Capacity Bounds on Point-to-Point Communication Using Molecules

Christopher Rose *Fellow, IEEE*, I.S. Mian and Mustafa Ozmen *Member, IEEE*

***Abstract*—Recent years have shown a rapid increase in the amount of study devoted to communication systems where molecules are information carriers. The reasons for such interest are varied, from seeking to understand the ubiquity of molecular communication in biology to the search for communication methods in media where electromagnetic and acoustic methods are inappropriate, to exploring the energy efficiency of methods where where some delivery latency can be allowed. With this tutorial on recent discrete molecular communication research we seek to organize the work into broad categories and thence under the umbrella of what can be called "inscribed matter communication" – where information is conveyed through assemblage, release and capture of matter as opposed to transmission of photons or phonons. We will begin by considering discrete passage of molecules between senders and receivers and argue that matter emission and detection even at Avogadrian levels is a subset of the discrete problem, all with a focus on point-to-point communication. In this way we hope to contextualize current work within a larger fundamental framework, illuminate the hard boundaries of what is known and then stimulate further research on this fascinating topic.**

I. Introduction

The earliest form of biological communication was arguably molecular. Primitive organisms navigated their environments seeking nutrients and avoiding the noxious. They sensed chemicals, moved along concentration gradients and up/down-regulated various cellular machines in ways that benefited survival. One could even argue that before the first organisms were organisms, chemical communication was the norm, but perhaps it is best to stop there before we lose sight of our goal here – the contention that molecular communication is ancient – and fall into the "it's turtles all the way down" trap about precedence [1].

In contrast, modern communication is predominantly electronic/acoustic and communication theory was developed primarily around these physical modalities. Nonetheless, communication and information theory are fundamental and can be applied to any system where information is exchanged. Our field therefore seeks to quantitatively understand the benefits and liabilities of molecular communication with both an eye toward exploiting such methods when electromagnetic or acoustic approaches are infeasible, and toward understanding why biological systems may have evolved the ways they did. Thus, everything old is new again as we consider molecular communication through a modern technological lens.

After such portentous rumblings about the utility and power of molecular communications research, it seems prudent to rein in expectations. We could certainly muse about fantastic potential applications and consider in detail the plethora of models articulated over the years since the papers which arguably launched the area [2]–[5] appeared. We could catalog the myriad biological instances of molecular communication as well. However, our purpose here is very simple. We seek only to provide an understandable, streamlined but unified framework for evaluating the capacity of *any* point-to-point molecular communication system. We also attempt to be rigorous but not so mathematically dense as to be discouraging [6], [7]. In the same vein, we do not provide a compendium of *all* work in the field, but rather, pick and choose some illustrative examples while also providing citations to more encyclopedic reviews such as [4], [8], [9]. Finally, we will not consider detailed biological systems, although we may gently point to some well known ones for illustration. Our hope is that this paper will help researchers to better see the connections between their work and readily identify bounds. And for those not yet part of the molecular communication fold we hope that the beauty, simplicity and potential scope of molecular communication systems will be enchanting enough lure you in.

To that end we will begin with a finest grain (micro) abstraction of point-to-point molecular communication applicable to *any* system. We will then explore a number of variations on this basic theme, but with the common feature that individual or small numbers of molecules are used as the information carriers. We will always assume perfect detection since via the data processing theorem [10] the capacity of any channel cannot be increased by post-processing. We will also briefly consider cases where the information particles themselves can carry information payloads. Then, since most laboratory-realizable systems use Avogadrian numbers of molecules[1] for signaling, we will consider (macro) models where concentration (as opposed to individual molecules) is detected. These macro systems are often analytically similar to more standard "X"-shift-keying communication systems and use similar signal corruption (noise) models. Nonetheless, we will always be careful to interpret such work through the lens of our micro models since, via the data processing theorem, the capacities of macro signaling systems must lie within the bounds dictated by micro models.

II. A Pictorial Overview

We feel the basic idea of molecular communication is most easily understood through a series of pictures to which we will add conceptual and mathematical complications as we go.

A. Particle-Based Systems

The simplest scenario is depicted in Fig. 1 where a chemical emitter releases identical molecules (or particles) into a

[1]We estimate that [11]–[13] expended on the order of 10^{17} to 10^{19} molecules per emission and thus Avogadrian (6.022×10^{23}) or more molecules over the course of 10-1000 kb messages.

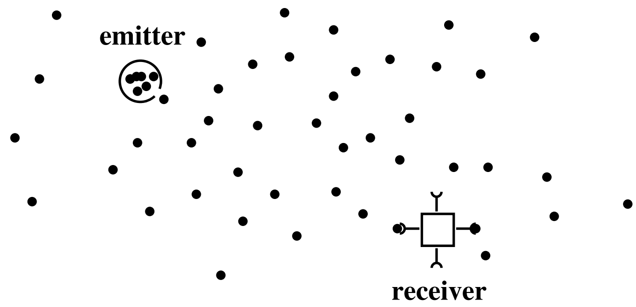

Fig. 1. A chemical source emits identical molecules/particles (solid circles) over time that waft through some medium and are detected at a receiver by specialized receptors. The fundamental control/detection abstraction quantities are the ordered departure times $\{T_i\}$ from the emitter, and the ordered detection times $\{S_j\}$ at the receiver as described in the text. Particles may be detected multiple times (or not at all) at the receiver, depending upon the (usually stochastic) particle transport method at play.

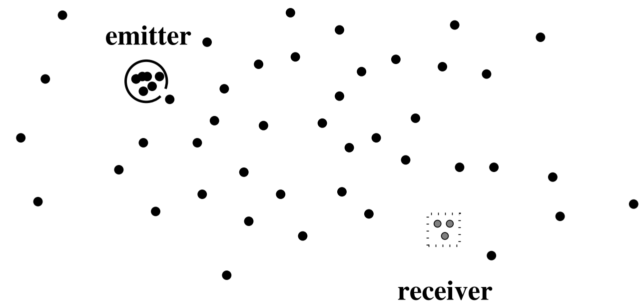

Fig. 2. Molecular communication system identical to Fig. 1 except that particles piercing the receiver boundary are perfectly detected and thence removed from the system (gray circles).

medium and those particles are transported to and detected by a receiver sensitive to that chemical. Loosely speaking, information could be conveyed by the timing and number of molecules released and it is tempting to immediately plunge in and model the multiplicity of transport and uptake mechanisms. However, we pause to note the first *inviolable* mathematical abstraction under an assumption of identical particles is a random set of release times (hence, uppercase)

$$\mathbf{T} = [T_1, T_2, \cdots, T_M] \tag{1}$$

and an ordered sequence of random detection times at the receiver

$$\vec{\mathbf{S}} = [\vec{S}_1, \vec{S}_2, \cdots, \vec{S}_{M'}]. \tag{2}$$

We assume this process of a "channel use" is repeated sequentially and independently so that asymptotic information theoretic capacity bounds based on mutual information may be applied.

It is *vital* to understand that no matter what sophisticated methods are applied, the capacity of the molecular communication channel between emitter and receiver is completely determined by how we structure the probability density of $\mathbf{T}$ and how we decode the corresponding $\vec{\mathbf{S}}$. That is, all molecular communication channels are, at their core, timing channels. Put another way, timing channels are the finest grain description of all molecular communication channels.

It is first important to note that the ordered arrivals S_i may or may not correspond (in index) to the ordered emissions T_i since owing to particle motion uncertainty, some particles may take longer to arrive at the receiver. This complication requires us to apply a bit of mathematical legerdemain [6], [7], but there is a more subtle issue as well. That is, it is even *more important* to notice that M' may or may not be equal to M for at least two possible reasons.

First, the depiction of Fig. 1 implicitly assumes operation in "free space" where particles are free to roam where they will. Depending upon the nature of stochastic particle motion, there may be nonzero probability that any given particle will *never* arrive at the detector (null recurrence [14]). Second, given what we know about detection in biological systems, particles (called "ligands" in the biological literature) can bind and *unbind* at the "receptor", a typically more complex molecule that binds preferentially to the ligand. Thus, a single particle may be detected multiple times at the receiver so that multiple $\vec{S}_j$ may correspond to a single T_m. Furthermore, a bound particle may block detection of otherwise detectable particles at the receiver, further complicating the system.

Once again, it is tempting to dive into the details of binding kinetics and concomitant detection issues, but for now we concentrate on repeated particle detection and blocking. 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. It can then be easily shown that subsequent detection times of the same particle convey no additional information because later re-detection times are independent of emission time given the first detection time [5]. Likewise, if a particle is blocked from detection at the receiver, we cannot gain information. Therefore we can imagine a model where all the detection times $\vec{s}_j$ are first detections, we ignore subsequent detections, effectively removing that particle from the system, and we assume that the first arrival of any particle near the receiver will be detected.

We now have the simplified model of Fig. 2 which provides an upper bound on molecular channel capacity without requiring us to restrict our attention to specific particle transport or detector uptake/collision models. Emitted particles pierce an imaginary detector boundary and their arrival times are recorded. They are thereafter removed from further consideration. Certainly were detected particles not removed, subsequent re-detections could not be differentiated from first detections and a loss of capacity would result. Likewise, receptor blocking would lead to capacity loss. Thus, this simplified model of Fig. 2 provides an upper bound for all systems where the transport characteristics are independent of particle release times, $\{T_m\}$.

Of course, the simplification of Fig. 2 does not solve the problem of "missing" detections owing to non-arrival of a particle at the receiver. We could imagine requiring that particles arrive in finite time, but that imposes a condition on the transport model statistics. We could instead simply imagine that emissions are erased in an independent identically distributed (i.i.d) way under the assumption that particles travel to the receiver independently. A significant variation on this

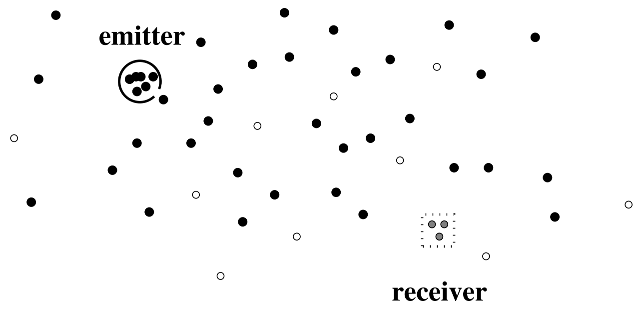

Fig. 3. Molecular communication system identical to Fig. 2 except that particles can be neutralized/degraded (open circles) before they reach the receiver.

theme is to assume particles have finite (and likely stochastic) lifetimes either through degradation in the medium or active and deliberate chemical gettering (removal/incapacitation of signaling particles) in the channel design [15]. This type of model augmentation is depicted in Fig. 3 where some particles (unfilled) have been rendered undetectable/extinct either through aging or gettering. We will later see that such particle deactivation can improve channel performance by reducing timing ambiguity at the expense of increased emission energy.

The issue of null recurrence (non-arrival) of particles can also be treated by an appeal to the physics of many practical molecular communication scenarios – compartmentalization. That is, free-space diffusion without drift is null-recurrent with particles potentially never appearing at the receiver. However, many systems have physical boundaries that force the probability of a particle non-arrival to zero. One can even combine boundaries and particle extinction. These two scenarios are depicted in Fig. 4.

Finally, as the natural culmination of point-to-point particle systems, one might consider a multiplicity of different particles wherein information is conveyed not only through timing but through the implicit information payload associated with particle identity. Such a system is depicted in Fig. 5. The same issues as associated with timing channels pertain, but there is also the issue of stitching together a message from particles that could arrive in arbitrary order [6], [16].

B. Concentration-Based Systems

Collectively, Fig. 1 through Fig.5 cover the full range of finest grain molecular communication models. However, since practical systems may operate through release of large numbers of molecules, one can also imagine a system wherein some time-varying amount of chemical is emitted, $C_e(t)$ and the receiver detects a time-varying concentration $C_r(t)$ as opposed to individual particle arrivals. A concentration-based system *cannot have higher capacity than the fine grained systems depicted in Fig. 1 through Fig.5* since time-varying concentrations are a deterministic function of the detailed particle release and capture timing. That is

$$C_r(t) = \frac{1}{V} \left| \{\vec{S}_i \in [t, t+\Delta]\} \right| \tag{3}$$

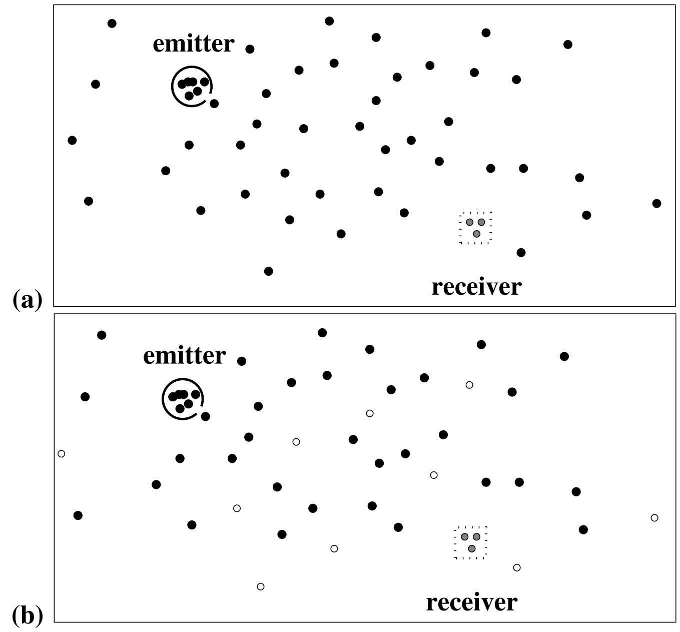

Fig. 4. Molecular communication system identical to Fig. 2 **(a)** and Fig. 3 **(b)** except compartmentalization results in finite-time detection of particles at the receiver.

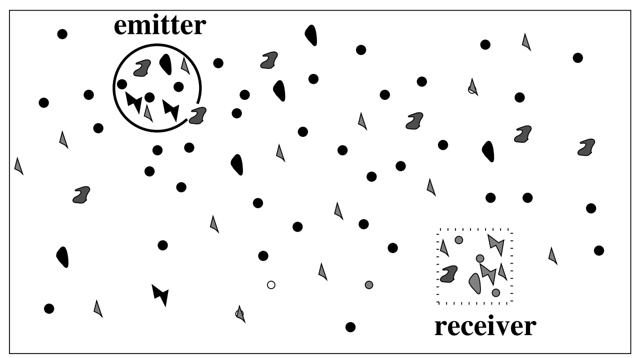

Fig. 5. Compartmentalized molecular communication system with different particles

where V is the volume sampled by the sensor, Δ is the integration interval and $|\cdot|$ indicates the cardinality of the enclosed set. In the limit of vanishing Δ it is obvious that

$$VC_r(t) = \lim_{\Delta \to 0} \left| \{\vec{S}_i \in [t, t+\Delta]\} \right| \tag{4}$$

is a point process that indicates the receiver arrivals $\{\vec{S}_i\}$.

A good deal of work in molecular communication considers such models owing to mathematically tractable descriptions of the transport medium [17] (i.e., various types of diffusion) and to date all experimental systems are concentration-based. Nonetheless these models must obey the same fundamental limits as precise timing models. Thus, all the results presented for particle systems must carry over after adjustment for the much higher particle intensity associated with concentration-based systems.

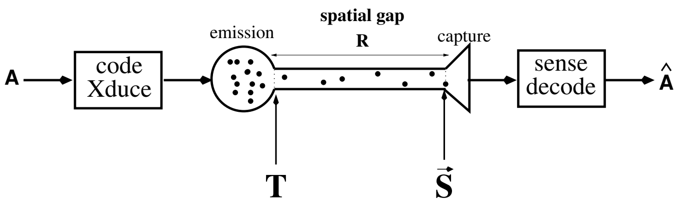

Fig. 6. General Molecular Communication Channel Abstraction: a message $\boldsymbol{A}$ is coded and then transduced into a set of particle emission times $\mathbf{T}$. These particles propagate over a spatial gap R through a transmission medium and are captured (exactly once) at corresponding times $\mathbf{S}$. Since the particles are identical, capture results in the ordered arrivals $\vec{\mathbf{S}}$. These ordered arrivals are sensed and decoded into the message estimate $\hat{\boldsymbol{A}}$.

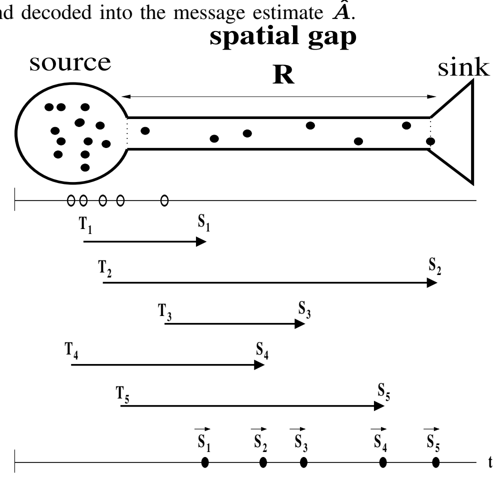

Fig. 7. A More Detailed Look at Fig. 6. A sender transmits an ensemble of particles ("inscribed matter") to a receiver across a spatial gap (of length R in the figure). The particles are released at (unordered) times $\{T_m\}$, propagate through a transmission medium and are captured at corresponding times $\{S_m\}$. For identical particles, the receiver sees ordered arrivals $\{\vec{S}_m\}$ which may differ in index from the unordered arrivals $\{S_m\}$. Particles themselves may or may not carry information payloads.

C. A High Level Abstraction

Overall, the models of Fig. 1 through Fig.5 can be abstracted as depicted in Fig. 6 This basic arrangement is a staple of the field and can be found in various forms in a variety of prior work (see [8] for a survey). A message is composed, coded into chemical emission patterns and released into a medium that 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. What we consider here will concern only the central portion of the diagram as shown in Fig. 7. The information carriage from $\mathbf{T}$ to $\vec{\mathbf{S}}$ which although modeled as particle release naturally includes concentration models in the limit of Avogradrian numbers of particles.

III. Mathematical Details

Using Fig. 2 through Fig.6 as our guides, we can now consider the mathematical modeling details. In the interests of clarity we try not to provide derivations since most of the details of what we present here are available elsewhere. Rather, we simply seek to limn the analytic motifs that recur in the study of molecular communication channels.

A. Identical-Particle Timing Channels

Consider Fig. 2 where identical particles are emitted and then captured (and removed) after transit to the receiver – and we admit the possibility that a particle may arrive only at time ∞. 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 each arrival corresponds to a single emission, we can define M-vectors $\mathbf{T}$, $\mathbf{D}$ and $\mathbf{S}$ accordingly:

$$\mathbf{S} = \mathbf{T} + \mathbf{D} \tag{5}$$

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

$$\vec{\mathbf{S}} = P_\Omega(\mathbf{S}) \tag{6}$$

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

$$\vec{\mathbf{S}} = P_\Omega\left(\mathbf{T} + \mathbf{D}\right) \tag{7}$$

and Ω is a discrete random variable associated with the channel. Equation (7) is the basic description of a "timing channel" wherein emissions $\mathbf{T}$ are constructed so that information can be extracted from $\vec{\mathbf{S}}$. The essence of this channel is abstracted in Fig. 7.

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, the most popular of which is the Levy distribution derived from diffusion characteristics [15] and the related additive inverse Gaussian channel model [18]. A more complete list of timing channel models can be found in [8] and references therein. In this paper we will not delve into the various possible models, but rather, we establish a (hopefully transparent) framework under which all such channels can be evaluated.

That said, "timing channel" is a bit of a misnomer since a sort of amplitude modulation is certainly possible with simultaneous (or temporally intense) release of multiple particles. Thus, with the understanding that timing channels encompass all other identical particle channel models, we will not mention amplitude modulation again until we consider ephemeral particle models as depicted in Fig. 3 where amplitude modulation is a more natural description.

It is also worthwhile noting that "timing channels" appear in a variety of contexts, most notably in the award-winning paper "Bits Through Queues" [19] and follow-on work [20], [21]. The difference between the bits through queues model and the particle timing channel is the lack of a queue – or perhaps better said, an infinity of servers [6]. That is, we explicitly assume particle detections at the receiver do not interfere with one another. Bits through queues channels implicitly assume particle-particle competition for service (detection).

Photonic channels with precisely timed arrival detection at the receiver are also a sort of timing (and concentration, via intensity detection) channel. Under an assumption of stochastic (but perfect) detection, release-time uncertainty at the transmitter and/or a scattering medium, the particle (photon) transit time between transmitter and receiver would be stochastic and the model of equation (7) would pertain. However, unavoidable thermal noise at photonic receivers results in missed (or worse yet, spurious) detections at the receiver. The missed detections could possibly be modeled as erasures, but there is no analog for spurious detection unless the detection process is considered – which in search of outer bounds we do not do here.

Finally, since the word "timing" has through repeated use here become a sort of bludgeon, it is perhaps important to note that like [22] and many others, we assume perfect synchronization between transmitter and receiver. However, we also note that so long as the "clocks" can be synchronized at the M-particle channel use level where M can be arbitrarily large, some types of random clock skew can probably be incorporated into the first-passage time distribution description. Otherwise, a synchronization overhead penalty of at least the clock skew entropy rate would be imposed. More precise treatments of the general problem can be found in [23]–[25].

Persistent Particles: Suppose particles persist until they are captured at the receiver – however long that takes. The maximum mutual information between the emission ensemble $\mathbf{T}$ the input and arrival ensemble $\vec{\mathbf{S}}$ is the natural definition of channel capacity. However, we would then require a signaling model that supports the usual asymptotically large block length and repeated independent sequential channel uses paradigm [10, (chapt 8 & 10)]. In addition, we must also pay attention to energy usage since lack of energy constraints can lead to unrealistic results. We therefore define a channel use as the launch and capture of M particles under an emission deadline constraint, τ, with the further constraint that

$$\lambda\tau = M \tag{8}$$

where λ, the particle launch average intensity, has units of particles per time. Equation (8) is implicitly a constraint on average power assuming a fixed per-particle energy cost for construction/sequestration/release/delivery. We also note that the signaling interval τ is now an explicit function of M as in

$$\tau = \tau(M) = \frac{M}{\lambda}$$

So, consider Fig. 8 where sequential M-particle transmissions – channel uses or symbol intervals – are depicted. We will assume a "guard interval" of some duration $\gamma(M, \epsilon)$ between successive transmissions so that all M particles are received before the beginning of the next channel use with probability $(1 - \epsilon)$ for arbitrarily small $\epsilon > 0$. This condition guarantees the (asymptotically) independent channel uses necessary for a patent information-theoretic channel

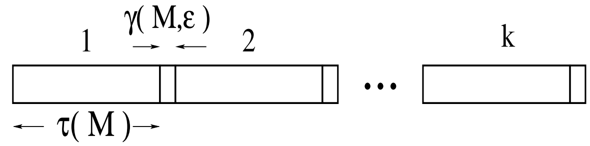

Fig. 8. Successive M-emission channel uses. For a given use of the particle timing channel, the sender emits M particles over the transmission interval $\tau(M) = \frac{M}{\lambda}$. $\gamma(M, \epsilon)$ is the waiting period (guard interval) before the next channel use.

description. We further require that the average emission rate, $M/(\tau(M) + \gamma(M, \epsilon))$ satisfies

$$\lim_{\epsilon \to 0} \lim_{M \to \infty} \frac{M}{\tau(M) + \gamma(M, \epsilon)} = \lambda \tag{9}$$

We then require that the last particle arrival time $\vec{S}_M$ occurs before the start of the next channel use with probability 1.

It has been shown [6] that if the expected value of the first-passage time D is finite, then the assumption of asymptotically independent channel uses holds. However, it has also been shown that if $E[D]$ is infinite, then it is impossible to provide such a signaling model and the usual channel capacity analysis is ill-posed. Interestingly, popular transport models such as simple drift-free diffusion have infinite first-passage times, so the capacity question is moot for such systems in the context of Fig. 1 and Fig. 2. However, in any practical system there are boundaries or compartments over which particles may roam as depicted in Fig. 4a which results in $E[D] < \infty$ and renders the system tractable from an information theoretic standpoint.

Assuming $E[D] < \infty$ we can now consider the mutual information $I[\vec{\mathbf{S}}; \mathbf{T}]$ as a measure of channel capacity. However, the ordering operation that produces $\vec{\mathbf{S}}$ renders this difficult in cases where $\mathbf{T}$ may be arbitrarily distributed. However, for nonsingular first-passage time distributions on $\mathbf{D}$, $\mathbf{S}$ will be continuous. It can then be shown that the differential entropy of $\vec{\mathbf{S}}$ is

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

We then note the equivalence

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

which allows us to write

$$h(\mathbf{S}|\mathbf{T}) = h(\Omega, \vec{\mathbf{S}}|\mathbf{T}) = h(\vec{\mathbf{S}}|\mathbf{T}) + H(\Omega|\vec{\mathbf{S}}, \mathbf{T}) \tag{10}$$

where, since Ω is a discrete random variable, $H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ is the discrete entropy of Ω given $\vec{\mathbf{S}}$ and $\mathbf{T}$. Equation (10) and leads immediately to

$$I(\vec{\mathbf{S}}; \mathbf{T}) = I(\mathbf{S}; \mathbf{T}) - \left(\log M! - H(\Omega|\vec{\mathbf{S}}, \mathbf{T})\right) \tag{11}$$

Equation (11) 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, $\left(\log M! - H(\Omega|\vec{\mathbf{S}}, \mathbf{T})\right)$. 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. Evaluating $H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ and then maximizing equation (11) requires some effort [6], [7]. Nonetheless, there are some useful and general results which we now state.

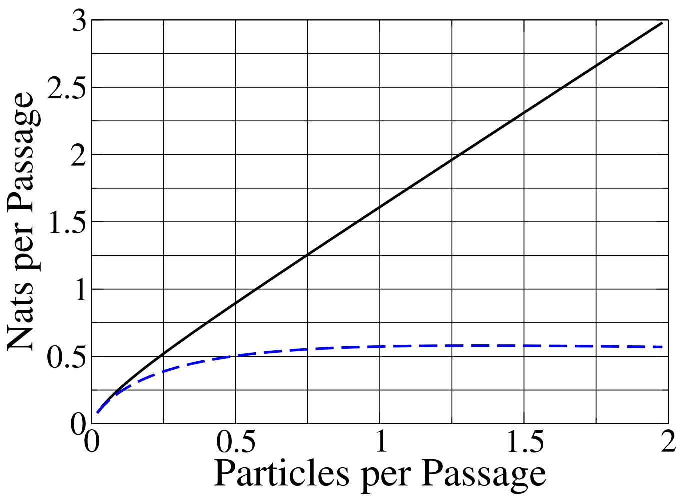

Fig. 9. Lower bound (dashed line: equation (14) and upper bound (solid line: equation (15)) for the persistent particle timing channel capacity C_t (in nats per passage time $1/\mu$) as a function of channel load ρ, the ratio of the particle emission rate λ to the particle uptake rate μ. Exponential first-passage assumed.

First, if we set the signaling/symbol interval to $\tau = M/\lambda$ where λ is the average rate at which particles are emitted (see equation (9)) then we can define a capacity per particle as

$$C_q = \lim_{M\to\infty} \frac{1}{M} \sup I(\vec{\mathbf{S}};\mathbf{T}) \tag{12}$$

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

$$C_t = \lambda C_q \tag{13}$$

It was shown in [19] that exponential first-passage minimizes the mutual information between T and S under a mean constraint on both D and T. So while the distribution on D that minimizes the mutual information between $\vec{\mathbf{S}}$ and $\mathbf{T}$ remains unknown at this time [7], exponential D is still a useful reference point. So, assume D is exponential with mean $1/\mu$ and define particle emission intensity as $\rho = \lambda/\mu$. In units of nats per mean first-passage time the channel capacity is at least

$$C_t(\rho) = \rho C_q(\rho) \geq \rho \left(\log \frac{1}{\rho} + \frac{1}{\rho} E[\ell \log \ell] \right) \tag{14}$$

where ℓ is a Poisson random variable with mean ρ. Equation (14) initially increases with ρ, reaches a maximum (which must be calculated numerically) and then settles to an asymptotic value of 0.5 nats. Likewise we have an upper bound, also in nats per mean first-passage time,

$$C_t = \rho C_q(\rho) \leq \rho \log \left(\frac{1}{\rho} + 4 \right) \tag{15}$$

The bounds given in equation (14) and equation (15) on C_t (taken from [6], [7]) are plotted in Fig. 9.

The lower and upper bound agree well for smaller values of ρ but diverge for $\rho \gg 0.25$. That said, an input distribution on $\mathbf{T}$ that does better than the lower bound has yet to be found. So, we suspect the lower bound lies close to (or is coincident with) the true upper bound. If this is true, then the lower bound is even more interesting in that it suggests particle emission much in excess of the mean first-passage time may reduce capacity. *It is important to note* that first-passage time uncertainty (jitter, or entropy) produces disordered particles. That is, the mean first-passage time is only a measure of channel *latency* – the "propagation delay" so to speak – and does not itself impact particle order uncertainty. However, for exponential first-passage, the uncertainty is an explicit function of the mean first-passage time $1/\mu$ so here we use mean first-passage time and timing jitter/entropy interchangeably.

Perhaps most important of all, if we assume equation (14) lies close to the capacity upper bound, then rates on the order of 0.5 nats per passage time are the best we can do with identical particle channels with exponential passage. Since timing channels are finest grain, all other coarser models must be similarly capacity-constrained. It is therefore noteworthy that in [17] (figures 4 and 5) rates on the order of 3 kb/s were derived for emitter-receiver separations in the range of $50 - 500\mu$m in a diffusive medium. This discrepancy is worth exploring in a little more detail.

The mean first-passage time in a constrained 1-D system where the particle starts at $x = 0$ and is absorbed at $x = \pm R$ is $\frac{R^2}{2\mathcal{D}}$ where $\mathcal{D}$ is the diffusion coefficient. The standard deviation (jitter) of the first-passage time is on the order of its mean, and it is this jitter that determines the arrival uncertainty. The diffusion coefficient of water ($10^{-9}\mathrm{m}^2/\mathrm{s}$) was used in [17] with emitter-source separations ranging from 50μm to 500μm which would imply mean first-passage times in the range $E[D] \in (1.25, 125)$s with comparable standard deviations. Thus, the kilobit range capacities reported in [17] exceed the capacities predicted by equation (14) by many orders of magnitude. This disparity could be a result of the way capacity was defined in [17] without stipulation of a signaling interval that allows the independent (ISI-free) channel uses necessary for application of the channel coding theorem [10]. It could also be a result of the specific first-passage time distribution – it is well-known that exponential first-passage produces min-max $I(S;T)$ [19] while non-exponential first-passage *may* allow linear growth in capacity [26]. Regardless, if the achievable upper bound truly diverges from the lower bound as in Fig. 9, then perhaps such high (or even higher) rates are accessible through the increasingly intense particle emissions implicit in [17].

Ephemeral Particles: A fundamental difficulty with identical particle channels is particle persistence which leads to intersymbol interference (ISI). We thus took care to define signaling intervals that displayed zero ISI in the limit so that standard information-theoretic techniques could be applied to derive capacity. However, what if we could guarantee that particles that did not arrive at the receiver by a certain time would disappear from the system thereby limiting the temporal extent of ISI?

A number of previous studies have, in various ways, imposed exactly this sort of condition either by limiting the emission period and extending the detection period so that

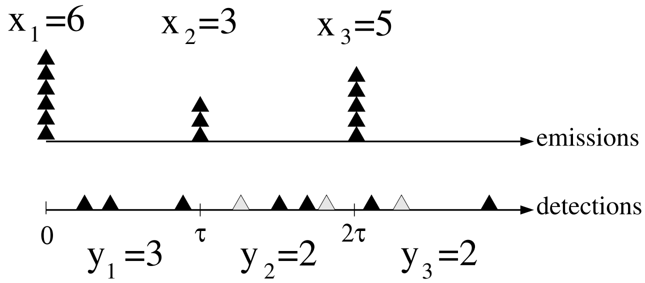

Fig. 10. Depiction of the particle-intensity channel [15]). x_i particles are emitted at the start of each symbol interval and y_i particles detected. If the transit time of a particle exceeds τ, that particle is "retired" and cannot be detected so that the possibility of intersymbol interference is precluded.

only particles emitted in interval k would arrive during interval k, or by assuming particle lifetimes through design or explicit injection of gettering agents [27]–[29]. Extending the detection period is simply a coding restriction on the persistent particle timing channel and therefore *must* have lower capacity. The same is true of erasures. Imposing strict lifetime deadlines, however, provides a type of implicit side-information to the receiver that can greatly increase capacity.

Following [15], suppose we design an emission strategy where we periodically release different numbers of particles in a burst at the beginning of each signaling interval and then record how many arrive during the associated detection interval. The basic idea is illustrated in Fig. 10 where three particle bursts (emissions) and their associated detections are depicted. Particles that arrive within τ of their release time are detected (black) while those whose transit time exceeds τ are rendered undetectable (gray). [15] also considers the possibility of release imperfection (attempting to emit m particles but emitting m' instead) and detection imperfection (mistaking the number of valid arrivals). However, we will ignore these complications since they can only serve to decrease capacity and as with the timing channel, we seek outer bounds.

This channel discipline lends itself to a particularly simple description of Fig. 11 where as many as $X = M$ particles could be released, each having a probability p of reaching the receiver within travel time τ, but only $Y \leq X$ are detected. Furthermore and perhaps most important, the survival probability is a monotonically increasing function of τ since it is the cumulative distribution function (CDF) of the particle transit time distribution. If τ is small, then the channel can be used rapidly, but fewer of the particles are likely to arrive on time which decreases the per-channel-use capacity (which, incidentally, is the reason that all particles are released at the start of an interval, maximizing the number of particles that arrive in time). If τ is large, then the number of input particles is more easily distinguishable which increases per-channel-use capacity, but the channel is used less frequently.

We can compute the ephemeral particle channel capacity by calculating the capacity of the channel in Fig. 11 and dividing by τ since all channel uses are, by design, independent. However, since no closed form exists we can consider a related suboptimal Z-channel wherein either 0 or M particles are

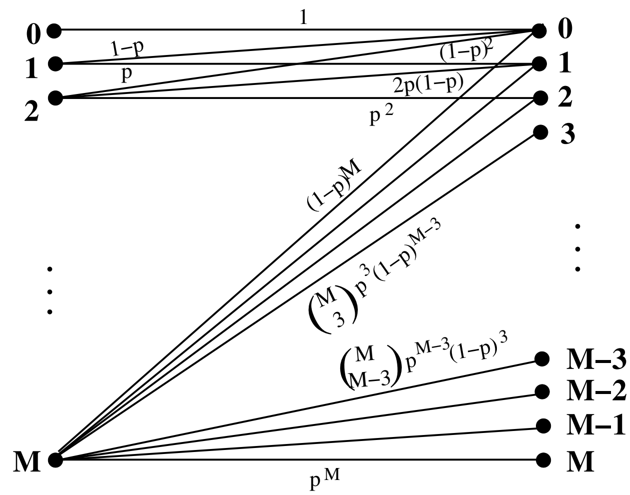

Fig. 11. Particle intensity channel diagram with survival probably p. p is a monotonically decreasing function of the symbol interval τ.

released and the arrival of any particles constitutes a "1" while no arrivals are interpreted as "0" [15]. The Z-channel crossover probability is therefore $(1-p)^M$ and the capacity per channel use is

$$C_Z(p) = \log\left(1 + (1-(1-p)^M)(1-p)^{M\frac{(1-p)^M}{1-(1-p)^M}}\right)$$

So that we can relate the result to channels with persistent particles, let us assume exponential first-passage of the particles between emitter and receiver with mean $1/\mu$ so that

$$p(\tau) = 1 - e^{-\mu\tau}$$

The mean particle intensity is

$$\lambda = \frac{M(1-\alpha)}{\tau} \tag{16}$$

where α is the capacity-optimizing probability of sending no particles. If we define

$$a = (1-p(\tau))^M$$

the probability of "crossover", then α can be written as

$$\alpha = 1 - \left(\frac{1}{1-a}\right)\frac{1}{\left(1+\frac{1}{1-a}\frac{1}{a^{\frac{a}{1-a}}}\right)} \tag{17}$$

The capacity of the system is

$$C_E(\tau) = \frac{1}{\tau}\log(1+(1-a)a^{\frac{a}{1-a}}) \tag{18}$$

Now, deviating from the course taken in [15], suppose $\mu\tau$ is small enough that the probability of a particle reaching the emitter is small

$$p(\tau) \approx \mu\tau \ll \frac{1}{M} \tag{19}$$

We then have

$$a = e^{-M\mu\tau} \approx 1 - M\mu\tau$$

and

$$1 - a \approx M\mu\tau$$

so that

$$\begin{aligned} C_E(\tau) &\approx \tfrac{1}{\tau}\log\left(1 + M\mu\tau(1-M\mu\tau)^{\frac{1-M\mu\tau}{M\mu\tau}}\right) \\ &\approx \tfrac{1}{\tau}\log\left(1 + M\mu\tau(1-M\mu\tau)^{\frac{1}{M\mu\tau}}\right) \end{aligned}$$

which in the limit of small τ becomes

$$C_E(\tau) \approx \frac{\mu M}{e} \tag{20}$$

Equation (20) implies that *capacity increases linearly with* M for very short symbol intervals so long as $p(\tau)$ approaches zero linearly or sublinearly in small τ. $p(\tau)$ is indeed linear in small τ for both exponential passage and Brownian motion, with and without drift. Therefore, in contrast to persistent particle timing channels where increasing particle intensity might not increase capacity, engineering finite particle lifetime can confer impressive capacity increases limited only by the number of particles available for injection during symbol intervals and the rate at which $p(\tau)$ approaches zero in τ.

B. Non-Identical Particle Channels

So far, we have only considered timing (including emission intensity) as the information carrier. However, it is possible that the particle also carries information, much as a "packet" carries information over the Internet, essentially by virtue of its identity as depicted in Fig. 5. So, imagine we have Q identifiable particles from which to choose and we can emit these particles into the medium at arbitrary times. How much information can be conveyed?

Persistent Particles: If the particles are persistent, one might consider avoiding timing altogether and assemble messages as combinations of particles. With particle intensity λ and observation interval τ we will have on average $M = \lambda\tau$ particles. The number of possible distinguishable combinations of M particles choosing from a library of size Q is $\binom{M+Q-1}{Q-1}$ and if we allow between zero and M particles on $(0, \tau)$ there are

$$N = \frac{M+1}{Q}\binom{M+Q}{Q-1}$$

distinct combinations. The data rate is then

$$R = \frac{1}{\tau}\log\left(\frac{M+1}{Q}\binom{M+Q}{Q-1}\right)$$

where we assume $M \propto \tau$. It is easy to see that $R \to 0$ as $\tau \to \infty$ which suggests that smaller intervals must be used sequentially which again raises the specter of ISI.

So rather than worry about re-engineering guard intervals between symbols to combat ISI, suppose that we can use timing information and the channel discipline previously developed for it. Further, assume each particle can carry $\log_2 Q$ bits of information and that we wish to string particles together into a message. To recover the original message at the receiver, *each particle must be uniquely identifiable* otherwise message reconstruction is impossible. The simplest approach is to add a sequence number [6], [16]. Given M particles per channel use, we could obviously append $\log M$ bits to each particle and likely this would be sufficiently efficient. Alternatively, we might also consider gross structural differences in the particles – sending particles of distinct lengths $1, 2, \cdots, K$ where $M = K(K+1)/2$, for instance or some other clever structural embedding. Nonetheless, it seems worth noting that $H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ provides exactly the measure of essential particle "overhead" or "side-information" (of any form) necessary to maintain proper sequence.

Consider that operation of the timing channel involves construction of deterministic blocklength-N codewords $\{\mathbf{t}_1, \cdots, \mathbf{t}_N\}$ where each $\mathbf{t}_n$ represents the emission schedule for M particles (a channel use). If capacity is not exceeded, the receiver can reliably recover the information embedded in the codewords and since we generally assume the receiver has access to the coding method, a correctly decoded message implies knowledge of the constituent codewords $\{\mathbf{t}_1, \cdots, \mathbf{t}_N\}$. However, the channel imposes residual uncertainty about the mapping $\mathbf{S} \to \vec{\mathbf{S}}$ – the ambiguity about which $\vec{S}_i$ is associated with which S_j. For this reason, the payload-inscribed particles cannot yet be correctly strung together to recover the message.

However, given the observed arrivals $\vec{\mathbf{S}}$ and the correctly decoded $\mathbf{t}$, $H(\Omega|\vec{\mathbf{S}}, \mathbf{t})$ is *the definition* of the uncertainty about that ordering, Ω. Likewise, the average uncertainty is $H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$. Thus, the source coding theorem implies that at least $H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ bits must be used, on average, to resolve the mapping ambiguity.

It is important to note that we have not actually provided a *method* for message reconstruction, only a lower bound on the amount of "side information" necessary at the receiver to assure proper reconstruction. However, as a practical matter, the quantity $\frac{1}{M}H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ does provide some guidance. In the worst case where the order of particle arrival is completely random, $H(\Omega|\vec{\mathbf{S}}, \mathbf{T}) = \log M!$ which amounts to each packet carrying a header of size $\frac{1}{M}\log M! \approx \log M$ for large M – essentially numbering the packets from 1 to M. If $\frac{1}{M}H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ is much smaller, then cyclic packet numbers could be useful since smaller $\frac{1}{M}H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ implies that packets are unlikely to arrive grossly out of order. The sequence header could then be commensurately smaller. In either case, the total amount information necessary to resolve the ordering Ω is $\frac{1}{M}H(\Omega|\vec{\mathbf{S}}, \mathbf{T})$ per packet on average.

Ephemeral Particles: To our knowledge there has been no work on sequencing for payload-bearing particles with finite lifetimes, and given the improvements finite lifetime affords timing channels, sequencing for ephemeral particles might be an interesting line of investigation. For instance, one could imagine fountain-like codes [30] given the fact that increased particle intensity leads to higher probability of capture within a given symbol interval. However, the same sequencing issues and associated overhead would apply, and under the simultaneous release discipline, timing information could not be used to reduce the necessary sequencing information that must be carried by each particle. However, since the symbol interval is crisply defined by particle lifetime and only finite information can be sent during this fixed interval, the use of explicit sequence numbers would be a good low overhead solution.

IV. Energy Considerations

Just as for more typical communication channels, capacity evaluation without consideration of energy can lead to erroneous conclusions. So we posit simple but reasonable energy requirements for molecular communication – although unlike [31] we assume that actual transport is "for free" through any number of stochastic collision processes. Suppose the construction cost for a particle without a payload is c_0 Joules and with a payload c_1 Joules. Symbolic string particles incur a "per character" cost which we define as Δc_1 per character per particle. For example, adding a nucleotide to double-stranded DNA requires 2 ATP (1.6×10^{-19} J) while adding an amino acid to a protein requires 4 ATP (3.2×10^{-19} J) [32]. There may also be other energy involved in sequestration, release and/or particle transport across a gap. However, the key assumption is constant energy expenditure per particle. Without considering the details as in [33] we will denote the combination of these and any other relevant energies as c_e Joules per particle. Thus, the power required for the timing-only channel in Joules per first-passage time is

$$\mathcal{P}_T = \rho(c_0 + c_e) \tag{21}$$

which we rewrite as

$$\rho = \frac{\mathcal{P}_T}{c_0 + c_e} \tag{22}$$

The ephemeral-particle channel average energy is also

$$\mathcal{P}_E = \rho(c_0 + c_e) \tag{23}$$

so that

$$\rho = \frac{\mathcal{P}_E}{c_0 + c_e} \tag{24}$$

except that ρ is defined using λ as in equation (16). For the timing-plus-payload channel we have

$$\mathcal{P}_{T+P} = \rho\left(c_1 + c_e + \left(\frac{H(\Omega|\mathbf{T},\vec{\mathbf{S}})}{\log b} + K\right)\Delta c_1\right) \tag{25}$$

where K is the string length of information-laden particles, and b is the alphabet size used to construct the strings so that each particle carries a payload of $Q = K \log b$ bits. $H(\Omega|\mathbf{T},\vec{\mathbf{S}})$ is the emission order uncertainty induced by independent particle passage through the channel. We then have

$$\rho = \frac{\mathcal{P}_{T+P}}{\left(c_1 + c_e + \left(\frac{H(\Omega|\mathbf{T},\vec{\mathbf{S}})}{\log b} + K\right)\Delta c_1\right)} \tag{26}$$

The capacity of the timing-only channel in bits per mean first-passage time is

$$\mathcal{C}_T = \rho C_q(\rho) \tag{27}$$

Similarly, the capacity of the particle-timing+payload channel is

$$\mathcal{C}_{T+P} = \rho\left(C_q(\rho) + K\log b\right) \tag{28}$$

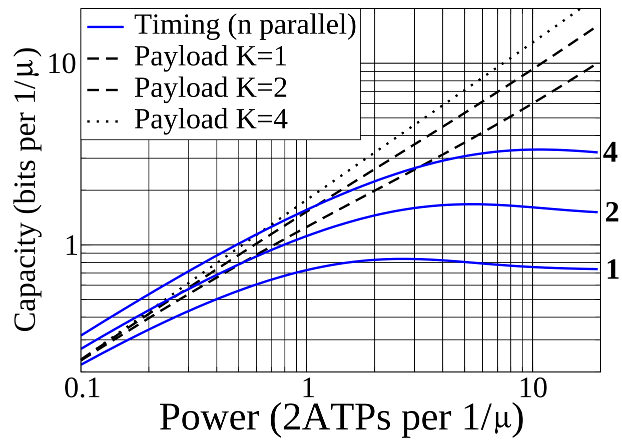

Fig. 12. Lower bounds for the capacities of the particle timing ($\mathcal{C}_T$), and particle timing plus particle payload ($\mathcal{C}_{P+T}$) channels as a function of power budget ($\mathcal{P}_T$ and $\mathcal{P}_{P+T}$) for DNA string particles with exponential first-passage times. Capacity is in units of bits per first-passage time $1/\mu$. Power is in units of 2-ATP (1.6×10^{-19} J) per first-passage time $1/\mu$ and a nucleotide residue is assumed to carry 2-bits of information. Solid lines: aggregate capacity of $n = 1, 2, 4$ separate (independent or parallel) particle timing channels where DNA string particles carry no information payload. Notice the maxima at powers $\approx 3, 5$ and 10 respectively, consonant with equation (14) and Fig. 9. Dashed/Dotted lines: aggregate capacity of particle timing plus particle payload channels for DNA string particles of different lengths, $K = 1, 2, 4$-residue payloads. Exponential first-passage assumed.

and the capacity for the ephemeral particle channel for very small symbol intervals is then, via equation (20) and equation (16)

$$\mathcal{C}_E \approx \frac{M}{e} = \frac{\mu\tau}{e}\frac{\rho}{(1-\alpha)} \approx \frac{\mathcal{P}_E}{10(c_0 + c_e)M} \tag{29}$$

if we set $\mu\tau = \frac{1}{10M}$, bettering the assumption of equation (19).

In Fig. 12 and Fig. 13 we plot $\mathcal{C}_T$, $\mathcal{C}_{P+T}$ and $\mathcal{C}_E$ in bits per first-passage time ($1/\mu$) as a function of power budget $\mathcal{P}$ in assuming DNA-based particles. For particle timing plus payload signaling we show plots for $K = 1, 2, 4$ DNA-residue particles. For timing-only signaling we also include plots where different identifiable particles (different molecule types or physically separate channels) are used (*i.e.*, $n = 1, 2, 4$ parallel timing channels as shown) for comparison with payload channels. We have assumed costs $c_0 = \Delta c_1 = 2$ ATP – which is why the power unit is 2 ATP per first-passage time. Furthermore, we assume $c_1 = c_e = \Delta c_1$ since it seems likely that the absolute minimum energy for particle release, c_e, in a purely diffusive channel is probably comparable to the cost of creating (or breaking) the covalent bond used to append a nucleotide residue. If we assume $1/\mu = 1$ms, then the ordinate of Fig. 12 and Fig. 13 are in kbit/s and the abscissa is in units of 1.6×10^{-16}W. If $1/\mu = 1\mu s$, (as might be the case for smaller gaps in a nano-system) the ordinate is in Mbit/s and the abscissa is in units of 1.6×10^{-13}W.

These data rates are many many orders of magnitude larger

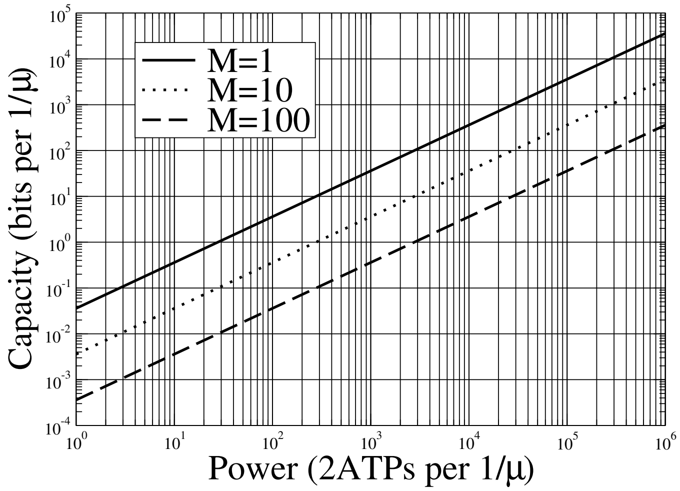

Fig. 13. Capacity of the ephemeral particle ($\mathcal{C}_E$) channel as a function of power budget ($\mathcal{P}_E$) for DNA string particles with exponential first-passage times. Capacity is in units of bits per first-passage time $1/\mu$. Power is in units of 2-ATP (1.6×10^{-19} J) per first-passage time $1/\mu$. Exponential first-passage assumed. Plots for $M = 1, 10, 100$ shown.

than the fractional bit/second data rates previously reported for simple demonstrations of molecule communication [12] or even the 10's of bits per second rates reported in [13]. And the predicted power efficiencies are startling. Comparison of our results to [12], [13] and others would be relatively straightforward if first-passage time jitter/entropy for the experimental setup were provided, although in both [12] and [13] very large numbers of molecules were released with each alcohol "puff" so precise timing at the molecular level was not attempted.

V. Discussion & Conclusion

We have reviewed a general and finest grain (micro) fundamental mathematical framework for molecular communication channels. We have emphasized that the bounds derived for such micro systems must be obeyed by any molecular communication system including those that use Avogadrian (macro) numbers of signaling particles. Along the way we have presented bounds and made comparisons between different systems. We now conclude with a slightly deeper dive into these comparisons as well, some gentle suggestions for future work.

Capacity Bounds and Coding Methods: The upper bound on capacity C_t, the timing capacity for persistent identical particles, is tight for low particle load ρ but diverges for large ρ. However, no empirical distributions that provide rates higher than the lower bound have yet been found [6], [7]. So, does the capacity of the timing-only channel truly flatten with increasing ρ as in Fig. 12 and Fig.9, or is there a benefit to increasing the intensity of timing particle release as suggested in [17]? Intriguingly, this rate flattening of the lower bound seems to comport almost exactly with a result in [21, section III] which considers a related system where only a finite number of particles can be simultaneously in flight (a finite "server" system in the parlance of [19], [21]). More careful exploration of these parallels may reveal launch densities that cause C_t to grow without bound with ρ even with persistent particles.

Since exponential first-passage is not the worst case corruption for such systems [7], we cannot be assured of minmax and maxmin performance bounds and so can only use what we have seen so far as a guide. Questions such as "what *is* the minmax capacity of the molecular timing channel?" and "How much better than exponential might other first-passage densities imposed by various physical channels be?" remain open.

However, we did unequivocally see that systems that use particles with finite lifetimes can have much higher data rates albeit at the expense of increased energy. This simple idea suggests one could even imagine channels with chemically reactive species in which emitted particles elicited spatially structured propagation of detectable reaction products [34]–[36]. But perhaps even more interesting, certain biological systems (neuromuscular synapes) that arguably require high and reliable data rates seem to employ finite particle lifetimes through gettering (secretion of cholinesterase to getter acetylcholine). One wonders what other biological phenomena can be identified (or predicted) based on the analytically derived characteristics of particle channels.

Precise Timing, Fuzzy Timing and Concentration: It is important to quantify the relationship between our fine grain timing model and other less temporally precise ones [4], [5], [12], [17], [37]–[43]. The timing model described here seems to imply infinitely precise control over the release times $\mathbf{T}$ and infinitely precise measurement of the arrival times $\vec{\mathbf{S}}$. However, imprecision in both times can be incorporated easily into the transit time vector $\mathbf{D}$. Thus, applying a precise timing model to the "fuzzier" release and detection times associated with practical/real systems is straightforward. That is, first-passage time jitter already imposes limits on timing precision. So long as timing precision is significantly better than first-passage time jitter, the bounds provided here (and developed in [6], [7]) will be moderately tight. In addition, we are hopeful that the upper bound of equation (15) shown in Fig. 9 will be useful for evaluating molecular timing channel capacity for arbitrary first-passage time distributions since it requires only knowledge of the timing channel capacity coupled to average properties of the corresponding input distribution.

As previously discussed and precisely stated in equation (3) and equation (4), concentration is derived from counting arrivals within temporal windows. The data processing theorem [10] indicates that our precise timing model **must** undergird all concentration based methods which, even with perfect concentration detection, *cannot possibly exceed the capacity of the fine grain timing model presented here*. Given the asymptotic nature of analysis, an individual emission schedule $\mathbf{t}$ for large M is *exactly* a temporal emission concentration profile as time resolution coarsens. Thus, micro results provide crisp upper bounds on the capacities derived from concentration-based models. Likewise, concentration-based results inform micro level particle systems. For instance, do the startlingly large kb/s results of [17] suggest that equation (15) is an achievable upper bound? Or does the failure to find input distributions that approach this bound suggest some discrepancy between the assumptions behind analysis in [17] versus those behind the lower bound of equation (14)? We suspect the answer lies in the particular first-passage density and that exponential first-passage may preclude growth of capacity with increasing particle intensity.

Identifiable Particles Without Payload: We considered the possibility of uniquely identifying each of M emitted particles with a sequence number of length $\log M$ bits. We treat this scenario as distinct from ensemble timing channel coding which resolves residual ordering ambiguity because if the particles are individually identifiable, the potential emission schedules are not constrained to ensemble timing channel coding. Thus, the M identifiable particles constitute M parallel single-particle timing channels, which for exponential first-passage have aggregate capacity $M \log(1 + \frac{M}{\rho e})$ [7].

However, because each particle requires $\log M$ bits of sequencing information, ρ is limited by the power budget $\mathcal{P}$ (in units of energy per first-passage, $1/\mu$)

$$\rho \log M \leq \mathcal{P} \tag{30}$$

Following Fig. 8 we have $\lambda\tau(M) = M$ so the capacity in nats per first-passage time is

$$C = \rho \log\left(1 + \frac{M}{\rho e}\right) \leq \rho \log\left(1 + \frac{e^{\frac{\mathcal{P}}{\rho}}}{\rho e}\right) \tag{31}$$

with the inequality owed to equation (30). However, in the limit of $M \to \infty$ we have $\rho \to 0$ so we have

$$\lim_{\rho \to 0} C = \mathcal{P} \tag{32}$$

in units of nats per first-passage time (and assuming unit per-bit cost of the particle identfier string). Thus, the identifiable particle timing channel capacity exceeds the identical particle timing channel lower bound with increasing power budget and, more importantly, scales linearly in power.

Particle Corruption and Receptor Noise: The potential for lost or corrupted particles and potential binding noise at receptor sites must eventually be considered. However, as previously stated, particle erasure (particles that do not arrive) or payload particle corruption (particles that are altered in passage) or receptor noise (particles bind stochastically to the receptor) cannot increase capacity (data processing theorem). Thus, the results here provide upper bounds. Nonetheless it is worth considering how the analytic machinery developed previously might be modified to accommodate such impediments.

First, consider alteration of payload-carrying particles *en route*. If the corruption is i.i.d. for each particle, then error correcting codes can be applied individually, or to the particle ensemble. The resulting overall channel capacity will be degraded by the coding overhead necessary to preserve payload message integrity (including the sequencing information).

Then consider particle erasure where a particle never arrives (and is assumed to not arrive in a later signaling interval). Since each signaling interval uses M particles, we will know whether particles get "lost" in transit and can arbitrarily assign a faux arrival time to such particles. However, the problem this poses for the analysis is two-fold. First, particles released later in the signaling interval are more likely to be lost which implies that the first-passage density is not identical for each particle. Second, the first-passage density for each particle would then contain a singularity equal to the probability of loss, which violates a key assumption (hypersymmetry [6]) that drives the analysis. That said, an erasure channel approach where arriving particles were deleted randomly could be pursued, providing a worst case scenario since the information associated with erasures being more likely for later emissions would be absent. However, the capacity for such systems would necessarily be *lower* – erasure does not confer the advantage of deliberately ephemeral particles.

Interference and Multiple Users: Multi-user communication in a molecular setting is a critical question, and a better understanding of the single-user channel will certainly help with multi-user studies where transmissions interfere. There is some prior work that may provide an information-theoretic foundation [44] similar to how the work described here builds on [19], but the multi-user molecular signaling problem has not yet been rigorously considered. Of particular interest would be a version of MIMO since Fig. 12 shows capacity benefits from parallel channels. One could imagine apposed arrays of emitters and receivers which could be engineered to collaborate to encode and decode information in a variety of ways owing to the differing particle transport channel properties between spatially distinct emitter/receiver pairs.

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 useful discussions over many years.

REFERENCES

[1] Stephen Hawking. *A Brief History of Time*. Bantam Books, 1988. "*A well-known scientist (some say it was Bertrand Russell) once gave a public lecture on astronomy. He described how the earth orbits around the sun and how the sun, in turn, orbits around the center of a vast collection of stars called our galaxy. At the end of the lecture, a little old lady at the back of the room got up and said: 'What you have told us is rubbish. The world is really a flat plate supported on the back of a giant tortoise.' The scientist gave a superior smile before replying, 'What is the tortoise standing on?' 'You're very clever, young man, very clever,' said the old lady. 'But it's turtles all the way down!'*".

[2] T. Nakano, T. Suda, M. Moore, R. Egashira, A. Enomoto, and K. Arima. Molecular Communication for Nanomachines Using Intercellular Calcium Signaling. In *5th IEEE Conference on Nanotechnology, 2005.*, pages 478–481 vol. 2, July 2005.

[3] Ian F. Akyildiz, Fernando Brunetti, and Cristina Blázquez. Nanonetworks: A new communication paradigm. *Comput. Netw.*, 52(12):2260–2279, August 2008.

[4] I.F. Akyildiz, J.M. Jornet, and M. Pierobon. Nanonetworks: A New Frontier in Communications. *Communications of the ACM*, 54(11):84–89, 2011.

[5] A.W. Eckford. Nanoscale communication with Brownian motion. In *CISS'07*, pages 160–165, 2007. Baltimore.

[6] C. Rose and I. S. Mian. Inscribed Matter Communication: Part I. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2(2):209–227, Dec 2016.

[7] C. Rose and I. S. Mian. Inscribed Matter Communication: Part II. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2(2):228–239, Dec 2016.

[8] 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*, 18(3):1887–1919, 2016. 3rd Quarter, 2016.

[9] T. Nakano, M. Moore, F. Wei, A. V. Vasilakos, and J. Shuai. Molecular communication and networking: Opportunities and challenges. *IEEE Trans. Nanobiosci.*, 11(2):135 – 148, June 2012.

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

[11] N. Farsad, W. Guo, and A.W. Eckford. Tabletop Molecular Communication: Text Messages Through Chemical Signals. *PLOS ONE*, 8(12), 2013.

[12] It's the Alcohol Talking. *The Economist*, March 2014. Technology Quarterly: monitor.

[13] M. Ozmen, E. Kennedy, J. Rose, P. Shakya, J.K. Rosenstein, and C. Rose. High Speed Chemical Vapor Communication Using Photoionization Detectors. In *Proc. IEEE Globecom 2018*, December 2018. Abu Dhabi.

[14] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, third edition, 1991.

[15] 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)*, pages 2468–2472, June 2017.

[16] T. Furubayashi, T. Nakano, A. W. Eckford, Y. Okaie, and T. Yomo. Packet Fragmentation and Reassembly in Molecular Communication. *IEEE Trans. Nanobiosci.*, 15(3):284 – 288, Apr. 2016.

[17] M. Pierobon and I. Akyildiz. Capacity of a Diffusion-based Molecular Communication System with Channel Memory and Molecular Noise. *IEEE Transactions on Information Theory*, 59(2):942–954, 2013.

[18] 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*, 58(7):4678–4692, July 2012.

[19] V. Anantharam and S. Verdu. Bits Through Queues. *IEEE Transactions on Information Theory*, 42(1):4–18, January 1996.

[20] R. Sundaresan and S. Verdú. Robust Decoding for Timing Channels. *IEEE Transactions on Information Theory*, 46(2), 2000.

[21] R. Sundaresan and S. Verdú. Capacity of Queues Via Point-Process Channels. *IEEE Transactions on Information Theory*, 52(6), 2006.

[22] N. Farsad, Y. Murin, A. W. Eckford, and A. Goldsmith. On the Capacity of Diffusion-Based Molecular Timing Channels. In *IEEE International Symposium on Information Theory 2016*, pages 1023–1027, July 2016.

[23] V. Chandar, A. Tchamkerten, and D. Tse. Asynchronous capacity per unit cost. *IEEE Transactions on Information Theory*, 59(3):1213–1226, 2013.

[24] R. Gallager. Basic limits on protocol information in data communication networks. *IEEE Transactions on Information Theory*, 22(4):385–398, 1976.

[25] R.L. Dobrushin. Shannon's theorems for channels with synchronization errors. *Problemy Peredachi Informatsii* , 3(4):18–36, 1967.

[26] C. Rose and I.S. Mian. A General Upper Bound on Point-to-Point Particle Timing Channel Capacity Under Constant Particle Emission Intensity. In *IEEE International Symposium on Information Theory (ISIT'19)*, July 2019. (to appear).

[27] A. Noel, K. C. Cheung, and R. Schober. Improving receiver performance of diffusive molecular communication with enzymes. *IEEE Trans. Nanobiosci.*, 13(1):31–43, Mar. 2014.

[28] N. Farsad and A. Goldsmith. A molecular communication system using acids bases and hydrogen ions. In *Proc. IEEE SPAWC*, pages 1–6, Jul. 2016.

[29] V. Jamali, A. Ahmadzadeh, N. Farsad, and R. Schober. Constant-Composition Codes for Maximum Likelihood Detection Without CSI in Diffusive Molecular Communications. *IEEE Trans. Comm.*, 66(5):1981 – 1995, May 2018.

[30] John W. Byers, Michael Luby, Michael Mitzenmacher, and Ashutosh Rege. A digital fountain approach to reliable distribution of bulk data. *SIGCOMM Comput. Commun. Rev.*, 28(4):56–67, October 1998.

[31] C. Rose and G. Wright. Inscribed Matter As An Energy-Efficient Means Of Communication With An Extraterrestrial Civilization. *Nature*, 431:47–49, 2004.

[32] D.L. Nelson and M.M. Cox. *Lehninger Principles of Biochemistry*. Freeman, 2005. 4th Ed.

[33] M.S. Kurana, H.B. Yilmaz, T. Tugcu, and B. Ozerman. Energy model for communication via diffusion in nanonetworks. *Nano Communication Networks*, 1:86–95, 2010.

[34] Kapral, R. and Showalter, K. *Chemical Waves and Patterns*. Springer, 1995.

[35] B.H. Gilding and R. Kersner. *Travelling Waves in Nonlinear Diffusion Convection Reaction*. Birkhauser, 2004.

[36] P.C Fife. *Mathematical Aspects of Reacting and Diffusing Systems*. Springer, 1979.

[37] B. L. Bassler. How bacteria talk to each other: regulation of gene expression by quorum sensing. *Curr Opin Microbiol*, 2(6):582–587, Dec 1999.

[38] B.L. Bassler. Small talk. Cell-to-cell communication in bacteria. *Cell*, 109(4):421–424, May 2002.

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

[40] T. Nakano, A. W. Eckford, and T. Haraguchi. *Molecular Communication*. Cambridge University Press, 2013.

[41] A. Einolghozati, M. Sardari, and F. Fekri. Relaying in diffusion-based molecular communication. In *IEEE International Symposium on Information Theory 2013*, pages 2959–2963, 2013.

[42] 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*, 12(12):6096–6105, 2013.

[43] A. Einolghozati, M. Sardari, and F. Fekri. Decode and Forward Relaying in Diffusion-based Molecular Communication Between Two Populations of Biological Agents. In *Information Conference on Communications, 2014 IEEE ICC*, pages 3975–3980, 2014.

[44] G. C. Ferrante, T. Q. S. Quek, and M. Z. Win. An achievable rate region for superposed timing channels. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 365–369, July 2016.