IEEE Open Journal of the Communications Society

Received 28 November 2023; revised 9 January 2024; accepted 23 January 2024. Date of publication 29 January 2024; date of current version 16 February 2024.

Digital Object Identifier 10.1109/OJCOMS.2024.3359186

Deterministic K-Identification for MC Poisson Channel With Inter-Symbol Interference

MOHAMMAD JAVAD SALARISEDDIGH [1,2], VAHID JAMALI [3] (Senior Member, IEEE), UZI PEREG [4] (Member, IEEE), HOLGER BOCHE [2,5] (Fellow, IEEE), CHRISTIAN DEPPE [2,6] (Member, IEEE), AND ROBERT SCHOBER [7] (Fellow, IEEE)

(Invited Paper)

[1]School of Computation, Information and Technology, Department of Electrical and Computer Engineering, Technical University of Munich, 80333 Munich, Germany

[2]BMBF Research Hub 6G-life, Technical University of Munich, 80333 Munich, Germany

[3]Department of Electrical Engineering and Information Technology, Technical University of Darmstadt, 64289 Darmstadt, Germany

[4]Viterbi Faculty of Electrical and Computer Engineering, Technion—Israel Institute of Technology, Haifa 3200003, Israel

[5]Chair of Theoretical Information Technology, Technical University of Munich, 80333 Munich, Germany

[6]Institute for Communications Technology, Technical University of Braunschweig, 38106 Braunschweig, Germany

[7]Department of Electrical Engineering, Electronics and Information Technology, Friedrich–Alexander-University Erlangen–Nürnberg, 91054 Erlangen, Germany

CORRESPONDING AUTHOR: M. J. Salariseddigh (e-mail: mjss@tum.de)

The work of Mohammad Javad Salariseddigh was supported in part by the German Research Foundation (DFG) under Grant DE 1915/2-1, and in part by the 6G-Life Project under Grant 16KISK002. The work of Vahid Jamali was supported in part by the German Federal Ministry of Education and Research (BMBF) under Grant 16KIS1992, and in part by the LOEWE Initiative, Hesse, Germany, within the emergenCITY Center. The work of Uzi Pereg was supported in part by the Israel Science Foundation (ISF) under Grant 939/23 and Grant 2691/23; in part by the Israel VATAT Junior Faculty Program for Quantum Science and Technology under Grant 86636903; in part by the Chaya Career Advancement Chair under Grant 8776026; in part by the German-Israeli Project Cooperation (DIP) under Grant 2032991; in part by the Helen Diller Quantum Center at the Technion; in part by BMBF under Grant 16KIS1005 and Grant 16KISQ028; and in part by the 6G-Life Project under Grant 16KISK002. The work of Holger Boche was supported in part by BMBF under Grant 16KIS1003K; in part by the National Initiative for "Molecular Communications" (MAMOKO) under Grant 16KIS0914; and in part by the 6G-Life Project under Grant 16KISK002. The work of Christian Deppe was supported in part by BMBF under Grant 16KIS1005 and Grant 16KISQ028, and in part by the 6G-Life Project under Grant 16KISK002. The work of Robert Schober was supported by MAMOKO under Grant 16KIS0913.

This paper was presented in part at the IEEE International Conference on Communications (ICC 2023) [1] [DOI: 10.1109/ICC45041.2023.10278856].

ABSTRACT Various applications of molecular communications (MCs) 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 K-identification (DKI) for the discrete-time *Poisson* channel (DTPC) with inter-symbol 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 DKI capacity of the DTPC with ISI when the size of the target message set K and the number of ISI channel taps L may grow with the codeword length n. As a key finding, we establish that for deterministic encoding, assuming that K and L both grow sub-linearly in n, i.e., $K = 2^{\kappa \log n}$ and $L = 2^{l \log n}$ with $\kappa + 4l \in [0, 1)$, where $\kappa \in [0, 1)$ is the identification target rate and $l \in [0, 1/4)$ is the ISI rate, then the number of different messages that can be reliably identified scales super-exponentially in n, i.e., $\sim 2^{(n \log n)R}$, where R is the DKI coding rate. Moreover, since l and κ must fulfill $\kappa + 4l \in [0, 1)$, we show that optimizing l (or equivalently the symbol rate) leads to an effective identification rate [bits/s] that scales sub-linearly with n. This result is in contrast to the typical transmission rate [bits/s] which is independent of n.

INDEX TERMS Channel capacity, deterministic identification, inter-symbol interference, molecular communication, Poisson channel.

I. INTRODUCTION

MOLECULAR communication (MC) is a new communication concept where messages are embedded in the properties of molecules [2], [3]. In contrast to conventional electromagnetic-based (EM) communication systems, which embed information into the properties of EM waves such as their amplitude, frequency, and phase, MC systems embed information into the properties of molecules such as their concentration [4], type [5], time of release [6], and spatial release pattern [7]. A parallel growing and related field to MC is synthetic biology [8], which provides tools for realizing the hardware components needed for MC systems. In [9], [10], the realization of MC systems using synthetic biology techniques is discussed and biological components are investigated which can potentially serve as the main building blocks of synthetic MC systems, i.e., as transmitter, receiver, and signaling particles. This promising vision for realizing synthetic MC systems has motivated the research community to establish theoretical frameworks for their modeling, design, and analysis. Examples include channel modeling [11], modulation and detection design [12], and information-theoretical performance characterization [13]. Advanced synthetic MC systems are expected to facilitate the realization of the Internet of Bio-Nano Things [14], [15] capable of performing sophisticated tasks such as sensing, computing, and networking inside the human body.

A. POISSON CHANNEL WITH INTER-SYMBOL INTERFERENCE

In the context of MC systems, information can be encoded in the concentration (rate) of molecules released by the transmitter and be decoded based on the number of molecules observed at the receiver. Assuming that the release, propagation, and reception of different molecules are independent of each other, the number of molecules observed by molecule counting receivers in MC systems are characterized by the Binomial distribution. However, the Binomial channel model and in particular its probability and cumulative distribution functions make theoretical analysis cumbersome. Fortunately, in most MC applications, the number of released molecules is quite large which allows approximating the Binomial channel model by the Gaussian or Poisson channel models [16]. Specifically, when the number of molecules emitted by the transmitter, N, is large but the probability of successful observation of one molecule at the receiver, p, is small (such that Np is still small), the Poisson distribution can be shown to result as the limiting case of the Binomial distribution[1] [11, Sec. IV].

Diffusive MC channels are inherently dispersive since molecules do not quickly fade away and stay in the channel for a long time. This leads to a long tail of channel impulse response (CIR) and causes inter-symbol interference (ISI) [13]. The number of relevant channel memory taps, L, depends on the relative length of the CIR and the symbol duration and hence is a function of the symbol rate. Motivated by the above discussions, we focus on investigating the fundamental performance limits of the discrete-time Poisson channel (DTPC) with ISI in this paper.

[1]In the literature, this result is also known as the *Poisson limit theorem* or the *law of rare events* [17].

B. INFORMATION THEORETICAL ANALYSIS OF MC SYSTEMS

Despite the recent theoretical and technological advancements in the field of MC, the information-theoretical performance limits of DTPC MC systems with and without ISI are still not fully understood [13]. In fact, finding an analytic expression for the transmission rate (TR) capacity of the DTPC with ISI under an average power constraint is still an open problem [13], [18], [19]. Nevertheless, for characterizing the TR capacity for the DTPC, a number of approaches have been explored and several bounds and asymptotic results for the DTPC with ISI have been established. For instance, analytical lower and upper bounds on the TR capacity of the DTPC with input constraints and ISI are provided in [20]. Bounds on the TR capacity of the DTPC with ISI are developed in [21], [22]. The design of optimal codes for the DTPC with ISI is studied in [23], [24]. In [25], the impact of ISI on the transmission performance over a diffusive MC channel is investigated. Nonetheless, the DTPC with ISI has been mostly studied for the TR problem in the existing literature. On the other hand, in [26], [27], deterministic identification (DI) for the DTPC *without* ISI is studied, where bounds on the DI capacity are established. To the best of the authors' knowledge, for the DTPC *with* ISI, the fundamental performance limits for the DI problem, have not been investigated in the literature, yet, except in the conference version of this paper [1].

C. APPLICATIONS OF THE K-IDENTIFICATION PROBLEM FOR MC SCENARIOS

Numerous envisioned applications of MCs under the umbrella of future generation (XG) communication networks [28], [29] give rise to event-triggered communication scenarios,[2] where TR capacity may not be the appropriate performance metric. In particular, in event-detection, object-finding or alarm-prompt scenarios, where the receiver has to decide about the occurrence of a specific event or the presence of an object with a reliable Yes / No answer, the so-called K-identification capacity is the relevant performance measure [31]. More specifically, in the K-identification problem, it is assumed that the receiver is interested in a subset of size K of the message set, $\mathcal{M} = \{1, \ldots, M\}$, referred to as the target message set. Since $\mathcal{M}$ has cardinality M, there are in total $\binom{M}{K}$ possible target message sets or subsets of size K. For each *inclusion test*, the receiver chooses an arbitrary message from the message set and checks whether or not it belongs to a given target message set. The error criteria imposed on the corresponding K-identification codes dictate that such an *inclusion test* must be reliable no matter which specific target message set is considered.

[2]Such communication systems are also known as post-Shannon communication systems in the literature [29]. A detailed discussion of the potential of MC and post-Shannon communication for the sixth generation (6G) of communication systems can be found in [30].

TABLE 1. Mathematical notations used throughout this paper.

Symbols	Description
$\mathbb{X}, \mathbb{Y}, \mathbb{Z}, \ldots$ (blackboard bold letters)	Alphabet sets
$\mathbb{X}^c$	Complement of set $\mathbb{X}$
$x, y, z, \ldots$ (lower case letters)	Constants and values of random variables (RVs)
$X, Y, Z, \ldots$ (upper case letters)	RVs
$\mathbf{x}, \mathbf{y}, \ldots$ (lower case bold symbols)	Row vectors
$\mathbf{1}_n$	All-ones row vector of size n
$[\![M]\!] \triangleq \{1, 2, \ldots, M\}$	Set of consecutive natural numbers from 1 through M
$\mathbb{N}_0 \triangleq \{0, 1, 2, \ldots\}$	Set of whole numbers
$\mathbb{R}_+$	Set of non-negative real numbers
$x! \triangleq x \times (x-1) \times \cdots \times 1$	Factorial for non-negative integer x
$\Gamma(x) = (x-1)! \triangleq (x-1) \times (x-2) \times \cdots \times 1$	Gamma function for non-negative integer x
$f(n) = o(g(n))$ (small O notation)	$f(n)$ is dominated by $g(n)$ asymptotically, i.e., $\lim_{n\to\infty} f(n)/g(n) = 0$
$f(n) = O(g(n))$ (big O notation)	$\lvert f(n)\rvert$ is bounded above by $g(n)$ (up to constant factor) asymptotically, i.e., $\limsup_{n\to\infty} \lvert f(n)\rvert/g(n) < \infty$
$\mathbb{E}[X]$	Statistical expectation of RV X
$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$	Covariance of two real-valued RVs X and Y with finite second moments
$\lVert\mathbf{x}\rVert_1, \lVert\mathbf{x}\rVert, \lVert\mathbf{x}\rVert_\infty$	ℓ_1-norm, ℓ_2-norm, ℓ_∞-norm
$\mathcal{S}_{\mathbf{x}_0}(n, r) = \{\mathbf{x} \in \mathbb{R}_+^n : \lVert\mathbf{x} - \mathbf{x}_0\rVert \leq r\}$	n-dimensional hyper sphere of radius r centered at $\mathbf{x}_0$ with respect to the ℓ_2-norm
$\mathbb{Q}_{\mathbf{0}}(n, A) = \{\mathbf{x} \in \mathbb{R}_+^n : 0 \leq x_t \leq A, \forall t \in [\![n]\!]\}$	n-dimensional cube with center $(A/2, \ldots, A/2)$ and a corner at the origin, i.e., $\mathbf{0} = (0, \ldots, 0)$, whose edges have length A
$H(z) \triangleq -z\log(z) - (1-z)\log(1-z)$	Binary entropy function for $z \in [0, 1]$
$\mathbb{C}_{\mathrm{TR}}$	Message transmission capacity of a channel
$\mathcal{P}$	Poisson channel with ISI

Concrete examples of the K-identification problem in the context of MC can be found in communication scenarios featuring event/object recognition tasks. In particular, for targeted drug delivery [2], [32], [33], [34], [35] where, e.g., a nano-device's objective may be to identify whether or not a specific biomarker present around a target tissue belongs to a certain category of cancers; in health monitoring [36], [37], where, e.g., one may be interested in finding to which group/set of diseases a target bacteria belongs to. Moreover, K-identification problems may find applications in natural MC systems. For example, in natural olfactory MC systems [38], [39], [40], where the communication goal may involve the *inclusion* of a specific type of secreted odor/pheromone into a target group of K-odors corresponding to a specific identification task for foraging, mating, etc.

Besides, in the context of molecular modeling [41], a computational representation of an MC unit, called a *digital twin* [42], may be required. In order to manage complex tasks (e.g., prediction of future behavior) and perform reliable computational functions (e.g., real-time simulation) on the digital twin, it has to continually remain consistent with its real counterpart [42]. Such a virtual copy of a target MC unit allows experts to accomplish and evaluate their subsequent computational tasks in a more reliable manner. Therefore, it is crucial for the *digital twin* to verify/identify whether or not it is consistent with the real MC unit. Examples include the creation of a functioning Human brain at the molecular level [43] and real-time calibration between an operating nano-scale communication system and its digital twin [44].

D. CONTRIBUTIONS

In this paper, we study the problem of deterministic K-identification (DKI) over the DTPC with ISI under average and peak molecule release rate constraints which account for the restricted molecule production / release rates by the transmitter. In particular, this work makes the following contributions:

- *Generalized DKI and ISI model:* In this paper, we study the DTPC, where the ISI memory length, L, and the size of the identification set, K, may scale with the codeword length, n. As special cases, this model includes the ISI-free channel ($L = 1$), the ISI channel with constant L, DI ($K = 1$), and DKI with constant K. For a given MC channel, scaling L implies a higher symbol rate. Therefore, the proposed generalized model allows us to investigate whether large codeword lengths enable reliable identification even if the symbol rate is increased (or similarly K is increased). To the best of the authors' knowledge, such a generalized DKI and ISI model has not been studied in the literature, yet.
- *Codebook scale:* We establish that the codebook size for K-identification for the DTPC with ISI for deterministic encoding scales in n similar as for the memoryless DTPC [27], [45], namely super-exponentially in the codeword length, i.e., $\sim 2^{(n\log n)R}$, where R is the DKI

coding rate, even when the size of the target message set K and the number of ISI channel taps L, both grow sub-linearly in n, i.e., $K = 2^{\kappa \log n}$ and $L = 2^{l \log n}$, respectively, where $\kappa + 4l \in [0, 1)$, has to hold, and $\kappa \in [0, 1)$ is called the identification target rate and $l \in [0, 1/4)$ is the ISI rate. This result reveals that the set of target messages for identification and the ISI memory can indeed scale with n without affecting the scale of the codebook, confirming the result for the standard identification problem for the memoryless DTPC (i.e., $K = L = 1$) [45].

- *Capacity bounds:* We derive DKI capacity bounds for constant $K \geq 1$ and growing $K = 2^{\kappa \log n}$ for the dispersive DTPC with constant $L \geq 1$ and growing ISI $L = 2^{l \log n}$, respectively. We show that for constant K and L, the proposed lower and upper bounds on R are independent of K and L, whereas for growing target message set or growing number of ISI taps, they are functions of the target identification rate κ and ISI rate l, respectively. Moreover, we show that optimizing the ISI rate l (or equivalently the symbol rate) leads to an effective identification rate [bits/s] that scales super-linearly with n. This result is in contrast to the typical transmission rate in [bits/s] for which the rate is independent of n.
- *Technical novelty in the capacity proof:* To obtain the proposed lower bound on the DKI capacity, we analyze the input space imposed by the input constraints and exploit it for an appropriate sphere packing (non-overlapping spheres with identical radius), namely we consider the packing of hyper spheres inside a larger hypercube, whose radius grows in the codeword length n, the target identification rate κ, and the ISI rate l, i.e., $\sim n^{(1+\kappa+4l)/4}$. Unlike the existing construction for Gaussian channels [46], [47], where the radius of spheres vanishes for asymptotic codeword length n, i.e., $n \to \infty$, here, the radius of the hyper spheres tends to infinity with a polynomial growth, i.e., $\sim n^{(1+\kappa+4l)/4}$. This packing incorporates the impact of the size of the target message set and the ISI as functions of κ and l, respectively. For derivation of the upper bound on the DKI capacity, we assume that an arbitrary sequence of codes with vanishing error probabilities is given. Then, for such a sequence of codes, we prove that a certain minimum distance between the codewords is asserted. Unlike the previous construction for Gaussian [46], [47] and memoryless channels, here this distance converges to zero more rapidly for the asymptotic codeword length n and depends on the target identification rate and the ISI rate and decreases as K and L grow, respectively.

E. ORGANIZATION

The remainder of this paper is structured as follows. Section II reviews previous results for the DI, RI, and DKI problems and includes background information. In Section III, the system model is presented and the required preliminaries regarding DKI codes are established. Section IV provides the main contributions and results on the K-identification capacity of the DTPC with ISI. Finally, Section V concludes the paper with a summary and directions for future research.

The notations adopted throughout this paper are summarized in Table 1. Moreover, all logarithms are to base two.

II. BACKGROUND ON IDENTIFICATION PROBLEM

In this section, we establish the required background for our work and introduce the identification problem. Furthermore, we review relevant previous results on the randomized-encoder identification (RI), DI, and DKI capacities for different channels.

A. IDENTIFICATION PROBLEM

In Shannon's communication paradigm [48], a sender, Alice, encodes her message in a manner that will allow the receiver, Bob, to reliably recover the message. In other words, the receiver's task is to determine which message was sent. In contrast, in the identification setting, the coding scheme is designed to accomplish a different objective [31]. The decoder's main task is to determine whether a *particular* message was sent or not, while the transmitter does not know which message the decoder is interested in.

Randomized identification: Ahlswede and Dueck [31] introduced an RI scheme, in which the codewords are tailored according to their corresponding random source (distribution). Note that such an approach cannot increase the TR capacity for Shannon's message transmission task [49]. On the other hand, Ahlswede and Dueck [31] established that given local randomness at the encoder, reliable identification is accomplished with a codebook size that is double-exponential in the codeword length n, i.e., $\sim 2^{2^{nR}}$ [31], where R is the coding rate. This behavior differs radically from the conventional message transmission setting, where the codebook size grows only exponentially with the codeword length, i.e., $\sim 2^{nR}$. Therefore, RI yields an exponential gain in the codebook size compared to the transmission problem. The construction of RI codes has been considered in previous works [50], [51]. For example, in [50], a binary code is constructed based on a three-layer concatenated constant-weight code.

Deterministic identification: The realization of RI codes can be challenging in practice since they require the implementation of probability distribution functions. Therefore, from a practical point of view, it is of interest to consider the case where the codewords are not selected based on distributions but rather by means of a deterministic mapping from the message set to the input space. This is known as DI in which the encoder is a deterministic function.

K-Identification framework: For the standard DI or RI problems [31], [47], the receiver is interested in identifying a *single* message, that is, it selects an arbitrary message called target message and using a decision rule (decoder) decides

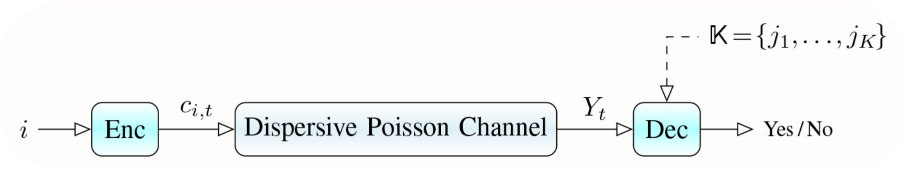

FIGURE 1. End-to-end transmission chain for DKI communication in a generic MC system modelled as a Poisson channel with ISI. The transmitter maps message i onto a codeword $\mathbf{c}_i = (c_{i,t})|_{t=1}^n$. The receiver selects an arbitrary target message set $\mathbb{K} = \{j_1, \ldots, j_K\}$ and, given the channel output vector $\mathbf{Y} = (Y_t)|_{t=1}^n$, asks whether or not the sent message i belongs to target message set $\mathbb{K}$.

whether or not this target message is *identical* to the sent message. For the K-identification problem [52], the receiver selects a subset of K messages from the message set called target message set (denoted by $\mathbb{K}$) and in contrast to the standard DI or RI problems, it decides whether or not the sent message *belongs to* this target message set. We note that such a target message set can be in general any arbitrary subset of the message set of size K, where the total possible number of such subsets is $\binom{M}{K}$. The K-identification scenario may be interpreted as a *generalization* of the standard DI or RI problems in the sense that the target message at the receiver is substituted with a set of K target messages, where $K \geq 1$. That is, the DKI for the special case where $K = 1$ corresponds to the standard DI problem considered in [45], [53].

B. RELATED WORK ON DI CAPACITY

In the deterministic coding setup for identification, for discrete memoryless channels (DMCs), the codebook size grows only exponentially in the codeword length, similar to the conventional transmission problem [31], [54], [55], [56], [57], [58]. However, the achievable identification rates are significantly higher compared to the transmission rates [47], [55]. Compared to RI codes, DI codes often have the advantage of simpler implementation and simulation [59], [60] and explicit construction [61]. In [47], [55], DI for DMCs with an average power constraint is considered and a full characterization of the DI capacity is provided. In [46], [47], Gaussian channels with fast and slow fading and subject to an average power constraint are studied and the codebook size is shown to scale as $2^{(n \log n)R}$. DI is also studied in [62] for Gaussian channels in the presence of feedback and in [63] for general continuous-time channels with infinite alphabets. Furthermore, DI for MC channels modelled as DTPC without ISI and the Binomial channel is studied in [26], [27], [45], [53], where the scale of the size of the codebook is shown to be $2^{(n \log n)R}$.

C. RELATED WORK ON DKI CAPACITY

Randomized K-identification for the DMC is studied in [52] where assuming $K = 2^{\kappa n}$, the set of all achievable pairs of the identification coding rate R and the target identification rate κ, is shown to contain $\{(R, \kappa) : R, \kappa \geq 0;\ R + 2\kappa \leq \mathbb{C}_{\text{TR}}\}$. Assuming $K = 2^{\kappa \log n}$, the DKI for slow fading channels, denoted by $\mathcal{G}_{\text{slow}}$, subject to an average power constraint and a codebook size of super-exponential scale, i.e., $\sim 2^{(n \log n)R}$, is studied in [64], [65] and the following bounds on the DKI capacity are derived: $(1 - \kappa)/4 \leq \mathbb{C}_{\text{DKI}}(\mathcal{G}_{\text{slow}}, M, K) \leq 1 + \kappa$, for $0 \leq \kappa < 1$. Also, a full characterization of the DKI capacity for the binary symmetric channel subject to a Hamming weight constraint is established in [66]. On the other hand, to the best of the authors' knowledge, the DKI capacity of the DTPC with ISI (which is relevant for MC systems) has not been studied in the literature, yet, and hence it is the main focus of this paper.

III. SYSTEM MODEL AND K-IDENTIFICATION CODING

In this section, we present the adopted system model and establish some preliminaries regarding DKI coding.

A. SYSTEM MODEL

In this paper, we consider a K-identification-focused communication setup, where the decoder aims to accomplish the following task: Determining whether or not a specific received message belongs to a target set of messages of size K; see Figure 1. To achieve this objective, a coded communication between the transmitter and the receiver over n uses of the MC channel is established by modulating the molecule concentration. We assume that the transmitter releases molecules with rate x_t (molecules/second) during T_R seconds at the beginning of each symbol interval having a length of T_S seconds [13]. These molecules propagate through the channel via diffusion and/or advection, and may even experience degradation in the channel via enzymatic reactions [11]. The ISI of the channel is modelled by a length L sequence of probability values, i.e., $\mathbf{p} = (p_0, p_1, \ldots, p_{L-1})$, where value $p_l \in (0, 1]$ denotes the probability that a given molecule released by the transmitter for the t-th channel use is observed at the receiver during time slot $t + l$. Further, let $\boldsymbol{\rho} \triangleq (\rho_0, \ldots, \rho_{L-1})$, where $\rho_l \triangleq p_l T_R$.

We assume a counting-type receiver.[3] Examples of such receivers include the transparent receiver, which counts

[3]We note that there are different types of receivers in MC systems including timing receivers, counting receivers, concentration-based receivers, and receivers using secondary/indirect signals, see [3], [11] for a comprehensive review. We adopt counting receivers in this paper, since on the one hand, they are not as complex as timing receivers, and on the other hand, they are more accurate than concentration-based receivers or receivers that employ secondary/indirect signals.

the molecules that are at a given time within its sensing volume [67], the fully absorbing receiver, which absorbs and counts the molecules hitting its surface within a given time interval [68], and the reactive receiver which counts the molecules bound to the ligand proteins on its surface at a given time [69]. The value of p_l depends on parameters such as the diffusion coefficient of the molecules, D, the propagation environment (e.g., diffusion, advection, and reaction processes), the distance between transmitter and receiver, d, and the type of reception mechanism (e.g., transparent, absorbing, or reactive receiver); see [11, Sec. III] for the characterization of p_l for various setups. For instance, assuming instantaneous release (i.e., $T_R \to 0$) of molecules by a point-source transmitter, molecule propagation via diffusion in an unbounded three-dimensional environment, and assuming a uniform concentration approximation of molecules within the reception volume of a transparent receiver, p_l can be obtained as [11]

$$p_l = \frac{V_{\text{rx}}}{(4\pi D\tau_l)^{3/2}} \cdot e^{-d^2/(4D\tau_l)}, \tag{1}$$

where V_{rx} is the reception volume size and $\tau_l \triangleq lT_S + \bar{\tau}$ with $l \in \{0, \ldots, L-1\}$, denotes the sampling time at the receiver, where $\bar{\tau}$ is a constant time offset between the release time by the transmitter and sample time at the receiver within each symbol duration.

Assuming that the release, propagation, and reception of individual molecules are statistically identical but independent of each other, the received signal follows Poisson statistics when the number of released molecules is large, i.e., $\lfloor x_t T_R \rfloor \gg 1$ [11, Sec. IV]. We assume that $X \in \mathbb{R}_{\geq 0}$ and $Y \in \mathbb{N}_0$ denote RVs modeling the rate of molecule release by the transmitter and the number of molecules observed at the receiver, respectively. The channel output Y is related to the channel input X according to

$$Y_t = \text{Pois}\left(X_t^{\rho} + \lambda\right), \tag{2}$$

where

$$X_t^{\rho} \triangleq \sum_{l=0}^{L-1} \rho_l X_{t-l}, \tag{3}$$

is the mean number of observed molecules at the receiver after the release of molecules at the time t. The constant $\lambda \in \mathbb{R}_{>0}$ is the mean number of observed interfering molecules originating from external noise sources which employ the same type of molecule as the considered MC system. Let

$$\mathbf{x}_t^* \stackrel{\text{def}}{=} (x_{t-L+1}, \ldots, x_t)$$

be the vector of the L most recently released symbols. Considering the Poisson distribution provided in (2), the letter-wise conditional distribution of the output of the DTPC with ISI $\mathcal{P}$ is given by

$$V(Y_t|\mathbf{x}_t^*) = e^{-\left(X_t^{\rho}+\lambda\right)} \left(X_t^{\rho} + \lambda\right)^{Y_t} / (Y_t!). \tag{4}$$

Standard transmission schemes employ strings of letters (symbols) of length n, referred to as codewords, that is, the encoding scheme uses the channel in n consecutive symbol intervals to transmit one message. Since the channel is dispersive, each output symbol is influenced by the L most recent input symbols. As a consequence, the receiver observes a string of length $\bar{n} = n + L - 1$, referred to as output vector (received signal). Since the ISI of the channel, characterized by $\mathbf{p} = (p_0, p_1, \ldots, p_{L-1})$, has length L, we assume that different channel uses given any L previous input symbols are statistically independent. Therefore, for n channel uses, the transition probability distribution is given by

$$V^{\bar{n}}(\mathbf{y}|\mathbf{x}) = \prod_{t=1}^{\bar{n}} V(Y_t|\mathbf{x}_t^*) = \prod_{t=1}^{\bar{n}} \frac{e^{-\left(X_t^{\rho}+\lambda\right)} \left(X_t^{\rho} + \lambda\right)^{Y_t}}{Y_t!}, \tag{5}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_{\bar{n}})$ denote the transmitted codeword and the received signal, respectively.

We assume that $x_t = 0$ for $t > n$ or $t < 0$. We impose peak and average molecule release rate constraints on the codewords as follows

$$0 \leq x_t \leq P_{\max} \quad \text{and} \quad \frac{1}{n}\sum_{t=1}^{n} x_t \leq P_{\text{avg}}, \tag{6}$$

respectively, $\forall t \in [\![n]\!]$, where $P_{\max} > 0$ 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. Imposing such constraints on the rate of the released molecules is motivated by the fact that there are a *finite* and limited number of signalling molecules contained in the molecule reservoir and the constraints guarantee that, for large number of channel uses, the number of stored molecules suffices. Unlike the classical average power constraint imposed on the input of the Gaussian channel which is a non-linear function of the symbols signifying the signal (symbol) energy, here for the DTPC with ISI, the average constraint is a linear function of the symbols signifying the number of released molecules normalized by the codeword length [13].

B. DKI CODING FOR THE POISSON CHANNEL WITH ISI

The definition of a DKI code for the Poisson channel with ISI, $\mathcal{P}$, is given below.

Definition 1 (ISI-Poisson DKI Code): An $(n, M(n, R), K(n, \kappa), L(n, l), e_1, e_2)$ DKI code for a Poisson channel with ISI, $\mathcal{P}$, under average and peak molecule release rate constraints of $P_{\text{ave}} > 0$, and $P_{\max} > 0$, respectively, and for integers $M(n, R)$, $K(n, \kappa)$, and $L(n, l)$, where n, R, κ, and l are the codeword length, the DKI coding rate, the target identification rate, and the ISI rate, respectively, is defined as a system $(\mathcal{C}, \mathscr{T})$, which consists of a codebook $\mathcal{C} = \{\mathbf{c}_i\} \subset \mathbb{R}_+^n$, with $i \in [\![M]\!]$, such that

$$0 \leq c_{i,t} \leq P_{\max} \quad \text{and} \quad \frac{1}{n}\sum_{t=1}^{n} c_{i,t} \leq P_{\text{avg}}, \tag{7}$$

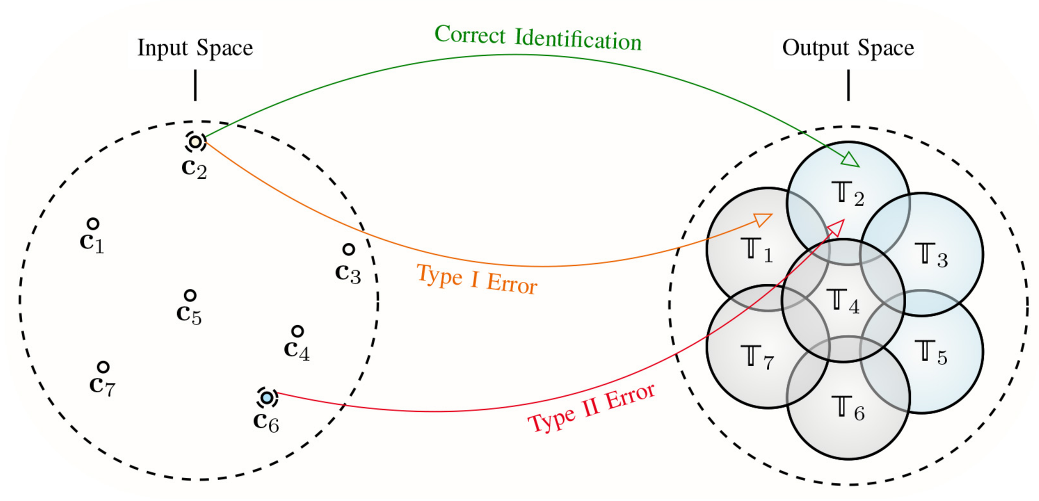

FIGURE 2. Depiction of a DKI setting with $K = 3$ and target message set $\mathbb{K} = \{2, 3, 5\}$ colored in blue. In the correct identification event, the channel output is detected in the union of the individual decoders $\mathbb{T}_j$, where j belongs to the target message set $\mathbb{K}$. A type I error event occurs if the channel output is observed in the *complement* of the union of the individual decoders to which the index of the codeword belongs to. A type II error event occurs if the channel output is detected in the union of the individual decoders $\mathbb{T}_j$, with $j \in \mathbb{K}$, but the index of the sent codeword does *not* belong to $\mathbb{K}$.

$\forall i \in [\![M]\!]$, $\forall t \in [\![n]\!]$ and a decoder $\mathscr{T}_{\mathbb{K}} \subseteq \mathbb{N}_0^{\bar{n}}$, where $\mathbb{K}$ is an arbitrary subset of size K. That is $\mathbb{K} \in \{\mathbb{G} \subseteq [\![M]\!]; |\mathbb{G}| = K\}$. Given a message $i \in [\![M]\!]$, the encoder sends $\mathbf{c}_i$, and the decoder's task is to perform a binary hypothesis test: Was a target message $j \in \mathbb{K}$ sent or not? There exist two types of errors that may happen[4] (see Figure 2):

Type I error: Rejection of the *correct* message, $i \in \mathbb{K}$.

Type II error: Acceptance of a *wrong* message, $i \notin \mathbb{K}$.

The associated error probabilities of the DKI code reads

$$P_{e,1}(i, \mathbb{K}) = \Pr\left(\mathbf{Y} \in \mathscr{T}_{\mathbb{K}}^c \,\middle|\, \mathbf{x} = \mathbf{c}_i\right) = 1 - \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}}} V^{\bar{n}}\left(\mathbf{y} \,\middle|\, \mathbf{c}_i\right), \quad i \in \mathbb{K}, \tag{8}$$

$$P_{e,2}(i, \mathbb{K}) = \Pr\left(\mathbf{Y} \in \mathscr{T}_{\mathbb{K}} \,\middle|\, \mathbf{x} = \mathbf{c}_i\right) = \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}}} V^{\bar{n}}\left(\mathbf{y} \,\middle|\, \mathbf{c}_i\right), \quad i \notin \mathbb{K}, \tag{9}$$

and $\forall e_1, e_2 > 0$ fulfill the bounds $P_{e,1}(i, \mathbb{K}) \leq e_1, \forall i \in \mathbb{K}$, and $P_{e,2}(i, \mathbb{K}) \leq e_2, \forall i \notin \mathbb{K}$.

Note that correct K-identification implies that neither type-I nor type-II errors occur. In this paper, we are interested in the asymptotic case when arbitrary small error probabilities are achievable for sufficiently large codeword length n.

Definition 2 (DKI Coding/Target Identification/ISI Rates): The size of the codebook $M(n, R)$, the size of the target message set $K(n, \kappa)$, and the number of ISI taps $L(n, l)$ are sequences of monotonically non-decreasing functions in codeword length n with R, κ, and l denoting the DKI coding rate, target identification rate, and ISI rate, respectively. In particular, we consider the following functions in this paper:

$$M(n, R) = 2^{(n \log n)R}, K(n, \kappa) = 2^{\kappa \log n}, L(n, l) = 2^{l \log n}.$$

Definition 3 (Achievable Rate Region): The triple of rates (R, κ, l) is called *achievable* if for every $e_1, e_2 > 0$ and sufficiently large n, there exists an $(n, M(n, R), K(n, \kappa), L(n, l), e_1, e_2)$-ISI-Poisson DKI code. Then, the set of all achievable rate triples (R, κ, l) is referred to as the achievable rate region for $\mathcal{P}$.

Definition 4 (Capacity Region/Capacity): The operational DKI capacity region of the ISI-Poisson channel, $\mathcal{P}$, is defined as the closure of all achievable rate triples (R, κ, l). The supremum of the identification coding rate R is called the identification capacity and is denoted by $\mathbb{C}_{\text{DKI}}(\mathcal{P}, M, K, L)$.

IV. DKI CAPACITY OF THE POISSON CHANNEL WITH ISI

In this section, we first present our main results, i.e., lower and upper bounds on the achievable DKI rates for $\mathcal{P}$. Subsequently, we provide detailed proofs of these bounds.

A. MAIN RESULTS

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

Theorem 1: Consider the DTPC with ISI, $\mathcal{P}$, and assume that both the target message set and the number of ISI channel taps grow *sub-linearly* with the codeword length, i.e.,

$$K(n, \kappa) = 2^{\kappa \log n} \quad \text{and} \quad L(n, l) = 2^{l \log n},$$

respectively, where $\kappa \in [0, 1), l \in [0, 1/4)$, and $\kappa + 4l \in [0, 1)$. Then, the DKI capacity of $\mathcal{P}$ subject to average and peak molecule release rate constraints of the form $n^{-1} \sum_{t=1}^{n} c_{i,t} \leq P_{\text{ave}}$ and $0 \leq c_{i,t} \leq P_{\max}$, respectively,

[4] The error requirement as imposed by the DKI code definition applies to all possible choices of the set $\mathbb{K}$, i.e., $\binom{M}{K}$ cases; see [70, p. 140] for further details on K-identification codes.

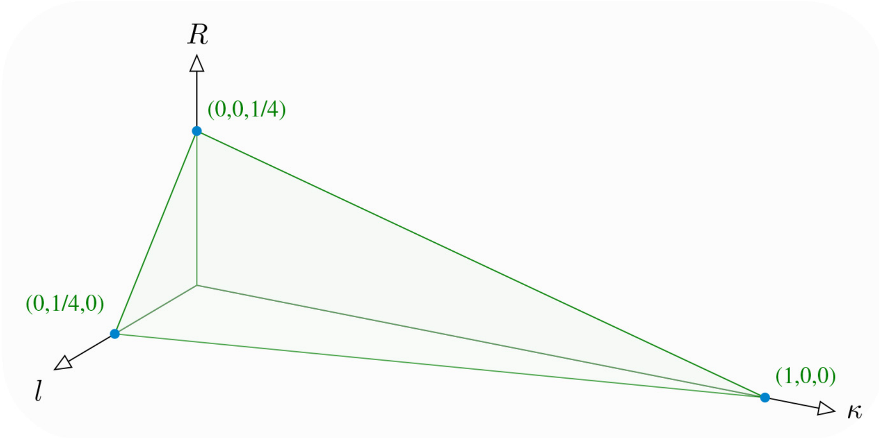

FIGURE 3. Illustration of achievable rate region for triple rates (κ, l, R) for the DTPC with ISI, $\mathcal{P}$. The DKI capacity region of $\mathcal{P}$ includes the entire depicted convex tetrahedron with vertices at $(\kappa = 0, l = 0, R = 0)$, $(\kappa = 1,0,0)$, $(0, l = 1/4, 0)$, and $(0,0,R = 1/4)$. The plane formed by the three extreme points marked in blue is characterized by $\kappa + 4l + 4R - 1 = 0$, which is derived by considering the equality case in the lower bound on the DKI capacity in Theorem 1. The subspace inscribed by this plane, and three other equations namely, $0 \leq \kappa < 1$, $0 \leq l < 1/4$, and $R \geq 0$, defines the entire rate region.

with $i \in [\![M]\!]$ and a codebook of *super-exponential* scale, i.e., $M(n, R) = 2^{(n \log n)R}$, is bounded by

$$\frac{1 - (\kappa + 4l)}{4} \leq \mathbb{C}_{\text{DKI}}\big(\mathcal{P}, M, K, L\big) \leq \frac{3}{2} + \kappa + l.$$

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

In the following, we highlight some insights obtained from Theorem 1 and its proof.

Rate region: Theorem 1 unveils the feasible region for three different rates, namely, the DKI achievable rate R, the ISI rate l, and the target identification rate κ. The geometric structure for all possible triples (κ, l, R) obtained from Theorem 1 is shown in Figure 3. A tetrahedron characterizes the feasible triple vectors (κ, l, R) for which a communication system can accomplish the task of K-identification for a DTPC with L ISI taps at a DKI achievable rate of at least R, where $K = 2^{\kappa \log n}$ and $L = 2^{l \log n}$.

The cross section of the tetrahedron with plane $R = 0$ determines the feasible region for rate pairs (κ, l); see Figure 4. This region can also be derived by the following argument: Since the target identification rate κ, the ISI rate l, and the lower bound on the DKI capacity given in Theorem 1 are non-negative rate values, we obtain $0 \leq \kappa < 1$, $0 \leq l < 1/4$, and $0 \leq \kappa + 4l \leq 1$. The first two constraints involving only κ and l, respectively, yield a rectangle having two of its corners at the origin $(0, 0)$ and $(1, 1/4)$, and the third joint constraint on κ and l, i.e., $0 \leq \kappa + 4l \leq 1$ excludes some of the rate pairs (κ, l) from such a rectangle for which the corresponding lower bound on the DKI capacity would be a strictly negative value. In addition, we note that more sophisticated coding schemes may result in an achievable region for rate pairs (κ, l) beyond the blue line in Figure 4.

Adopted decoder: Before going through the details of the achievability proof, we will present some insight into the proposed decoder. In particular, in the proposed achievable scheme, we adopt a distance decoder that decides in favour of a candidate codeword based on the distance between the received vector and expected value of the received vector if such a candidate codeword was really sent by the transmitter. More specifically, upon observing an output sequence $\mathbf{y}$ at the receiver, the decoder declares that message j was sent if the following condition is met

$$\Big| \, \|\mathbf{y} - \mathbb{E}\big(\mathbf{Y}|\mathbf{c}_j\big)\|^2 - \big\|\mathbf{y}\big\|_1 \Big| \leq \bar{n}\delta_n, \tag{10}$$

where δ_n is referred to as a decoding threshold and $\mathbf{c}_j = [c_{j,1}, \ldots, c_{j,n}]$ is the codeword associated with message j. Unlike the distance decoder used for Gaussian channels [46], which includes only the distance term $\|\mathbf{y} - \mathbb{E}(\mathbf{Y}|\mathbf{c}_j)\|$, the proposed decoder provided in (10) requires subtraction of an additional correction term $\big\|\mathbf{y}\big\|_1$. This correction term stems from the fact that the noise in the DTPC with ISI is signal (input codeword) dependent [11]. Therefore, the variance of $\|\mathbf{y} - \mathbb{E}(\mathbf{Y}|\mathbf{c}_j)\|$ depends on the adopted codeword $\mathbf{c}_j$ which implies that, unlike for the Gaussian channel, here the radius of the decoding region is not constant for all the codewords. To account for this fact, we include the correction term $\big\|\mathbf{y}\big\|_1$.

Corollary 1 (DI Capacity of the ISI-free DTPC): The lower and upper bounds on the DKI capacity of the DTPC with ISI, $\mathcal{P}$, for the asymptotic range of $l, \kappa \to 0$ converges to their maximum possible and minimum possible values, i.e., $1/4$ and $3/2$, respectively. Specifically, for $L = K = 1$, i.e., $l = \kappa = 0$, Theorem 1 recovers the results for the memoryless standard DI problem studied in [26], [45]:

$$\frac{1}{4} \leq \mathbb{C}_{\text{DKI}}\big(\mathcal{P}, M, K = 1, L = 1\big) \leq \frac{3}{2}. \tag{11}$$

Proof: The proof follows directly by substituting the extreme values of l and κ in the capacity results in Theorem 1. ∎

Remark 1: The lower bound on the DKI capacity in Theorem 1 suggests that by considering a dispersive communication system or allowing the receiver to identify its

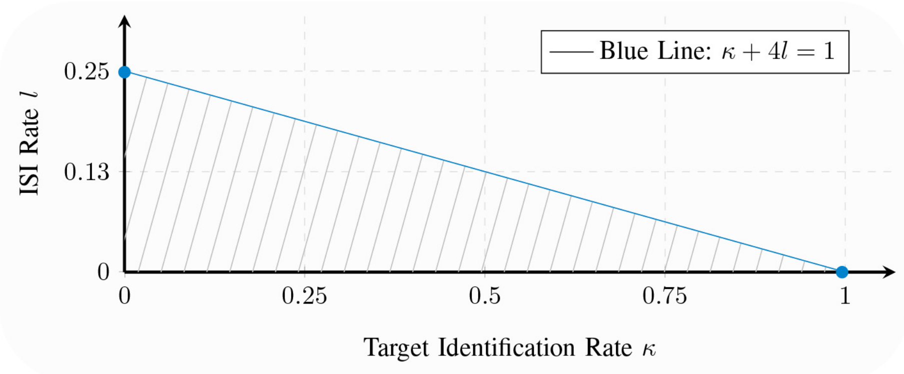

FIGURE 4. Illustration of achievable rate region for rate pairs (κ,l) for the DTPC with ISI, $\mathcal{P}$, which is the set of all points inscribed by the blue line (obtained by equating the lower bound given in Theorem 1 to zero) as well as the horizontal and vertical axes. The extreme points (0,0.25) and (1, 0) correspond to the DI with maximum possible number of ISI taps [1] and the DKI with the maximum size of the target message set, respectively. Walking on the blue line towards each of the extreme cases exemplifies the trade-off between the target identification rate and the ISI rate. The origin (0, 0) corresponds to the standard identification scheme (i.e., DI for the DTPC without ISI), where the set of target messages has only one element and the channel is memoryless [26], [27].

favourite message among a larger set of target messages, a penalty on the value of the lower bound is incurred. However, an increase in the exponent l for the number of ISI channel taps has a four times larger impact on the proposed lower bound. Another observation is that for a communication setting, where a fixed given lower bound on the identification performance in terms of the maximum achievable rate is required, there is a trade-off between the target identification rate and the ISI rate.

Corollary 2 (Effective Identification Rate): Let us assume that the physical length of the CIR interval is fixed and given by T_{cir}. Further, assume that the L ISI taps span the CIR interval, T_{cir}. Then, the following relation between the symbol duration, T_S, and the number of ISI taps, L holds:

$$T_S = T_{\text{cir}}/L = T_{\text{cir}} 2^{-l \log n}, \tag{12}$$

for some $l \in [0, 1/4)$ with $\kappa + 4l \in [0, 1)$. Now, let the effective identification rate, $\bar{R}_{\text{eff}}$, be defined as follows

$$\bar{R}_{\text{eff}} \overset{\text{def}}{=} \frac{\log M(n, R)}{nT_S} \tag{13}$$

(in bits/s). Then, the effective identification rate subject to average and peak molecule release rate constraints is bounded by

$$\frac{\big(1 - (\kappa + 4l)\big) n^l \log n}{4T_{\text{cir}}} \leq \bar{R}_{\text{eff}} \leq \frac{\big(3 + 2(\kappa + l)\ \big) n^l \log n}{2T_{\text{cir}}}. \tag{14}$$

Proof: The proof follows directly by substituting the capacity results in Theorem 1 into the definition of the effective rate and performing some mathematical simplifications. ■

Remark 2: Theorem 1 assumes that the number of ISI taps $L(n, l)$ scales sub-linearly in the codeword length n, i.e., $\sim 2^{l \log n}$. More specifically, L used in Theorem 1 may comprise the following three different cases:

1) *ISI-free, $L = 1$:* This case corresponds to an ISI-free setup, which is valid when the symbol duration is large ($T_S \geq T_{\text{cir}}$), and implies $L = 1$ and $l = 0$. Thereby, $\bar{R}_{\text{eff}}$ scales logarithmically with the codeword length n. This is in contrast to the transmission setting, where $\bar{R}_{\text{eff}}$ is independent of n (e.g., the well-known Shannon formula for the Gaussian channel). This result is known in the identification literature [31], [45].
2) *Constant $L > 1$:* When T_S is constant and $T_S < T_{\text{cir}}$, we have a constant $L > 1$, which implies $l \to 0$ as $n \to \infty$. Surprisingly, our capacity result in Theorem 1 reveals that the bounds for the DTPC with memory are in fact identical to those for the memoryless DTPC given in [45].
3) *Growing L:* Our capacity result shows that reliable identification is possible even when L scales with the codeword length as $\sim 2^{l \log n}$. Moreover, the impact of ISI rate l is reflected in the capacity lower and upper bounds in Theorem 1, where the bounds respectively decrease and increase in l. While the upper bound on $\bar{R}_{\text{eff}}$ increases in l, too, the lower bound in (14) suggests a trade-off in terms of l, which is investigated in Corollary 3.

Corollary 3 (Optimum ISI Rate): The lower bound given in Corollary 2 is maximized for the following ISI rate

$$l_{\max}(n) = \frac{1}{4}\Big(1 - \kappa - \frac{4}{\ln n}\Big), \tag{15}$$

where $n \in \mathbb{N}$. Moreover, the maximum ISI rate, $l_{\max}$, provided in (15) yields the following lower bound on the effective identification rate, $\bar{R}_{\text{eff}}(n)$:

$$\bar{R}_{\text{eff}}(n) \geq \frac{\log e}{eT_{\text{cir}}} \cdot n^{\frac{1}{4}(1-\kappa)}. \tag{16}$$

Thereby, the normalized effective identification rate is lower bounded as follows

$$\liminf_{n\to\infty} \frac{\bar{R}_{\text{eff}}(n)}{n^{\frac{1}{4}(1-\kappa)}} \geq \frac{\log e}{eT_{\text{cir}}}. \tag{17}$$

Proof: The proof follows from differentiating the lower bound in Corollary 2 with respect to l and equating the result to zero. ■

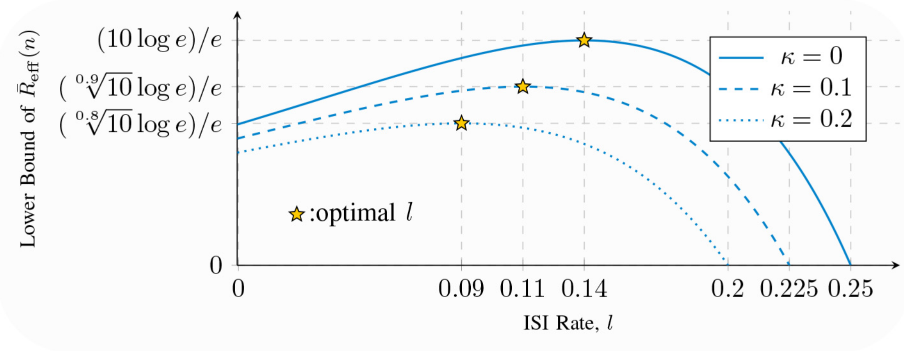

FIGURE 5. Illustration of the lower bound on the effective identification rate $\bar{R}_{\text{eff}}$ provided in (14) for target identification rates $\kappa = 0, 0.1, 0.2$ and codeword length $n = 10^4$. The ISI rate l that yields the maximum value for each value of l is marked by a yellow star and coincides with the optimal $l_{\max}$ provided in (15).

The effective identification rate $\bar{R}_{\text{eff}}$ [bits/s] in (13) is the product of two terms, namely the identification rate per symbol $\log M(n, R)/n$ [bits/symbol] (which decreases with l for the lower bound provided in Theorem 1) and the symbol rate $1/T_S$ [symbol/s] (which increases with l). The above corollary reveals that in order to maximize $\bar{R}_{\text{eff}}$, it is optimal to set the trade-off for l such that the identification rate, i.e.,

$$\frac{\log M(n, R)}{n} = \frac{\left(1 - (\kappa + 4l_{\max})\right) \log n}{4} = \log e, \qquad (18)$$

becomes independent of n but the symbol rate scales polynomially with fractional exponent in n, i.e.,

$$1/T_S = n^{\frac{1}{4}(1-\kappa)}/T_{\text{cir}} = 2^{O(\log n)}. \qquad (19)$$

As a result, in contrast to the typical transmission setting, where the effective rate is independent of n, here, the effective identification rate $\bar{R}_{\text{eff}}$ for the optimal l grows sub-linearly in n. Moreover, the sub-linear increase of the effective rate in n is faster compared to the typical scenario, where T_S (and hence L) is fixed and $l = 0$, and the effective rate, i.e., $((1 - \kappa) \log n)\ /4$ increases logarithmically in n. Fig. 5 shows the lower bound on the effective identification rate $\bar{R}_{\text{eff}}$ in (14) for target identification rates $\kappa = 0, 0.1, 0.2$ and codeword length $n = 10^4$. Note that n should be large since our capacity results are valid asymptotically. As expected, each curve in Fig. 5 has a unique maximum at an ISI rate l that coincides with $l_{\max}$ in (15).

In addition, based on Theorem 1, we can distinguish the following three cases in terms of K:

- *DI, $K = 1$:* This case accounts for a standard identification setup ($\kappa = 0$), i.e., the degenerate case where the target message set has only one element, namely, $\mathbb{K} = \{i\}$, with $i \in [\![M]\!]$ and $|\mathbb{K}| = K = 1$. Therefore, the identification setup in the deterministic [47] and randomized regimes [31] can be regarded as a special case of K-identification considered in this paper.
- *Constant $K > 1$:* Constant $K > 1$ implies $\kappa \to 0$ as $n \to \infty$. Our DKI capacity result in Theorem 1 reveals that the bounds on the DKI achievable rate are identical to those for $K = 1$.
- *Growing K:* The DKI capacity bounds in Theorem 1 suggest that reliable identification is possible even when K scales with the codeword length as $\sim 2^{\kappa \log n}$, for some $\kappa \in [0, 1)$ and $\kappa + 4l \in [0, 1)$.

In the following, we provide the proof of Theorem 1, namely the achievability proof in Section IV-B and the converse proof in Section IV-C.

B. LOWER BOUND (ACHIEVABILITY PROOF)

The achievability proof consists of the following two steps.

- *Step 1:* We propose a codebook construction and derive an analytical lower bound on the corresponding codebook size using inequalities for the sphere packing density.
- *Step 2:* We prove that this codebook leads to an *achievable* rate by proposing a decoder and showing that the corresponding type I and type II error probabilities vanish as $n \to \infty$.

A DKI code for the DTPC, $\mathcal{P}$, is constructed as follows.

Input constraint adaptation: We restrict ourselves to codewords that meet the condition $0 \leq c_{i,t} \leq P_{\text{ave}}, \forall\, i \in [\![M]\!], \forall\, t \in [\![n]\!]$, which ensures that both constraints in (7) are met for $P_{\text{ave}} > P_{\max}$ and $P_{\text{ave}} \leq P_{\max}$:

1) *$P_{\text{ave}} > P_{\max}$* : In this case, the condition $0 \leq c_{i,t} \leq P_{\max}, \forall\, i \in [\![M]\!], \forall\, t \in [\![n]\!]$, yields $n^{-1} \sum_{t=1}^{n} c_{i,t} \leq P_{\text{ave}}$. In this case, the average constraint trivially holds and we exclude this scenario from the analysis.
2) *$P_{\text{ave}} \leq P_{\max}$* : Then, the condition $0 \leq c_{i,t} \leq P_{\text{ave}}, \forall\, i \in [\![M]\!], \forall\, t \in [\![n]\!]$, implies both $0 \leq c_{i,t} \leq P_{\max}$ and $n^{-1} \sum_{t=1}^{n} c_{i,t} \leq P_{\text{ave}}$.

Thus, for the construction of the codebook in the next steps, we only require that $0 \leq c_{i,t} \leq P_{\text{ave}}, \forall\, i \in [\![M]\!], \forall\, t \in [\![n]\!]$.

Convoluted codebook construction: In the following, instead of directly constructing the original codebook $\mathcal{C} = \{\mathbf{c}_i\} \subset \mathbb{R}_+^n$, with $i \in [\![M]\!]$, we present a construction of a codebook called convoluted codebook and show that the original codebook can be uniquely reconstructed for a convoluted codebook. In particular, the convoluted codebook is denoted by $\mathcal{C}^{\rho} = \{\mathbf{c}_i^{\rho}\} \subset \mathbb{R}_+^n$, with $i \in [\![M]\!]$, where each $\mathbf{c}_i^{\boldsymbol{\rho}} \triangleq (\mathbf{c}_{i,1}^{\rho}, \ldots, \mathbf{c}_{i,n}^{\rho})$ is referred to as a convoluted codeword whose symbols are formed as a linear combination (convolution) of the L most recent symbols of codeword $\mathbf{c}_i \triangleq (\mathbf{c}_{i,1}, \ldots, \mathbf{c}_{i,n})$ and CIR vector $\boldsymbol{\rho}$, i.e.,

$$c_{i,t}^{\boldsymbol{\rho}} \triangleq \sum_{l=0}^{L-1} \rho_l c_{i,t-l}. \tag{20}$$

Observe that the convoluted symbol $c_{i,t}^{\boldsymbol{\rho}}$ represents the expected value of the signal observed at the receiver after the release of $c_{i,t}$ molecules by the transmitter. The proposed convoluted codebook construction is motivated by the structure of the ISI channel and the choice of the distance decoder given in (10). More specifically, the term $\mathbb{E}(\mathbf{Y} \,|\, \mathbf{c}_j)$ for $j \in [\![M]\!]$ given in (10) is the center of the distance decoder and includes the convoluted codeword, i.e., $\mathbf{c}_j^{\boldsymbol{\rho}}$.

In order to use the convoluted codebook, we have to show that the original codewords $\mathbf{c}_i$ can be uniquely derived from the convoluted codewords $\mathbf{c}_i^{\boldsymbol{\rho}}$, i.e., there is a one-to-one mapping between the convoluted and the original codebooks. To show this, let us first define the set of feasible original and convoluted codewords, respectively, as:

$$\begin{aligned}\mathbb{C}_0 &= \mathbb{Q}_{\mathbf{0}}(n, P_{\text{ave}}) \\ &\triangleq \left\{\mathbf{c}_i \in \mathbb{R}^n : \; 0 \le c_{i,t} \le P_{\text{ave}}, \forall\, i \in [\![M]\!], \forall\, t \in [\![n]\!]\right\}\end{aligned} \tag{21}$$

$$\mathbb{C}_0^{\boldsymbol{\rho}} \triangleq \left\{\mathbf{c}_i^{\boldsymbol{\rho}} \in \mathbb{R}^n : \; c_{i,t}^{\boldsymbol{\rho}} \triangleq \sum_{l=0}^{L-1} \rho_l c_{i,t-l}, \;\; \mathbf{c}_i \in \mathbb{C}_0, \forall\, i \in [\![M]\!]\right\}. \tag{22}$$

Unfortunately, unlike the feasible set of the original codewords $\mathbb{C}_0$, the feasible set of the convoluted codewords $\mathbb{C}_0^{\boldsymbol{\rho}}$ lacks the simple structure and geometry needed for the calculation of the volume and rate analysis. To cope with this issue, we target a *subset* of $\mathbb{C}_0^{\boldsymbol{\rho}}$ that enjoys a suitable structure with well-known geometry and analytic volume formula, namely the following hyper cube:

$$\begin{aligned}&\mathbb{Q}_{\mathbf{0}}\left(n, \bar{P}_{\text{ave}}\right) \\ &\quad = \{\mathbf{c}_i^{\boldsymbol{\rho}} : \; 0 \le c_{i,t}^{\boldsymbol{\rho}} \le \bar{P}_{\text{ave}}, \forall\, i \in [\![M]\!], \forall\, t \in [\![n]\!]\},\end{aligned} \tag{23}$$

where

$$\bar{P}_{\text{ave}} \triangleq \min_{\substack{i \in [\![M]\!]; \\ \mathbf{c}_i^{\boldsymbol{\rho}} \in \mathbb{C}_1 \cap \mathbb{C}_2^c}} \; \min_{\substack{t \in [\![n]\!]; \\ t-L+1 \le \bar{t} \le t}} c_{i,t}^{\boldsymbol{\rho}}, \tag{24}$$

where $\bar{t}$ is a specific symbol index for which the corresponding input symbol yields a non-zero number of released molecules from the transmitter, i.e., $\lfloor T_R c_{i,\bar{t}} \rfloor \ge 1$. Moreover, sets $\mathbb{C}_1$ and $\mathbb{C}_2$ are given by

$$\begin{aligned}\mathbb{C}_1 &= \mathbb{Q}_{\mathbf{0}}\left(n, P'_{\text{ave}}\right) \\ &\triangleq \left\{\mathbf{c}_i^{\boldsymbol{\rho}} \in \mathbb{R}^n \; : \; 0 \le c_{i,t}^{\boldsymbol{\rho}} \le P'_{\text{ave}}, \forall\, i \in [\![M]\!], \forall\, t \in [\![n]\!]\right\}, \\ \mathbb{C}_2 &= \left\{\mathbf{c}_i^{\boldsymbol{\rho}} \in \mathbb{R}^n \; : \; c_{i,t} \ge 0, \forall\, i \in [\![M]\!], \forall\, t \in [\![n]\!]\right\},\end{aligned} \tag{25}$$

where $P'_{\text{ave}} \triangleq \rho_0 P_{\text{ave}}$.

Next, we have to show that the volume of $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$ is non-zero (i.e., $\bar{P}_{\text{ave}}$ is bounded away from zero) and $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}}) \subseteq \mathbb{C}_0^{\boldsymbol{\rho}}$. The former follows from the fact that $\bar{P}_{\text{ave}}$ tends to zero only if all symbols of at least one of the original codewords are arbitrary close to zero. Such a single all-zero codeword can be excluded without affecting the rate analysis. To prove $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}}) \subseteq \mathbb{C}_0^{\boldsymbol{\rho}}$, we show that the original codeword $\mathbf{c}_i$ obtained from $\mathbf{c}_i^{\boldsymbol{\rho}} \in \mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$ belongs to $\mathbb{C}_0$, namely the extracted original symbols must meet $0 \le c_{i,t} \le P_{\text{ave}}$. We first show that $c_{i,t} \ge 0$ holds via contradiction. In other words, we assume $\mathbf{c}_i^{\boldsymbol{\rho}} \in \mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$ but the corresponding original codeword meets $\mathbf{c}_i \in \mathbb{C}_2^c$. This already contradicts the fact that $\bar{P}_{\text{ave}} > 0$, see (24). To show $c_{i,t} \le P_{\text{ave}}$, we use the following chain of inequalities assuming $\mathbf{c}_i^{\boldsymbol{\rho}} \in \mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$:

$$\begin{aligned}\rho_0 c_{i,1} &\le \bar{P}_{\text{ave}} \le P'_{\text{ave}} \\ \rho_0 c_{i,2} + \rho_1 c_{i,1} &\le \bar{P}_{\text{ave}} \le P'_{\text{ave}} \\ \rho_0 c_{i,3} + \rho_1 c_{i,2} + \rho_2 c_{i,1} &\le \bar{P}_{\text{ave}} \le P'_{\text{ave}} \\ &\vdots \\ \rho_0 c_{i,n} + \rho_1 c_{i,n-1} + \ldots + \rho_{L-1} c_{i,n-L+1} &\le \bar{P}_{\text{ave}} \le P'_{\text{ave}},\end{aligned} \tag{26}$$

where $\bar{P}_{\text{ave}} \le P'_{\text{ave}}$ holds since $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}}) \subset \mathbb{C}_1$, see (24). The above inequalities can be rewritten as follows

$$\begin{aligned}c_{i,1} &\le P'_{\text{ave}}/\rho_0 = P_{\text{ave}} \\ c_{i,2} &\le \frac{P'_{\text{ave}} - \rho_0 c_{i,1}}{\rho_0} \le P'_{\text{ave}}/\rho_0 = P_{\text{ave}} \\ &\vdots \\ c_{i,n} &\le \frac{P'_{\text{ave}} - \sum_{t=1}^{K-1} \rho_l c_{i,t-l}}{\rho_0} \le P'_{\text{ave}}/\rho_0 = P_{\text{ave}},\end{aligned} \tag{27}$$

where we used the fact that $c_{i,t} \ge 0$. Hence, condition $\|\mathbf{c}_i\|_\infty \le P_{\text{ave}}$ holds for the extracted original codewords. In summary, we showed that for convoluted codewords $\mathbf{c}_i^{\boldsymbol{\rho}} \in \mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$, there is a unique feasible original codeword $\mathbf{c}_i \in \mathbb{Q}_{\mathbf{0}}(n, P_{\text{ave}})$. Therefore, the rate analysis of the convoluted codebook is also valid for the original codebook.

Calculation of the codebook size/rate: We use a packing arrangement of non-overlapping hyper spheres of radius $r_0 = \sqrt{n\epsilon_n}$ in a hyper cube with edge length $\bar{P}_{\text{ave}}$, where

$$\epsilon_n = \frac{3a}{4n^{\frac{1}{2}(1-(b+\kappa+4l))}}, \tag{28}$$

and $a > 0$ is a non-vanishing fixed constant, $0 < b < 1$ is an arbitrarily small constant, and $0 \le \kappa + 4l < 1$.

Let $\mathscr{S}$ denote a sphere packing, i.e., an arrangement of M non-overlapping spheres $\mathcal{S}_{\mathbf{c}_i^{\boldsymbol{\rho}}}(n, r_0)$, $i \in [\![M]\!]$, that are packed inside the larger cube $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$ with edge length $\bar{P}_{\text{ave}}$, see Figure 6. As opposed to standard sphere packing coding techniques [71], the spheres are not necessarily entirely contained within the cube. That is, we only require that the centers of the spheres are inside $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$, the spheres are disjoint from each other, and they have a non-empty intersection with $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$. The packing density $\Delta_n(\mathscr{S})$

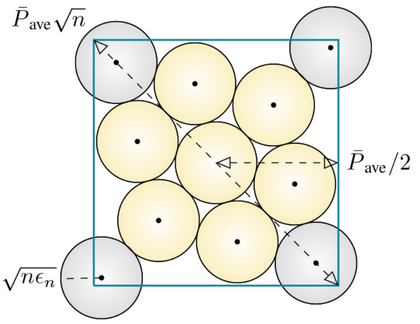

FIGURE 6. Illustration of a saturated sphere packing inside a cube, where small spheres of radius $r_0 = \sqrt{n\epsilon_n}$ cover a larger cube. Dark gray colored spheres are not entirely contained within the larger cube, and yet they contribute to the packing arrangement. As we assign a codeword to each sphere center, the 1-norm and arithmetic mean of a codeword are bounded by $\bar{P}_{\text{ave}}$ as required.

is defined as the ratio of the saturated packing volume to the cube volume $\text{Vol}(\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}}))$, i.e.,

$$\Delta_n\big(\mathscr{S}\big) \triangleq \frac{\text{Vol}\Big(\bigcup_{i=1}^{M} \mathcal{S}_{\mathbf{c}_i^{\rho}}(n, r_0)\Big)}{\text{Vol}\Big(\mathbb{Q}_{\mathbf{0}}\Big(n, \bar{P}_{\text{ave}}\Big)\Big)}. \tag{29}$$

Sphere packing $\mathscr{S}$ 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 for the Minkowski–Hlawka theorem for saturated packings [71]. Specifically, consider the saturated packing arrangement of

$$\bigcup_{i=1}^{M(n,R)} \mathcal{S}_{\mathbf{c}_i^{\rho}}\big(n, \sqrt{n\epsilon_n}\big) \tag{30}$$

spheres with radius $r_0 = \sqrt{n\epsilon_n}$ embedded within cube $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$. Then, for such an arrangement, we have the following lower [72, Lemma 2.1] and upper bounds [71, eq. (45)] on the packing density

$$2^{-n} \leq \Delta_n\big(\mathscr{S}\big) \leq 2^{-0.599n}. \tag{31}$$

In particular, in our subsequent analysis, we employ the lower bound given in (31), which can be proved as follows: For the saturated packing arrangement given in (30), there cannot be a point in the larger cube $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$ with a distance of more than $2r_0$ from all sphere centers. Otherwise, a new sphere could be added which contradicts the assumption that the union of $M(n, R)$ spheres with radius $\sqrt{n\epsilon_n}$ is saturated. Now, if we double the radius of each sphere, the spheres with radius $2r_0$ cover thoroughly the entire volume of $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$, that is, each point inside the hyper cube $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$ belongs to at least one of the small spheres. In general, the volume of a hyper sphere of radius r is given by [71, eq. (16)]

$$\text{Vol}\big(\mathcal{S}_{\mathbf{x}}(n, r)\big) = \frac{\pi^{\frac{n}{2}}}{\Gamma\Big(\frac{n}{2}+1\Big)} \cdot r^n. \tag{32}$$

Hence, if the radius of the small spheres is doubled, the volume of $\bigcup_{i=1}^{M(n,R)} \mathcal{S}_{\mathbf{c}_i^{\rho}}(n, \sqrt{n\epsilon_n})$ is increased by 2^n. Since the spheres with radius $2r_0$ cover $\mathbb{Q}_{\mathbf{0}}(n, \bar{P}_{\text{ave}})$, it follows that the original r_0-radius packing[5] has a density of at least 2^{-n}. We assign a convoluted codeword to the center $\mathbf{c}_i^{\rho}$ of each small hyper sphere. The convoluted codewords satisfy the input constraint as $0 \leq c_{i,t}^{\rho} \leq P'_{\text{ave}}, \forall t \in [\![n]\!], \forall i \in [\![M]\!]$, which is equivalent to

$$\left\|\mathbf{c}_i^{\rho}\right\|_{\infty} \leq \bar{P}_{\text{ave}}. \tag{33}$$

Since the volume of each sphere is equal to $\text{Vol}(\mathcal{S}_{\mathbf{c}_1^{\rho}}(n, r_0))$ and the centers of all spheres lie inside the cube, the total number of spheres is bounded from below by

$$\begin{aligned} M &= \frac{\text{Vol}\Big(\bigcup_{i=1}^{M} \mathcal{S}_{\mathbf{c}_i^{\rho}}(n, r_0)\Big)}{\text{Vol}\Big(\mathcal{S}_{\mathbf{c}_1^{\rho}}(n, r_0)\Big)} = \frac{\Delta_n\big(\mathscr{S}\big) \cdot \text{Vol}\Big(\mathbb{Q}_{\mathbf{0}}\Big(n, \bar{P}_{\text{ave}}\Big)\Big)}{\text{Vol}\Big(\mathcal{S}_{\mathbf{c}_1^{\rho}}(n, r_0)\Big)} \\ &\geq 2^{-n} \cdot \frac{P_{\text{ave}}^n}{\text{Vol}\Big(\mathcal{S}_{\mathbf{c}_1^{\rho}}(n, r_0)\Big)}, \end{aligned} \tag{34}$$

where the inequality holds by (31). The bound in (34) can be written as follows

$$\begin{aligned} \log M &\geq \log\Big(\bar{P}_{\text{ave}}^n \,/\, \text{Vol}\big(\mathcal{S}_{\mathbf{c}_1^{\rho}}(n, r_0)\big)\Big) - n \\ &\geq n \log\Big(\bar{P}_{\text{ave}}/\sqrt{\pi}\, r_0\Big) + \log\big(\Gamma\big(n/2+1\big)\big) - n, \end{aligned} \tag{35}$$

where the last inequality exploits (32). The above bound can be further simplified as follows

$$\log M \geq n \log\Big(\bar{P}_{\text{ave}}/\sqrt{\pi}\, r_0\Big) + \log\big(\lfloor n/2 \rfloor !\big) - n, \tag{36}$$

where the equality exploits the following relation:

$$\Gamma\big(n/2+1\big) \overset{(a)}{=} \frac{n}{2}\Gamma\big(n/2\big) \overset{(b)}{\geq} \lfloor n/2 \rfloor \Gamma\big(\lfloor n/2 \rfloor\big) \overset{(c)}{\triangleq} \lfloor n/2 \rfloor !. \tag{37}$$

In the above equation, (a) holds by the recurrence relation of the Gamma function [73] for real $n/2$, (b) follows from $\lfloor n/2 \rfloor \leq n/2$, the monotonicity of the Gamma function [73] for $\lfloor n/2 \rfloor \geq 1.46 \equiv n \geq 4$, and (c) holds since for positive integer $\lfloor n/2 \rfloor$, we have $\Gamma(\lfloor n/2 \rfloor) = (\lfloor n/2 \rfloor - 1)!$, cf. [73]. Next, we proceed to simplify the factorial term given in (36). To this end, we exploit *Stirling's approximation*, i.e., $\log n! = n \log n - n \log e + o(n)$ [74, p. 52] with the substitution of $n = \lfloor n/2 \rfloor$, where $\lfloor n/2 \rfloor \in \mathbb{Z}$. Thereby, we obtain

$$\begin{aligned} \log M \geq\; & n \log \bar{P}_{\text{ave}} - n \log r_0 + \lfloor n/2 \rfloor \log\big(\lfloor n/2 \rfloor\big) \\ & - \lfloor n/2 \rfloor \log e + o\big(\lfloor n/2 \rfloor\big) - n, \end{aligned} \tag{38}$$

[5] We note that the proposed proof of the lower bound in (31) is non-constructive in the sense that, while the existence of the respective saturated packing is proved, no systematic construction method is provided.

Therefore, for $r_0 = \sqrt{n\epsilon_n} = \sqrt{a' n^{\frac{1+b+\kappa+4l}{4}}}$, where $a' \triangleq 3a/4$, we have

$$\begin{aligned}\log M &\overset{(a)}{\geq} n \log \frac{\bar{P}_{\text{ave}}}{\sqrt{\pi a'}} - \left(\frac{1+b+\kappa+4l}{4}\right) n \log \sqrt{an} \\ &\quad + (n/2-1)\log\left(n/2-1\right) - \lfloor n/2 \rfloor \log e \\ &\quad + o\left(n/2-1\right) - n \\ &\overset{(b)}{\geq} n \log \frac{\bar{P}_{\text{ave}}}{\sqrt{\pi a'}} - \left(\frac{1+b+\kappa+4l}{4}\right) n \log \sqrt{an} \\ &\quad + \frac{1}{2} n \log n - 2n - \log n - \frac{n}{2}\log e + o\left(n/2\right) \\ &= \left(\frac{1-(b+\kappa+4l)}{4}\right) n \log n \\ &\quad + n\left(\log \bar{P}_{\text{ave}}/\sqrt{\pi a' e}\right) + O(n), \end{aligned} \tag{39}$$

where (a) follows from $\left\lfloor \frac{n}{2} \right\rfloor > \frac{n}{2} - 1$ and (b) holds since $\log(t-1) \geq \log t - 1$ for $t \geq 2$ and $\left\lfloor \frac{n}{2} \right\rfloor \leq \frac{n}{2}$ for integer n. Observe that the dominant term in (39) is of order $n \log n$. Hence, to obtain a finite value for the lower bound on the rate, R, (39) reveals that the scaling law of M is $2^{(n \log n)R}$. Therefore, we obtain

$$\begin{aligned} R \geq \frac{1}{n \log n}&\left[\left(\frac{1-(b+\kappa+4l)}{4}\right) n \log n \right. \\ &\left. + n\left(\log \frac{\bar{P}_{\text{ave}}}{\sqrt{\pi a' e}}\right) + O(n)\right], \end{aligned} \tag{40}$$

which tends to $(1-(\kappa+4l))/4$ when $n \to \infty$ and $b \to 0$.

Encoder: Given message $i \in [\![M]\!]$, transmit $\mathbf{x} = \mathbf{c}_i$.

Proposed decoder: In order to analyze the error performance of the proposed codebook, we need to adopt a decoder which is introduced next. Before we proceed, for the sake of a concise analysis, we introduce the following conventions. Let:

- $Y_t(i) \sim \text{Pois}(c_{i,t}^{\boldsymbol{\rho}} + \lambda)$ denote the channel output at time t given that $\mathbf{x} = \mathbf{c}_i$.
- The output vector is defined as the vector of symbols, i.e., $\mathbf{Y}(i) = (Y_1(i), \ldots, Y_{\bar{n}}(i))$.
- $\bar{y}_t(i) \triangleq y_t(i) - (c_{i,t}^{\boldsymbol{\rho}} + \lambda)$, where $y_t(i)$ is a realization of $Y_t(i)$.

Furthermore, let

$$\delta_n \triangleq 4\epsilon_n/3 = 4a/\left(3n^{\frac{1}{2}(1-(b+\kappa+4l))}\right), \tag{41}$$

where $0 < b < 1$ is an arbitrarily small constant and $0 \leq \kappa + 4l < 1$ with κ and l being the identification target rate and the ISI rate, respectively. To identify whether a message $j \in [\![M]\!]$ was sent, the decoder checks whether the channel output $\mathbf{y}$ belongs to the following decoding set,

$$\mathbb{T}_j = \left\{ \mathbf{y} \in \mathbb{N}_0^{\bar{n}} : \left| T\left(\mathbf{y}, \mathbf{c}_j\right) \right| \leq \delta_n \right\}, \tag{42}$$

where

$$T\left(\mathbf{y}; \mathbf{c}_j\right) = \frac{1}{\bar{n}} \sum_{t=1}^{\bar{n}} \left(y_t - \left(c_{j,t}^{\boldsymbol{\rho}} + \lambda\right)\right)^2 - y_t, \tag{43}$$

is referred to as the *decoding metric* evaluated for observation vector $\mathbf{y}$ and codeword $\mathbf{c}_j$. Finally, let $e_1, e_2 > 0$ and $\zeta_0, \zeta_0', \zeta_1, \zeta_1' > 0$ be arbitrarily small constants.

Error analysis: In the following, we exploit Chebyshev's inequality in order to establish upper bounds for the type I and type II error probabilities.

Type I error analysis: Consider the type I errors, i.e., the transmitter sends $\mathbf{c}_i$, yet $\mathbf{Y} \notin \mathbb{T}_i$. For every $i \in [\![M]\!]$, the type I error probability is bounded as

$$\begin{aligned} P_{e,1}\left(i, \mathbb{K}\right) &= \Pr\left(\mathbf{Y}(i) \in \mathscr{T}_{\mathbb{K}}^c\right) \\ &= \Pr\left(\mathbf{Y}(i) \in \left(\bigcup_{j \in \mathbb{K}} \mathbb{T}_j\right)^c\right) \\ &\overset{(a)}{=} \Pr\left(\mathbf{Y}(i) \in \bigcap_{j \in \mathbb{K}} \mathbb{T}_j^c\right) \\ &\overset{(b)}{\leq} \Pr\left(\mathbf{Y}(i) \in \mathbb{T}_i^c\right) \\ &= \Pr\left(\left|T\left(\mathbf{Y}(i), \mathbf{c}_j\right)\right| > \delta_n\right), \end{aligned} \tag{44}$$

where (a) holds by *De Morgan*'s law for a finite number of union of sets [75], i.e., $(\bigcup_{j\in\mathbb{K}} \mathbb{T}_j)^c = \bigcap_{j\in\mathbb{K}} \mathbb{T}_j^c$ and (b) follows since $\bigcap_{j\in\mathbb{K}} \mathbb{T}_j^c \subset \mathbb{T}_j$.

In order to bound $P_{e,1}(i, \mathbb{K})$ in (44), we apply Chebyshev's inequality, namely

$$\begin{aligned} \Pr&\left(\left|T\left(\mathbf{Y}(i), \mathbf{c}_i\right) - \mathbb{E}\left[T\left(\mathbf{Y}(i), \mathbf{c}_i\right)\right]\right| > \delta_n\right) \\ &\leq \frac{\text{Var}\left[T\left(\mathbf{Y}(i), \mathbf{c}_i\right)\right]}{\delta_n^2}. \end{aligned} \tag{45}$$

First, we calculate the expectation of the *decoding metric* as follows

$$\begin{aligned} &\mathbb{E}\left[T\left(\mathbf{Y}(i), \mathbf{c}_i\right)\right] \\ &\overset{(a)}{=} \frac{1}{\bar{n}} \sum_{t=1}^{\bar{n}} \mathbb{E}\left[\left(Y_t(i) - \left(c_{i,t}^{\boldsymbol{\rho}} + \lambda\right)\right)^2\right] - \mathbb{E}\left[Y_t(i)\right] \\ &\overset{(b)}{=} \frac{1}{\bar{n}} \sum_{t=1}^{\bar{n}} \text{Var}\left[Y_t(i)\right] - \left(c_{i,t}^{\boldsymbol{\rho}} + \lambda\right) \\ &\overset{(c)}{=} \frac{1}{\bar{n}} \sum_{t=1}^{\bar{n}} \left(c_{i,t}^{\boldsymbol{\rho}} + \lambda\right) - \left(c_{i,t}^{\boldsymbol{\rho}} + \lambda\right) = 0, \end{aligned} \tag{46}$$

where (a) follows from the linearity of expectation, (b) holds since $\mathbb{E}[(Y_t(i) - \mathbb{E}[Y_t(i)])^2] = \text{Var}[Y_t(i)]$ and $\mathbb{E}[Y_t(i)] = c_{i,t}^{\boldsymbol{\rho}} + \lambda$, and (c) follows since $\text{Var}[Y_t(i)] = \mathbb{E}[Y_t(i)] = c_{i,t}^{\boldsymbol{\rho}} + \lambda$. Second, in order to compute the upper bound in (45), we proceed to compute the variance of the *decoding metric*. Let us define

$$\psi_{\text{Var}} \triangleq \sum_{t=1}^{\bar{n}} \text{Var}\left[\bar{Y}_t^2(i) - Y_t(i)\right]. \tag{47}$$

Since, conditioned on $\mathbf{c}_i$, the channel outputs conditioned on the L most recent input symbols are uncorrelated, we obtain

$$\mathrm{Var}\Big[T\big(\mathbf{Y}(i);\mathbf{c}_i\big)\Big] = \frac{\psi_{\mathrm{Var}}}{\bar{n}^2}. \tag{48}$$

Next, we proceed to establish an upper bound $\psi_{\mathrm{Var}}^{\mathrm{UB}}$ for ψ_{Var}. To this end, let us define

$$\begin{aligned}\psi_{\mathrm{Var}} &\triangleq \mathrm{Var}\big[\bar{Y}_t^2(i) - Y_t(i)\big]\\ &\overset{(a)}{=} \mathrm{Var}\Big[Y_t^2(i) - \big(2(c_{i,t}^{\boldsymbol{\rho}} + \lambda) + 1\big)Y_t(i)\Big]\\ &\overset{(b)}{=} \mathrm{Var}\big[Y_t^2(i)\big] + \big(2(c_{i,t}^{\boldsymbol{\rho}} + \lambda) + 1\big)^2 \mathrm{Var}\big[Y_t(i)\big]\\ &\quad - \big(4(c_{i,t}^{\boldsymbol{\rho}} + \lambda) + 2\big)\,\mathrm{Cov}\Big[Y_t^2(i), Y_t(i)\Big],\end{aligned} \tag{49}$$

where (a) holds since $\bar{Y}_t(i) \triangleq Y_t(i) - (c_{i,t}^{\boldsymbol{\rho}} + \lambda)$ and the decomposition in (b) follows from the following identity for constants a and b:

$$\mathrm{Var}\big[aX - bY\big] = a^2\mathrm{Var}[X] + b^2\mathrm{Var}[Y] - 2ab\,\mathrm{Cov}[X, Y]. \tag{50}$$

Next, let us define

$$\begin{aligned}\psi_{\mathrm{Cov}} &\triangleq \big(4(\bar{L}P_{\mathrm{avg}} + \lambda) + 2\big)\sqrt{\exp(8/\lambda)\big(\bar{L}P_{\mathrm{avg}} + \lambda\big)}\\ &= O\Big(L^{3/2}\Big),\end{aligned} \tag{51}$$

with $\bar{L} \triangleq LT_R$. Now, we proceed to establish an upper bound on (49) as follows

$$\begin{aligned}\psi_{\mathrm{Var}} &\overset{(a)}{\leq} \mathbb{E}\Big[Y_t^4(i)\Big] + \big(2(c_{i,t}^{\boldsymbol{\rho}} + \lambda) + 1\big)^2 (c_{i,t}^{\boldsymbol{\rho}} + \lambda)\\ &\quad + \big(4(c_{i,t}^{\boldsymbol{\rho}} + \lambda) + 2\big)\sqrt{\mathbb{E}\Big[Y_t^4(i)\Big]\mathrm{Var}\big[Y_t(i)\big]}\\ &\overset{(b)}{\leq} \big(\bar{L}P_{\mathrm{avg}} + \lambda\big)^4 \exp(8/\lambda) + \big(2(\bar{L}P_{\mathrm{avg}} + \lambda) + 1\big)^2 + \psi_{\mathrm{Cov}},\end{aligned} \tag{52}$$

where (a) follows from the triangle inequality, i.e., $\alpha - \beta \leq |\alpha - \beta| \leq |\alpha| + |\beta|$ for real a and b, $\mathrm{Var}[Y_t^2(i)] \leq \mathbb{E}[Y_t^4(i)], \mathrm{Var}[Y_t(i)] = c_{i,t}^{\boldsymbol{\rho}} + \lambda$, and $\mathrm{Cov}[X, Y] \leq \sqrt{\mathrm{Var}[X] \cdot \mathrm{Var}[Y]}$ for RVs with finite variances, (b) follows from $c_{i,t} \leq P_{\mathrm{avg}}, \forall i \in [\![M]\!], \forall t \in [\![n]\!]$, for a Poisson RV $Y_t(i) \sim \mathrm{Pois}(\lambda)$, an upper bound on the non-centered moments:

$$\mathbb{E}\Big[Y_t^k(i)\Big] \leq \mathbb{E}^k\big[Y_t(i)\big] \cdot \exp\Big(k^2/2\,\mathbb{E}\big[Y_t(i)\big]\Big), \tag{53}$$

(see [76, Th. 1]), and (51). Thereby, exploiting (45)-(49), we can establish the following upper bound on the type I error probability given in (44):

$$\begin{aligned}P_{e,1}(i, \mathbb{K}) &= \Pr\Big(\big|T\big(\mathbf{Y}(i), \mathbf{c}_j\big)\big| > \delta_n\Big)\\ &\overset{(a)}{\leq} \frac{\Big(\bar{L}P_{\mathrm{avg}} + \lambda\Big)^4 \exp(8/\lambda) + \Big(2\Big(\bar{L}P_{\mathrm{avg}} + \lambda\Big) + 1\Big)^2 + \psi_{\mathrm{Cov}}}{n\delta_n^2}\\ &\overset{(b)}{=} \frac{9\Big(\Big(\bar{L}P_{\mathrm{avg}} + \lambda\Big)^4 \exp(8/\lambda) + \Big(2\Big(\bar{L}P_{\mathrm{avg}} + \lambda\Big) + 1\Big)^2 + \psi_{\mathrm{Cov}}\Big)}{16a^2 n^{b+\kappa+4l}}\\ &= \frac{O\big(L^4\big)}{n^{b+\kappa+4l}} = \frac{O(1)}{n^{b+\kappa}}\\ &\leq e_1,\end{aligned} \tag{54}$$

for sufficiently large n and arbitrarily small e_1, where (a) follows from (45), (104), (52), and (b) follows from (41).

Type II error analysis: Next, we consider type II errors, i.e., when $\mathbf{Y}(i) \in \mathscr{T}_{\mathbb{K}}$ while the transmitter sent $\mathbf{c}_i$ with $i \notin \mathbb{K}$. Then, for each of the $\binom{M}{K}$ possible cases of $\mathbb{K}$, where $i \notin \mathbb{K}$, the type II error probability is bounded as

$$\begin{aligned}P_{e,2}\big(i, \mathbb{K}\big) &= \Pr\big(\mathbf{Y}(i) \in \mathbb{T}_{\mathbb{K}}\big)\\ &= \Pr\Bigg(\mathbf{Y}(i) \in \bigcup_{j\in\mathbb{K}} \mathbb{T}_j\Bigg)\\ &= \Pr\Bigg(\bigcup_{j\in\mathbb{K}} \Big\{|T\big(\mathbf{Y}(i), c_j\big)| \leq \delta_n\Big\}\Bigg)\\ &\leq \sum_{j=1}^{|\mathbb{K}|} \Pr\Big(\Big|T\big(\mathbf{Y}(i);\mathbf{c}_j\big)\Big| \leq \delta_n\Big)\\ &\leq |\mathbb{K}| \cdot \Bigg[\max_{1\leq j\leq K} \Pr\Big(\Big|T\big(\mathbf{Y}(i);\mathbf{c}_j\big)\Big| \leq \delta_n\Big)\Bigg],\end{aligned} \tag{55}$$

where $T(\mathbf{Y}(i);\mathbf{c}_j)$ is a random variable modeling the decoding metric in (43), i.e.,

$$T\big(\mathbf{Y}(i);\mathbf{c}_j\big) = \frac{1}{\bar{n}} \sum_{t=1}^{\bar{n}} \big(Y_t(i) - (c_{j,t}^{\boldsymbol{\rho}} + \lambda)\big)^2 - Y_t(i). \tag{56}$$

Next, we establish an upper bound on the RHS of (55), while we assume that j can be an arbitrary value from set $[\![K]\!]$. Further, let

$$\tilde{j} \triangleq \underset{1\leq j\leq K}{\arg\max} \Pr\Big(\Big|T\big(\mathbf{Y}(i);\mathbf{c}_j\big)\Big| \leq \delta_n\Big). \tag{57}$$

We note that if our analysis gives an upper bound on $\Pr(|T(\mathbf{Y}(i);\mathbf{c}_j)| \leq \delta_n)$ for arbitrary $j \in [\![K]\!]$, then the same upper bound is valid for $\Pr(|T(\mathbf{Y}(i);\mathbf{c}_{\tilde{j}})| \leq \delta_n)$. That is, we immediately obtain an upper bound for $\max_{1\leq j\leq K} \Pr(|T(\mathbf{Y}(i);\mathbf{c}_j)| \leq \delta_n)$ in (55).

Observe that (56) for $j = \tilde{j}$ can be rewritten as follows

$$\begin{aligned}&T\Big(\mathbf{Y}(i);\mathbf{c}_{\tilde{j}}\Big)\\ &= \frac{1}{\bar{n}} \sum_{t=1}^{\bar{n}} \underbrace{\big(Y_t(i) - (c_{i,t}^{\boldsymbol{\rho}} + \lambda) + (c_{i,t}^{\boldsymbol{\rho}} - c_{\tilde{j},t}^{\boldsymbol{\rho}})\big)^2}_{\triangleq \phi_{i,\tilde{j},t}} - Y_t(i).\end{aligned} \tag{58}$$

Observe that $\phi_{i,\tilde{j},t}$ in (58) can be expressed as

$$\phi_{i,\tilde{j},t} = \bar{Y}_t(i)^2 + \psi_{i,\tilde{j},t}^2 + 2\bar{Y}_t(i)\psi_{i,\tilde{j},t}, \tag{59}$$

where

$$\bar{Y}_t(i) = Y_t(i) - \big(c_{i,t}^{\boldsymbol{\rho}} + \lambda\big) \quad \text{and} \quad \psi_{i,\tilde{j},t} = c_{i,t}^{\boldsymbol{\rho}} - c_{j,t}^{\boldsymbol{\rho}}. \tag{60}$$

Then, define the following events

$$\mathcal{E}_{i,\tilde{j}} = \left\{ \left| \sum_{t=1}^{\bar{n}} \left(\bar{Y}_t(i) + \psi_{i,\tilde{j},t}\right)^2 - Y_t(i) \right| \leq \bar{n}\delta_n \right\},$$
$$\mathcal{E}'_{i,\tilde{j}} = \left\{ \sum_{t=1}^{\bar{n}} \left(\bar{Y}_t(i) + \psi_{i,\tilde{j}}, t\right)^2 - Y_t(i) \leq \bar{n}\delta_n \right\},$$
$$\mathcal{E}_{i,\tilde{j}}{}'\prime = \left\{ \left| \sum_{t=1}^{\bar{n}} \bar{Y}_t(i)\psi_{i,\tilde{j},t} \right| > \bar{n}\delta_n/2 \right\},$$
$$\mathcal{E}_{i,\tilde{j}}{}'\prime' = \left\{ \sum_{t=1}^{\bar{n}} \bar{Y}_t(i)^2 + \psi^2_{i,\tilde{j},t} - Y_t(i) \leq 2\bar{n}\delta_n \right\}. \quad (61)$$

Hence,

$$\begin{aligned} P_{e,2}(i,\mathbb{K}) &\leq K \cdot \Pr\left(\mathcal{E}_{i,\tilde{j}}\right) \\ &= K \cdot \Pr\left(\left| \sum_{t=1}^{\bar{n}} \left(\bar{Y}_t(i) + \psi_{i,\tilde{j},t}\right)^2 - Y_t(i) \right| \leq \bar{n}\delta_n \right) \\ &\overset{(a)}{\leq} K \cdot \Pr\left(\sum_{t=1}^{\bar{n}} \left(\bar{Y}_t(i) + \psi_{i,\tilde{j},t}\right)^2 - Y_t(i) \leq \bar{n}\delta_n \right) \\ &= K \cdot \Pr\left(\mathcal{E}'_{i,\tilde{j}}\right), \end{aligned} \quad (62)$$

where (a) holds since $\alpha - \beta \leq |\alpha - \beta|$ for real α, β. Now, we apply the law of total probability to event $\mathcal{E}'_{i,\tilde{j}}$ with respect to the pair of $(\mathcal{E}_{i,\tilde{j}}{}'\prime, \mathcal{E}_{i,\tilde{j}}{}'\prime^c)$, and obtain the following upper bound on the type II error probability,

$$\begin{aligned} P_{e,2}(i,\mathbb{K}) &\leq K \cdot \Pr\left(\mathcal{E}'_{i,\tilde{j}}\right) \\ &= K \cdot \left[\Pr\left(\mathcal{E}'_{i,\tilde{j}} \cap \mathcal{E}_{i,\tilde{j}}{}'\prime\right) + \Pr\left(\mathcal{E}'_{i,\tilde{j}} \cap \mathcal{E}_{i,\tilde{j}}{}'\prime^c\right) \right] \\ &\overset{(a)}{\leq} K \cdot \left[\Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime\right) + \Pr\left(\mathcal{E}'_{i,\tilde{j}} \cap \mathcal{E}_{i,\tilde{j}}{}'\prime^c\right) \right] \\ &\overset{(b)}{=} K \cdot \left[\Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime\right) + \Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime'\right) \right], \end{aligned} \quad (63)$$

where (a) follows from $\mathcal{E}'_{i,\tilde{j}} \cap \mathcal{E}_{i,\tilde{j}}{}'\prime \subset \mathcal{E}_{i,\tilde{j}}{}'\prime$ and (b) holds since the event $\mathcal{E}'_{i,\tilde{j}} \cap \mathcal{E}_{i,\tilde{j}}{}'\prime^c$ yields event $\mathcal{E}_{i,\tilde{j}}{}'\prime'$, with the following argument. Observe that,

$$\begin{aligned} &\Pr\left(\mathcal{E}'_{i,\tilde{j}} \cap \mathcal{E}_{i,\tilde{j}}{}'\prime^c\right) \overset{(a)}{\leq} \Pr\left(\sum_{t=1}^{\bar{n}} \bar{Y}_t(i)^2 + \psi^2_{i,\tilde{j},t} - Y_t(i) \leq 2\bar{n}\delta_n \right) \\ &= \Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime'\right), \end{aligned} \quad (64)$$

where (a) holds since given the complementary event $\mathcal{E}_{i,\tilde{j}}{}'\prime^c$, we obtain

$$-\bar{n}\delta_n/2 \leq \sum_{t=1}^{\bar{n}} \bar{Y}_t(i)\psi_{i,\tilde{j},t} \leq \bar{n}\delta_n/2,$$

which implies that $-2\sum_{t=1}^{\bar{n}} \bar{Y}_t(i)\psi_{i,\tilde{j},t} \leq \bar{n}\delta_n$. That is, event $\mathcal{E}'_{i,\tilde{j}} \cap \mathcal{E}_{i,\tilde{j}}{}'\prime^c$ *yields* the event

$$\sum_{t=1}^{\bar{n}} \bar{Y}_t(i)^2 + \psi^2_{i,\tilde{j},t} - Y_t(i) \leq 2\bar{n}\delta_n.$$

Now, we establish an upper bound on $\Pr(\mathcal{E}_{i,\tilde{j}}{}'\prime)$ by exploiting Chebyshev's inequality:

$$\begin{aligned} \Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime\right) &= \Pr\left(\left| \sum_{t=1}^{\bar{n}} \bar{Y}_t(i)\psi_{i,\tilde{j},t} \right| > \bar{n}\delta_n/2 \right) \\ &\leq \frac{\mathrm{Var}\left[\sum_{t=1}^{\bar{n}} \bar{Y}_t(i)\psi_{i,\tilde{j},t}\right]}{(\bar{n}\delta_n)^2} \\ &= \frac{\sum_{t=1}^{\bar{n}} \mathrm{Var}\left[\bar{Y}_t(i)\psi_{i,\tilde{j},t}\right]}{(\bar{n}\delta_n)^2}, \end{aligned} \quad (65)$$

where the last equality holds since the variance of the sum of uncorrelated RVs is the sum of the respective variances. Thereby,

$$\begin{aligned} \Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime\right) &\leq \frac{\sum_{t=1}^{\bar{n}} \psi^2_{i,\tilde{j},t}\mathrm{Var}\left[\bar{Y}_t(i)\right]}{(\bar{n}\delta_n)^2} \\ &= \frac{\sum_{t=1}^{\bar{n}} \left(c^{\rho}_{i,t} - c^{\rho}_{\tilde{j},t}\right)^2 \mathrm{Var}\left[\bar{Y}_t(i)\right]}{(\bar{n}\delta_n)^2} \\ &\overset{(a)}{\leq} \frac{\sum_{t=1}^{\bar{n}} \left(c^{\rho}_{i,t} + c^{\rho}_{\tilde{j},t}\right)^2 \mathrm{Var}\left[\bar{Y}_t(i)\right]}{(\bar{n}\delta_n)^2} \\ &\overset{(b)}{=} \frac{\sum_{t=1}^{\bar{n}} \left(c^{\rho}_{i,t} + c^{\rho}_{\tilde{j},t}\right)^2 \left(c^{\rho}_{i,t} + \lambda\right)}{(\bar{n}\delta_n)^2} \\ &\overset{(c)}{\leq} \frac{\|\mathbf{c}^{\rho}_i + \mathbf{c}^{\rho}_{\tilde{j}}\|^2 \left(\bar{L}P_{\text{ave}} + \lambda\right)}{(\bar{n}\delta_n)^2}, \end{aligned} \quad (66)$$

where (a) exploits the triangle inequality, i.e., $|c^{\rho}_{i,t} - c^{\rho}_{\tilde{j},t}| \leq |c^{\rho}_{i,t} + c^{\rho}_{\tilde{j},t}|$, (b) follows since $\mathrm{Var}[\bar{Y}_t(i)] = c^{\rho}_{i,t} + \lambda, \forall t \in [\![\bar{n}]\!]$, and (c) follows since $c^{\rho}_{i,t} \leq \bar{L}P_{\text{ave}} + \lambda$. Now, observe that

$$\begin{aligned} \|\mathbf{c}^{\rho}_i + \mathbf{c}^{\rho}_{\tilde{j}}\|^2 &\overset{(a)}{\leq} \left(\left\|\mathbf{c}^{\rho}_i\right\| + \left\|\mathbf{c}^{\rho}_{\tilde{j}}\right\| \right)^2 \\ &\overset{(b)}{\leq} \left(\sqrt{n}\left\|\mathbf{c}^{\rho}_i\right\|_{\infty} + \sqrt{n}\left\|\mathbf{c}^{\rho}_{\tilde{j}}\right\|_{\infty} \right)^2 \\ &\overset{(c)}{\leq} \left(\sqrt{n}\bar{L}P_{\text{avg}} + \sqrt{n}\bar{L}P_{\text{avg}} \right)^2 \\ &= 4\bar{L}^2 nP^2_{\text{avg}}, \end{aligned} \quad (67)$$

where (a) holds by the triangle inequality, (b) follows since $\|\cdot\| \leq \sqrt{n}\|\cdot\|_{\infty}$, and (c) is valid by the definition of $\mathbf{c}^{\rho}_i$, i.e., $\mathbf{c}^{\rho}_i = \sum_{l=0}^{L-1} \rho_l c_{i,t-l}$, and (33). Hence,

$$\begin{aligned} \Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime\right) &\leq \frac{\|\mathbf{c}^{\rho}_i + \mathbf{c}^{\rho}_{\tilde{j}}\|^2 \left(\bar{L}P_{\text{ave}} + \lambda\right)}{(\bar{n}\delta_n)^2} \\ &\leq \frac{4\bar{L}^2 P^2_{\text{avg}} \left(\bar{L}P_{\text{ave}} + \lambda\right)}{n\delta_n^2} \\ &= \frac{9\bar{L}^3 P^2_{\text{avg}} \left(P_{\text{avg}} + \lambda\right)}{4a^2 n^{b+\kappa+4l}} \end{aligned}$$

$$= \frac{O\left(L^3\right)}{n^{b+\kappa+4l}} \triangleq \zeta_0. \tag{68}$$

We now proceed with bounding $\Pr(\mathcal{E}_{i,\tilde{j}}{}'\prime\prime)$ as follows. Based on the convoluted codebook construction, each convoluted codeword is surrounded by a sphere of radius $\sqrt{n\epsilon_n}$, that is

$$\left\|\mathbf{c}_i^{\rho} - \mathbf{c}_{\tilde{j}}^{\rho}\right\|^2 \geq 4n\epsilon_n = 3\bar{n}\delta_n, \tag{69}$$

where the last equality exploits (41). Thus, we can establish the following upper bound for event $\mathcal{E}_{i,\tilde{j}}{}'\prime$:

$$\begin{aligned}
\Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime\right) &= \Pr\left(\sum_{t=1}^{\bar{n}} \bar{Y}_t(i)^2 + \psi_{i,\tilde{j},t}^2 - Y_t(i) \leq 2\bar{n}\delta_n\right) \\
&= \Pr\left(\sum_{t=1}^{\bar{n}} \bar{Y}_t(i)^2 - Y_t(i) \leq 2\bar{n}\delta_n - \psi_{i,\tilde{j},t}^2\right) \\
&\overset{(a)}{\leq} \Pr\left(\sum_{t=1}^{\bar{n}} \bar{Y}_t(i)^2 - Y_t(i) \leq 2\bar{n}\delta_n - 3\bar{n}\delta_n\right) \\
&\overset{(b)}{\leq} \frac{\mathrm{Var}\left[\sum_{t=1}^{\bar{n}} \bar{Y}_t(i)^2 - Y_t(i)\right]}{\bar{n}^2\delta_n^2} \\
&\overset{(c)}{=} \frac{\mathrm{Var}\left[T\left(\mathbf{Y}(i), \mathbf{c}_i\right)\right]}{\delta_n^2} \\
&\overset{(d)}{=} \frac{9\left(\left(\bar{L}P_{\text{avg}} + \lambda\right)^4 \exp(8/\lambda) + \left(2\left(\bar{L}P_{\text{avg}} + \lambda\right) + 1\right)^2 + \psi_{\text{Cov}}\right)}{16a^2 n^{b+\kappa+4l}} \\
&\triangleq \zeta_1,
\end{aligned} \tag{70}$$

where (a) follows from (69), (b) holds from applying Chebyshev's inequality, (c) follows from similar arguments as provided for the type I error probability, i.e., the calculations provided in (47) and (48), (d) holds by (52).

To sum up, recalling (68), we obtain

$$K\zeta_0 = \frac{9K\bar{L}^3 P_{\text{avg}}^2\left(P_{\text{avg}} + \lambda\right)}{4a^2 n^{b+\kappa+4l}} \overset{(a)}{=} \frac{O\left(L^3\right)}{n^{b+4l}} \overset{(b)}{=} \frac{O(1)}{n^{b+l}} \triangleq \zeta_0', \tag{71}$$

where (a) exploits $K = n^{\kappa}$ and (b) holds as $L = n^l$. On the other hand, recalling (70), we obtain

$$\begin{aligned}
K\zeta_1 &= \frac{9n^{\kappa}\left(\left(\left(\bar{L}P_{\text{avg}} + \lambda\right)^4 \exp(8/\lambda) + \left(2\left(\bar{L}P_{\text{avg}} + \lambda\right) + 1\right)^2 + \psi_{\text{Cov}}\right)\right)}{16a^2 n^{b+\kappa+4l}} \\
&\overset{(a)}{=} \frac{O\left(L^4\right)}{n^{b+4l}} \overset{(b)}{=} \frac{O(1)}{n^b} \triangleq \zeta_1',
\end{aligned} \tag{72}$$

where (a) exploits $K = n^{\kappa}$ and (b) holds as $L = n^l$. Therefore, recalling (63) and (68), and (70) we obtain

$$\begin{aligned}
P_{e,2}(i, \mathbb{K}) &\leq K \cdot \left[\Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime\right) + \Pr\left(\mathcal{E}_{i,\tilde{j}}{}'\prime\right)\right] \\
&\leq K \cdot \left[\zeta_0 + \zeta_1\right] \\
&= \zeta_0' + \zeta_1' \\
&\leq e_2,
\end{aligned} \tag{73}$$

hence, $P_{e,2}(i, \mathbb{K}) \leq e_2$ holds for sufficiently large n and arbitrarily small $e_2 > 0$.

We have thus shown that for every $e_1, e_2 > 0$ and sufficiently large n, there exists an $(n, M(n,R), K(n,\kappa), L(n,l), e_1, e_2)$-ISI-Poisson DKI code.

Remark 3: In the error analysis, we established upper bounds on the type I (cf. (54)) and type II error probabilities (cf. (71) and (72)). These results reveal that the fastest scales for the size of the target message set $K(n,\kappa)$ and the number of ISI taps $L(n,l)$ which ensure the vanishing of the type I and type II error probabilities as $n \to \infty$, are allowed to be defined as follows:

$$K(n,\kappa) = 2^{\kappa \log n} = n^{\kappa} \quad \text{and} \quad L(n,l) = 2^{l \log n} = n^l.$$

C. UPPER BOUND (CONVERSE PROOF)

Before we start with the converse proof, for the sake of a concise presentation of the analysis, we introduce the following notations. Let:

- $I_t^{\mathbf{x}} \triangleq \lambda + \sum_{l=1}^{L-1} \rho_l x_{t-l}$.
- $d_{i,t} = \rho_0 c_{i,t} + I_t^{\mathbf{c}_i}, \forall t \in [\![n]\!]$.

The converse proof consists of the following two main steps.

- *Step 1:* First, we show in Lemma 1 that for any achievable DKI rate (for which the type I and type II error probabilities vanish as $n \to \infty$), the distance between any selected entry of one codeword and any entry of another codeword is at least larger than a threshold.
- *Step 2:* Employing Lemma 1, we then derive an upper bound on the codebook size of DKI codes.

We start with the following lemma on the ratio of $d_{i_2,t}/d_{i_1,t}$ for two distinct messages i_1 and i_2, with $i_1, i_2 \in [\![M]\!]$.

Lemma 1 (Shifted Symbol Distance): Suppose that $R > 0$ is an achievable DKI rate for the DTPC with ISI, $\mathcal{P}$. Consider a sequence of $(n, M(n,R), K(n,\kappa), L(n,l), e_1^{(n)}, e_2^{(n)})$-ISI-Poisson codes $(\mathcal{C}^{(n)}, \mathscr{T}^{(n)})$, where

$$K(n,\kappa) = 2^{\kappa \log n} \quad , \quad L(n,l) = 2^{l \log n},$$

with $\kappa, l \in [0,1)$ such that $e_1^{(n)}$ and $e_2^{(n)}$ tend to zero as $n \to \infty$. Then, given a sufficiently large n, the codebook $\mathcal{C}^{(n)}$ satisfies the following property. For every pair of codewords, $\mathbf{c}_{i_1}$ and $\mathbf{c}_{i_2}$, there exists at least one letter $t \in [\![n]\!]$ such that

$$\left|1 - \frac{d_{i_2,t}}{d_{i_1,t}}\right| > \theta_n, \tag{74}$$

for all $i_1, i_2 \in [\![M]\!]$, such that $i_1 \neq i_2$, with

$$\theta_n \triangleq \frac{P_{\max}}{KLn^{1+b}} = \frac{P_{\max}}{n^{1+b+l+\kappa}}, \tag{75}$$

where $b > 0$ is an arbitrarily small constant.

Proof: The method of proof is by contradiction, namely, we assume that the condition given in (74) is violated and then we show that this leads to a contradiction, namely the sum of the type I and type II error probabilities converges to one, i.e., $\lim_{n\to\infty} [P_{e,1}(i_1, \mathbb{K}) + P_{e,2}(i_2, \mathbb{K})] = 1$ for some $\mathbb{K} \subseteq [\![M]\!]$, where $i_1 \in \mathbb{K}$ and $i_2 \notin \mathbb{K}$.

Let $e_1, e_2 > 0$ and $\eta_0, \eta_1, \eta_2, \delta > 0$ be arbitrarily small constants. Assume to the contrary that there exist two messages i_1 and i_2, where $i_1 \neq i_2$, meeting the error constraints in (8) and (9), such that $\forall t \in [\![n]\!]$, we have

$$\left|1 - \frac{d_{i_2,t}}{d_{i_1,t}}\right| \leq \theta_n. \quad (76)$$

In order to show contradiction, we bound the sum of the two error probabilities, $P_{e,1}(i_1, \mathbb{K}) + P_{e,2}(i_2, \mathbb{K})$, from below. Then, observe that

$$P_{e,1}(i_1, \mathbb{K}) + P_{e,2}(i_2, \mathbb{K}) = \left[1 - \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1})\right] + \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_2}). \quad (77)$$

To bound the error, let us define

$$\mathbb{F}_{i_1} = \left\{\mathbf{y} \in \mathscr{T}_{\mathbb{K}} : \bar{n}^{-1} \sum_{t=1}^{\bar{n}} Y_t - I_t^{\mathbf{c}_{i_1}} \leq \rho_0 P_{\max} + \delta\right\}, \quad (78)$$

where $\mathscr{T}_{\mathbb{K}} \subseteq \mathbb{N}_0^{\bar{n}}$ is the decoding set adopted[6] for the set of target messages $\mathbb{K}$.

Now, consider the sum inside the bracket in (77),

$$\sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) = \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}} \cap \mathbb{F}_{i_1}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) + \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}} \cap \mathbb{F}_{i_1}^c} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}), \quad (79)$$

where the equality follows from applying the law of total probability on $\mathscr{T}_{\mathbb{K}}$ with respect to $(\mathbb{F}_{i_1}, \mathbb{F}_{i_1}^c)$.

Now, we proceed to establish an upper bound on the RHS sum in (79) as follows

$$\sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}} \cap \mathbb{F}_{i_1}^c} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) = \Pr\left(\mathscr{T}_{\mathbb{K}} \cap \mathbb{F}_{i_1}^c\right) \leq \Pr\left(\bar{n}^{-1} \sum_{t=1}^{\bar{n}} Y_t(i_1) - I_t^{\mathbf{c}_{i_1}} > \rho_0 P_{\max} + \delta\right). \quad (80)$$

Next, we apply Chebyshev's inequality to the probability term in (80) and obtain

$$\begin{aligned}
&\sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}} \cap \mathbb{F}_{i_1}^c} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) \\
&\overset{(a)}{\leq} \Pr\left(\bar{n}^{-1} \sum_{t=1}^{\bar{n}} Y_t(i_1) - \bar{n}^{-1} \sum_{t=1}^{\bar{n}} \mathbb{E}\left[Y_t(i_1)\right] > \rho_0 P_{\max} + \delta\right) \\
&\overset{(b)}{\leq} \frac{\mathrm{Var}\left[\bar{n}^{-1} \sum_{t=1}^{\bar{n}} Y_t(i_1)\right]}{\left(\rho_0 P_{\max} + \delta\right)^2} \\
&\overset{(c)}{=} \frac{\bar{n}^{-2} \sum_{t=1}^{\bar{n}} \rho_0 c_{i_1,t} + I_t^{\mathbf{c}_{i_1}}}{\left(\rho_0 P_{\max} + \delta\right)^2} \\
&\overset{(d)}{\leq} \frac{T_R P_{\max} + \lambda + (L-1)\, T_R P_{\max}}{n\delta^2} \\
&\leq \frac{L T_R P_{\max} + \lambda}{n\delta^2} = \frac{O(L)}{n\delta^2} \\
&\overset{(e)}{=} \frac{O(1)}{n^{1-l}\delta^2} \triangleq \eta_0,
\end{aligned} \quad (81)$$

for sufficiently large n, where (a) holds since $\mathbb{E}[Y_t(i_1)] = I_t^{\mathbf{c}_{i_1}}$, for inequality (b), we exploited Chebyshev's inequality, and for equality (c), we used the fact that $\mathrm{Var}[Y_t(i_1)] = \mathbb{E}[Y_t(i_1)] = \rho_0 c_{i_1,t} + I_t^{\mathbf{c}_{i_1}}$, $\forall t \in [\![n]\!]$. Inequality (d) employs $c_{i_1,t} \leq P_{\max}$, $\forall i_1 \in [\![M]\!]$, $\forall t \in [\![n]\!]$, $\rho_0 \leq T_R$, $n \leq \bar{n}$ and (e) exploits $L = n^l$. Thereby, recalling (79) and (81), we obtain

$$\begin{aligned}
\sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1})
&\leq \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}} \cap \mathbb{F}_{i_1}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) + \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}} \cap \mathbb{F}_{i_1}^c} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) \\
&\leq \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}} \cap \mathbb{F}_{i_1}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) + \eta_0.
\end{aligned} \quad (82)$$

Next, recalling the sum of error probabilities in (77), where $i_1 \in \mathbb{K}$ and $i_2 \notin \mathbb{K}$, we obtain

$$\begin{aligned}
&P_{e,1}(i_1, \mathbb{K}) + P_{e,2}(i_2, \mathbb{K}) \\
&= \left[1 - \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1})\right] + \sum_{\mathbf{y} \in \mathscr{T}_{\mathbb{K}}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_2}) \\
&\overset{(a)}{\geq} 1 - \eta_0 - \sum_{\mathbb{F}_{i_1}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) + \sum_{\mathscr{T}_{\mathbb{K}}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_2}) \\
&\overset{(b)}{\geq} 1 - \eta_0 - \sum_{\bigcup_{i_1 \in \mathbb{K}} \mathbb{F}_{i_1}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) + \sum_{\bigcup_{i_1 \in \mathbb{K}} \mathbb{F}_{i_1}} V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_2}) \\
&\geq 1 - \eta_0 - \sum_{\bigcup_{i_1 \in \mathbb{K}} \mathbb{F}_{i_1}} \left[V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) - V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_2})\right],
\end{aligned} \quad (83)$$

where (a) holds by (82) and (b) follows since $\mathbb{F}_{i_1} \subset \bigcup_{i_1 \in \mathbb{K}} \mathbb{F}_{i_1} \subset \mathscr{T}_{\mathbb{K}}$. Now, let us focus on the summand in the square brackets in (83). Employing (5), we have

$$\begin{aligned}
&V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) - V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_2}) \\
&= V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) \cdot \left[1 - V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_2}) / V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1})\right] \\
&= V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) \cdot \left[1 - \prod_{t=1}^{\bar{n}} e^{-\left(d_{i_2,t} - d_{i_1,t}\right)} \left(\frac{d_{i_2,t}}{d_{i_1,t}}\right)^{Y_t}\right] \\
&= V^{\bar{n}}(\mathbf{y} \mid \mathbf{c}_{i_1}) \cdot \left[1 - \prod_{t=1}^{\bar{n}} e^{-\theta_n d_{i_1,t}} (1 - \theta_n)^{Y_t}\right],
\end{aligned} \quad (84)$$

where for the last inequality, we exploited

$$d_{i_2,t} - d_{i_1,t} \leq \left|d_{i_2,t} - d_{i_1,t}\right| \leq \theta_n d_{i_1,t} \quad (85)$$

and

$$1 - \frac{d_{i_2,t}}{d_{i_1,t}} \leq \left|1 - \frac{d_{i_2,t}}{d_{i_1,t}}\right| \leq \theta_n, \quad (86)$$

[6] We note that in the achievability proof given in Section IV-B we impose a specific structure on the decoding set $\mathscr{T}_{\mathbb{K}}$, namely, we defined $\mathscr{T}_{\mathbb{K}}$ to be the union of the individual decoding set corresponding to messages that belong to set $\mathbb{K}$, i.e., $\mathscr{T}_{\mathbb{K}} = \bigcup_{i_1 \in \mathbb{K}} \mathbb{T}_{i_1}$. In contrast, in the converse proof, we do not impose any structure on $\mathscr{T}_{\mathbb{K}}$ and treat the decoding set $\mathscr{T}_{\mathbb{K}}$ as a general choice $\mathscr{T}_{\mathbb{K}} \subseteq \mathbb{N}_0^{\bar{n}}$.

which holds by (76). Now, we bound the product term inside the bracket in (84) for space $\mathbf{y} \in \bigcup_{i_1 \in \mathbb{K}} \mathbb{F}_{i_1}$ as follows:

$$\begin{aligned}
&\prod_{t=1}^{\bar{n}} e^{-\theta_n d_{i_1,t}} (1-\theta_n)^{Y_t} = e^{-\theta_n \sum_{t=1}^{\bar{n}} d_{i_1,t}} \cdot (1-\theta_n)^{\sum_{t=1}^{\bar{n}} Y_t} \\
&\overset{(a)}{\geq} e^{-\bar{n}\theta_n \left(\rho_0 P_{\max} + \bar{n}^{-1} \sum_{t=1}^{\bar{n}} I_t^{\mathbf{c}_{i_1}}\right)} \\
&\quad \cdot (1-\theta_n)^{\bar{n}\left(\rho_0 P_{\max} + \bar{n}^{-1} \sum_{t=1}^{\bar{n}} I_t^{\mathbf{c}_{i_1}} + \delta\right)} \\
&= e^{\bar{n}\theta_n \delta} \cdot e^{-\bar{n}\theta_n \left(\rho_0 P_{\max} + \bar{n}^{-1} \sum_{t=1}^{\bar{n}} I_t^{\mathbf{c}_{i_1}} + \delta\right)} \\
&\quad \cdot (1-\theta_n)^{\bar{n}\left(\rho_0 P_{\max} + \bar{n}^{-1} \sum_{t=1}^{\bar{n}} I_t^{\mathbf{c}_{i_1}} + \delta\right)} \\
&\overset{(b)}{\geq} e^{\bar{n}\theta_n \delta} \cdot e^{-\bar{n}\theta_n \left(\rho_0 P_{\max} + \bar{n}^{-1} \sum_{t=1}^{\bar{n}} I_t^{\mathbf{c}_{i_1}} + \delta\right)} \\
&\quad \cdot (1-\bar{n}\theta_n)^{\rho_0 P_{\max} + \bar{n}^{-1} \sum_{t=1}^{\bar{n}} I_t^{\mathbf{c}_{i_1}} + \delta} \\
&\overset{(c)}{=} e^{\bar{n}\theta_n \delta} \cdot f(\bar{n}\theta_n) \geq e^{n\theta_n \delta} \cdot f(\bar{n}\theta_n) \overset{(d)}{>} f(\bar{n}\theta_n) \\
&\overset{(e)}{\geq} 1 - 3\Big(\rho_0 P_{\max} + \sum_{t=1}^{\bar{n}} I_t^{\mathbf{c}_{i_1}} + \delta\Big)\bar{n}\theta_n \\
&\overset{(f)}{\geq} 1 - \frac{3\big(T_R P_{\max} + \lambda + (L-1)T_R P_{\max} + \delta\big)P_{\max}}{n^{b+l+\kappa}} \cdot \frac{\bar{n}}{n} \\
&= 1 - \frac{O(L)}{n^{b+l+\kappa}} \cdot \left(1 + \frac{O(L)}{n}\right) \\
&= 1 - \frac{O(1)}{n^{b+\kappa}} - \frac{O\left(L^2\right)}{n^{1+b+l+\kappa}} \\
&\overset{(g)}{=} 1 - \left(\frac{O(1)}{n^{b+\kappa}} + \frac{O(1)}{n^{1+b+\kappa-l}}\right) \\
&\overset{(h)}{=} 1 - \eta_1,
\end{aligned} \tag{87}$$

for sufficiently large n. We used the following facts for the above inequalities:

- Inequality (a) follows since

$$d_{i_1,t} \leq \rho_0 P_{\max} + I_t^{\mathbf{c}_{i_1}}, \ \forall\, t \in [\![n]\!], \tag{88}$$

  and

$$\sum_{t=1}^{\bar{n}} Y_t \leq \bar{n}\Big(\rho_0 P_{\max} + \bar{n}^{-1} \sum_{t=1}^{\bar{n}} I_t^{\mathbf{c}_{i_1}} + \delta\Big), \tag{89}$$

  where the latter inequality follows from $\mathbf{y} \in \bigcup_{i_1 \in \mathbb{K}} \mathbb{F}_{i_1}$, cf. (78).
- For (b), we used Bernoulli's inequality [77, Ch. 3]:

$$(1-x)^r \geq 1 - rx\,, \ \ \forall x > -1\,, \ \forall r > 0. \tag{90}$$

- For (c), we used the following definition:

$$f(x) = e^{-cx}(1-x)^c, \tag{91}$$

  with $c = \rho_0 P_{\max} + \bar{n}^{-1} \sum_{t=1}^{\bar{n}} I_t^{\mathbf{c}_{i_1}} + \delta$.
- For (d), we used the fact that

$$e^{n\theta_n \delta} = e^{P_{\max}\delta / n^{b+l+\kappa}} > 1. \tag{92}$$

- For (e), we used the Taylor expansion

$$f(\bar{n}\theta_n) = 1 - 2c\bar{n}\theta_n + O\Big((\bar{n}\theta_n)^2\Big) \tag{93}$$

  to obtain the upper bound $f(\bar{n}\theta_n) \geq 1 - 3c\bar{n}\theta_n$ for sufficiently small values of $\bar{n}\theta_n$, i.e.,

$$\begin{aligned}
\bar{n}\theta_n &= \frac{P_{\max}}{n^{1+b+l+\kappa}} \cdot (n + L - 1) \ = \frac{P_{\max}}{n^{b+l+\kappa}} \cdot \frac{\bar{n}}{n} \\
&= \frac{P_{\max}}{n^{b+l+\kappa}} \cdot \left(1 + \frac{O(L)}{n}\right) = \frac{P_{\max}}{n^{b+l+\kappa}} + \frac{O(1)}{n^{b+l+\kappa}}.
\end{aligned} \tag{94}$$

- Inequality (f) exploits (75).
- Equality (g) employs $L = n^l$, with $l \in [0, 1)$.
- Finally, (h) follows from

$$\frac{O(1)}{n^{b+\kappa}} + \frac{O(1)}{n^{1+b+\kappa-l}} \triangleq \eta_1.$$

Thereby, (84) can then be written as follows

$$\begin{aligned}
&V^{\bar{n}}\big(\mathbf{y} \,|\, \mathbf{c}_{i_1}\big) - V^{\bar{n}}\big(\mathbf{y} \,|\, \mathbf{c}_{i_2}\big) \\
&\leq V^{\bar{n}}\big(\mathbf{y} \,|\, \mathbf{c}_{i_1}\big) \cdot \left[1 - e^{-\theta_n \sum_{t=1}^{\bar{n}} d_{i_1,t}} \cdot (1-\theta_n)^{\sum_{t=1}^{\bar{n}} Y_t}\right] \\
&\leq \eta_1 \cdot V^{\bar{n}}\big(\mathbf{y} \,|\, \mathbf{c}_{i_1}\big).
\end{aligned} \tag{95}$$

Next, recalling the definition of an ISI-Poisson DKI code given in (1), we focus on the underlying assumptions stated in Lemma 1 on the properties of a given sequence of ISI-Poisson DKI codes $(\mathcal{C}^{(n)}, \mathscr{T}^{(n)})$. Such a code sequence has five parameters $(n, M(n,R), K(n,\kappa), L(n,l), e_1^{(n)}, e_2^{(n)})$, and endows the following property:

For each general choice (arrangement) of the target message set $\mathbb{K} \subset [\![M]\!]$ of size K, the upper bound on the type I and type II error probabilities, i.e., $e_1^{(n)}$ and $e_2^{(n)}$, respectively, tends to zero as n tends to infinity. That is,

$$\lim_{n\to\infty} \Big[P_{e,1}\big(i_1, \mathbb{K}\big) + P_{e,2}\big(i_2, \mathbb{K}\big)\Big] = 0, \ \forall \mathbb{K} \subset [\![M]\!]. \tag{96}$$

Next, let $\mathbb{K}(i_1, i_2)$ denote a specific class of the target message sets $\mathbb{K}$, where $i_1 \in \mathbb{K}$ and $i_2 \notin \mathbb{K}$, i.e.,

$$\mathbb{K}(i_1, i_2) \triangleq \big\{\mathbb{K} \subseteq [\![M]\!];\ |\mathbb{K}| = K;\ i_1 \in \mathbb{K},\ i_2 \notin \mathbb{K}\big\}. \tag{97}$$

Observe that the above set cannot be empty (i.e., $|\mathbb{K}(i_1, i_2)| \geq 1$), that is, there exists at least one arrangement $\mathbb{K}$ belonging to $\mathbb{K}(i_1, i_2)$, where $i_1 \in \mathbb{K}$, $i_2 \notin \mathbb{K}$. This holds true since according to Lemma 1 the two messages i_1 and i_2 are distinct, i.e., $i_1 \neq i_2$. Thereby, for every set $\mathbb{K} \in \mathbb{K}(i_1, i_2)$, we have the following upper bounds on the type I and type II error probabilities

$$\begin{aligned}
P_{e,1}\big(i_1, \mathbb{K}\big) &= V^{\bar{n}}\big(\mathscr{T}_{\mathbb{K}}^c \,|\, x^n = \mathbf{c}_{i_1}\big) \leq e_1^{(n)}, \\
P_{e,2}\big(i_2, \mathbb{K}\big) &= V^{\bar{n}}\big(\mathscr{T}_{\mathbb{K}} \,|\, x^n = \mathbf{c}_{i_2}\big) \leq e_2^{(n)}.
\end{aligned} \tag{98}$$

Hence,

$$\begin{aligned}
e_1^{(n)} + e_2^{(n)} &\geq P_{e,1}\big(i_1, \mathbb{K}\big) + P_{e,2}\big(i_2, \mathbb{K}\big) \\
&\overset{(a)}{\geq} 1 - \eta_0 - \sum_{\bigcup_{i_1 \in \mathbb{K}} \mathbb{F}_{i_1}} \Big[V^{\bar{n}}\big(\mathbf{y} \,|\, \mathbf{c}_{i_1}\big) - V^{\bar{n}}\big(\mathbf{y} \,|\, \mathbf{c}_{i_2}\big)\Big]
\end{aligned}$$

$$\overset{(b)}{\geq} 1-\eta_0-\eta_1 \sum_{\bigcup_{i_1\in\mathbb{K}} \mathbb{F}_{i_1}} V^{\bar{n}}\big(\mathbf{y}\,|\,\mathbf{c}_{i_1}\big)$$
$$\overset{(c)}{\geq} 1-\eta_0-\eta_1 \sum_{i_1\in\mathbb{K}}\sum_{\mathbb{F}_{i_1}} V^{\bar{n}}\big(\mathbf{y}\,|\,\mathbf{c}_{i_1}\big)$$
$$\overset{(d)}{\geq} 1-\eta_0-\eta_1\cdot|\mathbb{K}|$$
$$\overset{(e)}{\geq} 1-\eta_0-\frac{KO(1)}{n^{b+\kappa}}$$
$$\overset{(e)}{=} 1-\eta_0-\eta_2, \tag{99}$$

where (a) follows from (83), and (b) holds by (95), (c) exploits the union bound, (d) follows since $\sum_{\mathbb{F}_{i_1}} V^{\bar{n}}(\mathbf{y}\,|\,\mathbf{c}_{i_1}) = \Pr(\mathbb{F}_{i_1}) \leq 1$, (e) holds since $|\mathbb{K}| = K$, and (f) follows from $O(n^{-b}) \triangleq \eta_2$.

Therefore, $e_1^{(n)}+e_2^{(n)} \geq 1-\eta_0-\eta_2$ which is a contradiction to (96). In other words, Lemma 1 states that every given sequence of ISI-Poisson DKI codes $(\mathcal{C}^{(n)}, \mathcal{T}^{(n)})$ with the parameters $(n, M(n,R), K(n,\kappa) = 2^{\kappa\log n}, L(n,l) = 2^{l\log n}, e_1^{(n)}, e_2^{(n)})$ endows the following property: For an arbitrary (general) choice of $\mathbb{K}$ of size $K(n,\kappa)$, the upper bounds on the type I and type II error probabilities vanish, i.e., $e_1^{(n)}$ and $e_2^{(n)}$ tend to zero as $n\to\infty$. However, we show that there exist some particular choices for $\mathbb{K}$ denoted by $\mathbb{K}(i_1,i_2)$ whose elements satisfy the following property: The sum of the corresponding upper bounds on the type I and type II errors is lower bounded by one, i.e., $e_1^{(n)}$ and $e_2^{(n)}$ do not vanish. This is clearly a contradiction and implies that the inequality given in (76) does not hold. This completes the proof of Lemma 1. ■

Next, we use Lemma 1 to prove the upper bound on the DKI capacity. Observe that since

$$d_{i,t} = \rho_0 c_{i,t} + I_t^{\mathbf{c}_i} > \lambda, \tag{100}$$

Lemma 1 implies

$$\rho_0\big|c_{i_1,t}-c_{i_2,t}\big| = \big|d_{i_1,t}-d_{i_2,t}\big| \overset{(a)}{>} \theta_n d_{i_1,t} \overset{(b)}{>} \lambda\theta_n, \tag{101}$$

where (a) follows from (74) and (b) holds by (100). Now, since $\|\mathbf{c}_{i_1}-\mathbf{c}_{i_2}\| \geq |c_{i_1,t}-c_{i_2,t}|$, we deduce that the distance between every pair of codewords satisfies

$$\|\mathbf{c}_{i_1}-\mathbf{c}_{i_2}\| > \lambda\theta_n/\rho_0. \tag{102}$$

Thus, we can define an arrangement of non-overlapping spheres $\mathcal{S}_{\mathbf{c}_i}(n,\lambda\theta_n/2\rho_0)$, i.e., spheres of radius $r_0 = \lambda\theta_n/2\rho_0$ that are centered at the codewords $\mathbf{c}_i$. Since all codewords belong to a hyper cube $\mathbb{Q}_{\mathbf{0}}(n,P_{\max})$ with edge length $P_{\max}$, it follows that the number of packed small spheres, i.e., the number of codewords M, is bounded by

$$M = \frac{\mathrm{Vol}\big(\bigcup_{i=1}^{M}\mathcal{S}_{\mathbf{c}_i}(n,r_0)\big)}{\mathrm{Vol}\big(\mathcal{S}_{\mathbf{c}_1}(n,r_0)\big)} = \frac{\Delta_n\big(\mathscr{S}\big)\cdot\mathrm{Vol}\big(\mathbb{Q}_{\mathbf{0}}(n,P_{\max})\big)}{\mathrm{Vol}\big(\mathcal{S}_{\mathbf{c}_1}(n,r_0)\big)}$$
$$\leq 2^{-0.599n}\cdot\frac{P_{\max}^n}{\mathrm{Vol}\big(\mathcal{S}_{\mathbf{c}_1}(n,r_0)\big)}, \tag{103}$$

where the last inequality follows from (31). Thereby,

$$\log M \leq \log\left(\frac{P_{\max}^n}{\mathrm{Vol}\big(\mathcal{S}_{\mathbf{c}_1}(n,r_0)\big)}\right) - 0.599n$$
$$= n\log(P_{\max}) - \log\big(\mathrm{Vol}\big(\mathcal{S}_{\mathbf{c}_1}(n,r_0)\big)\big) - 0.599n$$
$$\overset{(a)}{=} n\log P_{\max} - n\log r_0 - n\log\sqrt{\pi} + \log\big(\Gamma\big(n/2+1\big)\big) \tag{104}$$

where (a) exploits (32). Next, we proceed to establish an upper bound on the last term in (104). Observe that

$$\Gamma\big(n/2+1\big) \overset{(a)}{=} (n/2)\Gamma\big(n/2\big)$$
$$\overset{(b)}{<} \big(\lfloor n/2\rfloor+1\big)\Gamma\Big(\lfloor n/2\rfloor+1\Big)$$
$$\overset{(c)}{=} \big(\lfloor n/2\rfloor+1\big)!, \tag{105}$$

where (a) holds by the recurrence relation of the Gamma function [73] for real $n/2$, (b) follows since $n/2 < \lfloor n/2\rfloor+1$ for real $n/2$, and (c) holds since for positive integer $\lfloor n/2\rfloor$, we have $\Gamma(\lfloor n/2\rfloor+1) = (\lfloor n/2\rfloor)!$, cf. [73]. Next, we proceed to simplify the factorial term given in (105). To this end, we exploit *Stirling's approximation*, i.e., $\log n! = n\log n - n\log e + o(n)$ [74, p. 52] with the substitution of $n = \lfloor n/2\rfloor + 1$, where $\lfloor n/2\rfloor \in \mathbb{Z}$. Thereby, we obtain

$$\log\big(\Gamma\big(n/2+1\big)\big)$$
$$< \big(\lfloor n/2\rfloor+1\big)\log\big(\lfloor n/2\rfloor+1\big) - \big(\lfloor n/2\rfloor+1\big)\log e + o\big(\lfloor n/2\rfloor\big)$$
$$\overset{(a)}{\leq} \big(n/2+1\big)\log\big(n/2+1\big) - \big(n/2\big)\log e + o\big(\lfloor n/2\rfloor\big), \tag{106}$$

where (a) follows from $\left\lfloor\frac{n}{2}\right\rfloor \leq \frac{n}{2}$ and $\left\lfloor\frac{n}{2}\right\rfloor > \frac{n}{2}-1$, for integer n. Therefore, merging (104)–(106), we obtain

$$\log M \leq n\log P_{\max} - n\log r_0 - n\log\sqrt{\pi}$$
$$+ \big(n/2+1\big)\log\big(n/2+1\big) - \big(n/2\big)\log e + o\big(\lfloor n/2\rfloor\big)$$
$$= n\log P_{\max} - n\log\big(\lambda P_{\max}/(2\rho_0)\big)$$
$$+ (1+b+l+\kappa)\,n\log n$$
$$- n\log\sqrt{\pi} + \big(n/2+1\big)\log\big(n/2+1\big)$$
$$- \big(n/2\big)\log e + o\big(\lfloor n/2\rfloor\big), \tag{107}$$

where for the equality we used

$$r_0 = \frac{\lambda\theta_n}{2\rho_0} = \frac{\lambda P_{\max}}{2\rho_0 n^{1+b+l+\kappa}}. \tag{108}$$

The dominant term in (104) is again of order $n\log n$. Hence, to ensure a finite value for the upper bound of the rate, R, (104) induces the scaling law of M to be $2^{(n\log n)R}$. By setting $M(n,R) = 2^{(n\log n)R}$, we obtain

$$R \leq \frac{1}{n\log n}\Bigg[\left(\frac{1}{2}+(1+b+\kappa+l)\right)n\log n$$
$$- n\left(\frac{1}{2}+\log\Big(\lambda\sqrt{\pi e}/(2\rho_0)\Big)\right) + o(n)\Bigg], \tag{109}$$

which tends to $\frac{3}{2}+l+\kappa$ as $n\to\infty$ and $b\to 0$. This completes the proof of Theorem 1.

V. SUMMARY AND FUTURE DIRECTIONS

In this paper, we studied the deterministic K-identification problem for MC channels. In particular, we considered MC systems with molecule counting receivers, modeled by the DTPC with ISI. For this setting, we derived lower and upper bounds on the DKI capacity subject to average and peak molecule release rate constraints for a codebook size of $M(n, R) = 2^{(n \log n)R} = n^{nR}$. Our results revealed that the super-exponential scale of n^{nR} is the appropriate scale for the DKI capacity of the DTPC with ISI. This was proved by finding a suitable sphere packing arrangement embedded in a hyper cube. In particular, in the rate analysis, we established a lower bound for the logarithm of the number of codewords, whose fastest growing term has order $n \log n$; cf. (39) and (40). This observation dictates, that in order to obtain a finite and positive value for the DKI capacity, the codebook size should scale as $M(n, R) = 2^{(n \log n)R}$. We note that this scale is distinctly different from the ordinary scales in transmission and RI settings, where the codebook size grows exponentially and double exponentially, respectively.

The results presented in this paper can be extended in several directions, some of which are listed in the following as potential topics for future research:

- *Continuous alphabet conjecture:* Our observations for the codebook size of the DTPC with ISI, DTPC without ISI [26], [45], Binomial channel [53], [78], and Gaussian channel with fading [46] lead us to conjecture that the codebook size for any continuous alphabet channel is a super-exponential function, i.e., $2^{(n \log n)R}$. However, a formal proof of this conjecture remains unknown.
- *Multiuser and multi-antenna systems:* This study has focused on a point-to-point system and may be extended to multi-user scenarios (e.g., broadcast and multiple access channels) or multiple-input multiple-output channels which are relevant in complex MC nano-networks.
- *Finite codeword length coding:* The identification capacity results in this paper reveal the performance limits of DTPC with ISI for the asymptotic regime when the length of codewords can be arbitrarily large. However, the codeword length is finite in practice, particularly for MC applications, where large encoding/decoding delays cannot be afforded. Therefore, the study of the non-asymptotic DKI capacity of the DTPC is an important direction for future work.
- *Explicit code construction:* Our main focus in this paper was the establishment of fundamental performance limits for the DKI problem for the DTPC with ISI, where an explicit code construction was not considered. In fact, the proposed achievable scheme proves the existence of a code without providing a systematic approach to construct it. Hence, interesting directions for future research include the explicit construction of identification codes and the development of low-complexity encoding and decoding schemes for practical applications. The efficiency of these designs can be evaluated against the performance bounds derived in this paper.

REFERENCES

[1] M. J. Salariseddigh, V. Jamali, U. Pereg, H. Boche, C. Deppe, and R. Schober, "Deterministic identification for MC ISI-Poisson channel," in *Proc. IEEE Int. Conf. Commun.*, 2023, pp. 6108–6113.

[2] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[3] 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, 3rd Quart., 2016.

[4] V. Jamali, A. Ahmadzadeh, C. Jardin, H. Sticht, and R. Schober, "Channel estimation for diffusive molecular communications," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4238–4252, Oct. 2016.

[5] Y. Tang, F. Ji, M. Wen, Q. Wang, and L.-L. Yang, "Enhanced molecular type permutation shift keying for molecular communication," *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2722–2726, Dec. 2021.

[6] Q. Li, "A novel time-based modulation scheme in time-asynchronous channels for molecular communications," *IEEE Trans. Nanobiosci.*, vol. 19, no. 1, pp. 59–67, Jan. 2020.

[7] Y. Huang, M. Wen, L.-L. Yang, C.-B. Chae, and F. Ji, "Spatial modulation for molecular communication," *IEEE Trans. Nanobiosci.*, vol. 18, no. 3, pp. 381–395, Jul. 2019.

[8] M. Magarini and P. Stano, "Synthetic cells engaged in molecular communication: An opportunity for modelling Shannon-and semantic-information in the chemical domain," *Front. Comms. Net*, vol. 2, Sep. 2021, Art. no. 724597.

[9] C. A. Söldner et al., "A survey of biological building blocks for synthetic molecular communication systems," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2765–2800, 4th Quart., 2020.

[10] D. Bi, A. Almpanis, A. Noel, Y. Deng, and R. Schober, "A survey of molecular communication in cell biology: Establishing a new hierarchy for interdisciplinary applications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1494–1545, 3rd Quart., 2021.

[11] 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. 1256–1301, Jul. 2019.

[12] M. Kuscu, E. Dinc, 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, Jul. 2019.

[13] A. Gohari, M. Mirmohseni, and M. Nasiri-Kenari, "Information theory of molecular communication: Directions and challenges," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 2, no. 2, pp. 120–142, Dec. 2016.

[14] I. F. Akyildiz, M. Pierobon, S. Balasubramaniam, and Y. Koucheryavy, "The Internet of Bio-Nano Things," *IEEE Commun. Mag.*, vol. 53, pp. 32–40, Mar. 2015.

[15] C. Lee, B.-H. Koo, C.-B. Chae, and R. Schober, "The Internet of Bio-Nano Things in blood vessels: System design and prototypes," *J. Commun. Netw.*, vol. 25, no. 2, pp. 222–231, Apr. 2023.

[16] H. B. Yilmaz and C.-B. Chae, "Arrival modelling for molecular communication via diffusion," *Electron. Lett.*, vol. 50, no. 23, pp. 1667–1669, Nov. 2014.

[17] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*. Boston, MA, USA: McGraw-Hill, 2002.

[18] J. Cao, S. Hranilovic, and J. Chen, "Capacity-achieving distributions for the discrete-time poisson channel—Part I: General properties and numerical techniques," *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 194–202, Jan. 2014.

[19] J. Cao, S. Hranilovic, and J. Chen, "Capacity-achieving distributions for the discrete-time Poisson channel—Part II: Binary inputs," *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 194–202, Jan. 2014.

[20] F. Ratti, F. Vakilipoor, H. Awan, and M. Magarini, "Bounds on the constrained capacity for the diffusive Poisson molecular channel with memory," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 7, no. 2, pp. 100–105, Jun. 2021.

[21] G. Aminian, H. Arjmandi, A. Gohari, M. N. Kenari, and U. Mitra, "Capacity of LTI-Poisson channel for diffusion based molecular communication," in *Proc. IEEE Int. Conf. Commun.*, 2015, pp. 1060–1065.
[22] G. Aminian, H. Arjmandi, A. Gohari, M. Nasiri-Kenari, and U. Mitra, "Capacity of diffusion-based molecular communication networks over lti-Poisson channels," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 1, no. 2, pp. 188–201, Jun. 2015.
[23] N. Ahmadypour and A. Gohari, "Transmission of a bit over a discrete Poisson channel with memory," *IEEE Trans. Inf. Theory*, vol. 67, no. 7, pp. 4710–4727, Jul. 2021.
[24] A. O. Kislal, H. B. Yilmaz, A. E. Pusane, and T. Tugcu, "ISI-aware channel code design for molecular communication via diffusion," *IEEE Trans. Nanobiosci.*, vol. 18, no. 2, pp. 205–213, Apr. 2019.
[25] R. Mosayebi, H. Arjmandi, A. Gohari, M. Nasiri-Kenari, and U. Mitra, "Receivers for diffusion-based molecular communication: Exploiting memory and sampling rate," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 12, pp. 2368–2380, Dec. 2014.
[26] M. J. Salariseddigh, U. Pereg, H. Boche, C. Deppe, and R. Schober, "Deterministic identification over Poisson channels," in *Proc. IEEE Glob. Commun. Conf. (GC)*, 2021, pp. 1–6.
[27] M. J. Salariseddigh, V. Jamali, U. Pereg, H. Boche, C. Deppe, and R. Schober, "Deterministic identification for molecular communications over the Poisson channel," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 9, no. 4, pp. 408–424, Dec. 2023.
[28] W. Haselmayr et al., "Integration of molecular communications into future generation wireless networks," in *Proc. 6G Wireless Summit.*, 2019, pp. 1–2.
[29] J. A. Cabrera, H. Boche, C. Deppe, R. F. Schaefer, C. Scheunert, and F. H. Fitzek, "6G and the post-shannon theory," in *Shaping Future 6G Networks: Needs, Impacts and Technologies*, N. O. Frederiksen and H. Gulliksen, Eds. Hoboken, NJ, USA: Wiley, 2021.
[30] P. Schwenteck, G. T. Nguyen, H. Boche, W. Kellerer, and F. H. P. Fitzek, "6G perspective of mobile network operators, manufacturers, and verticals," *IEEE Netw. Lett.*, vol. 5, no. 3, pp. 169–172, Sep. 2023.
[31] R. Ahlswede and G. Dueck, "Identification via channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 15–29, Jan. 1989.
[32] R. H. Muller and C. M. Keck, "Challenges and solutions for the delivery of biotech drugs–A review of drug nanocrystal technology and lipid nanoparticles," *J. Biotechnol.*, vol. 113, nos. 1–3, pp. 151–170, 2004.
[33] R. Jain, "Transport of molecules, particles, and cells in solid tumors," *Annu. Biomed. Eng. Rev.*, vol. 1, no. 1, pp. 241–263, 1999.
[34] S. Wilhelm, A. J. Tavares, Q. Dai, S. Ohta, J. Audet, H. F. Dvorak, and W. C. Chan, "Analysis of nanoparticle delivery to tumours," *Nat. Rev. Mater.*, vol. 1, no. 5, pp. 1–12, 2016.
[35] S. K. Hobbs et al., "Regulation of transport pathways in tumor vessels: Role of tumor type and microenvironment," *Proc. Nat. Acad. Sci.*, vol. 95, no. 8, pp. 4607–4612, 1998.
[36] T. Nakano, T. Suda, Y. Okaie, M. J. Moore, and A. Vasilakos, "Molecular communication among biological nanomachines: A layered architecture and research issues," *IEEE Trans. Nanobiosci.*, vol. 13, no. 3, pp. 169–197, Sep. 2014.
[37] S. Ghavami, "Anomaly detection in molecular communications with applications to health monitoring networks," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 6, no. 1, pp. 50–59, Jul. 2020.
[38] L. B. Buck, "Unraveling the sense of smell (nobel lecture)," *Angew. Chem. Int. Ed.*, vol. 44, no. 38, pp. 6128–6140, 2005.
[39] A. Buettner, *Springer Handbook of Odor*. Cham, Switzerland: Springer, 2017.
[40] V. Jamali, H. M. Loos, A. Buettner, R. Schober, and H. Vincent Poor, "Olfaction-inspired MCs: Molecule mixture shift keying and cross-reactive receptor arrays," *IEEE Trans. Commun.*, vol. 71, no. 4, pp. 1894–1911, Apr. 2023.
[41] M. Baaden, "Deep inside molecules—Digital twins at the nanoscale," *Virtual Real. Intell. Hardw.*, vol. 4, no. 4, pp. 324–341, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2096579622000171
[42] B. R. Barricelli, E. Casiraghi, and D. Fogli, "A survey on digital twin: Definitions, characteristics, applications, and design implications," *IEEE Access*, vol. 7, pp. 167653–167671, 2019.
[43] H. Markram, "The blue brain project," *Nat. Rev. Neurosci.*, vol. 7, no. 2, pp. 153–160, 2006.
[44] J. Chen, C. Yi, S. D. Okegbile, J. Cai, S. Xuemin, and Shen, "Networking architecture and key supporting technologies for human digital twin in personalized healthcare: A comprehensive survey," 2023, *arXiv:2301.03930*.
[45] M. J. Salariseddigh, U. Pereg, H. Boche, C. Deppe, V. Jamali, and R. Schober, "Deterministic identification for molecular communications over the Poisson channel," 2022, *arXiv:2203.02784*.
[46] M. J. Salariseddigh, U. Pereg, H. Boche, and C. Deppe, "Deterministic identification over fading channels," in *Proc. IEEE Inf. Theory Wksp. (ITW)*, 2021, pp. 1–5.
[47] M. J. Salariseddigh, U. Pereg, H. Boche, and C. Deppe, "Deterministic identification over channels with power constraints," *IEEE Trans. Inf. Theory*, vol. 68, no. 1, pp. 1–24, Jan. 2022.
[48] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.
[49] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Zs. Wahrscheinlichkeitstheorie Verw. Geb.*, vol. 44, no. 2, pp. 159–175, 1978.
[50] S. Verdu and V. Wei, "Explicit construction of optimal constant-weight codes for identification via channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 1, pp. 30–36, Jan. 1993.
[51] O. Günlü, J. Kliewer, R. F. Schaefer, and V. Sidorenko, "Code constructions and bounds for identification via channels," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 1486–1496, Mar. 2022.
[52] R. Ahlswede, "General theory of information transfer: Updated," *Discrete Appl. Math.*, vol. 156, no. 9, pp. 1348–1388, 2008.
[53] M. J. Salariseddigh, V. Jamali, H. Boche, C. Deppe, and R. Schober, "Deterministic identification for MC binomial channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2023, pp. 448–453.
[54] R. Ahlswede and N. Cai, "Identification without randomization," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2636–2642, Nov. 1999.
[55] M. J. Salariseddigh, U. Pereg, H. Boche, and C. Deppe, "Deterministic identification over channels with power constraints," in *Proc. IEEE Int. Conf. Commun.*, 2021, pp. 1–6.
[56] M. J. Salariseddigh, U. Pereg, H. Boche, and C. Deppe, "Deterministic identification over channels with power constraints," 2020, *arXiv:2010.04239*.
[57] J. JáJá, "Identification is easier than decoding," in *Proc. Annu. Symp. Found. Comput. Sci.*, 1985, pp. 43–50.
[58] M. V. Burnashev, "On the method of types and approximation of output measures for channels with finite alphabets," *Prob. Inf. Trans.*, vol. 36, no. 3, pp. 195–212, 2000.
[59] Z. Brakerski, Y. T. Kalai, and R. R. Saxena, "Deterministic and efficient interactive coding from hard-to-decode tree codes," in *Proc. 61st IEEE Annu. Symp. Found. Comput. Sci.*, 2020, pp. 446–457.
[60] R. L. Bocchino, V. Adve, S. Adve, and M. Snir, "Parallel programming must be deterministic by default," *Usenix HotPar*, vol. 6, no. 10.5555, pp. 1855591–1855595, 2009.
[61] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.
[62] M. Wiese, W. Labidi, C. Deppe, and H. Boche, "Identification over additive noise channels in the presence of feedback," *IEEE Trans. Inf. Theory*, vol. 69, no. 11, pp. 6811–6821, Nov. 2023.
[63] M. V. Burnashev, "On identification capacity of infinite alphabets or continuous-time channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2407–2414, Nov. 2000.
[64] M. Spahovic, M. J. Salariseddigh, and C. Deppe, "Deterministic k-identification for slow fading channels," in *Proc. IEEE Inf. Theory Wksp. (ITW)*, 2023, pp. 353–358.
[65] M. J. Salariseddigh, M. Spahovic, and C. Deppe, "Deterministic k-identification for slow fading channel," 2022, *arXiv:2212.02732*.
[66] O. Dabbabi, M. J. Salariseddigh, C. Deppe, and H. Boche, "Deterministic k-identification for binary symmetric channel," 2023, *arXiv:2305.04260*.
[67] H. Unterweger et al., "Experimental molecular communication testbed based on magnetic nanoparticles in duct flow," in *Proc. IEEE Int. Works. Sig. Process. Adv. Wireless Commun.*, 2018, pp. 1–5.

[68] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C.-B. Chae, "Three-dimensional channel characteristics for molecular communications with an absorbing receiver," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 929–932, Jun. 2014.
[69] A. Ahmadzadeh, H. Arjmandi, A. Burkovski, and R. Schober, "Comprehensive reactive receiver modeling for diffusive molecular communication systems: Reversible binding, molecule degradation, and finite number of receptors," *IEEE Trans. Nanobiosci.*, vol. 15, no. 7, pp. 713–727, Oct. 2016.
[70] R. Ahlswede, *Identification and Other Probabilistic Models*, A. Ahlswede, I. Althöfer, C. Deppe, and U. Tamm, Eds. Cham, Switzerland: Springer, 2021.
[71] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York, NY, USA: Springer, 2013.
[72] H. Cohn, "Order and disorder in energy minimization," in *Proc. Int. Congr. Math.*, 2010, pp. 2416–2443.
[73] R. Beals and R. Wong, *Special Functions: A Graduate Text*, vol. 126. Cambridge, U.K.: Cambridge Univ. Press, 2010.
[74] W. Feller, *An Introduction to Probability Theory and Its Applications*. New York, NY, USA: Wiley, 1966.
[75] I. M. Copi, C. Cohen, and K. McMahon, *Introduction to Logic*. New York, NY, USA: Routledge, 2016.
[76] T. D. Ahle, "Sharp and simple bounds for the raw moments of the binomial and Poisson distributions," *Stat. Probab. Lett.*, vol. 182, Mar. 2022, Art. no. 109306.
[77] D. Mitrinovic, J. Pecaric, and A. Fink, *Classical and New Inequalities in Analysis*, vol. 61. Dordrecht, The Netherlands: Springer, 2013.
[78] M. J. Salariseddigh, V. Jamali, H. Boche, C. Deppe, and R. Schober, "Deterministic identification for MC binomial channel," 2023, *arXiv:2304.12493*.