

Capacity of First Arrival Position Channel in Diffusion-Based Molecular Communication

Yen-Chi Lee

Hon Hai (Foxconn) Research Institute

Taipei, Taiwan

yen-chi.lee@foxconn.com

Min-Hsiu Hsieh

Hon Hai (Foxconn) Research Institute

Taipei, Taiwan

min-hsiu.hsieh@foxconn.com

Abstract—In [1], the impulse response of the first arrival position (FAP) channel of 2D and 3D spaces in molecular communication (MC) is derived, but its Shannon capacity remains open. The main difficulty of depicting the FAP channel capacity comes from the fact that the FAP density becomes a multi-dimensional Cauchy distribution when the drift velocity approaches zero. As a result, the commonly used techniques in maximizing the mutual information no longer work because the first and second moments of Cauchy distributions do not exist.

Our main contribution in this paper is a complete characterization of the zero-drift FAP channel capacity for the 2D and 3D spaces. The capacity formula for FAP channel turns out to have a similar form compared to the Gaussian channel case (under second-moment power constraint). It is also worth mentioning that the capacity value of 3D FAP channel is twice as large as 2D FAP channel. This is an evidence that the FAP channel has larger capacity as the spatial dimension grows. Finally, our technical contributions are the application of a modified logarithmic constraint as a replacement of the usual power constraint, and the choice of output signal constraint as a substitution to input signal constraint in order to keep the resulting formula concise.

Index Terms—Molecular communication (MC), diffusion, Brownian motion, first arrival position (FAP), channel capacity, Cauchy distribution, Lorentz distribution, alpha-stable distribution, logarithmic constraint.

I. INTRODUCTION

Molecular communication (MC) is a communication paradigm based on the exchange of molecules [2], [3]. Due to the nano-scale feasibility and bio-compatibility, MC is a promising communication approach for nano-networks [4], [5]. In MC systems, tiny message molecules (MM) operate as information carriers. A propagation mechanism is necessary for transporting MMs to the receiver (Rx), and this mechanism can be diffusion-based [6], flow-based [7], or an engineered transport system like molecular motors [8], [9]. Among these different propagation mechanisms, diffusion-based MC, sometimes in combination with a drift field, has been the most prevalent approach for both MC theoretical research and practical implementation.

The reception mechanism of a MC receiver can be categorized, see [10], into two classes: i) passive reception, and ii) active reception. We consider a common type of active reception called the fully-absorbing Rx [11], and assume that the Rx has the ability to measure the time [12] and the position [13] at which the MM first reaches the Rx. In order to

explore the Shannon capacity [14] of this new type of position channels, we assume the simplest geometric structure of the receiver, namely an infinite large¹ receiving plane [15]. The MMs will be removed when they first arrive at the receiving plane [11], [13].

The first paper in MC society promoting the first arrival position (FAP) as an information carrying property is [16]. Although most works in MC consider the first arrival time (FAT)² for absorbing receivers, there are at least two reasons why FAP is preferable.

- For each independent channel use (i.e. for each transmission of a single MM), the FAT information is only one-dimensional (1D), while the FAP-type modulation could have higher dimensions (say $n - 1$) to carry information when considering n -dimensional spaces. Hence, the capacity of FAP channel could be larger than FAT channel per single channel use in high dimensional cases (see [1]).
- The second reason is about the time efficiency. When considering multiple-MM transmission, the MMs may arrive out of order due to the randomness of the diffusion phenomenon, causing the cross-over effects, see [10], [18]. In order to prevent the MC system from cross-over effects, the transmission time of each symbol cannot be too short.³ Consequently, for applications in which the time efficiency plays an important role, the FAP-type modulation is arguably a better solution.

Before exploring the general capacity that allows transmission of multiple MMs, a clever way is to have a deeper understanding of single-MM transmission first. In the remaining of this paper, we will mainly focus on one-shot transmission [19] (i.e. transmission using a single MM). In order to explore the channel capacity of FAP or FAT channels, the first step is to provide quantitative descriptions about the channel impulse response. The one-shot FAT channel can be described as a time-invariant additive channel [12]:

$$t_{\text{out}} = t_{\text{in}} + t_n, \quad (1)$$

¹This geometrical assumption can also be regarded as an approximate model provided that the transmission distance is short compared to the receiver size.

²One can refer to [12], [17] for more details concerning FAT-type modulation and its channel characteristics.

³This can be understood by roughly thinking that there should be some “guard interval” between two consecutive timing symbols.

where t_{in} is the releasing time, t_{out} is the arriving time, and t_n is the random time delay due to the propagation mechanisms. The authors of [12] showed that t_n follows the inverse Gaussian distribution, so channel (1) is known as additive inverse Gaussian noise (AIGN) channel in MC. Later in [12], [17], some bounds on capacity of AIGN channel is derived, and the capacity-achieving input time distribution was also characterized.

As for FAP channels in n -dimensional spaces, the one-shot channel model can be also written in an additive vector form [1]:

$$\mathbf{x}_{\text{out}} = \mathbf{x}_{\text{in}} + \mathbf{x}_n, \quad (2)$$

where \mathbf{x}_{in} is the releasing position, \mathbf{x}_{out} is the arriving position, and \mathbf{x}_n is the random position bias due to the propagation mechanisms. Note that \mathbf{x}_{in} , \mathbf{x}_{out} and \mathbf{x}_n are all Euclidean vectors in \mathbb{R}^{n-1} . Although the density function of \mathbf{x}_n is obtained in [1], [13], [16], the channel capacity of such kind of position channel remains open at the time of this writing. In this paper, we shall provide the zero-drift FAP channel capacity formulas in 2D and 3D spaces under a newly proposed logarithmic constraint.

The remainder of this paper is structured as follows. In Section II, we briefly review the concept of mutual information and Shannon channel capacity description. In Section III, the channel model under consideration is depicted and the α -power constraint is introduced. The main results of this work are presented in Section IV. Finally, we conclude in Section V.

II. PRELIMINARY

We briefly review some background knowledge for mutual information and channel capacity. Let X be a random variable with a probability density function f whose support is denoted by \mathcal{X} . The (differential) entropy $h(X)$ is defined as:

$$h(X) := \mathbb{E}_X[-\ln f(X)] = - \int_{\mathcal{X}} [\ln f(x)] f(x) dx. \quad (3)$$

Based on the definition of entropy, the mutual information between the input and output of the additive channel

$$Y = X + N \quad (4)$$

is given by

$$I(X; Y) = h(Y) - h(Y|X) \quad (5)$$

$$= h(Y) - h(X + N|X) \quad (6)$$

$$= h(Y) - h(X|X) - h(N|X) \quad (7)$$

$$= h(Y) - h(N). \quad (8)$$

Here we have assumed that the noise N is independent of signal X . Recall that the (Shannon) channel capacity is defined as:

$$C = \sup_{f_X(x)} I(X; Y). \quad (9)$$

The supremum is taken over all possible distributions of X satisfying some chosen constraint. For point-to-point communication or single-input-single-output (SISO) scenarios, sometimes

the output constraints are used instead of the input constraints, yielding

$$C = \sup_{f_Y(x)} I(X; Y). \quad (10)$$

The selection of constraint depends heavily on the noise distribution. To get some feeling about this, let us study a simple example before exploring the FAP channels. For additive Gaussian channel (i.e. the noise is Gaussian distributed), the most familiar power measure is the second moment constraint. We can prescribe our signal set to satisfy:

$$\mathbb{E}[X^2] = \int_{\mathcal{X}} x^2 f(x) dx \leq P. \quad (11)$$

The constant $P \geq 0$ is the maximum power we allowed for all input signals. Using this power constraint (11), the channel capacity (9) can be written as:

$$C = \sup_{f_X(x): \mathbb{E}[X^2] \leq P} I(X; Y). \quad (12)$$

Following a standard textbook derivation (see [20]), one can see that the supremum value of $I(X; Y)$ is

$$\frac{1}{2} \ln \left(\frac{\sigma^2 + P}{\sigma^2} \right), \quad (13)$$

and the capacity attaining distribution is normal distributed with variance P , namely, $X \sim \mathcal{N}(0, P)$.

For later comparison to the capacity formula of zero-drift FAP channels, we let $A^2 := \sigma^2 + P$ and rewrite Eq. (13) in the following way:

$$C_G = C(A, \sigma) = \frac{1}{2} \ln \left(\frac{A^2}{\sigma^2} \right) = \ln \left(\frac{A}{\sigma} \right), \quad (14)$$

where the subscript G stands for Gaussian.

III. CHANNEL MODEL

We consider an additive vector channel:

$$\mathbf{Y} = \mathbf{X} + \mathbf{N} \quad (15)$$

where the noise \mathbf{N} is (multivariate) Cauchy distributed.⁴ Because the Cauchy distributions are heavy-tailed, traditional signal processing theory, which is tailored for finite second moment signals, does not apply directly to our channel model (15).

Perhaps the most important type of heavy-tailed probability models is the alpha-stable family (see [21]) of distributions, which have been recently found to accurately simulate the multiple access interference [22], and also the co-channel interference originated from a field of Poisson distributed interferers [23]. For alpha-stable channels, the second moment is no longer considered to be a suitable power measure because all related quantities become infinite.

The problem of finding the channel capacity for alpha-stable additive noise (with $\alpha \geq 1$) with an r -th absolute moment input constraint was solved by [24] for the case of symmetric

⁴For a detailed characterization of Cauchy density functions, one can refer to Appendix A.

alpha-stable noise, and the optimal capacity-achieving input was found to be discrete and compactly supported. An alternative notion of power, called the geometric power, was introduced by [25] for heavy-tailed distributions. In this paper, we consider another power characterization framework for α -stable family called the α -power (see [26]). It is a relative power measure $P_\alpha(\mathbf{X})$ satisfying the following:

- P1) $P_\alpha(\mathbf{X}) \geq 0$, with equality if and only if $\mathbf{X} = 0$ almost surely.
P2) $P_\alpha(k\mathbf{X}) = |k|P_\alpha(\mathbf{X})$, for any $k \in \mathbb{R}$.

For α -stable distributions, it is known (see [26]) that

$$\begin{cases} \mathbb{E}_{\mathbf{X}} [\ln(1 + \|\mathbf{X}\|^2)] < \infty, & \alpha < 2; \\ \mathbb{E}_{\mathbf{X}} [\|\mathbf{X}\|^2] < \infty, & \alpha = 2. \end{cases} \quad (16)$$

We suggest to adopt a power measure $P_1(\cdot)$ for multivariate Cauchy-like distributions as follows:

$$\mathbb{E}_{\mathbf{X}} \left[\ln \left(1 + \left\| \frac{\mathbf{X}}{P_1(\mathbf{X})} \right\|^2 \right) \right] = w_2 \left(\frac{1+p}{2}; \frac{p}{2} \right), \quad (17)$$

where the constant evaluation function $w_2(t; \alpha)$ can be written explicitly as the difference of two digamma functions:

$$w_2(t; a) = \psi(t) - \psi(t - a), \quad \text{for } t > a. \quad (18)$$

The values of w_2 function when $p = 1$ and $p = 2$ are $2 \ln(2)$ and $2 \ln(e)$ respectively.

IV. MAIN RESULTS

We consider the FAP communication channel encountered in diffusion-based MC systems. In this main section, we first show in Section IV-A that when the drift velocity in the fluid medium approaches zero, both the 2D and 3D FAP channel reduce to univariate and bivariate Cauchy distributions respectively. This result is new to the MC society.

Next in Section IV-B and Section IV-C, we introduce a relative power measure called the α -power (see [26]) into the MC field. We write down explicitly the signal space needed to characterize the no-drift FAP channels. Under this newly proposed logarithmic constraint on the output signal space, we derive the Shannon capacity for FAP channel both in 2D and 3D spaces in closed-form. The conclusions about the FAP channel capacity are summarized in Theorem 1 and Theorem 2. Notice that in [27], the authors derived the capacity formula for univariate Cauchy channel under logarithmic constraint imposed on the input signal space. However, their result is not easy to extend to high dimensional cases because a complicated integration process was involved, see [27, Appendix I].

A. FAP Channel Reduces to Cauchy Channel Under Zero Drift Condition

1) *In 2D Space:* From [1] we know that the FAP density function in 2D space, with a drift velocity $\mathbf{v} = (v_1, v_2)$, can be written as:

$$f_{Y|X}(\mathbf{y}|\mathbf{x}) = \frac{|\mathbf{v}|\lambda}{\sigma^2\pi} \exp\left\{\frac{-v_2\lambda}{\sigma^2}\right\} \exp\left\{\frac{-v_1(x_1 - y_1)}{\sigma^2}\right\} \cdot \frac{K_1\left(\frac{|\mathbf{v}|}{\sigma^2}\sqrt{(x_1 - y_1)^2 + \lambda^2}\right)}{\sqrt{(x_1 - y_1)^2 + \lambda^2}}, \quad (19)$$

where $\mathbf{x} = (x_1, \lambda)$, $\mathbf{y} = (y_1, 0)$, and λ is the *transmission distance* between transmitter (Tx) and Rx. Note that in Eq. (19), σ^2 is the microscopic diffusion coefficient which can be related to the macroscopic diffusion coefficient D through the relation $\sigma^2 = 2D$, see [5]. Note also that the special function $K_1(\cdot)$ in Eq. (19) is the modified Bessel function of the second kind (see [28]) with order $\nu = 1$.

Next we show that when drift velocity approaches zero, the 2D FAP density reduces to a univariate Cauchy distribution. We use a limit property about the special function $K_1(\cdot)$ from [29]:

$$\lim_{x \rightarrow 0} xK_1(x) = 1. \quad (20)$$

With this limit property, we can do the following calculation:

$$\begin{aligned} & f_{Y|X; \mathbf{v}=\mathbf{0}}(\mathbf{y}|\mathbf{x}) \\ &= \lim_{|\mathbf{v}| \rightarrow \mathbf{0}} \frac{|\mathbf{v}|\lambda}{\sigma^2\pi} \cdot \frac{K_1\left(\frac{|\mathbf{v}|}{\sigma^2}\sqrt{(x_1 - y_1)^2 + \lambda^2}\right)}{\sqrt{(x_1 - y_1)^2 + \lambda^2}} \\ &= \frac{\lambda}{\pi} \lim_{|\mathbf{v}| \rightarrow \mathbf{0}} \left[\frac{|\mathbf{v}|}{\sigma^2} \sqrt{(x_1 - y_1)^2 + \lambda^2} \right. \\ & \quad \left. \cdot \frac{K_1\left(\frac{|\mathbf{v}|}{\sigma^2}\sqrt{(x_1 - y_1)^2 + \lambda^2}\right)}{(x_1 - y_1)^2 + \lambda^2} \right] \\ &= \frac{\lambda}{\pi} \frac{1}{(x_1 - y_1)^2 + \lambda^2}. \end{aligned} \quad (21)$$

Notice that in the last equality of Eq. (21), we have used the limit property Eq. (20). This calculation result can be interpreted as follows: let x_1 be the input position and y_1 be the output position. The displacement of position is an additive noise n , following the density function:

$$\frac{\lambda}{\pi} \frac{1}{n^2 + \lambda^2}, \quad (22)$$

which is Cauchy distributed. (See Appendix A.)

2) *In 3D Space*: From [1], [16] we know that the FAP density function in 3D diffusion channel, with a drift velocity $\mathbf{v} = (v_1, v_2, v_3)$, can be written as:

$$f_{Y|X}(\mathbf{y}|\mathbf{x}) = \frac{\lambda}{2\pi} \exp\left\{-\frac{v_3\lambda}{\sigma^2}\right\} \exp\left\{\frac{v_1}{\sigma^2}(y_1 - x_1) + \frac{v_2}{\sigma^2}(y_2 - x_2)\right\} \cdot \exp\left\{-\frac{|\mathbf{v}|}{\sigma^2}\|\mathbf{y} - \mathbf{x}\|\right\} \left[\frac{1 + \frac{|\mathbf{v}|}{\sigma^2}\|\mathbf{y} - \mathbf{x}\|}{\|\mathbf{y} - \mathbf{x}\|^3}\right], \quad (23)$$

where σ^2 is the microscopic diffusion coefficient, the scale parameter λ is the *transmission distance* between Tx and Rx plane, v_3 is the longitudinal component of drift velocity in the direction parallel to the transmission direction, and v_1, v_2 are the components perpendicular (or transverse) to the transmission direction. Here we use the symbol $\|\cdot\|$ to represent the Euclidean norm. Namely, $\|\mathbf{y} - \mathbf{x}\| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \lambda^2}$, where $\mathbf{x} = (x_1, x_2, \lambda)$, $\mathbf{y} = (y_1, y_2, 0)$ are position vectors in \mathbb{R}^3 .

Next we show that when drift velocity approaches zero, the 3D FAP density reduces to a bivariate Cauchy distribution. Note that when $\mathbf{v} = \mathbf{0}$, all the exponential terms in Eq. (23) become $e^0 = 1$, so that we have:

$$\begin{aligned} f_{Y|X; \mathbf{v}=\mathbf{0}}(\mathbf{y}|\mathbf{x}) &= \frac{\lambda}{2\pi} \lim_{|\mathbf{v}| \rightarrow \mathbf{0}} \frac{1 + \frac{|\mathbf{v}|}{\sigma^2}\|\mathbf{y} - \mathbf{x}\|}{\|\mathbf{y} - \mathbf{x}\|^3} \\ &= \frac{\lambda}{2\pi} \frac{1}{\|\mathbf{y} - \mathbf{x}\|^3} \\ &= \frac{\lambda}{2\pi} \frac{1}{((y_1 - x_1)^2 + (y_2 - x_2)^2 + \lambda^2)^{3/2}}. \end{aligned} \quad (24)$$

That is, if we regard the position channel as:

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \quad (25)$$

then \mathbf{n} follows a bivariate Cauchy distribution, see Eq. (57) in Appendix A.

B. Channel Capacity for 2D FAP Channel Under Logarithmic Constraint

As mentioned in Section IV-A, a 2D FAP channel reduces to a Cauchy channel when the drift velocity approaches zero. Since we are considering a point-to-point (or SISO) communication scenario, the input constraint is actually equivalent to the output constraint. We choose to adopt output constraint for convenience.

It is known that a Cauchy distribution $X \sim \text{Cauchy}(0, k)$ maximizes the entropy among all RVs⁵ that satisfy

$$\mathbb{E}_X \ln \left[1 + \left(\frac{X}{k}\right)^2 \right] = 2 \ln(2), \quad (26)$$

or equivalently,

$$\mathbb{E}_X \ln [1 + X^2] = 2 \ln(1 + k). \quad (27)$$

⁵This is the *principle of maximum entropy* for Cauchy distribution, see [30].

The corresponding maximum entropy value is $\ln(4\pi k)$, see Appendix A. Note that the parameter k restricts the dispersion of all random variables X under consideration.

Consider a 2D FAP zero-drift channel (as shown in Section IV-A), the input-output relation can be written as

$$Y = X + N, \quad (28)$$

where $N \sim \text{Cauchy}(0, \lambda)$. For continuous-variable channel capacity problem, we need to specify a family of distributions that are under consideration in order to prevent the capacity value from being infinite. Instead of writing down explicitly the constraint equality for X or Y , we define

$$\mathcal{D}(A) = \left\{ \text{distributions } Y \mid \exists k \in [\lambda, A] \text{ such that } \mathbb{E}_Y \ln \left[1 + \left(\frac{Y}{k}\right)^2 \right] = 2 \ln(2) \right\}. \quad (29)$$

The parameter A appeared in Eq. (29) indicates the ‘‘largest allowed dispersion’’ for the distributions in $\mathcal{D}(A)$. For later use, we also define

$$\mathcal{D}_k = \left\{ \text{distributions } Y \mid \mathbb{E}_Y \ln \left[1 + \left(\frac{Y}{k}\right)^2 \right] = 2 \ln(2) \right\}, \quad (30)$$

so that we can write $\mathcal{D}(A) = \bigcup_{k \in [\lambda, A]} \mathcal{D}_k$. Now we can state the main theorem.

Theorem 1. *The capacity of 2D FAP channel (28) is*

$$C_{2\text{D}, \text{FAP}} = C(A, \lambda) = \ln \left(\frac{A}{\lambda} \right) \quad (31)$$

under the output logarithmic constraint

$$Y \in \mathcal{D}(A) \quad (32)$$

for some prescribed dispersion level A , where $A \geq \lambda$. In addition, the corresponding capacity achieving output distribution is

$$Y^* \sim \text{Cauchy}(0, A), \quad (33)$$

or equivalently,

$$X^* \sim \text{Cauchy}(0, A - \lambda). \quad (34)$$

Because the proof of Theorem 1 is very similar to the proof of Theorem 2. We settle this proof into Appendix B and move directly to the 3D case.

C. Channel Capacity for 3D FAP Channel Under Logarithmic Constraint

It is known (see [30], [31]) that a bivariate Cauchy distribution $\mathbf{X} \sim \text{Cauchy}_2(0, \Sigma = \text{diag}(k^2, k^2))$ maximizes the entropy among all bivariate RVs that satisfy

$$\mathbb{E}_{\mathbf{X}} \ln \left[1 + \left\| \frac{\mathbf{X}}{k} \right\|^2 \right] = 2 \ln(e), \quad (35)$$

The corresponding maximum entropy value is $\ln(2\pi e^3 k^2)$, see Appendix A. Note that the parameter k restricts the dispersion of all random vectors \mathbf{X} under consideration.

Consider a 3D FAP zero-drift channel (as shown in Section IV-A), the input-output relation can be written as

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}, \quad (36)$$

where $\mathbf{N} \sim \text{Cauchy}_2(\mathbf{0}, \Sigma = \text{diag}(\lambda^2, \lambda^2))$. We define

$$\mathcal{D}_{\text{bi}}(A) = \left\{ \text{distributions } \mathbf{Y} \mid \exists k \in [\lambda, A] \right. \\ \left. \text{such that } \mathbb{E}_{\mathbf{Y}} \ln \left[1 + \left\| \frac{\mathbf{Y}}{k} \right\|^2 \right] = 2 \ln(e) \right\}. \quad (37)$$

The parameter A appeared in Eq. (37) indicates the ‘‘largest allowed dispersion’’ for the distributions in $\mathcal{D}_{\text{bi}}(A)$. For later use, we also define

$$\mathcal{D}_{k,\text{bi}} = \left\{ \text{distributions } \mathbf{Y} \mid \mathbb{E}_{\mathbf{Y}} \ln \left[1 + \left\| \frac{\mathbf{Y}}{k} \right\|^2 \right] = 2 \ln(e) \right\}, \quad (38)$$

so that we can write

$$\mathcal{D}_{\text{bi}}(A) = \bigcup_{k \in [\lambda, A]} \mathcal{D}_{k,\text{bi}}. \quad (39)$$

Now we can state the main theorem.

Theorem 2. *The capacity of 3D FAP channel (36) is*

$$C_{\text{3D,FAP}} = C(A, \lambda) = 2 \ln \left(\frac{A}{\lambda} \right) \quad (40)$$

under the output logarithmic constraint

$$\mathbf{Y} \in \mathcal{D}_{\text{bi}}(A) \quad (41)$$

for some prescribed dispersion level A , where $A \geq \lambda$. In addition, the corresponding capacity achieving output distribution is

$$\mathbf{Y}^* \sim \text{Cauchy}_2(0, \text{diag}(A^2, A^2)). \quad (42)$$

Proof. For the additive channel we are considering, the mutual information can be written as

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{N}). \quad (43)$$

The issue of finding the channel capacity turns out to be an optimization problem:

$$C(A, \lambda) = \sup_{\mathbf{Y} \in \mathcal{D}_{\text{bi}}(A)} I(\mathbf{X}; \mathbf{Y}). \quad (44)$$

The calculation is as follows. We have

$$C(A, \lambda) = \sup_{\mathbf{Y} \in \mathcal{D}_{\text{bi}}(A)} \left\{ h(\mathbf{Y}) - h(\mathbf{N}) \right\} \quad (45)$$

$$= \left\{ \sup_{\mathbf{Y} \in \mathcal{D}_{\text{bi}}(A)} h(\mathbf{Y}) \right\} - \ln(2\pi e^3 \lambda^2), \quad (46)$$

where

$$\sup_{\mathbf{Y} \in \mathcal{D}_{\text{bi}}(A)} h(\mathbf{Y}) = \sup_{k \in [\lambda, A]} \left\{ \sup_{\mathbf{Y} \in \mathcal{D}_{k,\text{bi}}} h(\mathbf{Y}) \right\} \quad (47)$$

$$= \sup_{k \in [\lambda, A]} \ln(2\pi e^3 k^2) \quad (48)$$

$$= \ln(2\pi e^3 A^2). \quad (49)$$

Combining equations (46) and (49) yields

$$C(A, \lambda) = \ln(2\pi e^3 A^2) - \ln(2\pi e^3 \lambda^2) \\ = \ln \left(\frac{A^2}{\lambda^2} \right) = 2 \ln \left(\frac{A}{\lambda} \right). \quad (50)$$

Notice that the equalities (47)-(49) hold if and only if \mathbf{Y} distributes as $\text{Cauchy}_2(0, \text{diag}(A^2, A^2))$. Hence, Theorem 2 is proved. \square

V. CONCLUSIONS

We consider the FAP communication channel encountered in diffusion-based MC systems. In this paper, we first show that when the drift velocity in the fluid medium approaches zero, both the 2D and 3D FAP channel reduces to univariate and bivariate Cauchy distribution respectively. This conclusion is new to the MC society.

Although in [1], the impulse response of the FAP channel is discussed in detail for 2D and 3D spaces, the Shannon capacity for FAP channel remains open at the time of this writing. The main difficulty of the capacity characterization problem of Cauchy-like channel is that, unlike commonly encountered distributions such as Gaussian or exponential, the Cauchy distribution is heavy-tailed and belongs to a broader family called α -stable distributions.

For α -stable distributions, usually the second moment does not exist, so we cannot use traditional energy constraints (such as the variance of RV) to depict our signal space. To tackle with this problem, we introduce a relative power measure called the α -power (see [26]) into the MC field. We write down explicitly the signal space needed to characterize the no-drift FAP channel capacity. Under this newly proposed logarithmic constraint, we derive the Shannon capacity for FAP channel both in 2D and 3D spaces in closed-form. Based on the capacity formulas derived in Theorem 1 and Theorem 2, we can see that: under the same value A , the channel capacity of 3D FAP channel is twice as large as the capacity of 2D FAP channel. In some sense, this demonstrates the spirit that FAP channel can carry more information when the spatial dimension n of the diffusion process becomes higher, as mentioned in the introductory section.

REFERENCES

- [1] Y.-C. Lee, J.-M. Wu, and M.-H. Hsieh, ‘‘First arrival position in molecular communication via generator of diffusion semigroup,’’ *arXiv preprint arXiv:2201.04476*, 2022.
- [2] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular communication*. Cambridge University Press, 2013.
- [3] P.-C. Yeh, K.-C. Chen, Y.-C. Lee, L.-S. Meng, P.-J. Shih, P.-Y. Ko, W.-A. Lin, and C.-H. Lee, ‘‘A new frontier of wireless communication theory: diffusion-based molecular communications,’’ *IEEE Wireless Communications*, vol. 19, no. 5, pp. 28–35, 2012.

- [4] I. F. Akyildiz, F. Brunetti, and C. Blázquez, “Nanonetworks: A new communication paradigm,” *Computer Networks*, vol. 52, no. 12, pp. 2260–2279, 2008.
- [5] 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.
- [6] M. Pierobon and I. F. 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, 2012.
- [7] S. Kadloor, R. S. Adve, and A. W. Eckford, “Molecular communication using brownian motion with drift,” *IEEE Transactions on NanoBioscience*, vol. 11, no. 2, pp. 89–99, 2012.
- [8] M. Moore, A. Enomoto, T. Nakano, R. Egashira, T. Suda, A. Kayasuga, H. Kojima, H. Sakakibara, and K. Oiwa, “A design of a molecular communication system for nanomachines using molecular motors,” in *Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW’06)*. IEEE, 2006, pp. 6–pp.
- [9] M. Gregori and I. F. Akyildiz, “A new nanonetwork architecture using flagellated bacteria and catalytic nanomotors,” *IEEE Journal on selected areas in communications*, vol. 28, no. 4, pp. 612–619, 2010.
- [10] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel, and R. Schober, “Channel modeling for diffusive molecular communication—a tutorial review,” *Proceedings of the IEEE*, vol. 107, no. 7, pp. 1256–1301, 2019.
- [11] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C.-B. Chae, “Three-dimensional channel characteristics for molecular communications with an absorbing receiver,” *IEEE Communications Letters*, vol. 18, no. 6, pp. 929–932, 2014.
- [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, 2012.
- [13] N. Pandey, R. K. Mallik, and B. Lall, “Molecular communication: The first arrival position channel,” *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 508–511, 2018.
- [14] C. E. Shannon, “A mathematical theory of communication,” *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [15] S. S. Andrews, “Accurate particle-based simulation of adsorption, desorption and partial transmission,” *Physical biology*, vol. 6, no. 4, p. 046015, 2009.
- [16] Y.-C. Lee, C.-C. Chen, P.-C. Yeh, and C.-H. Lee, “Distribution of first arrival position in molecular communication,” in *2016 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2016, pp. 1033–1037.
- [17] H. Li, S. M. Moser, and D. Guo, “Capacity of the memoryless additive inverse gaussian noise channel,” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 12, pp. 2315–2329, 2014.
- [18] A. W. Eckford, “Molecular communication: Physically realistic models and achievable information rates,” *arXiv preprint arXiv:0812.1554*, 2008.
- [19] Y. Murin, N. Farsad, M. Chowdhury, and A. Goldsmith, “Optimal detection for one-shot transmission over diffusion-based molecular timing channels,” *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 4, no. 2, pp. 43–60, 2018.
- [20] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.
- [21] R. D. Pierce, “Application of the positive alpha-stable distribution,” in *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics*. IEEE, 1997, pp. 420–424.
- [22] H. El Ghannudi, L. Clavier, N. Azaoui, F. Septier, and P.-A. Rolland, “ α -stable interference modeling and cauchy receiver for an ir-uwbb ad hoc network,” *IEEE Transactions on Communications*, vol. 58, no. 6, pp. 1748–1757, 2010.
- [23] K. Gulati, B. L. Evans, J. G. Andrews, and K. R. Tinsley, “Statistics of co-channel interference in a field of poisson and poisson-poisson clustered interferers,” *IEEE Transactions on Signal Processing*, vol. 58, no. 12, pp. 6207–6222, 2010.
- [24] J. Fahs and I. Abou-Faycal, “On the capacity of additive white alpha-stable noise channels,” in *2012 IEEE International Symposium on Information Theory Proceedings*. IEEE, 2012, pp. 294–298.
- [25] J. G. Gonzalez, D. W. Griffith, and G. R. Arce, “Zero-order statistics: a signal processing framework for very impulsive processes,” in *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics*. IEEE, 1997, pp. 254–258.
- [26] J. Fahs and I. Abou-Faycal, “Information measures, inequalities and performance bounds for parameter estimation in impulsive noise environments,” *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 1825–1844, 2017.
- [27] —, “A cauchy input achieves the capacity of a cauchy channel under a logarithmic constraint,” in *2014 IEEE International Symposium on Information Theory*. IEEE, 2014, pp. 3077–3081.
- [28] W. W. Bell, *Special functions for scientists and engineers*. Courier Corporation, 2004.
- [29] Z.-H. Yang and Y.-M. Chu, “On approximating the modified bessel function of the second kind,” *Journal of inequalities and applications*, vol. 2017, no. 1, pp. 1–8, 2017.
- [30] J. N. Kapur, *Maximum-entropy models in science and engineering*. John Wiley & Sons, 1989.
- [31] S. Kotz and S. Nadarajah, *Multivariate t-distributions and their applications*. Cambridge University Press, 2004.
- [32] M. Shao and C. L. Nikias, “Signal processing with fractional lower order moments: stable processes and their applications,” *Proceedings of the IEEE*, vol. 81, no. 7, pp. 986–1010, 1993.

APPENDIX A

DENSITY FUNCTION AND ENTROPY OF MULTIVARIATE CAUCHY DISTRIBUTION

The Cauchy distribution, named after Augustin Cauchy, is a continuous probability distribution often used in statistics as a canonical example of distribution since both its expected value and its variance are undefined. It is also known, especially among physicists, as the Lorentz distribution. In mathematics, it is closely related to the fundamental solution for the Laplace equation in the half-plane, and it is one of the few distributions that is stable (see [21], [24]) and has a density function that can be expressed analytically. (Other examples are normal distribution and Lévy distribution.)

The univariate⁶ Cauchy distribution has the probability density function (PDF) which can be expressed as:

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma} \frac{\gamma^2}{(x - x_0)^2 + \gamma^2} = \frac{1}{\pi\gamma} \frac{1}{1 + \left(\frac{x-x_0}{\gamma}\right)^2}, \quad (51)$$

where $x_0 \in \mathbb{R}$ is the location parameter, and $\gamma > 0$ is the scaling parameter (see [32]). For the purpose of exploring the channel capacity, since the location parameter x_0 is irrelevant to the entropy, we may assume without loss of generality that $x_0 = 0$, yielding the so called *symmetry Cauchy* with PDF as:

$$f(x; \gamma) = \frac{\gamma}{\pi} \frac{1}{x^2 + \gamma^2} = \frac{1}{\pi\gamma} \frac{1}{1 + \left(\frac{x}{\gamma}\right)^2}. \quad (52)$$

The entropy of Cauchy distribution can be evaluated by direct calculation. From Eq. (51), we have

$$h(X) = \ln(4\pi\gamma) \quad (53)$$

whenever $X \sim \text{Cauchy}(x_0, \gamma)$, and x_0 can be chosen arbitrarily.

As for multivariate Cauchy distribution, our notation system mainly follows [31]. Consider a p -dimensional Euclidean random vector $\mathbf{X} = (X_1, \dots, X_p)^\top$ which follows multivariate Cauchy distribution. We use the notation $\mathbf{X} \sim \text{Cauchy}_p(\boldsymbol{\mu}, \Sigma)$ to specify the parameters, where $\boldsymbol{\mu} \in \mathbb{R}^{p \times 1}$ is the location vector, and $\Sigma \in \mathbb{R}^{p \times p}$ is the scale matrix describing the

⁶For univariate Cauchy, people usually suppress the word “univariate”, and simply call it Cauchy.

shape of the distribution. The PDF of multivariate Cauchy distribution is given by the following formula:

$$f_{\mathbf{X}}^{(p)}(\mathbf{x}; \boldsymbol{\mu}, \Sigma) = \frac{\Gamma(\frac{1+p}{2})}{\Gamma(\frac{1}{2})\pi^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}[1 + (\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})]^{\frac{1+p}{2}}}. \quad (54)$$

Note that Σ is by nature a positive-definite square matrix. For our later purpose, we mainly work in the case $p = 2$. We provide the PDF of this special case here for convenience:

$$f_{\mathbf{X}}^{(2)}(\mathbf{x}; \boldsymbol{\mu}, \Sigma) = \frac{\Gamma(\frac{3}{2})}{\pi^{\frac{3}{2}}|\Sigma|^{\frac{1}{2}}[1 + (\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})]^{\frac{3}{2}}}. \quad (55)$$

Eq. (55) is the so-called bivariate Cauchy distribution.

Similar to the univariate case, we can without loss of generality set $\boldsymbol{\mu} = \mathbf{0}$ in equations (54) and (55) for the purpose of entropy and channel capacity analysis. When $\boldsymbol{\mu} = \mathbf{0}$, the multivariate Cauchy is called *central*. For the case that the Cauchy is central, we abuse the notations \mathbf{R} and Σ to mean the same scale matrix. We shall assume all Cauchy distributions considered in this paper to be central from now on. In addition, for the special case that Σ is a diagonal matrix, namely

$$\Sigma = \text{diag}(\gamma^2, \gamma^2) = \begin{bmatrix} \gamma^2 & 0 \\ 0 & \gamma^2 \end{bmatrix}, \quad (56)$$

the central bivariate Cauchy density function can be simplified into a more concise form:

$$\frac{1}{2\pi} \frac{\gamma}{(n_1^2 + n_2^2 + \gamma^2)^{3/2}}, \quad (57)$$

where $\mathbf{n} = (n_1, n_2)$ is the Cauchy random vector.

Finally, the (differential) entropy of p -variate central Cauchy distribution can be found in [4]. We briefly state the results here for later use. Suppose the scale matrix \mathbf{R} of a p -variate Cauchy \mathbf{X} is prescribed, the entropy of \mathbf{X} can be expressed as

$$h(\mathbf{X}; \mathbf{R}) = \frac{1}{2} \ln |\mathbf{R}| + \Phi(p), \quad (58)$$

where $|\mathbf{R}|$ represents the determinant of matrix \mathbf{R} ; $\Phi(p)$ is a constant depending only on dimension p , and it can be evaluated through

$$\begin{aligned} \Phi(p) = \ln & \left[\frac{\pi^{\frac{p}{2}}}{\Gamma(\frac{p}{2})} B\left(\frac{p}{2}, \frac{1}{2}\right) \right] \\ & + \frac{1+p}{2} \left[\psi\left(\frac{1+p}{2}\right) - \psi\left(\frac{1}{2}\right) \right]. \end{aligned} \quad (59)$$

In the above expression, $B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the so-called *beta function*, and

$$\psi(t) := \frac{d}{dt} [\ln \Gamma(t)] \quad (60)$$

is known as the *digamma function*.

APPENDIX B PROOF OF THEOREM 1

In this appendix section, we prove Theorem 1 step by step. For the additive channel

$$Y = X + N \quad (61)$$

we are considering, the mutual information can be written as $I(X; Y) = h(Y) - h(N)$ according to Section II. The issue of finding the capacity of channel (28) turns out to be an optimization problem:

$$C(A, \lambda) = \sup_{Y \in \mathcal{D}(A)} I(X; Y). \quad (62)$$

The calculation is as follows. We have

$$C(A, \lambda) = \sup_{Y \in \mathcal{D}(A)} \{h(Y) - h(N)\} \quad (63)$$

$$= \left\{ \sup_{Y \in \mathcal{D}(A)} h(Y) \right\} - \ln(4\pi\lambda), \quad (64)$$

where

$$\sup_{Y \in \mathcal{D}(A)} h(Y) = \sup_{k \in [\lambda, A]} \left\{ \sup_{Y \in \mathcal{D}_k} h(Y) \right\} \quad (65)$$

$$= \sup_{k \in [\lambda, A]} \ln(4\pi k) \quad (66)$$

$$= \ln(4\pi A). \quad (67)$$

Combining equations (64) and (67) yields

$$C(A, \lambda) = \ln(4\pi A) - \ln(4\pi\lambda) = \ln\left(\frac{A}{\lambda}\right). \quad (68)$$

Notice that the equalities (65)-(67) hold if and only if Y distributes as Cauchy(0, A). Hence, Theorem 1 is proved.

APPENDIX C

SOME KNOWN FACTS ABOUT CAUCHY DISTRIBUTION

For the probability density function formula and entropy evaluation of Cauchy distribution, please refer to Appendix A. In this appendix section, we recall three important properties of Cauchy distributions.

The first and second property is about the independent sum of two Cauchy distributions. Briefly speaking, univariate symmetry Cauchy distribution is closed under independent sum, as the following lemma states.

Lemma 1. Letting $U \sim \text{Cauchy}(0, \sigma)$, $V \sim \text{Cauchy}(0, \tau)$ be two independent Cauchy random variables, we have

$$Z = U + V \sim \text{Cauchy}(0, \sigma + \tau). \quad (69)$$

As for the independent sum property of bivariate Cauchy distributions, we have the following lemma.

Lemma 2. Letting $\mathbf{U} \sim \text{Cauchy}_2(\mathbf{0}, \Sigma_1 = \text{diag}(\sigma^2, \sigma^2))$, $\mathbf{V} \sim \text{Cauchy}_2(\mathbf{0}, \Sigma_2 = \text{diag}(\tau^2, \tau^2))$ be two independent bivariate Cauchy random vectors, we have

$$\mathbf{Z} = \mathbf{U} + \mathbf{V} \sim \text{Cauchy}_2(\mathbf{0}, \Sigma_3 = \text{diag}((\sigma + \tau)^2, (\sigma + \tau)^2)). \quad (70)$$

Notice that in Lemma 2, we have used the notation $\text{diag}(a, b) = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ for conveniently representing diagonal matrices.

The third property is about linear combination of components of Cauchy vector \mathbf{X} . Let $\mathbf{v} \in \mathbb{R}^{p \times 1}$ be an arbitrary constant vector, then the following lemma holds.

Lemma 3. If $\mathbf{X} \sim \text{Cauchy}_p(\boldsymbol{\mu}, \Sigma)$, then

$$\mathbf{v}^\top \mathbf{X} \sim \text{Cauchy}(\mathbf{v}^\top \boldsymbol{\mu}, \mathbf{v}^\top \Sigma \mathbf{v}), \quad (71)$$

where $\boldsymbol{\mu}$ is the location vector of \mathbf{X} and Σ is the scale matrix of \mathbf{X} .

This figure "new2D.png" is available in "png" format from:

<http://arxiv.org/ps/2201.11383v1>

This figure "new3D.png" is available in "png" format from:

<http://arxiv.org/ps/2201.11383v1>