A General Upper Bound on Point-to-Point Particle Timing Channel Capacity Under Constant Particle Emission Intensity

Christopher Rose
School of Engineering
Brown University
Email: Christopher_Rose@brown.edu

I. Saira Mian
Department of Computer Science
University College London
Email: S.Mian@cs.ucl.ac.uk

***Abstract*—The past decade has produced a large body of work on communication channels which use chemicals to communicate. Some work uses a finest grain model wherein the arrival times of individual emitted particles convey information. Others consider a related particle intensity system where the time is binned at the transmitter/receiver and the number of particles released and counted conveys information. Still others consider a macroscopic model that uses an Avogradrian number of particles and thus concentration as the information carrier. However, given the myriad emission, carrier transport and uptake/sensing methods studied, it has been difficult to precisely relate timing, intensity and concentration results. Here we attempt a partial unification through a simple upper bound applicable to any finite-mean first-passage time transport mechanism under an assumption of constant particle emission intensity $\bar{\lambda}$. Our result is expressed in terms of three quantities: the particle emission rate $\bar{\lambda}$, the average particle uptake rate μ, and the entropy of the first-passage time distribution, $h(D)$.**

***Index Terms*—molecular channel capacity, molecular signaling, timing channel capacity**

I. Introduction

Three abstractions of communication channels which use chemicals to communicate have been studied widely (see [1], [2] and references therein for a history and survey of molecular communication). These consider *particle timing*: the finest grain model wherein the arrival times of individual emitted particles convey information [3]–[6], *particle intensity*: a related model where the time is binned at the transmitter/receiver and the number of particles released and counted conveys information [7], and *particle concentration*: a macroscopic model that uses an Avogradrian number of particles and concentration imparts information [8]–[10].

Since the intensity and concentration abstractions are "processed" versions of the timing abstraction, timing capacity must upper bound that of the others via the data processing theorem [11]. However, given the myriad carrier transport and uptake/sensing methods studied, it has been difficult to precisely relate timing, intensity and concentration results. Here we attempt a partial unification through a simple upper bound applicable to any finite-mean first-passage time transport mechanism under an assumption of constant particle emission intensity. Our result is expressed in terms of three quantities: the particle emission rate $\bar{\lambda}$, the average particle uptake rate μ, and the entropy of the first-passage time distribution, $h(D)$.

II. Molecular Channel Modeling Overview

A point to point molecular communication channel model is illustrated in FIGURE 1. This basic arrangement is a staple of the field and can be found in various forms in a variety of prior work (see [2] for a survey). A message is composed, coded into chemical emission patterns and released into a medium which transports the chemicals to a sensor whose output is interpreted to reconstruct the message. There are many variations on the transport, transduction and sensing methods, but this basic model is generally accepted.

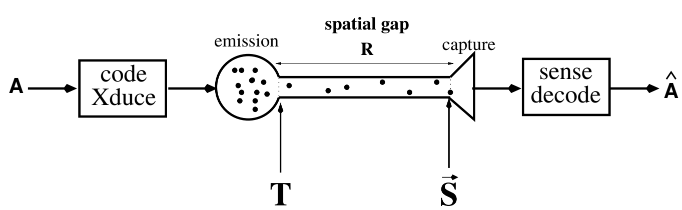

Fig. 1. **General Molecular Communication Channel Abstraction:** a message $\boldsymbol{A}$ is coded and then transduced into a set of particle emission times $\mathbf{T}$. These particles propagate over a spatial gap R through a transmission medium and are captured (exactly once) at corresponding times $\mathbf{S}$. Since the particles are identical, capture results in the ordered arrivals $\vec{\mathbf{S}}$. These ordered arrivals are sensed and and decoded into the message estimate $\hat{\boldsymbol{A}}$

To derive our bounds we will use only the central portion of FIGURE 1 as depicted in FIGURE 2. All such models are defined by three random variables: emission time T_m, transit (first-passage) time D_m and arrival time S_m, and related by

$$S_m = T_m + D_m$$

Since it is reasonable to assume that the action of the transport medium on particles is independent of the number (to within reason), position and release times of particles, we assume the $\{D_m\}$ are independent and identically distributed (i.i.d.) with density $g(\cdot)$.

As each arrival corresponds to a single emission, we can define M-vectors $\mathbf{T}$, $\mathbf{D}$ and $\mathbf{S}$ accordingly:

$$\mathbf{S} = \mathbf{T} + \mathbf{D} \tag{1}$$

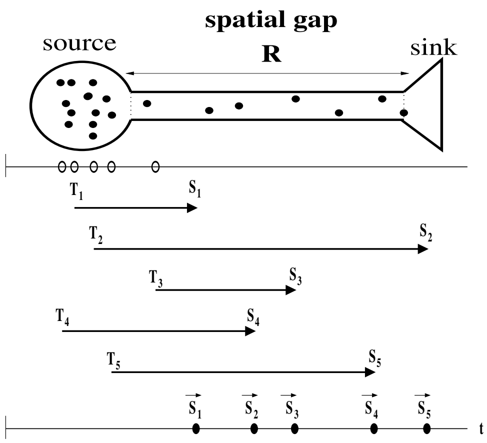

Fig. 2. **Core Molecular Communication Channel:** an ensemble of particles is emitted and transported across a spatial gap. The particles are released at (unordered) times $\{T_m\}$, propagate through a transmission medium and are captured at corresponding times $\{S_m\}$. For identical particles, the receiver sees ordered arrivals $\{\vec{S}_m\}$ which may differ in index from the unordered arrivals $\{S_m\}$.

However, since the particles are identical, the receiver sees only an ordered set of arrivals

$$\vec{\mathbf{S}} = P_\Omega(\mathbf{S}) \tag{2}$$

where $P_\Omega(\cdot)$, $\Omega = 1, 2, \ldots, M!$, is a permutation operator and Ω is that permutation index which produces ordered $\vec{\mathbf{S}}$ from the argument $\mathbf{S}$. Thus,

$$\vec{\mathbf{S}} = P_\Omega\left(\mathbf{T} + \mathbf{D}\right) \tag{3}$$

and Ω is a discrete random variable associated with the channel. Equation (3) is the basic description of a "timing channel" wherein emissions $\mathbf{T}$ are constructed so that information can be extracted from $\vec{\mathbf{S}}$.

In addition to being finest grain, this abstraction is particularly useful because it distills any number of emitter/receiver geometries and transport medium properties to a single random variable – the first-passage time D. There are many models for the transport mechanism and resultant first-passage time distributions [2], the most popular of which is the Lévy distribution derived from diffusion characteristics [7] and the related additive inverse Gaussian channel model [12].

It must be emphasized that any uncertainty introduced in the coding/transduction of $\mathbf{T}$ and any degradation introduced by the sensor/decoding which detects $\vec{\mathbf{S}}$ can only, via the data processing theorem [11], decrease channel capacity. For instance, uncertainty in the number or concentration of released chemicals will decrease capacity. Likewise it can easily be shown that repeated detection of the same particle at the receiver conveys no additional information because later re-detection times are independent of emission time given the first detection time [3]. Likewise, if a particle is blocked from detection at the receiver, we cannot gain information. Therefore we assume all emission times are perfectly executed and the detection times $\vec{\mathbf{S}}$ are first detections (effectively removing each particle from the system upon detection). In this way a capacity analysis of the channel in FIGURE 2 provides an upper bound on the channel capacity of FIGURE 1.

III. Capacity Upper Bound

We will assume a signaling interval of duration duration $\tau(M) = \tau = M/\bar{\lambda}$ and a guard interval $\gamma(M, \epsilon)$ as depicted in FIGURE 3. ϵ is the probability that emissions during the signaling interval arrive after the guard interval. It was proven in [6] that so long as $E[D] = 1/\mu < \infty$, for any chosen ϵ we can always find a $\gamma(M, \epsilon)$ such that the ratio $\gamma(M, \epsilon)/\tau(M)$ goes to zero for large M. Put another way, if $E[D]$ is finite, there exists a guard interval whose relative duration goes to zero with large M. If $E[D] = \infty$ then no such guard interval exists. Thus $E[D] = \infty$ implies the usual channel capacity analysis is ill-posed [5]. Interestingly, popular transport models such as simple drift-free diffusion have infinite first-passage times, so the capacity question is moot for such systems. However, in any practical system there are usually physical limits which impose $E[D] < \infty$ and renders the system tractable from an information theoretic standpoint.

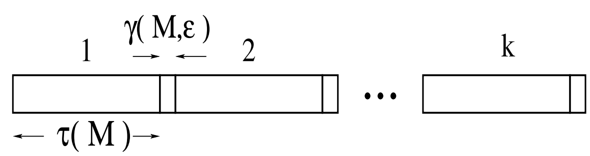

Fig. 3. **Successive M-Emission Channel Uses.** For a given use of the particle timing channel, the sender emits M particles over the transmission interval $\tau(M) = \frac{M}{\bar{\lambda}}$. $\gamma(M, \epsilon)$ is the waiting period (guard interval) before the next channel use. ϵ is the probability that at least one particle arrives after the guard interval. A $\gamma(M, \epsilon)$ sublinear in M (so that the ratio of the guard interval to the signaling interval approaches zero) can be found iff $E[D] < \infty$ [5]

Assuming $E[D] = 1/\mu < \infty$ the maximum of the mutual information $I[\vec{\mathbf{S}}; \mathbf{T}]$ is the channel capacity, but the ordering operation which produces $\vec{\mathbf{S}}$ renders this maximization difficult. However, if we permit only nonsingular first-passage time distributions on $\mathbf{D}$, the distribution of $\mathbf{S}$ will be continuous and this allows us to "fold" the density on $\mathbf{S}$ to obtain the density on $\vec{\mathbf{S}}$. In turn, this allows us to consider only "hypersymmetric" distributions on $\mathbf{T}$ [5], [6] where $f_{\mathbf{T}}(\mathbf{t})$ is invariant under permutations of $\mathbf{t}$. It can then be shown that

$$h(\vec{\mathbf{S}}) = h(\mathbf{S}) - \log M!$$

We then note the equivalence

$$\{\Omega, \vec{\mathbf{S}}\} \Leftrightarrow \mathbf{S}$$

which allows us to write

$$h(\mathbf{S}|\mathbf{T}) = h(\Omega, \vec{\mathbf{S}}|\mathbf{T}) = h(\vec{\mathbf{S}}|\mathbf{T}) + H(\Omega|\vec{\mathbf{S}}, \mathbf{T}) \tag{4}$$

and leads immediately to

$$\begin{aligned} I(\vec{\mathbf{S}}; \mathbf{T}) &= h(\vec{\mathbf{S}}) - h(\vec{\mathbf{S}}|\mathbf{T}) \\ &= h(\mathbf{S}) - h(\mathbf{S}|\mathbf{T}) - \Big(\log M! - H(\Omega|\vec{\mathbf{S}}, \mathbf{T})\Big) \\ &= I(\mathbf{S}; \mathbf{T}) - \Big(\log M! - H(\Omega|\vec{\mathbf{S}}, \mathbf{T})\Big) \end{aligned} \tag{5}$$

Now, if we set the signaling/symbol interval to $\tau = M/\bar{\lambda}$ where $\bar{\lambda}$ is the average rate at which particles are emitted, then we can define a capacity per particle [5] as

$$\mathcal{C}_q = \lim_{M\to\infty} \frac{1}{M} \sup I(\vec{\mathbf{S}};\mathbf{T}) \tag{6}$$

Furthermore, if we set our time base to units of mean first-passage time $E[D] = 1/\mu$, the capacity in bits per passage time is

$$\mathcal{C}_t = \rho \mathcal{C}_q \tag{7}$$

where $\rho \equiv \frac{\bar{\lambda}}{\mu}$.

Equation (5) is satisfying in that $I(\vec{\mathbf{S}};\mathbf{T})$ is expressible as the mutual information when emission-arrival correspondence is known ($I(\mathbf{S};\mathbf{T})$) less a penalty imposed by particle indistinguishability ($\log M! - H(\Omega|\vec{\mathbf{S}},\mathbf{T})$). That is, $H(\Omega|\vec{\mathbf{S}},\mathbf{T})$ is the average amount of disorder between $\mathbf{S}$ and $\mathbf{T}$ imposed by the channel.

However, there is also tension between maximizing $I(\vec{\mathbf{S}};\mathbf{T})$ and $H(\Omega|\vec{\mathbf{S}},\mathbf{T})$. That is, $H(\Omega|\vec{\mathbf{S}},\mathbf{T})$ increases when the distribution on $\mathbf{T}$ is clustered/structured. Consider that if all M emissions occurred at once, the mapping between $\vec{\mathbf{S}}$ and $\mathbf{T}$ would be arbitrary and $\log M! - H(\Omega|\vec{\mathbf{S}},\mathbf{T}) = 0$. In contrast if the T_m are correlated, then the $\mathbf{S}$ are also correlated resulting in a reduction of $h(\mathbf{S})$. Similarly, if the clustering is owed to emission intensity modulation with time (even while maintaining independence of the T_m) then the density of each S_m will be relatively concentrated in certain regions of the signaling interval and the upper limit of $h(S_m)$ reduced.

For analytic simplicity, we will assume that emission times $\{T_m\}$ are independent with probability density $f_T(\cdot)$. This also comports well with the physics of concentration-based models where identical particles are released *en masse* but with temporal imprecision. For large τ and M, the emission independence assumption implies that on any small (relative τ) interval the ordered emissions $\vec{\mathbf{T}}$ are approximately Poisson with time-varying rate

$$\lambda(t) = M f_T(t) \tag{8}$$

A. An Upper Bound on $I(\mathbf{S};\mathbf{T})$

First consider that owing to the presumed hypersymmetry of $\mathbf{T}$ we have

$$I(\mathbf{S};\mathbf{T}) \le M I(S;T) = M(h(S) - h(D))$$

where

$$f_S(s) = (f_T * g)(s)$$

with equality if the T_m are i.i.d. so that

$$\max_{f_{\mathbf{T}}} I(\mathbf{S};\mathbf{T}) \le M \max_{f_T} I(S;T) = M \max_{f_T} h(S) - M h(D)$$

Then, since the probability of arrivals beyond $\tau(M) + \gamma(M,\epsilon)$ is vanishingly small we must have

$$h(S) \le \log\left(\tau + O(M^{-(1+\delta)})\right)$$

where $\delta > 0$. This implies that the mutual information between S and T is strictly upper bounded by

$$I(S;T) \le \log\left(\tau + O(M^{-(1+\delta)})\right) - h(D)$$

It is easily shown that an asymptotically uniform density on S can be achieved with a uniform distribution on T as $\tau \to \infty$. Thus, for $s \gg E[D]$, $f_S(s) \approx \frac{1}{\tau}$ and the mutual information $I(S;T)$ is asymptotically maximized by uniform $f_T(t)$ so that

$$\max_{f_T()} I(S;T) \le \log \tau - h(D) \tag{9}$$

It is worth emphasizing that the bound of equation (9) holds for *any non-singular finite-mean first-passage density* $g(\cdot)$.

B. An Upper Bound for $H(\Omega|\vec{\mathbf{S}},\mathbf{T})$

$H(\Omega|\vec{\mathbf{S}},\mathbf{T})$ is the discrete entropy of the correct mapping between $\mathbf{T}$ and $\vec{\mathbf{S}}$ given both $\mathbf{T}$ and $\vec{\mathbf{S}}$. As the mapping is one-to-one, there are exactly $M!$ possibilities for Ω only some of which are feasible. In [5] and [6], explicit expressions for $H(\Omega|\vec{\mathbf{S}},\mathbf{T})$ were found for exponential first-passage distributions. Here we derive an upper bound on $H(\Omega|\vec{\mathbf{S}},\mathbf{T})$ for *any* finite-mean first-passage density under an assumption of constant particle emission intensity $\bar{\lambda}$.

We begin by reviewing the initial development of [5] (which we repeat here in part for clarity) where it was shown that if we define $|\Omega|_{\vec{\mathbf{s}},\mathbf{t}}$ as the number of permissible mappings of $\vec{\mathbf{s}} \to \mathbf{t}$, then

$$H(\Omega|\vec{\mathbf{S}},\mathbf{T}) \le E_{\vec{\mathbf{S}},\mathbf{T}}\left[|\Omega|_{\vec{\mathbf{S}},\mathbf{T}}\right] \tag{10}$$

with equality (assuming $g(\cdot)$ is nonzero everywhere) iff the first-passage density $g(\cdot)$ is exponential. Additionally, if we allow for $g(x)$ which can be identically zero on regions of nonzero measure ($x \ge 0$), equality in equation (10) also holds if $g(\cdot)$ is binary, taking on only two values 0 and c – the simplest example being a compact uniform distribution.

In counting admissible mappings, we note that knowledge of $\mathbf{T}$ implies knowledge of $\vec{\mathbf{T}}$. This allows us to define contiguous "bins" $\mathcal{B}_k = \{t | t \in [\vec{T}_k, \vec{T}_{k+1})\}$, $k = 1, 2, \ldots, M$ ($\vec{T}_{M+1} \equiv \infty$) and then define σ_m as bin occupancies. That is, $\sigma_m = r$ if there are exactly r arrivals in $\mathcal{B}_m$. The benefit of this approach is that the σ_m do not depend on whether $\vec{\mathbf{s}}$ or $\mathbf{s}$ is used to count the arrivals. Thus, expectations can be taken over $\mathbf{S}$ whose components are mutually independent given the $\mathbf{T}$ and no order distributions for $\vec{\mathbf{S}}$ need be derived. Put another way, hypersymmetry of $\mathbf{T}$ implies

$$|\Omega|_{\vec{\mathbf{S}},\mathbf{T}} = |\Omega|_{\vec{\mathbf{S}},\vec{\mathbf{T}}} = |\Omega|_{\mathbf{S},\vec{\mathbf{T}}} = |\Omega|_{\mathbf{S},\mathbf{T}} \tag{11}$$

To determine the random variable $|\Omega|_{\vec{\mathbf{S}},\mathbf{T}}$ we define

$$\eta_m = \sum_{j=1}^{m} \sigma_j$$

the total number of arrivals of $\vec{\mathbf{S}}$ up to and including bin $\mathcal{B}_m$. Clearly η_m is monotonically increasing in m with $\eta_0 = 0$ and $\eta_M = M$. We then observe that the σ_m arrivals on $[\vec{T}_m, \vec{T}_{m+1})$ can be assigned to any of the $\vec{T}_1, \vec{T}_2, \ldots, \vec{T}_m$ known emission times *except* for those η_{m-1} previously assigned. The number of possible new assignments is $(m-\eta_{m-1})!/(m-\eta_m)!$ which when applied iteratively leads to

$$|\Omega|_{\vec{\mathbf{S}},\mathbf{T}} = \prod_{m=1}^{M} \frac{(m-\eta_{m-1})!}{(m-\eta_m)!} = \prod_{m=1}^{M-1} (m+1-\eta_m) \quad (12)$$

which implies

$$H(\Omega|\vec{\mathbf{S}},\mathbf{T}) \le E_{\vec{\mathbf{S}},\mathbf{T}}\left[\sum_{m=1}^{M-1} \log(m+1-\eta_m)\right] \quad (13)$$

We then define the random variable

$$X_i^{(m)} = \begin{cases} 1 & S_i < \vec{T}_{m+1} \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1, 2, \ldots, m$. The PMF of $X_i^{(m)}$ is then

$$p_{X_i^{(m)}}(x) = \begin{cases} G(\vec{T}_{m+1} - \vec{T}_i) & x = 1 \\ \bar{G}(\vec{T}_{m+1} - \vec{T}_i) & x = 0 \end{cases} \quad (14)$$

where we note that for a given m, $X_i^{(m)}$ and $X_j^{(m)}$ are independent, $i \ne j$, and $G(\cdot)$ is the CDF of the causal first-passage density $g(\cdot)$. $\bar{G}(\cdot) = 1 - G(\cdot)$ is the corresponding CCDF. We can then write

$$\eta_m = \sum_{i=1}^{m} X_i^{(m)}$$

so that

$$\begin{aligned} H(\Omega|\vec{\mathbf{S}},\mathbf{T}) &\le E_{\mathbf{S},\mathbf{T}}\left[\sum_{m=1}^{M-1} \log\left(m+1-\sum_{i=1}^{m} X_i^{(m)}\right)\right] \\ &= E_{\mathbf{S},\mathbf{T}}\left[\sum_{m=1}^{M-1} \log\left(1+\sum_{i=1}^{m} (1-X_i^{(m)})\right)\right] \quad (15) \end{aligned}$$

Via Jensen's inequality we have

$$H(\Omega|\vec{\mathbf{S}},\mathbf{T}) \le \sum_{m=1}^{M-1} \log\left(1+\sum_{i=1}^{m} E_{\mathbf{S},\mathbf{T}}\left[(1-X_i^{(m)})\right]\right) \quad (16)$$

Now consider that $X_i^{(m)}$ is a function of S_i, $\vec{T}_{m+1}$ and $\vec{T}_i$. Given $\vec{T}_{m+1}$ and $\vec{T}_i$, we have

$$E_{\mathbf{S}|\mathbf{T}}\left[1 - X_i^{(m)}\right] = \bar{G}(\vec{T}_{m+1} - \vec{T}_i) = \bar{G}(\Delta_{m+1-i}) \quad (17)$$

where $\Delta_{m+1-i} \equiv \vec{T}_{m+1} - \vec{T}_i$. If $f_T(\cdot)$ is uniform, then on any interval $\ll \tau$ the $\vec{\mathbf{T}}$ are approximately samples of a Poisson process with constant rate $\lambda(t) = \bar{\lambda}$. The Δ_{ji} are then approximately Erlang with parameter $\bar{\lambda}$.

$$f_{\Delta_{ji}}(x) = \frac{\bar{\lambda}(\bar{\lambda}x)^{j-i-1} e^{-\bar{\lambda}x}}{(j-i-1)!}$$

so that

$$E_{\mathbf{T}}\left[\bar{G}(\Delta_{m+1-i})\right] = \int_0^\infty \frac{\bar{\lambda}(\bar{\lambda}x)^{m-i} e^{-\bar{\lambda}x}}{(m-i)!} \bar{G}(x)dx \quad (18)$$

We then note that

$$\sum_{i=1}^{m} \frac{\bar{\lambda}(\bar{\lambda}x)^{m-i} e^{-\bar{\lambda}x}}{(m-i)!} = \bar{\lambda} e^{-\bar{\lambda}x} \sum_{k=0}^{m-1} \frac{(\bar{\lambda}x)^k}{k!} \le \bar{\lambda} \quad (19)$$

so that

$$\begin{aligned} H(\Omega|\vec{\mathbf{S}},\mathbf{T}) &\le \sum_{m=1}^{M-1} \log\left(1+\int_0^\infty \bar{\lambda}\bar{G}(x)dx\right) \\ &= (M-1)\log(1+\rho) \quad (20) \end{aligned}$$

where $\rho = \bar{\lambda}/\mu$. Equation (20) (a more simply-derived special case of Theorem$\sim$ 9 in [6]) is a *general* upper bound on $H(\Omega|\vec{\mathbf{S}},\mathbf{T})$ for *any* first-passage time density $g(\cdot)$ with mean $1/\mu$ at constant particle emission rate $\bar{\lambda}$.

C. An Upper Bound on $\mathcal{C}_t$

Through equation (5), equation (9) and equation (20) we have established an upper bound on channel capacity at constant particle emission rate. That is, combining equation (5), equation (9), equation (20) and using Stirling's approximation with the identity $\bar{\lambda} = M/\tau$ we have

$$\begin{aligned} I(\vec{\mathbf{S}};\mathbf{T}) = (M-1)\log(1+\rho) - M(\log\frac{\bar{\lambda}}{e} + h(D)) \\ - \frac{1}{2}\log 2\pi M \quad (21) \end{aligned}$$

Applying equation (6) and equation (7) leads to our main results:

Theorem 1: **Constant Emission Rate Capacity Bound:** *The upper bounds on $\mathcal{C}_q$ the per-particle capacity and $\mathcal{C}_t$ the per-first-passage time capacity of the point-to-point particle timing channel with constant emission rate are given by*

$$\mathcal{C}_q \le \log(1+\rho) - \log\frac{\bar{\lambda}}{e} - h(D) \quad (22)$$

and

$$\mathcal{C}_t \le \rho\left(\log(1+\rho) - \log\frac{\bar{\lambda}}{e} - h(D)\right) \quad (23)$$

where $\bar{\lambda}$ is the rate of particle emission, $h(D)$ is the entropy of the first-passage density $g(\cdot)$ with mean $1/\mu$ and $\rho = \bar{\lambda}/\mu$.

Proof: Theorem 1 See exposition leading to Theorem 1. •

If D is exponential with parameter μ, then $h(D) = 1 - \log\mu$ and we have

$$\mathcal{C}_t^{(e)} \le \rho\log\left(1+\frac{1}{\rho}\right) \quad (24)$$

If D is uniform with mean $1/\mu$, then $h(D) = \log\frac{2}{\mu}$ and we have

$$\begin{aligned} \mathcal{C}_t^{(u)} &\le \rho\left(\log\left(1+\frac{1}{\rho}\right) + \log\left(\frac{e}{2}\right)\right) \\ &= \mathcal{C}_t^{(e)} + \rho\log\left(\frac{e}{2}\right) \quad (25) \end{aligned}$$

IV. Discussion and Conclusion

In FIGURE 4 we plot the upper bound of equation (23) and the lower bound from [5] for exponential first-passage with mean $1/\mu$. The upper bound asymptotes to 1 which is

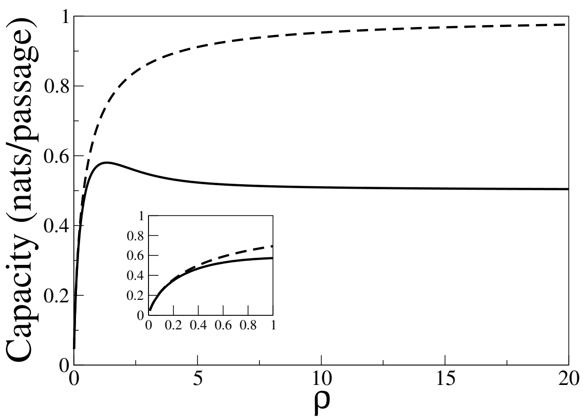

Fig. 4. **Upper and Lower Bounds on $\mathcal{C}_t$ for Exponential First-Passage:** $\rho = \frac{\lambda}{\mu}$, capacity in nats per first-passage time $(1/\mu)$. Upper bound approaches 1 for large ρ. INSET: detail $\rho \in (0, 1)$ showing close apposition of upper and lower bounds.

consistent with our previous inability to find input distributions $f_{\mathbf{T}}(\cdot)$ that show unbounded capacity with increasing ρ. However, since an exponential distribution maximizes entropy subject to a mean constraint [13], exponential first-passage minimizes the upper bound of equation (9). Thus, first-passage densities with the same mean but lower entropy would seem to allow for linear growth in particle intensity ρ as suggested by equation (25).

Of course, it must be noted that the upper bound for $H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ of equation (20) reflects only the mean first-passage time and not the amount of disorder imposed by different first-passage time distributions. That is, a distribution with a given mean could have arbitrarily small entropy which would tend to decrease $H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ and thus increase the penalty $M! - H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ in equation (5). So, there are likely tighter upper bounds on $H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ which explicitly include entropy, $h(D)$, in addition to the first-passage time mean $1/\mu$.

Our results do *not* directly apply to channels where particles never arrive at the receiver. Nonetheless, it is worthwhile to consider specific *methods* of non-arrival. If non-arrival results from excessive transit time (a fixed particle lifetime or a heavy-tailed first-passage distribution with infinite mean), then our results do not apply. However, in cases where particles are assumed to randomly "denature" independent of transit time (e.g., [7]), our results *do* provide an upper bound since random particle erasure can be interposed between reception (first-passage) and detection.

By their finest grain nature, the timing channel bounds *must* apply to all cases where independently emitted particles are used to convey information including most especially concentration channels. However, concentration channels stipulate time-varying particle emission rate while we have imposed a constant rate. Nonetheless, some insight can be gained by considering that if $\lambda(t)$ varies slowly with respect to the rate μ, we might credibly replace ρ in equation (20) with $\rho(t)$ and model such a system as successive channel uses at constant (but differing) rates $\rho_\ell = \rho(\ell\Delta t)$ with $E[\rho_\ell] = \rho$. Since equation (23) is *at most* linear in ρ, using $\rho_\ell \neq \rho$ confers no advantage and will at high rates be a disadvantage. So at least for slowly varying $\lambda(t)$ it seems likely that concentration channels will obey the bounds of equation (23).

Finally, we have not considered the effect of correlations between the T_m (as opposed to emission rate modulation), so the question of whether correlation can be used to increase the sum $H(\Omega|\vec{\mathbf{S}}, \mathbf{T}) + h(\mathbf{S})$ even while it reduces $h(\mathbf{S})$ is still open.

Overall, it is our hope that this work helps unify the field of molecular communication by allowing researchers to not only quickly compare results across different channel types, but also by supplying a simple-to-use sanity-check on capacity results.

Acknowledgments

This work was supported in part by NSF Grant CDI-0835592 as well as by funding from the Defense Advanced Research Projects Agency (DARPA). The views, opinions and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. Profound thanks are owed to I. Akyildiz, V. Anantharam, A. Eckford, M. Effros, N. Farsad, F. Fekri, A. Goldsmith, U. Mitra, V. Poor and S. Verdú for their time and generosity over many many years.

References

[1] T. Suda and T. Nakano, "Molecular communication: A personal perspective," *IEEE Trans Nanobioscience*, vol. 17, pp. 424–432, 2018. [Online]. Available: https://doi.org/10.1109/TNB.2018.2851951

[2] N. Farsad, H. B. Yilmaz, A. Eckford, C. B. Chae, and W. Guo, "A Comprehensive Survey Of Recent Advancements In Molecular Communication," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1887–1919, 2016, 3rd Quarter, 2016.

[3] A. Eckford, "Nanoscale communication with Brownian motion," in *CISS'07*, 2007, pp. 160–165, baltimore.

[4] C. Rose and I. Mian, "Signaling with identical tokens: Lower bounds with energy constraints," in *IEEE International Symposium on Information Theory (ISIT'13)*, July 2013, pp. 1839–1843.

[5] C. Rose and I. S. Mian, "Inscribed Matter Communication: Part I," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 2, no. 2, pp. 209–227, Dec 2016.

[6] ——, "Inscribed Matter Communication: Part II," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 2, no. 2, pp. 228–239, Dec 2016.

[7] N. Farsad, C. Rose, M. Medard, and A. Goldsmith, "Capacity of Molecular Channels With Imperfect Particle-Intensity Modulation and Detection," in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 2468–2472.

[8] M. Pierobon and I. Akyildiz, "Capacity of a Diffusion-based Molecular Communication System with Channel Memory and Molecular Noise," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 942–954, 2013.

[9] A. Einolghozati, M. Sardari, A. Beirami, and F. Fekri, "Capacity of Discrete Molecular Diffusion Channels," in *IEEE International Symposium on Information Theory (ISIT) 2011*, July 2011, pp. 603–607.

[10] A. Einolghozati, M. Sardari, and F. Fekri, "Design and Analysis of Wireless Communication Systems Using Diffusion-Based Molecular Communication Among Bacteria," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 6096–6105, 2013.

[11] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[12] K. V. Srinivas, A. W. Eckford, and R. S. Adve, "Molecular Communication in Fluid Media: the additive inverse gaussian noise channel," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4678–4692, July 2012.

[13] V. Anantharam and S. Verdu, "Bits Through Queues," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 4–18, January 1996.