Achievable Information Rate Analysis in Diffusive Channels with Memory and Markov Source

Farhad Vakilipoor, *Graduate Student Member, IEEE*, Luca Barletta, *Member, IEEE*, Stefano Bregni, *Member, IEEE* and Maurizio Magarini, *Member, IEEE*

Abstract—**This paper explores the Achievable Information Rate (AIR) of diffusive Molecular Communication (MC) channels featuring a fully absorbing receiver that counts the absorbed particles during symbol time intervals (STIs) and resets the counter at the start of each interval. The MC channel, influenced by memory effect, experiences inter-symbol interference (ISI) arising from the delayed arrival of molecules. The channel's memory is quantified as an integer multiple of the STI and a single-sample memoryless detector is employed to mitigate complexity in computing the mutual information (MI). To maximize MI, the detector threshold is optimized under Gaussian approximation of its input. The channel's MI is calculated, considering the influence of ISI, in the context of binary concentration shift keying modulation. Two distinct scenarios were considered; independent and correlated source-generated symbols, the latter modeled as a first-order Markov process. For each communication scenario, two degrees of knowledge: ISI-Aware and ISI-Unaware were considered. Remarkably, it is demonstrated that employing a correlated source enables the attainment of higher capacity. The results indicate that the capacity-achieving input distribution is not necessarily uniform. Notably, when the STI is small, corresponding to the case of strong ISI, the maximum AIR is not achieved through equiprobable symbol transmission.**

Index Terms—**Diffusion, molecular communication, channel capacity, channel memory, achievable information rate.**

I. Introduction

Molecular Communication (MC) is an interdisciplinary communication paradigm that relies on particle propagation as a means of information transmission. MC has natural and artificial forms. Natural MC, which has evolved over millions of years, has great potential for investigating information exchange in biological systems. On the other hand, artificial MC is a human field that studies communication systems based on the principles of natural MC. One of the advantages of MC is its potential for use in environments where electromagnetic communication is not possible or desirable, such as in targeted drug delivery, nanomedicine, and implantable devices, for which electromagnetic radiation can be harmful or interfere [?], [?]. Various aspects of MC systems have been studied, including active vs. passive receivers, instantaneous vs. continuous release of molecules, and for different boundary conditions of the physical channel [?]. In order to better understand this novel communication paradigm, an analysis from the information theory perspective would give new insights into MC and even improve the system performance in artificial counterparts.

A. Related Literature

Channel capacity serves as a fundamental metric to quantify the ability of a communication system to transmit information reliably from a sender to a receiver, as established by Shannon in his seminal work [?]. When it comes to MC, analyzing channel capacity becomes a necessary undertaking due to various factors, such as inter-symbol interference (ISI) caused by memory effects, energy constraints, slow propagation, and distinctive statistical characteristics [?].

One approach to investigate the capacity limits of molecular communication channels is to encode information through the timing of particle releases, as explored extensively by Rose *et al.* [?]. First, they illustrated that any MC channel can be contextualized either from the timing perspective or the type of particles. Then, they mainly focused on obtaining upper and lower bounds on the capacity of MC timing channels. Lastly, they applied their theory obtained from the particle counting process to DNA and protein sequences. In another study [?], an MC timing channel is introduced, where particles decay after a finite time, and upper and lower bounds on the associated capacity are derived.

Another approach in MC involves encoding information based on the number of particles released at the transmitter and decoding based on the number of received particles during the symbol time interval (STI) [?].

Early investigations into MC channel capacity in the presence of particle intensity considered receivers that do not interact with information particles (IP), which are referred to as transparent receivers [?]. However, in practical scenarios, receivers commonly engage in a reaction process, binding with the IP through a Markov process.

The concentration-based channel, coupled with a ligand receptor-equipped receiver, has been the subject of investigation in previous works, such as those by Einolghozati et al. [?] and Tahmasbi et al. [?]. These studies employ a Markov chain model to capture the reception of molecules by ligand receptors and analyze the channel capacity in terms of bits per channel use. Furthermore, the findings

F. Vakilipoor, L.L. Barletta, S. Bregni and M. Magarini are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, I-20133 Milano, Italy e-mail: farhad.vakilipoor@polimi.it, luca.barletta@polimi.it, stefano.bregni@polimi.it, maurizio.magarini@polimi.it.

performance of correlated sources in terms of AIRs, shown in this work, gives an insight into how to design codes for the molecular diffusive channel and, more in general, for additive Gaussian channels whose variance depends on the transmitted symbols.

We would like to point out that our approach and methodology are valid, not necessarily in MC studies. Actually, it is applicable to any channel with additive memory property under Gaussian statistics when the source is either independent or correlated.

C. Outline

The paper is structured as follows: Sec. ?? introduces the system model, including the calculation of memory, as well as the CIR. Furthermore, the suitability of the Gaussian approximation for channel modeling is discussed, followed by an examination of the transition probabilities of the channel. In Sec. ??, we provide a detailed explanation and formulation of the memoryless channel capacity, AIR, and MI for both the independent and correlated sources. For each source type, two MIs are derived, corresponding to the ISI-Aware and ISI-Unaware scenarios. Section ?? presents numerical results illustrating the capacity and AIR for the four distinct scenarios, considering various STIs and input probabilities. Finally, Sec. ?? concludes the paper by offering final remarks.

D. Notations

The RVs are represented by uppercase italic letters (X), while their realizations are denoted by lowercase italic letters (x). The vector $(x_r, \ldots, x_v)$ is expressed as x_r^v. Specifically, the presence of a superscript indicates that the variable is a vector, while the subscript indicates the index of the first element, and the superscript indicates the index of the last element in the vector. If there is only a subscript, it denotes a single variable with the corresponding index. Additionally, the joint probability of the vector $(x_r, \ldots, x_{v-1})$ and x_v can be written as $P_{X_r^{v-1}, X_v}(x_r^{v-1}, x_v) = P_{X_r^v}(x_r^v)$. The Hamming weight operator applied to a binary vector $x_r^{r+n} \in \{0,1\}^{n+1}$ is denoted as $w_H(x_r^{r+n})$, which counts the number of occurrences of "1" in vector x_r^{r+n}. The operator $\{f\}^+$ is defined as the $\max\{0, f\}$. The Q function and complementary error function are defined as

$$Q(z) = \frac{1}{2}\operatorname{erfc}\left(\frac{z}{\sqrt{2}}\right) = \frac{1}{\sqrt{2\pi}}\int_z^\infty e^{-\frac{y^2}{2}} dy, \quad z \in \mathbb{R} \ . \quad (1)$$

The binary entropy function $H_2 : [0,1] \to [0,1]$ is defined as

$$H_2(x) \triangleq -x\log_2(x) - (1-x)\log_2(1-x) . \quad (2)$$

II. System Model and Analysis

In this section, we undertake a comprehensive characterization of the transmitter, receiver, and the propagation dynamics of the IPs. Subsequently, we proceed to quantify the memory of the system, representing it as an integer multiple of the STI. Following the characterization of the CIR, we proceed to model the received signal and approximate its statistical behavior by employing a Gaussian distribution.

A. Propagation Aspects

This work considers a communication system made of a point transmitter, a diffusion-based channel, and an FA spherical receiver. At the beginning of each STI of duration T_{sym} where "1" is sent, the transmitter sends a pulse corresponding to the instantaneous release of N_{T} IPs. The receiver counts the number of particles absorbed within each STI and resets the counter at the beginning of the next interval. We believe that this mechanism is not far from reality [?]. The IPs diffuse through the medium between transmitter and receiver with constant diffusion coefficient D [$\mu m^2/s$]. In practice, the value of D depends on the temperature, viscosity of the fluid, and the Stokes' radius of the molecule [?].

The receiver's absorption property stems from the reaction between receiver and IPs. In effect, the counting process is tantamount to measure the concentration of desired particles at the receiver, resulting from the interaction between its surface and particles. In a biological environment, enzymes can be secreted by the receiver to eliminate effects resulting from past reactions, thus enabling resetting [?].

The propagation of diffusive particles is governed by the Fick's second law [?], which relates the time derivative of the flux to the Laplacian of the concentration of molecules $c(d,t)$ at a given distance d and time t, as

$$\frac{\partial c(d,t)}{\partial t} = D\nabla^2 c(d,t) . \quad (3)$$

The initial and boundary conditions of (??) vary depending on the MC system characteristics. Authors in [?] specified the boundary and initial conditions for an impulsive release of molecules, an unbounded environment, and an FA spherical receiver. They obtained the expression for the hitting rate of molecules onto the receiver surface, as a function of the distance d between the transmitter and the center of the receiver with radius R at time t. Then, assuming the independent random movement of the particles and the homogeneity of the medium, they derived the expected cumulative number of absorbed particles as

$$N(t) = \frac{N_{\text{T}} R}{d}\operatorname{erfc}\left(\frac{d-R}{2\sqrt{Dt}}\right) . \quad (4)$$

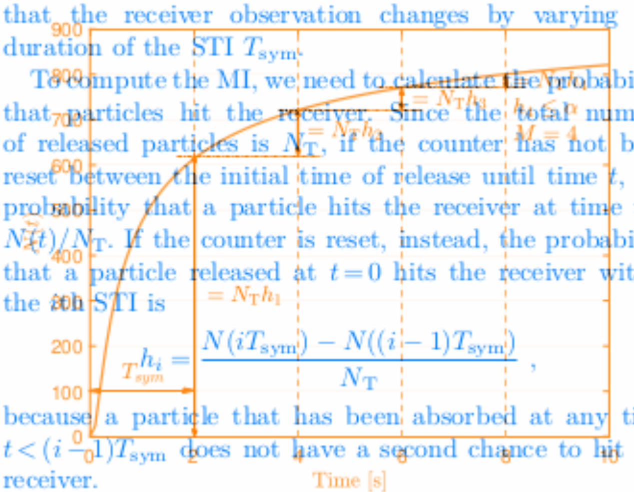
In our study, we consider a binary concentration shift keying (BCSK) modulation, where IPs release corresponds to "1" and no release corresponds to "0". At the receiver, the number of absorbed particles is counted and reset at the beginning of the next interval. At the end of each STI, the receiver returns a single sample, representing the total number of particles absorbed during that interval. Assuming that the receiver resets the counter right at the beginning of STIs (*i.e.* perfect synchronization between transmission and reset intervals at the receiver), we expect that the receiver observation changes by varying the duration of the STI T_{sym}.

To compute the MI, we need to calculate the probability that particles hit the receiver. Since the total number of released particles is N_{T}, if the counter has not been reset between the initial time of release until time t, the probability that a particle hits the receiver at time t is $N(t)/N_{\text{T}}$. If the counter is reset, instead, the probability that a particle released at $t = 0$ hits the receiver within the ith STI is

$$h_i = \frac{N(iT_{\text{sym}}) - N((i-1)T_{\text{sym}})}{N_{\text{T}}} , \quad (5)$$

because a particle that has been absorbed at any time $t < (i-1)T_{\text{sym}}$ does not have a second chance to hit the receiver.

B. Memory Duration Characterization



that the receiver observation changes by varying the duration of the STI T_{sym}.

To compute the MI, we need to calculate the probability that particles hit the receiver. Since the total number of released particles is N_T, if the counter has not been reset between the initial time of release until time t, the probability that a particle hits the receiver at time t is $N(t)/N_T$. If the counter is reset, instead, the probability that a particle released at $t=0$ hits the receiver within the ith STI is

$$h_i = \frac{N(iT_{\text{sym}}) - N((i-1)T_{\text{sym}})}{N_T}, \tag{5}$$

because a particle that has been absorbed at any time $t < (i-1)T_{\text{sym}}$ does not have a second chance to hit the receiver.

Fig. 1: Expected cumulative number of absorbed particles over time without resetting (system parameters as in Tab. ??, $T_{\text{sym}} = 2$ s, $M = 4$). The expected number of absorbed particles within each interval, when the counter is reset, is highlighted by vertical double arrows.

B. Memory Duration Characterization

When studying slow diffusive communication, it is important to quantify the effect of channel memory. To compute MI and transition probabilities between input and detected output, we need to account for all possible combinations of the preceding symbol sequence. If channel memory spans M symbols, there are 2^M different possible sequences to consider. The memory length depends on the transmission rate of symbols. In our model, it should be as small as possible, because evaluating 2^M combinations can make computation impractical. Moreover, due to the differential nature of (??) and asymptotical convergence of (??), the probability of a particle being absorbed a long time after release eventually tends to 0.

To obtain an estimate M of the effective memory length in terms of STIs, being not unnecessarily long or so short to miss the effect of the released particles, we define

$$M = \left\lceil \frac{T_\alpha}{T_{\text{sym}}} \right\rceil, \tag{6}$$

where T_α is the time required to reach some negligible hitting probability α, as given by

$$\frac{R}{d}\left(\operatorname{erfc}\left(\frac{d-R}{2\sqrt{D(T_\alpha + T_{\text{sym}})}}\right) - \operatorname{erfc}\left(\frac{d-R}{2\sqrt{DT_\alpha}}\right)\right) = \alpha. \tag{7}$$

Note that (??) is a transcendental equation with unknown T_α. We do not know an explicit solution for such equation. Hence, we solve it numerically by the regula-falsi method [?].

For example, Fig. ?? plots the expected cumulative number of absorbed particles over time without resetting the counter (blue curve), computed by (??) for a diffusive MC system modeled as above, with parameters set as in Tab. ??, and with STI $T_{\text{sym}} = 2$ s. Here, the memory length results in $M = 4$. The expected number of absorbed particles within each STI, when resetting the counter at its beginning, is also highlighted as the difference of values at interval boundaries. On the other hand, Fig. ?? plots the distribution (??) of probability h_i that particles are absorbed by a resetting receiver within the ith interval for different values of T_{sym}. We observe that the memory length resulting from (??) increases with T_{sym} when measured in time units (T_α), but decreases in terms of STIs (M). The vector $h_1^M = (h_1, \ldots, h_M)$ represents the CIR of the system.

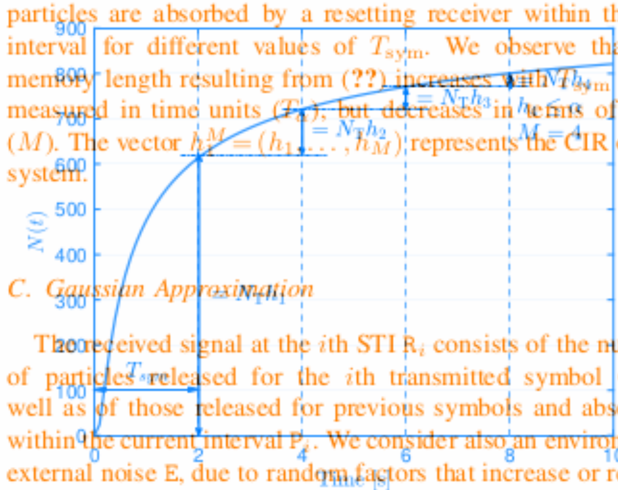


Fig. 1: Expected cumulative number of absorbed particles over time without resetting (blue curve) (system parameters as in Tab. ??, $T_{\text{sym}} = 2$ s, $M = 4$). The expected number of absorbed particles within each interval, when the counter is reset, is highlighted by vertical double arrows.

C. Gaussian Approximation

The received signal at the ith STI $\mathtt{R}_i$ consists of the number of particles released for the ith transmitted symbol $\mathtt{C}_i$ as well as of those released for previous symbols and absorbed within the current interval $\mathtt{P}_i$. We consider also an environment external noise $\mathtt{E}$, due to random factors that increase or reduce the number of particles that the receiver counts in any interval. For example, a negative value of $\mathtt{E}$ expresses the effect of extraneous molecules that unbind absorbed IPs. As obvious, $\mathtt{E}$ is independent of $\mathtt{C}_i$ and $\mathtt{P}_i$. In conclusion, the observation at the ith interval is the superposition of the current signal, of previously transmitted symbols and of external noise, that is,

$$\mathtt{R}_i = \mathtt{C}_i + \mathtt{P}_i + \mathtt{E}. \tag{8}$$

Next, we want to show that we can use a Gaussian model to describe the randomness in the number of absorbed particles in STIs. Due to the nature of the absorption phenomenon and the temporal correlation of the absorption in different STIs, the number of particles absorbed in each STI follows the multinomial distribution. Consider each STI as a bin. Hence, we have M bins that correspond to the STIs in which a particle can reach the receiver. There is also an extra bin that represents the scenario of a particle that has not been absorbed. So in the end, for our statistical model, there are $M+1$ bins. We can write the probability of a particle falling into the ith of the first M bins as h_i and the probability corresponding to the last extra bin as $h_{M+1} = 1 - \sum_{i=1}^{M} h_i$. Let $N_1^{M+1} = (N_1, \ldots, N_{M+1})$ be the number of particles that fell into each bin. Then, N_1^{M+1} is multinomial-distributed over N_T trials and bin probabilities h_1^{M+1}. We can compute the entries of the covariance matrix of N_1^{M+1} as follows.

$$\operatorname{Var}(N_i) = N_T h_i (1 - h_i), \quad i \in \{1, \ldots, M+1\}, \tag{9}$$

$$\operatorname{Cov}(N_i, N_j) = -N_T h_i h_j, \quad (i \neq j) \tag{10}$$

If h_i and h_j are much smaller than 1, then we have

$$\frac{\operatorname{Var}(N_i)}{|\operatorname{Cov}(N_i, N_j)|} = \frac{1 - h_i}{h_j} \gg 1. \tag{11}$$

By Central Limit Theorem (CLT), for N_T sufficiently large, the random vector N_1^{M+1} is approximately Gaussian distributed. However, our focus is on characterizing the joint distribution of the first M bins. so the vector N_1^M is approximately Gaussian distributed with mean vector $N_T h_1^M$ and covariance matrix that is approximately diagonal thanks to (??). This is equivalent to having virtually independent and Gaussian-distributed marginals as follows

$$N_i \sim \mathcal{N}\left(N_T h_i, N_T h_i (1 - h_i)\right) \quad i \in \{1, \ldots, M\} \tag{12}$$

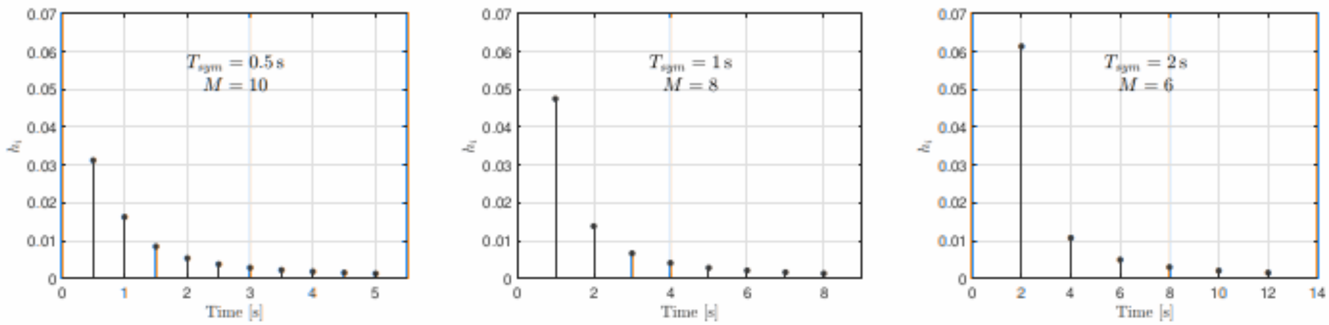


Fig. 2: Distribution (??) of the probability that particles are absorbed by a resetting receiver within the ith interval for $T_{\rm sym} = 0.5, 1, 2$ s. The channel memory length, resulting from (??) for $h \geq 0.001$, increases with $T_{\rm sym}$ when measured in time units (T_c), but decreases in terms of STIs (M).

the ith of the first M bins as h_i, and the probability corresponding to the last extra bin as $h_{M+1} = 1 - \sum_{i=1}^{M} h_i$. Let $N_1^{M+1} = (N_1, \ldots, N_{M+1})$ be the number of particles that fall into each bin. Then N_1^{M+1} is multinomial distributed over N_T trials and with probabilities h_1^{M+1}. We can compute the entries of the covariance matrix of N_1^{M+1} as follows

$$\mathrm{Var}(N_i) = N_{\rm T} h_i (1 - h_i) , \qquad i \in \{1, \ldots, M+1\} , \tag{9}$$

$$\mathrm{Cov}(N_i, N_j) = -N_{\rm T} h_i h_j , \qquad (i \neq j) . \tag{10}$$

If h_i and h_j are much smaller than 1, then we have

$$\frac{\mathrm{Var}(N_i)}{|\mathrm{Cov}(N_i, N_j)|} = \frac{1 - h_i}{h_j} \gg 1 . \tag{11}$$

By Central Limit Theorem (CLT), for $N_{\rm T}$ sufficiently large, the random vector N_1^{M+1} is approximately Gaussian distributed. However, our focus is on characterizing the joint distribution of the first M bins: so the vector N_1^M is approximately Gaussian distributed with mean vector $N_{\rm T} h_1^M$ and covariance matrix that is approximately diagonal thanks to (11). This is equivalent to having virtually independent and Gaussian-distributed marginals as follows

$$N_i \sim \mathcal{N}\Big(N_{\rm T} h_i, N_{\rm T} h_i (1 - h_i)\Big), \qquad i \in \{1, \ldots, M\}. \tag{12}$$

In practice, to be this approximation valid, the probability that the Gaussian distribution generates negative values should be negligible. That is, the model parameters should be chosen to have the mean μ and standard deviation σ of the Gaussian distribution satisfying e.g. $\mu > 3\sigma$, which implies

$$\frac{N_{\rm T} h_i}{1 - h_i} > 9 . \tag{13}$$

One of the necessary parts of a communication system is the detector that attempts to recover the transmitted symbol from the received signal. In this paper, we consider a memoryless binary detector with the rule

$$\hat{S}_i = \begin{cases} 1 & \text{if } \mathtt{R}_i \geq \tau , \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

Note that, for each realization of the channel, such as input symbol distribution, and STIs, we always look for the threshold τ that maximizes the MI.

Let $s_i \in \{0, 1\}$ denote the transmitted symbol associated with the ith STI and $g(\omega; \mu, \sigma^2)$ be a Gaussian probability density function (pdf) with mean μ and variance σ^2, where $g(\omega; 0, 0) = \delta(\omega)$ is the Dirac delta function. Then, the pdf of the current signal conditioned on a specific realization of the current transmitted symbol ($S_i = s_i$) can be written as

$$f_{\mathtt{C}_i|S_i=s_i}(\omega) = g\Big(\omega; s_i N_{\rm T} h_1, s_i N_{\rm T} h_1 (1 - h_1)\Big) . \tag{15}$$

Let the vector $s_{i-M+1}^{i-1} \in \{0, 1\}^{M-1}$ be a realization of $(S_{i-M+1}, \ldots, S_{i-1})$, that is the $M-1$ symbols preceding the ith interval. The conditional pdf of particles released in the past $M-1$ intervals and absorbed within the ith interval, given the sequence of preceding symbols, is

$$f_{\mathtt{P}_i|S_{i-M+1}^{i-1}=s_{i-M+1}^{i-1}}(\omega) = g\left(\omega; N_{\rm T} \sum_{j=2}^{M} s_{i-j+1} h_j, N_{\rm T} \sum_{j=2}^{M} s_{i-j+1} h_j (1 - h_j)\right) . \tag{16}$$

The external noise is assumed to follow a time-independent Gaussian distribution with pdf

$$f_{\mathtt{E}}(\omega) = g\Big(\omega; \mu_{\mathtt{E}}, \sigma_{\mathtt{E}}^2\Big). \tag{17}$$

To obtain the conditional pdf of the received signals, two pdfs need to be considered first. The first is the pdf of the number of particles received in the ith interval but released in previous intervals given the previously transmitted symbols, including the environment noise, i.e. $(\mathtt{P}_i + \mathtt{E})|s_{i-M+1}^{i-1}$. The second is the pdf of particles received in the ith interval, released in the current and previous intervals given the previously transmitted symbols and the current symbol, including the external noise, that is $(\mathtt{C}_i + \mathtt{P}_i + \mathtt{E})|s_{i-M+1}^{i}$. As all the involved RVs are Gaussian and conditionally independent, their sum results in a Gaussian RV with a mean and variance that is the sum of means and variances, respectively, and we obtain

$$f_{\mathtt{P}_i+\mathtt{E}|s_{i-M+1}^{i-1}}(\omega)=g\left(\omega;\mu_{\mathtt{E}}+N_{\mathrm{T}}\sum_{j=2}^{M}s_{i-j+1}h_j,\ \sigma_{\mathtt{E}}^2+N_{\mathrm{T}}\sum_{j=2}^{M}s_{i-j+1}h_j(1-h_j)\right), \tag{18}$$

$$f_{\mathtt{C}_i+\mathtt{P}_i+\mathtt{E}|s_{i-M+1}^{i}}(\omega)=g\left(\omega;\mu_{\mathtt{E}}+N_{\mathrm{T}}\sum_{j=1}^{M}s_{i-j+1}h_j,\ \sigma_{\mathtt{E}}^2+N_{\mathrm{T}}\sum_{j=1}^{M}s_{i-j+1}h_j(1-h_j)\right). \tag{19}$$

Thus, the channel transition probabilities given a specific sequence of symbols can be written as

$$P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(1|s_{i-M+1}^{i-1},0)=\Pr(\mathtt{P}_i+\mathtt{E}\geq\tau|s_{i-M+1}^{i-1})=Q\left(\frac{\tau-\mu_{\mathtt{E}}-N_{\mathrm{T}}\sum_{j=2}^{M}s_{i-j+1}h_j}{\sqrt{\sigma_{\mathtt{E}}^2+N_{\mathrm{T}}\sum_{j=2}^{M}s_{i-j+1}h_j(1-h_j)}}\right), \tag{20}$$

$$P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(0|s_{i-M+1}^{i-1},0)=\Pr(\mathtt{P}_i+\mathtt{E}<\tau|s_{i-M+1}^{i-1})=1-\Pr(\mathtt{P}_i+\mathtt{E}\geq\tau|s_{i-M+1}^{i-1}), \tag{21}$$

$$P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(1|s_{i-M+1}^{i-1},1)=\Pr(\mathtt{C}_i+\mathtt{P}_i+\mathtt{E}\geq\tau|s_{i-M+1}^{i})=Q\left(\frac{\tau-\mu_{\mathtt{E}}-N_{\mathrm{T}}\sum_{j=1}^{M}s_{i-j+1}h_j}{\sqrt{\sigma_{\mathtt{E}}^2+N_{\mathrm{T}}\sum_{j=1}^{M}s_{i-j+1}h_j(1-h_j)}}\right), \tag{22}$$

$$P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(0|s_{i-M+1}^{i-1},1)=\Pr(\mathtt{C}_i+\mathtt{P}_i+\mathtt{E}<\tau|s_{i-M+1}^{i})=1-\Pr(\mathtt{C}_i+\mathtt{P}_i+\mathtt{E}\geq\tau|s_{i-M+1}^{i}). \tag{23}$$

III. Achievable Information Rate Analysis

In this section, we define and derive the MI, AIR, and capacity for two different types of sources and two degrees of knowledge about the preceding transmitted symbols at the receiver side. The sources under investigation are a correlated source that generates symbols based on the first-order Markov process, and a source that generates symbols independently. For each source, we derive two MIs that correspond to different degrees of knowledge about the ISI at the receiver side. The first case, termed ISI-Aware scenario, assumes that the receiver knows the previously transmitted symbols within the memory intervals, which is equivalent to knowing the terms contributing to ISI. The second scenario assumes that the receiver is uncertain about the previously transmitted symbols, corresponding to not knowing the ISI terms. This is referred as ISI-Unaware scenario. Before delving into these four scenarios, we need to establish the formalism for calculating the MI, AIR, and channel capacity. Actually, a unique feature of the CIR introduced in Sec. ?? and shown in Fig. ??, was its variation based on T_{sym}. Calculation of the AIR facilitates the exploration of optimal T_{sym} that allows us to transmit information at the highest possible rate per unit of time.

The channel capacity, C, is defined as the maximization of the AIR over the input distributions, P_S:

$$C=\max_{P_S}\mathrm{AIR}. \tag{24}$$

The AIR is defined as the MI between the input process $S=(S_1,S_2,\dots)$ and output process $\hat{S}=(\hat{S}_1,\hat{S}_2,\dots)$ divided by the STI T_{sym}:

$$\mathrm{AIR}=\frac{I(S;\hat{S})}{T_{\mathrm{sym}}}. \tag{25}$$

The MI between the input process $S=(S_1,S_2,\dots)$ and output process $\hat{S}=(\hat{S}_1,\hat{S}_2,\dots)$ of time-invariant channel with memory can be written as [?]

$$I(S;\hat{S})=\lim_{n\to\infty}\frac{1}{n}I(S_1,\dots,S_n;\hat{S}_1,\dots,\hat{S}_n)=\lim_{n\to\infty}\frac{1}{n}I(S_1^n;\hat{S}_1^n). \tag{26}$$

The expression given in (??) evaluates the average of the MI per channel use. In more detail, the MI is calculated between the input-output processes across an infinitely long symbols sequence and then divided by the length of the sequence.

The MI can be written as the difference between the entropy of the input sequence and the entropy of the input given the output sequence

$$I(S^n;\hat{S}^n)=H(S^n)-H(S^n|\hat{S}^n). \tag{27}$$

From the chain rule of entropy, we can write (??) as

$$I(S_1^n;\hat{S}_1^n)=\sum_{i=1}^{n}H(S_i|S_1^{i-1})-H(S_i|S_1^{i-1},\hat{S}_1^n). \tag{28}$$

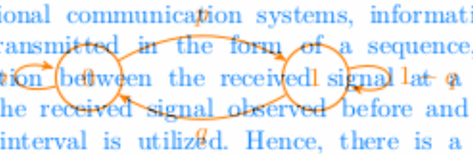
In conventional communication systems, information is typically transmitted in the form of a sequence, and the correlation between the received signal at a given time and the received signal observed before and after that time interval is utilized. Hence, there is a dual-side correlation exploitation regardless of time. However, in this study, we employ the causality assumption for detection in the communication system. Consequently, we discard the estimated samples following the ith interval (i.e., current time).

$$\sum_{i=1}^{n}H(S_i|S_1^{i-1})-H(S_i|S_1^{i-1},\hat{S}_i)\leq I(S_1^n;\hat{S}_1^n). \tag{29}$$

Once the previously transmitted symbols, S_1^{i-1}, are known, there is no longer any uncertainty in the previously detected symbols, $\hat{S}_1^{i-1}$. Consequently, we can safely discard the previously detected symbols as they contain redundant information.

A. Correlated Source

One of the primary obstacles ... analyzing the impact of bursty symbol transmission. To delve into the MI of the communication system when symbols are transmitted in a bursty fashion, we consider a source with memory that generates symbols while taking into account temporal correlations between them. In this study, we employ a first-order Markov process to model the correlated source, as illustrated in the following representation



where the transition probabilities are $1-p=P_{S_i|S_{i-1}}(0|0)$, $p=P_{S_i|S_{i-1}}(1|0)$, $q=P_{S_i|S_{i-1}}(0|1)$, and $1-q=P_{S_i|S_{i-1}}(1|1)$. Mathematically, the temporal evolution of the Markov source shown in the preceding representation can be expressed as

$$\begin{bmatrix}P_{S_i}(0)\\P_{S_i}(1)\end{bmatrix}=\begin{bmatrix}1-p & q\\ p & 1-q\end{bmatrix}\begin{bmatrix}P_{S_{i-1}}(0)\\P_{S_{i-1}}(1)\end{bmatrix}. \tag{30}$$

Accordingly, one can compute the asymptotical/stationary probability of the two symbols (i.e., the probability of "1"s and "0"s) generated by the source in a sufficiently long sequence as

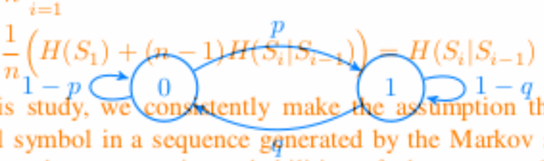
$$\pi_0=\frac{q}{p+q}, \tag{31}$$

$$\pi_1=\frac{p}{p+q}. \tag{32}$$

redundant information.

A. Correlated Source

One of the primary obstacles in communication systems involving memory lies in analyzing the impact of bursty symbol transmission. To delve into the MI of the communication system when symbols are transmitted in a bursty fashion, we consider a source with memory that generates symbols while taking into account temporal correlations between them. In this study, we employ a first-order Markov process to model the correlated source, as illustrated in the following representation



where the transition probabilities are $1-p = P_{S_i|S_{i-1}}(0|0)$, $p = P_{S_i|S_{i-1}}(1|0)$, $q = P_{S_i|S_{i-1}}(0|1)$, and $1-q = P_{S_i|S_{i-1}}(1|1)$. Mathematically, the temporal evolution of the Markov source shown in the preceding representation can be expressed as

$$\begin{bmatrix} P_{S_i}(0) \\ P_{S_i}(1) \end{bmatrix} = \begin{bmatrix} 1-p & q \\ p & 1-q \end{bmatrix} \begin{bmatrix} P_{S_{i-1}}(0) \\ P_{S_{i-1}}(1) \end{bmatrix} . \tag{30}$$

Accordingly, one can compute the asymptotical/stationary probability of the two symbols (i.e. the probability of "1"s and "0"s) generated by the source in a sufficiently long sequence as

$$\pi_0 = \frac{q}{p+q} , \tag{31}$$

$$\pi_1 = \frac{p}{p+q} . \tag{32}$$

In the case of the first-order Markov source, we can simplify the sum of the entropy of the sequence generated by the source as

$$\begin{aligned} \sum_{i=1}^{n} H(S_i|S_1^{i-1}) &= H(S_1) + \sum_{i=2}^{n} H(S_i|S_1^{i-1}) \\ &= H(S_1) + (n-1)H(S_i|S_1^{i-1}) \\ &= H(S_1) + (n-1)H(S_i|S_{i-1}) . \end{aligned} \tag{33}$$

Substituting (??) into the definition of entropy of the process (??) and taking the limit, we can write the average entropy of the source as

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} H(S_i|S_1^{i-1}) = \\ &\lim_{n\to\infty} \frac{1}{n} \Big(H(S_1) + (n-1)H(S_i|S_{i-1}) \Big) = H(S_i|S_{i-1}) . \end{aligned} \tag{34}$$

In this study, we consistently make the assumption that the initial symbol in a sequence generated by the Markov source follows the asymptotic probabilities of the source. Consequently, we express the entropy of the Markov source as the mean of the two conditional entropies.

$$\begin{aligned} H(S_i|S_{i-1}) = &P_{S_{i-1}}(0)H(S_i|S_{i-1}=0) + \\ &P_{S_{i-1}}(1)H(S_i|S_{i-1}=1) = \pi_0 H_2(p) + \pi_1 H_2(q) . \end{aligned} \tag{35}$$

1) ISI-Aware: In this scenario, we consider the presence of a receiver which has a comprehensive knowledge about the previously transmitted symbols. This knowledge is tantamount to being aware of the factors responsible for ISI. Analogous to the source entropy, we can express the sum of conditional entropy of the transmitted symbol given the previously transmitted sequence and the currently estimated symbol as the sum of individual entropies.

$$\begin{aligned} &\sum_{i=1}^{n} H(S_i|S_1^{i-1}, \hat{S}_i) = \\ &H(S_1|\hat{S}_1) + H(S_2|S_1^2, \hat{S}_2) + \cdots + \sum_{i=M}^{n} H(S_i|S_1^{i-1}, \hat{S}_i) = \\ &H(S_1|\hat{S}_1) + H(S_2|S_1^2, \hat{S}_2) + \cdots + (n-M)H(S_i|S_1^{i-1}, \hat{S}_i) \end{aligned} \tag{36}$$

Applying the limit to take the average results in

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} H(S_i|S_1^{i-1}, \hat{S}_i) = \\ &\lim_{n\to\infty} \frac{1}{n} \Big(H(S_1|\hat{S}_1) + H(S_2|S_1^2, \hat{S}_2) + \cdots \\ &+ (n-M)H(S_i|S_1^{i-1}, \hat{S}_i) \Big) = H(S_i|S_1^{i-1}, \hat{S}_i) . \end{aligned} \tag{37}$$

Considering the finite length of the memory interval as previously defined in Sec. ??, we can disregard the level of surprise associated with symbols transmitted significantly earlier. Consequently, we discard the symbols transmitted prior to the effective memory interval.

$$H(S_i|S_1^{i-1}, \hat{S}_i) \le H(S_i|S_{i-M+1}^{i-1}, \hat{S}_i) . \tag{38}$$

The entropy of the current symbol, conditioned on the previously transmitted symbols and the currently estimated symbol, can be expressed as the marginalization over the realizations of the previously transmitted symbols and the currently estimated symbol, yielding

$$\begin{aligned} &H(S_i|S_{i-M+1}^{i-1}, \hat{S}_i) = \\ &\sum_{\forall s_{i-M+1}^{i-1}, \hat{s}_i} P_{S_{i-M+1}^{i-1}, \hat{S}_i}(s_{i-M+1}^{i-1}, \hat{s}_i) H(S_i|s_{i-M+1}^{i-1}, \hat{s}_i) , \end{aligned} \tag{39}$$

$$\begin{aligned} &H(S_i|s_{i-M+1}^{i-1}, \hat{s}_i) = \\ &-\sum_{\forall s_i} P_{S_i|S_{i-M+1}^{i-1}, \hat{S}_i}(s_i|s_{i-M+1}^{i-1}, \hat{s}_i) \times \\ &\log_2 \Big(P_{S_i|S_{i-M+1}^{i-1}, \hat{S}_i}(s_i|s_{i-M+1}^{i-1}, \hat{s}_i) \Big) . \end{aligned} \tag{40}$$

The conditional probability of realizations in (??) can be computed as follows (see Appendix ??)

$$\begin{aligned} &P_{S_i|S_{i-M+1}^{i-1}, \hat{S}_i}(s_i|s_{i-M+1}^{i-1}, \hat{s}_i) = \\ &\frac{P_{\hat{S}_i|S_{i-M+1}^{i}}(\hat{s}_i|s_{i-M+1}^{i}) P_{S_i|S_{i-1}}(s_i|s_{i-1})}{\sum_{x\in\{0,1\}} P_{\hat{S}_i|S_{i-M+1}^{i}}(\hat{s}_i|s_{i-M+1}^{i-1}, x) P_{S_i|S_{i-1}}(x|s_{i-1})} \end{aligned} \tag{41}$$

where the conditional and joint probabilities of a specific estimated symbol and previously transmitted symbols are

$$\begin{aligned} &P_{\hat{S}_i|S_{i-M+1}^{i-1}}(\hat{s}_i|s_{i-M+1}^{i-1}) = \\ &\sum_{\forall s_i} P_{\hat{S}_i|S_{i-M+1}^{i}}(\hat{s}_i|s_{i-M+1}^{i-1}, s_i) P_{S_i|S_{i-1}}(s_i|s_{i-1}) , \end{aligned} \tag{42}$$

$$\begin{aligned} &P_{\hat{S}_i, S_{i-M+1}^{i-1}}(\hat{s}_i, s_{i-M+1}^{i-1}) = \\ &P_{\hat{S}_i|S_{i-M+1}^{i-1}}(\hat{s}_i|s_{i-M+1}^{i-1}) P_{S_{i-M+1}^{i-1}}(s_{i-M+1}^{i-1}) . \end{aligned} \tag{43}$$

Based on (??), the calculation of the probability for a given symbol sequence generated by the source requires knowledge of the Markov model. Since we are utilizing a Markov source as described in (??), the probability of a specific sequence can be determined by traversing the sequence through the Markov model. It is important to note that the probability of the first element in a sequence is assumed to correspond to the asymptotic probability of the Markov source for that particular symbol realization.

$$P_{S_r^u}(s_r^u) = \pi_0^{1-s_r} \pi_1^{s_r} \prod_{j=r+1}^{u} P_{S_j|S_{j-1}}(s_j|s_{j-1}) . \tag{44}$$

In the end, the MI associated with the correlated source and ISI awareness is (??). Note that the MI cannot have a negative value, and what we are computing in this paper is equivalent to lower bounds to the actual MIs due to the assumptions such as causality, effective memory, etc. Hence, we only take into account the positive values of MIs.

2) ISI-Unaware: In this particular scenario, we make the assumption that the receiver does not have any knowledge regarding the symbols transmitted prior to the current time. This assumption is equivalent to loosening the bound on MI. Consequently, by disregarding the information pertaining to

$$I_{\text{ISIA}}^{\text{CRR}} = \Bigg\{ \pi_0 H_2(p) + \pi_1 H_2(q) + \sum_{\forall s_{i-M+1}^{i-1}, \hat{s}_i} \Bigg[\sum_{\forall s_i} \Big[P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},s_i) P_{S_i|S_{i-1}}(s_i|s_{i-1}) \Big] \pi_0^{1-s_{i-M+1}} \times \pi_1^{s_{i-M+1}} \prod_{j=i-M+2}^{i-1} \Big[P_{S_j|S_{j-1}}(s_j|s_{j-1}) \Big] \sum_{\forall s_i} \Bigg[\frac{P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},s_i) P_{S_i|S_{i-1}}(s_i|s_{i-1})}{\sum_{x\in\{0,1\}} P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},x) P_{S_i|S_{i-1}}(x|s_{i-1})} \times \log_2\Bigg(\frac{P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},s_i) P_{S_i|S_{i-1}}(s_i|s_{i-1})}{\sum_{x\in\{0,1\}} P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},x) P_{S_i|S_{i-1}}(x|s_{i-1})} \Bigg) \Bigg] \Bigg] \Bigg\}^{+} . \quad (45)$$

The conditional probability of realizations in (??) can be computed as follows (see Appendix ??)

$$P_{S_i|S_{i-M+1}^{i-1},\hat{S}_i}(s_i|s_{i-M+1}^{i-1},\hat{s}_i) = \frac{P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},s_i) P_{S_i|S_{i-1}}(s_i|s_{i-1})}{\sum_{x\in\{0,1\}} P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},x) P_{S_i|S_{i-1}}(x|s_{i-1})} , \quad (41)$$

where the conditional and joint probabilities of a specific estimated symbol and previously transmitted symbols are

$$P_{\hat{S}_i|S_{i-M+1}^{i-1}}(\hat{s}_i|s_{i-M+1}^{i-1}) = \sum_{\forall s_i} P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},s_i) P_{S_i|S_{i-1}}(s_i|s_{i-1}) , \quad (42)$$

$$P_{\hat{S}_i,S_{i-M+1}^{i-1}}(\hat{s}_i, s_{i-M+1}^{i-1}) = P_{\hat{S}_i|S_{i-M+1}^{i-1}}(\hat{s}|s_{i-M+1}^{i-1}) P_{S_{i-M+1}^{i-1}}(s_{i-M+1}^{i-1}) . \quad (43)$$

Based on (??), the calculation of the probability for a given symbol sequence generated by the source requires knowledge of the Markov model. Since we are utilizing a Markov source as described in (??), the probability of a specific sequence can be determined by traversing the sequence through the Markov model. It is important to note that the probability of the first element in a sequence is assumed to correspond to the asymptotic probability of the Markov source for that particular symbol realization:

$$P_{S_r^v}(s_r^v) = \pi_0^{1-s_r} \pi_1^{s_r} \prod_{j=r+1}^{v} P_{S_j|S_{j-1}}(s_j|s_{j-1}) . \quad (44)$$

In the end, the MI associated with the correlated source and ISI awareness is (??). Note that the MI cannot have a negative value, and what we are computing in this paper is equivalent to lower bounds to the actual MIs due to the assumptions such as causality, effective memory, etc. Hence, we only take into account the positive values of MIs.

2) ISI-Unaware: In this particular scenario, we make the assumption that the receiver does not have any knowledge regarding the symbols transmitted prior to the current time. This assumption is equivalent to loosening the bound on MI. Consequently, by disregarding the information pertaining to previously transmitted symbols, we can establish the following inequality

$$H(S_i|S_{i-M+1}^{i-1}, \hat{S}_i) \leq H(S_i|\hat{S}_i) , \quad (46)$$

and from the definition of average conditional entropy, we write

$$H(S_i|\hat{S}_i) = -\sum_{\forall s_i,\hat{s}_i} P_{S_i,\hat{S}_i}(s_i,\hat{s}_i) \log_2\left(P_{S_i|\hat{S}_i}(s_i|\hat{s}_i)\right) , \quad (47)$$

$$P_{S_i,\hat{S}_i}(s_i,\hat{s}_i) = P_{\hat{S}_i|S_i}(\hat{s}_i|s_i) P_{S_i}(s_i) . \quad (48)$$

The conditional probability of the detected symbol given the transmitted symbol is obtained by marginalizing over the previously transmitted symbols using the Bayes theorem.

$$P_{\hat{S}_i|S_i}(\hat{s}_i|s_i) = \sum_{\forall s_{i-M+1}^{i-1}} P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},s_i) \frac{P_{S_{i-M+1}^{i}}(s_{i-M+1}^{i})}{P_{S_i}(s_i)} . \quad (49)$$

Substituting (??) into (??), the term corresponding to the currently transmitted symbol cancel out, and we obtain

$$P_{S_i,\hat{S}_i}(s_i,\hat{s}_i) = \sum_{\forall s_{i-M+1}^{i-1}} P_{\hat{S}_i|S_{i-M+1}^{i-1},S_i}(\hat{s}_i|s_{i-M+1}^{i-1},s_i) P_{S_{i-M+1}^{i}}(s_{i-M+1}^{i}) . \quad (50)$$

To compute the conditional probability of the current transmitted symbol given the estimated one, we apply the Bayes rule

$$P_{S_i|\hat{S}_i}(s_i|\hat{s}_i) = \frac{P_{S_i,\hat{S}_i}(s_i,\hat{s}_i)}{P_{\hat{S}_i}(\hat{s}_i)} . \quad (51)$$

The probability of the estimated symbol can be computed from the marginalization of the joint probability over all possible realizations of the transmitted symbols.

$$P_{\hat{S}_i}(\hat{s}_i) = \sum_{\forall s_i} P_{S_i,\hat{S}_i}(s_i,\hat{s}_i) . \quad (52)$$

In the end, one can compute the MI corresponding to the correlated source with ISI unawareness on the receiver side as (??).

B. Independent Source

Another type of source that is considered in this paper from a statistical perspective is the one where symbols are generated independently with specific probabilities. Let $\lambda_1 = P_{S_i}(1)$ and $\lambda_0 = 1 - \lambda_1 = P_{S_i}(0)$ denote the probabilities of transmitting symbols "1" and "0", respectively. In this scenario, there is no temporal dependency between the symbols generated by the source, *i.e.*,

$$P_{S_1^n}(s_1^n) = \prod_{i=1}^{n} P_{S_i}(s_i) . \quad (54)$$

As a result, we can discard the conditioning on the previously transmitted symbol, and the entropy of the source simplifies to a binary entropy function.

$$H(S_i|S_1^{i-1}) = H(S_i) = H_2(\lambda_0) . \quad (55)$$

1) ISI-Aware: Similarly to the scenario with the correlated source, we can rely on Eqns. (??)-(??). Noting that the source is independent, a main difference in this case in comparison to the correlated source scenario is the probability of a particular sequence of symbols, and we write it as

$$P_{S_r^v}(s_r^v) = \lambda_0^{v-r+1-w_H(S_r^v)} \lambda_1^{w_H(S_r^v)} . \quad (56)$$

The MI for the case of an independent source with knowledge about the previously transmitted symbols (*i.e.*, ISI-Aware) is (??).

2) ISI-Unaware: Without the knowledge of previously transmitted symbols, the equations derived in Sec. ?? remain applicable. Nevertheless, it is necessary to calculate the probability of each specific sequence using (??), considering the independent nature of symbol generation by the source. Consequently, the MI of the independent source, under the condition of unknown previously transmitted symbols, can be extracted (??).

IV. NUMERICAL EVALUATION AND RESULTS

We present a selection of results that illustrate the superiority of correlated sources in achieving higher capacity. It should be noted that the optimal input distribution for achieving capacity may not be uniform. The numerical evaluation was conducted using system parameters listed in Table ??, obtained from [?], with the exception of the external noise and α. We intentionally selected noise standard deviation σ_{ext} and mean μ_{ext} such that there are instances where the values of E

$$I_{\text{ISIU}}^{\text{CRR}} = \Bigg\{ \pi_0 H_2(p) + \pi_1 H_2(q) + \sum_{\forall s_i, \hat{s}_i} \Bigg[\sum_{\forall s_{i-M+1}^{i-1}} \Bigg[P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i|s_{i-M+1}^{i-1}, s_i) \pi_0^{1-s_{i-M+1}} \pi_1^{s_{i-M+1}} \prod_{j=i-M+2}^{i} P_{S_j|S_{j-1}}(s_j|s_{j-1}) \Bigg] \times$$

$$\log_2 \Bigg(\frac{\sum_{\forall s_{i-M+1}^{i-1}} \Big[P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i|s_{i-M+1}^{i-1}, s_i) \pi_0^{1-s_{i-M+1}} \pi_1^{s_{i-M+1}} \prod_{j=i-M+2}^{i} P_{S_j|S_{j-1}}(s_j|s_{j-1}) \Big]}{\sum_{x \in \{0,1\}} \Big[\sum_{\forall s_{i-M+1}^{i-1}} P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i|s_{i-M+1}^{i-1}, x) P_{S_i|S_{i-1}}(x|s_{i-1}) \pi_0^{1-s_{i-M+1}} \pi_1^{s_{i-M+1}} \prod_{j=i-M+2}^{i-1} P_{S_j|S_{j-1}}(s_j|s_{j-1}) \Big]} \Bigg) \Bigg] \Bigg\}^{+}. \quad (53)$$

B. Independent Source

Another type of source that is considered in this paper from a statistical perspective is the one where symbols are generated independently with specific probabilities. Let $\lambda_1 = P_{S_i}(1)$ and $\lambda_0 = 1 - \lambda_1 = P_{S_i}(0)$ denote the probabilities of transmitting symbols "1" and "0", respectively. In this scenario, there is no temporal dependency between the symbols generated by the source, i.e.,

$$P_{S_1^n}(s_1^n) = \prod_{i=1}^{n} P_{S_i}(s_i). \quad (54)$$

As a result, we can discard the conditioning on the previously transmitted symbol, and the entropy of the source simplifies to a binary entropy function.

$$H(S_i|S_1^{i-1}) = H(S_i) = H_2(\lambda_0). \quad (55)$$

1) ISI-Aware: Similarly to the scenario with the correlated source, we can rely on Eqns. (??)-(??). Noting that the source is independent, a main difference in this case in comparison to the correlated source scenario is the probability of a particular sequence of symbols, and we write it as

$$P_{S_r^v}(s_r^v) = \lambda_0^{v-r+1-w_H(s_r^v)} \lambda_1^{w_H(s_r^v)}. \quad (56)$$

The MI for the case of an independent source with knowledge about the previously transmitted symbols (i.e., ISI-Aware) is (??).

2) ISI-Unaware: Without the knowledge of previously transmitted symbols, the equations derived in Sec. ?? remain applicable. Nevertheless, it is necessary to calculate the probability of each specific sequence using (??). On the other hand, the independent nature of symbol generation by the source... Consequently, the MI of the independent source, under the condition of unknown previously transmitted symbols, can be extracted (??).

IV. Numerical Evaluation and Results

We present a selection of results that illustrate the superiority of correlated sources in achieving higher capacity. It should be noted that the optimal input distribution for achieving capacity may not be uniform. The numerical evaluation was conducted using system parameters listed in Table ??, obtained from [?], with the exception of the external noise and α. We intentionally selected noise standard deviation σ_{ext} and mean μ_{ext} such that there are instances where the values of E become negative, indicating that the external noise impedes IPs absorption. The parameter α is chosen to ensure the validity of the last sample of the CIR as (??).

Figure ?? illustrates the channel capacity (??) for various STIs ($0.2 \leq T_{\text{sym}} \leq 1.5$) in different scenarios. These scenarios include ISI-Aware with a correlated source (blue curve with square marker), ISI-Unaware with a correlated source (blue curve with triangle marker), ISI-Aware with an independent source (red curve with square marker), and ISI-Unaware with independent source (red curve with triangle marker).

As expected, the capacity with ISI awareness is generally higher than that with ISI unawareness. Interestingly, the correlated source achieves a higher capacity compared to the independent source. Normally, it is expected that hiring an independent source results in higher capacities in

TABLE I: System parameters

Variable	Definition	Value
N_T	Number of released molecules	10^4
R_r	Radius of the receiver R_r	1 μm
d	Distance between transmitter and center of receiver	10 μm
α	Minimum acceptable probability	0.001
μ_{ext}	Mean of the external noise signal	50
σ_{ext}	Standard deviation of the external noise signal	50
D	Diffusion coefficient for the signaling molecule	79.4 $\mu m^2/s$

TABLE I: System parameters

Variable	Definition	Value
N_T	Number of released molecules	10^4
R_r	Radius of the receiver R_r	1 μm
d	Distance between transmitter and center of receiver	10 μm
α	Minimum acceptable probability	0.001
μ_{ext}	Mean of the external noise signal	50
σ_{ext}	Standard deviation of the external noise signal	50
D	Diffusion coefficient for the signaling molecule	79.4 $\mu m^2/s$

become negative, indicating that the external noise impedes IPs absorption. The parameter α is chosen to ensure the validity of the last sample of the CIR as (??).

Figure ?? illustrates the channel capacity (??) for various STIs ($0.2 \leq T_{\text{sym}} \leq 1.5$) in different scenarios. These scenarios include ISI-Aware with a correlated source (blue curve with square marker), ISI-Unaware with a correlated source (blue curve with triangle marker), ISI-Aware with an independent source (red curve with square marker), and ISI-Unaware with independent source (red curve with triangle marker).

As expected, the capacity with ISI awareness is generally higher than that with ISI unawareness. Interestingly, the correlated source achieves a higher capacity compared to the independent source. Normally, it is expected that hiring an independent source results in higher capacities in communication systems. However, in this scenario, due to the high ISI effect, the correlated source allows us to tackle the problem of the ISI, and the reduction of the source entropy compared to the independent source is worth it. Specifically, the maximum channel capacity in ISI-Aware scenario with correlated source, $C_{\text{ISIA}}^{\text{CRR}}$, is 1.50 [bit/s] at $T_{\text{sym}} = 0.40$ s, with the input probability distribution $p = 0.60$ and $q = 0.62$. On the other hand, the maximum capacity for the ISI-Aware scenario with independent source, $C_{\text{ISIA}}^{\text{IND}}$, is 1.43 [bit/s] at $T_{\text{sym}} = 0.45$ s, with an input probability distribution of $\lambda_0 = 0.52$.

Comparing the two maximum capacities in the ISI-Aware scenario, we observe that the independent source achieves its maximum at a higher T_{sym} compared to the correlated source.

Moving on to the ISI-Unaware scenario, the maximum capacity for the correlated source, $C_{\text{ISIU}}^{\text{CRR}}$, is 1.24 [bit/s] at $T_{\text{sym}} = 0.57$ s, with an optimum input probability distribution at $p = 0.60$ and $q = 0.60$. The maximum capacity for the independent source, $C_{\text{ISIU}}^{\text{IND}}$, is 1.18 [bit/s] at $T_{\text{sym}} = 0.60$ s with an input probability distribution of $\lambda_0 = 0.50$. We also observe a slight shift in the STI corresponding to the maximum capacity in both ISI-Unaware cases.

Interestingly, by increasing the STI and consequently reducing the impact of ISI, all capacities overlap. Therefore, regardless of the source type or the ISI knowledge, the same performance can be achieved. This overlap occurs because as T_{sym} increases, the effect of ISI diminishes, rendering the knowledge of previously transmitted symbols less valuable. It is important to note that the significant difference between the ISI-Aware cases corresponding to different types of receivers is observed only within a certain range of STIs (0.3 to 0.85). The same observation applies to the other two curves representing ISI unawareness.

Figure ?? illustrates the AIR across the input distribution space represented by ... in the context of ISI awareness with a correlated source. The black hexagram marker (??) indicates the capacity associated with ... corresponding to $p = 0.60$ and $q = 0.65$, with a value of $C_{\text{ISIA}}^{\text{CRR}} = 1.42$. By analyzing the expression for stationary probabilities (??) ... p ... equivalently ... indicates a preference to avoid consecutive transmission of ... it is preferable to avoid generating identical symbols, particularly for high ...

$$I_{\text{ISIA}}^{\text{IND}} = \Bigg\{ H_2(\lambda_0) + \sum_{\forall s_{i-M+1}^{i-1}, \hat{s}_i} \Bigg[\lambda_0^{M-1-w_H(s_{i-M+1}^{i-1})} \lambda_1^{w_H(s_{i-M+1}^{i-1})} \sum_{\forall s_i} \Big[\lambda_0^{1-s_i} \lambda_1^{s_i} P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i|s_{i-M+1}^{i-1}, s_i) \Big] \times$$

$$\sum_{\forall s_i} \Bigg[\frac{\lambda_0^{1-s_i} \lambda_1^{s_i} P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i|s_{i-M+1}^{i-1}, s_i)}{\sum_{x \in \{0,1\}} \lambda_0^{1-x} \lambda_1^{x} P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i|s_{i-M+1}^{i-1}, x)} \log_2 \Bigg(\frac{\lambda_0^{1-s_i} \lambda_1^{s_i} P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i|s_{i-M+1}^{i-1}, s_i)}{\sum_{x \in \{0,1\}} \lambda_0^{1-x} \lambda_1^{x} P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i|s_{i-M+1}^{i-1}, x)} \Bigg) \Bigg] \Bigg] \Bigg\}^{+}. \quad (57)$$

$$I_{\text{ISIU}}^{\text{IND}} = \Bigg\{ H_2(\lambda_0) + \sum_{\forall s_i, \hat{s}_i} \Bigg[\sum_{\forall s_{i-M+1}^{i-1}} \Big[\lambda_0^{M-1-w_H(s_{i-M+1}^{i-1})} \lambda_1^{w_H(s_{i-M+1}^{i-1})} P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i | s_{i-M+1}^{i-1}, s_i) \Big] \times \log_2 \Bigg(\frac{\sum_{\forall s_{i-M+1}^{i-1}} \lambda_0^{M-w_H(s_{i-M+1}^{i-1})-s_i} \lambda_1^{w_H(s_{i-M+1}^{i-1})+s_i} P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i | s_{i-M+1}^{i-1}, s_i)}{\sum_{x\in\{0,1\}} \sum_{\forall s_{i-M+1}^{i-1}} \lambda_0^{M-w_H(s_{i-M+1}^{i-1})-x} \lambda_1^{w_H(s_{i-M+1}^{i-1})+x} P_{\hat{S}_i|S_{i-M+1}^{i-1}, S_i}(\hat{s}_i | s_{i-M+1}^{i-1}, s_i)} \Bigg) \Bigg] \Bigg\}^{+}. \tag{58}$$

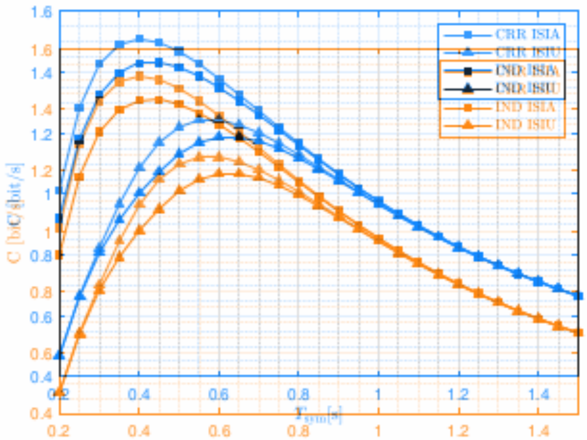


Fig. 3: Capacity, C, as a function of STI for four different scenarios as ISI-Aware with correlated source, CRR ISIA, (blue curve with square marker), ISI-Unaware with correlated source, CRR ISIU, (blue curve with triangle marker), ISI-Aware with independent source, IND ISIA, (red curve with square marker), and ISI-Unaware with independent source, IND ISIU, (red curve with triangle marker).

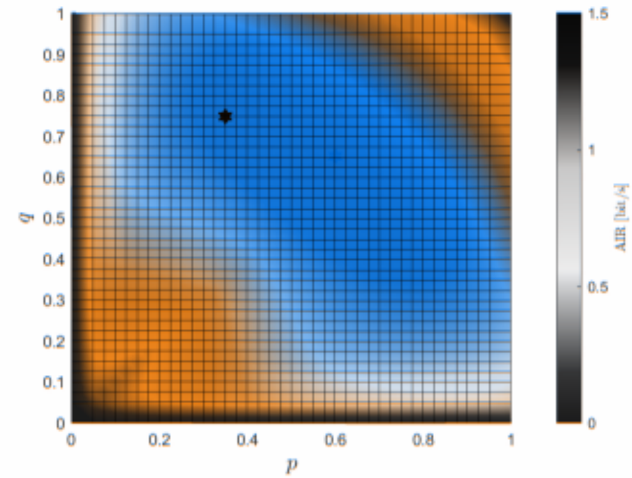


Fig. 4: AIR as a function of the correlated source transition probabilities, p and q, with $T_{\text{sym}} = 0.3$ s in ISI-Aware scenario. Capacity is obtained at $p = 0.60$ and $q = 0.65$, and it is $C_{\text{ISIA}}^{\text{CRR}} = 1.42$.

communication systems. However, in this unique scenario, due to the high ISI effect, the correlated source allows us to tackle the problem of the ISI, and the reduction of the source entropy compared to the independent one is worth it. Specifically, the maximum channel capacity in ISI-Aware scenario with correlated source, $C_{\text{ISIA}}^{\text{CRR}}$, is 1.50 [bit/s] at $T_{\text{sym}} = 0.40$ s, with the input probability distribution of $p = 0.60$ and $q = 0.62$. On the other hand, the maximum capacity for the ISI-Aware scenario with independent source, $C_{\text{ISIA}}^{\text{IND}}$, is 1.43 [bit/s] at $T_{\text{sym}} = 0.45$ s, with an input probability distribution of $\lambda_0 = 0.52$.

Comparing the two maximum capacities in the ISI-Aware scenario, we observe that the independent source achieves its maximum at a higher T_{sym} compared to the correlated source.

Moving on to the ISI-Unaware scenario, the maximum capacity for the correlated source, $C_{\text{ISIU}}^{\text{CRR}}$, is 1.24 [bit/s] at $T_{\text{sym}} = 0.57$ s, with an optimum input probability distribution at $p = 0.60$ and $q = 0.60$. The maximum capacity for the independent source, $C_{\text{ISIA}}^{\text{IND}}$, is 1.18 [bit/s] at $T_{\text{sym}} = 0.60$ s, with an input probability distribution of $\lambda_0 = 0.50$. We also observe a slight shift in the STI corresponding to the maximum capacity in both ISI-Unaware cases.

Interestingly, by increasing the STI and consequently reducing the impact of ISI, all capacities overlap. Therefore, regardless of the source type or the ISI knowledge, the same performance can be achieved. This overlap occurs because as T_{sym} increases, the effect of ISI diminishes, rendering the knowledge of previously transmitted symbols less valuable. It is important to note that the significant difference between the ISI-Aware and ISI-Unaware corresponding to different types of receivers is observed only within a certain range of STIs (0.3 to 0.85). The same observation applies to the other two curves representing ISI unawareness.

Figure ?? illustrates the AIR across the input distribution space, represented by color, in the context of ISI awareness with the correlated source. The black hexagram indicates the capacity point associated with this scenario, corresponding to $p = 0.60$ and $q = 0.65$, with a value of $C_{\text{ISIA}}^{\text{CRR}} = 1.42$.

By analyzing the expressions for stationary probabilities (??) and (??), we can infer that when p and q are equal to each other, the AIR is equivalent to the capacity to transmit an infinitely long sequence with equiprobable symbols. However, as p and q approach 1, it indicates a preference to avoid consecutive transmission of the same symbol. In the provided example, we observe that the optimal

[1]Please note that in Figs. ??, ??, ??, ??, ??, and ?? a deliberately consistent color spectrum indicating the range of AIR from 0 to 1.5 was employed. This choice was made to facilitate enhanced visual comparison between the figures, promoting general perception.

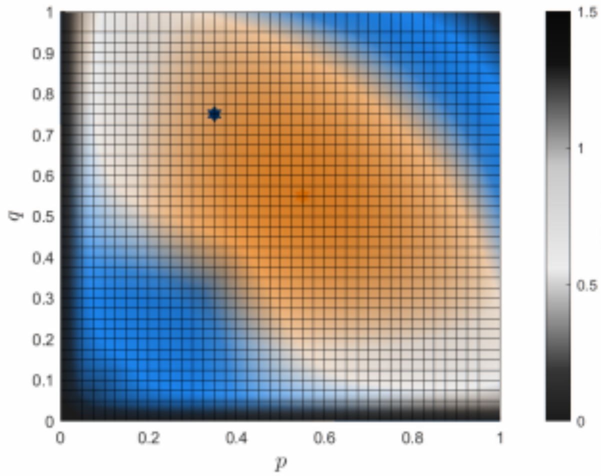


Fig. 5: AIR as a function of the correlated source transition probabilities with $T_{\text{sym}} = 0.3$ s in ISI-Unaware scenario. Capacity is achieved at $p = 0.35$ and $q = 0.75$, with a value of $C^{\text{CRR}}_{\text{ISIU}} = 0.82$.

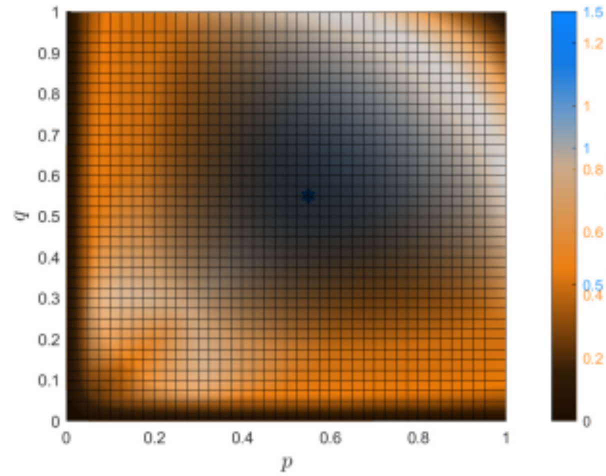


Fig. 6: AIR as a function of the correlated source transition probabilities with $T_{\text{sym}} = 0.3$ s in ISI-Aware scenario. Capacity is achieved at $p = q = 0.55$, with a value of $C^{\text{CRR}}_{\text{ISIA}} = 1.27$.

Fig. 7: AIR as a function of the correlated source transition probabilities with $T_{\text{sym}} = 0.7$ s in ISI-Unaware scenario. Capacity is $C^{\text{CRR}}_{\text{ISIU}}$ obtained with $p = 0.57$ and $q = 0.57$.

input distribution is when the input distribution of the correlated source is asymptotically equiprobable, but it is preferable to avoid generating consecutive identical symbols, particularly for the transmission of "1". This observation is supported by the fact that q is slightly higher than p, indicating a lesser desire for transmitting two successive "1" symbols. Of course, when both p and q are close to 1, little information is transmitted since, after transmitting a "1", a "0" follows with high probability, and vice versa.

Figure ??[1] displays the AIR in the context of ISI unawareness, where symbols are generated by a correlated source, with $T_{\text{sym}} = 0.3$ s. The capacity achieved in this scenario is $C^{\text{CRR}}_{\text{ISIU}} = 0.82$, with the correlated source input distribution of $p = 0.35$ and $q = 0.75$.

An intriguing observation is that in this case, the capacity is not attained with a stationary equiprobable distribution. Rather, it is achieved when the probability of having a "0" after a "1", and the likelihood of generating successive "0" symbols, is higher compared to generating "1". This optimal strategy can be readily understood since there is no knowledge of ISI, and with an STI of $T_{\text{sym}} = 0.3$ s, where ISI is prominent, it is preferable to avoid generating "1"s.

Another noteworthy observation in the figure is that when $p \approx 0.7$ and $q \approx 0.3$, a similar AIR value to the capacity is observed. This region of the correlated source input space signifies the preference for generating more "1"s. In comparison to the ISI-Aware case shown in Fig. ??, we can observe that the AIR is lower in the ISI-Unaware scenario as expected but with the same color pattern.

Figure ?? depicts the AIR for the same scenario as shown in Fig. ??, but here the STI is $T_{\text{sym}} = 0.7$ s. In this case, the capacity is achieved with $p = q = 0.55$, indicating a preference for almost equiprobable input distributions.

By comparing the capacity-achieving input distribution in Figs. ?? and ?? we can observe that the optimal values of p and q converge to 0.5 as the STI increases. With p and q closer to 0.5, the source behaves more similarly to a source that emits independent and uniformly distributed symbols.

Figure ?? presents the AIR for a similar scenario as depicted in Fig. ??, with STI $T_{\text{sym}} = 0.7$ s. The capacity is obtained when the input distribution is equiprobable, characterized by $p = q = 0.57$, indicating a preference for equal probabilities from a stationarity perspective.

Comparing this capacity with the one shown in Fig. ??, we observe a shift in the optimal input distribution. In this case, the capacity is achieved with an equiprobable distribution that slightly avoids generating the same symbols successively, whereas in Fig. ??, the avoidance of generating "1"s was a little more preferred. This observation can be explained by the increased STI T_{sym}, which leads to a relatively reduced impact of ISI.

In Fig. ??, we present the difference between the AIRs illustrated in Fig. ?? and Fig. ?? to analyze the disparity between the ISI-Aware scenario and the ISI-Unaware case for an STI of $T_{\text{sym}} = 0.3$ s. Consistent with the theoretical prediction stated in (??), the difference between the two AIRs is non-negative.

There are two prominent regions where the difference between the AIRs is significant. The first region, located in the upper right side of the figure, demonstrates that the performance of the ISI-Unaware scenario tends to approach 0, whereas the ISI-Aware case exhibits a higher AIR in that region. A similar observation can be made for the bottom left region of the figure.

Since the independent source's input space can be spanned by a single variable, λ_0, we can examine the AIR values of the independent source as a function of the STI, T_{sym}, and the probability of transmitting a "0", λ_0. Figure ?? illustrates the AIR values in the ISI-Aware scenario, where the source generates symbols independently, and the receiver is aware of the previously transmitted symbols. As we also observed in Fig. ??, the highest capacity is achieved at smaller T_{sym}s.

To gain further insight, in Fig. ?? we present a cross-sectional view of Fig. ?? focusing on specific STIs of $T_{\text{sym}} \in \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ s. The histogram markers indicate the capacity corresponding to each STI. It is evident that as T_{sym} increases, the capacity distribution shifts in the direction of the equiprobable input distribution. In

[1] Please note that in Figs. ??, ??, ??, ??, ??, and ?? a deliberately consistent color spectrum indicating the range of AIR from 0 to 1.5 was employed. This choice was made to facilitate enhanced visual comparisons between the figures, promoting a more comprehensive

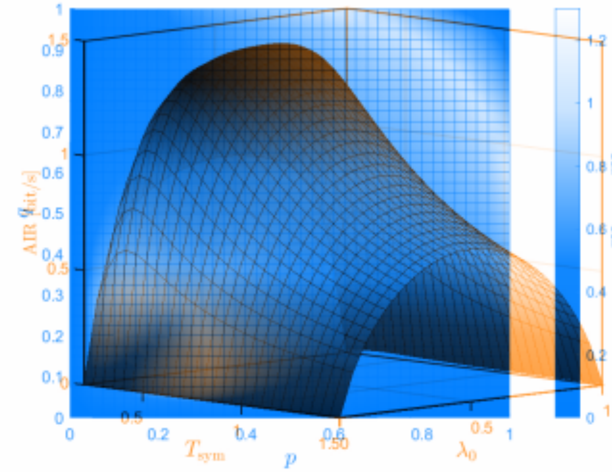


Fig. 9: Difference between the AIR values corresponding to Fig. ?? and Fig. ?? over the input distribution space.

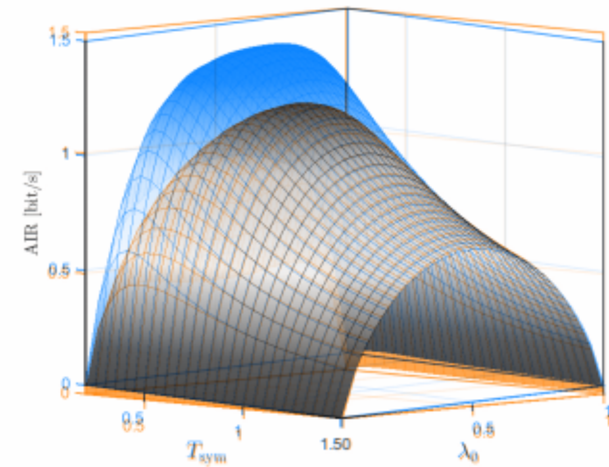


Fig. 11: AIR as a function of the independent source probability of transmitting "0", λ_0, and STI, T_{sym}, in ISI-Aware scenario.

this case, the capacity is achieved with an equiprobable distribution that slightly avoids generating the same symbols successively, whereas in Fig. ??, the avoidance of generating "1"s was a little more preferred. This observation can be explained by the increased STI T_{sym}, which leads to a relatively reduced impact of ISI.

In Fig. ??, we present the difference between the AIRs illustrated in Fig. ?? and Fig. ?? to analyze the disparity between the ISI-Aware scenario and the ISI-Unaware case for an STI of $T_{\text{sym}} = 0.3$ s. Consistent with the theoretical prediction stated in (??), the difference between the two AIRs is non-negative.

There are two prominent regions where the difference between the AIRs is significant. The first region, located in the upper right side of the figure, demonstrates that the performance of the ISI-Unaware scenario tends to approach 0, whereas the ISI-Aware case exhibits a higher AIR in that region. A similar observation can be made for the bottom left region of the figure.

Since the independent source's input space can be spanned by a single variable, λ_0, we can examine the AIR values of the independent source as a function of the STI, T_{sym}, and the probability of transmitting a "0", λ_0. Figure ?? illustrates the AIR values in the ISI-Aware scenario, where the source generates symbols independently, and the receiver is aware of the previously transmitted symbols. As we also observed in Fig. ??, the highest capacity is achieved when $T_{\text{sym}} = 0.45$ s.

To gain further insight, in Fig. ?? we present a cross-sectional view of Fig. ??, focusing on specific STIs ($T_{\text{sym}} \in \{0.30, 0.45, 0.60, 0.75, 0.90\}$ s). The hexagram markers indicate the capacity corresponding to each STI. It is evident that as T_{sym} increases, the input distribution associated with the channel capacity gradually approaches an equiprobable input distribution (i.e., $\lambda_0 = 0.5$). However, the maximum possible capacity achieved with equiprobability of input symbols ...

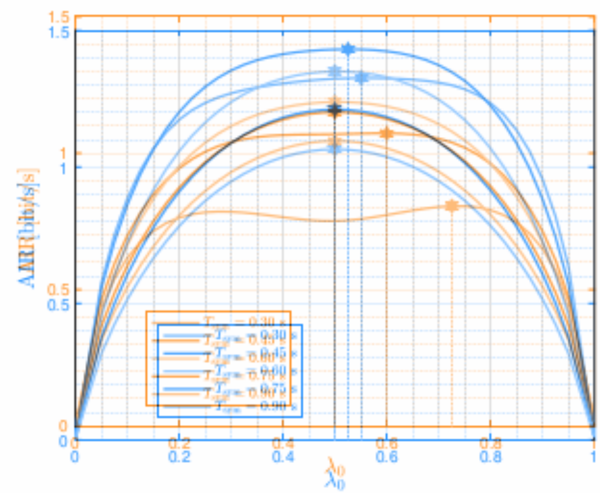


Fig. 10: AIR values corresponding to $T_{\text{sym}} \in \{0.30, 0.45, 0.60, 0.75, 0.90\}$ s when the source is of the independent type, and ISI-Aware scenario holds. Hexagram markers indicate the capacity associated with each T_{sym}.

Figure ?? shows the AIR in a similar fashion as in Fig. ??, but assumes no knowledge on ISI. Compared to Fig. ??, we can observe that the AIR drops. However, the shape of the manifold remains similar. The cross-sectional view of Fig. ?? is depicted in Fig. ?? for the same set of T_{sym} as for Fig. ??. The maximum possible capacity achieved with equiprobability of input symbols is at $T_{\text{sym}} = 0.6$ s. The reason for the complex shape of the AIR curves in Fig. ?? is that each λ_0 is associated with a different channel, which depends on the specific

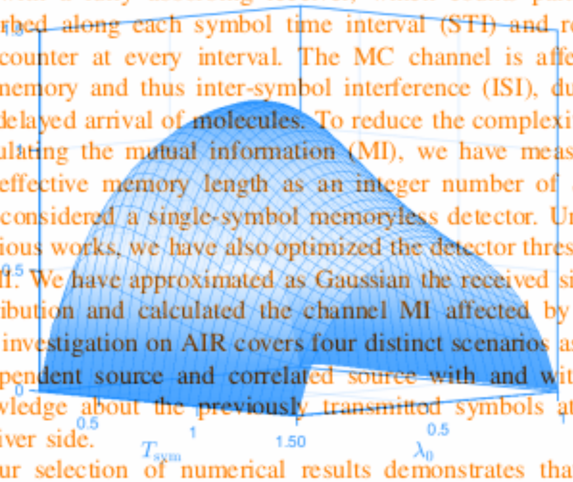


Fig. 11: AIR as a function of the independent source probability of transmitting "0" λ_0 and of the STI T_{sym} in ISI-Unaware scenario.

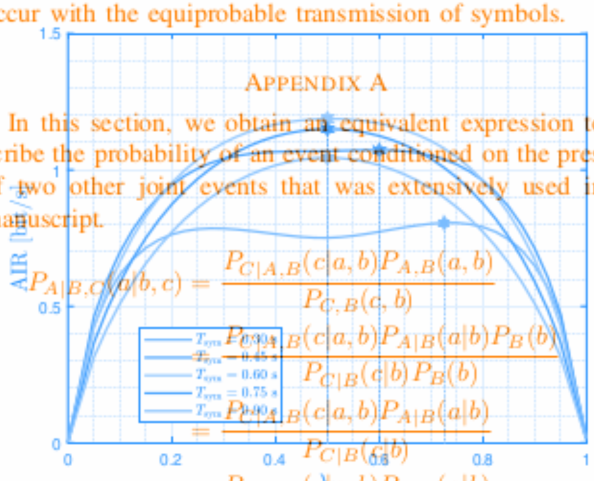


Fig. 12: AIR values corresponding to $T_{\mathrm{sym}} \in \{0.30, 0.45, 0.60, 0.75, 0.90\}$ s when the source is of the independent type, and ISI-Unaware scenario holds. Hexagram markers indicate the capacity associated with each T_{sym}.

optimum detector threshold, τ. The curve for $T_{\mathrm{sym}} = 0.3$ s exhibits two local maxima. The maximum at $\lambda_0 \approx 0.28$ suggests transmitting fewer "0"s is beneficial, which may seem counter-intuitive given the higher ISI associated with faster transmission rates. However, the other maximum at $\lambda_0 \approx 0.75$ suggests transmitting more "0"s is optimal. This observation is sensible because the ISI increases with the transmission rate. By transmitting "1" less frequently, the ISI is reduced, yielding an improvement in the AIR. As expected, the maximum associated with $\lambda_0 \approx 0.75$ is higher than the one associated with $\lambda_0 \approx 0.28$.

V. Conclusions

We have investigated the Achievable Information Rate (AIR) of a diffusive molecular communication (MC) channel with a fully absorbing receiver, which counts particles absorbed along each symbol time interval (STI) and resets the counter at every interval. The MC channel is affected by memory and thus inter-symbol interference (ISI), due to the delayed arrival of molecules. To reduce the complexity in calculating the mutual information (MI), we have measured the effective memory length as an integer number of STIs and considered a single-symbol memoryless detector. Unlike previous works, we have also optimized the detector threshold to MI. We have approximated as Gaussian the received signal distribution and calculated the channel MI affected by ISI. Our investigation on AIR covers four distinct scenarios as the independent source and correlated source with and without knowledge about the previously transmitted symbols at the receiver side.

Our selection of numerical results demonstrates that, in general, with correlated source, we can achieve higher capacity. The optimal input probability distribution achieving the capacity may not be uniform. In particular, when the STI T_{sym} is small, thus implying strong ISI, the maximum AIR does not occur with the equiprobable transmission of symbols.

Appendix A

In this section, we obtain an equivalent expression to describe the probability of an event conditioned on the presence of two other joint events that was extensively used in the manuscript.

$$
\begin{aligned}
P_{A|B,C}(a|b,c) &= \frac{P_{C|A,B}(c|a,b)P_{A,B}(a,b)}{P_{C,B}(c,b)} \\
&= \frac{P_{C|A,B}(c|a,b)P_{A|B}(a|b)P_B(b)}{P_{C|B}(c|b)P_B(b)} \\
&= \frac{P_{C|A,B}(c|a,b)P_{A|B}(a|b)}{P_{C|B}(c|b)} \\
&= \frac{P_{C|A,B}(c|a,b)P_{A|B}(a|b)}{\sum_{\forall y} P_{C,A|B}(c,y|b)} \\
&= \frac{P_{C|A,B}(c|a,b)P_{A|B}(a|b)}{\sum_{\forall y} P_{C|A,B}(c|y,b)P_{A|B}(y|b)}
\end{aligned} \tag{59}
$$

nel with a fully absorbing receiver, which counts particles absorbed along each symbol time interval (STI) and resets the counter at every interval. The MC channel is affected by memory and thus inter-symbol interference (ISI), due to the delayed arrival of molecules. To reduce the complexity in calculating the mutual information (MI), we have measured the effective memory length as an integer number of STIs and considered a single-symbol memoryless detector. Unlike previous works, we have also optimized the detector threshold to MI. We have approximated as Gaussian the received signal distribution and calculated the channel MI affected by ISI. Our investigation on AIR covers four distinct scenarios as the independent source and correlated source with and without knowledge about the previously transmitted symbols at the receiver side.

Our selection of numerical results demonstrates that, in general, with correlated source, we can achieve higher capacity. The optimal input probability distribution achieving the capacity may not be uniform. In particular, when the STI T_{sym} is small, thus implying strong ISI, the maximum AIR does not occur with the equiprobable transmission of symbols.

Appendix A

In this section, we obtain an equivalent expression to describe the probability of an event conditioned on the presence of two other joint events that was extensively used in the manuscript.