Communication over Diffusion-Based Molecular Timing Channels

Yonathan Murin, Nariman Farsad, Mainak Chowdhury, and Andrea Goldsmith
Department of Electrical Engineering, Stanford University, USA

***Abstract*—This work studies communication over diffusion-based molecular timing (DBMT) channels. The transmitter *simultaneously* releases *multiple* small information particles, where the information is encoded in the time of release. The receiver decodes the transmitted information based on the random time of arrival of the information particles, which is represented as an additive noise channel. For a DBMT channel, without flow, this noise follows the Lévy distribution. Under this channel model, the maximum-likelihood (ML) detector is derived and shown to have high computational complexity. It is further shown that for *any* additive noise channel with α-stable noise, $\alpha < 1$, such as the DBMT channel, a linear receiver is *not able to take advantage* of the release of multiple information particles. Thus, instead of the common low complexity linear approach, a new detector, which is based on the first arrival (FA) among all the transmitted particles, is derived. Numerical simulations indicate that for a small to medium number of released particles, the performance of the FA detector is very close to the performance of the ML detector.**

I. Introduction

Molecular communications (MC) is an emerging field in which nano-scale devices communicate with each other via chemical signaling, based on exchanging small *information particles* [1]. To embed information in these particles one may use the particle's type, concentration, or the time of release, as described in [2] and references therein. Particles can be transported from the transmitter to the receiver via diffusion, active transport, bacteria, and flow. Although this new field is still in its infancy, several basic experimental systems serve as a proof of concept for transmitting short messages at low bit rates [3], [4].

This work focuses on receiver design for MC systems where information is modulated through the *time of release of the information particles*, which is reminiscent of pulse position-modulation. A common assumption, which is accurate for many sensors, is that after some time duration the particle is absorbed by the receiver and removed from the environment. In this case, the random delay until the particle arrives at the receiver can be represented as an additive noise term. For diffusion-based channels *without flow*, this additive noise is Lévy-distributed [5], [6], [7], while for diffusion-based channels *with flow*, this additive noise follows an inverse Gaussian (IG) distribution [8].

At first glance, the cases of diffusion with and without flow may seem similar; however, a closer look reveals a fundamental difference which stems from the different properties of the additive noise. The Lévy distribution has a heavy tail, namely, its tail decays polynomially, while the tail of the IG distribution, similarly to the standard Gaussian distribution, decays exponentially. Thus, traditional linear detection and signal processing techniques, which work well in the presence of Gaussian or IG noise, will not necessarily be efficient in the presence of a Lévy-distributed noise. The observation that channels with additive heavy-tailed noise require different detection methods was already stated in [9] based on numerical simulations. In this work we provide a rigorous proof that linear processing *cannot improve* the detection performance in channels corrupted by additive α-stable noise [10], [11], with $\alpha < 1$. In particular, linear processing cannot improve the detection performance in DBMT channels, as the Lévy distribution is α-stable with $\alpha = \frac{1}{2}$.

Apart from the fact that the tails of the additive noise decay slowly, in the considered diffusion-based timing channel ordering in time is not preserved, namely, information particles from consecutive channel uses may arrive out of order [12]. This gives rise to inter-symbol interference (ISI). In this work, however, we focus on settings without ISI such as communication systems in which consecutive transmissions are far enough apart, or a nano-scale sensor which occasionally sends a limited number of bits, modulated in a single channel use, to a centralized molecular controller, and then remains silent for a long period. Hence, we study independent consecutive channel uses without ISI such that each transmission can be analyzed separately.

In this paper we study transmission over a diffusion-based molecular timing (DBMT) channel without flow, assuming that consecutive channel uses are independent and identically distributed (i.i.d). We consider an MC system in which the information is encoded in the time of release of the information particles, where this time is selected out of a set with *finite cardinality*. At each transmission M *information particles are simultaneously* released at the time corresponding to the current symbol, while the receiver's objective is to detect this transmission time. Note that M is constant and does not change from one transmission to the next, i.e., information is not encoded in the number of particles.

We derive the ML detection rule which, as expected, requires relatively high computational complexity and therefore is not necessarily suitable for nano-scale MC systems. This motivates studying detectors with lower complexity. A common approach in traditional electromagnetic (EM) communication, which was also proposed in [8] for an MC system, is to use a linear detector. We show that for *any* $\alpha-$stable additive noise with $\alpha < 1$, and in particular for Lévy-distributed noise, linear processing increases the dispersion of the noise compared to the case of a *single particle*. As higher dispersion is reminiscent of higher variance[1], this increased dispersion degrades the probability of correct detection, compared to the case of a *single particle*. To the best of our knowledge this is the first proof for the destructive effect of linear processing in a MC system.

Since linear detectors cannot take advantage of the multiple transmitted particles, we derive a new detector which is based on the first arrival (FA) among the M information particles. We show that the conditional probability densities of the FA concentrate around the possible transmitted signal, thus, increasing the probability of correct detection, compared to the case of a single particle. Via numerical simulations we show that for small-to-medium values of M, the performance of the proposed FA detector is very close to that of the optimal detector. On the other hand, for large values of M, ML significantly outperforms the FA detector, which agrees with the fact that the FA is *not* a sufficient statistic for the considered detection problem.

The rest of this paper is organized as follows. The problem formulation is presented in Section II. The ML detector and linear detection are studied in Section III. The FA detector is derived in Section IV. Numerical results are presented in Section V, and concluding remarks are provided in Section VI.

Notation: We denote the set of real numbers by $\mathcal{R}$ and the set of positive real numbers by $\mathcal{R}^+$. Other than these sets, we denote sets with calligraphic letters, e.g., $\mathcal{X}$. We denote RVs with upper case letters, e.g., X, and their realizations with lower case letters, e.g., x. An RV takes values in the set $\mathcal{X}$, and we use $|\mathcal{X}|$ to denote the cardinality of a finite set. We use $f_Y(y)$ to denote the probability density function (PDF) of a continuous RV Y on $\mathcal{R}$, $f_{Y|X}(y|x)$ to denote the conditional PDF of Y given X, and $F_{Y|X}(y|x)$ to denote the conditional cumulative distribution function (CDF). We denote vectors with boldface letters, e.g., $\mathbf{x}$, where the k^{th} element of a vector $\mathbf{x}$ is denoted by x_k. Finally, we use $\mathsf{erfc}(\cdot)$ to denote the complementary error function given by $\mathsf{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-u^2}du$ and $\log(\cdot)$ to denote the natural logarithm.

II. Problem Formulation

A. System Model

We assume that the information particles themselves are *identical and indistinguishable* at the receiver. Therefore, the receiver can only use the time of arrival to decode the intended message. The information particles propagate from the transmitter to the receiver through some random propagation mechanism (e.g. diffusion). We make the following assumptions about the system:

A1) The transmitter perfectly controls the release time of each information particle, and the receiver perfectly measures the arrival times of the information particles. Moreover, the transmitter and the receiver are perfectly synchronized in time.

A2) An information particle which arrives at the receiver is absorbed and removed from the propagation medium.

A3) All information particles propagate independently of each other, and their trajectories are random according to an i.i.d. random process.[2]

Note that these assumptions are consistent with those made in all previous works, in order to make the models tractable (see [6] and the references therein).

Let $\mathcal{X}$ be a finite set of constellation points on the real line: $\mathcal{X} \triangleq \{\xi_0, \xi_1, \dots, \xi_{L-1}\}$, $0 \le \xi_0 \le \cdots \le \xi_{L-1}$, and let $\xi_{L-1} < T_s < \infty$ denote the symbol duration. The k^{th} transmission takes place at time $(K-1)T_s + X_k, X_k \in \mathcal{X}, k = 1, 2, \dots, K$. At this time, M information particles are *simultaneously* released into the medium by the transmitter. The transmitted information is encoded in the sequence $\{(K-1)T_s + X_k\}_{k=1}^K$, which is assumed to be independent of the random propagation time of *each* of the information particles. Let $\mathbf{Y}_k$ denote an M-length vector consisting of the times of arrival of each of the information particles released at time $(k-1)T_s + X_k$. It follows that $Y_{k,m} > X_k, m = 1, 2, \dots, M$. Thus, we obtain the following additive noise channel model:

$$Y_{k,m} = (k-1)T_s + X_k + Z_{k,m}, \tag{1}$$

where $Z_{k,m}$, is a random noise term representing the propagation time of the m^{th} particle of the k^{th} transmission. Assumption **A3**) implies that all the RVs $Z_{k,m}$ are independent.

In the channel model (1), particles may arrive out of order, which results in a channel with memory. In this work, however, we assume that each information particle arrives before the next transmission takes place. This assumption can be formally stated as:

A4) T_s is a fixed constant chosen to be large enough such that the transmission times X_k obey $Y_{k,m} \le kT_s$ with high probability.[3]

With this assumption, we obtain an i.i.d. memoryless channel model which can be written as:

$$Y_m = X + Z_m, \quad m = 1, 2, \dots, M. \tag{2}$$

Assumption **A4**) implies that T_s is chosen such that consecutive transmissions are sufficiently separated in time. We further note that the model (2) also represents well the setting of a nano-scale sensor which infrequently sends a symbol (which conveys a limited number of bits) to a centralized molecular controller, and then remains silent for a long period. Thus, the effective communication channel is memoryless.

To simplify the presentation, we restrict our attention to the case of binary modulations, i.e., $\mathcal{X} = \{0, \Delta\}$, where the constellation points are sent with equal probability. We note that all the results and techniques derived in this work can be easily extended to more than two elements in the set $\mathcal{X}$, and to

[1] Recall that the variance of $\alpha-$stable, $\alpha < 1$, random variable (RV) is ∞.

[2] This is a reasonable assumption for many different propagation schemes in molecular communication such as diffusion in dilute solutions, i.e., when the number of particles released is much smaller than the number of molecules of the solutions.

[3] Formally, let η be arbitrarily high probability, then we choose T_s such that $\Pr\{Y_{k,m} < kT_s\} > \eta, k = 1, 2, \dots, K, m = 1, 2, \dots, M$.

unequal a-priori probabilities. Let $\hat{X}$ denote the estimation of X at the receiver. Our objective is to design a (simple) receiver which minimizes the probability of error $P_\varepsilon = \Pr\{X \neq \hat{X}\}$. Note that the above description of communication over an MT channel is fairly general and can be applied to different propagation mechanisms as long as Assumptions **A1)**–**A4)** are not violated. Next, we describe the DBMT channel.

B. The DBMT Channel

In diffusion-based propagation, the released particles follow a random Brownian path from the transmitter to the receiver. In this case, to specify the random additive noise term Z_m in (2), we define a Lévy-distributed RV as follows:

Definition 1. Let Z be Lévy-distributed with location parameter μ and scale parameter c [10]. Then, its PDF given by:

$$f_Z(z) = \begin{cases} \sqrt{\frac{c}{2\pi(z-\mu)^3}} \exp\left(-\frac{c}{2(z-\mu)}\right), & z > \mu \\ 0, & z \leq \mu \end{cases}, \quad (3)$$

and its CDF given by:

$$F_Z(z) = \begin{cases} \mathsf{erfc}\left(\sqrt{\frac{c}{2(z-\mu)}}\right), & z > \mu \\ 0, & z \leq \mu \end{cases}. \quad (4)$$

Let d denote the distance between the transmitter and the receiver, and D denote the diffusion coefficient of the information particles in the propagation medium. Following along the lines of the derivations in [8, Sec. II], and using the results of [13, Sec. 2.6.A], it can be shown that for 1-dimensional pure diffusion, the propagation time of each of the information particles follows a Lévy distribution, denoted in this work by $\sim \mathscr{L}(\mu, c)$ with $c = \frac{d^2}{2D}$ and $\mu = 0$. Thus, $Z_m \sim \mathscr{L}(0, c), m = 1, 2, \ldots, M$.

The Lévy distribution belongs to the class of stable distributions, discussed in the next subsection. For a detail description we refer the reader to [10], [11].

C. Stable Distributions

Definition 2. An RV X has a stable distribution if for two independent copies of X, X_1 and X_2, and for any constants $a_1, a_2 \in \mathcal{R}^+$, there exists constants $a_3 \in \mathcal{R}^+$ and $a_4 \in \mathcal{R}$ such that $a_1X_1 + a_2X_2 \stackrel{d}{=} a_3X + a_4$, where $\stackrel{d}{=}$ denotes equality in distribution.

Stable distributions can also be defined via their characteristic function:

Definition 3. Let $\mu \in \mathcal{R}, c \in \mathcal{R}^+, 0 < \alpha \leq 2$, and $-1 \leq \beta \leq 1$. Further define:

$$\Phi(t, \alpha) \triangleq \begin{cases} \tan\left(\frac{\pi\alpha}{2}\right), & \alpha \neq 1 \\ -\frac{2}{\pi}\log(|t|), & \alpha = 1 \end{cases}.$$

Then, the characteristic function of a stable RV X, with location parameter μ, scale (or dispersion) parameter c, characteristic exponent α, and skewness parameter β, is given by:

$$\varphi(t; \mu, c, \alpha, \beta) = \exp\left\{j\mu t - |ct|^\alpha (1 - j\beta \mathsf{sgn}(t)\Phi(t, \alpha))\right\}. \quad (5)$$

In the following, we use the notation $\mathscr{S}(\mu, c, \alpha, \beta)$ to represent a stable distribution with the parameters μ, c, α, and β. Apart from several special cases, stable distributions do not have closed-form PDFs. The exceptional cases are the Gaussian distribution ($\alpha = 2$), the Cauchy distribution ($\alpha = 1$), and the case of $\alpha = \frac{1}{2}$ with general β, which was very recently derived in [5, Theorem 2]. Note that the Lévy distribution is a special case of the results of [5] with $\beta = 1$. Finally, we note that all stable distributions, apart from the case $\alpha = 2$, have infinite variance, and all stable distributions with $\alpha \leq 1$ also have infinite mean. In fact, this statement can be generalized to moments of order $p \leq \alpha$, see [14].

Next, we study ML and linear detection for transmission over the DBMT channel.

III. Transmission over the DBMT Channel: ML and Linear Detection

We begin this section with the relatively simple case in which a single information particle is released, i.e., $M = 1$. For this setup, the decision rule which minimizes the probability of error, and the minimal probability of error, are given in the following proposition:

Proposition 1. The decision rule which minimizes the probability of error when $M = 1$, is given by:

$$\hat{X}_{\text{ML}}(y_1) = \begin{cases} 0, & y_1 < \theta \\ \Delta, & y_1 \geq \theta, \end{cases} \quad (6)$$

where θ is the unique solution, in the interval $[\Delta, \Delta + \frac{c}{3}]$, of the following equation in y_1:

$$y_1(y_1 - \Delta)\log\left(\frac{y_1}{y_1 - \Delta}\right) = \frac{c\Delta}{3}, \quad y_1 > \Delta > 0. \quad (7)$$

Furthermore, the probability of error of this decision rule is given by:

$$P_\varepsilon = 0.5\left(1 - \mathsf{erfc}\left(\sqrt{\frac{c}{2\theta}}\right) + \mathsf{erfc}\left(\sqrt{\frac{c}{2(\theta - \Delta)}}\right)\right). \quad (8)$$

Proof: The proof is provided in [15]. ■

The probability of error in molecular communications can be reduced by transmitting multiple information particles for each symbol [8], namely, using $M > 1$ particles for each transmission.[4] In fact, in [6] we showed that the capacity of the DBMT channel scales linearly with M. Yet, in [6] we did not provide analysis of the probability of error, nor an efficient decoding method. In this section we first present the ML detector for the DBMT channel, and then discuss more practical detection approaches.

A. ML Detection for $M > 1$

Let $\mathbf{y} = \{y_m\}_{m=1}^M$. The following proposition characterizes the ML detector based on the channel output $\mathbf{y}$:

Proposition 2. The decision rule which minimizes the probability of error for $M \geq 1$, is given by:

[4] As we assume that the transmitter and the receiver are perfectly synchronized, the best strategy is to simultaneously release M molecules. Releasing the M molecules in different times can only increase the ambiguity at the receiver and therefore increase the probability of error [8, Sec IV.C].

$$\hat{X}_{\text{ML}}(\mathbf{y}) = \begin{cases} 0, & \sum_{m=1}^{M} \log\left(\frac{y_m - \Delta}{y_m}\right) + \frac{c\Delta}{3}\frac{1}{y_m(y_m-\Delta)} > 0 \\ \Delta, & \text{otherwise} \end{cases}. \quad (9)$$

Proof: A proof outline is provided in [15]. ∎

Although the above ML detector minimizes the probability of error for equiprobable signaling, it lacks an exact performance analysis and is relatively complicated to compute in nano-scale devices; this in particular holds for the $\log(\cdot)$ operation. In traditional EM communication the common approach is to apply a *linear signal processing* based on the sequence $\mathbf{y}$. The complexity of such a receiver is significantly lower compared to that of the ML detector, and for an additive white Gaussian noise (AWGN) channel this approach is known to be optimal [16, Ch. 3.3]. In fact, even in non-Gaussian settings such as transmission over a timing channel with drift [8, Sec. IV.C.2], modeled by the additive IG noise (AIGN) channel, the linear approach yields significant performance gains compared to the transmission of a single particle (even though it is not necessarily the optimal detector). We note that linear techniques are favored for signal processing problems involving processes with *finite second-order moment* [14, Sec. IV].[5] However, the second moment of the additive Lévy noise is infinite.

In the next subsection we argue that when the transport mechanism is based only on diffusion, then such a linear receiver in fact *degrades* the performance *compared to the transmission of a single particle*. The sub-optimality of linear signal processing of signals corrupted by α-stable additive noise was already observed in [9, Ch. 10.4.6], yet, to the best of our knowledge, the analysis in the next subsection is the first to rigorously show that linear signal processing is not only sub-optimal, but can also degrade the performance.

B. Linear Detection for $M > 1$

In this subsection we consider linear detection of signals transmitted over an additive channel corrupted by an α-stable noise with characteristic exponent smaller than unity, namely, we use the channel model (2), with the minor change that $Z_m \sim \mathscr{S}(0, c, \beta, \alpha), \alpha < 1$.[6] Thus, the results presented in this subsection also hold for the Lévy-distributed noise. Let $\{w_m\}_{m=1}^{M}$, $w_m \in \mathcal{R}^+, \sum_{m=1}^{M} w_m = 1$ be a set of coefficients, and consider ML detection based on $Y_{\text{LIN}} \triangleq \sum_{m=1}^{M} w_m Y_m$:

$$\hat{X}_{\text{LIN}} = \underset{x \in \{0,\Delta\}}{\operatorname{argmax}} \, f_{Y_{\text{LIN}}|X}(y_{\text{LIN}}|X = x). \quad (10)$$

Let $P_{\varepsilon,\text{LIN}}$ denote the probability of error of the detector $\hat{X}_{\text{LIN}}$. We now have the following theorem:

Theorem 1. The probability of error of the linear detector is higher than the probability of error of the detector in (6), namely, $P_{\varepsilon,\text{LIN}} \geq P_{\varepsilon}$, where P_{ε} is given in (8).

Proof: We show that given $X = x$, $Y_{\text{LIN}} \sim \mathscr{S}(x, c_{\text{LIN}}, \alpha, \beta)$, with $c_{\text{LIN}} \geq c$. Note that when $X = x$ is given, then the Y_m's are independent. Therefore, the characteristic function of Y_{LIN}, given $X = x$, is given by:

$$\begin{aligned} &\varphi_{Y_{\text{LIN}}|X=x}(t) \\ &\overset{(a)}{=} \prod_{m=1}^{M} \exp\left\{jxw_m t - |cw_m t|^{\alpha}(1 - j\beta \mathsf{sgn}(t)\Phi(t,\alpha))\right\} \\ &\overset{(b)}{=} \exp\left\{jxt - \left(\sum_{m=1}^{M} cw_m^{\alpha}\right)|t|^{\alpha}(1 - j\beta \mathsf{sgn}(t)\Phi(t,\alpha))\right\} \\ &\overset{(c)}{=} \exp\left\{jxt - |c_{\text{LIN}} t|^{\alpha}(1 - j\beta \mathsf{sgn}(t)\Phi(t,\alpha))\right\}, \end{aligned} \quad (11)$$

where (a) follows from the fact that $w_m > 0$ and from the fact that $\Phi(t,\alpha)$ is independent of t, for $\alpha < 1$; (b) follows from the fact that $\sum_{m=1}^{M} w_m = 1$; and (c) follows by defining $c_{\text{LIN}} = c \cdot \left(\sum_{m=1}^{M} w_m^{\alpha}\right)^{\frac{1}{\alpha}}$. Therefore, given $X = x$, we have $Y_{\text{LIN}} \sim \mathscr{S}(x, c_{\text{LIN}}, \alpha, \beta)$. Since $w_m \leq 1$, we have $\left(\sum_{m=1}^{M} w_m^{\alpha}\right)^{\frac{1}{\alpha}} \geq 1$, and therefore $c_{\text{LIN}} \geq c$. Finally, as c is the dispersion of the distribution, and since stable distributions are unimodal [11, Ch. 2.7], it follows that the probability of error increases with c. Therefore, we conclude that $P_{\varepsilon,\text{LIN}} \geq P_{\varepsilon}$. ∎

As the Lévy distribution is a special case of the family $\mathscr{S}(0, c, \beta, \alpha), \alpha < 1$, it follows that the linear detector degrades the performance compared to the case of $M = 1$. This is numerically demonstrated in Section V.

Remark 1. The difference between the AIGN channel (or the AWGN channel) and the channel considered in this paper stems from the fact that for the AIGN, (weighted) averaging can decrease the noise variance, namely, the tails of the noise. On the other hand, in the case of the Lévy distribution, averaging leads to a heavier tail, and therefore to a higher probability of error.

Remark 2. In order to implement the ML detector (9), the receiver must wait for *all* particles to arrive. However, as the Lévy distribution has heavy tails, this may result in very long reception intervals. In fact, the average waiting time of such a receiver will be infinite. The detector in (10) can be implemented with shorter waiting intervals, yet, these intervals are significantly longer then the ones required in the single-particle detection problem.

In the next section we present a simple detector. This detector requires a short reception interval and achieves performance very close to that achieved by the ML detector.

IV. Transmission over the DBMT Channel for $M > 1$: FA Detection

The detector proposed in this section detects the transmitted symbol based only on the FA among the M particles, namely, it waits for the first particle to arrive and then applies ML detection based on this arrival. In terms of complexity, the FA detector simply compares the first arrival to a threshold; this is in contrast to the complicated ML detector in (9).

Let $y_{\text{FA}} \triangleq \min\{y_1, y_2, \ldots, y_M\}$. In [15] we show that the PDF of Y_{FA} is more concentrated around the transmitted symbol than the original Lévy distribution, which is reminiscent

[5] Note that the IG distribution considered in [8] has a finite variance.

[6] Timing channels with $\alpha < 1$ and $\beta \neq \frac{1}{2}$ were studied in [5].

$$y_{\text{FA}}(y_{\text{FA}}-\Delta)\left(\log\left(\frac{y_{\text{FA}}}{y_{\text{FA}}-\Delta}\right)+\frac{2(M-1)}{3}\log\left(\frac{1-\mathsf{erfc}\left(\sqrt{\frac{c}{2(y_{\text{FA}}-\Delta)}}\right)}{1-\mathsf{erfc}\left(\sqrt{\frac{c}{2y_{\text{FA}}}}\right)}\right)\right)=\frac{c\Delta}{3}. \tag{13}$$

of the lower variance achieved by averaging in AIGN and AWGN channels. The FA detector is presented in the following theorem:

Theorem 2. The decision rule which minimizes the probability of error, based on y_{FA}, is given by:

$$\hat{X}_{\text{FA}}(y_{\text{FA}})=\begin{cases}0, & y_{\text{FA}}<\theta_M\\ \Delta, & y_{\text{FA}}\geq\theta_M,\end{cases} \tag{12}$$

where $\Delta\leq\theta_M\leq\theta_{M-1}, \theta_1=\theta$, is the solution, in $y_{\text{FA}}\geq\Delta>0$, of (13) at the top of the page. Furthermore, the probability of error of the FA detector is given by:

$$P_{\varepsilon,\text{FA}}=0.5\left(\left(1-\mathsf{erfc}\left(\sqrt{\frac{c}{2\theta_M}}\right)\right)^M+1-\left(1-\mathsf{erfc}\left(\sqrt{\frac{c}{2(\theta_M-\Delta)}}\right)\right)^M\right). \tag{14}$$

Proof: Let $F_{Y|X}(y|x)$ denote the CDF of y_m given X. Assumption **A3**) implies that given X, the channel outputs $Y_1,Y_2,\ldots,Y_M$ are independent. Hence, using results from order statistics we write:

$$\begin{aligned}F_{Y_{\text{FA}}|X}(y|x)&=1-(\Pr\{Y>y|X=x\})^M\\&\overset{(a)}{=}1-\left(1-\mathsf{erfc}\left(\sqrt{\frac{c}{2(y-x)}}\right)\right)^M\end{aligned} \tag{15}$$

where (a) follows from (4). Next, to obtain the PDF of Y_{FA} given X, we write:

$$\begin{aligned}f_{Y_{\text{FA}}|X}(y|x)&=\frac{\partial F_{Y_{\text{FA}}|X}(y|x)}{\partial y}\\&=M\cdot f_{Y|X}(y|x)\cdot\left(1-\mathsf{erfc}\left(\sqrt{\frac{c}{2(y-x)}}\right)\right)^{M-1}.\end{aligned} \tag{16}$$

The ML decision rule is now obtained by simply comparing $f_{Y_{\text{FA}}|X}(y|x=0)$ and $f_{Y_{\text{FA}}|X}(y|x=\Delta)$. Plugging in the density of (16), and applying some algebraic manipulations we obtain (13).

To show that $\theta_M\leq\theta_{M-1}$ we first note that by plugging (16) into the ML decision rule, it follows that θ_M is the solution of the following equation:

$$\frac{f_{Y|X}(y_{\text{FA}}|x=0)}{f_{Y|X}(y_{\text{FA}}|x=\Delta)}=\left(\frac{1-\mathsf{erfc}\left(\sqrt{\frac{c}{2(y_{\text{FA}}-\Delta)}}\right)}{1-\mathsf{erfc}\left(\sqrt{\frac{c}{2y_{\text{FA}}}}\right)}\right)^{M-1}. \tag{17}$$

Now, for $M=1$, the RHS of (17) equals 1, and $\theta_1\in[\Delta,\Delta+\frac{c}{3}]$. Thus, the LHS of (17) is equal to 1 in this interval. An explicit evaluation of the derivative of the LHS of (17) shows that in this range the derivative is negative, and therefore the LHS of (17) decreases with y_{FA}, independently of M. On the other hand, the RHS of (17) increases with M for all $y_{\text{FA}}\geq\Delta$. Thus, we conclude that the solution of (17) decreases with M.

Regarding the probability of error, we first note that for $y_{\text{FA}}<\Delta$, due to the causality of the arrival time, X must be equal to 0, and therefore the probability of error is zero. For $y_{\text{FA}}\geq\Delta$ we use the fact that the constellation points are equiprobable, and the CDF in (15) to obtain (14). ■

Remark 3. The FA detection framework can be directly extended to the case of constellations with more than two elements, i.e., $L>2$. In such cases the detection will be based on $L-1$ thresholds, which define the L constellation points. Furthermore, as the conditional PDFs concentrate near x when M increases, we conclude that by increasing M one can support larger L for a given target probability of error (note that when $L>2$, then $\Pr\{X\neq\hat{X}\}$ refers to the *symbol* error probability).

Remark 4. In [15] we analytically compare the ML and FA detectors. For small to medium values of M we show that the FA and ML detectors achieve almost the same probability of error. Then, we use error exponent analysis to show that for large values of M the ML detector outperforms the FA detector. This is demonstrated in the following section.

V. Numerical Results

We begin our numerical evaluations with the probability of error for the different detectors. Fig. 1 depicts the probability of error versus different values of Δ, for $M=1,2,3$, for the ML, FA, and linear detectors. Throughout this section 10^6 trials were carried out for each Δ point. When $M=1$ all the detectors are identical. For larger values of M, the probability of error of the ML detector was evaluated numerically, while the probability of error of the FA detector was calculated using (14). For the linear detector we assumed $w_m=\frac{1}{M}$ which leads to $c_{\text{LIN}}=Mc$. It can be observed that the probability of error decreases with Δ and with M (for the ML and FA detectors). Moreover, as stated in Remark 4, Fig. 1 shows that the ML and FA detectors are practically indistinguishable for small values of M. Finally, note that the linear detector indeed degrades the performance as M increases.

Fig. 2 depicts the probability of error versus the number of released particles M, for the ML and FA detectors, $\Delta=0.2, 0.5$, and $c=2$. Here, 10^6 trials were carried out for each M point. It can be observed that for small values of M, as indicated by Fig. 1, the FA and ML are indistinguishable. On the other hand, when M increases, the superiority of the ML detector is revealed, e.g., $M\approx 50$. This further supports the results reported in Remark 4.

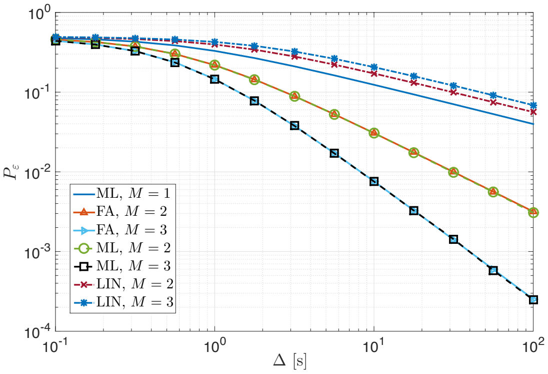

Fig. 1: P_ε vs. Δ, for $c = 1$ and $M = 1, 2, 3$.

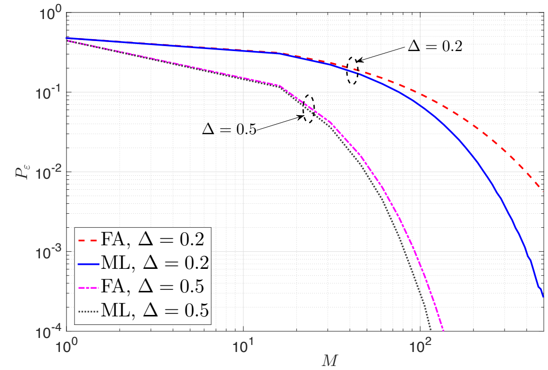

Fig. 2: P_ε vs. M, for $c = 2$ and $\Delta = 0.2, 0.5$.

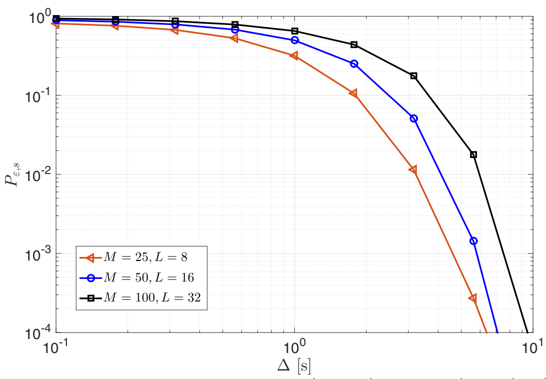

Fig. 3: $P_{\varepsilon,s}$ vs. Δ, for $c=1$ and (M, L) pairs: $(25, 8)$, $(50, 16)$, $(100, 32)$.

Finally, Fig. 3 demonstrates that for a given Δ and L, by using M large enough, one can theoretically achieve any target probability of error. For instance, Fig. 3 shows that a *symbol* probability of error of 10^{-3} can be achieved with $\Delta \approx 7s$, for different pairs of L and M. We conclude that by using large values of M, the transmitter can send short messages using a single-shot transmission with relatively small values of Δ.

VI. Conclusions

We studied communication over DBMT channels assuming that multiple information particles are simultaneously released at each transmission. We first derived the ML detector and argued that it is impractical for nano-scale devices due to its high complexity. Then, we considered the linear detection framework, and showed that when the noise is stable with characteristic exponent smaller then unity, then linear processing increases the noise dispersion, which results in a higher probability of error. To take advantage of the multiple transmitted particles, we then derived the FA detector and showed that for low to medium values of M it achieves a probability of error very close to that of the ML detector. On the other hand, since the first arrival is not a sufficient statistic for the detection problem, when M is large the ML detector strictly outperforms the FA detector. Our analysis indicates that the FA detector has a nice concentration property of the conditional densities, which implies that by using M large enough one can use large constellations, thus, conveying several bits in each transmission. This property is very attractive for molecular nano-scale sensors that are required to send a limited number of bits and then remain silent for a long period of time.

References

[1] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular communication*. Cambridge University Press, 2013.

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

[3] N. Farsad, W. Guo, and A. W. Eckford, "Tabletop molecular communication: Text messages through chemical signals," *PLOS ONE*, vol. 8, no. 12, p. e82935, Dec 2013.

[4] C. Lee, B. Koo, N.-R. Kim, H. B. Yilmaz, N. Farsad, A. Eckford, and C.-B. Chae, "Molecular MIMO communication link," in *IEEE Conference on Computer Communications*, Apr. 2015, pp. 1–2.

[5] N. Farsad, W. Guo, C.-B. Chae, and A. W. Eckford, "Stable distributions as noise models for molecular communication," in *IEEE Global Communications Conference*, 2015.

[6] N. Farsad, Y. Murin, A. Eckford, and A. Goldsmith, "Capacity limits of diffusion-based molecular timing channels," *IEEE Transactions on Information Theory*, Feb. 2016, submitted to, available at http://arxiv.org/abs/1602.07757.

[7] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C.-B. Chae, "Three-dimensional channel characteristics for molecular communications with an absorbing receiver," *IEEE Communinications Letters*, vol. 18, no. 6, pp. 929–932, 2014.

[8] K. V. Srinivas, A. Eckford, and R. Adve, "Molecular communication in fluid media: The additive inverse gaussian noise channel," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4678–4692, Jul. 2012.

[9] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, 1st ed. John Wiley & Sons, 1995.

[10] J. P. Nolan, *Stable Distributions - Models for Heavy Tailed Data*. Boston: Birkhauser, 2015, in progress, Chapter 1 online at academic2.american.edu/∼jpnolan.

[11] V. Zolotarev, *One-dimensional stable distributions*. American Mathematical Soc., 1986, vol. 65.

[12] A. W. Eckford, "Nanoscale communication with brownian motion," in *Proc. of 41st Annual Conference on Information Sciences and Systems*, Baltimore, MD, 2007, pp. 160–165.

[13] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag, 1991.

[14] J. G. Gonzales, J. L. Paredes, and G. R. Arce, "Zero-order statistics: A mathematical framework for the processing and characterization of very impulsive signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3839–3851, Oct. 2007.

[15] Y. Murin, N. Farsad, M. Chowdhury, and A. Goldsmith, "Optimal detection for diffusion-based molecular timing channels," in preparation, available at http://web.stanford.edu/∼moriny.

[16] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, 1st ed. Cambridge University Press, 2005.