \( \varpi > 0 \) holds for each summand, (53) is nonnegative. However, the condition \( \varpi > 0 \) is not always valid for \( a_1 \in \{0, 1\} \) and \( a_2 \in \{0, 1\} \). Thus, for a fixed \( \xi \), \( \partial^2 Q^T_{\text{rc}}[j]/\partial \xi^2 \) is not always nonnegative, which means that the Hessian of \( Q^T_{\text{rc}}[j] \) with respect to \( S_1 \) is not always positive semidefinite.

**APPENDIX D
Proof of Lemma 3**

Using (36), we simplify (52) as

\[
\frac{\partial Q^T_{\text{rc}}[j]}{\partial S_1} = \frac{\exp(-\Phi_1 - 2\nu S_1)}{\Phi_1 \Phi_2 \Gamma\left(\lfloor \xi \rfloor\right)} \left( \alpha(P_1 - 1) + \beta P_1 \right) \nu \times \left( \exp(2\nu S_1) \Phi_1^{\lfloor \xi \rfloor} \Phi_2 - \exp(N \nu) \Phi_2^{\lfloor \xi \rfloor} \Phi_1 \right),
\]

where \( \Phi_1 = N(\nu + \sigma) - \nu S_1 \) and \( \Phi_2 = N \sigma + \nu S_1 \). It can be shown that \( S_1 = \frac{N}{2} \) is the one of the solutions of (55). Hence, \( Q^T_{\text{rc}}[j] \) has a local minimum or maximum when \( S_1 = \frac{N}{2} \). We then apply (36) and \( S_1 = \frac{N}{2} \) to (53) to obtain the second derivative of \( Q^T_{\text{rc}}[j] \) at \( S_1 = \frac{N}{2} \). By doing so, we have

\[
\frac{\partial^2 Q^T_{\text{rc}}[j]}{\partial S_1^2} \bigg|_{S_1=\frac{N}{2}} = \frac{4 \exp(-\frac{1}{2} N(\nu + 2\sigma))}{N^2(\nu + 2\sigma)^2 \Gamma\left(\lfloor \xi \rfloor\right)} \times \nu^2(2\nu(\nu + 2\sigma))\lfloor \xi \rfloor \Gamma\left(\lfloor \xi \rfloor\right),
\]

where all terms are nonnegative except for \( \lambda \). Hence if \( \lambda > 0 \), (56) is nonnegative and \( Q^T_{\text{rc}}[j] \) achieves a local minimum at \( S_1 = \frac{N}{2} \); otherwise, it achieves a local maximum.

**APPENDIX E
Proof of Theorem 3**

Based on Lemma 3, \( Q^T_{\text{rc}}[j] \) achieves a local minimal value at \( S_1 = \frac{N}{2} \) when \( \lambda = 0 \). Based on Lemma 1, the approximation of \( Q_{\text{rc}}[j] \) by \( Q^T_{\text{rc}}[j] \) is tight when \( \xi = \xi_{\text{rc}}[j] \). Thus, we can prove that \( Q_{\text{rc}}[j] \) always achieves a local minimal value at \( S_1 = \frac{N}{2} \) by proving that \( \lambda \) always holds when \( \xi = \xi_{\text{rc}}[j] \). That is to say, we need to prove \( \lambda(\xi_{\text{rc}}[j]) > 0 \). Based on the proof of Lemma 2, we also recall that \( \xi_{\text{rc}}[j] \) is the solution to \( P_1 \psi_1(\xi) - (1 - P_1) \psi_2(\xi) = \partial Q^T_{\text{rc}}[j]/\partial \xi = 0 \) if \( P_1 = \frac{1}{2} \). Thus, \( \xi_{\text{rc}}[j] \) satisfies the condition: \( \psi_1(\xi_{\text{rc}}[j]) = \psi_2(\xi_{\text{rc}}[j]) = 0 \). Applying \( S_1 = N/2 \) to (50), we write \( \psi_b(\xi_{\text{rc}}[j]) \), where \( b \in \{0, 1\} \), using \( \nu \) and \( \sigma \) as

\[
\psi_b(\xi_{\text{rc}}[j]) = (\xi_{\text{rc}}[j])^{-1} \left[ \Pr(0, 0|b) \exp(-\nu N) \left( \sigma N \right)^{\xi_{\text{rc}}[j]} \right]
\]

\[
+ \Pr(0, 1|b) \exp(-\nu N) \left( \sigma + \nu N \right)^{\xi_{\text{rc}}[j]} - \Pr(1, 0|b) \exp(-\nu N) \left( \sigma N + \nu N \right)^{\xi_{\text{rc}}[j]},
\]

where \( \psi_b(\xi_{\text{rc}}[j]) = \Pr(\gamma_{\text{rc}}[j]|b) W_{\text{tx}}[j] = b, W_{j}^{-1} \). In (57), we have \( \Pr(0, 1|0) = \Pr(1, 0|0) = \alpha \) and \( \Pr(0, 1|1) = \Pr(1, 0|1) = \beta \) based on (36). We then approximate \( \Pr(0, 0|0) = \Pr(1, 1|1) \approx 1 \) and \( \Pr(0, 0|1) = \Pr(1, 1|0) \approx 0 \), which is tight when the error probability of the TX–RX link is small.
Using these approximations, α, and β, we rewrite \(\psi(\zeta_{SC})\) and \(\psi_{1}(\zeta_{SC})\) as
\[
\psi(\zeta_{SC}) \approx (\zeta_{SC})^{-1} \exp(-\sigma N) (\sigma N)^{\zeta_{SC}} + 2\alpha (\zeta_{SC})^{-1} \exp(-\sigma N - \nu N/2) \times (\sigma N + \nu N/2)^{\zeta_{SC}} \tag{58}
\]
and
\[
\psi_{1}(\zeta_{SC}) \approx (\zeta_{SC})^{-1} \exp(-\sigma N - \nu N) \times (\sigma N + \nu N)^{\zeta_{SC}} + 2\beta (\zeta_{SC})^{-1} \exp(-\sigma N - \nu N/2) \times (\sigma N + \nu N/2)^{\zeta_{SC}} \tag{59}
\]
respectively. Using \(\psi_{1}(\zeta_{SC}) - \psi_{2}(\zeta_{SC}) = 0\) and some basic manipulations, we obtain
\[
\beta - \alpha = \frac{1}{2} \exp\left(-\frac{\nu N}{2}\right) \left(\frac{\sigma N + \nu N}{\sigma N + \nu N/2}\right)^{\zeta_{SC}} \tag{60}
\]
Applying (60) and \(P_1 = \frac{1}{2}\) to (35), we have
\[
Y(\zeta_{SC}) = \frac{\theta_1 \theta_2}{4} \exp\left(-\frac{\nu N}{2}\right) \left(\frac{\sigma N + \nu N}{\sigma N + \nu N/2}\right)^{\zeta_{SC}} \tag{61}
\]
where \(\theta_1 = \left(\exp(\nu N) / (\sigma N + \nu N)\right)^{\zeta_{SC}} - 1\) and \(\theta_2 = \left(2 + N(\nu + 2\sigma - 2\zeta_{SC})\right)^{\zeta_{SC}}\). To prove \(Y(\zeta_{SC}) > 0\), we only need to prove \(\theta_1 \theta_2 > 0\), since all other terms in (61) are nonnegative. If \(\theta_1 > 0\), we have \(\zeta_{SC} < \nu N/\log((\sigma N + \nu N)/\nu N)\). Applying \(x \leq x - 1\), when \(x > 0\), we further lower-bound \(\zeta_{SC}\) by \(\frac{\sigma N + \nu N}{\sigma N + \nu N/2}\), which leads to \(\theta_2 < 0\). If \(\theta_1 < 0\), we have \(\zeta_{SC} > \nu N/\log((\sigma N + \nu N)/\nu N)\). Applying \(x \geq 1 - 1/x\), when \(x > 0\), we further upper-bound \(\zeta_{SC}\) by \(\sigma N + \nu N\), which leads to \(\theta_2 < 0\) if \(\nu N/2 > 1\). Although the validity of \(\nu N/2 > 1\) depends on the value of \(\nu N\), it is generally valid. This is because \(\nu N/2 > 1\) means that at least one signaling molecule is expected to be processed at the FC if the decision at RXk is “1” and it is a reasonable condition to be satisfied. Since \(\theta_1\) and \(\theta_2\) are always both negative or positive, \(\theta_1 \theta_2 > 0\) holds, which leads to \(Y(\zeta_{SC}) > 0\). Therefore, \(Q_{SC}\) achieves a local minimum at \(S_1 = N/2\) when \(K = 2\) in a symmetric topology.

REFERENCES

[1] Y. Fang, A. Noel, N. Yang, A. W. Eckford, and R. A. Kennedy, “Maximum likelihood detection for cooperative molecular communication,” in Proc. IEEE ICC, May 2018, pp. 1–7.

[2] N. Farsad, H. B. Yilmaz, A. Eckford, C. B. Chae, and W. Guo, “A comprehensive survey of recent advancements in molecular communication,” IEEE Commun. Surveys Tuts., vol. 18, no. 3, pp. 1887–1919, Aug. 2016.

[3] T. Nakano and J. Q. Liu, “Design and analysis of molecular relay channels: An information theoretic approach,” IEEE Trans. Nanobiosci., vol. 9, no. 3, pp. 213–221, Sep. 2010.

[4] B. Atakan and O. B. Akan, “On molecular multiple-access, broadcast, and relay channels in nanonetworks,” in Proc. ICST BIONETICS, Nov. 2008, pp. 16:1–16:8.

[5] T. Nakano, Y. Okai, and A. V. Vasilakos, “Transmission rate control for molecular communication among biological nanomachines,” IEEE J. Select. Areas Commun., vol. 31, no. 12, pp. 835–846, Dec. 2013.

[6] C. T. Chou, “Extended master equation models for molecular communication networks,” IEEE Trans. Nanobiosci., vol. 12, no. 2, pp. 79–92, June 2013.

[7] B. H. Koo, C. Lee, H. B. Yilmaz, N. Farsad, A. W. Eckford, and C.-B. Chae, “Molecular MIMO-FC: From theory to prototype,” IEEE J. Select. Areas Commun., vol. 34, no. 3, pp. 600–614, Mar. 2016.

[8] Y. Fang, A. Noel, N. Yang, A. W. Eckford, and R. A. Kennedy, “Distributed cooperative detection for multi-receiver molecular communication,” in Proc. IEEE GLOBECOM, Dec. 2016, pp. 1–7.

[9] ——, “Convex optimization of distributed cooperative detection in multi-receiver molecular communication,” IEEE Trans. Mol. Bio. Multi-Scale Commun., vol. 3, no. 1, pp. 1–12, Apr. 2017.

[10] Y. Fang, A. Noel, Y. Wang, and N. Yang, “Simplified cooperative detection for multi-receiver molecular communication,” in Proc. IEEE ITW, Nov. 2017, pp. 1–5.

[11] J. G. Proakis, Digital Communication, 4th ed. New York: McGraw-Hill, 2000.

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

[13] A. Noel, K. C. Cheung, and R. Schober, “Optimal receiver design for diffusive molecular communication with flow and additive noise,” IEEE Trans. Nanobiosci., vol. 13, no. 3, pp. 350–362, Sept. 2014.

[14] M. U. Mahfuz et al., “A comprehensive analysis of strength-based optimum signal detection in concentration-encoded molecular communication with spike transmission,” IEEE Trans. Nanobiosci., vol. 14, no. 1, pp. 67–83, Jan. 2015.

[15] A. Singhal, R. K. Mallik, and B. Lall, “Performance analysis of amplitude modulation schemes for diffusion-based molecular communication,” IEEE Trans. Wireless Commun., vol. 14, no. 10, pp. 5681–5691, Oct. 2015.

[16] S. Ghavami and F. Lahouti, “Abnormality detection in correlated gaussian molecular nano-networks: Design and analysis,” IEEE Trans. Nanobiosci., vol. 16, no. 3, pp. 383–396, Jun. 2017.

[17] T. C. Mai, M. Egan, T. Q. Duong, and M. D. Renzo, “Event detection in molecular communication networks with anomalous diffusion,” IEEE Commun. Lett., vol. 21, no. 6, pp. 1249–1252, June 2017.

[18] R. Mosayebi, V. Jamali, N. Ghoroghchian, R. Schober, M. Nasiri-Kenari, and M. Mehrab, “Cooperative abnormality detection via diffusive molecular communications,” IEEE Trans. Nanobiosci., vol. 16, no. 8, pp. 828–842, Dec. 2017.

[19] U. Rogers and M. S. Koh, “Parallel molecular distributed detection with brownian motion,” IEEE Trans. Nanobiosci., vol. 15, no. 8, pp. 871–880, Dec. 2016.

[20] A. P. de Silva et al., “Molecular logic and computing,” Nature Nanotech., vol. 2, no. 7, pp. 399–410, Jul. 2007.

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

[22] A. Ahmadzadeh, A. Noel, A. Burkovski, and R. Schober, “Amplify-and-forward relaying in two-hop diffusion-based molecular communication networks,” in Proc. IEEE GLOBECOM, Dec. 2015, pp. 1–7.

[23] P. Siuti, J. Yazbek, and T. K Lu, “Synthetic circuits integrating logic and memory in living cells,” Nature nanotechnology, vol. 31, pp. 448–452, Feb. 2013.

[24] O. Mondragon-Palomino, T. Danino, J. Selimkhah, L. Tsimring, and J. Hasty, “Entrainment of a population of synthetic genetic oscillators,” Sci., vol. 333, no. 6047, pp. 1315–1319, Sep. 2011.

[25] H. ShahMohammadian, G. G. Messier, and S. Magierowski, “Blind synchronization in diffusion-based molecular communication channels,” IEEE Commun. Lett., vol. 17, no. 11, pp. 2156–2159, Nov. 2013.

[26] S. Abadal et al., “Bio-inspired synchronization for nanonetwork communication,” in Proc. IEEE GLOBECOM, Dec. 2011, pp. 1–5.

[27] V. Lu, M. D. Higgins, and M. S. Leeson, “Comparison of channel coding schemes for molecular communications systems,” IEEE Trans. Commun., vol. 63, no. 11, pp. 3991–4001, Nov. 2015.
[28] A. Noel, K. C. Cheung, and R. Schober, “Using dimensional analysis to assess scalability and accuracy in molecular communication,” in Proc. IEEE ICC, June 2013, pp. 818–823.

[29] M. Pierobon and I. F. Akyildiz, “Diffusion-based noise analysis for molecular communication in nanonetworks,” IEEE Trans. Signal Processing, vol. 59, no. 6, pp. 2532–2547, Jun. 2011.

[30] R. Mosayebi et al., “Receivers for diffusion-based molecular communication: Exploiting memory and sampling rate,” IEEE J. Select. Areas Commun., vol. 32, no. 12, pp. 2368–2380, Dec. 2014.

[31] A. Ahmadzadeh, A. Noel, and R. Schober, “Analysis and design of multi-hop diffusion-based molecular communication networks,” IEEE Trans. Mol. Biol. Multi-Scale Commun., vol. 1, no. 2, pp. 144–157, June 2015.

[32] S. S. Andrews and D. Bray, “Stochastic simulation of chemical reactions with spatial resolution and single molecule detail,” Physical Biology, vol. 1, no. 3, pp. 135–151, Aug. 2004.

[33] G. Aminian, H. Ghourchian, A. Gohari, M. Mirmohseni, and M. Nasirikenari, “On the capacity of signal dependent noise channels,” in Proc. IEEE IWCMC, May 2017, pp. 1–6.