Non-Coherent Detection for Diffusive Molecular Communications

Vahid Jamali†, Nariman Farsad‡, Robert Schober†, and Andrea Goldsmith‡

†Friedrich-Alexander University (FAU), Erlangen, Germany

‡Stanford University, Stanford, California, USA

Abstract

We study non-coherent detection schemes for molecular communication (MC) systems that do not require knowledge of the channel state information (CSI). In particular, we first derive the optimal maximum likelihood (ML) multiple-symbol (MS) detector for MC systems. As a special case of the optimal MS detector, we show that the optimal ML symbol-by-symbol (SS) detector can be equivalently written in the form of a threshold-based detector, where the optimal decision threshold is constant and depends only on the statistics of the MC channel. The main challenge of the MS detector is the complexity associated with the calculation of the optimal detection metric. To overcome this issue, we propose an approximate MS detection metric which can be expressed in closed form. To reduce complexity even further, we develop a non-coherent decision-feedback (DF) detector and a suboptimal blind detector. Finally, we derive analytical expressions for the bit error rate (BER) of the optimal SS detector, as well as upper and lower bounds for the BER of the optimal MS detector. Simulation results confirm the analysis and reveal the effectiveness of the proposed optimal and suboptimal detection schemes compared to a benchmark scheme that assumes perfect CSI knowledge, particularly when the number of observations used for detection is sufficiently large.

Index Terms

Molecular communications, channel state information, non-coherent detection, blind detection, and Poisson channel.

I. Introduction

Molecular communication (MC) has recently emerged as a bio-inspired approach for synthetic communication systems having nano/micrometer scale dimensions [2]–[4]. Unlike conventional

wireless communication systems that employ electromagnetic waves to convey information, MC systems encode information in the number, type, or time of release of signalling molecules. Calcium signaling of neurons and the exchange of autoinducers by bacteria in quorum sensing are among the many examples of MC in nature [3], [5].

The effect of the diffusive MC channel on a concentration-based MC system is reflected in *i)* the *probability* that a molecule released by the transmitter is observed at the receiver, and *ii)* the *expected* number of interfering molecules observed at the receiver. Hence, *i)* and *ii)* constitute the channel state information (CSI) of the MC channel. The CSI of an MC system depends on parameters such as the diffusion coefficient of the information molecules, the velocity of the flow in the MC channel, the concentration of enzyme in the environment, the distance between the transmitter and the receiver, and the type of the adopted receiver, among other factors, see [6, Chapters 3 and 4] and [7, Chapter 4]. Most existing works on MC assume that the CSI is perfectly known at the receiver for reliable detection of the transmitted information bits [8]–[11]. In reality, the CSI is not known a priori and has to be estimated. To this end, a training-based CSI estimation framework was developed in [12] where several optimal and suboptimal estimators were proposed. However, we note that the CSI of MC systems may change over time due to factors such as variations in the velocity of the flow, the temperature (which leads to variations in the diffusion coefficient and the enzyme concentration), and the distance between the transmitter and the receiver (e.g. if they are suspended in a fluid) [7], [13], [14]. Therefore, the CSI acquisition has to be conducted repeatedly to keep track of CSI variations. Detection schemes requiring the acquisition of the CSI are suitable options only if the coherence time of the MC channel is sufficiently large (e.g. when the temperature, the flow, and the positions of the transmitter and the receiver change very slowly) such that the corresponding training overhead is tolerable. On the other hand, for the case when the MC channel changes rapidly, CSI estimation either entails a large overhead or results in a low CSI estimation quality. In this case, directly detecting the data symbols without spending any resources on CSI acquisition is an attractive option which is referred to as *non-coherent* detection.

In this paper, our focus is the design of optimal and suboptimal non-coherent detection schemes that do not require knowledge of the instantaneous CSI. In particular, we first derive the optimal maximum likelihood (ML) multiple-symbol (MS) detector. As a special case of the optimal MS detector, we show that the optimal ML symbol-by-symbol (SS) detector can be equivalently written in the form of a threshold-based detector where the optimal decision threshold is constant and

depends only on the statistics of the MC channel. One of the main challenges in implementing the MS detector is the computation of the detection metrics, which requires the evaluation of expectations with respect to the probability density function (PDF) of the CSI. The PDF of the CSI depends on the considered MC environment and a general analytical expression is not known. In practice, the PDF of the CSI for a particular MC channel can be estimated using empirical measurements of the CSI. To obtain a closed-form expression for the MS detection metric, we approximate the PDF of the CSI by the Gamma distribution. We can show that the Gamma distribution accurately matches the exact distribution of the CSI for several examples of stochastic MC channels. Additionally, to reduce complexity even further, we develop a non-coherent decision-feedback (DF) detector and a suboptimal blind detector. Furthermore, we derive analytical expressions for the bit error rate (BER) of the optimal SS detector, and an upper bound and a lower bound for the BER of the optimal MS detector. Our simulation results reveal the effectiveness of the proposed optimal and suboptimal non-coherent detectors and show that their performance approaches that of a benchmark scheme that assumes perfect CSI as the number of symbols used for detection increases.

In contrast to MC, for conventional wireless communication, there is a rich literature on non-coherent MS detection, see e.g. [15], [16], and the references therein. In particular, in [15], MS differential detection without CSI was presented for radio frequency (RF) communications, and in [16], non-coherent MS detection for a photon-counting receiver was studied for optical communications. We note that the detection problem and the resulting detection strategies developed for conventional wireless communications are not applicable to the corresponding MC detection problem due to the fundamental differences in the two systems. The problem of non-coherent data detection in MC was considered before in [17], [18] and heuristic low-complexity non-coherent symbol-by-symbol detectors were proposed. However, to the best of the authors' knowledge, the design of *non-coherent MS detectors*, which is considered in this paper and its conference version [1], has not yet been considered in the MC literature. In addition to the detectors proposed in [1], this paper derives the BER of the optimal SS detector along with lower and upper bounds for the BER of the optimal MS detector, and validates these results with simulations.

The remainder of this paper is organized as follows. In Section II, the system model and the CSI model are introduced. The proposed detectors are presented in Section III and their performance is analyzed in Section IV. Numerical results are presented in Section V, and conclusions are drawn in Section VI.

Notation: We use the following notation throughout this paper: $\mathbb{E}_x\{\cdot\}$ denotes expectation with respect to random variable (RV) x and $|\cdot|$ represents the cardinality of a set. Bold letters are used to denote vectors, $\mathbf{a}^T$ represents the transpose of vector $\mathbf{a}$, and $\mathbf{a} \geq \mathbf{0}$ specifies that all elements of vector $\mathbf{a}$ are non-negative. Moreover, $\mathbb{N}$ and $\mathbb{R}^+$ denote the sets of non-negative integer numbers and positive real numbers, respectively. $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ denote the floor and ceiling functions which map a real number to the largest previous and the smallest following integer number, respectively. $\Gamma(\cdot)$ is the Gamma function and $\mathbf{1}\{\cdot\} \in \{0,1\}$ is an indicator function which is equal to one if the argument is true and equal to zero if it is not true. Moreover, $\mathrm{Poiss}(\lambda)$ denotes a Poisson RV with mean λ, $\mathrm{Bin}(n,p)$ denotes a binomial RV for n trials and success probability p, $\mathcal{N}(\mu,\sigma^2)$ denotes a Gaussian RV with mean μ and variance σ^2, and $\mathrm{Gamma}(\alpha,\beta)$ denotes a Gamma distributed RV with scale parameter α and rate parameter β.

II. SYSTEM MODEL

In this section, we introduce the MC channel model and the CSI model used in this paper.

A. Channel Model

We consider an MC system consisting of a transmitter, a channel, and a receiver, see Fig. 1. At the beginning of each symbol interval, the transmitter releases either N^{TX} or zero molecules corresponding to the binary bits 1 and 0, respectively, i.e., the modulation format used is ON-OFF keying (OOK) [2]. In this paper, we assume that the transmitter emits only one type of molecule. The released molecules diffuse through the fluid medium between the transmitter and the receiver. The receiver counts the number of observed molecules in each symbol interval. We note that this is a rather general model for the MC receiver which includes well-known receivers such as the transparent receiver [11], the absorbing receiver [19], and the reactive receiver [20]. In particular, the number of observed molecules at the receiver in symbol interval k, denoted by $r[k]$, is given by [12], [21]–[23]

$$r[k] = c_{\mathbf{s}}[k] + c_{\mathbf{n}}[k], \tag{1}$$

where $c_{\mathbf{s}}[k]$ is the number of molecules observed at the receiver in symbol interval k due to the release of $s[k]N^{\mathrm{Tx}}$ molecules by the transmitter at the beginning of symbol interval k, with $s[k] \in \{0,1\}$. We assume that the binary information bits are equiprobable, i.e., $\mathrm{Pr}\{s[k]=1\} = \mathrm{Pr}\{s[k]=0\} = 0.5$. Moreover, $c_{\mathbf{n}}[k]$ is the number of interfering noise molecules comprising
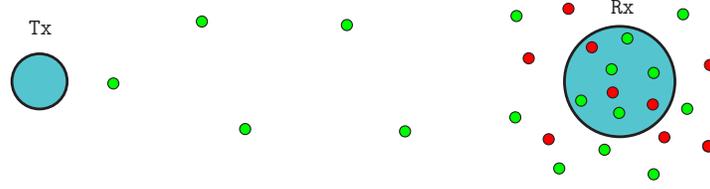
Fig. 1. Schematic illustration of the considered MC system where the molecules released by the transmitter in a given symbol interval are shown in green color whereas the noise molecules are shown in red color.

residual inter-symbol interference (ISI), multiuser interference (caused by other MC links), and external noise (originating from natural sources) observed by the receiver in symbol interval k.

The MC channel is dispersive due to the diffusive propagation of the molecules [24]. The ISI-free communication model in (1) implies that the symbol intervals are chosen sufficiently large such that the channel impulse response (CIR) fully decays to zero within one symbol interval. We note that enzymes [24] and reactive information molecules, such as acid/base molecules [25], may be used to speed up the decaying of the CIR as a function of time. Nevertheless, since the length of the symbol intervals is finite, some residual ISI always exists. Throughout this paper, we assume that the effect of the residual ISI is included in $c_{\mathrm{n}}[k]$ and is sufficiently small compared to the other components of $c_{\mathrm{n}}[k]$ such that $c_{\mathrm{n}}[k]$ is (approximately) independent of the signal $c_{\mathrm{s}}[k]$. We now describe the underlying models and associated distributions of RVs $c_{\mathrm{s}}[k]$ and $c_{\mathrm{n}}[k]$.

Information Molecules $(c_{\mathrm{s}}[k])$: The movements of individual molecules are assumed to be independent from each other. We further assume that the observations of different molecules at the receiver are independent[1]. Therefore, from a probabilistic point of view, we can assume that each molecule released by the transmitter in a given symbol interval is observed at the receiver in the same symbol interval with a certain probability, denoted by p_{s}. Since any given molecule released by the transmitter is either observed by the receiver or not, a binary state model applies and the number of observed molecules follows the binomial distribution $\mathrm{Bin}(N^{\mathrm{Tx}}, p_{\mathrm{s}})$. Moreover, assuming that N^{Tx} is very large while $N^{\mathrm{Tx}} p_{\mathrm{s}} \triangleq \bar{c}_{\mathrm{s}}$ is relatively small, the binomial distribution $\mathrm{Bin}(N^{\mathrm{Tx}}, p_{\mathrm{s}})$ converges to the Poisson distribution $\mathrm{Poiss}\,(\bar{c}_{\mathrm{s}})$ [26]. Another approximation of the binomial distribution $\mathrm{Bin}(N^{\mathrm{Tx}}, p_{\mathrm{s}})$ is the Gaussian distribution $\mathcal{N}(N^{\mathrm{Tx}} p_{\mathrm{s}}, \sqrt{N^{\mathrm{Tx}} p_{\mathrm{s}}(1 - p_{\mathrm{s}})})$ which holds for very large N^{Tx} when $\bar{c}_{\mathrm{s}}$ is relatively large [26]. We note that the assumptions for the

---

[1]This assumption holds for the transparent and absorbing receiver models. For this assumption to hold for the reactive receiver model in [20], the number of receptors on the surface of the receiver has to be much larger than the average number of molecules around the receiver which is a valid assumption for typical biological cells in nature, see [6, Page 33].

Poisson approximation are more justified for MC than those of the Gaussian approximation since the number of released molecules is typically very large (on the order of hundreds or a few thousands) but, typically, only few molecules reach the receiver. Therefore, we adopt the Poisson approximation in this paper, i.e., $c_{\mathrm{s}}[k] \sim \mathrm{Poiss}\,(\bar{c}_{\mathrm{s}}s[k])$. The accuracy of the Poisson distribution in modeling the number of molecules observed at the receiver was verified in [21], [8], [10] and compared with the accuracy of the corresponding binomial and Gaussian models[2].

Noise Molecules ($c_{\mathrm{n}}[k]$): Noise molecules originate from interfering natural or synthetic sources [27]. Suppose there are in total $N^{\mathtt{noise}}$ noise molecules in the environment and they are uniformly distributed in space and time, which is a reasonable assumption if a priori knowledge about the locations and activities of the noise sources is not available. In addition, the probability that at any given time, any noise molecule is observed at the receiver is assumed to be the same for all noise molecules and is denoted by p_{n}. Since at the sampling time, each noise molecule is either observed at the receiver or not, again, a two state model is valid and $c_{\mathrm{n}}[k]$ follows a binomial distribution, i.e., $c_{\mathrm{n}}[k] \sim \mathrm{Bin}(N^{\mathtt{noise}}, p_{\mathrm{n}})$. Moreover, the probability that a given noise molecule is observed at the receiver is again expected to be very small, i.e., $p_{\mathrm{n}} \to 0$ holds. Therefore, $N^{\mathtt{noise}} \to \infty$ has to hold such that the average number of noise molecules observed at the receiver, denoted by $\bar{c}_{\mathrm{n}} = N^{\mathtt{noise}}p_{\mathrm{n}}$, is non-zero. Considering this, $c_{\mathrm{n}}[k]$ follows also a Poisson distribution, i.e., $c_{\mathrm{n}}[k] \sim \mathrm{Poiss}\,(\bar{c}_{\mathrm{n}})$.

Remark 1: The channel model in (1) can be generalized to the case of multiple-sample detection if the following sum detector is employed:

$$r[k] = \sum_{m=1}^{M} y[k,m] = \sum_{m=1}^{M} c_{\mathrm{s}}[k,m] + \sum_{m=1}^{M} c_{\mathrm{n}}[k,m] \triangleq c_{\mathrm{s}}[k] + c_{\mathrm{n}}[k], \tag{2}$$

where M denotes the number of samples per symbol interval and $y[k,m]$ is the number of molecules observed at the receiver in the m-th sample of symbol interval k. Moreover, $c_{\mathrm{s}}[k,m]$ is the number of molecules observed at the receiver for the m-th sample of symbol interval k due to the release of $s[k]N^{\mathrm{Tx}}$ molecules by the transmitter at the beginning of symbol interval k, and $c_{\mathrm{n}}[k,m]$ is the number of noise molecules observed at the receiver for the m-th sample of symbol interval k. Since $c_{\mathrm{s}}[k,m]$ and $c_{\mathrm{n}}[k,m]$ are Poisson RVs, $c_{\mathrm{s}}[k]$ and $c_{\mathrm{n}}[k]$ follow Poisson

---

[2]For instance, the analytical framework developed in [21] can be used to show that for $N^{\mathrm{Tx}} = 1000$, if $p_{\mathrm{s}} \leq 0.115$ holds, the Poisson distribution more accurately approximates the binomial distribution in terms of the root mean squared error (RMSE) of the cumulative distribution function (CDF), whereas, if $p_{\mathrm{s}} > 0.115$ holds, the Gaussian approximation is a better fit. For MC systems, if $N^{\mathrm{Tx}} = 1000$ molecules are released by the transmitter, typically we expect to observe much fewer than $N^{\mathrm{Tx}}p_{\mathrm{s}} = 113$ molecules at the receiver. Hence, the Poisson approximation is more accurate compared to the Gaussian approximation in this case.

distributions with means $\bar{c}_{\mathrm{s}} = \sum_{m=1}^{M} \bar{c}_{\mathrm{s}}^{(m)}$ and $M\bar{c}_{\mathrm{n}}$, respectively, where $\bar{c}_{\mathrm{s}}^{(m)} = \mathbb{E}\{c_{\mathrm{s}}[k,m]\}$ and $\bar{c}_{\mathrm{n}} = \mathbb{E}\{c_{\mathrm{n}}[k,m]\}$. We note that the sum detector in (2) includes the well-known peak value [10] and energy [28] detectors as special cases when only one sample at the peak concentration is taken and (ideally infinitely) many samples per symbol interval are taken, respectively.

Remark 2: Unlike the conventional linear input-output model for channels in wireless communications [15], the channel model in (1) is not linear since $s[k]$ does not affect the observation $r[k]$ directly but rather via the Poisson RV $c_{\mathrm{s}}[k]$. However, the *expectation* of the received signal is linearly dependent on the transmitted signal, i.e.,

$$\bar{r}[k] = \mathbb{E}\{r[k]\} = \bar{c}_{\mathrm{s}} s[k] + \bar{c}_{\mathrm{n}}. \tag{3}$$

We note that for a given $s[k]$, in general, the actual number of molecules observed at the receiver, $r[k]$, will differ from the expected number of observed molecules, $\bar{r}[k]$, due to the intrinsic noisiness of diffusion.

B. CSI Model

The state of the diffusive MC channel specified in (1) can be captured by p_{s} and $\bar{c}_{\mathrm{n}}$. Since we employ OOK signaling and N^{Tx} is assumed to be fixed, without loss of generality, we refer to the vector $\bar{\mathbf{c}} = [\bar{c}_{\mathrm{s}}, \bar{c}_{\mathrm{n}}]^T$ (instead of $[p_{\mathrm{s}}, \bar{c}_{\mathrm{n}}]^T$) as the CSI of the considered MC system in the remainder of this paper. Most existing detection schemes in MC assume that knowledge of the CSI is available at the receiver [8], [10], [11]. In contrast, in this paper, we directly detect a block of transmitted symbols based on the corresponding received observations without spending any resources on CSI acquisition at the receiver. Let $\mathbf{s} = [s[1], s[2], \ldots, s[K]]^T$ and $\mathbf{r} = [r[1], r[2], \ldots, r[K]]^T$ denote the vectors of the transmitted symbols and the received observations over a block of K symbol intervals, respectively. We assume that the CSI remains unchanged over one block of transmitted symbols, but may change from one block to the next (e.g., due to a change in MC channel parameters such as the temperature, the velocity of the flow, etc.). To model this, we assume that CSI, $\bar{\mathbf{c}}$, is an RV that takes its values in each block according to PDF $f_{\bar{\mathbf{c}}}(\bar{\mathbf{c}})$. Furthermore, we assume that RVs $\bar{c}_{\mathrm{s}}$ and $\bar{c}_{\mathrm{n}}$ are independent, i.e., $f_{\bar{\mathbf{c}}}(\bar{\mathbf{c}}) = f_{\bar{c}_{\mathrm{s}}}(\bar{c}_{\mathrm{s}}) f_{\bar{c}_{\mathrm{n}}}(\bar{c}_{\mathrm{n}})$ where $f_{\bar{c}_{\mathrm{s}}}(\bar{c}_{\mathrm{s}})$ and $f_{\bar{c}_{\mathrm{n}}}(\bar{c}_{\mathrm{n}})$ are the marginal PDFs of $\bar{c}_{\mathrm{s}}$ and $\bar{c}_{\mathrm{n}}$, respectively. For future reference, in the rest of this work, we define $f_{\mathbf{r}}(\mathbf{r}|\bar{\mathbf{c}}, \mathbf{s}) = \prod_k f_{r[k]}(r[k]|\bar{\mathbf{c}}, s[k])$ as the PDF of the observation vector $\mathbf{r}$ conditioned on both CSI $\bar{\mathbf{c}}$ and transmitted symbol vector $\mathbf{s}$, and we define $f_{\mathbf{r}}(\mathbf{r}|\mathbf{s}) = \prod_k f_{r[k]}(r[k]|s[k])$ as the PDF of $\mathbf{r}$ conditioned only on $\mathbf{s}$. We note that although the proposed non-coherent detection

schemes do not require knowledge of the instantaneous CSI, $\bar{\mathbf{c}}$, we assume that statistical CSI, i.e., $f_{\bar{\mathbf{c}}}(\bar{\mathbf{c}})$, is available for the design of the proposed non-coherent MS, SS, and DF detectors. Since obtaining the CSI statistics might be difficult for some practical systems, we also propose a suboptimal blind detector that does not require statistical CSI knowledge.

Remark 3: We note that for the conventional channel model used in wireless communications, the noise power depends largely on the characteristics of the receiver [15], [16]. For instance, in RF communications, the noise variance is determined by the receiver thermal noise [15], and in optical communications, the power of the noise depends on the sensitivity of the photo-detector to the background radiation [16]. Therefore, when modeling these systems, it is commonly assumed that the noise power is constant. Hence, the channel gain is typically referred to as the CSI and the noise power is assumed to be fixed and known. On the other hand, the noise molecules in MC originate from interfering natural or synthetic sources. Thus, the mean of the noise is in general dependent on the properties of the MC channel in addition to the type of the adopted receiver. Hence, it is reasonable to treat both $\bar{c}_{\mathbf{s}}$ and $\bar{c}_{\mathbf{n}}$ as the CSI of the MC system.

III. OPTIMAL AND SUBOPTIMAL NON-COHERENT DETECTION

In this section, we first present the optimal coherent detector which provides a performance upper bound for any non-coherent detector. Subsequently, we present the proposed non-coherent detectors, namely the optimal non-coherent MS, SS, and DF detectors. Furthermore, motivated by the structure of the optimal coherent detector, we develop a simple blind detector.

A. Coherent ML Detector

As a performance upper bound, we consider the optimal detector for perfect CSI knowledge. Moreover, since the observations in different symbol intervals are independent, without loss of optimality, the considered benchmark scheme performs symbol-by-symbol detection. Thus, the optimal ML detector is given by

$$
\begin{aligned}
\hat{s}^{\mathrm{ML}}[k] &= \operatorname*{argmax}_{s[k]\in\{0,1\}} f_{r[k]}\big(r[k]\big|\bar{\mathbf{c}}, s[k]\big) \\
&= \operatorname*{argmax}_{s[k]\in\{0,1\}} \frac{\big(\bar{c}_{\mathbf{s}}s[k] + \bar{c}_{\mathbf{n}}\big)^{r[k]} \exp\big(-\bar{c}_{\mathbf{s}}s[k] - \bar{c}_{\mathbf{n}}\big)}{r[k]!},
\end{aligned}
\tag{4}
$$

where $f_{r[k]}\big(r[k]\big|\bar{\mathbf{c}}, s[k]\big)$ is the Poisson distribution function. The ML detector can be rewritten in the form of a threshold-based detector as follows [22]

$$\hat{s}^{\texttt{ML}}[k] = \begin{cases} 1, & \text{if } r[k] \geq \xi^{\texttt{ML}}(\bar{\mathbf{c}}) \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where the computation of the decision threshold, $\xi^{\texttt{ML}}(\bar{\mathbf{c}}) = \frac{\bar{c}_{\mathbf{s}}}{\ln(1+\frac{\bar{c}_{\mathbf{s}}}{\bar{c}_{\mathbf{n}}})}$, requires CSI knowledge.

B. Non-Coherent ML Detector

We first develop the optimal non-coherent MS and SS detectors. Following this development, we propose suboptimal decision metrics to cope with the complexity of evaluating the optimal MS and SS metrics for this detector.

1) Optimal MS Detector: We first formulate the ML problem for non-coherent detection of a block of symbols. Note that, in this case, joint detection of multiple symbols is beneficial, despite the fact that the symbols are independent, since observing multiple symbols provides more information about the MC channel than observing just one symbol. The optimal ML MS detector is mathematically given by

$$\hat{\mathbf{s}}^{\texttt{MS}} = \operatorname*{argmax}_{\mathbf{s} \in \mathcal{A}} f_{\mathbf{r}}(\mathbf{r}|\mathbf{s}) = \operatorname*{argmax}_{\mathbf{s} \in \mathcal{A}} \int_{\bar{\mathbf{c}} \geq \mathbf{0}} f_{\mathbf{r}}(\mathbf{r}|\bar{\mathbf{c}}, \mathbf{s}) f_{\bar{\mathbf{c}}}(\bar{\mathbf{c}}) \mathrm{d}\bar{\mathbf{c}}$$

$$= \operatorname*{argmax}_{\mathbf{s} \in \mathcal{A}} \int_{\bar{c}_{\mathbf{s}} \geq 0} \int_{\bar{c}_{\mathbf{n}} \geq 0} \prod_{k=1}^{K} \frac{(\bar{c}_{\mathbf{s}}s[k] + \bar{c}_{\mathbf{n}})^{r[k]} \exp\left(-\bar{c}_{\mathbf{s}}s[k] - \bar{c}_{\mathbf{n}}\right)}{r[k]!} \times f_{\bar{c}_{\mathbf{s}}}(\bar{c}_{\mathbf{s}}) f_{\bar{c}_{\mathbf{n}}}(\bar{c}_{\mathbf{n}}) \mathrm{d}\bar{c}_{\mathbf{s}} \mathrm{d}\bar{c}_{\mathbf{n}}, \tag{6}$$

where $\mathcal{A}$ is the set of all 2^K possible binary sequences of length K. In (6), we employ the multivariate Poisson distribution function $f_{\mathbf{r}}(\mathbf{r}|\bar{\mathbf{c}}, \mathbf{s})$ and exploit the fact that the observations in different symbol intervals are independent and that RVs $\bar{c}_{\mathbf{s}}$ and $\bar{c}_{\mathbf{n}}$ are independent. Before presenting the optimal MS detector as a solution of (6) in the following theorem, we introduce some auxiliary variables. For a given hypothetical sequence $\mathbf{s}$, let $\mathcal{K}_1$ and $\mathcal{K}_0$ denote the sets of indices k for which $s[k] = 1$ and $s[k] = 0$ holds, respectively. Moreover, we define $\mathbf{n} = [n_1, n_0]^T$ where n_1 and n_0 are the number of ones and zeros in the given hypothetical sequence $\mathbf{s}$. Additionally, for a given observation vector $\mathbf{r}$, we define $\mathbf{N} = [N_1, N_0]^T$ where $N_1 = \sum_{k \in \mathcal{K}_1} r[k]$ and $N_0 = \sum_{k \in \mathcal{K}_0} r[k]$.

Theorem 1: The optimal non-coherent MS detector as the solution to (6) selects a sequence whose "1" elements correspond to the $\zeta^{\texttt{MS}}(\mathbf{r})$ largest elements of $\mathbf{r}$. Moreover, the optimal threshold

$\zeta^{\text{MS}}(\mathbf{r})$ is obtained as

$$\zeta^{\text{MS}}(\mathbf{r}) = \underset{n_1 \in \{0,1,\dots,K\}}{\arg\max} \Lambda^{\text{MS}}(\mathbf{s}, \mathbf{r}), \tag{7}$$

where $\Lambda^{\text{MS}}(\mathbf{s}, \mathbf{r})$ is the MS detection metric given by

$$\Lambda^{\text{MS}}(\mathbf{s}, \mathbf{r}) = \mathbb{E}_{\bar{\mathbf{c}}}\Big\{ (\bar{c}_{\mathbf{s}} + \bar{c}_{\mathbf{n}})^{N_1} \, \bar{c}_{\mathbf{n}}^{N_0} \, e^{-n_1 \bar{c}_{\mathbf{s}} - K\bar{c}_{\mathbf{n}}} \Big\}$$

$$= \sum_{i=0}^{N_1} \binom{N_1}{i} \mathbb{E}_{\bar{c}_{\mathbf{s}}} \big\{ \bar{c}_{\mathbf{s}}^{N_1-i} e^{-n_1 \bar{c}_{\mathbf{s}}} \big\} \, \mathbb{E}_{\bar{c}_{\mathbf{n}}} \big\{ \bar{c}_{\mathbf{n}}^{N_0+i} e^{-K\bar{c}_{\mathbf{n}}} \big\}. \tag{8}$$

Furthermore, as $K \to \infty$, we obtain $\zeta^{\text{MS}}(\mathbf{r}) \to K \times \Pr\{s[k] = 1\} = \frac{K}{2}$.

Proof: The proof is provided in Appendix A. ∎

Remark 4: The complexity of the proposed MS detector in Theorem 1 is significantly smaller than that of the full search required in (6). In particular, the complexity of the full search in (6) grows exponentially in K, i.e., $|\mathcal{A}| = 2^K$, whereas the complexity of the search in Theorem 1 is linearly increasing in K, i.e., there are $K + 1$ possibilities. Furthermore, for each search step, we have to calculate the metric $\Lambda^{\text{MS}}(\mathbf{s}, \mathbf{r})$ which is a function of the statistical CSI, but not of the instantaneous CSI, since $\bar{c}_{\mathbf{s}}$ and $\bar{c}_{\mathbf{n}}$ are averaged out in $\Lambda^{\text{MS}}(\mathbf{s}, \mathbf{r})$, cf. (8). Moreover, all expectations in the MS detection metric are in the form of $\mathbb{E}_x \{ x^a e^{-bx} \}$ where $x \in \{\bar{c}_{\mathbf{s}}, \bar{c}_{\mathbf{n}}\}$ and a and b are constants. Therefore, if these expectations for all required a and b can be computed offline, they can be stored at the receiver and used for online data detection.

As a special case of the optimal MS detector in Theorem 1, in the following corollary we present the optimal SS detector which, unlike the general form in Theorem 1, lends itself to a simple threshold-based detection.

Corollary 1: The optimal non-coherent SS detector as the solution to (6) for $K = 1$ can be written in the form of the following threshold-based detector:

$$\hat{s}^{\text{SS}}[k] = \begin{cases} 1, & \text{if } r[k] \geq \xi^{\text{SS}} \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

where threshold ξ^{SS} is given by

$$\xi^{\text{SS}} = \Big\lceil \xi \in \mathbb{R}^+ \big| \mathbb{E}_{\bar{\mathbf{c}}} \big\{ (\bar{c}_{\mathbf{s}} + \bar{c}_{\mathbf{n}})^{\xi} \, e^{-\bar{c}_{\mathbf{s}} - \bar{c}_{\mathbf{n}}} \big\} = \mathbb{E}_{\bar{c}_{\mathbf{n}}} \big\{ \bar{c}_{\mathbf{n}}^{\xi} e^{-\bar{c}_{\mathbf{n}}} \big\} \Big\rceil$$

$$= \min \Big\{ \xi \in \mathbb{N} \big| \sum_{i=0}^{\xi} \binom{\xi}{i} \mathbb{E}_{\bar{c}_{\mathbf{s}}} \big\{ \bar{c}_{\mathbf{s}}^{\xi-i} e^{-\bar{c}_{\mathbf{s}}} \big\} \mathbb{E}_{\bar{c}_{\mathbf{n}}} \big\{ \bar{c}_{\mathbf{n}}^{i} e^{-\bar{c}_{\mathbf{n}}} \big\} > \mathbb{E}_{\bar{c}_{\mathbf{n}}} \big\{ \bar{c}_{\mathbf{n}}^{\xi} e^{-\bar{c}_{\mathbf{n}}} \big\} \Big\}. \tag{10}$$

Proof: The proof is provided in Appendix B. ∎

Corollary 1 reveals that, similar to the coherent ML detector under perfect CSI knowledge in (5), the optimal non-coherent SS detector has a threshold-based structure. However, the threshold for the optimal SS detector depends only on the statistics of the MC channel. Therefore, ξ^{SS} remains fixed as long as the statistics of the MC channel do not change. Hence, ξ^{SS} can be obtained offline once and thereafter used for online detection. Consequently, the optimal SS detector in Corollary 1 is considerably less complex than the equivalent detector obtained from Theorem 1 for $K = 1$, which requires the online computation of threshold $\zeta^{\text{MS}}(r[k])$ or, equivalently, detection metrics $\Lambda^{\text{MS}}(s[k] = 1, r[k])$ and $\Lambda^{\text{MS}}(s[k] = 0, r[k])$ in each symbol interval.

2) Suboptimal ML Metric: The main challenge of the optimal MS detector is the calculation of the detection metric in (8). In particular, the detection metric in (8) involves expectations which require the PDF of the CSI, i.e., $f_x(x)$, $x \in \{\bar{c}_{\text{s}}, \bar{c}_{\text{n}}\}$. These PDFs depend on the considered MC environment and general analytical expressions for $f_x(x)$, $x \in \{\bar{c}_{\text{s}}, \bar{c}_{\text{n}}\}$, are not known. In practice, for a particular MC channel, these PDFs can be obtained using empirical measurements of $\bar{c}_{\text{s}}$ and $\bar{c}_{\text{n}}$. However, the empirical PDFs might not lend themselves to a simple analytical form. Therefore, one convenient approach for obtaining a mathematical expression for the detection metric in (8) is to assume a particular parametric model for $f_x(x)$ and to adjust the parameters of the model to match the simulation/experimental data. In general, it is desirable that a parametric PDF model, $f_x^{\text{para}}(x)$, meets the following criteria:

- $f_x^{\text{para}}(x)$ should meet any inherent constraints of the RV x. Such constraints might include that x is non-negative or non-positive, or that it is discrete, continuous, or mixed.
- $f_x^{\text{para}}(x)$ should facilitate analysis and ideally lead to closed-form expressions for metric associated with x under consideration, which in this case is the detection metric (8).
- $f_x^{\text{para}}(x)$ should be flexible enough to provide an accurate approximation of the exact distribution for a wide range of system parameters.

For the problem at hand, the PDF of the CSI has to be supported only over the non-negative range, i.e., $f_x(x) = 0$ for $x < 0$, since $\bar{c}_{\text{s}}$ and $\bar{c}_{\text{n}}$ assume only non-negative values. In addition, for the purposes of this paper, the adopted PDF $f_x^{\text{para}}(x)$ should lead to a sufficiently simple detection metric (8). We have investigated several well-known distributions as parametric models for the PDF of the CSI, including the Nakagami [29], inverse Gaussian [30], Levy [31], and Gamma

distributions [29]. We found that the Gamma distribution,

$$
f_x^{\texttt{gamma}}(x) = \begin{cases} \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, & \text{if } x \geq 0 \\[2mm] 0, & \text{otherwise} \end{cases}
\tag{11}
$$

with parameters $\alpha, \beta > 0$ [29], is the best fit to the PDF of the CSI, $f_x(x)$, $x \in \{\bar{c}_{\mathsf{s}}, \bar{c}_{\mathsf{n}}\}$, for different system realizations. Moreover, it leads to a simple detection metric. We emphasize that an *experimentally* verified stochastic channel model for MC systems has not been reported yet. Despite this shortcoming, our motivation for adopting the Gamma distribution is as follows. First, for the stochastic MC channel used for simulation in Section V, the Gamma distribution can accurately model the randomness of the CSI introduced by random variations of the underlying MC channel parameters, e.g., the flow velocity, the enzyme concentration, the diffusion coefficient, etc. In fact, in Section V, we show that the BERs of the optimal MS detector using the exact PDF of the CSI and the Gamma distribution perfectly match. Secondly, for the Gamma distribution, the ML decision metrics can be calculated in closed form. In particular, the expectations in the detection metric in (8) are of the form $\mathbb{E}_x\left\{x^a e^{-bx}\right\}$ where $x \in \{\bar{c}_{\mathsf{s}}, \bar{c}_{\mathsf{n}}\}$ and a and b are constants. Therefore, using the Gamma distribution, $\mathbb{E}_x\left\{x^a e^{-bx}\right\}$ can be expressed as

$$
\begin{aligned}
\mathbb{E}_x\left\{x^a e^{-bx}\right\} &= \int_{x=0}^{\infty} x^a e^{-bx} \times \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \mathrm{d}x = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_{x=0}^{\infty} x^{a+\alpha-1} e^{-(b+\beta)x} \mathrm{d}x \\
&= \frac{\beta^\alpha \Gamma(a+\alpha)}{\Gamma(\alpha)(b+\beta)^{a+\alpha}} \underbrace{\int_{x=0}^{\infty} \frac{(b+\beta)^{a+\alpha} x^{a+\alpha-1} e^{-(b+\beta)x}}{\Gamma(a+\alpha)} \mathrm{d}x}_{=1} = \frac{\Gamma(a+\alpha)\beta^\alpha}{\Gamma(\alpha)(b+\beta)^{a+\alpha}}.
\end{aligned}
\tag{12}
$$

The parameters α and β have to be properly chosen such that the resulting Gamma distribution well approximates the exact distribution or, if only the measurement data is available, the histogram of the measurement data. To this end, we adopt the weighted mean square error as a criterion to be minimized for the optimal choice of α and β. In particular, the optimal α^* and β^* are obtained as

$$
(\alpha^*, \beta^*) = \operatorname*{argmin}_{\alpha, \beta > 0} \int_{x=0}^{\infty} w(x) \left| f_x(x) - f_x^{\texttt{gamma}}(x) \right|^2 \mathrm{d}x,
\tag{13}
$$

where $w(x) \geq 0$, $\forall x$, is an appropriately chosen weight function that can be used to prioritize the accuracy of the approximation in a desired range of x. Using the Gamma distribution with the optimized parameters, the detection metrics required for the MS and SS detectors are given in closed form based on (12).

Remark 5: Since the feasible sets of α and β are semi-infinite, i.e., $\alpha, \beta \in (0, +\infty)$, performing a full search to find the optimal α^* and β^* is not possible. To overcome this challenge, we first note that for a Gamma distribution with mean μ_x and variance σ_x^2, the parameters (α, β) are *uniquely* obtained as

$$(\alpha, \beta) = \left(\frac{\mu_x^2}{\sigma_x^2}, \frac{\mu_x}{\sigma_x^2} \right). \tag{14}$$

Since the optimal parameters (α^*, β^*) from (13) are expected to lead to a Gamma distribution that has a mean and a variance that are close to those of the exact distribution, we can efficiently bound the search for the optimal values to intervals $\alpha \in [(1-\delta)\frac{\bar{\mu}_x^2}{\bar{\sigma}_x^2}, (1+\delta)\frac{\bar{\mu}_x^2}{\bar{\sigma}_x^2}]$ and $\beta \in [(1-\delta)\frac{\bar{\mu}_x}{\bar{\sigma}_x}, (1+\delta)\frac{\bar{\mu}_x}{\bar{\sigma}_x}]$, where $\bar{\mu}_x$ and $\bar{\sigma}_x^2$ are the mean and the variance of the exact distribution, respectively, and $\delta \geq 0$ determines how large the search intervals are.

C. Non-Coherent DF Detector

As will be shown in Section V, the proposed optimal MS detector significantly outperforms the optimal SS detector, particularly when the number of jointly detected symbols is large. However, the gain obtained by the optimal MS detector in this case comes at the expense of a large detection delay. In particular, we have to wait until all K symbols in one detection window arrive before detection can be performed. In order to mitigate this problem and to further reduce complexity while still exploiting the benefits of MS detection, we consider a DF detector in the following. In particular, the DF detector instantly detects the received symbols while exploiting the memory of the previously detected symbols assuming that these symbols were *correctly* detected.

To characterize the DF detector, let us first define $\mathbf{r}_k = [r[k-1], \ldots, r[k-K+1]]^T$ and $\hat{\mathbf{s}}_k = [\hat{s}[k-1], \hat{s}[k-2], \ldots, \hat{s}[k-K+1]]^T$ as the observation vector and the detected symbol vector for the previous $K-1$ symbol intervals. The DF detector is derived in a similar manner as the optimal MS detector, cf. Appendix A. In particular, given the $K-1$ previously detected symbols in $\hat{\mathbf{s}}_k$, we evaluate the detection metric in (8) for the two possibilities $s[k] = 1$ and $s[k] = 0$ which leads to the detection metrics $\Lambda^{\mathtt{MS}}(\mathbf{s}_1, \mathbf{r})$ and $\Lambda^{\mathtt{MS}}(\mathbf{s}_0, \mathbf{r})$, respectively, where $\mathbf{s}_0 = [0, \hat{\mathbf{s}}_k^T]^T$, $\mathbf{s}_1 = [1, \hat{\mathbf{s}}_k^T]^T$, and $\mathbf{r} = [r[k], \mathbf{r}_k^T]^T$. This leads to the following decision rule.

Proposed DF Detector: The non-coherent DF detector for the MC system considered in this

paper is given by

$$\hat{s}^{\text{DF}}[k] = \begin{cases} 1, & \text{if } \Lambda^{\text{MS}}(\mathbf{s}_1, \mathbf{r}) \geq \Lambda^{\text{MS}}(\mathbf{s}_0, \mathbf{r}) \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

Since the above DF detector uses the MS detection metric in (8), we can also employ (12) to obtain an approximate metric based on the Gamma distribution proposed in Subsection III-A.

Given $\mathbf{r}_k$ and $\hat{\mathbf{s}}_k$, the ratio $\Lambda^{\text{MS}}(\mathbf{s}_1, \mathbf{r})/\Lambda^{\text{MS}}(\mathbf{s}_0, \mathbf{r})$ is a monotonically increasing function of $r[k]$. Since $\Lambda^{\text{MS}}(\mathbf{s}_1, \mathbf{r}) < \Lambda^{\text{MS}}(\mathbf{s}_0, \mathbf{r})$ holds for $r[k] = 0$, there exists a unique threshold for $r[k]$, denoted by $\xi^{\text{DF}}(\hat{\mathbf{s}}_k, \mathbf{r}_k)$, above which $\Lambda^{\text{MS}}(\mathbf{s}_1, \mathbf{r}) > \Lambda^{\text{MS}}(\mathbf{s}_0, \mathbf{r})$ holds. Hence, the non-coherent DF detector can be equivalently written in the following threshold-based form

$$\hat{s}^{\text{DF}}[k] = \begin{cases} 1, & \text{if } r[k] \geq \xi^{\text{DF}}(\hat{\mathbf{s}}_k, \mathbf{r}_k) \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

The adaptive DF detection threshold $\xi^{\text{DF}}(\hat{\mathbf{s}}_k, \mathbf{r}_k)$ is given by

$$\xi^{\text{DF}}(\hat{\mathbf{s}}_k, \mathbf{r}_k) = \min\left\{ \xi \in \mathbb{N} \big| \Lambda^{\text{MS}}(\mathbf{s}_1, \mathbf{r}^\xi) > \Lambda^{\text{MS}}(\mathbf{s}_0, \mathbf{r}^\xi) \right\}, \tag{17}$$

where $\mathbf{r}^\xi = [\xi, \mathbf{r}_k^T]^T$. We will employ the threshold detection form of the proposed DF detector in (16) for the performance analysis provided in Section IV-C.

D. Suboptimal Detector Based on Blind CSI Estimation

The non-coherent detectors developed so far require *statistical* knowledge of the CSI which might be difficult to acquire for some MC systems deployed in practice. Therefore, in the following, we propose a suboptimal detector that does not need statistical CSI knowledge. The main idea behind the simple detector, which we propose in this subsection, is to first estimate the CSI based on the symbols received in the considered detection window in order to approximate the optimal ML threshold which is denoted by $\xi_{\text{BL}}^{\text{ML}}$. Subsequently, symbol-by-symbol detection can be performed based on the approximated threshold $\xi_{\text{BL}}^{\text{ML}}$. We note that the channel estimator is blind in the sense that a training sequence is not used.

For a given observation block $\mathbf{r}$, let $\widetilde{\mathcal{K}}_1$ ($\widetilde{\mathcal{K}}_0$) denote the sets of indices k for the $\lceil \frac{K}{2} \rceil$-th largest ($\lfloor \frac{K}{2} \rfloor$-th smallest) $r[k]$ in the block. The proposed suboptimal detector is formally presented in the following.

Proposed Blind ML-Based Detector: The proposed blind detector for ON-OFF keying modulation in diffusive MC is given by

$$\hat{s}_{\text{BL}}^{\text{ML}}[k] = \begin{cases} 1, & \text{if } r[k] \geq \xi_{\text{BL}}^{\text{ML}}(\mathbf{r}) \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

where $\xi_{\text{BL}}^{\text{ML}}(\mathbf{r}) = \frac{\hat{\bar{c}}_{\text{s}}}{\ln(1+\frac{\hat{\bar{c}}_{\text{s}}}{\hat{\bar{c}}_{\text{n}}})}$ is the detection threshold. Hereby, the CSI estimates $\hat{\bar{c}}_{\text{s}}$ and $\hat{\bar{c}}_{\text{n}}$ are obtained as

$$\hat{\bar{c}}_{\text{s}} = \frac{1}{\lceil \frac{K}{2} \rceil} \sum_{k \in \widetilde{\mathcal{K}}_1} [r[k] - \hat{\bar{c}}_{\text{n}}] \tag{19a}$$

$$\hat{\bar{c}}_{\text{n}} = \frac{1}{\lfloor \frac{K}{2} \rfloor} \sum_{k \in \widetilde{\mathcal{K}}_0} r[k]. \tag{19b}$$

The CSI estimates $\hat{\bar{c}}_{\text{s}}$ and $\hat{\bar{c}}_{\text{n}}$ in (19) correspond to a simple averaging over the expected positions of $s[k] = 1$ and $s[k] = 0$, respectively. Since the noise molecules are always present whereas the signal molecules are present only when $s[k] = 1$ holds, the *expected* number of molecules observed at the receiver in positions with $s[k] = 1$ is larger than that in positions with $s[k] = 0$. Therefore, we first compute $\hat{\bar{c}}_{\text{n}}$ directly from (19b) and then $\hat{\bar{c}}_{\text{s}}$ from (19a). Having the estimated CSI $(\hat{\bar{c}}_{\text{s}}, \hat{\bar{c}}_{\text{n}})$, we can compute the ML threshold $\xi_{\text{BL}}^{\text{ML}}$ and perform detection based on (9). We note that although $s[k] = 1$ and $s[k] = 0$ are equiprobable, the number of ones and zeros in each detection window may not be exactly $\frac{K}{2}$. Therefore, assuming that the $\lceil \frac{K}{2} \rceil$-th largest elements ($\lfloor \frac{K}{2} \rfloor$-th smallest elements) of $\mathbf{r}$ correspond to the positions of $s[k] = 1$ ($s[k] = 0$) leads to an inherent CSI estimation error.

IV. PERFORMANCE ANALYSIS

In this section, we derive exact expressions and performance bounds for the BERs of the proposed detection schemes. In particular, the average BER, denoted by $\bar{P}_e$, can be written mathematically as

$$\bar{P}_e = \mathbb{E}_{\bar{\mathbf{c}}}\{P_e(\bar{\mathbf{c}})\} = \int_{\bar{c}_{\text{s}} \geq 0} \int_{\bar{c}_{\text{n}} \geq 0} \Pr\{\hat{s}[k] \neq s[k]|\bar{\mathbf{c}}\} f_{\bar{c}_{\text{s}}}(\bar{c}_{\text{s}}) f_{\bar{c}_{\text{n}}}(\bar{c}_{\text{n}}) \mathrm{d}\bar{c}_{\text{s}} \mathrm{d}\bar{c}_{\text{n}}, \tag{20}$$

where $P_e(\bar{\mathbf{c}}) \triangleq \Pr\{\hat{s}[k] \neq s[k]|\bar{\mathbf{c}}\}$ is the BER conditioned on a particular state of the MC channel characterized by the CSI $\bar{\mathbf{c}}$. In the following, we derive the conditional BER of the SS detector and conditional bounds for the MS detector.

A. BER of the Optimal SS Detector

We first derive the exact expression for the BER of the optimal SS detector. In particular, for a given $\bar{\mathbf{c}}$, the BER of the optimal SS detector is obtained as

$$
\begin{aligned}
P_e^{\text{SS}}(\bar{\mathbf{c}}) &= \Pr\left\{\hat{s}[k]=0|\bar{\mathbf{c}}, s[k]=1\right\}\Pr\left\{s[k]=1\right\} \\
&\quad + \Pr\left\{\hat{s}[k]=1|\bar{\mathbf{c}}, s[k]=0\right\}\Pr\left\{s[k]=0\right\} \\
&= \frac{1}{2}\Pr\left\{r[k]<\xi^{\text{SS}}|\bar{\mathbf{c}}, s[k]=1\right\} + \frac{1}{2}\Pr\left\{r[k]\geq\xi^{\text{SS}}|\bar{\mathbf{c}}, s[k]=0\right\} \\
&\overset{(a)}{=} \frac{1}{2}\sum_{r[k]=0}^{\lfloor\xi^{\text{SS}}\rfloor}\frac{(\bar{c}_{\mathsf{s}}+\bar{c}_{\mathsf{n}})^{r[k]}\exp\left(-\bar{c}_{\mathsf{s}}-\bar{c}_{\mathsf{n}}\right)}{r[k]!} + \frac{1}{2}\sum_{r[k]=\lceil\xi^{\text{SS}}\rceil}^{\infty}\frac{\bar{c}_{\mathsf{n}}^{r[k]}\exp\left(-\bar{c}_{\mathsf{n}}\right)}{r[k]!} \\
&\overset{(b)}{=} \frac{1}{2}+\frac{1}{2}\left[Q\left(\lceil\xi^{\text{SS}}\rceil, \bar{c}_{\mathsf{s}}+\bar{c}_{\mathsf{n}}\right) - Q\left(\lceil\xi^{\text{SS}}\rceil, \bar{c}_{\mathsf{n}}\right)\right],
\end{aligned} \tag{21}
$$

where $Q(x,y) = \frac{\Gamma(x,y)}{\Gamma(x)}$ is the regularized Gamma function and $\Gamma(\cdot,\cdot)$ is the upper incomplete Gamma function [29]. For equality (a), we used the fact that $r[k]$ conditioned on $\bar{\mathbf{c}}$ and $s[k]$ follows a Poisson distribution, and for (b), we exploited the expression for the CDF of a Poisson RV with mean λ given by $F_{r[k]}(\xi) = \Pr\{r[k]\leq\xi\} = Q(\lfloor\xi+1\rfloor, \lambda)$ which leads to $\Pr\{r[k]<\xi\} = Q(\lceil\xi\rceil, \lambda)$ [29].

Remark 6: The BER expression in (21) can be also used to compute the BER of the benchmark scheme with perfect CSI knowledge given in Subsection III.A after replacing ξ^{SS} with $\xi^{\text{ML}}(\bar{\mathbf{c}})$.

B. BER Upper Bound for the Optimal MS Detector

Unfortunately, the derivation of an exact expression for the BER of the optimal MS detector is a difficult, if not impossible, task due to the multi-dimensional structure of the statistics of the observation vector. Hence, in this subsection, we use the union bound to arrive at an upper bound on the BER of the optimal MS detector.

The pairwise error probability (PEP), denoted by $P(\mathbf{s}\to\hat{\mathbf{s}})$, is defined as the probability that, assuming $\mathbf{s}$ is transmitted, $\hat{\mathbf{s}}$ is detected. We note that, due to the structure of the optimal MS detector, the error probabilities for all transmitted sequences which have the same number of ones are identical. Hence, without loss of generality, we consider only $K+1$ sequences $\mathbf{s}$ whose first $n_1 \in \{0,\ldots,K\}$, bits are ones and which are collected in set $\tilde{\mathcal{A}}$. Using the PEP, the BER is upper bounded based on the union bound as follows

$$
P_e^{\text{MS}}(\bar{\mathbf{c}}) \leq \frac{1}{K}\sum_{\mathbf{s}\in\hat{\mathcal{A}}}\sum_{\hat{\mathbf{s}}\in\{\mathcal{A}\backslash\mathbf{s}\}} h\left(\mathbf{s}, \hat{\mathbf{s}}\right) P^{\text{MS}}(\mathbf{s}\to\hat{\mathbf{s}})\Pr\{\mathbf{s}\}, \tag{22}
$$

where $h(\mathbf{s}, \hat{\mathbf{s}})$ denotes the Hamming distance between $\mathbf{s}$ and $\hat{\mathbf{s}}$ and $\Pr\{\mathbf{s}\} = \frac{1}{2^K}\binom{K}{n_1}$ where n_1 is the number of ones in $\mathbf{s}$. We note that in order to calculate the PEP for a given $\mathbf{s}$ and $\hat{\mathbf{s}}$, a K-dimensional summation with respect to the different possibilities for the observation vector $\mathbf{r}$ is needed, which is a computationally challenging task. However, as can be seen from (8), the detection metric is a function solely of n_1, N_1, and $N = N_1 + N_0$. Thus, we introduce the new notation $\Lambda^{\text{MS}}(n_1, N_1, N)$ to indicate this. Let $\Lambda^{\text{MS}}(n_1, N_1, N)$ and $\Lambda^{\text{MS}}(\hat{n}_1, \hat{N}_1, N)$ be the detection metrics for hypothesis sequences $\mathbf{s}$ and $\hat{\mathbf{s}}$, respectively, for a given $\mathbf{r}$. We note that, given $\mathbf{s}$ and $\hat{\mathbf{s}}$, the values of n_1 and $\hat{n}_1$ are known, respectively, and for a given $\mathbf{r}$, the value of N is the same for both $\mathbf{s}$ and $\hat{\mathbf{s}}$. Hence, unlike the expression in (22), which requires a K-dimensional summation with respect to the elements of $\mathbf{r}$, it suffices to consider a three-dimensional summation with respect to N_1, $\hat{N}_1$, and N. In particular, in order to calculate $P^{\text{MS}}(\mathbf{s} \to \hat{\mathbf{s}})$, we have to find the probabilities of the observation events $(N_1, \hat{N}_1, N)$, denoted by $\Pr\{N_1, \hat{N}_1, N | \bar{\mathbf{c}}\}$, for which $\Lambda^{\text{MS}}(n_1, N_1, N) < \Lambda^{\text{MS}}(\hat{n}_1, \hat{N}_1, N)$ holds. This leads to

$$
\begin{aligned}
P^{\text{MS}}(\mathbf{s} \to \hat{\mathbf{s}}) &= \Pr\left\{\Lambda^{\text{MS}}(\mathbf{s}, \mathbf{r}) < \Lambda^{\text{MS}}(\hat{\mathbf{s}}, \mathbf{r})\right\} \\
&= \sum_N \sum_{N_1} \sum_{\hat{N}_1} \mathbf{1}\left\{\Lambda^{\text{MS}}(n_1, N_1, N) < \Lambda^{\text{MS}}(\hat{n}_1, \hat{N}_1, N)\right\} \Pr\left\{N_1, \hat{N}_1, N | \bar{\mathbf{c}}\right\}. \quad (23)
\end{aligned}
$$

Conditioned on $\bar{\mathbf{c}}$, variables N_1, $\hat{N}_1$, and N are *correlated* Poisson RVs. Hence, in order to compute $\Pr\{N_1, \hat{N}_1, N | \bar{\mathbf{c}}\}$, we divide the observation vector $\mathbf{r}$ into the following four partitions:

$\mathcal{S}_1$: Positions of $\mathbf{r}$ where the corresponding elements of both $\mathbf{s}$ and $\hat{\mathbf{s}}$ are one.

$\mathcal{S}_2$: Positions of $\mathbf{r}$ where the corresponding elements of $\mathbf{s}$ and $\hat{\mathbf{s}}$ are one and zero, respectively.

$\mathcal{S}_3$: Positions of $\mathbf{r}$ where the corresponding elements of $\mathbf{s}$ and $\hat{\mathbf{s}}$ are zero and one, respectively.

$\mathcal{S}_4$: Positions of $\mathbf{r}$ where the corresponding elements of both $\mathbf{s}$ and $\hat{\mathbf{s}}$ are zero.

The sum of observed molecules in the positions specified by $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, and $\mathcal{S}_4$ are denoted by M_1, M_2, M_3, and M_4, respectively. It follows that, conditioned on $\bar{\mathbf{c}}$, variables M_1, M_2, M_3, and M_4 are *independent* Poisson RVs with means $\lambda_1 = \upsilon(\bar{c}_{\mathbf{s}} + \bar{c}_{\mathbf{n}})$, $\lambda_2 = (n_1 - \upsilon)(\bar{c}_{\mathbf{s}} + \bar{c}_{\mathbf{n}})$, $\lambda_3 = (\hat{n}_1 - \upsilon)\bar{c}_{\mathbf{n}}$, and $\lambda_4 = (K + \upsilon - n_1 - \hat{n}_1)\bar{c}_{\mathbf{n}}$, respectively, where $\upsilon = |\mathcal{S}_1|$. Moreover, we have $N_1 = M_1 + M_2$, $\hat{N}_1 = M_1 + M_3$, and $N = M_1 + M_2 + M_3 + M_4$. Using these results, we obtain

$$
\begin{aligned}
\Pr\left\{N_1, \hat{N}_1, N | \bar{\mathbf{c}}\right\} = \sum_{i=0}^{\min\{N_1, \hat{N}_1\}} & \Pr\left\{M_1 = i | \bar{\mathbf{c}}\right\} \Pr\left\{M_2 = N_1 - i | \bar{\mathbf{c}}\right\} \\
& \times \Pr\left\{M_3 = \hat{N}_1 - i | \bar{\mathbf{c}}\right\} \Pr\left\{M_4 = N_1 + \hat{N}_1 - i | \bar{\mathbf{c}}\right\}
\end{aligned}
$$

$$= \sum_{i=0}^{\min\{N_1,\hat{N}_1\}} \frac{\lambda_1^i e^{-i\lambda_1}}{i!} \times \frac{\lambda_2^{N_1-i} e^{-(N_1-i)\lambda_2}}{(N_1-i)!}$$

$$\times \frac{\lambda_3^{\hat{N}_1-i} e^{-(\hat{N}_1-i)\lambda_3}}{(\hat{N}_1-i)!} \times \frac{\lambda_4^{N_1+\hat{N}_1-i} e^{-(N_1+\hat{N}_1-i)\lambda_4}}{(N_1+\hat{N}_1-i)!}. \tag{24}$$

Using (20) and (22)-(24), the average PEP can be analytically computed.

Remark 7: In Section V, we employ (22) as an upper bound for the BER of the optimal MS detector and verify its tightness by comparing it with the exact BER obtained via simulation. We note that the computation of the *average* PEP in an analytical form based on (20) and (22)-(24) is a computationally challenging task due to the numerical integration required for evaluating (20). In order to reduce the complexity of the evaluation of the average PEP, one can employ a hybrid approach where, for a given CSI $\bar{\mathbf{c}}$, the PEP bound in (22) is calculated numerically and the expectation in (20) is performed via Monte Carlo simulation.

C. BER Lower Bound for the MS/DF Detectors

In this subsection, we derive the BER of a genie-aided DF detector, i.e., we assume *error-free* decision feedback, which constitutes a lower bound on the BERs of the MS/DF detectors. Similar to the MS detection metric, the DF detection threshold $\xi^{\mathrm{DF}}(\hat{\mathbf{s}}_k, \mathbf{r}_k)$ is also a function of only $n_{k,1}$, $N_{k,1}$, $N_{k,0}$, i.e., we may use the notation $\xi^{\mathrm{DF}}(n_{k,1}, N_{k,1}, N_{k,0})$. Exploiting this observation, the BER of the genie-aided DF detector conditioned on $\bar{\mathbf{c}}$ is written as

$$P_e^{\mathrm{DF}}(\bar{\mathbf{c}}) = \sum_{N_{k,0}} \sum_{N_{k,1}} \sum_{n_{k,1}} \Pr\left\{\hat{s}[k] \neq s[k] | \bar{\mathbf{c}}, n_{k,1}, N_{k,1}, N_{k,0}\right\} \Pr\left\{n_{k,1}, N_{k,1}, N_{k,0} | \bar{\mathbf{c}}\right\}$$

$$= \sum_{N_{k,0}} \sum_{N_{k,1}} \sum_{n_{k,1}} \Pr\left\{\hat{s}[k] \neq s[k] | \bar{\mathbf{c}}, n_{k,1}, N_{k,1}, N_{k,0}\right\} \Pr\left\{N_{k,1}, N_{k,0} | \bar{\mathbf{c}}, n_{k,1}\right\} \Pr\left\{n_{k,1} | \bar{\mathbf{c}}\right\}. \tag{25}$$

In (25), $n_{k,1}$ denotes the number of ones in the $K-1$ previously detected symbols and is a binomial RV. Moreover, $N_{k,1}$ and $N_{k,0}$ conditioned on $\bar{\mathbf{c}}$ and $n_{k,1}$ are independent Poisson RVs with means $n_{k,1}(\bar{c}_{\mathsf{s}} + \bar{c}_{\mathsf{n}})$ and $(K - n_{k,1})\bar{c}_{\mathsf{n}}$, respectively. Therefore, $\Pr\left\{n_{k,1} | \bar{\mathbf{c}}\right\}$, $\Pr\left\{N_{k,1} | \bar{\mathbf{c}}, n_{k,1}\right\}$, and $\Pr\left\{N_{k,0} | \bar{\mathbf{c}}, n_{k,1}\right\}$ can be computed as

$$\Pr\left\{n_{k,1} | \bar{\mathbf{c}}\right\} = \frac{1}{2^{K-1}} \binom{K-1}{n_{k,1}} \tag{26a}$$

$$\Pr\left\{N_{k,1} | \bar{\mathbf{c}}, n_{k,1}\right\} = \frac{n_{k,1}^{N_{k,1}} (\bar{c}_{\mathsf{s}} + \bar{c}_{\mathsf{n}})^{N_{k,1}} e^{-n_{k,1}(\bar{c}_{\mathsf{s}} + \bar{c}_{\mathsf{n}})}}{N_{k,1}!} \tag{26b}$$

$$\Pr\left\{N_{k,0} | \bar{\mathbf{c}}, n_{k,1}\right\} = \frac{(K - n_{k,1})^{N_{k,0}} \bar{c}_{\mathsf{n}}^{N_{k,0}} e^{-(K-n_{k,1})\bar{c}_{\mathsf{n}}}}{N_{k,0}!}. \tag{26c}$$

Furthermore, $\Pr\left\{\hat{s}[k] \neq s[k] | \bar{\mathbf{c}}, n_{k,1}, N_{k,1}, N_{k,0}\right\}$ for the genie-aided DF detector is given by

$$
\begin{aligned}
\Pr\left\{\hat{s}[k] \neq s[k] | \bar{\mathbf{c}}, n_{k,1}, N_{k,1}, N_{k,0}\right\} &= \frac{1}{2}\Pr\left\{r[k] < \xi^{\mathrm{DF}}(n_{k,1}, N_{k,1}, N_{k,0}) | \bar{\mathbf{c}}, s[k] = 1\right\} \\
&\quad + \frac{1}{2}\Pr\left\{r[k] \geq \xi^{\mathrm{DF}}(n_{k,1}, N_{k,1}, N_{k,0}) | \bar{\mathbf{c}}, s[k] = 0\right\} \\
&= \frac{1}{2} + \frac{1}{2}\Big[Q\left(\lceil\xi^{\mathrm{DF}}(n_{k,1}, N_{k,1}, N_{k,0})\rceil, \bar{c}_{\mathbf{s}} + \bar{c}_{\mathbf{n}}\right) \\
&\quad - Q\left(\lceil\xi^{\mathrm{DF}}(n_{k,1}, N_{k,1}, N_{k,0})\rceil, \bar{c}_{\mathbf{n}}\right)\Big].
\end{aligned}
\tag{27}
$$

V. NUMERICAL RESULTS

In this section, we first present the stochastic MC channel model and the adopted system parameters used for simulation. Subsequently, we evaluate the performance of the proposed non-coherent detectors where we employ the optimal coherent detector as a benchmark scheme.

A. Stochastic MC Channel Model

In this section, we present the stochastic MC channel model used for the simulation results. However, we emphasize that the proposed non-coherent detection framework is valid for any given $f_{\bar{c}_{\mathbf{s}}}(\bar{c}_{\mathbf{s}})$ and $f_{\bar{c}_{\mathbf{n}}}(\bar{c}_{\mathbf{n}})$ and is not limited to the stochastic MC channel model used here for simulation.

Let us assume a point source with impulsive molecule release, a fully transparent spherical receiver with volume V^{RX}, and an unbounded environment with diffusion coefficient D. Moreover, we denote the distance between the transmitter and the receiver by d. In addition, we assume that there is steady uniform flow (or drift) with parallel and perpendicular velocity components, denoted by $v_{\parallel}$ and $v_{\perp}$, respectively, with respect to the direction from the transmitter to the receiver. Furthermore, the signaling molecules may react with enzyme molecules, which are present in the MC environment, and degrade into a form that cannot be detected by the receiver. We assume a uniform and constant concentration of the enzyme, denoted by $\bar{c}_{\mathbf{e}}$, and a first order reaction mechanism between the signaling and enzyme molecules with constant reaction rate κ [10]. Based on the aforementioned assumptions, the *expected* number of molecules observed at the receiver as a function of time, denoted by $\bar{c}_{\mathbf{s}}(t)$, is given by [10]

$$
\bar{c}_{\mathbf{s}}(t) = \frac{N^{\mathrm{TX}}V^{\mathrm{RX}}}{(4\pi Dt)^{3/2}}\exp\left(-\kappa\bar{c}_{\mathbf{e}}t - \frac{(d - v_{\parallel}t)^2 + (v_{\perp}t)^2}{4Dt}\right).
\tag{28}
$$

Furthermore, assuming a peak observation detector, the sample time from the beginning of each symbol interval is chosen as $t^{\mathtt{max}} = \underset{t>0}{\mathrm{argmax}}\,\bar{c}_{\mathbf{s}}(t)$ which leads to $\bar{c}_{\mathbf{s}} = \underset{t>0}{\max}\,\bar{c}_{\mathbf{s}}(t)$.

TABLE I
DEFAULT VALUES OF THE NUMERICAL PARAMETERS [5], [10].

Variable	Definition	Value
V^{RX}	Receiver volume	$\frac{4}{3}\pi 50^3$ nm^3 (a sphere with radius 50 nm)
d	Distance between the transmitter and the receiver	500 nm
D	Diffusion coefficient for the signaling molecule	4.365×10^{-10} m$^2 \cdot$ s^{-1}
$\bar{c}_{\text{e}}$	Enzyme concentration	10^5 molecule $\cdot$ μm^3 (approx. 1.66 micromolar)
κ	Rate of molecule degradation reaction	2×10^{-19} m$^3 \cdot$ molecule$^{-1} \cdot$ s^{-1}
$(v_\parallel, v_\perp)$	Components of flow velocity	$(10^{-3}, 10^{-3})$ m $\cdot$ s^{-1}

The channel parameters undergo random variations that lead to random variations in $\bar{c}_{\text{s}}$. For instance, the flow velocity components, $v_\parallel$ and $v_\perp$, may vary over time or the diffusion coefficient, D, and enzyme concentration, $\bar{c}_{\text{e}}$, may change due to variations in the environment temperature. To capture these effects, we assume that the channel parameters in each detection window are realizations of RVs according to $z = z^{\text{def}}(1 + \sigma_z \mathcal{N}(0, 1))$, $z \in \{D, v_\parallel, v_\perp, \bar{c}_{\text{e}}\}$ where z^{def} denotes the mean value of parameter z and $\sigma_z z^{\text{def}}$ is its standard deviation, which determines how much the parameter may deviate from the mean[3]. As $\sigma_z \to 0$, $\forall z$, the respective MC channel becomes deterministic, and for large σ_z, the corresponding MC channel is highly stochastic. Substituting the Gaussian RVs $z \in \{D, v_\parallel, v_\perp, \bar{c}_{\text{e}}\}$ into (28) may not lead to a closed-form analytical expression for $f_{\bar{c}_{\text{s}}}(\bar{c}_{\text{s}})$. Therefore, we employ Monte Carlo simulation to determine a histogram for $\bar{c}_{\text{s}}$ as the "true" distribution, $f_{\bar{c}_{\text{s}}}(\bar{c}_{\text{s}})$, and to verify the effectiveness of the proposed Gamma distribution to approximate the true distribution. For the results provided in this section, we assume that the variation of the mean of the noise $\bar{c}_{\text{n}}$ is modeled similarly to the variation of the mean of the signal $\bar{c}_{\text{s}}$. In particular, we choose $\bar{c}_{\text{n}} = \text{SNR}^{-1} X$ where X is an RV whose PDF is identical to that of $\bar{c}_{\text{s}}$ and SNR is a constant analogous to the signal-to-noise ratio (SNR) in conventional wireless systems. Furthermore, if $\bar{c}_{\text{s}} \sim \text{Gamma}(\alpha, \beta)$ holds, we obtain $\bar{c}_{\text{n}} \sim \text{Gamma}(\alpha, \text{SNR}\beta)$ [29].

B. Simulation Parameters, DF Window Size, and Benchmark Scheme

Simulation Parameters: The default values of the channel parameters are given in Table I. Moreover, we consider the following three scenarios for the stochastic MC channel: Scenario 1: $(\sigma_D, \sigma_{v_\parallel}, \sigma_{v_\perp}, \sigma_{\bar{c}_{\text{e}}}) = (0.1, 0.5, 0.5, 0.1)$, Scenario 2: $(\sigma_D, \sigma_{v_\parallel}, \sigma_{v_\perp}, \sigma_{\bar{c}_{\text{e}}}) = (0.2, 1, 1.5, 0.1)$, and Scenario 3: $(\sigma_D, \sigma_{v_\parallel}, \sigma_{v_\perp}, \sigma_{\bar{c}_{\text{e}}}) = (0.1, 1.5, 0.5, 0.2)$. Each of these scenarios leads to a pair of PDFs for $\bar{c}_{\text{s}}$ and $\bar{c}_{\text{n}}$. In particular, we choose $\bar{c}_{\text{n}}$ such that for the default parameters in Table I

---

[3]We note that a Gaussian RV may assume negative values whereas D and $\bar{c}_{\text{e}}$ are non-negative parameters. Therefore, we assume small values for σ_D and $\sigma_{\bar{c}_{\text{e}}}$, and omit those realizations for which $z \in \{D, \bar{c}_{\text{e}}\}$ is negative.
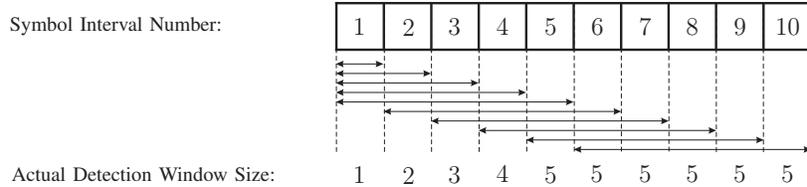
Fig. 2. Schematic illustration of the different detection window sizes required for the DF detection for $K = 5$ and $B = 10$. In fact, the DF detector employs the desired detection window size of $K = 5$ only for the last $B - K + 1 = 6$ symbol intervals in each block of B symbols where the CSI in assumed to be constant.

and $N^{\mathtt{Tx}} = 10^4$, we obtain $\mathbb{E}\{\bar{c}_{\mathtt{n}}\} = \mathbb{E}\{\bar{c}_{\mathtt{s}}\}$, i.e., $\mathtt{SNR} = 1$ holds for this case. To obtain different SNRs, we change the number of molecules released by the transmitter. We used Monte Carlo simulation to obtain simulation results. Thereby, we first generated $N = 10^6$ realizations of the CSI vector $\bar{\mathbf{c}}$. Then, for each $\bar{\mathbf{c}}$, we generated the number of molecules observed at the receiver in each symbol interval based on the channel model in (1) where RVs $c_{\mathtt{s}}[k]$ and $c_{\mathtt{n}}[k]$ are Poisson distributed with means $\bar{c}_{\mathtt{s}}$ and $\bar{c}_{\mathtt{n}}$, respectively. For the analytical results, we numerically evaluate the expression in (20).

DF Window Size: In order to employ the proposed DF detector with window size K, knowledge of the $K - 1$ previously detected symbols is required. However, at the beginning of each transmission interval, i.e., for symbol intervals $k < K$, such knowledge is not available. Therefore, for the results presented in this section, we assume that the DF detector detects the first symbol without knowledge of the previous symbols, i.e., with window size one. Then, it detects the second symbol based on the knowledge of the detected symbol in the first symbol interval, i.e., with window size two. We continue this process until symbol interval K where, for all the following symbols $k \geq K$, the previous $K - 1$ detected symbols are available and a fixed window size of K can be used for DF detection. Because of the overlapping detection windows, we implicitly have to assume that the CSI is constant for $B \geq K$ symbol intervals such that a DF detector with window size K can make $B - K + 1$ symbol decisions. For the results shown in this section, we consider the BER for only the $B - K + 1$ symbol intervals for which the DF detector actually employs a detection window of size K. Fig. 2 schematically shows the different detection window sizes used for DF detection assuming $K = 5$ and $B = 10$.

Benchmark Scheme: We note that the non-coherent detector in [17] was designed assuming an additive white Gaussian noise (AWGN) channel model. As discussed in Subsection II-A, the Poisson distribution for the observed number of molecules at the receiver is a more accurate model for the diffusive MC channel than the Gaussian distribution [22]. Additionally, the signal model

assumed in [17] is different from that considered in this paper. Therefore, we have not included the detection scheme in [17] as a benchmark in this section as a direct comparison would not be fair. Instead we use the optimal coherent detector, cf. Subsection III.A, as a benchmark scheme which in fact constitutes a performance upper bound for any non-coherent detector.

C. Performance Evaluation

In this subsection, we first verify the accuracy of the proposed Gamma distribution, cf. Subsection III-B, in Fig. 3, and our analytical derivations, cf. Section IV, in Fig. 4. Subsequently, in Figs. 5-8, we evaluate the performance of the proposed non-coherent detectors for different system parameters.

Fig. 3 shows the histogram of the CSI, $\bar{c}_\mathbf{s}$, obtained by Monte Carlo simulation, and the corresponding Gamma PDF approximation for the three considered stochastic scenarios. Additionally, the result for the case when all the underlying channel parameters in (28) assume their nominal values given in Table I, i.e., when the channel is deterministic, are shown. The optimal parameters of the Gamma distribution are also shown in Fig. 3 and found using the search procedure presented in Subsection III-A and Remark 5 with $w(x) = 1$, $\forall x$, and $\delta = 0.5$. For Scenario 3, Nakagami, inverse Gaussian, and Levy distribution approximations are also plotted with their parameters optimized based on a similar search procedure as proposed in Subsection III-A for the Gamma PDF. We visually observe a very close match between the histogram (exact PDF) and the Gamma PDF approximation for all three scenarios. Moreover, Fig. 3 suggests that for Scenario 3, the Gamma distribution is a better match to the exact distribution than the Nakagami, inverse Gaussian, and Levy distributions. Note that although the Nakagami distribution seems to also accurately match the exact distribution in Fig. 3, it does not lead to a closed-form expression for the ML detection metric, and hence, at least for the purpose of this paper, the Nakagami distribution is not a good option. We note that the variances of the channel parameters $(\sigma_D, \sigma_{v_\parallel}, \sigma_{v_\perp}, \sigma_{\bar{c}_\mathbf{e}})$ are larger for Scenario 3 compared to Scenario 1, i.e., the underlying channel parameters for Scenario 3 are more random compared to Scenario 1. From Fig. 3, we observe that as the randomness in the MC channel increases, i.e., from Scenario 1 to Scenario 3, the mean of the CSI decreases and its variance increases.

In Fig. 4, we verify the simulation results using the performance analysis developed in Section IV. In particular, we show the BER versus the SNR in dB for Scenario 1 and $K = 10$. For the symbol-by-symbol detectors, i.e., the optimal coherent ML (C-ML) detector and the optimal
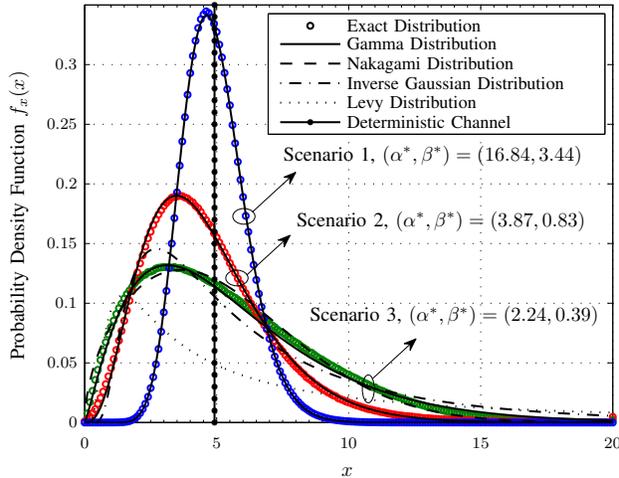
Fig. 3. The histogram of $\bar{c}_\mathsf{s}$ (exact PDF) and the Gamma PDF for Scenarios 1-3. Additionally, the Nakagami, inverse Gaussian, and Levy distributions are plotted only for Scenario 3 to keep the figure readable.
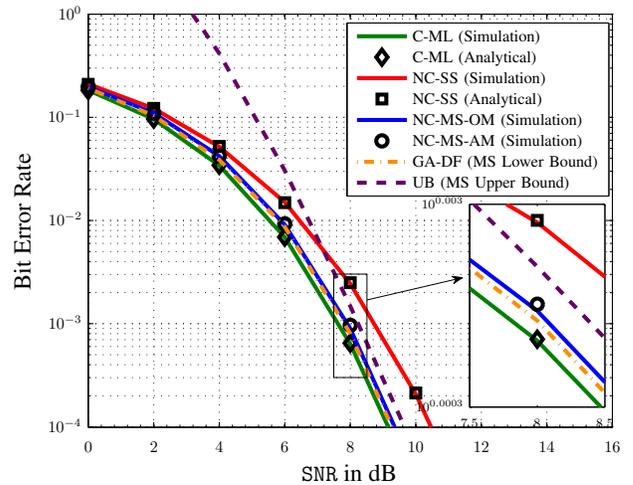
Fig. 4. Bit error rate versus the SNR in dB for Scenario 1 and $K = 10$.

non-coherent SS (NC-SS) detector, we observe that the analytical results obtained from (21) match perfectly with the simulation results. The accuracy of the proposed Gamma distribution to model the considered stochastic MC channel was verified in Fig. 3 by showing the true and approximated PDFs. In Fig. 4, we verify the Gamma approximation in terms of the resulting BER performance. In particular, we show results where the optimal MS detection metric is obtained via expectation over 10^6 realization of the true CSI, i.e., Monte Carlo simulation, and analytically using the proposed Gamma distribution, cf. (12). From Fig. 4, we observe that the BERs of the non-coherent MS detectors which employ the optimal metric (NC-MS-OM) and approximated metric (NC-MS-AM) are almost identical. Furthermore, we show the union bound (UB) in (22) and the BER of the genie-aided DF (GA-DF) detector in (25) as an upper bound and a lower bound for the BER of the optimal MS detector, respectively. Note that the union bound becomes a tight upper bound for large SNRs whereas the BER of the genie-aided DF detector is a tight lower bound for all SNR values considered in Fig. 4.

In Fig. 5, we show the BER versus $\kappa = \frac{B}{K}$, see Subsection V.B, for the three considered scenarios, $K = 5$, and $\mathtt{SNR} = 10$ dB. Note that the actual CSI in (28) is used for simulation whereas the approximated Gamma distribution is employed for calculation of the detection metrics for the proposed MS and DF detectors by using (12), where the corresponding curves in Fig. 5 are denoted by NC-MS-AM and NC-DF-AM, respectively. We observe that the optimal MS detector with detection window size B outperforms the DF detector for all $\kappa \geq 1$ and approaches the
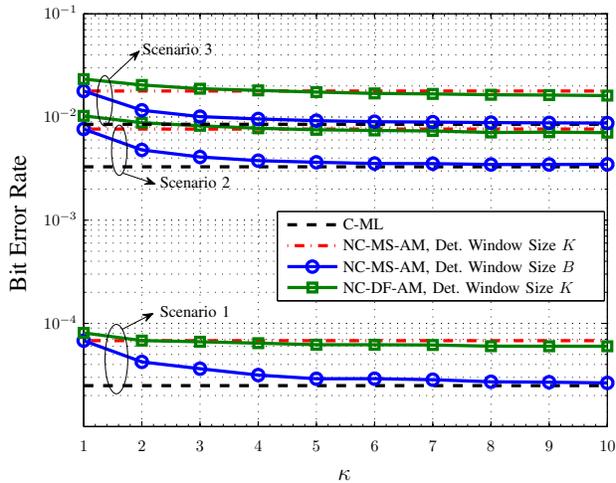
Fig. 5. Bit error rate versus κ for Scenarios 1-3, $K = 5$, and SNR $= 10$ dB. All curves were obtained via simulation.
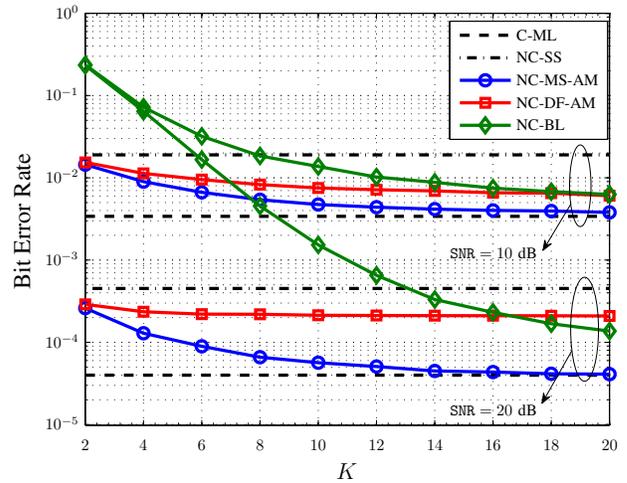
Fig. 6. Bit error rate versus the size of the detection window for Scenarios 2 and SNR $= \{10, 20\}$ dB. All curves were obtained via simulation.

performance of the coherent ML detector as $\kappa \to \infty$. For small κ, the BER of the DF detector is higher than the BER of the optimal MS detector with the same detection window size K, however, as κ increases, the DF detector outperforms the optimal MS detector with detection window size K by a small margin. This is due to the fact that the DF detector implicitly exploits the property that the CSI is fixed for $B \geq K$ symbol intervals whereas the decision of the optimal MS detector is independent for each detection window size K. Moreover, from Fig. 5, we observe that the BER increases for all considered detectors as the MC channel becomes more stochastic, i.e., from Scenario 1 to Scenario 2 to Scenario 3.

In Fig. 6, we show the BER versus the detection window size K for Scenario 2 and SNR $\in \{10, 20\}$ dB. We assume $\kappa = 1$ for the DF detector. Note that the BERs of the proposed detectors decrease and finally converge to the lower bound provided by the coherent ML detector as $K \to \infty$. Furthermore, the gap between the BER of the optimal non-coherent MS detector and the coherent ML detector is small for $K = 20$, which reveals the effectiveness of the optimal MS detector, although no resources are spent for training and CSI acquisition. Moreover, the gap between the BER of the proposed optimal MS detector and the proposed suboptimal non-coherent blind (NC-BL) detector decreases for larger values of K, which confirms the effectiveness of the proposed suboptimal blind detector for large detection window sizes. Furthermore, as expected, the performance of all detection schemes in Fig. 6 is better for SNR $= 20$ dB compared to SNR $= 10$ dB.

In Fig. 7, we plot the BER versus the SNR in dB for Scenario 2 and $K \in \{5, 10\}$. In this figure,

we observe that as the SNR increases, the BER improves for all considered detectors. We note that as $\mathtt{SNR} \to \infty$, the BER of the proposed blind detector saturates to a certain error floor. This is due to the fact that for the *blind* CSI estimator in (19), we assume that the percentages of ones and zeros in a given detection window are exactly 50%, which is not always true, especially for small values of K. This introduces an inherent CSI estimation error and leads to the aforementioned error floor for the proposed blind detector. In contrast, none of the other detectors has a BER error floor. Note that as K increases, the BER of the optimal non-coherent MS detector decreases and approaches the BER of the coherent ML detector. We can also observe from Fig. 7 that the optimal MS detector outperforms the proposed suboptimal blind detector, particularly for small K, but the gap between the BERs of these two detectors decreases as K increases.

Recall that we assume ISI-free transmission in the system model adopted in this paper. However, although the ISI may be considerably reduced using, e.g., the methods in [24] and [25], some residual ISI always exists as the length of the symbol intervals is finite. Therefore, in the system model, we assumed that the effect of the residual ISI is included in $c_\mathtt{n}[k]$ and is sufficiently small compared to the other components in $c_\mathtt{n}[k]$ such that $c_\mathtt{n}[k]$ is (approximately) independent of the signal component $c_\mathtt{s}[k]$. In Fig. 8 we study the BER performance loss if this assumption does not hold. To this end, we consider a finite symbol interval length, denoted by $T_\mathtt{symb}$, and assume that the contribution of the ISI from the previous symbol intervals is present. The detectors treat the ISI as noise, i.e., $\bar{c}_\mathtt{n} = \bar{c}_\mathtt{n}^\mathtt{ext} + \bar{c}_\mathtt{n}^\mathtt{ISI}$ where $\bar{c}_\mathtt{n}^\mathtt{ext}$ is the mean of the external noise and $\bar{c}_\mathtt{n}^\mathtt{ISI}$ is the expected ISI given by

$$\bar{c}_\mathtt{n}^\mathtt{ISI} = \mathbb{E}\left\{\sum_{l=2}^{\infty} s[k-l+1]\bar{c}_{\mathtt{s},l}\right\} = \frac{1}{2}\sum_{l=2}^{\infty} \bar{c}_{\mathtt{s},l}, \tag{29}$$

where $\bar{c}_{\mathtt{s},l}$ is the l-th channel tap, i.e., if the transmitter releases $N^\mathtt{TX}$ molecules in symbol interval k, $\bar{c}_{\mathtt{s},l}$ is the *expected* number of molecules arriving at the destination in symbol interval $k+l-1$. Therefore, $\bar{c}_{\mathtt{s},1}$ is the channel tap for the desired signal and $\bar{c}_{\mathtt{s},l}$, $l \geq 2$, are channel taps that create ISI.

In Fig. 8, we show the BER versus $\mathtt{SNR}^\mathtt{ext} = \frac{\bar{c}_{\mathtt{s},1}}{\bar{c}_\mathtt{n}^\mathtt{ext}}$ in dB for Scenario 3, $K = 10$, and symbol intervals $T_\mathtt{symb} \in \{2, 5, 10\} \times T_\mathtt{max}$ where $T_\mathtt{max} = \mathop{\mathrm{argmax}}\limits_{t>0} \bar{c}_\mathtt{s}(t)$ and $\bar{c}_\mathtt{s}(t)$ is obtained from (28) assuming the nominal values of the MC system parameters in Table I. We observe from Fig. 8 that as $\mathtt{SNR}^\mathtt{ext} \to \infty$, the BERs of all detectors saturate to a BER floor if ISI is present. Moreover, due to the severe mismatch between the system model assumed for development of the
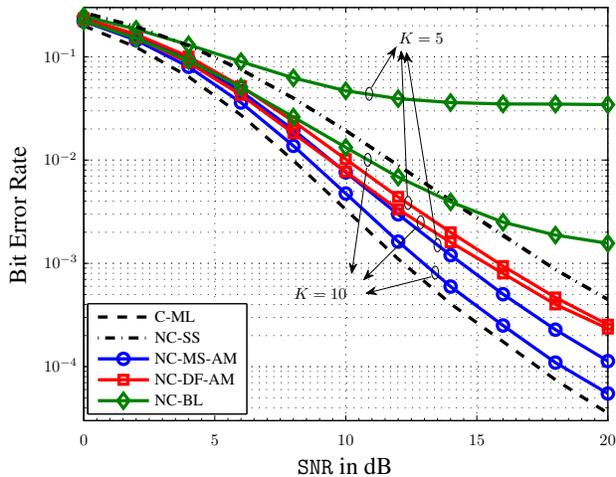
Fig. 7. Bit error rate versus the SNR in dB for Scenario 2 and $K \in \{5, 10\}$. All curves were obtained via simulation.
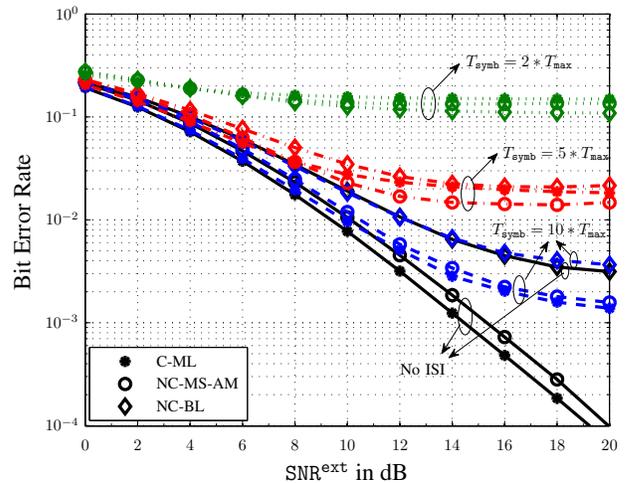
Fig. 8. Bit error rate versus the $\texttt{SNR}^{\texttt{ext}}$ in dB for Scenario 3 and $K = 10$. All curves were obtained via simulation.

considered detectors and the actual simulated MC channel for the strong ISI case, the relative performance of these detectors is not obvious. In fact, we observe from Fig. 8 that for the considered system parameters and $T_{\texttt{symb}} = 2T_{\texttt{max}}$, the performance of the coherent ML detector considerably deteriorates compared to that of the MS detector and, similarly, the performance of the MS detector degrades compared to that of the blind detector. This suggests that the coherent ML detector is more sensitive to the considered system model mismatch compared to the MS detector and, similarly, the MS detector is more sensitive compared to the blind detector. Finally, Fig. 8 shows that as the ISI becomes weaker, i.e., as $T_{\texttt{symb}}$ increases, the BERs of all detectors approach the respective BERs for the ISI-free case[4].

VI. CONCLUSIONS

We have derived the optimal non-coherent MS and SS detectors as well as a non-coherent DF detector which do not require instantaneous CSI knowledge. As compared to the coherent detectors previously studied, the proposed non-coherent detectors may be preferable in practical scenarios since the complexity and the overhead associated with CSI acquisition are avoided. In order to further reduce the complexity of our detectors, we proposed an approximate detection metric and a low-complexity suboptimal blind detector. We further derived an analytical expression for the BER of the optimal SS detector, and a lower bound and an upper bound for the BER of the optimal

---

[4]We note that optimal non-coherent detection for the case when strong ISI is present is an important research problem which is beyond the scope of this paper and left for future work.

MS detector. Simulation results confirmed the analysis and showed that the proposed optimal MS detector outperforms the suboptimal blind detector, particularly for small detection window sizes. However, as the size of the detection window increases, the performances of the proposed optimal and suboptimal detectors converge to that of the benchmark coherent ML detector which requires perfect CSI. This demonstrates the effectiveness of the proposed detection schemes.

APPENDIX A

PROOF OF THEOREM 1

Using sets $\mathcal{K}_1$ and $\mathcal{K}_0$, (6) can be simplified to

$$
\mathbf{s}^{\text{MS}} = \operatorname*{argmax}_{\mathbf{s} \in \mathcal{A}} \int_{\bar{c}_{\mathbf{s}} \geq 0} \int_{\bar{c}_{\mathbf{n}} \geq 0} (\bar{c}_{\mathbf{s}} + \bar{c}_{\mathbf{n}})^{\sum_{k \in \mathcal{K}_1} r[k]} \bar{c}_{\mathbf{n}}^{\sum_{k \in \mathcal{K}_0} r[k]}
$$
$$
\times \exp\left(-|\mathcal{K}_1|\bar{c}_{\mathbf{s}} - K\bar{c}_{\mathbf{n}}\right) f_{(\bar{c}_{\mathbf{s}}, \bar{c}_{\mathbf{n}})}(\bar{c}_{\mathbf{s}}, \bar{c}_{\mathbf{n}}) \mathrm{d}\bar{c}_{\mathbf{s}} \mathrm{d}\bar{c}_{\mathbf{n}}, \tag{30}
$$

where the term $r[k]!$ in (6) is removed in (30) since it does not affect the MS detection result. Moreover, the optimal MS detector can be written equivalently in expectation form as

$$
\mathbf{s}^{\text{MS}} = \operatorname*{argmax}_{\mathbf{s} \in \mathcal{A}} \mathbb{E}_{\bar{\mathbf{c}}} \left\{ \underbrace{(\bar{c}_{\mathbf{s}} + \bar{c}_{\mathbf{n}})^{\sum_{k \in \mathcal{K}_1} r[k]} \bar{c}_{\mathbf{n}}^{\sum_{k \in \mathcal{K}_0} r[k]}}_{A(\mathbf{s}, \mathbf{r})} \underbrace{\exp\left(-|\mathcal{K}_1|\bar{c}_{\mathbf{s}} - K\bar{c}_{\mathbf{n}}\right)}_{B(\mathbf{s}, \mathbf{r})} \right\}. \tag{31}
$$

We can conclude the following properties from the MS detection metric $\Lambda^{\text{MS}}(\mathbf{s}, \mathbf{r}) = \mathbb{E}_{(\bar{c}_{\mathbf{s}}, \bar{c}_{\mathbf{n}})}\{A(\mathbf{s}, \mathbf{r}) B(\mathbf{s}, \mathbf{r})\}$. First, with respect to the data sequence $\mathbf{s}$, term $B(\mathbf{s}, \mathbf{r})$ is only a function of the number of ones in $\mathbf{s}$. Second, for a given number of ones in $\mathbf{s}$, term $A(\mathbf{s}, \mathbf{r})$ is maximized if the ones in $\mathbf{s}$ correspond to the largest elements of the observation vector $\mathbf{r}$. Note that these two properties hold for any given CSI $\bar{\mathbf{c}}$. Hence, they also hold after the expectation operation with respect to $\bar{\mathbf{c}}$, i.e., $\mathbb{E}_{\bar{\mathbf{c}}}\{A(\mathbf{s}, \mathbf{r})B(\mathbf{s}, \mathbf{r})\}$. Therefore, we can avoid searching over all $\mathbf{s} \in \mathcal{A}$, and instead, find the optimal threshold $n_1 \in \{0, 1, \ldots, K\}$ which sets the elements of $\mathbf{s}$ corresponding to the n_1 largest elements of $\mathbf{r}$ to one and the remaining elements to zero. Moreover, we can further simplify the MS detection metric in (31) as

$$
\Lambda^{\text{MS}}(\mathbf{s}, \mathbf{r}) \overset{(a)}{=} \mathbb{E}_{(\bar{c}_{\mathbf{s}}, \bar{c}_{\mathbf{n}})} \left\{ \left[\sum_{i=0}^{N_1} \binom{N_1}{i} \bar{c}_{\mathbf{s}}^{N_1-i} \bar{c}_{\mathbf{n}}^{i} \right] \bar{c}_{\mathbf{n}}^{N_0} \exp\left(-|\mathcal{K}_1|\bar{c}_{\mathbf{s}} - K\bar{c}_{\mathbf{n}}\right) \right\}
$$
$$
\overset{(b)}{=} \sum_{i=0}^{N_1} \binom{N_1}{i} \mathbb{E}_{\bar{c}_{\mathbf{s}}} \left\{ \bar{c}_{\mathbf{s}}^{N_1-i} e^{-n_1 \bar{c}_{\mathbf{s}}} \right\} \mathbb{E}_{\bar{c}_{\mathbf{n}}} \left\{ \bar{c}_{\mathbf{n}}^{N_0+i} e^{-K\bar{c}_{\mathbf{n}}} \right\}, \tag{32}
$$

where in equality (a), we employ the Binomial expansion, i.e., $(x + y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$ with $\binom{n}{i} = \frac{n!}{i!(n-i)!}$, and in equality (b), we use $n_1 = |\mathcal{K}_1|$ and the assumption that RVs $\bar{c}_{\mathsf{s}}$ and $\bar{c}_{\mathsf{n}}$ are independent. Furthermore, recalling that $\Pr\{s[k] = 1\} = \Pr\{s[k] = 0\} = 0.5$ holds, the optimal threshold, denoted by $\zeta^{\mathtt{MS}}(\mathbf{r})$, approaches $K/2$ as $K \to \infty$. This concludes the proof.

APPENDIX B

PROOF OF COROLLARY 1

For the special case of the symbol-by-symbol detection, we obtain two detection metrics from (31) corresponding to $s[k] = 1$ and $s[k] = 0$, respectively, i.e.,

$$
\Lambda(s[k]) = \begin{cases} \mathbb{E}_{(\bar{c}_{\mathsf{s}}, \bar{c}_{\mathsf{n}})} \left\{ (\bar{c}_{\mathsf{s}} + \bar{c}_{\mathsf{n}})^{r[k]} e^{-(\bar{c}_{\mathsf{s}} + \bar{c}_{\mathsf{n}})} \right\}, & \text{if } s[k] = 1 \\ \mathbb{E}_{\bar{c}_{\mathsf{n}}} \left\{ \bar{c}_{\mathsf{n}}^{r[k]} e^{-\bar{c}_{\mathsf{n}}} \right\}, & \text{otherwise} \end{cases}
\tag{33}
$$

Now, using Lemma 1, we first show that $\frac{\Lambda(s[k]=1)}{\Lambda(s[k]=0)}$ is a monotonically increasing function of $r[k]$. Moreover, for $r[k] = 0$ and $r[k] \to \infty$, we obtain that $\Lambda(s[k] = 1) < \Lambda(s[k] = 0)$ and $\Lambda(s[k] = 1) > \Lambda(s[k] = 0)$ hold, respectively. Therefore, there exists a unique threshold for $r[k]$ below which $\Lambda(s[k] = 1) < \Lambda(s[k] = 0)$ holds.

Lemma 1: If $\frac{f_n(x)}{g_m(x)}$ is a monotonically increasing function of x for all possible pairs of (m, n), then function $\frac{\sum_n f_n(x)}{\sum_m g_m(x)}$ is also monotonically increasing in x.

Proof: Please refer to Appendix C. ∎

In order to apply the result of Lemma 1 to $\frac{\Lambda(s[k]=1)}{\Lambda(s[k]=0)}$, we first note that the expectation terms $\mathbb{E}_x\{f(x)\}$ can be written in summation form as $\sum_n \Pr\{x_n\} f(x_n)$ by discretizing the domain of x into a set $\{x_1, x_2, \dots\}$. Therefore, we have $\Lambda(s[k] = 1) \triangleq \sum_n f_n(r[k]) = \sum_n \Pr\{\bar{c}_{\mathsf{s}}\} \Pr\{\bar{c}_{\mathsf{n}}\} (\bar{c}_{\mathsf{s}} + \bar{c}_{\mathsf{n}})^{r[k]} e^{-(\bar{c}_{\mathsf{s}} + \bar{c}_{\mathsf{n}})}$ and $\Lambda(s[k] = 0) \triangleq \sum_m g_m(r[k]) = \sum_n \Pr\{\bar{c}_{\mathsf{n}}\} \bar{c}_{\mathsf{n}}^{r[k]} e^{-\bar{c}_{\mathsf{n}}}$. Now, we have to show that $\frac{f_n(r[k])}{g_m(r[k])} = \frac{\Pr\{\bar{c}_{\mathsf{s}}\} \Pr\{\bar{c}_{\mathsf{n}}\} (\bar{c}_{\mathsf{s}} + \bar{c}_{\mathsf{n}})^{r[k]} e^{-(\bar{c}_{\mathsf{s}} + \bar{c}_{\mathsf{n}})}}{\Pr\{\bar{c}_{\mathsf{n}}\} \bar{c}_{\mathsf{n}}^{r[k]} e^{-\bar{c}_{\mathsf{n}}}} = \Pr\{\bar{c}_{\mathsf{s}}\} \left(1 + \frac{\bar{c}_{\mathsf{s}}}{\bar{c}_{\mathsf{n}}}\right)^{r[k]} e^{-\bar{c}_{\mathsf{s}}}$ is a monotonically increasing function of $r[k]$ for all $\bar{c}_{\mathsf{s}}$ and $\bar{c}_{\mathsf{n}}$ which straightforwardly holds.

To further simplify the SS detection metric under hypothesis $s[k] = 1$, we employ the Binomial expansion which leads to

$$
\Lambda(s[k] = 1) = \sum_{i=0}^{r[k]} \binom{r[k]}{i} \mathbb{E}_{\bar{c}_{\mathsf{s}}} \left\{ \bar{c}_{\mathsf{s}}^{r[k]-i} e^{-\bar{c}_{\mathsf{s}}} \right\} \mathbb{E}_{\bar{c}_{\mathsf{n}}} \left\{ \bar{c}_{\mathsf{n}}^i e^{-\bar{c}_{\mathsf{n}}} \right\}.
\tag{34}
$$

The above results are concisely stated in Corollary 1 which concludes the proof.

APPENDIX C

PROOF OF LEMMA 1

If $\frac{f_n(x)}{g_m(x)}$ is a monotonically increasing function of x, its derivative with respect to x has to be positive, i.e.,

$$\frac{\partial}{\partial x}\left(\frac{f_n(x)}{g_m(x)}\right) = \frac{f_n'(x)g_m(x) - f_n(x)g_m'(x)}{(g_m(x))^2} > 0, \quad \forall n, m. \tag{35}$$

In other words, $f_n'(x)g_m(x) - f_n(x)g_m'(x) > 0$ has to hold for $\forall n, m$. Using this result, the sufficient condition for $\frac{\sum_n f_n(x)}{\sum_m g_m(x)}$ to be a monotonically increasing function of x can be shown as

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\sum_n f_n(x)}{\sum_m g_m(x)}\right) &= \frac{\sum_n f_n'(x) \times \sum_m g_m(x) - \sum_n f_n(x) \times \sum_m g_m'(x)}{(\sum_m g_m(x))^2} \\
&\overset{(a)}{=} \frac{\sum_n \sum_m f_n'(x)g_m(x) - \sum_n \sum_m f_n(x)g_m'(x)}{(\sum_m g_m(x))^2} \\
&= \frac{\sum_n \sum_m [f_n'(x)g_m(x) - f_n(x)g_m'(x)]}{(\sum_m g_m(x))^2} \overset{(b)}{>} 0,
\end{aligned} \tag{36}$$

where equality (a) follows the sum-product rule, i.e., $\sum_n x_n \sum_m y_m = \sum_n \sum_m x_n y_m$, and inequality (b) follows from (35). This completes the proof.

REFERENCES

[1] V. Jamali, N. Farsad, R. Schober, and A. Goldsmith, "Non-Coherent Multiple-Symbol Detection for Diffusive Molecular Communications," in *Proc. ACM NanoCom*, Sept. 2016.

[2] N. Farsad, H. Yilmaz, A. Eckford, C. Chae, and W. Guo, "A Comprehensive Survey of Recent Advancements in Molecular Communication," *IEEE Commun. Surveys Tutorials*, vol. 18, no. 3, pp. 1887–1919, third quarter 2016.

[3] I. Akyildiz, F. Brunetti, and C. Blazquez, "Nanonetworks: A New Communication Paradigm," *Comput. Net.*, vol. 52, pp. 2260–2279, Apr. 2008.

[4] I. F. Akyildiz, M. Pierobon, S. Balasubramaniam, and Y. Koucheryavy, "The Internet of Bio-Nano Things," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 32–40, Mar. 2015.

[5] B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Essential Cell Biology*. New York, NY: Garland Science, 4th ed., 2014.

[6] H. C. Berg, *Random Walks in Biology*. Princeton University Press, 1993.

[7] P. Nelson, *Biological Physics: Energy, Information, Life*. Freeman, 1st ed., 2008.

[8] H. Arjmandi, A. Gohari, M. Kenari, and F. Bateni, "Diffusion-Based Nanonetworking: A New Modulation Technique and Performance Analysis," *IEEE Commun. Lett.*, vol. 17, no. 4, pp. 645–648, Apr. 2013.

[9] V. Jamali, M. Ahmadzadeh, and R. Schober, "On the Design of Matched Filters for Molecule Counting Receivers," *IEEE Commun. Lett.*, 2017.

[10] A. Noel, K. 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.

[11] M. Mahfuz, D. Makrakis, and H. Mouftah, "A Comprehensive Study of Sampling-Based Optimum Signal Detection in Concentration-Encoded Molecular Communication," *IEEE Trans. NanoBiosci.*, vol. 13, no. 3, pp. 208–222, Sept. 2014.

[12] V. Jamali, A. Ahmadzadeh, C. Jardin, C. Sticht, and R. Schober, "Channel Estimation for Diffusive Molecular Communications," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4238–4252, Oct. 2016.

[13] M. Ahmadzadeh, V. Jamali, A. Noel, and R. Schober, "Diffusive Mobile Molecular Communications Over Time-Variant Channels," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1265–1268, Jun. 2017.

[14] M. Ahmadzadeh, V. Jamali, and R. Schober, "Statistical Analysis of Time-Variant Channels in Diffusive Mobile Molecular Communications," *Accepted for presentation in IEEE Globecom*, 2017. [Online]. Available: https://arxiv.org/abs/1704.06298

[15] D. Divsalar and M. K. Simon, "Multiple-Symbol Differential Detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, Mar 1990.

[16] M. Riediger, R. Schober, and L. Lampe, "Multiple-Symbol Detection for Photon-Counting MIMO Free-Space Optical Communications," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5369–5379, Dec. 2008.

[17] B. Li, M. Sun, S. Wang, W. Guo, and C. Zhao, "Low-Complexity Noncoherent Signal Detection for Nanoscale Molecular Communications," *IEEE Trans. on NanoBiosci.*, vol. 15, no. 1, pp. 3–10, Jan. 2016.

[18] ——, "Local Convexity Inspired Low-Complexity Noncoherent Signal Detector for Nanoscale Molecular Communications," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2079–2091, May 2016.

[19] A. Akkaya, H. Yilmaz, C. Chae, and T. Tugcu, "Effect of Receptor Density and Size on Signal Reception in Molecular Communication via Diffusion With an Absorbing Receiver," *IEEE Commun. Lett.*, vol. 19, no. 2, pp. 155–158, Feb. 2015.

[20] A. Ahmadzadeh, H. Arjmandi, A. Burkovski, and R. Schober, "Comprehensive Reactive Receiver Modeling for Diffusive Molecular Communication Systems: Reversible Binding, Molecule Degradation, and Finite Number of Receptors," *IEEE Trans. NanoBiosci.*, vol. 15, no. 7, pp. 713–727, Oct 2016.

[21] H. B. Yilmaz and C. B. Chae, "Arrival Modelling for Molecular Communication via Diffusion," *Electron. Lett.*, vol. 50, no. 23, pp. 1667–1669, Nov. 2014.

[22] R. Mosayebi, H. Arjmandi, A. Gohari, M. Nasiri-Kenari, and U. Mitra, "Receivers for Diffusion-Based Molecular Communication: Exploiting Memory and Sampling Rate," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 12, pp. 2368–2380, Dec. 2014.

[23] V. Jamali, Ahmadzadeh, and R. Schober, "Symbol Synchronization for Diffusive Molecular Communications," in *Proc. IEEE ICC*, May 2017.

[24] A. Noel, K. Cheung, and R. Schober, "Improving Receiver Performance of Diffusive Molecular Communication with Enzymes," *IEEE Trans. NanoBiosci.*, vol. 13, no. 1, pp. 31–43, Mar. 2014.

[25] N. Farsad and A. Goldsmith, "A Novel Molecular Communication System Using Acids, Bases and Hydrogen Ions," *Available online on arXiv*, 2015. [Online]. Available: http://arxiv.org/abs/1511.08957

[26] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*. Taylor & Francis, 2014, vol. 2.

[27] A. Noel, K. Cheung, and R. Schober, "A Unifying Model for External Noise Sources and ISI in Diffusive Molecular Communication," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 12, pp. 2330–2343, Dec. 2014.

[28] X. Wang, M. Higgins, and M. Leeson, "Distance Estimation Schemes for Diffusion Based Molecular Communication Systems," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 399–402, Mar. 2015.

[29] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. 7th ed. Academic, 2007.

[30] K. V. Srinivas, A. W. Eckford, and R. S. Adve, "Molecular Communication in Fluid Media: The Additive Inverse Gaussian Noise Channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4678–4692, Jul. 2012.

[31] N. Farsad, W. Guo, C. B. Chae, and A. Eckford, "Stable Distributions as Noise Models for Molecular Communication," in *Proc. IEEE Globecom*, Dec. 2015, pp. 1–6.