

On the Capacity of Diffusion-Based Molecular Timing Channels With Diversity

Nariman Farsad, Yonathan Murin, Milind Rao, and Andrea Goldsmith
Electrical Engineering, Stanford University, USA

Abstract—This work introduces a class of molecular timing (MT) channels, where information is modulated on the release timing of *multiple indistinguishable* information particles, with finite life time, and decoded from the times of arrival at the receiver. The capacity of the MT channel, as well as an upper bound on this capacity, are derived for the case where information particles are released simultaneously by the transmitter. The paper is concluded by outlining two possible lower bounds on the capacity.

I. INTRODUCTION

In molecular communication information is modulated on different properties of small particles (e.g., concentration, the type, the number, or the time of release) that are released by the transmitter [1]. The information particles are transported from the sender to a receiver through different means such as: diffusion, active transport, bacteria, and flow [1]. Several experimental platforms have been developed in recent years that are capable of transmitting short messages [2].

In this work, we consider the molecular timing channel (MT) presented in [3], where information is modulated on the *time of release of the information particles*, and extend the results to the case of *multiple particles*. Communication based on the *time of release* may be used in the brain at the synaptic cleft, where two chemical synapses communicate over a chemical channel [4]. The released information particles randomly propagate from the transmitter to the receiver, which results in a random delay in time until detection at the receiver. A common assumption, which is accurate for many sensors, is that a particle is absorbed and then removed from the environment as part of the detection process [3]. Thus, the random delay until the particle first arrives at the receiver can be represented as an additive noise term.

One may observe some similarities between the timing channel considered in this work and the timing channel considered in [5], which studied the transmission of bits through queues. Yet, the problem formulation and the noise models are fundamentally different. In [5], the queue induces an order on the channel output (i.e. arrival times), namely, the first arrival time corresponds to the first channel use, the second arrival corresponds to the second channel use, and so on. On the other hand, in molecular channels with *indistinguishable* particles, *order may not be preserved*, as was observed in [6].

To account for the lack of ordering, in [3] we considered an MT channel where the transmitter encodes messages over a finite time interval, called the *symbol interval*, by releasing a single particle at a corresponding time in that interval. Furthermore, the released information particles have a finite lifetime, called the *particle's lifetime*, after which they spon-

taneously dissipate. Note that many particles naturally degrade over time and the speed of this process can be controlled through chemical reactions, e.g., through enzymes [7]. Using this scheme, a *single channel use interval* is the sum of the symbol interval and the particle's lifetime, and the channel can then be used *sequentially* without intersymbol interference.

Some of the other previous papers on molecular timing channels focused on the additive inverse Gaussian noise (AIGN) channel, which features a positive drift from the transmitter to the receiver [8]–[11]. In this case, the first time of arrival over a one-dimensional space follows the inverse Gaussian distribution, giving the channel its name. In these papers, upper and lower bounds on the maximal mutual information between the input and output of the AIGN channel, denoted in this work by *capacity per channel use*, were provided for different input and output constraints. However, it is not clear what the associated capacity is in *bits per second*.

In [3] we studied the case where a single particle is released during the symbol interval, and derived tight upper and lower bounds on the capacity of the diffusion-based MT (DBMT) channel. In the current work, we extend the results of [3] to the case where m particles are released during the symbol interval. In particular, we present the channel capacity expression for the case where all m particles are released *simultaneously* by the transmitter, and derive an upper bound on the capacity of this channel. We also present a guideline for deriving two lower bounds, one for the case when the receiver observes a linear combination of arrival times, and one for the case where the receiver uses the time of the first of the m particles that arrive for detection.

The rest of this paper is organized as follows. The system model and the problem formulation are presented in Section II. The capacity of the DBMT channel with multiple particles is studied in Section III: The capacity and an upper bound on capacity are derived. In addition, two different approaches for deriving a lower bound are discussed. Concluding remarks are provided in Section IV.

Notation: We denote the set of real numbers by \mathbb{R} , the set of positive real numbers by \mathbb{R}_+ , and the set of positive natural numbers by \mathbb{N} . Other than these sets, we denote sets with calligraphic letters, e.g., \mathcal{J} , where $|\mathcal{J}|$ denotes the cardinality of the set \mathcal{J} . $[n]$ denotes the set $\{1, \dots, n\}$. We denote random variables (RVs) with upper case letters, e.g., X , Y , their realizations with the corresponding lower case letters, e.g., x , y , and vectors with boldface letters, e.g., \mathbf{X} , \mathbf{Y} . The i^{th} element of a vector \mathbf{X} is denoted by $\mathbf{X}[i]$. We use $f_Y(y)$ to

denote the probability density function (PDF) of a continuous RV Y on \mathbb{R} , $f_{Y|X}(y|x)$ to denote the conditional PDF of Y given X , and $F_Y(y)$ to denote the cumulative distribution function (CDF). We use $h(\cdot)$ to denote the entropy of a continuous RV and $I(\cdot; \cdot)$ to denote the mutual information between two continuous RVs, as defined in [12, Ch. 8]. Finally, $X \leftrightarrow Y \leftrightarrow Z$ is used to denote a Markov chain formed by the RVs X, Y, Z as defined in [12, Ch. 2.8].

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. The Molecular Timing Channel

We consider a molecular communication channel in which information is modulated on the time of release of the information particles. The information particles themselves are assumed to be *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). To develop our model, we make the following assumptions about the system:

- A1)** The transmitter and receiver are perfectly synchronized in time. The transmitter perfectly controls the release time of the particles, while the receiver perfectly measures the arrival times.
- 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 independent and identically distributed (i.i.d.) process.

Note that these assumption have been traditionally considered in all previous works [6], [8]–[11] to make the models tractable. Next, we formally define the channel.

Let $\mathbf{T}_{x,k} \in \mathbb{R}_+^m, k \in [K]$, denote the times of the k^{th} transmissions for the $m \in \mathbb{N}$ indistinguishable particles released into the medium by the transmitter. The transmitted information is encoded in the sequence of times $\mathbf{T}_{x,k}, k \in [K]$, where $\mathbf{T}_{x,k}$ are assumed to be independent of the random propagation time of *each* of the information particles. Let $\mathbf{T}_{y,k}$ be an m -length vector consisting of the times of arrival of each of the information particles, i.e., $\mathbf{T}_{y,k}[i]$ is the arrival time of the i^{th} particle released at time $\mathbf{T}_{x,k}[i]$. Therefore, we have $\mathbf{T}_{y,k}[i] \geq \mathbf{T}_{x,k}[i], i \in [m]$. Thus, we obtain the following vector additive noise channel model:¹

$$\mathbf{T}_{y,k} = \mathbf{T}_{x,k} + \mathbf{T}_{n,k}, \quad (1)$$

where $\mathbf{T}_{n,k}[i], i \in [m]$, is a random noise term representing the propagation time of the i^{th} particle of the k^{th} transmission. Note that assumption **A3**) implies that all the elements of $\mathbf{T}_{n,k}$ are independent.

One of the main challenges of the channel in (1) is that the particles from different channel uses may arrive out of order, which results in channel memory. To resolve this issue, we make two assumptions. First, we assume that at the beginning of each transmission there is a finite time interval called the

¹Here we place no assumption on the particles' lifetimes, namely, they do not dissipate before arriving at the receiver.

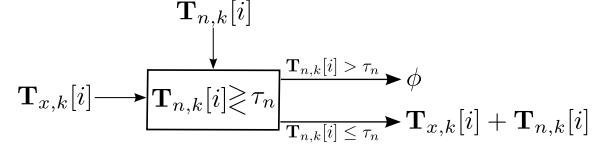


Fig. 1. The MT channel in (2). The channel input is $\mathbf{T}_{x,k}[i]$, while the channel output depends on the condition $\mathbf{T}_{n,k}[i] \geq \tau_n$.

symbol interval over which the transmitter can choose times to release the information particles for that transmission. Second, we assume that information particles have a finite lifetime, i.e., they dissipate immediately after this finite interval, denoted by the *particle's lifetime*. By setting the channel use interval to be a concatenation of the symbol interval and the particle's lifetime, we ensure that order is preserved and obtain a memoryless channel.

Let $\tau_x < \infty$ be the symbol interval, and $\tau_n < \infty$ be the particle's lifetime (i.e. each transmission interval is equal to $\tau_x + \tau_n$). Then our two assumptions can be formally stated as:

- A4)** The release times obey:

$$(k-1) \cdot (\tau_x + \tau_n) \leq \mathbf{T}_{x,k}[i] \leq (k-1) \cdot (\tau_x + \tau_n) + \tau_x,$$

- A5)** The information particles dissipate and are never received if $\mathbf{T}_{n,k}[i] \geq \tau_n, i \in [m]$.

The first assumption can be justified by noting that the transmitter can choose its release interval, while the second assumption can be justified by designing the system such that information particles are degraded in the environment after a finite time (e.g. using chemical reactions) [7], [13]. The resulting channel, which we call the *molecular timing (MT) channel*, is given by:

$$\tilde{\mathbf{Y}}_k[i] = \begin{cases} \mathbf{T}_{y,k}[i] = \mathbf{T}_{x,k}[i] + \mathbf{T}_{n,k}[i], & \mathbf{T}_{n,k}[i] \leq \tau_n \\ \phi, & \mathbf{T}_{n,k}[i] > \tau_n \end{cases}, \quad (2)$$

where ϕ is the *empty symbol* (i.e., a symbol indicating nothing has arrived), $\mathbf{T}_{x,k}$ is the channel input, i.e., the k^{th} release timing vector, $\mathbf{T}_{y,k}[i]$ is the arrival time of the i^{th} information particle at the receiver (if it arrives), and $\tilde{\mathbf{Y}}_k$ is an m -length vector of channel outputs at the k^{th} channel use interval. The i^{th} element of the MT channel (2) is depicted in Fig. 1. We emphasize that the receiver observes a *sorted* version of the channel output $\tilde{\mathbf{Y}}_k$, which we denote by \mathbf{Y}_k . Next, we formally define the capacity of the MT channel with input $\mathbf{T}_{x,k}$ and output \mathbf{Y}_k .

B. Capacity Formulation for the MT Channel

Let $\mathcal{A}_k \triangleq [(k-1) \cdot (\tau_x + \tau_n), (k-1) \cdot (\tau_x + \tau_n) + \tau_x]$ and $\mathcal{B}_k \triangleq \{[(k-1) \cdot (\tau_x + \tau_n), k \cdot (\tau_x + \tau_n)] \cup \phi\}$ for $k \in [K]$. We now define a code for the MT channel (2) as follows:

Definition 1 (Code): A (K, R, τ_x, τ_n) code for the MT channel (2), with code length K and code rate R , consists of a message set $\mathcal{W} = \{1, 2, \dots, 2^{K(\tau_x + \tau_n)R}\}$, an encoder function $\varphi^{(K)} : \mathcal{W} \mapsto \mathcal{A}_1^m \times \mathcal{A}_2^m \times \dots \times \mathcal{A}_K^m$, and a decoder function $\nu^{(K)} : \mathcal{B}_1^m \times \mathcal{B}_2^m \times \dots \times \mathcal{B}_K^m \mapsto \mathcal{W}$.

Remark 1: Observe that since we consider a timing channel, similarly to [5], the codebook size is a function of $\tau_x + \tau_n$, and $K(\tau_x + \tau_n)$ is the maximal time that it takes to transmit

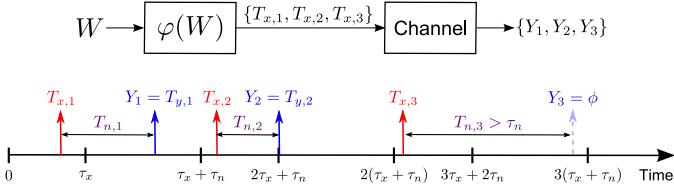


Fig. 2. Illustration of the encoding procedure of Definition 1 for $K = 3$ and $m = 1$. Red pulses correspond to transmission times, while blue pulses correspond to arrival times at the receiver.

a message using a (K, R, τ_x, τ_n) code. Furthermore, note that the above encoder maps the message $W \in \mathcal{W}$ into K m -dimensional vectors of time indices, $\mathbf{T}_{x,k}, k \in [K]$, where $\mathbf{T}_{x,k} \in \mathcal{A}_k^m$, while the decoder decodes the transmitted message using the (sorted) $K \times m$ channel outputs $\mathbf{Y}_k, k \in [K]$, where $\mathbf{Y}_k \in \mathcal{B}_k^m$. We emphasize that this construction prevents intersymbol interference, namely, the m particles transmitted at the interval \mathcal{A}_k either arrive *before* the m particles transmitted at the interval \mathcal{A}_{k+1} or never arrive. Thus, we obtain K identical and independent channels (per channel use interval). However, note that within the channel use interval the arrivals of the m particles are not ordered. Finally, we note that this construction was not used in [5] since, when transmitting bits through queues, the channel itself forces an ordering.

The encoding and transmission through the channel are illustrated in Fig. 2 for the case of $K = 3$ and $m = 1$. The encoder produces three release times $\{T_{x,1}, T_{x,2}, T_{x,3}\}$ which obey $T_{x,k} \in \mathcal{A}_k, k = 1, 2, 3$. In each time index a single particle is released to the channel which adds a random delay according to (2). The channel outputs are denoted by $\{Y_1, Y_2, Y_3\}$. It can be observed that while $Y_1 = T_{y,1} = T_{x,1} + T_{n,1}$ and $Y_2 = T_{y,2} = T_{x,2} + T_{n,2}$, $Y_3 = \phi$ since $T_{n,3} > \tau_n$ and therefore the third particle does not arrive.

Definition 2 (Probability of Error): The average probability of error of a (K, R, τ_x, τ_n) code is defined as:

$$P_e^{(K)} \triangleq \Pr \{ \nu(\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_K) \neq W \},$$

where the message W is selected uniformly from the message set \mathcal{W} .

Definition 3 (Achievable Rate): A rate R is called achievable if for any $\epsilon > 0$ and $\delta > 0$ there exists some blocklength $K_0(\epsilon, \delta)$ such that for every $K > K_0(\epsilon, \delta)$ there exists a $(K, R - \delta, \tau_x, \tau_n)$ code with $P_e^{(K)} < \epsilon$.

Definition 4 (Capacity): The capacity C is the supremum of all achievable rates.

Remark 2: Note that even though we consider a timing channel, we define the capacity in terms of bits per time unit [5, Definition 2]. This is in contrast to the works [8]–[11] which defined the capacity as the maximal number of bits that can be conveyed through the channel *per channel use*.

Note that this definition of capacity C for the MT channels is fairly general and can be applied to different propagation mechanism as long as Assumptions **A1**–**A5** are not violated. Our objective in this paper is to characterize the capacity of the MT channel for diffusion-based propagation.

C. The Diffusion-Based MT Channel

In diffusion-based propagation, the released information particles follow a random Brownian path from the transmitter to the receiver. In this case, to specify the random additive noise term $\mathbf{T}_{n,k}[i]$ in (2), we define a Lévy-distributed RV as follows:

Definition 5 (Lévy Distribution): Let the RV Z be a Lévy-distributed RV with location parameter μ and scale parameter c . Then, its PDF is 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 is given by

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

Throughout the paper, we use the notation $Z \sim \mathcal{L}(\mu, c)$ to indicate a Lévy RV with parameters μ and c .

Let r 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 [14, 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, and therefore the noise in (2) is distributed as $T_{n,k} \sim \mathcal{L}(0, c)$ with $c = \frac{r^2}{2d}$. In this case, we call the diffusion-based MT channel in (2) the *DBMT channel*.

Remark 3: In [15] it is shown that for an infinite, three-dimensional homogeneous medium without flow, and a spherically absorbing receiver, the first arrival time follows a scaled Lévy distribution. Thus, the results presented in this paper can be extended to 3-D space.

III. THE CAPACITY OF THE DBMT CHANNEL WITH DIVERSITY

In [3], we defined the capacity of the MT channel, for $m = 1$, and provided upper and lower bounds on the capacity for this case. In this section, we extend the results of [3] to the case where $m > 1$. To simplify the analysis, we consider the special case wherein every symbol interval has all its particles released *simultaneously*. In this case the lack of intra-symbol ordering has no effect.

A natural question that arises from the work [3] is: *Can the capacity be increased by releasing multiple particles, namely, using $m > 1$?* and, if the answer is positive then how does the capacity scale with m ? In [10, Sec. IV.C] and [16, Sec. IV] it is shown that by releasing multiple particles one can reduce the probability of error; yet, it is not clear if and how the capacity scales with the number of particles which are simultaneously released in each transmission interval \mathcal{A}_k (see Section II-B for detailed definitions). We note that the highest possible scaling is linear, as the capacity in the considered setting is upper bounded by the the setup in which the particles are distinguishable.

$$C_m(\tau_n) = \max_{\tau_x, \mathcal{F}(\tau_x)} \left\{ \frac{1}{\tau_x + \tau_n} \sum_{\mathcal{J}=[\bar{m}]:\bar{m} \in [m]} I(\mathbf{T}_x[\mathcal{J}]; \mathbf{T}_y[\mathcal{J}] | \mathbf{T}_n[\mathcal{J}] \leq \tau_n) \cdot v(p, m, |\mathcal{J}|) \right\}. \quad (6)$$

We begin our analysis by noting that as the particles are released simultaneously, we have $\mathbf{T}_{x,k}[i] = T_{x,k}, i \in [m], k \in [K]$. We further define the set $\mathcal{J}_k \triangleq \{j : \mathbf{T}_{n,k}[j] \leq \tau_n\}, k \in [K]$, which is the set of the indices of all particles which arrive within the interval $[(k-1) \cdot (\tau_x + \tau_n), k \cdot (\tau_x + \tau_n)]$. Clearly, $|\mathcal{J}_k| \leq m$. Note that if there exists $l \in [m]$ such that $l \notin \mathcal{J}_k$, then the output of the channel for the l^{th} particle by (2) is ϕ , and therefore this particle does not convey information over the channel. More precisely, let $\mathbf{Y}_{k,\mathcal{J}_k}$ denote the vector $\mathbf{Y}_k[j], j \in \mathcal{J}_k$, and $\mathbf{Y}_{k,\mathcal{J}_k^c}$ denote the vector $\mathbf{Y}_k[l], l \notin \mathcal{J}_k$. We now write:

$$\begin{aligned} I(T_{x,k}; \mathbf{Y}_k) &= I(T_{x,k}; \mathbf{Y}_{k,\mathcal{J}_k}, \mathbf{Y}_{k,\mathcal{J}_k^c}) \\ &= I(T_{x,k}; \mathbf{Y}_{k,\mathcal{J}_k}). \end{aligned}$$

Since all the particles are statistically indistinct, the term $I(T_{x,k}; \mathbf{Y}_{k,\mathcal{J}_k})$ depends on $|\mathcal{J}_k|$ and not on the specific indices of the set \mathcal{J}_k . In fact, one can re-label the transmitted particles such that the first $|\mathcal{J}_k|$ are the particles that arrive within the interval $[(k-1) \cdot (\tau_x + \tau_n), k \cdot (\tau_x + \tau_n)]$. Therefore, in the following we slightly abuse the notation and let $\mathcal{J}_k = \{1, 2, \dots, |\mathcal{J}_k|\}$. We define $\mathbf{T}_{y,k}[\mathcal{J}_k] \triangleq [\mathbf{T}_{y,k}[1], \mathbf{T}_{y,k}[2], \dots, \mathbf{T}_{y,k}[|\mathcal{J}_k|]]$, while $\mathbf{T}_{n,k}[\mathcal{J}_k]$ is defined in a similar manner. Finally, we define $\mathbf{T}_{x,k}[\mathcal{J}_k]$ to be a vector of length $|\mathcal{J}_k|$ with all its elements equal to the repeated values $T_{x,k}$. With this notation we now define a channel equivalent to (2):

$$\mathbf{Y}_k = \begin{cases} \phi, & |\mathcal{J}_k| = 0 \\ \mathbf{T}_{y,k}[\mathcal{J}_k] = \mathbf{T}_{x,k}[\mathcal{J}_k] + \mathbf{T}_{n,k}[\mathcal{J}_k], & |\mathcal{J}_k| > 0 \end{cases}. \quad (5)$$

Let $C_m(\tau_n)$ denote the capacity of the DBMT channel with diversity in (2), and therefore also the capacity of the channel (5). In addition, let $p \triangleq F_{T_n}(\tau_n)$, and define the function $v(p, m, i) \triangleq \binom{m}{i} p^i (1-p)^{m-i}, i \in [m]$. The following theorem characterizes $C_m(\tau_n)$:

Theorem 1: $C_m(\tau_n)$ is given by (6) at the top of the page, where the condition $\mathbf{T}_n[\mathcal{J}] \leq \tau_n$ reads $\mathbf{T}_n[j] \leq \tau_n, \forall j \in \mathcal{J}, \mathbf{T}_n[l] > \tau_n, \forall l \notin \mathcal{J}$.

Proof: Based on the proof of [3, Theorem 1], one can show that the capacity of the channel (2), and therefore also the channel (5), in bits per second, is given by:

$$C(\tau_n) = \max_{\tau_x, \mathcal{F}(\tau_x)} \frac{I(T_x; \mathbf{Y})}{\tau_x + \tau_n}.$$

Next, we note that since the propagation of the different particles is independent, see assumption **A3**), $|\mathcal{J}|$ follows a binomial distribution, i.e., $|\mathcal{J}| \sim \mathcal{B}(m, F_{T_n}(\tau_n))$.

Furthermore, as $|\mathcal{J}|$ is a function of only the received symbol set \mathbf{Y} , we have the Markov chain $\mathbf{T}_x \leftrightarrow \mathbf{Y} \leftrightarrow |\mathcal{J}|$.

Thus, we write:

$$I(T_x; \mathbf{Y}) = I(\mathbf{T}_x; \mathbf{Y}) \quad (7)$$

$$= I(\mathbf{T}_x; \mathbf{Y}, |\mathcal{J}|) \quad (8)$$

$$= I(\mathbf{T}_x; \mathbf{Y} | |\mathcal{J}|) \quad (9)$$

$$\begin{aligned} &= \sum_{j=0}^M \Pr\{|\mathcal{J}|=j\} \cdot I(\mathbf{T}_x[\mathcal{J}]; \mathbf{T}_y[\mathcal{J}] | |\mathcal{J}|=j) \\ &= \sum_{j=1}^M \Pr\{|\mathcal{J}|=j\} \cdot I(\mathbf{T}_x[\mathcal{J}]; \mathbf{T}_y[\mathcal{J}] | |\mathcal{J}|=j), \end{aligned} \quad (10)$$

where (7) follows from the fact that \mathbf{T}_x is simply a vector which contains T_x multiple times; (8) follows from the Markov chain $\mathbf{T}_x \leftrightarrow \mathbf{Y} \leftrightarrow |\mathcal{J}|$; (9) follows from the fact that \mathbf{T}_x is independent of $|\mathcal{J}|$; and, (10) follows by noting that when $|\mathcal{J}| = 0$ no information goes through the channel.

Finally, we note that the condition $\mathbf{T}_n[\mathcal{J}] \leq \tau_n, |\mathcal{J}| = j$ is equivalent to the condition $|\mathcal{J}| = j$, and since $|\mathcal{J}| \sim \mathcal{B}(m, F_{T_n}(\tau_n))$ then $\Pr\{|\mathcal{J}| = j\} = v(p, m, j)$. ■

Similarly to the single-particle case studied in [3], obtaining an exact expression for the capacity is highly complicated, thus, we present an upper bound and outline two different approaches for obtaining a lower bound. Let X be a continuous RV with PDF $f_X(x)$ and CDF $F_X(x)$, and let τ be a real constant. In [3, Thm. 2] we provide a general expression for $h(X | X \leq \tau)$. For the specific case of a Lévy-distributed RV, $h(X | X \leq \tau)$ can be calculated using the result of [3, Lemma 1].

The upper bound on capacity is now given in the following theorem.

Theorem 2: The capacity of the DBMT channel with diversity is upper bounded by $C_m(\tau_n) \leq C_m^{\text{ub}}(\tau_n)$, where $C_m^{\text{ub}}(\tau_n)$ is given by:

$$C_m^{\text{ub}}(\tau_n) \triangleq m \cdot F_{T_n}(\tau_n) \max_{\tau_x} \frac{\log(\tau_x + \tau_n) - h(T_n | T_n \leq \tau_n)}{\tau_x + \tau_n}. \quad (11)$$

Remark 4: Note that for $m=1$ the upper bound of Thm. 2 specialize to the upper bound presented in [3, Thm. 3].

Proof: First, we note that the conditional mutual information in (6) can be written as:

$$\begin{aligned} &I(\mathbf{T}_x[\mathcal{J}]; \mathbf{T}_y[\mathcal{J}] | \mathbf{T}_n[\mathcal{J}] \leq \tau_n) \\ &= h(\mathbf{T}_y[\mathcal{J}] | \mathbf{T}_n[\mathcal{J}] \leq \tau_n) - h(\mathbf{T}_n[\mathcal{J}] | \mathbf{T}_n[\mathcal{J}] \leq \tau_n). \end{aligned} \quad (12)$$

Next, we explicitly evaluate $h(\mathbf{T}_n[\mathcal{J}] | \mathbf{T}_n[\mathcal{J}] \leq \tau_n)$ and bound $h(\mathbf{T}_y[\mathcal{J}] | \mathbf{T}_n[\mathcal{J}] \leq \tau_n)$. From assumption **A3**) we

have:

$$\begin{aligned} h(\mathbf{T}_n[\mathcal{J}] | \mathbf{T}_n[\mathcal{J}] \leq \tau_n) &= \sum_{j=1}^{|\mathcal{J}|} h(\mathbf{T}_n[j] | \mathbf{T}_n[j] \leq \tau_n) \\ &= |\mathcal{J}| \cdot h(T_n | T_n \leq \tau_n). \end{aligned} \quad (13)$$

We now write:

$$h(\mathbf{T}_y[\mathcal{J}] | \mathbf{T}_n[\mathcal{J}] < \tau_n) \leq |\mathcal{J}| \log(\tau_x + \tau_n), \quad (14)$$

where (14) is due to the fact that the uniform distribution maximizes entropy over a finite interval. Therefore, (6) can be upper bounded by:

$$\sum_{j=1}^m (\log(\tau_x + \tau_n) - h(T_n | T_n \leq \tau_n)) \cdot j \cdot v(p, m, j). \quad (15)$$

Finally, using the expression for the mean of a Binomial RV [17, Ch. 16.2.3.1], we write:

$$\sum_{j=1}^m j \cdot v(p, m, j) = \sum_{j=1}^m j \cdot \binom{m}{j} p^j (1-p)^{m-j} = mp. \quad (16)$$

Combining (16) with (15) and recalling that $p = F_{T_n}(\tau_n)$ we obtain the upper bound in (11). \blacksquare

The channels (2) and (5) have a single input $T_{x,k}$ and multiple outputs \mathbf{Y}_k . Thus, by simultaneously releasing $m > 1$ particles we achieve *receive diversity*. As the propagation of all particles is independent and identically distributed, the channel (2) can also be viewed as a single-input-multiple-output (SIMO) channel in which all the channel outputs experience an independent and identical propagation law. While the upper bound in (11) scales linearly with m , we currently do not have a lower with the same scaling. Thus, a linear scaling cannot be concluded. In the next paragraphs we outline two plausible approaches for deriving lower bounds. The exact analysis is left for future work.

We plan to use two different approaches for deriving lower bounds on the capacity of the channel in (5). The first approach is to analyze a receiver that uses a linear combination of the arrival times to detect the channel input. This case is equivalent to an additive noise channel, where the channel input is $X = |\mathcal{J}| \cdot T_x$ and the additive noise is $N = \sum_{i \in \mathcal{J}} \mathbf{T}_n[i]$. As $m \rightarrow \infty$, N becomes Gaussian due to central limit theorem and a lower bound can be formulated.

Another approach is to assume that the receiver detects based on the first arrival time between all $|\mathcal{J}|$ arrivals. In this case the channel output is $\min_{i \in \mathcal{J}} \mathbf{T}_y[i]$. In [16], we showed that the performance of such a detector is very close to the performance of the maximum likelihood detector. In fact, it was shown that when the particle lifetime is infinite, the linear detector would actually degrade the performance compared to choosing any single arrival times at random. This is due to the heavy tails associated with the Lévy distribution. Although with a finite particle life time it is not clear if the first arrival detector would still outperform the linear one, we believe that such a lower bound can be useful. The main challenge in this approach is deriving an expression for the noise term which is given by $N = \min_{i \in \mathcal{J}} \mathbf{T}_n[i]$.

IV. CONCLUSIONS

In this work we considered MT channels, where the information is modulated on the release time of multiple indistinguishable particles. We presented the capacity expression for the case when the particles are all released simultaneously, and derived an upper bound on the capacity of this channel. We showed that the upper bound increases linearly with the number of released particles for the DBMT channel. This is analogous to receive diversity as each particle propagates to the receiver independently. We also outlined two different approaches to finding a lower bound on the capacity using linear and first arrival detector. As part of future work, we intend to derive these lower bounds and investigate which detector performs better for channels with finite particle lifetime. We also plan to find tighter upper bounds on the capacity and explore how capacity scales with the number of particles.

REFERENCES

- [1] 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, thirdquarter 2016.
- [2] 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.
- [3] N. Farsad, Y. Murin, A. Eckford, and A. Goldsmith, "On the capacity of diffusion-based molecular timing channels," *IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2016, to be presented. [Online]. Available: <http://arxiv.org/abs/1604.06409>
- [4] A. Borst and F. E. Theunissen, "Information theory and neural coding," *Nature Neuroscience*, vol. 2, no. 11, pp. 947–957, Nov. 1999.
- [5] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 4–18, Jan 1996.
- [6] C. Rose and I. S. Mian, "Signaling with identical tokens: Upper bounds with energy constraints," in *IEEE International Symposium on Information Theory*, 2014, pp. 1817–1821.
- [7] W. Guo, T. Asyharti, N. Farsad, H. B. Yilmaz, B. Li, A. Eckford, and C.-B. Chae, "Molecular communications: channel model and physical layer techniques," *IEEE Wireless Communications*, 2015, to appear.
- [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, 2012.
- [9] H.-T. Chang and S. M. Moser, "Bounds on the capacity of the additive inverse gaussian noise channel," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*. IEEE, 2012, pp. 299–303.
- [10] A. W. Eckford, K. Srinivas, and R. S. Adve, "The peak constrained additive inverse gaussian noise channel," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*. IEEE, 2012, pp. 2973–2977.
- [11] H. Li, S. Moser, and D. Guo, "Capacity of the memoryless additive inverse gaussian noise channel," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 12, pp. 2315–2329, Dec 2014.
- [12] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*, 2nd ed. Wiley-Interscience, 2006.
- [13] A. Noel, K. Cheung, and R. Schober, "Improving receiver performance of diffusive molecular communication with enzymes," *IEEE Transactions on NanoBioscience*, vol. 13, no. 1, pp. 31–43, March 2014.
- [14] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag, 1991.
- [15] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C.-B. Chae, "Three-dimensional channel characteristics for molecular communications with an absorbing receiver," *IEEE Communications Letters*, vol. 18, no. 6, pp. 929–932, 2014.
- [16] Y. Murin, N. Farsad, M. Chowdhury, and A. Goldsmith, "Optimal detection for diffusion-based molecular timing channels," 2016, in preparation, available at <http://web.stanford.edu/~morin/>.
- [17] G. I.N. Bronshtein, K.A. Semendyayev and H. Muehlig, *Handbook of Mathematics*, 5th ed. Springer-Verlag, 2007.