Isotropic and Non-Isotropic Signaling in Multivariate α-Stable Noise

Malcolm Egan *

CITI Laboratory, Univ Lyon, INSA Lyon, Inria, Villeurbanne, France

A wide range of communication systems are corrupted by non-Gaussian noise, ranging from wireless to power line. In some cases, including interference in uncoordinated OFDM-based wireless networks, the noise is both impulsive and multivariate. At present, little is known about the information capacity and corresponding optimal input distributions. In this paper, we derive upper and lower bounds of the information capacity by exploiting non-isotropic inputs. For the special case of sub-Gaussian α-stable noise models, a numerical study reveals that isotropic Gaussian inputs can remain a viable choice, although the performance depends heavily on the dependence structure of the noise.

Keywords: α-stable distributions, heavy-tailed noise, non-isotropic inputs, information capacity, discrete inputs

1 INTRODUCTION

In many communication systems, additive Gaussian noise is the dominant form of signal corruption due to thermal fluctuations in the electronic devices comprising the receiver. Nevertheless, additive non-Gaussian noise has also been observed to play an important role in power line (Zimmermann and Dostert, 2002) and molecular communications (Farsad et al., 2015). Even in wireless communications, interference from uncoordinated transmitters, such as in the Internet of Things (IoT), has been suggested to admit non-Gaussian statistics (Clavier et al., 2021b). Another form of wireless communications where non-Gaussian noise arises is in underwater communications (Chitre et al., 2004).

A particularly important family of non-Gaussian noise models are impulsive, where the probability of large amplitude noise is significantly higher than predicted by corresponding Gaussian models; that is, impulsive noise is heavy-tailed. A key property of impulsive noise is that higher-order moments are often infinite or undefined, arising in Student's t (Hall, 1966), generalized Gaussian (Dytso et al., 2018), and α-stable models (Middleton, 1977; Sousa, 1992; Ilow and Hatzinakos, 1998; Gulati et al., 2010; Pinto and Win, 2010).

Of all impulsive noise families, one of the most ubiquitous are the α-stable models. As a generalization of Gaussian models admitting the key property known as stability under convolution, these models arise via several mechanisms. The first mechanism, relevant for molecular communications, is *via* the distribution of the first hitting time of the standard Wiener process (Farsad et al., 2015). The second mechanism is *via* the generalized central limit theorem, which characterizes the behavior of partial sums of n independent and identically distributed random variables under the scaling $n^{-\frac{1}{\alpha}}$ (Mahmood et al., 2014).

The third mechanism to obtain α-stable models relevant for interference in wireless communication systems was first identified by Middleton (Middleton, 1977) and further clarified in (Sousa, 1992; Ilow and Hatzinakos, 1998). In particular, given uncoordinated transmitting devices located according to a homogeneous Poisson point process on the plane, the interference under power-law path loss converges almost surely to an α-stable random variable by identification with the LePage series (Samorodnitsky and Taqqu, 1994). This third mechanism has recently seen application in interference studies for the IoT (Egan et al., 2018). Indeed both theory and recent experimental

data (Lauridsen et al., 2017) in the 868 MHz band, utilized by SigFox and LoRa devices, has indicated the presence of heavy-tailed interference which may be modeled via α-stable models (Clavier et al., 2021b).

Despite the utility of α-stable noise models in communications, the vast majority of work has focused on real-valued noise. In this setting, information capacity bounds have been derived in (de Freitas et al., 2017) and the structure of optimal input distributions characterized in (Fahs and Abou-Faycal, 2017). The design of symbol detection strategies and their performance has been addressed in (Niranjayan and Beaulieu, 2009; Ghannudi et al., 2010; Clavier et al., 2021a) and decoding algorithms developed in (Gu and Clavier, 2012; Mestrah et al., 2020). Noise parameter estimation algorithms have also been developed in (Kuruoglu, 2001) and power control strategies in (Freitas et al., 2018).

On the other hand, baseband signals in wireless communications are typically complex-valued for which few signal processing strategies and studies of performance analysis have been developed, with notable exceptions in (Gulati et al., 2010; Mahmood et al., 2014) for the narrowband case. The situation is further complicated when transmissions utilize orthogonal frequency division multiplexing (OFDM), where signals are transmitted over multiple subcarriers. In such cases, the noise forms a random vector and real-valued α-stable models are insufficient. Nevertheless, it has recently been shown that multivariate α-stable models can naturally arise from statistical analysis of interference in complex baseband signals over multiple subcarriers (Egan et al., 2018). However, little is known about performance limits or optimal signaling strategies in the presence of multivariate α-stable noise. In particular, the information capacity remains an open question in such channels, which is useful for selecting coding rates—*via* the noisy channel coding theorem—and in designing resource allocation strategies (Freitas et al., 2018).

In this paper, as a step towards resolving these open questions, we study the information capacity and signaling in multivariate symmetric α-stable noise channels with $1 < \alpha < 2$. We first return to the question of the information capacity in real-valued symmetric α-stable noise channels, where we establish new upper and lower bounds that are tighter and more general than those given in (de Freitas et al., 2017). In particular, bounds are also given for power-constrained inputs as well as fractional moment constraints. In the case of a power constraint, we establish that the information capacity is within a constant of the information capacity for the Gaussian noise channel and that Gaussian inputs yield this behavior.

We then turn to the case of multivariate symmetric α-stable noise. We show that there exists a unique optimal input achieving the information capacity and also derive a general upper bound, which is applicable to all multivariate symmetric α-stable noise channels subject to fractional moment and power constraints. We then derive a general lower bound applicable for fractional moment constraints with exponent $r < \alpha$. In the case of sub-Gaussian α-stable models, we also obtain a lower bound on the information capacity subject to a power constraint.

Our bounds suggest, at least from an analytical point of view, that it is desirable to match the dependence structure of the input distribution to that of the noise. Indeed, our lower bounds are obtained with non-isotropic inputs, often matched to the dependence structure of the noise distribution. To study the performance of non-isotropic inputs, we consider communication in sub-Gaussian α-stable noise subject to a power constraint, and numerically study the behavior of the bounds. In this particular case, we observe that isotropic Gaussian inputs nearly achieve the capacity upper bound, suggesting that matching the input to the dependence structure of the noise is not always desirable.

1.1 Notation

Vectors are denoted by bold lowercase letters and random vectors by bold uppercase letters, respectively (e.g., $\mathbf{x}$, $\mathbf{X}$). We denote the distribution of a random vector $\mathbf{X}$ by $P_{\mathbf{X}}$. If $\mathbf{X}$, $\mathbf{Y}$ are two random vectors equal in distribution, then we write $\mathbf{X} \overset{d}{=} \mathbf{Y}$.

Let $\mathbf{z} \in \mathbb{R}^d$, then $\|\mathbf{z}\|_r$, $1 \le r \le 2$ is given by

$$\|\mathbf{z}\|_r = \left(\sum_{i=1}^{d} |z_i|^r \right)^{\frac{1}{r}} \tag{1}$$

and $\|\cdot\|_r$ is called the r-norm on $\mathbb{R}^d$. For two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $\mathbf{a} \ge \mathbf{b}$ indicates that $a_i \ge b_i$, $i = 1, \ldots, n$.

Let $f \colon \mathbb{R} \to \mathbb{R}$ and $g \colon \mathbb{R} \to \mathbb{R}$. We use the Landau notation where $f(x) = o(g(x))$ if $\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0$ and $f(x) = O(g(x))$ if there exists $C > 0$ and $x_0 \in \mathbb{R}$ such that $f(x) \le Cg(x)$, $\forall x \ge x_0$.

2 PROBLEM FORMULATION

In this section, we detail the problem of characterizing the information capacity and optimal input distributions in multivariate symmetric α-stable noise channels ($1 < \alpha < 2$). To this end, we first recall preliminary definitions and properties of scalar and multivariate α-stable models that will be used in the sequel. For further details, we refer the reader to (Samorodnitsky and Taqqu, 1994).

2.1 α-Stable Models

The probability density function of an α-stable random variable is described by four parameters: the exponent $0 < \alpha \le 2$; the scale parameter $\gamma \in \mathbb{R}_+$; the skew parameter $\beta \in (-1, 1)$; and the shift parameter $\delta \in \mathbb{R}$. If X has an α-stable distribution, then we write $X \sim S_\alpha(\gamma, \beta, \delta)$. In the case $\beta = \delta = 0$, X is said to be a symmetric α-stable random variable.

In general, α-stable random variables do not have closed-form probability density functions. Instead, they are more compactly represented by the characteristic function, given by (Samorodnitsky and Taqqu, 1994, Eq. 1.1.6)

$$
\mathbb{E}[e^{i\theta X}]
$$

$$
= \begin{cases}
\exp\left\{ -\gamma^\alpha |\theta|^\alpha \left(1 - i\beta\,(\mathrm{sign}\,\theta) \tan \dfrac{\pi\alpha}{2} \right) + i\delta\theta \right\}, \\
\qquad\qquad\qquad \alpha \ne 1 \\[6pt]
\exp\left\{ -\gamma|\theta| \left(1 + i\beta \dfrac{2}{\pi}\,(\mathrm{sign}\,\theta)\log|\theta| \right) + i\delta\theta \right\}, \\
\qquad\qquad\qquad \alpha = 1
\end{cases}
\tag{2}
$$

Observe that in the special case $\alpha = 2$, the α-stable distribution is Gaussian. As such, the family of α-stable distribution generalize the family of Gaussian distributions. In fact, like Gaussian models, if $X^{(1)}$ and $X^{(2)}$ are independent copies of an α-stable random variable X, then for $a, b > 0$, there exists constants $c > 0, d \in \mathbb{R}$ such that

$$aX^{(1)} + bX^{(2)} \stackrel{d}{=} cX + d. \tag{3}$$

More precisely, the following property holds (Samorodnitsky and Taqqu, 1994).

Property 1. Suppose Z_1, Z_2 are independent with $Z_1 \sim S_\alpha(\gamma_1, \beta_1, \delta_1)$ and $Z_2 \sim S_\alpha(\gamma_2, \beta_2, \delta_2)$. Then, $Z_1 + Z_2 \sim S_\alpha(\gamma, \beta, \delta)$, where

$$\begin{aligned} \gamma &= (\gamma_1^\alpha + \gamma_2^\alpha)^{\frac{1}{\alpha}}, \\ \beta &= \frac{\beta_1\gamma_1^\alpha + \beta_2\gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha} \\ \delta &= \delta_1 + \delta_2. \end{aligned} \tag{4}$$

When $\beta = \delta = 0$ in **Eq. 2**, the resulting α-stable distribution is said to be symmetric. An important alternative characterization of symmetric α-stable random variables is *via* the LePage series.

Theorem 1. [Theorem 1.4.2 (Samorodnitsky and Taqqu, 1994)]. Suppose $0 < \alpha < 2$, $(\Gamma_i)_{i=1}^\infty$ is a homogeneous Poisson point process with intensity 1, and $(W_i)_{i=1}^\infty$ are symmetric, independent and identically distributed random variables satisfying $\mathbb{E}[|W_i|^\alpha] < \infty$. Then,

$$\sum_{i=1}^\infty \Gamma_i^{-\frac{1}{\alpha}} W_i \tag{5}$$

converges almost surely to a random variable $X \sim S_\alpha\left(\left(C_\alpha^{-1}\mathbb{E}[|W_1|^\alpha]\right)^{1/\alpha}, 0, 0\right)$, where

$$C_\alpha = \begin{cases} \dfrac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)}, & \alpha \neq 1 \\ 2/\pi, & \alpha = 1. \end{cases} \tag{6}$$

In the multivariate setting, we consider random vectors $\mathbf{X}$ in $\mathbb{R}^d \sim d > 1$. Analogously to the scalar case ($d = 1$), a random vector $\mathbf{X} \in \mathbb{R}^d$ is a symmetric α-stable random vector if for all a, $b > 0$ there exists $c > 0$ such that

$$a\mathbf{X}^{(1)} + b\mathbf{X}^{(2)} \stackrel{d}{=} c\mathbf{X}, \tag{7}$$

where $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are independent copies of $\mathbf{X}$.

A sufficient condition for a random vector $\mathbf{X}$ in $\mathbb{R}^d$ to be a symmetric α-stable random vector is that all linear combinations of the elements of $\mathbf{X}$ are symmetric α-stable (Samorodnitsky and Taqqu, 1994). In general, d-dimensional symmetric α-stable random vectors can be represented via their characteristic function, given by (Samorodnitsky and Taqqu, 1994)

$$\mathbb{E}[e^{i\boldsymbol{\theta}\cdot\mathbf{X}}] = \exp\left(-\int_{\mathbb{S}^{d-1}} \left|\sum_{k=1}^d \theta_k s_k\right|^\alpha \Gamma(d\mathbf{s})\right), \tag{8}$$

where Γ is the unique symmetric measure on the surface of the d-dimensional unit sphere.

In the case that a d-dimensional α-stable random vector $\mathbf{X}$ is truly d-dimensional, there exists a joint probability density function $p_\mathbf{X}(\cdot)$ on $\mathbb{R}^d$. Note that a simple necessary and sufficient condition for $\mathbf{X}$ to be truly d-dimensional is for the support of the spectral measure to span $\mathbb{R}^d$ (Byczkowski et al., 1993). This condition means that degenerate α-stable random vectors (e.g., when $X_i = X_j$ for some $i \neq j$, $i, j \in \{1, \ldots, d\}$) are not considered.

A key family of truly α-stable random vectors are the sub-Gaussian α-stable random vectors, defined as follows.

Definition 1. Any vector $\mathbf{X}$ satisfying $\mathbf{X} \stackrel{d}{=} (A^{1/2}G_1, \ldots, A^{1/2}G_d)$, where

$$A \sim S_{\alpha/2}\left((\cos\pi\alpha/4)^{2/\alpha}, 1, 0\right), \tag{9}$$

and $\mathbf{G} = [G_1, \ldots, G_d]^T \sim \mathcal{N}(\mathbf{0}, \Sigma)$ is called a sub-Gaussian α-stable random vector in $\mathbb{R}^d$ with underlying Gaussian vector $\mathbf{G}$. If $\Sigma = \sigma^2\mathbf{I}$ with $\sigma > 0$, then $\mathbf{X}$ is said to be an isotropic sub-Gaussian α-stable random vector.

2.2 The Information Capacity Problem

Consider the memoryless, stationary, linear and point-to-point communication channel

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}, \tag{10}$$

where $\mathbf{N}$ is a truly symmetric α-stable random vector with $1 < \alpha < 2$, admitting a multivariate probability density function $p_\mathbf{N}(\cdot)$, with $\mathbf{X}$ and $\mathbf{N}$ independent. The random vector $\mathbf{X}$ is defined on[1] $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with probability measure $\mu_\mathbf{X}$ lying in the set $\Lambda_\mathbf{X}(\mathbf{P}, r)$, where

$$\Lambda_\mathbf{X}(\mathbf{P}, r) = \{\mu_\mathbf{X}: (\mathbb{E}_{\mu_\mathbf{X}}[|X_1|^r], \ldots, \mathbb{E}[|X_d|^r]) \succeq \mathbf{P}\}, \tag{11}$$

for a given $\mathbf{P} \in \mathbb{R}_+^d$ and $1 \leq r \leq 2$. As such, the set $\Lambda_\mathbf{X}(\mathbf{P}, r)$ corresponds to the set of inputs satisfying element-wise moment constraints. Note that by virtue of $\mathbf{N}$ admitting a probability density function, $\mathbf{Y}$ also admits a probability density function $p_\mathbf{Y}(\cdot)$.

The main focus of this paper is to investigate the information capacity and corresponding optimal inputs for communication channels of the form **Eq. 10**. To this end, let $\mathcal{P}(\mathbb{R}^d)$ be the set of probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ equipped with the topology of weak convergence (Billingsley, 1999).

$$C(\mathbf{P}, r) = \sup_{\mu_\mathbf{X} \in \Lambda_\mathbf{X}(\mathbf{P}, r)} I(\mathbf{X}; \mathbf{Y}) \tag{12}$$

where the mutual information $I(\mathbf{X}; \mathbf{Y})$ is given by

---

[1]The notation $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel σ-algebra on $\mathbb{R}^d$.

$$I(\mathbf{X}; \mathbf{Y}) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p_N(\mathbf{y} - \mathbf{x}) \log \frac{p_N(\mathbf{y} - \mathbf{x})}{p_Y(\mathbf{y})} d\mathbf{y} \mu_{\mathbf{X}}(d\mathbf{x}). \quad (13)$$

In the case an optimal input exists, it satisfies

$$\mu_{\mathbf{X}}^* \in \arg \max_{\mu_{\mathbf{X}} \in \Lambda_{\mathbf{X}}(\mathbf{P}, r)} I(\mathbf{X}; \mathbf{Y}). \quad (14)$$

Note that, by a generalization of Shannon's noisy channel coding theorem for vector non-Gaussian channels (Han, 2003), the information capacity may be interpreted as the maximum achievable rate with asymptotically zero average probability of error.

In the remainder of this paper, we address the following questions:

(i) What is the value of $C(\mathbf{P}, r)$ for varying $\mathbf{P}$ and r?
(ii) Does an optimal input $\mu_{\mathbf{X}}^* \in \Lambda_{\mathbf{X}}(\mathbf{P}, r)$ exist?
(iii) What is the structure of nearly optimal inputs?

In this work, we will allow $\mu_{\mathbf{X}}$ to be non-isotropic; that is, for all $d \times d$ orthogonal matrices $\mathbf{O} \in \mathcal{O}(d)$, it does not hold that

$$\mathbf{OX} \overset{d}{=} \mathbf{X}, \quad (15)$$

where $\mathbf{X}$ has probability measure $\mu_{\mathbf{X}}$. In the following section, we begin with scalar channels—which have not previously been comprehensively studied—before considering more general vector channels in **Section 4**.

3 SCALAR CHANNELS

Before turning to multivariate α-stable noise channels, we first consider the scalar case. In particular, we first improve on the capacity bounds in (de Freitas et al., 2017) and in the process develop techniques that will be generalized to the multivariate setting in the sequel. To begin, we specialize the problem in **Eq. 12** to the scalar case: a stationary and memoryless scalar additive symmetric α-stable noise channel is given by

$$Y = X + N, \quad (16)$$

where the noise N is a symmetric α-stable random variable with scale parameter γ_N, admitting a probability density function $p_N(\cdot)$, with X and N independent. The input random variable X is required to satisfy the constraint

$$\mathbb{E}[|X|^r] \le P, \quad (17)$$

where $1 \le r \le 2$. In terms of the probability measure of X, the constraint can be written as

$$\Lambda_X(P, r) = \{\mu_X \in \mathcal{P}(\mathbb{R}): \mathbb{E}_{\mu_X}[|X|^r] \le P\}. \quad (18)$$

In this case, the information capacity of the channel (16) is defined as

$$C(P) = \sup_{\mu \in \Lambda_X(P, r)} I(X; Y) \quad (19)$$

It follows from (Fahs and Abou-Faycal, 2016) that an optimal solution of **Eq. 19** exists and is unique. Indeed, the optimal input is known to be discrete (Fahs and Abou-Faycal, 2017).

3.1 Capacity Upper Bounds
In (de Freitas et al., 2017), an upper bound on $C(P, r)$ was established when $r = 1$ and $1 < \alpha < 2$.

Theorem 2. *Let* $\lambda > 0$ *and* $r = 1$. *For the channel* (16), *the capacity* $C(P, 1)$ *in* (19) *is upper bounded by*

$$C(P, 1) \le \log \left(\frac{2\Gamma\left(\frac{1}{\alpha}\right)}{\lambda \gamma \alpha \pi} \right) + \lambda \left(\frac{2\gamma \Gamma\left(1 - \frac{1}{\alpha}\right)}{\pi} + P \right). \quad (20)$$

It was shown in (de Freitas et al., 2017) that this bound was tight for moderate values of P and appropriate values of λ, but quickly diverged. An asymptotic upper bound, that is, the upper bound is only guaranteed to hold as $P \to \infty$, was established. In the following theorem, we establish an upper bound which holds for all $1 \le r \le 2$ and $p > 0$.

Theorem 3. *Let* $1 \le r \le 2$ *and* $p > 0$. *For the channel* (16), *the capacity* $C(P, r)$ *in* (19) *is upper bounded by*

$$C(P, r) \le C_{UB}(P, r)$$
$$= \log \left(2 \left(P^{\frac{1}{r}} + \frac{2\Gamma\left(1 - \frac{1}{\alpha}\right)}{\pi} \gamma_N \right) e \right) - h(N). \quad (21)$$

Proof. Note that under the constraint that $\mathbb{E}[|Y|] = c_Y$, $c_Y > 0$, the entropy is maximized by the Laplace distribution (Cover and Thomas, 2006). This yields a bound of

$$h(Y) \le \log(2c_Y e). \quad (22)$$

By the triangle inequality,

$$\mathbb{E}[|Y|] \le \mathbb{E}[|X|] + \mathbb{E}[|N|]. \quad (23)$$

We also have by (Zolotarev, 1957)

$$\mathbb{E}[|N|] = \frac{2\Gamma\left(1 - \frac{1}{\alpha}\right)}{\pi} \gamma_N. \quad (24)$$

All that remains is to obtain $\mathbb{E}[|X|^r]$. By construction, $\mathbb{E}[|X|^r] \le P$. Using Hölder's inequality then yields

$$\mathbb{E}[|X|] \le (\mathbb{E}[|X|^r])^{\frac{1}{r}} = P^{\frac{1}{r}}. \quad (25)$$

Substituting **Eqs 23–25** into **Eq. 22** gives

$$C(P) \le \log \left(2 \left(P^{\frac{1}{r}} + \frac{2\Gamma\left(1 - \frac{1}{\alpha}\right)}{\pi} \gamma_N \right) e \right) - h(N), \quad (26)$$

as required.

3.2 Capacity Lower Bounds
We now turn to lower bounding $C(P, r)$. We first consider the case where $1 \le r < \alpha$.

Theorem 4. Let $1 \le r < \alpha$. For the channel (16), the capacity $C(P, r)$ in (19) is lower bounded by

$$C(P,r) \ge C_{LB}(P,r)$$

$$= \frac{1}{\alpha}\log\left(1 + \frac{P^{\frac{\alpha}{r}}}{\gamma_N^\alpha}\left(\frac{\alpha\sqrt{\pi}\,\Gamma\left(-\frac{r}{2}\right)}{2^{r+1}\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(-\frac{r}{\alpha}\right)}\right)^{\frac{1}{r}}\right). \quad (27)$$

Proof. Let $X \sim S_\alpha(\gamma_X, 0, 0)$ with $\gamma_X \in \mathbb{R}_+$. Consider the random variable $U \sim S_\alpha(1, 0, 0)$. By the scaling and translation properties of α-stable random variables, we can write

$$\begin{aligned} X &\overset{d}{=} \gamma_X U \\ N &\overset{d}{=} \gamma_N U. \end{aligned} \quad (28)$$

By the stability property

$$Y = X + N \sim S_\alpha\left((\gamma_X^\alpha + \gamma_N^\alpha)^{\frac{1}{\alpha}}, 0, 0\right) \quad (29)$$

and hence

$$Y \overset{d}{=} \gamma_Y U, \quad (30)$$

where $\gamma_Y = (\gamma_X^\alpha + \gamma_N^\alpha)^{\frac{1}{\alpha}}$.

We then have

$$\begin{aligned} I(X;Y) &= h(Y) - h(Y|X) \\ &= h(\gamma_Y U) - h(\gamma_N U) \\ &= h(U) + \log(\gamma_Y) - h(U) - \log(\gamma_N) \\ &= \log\left(\frac{(\gamma_X^\alpha + \gamma_N^\alpha)^{\frac{1}{\alpha}}}{\gamma_N}\right) \\ &= \frac{1}{\alpha}\log\left(1 + \frac{\gamma_X^\alpha}{\gamma_N^\alpha}\right). \end{aligned} \quad (31)$$

Using (Shao and Nikias, 1993, Theorem 4)

$$\mathbb{E}[|X|^r] = \frac{2^{r+1}\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(-\frac{r}{\alpha}\right)}{\alpha\sqrt{\pi}\Gamma\left(\frac{-r}{2}\right)}\gamma_X^r. \quad (32)$$

and the constraint $\mathbb{E}[|X|^r] \le P$, it follows that

$$C(P) \ge \frac{1}{\alpha}\log\left(1 + \frac{P^{\frac{\alpha}{r}}}{\gamma_N^\alpha}\left(\frac{\alpha\sqrt{\pi}\,\Gamma\left(-\frac{r}{2}\right)}{2^{r+1}\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(-\frac{r}{\alpha}\right)}\right)^{\frac{1}{r}}\right). \quad (33)$$

Remark 1. When $r = 1$, Theorem 4 specializes to the lower bound in (de Freitas et al., 2017).

Since $1 \le r < \alpha$ and $\alpha < 2$, it follows that Theorem 4 does not apply in the important case where the input X is constrained to satisfy $\mathbb{E}[X^2] \le P$. In the following theorem, we establish a lower bound in this setting.

Theorem 5. Let $r = 2$. For the channel (16), the capacity $C(P, 2)$ in (19) is lower bounded by

$$\begin{aligned} C(P,2) &\ge C_{LB}(P,2) \\ &= \int_0^\infty \frac{1}{2}p_A(a)\log(2\pi e(P+a))\mathrm{d}a - h(N), \end{aligned} \quad (34)$$

where $p_A(a)$ is the probability density function of a totally skewed $\alpha/2$-stable random variable with scale parameter $\cos^{2/\alpha}(\frac{\pi}{4}\alpha)$.
Proof. Let $X \sim \mathcal{N}(0, P)$ and $G \sim \mathcal{N}(0, 2\gamma_N^2)$. Then,

$$\begin{aligned} h(Y) &= h(X + N) \\ &= h(X + A^{1/2}G) \\ &\ge h(X + A^{1/2}G|A) \\ &= h(G_A|A), \end{aligned} \quad (35)$$

where $G_A \sim \mathcal{N}(0, P + A\gamma_N^2)$. Hence,

$$h(Y) \ge \int_0^\infty \frac{1}{2}p_A(a)\log(2\pi e(P + 2a\gamma_N^2))\mathrm{d}a. \quad (36)$$

As such,

$$C(P) \ge \int_0^\infty \frac{1}{2}p_A(a)\log(2\pi e(P + 2a\gamma_N^2))\mathrm{d}a - h(N), \quad (37)$$

as required.

A key question is whether the capacity bounds we have established so far are tight. To this end, we make the following observation.

Corollary 1. Let $1 \le r < \alpha$. For the channel (16), the capacity $C(P, r)$ in (19) satisfies

$$C(P,r) = \frac{1}{r}\log P + O(1), \quad P \to \infty. \quad (38)$$

Proof. Observe that

$$\begin{aligned} C_{LB}(P,r) &= \frac{1}{r}\log P + O(1), \quad P \to \infty, \\ C_{UB}(P,r) &= \frac{1}{r}\log P + O(1), \quad P \to \infty. \end{aligned} \quad (39)$$

Since

$$C_{LB}(P,r) \le C \le C_{UB}(P,r), \quad (40)$$

the corollary follows.
By the same argument, we also have the following corollary.

Corollary 2. Let $r = 2$. For the channel (16), the capacity $C(P, 2)$ in (19) satisfies

$$C(P,2) = \frac{1}{2}\log P + O(1), \quad P \to \infty. \quad (41)$$

As a consequence, for sufficiently large values of P and $r = 2$, the rate achievable using a Gaussian input is within a constant of the capacity $C(P, 2)$. A further observation, which will be useful in the sequel, is that matching the input distribution to the noise distribution yields a rate that forms a good approximation of the capacity. Finally, the
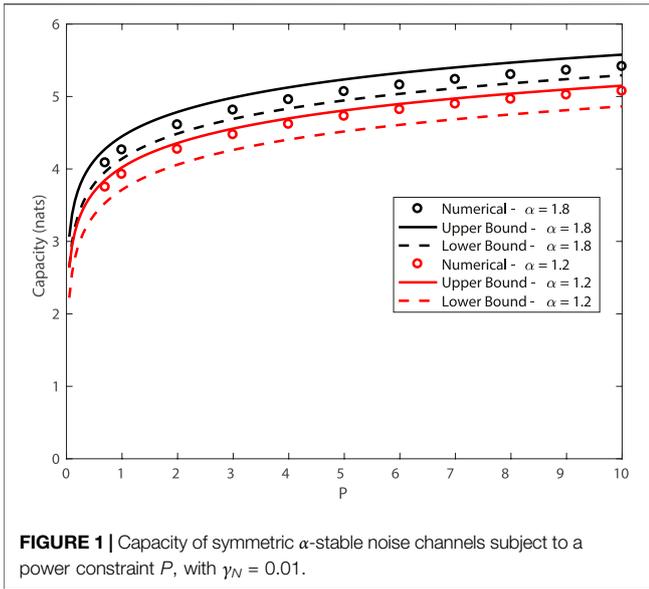
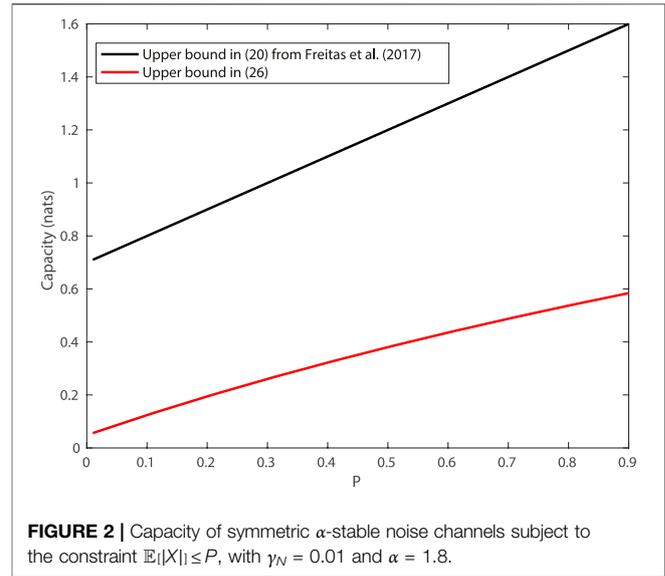**FIGURE 1** | Capacity of symmetric α-stable noise channels subject to a power constraint P, with $\gamma_N = 0.01$.

FIGURE 2 | Capacity of symmetric α-stable noise channels subject to the constraint $\mathbb{E}_{[}|X_{]}| \leq P$, with $\gamma_N = 0.01$ and $\alpha = 1.8$.

capacity of symmetric α-stable noise channels is within a constant of the capacity for an additive Gaussian noise channel.

3.3 Numerical Results

In order to further verify the tightness of the upper bound in Theorem 3, we compare the bound with the numerical computation of the capacity *via* the Blahut-Arimoto algorithm (Blahut, 1972; Arimoto, 1972). **Figure 1** plots the power against the information capacity for varying and α. The scale parameter is set as $\gamma_N = 0.01$. Observe that the upper bound and the numerical approximation are in good agreement. Note that the lower bound is obtained based on a Gaussian input is also in good agreement, despite the fact that the optimal input is discrete (Fahs and Abou-Faycal, 2017).

Figure 2 plots the capacity upper bounds in Theorem 2 [from (de Freitas et al., 2017)] and our new upper bound in Theorem 3 in the case of the constraint $\mathbb{E}_{[}|X| \leq P$ with $\gamma_N = 0.01$ and $\alpha = 1.8$. The parameter λ required in the bound from Theorem 2 corresponds to $\lambda = 1$. Observe that for all plotted values of P, the new bound in Theorem 3 is below that of Theorem 2, implying that the new bound is tighter. Note that the improvement over the bound in (de Freitas et al., 2017) is already evident from the form of the bounds for large P, due to the fact that P dominates $\log P$.

4 VECTOR CHANNELS

In this section, we return to the general problem in **Eq. 12** for vector channels with $d > 1$.

4.1 Existence and Uniqueness of Optimal Inputs

While existence and uniqueness of optimal inputs is well understood in the scalar case (Fahs and Abou-Faycal, 2016), it has not yet been established in the vector case. We prove this result in the following theorem by utilizing the theory of weak convergence (Billingsley, 1999).

Theorem 6. For the optimization problem in (12), there exists a unique input probability measure $\mu^\star$ corresponding to an input random vector $\mathbf{X}^\star$ on $(\mathbb{R}^d, \mathcal{B}^d)$ such that $C(\mathbf{P}, r) = I(\mathbf{X}^\star; \mathbf{Y})$.

Proof. The proof proceeds in three steps: (i) weak compactness of the constraint set $\Lambda_{\mathbf{X}}(\mathbf{P}, r)$; (ii) weak continuity of $I(\mathbf{X}; \mathbf{Y})$ on $\Lambda_{\mathbf{X}}(\mathbf{P}, r)$, yielding existence of $\mu_{\mathbf{X}}^*$; and (iii) uniqueness of $\mu_{\mathbf{X}}^*$.

(i) For any $\epsilon > 0$, there exists $\mathbf{a}_\epsilon = [a_{1,\epsilon}, \ldots, a_{d,\epsilon}]^T > 0$ such that for all $\mu \in \Lambda_{\mathbf{X}}(\mathbf{P}, r)$

$$\Pr\left(|X_1|^r > a_{1,\epsilon}, \ldots, |X_d|^r > a_{d,\epsilon}\right) < \epsilon. \quad (42)$$

The inequality in **Eq. 42** holds as a consequence of the generalized Markov inequality in (Marshall, 1984, Example 2.3). In more detail,

$$\Pr\left(|X_1|^r > a_{1,\epsilon}, \ldots, |X_d|^r > a_{d,\epsilon}\right) \leq \min_{i=1,\ldots,d} \frac{\mathbb{E}[|X_i|^r]}{a_{i,\epsilon}}$$
$$\leq \min_{i=1,\ldots,d} \frac{c_i}{a_{i,\epsilon}} < \epsilon. \quad (43)$$

Now, choose $\mathcal{K}_\epsilon = [-a_{1,\epsilon}, a_{1,\epsilon}] \times \cdots \times [-a_{d,\epsilon}, a_{d,\epsilon}]$. Then, $\mathcal{K}_\epsilon$ is compact subset of $\mathbb{R}^d$ and $\mu(\mathcal{K}_\epsilon) \geq 1 - \epsilon$ for all $\mu \in \Lambda_{\mathbf{X}}(\mathbf{P}, r)$. Hence, $\Lambda_{\mathbf{X}}(\mathbf{P}, r)$ is tight.

To establish closure, we apply a variation of the Portmanteau theorem (Billingsley, 1999). Let $\{\mu_n\}_{n=1}^\infty$ be a weakly convergent sequence in $\Lambda_{\mathbf{X}}(\mathbf{P}, r)$ with limit μ_0. By a consequence of the Portmanteau theorem, it follows that

$$\left[\mathbb{E}_{\mu_0}[|X_1|^r], \ldots, \mathbb{E}[|X_d|^r]\right]^T$$
$$= \left[\int |x_1|^r d\mu_0(\mathbf{x}), \ldots, \int |x_d|^r d\mu_0(\mathbf{x})\right]^T$$
$$\leq \left[\liminf_{n \to \infty} \int |x_1|^r d\mu_n(\mathbf{x}), \ldots, \liminf_{n \to \infty} \int |x_d|^r d\mu_n(\mathbf{x})\right]^T$$
$$\leq \mathbf{P}. \quad (44)$$

Hence, $\mu_0 \in \Lambda_{\mathbf{X}}(\mathbf{P}, r)$. Since the choice of sequence is arbitrary, it follows that $\Lambda_{\mathbf{X}}(\mathbf{P}, r)$ is closed in the topology of weak

convergence. Since $\Lambda_{\mathbf{X}}(\mathbf{P}, r)$ is tight and closed in the topology of weak convergence, it then follows by Prokhorov's theorem (Billingsley, 1999) that $\Lambda_{\mathbf{X}}(\mathbf{P}, r)$ is compact.

(ii) The second step is to establish that $I(\mathbf{X}; \mathbf{Y})$ is weakly continuous on $\Lambda_{\mathbf{X}}(\mathbf{P}, r)$. In particular, we need to show that for any weakly convergent sequence of probability measures $(\mu_n)_{n=1}^{\infty}$ with limit μ_0

$$\lim_{n\to\infty} -\int p_{\mathbf{Y}_n}(\mathbf{y})\log p_{\mathbf{Y}_n}(\mathbf{y})d\mathbf{y} = -\int p_{\mathbf{Y}_0}(\mathbf{y})\log p_{\mathbf{Y}_0}(\mathbf{y})d\mathbf{y}, \quad (45)$$

where $\mathbf{Y}_n$ is the output corresponding to an input $\mathbf{X}_n$ with probability measure μ_n. Note that $\mathbf{Y}_n = \mathbf{X}_n + \mathbf{N}$ admits a probability density function since $\mathbf{N}$ is truly d-dimensional.

Observe that if the limit and the integral in **Eq. 45** can be swapped, the result follows from the definition of weak convergence if the probability density function of $\mathbf{N}$, $p_{\mathbf{N}}$, is bounded and continuous. Note that this is indeed the case since the characteristic function of $\mathbf{N}$, $\Phi_{\mathbf{N}}(\mathbf{t}) = \mathbb{E}[e^{i\mathbf{t}\cdot\mathbf{N}}]$ is integrable when $\mathbf{N}$ is truly d-dimensional.

To complete the proof, we must justify swapping of the limit and integral in **Eq. 45**. Let $1 < s < \alpha$. We need to establish that for all $n \geq 0$ and any $\delta > 0$, there exists $R(\delta) > 0$ such that

$$\left| \int_{\|\mathbf{y}\|_s > R(\delta)} p_{\mathbf{Y}_n}(\mathbf{y})\log p_{\mathbf{Y}_n}(\mathbf{y})d\mathbf{y} \right| < \delta. \quad (46)$$

To proceed, let

$$q(\mathbf{y}) = \frac{1}{\pi^d} \frac{1}{\prod_{i=1}^{d}(1 + y_i^2)}, \quad \mathbf{y} \in \mathbb{R}^d. \quad (47)$$

which is a Cauchy density on $\mathbb{R}^d$. Observe that

$$-\int_{\|\mathbf{y}\|_s > R(\delta)} p_{\mathbf{Y}_n}(\mathbf{y})\log p_{\mathbf{Y}_n}(\mathbf{y})d\mathbf{y}$$
$$= -\int_{\|\mathbf{y}\|_s > R(\delta)} p_{\mathbf{Y}_n}(\mathbf{y})\log q(\mathbf{y})d\mathbf{y}$$
$$- \int_{\|\mathbf{y}\|_s > R(\delta)} p_{\mathbf{Y}_n}(\mathbf{y})\log \frac{p_{\mathbf{Y}_n}(\mathbf{y})}{q(\mathbf{y})}d\mathbf{y} \quad (48)$$
$$\leq \int_{\|\mathbf{y}\|_s > R(\delta)} p_{\mathbf{Y}_n}(\mathbf{y})\left(\log\pi^d + \log\prod_{i=1}^{d}(1 + y_i^2) \right)d\mathbf{y}$$
$$+ \frac{1}{e}\int_{\|\mathbf{y}\|_s > R(\delta)} q(\mathbf{y})d\mathbf{y},$$

where the last term follows from the fact that $a\log a \geq -\frac{1}{e}$, $a > 0$.

Note that by the Markov inequality,

$$\int_{\|\mathbf{y}\|_s > R(\delta)} \log\pi^d\, p_{\mathbf{Y}_n}(\mathbf{y})d\mathbf{y} \quad \leq d\log\pi\frac{\mathbb{E}[\|\mathbf{Y}\|_s]}{R(\delta)}$$
$$\leq d\log\pi\frac{L}{R(\delta)}, \quad (49)$$

which tends to zero as $R(\delta) \to \infty$. Here, $L < \infty$ since by the Jensen and Hölder inequalities

$$\mathbb{E}[\|\mathbf{X}\|_s] = \mathbb{E}\left[\left(\sum_{i=1}^{d}|X_i|^s \right)^{1/s} \right] \leq \left(\mathbb{E}\left[\sum_{i=1}^{d}|X_i|^s \right] \right)^{1/s}$$
$$\leq \left(\sum_{i=1}^{d}(\mathbb{E}[|X_i|^r])^{s/r} \right)^{1/s} \quad (50)$$

which implies

$$\mathbb{E}[\|\mathbf{Y}\|_s] \leq \mathbb{E}[\|\mathbf{X}\|_s] + \mathbb{E}[\|\mathbf{N}\|_s] \leq \left(\sum_{i=1}^{d}(\mathbb{E}[|X_i|^r])^{s/r} \right)^{1/s}$$
$$+ \mathbb{E}[\|\mathbf{N}\|_s] < L < \infty, \quad (51)$$

since the probability measure μ corresponding to $\mathbf{X}$ lies in $\Lambda_{\mathbf{X}}(\mathbf{P}, r)$. Similarly,

$$\frac{1}{e}\int_{\|\mathbf{y}\|_s > R(\delta)} q(\mathbf{y})d\mathbf{y}$$
$$\leq \frac{1}{eR(\delta)}\mathbb{E}_q[\|\mathbf{Y}\|_s], \quad (52)$$

which tends to zero as $R(\delta) \to \infty$.

Moreover,

$$\int_{\|\mathbf{y}\|_s > R(\delta)} p_{\mathbf{Y}_n}(\mathbf{y})\log\prod_{i=1}^{d}(1 + y_i^2)d\mathbf{y}$$
$$\leq 2\int_{\|\mathbf{y}\|_s > R(\delta)} p_{\mathbf{Y}_n}(\mathbf{y})\log\prod_{i=1}^{d}(1 + |y_i|)d\mathbf{y}$$
$$\leq 2\sup_{\|\mathbf{y}\|_s > R(\delta)} \left\{ \frac{\log\prod_{i=1}^{d}(1 + |y_i|)}{\|\mathbf{y}\|_s} \right\}\int_{\|\mathbf{y}\|_s > R(\delta)} \|\mathbf{y}\|_s p_{\mathbf{Y}_n}(\mathbf{y})d\mathbf{y}. \quad (53)$$

Note that

$$\frac{\log\prod_{i=1}^{d}(1 + |y_i|)}{\|\mathbf{y}\|_s} \leq \frac{\sum_{i=1}^{d}|y_i|}{\|\mathbf{y}\|_s} \leq 1, \quad (54)$$

by Hölder's inequality. As such, **Eq. 53** is finite and tends to zero as $R(\delta) \to \infty$.

After an application of the dominated convergence theorem, for any $\delta > 0$

$$\lim_{n\to\infty} -\int_{\|\mathbf{y}\|_r \leq R(\delta)} p_{\mathbf{Y}_n}(\mathbf{y})\log p_{\mathbf{Y}_n}(\mathbf{y})d\mathbf{y}$$
$$= -\int_{\|\mathbf{y}\|_r \leq R(\delta)} p_{\mathbf{Y}_0}(\mathbf{y})\log p_{\mathbf{Y}_0}(\mathbf{y})d\mathbf{y}. \quad (55)$$

Since the identities in **Eqs 49, 52, 53, 55** hold for all $\delta > 0$, weak continuity of $I(\mathbf{X}; \mathbf{Y})$ follows by taking $\delta \to 0$ (and hence $R(\delta) \to \infty$). The existence part of Theorem 6 then holds by applying the extreme value theorem.

(iii) The uniqueness of the optimal input follows from the fact that the entropy $h(\mathbf{Y})$ is a strictly concave function of the distribution $P_{\mathbf{Y}}$. By the fact that the characteristic function of $\mathbf{N}$ is strictly positive, $P_{\mathbf{Y}}$ is a one-to-one function of the

distribution $P_{\mathbf{X}}$. Hence, $h(\mathbf{Y})$ is a strictly concave function of $P_{\mathbf{X}}$. As the mutual information can be written as

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{N}) \qquad (56)$$

it follows that $I(\mathbf{X}; \mathbf{Y})$ is a strictly concave function of $P_{\mathbf{X}}$ since $h(\mathbf{N})$ does not depend on $P_{\mathbf{X}}$. Since this holds for any input lying in $\Lambda(\mathbf{P}, r)$, it then follows that the optimal input probability measure $\mu_{\mathbf{X}}^*$ is unique.

4.2 Capacity Upper Bound

We now obtain a general upper bound on the capacity in multivariate α-stable noise, which holds for constraints with $1 \leq r \leq 2$.

Theorem 7. Let $1 < \alpha < 2$, $1 \leq r \leq 2$, $\mathbf{P} \geq \mathbf{0}$ and γ_{N_i} be the scale parameter for the *ith* element of $\mathbf{N}$. The capacity $C(\mathbf{P}, r)$ in (12) is upper bounded by

$$C(\mathbf{P}, r) \leq C_{UB}(\mathbf{P}, r)$$
$$= \sum_{i=1}^{d} \log\left(2\left(P_i^{\frac{1}{r}} + \frac{2\Gamma\left(1 - \frac{1}{\alpha}\right)}{\pi}\gamma_{N_i}\right)e\right) - h(\mathbf{N}). \qquad (57)$$

Proof. Recall that

$$h(\mathbf{Y}) \leq \sum_{i=1}^{d} h(Y_i). \qquad (58)$$

For each term $h(Y_i)$, the same argument as Theorem 3 yields

$$h(Y_i) \leq \log\left(2\left(P_i^{\frac{1}{r}} + \frac{2\Gamma\left(1 - \frac{1}{\alpha}\right)}{\pi}\gamma_{N_i}\right)e\right). \qquad (59)$$

The result then follows since for all $\mathbf{X}$ with probability measure $\mu_{\mathbf{X}} \in \Lambda(\mathbf{P}, r)$,

$$I(\mathbf{X}; \mathbf{Y}) \leq \sum_{i=1}^{d} h(Y_i) - h(\mathbf{N}). \qquad (60)$$

As for the scalar case, the term $h(\mathbf{N})$ is not available in closed-form and must be numerically evaluated. In the numerical study in **Section 4.4**, $h(\mathbf{N})$ will be estimated via nearest neighbor methods.

4.3 Capacity Lower Bounds

We now generalize the results in **Section 3.2** to the case of vector channels. As for scalar channels, we consider the two cases: $1 \leq r < \alpha$; and $r = 2$.

Theorem 8. Let $1 < \alpha < 2$, $1 \leq r < \alpha$ and $\mathbf{P} \geq \mathbf{0}$. The capacity $C(\mathbf{P}, r)$ in (12) is lower bounded by

$$C(\mathbf{P}, r) \geq C_{LB}(\mathbf{P}, r)$$
$$= \frac{1}{\alpha}\sum_{i=1}^{d} \log\left(1 + \gamma_{X_i}^{\alpha}\right), \qquad (61)$$

where

$$\gamma_{X_i} = \left(\frac{P_i \alpha \sqrt{\pi}\, \Gamma\left(-\frac{r}{2}\right)}{2^{r+1}\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(-\frac{r}{\alpha}\right)\gamma_{N_i}^{r}}\right)^{\frac{1}{r}}. \qquad (62)$$

Proof. Let

$$\mathbf{A} = \text{diag}\left(\gamma_{X_1}, \ldots, \gamma_{X_d}\right)$$
$$\mathbf{A}^{\alpha} = \text{diag}\left(\gamma_{X_1}^{\alpha}, \ldots, \gamma_{X_d}^{\alpha}\right) \qquad (63)$$

and set $\mathbf{X} \stackrel{d}{=} \mathbf{A}\mathbf{N}$. We then have

$$\begin{aligned}
I(\mathbf{X}; \mathbf{Y}) &= h(\mathbf{Y}) - h(\mathbf{N}) \\
&= h(\mathbf{X} + \mathbf{N}) - h(\mathbf{N}) \\
&= h\left((\mathbf{A}^{\alpha} + \mathbf{I})^{\frac{1}{\alpha}}\mathbf{N}\right) - h(\mathbf{N}) \\
&= \log\det(\mathbf{A}^{\alpha} + \mathbf{I})^{\frac{1}{\alpha}} + h(\mathbf{N}) - h(\mathbf{N}) \\
&= \frac{1}{\alpha}\sum_{i=1}^{d}\log(1 + \gamma_{X_i}^{\alpha}).
\end{aligned} \qquad (64)$$

In order to ensure that $\mu_{\mathbf{X}} \in \Lambda_{\mathbf{X}}(\mathbf{P}, r)$, we recall (Shao and Nikias, 1993, Theorem 4)

$$\begin{aligned}
\mathbb{E}[|X_i|^{r}] &= \mathbb{E}[|\gamma_{X_i}N_i|^{r}] \\
&= \gamma_{X_i}^{r}\frac{2^{r+1}\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(-\frac{r}{\alpha}\right)}{\alpha\sqrt{\pi}\Gamma\left(\frac{-r}{2}\right)}\gamma_{N_i}^{r} \leq P_i.
\end{aligned} \qquad (65)$$

It then follows that

$$\gamma_{X_i} = \left(\frac{P_i \alpha \sqrt{\pi}\, \Gamma\left(-\frac{r}{2}\right)}{2^{r+1}\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(-\frac{r}{\alpha}\right)\gamma_{N_i}^{r}}\right)^{\frac{1}{r}}, \qquad (66)$$

as required.

Theorem 9. Let $1 < \alpha < 2$, $r = 2$, and $\mathbf{P} \geq \mathbf{0}$ For the channel (10) with $\mathbf{N}$ sub-Gaussian α-stable with underlying random vector $\mathbf{G} \sim \mathcal{N}(\mathbf{0}, \Sigma)$, the capacity $C(\mathbf{P}, 2)$ in (12) is lower bounded by

$$C(\mathbf{P}, 2) \geq C_{LB}(\mathbf{P}, 2)$$
$$= \max\left\{0, \int_0^{\infty}\frac{1}{2}p_A(a)\log\det\left(2\pi e\left(\Sigma_{\mathbf{X}} + a\Sigma\right)\right)da - h(\mathbf{N})\right\}, \qquad (67)$$

where $\Sigma_{\mathbf{X}}$ is any positive definite matrix with diagonal elements

$$\Sigma_{\mathbf{X},ii} = P_i, \quad i = 1, \ldots, d \qquad (68)$$

and $A \sim S_{\alpha/2}(\cos^{2/\alpha}(\frac{\pi}{4}\alpha), 1, 0)$.

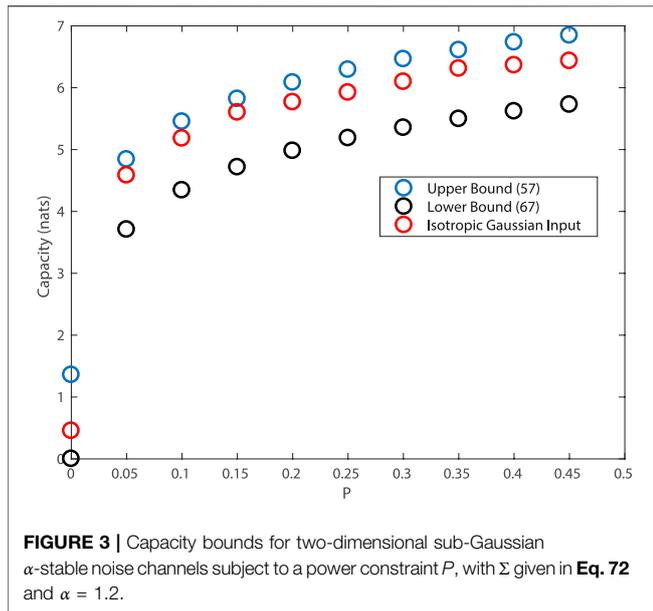Proof. Let $X \sim \mathcal{N}(0, \Sigma_{\mathbf{X}})$ and $G \sim \mathcal{N}(0, \Sigma)$. Then,

FIGURE 3 | Capacity bounds for two-dimensional sub-Gaussian α-stable noise channels subject to a power constraint P, with Σ given in **Eq. 72** and $\alpha = 1.2$.

FIGURE 4 | Capacity bounds for two-dimensional sub-Gaussian α-stable noise channels for varying noise dependence ρ subject to a power constraint $p = 0.01$, with Σ given in **Eq. 73** and $\alpha = 1.2$.

$$
\begin{aligned}
h(\mathbf{Y}) &= h(\mathbf{X} + \mathbf{N}) \\
&= h(\mathbf{X} + A^{1/2}\mathbf{G}) \\
&\geq h(\mathbf{X} + A^{1/2}\mathbf{G} | A) \\
&= h(\mathbf{G}_A | A),
\end{aligned}
\tag{69}
$$

where $\mathbf{G}_A \sim \mathcal{N}(0, \Sigma_\mathbf{X} + A\Sigma)$. Hence,

$$
h(\mathbf{Y}) \geq \int_0^\infty \frac{1}{2} p_A(a) \log \det(2\pi e (\Sigma_\mathbf{X} + a\Sigma)) \mathrm{d}a.
\tag{70}
$$

In order to satisfy the power constraints, we require

$$
\Sigma_{\mathbf{X},ii} \leq P_i, \quad i = 1, \ldots, d,
\tag{71}
$$

which yields the desired result.

4.4 Numerical Results

In this section, we study the behavior of the bounds in the case of two-dimensional sub-Gaussian α-stable noise, where inputs are subject to a power constraint.

Figure 3 plots the capacity bounds in the previous section for varying values of P in the presence of sub-Gaussian α-stable noise, with $\alpha = 1.2$ and

$$
\Sigma = 2 \cdot 0.01^2 \cdot \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}.
\tag{72}
$$

In order to compute the entropy $h(\mathbf{N})$, the 1-nearest neighbor method (Berrett et al., 2019) was used. Observe there is roughly a gap of approximately one nat between the upper bound in Theorem 7 and the lower bound in Theorem 9 with $\Sigma_\mathbf{X}$ chosen to proportional to Σ.

The third curve in red corresponds to the case of a two-dimensional isotropic Gaussian input where each component has variance P. Observe that the mutual information obtained with this input is close to the upper bound. This suggests that for sub-Gaussian α-stable noise channels, Gaussian inputs
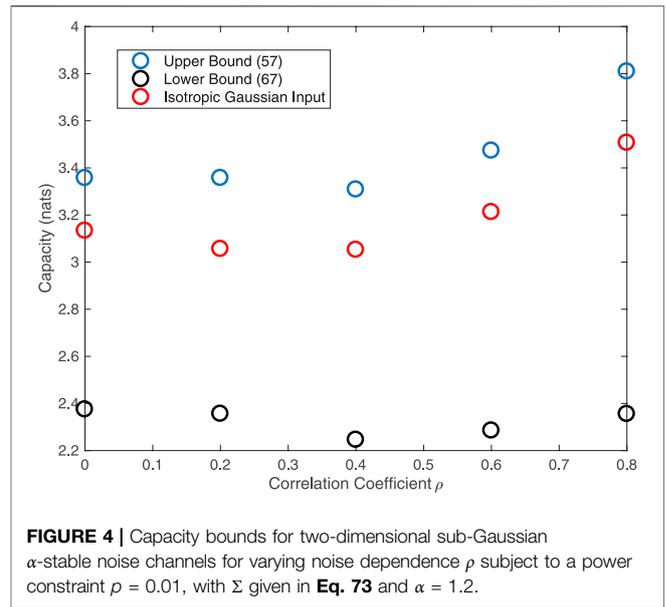
perform well and, moreover, independent components are desirable. This can be understood by an inspection of Theorem 9, where choosing $\Sigma_\mathbf{X}$ to be diagonal maximizes the determinant when $\Sigma = \mathbf{0}$.

Figure 4 plots the capacity bounds in the previous section subject to a power constraint $p = 0.01$ for varying values of ρ in the presence of sub-Gaussian α-stable noise, with $\alpha = 1.2$ and

$$
\Sigma = 2 \cdot 0.01^2 \cdot \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.
\tag{73}
$$

The results are consistent with **Figure 3**, with the isotropic Gaussian input performing well for all values of ρ. We also observe that the curves also increase for sufficient large values of ρ, suggesting that increasing the dependence can lead to performance improvements. This is relevant for communication systems, such as in (Zheng et al., 2019, 2020), where noise is dominated by interference, which may be modified via changes to access policies.

5 CONCLUSION

Multivariate α-stable models have been suggested to capture the heavy-tailed nature of interference in OFDM-based wireless communication systems. In this paper, we studied the capacity of fractional moment and power constrained signaling in the presence of such noise. By considering non-isotropic inputs, we obtained upper and lower bounds, which provide insights into the behavior of the capacity and its relation to Gaussian noise models. *Via* a numerical study in two-dimensional channels with sub-Gaussian α-stable noise, we compared the performance of isotropic and non-isotropic Gaussian inputs. This suggests, at least for this special case, isotropic Gaussian inputs remain a desirable choice.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

ME formulated the problem, carried out the mathematical analysis and numerical study, and wrote the paper.

REFERENCES

Arimoto, S. (1972). An Algorithm for Computing the Capacity of Arbitrary Discrete Memoryless Channels. *IEEE Trans. Inform. Theor.* 18, 14–20. doi:10.1109/tit.1972.1054753

Berrett, T., Samworth, R., and Yuan, M. (2019). Efficient Multivariate Entropy Estimation via *K*-Nearest Neighbour Distances. *Ann. Stat.* 47, 288–318. doi:10.1214/18-aos1688

Billingsley, P. (1999). *Convergence of Probability Measures*. John Wiley & Sons.

Blahut, R. (1972). Computation of Channel Capacity and Rate-Distortion Functions. *IEEE Trans. Inform. Theor.* 18, 460–473. doi:10.1109/tit.1972.1054855

Byczkowski, T., Nolan, J. P., and Rajput, B. (1993). Approximation of Multidimensional Stable Densities. *J. Multivariate Anal.* 46, 13–31. doi:10.1006/jmva.1993.1044

Chitre, M., Potter, J., and Heng, O. (2004). "Underwater Acoustic Channel Characterisation for Medium-Range Shallow Water Communications," in *MTS/IEEE Techno-Ocean'04* (IEEE).

Clavier, L., Pedersen, T., Larrad, I., Lauridsen, M., and Egan, M. (2021b). Experimental Evidence for Heavy Tailed Interference in the IoT. *IEEE Commun. Lett.* 25, 692–695. doi:10.1109/lcomm.2020.3034430

Clavier, L., Peters, G., Septier, F., and Nevat, I. (2021a). Impulsive Noise Modeling and Robust Receiver Design. *EURASIP J. Wireless Commun. Networking* 13, 1. doi:10.1186/s13638-020-01868-1

Cover, T., and Thomas, J. (2006). *Elements of Information Theory*. Second Edition. Hoboken, NJ: John Wiley & Sons.

de Freitas, M. L., Egan, M., Clavier, L., Goupil, A., Peters, G. W., and Azzaoui, N. (2017). Capacity Bounds for Additive Symmetric α -Stable Noise Channels. *IEEE Trans. Inform. Theor.* 63, 5115–5123. doi:10.1109/tit.2017.2676104

Dytso, A., Bustin, R., Vincent Poor, H., and Shamai, S. (2018). Analytical Properties of Generalized Gaussian Distributions. *J. Stat. Distributions Appl.* 5, 1–40. doi:10.1186/s40488-018-0088-5

Egan, M., Clavier, L., Zheng, C., de Freitas, M., and Gorce, J.-M. (2018). Dynamic Interference for Uplink SCMA in Large-Scale Wireless Networks without Coordination. *EURASIP J. Wireless Commun. Networking* 213, 1. doi:10.1186/s13638-018-1225-z

Egan, M. (2019). "On Capacity Sensitivity in Additive Vector Symmetric *α*-stable Noise Channels," in Proc. IEEE Wireless Communications and Networking Conference Workshop (WCNCW) (IEEE). doi:10.1109/wcncw.2019.8902901

Fahs, J., and Abou-Faycal, I. (2017). On Properties of the Support of Capacity-Achieving Distributions for Additive Noise Channel Models with Input Cost Constraints. *IEEE Trans. Inf. Theor.* 64, 1178–1198. doi:10.1109/TIT.2017.2771815

Fahs, J., and Abou-Faycal, I. (2016). On the Finiteness of the Capacity of Continuous Channels. *IEEE Trans. Commun.* 64, 166–173. doi:10.1109/tcomm.2015.2503403

Farsad, N., Guo, W., Chae, C.-B., and Eckford, A. (2015). "Stable Distributions as Noise Models for Molecular Communication," in Proc. IEEE Global Communications Conference (IEEE). doi:10.1109/glocom.2015.7417583

Freitas, M., Egan, M., Clavier, L., Savard, A., and Gorce, J.-M. (2018). "Power Control in Parallel Symmetric *α*-stable Noise Channels," in IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC) (IEEE).

Ghannudi, H. E., Clavier, L., Azzaoui, N., Septier, F., and Rolland, P.-a. (2010). α-Stable Interference Modeling and Cauchy Receiver for an IR-UWB Ad Hoc Network. *IEEE Trans. Commun.* 58, 1748–1757. doi:10.1109/tcomm.2010.06.090074

Gu, W., and Clavier, L. (2012). Decoding Metric Study for Turbo Codes in Very Impulsive Environment. *IEEE Commun. Lett.* 16, 256–258. doi:10.1109/lcomm.2011.112311.111504

Gulati, K., Evans, B. L., Andrews, J. G., and Tinsley, K. R. (2010). Statistics of Co-channel Interference in a Field of Poisson and Poisson-Poisson Clustered Interferers. *IEEE Trans. Signal. Process.* 58, 6207–6222. doi:10.1109/tsp.2010.2072922

Hall, H. (1966). *A New Model for Impulsive Phenomena: Application to Atmospheric-Noise Communication Channels*. Stanford, CA: Tech. Rep. 3412-8, Stanford Electronics Labs, Stanford University.

Han, T. (2003). *Information-Spectrum Methods in Information Theory*. Berlin Heidelberg: Springer-Verlag.

Ilow, J., and Hatzinakos, D. (1998). Analytic Alpha-Stable Noise Modeling in a Poisson Field of Interferers or Scatterers. *IEEE Trans. Signal. Process.* 46, 1601–1611. doi:10.1109/78.678475

Kuruoglu, E. E. (2001). Density Parameter Estimation of Skewed α-stable Distributions. *IEEE Trans. Signal. Process.* 49, 2192–2201. doi:10.1109/78.950775

Lauridsen, M., Vejlgaard, B., Kovács, I., Nguyen, H., and Mogensen, P. (2017). "Interference Measurements in the European 868 MHz ISM Band with Focus on LoRa and SigFox," in IEEE Wireless Communications and Networking Conference (WCNC) (IEEE). doi:10.1109/wcnc.2017.7925650

Mahmood, A., Chitre, M., and Armand, M. A. (2014). On Single-Carrier Communication in Additive white Symmetric Alpha-Stable Noise. *IEEE Trans. Commun.* 62, 3584–3599. doi:10.1109/tcomm.2014.2351819

Marshall, A. W. (1984). Markov's Inequality for Random Variables Taking Values in a Linear Topological Space. *Inequalities Stat. Probab.* 5, 104–108. doi:10.1214/lnms/1215465634

Mestrah, Y., Savard, A., Goupil, A., Gellé, G., and Clavier, L. (2020). An Unsupervised Llr Estimation with Unknown Noise Distribution. *EURASIP J. Wireless Commun. Networking* 26, 1. doi:10.1186/s13638-019-1608-9

Middleton, D. (1977). Statistical-physical Models of Electromagnetic Interference. *IEEE Trans. Electromagn. Compat.* EMC-19, 106–127. doi:10.1109/temc.1977.303527

Niranjayan, S., and Beaulieu, N. (2009). The BER Optimal Linear Rake Receiver for Signal Detection in Symmetric Alpha-Stable Noise. *IEEE Trans. Commun.* 57, 3585–3588. doi:10.1109/tcomm.2009.12.0701392

Pinto, P. C., and Win, M. Z. (2010). Communication in a Poisson Field of Interferers--Part I: Interference Distribution and Error Probability. *IEEE Trans. Wireless Commun.* 9, 2176–2186. doi:10.1109/twc.2010.07.060438

Samorodnitsky, G., and Taqqu, M. (1994). *Stable Non-gaussian Random Processes*. New York, NY: CRC Press.

Shao, M., and Nikias, C. L. (1993). Signal Processing with Fractional Lower Order Moments: Stable Processes and Their Applications. *Proc. IEEE* 81, 986–1010. doi:10.1109/5.231338

Sousa, E. S. (1992). Performance of a Spread Spectrum Packet Radio Network Link in a Poisson Field of Interferers. *IEEE Trans. Inform. Theor.* 38, 1743–1754. doi:10.1109/18.165447

Zheng, C., Egan, M., Clavier, L., Pedersen, T., and Gorce, J.-M. (2020). "Linear Combining in Dependent α-stable Interference," in IEEE International Conference on Communications (ICC) (IEEE). doi:10.1109/icc40277.2020.9148724

Zheng, C., Egan, M., Clavier, L., Peters, G., and Gorce, J.-M. (2019). "Copula-based Interference Models for IoT Wireless Networks," in Proc. IEEE International Conference on Communications (ICC) (IEEE). doi:10.1109/icc.2019.8761783

Zimmermann, M., and Dostert, K. (2002). Analysis and Modeling of Impulsive Noise in Broad-Band Powerline Communications. *IEEE Trans. Electromagn. Compat.* 44, 249–258. doi:10.1109/15.990732

Zolotarev, V. M. (1957). Mellin-stieltjes Transforms in Probability Theory. *Theor. Probab. Appl.* 2, 433–460. doi:10.1137/1102031

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.