Deterministic Identification for MC ISI-Poisson Channel

Mohammad J. Salariseddigh∗, Vahid Jamali†, Uzi Perez‡, Holger Boche§, Christian Deppe¶, and Robert Schober∥

∗ Technical University of Munich (TUM) † Technical University of Darmstadt ‡ Technion; Helen Diller Quantum Center
§ BMBF Research Hub 6G-Life, TUM; Munich Center for Quantum Science and Technology (MCQST); Munich Quantum Valley (MQV)
¶ BMBF Research Hub 6G-Life, TUM ∥ Friedrich-Alexander-University Erlangen-Nürnberg

Abstract—Several applications of molecular communications (MC) feature an alarm-prompt behavior for which the prevalent Shannon capacity may not be the appropriate performance metric. The identification capacity as an alternative measure for such systems has been motivated and established in the literature. In this paper, we study deterministic identification (DI) for the discrete-time Poisson channel (DTPC) with intersymbol interference (ISI) where the transmitter is restricted to an average and a peak molecule release rate constraint. Such a channel serves as a model for diffusive MC systems featuring long channel impulse responses and employing molecule counting receivers. We derive lower and upper bounds on the DI capacity of the DTPC with ISI when the number of ISI channel taps K may grow with the codeword length n (e.g., due to increasing symbol rate). As a key finding, we establish that for deterministic encoding, the codebook size scales as 2(log n R) assuming that the number of ISI channel taps scales as K = 2 log n R, where R is the coding rate and κ is the ISI rate. Moreover, we show that optimizing κ leads to an effective identification rate [bits/s] that scales linearly with n, which is in contrast to the typical transmission rate [bits/s] that is independent of n.

I. INTRODUCTION

Molecular communications (MC) is a promising paradigm for communication between nanomachines or different biological entities [1] and realizes the exchange of information via transmission, propagation, and reception of signaling molecules [2]. In the past decade, different aspects of synthetic MC have been considered in the literature, including channel modeling [3], modulation and detection design [4], biological building blocks for transceiver design [5], and information-theoretical and relevant mathematical foundations [6], [7].

One of the widely-accepted models for MC systems with molecule counting receivers is the discrete-time Poisson channel (DTPC) model with intersymbol interference (ISI) [6], [8]. The DTPC model with memory has been used to study the performance limits of MC systems. Analytical lower and upper bounds on Shannon’s transmission capacity of the DTPC with input constraints and memory are provided in [9]. Bounds on the transmission capacity of the DTPC with memory are developed in [10], [11]. Code design for the DTPC with memory under peak and average power constraints is studied in [12]. In [13], the impact of memory on the MC performance for a diffusion-based channel is characterized.

Numerous applications of MC within the scope of future generation wireless networks (XG) [14] require event-triggered communication systems. In such systems, Shannon’s message transmission capacity, as studied in [6], may not be the appropriate performance metric. In particular, in event-detection scenarios, where the receiver aims to determine the occurrence of a specific event in terms of a reliable Yes/No decision, the so-called identification capacity is the key performance measure [15]. Examples of the identification problem within the MC context include targeted drug delivery [1], cancer treatment [16], olfactory-inspired synthetic MC systems [17], and bionic nose [18].

The original coding scheme for the identification problem introduced by Ahlsweide and Dueck [15] employs a randomized encoder, where the codewords are selected from distributions. The codebook size for randomized identification (RI) grows double-exponentially in the codeword length n, i.e., ∼22nR [15], where R is the coding rate. Realization of RI codes entails extra complexity, which can be challenging in the context of MC [19]. In contrast, in the deterministic encoding setting of identification, also referred to as deterministic identification (DI) [20], the codewords are selected via a deterministic function from the messages. In our recent works [20]–[23], we target DI for channels with power constraint, including discrete memoryless channels (DMCs), Gaussian channels with fast and slow fading, and the memoryless discrete-time Poisson channel (DTPC), respectively. The codebook size for DMCs, grows exponentially in the codeword length [15], [22], where as for the Gaussian channel [21] and the DTPC [23] it scales as ∼2(2nR)H. However, these results are not applicable to MC systems suffering from ISI. In fact, to the best of the authors’ knowledge, the fundamental performance limits of DI for the DTPC with ISI has not been studied in the literature, yet.

In this paper, we consider MC systems employing molecule counting receivers with a large number of released molecules at the transmitter, see [3, Sec. IV]. Further, we assume that the received signal experiences ISI and follows the Poisson distribution. We formulate the problem of DI over the DTPC with memory under average and peak molecule release rate constraints to account for the limited molecule production/release rates of the transmitter. As our main objective, we investigate the fundamental performance limits of DI over the DTPC with ISI. In particular, this paper makes the following contributions:

• Generalized ISI Model: In MC systems, often the number of channel taps K can be large, particularly for non-degrading signalling molecules in bounded environments, which leads to a long channel impulse response (CIR). In addition, the value of K increases not only with the dispersiveness of the channel but also with the symbol rate. Therefore, it is of interest to investigate the asymptotic limits of the system for large symbol rates (leading to large K) and large codeword lengths n. To do so, we consider a generalized ISI model that captures the ISI-free channel (i.e., K = 1), ISI channels with constant K > 1, and ISI channels for which K increases with the codeword length n (e.g., due to increasing symbol rate). To the best of the authors’ knowledge, such a generalized ISI model has not
been studied in the literature, yet.

- **Codebook Scale**: We establish that the codebook size of the DTPC with ISI for deterministic encoding scales in $n$ similar to the memoryless DTPC [19], namely super-exponentially in the codeword length ($\sim 2^{(n \log n)R}$), even when the number of ISI taps scale as $K = 2^{n \log n}$ for any $\kappa \in [0, 1)$, which we refer to as the ISI rate. This observation suggests that memory does not change the scale of the codebook derived for memoryless DTPC [19] and Gaussian channels [21].

- **Capacity Bounds**: We derive DI capacity bounds for the DTPC with constant $K \geq 1$ and growing ISI $K = 2^{n \log n}$, respectively. We show that for constant $K$, the proposed lower and upper bounds on $R$ are independent of $K$, whereas for growing ISI, they are functions of the ISI rate $\kappa$. Moreover, we show that optimizing $\kappa$ leads to an effective identification rate [bits/s] that scales linearly with $n$, which is in contrast to the typical transmission rate [bits/s] that is independent of $n$.

- **Technical Novelty in the Capacity Proof**: To obtain the proposed lower bound, the existence of an appropriate sphere packing within the input space, for which the distance between the centers of the spheres does not fall below a certain value, is guaranteed. This packing incorporates the effect of ISI as a function of $\kappa$. In particular, we consider the packing of hyper spheres inside a larger hyper cube, whose radius grows in both the codeword length $n$ and the ISI rate $\kappa$, i.e., $\sim n^{\frac{1+\kappa}{2}}$. For derivation of the upper bound, we assume that for given sequences of codes with vanishing error probabilities, a certain minimum distance between the codewords is asserted, where this distance depends on the ISI rate and decreases as $K$ grows.

**Notations**: Calligraphic letters $X, Y, Z, \ldots$ are used for finite sets. Lower case letters $x, y, z, \ldots$ stand for constants and values of random variables, and upper case letters $X, Y, Z, \ldots$ stand for random variables. Lower case bold symbol $x$ and $y$ stand for row vectors. Bold symbol $1_n$ indicates the all-one row vector of size $n$. All logarithms and information quantities are base 2. The set of consecutive natural numbers from 1 through $M$ is denoted by $[M]$. The set of whole numbers is denoted by $\mathbb{N}_0 \triangleq \{0, 1, 2, \ldots\}$. The set of non-negative real numbers is denoted by $\mathbb{R}_+$. The gamma function for non-negative integer $x$ is denoted by $\Gamma(x)$ and is defined as $\Gamma(x) = (x-1)!$, where $(x-1)! \triangleq (x-1) \times (x-2) \times \cdots \times 1$. We use the notation, $f(n) = o(g(n))$, to indicate that $f(n)$ is dominated by $g(n)$ asymptotically. The $\ell_2$-norm and $\ell_\infty$-norm of vector $x$ are denoted by $||x||$ and $||x||_\infty$, respectively. Furthermore, we denote the $n$-dimensional hyper sphere of radius $r$ centered at $x_0$ with respect to the $\ell_2$-norm by $S_{x_0}(n, r) = \{x \in \mathbb{R}^n : ||x-x_0|| \leq r\}$. An $n$-dimensional cube with center $(\frac{A}{2}, \ldots, \frac{A}{2})$ and a corner at the origin, i.e., $0 = (0, \ldots, 0)$, whose edges have length $A$ is denoted by $Q_0(n, A) = \{x \in \mathbb{R}^n : 0 \leq x_i \leq A, \forall i \in [n]\}$.

### II. SYSTEM MODEL AND PRELIMINARIES

In this section, we present the system model and coding.

#### A. System Model

We consider the Poisson channel $P$ which arises as a channel model in the context of MC for molecular counting receivers [6]. Let $X \in \mathbb{R}_{\geq 0}$ and $Y \in \mathbb{N}_0$ denote random variables (RVs) modeling the rate of molecule release by the transmitter and the number of molecules observed at the receiver, respectively. We consider a stochastic release model, where for the $t$-th channel use, the transmitter releases molecules with rate $x_t$ (molecules/second) over a time slot of $T_s$ seconds into the channel [6]. These molecules propagate through the channel via diffusion and/or advection, and may even be degraded in the channel via enzymatic reactions [3]. The channel memory is modelled by a length $K$ sequence of probability values, i.e., $p = [p_0, p_1, \ldots, p_{K-1}]$. Let $p_k \defeq p_k T_s$ where the value $p_k \in (0, 1]$ denotes the probability that a given molecule released by the transmitter at the beginning time slot $t$, is observed at the receiver during time slot $t + k$. The relation of channel output $Y$ and input $X$ is given by

$$Y_t = \text{Poiss}(X^p_t + \lambda),$$  

where $X^p_t \defeq \sum_{k=0}^{K-1} p_k X_{t-k}$ is the mean number of observed molecules due to the release at the transmitter and the constant $\lambda \in \mathbb{R}_{>0}$ is the mean number of observed interfering molecules. Let $x^p_t \defeq (x_{t-K+1}, \ldots, x_t)$ be the vector of the $K$ most recently released symbols. The letter-wise channel law reads $V(y_t|x^p_t^t) = e^{-\sum_{t=1}^T \lambda^t} (\frac{\lambda^t}{y_t!})$. We assume that different channel uses given any $K$ previous input symbols are statistically independent, which is a valid assumption for, e.g., fully absorbing receivers [3]. Therefore, for $n$ channel uses, the transition probability law is given by

$$V^{\bar{n}}(y|x) = \prod_{t=1}^{\bar{n}} V(y_t|x^p_t) = \prod_{t=1}^{\bar{n}} e^{-\sum_{t=1}^{T} \lambda^t} (\frac{\lambda^t}{y_t!}).$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ denote the transmitted codeword and the received signal, respectively, with $\bar{n} = n + K - 1$. We assume that $x_t = 0$ when $t > n$ or at $0 < t < T$. The peak and average molecule release rate constraints on the codewords are

$$0 \leq x_t \leq P_{\text{max}} \quad \text{and} \quad \frac{1}{n} \sum_{t=1}^{n} x_t \leq P_{\text{avg}},$$

respectively, $\forall t \in [n]$, where $P_{\text{max}} > 0$ and $P_{\text{avg}} > 0$ constrain the rate of molecule release per channel use and over the entire $n$ channel uses in each codeword, respectively.

#### B. DI Coding for the DTPC

The definition of a DI code for the DTPC $P$ is given below.

**Definition 1** (ISI-Poisson DI Code). An $(n, M(n, R), K(n, \kappa), e_{1, e_2})$ DI code for a DTPC $P$ under average and peak molecule release rate constraints of $P_{\text{avg}}$ and $P_{\text{max}}$, respectively, and for integers $M(n, R)$ and $K(n, \kappa)$, respectively, where $n$ and $R$ are the codeword length and coding rate, respectively, is defined as a system $(C, \mathcal{F})$, which consists of a codebook $C = \{c_{i,t} \in M \subseteq \mathbb{R}^n_+ \}$ such that $0 \leq c_{i,t} \leq P_{\text{max}}$ and $\frac{1}{n} \sum_{t=1}^{n} c_{i,t} \leq P_{\text{avg}}, \forall i \in [M], \forall t \in [n]$.
and a collection of decoding regions $\mathcal{T} = \{T_i\}_{i \in [M]}$ with $\bigcup_{i=1}^{M} T_i \subset \mathbb{N}_0^n$. Given a message $i \in [M]$, the encoder transmits $c_i$, and the decoder’s aim is to answer the following question: Was a desired message $j$ sent or not? There are two types of errors that may occur: Rejection of the true message (type I) or acceptance of a false message (type II). The corresponding error probabilities of the DI code $(C, \mathcal{T})$ are given by

$$P_{e,1}(i) = 1 - \sum_{y \in T_i} V^n(y | c_i)$$

and satisfy the following bounds

$$P_{e,2}(i, j) = \sum_{y \in T_j} V^n(y | c_i),$$

and the following bounds $P_{e,1}(i) \leq e_1$ and $P_{e,2}(i, j) \leq e_2$, for all $i, j \in [M]$ and every $e_1, e_2 > 0$. A rate $R > 0$ is called achievable if for every $e_1, e_2 > 0$ and sufficiently large $n$, there exists an $(n, M(n, R), K(n, \kappa), e_1, e_2)$ DI code. The DI capacity of the DTDP $\mathcal{P}$ is defined as the supremum of all achievable rates, and is denoted by $C_{DI}(\mathcal{P}, M, K)$. \hfill $\Box$

### III. DI CAPACITY OF THE DTDP WITH MEMORY

Here, we present our main results, i.e., lower and upper bounds on the achievable DI rates for the DTDP $\mathcal{P}$ with ISI. The number of ISI taps can be obtained as $K = \lceil n \kappa \rceil$, where $T_{cir}$ and $T_s$ are the CIR length and the symbol duration, respectively. We note that the value of $T_{cir}$ (and hence $K$) can become quite large in bounded MC environments. The results provided in this section are valid for both constant $K$ as well as $K$ increasing with codeword length $n$ (due to decreasing $T_s$). We note that, in practice, the transmitter and receiver may not be able to support arbitrary small symbol duration $T_s$. For the results reported in this paper, we implicitly assume that the adopted $T_s$ can be supported by the considered MC system.

#### A. Main Results

The DI capacity theorem for DTDP $\mathcal{P}$ is stated below.

**Theorem 1.** Consider the DTDP with ISI $\mathcal{P}$ and assume that the number of ISI channel taps scales sub-linearly with codeword length $n$, i.e., $K(n, \kappa) = 2^n \kappa \log n$, where $\kappa \in (0, 1)$. Then the DI capacity of $\mathcal{P}$ subject to average and peak molecule release rate constraints of the form $n^{-1} \sum_{t=1}^{n} c_{i,t} \leq P_{ave}$ and $0 \leq c_{i,t} \leq P_{max}$, respectively, and a codebook of super-exponential scale, i.e., \( M(n, R) = 2^n (\log n)^R \), is bounded by

$$1 - \frac{\kappa}{4} \leq C_{DI}(\mathcal{P}, M, K) \leq \frac{3}{2} + \kappa. \quad (4)$$

**Proof:** The proof of Theorem 1 consists of two parts, namely the achievability and the converse proofs, which are provided in Sections III-B and III-C, respectively.

**Corollary 1.** (Effective Identification Rate) Assuming $T_s = T_{cir}/K = T_{cir}2^{-\kappa \log n}$, $\kappa \in (0, 1)$, the effective identification rate, defined as

$$R_{eff} \overset{\text{def}}{=} \frac{\log M(n, R)}{n T_s} \quad (5)$$

(in bits/symbol), under average and peak molecule release rate constraints is bounded by

$$\frac{(1-\kappa) n^{\kappa} \log n}{4 T_{cir}} \leq R_{eff} \leq \frac{(3+2\kappa) n^{\kappa} \log n}{2 T_{cir}}. \quad (6)$$

**Proof:** The proof follows directly by substituting the capacity results in Theorem 1 into the definition of the effective rate and making further mathematical simplifications. \hfill $\Box$

**Remark 1.** The result in Theorem 1 comprises the following three special cases in terms of $K$:

1) **$K = 1$:** This case accounts for an ISI-free setup ($\kappa = 0$), which is valid when the symbol duration is large ($T_s \geq T_{cir}$), and implies $K = 1$ and $\kappa = 0$. Thereby, $R_{eff}$ scales logarithmically with the codeword length $n$. This is in contrast to the transmission setting in which $R_{eff}$ is independent of $n$ (e.g., the well-known Shannon formula for the Gaussian channel). This result is known in the identification literature [15], [19].

2) **Constant $K > 1$:** When $T_s$ is constant and $T_{cir} < T_{cir}$, we have constant $K > 1$ which implies $\kappa \rightarrow 0$ as $n \rightarrow \infty$. Surprisingly, our capacity result in Theorem 1 reveals that the bounds for the DTDP with memory are in fact identical to those for the memoryless DTDP given in [19].

3) **Growing $K$:** Our capacity results reveal that reliable identification is possible even when $K$ scales with the codeword length as $\sim 2^{\kappa \log n}$. Moreover, the impact of ISI $\kappa$ is reflected in the capacity lower and upper bounds in (4), where the bounds respectively decrease and increase in $n$. While the upper bound on $R_{eff}$ increases in $\kappa$, too, the lower bound in (6) suggests a trade-off in terms of $\kappa$, which is investigated in the Corollary 2.

**Corollary 2 (Optimum ISI Rate).** The lower bound given in Corollary 1 is maximized for the following ISI rate $\kappa_{max}(n), n \in \mathbb{N}$, with

$$\kappa_{max}(n) = 1 - \frac{1}{\ln n}. \quad (7)$$

The above $\kappa_{max}$ gives the following lower bound on $R_{eff}(n)$:

$$R_{eff}(n) \geq \frac{\log e}{4 e T_{cir}} n. \quad (8)$$

Thereby,

$$\lim \inf_{n \rightarrow \infty} \frac{R_{eff}(n)}{n} \geq \frac{\log e}{4 e T_{cir}}. \quad (9)$$

**Proof:** The proof follows from differentiating the lower bound in Corollary 1 with respect to $\kappa$ and equating it to zero. \hfill $\Box$

The effective identification rate $R_{eff}$ [bits/s] in (5) consists of two terms, namely the identification rate per symbol $\log M(n, R)$ [bits/symbol] (which decreases with $\kappa$ for the lower bound in (4)) and the symbol rate $\frac{T_s}{n T_{cir}}$ [symbol/s] (which increases with $\kappa$). The above corollary reveals that in order to maximize $R_{eff}$, it is optimal to set the trade-off for $\kappa$ such that the identification rate (i.e., $\log M(n, R)$) $= \frac{\log e}{4 e T_{cir}}$ becomes independent of $n$ but the symbol rate (i.e., $\frac{T_s}{n T_{cir}}$) linearly scales with $n$. As a result, in contrast to the typical transmission settings where the effective rate is independent of $n$, here, the effective identification rate $R_{eff}$ for the optimal $\kappa$ linearly grows in $n$. 6110
B. Achievability

The achievability proof consists of the following two steps.

Step 1: First, we propose a codebook construction and derive an analytical lower bound on the corresponding codebook size using inequalities for sphere packing density.

Step 2: Then, to prove that this codebook leads to an achievable rate, we propose a decoder and show that the corresponding type I and type II error rates tend to zero.

**Codebook construction:** Let $\mathcal{A} = \min\{P_{ave}, P_{max}\}$. In the sequel, we restrict ourselves to codewords that meet the condition $0 \leq x_i \leq A, \forall t \in [n]$. Similar to our previous explanations for DTPC [19, Sec. III], this condition ensures both the average and the peak molecule release rate constraints in (3). Hence, in the following, we restrict our considerations to a hyper cube with edge length $A$.

We use a packing arrangement of non-overlapping hyper spheres of radius $r_0 = \sqrt{n}\theta_n$ in a hyper cube with edge length $A$, where $\theta_n = \alpha\sqrt{K/n^{1/2}}$, and $\alpha > 0$ is a non-vanishing fixed constant and $0 < b < 1$ is an arbitrarily small constant.

Let $\mathcal{S}$ denote a sphere packing, i.e., an arrangement of $M$ non-overlapping spheres $S_c(n,r_0), i \in [M]$, that are packed inside the larger cube $Q_0(n,A)$ with edge length $A$. We require that the centers of the spheres are inside $Q_0(n,A)$ and are disjoint from each other and have a non-empty intersection with $Q_0(n,A)$. The packing density $\Delta_n(\mathcal{S})$ is defined as the ratio of the saturated packing volume to the cube volume $\text{Vol}(Q_0(n,A))$, i.e.,

$$\Delta_n(\mathcal{S}) \triangleq \frac{\text{Vol}(\bigcup_{i=1}^{M} S_c(n,r_0))}{\text{Vol}(Q_0(n,A))}.$$  

Sphere packing $\mathcal{S}$ is called saturated if no spheres can be added to the arrangement without overlap. In particular, we use a packing argument that has a similar flavor as that observed in the Minkowski–Hlawka theorem for saturated packings [24]. Specifically, consider a saturated packing arrangement of $\bigcup_{i=1}^{M} S_c(n,\sqrt{n}\theta_n)$ spheres with radius $r_0 = \sqrt{n}\theta_n$ embedded within cube $Q_0(n,A)$. Then, for such an arrangement, we have the following lower [25, Lem. 2.1] and upper bounds [24, Eq. (45)] on the packing density $2^{-n} \leq \Delta_n(\mathcal{S}) \leq 2^{-0.599n}$.

In general, the volume of a hyper sphere of radius $r$ is given by [24, Eq. (16)] $\text{Vol}(S_c(n,r)) = \pi^{n/2} \cdot r^n / T(\frac{n}{2}+1)$. We assign a codeword to the center $c_i$ of each small sphere. The codewords satisfy the input constraint as $0 \leq c_i,t \leq A, \forall t \in [n], i \in [M]$, which is equivalent to $\|c_i\|_\infty \leq A$.

Since the volume of a generic center with sphere $c_i$ is equal to $\text{Vol}(S_c(n,r_0))$ and the centers of all spheres lie inside the cube, the total number of spheres is bounded from below by

$$M = \frac{\text{Vol}(\bigcup_{i=1}^{M} S_c(n,r_0))}{\text{Vol}(S_c(n,r_0))} \geq 2^{-n} \cdot \frac{A^n}{\text{Vol}(S_c(n,r_0))},$$  

where the inequality holds by $\Delta_n(\mathcal{S}) \geq 2^{-n}$.

Now (11) can be further simplified as follows log $M \geq n \cdot \log A \cdot n \cdot \log n + \frac{1}{4} n \cdot \log n + n \cdot \log e + o(n)$. Now, for $\log r_0 = \sqrt{n}\theta_n$, we obtain $\log M \geq n \cdot \log \sqrt{n} \cdot \frac{1}{4}(1+b) \cdot n \cdot \log n + \frac{1}{2} n \cdot \log n + n \cdot \log e + o(n) = \frac{1}{4}(1+b) \cdot n \cdot \log n + n \cdot (\log \frac{A}{\sqrt{n}A} + o(n))$, where the dominant term is of order $n \cdot \log n$. Hence, for obtaining a finite value for the lower bound of the rate, $R$, the scaling law of $M$ is induced to be $2^{(n \cdot \log n)R}$. Therefore, we have $R \geq \frac{1}{n \cdot \log n} \cdot \log n \cdot n \log n + n \cdot \log (\frac{A}{\sqrt{n}A} + o(n))$, which tends to $\frac{1}{4}$ when $n \rightarrow \infty$ and $b \rightarrow 0$.

**Encoding:** Given message $i \in [M]$, transmit $x = c_i$.

**Decoding:** Let $\tau_n = c_i^0 \rho_a = a c_i^0 \rho_a \cdot (\log (n+\beta))^{-1}$, where $0 < b < 1$ is an arbitrarily small constant and $0 < c < 2$ is a constant. Before we proceed, for the sake of brevity of analysis, we introduce the following conventions: 1) Let $Y(i) \sim \text{Pois}(e_i^0 + \lambda)$ denote the channel output at time $t$ given that $x = c_i$. Note that $e_i^0 = \sum_{k=0}^{K-1} \rho_k c_i,t-k$ is only one symbol but is a linear combination of all previous $K$ symbols weighted by coefficients $\rho_k$. 2) Let $Y(i) \sim \text{Pois}(e_i^0 + \lambda)$ denote $Y(i) = (Y_i(i), ..., Y_n(i))$ when $e_i^0 = (e_i^0, ..., e_i^n)$.

Let $I^* \triangleq \lambda + \sum_{k=1}^{K-1} \rho_k c_i,t-k$. 4) Let $y_i(i) \sim \text{Pois}(Y_i(i) - \rho_0 c_i,t + \lambda)$. To determine whether message $j \in [M]$ was sent, the decoder checks whether channel output $y$ belongs to the decoding set or not. We assume the following decoder:

$$T_j = \{ y \in \mathcal{Y} : |T(y;c_j)| \leq \tau_n \},$$  

where

$$T(y;c_j) = \frac{1}{n} \sum_{i=1}^{n} y_i(j) - (\rho_0 c_i,t + I^*_t) - (I^*_t - \lambda)^2$$

(13) is referred to as the decoding metric evaluated for observation vector $y$ and codeword $c_j$.

**Error Analysis:** Fix $e_1,e_2 > 0$ and let $\zeta_0, \zeta_1 > 0$ be arbitrarily small constants. Consider the type I errors, i.e., the transmitter sends $c_i$, yet $y \notin T_i$. For every $i \in [M]$, the type I error probability is bounded by

$$P_{e,1}(i) = \Pr[(T(Y(i);c_i)) > \tau_n],$$  

(14) where the condition means that $x = c_i$ was sent. In order to bound $P_{e,1}(i)$, we apply Chebyshev’s inequality, namely

$$\Pr \left( |T(Y(i);c_i) - \mathbb{E}[T(Y(i);c_i)]| > \tau_n \right) \leq \frac{\text{Var}[T(Y(i);c_i)]}{\tau_n^2}.$$  

Employing standard techniques, we can show that the expectation of the decoding metric is zero, see [26, Subsec. III-B] for detailed derivation.

Second, we proceed to compute the variance of the decoding metric. Let us define $\psi_{\text{Var}}^{\text{def}} = \sum_{i=1}^{m} \text{Var}[Y_i(i)]^2$. Observe that we have $\text{Var}[T(Y(i);c_i)] = \psi_{\text{Var}}^{\text{def}}$. Since, conditioned on $c_i$, the channel outputs conditioned on the $K$ most recent input symbols are independent. Now, based on [26, App. A] we can provide upper bounds $\psi_{\text{UB}}$ for $\psi_{\text{Var}}$. Hence, we can bound the type I error probability in (14) as follows

$$P_{e,1}(i) = \Pr \left( |T(Y(i);c_i)) > \tau_n \right) \leq \frac{\psi_{\text{UB}}}{n^2 \tau_n^2}$$

$$\leq \frac{6(A + \lambda)^3(1+e^2(1+(A+\lambda)+(A+\lambda)^2+(A+\lambda)^3))}{\tau_n^2} \leq e_1,$$

hence, $P_{e,1}(i) \leq e_1$ holds for sufficiently large $n$ and arbitrarily small $e_1 > 0$. Next, we address type II errors, i.e.,
when \( Y \in T_j \) while the transmitter sent \( c_i \). Then, for every \( i, j \in [M] \), the type II error probability is given by
\[
P_e(i,j) = \Pr \left( T(Y^i); c_j \right) \leq \tau_n,
\]
where
\[
\Pr \left( T(Y^i); c_j \right) = \beta - \alpha \text{ with }
\]
\[
\beta = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( T_i(i) + \rho_0 \left( c_i - c_j \right)^2 \right)
\]
and
\[
\alpha = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( \left( \rho_0 c_i + I_{c_i}^2 \right) + \left( I_{c_i}^2 - \lambda \right) \right)^2.
\]
Observe that term \( \beta \) can be expressed by \( \beta = \beta_1 + \beta_2 \) where
\[
\beta_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( T_i(i) + \rho_0 \left( c_i - c_j \right)^2 \right)
\]
and
\[
\beta_2 = \frac{2 \rho_0}{\sqrt{n}} \sum_{i=1}^{n} \left( c_i - c_j \right)^2 \left( \rho_0 \left( I_{c_i} \right) + \lambda \right).
\]
Then, define the following events \( \mathcal{H}_1 = \{ \beta_1 \leq \tau_n \} \), \( \mathcal{E}_n = \{ \beta_2 > \tau_n \} \) and \( \mathcal{E}_1 = \{ \beta_1 \leq \lambda \} \). Exploiting the triangle inequality, i.e., \( |\beta_1 - \beta_2| \leq |\beta_1 - \beta_2| \), we obtain the following upper bound on the type II error probability \( P_e(i,j) \).

C. Converse Proof

The proof of converse is based on the following two steps.

- **Step 1:** First, we show in Lemma 2 that for any achievable rate (for which the type I and type II error rates vanish as \( n \to \infty \)), the distance between any selected entry of one codeword with any entry of another codeword should be at least larger than a threshold.

- **Step 2:** Then, using Lemma 2, we derive an upper bound on the codebook size of achievable identification codes.

**Lemma 2.** Suppose that \( R \) is an achievable rate for the DTPC. Consider a sequence of \( (n, M(n, R), K(n, \kappa), e_1^{(n)}, e_2^{(n)}) \) codes \( (C^{(n)}, T^{(n)}) \) such that \( e_1^{(n)} \) and \( e_2^{(n)} \) tend to zero as \( n \to \infty \). Then, given a sufficiently large \( n \), codebook \( C^{(n)} \) satisfies the following property. For every pair of codewords, \( c_{i_1} \) and \( c_{i_2} \), there exists at least one letter \( t \in [n] \) such that
\[
1 - \rho_0 c_{i_1,t} + I_{c_{i_1,t}}^2 > \theta_n,
\]
for all \( i_1, i_2 \in [M] \), such that \( i_1 \neq i_2 \), with \( \theta_n > 0 \), is an arbitrarily small constant and \( I_{c_{i_1,t}}^2 \) is bounded by \( \lambda + n \rho_0 c_{i_1,t} + \kappa \), \( z \in \{1, 2\} \).

**Proof.** The proof is given in Appendix A.

Next, we use Lemma 2 to prove the upper bound on the DTC capacity. Observe that since \( d_{i,t} = 4 \rho_0 c_{i,t} + I_{c_{i,t}}^2 > \lambda \), Lemma 2 implies
\[
\rho_0 c_{i_1,t} + I_{c_{i_1,t}}^2 > \rho_0 c_{i_2,t} + I_{c_{i_2,t}}^2 > 4 \rho_0 c_{i_2,t} + \lambda \theta_n,
\]
where \( \lambda \theta_n > 0 \), which follows by (17). Now, since \( c_{i_1,t} - c_{i_2,t} \geq 0 \), we deduce that the distance between every pair of codewords satisfies \( d_{c_{i_1,t}, c_{i_2,t}} \geq \lambda \theta_n / \rho_0 \). Thus, we can arrange a definition of non-overlapping spheres \( S_{c_{i,t}}(n, \lambda \theta_n / \rho_0) \), i.e., spheres of radius \( \lambda \theta_n / \rho_0 \), that are centered at the codewords \( c_{i,t} \). Since the codewords all belong to a hypercube \( Q_0(n, P_{\max}) \) with edge length \( P_{\max} \), it follows that the number of packed small spheres, i.e., the number of codewords \( M \) is bounded by
\[
M = \frac{\text{Vol} \left( \bigcup_{i=1}^{M} S_{c_i}(n, r_0) \right)}{\text{Vol} \left( S_{c_1}(n, r_0) \right)} \leq 2^{-0.599n} \cdot \frac{P_{\max}}{\text{Vol} \left( S_{c_1}(n, r_0) \right)},
\]
for sufficiently large \( n \), where \( \rho_0 > 0 \) is an arbitrarily small constant, see [26, Subsec. III-B] for detailed derivation.

We now proceed with bounding \( \Pr(E_1) \) as follows. Based on the codebook construction, each codeword is surrounded by a sphere of radius \( \sqrt{n} \theta_n \), that is \( \| c_i - c_{i,t} \| \geq 4 \theta_n \). Thus, employing the Chebyshev's inequality and [26, Subsec. III-B], we can establish following bound for event \( E_1 \):
\[
Pr(E_1) \leq \frac{6 (A + \lambda)^2 \left( 1 + c^2 + (A + \lambda) + (A + \lambda)^2 \right)}{4(A - c)^2 \rho_0^2 a^2 n^{a + b}} \leq \zeta_1,
\]
for sufficiently large \( n \), where \( \zeta_1 > 0 \) is an arbitrarily small constant, see [26] for step by step derivation. Therefore
\[
P_e(i,j) \leq \Pr(E_0) + \Pr(E_1) \leq \zeta_0 + \zeta_1 \leq \zeta_2,
\]
holds for sufficiently large \( n \) and arbitrarily small \( \zeta_2 > 0 \). We have thus shown that for every \( e_1, e_2 > 0 \) and sufficiently large \( n \), there exists an \( (n, M(n, R), K(n, \kappa), e_1, e_2) \) code.
Assume to the contrary that there exist two messages $i_1$ and $i_2$, where by (2), the bracket is

$$
\left\{ y \in T_{i_1} : \frac{1}{n} \sum_{t=1}^{n} y_t \leq \rho_0 P_{\text{max}} + \epsilon_{i_1} + \delta \right\},
$$

(19)

Then, observe that $P_{e,1}(i_1) + P_{e,2}(i_2, i_1) \geq 1 - \sum_{\iota_i} |V_i^{\text{n}}(y_{c_{i_1}}) - V_i^{\text{n}}(y_{c_{i_2}})| - \eta$ where if $V_i^{\text{n}}(y_{c_{i_1}}) + \sum_{\iota_i} |V_i^{\text{n}}(y_{c_{i_1}}) - V_i^{\text{n}}(y_{c_{i_2}})| - \eta$ > 1 - $\sum_{\iota_i} |V_i^{\text{n}}(y_{c_{i_1}}) - V_i^{\text{n}}(y_{c_{i_2}})| - \eta$. Hence, we replaced $R_{i_1} \cap T_{i_1}$ by $R_{i_1}$ to enlarge the domain and for second inequality, we used $\sum_{\iota_i} |V_i^{\text{n}}(y_{c_{i_1}}) - V_i^{\text{n}}(y_{c_{i_2}})| \leq 1$. Since this is a contradiction, the assumption in (18) is false. This completes the proof of Lemma 2.

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 Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1887–1919, 2016.

[3] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel, and R. Schober, “Channel modeling for diffusive molecular communication - A tutorial review,” *Proc. IEEE*, vol. 107, no. 7, pp. 1302–1341, 2019.

[4] M. Kuscu, E. Dine, B. A. Bilgin, H. Ramezani, and O. B. Akan, “Transmitter and receiver architectures for molecular communications: A survey on physical design with modulation, coding, and detection techniques,” *Proc. IEEE*, vol. 107, no. 7, pp. 1302–1341, 2019.