UNIVERSITÉ LILLE 1 - SCIENCES ET TECHNOLOGIES

ÉCOLE DOCTORALE ED RÉGIONALE SPI 72

IEMN - UMR CNRS 8520

Wireless Communications in Dynamic Interference - modeling, capacity and applications

Author : Mauro LOPES DE FREITAS

Présentée et soutenue publiquement le 13/06/2018
pour l'obtenir le titre de ***Docteur***
*de l'*Université Lille 1 - Sciences et Technologies
Spécialité Micro et nano technologies, acoustique et télécommunications. Numéro d'ordre : 9999

Composition du jury (prévu):

Rapporteurs :	**Philippe Ciblat** Professeur, Telecom ParisTech et Université Paris-Saclay, France
	Marco Di Renzo Chargé de recherche CNRS, L2S, Centrale Supélec, France
Examinateurs :	**Mérouane Debbah** Directeur laboratoire R&D en mathématiques et algorithmes, Huawei Technologies, Professeur, Centrale Supélec, France
	Gareth W. Peters Chair Prof. in Statistics for Risk and Insurance, Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, UK
	Michèle Wigger Maître de Conférence, Telecom ParisTech, France
	Atika Rivenq Professeur, Université de Valenciennes et du Hainaut-Cambrésis, France
Directeur de thèse :	**Laurent Clavier** Professeur, IMT Lille Douai, France
Co-encadrant :	**Malcolm Egan** Chercheur Associé, INSA de Lyon, France

To my family and friends.

"Science is but a perversion of itself unless it has as its ultimate goal the betterment of humanity."

Nikola Tesla

"Prediction is very difficult, especially if it's about the future."

Niels Bohr

"Magic is just science that we don't understand yet."

Arthur C. Clarke

Acknowledgements

I would first like to express my special appreciation and thanks to my supervisor Professor Laurent CLAVIER for his encouragement throughout my entire candidature, who has been a mentor for me. I also would like to thank my advisor Malcolm Egan for his patience, and advices. This thesis would not have been possible without their support, technical expertise and willingness to share the knowledge they have gained in research.

Besides my advisors, I would also like to thank my committee members, professor Philippe CIBLAT, Chargé de recherche Marco DI RENZO, professor Mérouane DEBBAH, maître de conferénce Michèle WIGGER and professor Atika RIVENQ for serving as my committee members.

I want to thank all the people who have cooperated with me: Nourddine AZZAOUI, Alban GOUPIL, Gareth W. PETERS, Louis DORVILLE, Jean-Marie GORCE and Anne SAVARD.

I thank all the members of the CSAM group and people I have known in IRCICA: Nathalie ROLLAND, Christophe LOYEZ, Rédha KASSI, Bernard VERBEKE, Kamel GUERCHOUCHE, Roman IGUAL, Aymeric PASTRE, Viktor TOLDOV, Yan XIN, Ce ZHENG, Peggy STANKOWSKI, Michel SOULAGE, Nicolas DE ARAUJO MOREIRA, Umber NOREEN, Viktor TOLDOV, Roman IGUAL-PEREZ, Rahul VYAS, Laura GUÉRIN and etc.

My family is what I want to thank most for their unfailing emotional support. Words cannot express how grateful I am to my mother, father and Tatá for all of the sacrifices that they have made for supporting my study in foreign country. There is no way to express how much it meant to me the support of Izabel DE OLIVEIRA FERNANDES, my wife and love of my life. She has been my best friend and a great companion.

This work would not be possible without the support, encourage, and help of my previous teachers and supervisors Profs. Waldir SABINO, Eddie BATISTA and Wallace MARTINS. Finally, I thank Maria CONSOLAÇÃO SAID for showing me the importance of the curiosity to learn.

Abstract

Title: ***Wireless Communications in Dynamic Interference - modeling, capacity and applications***

This thesis focuses on the study of noise and interference exhibiting an impulsive behavior, an attribute that can be found in many contexts such as wireless communications or molecular communications. This interference is characterized by the presence of high amplitudes during short durations, an effect that is not well represented by the classical Gaussian model. In fact, these undesirable features lead to heavier tails in the distributions and can be modeled by the α-stable distribution. In particular, we study the impulsive behavior that occurs in large-scale communication networks that forms the basis for our model of dynamic interference. More precisely, such interference can be encountered in heterogeneous networks with short packets to be transmitted, as in the Internet of Things, when the set of active interferers varies rapidly.

The first part of this work is to study the capacity of α-stable additive noise channels, which is not well understood at present, except in the case of Cauchy noise ($\alpha = 1$) with a logarithmic constraint and Gaussian noise ($\alpha = 2$) with a power constraint. We derive lower and upper bounds for the capacity with an absolute moment (amplitude) constraint. We consider additive symmetric α-stable noise channels with $\alpha \in \,]\,1, 2\,]$. We then use an algorithm inspired by the Blahut-Arimoto algorithm in order to compare our bounds with a numerical approximation, which provides insight into the effect of noise parameters on the bounds. In particular, we find that our lower bound is in good agreement with the numerical approximation for α close to 2. We then extend the work to the capacity of the additive complex isotropic α-stable noise channel.

The second part consists in analyzing the impact of our bounds in practical contexts. We first study the case of parallel α-stable additive noise channels and provide insight into the effect of the index α on the achievable rate. We develop a new power allocation algorithm and show that our algorithm can significantly improve achievable rates over standard approaches assuming Gaussian noise. We then analyze the effect of slow fading. Finally, we derive the area spectral efficiency, i.e., the total rate per square meter. Our analysis suggests that, similar to the conventional Gaussian model, dense networks maximize the area spectral efficiency.

Résumé

Titre: ***Communications sans fil dans des interférences dynamiques - modélisation, capacité et applications***

Cette thèse se concentre sur l'étude du bruit et des interférences présentant un comportement impulsif, un attribut que l'on peut retrouver dans de nombreux contextes comme les communications sans fil ou les communications moléculaires. Cette interférence est caractérisée par la présence d'amplitudes élevées pendant des durées courtes, effet qui n'est pas bien représenté par le modèle gaussien classique. En fait, ces caractéristiques indésirables conduisent à des queues de distributions plus lourdes qui peuvent être modélisées par la distribution α-stable. En particulier, nous étudions le comportement impulsif qui se produit dans les réseaux de communication à grande échelle qui forme la base de notre modèle d'interférence dynamique. Plus précisément, une telle interférence peut se rencontrer dans des réseaux hétérogènes avec des paquets courts à transmettre, comme dans l'Internet des objets, lorsque l'ensemble des interférents actifs varie rapidement.

La première partie de ce travail est d'étudier la capacité des canaux de bruit α-stable, qui n'est pas bien comprise actuellement, sauf dans le cas du bruit de Cauchy ($\alpha = 1$) avec une contrainte logarithmique et du bruit gaussien ($\alpha = 2$) avec une contrainte de puissance. Nous calculons des bornes inférieures et supérieures pour la capacité avec une contrainte de moment de la valeur absolue (amplitude). Nous considérons les canaux à bruit symétrique additif α-stable avec $\alpha \in \,]\,1, 2\,]$. Nous utilisons ensuite un algorithme inspiré du Blahut-Arimoto afin de comparer les bornes proposées avec une approximation numérique, ce qui permet en particulier d'évaluer l'effet des paramètres de bruit sur les bornes. En particulier, nous trouvons que notre borne inférieure est en bon accord avec l'approximation numérique pour α proche de 2. Nous étendons ensuite le travail à la capacité de canaux à bruit additif complexe, isotrope α-stable.

La deuxième partie consiste à analyser l'impact de nos limites dans des contextes pratiques. Nous étudions d'abord le cas des canaux parallèles à bruit additif α-stable et donnons un aperçu de l'effet de l'indice α sur le débit atteignable. Nous développons un nouvel algorithme d'allocation de puissance et montrons que notre algorithme peut améliorer significativement débit atteignable en comparaison des approches standards qui supposent un bruit gaussien. Nous analysons ensuite l'effet des évanouissements lents. Enfin, nous obtenons l'efficacité spectrale par unité de surface,

c'est à dire le débit total par mètre carré. Notre analyse suggère que, de manière analogue au modèle gaussien conventionnel, les réseaux denses maximisent l'efficacité spectrale par unité de surface.

Contents

List of Figures

List of Tables

List of Abbreviations

ASαSN	Additive Symmetric α-stable noise
AIαSN	Additive Isotropic α-stable noise
AWCN	Additive white class-A noise
AWGN	Additive White Gaussian Noise
BAN	Body-Area Network
BS	Base Station
CLT	Central Limit Theorem
CPS	Cyber Physical System
CSI	Channel State Information
DSL	Digital Subscriber Line
D2D	Device to Device
H2H	Human-to-Human communication
i.i.d.	independent and identically distributed
IoT	Internet of Things
ISM	Industrial, scientific, and medical
LoRa	Long Range
MIMO	Multiple Input Multiple Output
M2M	Machine to Machine
NOMA	Non-orthogonal multiple access
OFDM	Orthogonal Frequency-Division Multiplexing
PLC	Power Line Communication
PPP	Poisson Point Process
r.v.	Random variable
SCMA	Sparse Code Multiple Access
SIR	Signal-to-interference ratio
SNR	Signal-to-noise ratio
WSN	Wireless Sensor Networks

Notation

Symbol	Description
i	Imaginary unit, $\sqrt{-1}$
$(\cdot)^T$	Matrix transpose
$\mathbf{X}$	Vector or matrix $\mathbf{X}$
$\mathbf{x} \preceq \mathbf{c}$	Each component of $\mathbf{x}$ satisfies $x_i \geq c_i$
$\mathbb{R}, \mathbb{C}, \mathbb{N}, \mathbb{Z}$	Set of Real, Complex, Natural and Integer numbers
$\mathbb{R}^n$	n-dimensional real vector space
$\mathbb{R}_{>0}$	$\{x \in \mathbb{R} \vert x > 0\}$
$\mathbb{R}_{\geq 0}$	$\{x \in \mathbb{R} \vert x \geq 0\}$
$\mathbb{R}_{<0}$	$\{x \in \mathbb{R} \vert x < 0\}$
$\mathbb{R}_{\leq 0}$	$\{x \in \mathbb{R} \vert x \leq 0\}$
$\mathbb{R}_{\neq 0}$	$\{x \in \mathbb{R} \vert x \neq 0\}$
$\mathbb{E}[\cdot]$	Expectation
$\mathbb{E}_{\mu}[\cdot]$	Expectation with respect to μ
$\Pr(\cdot), P[\cdot]$	Probability
$\xrightarrow[n\to\infty]{d}$	Converge in distribution when $n \to \infty$
$\overset{d}{=}$	Equality in distribution
$\{X_j\}_{j\in\mathbb{N}}$	Sequence in $\mathbb{N}$.
$\{X_j\}_{j=1}^{\infty}$	Sequence starting at $j = 1$ and ending at ∞
$\sim$	Has the probability distribution of
$\Gamma(\cdot)$	Gamma function
$\triangleq$	Defined as
$\vert \cdot \vert$	Absolute value
$\vert\mathbf{x}\vert$	Euclidean norm
$\ll$	Absolute continuous
$\lessapprox$	Asymptotically less than
$\circ$	Hadamard product
$\otimes$	Product measure
$\lesssim$	Approximately less than
$\Rightarrow$	Weak convergence
$\mathcal{CN}(\mu, \sigma^2)$	Circularly symmetric complex normal distribution with mean μ and variance σ^2
$\exp\{\cdot\}$	Exponential function
$\overset{a.s.}{\to}$	Converges almost surely to
$\Vert \cdot \Vert_{\text{TV}}$	Total variation

Introduction

TELECOMMUNICATION companies are struggling to coordinate limited resources, in the form of spectrum and time slots, with increasing number of users and data in modern wireless communications. This takes place because traditionally the data rate for voice is around 10 kbps [APY16], however, the new data requirements are some orders of magnitude larger. For instance, 4G technologies, such as WiMax and LTE-A are encountering problems to accommodate the amount of data [EKR14; AQM14].

Many alternatives have been proposed to overtake the data communication limitations. For instance, a new architecture of small cells has gained attention, which encompass femtocells, picocells, and microcells [HM12; WSKTEK15], that reduces the area where a Base Station (BS) is providing communication in order to support the modern capacity-hungry devices. Unfortunately, the area densification due to closer users increases the interference for all receivers present, limiting the improvement of these techniques.

To take advantage of the small cells, Device to Device (D2D) communication has emerged to cope with emerging data-hungry devices. It consists in a link between two users without the use of a Base Station, differently from the usual system in which all transmissions are forced to use the core network. Although, a limitation of this architecture remains on the range that devices can communicate with each other but keeping their sizes small and saving energy consumption. In the end, D2D also suffers from interference due to the high concentration of mobile users.

An important new paradigm arising that can be mentioned is the noteworthy Internet of Things (IoT). Multiple IoT incarnations examples are reviewed below [WCXDZ13; HC17]:

- **Machine to Machine (M2M)**: Generally, it is the counterpart of a Human-to-Human communication (H2H). In other words, it refers to a data communication with limited or without human intervention. Examples are smart sensors, computers and mobile devices. Basically this architecture has three principles, firstly, the amount of time to create smart and ubiquitous services is facilitated. Secondly, more autonomous applications can be achieved. Thirdly, the quantity of machines enables the use of a networked machine, which is more valuable than an isolated one.

- **Cyber Physical System (CPS)**: It is defined as collaborative computational elements controlling physical entities. M2M systems that have decision-make and autonomous control can be upgraded to CPS. They share the knowledge to control logistics and production systems, in contrast to traditional embedded systems that use standalone devices. E.g. sensor, communications tools and smart grids.

- **Wireless Sensor Networks (WSN)**: It is composed by distributed autonomous sensors and are the basic scenario of IoT. They are responsible for monitoring physical conditions and, for instance, creating a mesh topology in order to pass information in a cooperative way to a core network.

- **Body-Area Network (BAN)**: It appears due to the new demands on lightweight, small-size and ultra-low-power to monitor the human's physiology and actions.

Although the quantity of IoT devices is far from the stratospheric number of 50 billion by 2020 that Ericsson's former CEO Hans Vestburg [Nor16] predicted, 8.4 billion connected devices are estimated in 2017. In a more realistic view, 20.8 billions devices are believed to exist by 2020 [Gar17], not counting smartphones, tablets and computers. A number of nodes which increases significantly, orders of magnitude more than H2H, in a manner that, without any doubt, converges towards a new design where device-centric communication horn in human-centric communication.

Besides cellular networks, IoT devices are deployed in the license free industrial, scientific, and medical (ISM) bands. Long Range (LoRa) WAN [SWH17; TMIPWMC16] and Sigfox [NGK16] are two of the most common IoT connectivity technologies being deployed in the ISM band. The bands available are however not numerous and, unfortunately, they are shared by many technologies operating on the same or overlapping frequency spectrum. For instance at 2.4 GHz, we can find the standards 802.11b (Wi-Fi) [ILII00], 802.15.1 (Bluetooth) [GXR09] and 802.15.4 (e.g. ZigBee [Far08] and 6LoWPAN [Amm14]), resulting in a congested band. Alternatively, the millimeter-wave bands have gained attention, nonetheless, it is suitable to short-range and high-speed communications, restricting the usage for WSN devices. The strong potential of the IoT technology is emphasized due to the new use of located information and the expansion for spontaneous transmissions. However, it is important to notice that the concentration of devices poses a challenge for existing interference mitigation techniques.

From previous examples it is clear that the interference will become one of the main limitations to the systems performance and it is increasing the importance of interference models for the near future of wireless communications.

1.1 Focus of the Thesis and Overview of Contributions

In this thesis, we assess the impact of rapidly changing active transmitter sets—or *dynamic interference*—in large-scale, e.g., M2M. We consider a scenario where the network is has limited or no coordination between the devices. This setup is relevant for networks supporting the internet of things and in large-scale sensor networks, where transmitting devices are very simple and have limited ability to coordinate. We also assume that the active set of transmitters varies symbol-by-symbol, which contrasts with the Gaussian model where the active transmitter set is fixed. We show in this thesis that an impulsive interference is present, as such, the Gaussian model is not recommended anymore. However, little is known about achievable rates and optimal inputs in this scenario and the capacity of impulsive noise channels can significantly differ from the capacity of Gaussian noise channels.

Our thesis focuses on the challenges at the physical layer to better understand interference caused by simultaneous transmissions, which is a fundamental feature present in performance degradations to future systems. We assume that interference is modeled by an α-stable distribution. Such heavy-tailed distributions can be seen as an extension of the Gaussian distribution, a member of the family. Indeed, they share the stability property meaning that the sum of α-stable random variable (with the same α) is an α-stable random variable. This class of distribution allows heavy tails that are well suited to model rare events and, consequently, impulsive noise. With such an interference assumption and assuming an interference limited setting, i.e. the interference dominates the noise floor, we replace the traditional additive white Gaussian noise (AWGN) channel with the additive symmetric (or isotropic in the complex case) α-stable noise channel (ASαSN and AIαSN channels in the real and the complex cases respectively). We then address the question: what is the capacity of such a channel and what are the consequences for wireless communications?

1.2 Contributions of this thesis

- We derive lower and upper bounds for the $AS\alpha SN$ channel with $\alpha \in (1, 2]$ and an absolute moment constraint.
  - We numerically approximate the capacity via the Blahut-Arimoto algorithm. This algorithm requires truncation and discretization of the support of the noise distribution. We provide guidance for choices of step and support sizes to yield a good approximation.

  - We compare our bounds and the numerical approximation to gain insight into the effect of noise parameters. In particular, we show that the lower bound is a good approximation of the capacity obtained from the Blahut-Arimoto algorithm for α near 2.
  - We prove the existence and uniqueness for the optimal input with fractional moment constraints.
  - We study the medium SIR lower bound behavior for the achievable lower bound.
  - We also show the importance of the parametrization used for the input distribution, in which the lower bound results have different behavior if the terms are written in function of γ_X or $\mathbb{E}[|X|]$.

- We derive a complex extension of lower bounds for the $AIS\alpha SN$ with $\alpha < 2$ case with absolute moment constraints. Moreover, we prove the existence and uniqueness for the complex case. We also compare the behavior of achievable rates varying the α parameter, the Gaussian case ($\alpha = 2$) being the basis of comparison.

- We analyze the impact of dynamic interference on the performance of wireless communication systems:

  - We study a relaxation of amplitude and fractional moment constraints applied to parallel $AI\alpha SN$ channels and the approximation of truncated α-stable when fractional moment constraint is applied.
  - We derive an outage probability upper bound when slow fading and α-stable noise are present.
  - We study our achievable rates for the $AI\alpha SN$ channel subjected to a power constraint. We compare Gaussian and α-stable inputs, as well as their truncated versions to solve a power allocation problem in parallel symmetric α-stable noise channels.
  - We study the expected total rate per square meter by means of the effect of device density on the network performance.

1.3 Scientific production

This section summarizes the publications, conferences and collaborations based on work that was done during the period of Ph.D. candidature.

Journal

- M. de Freitas, M. Egan, L. Clavier, A. Goupil, G. W. Peters, and N. Azzaoui. "Capacity Bounds for Additive Symmetric α-Stable Noise Channels". In: *IEEE Transactions on Information Theory* (2017)
- M. Egan, L. Clavier, C. Zheng, M. de Freitas, J. Gorce. "Dynamic Interference in Uplink SCMA for Large-Scale Wireless Networks without Coordination". In: *EURASIP Journal on Wireless Communications and Networking* (2018 - under review)

Letter

- M. de Freitas, M. Egan, L. Clavier, A. Savard, and J. Gorce. "Power Control in Parallel Symmetric α-Stable Noise Channels". In: *IEEE Communications Letters* (2017 - under review)

Conference papers

- M. de Freitas, M. Egan, and L. Clavier. "Study of achievable Rates for Additive Symmetric α-Stable Noise Channels". In: *Gretsi* (2017)
- M. Egan, M. de Freitas, L. Clavier, A. Goupil, G. Peters, and N. Azzaoui. "Achievable rates for additive isotropic α-stable noise channels". In: *IEEE International Symposium on Information Theory*. 2016
- M. Egan, L. Clavier, M. de Freitas, L. Dorville, J.-M. Gorce, and A. Savard. "Wireless communication in dynamic interference". In: *IEEE Globecom* (2017 - Accepted for publication)

Workshop/others

- PhD Student Pitch Contest, In: *Journée doctorants IRCICA: Ma thèse en 180 secondes, IRCICA*, April 15, 2015
- Malcolm Egan, Mauro de Freitas, Laurent Clavier, Alban Goupil, Gareth W. Peters and Nourddine Azzaoui, "Wireless Network Design with Dynamic Interference", In: *Cost action CA15104, TD(16)01067, IRACON, 2nd MC meeting and first technical meeting*, Lille, France, 30 May-1 June, 2016
- Malcolm Egan, Laurent Clavier, Mauro de Freitas and Louis Dorville, "Communication in Dynamic Interference", In: *Cost action CA15104 IRACON, 4th MC meeting and 4th technical meeting*, Lund, Sweden, 29-31 May, 2017
- PhD Student Pitch Contest, In: *14th International Symposium on Wireless Communication Systems - ISWCS*, 28-31 August, 2017

Not related publications

Journal

- M. L. de Freitas, W. A. Martins, E. B. de Lima Filho, and W. S. da Silva Júnior. "New Designs for Reduced-Redundancy Transceivers". In: *Circuits, Systems, and Signal Processing* 36.5 (2017), pp. 2075–2101

Conference paper

- D. P. de Mello, M. L. de Freitas, L. C. Cordeiro, W. S. Júnior, I. V. de Bessa, B. Eddie Filho, and L. Clavier. "Verification of Magnitude and Phase Responses in Fixed-Point Digital Filters". In: *XXXV Simpósio Brasileiro de Telecomunicações e processamento de sinais - SBRT* (2017), pp. 1184–1188

Chapter 2

Impulsive interference and α-stable processes

This chapter introduces the dynamic interference characterization. For that purpose, noise and interference definitions are presented and principal models are reviewed. In particular, α-stable model and its properties are studied in univariate and multivariate cases. This is the model that we are going to consider in the rest of this manuscript.

2.1 Impulsive interference models

COMMUNICATION systems are limited by noise and interference, from internal and external sources. Hence, the study of models has gained importance and both theoretical and practical research have been developed. Noise is by definition an unwanted signal involving unpredictable perturbations that degrade the desired information or measurement. In order to facilitate the understanding, its source can be separated in many categories [Vas00], including:

- Acoustic noise - a disturbance in the audio frequency range, which may arise from moving objects, wind or rain, for instance.
- Electronic noise - examples are thermal noise, which has its origin from the random thermal motion of electrons, and the shot noise.
- Electrostatic noise - generated when a voltage is present and with or without current flow, such as, fluorescent lighting.
- Quantization noise - due to the package loss caused by a network congestion.
- Communication channel - distortion and fading, as consequence of non-ideal characteristics of communication channels.

In turn, interference refers to the unwanted signals added to the useful one, producing a crucial impact in wireless communications. It is formed when multiple uncoordinated links share a common communication channel. One scenario that experiences interference can be observed when a

communication contains N principal links and start transmitting their separate information to N receivers on the same medium, causing $N(N-1)$ interference links. Thus, it differs from thermal noise by being a special case of artificial noise generated by other signals.

In the context of cellular communication, the interference appears due to the technique of reusing frequencies and channels, known as frequency planning, which is common to improve efficiency, both for spectral and capacity. In turn, interference arises when access points share a common channel in local area networks.

The traditional way to limit the impact of interference consists in avoiding or mitigating interferences, in order, for instance, to improve some features such as capacity and coverage in heterogeneous networks that contains a large number of uncoordinated low-power nodes [CAG08; Gol05; CLCC11]. Convenient techniques have been created to exclude their effects. We can cite approaches like interference alignments at the Physical layer level to create systems without interference [EAPH13]. We can also give the example of Interference Cancellation [WAYDV07; And05]: we have several users that collides in a resource block but they are not seen as interference to the other signals but a simultaneous decoding of all the signals is implemented. We can also cite all the efforts put in the Medium Access Control layer like carrier sensing [YV05; JHMB05] in order to avoid simultaneous transmissions and, consequently the system performance, or in cellular network in 2G or 4G where orthogonal resource blocks are attributed to different users. However, the number of transmitting devices is continuously increasing, which necessitates an always higher spectral-spatial efficiency. This leads to an increased usage of resource blocks per m^2. Techniques including spatial separation of users [CBVVJOP09], using the same time-frequency resource, and Non Orthogonal Multiple Access have been proposed. The consequence is either a huge overload due to signalization information or an increased impact of interference. But trying to create systems without interference is not the optimal solution.

The reason is presented auspiciously in a work from Costa [Cos83] that has proved in fact that the elimination of interference is a sub-optimal strategy when the transmission is known at the transmitter, such as the channel state information, but not at the receiver. Alternatively, the optimal approach consists in creating codes that will take advantage of the interference. This contributes to show the importance of studies about fundamental characteristics in transmissions containing interferences. In particular in information theory approach, Costa and El Gamal [CEG87], Carleial [Car78] and Sato [Sat81] have demonstrated that a very strong interference is less harmful in comparison to an interference with a power close to the useful signal and comparable to a communication without interference. These characteristics are examples that again underline the necessity

of understanding the interference, its statistical properties and the theoretical limits that they induce on the communications.

To evaluate the effects of the noise and interference appearing in transmission systems, it is first essential to comprehend its main characteristics. Consequently in the following we are going to describe the main models proposed in the literature. A fundamental model choice is the Gaussian random variable (r.v.). It is motivated by the accumulated independent and identically distributed (i.i.d.) signals that compose the interference term, with the justification of the Central Limit Theorem (CLT). The beauty of this model is its simplicity. Simplicity in the mathematical approach: so many works have been done on this distribution that analytical expression exist for many derivations we can encounter in wireless communication. We would like here to underline the benefits of the stability property, which means that the addition of Gaussian r.v. is a Gaussian r.v.; simplicity in its parametrization ; simplicity in its use in wireless communication: much information is simply given by the Signal plus Noise ratio and the optimal receiver (obtained with the maximum log-likelihood) is a linear receiver, very easy to practically implement.

Many works have taken advantage of this well-known distribution. We can for instance cite the classic work of Shannon [Sha49] on capacity. The Additive White Gaussian Noise assumption allows to obtain a very basic formulation of this capacity, $C = B\log_2(1+S/N)$, where B is the bandwidth of the transmission and S/N the Signal to Noise Ratio. However, from a Information Theoretic point of view, results about capacity with interfering communications are much more difficult to obtain. Works were initiated by Shannon [Sha+61] and followed by [Ahl74]. This approach is specially used when leading with a fixed active set of transmitting devices, as will be clarified throughout this chapter.

Nevertheless, the drastic change in the environment in modern communication systems [Car10; PW10b] puts a limit on the validity of this classical model. In particular, we focus in this thesis on impulsive noises, which can be considered a fundamental limits in digital subscriber line (DSL) [KG95], wireless [PFFR09] and power line communication (PLC) [MGC05] for instance. Many works have shown in many contexts that the Gaussian approximation was not a good choice for the noise or the interference, then alternatively many models characterized by density functions that have heavier tails are assumed. The consequence is that large amplitude noise is more probable. The presence of this noise yield different results in comparison with Gaussian noise channels: in terms of capacity, in terms of performance, in terms of receiver design for instance. The mismatch has an impact on the communication design.

Extended research was carried out during the last years investigating new models in various scenarios to lead with interferences as alternative

to the standard Gaussian r.v. One of the first significant contributions comes for Middleton [Mid77b]. In the beginning of 2000, research on Ultra Wide Band communication also gave rise to many empirical modeling approaches. They are often based on pragmatic choices, that allow a good fit with generally simulated data and an analytical solution for the maximum likelihood detector [GCASR10; FH06; ECD08]. But many other communication contexts also gave rise to research on more appropriate models, as in underwater [ZQ06] and molecular communications [FGCE15], man-made and low frequency atmospheric noises [Raa10]. We can also mention an important amount of works in multiple users communications and *ad hoc* networks. The first can probably be traced back in 1992 [Sou92] but more recently, the use of stochastic geometry in network analysis [WA12; Car10; WPS09] has given a new insight in the interference modeling. One of the important distributions that come out of those work is certainly the α-stable.

In the following, we present three key interference models. First, the well-known Gaussian distribution is described, focusing on the simplicity of probability density function and characteristic function and their analytical forms. We also present the Central Limit Theorem which apparently could be invoked to model interference when it results from the sum of a large number of interferers. We show however that this model is not well suited for impulsiveness.

Next, two non-Gaussian models are described, namely Middleton and α-stable interference models, with a special attention to the latter. The Middleton model is divided into classes, their origin and differences are described. Finally, the main distribution used in this thesis is presented, the α-stable model. In particular, the physical mechanism that leads to this model is detailed. We also present several of its properties, which will be necessary for the rest of the thesis.

2.2 Gaussian

The Gaussian distribution is the most common noise model in wireless systems, it appears basically from external environment sources and the thermal vibration of atoms in conductors, known as thermal noise. Regarding the latter, one way to understand this model arises from the distribution that maximizes the entropy, which is a condition for thermal equilibrium. Alternatively, the approach to prove this model is to use the Central Limit Theorem (CLT), which is obtained by the superposition of a large number of independent contributions and is defined as

Definition 1 (Classical CLT). *Let $\{X_j\}_{j\in\mathbb{N}}$ be an i.i.d sequence and let the mean $\mu = \mathbb{E}[X_1]$ and variance $\sigma^2 = \mathbb{E}[(X_1 - \mu)^2] < \infty$, then*

$$\frac{1}{\sigma\sqrt{n}}\left(\sum_{j=1}^{n} X_j - n\mu\right) \xrightarrow[n\to\infty]{d} X \sim \mathcal{N}(0, \sigma^2) \tag{2.1}$$

In summary, the CLT imposes that regardless the X_j distribution, the sum tends to a normal if they are i.i.d. and have a finite variance.

The Gaussian model is convenient due to its analytical and tractable forms. Formally, the Gaussian noise pdf for a continuous random variable X is given by

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{2.2}$$

where μ is the mean and σ^2 the variance. Moreover, the characteristic function is represented by

$$\phi_G(\theta) = e^{i\mu\theta} e^{-\frac{1}{2}(\sigma\theta)^2}. \tag{2.3}$$

Fig. 2.1 represents the samples generated by a standard Gaussian noise $\mathcal{N}(0, 1)$. One may notice that the amplitudes are well framed, thus the presented model is inappropriate for impulsive behaviors. This is due to the

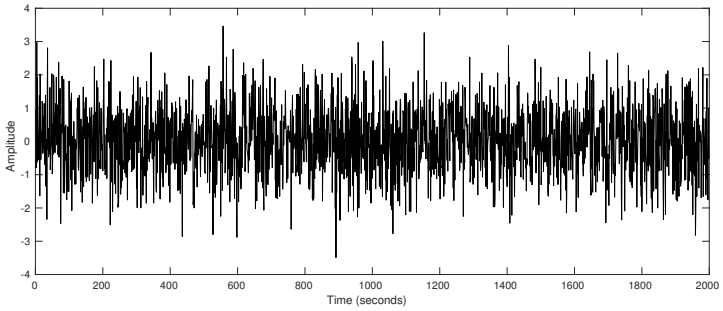


FIGURE 2.1: Samples of Gaussian distribution $\mathcal{N}(0, 1)$ over time. Window of 2000 seconds. The amplitudes are well behaved.

fast tail decay of the pdf depicted in Fig. 2.2 and that can be quantified as

Definition 2 (see [Gor41]). *Let $X \sim \mathcal{N}(\mu, \sigma^2)$ then*

$$\Pr(|X - \mu| > t) \leq \sqrt{\frac{2}{\pi}}\frac{\sigma}{t} e^{-\frac{t^2}{2\sigma^2}}, t > 0, \tag{2.4}$$

meaning that it decays exponentially. As such, the probability of large samples is small.

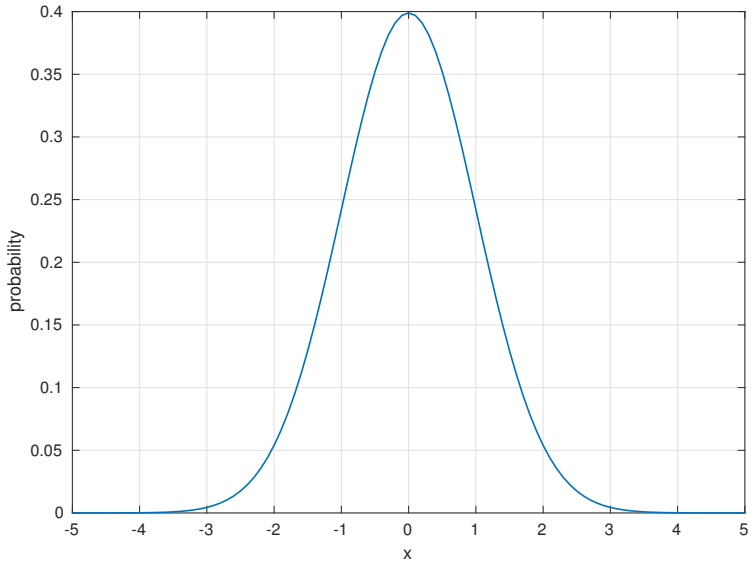


FIGURE 2.2: Gaussian probability density function $\mathcal{N}(0, 1)$

2.3 Middleton

In this section, a non-Gaussian model that copes with impulsive random noise is presented. The Middleton models have been proposed in different contexts, such as Multiple Input Multiple Output (MIMO) [CGETS09], Orthogonal Frequency-Division Multiplexing (OFDM) [III07] and Power Line Communications (PLC) [AP10]). This model gained popularity attributable to the invariance to noise waveform, noise source and propagation. The origin of this model can trace back to works from K. Furutsu and T. Ishida [FI60], A. Giordano and F. Haber [GH72] that created statistical treatment to atmospheric noise. The former separated its approach in two parts. Firstly, a Poisson noise model was presented, composed by the superposition of independent pulses. The waveforms may overlap and are identical, but the amplitudes, phases and duration may be random. They obtained the characteristic function of this impulsive model and, for the narrow band case, and also the density. Examples of this model include the ignition, precipitation and solar noises. In turn, their second model, a Poisson-Poisson noise, is an extension of the first, replacing the elementary impulses of Poisson Process by packets of Poisson noise, forming a representative example of the Atmospheric radio noise.

The works from Middleton [Mid77a; Mid99] obtained a more general result, based on series expansions, defined as an infinite weighted sum of Gaussian densities with decreasing weights for Gaussian densities with increasing variances. The model is divided in three classes:

- Class A is proposed when the noise bandwidth is smaller than the receiver bandwidth.

- The Class B is used for noise with a wider spectrum than the useful signal.
- Lastly, the Class C is the sum of the two previous.

Regarding Class A, the pdf can be written as

$$\Pr(x) = e^{-A} \sum_{m=0}^{+\infty} \frac{A^m}{m!\sqrt{2\pi\sigma_m^2}} e^{-\frac{x^2}{2\sigma_m^2}}, \tag{2.5}$$

in which

$$\sigma_m^2 = \left(\sigma_G^2 + \sigma_I^2\right) \frac{\left(\frac{m}{A} + \frac{\sigma_G^2}{\sigma_I^2}\right)}{1 + \frac{\sigma_G^2}{\sigma_I^2}}.$$

The impulsiviness is controlled by $A > 0$, which is called impulsive index, or more recently, overlap index. It concerns the mean length of an emission in seconds times the mean number of emissions per second. Smaller values produce more impulsive noise (although $A = 0$ degenerates into purely Gaussian), conversely, as A increases the noise tends towards the Gaussian noise. $\sigma_G^2 > 0$ and $\sigma_I^2 > 0$ represent the Gaussian and impulsive powers, respectively. The main appeal of this model is the possibility of interpret its parameters directly from a physical perspective.

Next, the zero mean Class B pdf can be defined using a infinity series as

$$\Pr(x) = \frac{1}{\pi\Omega} \sum_{m=0}^{+\infty} \frac{(-A)^m}{m!} \Gamma\left(\frac{1+\alpha m}{2}\right) \Psi\left(\frac{1+\alpha m}{2}; \frac{1}{2}; -\frac{x^2}{\Omega^2}\right), \tag{2.6}$$

the parameter $A > 0$ denotes again the intensity of the impulsive interference, $\Psi(\cdot;\cdot;\cdot)$ represents the confluent hypergeometric function. In addition, $0 < \alpha < 2$ controls the tails, such that smaller values correspond to heavy tails and, consequently, greater impulsiveness. Finally, Ω plays a similar role as the standard deviation in the Gaussian case.

The Fig. 2.3 gives a Middleton Class A with parameters $A = 0.3$, the background-to-impulsive noise ratio $\frac{\sigma_G^2}{\sigma_I^2} = 0.1$ and 10 terms of the summation in the Class A pdf are used, so that $2 \cdot 10^3$ samples are depicted. It has been shown that the approximation of its pdf needs just a few terms [Vas84], justifying the small number of terms considered.

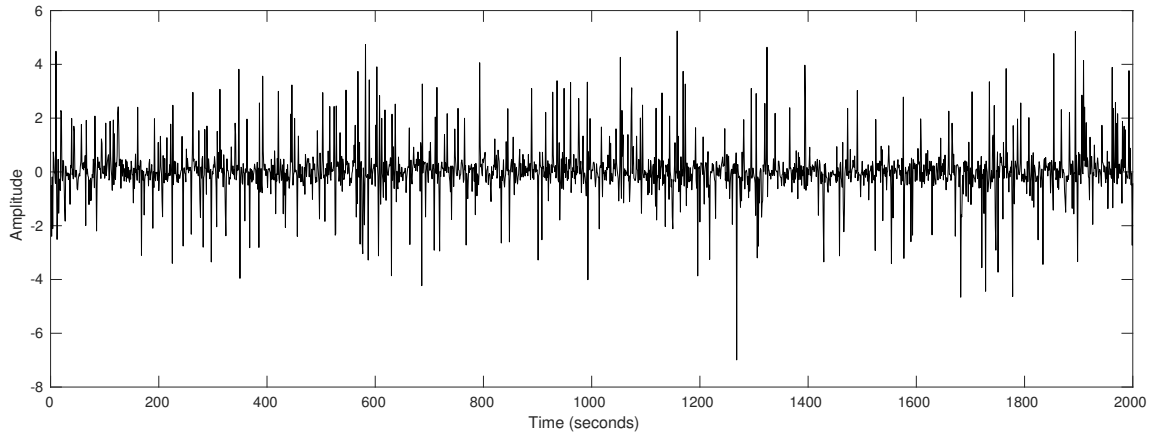


FIGURE 2.3: Middleton Class A samples over time. Window of 2000 seconds. Impulsive behavior is present

Next, Fig. 2.4 illustrates the probability density function in which different values for the impulsive index A and the background-to-impulsive noise ratio are given. One can see that the curves obtained with a smaller A have heavier pdf tails, while $\frac{\sigma_G^2}{\sigma_I^2}$ controls the spread or variety of possible values under the distribution.

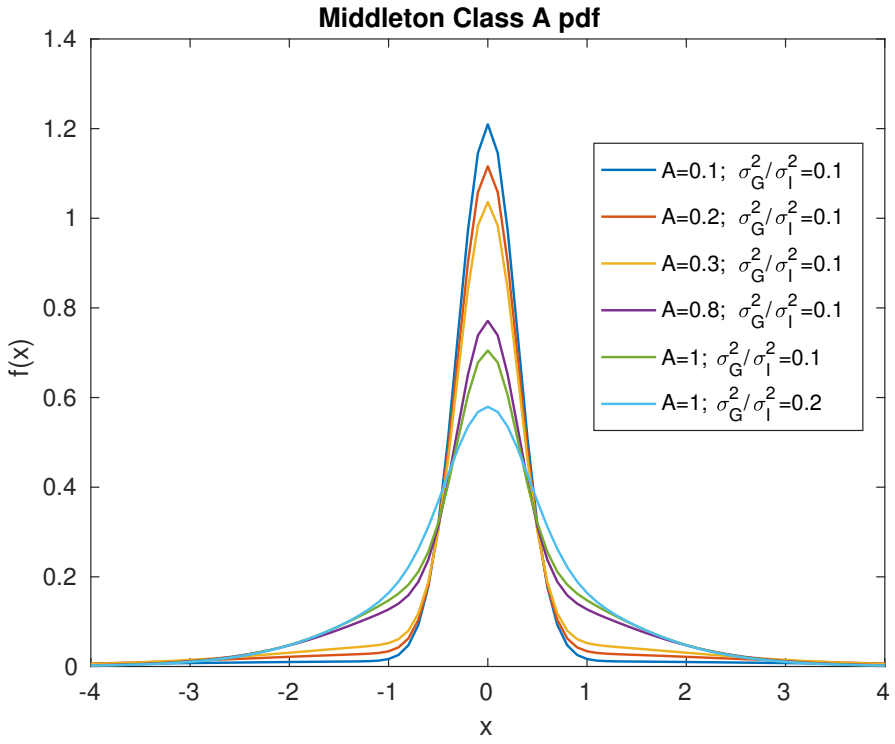


FIGURE 2.4: Middleton Class A probability density function containing different A and σ_G^2/σ_I^2 parameters

It has been proven difficult to work with Middleton models in many practical scenarios, due to the nature of infinity series. In particular, the individual terms of the summation can yield large values, turning the computational calculation problematic. Besides, the class B makes truncation

challenging in order to guarantee the production of positive quantities for the pdf. For that reason many approximations have been proposed, such as Gaussian mixture [GDK06], ϵ-contaminated noise [AB07; PLZB03; AALM17] and Bernoulli-Gaussian noise [HTLN12b; VTNH14]. In the next section and in the rest of this thesis, we focus on the α-stable distribution, which may be used to approximate the Middleton Class B model [NS95].

2.4 α-stable model

An important approach to modeling impulsive noise is via the α-stable distribution. They are variables belonging to an important class containing heavy-tailed probability density functions, which have been widely used to model impulsive signals [EFCGPA16]. They have—unlike Gaussian noise— infinite variance. They are known to be a good approximation of the true interference distribution when the radius of the network is large, there are no guard zones and the active interferer set changes rapidly [GEAT10; YP03; IH98]. Moreover, these distributions form an approximation of Middleton's noise models [Mid77a], which we have seen on previous section were derived for interference in wireless networks from a statistical physics perspective.

The challenge present in α-stable distribution is the lack of closed-form expressions for the probability density function. Therefore, for instance, the characterizations of achievable rates in the presence of α-stable noise have been arduous to obtain and, as consequence, its use has been limited in communication systems.

Recently, $AS\alpha SN$ channels have been directly shown to arise under certain conditions in the molecular timing channel [FGCE15]. Molecular communication is a method of communicating between nanoscale devices by encoding information into the number, type or release time of molecules, which then diffuse through a fluid to a receiving device. In the case of one-dimensional diffusion governed by Brownian motion without drift, the distribution for the diffusion time is Lévy. The $AS\alpha SN$ channel can then arise when information is encoded into the release time of each molecule and differential encoding is used [FGCE15]. More generally, in the presence of diffusion governed by Lévy processes, the tail of the diffusion time distribution decays as a polynomial [MK04] and hence α-stable distributions form natural approximate models.

In wireless networks, memoryless additive α-stable noise channel models have gained importance to characterize interference in reason of a study initialized by Pinto [PW10a; PW10b]. When fast-varying (symbol-by-symbol) active transmitter sets is present, assuming to be independent at each time t, and the set of active transmitting devices with locations governed by a homogeneous Poisson point process changing rapidly occur, the *dynamic interference* is produced, which means that the interference can be

characterized by a LePage series [ST94a]. This implies that the interference statistics are α-stable, as we will explain in detail in Section 2.5.1, after the presentation of α-stable concepts.

2.4.1 Definitions and some α-stable concepts

In this section, definitions of α-stable are exploited. In addition, we outline some characteristics of this distribution, such as the tail behavior, moments, the probability density, characteristic functions and the use of series to represent α-stable random variables, as well as, properties that will be used during the rest of this thesis.

The stable random variables follow the definition

Definition 3. *Formally, a stable random variable is defined as*

$$a'X_1 + a''X_2 \stackrel{d}{=} aX + b, \tag{2.7}$$

in which $a', a'' \in \mathbb{R}_{>0}$ and $a \in \mathbb{R}_{>0}$ and $b \in \mathbb{R}$. In addition, X_1 and X_2 are independent random variables sharing the same distribution X. In particular, when $b = 0$ it is strictly stable.

The Definition 3 justifies the stability notion, as the shape of X is conserved under addition.

The α-stable random variables are characterized by four parameters:

- the exponent $0 < \alpha \leq 2$: It controls the thickness of the tail of the distribution. In other words, as larger the value of α becomes, less rare events happen. When α decreases more impulsiveness is created. In particular when $\alpha = 2$ we return to a Gaussian random variable;
- the scale parameter $\gamma \in \mathbb{R}_{>0}$: Alternatively called dispersion, measures the width of the distribution. For the Gaussian case, γ is equivalent to half of this variance;
- the skew parameter or symmetry parameter $\beta \in [-1, 1]$; Particularly when $\beta = 1$ the distribution is totally skewed to right.
- the location parameter $\delta \in \mathbb{R}$: Determines the shift of the distribution. Let $p_X(x)$ be the probability density function of an α-stable random variable X,

  $$p_{X,\delta}(x) = p_X(x - \delta)$$

  thus, δ is the location parameter.

An α-stable random variable X is then represented by $X \sim S_\alpha(\gamma, \beta, \delta)$. Fig. 2.5 depicts the probability density function of a symmetric α-stable random variable with parameters $\mu = 0, \gamma = 1$, $\beta = 0$ and different α values. A stable random variable is called standard when $\mu = 0$ and $\gamma = 1$, but

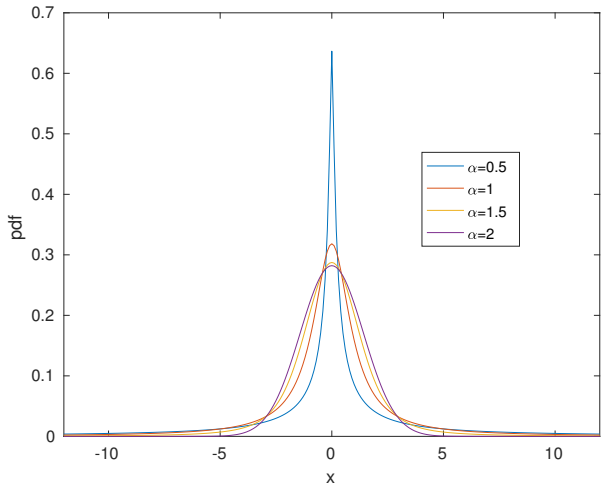


FIGURE 2.5: α-stable probability density function with $\mu = 0, \gamma = 1$ and $\beta = 0$ (symmetric). $\alpha = 0.5, 1, 1.5, 2$

it depends on the parametrization choice. As such, in Fig. 2.6 a standard α-stable pdf is presented, with $\alpha = 1.2, \mu = 0, \gamma = 1$ and three skewness values, namely, $\beta = 1$ (totally skewed to the right), $\beta = 0$ (symmetric) and $\beta = -1$ (totally skewed to the left).

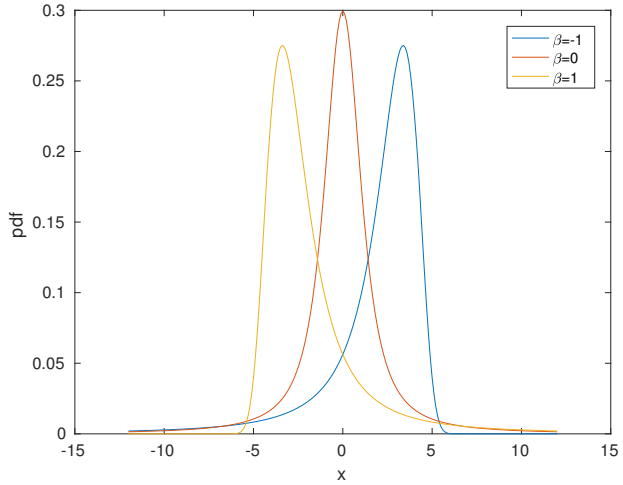


FIGURE 2.6: α-stable probability density function to represent the skewness. Contains the parameters $\alpha = 1.2, \mu = 0$ varying the skewness $\beta = -1, 0, 1$

Furthermore, in Fig 2.7, 2000 samples in time are generated from an α-stable distribution $S_{1.5}(0, 1, 0)$. When comparing with Fig 2.1, the impulsiveness behavior is easily confirmed.

In fact, a very important extension of Definition 3 is that any linear combination of stable random variables with characteristic exponent α has the same parameter α, which gives the following definition

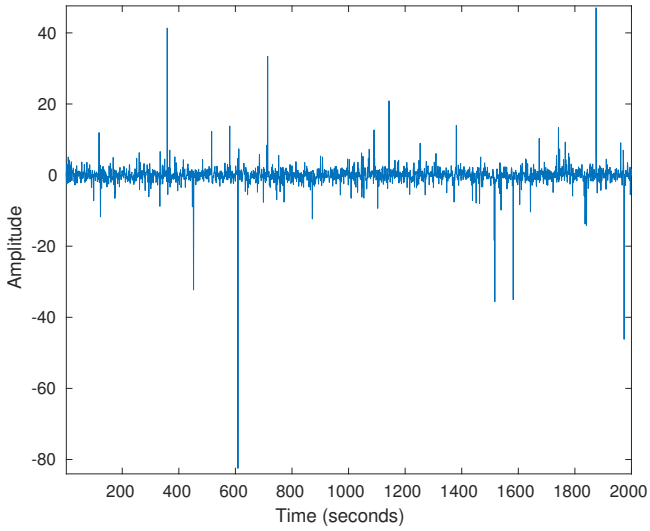


FIGURE 2.7: 2000 noise samples induced by an α-stable distribution with parameters $\alpha = 1.5, \beta = 0, \gamma = 1, \mu = 0$

Definition 4. *A random variable X is stable for all non-zero $n \in \mathbb{Z}_{\neq 0}$, if it exists constants $a_n \in \mathbb{R}_{>0}$ and $b \in \mathbb{R}$ such that*

$$\sum_{j=1}^{n} X_j \stackrel{d}{=} a_n X + b_n, \tag{2.8}$$

where $X_1, X_2, \ldots X_n$ are independent random variables with the same distribution X.

In addition, the constant a_n is obtained through the following lemma:

Lemma 1 ([Fel09, page 170]). *The norming constants are of the form $a_n = n^{1/\alpha}$, with $0 < \alpha < 2$.*

This result helps us to note that stable random variables are the only with a domain of attraction. In other words, an identically independently distributed random variable $\{X_j\}_{j \in \mathbb{N}}$, a sequence of real positive a_n and a real sequence b_n, such that

$$\frac{1}{a_n}\left(\sum_{j=1}^{n} X_j - b_n\right) \xrightarrow[n\to\infty]{d} X. \tag{2.9}$$

or, equivalently,

$$\lim_{n\to\infty} \Pr\left\{\frac{1}{a_n}\left(\sum_{j=1}^{n} X_j - b_n\right) < x\right\} = G(x) \tag{2.10}$$

where $G(x)$ is a non-degenerate random variable X, for all continuity points x of G. This case generalizes the well-known central limit theorem

Remark 1 (Classical CLT). *From 2.10, the Classical Central Limit Theorem is derived using $a_n = \sigma\sqrt{n}$ and $b_n = n\mu$*

In general, the distributions of α-stable random variables do not have closed form probability density functions, despite some rare cases (e.g. $\alpha = 1, 2$). Instead, they are usually represented by their characteristic function, given by

$$\mathbb{E}[e^{i\theta X}] = \begin{cases} \exp\left\{-\gamma^\alpha|\theta|^\alpha(1 - i\beta(\text{sign}(\theta))\tan\frac{\pi\alpha}{2}) + i\delta\theta\right\}, & \alpha \neq 1 \\ \exp\left\{-\gamma|\theta|(1 + i\beta\frac{2}{\pi}(\text{sign}(\theta))\log|\theta|) + i\delta\theta\right\}, & \alpha = 1 \end{cases}, \tag{2.11}$$

in which

$$\text{sign}(\theta) = \begin{cases} 1 & \text{if } \theta > 0 \\ 0 & \text{if } \theta = 0 \\ -1 & \text{if } \theta < 0 \end{cases} \tag{2.12}$$

Remark 2. *This form is presented in Samorodnitsky and Taqqu [ST94b, Definition 1.1.6], however this representation is not unique. See [CL97; Zol86; UZ99] for more options.*

An α-stable random variable is symmetric if $\beta = 0$ around $\delta = 0$ and is denoted by $S\alpha S$, for which a distribution of X is equal to $-X$. Therefore, a symmetric α-stable random variable is strictly stable, but the converse is not true in general. Clearly, if $(X - \delta)/\gamma^{1/\alpha}$ is standard with characteristic exponent α. The characteristic function for the symmetric α-stable random variable is given by

$$\begin{aligned} \phi(\theta) &= \mathbb{E}[e^{i\theta X}] \\ &= \exp\{-\gamma^\alpha|\theta|^\alpha\}, \quad \theta \in \mathbb{R}. \end{aligned} \tag{2.13}$$

By letting α receive values 0.5, 1 and 2, we obtain three special cases:

- Cauchy - Let $C \sim S_1(\gamma, 0, \delta) \triangleq S_1S(\gamma, \delta)$, then the probability density function is

$$f_C(x) = \frac{1}{\pi}\frac{\gamma}{\gamma^2 + (x - \delta)^2}, \quad -\infty < x < \infty \tag{2.14}$$

and with characteristic function

$$\phi_C(\theta) = \exp\{-\gamma|\theta| + i\delta t\} \tag{2.15}$$

- Gaussian - Let $G \sim S_2(\gamma, 0, \delta) \triangleq S_2S(\gamma, \delta)$, then the probability density function is

$$f_G(x) = \frac{1}{\sqrt{4\pi\gamma}} \exp\left[-\frac{(x-\delta)^2}{4\gamma}\right], \quad -\infty < x < \infty. \tag{2.16}$$

  As a consequence to obtain the normal distribution $\mathcal{N}(\mu, \sigma^2)$, we need to modify the dispersion in a way that $G \sim S_2S(\sigma/\sqrt{2}, \mu)$ to yield the density

$$f_G(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \tag{2.17}$$

  In turn, the characteristic function is

$$\phi_G(\theta) = \exp\left[-\frac{\sigma^2}{2}\theta^2 + i\mu t\right]. \tag{2.18}$$

- Lévy - Suppose $L \sim S_{1/2}(\gamma, 1, \delta)$, thus the density is

$$f_L(x) = \left(\frac{\gamma}{2\pi}\right)^{1/2} \frac{1}{(x-\delta)^{3/2}} \exp\left[-\frac{\gamma}{2(x-\delta)}\right]. \tag{2.19}$$

  where $x \in [\delta, \infty)$.

Despite these rare cases as already mentioned, the α-stable random variables do not have closed form. Alternatively, making use of the inverse Fourier Transform for a given characteristic function $\phi(t)$ such that $\int_{-\infty}^{\infty} |\phi(t)| < \infty$, the probability density function induced by $X \sim S\alpha S(\gamma, \beta, \delta)$ can be expressed as

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-i\theta x)\phi_X(\theta) d\theta. \tag{2.20}$$

In particular, when dealing with symmetric α-stable and $\delta = 0$, denoted as $S\alpha S(\gamma, 0)$, we obtain a simpler expression as

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\theta x} e^{-\gamma^\alpha |\theta|^\alpha} d\theta \\ &= \frac{1}{\pi} \int_{0}^{\infty} e^{-(\gamma^\alpha |\theta|^\alpha)} \cos(\theta x) d\theta \end{aligned} \tag{2.21}$$

where the last expression results from the fact that the function is real and even.

2.4.2 Tails and moments

The behavior of $\Pr(X > x)$ and $\Pr(X < -x)$ for large x, respectively right and left tails of an α-stable random variables are of the form

Proposition 1. *Assume $X \sim S\alpha(\gamma, \beta, \delta)$ with $0 < \alpha < 2$, Thus*

$$\begin{cases} \lim_{t\to\infty} t^\alpha \Pr(X > t) = \gamma^\alpha C_\alpha \frac{1+\beta}{2} \\ \lim_{t\to\infty} t^\alpha \Pr(X < -t) = \gamma^\alpha C_\alpha \frac{1-\beta}{2}, \end{cases} \tag{2.22}$$

where an α dependent constant C_α is given by

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x dx \right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\frac{\pi\alpha}{2})}, & \text{if } \alpha \neq 1, \\ \frac{2}{\pi}, & \text{if } \alpha = 1. \end{cases} \tag{2.23}$$

Proof. For more information see [ST94a, Property 1.2.15]. □

Definition 5 (Heavy tail)**.** *A real-valued random variable X have a distribution with heavy right tail if the probability tail $\Pr(X > x)$ decay more slowly than those of any exponential distribution, such that*

$$\lim_{x\to\infty} e^{\lambda x} \cdot \Pr(X > x) = \infty, \forall \lambda > 0. \tag{2.24}$$

A similar approach can be applied for heavy left tails. It is straightforward to show that α-stable densities have heavy tails by using (2.22) and knowing that polynomial expressions have a slower decay than exponential, as showed in Fig. 2.8, where the tails of a standard Gaussian and a $S_{1.5}S(\sqrt{0.5}, 0)$ r.v.'s are compared.

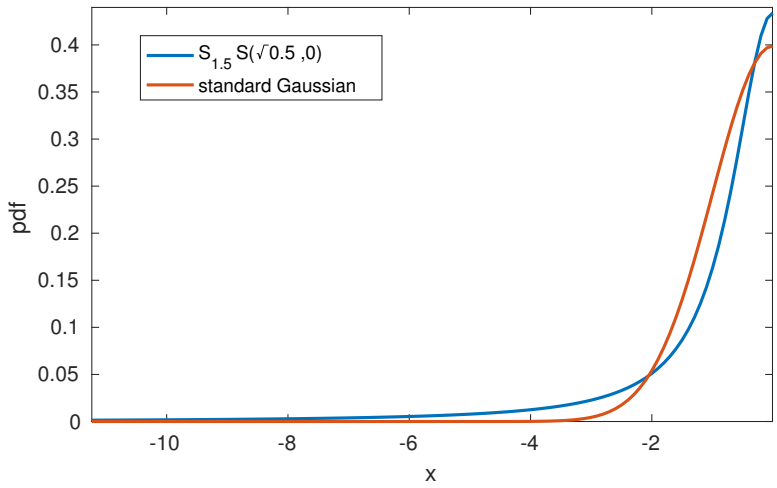


FIGURE 2.8: Comparison between the left tail of a standard Gaussian pdf and an α-stable pdf with $\alpha = 1.5$, $\beta = 0$, $\gamma = \sqrt{0.5}$ and $\mu = 0$.

This tail behavior gives the following consequence on the moments:

Proposition 2. *Let $X \sim S_\alpha(\gamma, \beta, \delta)$ with $0 < \alpha < 2$, then*

$$\begin{aligned} &\mathbb{E}[|X|^p] < \infty, \quad 0 < p < \alpha, \\ &\mathbb{E}[|X|^p] = \infty, \quad p \geq \alpha. \end{aligned} \tag{2.25}$$

This implies that the use of expectations suffers some restrictions. For instance, except the Gaussian case, α-stable random variables have infinite second order moment.

Proposition 3. *The fractional lower order moments (FLOM's) of a $S\alpha S(\gamma, 0)$ random variable have the form*

$$\mathbb{E}[|X|^p] = C(p, \alpha)\gamma^p, 0 < p < \alpha, \tag{2.26}$$

in which $C(p,\alpha) = \frac{2^{p+1}\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(-\frac{p}{\alpha}\right)}{\alpha\sqrt{\pi}\Gamma\left(-\frac{p}{2}\right)}$ and $\Gamma(\cdot)$ is the Gamma function defined for $x > 0$ as

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt \tag{2.27}$$

2.4.3 Lepage series to represent α-stable random variables

In this section we present a method regarding random variables using infinitely divisible series. These series are characterized by having an attraction domain and can represent, for instance, the Central Limit Theorem. The representation of processes without fixed points of discontinuity and containing independent increments having no Gaussian components was firstly been introduced by Fergusson and Klass [FK72], later by Lepage [LeP89], with extensions by Rosinski [Ros90]. Particularly, we study in this section a technique which consists in writing an α-stable process through an infinite sum of independent random variables and an arrival time induced by a Poisson process. To elucidate the following proposition we denote the sequences $(\epsilon_i)_{i\in\mathbb{N}}$, $(W_i)_{i\in\mathbb{N}}$ and $(\tau_i)_{i\in\mathbb{N}}$, such that

- The random variable sequence $\epsilon_1, \epsilon_2, \ldots$ is identical and independently distributed having a Rademacher distribution. This distribution is characterized by $\Pr(\epsilon_i = 1) = \Pr(\epsilon_i = -1) = \frac{1}{2}$, with a support $\epsilon_i \in \{-1, 1\}$.

- $W_1, W_2, \ldots$ are i.i.d random variables with finite absolute α^{th} moment.

- The $\tau_1, \tau_2, \ldots$ consist of a sequence of random variables representing the arrival times of a Poisson process with intensity 1. τ_i variables are dependent between each other and not identically distributed. On the other hand, they can be viewed as

$$\tau_i = \sum_{j=1}^{i} E_j \tag{2.28}$$

  where the E_j are, in turn, i.i.d random variables and follow an exponential process with mean 1.

Proposition 4 (Lepage series). *Let W a random variable with finite fractional order moment for $0 \leq \alpha < 2$, i.e. $\mathbb{E}[|W|^\alpha] < \infty$. Then the sum $\sum_i^\infty \epsilon_i \tau_i^{-1/\alpha} W_i \overset{a.s.}{\to} X \sim S\alpha S((C_\alpha^{-1} E[|W_1|^\alpha])^{1/\alpha}, 0)$, in which C_α is defined in Proposition 1*

The notation $X_n \overset{a.s.}{\to} X$ means that a sequence X_n converges almost surely towards X or $\Pr\left(\lim_{n\to\infty} X_n = X\right) = 1$. For the proof see [ST94a; JW93].

The reverse statement is also possible, a symmetric α-stable random variable has a Lepage series form, as shown below

Proposition 5. *Let X a symmetric α-stable random variable such that $X \sim S\alpha S(\mathbb{E}[|W|^\alpha]^{\frac{1}{\alpha}}, 0)$, then*

$$X \overset{d}{=} C_\alpha^{1/\alpha} \sum_i^\infty \epsilon_i \tau_i^{-1/\alpha} W_i$$

It is important to notice that this representation is not unique, due to W_i arbitrary as long as $\mathbb{E}[|W_1|^\alpha] < \infty$. The proof is detailed in [ST94b]

2.4.4 Properties

In this section we provide key properties of α-stable random variables. These properties are used throughout the remainder of the thesis.

The first property, known as the *stability property*, concerns the addition of two α-stable random variables with the same characteristic exponent α.

Property 1 (stability property). *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*

$$\gamma = (\gamma_1^\alpha + \gamma_2^\alpha)^{\frac{1}{\alpha}}, \tag{2.29}$$

$$\beta = \frac{\beta_1 \gamma_1^\alpha + \beta_2 \gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}, \tag{2.30}$$

$$\delta = \delta_1 + \delta_2. \tag{2.31}$$

The second property concerns the scaling of an α-stable random variable by a constant.

Property 2. *Let $a \in \mathbb{R}_{\neq 0}$, i.e., a non null constant, and $X \sim S_\alpha(\gamma, \beta, \delta)$. Thus,*

$$\begin{aligned} aX &\sim S_\alpha(|a|\gamma, \operatorname{sign}(a)\beta, a\delta), && \text{if } \alpha \neq 1. \\ aX &\sim S_1(|a|\gamma, \operatorname{sign}(a)\beta, a\delta - 2/\pi a(\ln|a|)\gamma\beta), && \text{if } \alpha = 1. \end{aligned} \tag{2.32}$$

Corollary 1. *Let $Z \sim S_\alpha(\gamma, \beta, \delta)$ and $a \in \mathbb{R}_{>0}$. If $1 < \alpha < 2$, then*

$$aZ \sim S_\alpha(a\gamma, \beta, a\delta). \tag{2.33}$$

The third property concerns the addition of an α-stable random variable with a constant.

Property 3. *If $Z \sim S_\alpha(\gamma, \beta, \delta)$, then $Z + \mu \sim S_\alpha(\gamma, \beta, \delta + \mu)$.*

The fourth property concerns the maximum value of α-stable probability density functions.

Property 4. *Let $Z \sim S_\alpha(\gamma, \beta, \delta)$ for $\alpha \in [1, 2)$, then*

$$p_Z(y) \leq \frac{\Gamma\left(\frac{1}{\alpha}\right)}{\gamma\alpha\pi}. \tag{2.34}$$

Proof. The characteristic function of a $S\alpha S$ random variable with $\delta = 0$ is $\phi_Z(t) = e^{-|\gamma\theta|^\alpha}$, $\gamma > 0$ and using (2.21), we have

$$\begin{aligned} p_Z(y) = F'_Z(y) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\theta y}\phi_Z(\theta)d\theta \\ &\leq \frac{1}{2\pi}\int_{\mathbb{R}} |e^{-i\theta y}\phi_Z(\theta)|d\theta \\ &\leq \frac{1}{2\pi}\int_{\mathbb{R}} |e^{-|\gamma\theta|^\alpha}|d\theta \\ &\leq \frac{1}{\pi}\int_{\mathbb{R}_{\geq 0}} e^{-(\gamma\theta)^\alpha}d\theta \\ &\leq \frac{\Gamma\left(\frac{1}{\alpha}\right)}{\gamma\alpha\pi}, \end{aligned} \tag{2.35}$$

in which $\Gamma(\cdot)$ is the Gamma function. □

The fifth property gives $\mathbb{E}[|Z|]$ for symmetric α-stable random variables, originally due to Zolotarev [Zol57] and can be seen from Proposition 3, for $p = 1$.

Property 5. *Let $Z \sim S_\alpha(\gamma, 0, 0)$, with $1 < \alpha \leq 2$. Then,*

$$\mathbb{E}[|Z|] = \frac{2\Gamma\left(1 - \frac{1}{\alpha}\right)}{\pi}\gamma. \tag{2.36}$$

The sixth property concerns the asymptotic behavior of the probability density function of symmetric α-stable distributions (see [ST94a; FN99]).

Property 6. *Let $Z \sim S_\alpha(\gamma, 0, \delta)$ with $1 < \alpha \leq 2$. Then, the probability density function of Z satisfies*

$$p_Z(z) \sim \frac{\alpha(1-\alpha)\gamma^\alpha}{\Gamma(2-\alpha)\cos\left(\frac{\pi\alpha}{2}\right)}|z|^{-\alpha-1} \text{ as } |z| \to \infty. \tag{2.37}$$

Fig. 2.9 plots the asymptotic version of the probability density function described in (2.37) and symmetric α-stable random variables with $\gamma = 1$

and $\mu = 0$ obtained numerically. One may notice that the asymptotic pdfs act as upper bounds, becoming rapidly similar to the α-stable r.v.'s as z increases.

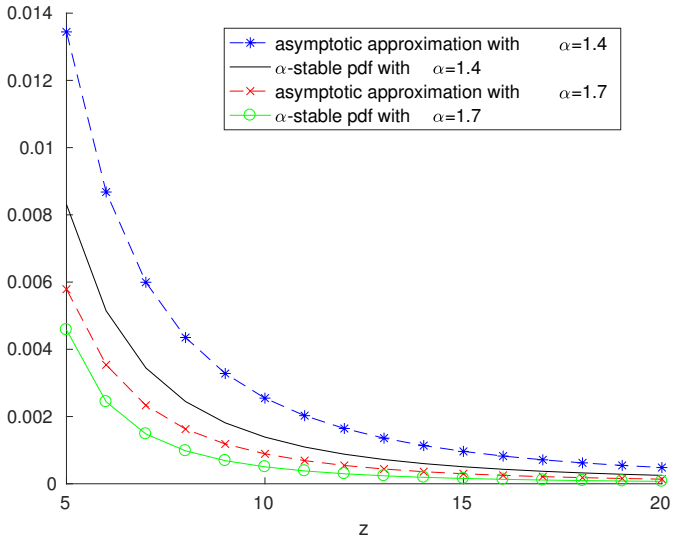


FIGURE 2.9: Comparison between the right tail of symmetric α-stable pdfs and the asymptotic pdf in (2.37) for $\beta = 0$, $\gamma = 1$ and $\mu = 0$.

Finally, the seventh property is an important extension to a random variable X is presented, in which a transform of $S\alpha' S$ into a $S\alpha S$ is given by

Property 7. *Let $X \sim S\alpha' S(\gamma, 0)$ with $0 < \alpha < \alpha' \leq 2$. Thus, for an $A \sim S_{\alpha/\alpha'}\left(\cos\left(\frac{\pi\alpha}{2\alpha'}\right)^{\alpha'/\alpha}, 1, 0\right)$, such that X and A are independent, then*

$$Z = A^{1/\alpha'} X \sim S\alpha S(\gamma, 0).$$

Proof.

$$\begin{aligned}\mathbb{E}\{\exp\{i\theta Z\}\} &= \mathbb{E}\{\exp\{A^{1/\alpha'} X\}\} \\ &= \mathbb{E}\{\mathbb{E}\{\exp\{i\theta a^{1/\alpha'} X\}|A = a\}\}\end{aligned} \tag{2.38}$$

After applying the characteristic function and Proposition 2, we obtain

$$\begin{aligned}\mathbb{E}\{\exp\{i\theta Z\}\} &= \mathbb{E}\{\mathbb{E}\{\exp\{-(\gamma^{\alpha'}|\theta|^{\alpha'} a^{\alpha'/\alpha'})\}|A = a\}\} \\ &= \mathbb{E}\{\exp\{-(\gamma^{\alpha'}|\theta|^{\alpha'} A)\}\}.\end{aligned} \tag{2.39}$$

Next, choosing $A \sim S_{\alpha/\alpha'}\left(\cos\left(\frac{\pi\alpha}{2\alpha'}\right)^{\alpha'/\alpha}, 1, 0\right)$ with $0 < \alpha < \alpha' \leq 2$, so that the Laplace transform of A, an α/α'-stable totally skewed to the right, is $\mathbb{E}\{\exp(-\gamma A)\} = \exp(-\gamma^{\alpha/\alpha'})$ [ST94b, Proposition 1.2.12], which is also called as Kohlrausch-Williams-Watts function [PG10]. Consequently, we

have

$$\begin{aligned}\exp\{i\theta Z\} &= \exp\{-(\gamma^{\alpha'}|\theta|^{\alpha'})^{\alpha/\alpha'}\} \\ &= \exp\{-\gamma^{\alpha}|\theta|^{\alpha}\},\end{aligned} \tag{2.40}$$

for $\gamma > 0$. □

This result implies that when $\alpha' = 2$ and X is a zero mean Gaussian random variable, the $S\alpha S$ random variables are conditionally Gaussian, i.e. $Z = A^{\frac{1}{2}}X \sim S\alpha S$.

2.4.5 Bivariate Isotropic Stable Distributions

So far we have considered the univariate stable distribution. In this section, it is extended to the multivariate distributions, specifically we focus on the bivariate isotropic case. Like the univariate stable distributions presented in the previous section, the multivariate stable distribution is defined by the stability property and the generalized Central Limit Theorem.

Consequently, the Definition 3 (page 16) can be extended to the multivariate case defined as

Definition 6. *Let $\boldsymbol{X} = (X_1, \ldots, X_d)$ in $\mathbb{R}^d$ a d-dimensional α-stable random vector. $\boldsymbol{X}$ is stable if $\forall (a, b) \in \mathbb{R}^2_{>0}$, $\exists\, c \in \mathbb{R}_{>0}, \boldsymbol{D} \in \mathbb{R}^d$ such that $a\boldsymbol{X}_1 + b\boldsymbol{X}_2 \overset{d}{=} c\boldsymbol{X} + \boldsymbol{D}$, where $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are i.i.d. copies of $\boldsymbol{X}$. If $\boldsymbol{D}$ is a null vector, it is said to be strictly stable.*

In fact, this stability property can be generalized to encompass linear combinations in manner that the following property is given

Property 8. *Let $\boldsymbol{X} \in \mathbb{R}^d$ be an α-stable random vector, thus*

- *All linear combinations $\sum_j^d b_j X_j$ are α-stable random variables.*
- *If the linear combinations are symmetric stable, this implies that $\boldsymbol{X}$ is a symmetric stable random vector in $\mathbb{R}^d$.*

Proof. See [ST94a, Theorem 2.1.5] for more details. □

The characteristic function plays a key role since α-stable random variables do not in general have closed-form probability density functions, similarly to the univariate case. As such, the characteristic function for the bivariate stable distribution case ($d = 2$) is given by

$$\phi(\boldsymbol{\theta}) = \begin{cases} \exp(i\boldsymbol{\theta}^T\boldsymbol{\delta} - \boldsymbol{\theta}^T K\boldsymbol{\theta}), & \text{if } \alpha = 2 \\ \exp(i\boldsymbol{\theta}^T\boldsymbol{\delta} - \int_S |\boldsymbol{\theta}^T\mathbf{s}|\mu(d\mathbf{s}) + i\beta_\alpha(\boldsymbol{\theta}), & \text{if } 0 < \alpha < 2, \end{cases}$$

in which

$$\beta_\alpha(\boldsymbol{\theta}) = \begin{cases} \tan\frac{\alpha\pi}{2}\int_S |\boldsymbol{\theta}^T\mathbf{s}|^\alpha \text{sign}|\boldsymbol{\theta}^T\mathbf{s}|\mu(d\mathbf{s}), & \text{if } \alpha \neq 1, 0 < \alpha < 2, \\ \int_S \boldsymbol{\theta}^T\mathbf{s}\log|\boldsymbol{\theta}^T\mathbf{s}|\mu(d\mathbf{s}), & \text{if } \alpha = 1, \end{cases} \tag{2.41}$$

where $\boldsymbol{\delta} = (\delta_1, \delta_2), \boldsymbol{\theta} = (\theta_1, \theta_2)$ and the $|\boldsymbol{\theta}| = \sqrt{\theta_1^2 + \theta_2^2}$. In addition, S is the unit circle, K represents a positive semidefinite symmetric matrix. The measure $\mu(\cdot)$ is the spectral measure.

Particularly, a notable example for the bivariate stable distribution is the rotationally invariant case, i.e., the random variable does not change under rotation, reflections or inversion operations. The stable distributions having such features are called *isotropic*. The following proposition expresses the necessary conditions

Proposition 6. *A complex variable* $\boldsymbol{X}$ *is isotropic if it satisfies the following conditions:*

- ***C1*** *: The random vector* $\mathbf{X} = (X_1, X_2)^T$ *is symmetric in* $\mathbb{R}^2$*; i.e.,* $\Pr(-\mathbf{X} \in A) = \Pr(\mathbf{X} \in A)$ *for all Borel sets* A *in* $\mathbb{R}^2$*.*
- ***C2*** *:* $e^{i\phi}X \stackrel{d}{=} X$ *for any* $\phi \in [0, 2\pi)$*.*

As a consequence, (2.41) is modified to the bivariate isotropic characteristic function using $\boldsymbol{\theta} = (\theta_1, \theta_2)$, yielding the following closed form

$$\begin{aligned}\phi(\theta_1, \theta_2) &= \exp\left\{i(\delta_1\theta_1 + \delta_2\theta_2) - \left(\gamma\sqrt{\theta_1^2 + \theta_2^2}\right)^{\alpha}\right\} \\ &= \exp\left\{i(\delta_1\theta_1 + \delta_2\theta_2) - \gamma^{\alpha}|\boldsymbol{\theta}|^{\alpha}\right\}, \end{aligned} \tag{2.42}$$

in which the Corollary 2 is applied for vector representation. where the parameter α is the characteristic exponent and γ represents the dispersion as in the univariate case. The δ_1 and δ_2 are the location parameters.

2.4.6 Multivariate sub-Gaussian stable distribution

The multivariate sub-Gaussian representation is presented in this section, which is a means of constructing a symmetric α-stable random variable vector by multiplying an $\alpha/2$-stable r.v. totally skewed to the right $(\beta = 1)$ and a normal random vector $\mathbf{G}$. Therefore, the multivariate sub-Gaussian can be expressed as

Definition 7. *A vector* $\boldsymbol{X}$ *is said to be sub-Gaussian* $S\alpha S$ *in* R^d *with underlying Gaussian vector* $\boldsymbol{G}$*, when*

$$\boldsymbol{X} = (A^{1/2}G_1, A^{1/2}G_2, \ldots, A^{1/2}G_d),$$

where $A \sim S_{\alpha/2}\left(\cos\left(\frac{\pi\alpha}{4}\right)^{2/\alpha}, 1, 0\right)$, $\boldsymbol{G} = (G_1, \ldots, G_d) \in \mathbb{R}^d$ *are zero mean Gaussian random vectors and* $\alpha < 2$*.* A *and* $\boldsymbol{G}$ *are independent.*

In other words, A is any positive r.v., but the product will be symmetric α-stable only if A is the given $\alpha/2$-stable random variable. In fact, differently from the linear spaces of Gaussian random variables, sub-Gaussian r.v.'s do not have (nondegenerate) i.i.d. elements [CS84, Lemma 2.1]. In

addition, **X** is completely parameterized by the statistics of the Gaussian random vector **G**. As a consequence, to check stationarity, it is necessary and sufficient to verify if the Gaussian term is stationary for instance.

Now, its characteristic function is presented as follows

Proposition 7. *The characteristic function of a sub-Gaussian symmetric α-stable random vector **X** is given by*

$$\mathbb{E}\left\{\exp\left[\sum_{k=1}^{d}\theta_k X_k\right]\right\} = \exp\left\{-\left|\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}R_{ij}\theta_i\theta_j\right|^{\frac{\alpha}{2}}\right\}, \tag{2.43}$$

in which $R_{ij} = \mathbb{E}[G_iG_j]$ is the covariance of the underlying zero mean Gaussian random variable $\boldsymbol{G} = (G_1, \ldots, G_d)$.

2.4.7 Complex SαS random variables

It is worth noting that the complex α-stable random variables are denoted as

$$X = X_1 + iX_2, \tag{2.44}$$

but, instead of working with the complex representation, one may reformulate the problem in terms of a real valued vector representation, as follows

Definition 8. *The complex random vector $X = (X_1, \ldots, X_d)$, where each element has the form $X_j = X_j^1 + iX_j^2$ for $j = 1, \ldots, d$ and X_j^1, X_j^2 are real random vectors, is α-stable if and only if there exists a random vector whose elements are real α-stable random variables $(X_1^1, X_1^2, \ldots, X_d^1, X_d^2) \in R^{2d}$.*

Therefore, plugging $d = 1$ into the previous definition we have

Corollary 2. *A complex random vector $X = X_1 + iX_2$ is α-stable if and only if there exists a random vector whose elements are real α-stable random variables $(X_1, X_2) \in R^2$.*

Now that a complex representation is depicted, a reformulation of the sub-Gaussian random vectors is possible and the key of this is given in [ST94a, Corollary 2.6.4], yielding

Theorem 1. *Let $0 < \alpha < 2$. A complex α-stable random variable $Z = Z_1 + iZ_2$ is isotropic if and only if there are two independent and identically distributed zero-mean Gaussian random variables G_1, G_2 with variance $\sigma_{\mathbf{N}}^2$ and a random variable $A \sim S_{\frac{\alpha}{2}}((\cos(\pi\alpha/4))^{2/\alpha}, 1, 0)$ independent of $(G_1, G_2)^T$ such that $(Z_1, Z_2)^T = A^{\frac{1}{2}}(G_1, G_2)^T$; i.e., $(Z_1, Z_2)^T$ is a sub-Gaussian random vector.*

2.4.8 Isotropic properties

The following property demonstrates the conditions necessary in order to maintain an isotropic random variable when multiplied by a matrix.

Property 9. *Let* $\mathbf{Z}$ *be a random vector induced by the isotropic* α*-stable random variable* Z*. Then,* $\mathbf{Z}' = \mathbf{V}\mathbf{Z}$ *is another random vector induced by an isotropic* α*-stable random variable* Z' *if and only if* $\mathbf{V}$ *satisfies* $\mathbf{V}\mathbf{V}^T = c\mathbf{I}$ *for some* $c \geq 0$*.*

Proof. To prove ($\Leftarrow$), consider a matrix $\mathbf{V} \in \mathbb{R}^2$. If $\mathbf{Z}'$ is a random vector induced by the isotropic α-stable random variable Z', then by Theorem 1

$$\mathbf{Z}' = \mathbf{V}\mathbf{Z} \stackrel{d}{=} A^{1/2}\mathbf{V}(G_1, G_2)^T. \tag{2.45}$$

For this condition to hold, $(G_1', G_2')^T = \mathbf{V}(G_1, G_2)^T$ must be i.i.d normal. Observe that the covariance of $(G_1', G_2')^T$

$$\text{cov}((G_1', G_2')^T) = \mathbf{V}\sigma\mathbf{I}\mathbf{V}^T = \sigma\mathbf{V}\mathbf{V}^T. \tag{2.46}$$

As such, for G_1', G_2' to be i.i.d normal, we require $\mathbf{V}\mathbf{V}^T = c\mathbf{I}$ for some $c \geq 0$. ($\Rightarrow$) follows immediately. □

Alternatively, the characteristic function of an isotropic α-stable random variable can be depicted as

Property 10. *The characteristic function of a random vector* $\mathbf{Z}$ *induced by an isotropic* α*-stable random variable* Z *(*$0 < \alpha < 2$*) is given by*

$$\phi_{\mathbf{Z}}(\boldsymbol{\theta}) = \mathbb{E}[e^{i(\theta_1 Z_1 + \theta_2 Z_2)}] = e^{-2^{-\alpha/2}\sigma_{\mathbf{Z}}^{\alpha}|\boldsymbol{\theta}|^{\alpha}}, \tag{2.47}$$

where $\sigma_{\mathbf{Z}}$ *corresponds to square root of the variance of the i.i.d Gaussian random variables in Theorem 1.*

Proof. The proof follows directly from Proposition 7 and the characteristic function of a multivariate Gaussian $\mathcal{N}(0, \mathbf{\Sigma})$, i.e. $\phi(\boldsymbol{\theta}) = e^{\frac{1}{2}\boldsymbol{\theta}^T\mathbf{\Sigma}\boldsymbol{\theta}}$, where $\mathbf{\Sigma}$ is the covariance matrix. Yielding

$$\begin{aligned}\phi_{\mathbf{Z}}(\boldsymbol{\theta}) &= \exp\left\{-\left|\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}R_{ij}\theta_i\theta_j\right|^{\frac{\alpha}{2}}\right\} \\ &= \exp\left\{-\left|\frac{1}{2}\sigma_{\mathbf{Z}}^2\sum_{j=1}^{d}\sigma_j^2\right|^{\frac{\alpha}{2}}\right\},\end{aligned} \tag{2.48}$$

where $|\boldsymbol{\theta}| = \sqrt{\sum_{j=1}^{d}\sigma_j^2}$ and using the fact that the multivariate Gaussian has zero mean and independent components. □

Next, we will investigate a model where the $AI\alpha SN$ channel naturally arises.

2.5 System Model and Dynamic Interference Characterization

In this section, an important system model is introduced to demonstrate that an α-stable interference can come from a realistic physical mechanism when the nature of interference sources and some propagation conditions are assumed.

Transmissions in a common band between networks providing multiple services and supporting standard cellular or WLAN communication and M2M ends up creating heterogeneous networks, due to several different quantities, type of data and symbol durations [AFGMAA15]. Particularly for small-cells and *ad hoc* networks, the heterogeneity appears from variations in transmit power constraints and non-uniform placement of base stations. We also remark that device heterogeneity can be captured under the assumption devices with a given protocol have the same probability for each device.

Dynamic interference can be induced by two key physical mechanisms. The first mechanism is any protocol where data is transmitted in non-contiguous blocks, in other words, interferers do not transmit data continuously. For instance, differently from the standard cellular services, which data transmissions typically vary between 1 KB and 2 MB per transmission for text and image transfers and up to 3 GB for video transfer [Tol15], M2M communications create a non-contiguous transmissions owing to a very short transmission, as rare as 1 MB per month [Dig09]. As a consequence, the active set of transmitting devices at each time can change rapidly.

As a result, in M2M networks with a rapidly changing active transmitter set, the Gaussian model does not represent the interference effects reliably, yielding an impact on performance guarantees and resource allocation, which is often based on the spectral efficiency of each link. In particular, it was shown that the interfering signal in each time slot is α-stable [ECFDGS17a]. This is specially the case when long range transmission are considered. Sensing the channel at the transmitter side does not efficiently represent the channel at the receiver side. In fact, this idea even lead to protocols based on a new ALOHA [Abr70] approach like in Sigfox [VLNKMS17]. The presence of strong colliding packets is then important and the Gaussian model is no longer adapted.

Finally, we can mention that Non-orthogonal multiple access (NOMA) [SKBNLH13]. An example is the promising strategy of Sparse code multiple access (SCMA) for OFDM systems [NB13], where users can transmit on a sparse subset of all frequency bands. This strategy leads to a non-Gaussian interference when considering change at every symbol [ECZFG18].

The second mechanism arises when there are multiple coexisting communication systems, such as IEEE 802.11 (Wi-Fi) and IEEE 802.15 (Zigbee, Bluetooth). The IEEE 802.11 frame is composed of a fixed header of 34 bytes and for a short payload of 250 bytes and data rate of 54 Mbps the on-air time is 42.07 microseconds. On the other hand, the IEEE 802.15 Zigbee frame is 40 bytes with data rate 250 kbps, leading to an on-air time of 1.25 ms. Moreover, Bluetooth is frequency hopping and is present in a 802.15 band only rarely and for a very short time. The result is that Bluetooth and Wi-Fi interferers are active for short periods of time relative to Zigbee transmissions, resulting in dynamic interference.

To illustrate the second mechanism, Fig. 2.10 shows the result of an experiment (detailed in [IP16, Section 2.5.2]) with coexisting Wi-Fi, Bluetooth and Zigbee transmissions. Observe that Bluetooth interference for very short periods of time is sufficient to corrupt a Zigbee transmission.

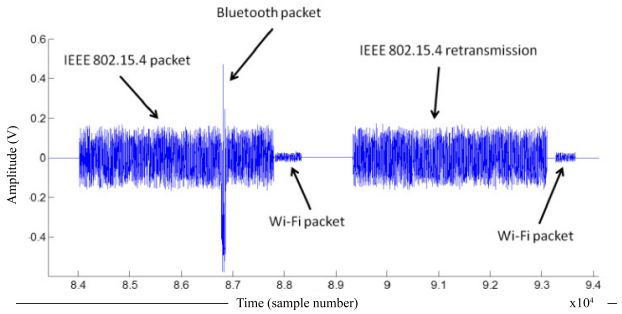


FIGURE 2.10: Coexistence of technologies in the 2.4-GHz band. Measurements made by a National Instruments USRP (detailed in [IP16, Section 2.5.2]).

In order to characterize the interference in this scenario, consider a large-scale wireless communication network consisting of K devices and K access points, where each device transmits data to a unique access point. The locations of the devices are governed by a homogeneous Poisson Point Process (PPP) [ABW10; LP17], denoted by Φ, with intensity λ. We assume that the network is uncoordinated, which forms a worst case model for large-scale M2M communication networks.

For an access point at the origin, A_0, served by device 0, the interference at time t from the other devices is given by

$$I_t = \sum_{k \in \Phi_t \setminus \{0\}} r_{k,t}^{-\eta/2} h_{k,t} x_{k,t}, \tag{2.49}$$

where η is the path loss exponent of the interfering links. $h_{k,t}$ is any circularly symmetric complex distributed random variable, which is assumed to be the Rayleigh fading coefficient $h_{k,t} \sim \mathcal{CN}(0,1)$ for the link from device k to the access point A_0. The baseband emission of each interferer k is denoted by $x_{k,t}$. We assume that the real and imaginary parts of $h_{k,t}x_{k,t}$ are symmetric and $e^{j\phi}h_{k,t}x_{k,t} \overset{d}{=} h_{k,t}x_{k,t}$ for all $\phi \in [0, 2\pi)$, which means

that $h_{k,t}x_{k,t}$ is isotropic. This is not a strong assumption and is satisfied, for instance, in the case of Rayleigh fading with circularly symmetric complex Gaussian baseband emissions.

The distance of the access point A_0 to device 0 is denoted by r_d, with distribution F_{r_d}. The signal received by the access point A_0 at time t is then given by

$$y_t = r_{d,t}^{-\eta/2} h_{d,t} x_{d,t} + I_t + N_t, \tag{2.50}$$

where $h_{d,t}$ is a circularly symmetric complex channel fading coefficient, e.g., a Rayleigh $h_{d,t} \sim \mathcal{CN}(0,1)$, and $x_{d,t}$ is the baseband emission for the typical user. The additive white zero-mean Gaussian noise $N_t \sim \mathcal{CN}(0,\sigma^2)$ corresponds to thermal noise at the access point.

In fact, the received signal by an access point A_0 in the presence of dynamic interference for the interference-limited setting, i.e., where I dominates over the noise, can be seen as the output of a memoryless additive noise channel. Therefore, the output y is converted to

$$y = r_d^{-\eta/2} h_d x_d + I, \tag{2.51}$$

where the time subscript was dropped due to the memoryless channel.

It is necessary to characterize the statistics of I in order to evaluate communication in dynamic interference. The foundation of this characterization is the theory of isotropic α-stable random variables, which we now review.

2.5.1 Interference Characterization

We show in this section that the interference I is, in fact, an isotropic α-stable random variable. For this purpose, consider a complex $I = I_1 + iI_2$, where I_1 and I_2 are the real and imaginary components, respectively, given that baseband signals are typical complex. The necessary conditions that I has to follow to be an isotropic complex α-stable are presented in Proposition 6 (page 27). In order to prove these conditions, let $z_k = h_k x_k$ and denote the real and imaginary parts as $z_{k,r}$ and $z_{k,i}$. The interference can then be written as

$$I = \sum_{k=1}^{\infty} r_k^{-\eta/2} (z_{k,r} + i z_{k,i}), \tag{2.52}$$

where each device in Φ_t is indexed by an integer $k = 1, 2, \ldots$ and we can ignore the effect of the serving device. One can notice this by Slivnyak's theorem that is equal to

Theorem 2 (Slivnyak-Mecke). *Let X be a PPP with intensity measure μ. Thus*

$$\mathbb{E}\left[\sum_{\varepsilon\in X} h(\varepsilon, X\backslash\{\varepsilon\})\right] = \int_{\mathbb{R}^d} \mathbb{E}\left[h(\varepsilon, X)\right]\mu(d\varepsilon), \tag{2.53}$$

where h is an arbitrary non-negative measurable function.

Proof. A direct proof is detailed in [MW03][Theorem 3.2]. □

This theorem is a means of showing that a property seen from a point at x does not depend on having a point x in Φ_t. In other words, the stationarity of Φ_t guarantees that the position of an arbitrary point does not change the probability.

Recall that the distances, $\{r_k\}_{k=1}^{\infty}$, are from points in a PPP to the origin. It follows that r_k^2 is an one-dimensional PPP with intensity $\lambda\pi$, which is obtained by changing the coordinates from Cartesian into polar and applying the mapping theorem(see [IH98, Proposition 1] for more details). By an application of the LePage series representation of symmetric α-stable random variables presented in Proposition 4 (page 23), it also follows that I converges almost surely to

$$I = Z_r + iZ_i, \tag{2.54}$$

where Z_r and Z_i are real and symmetric $4/\eta$-stable random variables.

By Proposition 8 (page 26), the induced random vector $\mathbf{I} = (Z_r, Z_i)^T$ is a symmetric $4/\eta$-random vector, which implies that condition **C1** presented in Proposition 6 (page 27) holds.

To show that condition **C2** holds, recall that $e^{i\phi}h_k x_k \overset{d}{=} h_k x_k$ for any $\phi \in [0, 2\pi)$. This implies that I is isotropic and hence I is an isotropic $4/\eta$-stable random variable.

In order to characterize the statistics of the interference I, all that remains is to obtain the parameter $\sigma_{\mathbf{N}}$ in the scale mixture representation stated in Theorem 1 (page 28). Again using the LePage series representation in Proposition 4 (page 23), the scale parameters of the real and imaginary parts of I are equal to $\left(\pi\lambda C_{\frac{\eta}{4}}^{-1}\mathbb{E}[|\mathrm{Re}(h_k x_k)|^{\frac{4}{\eta}}]\right)^{\frac{\eta}{4}}$. Using Property 10 (page 29), we then have

$$\sigma_{\mathbf{N}} = (\pi\lambda C_{\frac{\eta}{4}}^{-1}\mathbb{E}[|\mathrm{Re}(h_k x_k)|^{\frac{4}{\eta}}])^{\frac{\eta}{4}}, \tag{2.55}$$

where $C_{\frac{\eta}{4}}$ is given in (2.57).

In summary, the interference I is characterized as follows.

Proposition 8. *The interference I is an isotropic α-stable random variable, with $\alpha = \frac{4}{\eta}$ and parameter*

$$\sigma_{\mathbf{N}} = \left(\pi\lambda C_{\frac{4}{\eta}}^{-1}\mathbb{E}[|\mathrm{Re}(h_k x_k)|^{\frac{4}{\eta}}]\right)^{\frac{\eta}{4}}, \tag{2.56}$$

where

$$C_\alpha = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)}, & \text{if } \alpha \neq 1 \\ 2/\pi, & \text{if } \alpha = 1. \end{cases} \tag{2.57}$$

The main consequence of Proposition 8 (page 33) is that the channel in (2.51) is a memoryless additive isotropic α-stable noise ($AI\alpha SN$) channel. Unlike circularly symmetric complex Gaussian noise, the real and imaginary parts of I are not independent, a consequence of the dependency created in the sub-Gaussian representation in Theorem 1 and present in both components. Therefore, it is not possible to treat an $AI\alpha SN$ channel as parallel real α-stable noise channels. Instead, it is useful to view the $AI\alpha SN$ channel as a vector channel, where the real vector-valued noise is the vector induced by the isotropic α-stable interference I.

2.6 Chapter conclusion

Noises and interferences present in many modern communications systems were introduced. To cope with these features, different models for impulsive interference were depicted. Firstly, we started with the well-studied Gaussian distribution, which leads to an inappropriate characterization of impulsive scenarios due to the tail behavior that decays exponentially. The simplicity of its probability density function, characteristic function and the tail behavior as well as the reason that motivates the use of this model were outlined. Next, to encompass impulsive noises and interferences, non-Gaussian models were introduced. In a first step, the different classes of Middleton distributions were detailed to illustrate its physical perspective although the difficulty to use in practical systems. Lastly, we focus on the α-stable random variables, which will be necessary all along this thesis work. It has been suggested that in the class of heavy-tailed distributions the stable family are accurate solutions to model impulsive noise. Many definitions and properties were demonstrated in the univariate and bivariate cases. The tail was shown to be heavier when compared to Gaussian models and the series representation was used to study scenarios that give rise to α-stable interferences. Finally, the proof for the α-stable model in a communication setting (interference arising from an homogeneous Poisson Point Process with no guard zone) is presented.

Chapter 3

Capacity of Additive α-Stable Noise Channels

Capacity for continuous channels are in fact difficult to be address, a notorious exception is the expression derived by Shannon for the Gaussian case with a power constraint, which will be revisited. In this chapter, we are concerned with the capacity of an additive impulsive channel using α-stable models. A new proposed model allows the derivation of lower and upper bounds, as well as, the existence and uniqueness of the optimal input. An algorithm is proposed to compare the achievable bounds. Moreover, properties and parametrization of the lower bound are studied.

3.1 Problem formulation and related works.

THE central concern in communication systems is to transmit with fewer errors, using high speeds to guarantee an efficient communication on noisy channels. The capacity characterizes the maximum rate where the error probability can be arbitrary close to zero.

Fortunately, the channel capacity in the class of discrete memoryless channels is well understood as notably demonstrated by Shannon [Sha48]. Although, when we turn to the continuous channels, it has been proven difficult to characterize the capacity, one exception is the case of the linear additive white Gaussian noise (AWGN) constrained by power [Sha48]. To illustrate some results obtained, Table 3.1 shows the optimal input distributions that induce the capacity for some discrete and continuous cases given input constraints.

TABLE 3.1: Some known optimal distributions for distributions

Type of distribution	Constraints	Optimal distribution
Continuous and infinite support	$\mathbb{E}[X^2] < +\infty$	Gaussian distribution
Continuous and $x > 0$	$\mathbb{E}[X] < +\infty$	Exponential distribution
Continuous on $[a, b]$	-	Uniform distribution
Discrete and $x > 0$	$\mathbb{E}[X] < +\infty$	$\frac{1}{Z}\exp(-\lambda n)$

Focusing on impulsive models, approximations of Middleton noise have been known since Middleton's early work. The capacity of channels

with these simpler models have been derived and some examples may be mentioned. For the capacity of the Bernoulli-Gaussian channel, results divided into low and high input power regions were derived in [HTLN12a]. An optimal input to transform the channel output into Gaussian and the conditions of its existence have been shown. Its premise is the fact that Gaussian input maximizes the mutual information over all input distributions subjects to a power constraint $\mathbb{E}[X^2] < P$, as we will revisit Shannon's result in Section 3.1.2. In fact, they have proven that for high input power scenario, the Gaussian input is asymptomatically optimal in Bernoulli-Gaussian channels. For low power regions, bounds for the capacity of the Bernoulli-Gaussian channel were derived if assumed knowledge of the impulsive noise, as well as, a lower bound when using Gaussian inputs.

Another example is the alternative Gaussian mixture models that have been studied in [CE12]. They produce as result an outage probability performance study for many diversity combining receivers, as well as, signal-to-interference ratio (SIR) analytical expressions. Using a framework to analyze common interferences scenarios, they also proposed novel diversities to improve the outage probability.

Besides, when analyzing Middleton model itself, Wiklundh and Stenumgaard and Tullberg have shown an analytical expression by means of a non-physical perspective using aperiodic, stationary Markov chains to represent the additive white class-A noise (AWCN) [WST09], where an impulsive channel has a larger capacity than an AWGN channel. In addition, a MIMO case was studied in [NAHV14], in which a channel capacity was derived having different levels of channel state information (CSI).

Being the focus of this chapter, we turn the attention to $AS\alpha SN$ channels. There have recently been several new results characterizing the capacity of real additive α-stable noise channels subject to a range of constraints. Fahs and Abou-Faycal [FAF12] have shown that under the constraint $\mathbb{E}[|X|^r] \leq c,\ r > 1$ with $1 < \alpha \leq 2$ the second order moment does not seem adapted to measure the power and the optimal input is compactly supported and discrete. As a consequence, if a power infinite is allowed, the optimal input does not lead to a rate increase. In [FAF14], they characterized the capacity with following logarithm constraint for the input distribution

$$\mathbb{E}\left\{\ln\left[\left(\frac{A+\gamma}{A}\right)^2+\left(\frac{X}{A}\right)^2\right]\right\} \leq \ln 4, \tag{3.1}$$

where $A \geq \gamma$ and $\gamma > 0$ is the dispersion. They have shown that a Cauchy distributed input is optimal for Cauchy noise ($\alpha = 1$), which leads to a closed-form expression for the capacity. In a more recent work, [FAF16], they have presented results for more general constraints including $\mathbb{E}[|X|^r] \leq c,\ 0 \leq r < \alpha$, and shown that the optimal input distribution is

again compactly supported and discrete.

In [PW10c], an analytical capacity expression, nevertheless with a second order moment constraint and assuming that the transmitted signal is Gaussian, was obtained for asynchronous interferers scattered according to a spatial Poisson process in an infinity plane. The capacity channel has the form

$$C(G_0, \mathcal{P}) = \mathbb{E}_{\alpha_0} \left\{ \log_2(1 + \alpha_0^2 \Psi) | G_0, A \right\}, \tag{3.2}$$

where the r.v. A has a skewed stable distribution, the shadowing $G_0 \sim N(0,1)$, Ψ is the received signal-to-interference-plus-noise ratio and $\mathcal{P}$ the position of the interferers.

Although an expression without a closed-form for the error probability in the presence of α-stable interference was derived in [PW10a] and considering the obtention of numerical results via the Blahut-Arimoto algorithm for the capacity of the $AS\alpha SN$ channel studied containing a power constraint on the source [WKZ11], $AS\alpha SN$ channels are not well understood and there are currently no characterizations of the achievable rates for $AS\alpha SN$ channels with $0 < \alpha < 2$.

Before the study of α-stable Noise Channels, we outline the information measure machinery necessary to this end. Next, the capacity in AWGN noise channels is described, which will be adopted as a comparison metric.

3.1.1 Information Measures

In this section, we first introduce some notions of measure theory that can be taken into account in order to extend the class of random variables. Firstly, due to the differences that appeared because of the previous development of probability theory without the connection with measure theory, it should be pointed out a small vocabulary between a probability and the measures theory depicted in Table 3.2 [Fol13, Section 10.1], which we will use interchangeably

Analysts' Term	Probabilists' Term
Measure space $(X, \mathcal{M}, \mu)$	Sample space $(\Omega, \mathcal{B}, P)$
Measurable set	Event
Measurable real valued function f	Random variable X
σ-algebra	σ-field
Convergence in measure	Convergence in probability
Almost everywhere a.e.	Almost surely
Borel probability measure on $\mathbb{R}$	Distribution
Fourier transform of a measure	Characteristic function of a distribution

TABLE 3.2: Vocabulary between the probability and measures theories

Now, we wish to gather the basic definitions that are important to be mentioned for the work in the field of information measures. The σ-algebra on X is defined as follow

Definition 9 (σ-algebra)**.** *Let X be a nonempty set equipped with a collection $\mathcal{M}$ of subsets E, in which E is a proper subset of X, such that $\mathcal{M}$ is closed under complements and countable unions, in other words, whenever there is a sequence E_j of sets in $\mathcal{M}$, then $\cup_{j=1}^{\infty} E_j$ is also in $\mathcal{M}$. Moreover, whenever $E \in \mathcal{M}$, $X \setminus E$ is also in $\mathcal{M}$. Finally, $\emptyset \in \mathcal{M}$. Then, $\mathcal{M}$ is a σ-algebra on X.*

In fact, σ in σ-algebra concerns sum or union and, in turn, algebra refers to formal operations. It is worth noting that the minimal (trivial) σ-algebra concerns the empty $\emptyset$ and X sets. Furthermore, a measurable space regards a pair $(X, \mathcal{M})$.

Definition 10 (probability measure)**.** *Let $\mathcal{M}$ be a σ-algebra on X, such that $\mu : \mathcal{M} \to [0, 1]$ satisfies the conditions:*

1. $\mu(\emptyset) = 0$,
2. $\mu(X) = 1$, *known as normalization,*
3. *if $E_1, E_2, \ldots$ is a disjoint sequence (i.e. $E_i \cap E_j = \emptyset$, whenever $i \neq j$) in $\mathcal{M}$, then*

$$\mu\left\{\cup_{j=1}^{\infty} E_j\right\} = \sum_{j=1}^{\infty} \mu(E_j),$$

called countable additivity. Then μ is a probability measure on $\mathcal{M}$ and the triple $(X, \mathcal{M}, \mu)$ is called a probability measure space.

In addition, if there exists a measurable space $(X, \mathcal{M})$, any $E \in \mathcal{M}$ is a measurable set.

Definition 11 (measurable function)**.** *Given two measurable spaces $(X, \mathcal{M})$ and $(X', \mathcal{M}')$, a function $f : X \to X'$ is said $(\mathcal{M}, \mathcal{M}')$-measurable (or just measurable) if for every $E \in \mathcal{M}'$, then $f^{-1}(E) \in \mathcal{M}$.*

Definition 12 (Borel σ-algebra)**.** *Assume that $\mathcal{T}$ is the topology of X (the collection of all open subsets of X), in which X is a topological space. Thus, the σ-algebra generated by $\mathcal{T}$ that produces the smallest σ-algebra containing all open sets is called a Borel σ-algebra $\mathcal{B}$ on X. It is denoted as $\mathcal{B} = \mathcal{M}(\mathcal{T})$.*

For instance, $\mathcal{B}(\mathbb{R})$ represents the smallest σ-algebra containing the open intervals of $\mathbb{R}$. Moreover, a measure on $\mathcal{B}$ is called a Borel measure.

Definition 13 (random variable)**.** *Assume a probability measure space $(\Omega, \mathcal{M}, P)$ and a measurable space $(\mathbb{Y}, \mathcal{F})$. Thus a measurable function X is a $\mathbb{Y}$-valued random variable. Particularly, a measurable function called random variable has $\mathbb{Y} = \mathbb{R}$ and is a function from $(\Omega, \mathcal{M}, P)$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.*

In fact, assuming a random variable X on $(\Omega, \mathcal{M}, P)$ induces a measure μ on $(\mathbb{R}, \mathcal{B})$, which is said a distribution measure or just distribution and is represented as $\mu(B) = P[X \in B]$.

Definition 14. *Let $(X, \mathcal{M}, \mu)$ be a probability measure space and $A_j \in \mathcal{M}$, where $j = 1, \ldots, r$. Then, a measurable partition of X is $X = \cup_{j=1}^{r} A_j$.*

We first present the entropy for discrete measure given by

Definition 15 (Entropy)**.** *Let $\mathcal{A} \in \Omega_X$, in which Ω_X represents the collection of all partitions of X formed by finitely measurable subsets and $\mathcal{A}$ is the partition $X = \cup_{j=1}^{r} A_j$ so that the entropy can be denoted as*

$$h(\mathcal{A}) \triangleq -\sum_{j=1}^{r} p(A_j) \log(p(A_j)).$$

Regarding the unit measure, when using logarithms of base two, the entropy unit is bits, while natural logarithms corresponds to nats.

Definition 16 (Conditional entropy)**.** *The conditional entropy of $\mathcal{A}$ given $\mathcal{B}$ is $h(\mathcal{A}, \mathcal{B}) - h(\mathcal{B})$, where $\mathcal{B}$ is the partition $X = \cup_{j=1}^{s} B_j$.*

Definition 17 (Mutual information)**.** *The mutual information of $\mathcal{A}, \mathcal{B}$, where $\mathcal{A}, \mathcal{B} \in \Omega_X$ is given by*

$$I(\mathcal{A}, \mathcal{B}) \triangleq h(\mathcal{B}) - h(\mathcal{B}|\mathcal{A}).$$

After presenting the discrete definitions of entropy with discrete measures, the extension to more general cases is necessary for the next chapters of this thesis. As such, the measures can be classified as

- discrete, which contains countable values.
- continuous, when $\Pr\{X = x\} = 0$.
- mixed, as the name suggests, a mixed version of both.

However, simple replacing the summations by integrals into the previous definitions is not a sufficient operation. In fact, the discrete entropy diverges when smaller divisions are used. Firstly, we present the notion of absolute continuity as

Definition 18. *Let $(X, \mathcal{M}, \mu)$ and ν be a measure defined on $\mathcal{M}$. A measure ν is called absolutely continuous with respect to μ, written $\nu \ll \mu$, if $\forall E \in \mathcal{M} : \mu(E) = 0 \implies \nu(E) = 0$.*

In order to achieve the continuous case that will be used during the following chapters, the integral of a non-negative measurable function f concerning a measure μ is depicted, which is called expectation or expected value and is defined as

Definition 19. *Let $(X, \mathcal{M}, \mu)$ be a probability space and f be a measurable function of X, then the expectation of f over X is*

$$\mathbb{E}_\mu[f(x)] = \int_X f(x)d\mu = \int_X f(x)\mu(dx),$$

The right-hand side consists of a Lebesgue-Stieljes integration, covering discrete, continuous and mixed cases. In addition the Kullback-Leibler divergence (also known as relative entropy) can be represented as

Definition 20 (Relative entropy, Kullback–Leibler divergence)**.** *Let the measures P and Q having a common measurable space $(X, \mathcal{M})$, then*

$$D(P||Q) = \begin{cases} \int \log f dP = \int f \log f dQ, & \text{if } P \ll Q, \\ \infty, & \text{if } P \text{ is not absolutely continuous} \\ & \text{with respect to } Q. \end{cases}$$

Moreover, a Radon-Nikodym derivative $f = dP/dQ$ exists.

Proof. See [KL51] and [Csi92] for details. More information about Radon-Nikodym derivative in [May79]. □

It may be noticed that the relative entropy depicts the distance between two probability measures and is non-negative. However it is not generally symmetric under the positions of P and Q, i.e. $D(P||Q) \neq D(Q||P)$. As a consequence, it does not satisfy the triangle inequality (i.e., for any two real numbers x and y, $|x + y| \leq |x| + |y|$), restricting its use as a metric. To represent the discrete relative entropy, consider P and Q as discrete measures, such that $P \ll Q$ and then the Radon-Nikodym derivative dP/dQ guarantees the existence of probability mass functions $p(x)$ and $q(x)$, concluding the discrete case. In turn, if P and Q are probability measures on X and are absolutely continuous with respect to μ, then the probability density functions p and q exist, yielding

$$\begin{aligned} D(P||Q) = \mathbb{E}_Q\left[f \log(f)\right] &= \int f \log(f) dQ \\ &= \int \frac{dP}{dQ} \log\left(\frac{dP}{dQ}\right) dQ \\ &= \int p(x) \log\left(\frac{p(x)}{q(x)}\right) d\mu, \end{aligned} \tag{3.3}$$

where the expected value is presented in Definition 19.

Property 11. *Let two probability density functions p and q, then*

$$D(p||q) \geq 0$$

Proof.

$$\begin{aligned} D(p||q) &= \int p(x) \log \frac{p(x)}{q(x)} dx \\ &= \mathbb{E}_p \left[\log \left(\frac{p(x)}{q(x)} \right) \right] \\ &= \mathbb{E}_p \left[-\log \left(\frac{q(x)}{p(x)} \right) \right] \\ &\geq -\log \mathbb{E}_p \left[\frac{q(x)}{p(x)} \right] \\ &= -\log \int p(x) \frac{q(x)}{p(x)} dx \\ &= -\log \int q(x) dx \\ &= -\log 1 \\ &= 0, \end{aligned}$$

where the Jensen's inequality is used [Kuc09, Section 8.1]. □

Definition 21 (Mutual information [Bak79])**.** *Let $(X, \mathcal{M})$ and $(Y, \mathcal{F})$ be measurable spaces, μ_{XY} a probability measure on the product space $\mathcal{M} \times \mathcal{F}$, in addition, μ_X and μ_Y the projections of μ_{XY}. $\mu_X(E_1) = \mu_{XY}(E_1 \times Y)$; $\mu_Y(E_2) = \mu_{XY}(X \times E_2)$ and $\mu_X \otimes \mu_Y$ is the product measure. Then, if $\mu_{XY} \ll \mu_X \otimes \mu_Y$, the mutual information is defined as*

$$I(\mu_{XY}) \triangleq \int_{X \times Y} \log \left[\frac{d\mu_{XY}}{d\mu_X \otimes \mu_Y}(x, y) \right] d\mu_{XY}(x, y),$$

and $I(\mu_{XY}) = \infty$ otherwise.

Corollary 3. *Suppose X and Y be two continuous random variables with joint probability density function $p(x, y)$ and marginal probability density $p(x)$ and $p(y)$ respectively. Then, the mutual information $I(X, Y)$ is*

$$\begin{aligned} I(X;Y) &= D\left(p(x,y)||p(x)p(y)\right) \\ &= \int_Y \int_X p(x,y) \log \left[\frac{p(x,y)}{p(x)p(y)} \right] dxdy. \end{aligned}$$

Definition 22 (Differential entropy)**.** *The entropy is defined as the mutual information between a random variable X and itself, yielding*

$$\begin{aligned} h(X) &= I(X;X), \\ &= -\int_X p(x) \log(p(x)) dx \end{aligned}$$

which is possible to see in the form $I(X, X) = h(X) - h(X|X) = h(X)$, because $h(X|X) = 0$ for any r.v. X. In fact, it extends the Definition 15 to the continuous case. This seems to be similar to the discrete Shannon entropy, but differs since it can be positive or negative.

Property 12. *The differential entropy is invariant to translation, as such*

$$h(X + c) = h(X).$$

This property follows directly from the definition.

Property 13. *Let X be a continuous random variable, then the scaling property of a differential entropy can be written as*

$$h(aX) = h(X) + \log(|a|),$$

in which $a \neq 0$.

Definition 23 (Weak convergence [LW14])**.** *Let $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures on $(X, \mathcal{M})$. Then, μ_n converges weakly to a probability measure μ on $(X, \mathcal{M})$, written as $\mu_n \Rightarrow \mu$, if*

$$\int f d\mu_n \to \int f d\mu,$$

$\forall f \in C_b(X)$, where $C_b(X)$ denotes the set of all continuous and bounded functions $f : X \to \mathbb{R}$.

We preset the concept of tightness as follow

Definition 24 (Tightness [Bil99])**.** *Let a probability measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. It is called tight if for each ϵ there exists a compact set $\mathcal{K}_\epsilon$ such that $\mu(\mathcal{K}_\epsilon) > 1 - \epsilon$.*

Definition 25 (Weakly closed [LVG12])**.** *Let X be a normed linear space. A set $E \subset X$ is said to be weakly closed in X if all its weak limit points are in E. Thus if $\{\mu_n\}_{n\in\mathbb{N}} \subset E$, then $\mu_n \Rightarrow \mu_0 \in X$ implies $x_0 \in E$.*

Theorem 3 (Portmanteau theorem [Bil99])**.** *For probability measures $\{\mu_n\}_{n\in\mathbb{N}}$, μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, the following statements are equivalent:*

- *$\int f d\mu_n \to \int f d\mu$, for all $f \in C_b(X)$; i.e. $\mu_n \Rightarrow \mu$.*
- *$\liminf_{n\to\infty} \geq \int f d\mu$, for every lower semicontinuous f bounded from below.*

3.1.2 Additive Gaussian Noise Channels Capacity

In this section, the emblematic expression created in 1948 in Shannon's publication [Sha48] is revisited. The importance of this derivation for our further study remains on the use of two features, namely, the *stability property* and variance finiteness. The former allows the sum of two i.i.d. stable distributions and, consequently, of normal distributions and will be revisited in future chapters when deriving new achievable bounds. The latter is a limiting factor in α-stable models.

The result arises in the case that N is a Gaussian noise, so that with a channel input X, the channel output is

$$Y = X + N, \tag{3.4}$$

where the noise is independent of X. In fact, it can be shown that a Gaussian distribution for X allows to reach the maximum mutual information I, so called the capacity, which can be written as the following Shannon's formula

$$C = \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma_N^2}\right) \quad \text{in bit/s/Hz}, \tag{3.5}$$

in which the average power is $P = \sigma^2 = \mathbb{E}[X^2]$ and the noise power $\sigma_N^2 = \mathbb{E}[N^2]$. The equation (3.5) gives the maximum rate that a transmission can be reliable over a noisy communication channel containing a Gaussian noise, which leads to the following definition

Definition 26. *The capacity of the channel is described as the supremum of the achievable rates. A rate R is called achievable for a Gaussian channel containing a power constraint P if there exists a sequence of $(2^{nR}, n)$ codes with codewords containing the appropriate constrained power such that the maximal probability of error $\epsilon^{(n)} \to 0$.*

For more details see [CT06, Theorem 7.7.1].

The classical derivation assumes a maximal rate $C = \sup_X I(X;Y)$ bits per sample. As a result, we address in this section the solution of the following optimization problem

$$\begin{aligned} \sup_{f_X(x)} \quad & I(X;Y) \\ \text{subject to} \quad & f_X(x) = 1; \\ & \int_{\mathbb{R}} (x-m)^2 f_X(x)dx \leq \sigma^2, \end{aligned} \tag{3.6}$$

considering X any r.v., $f_X(x)$ its probability density function and m is the mean $\mathbb{E}[X]$. The first constraint $\int f_X(x)dx = 1$ is necessary to explicitly guarantee the probability support. X has a finite mean power, such that $\mathbb{E}[(X - \mathbb{E}[X])^2] \leq \sigma^2 < +\infty$ forms the second constraint.

The demonstration is in two steps: first, the random variable with finite second order moments that leads to the entropy maximization is presented. In the second part, the result is applied to show that the capacity will be obtained with a Gaussian source which allows reaching capacity in an explicit form.

Proof - Step 1

In this section we address the optimization of the entropy of X, as following

$$\begin{aligned} \underset{f_X(x)}{\arg\sup} \quad & h(X) \\ \text{subject to} \quad & f_X(x) = 1; \\ & \int_{\mathbb{R}} (x-m)^2 f_X(x) dx \leq \sigma^2, \end{aligned} \tag{3.7}$$

where a continuous source $f_X(.)$ is assumed, so that its entropy is given by $h(X) = -\int_{\Re} f_X(x) \log_2(f_X(x)) dx$ (see Definition 22, page 41). In order to solve the optimization problem, the first derivation part is written using the Lagrange Function [Ber99] with two constraints, constructed as

$$\begin{aligned} \Phi(f_X(x), \lambda_1, \lambda_2) &= h(x) + \lambda_1 \left(\int_{\Re} f_X(x) dx - 1 \right) \\ &\quad + \lambda_2 \left(\int_{\Re} (x-m)^2 f_X(x) dx - \sigma^2 \right) \\ &= -\int_{\Re} f_X(x) \log_2(f_X(x)) dx + \lambda_1 \left(\int_{\Re} f_X(x) dx - 1 \right) \\ &\quad + \lambda_2 \left(\int_{\Re} (x-m)^2 f_X(x) dx - \sigma^2 \right), \end{aligned} \tag{3.8}$$

where λ_1 and λ_2 are the Lagrange multipliers.

In order to maximize the previous Lagrange function, the derivative given by

$$\frac{d\Phi(f_X(x), \lambda_1, \lambda_2)}{df_X(x)} = -\ln(f_X(x)) - \log_2(e) + \lambda_1 + \lambda_2 (x-m)^2 \tag{3.9}$$

Next, letting this derivative be zero, yields

$$\begin{aligned} f_X(x) &= \exp - \left\{ \log_2(e) + \lambda_1 + \lambda_2 (x-m)^2 \right\} \\ &= \exp \left\{ -1 + \frac{\lambda_1 + \lambda_2 (x-m)^2}{\log_2(e)} \right\}. \end{aligned} \tag{3.10}$$

With the help of the two constraints, the equation can be rearranged. Firstly, a bound for the probability density function is applied as

- using $\int_{\Re} f_X(x)dx = 1$,

$$\int_{\Re} f_X(x)dx = \exp\left\{-1 + \frac{\lambda_1}{\log_2(e)}\right\} \int_{\Re} \exp\left\{\frac{\lambda_2 (x-m)^2}{\log_2(e)}\right\} dx$$

$$\text{it is necessary } \lambda_2 < 0 \text{ and setting } u = \sqrt{\frac{-\lambda_2}{\pi \log_2(e)}}(x-m)$$

$$= \exp\left\{-1 + \frac{\lambda_1}{\log_2(e)}\right\} \sqrt{-\frac{\pi \log_2(e)}{\lambda_2}} \int_{\Re} \exp\left\{-\pi u^2\right\} du$$

$$= \exp\left\{-1 + \frac{\lambda_1}{\log_2(e)}\right\} \sqrt{-\frac{\pi \log_2(e)}{\lambda_2}} = 1, \tag{3.11}$$

where $\int_{\Re} \exp\left\{-\pi u^2\right\} du$ is the Gaussian integral. Consequently

$$\exp\left\{-1 + \frac{\lambda_1}{\log_2(e)}\right\} = \sqrt{-\frac{\lambda_2}{\pi \log_2(e)}} \tag{3.12}$$

In turn, the variance is replaced as following

- using $\int_{\Re}(x-m)^2 f_X(x)dx = \sigma^2$:

$$\int_{\Re}(x-m)^2 f_X(x)dx = \exp\left\{-1 + \frac{\lambda_1}{\log_2(e)}\right\}$$

$$\int_{\Re}(x-m)^2 \exp\left\{\frac{\lambda_2 (x-m)^2}{\log_2(e)}\right\} dx$$

$$= \exp\left\{-1 + \frac{\lambda_1}{\log_2(e)}\right\} \sqrt{2\pi} \left(\frac{\log_2(e)}{-2\lambda_2}\right)^{\frac{3}{2}}$$

$$= \sigma^2. \tag{3.13}$$

Consequently

$$\exp\left\{-1 + \frac{\lambda_1}{\log_2(e)}\right\} \frac{\sqrt{\pi}}{2} \left(\frac{\log_2(e)}{-\lambda_2}\right)^{\frac{3}{2}} = \sigma^2. \tag{3.14}$$

From (3.12) and (3.14), we can get

$$\sqrt{-\frac{\lambda_2}{\pi \log_2(e)}} \frac{\sqrt{\pi}}{2} \left(\frac{\log_2(e)}{-\lambda_2}\right)^{\frac{3}{2}} = \sigma^2 \tag{3.15}$$

Then

$$\lambda_2 = -\frac{\log_2(e)}{2\sigma^2}. \tag{3.16}$$

From (3.12) and (3.16), yields

$$\exp\left\{-1 + \frac{\lambda_1}{\log_2(e)}\right\} = \sqrt{\frac{1}{2\pi\sigma^2}} \tag{3.17}$$

Finally, applying to (3.10), we have

$$\begin{aligned} f_X(x) &= \exp\left\{-1+\frac{\lambda_1}{\log_2(e)}\right\}\exp\left\{\frac{\lambda_2\,(x-m)^2}{\log_2(e)}\right\} \\ &= \sqrt{\frac{1}{2\pi\sigma^2}}\exp\left\{-\frac{\log_2(e)}{2\sigma^2}\frac{(x-m)^2}{\log_2(e)}\right\} \\ &= \sqrt{\frac{1}{2\pi\sigma^2}}\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}, \end{aligned} \tag{3.18}$$

which is exactly a Gaussian distribution with variance σ^2.

Step 2

In order to demonstrate that the mutual information is maximized when using an input induced by a Gaussian distribution, one may notice that $I(X;Y) = h(Y) - h(Y|X) = h(Y) - h(X+N|X) = h(Y) - h(N)$ due to the independence of noise and input. It is clear that $I(X;Y)$ is maximized (over $\Pr(X)$) if $h(Y)$ is maximized, knowing that $I(X;Y) \geq 0$ (see Property 11, page 40, and Corollary 3, page 41).

Recall that the channel output presented in (3.4) is formulated as $Y = X + N$, where N is assumed Gaussian and denoted by $\mathcal{N}(m, \sigma_N^2)$. By (3.7), $h(Y = X + N)$ is maximized by a Gaussian distribution. Consequently, X has to be also Gaussian due to the *stability property*, which indicates that the sum of two i.i.d normal random variables is normally distributed [Wal96, Section 34.5].

For this result, assume $h(X) = h(\phi(x))$ the entropy of a Gaussian distribution as follow

$$\begin{aligned} h(\phi(x)) &= \int \phi(x)\log_2\left(\phi(x)\right)dx \\ &= \int \phi(x)\log_2\left(\frac{1}{2\pi\sigma^2}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right)dx \\ &= \frac{1}{2}\log_2\left(2\pi\sigma^2\right)\int\phi(x)dx + \frac{1}{2\sigma^2\ln(2)}\int(x-\mu)^2\phi(x)dx \\ &= \frac{1}{2}\log_2\left(2\pi\sigma^2\right) + \frac{\log_2(e)}{2} \\ &= \frac{1}{2}\log_2\left(2\pi e\sigma^2\right). \end{aligned} \tag{3.19}$$

To conclude the proof, we can now calculate the capacity of the Gaussian channel given by $X \sim \mathcal{N}(0, P)$ and $N \sim \mathcal{N}(0, \sigma_N^2)$. Then, as X and N are

independent, we obtain $\text{var}(Y) = \text{var}(X + N) = P + \sigma_N^2$. Finally:

$$\begin{aligned} I(X,Y) &= h(Y) - h(N) \\ &= \frac{1}{2}\log_2\left(2\pi e(P+\sigma_N^2)\right) - \frac{1}{2}\log_2\left(2\pi e\sigma_N^2\right) \\ &= \frac{1}{2}\log_2\left(\frac{2\pi e(P+\sigma_N^2)}{2\pi e\sigma_N^2}\right) \\ &= \frac{1}{2}\log_2\left(1+\frac{P}{\sigma_N^2}\right), \end{aligned} \tag{3.20}$$

and the capacity problem is solved analytically.

The demonstration is important to understand that the Gaussian noise is, in fact, an exception in the continuous distributions. The finiteness of the second order moment helps to achieve the result, which is not always the case when using α-stable models to impulsive noise as will be clear in the following section.

3.2 Additive α-stable Noise Channels Capacity

A proposed model for interference is via the symmetric α-stable distributions. This approach leads to the additive symmetric α-stable noise ($AS\alpha SN$) channel given by

$$Y = X + N, \tag{3.21}$$

where the noise is distributed according to the symmetric α-stable distribution.

The difficulty in characterizing the capacity of $AS\alpha SN$ channels is in part due to the fact that a power constraint $\mathbb{E}[X^2] \leq P$ is typically imposed and, unlike the Gaussian case, the second moment of α-stable distributions is infinite for $\alpha < 2$. As such, even lower bounds are challenging since the stability property of α-stable random variables cannot be applied. Combined with the fact that the only analytical results for the capacity due to Fahs and Abou-Faycal were obtained using different constraints, this suggests that a more appropriate approach is to modify the constraint.

From a practical perspective, adopting non-second order constraints can be motivated in both wireless and molecular communication systems. In wireless networks, there are generally both amplitude [OU11; RQZ05] and power constraints; however, it is common practice in the presence of Gaussian noise to relax the amplitude constraint in order to obtain a tractable rate expression and gain design insights. A similar approach is also possible in the $AS\alpha SN$ channel, where a fractional moment constraint is considered and the amplitude constraint is relaxed. In the case of the molecular timing channel, information is encoded into the time a molecule is released. As

such, a first-order constraint is natural and has been studied under various noise models in [SEA12; LMG14].

In this chapter, we adopt the constraint

$$\mathbb{E}[|X|] \leq c, \quad c > 0. \tag{3.22}$$

A key feature of our choice of the constraint is to obtain a new, tractable upper and lower bounds for the capacity of the $AS\alpha SN$ channel with α in $(1, 2]$. In particular, we show that the capacity is lower bounded by

$$C \geq \frac{1}{\alpha} \log_2 \left(1 + \left(\frac{c}{\mathbb{E}[|N|]} \right)^{\alpha} \right), \tag{3.23}$$

obtained by matching the input and noise distributions.

We investigate the tightness of our bounds by numerically approximating the capacity using the Blahut-Arimoto algorithm [Ari72; Bla72], which provides evidence that our lower bound is, in fact, very tight for α near 2 ($\alpha = 2$ corresponds to the Gaussian noise channel). This is important as our lower bound is easy to work with, facilitating further analysis and optimization.

The remainder of this section is organized as follows. In Section 3.2.1, we define the capacity optimization problem. In Section 3.2.2, we prove existence and uniqueness of the optimal input distribution, and we derive upper and lower bounds on the capacity. In Section 3.2.5, we numerically compute the capacity using the Blahut-Arimoto algorithm and make a comparison with our upper and lower bounds. To do this, we provide parameters that yield accurate approximations within ≈ 0.01 bits. In Section 3.2.6, more properties regarding the lower bounds are depicted, such as the bending point and the parametrization choice.

3.2.1 Problem Formulation

The channel considered is additive with symmetric α-stable noise, N, given by (3.21). We constraint our study in this chapter to real case, X and N are real valued, and symmetric α-stable distributions with $1 < \alpha < 2$. In that case, we remind that the characteristic function given by

$$\begin{aligned} \phi_N(t) &= \mathbb{E}[e^{iNt}] \\ &= \exp\left(-\gamma^{\alpha}|t|^{\alpha}\right), \ t \in \mathbb{R}. \end{aligned} \tag{3.24}$$

As can be observed, the characteristic function is defined by only two parameters, the exponent α, restricted in the interval $(1, 2]$, and the dispersion γ.

Capacity Optimization Problem

Let $\mathcal{B}(\mathbb{R})$ be the Borel σ-algebra on $\mathbb{R}$ and let $\mathcal{P}$ denote the collection of Borel probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ equipped with the topology of weak convergence (see Definition 23, page 42). We define the capacity of the $AS\alpha SN$ channel as the solution to the following optimization problem.

$$\begin{aligned} \underset{\mu \in \mathcal{P}}{\text{maximize}} \quad & I(X;Y) \\ \text{subject to} \quad & \mathbb{E}_{\mu}[|X|] \leq c, \end{aligned} \tag{3.25}$$

where $I(X;Y)$ is the mutual information of the channel in (3.21), and μ is the probability measure of X. If it exists, we denote the probability density function of X as p_X and the probability density function of Y as $p_Y(\cdot;\mu)$, which is parameterized by the input probability measure μ.

In the following section, we show this choice for the constraint in (3.25) leads to tractable upper and lower bounds, unlike other standard choices such as $\mathbb{E}[X^2] \leq P$.

3.2.2 Properties of the Capacity

In this section, we prove several properties of the capacity defined by (3.25). To begin, we show that there exists a unique solution to (3.25). This forms the basis for new upper and lower bounds, which we derive in Sections 3.2.2 and 3.2.2, respectively.

Existence and Uniqueness

Denote $\Lambda(c)$ as the set of probability measures μ with support $\mathbb{R}$, such that $\mathbb{E}_{\mu}[|X|] \leq c$. We first prove that the capacity achieving distribution μ^* exists. To do this, we need to show that $\Lambda(c)$ is compact in $\mathcal{P}$. In turn, if $I(X;Y)$ is continuous on $\Lambda(c)$, the extreme value theorem then implies that the mutual information achieves its supremum on $\Lambda(c)$.

We first need to show that $I(X;Y)$ is continuous on $\Lambda(c)$.

Theorem 4. *$I(X;Y)$ is continuous on $\Lambda(c)$.*

Proof. Suppose $\mu_k \Rightarrow \mu$, which means that μ_k converges to μ weakly. Now, the mutual information $I(X_k; X_k + Z)$, where $Z \sim S_{\alpha}(\gamma, 0, 0)$ can be written as

$$\begin{aligned} I(X_k; X_k + Z) &= h(X_k + Z) - h(X_k + Z | X_k) \\ &= h(X_k + Z) - h(Z). \end{aligned} \tag{3.26}$$

Note that $h(Z)$ is bounded, which follows by applying Property 4 (page 24). It also means that $I(X_k; X_k + Z)$ is bounded by using the condition that $\mathbb{E}_{\mu_k}[|X_k|] \leq c$ for all $\mu_k \in \Lambda(c)$ and applying the following lemma [WV12, Lemma 3]

Lemma 2. *Let Z have a density with $h(Z) > -\infty$. Let $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be an increasing continuous function that satisfies the following conditions:*

1. $\int_{\mathbb{R}_{\geq 0}} \exp\{-\psi(x)\}dx < \infty$
2. *For any $0 \leq \lambda \leq 1$, there exists $a_\lambda, b_\lambda, c_\lambda \geq 0$, such that $\psi(\lambda x + (1-\lambda)y) \leq a_\lambda \psi(x) + b_\lambda \psi(y) + c_\lambda, \forall x, y \geq 0$ If $\mathbb{E}[\psi(|X|)] < \infty$ and $\mathbb{E}[\psi(|Z|)] < \infty$*

then

1. *$I(X, Z, snr) = I(X; \sqrt{snr}X + Z) < \infty$ for all $snr \geq 0$;*
2. *$snr \mapsto I(X, Z, snr)$ is continuous on $\mathbb{R}_{\geq 0}$.*

Now consider

$$\lim_{k\to\infty} h(X_k + Z) = -\lim_{k\to\infty} \int_{-\infty}^{\infty} p_{Y_k}(x) \log p_{Y_k}(x) dx. \tag{3.27}$$

Since $I(X_k; X_k + Z)$ and $h(Z)$ are bounded, it follows that $h(X_k + Z)$ is also bounded.

Let $Y_k = X_k + Z$. Then,

$$p_{Y_k}(x) = \int_{-\infty}^{\infty} p_Z(x - y)\mu_k(dy). \tag{3.28}$$

In addition, we present the dominated convergence theorem given by

Theorem 5 (Lebesgue dominated convergence theorem [BR95])**.** *Suppose $f_n : \mathbb{R} \to [-\infty, \infty]$ are (Lebesgue) measurable functions such that the pointwise limit $f(x) = \lim_{n\to\infty} f_n(x)$ exists. Assume there is an integrable $g : \mathbb{R} \to [0, \infty]$ with $|f_n(x)| \leq g(x)$ for each $x \in \mathbb{R}$. Then f is integrable as is f_n for each n, and*

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n d\mu = \int_{\mathbb{R}} \lim_{n\to\infty} f_n d\mu = \int_{\mathbb{R}} f d\mu.$$

By applying the dominated convergence theorem, we can swap the integral and the limit. Analyzing (3.27), this means that to prove the desired result, we need to use (3.28) to show that

$$\lim_{k\to\infty} \int_{-\infty}^{\infty} p_Z(x - y)\mu_k(dy) = \int_{-\infty}^{\infty} p_Z(x - y)\mu(dy). \tag{3.29}$$

To do this, note that p_Z is bounded and continuous, since the probability density function of $S\alpha S$ random variables is absolutely continuous. Applying the definition of weak convergence (see Definition 23, page 42) then concludes the proof. □

We now turn to showing that $\Lambda(c)$ is compact in $\mathcal{P}$.

Theorem 6. *The set of probability measures $\Lambda(c)$ is compact in the topology of weak convergence. Moreover, the capacity achieving probability measure μ^* exists.*

Proof. By Prokhorov's theorem, $\Lambda(c)$ is compact if it is tight (see Definition 24, page 42) and closed (see Definition 25, page 42) in the topology of weak convergence [Bil99; Sha11]. Now, to see that $\Lambda(c)$ is tight, observe that for any $\epsilon > 0$, there exists an $a_\epsilon > 0$ such that for all $\mu \in \Lambda(c)$,

$$\Pr(|X| \geq a_\epsilon) \leq \frac{\mathbb{E}_\mu[|X|]}{a_\epsilon} \leq \frac{c}{a_\epsilon} < \epsilon \tag{3.30}$$

by Markov's inequality and the condition $\mathbb{E}_\mu[|X|] \leq c$. Choose $\mathcal{K}_\epsilon = [-a_\epsilon, a_\epsilon]$, then $\mathcal{K}_\epsilon$ is compact on $\mathbb{R}$ and $\mu(\mathcal{K}_\epsilon) > 1 - \epsilon$ for all $\mu \in \Lambda(c)$. As such, $\Lambda(c)$ is tight.

To show that $\Lambda(c)$ is closed, let $\{\mu_n\}_{n=1}^{\infty}$ be a convergent sequence in $\Lambda(c)$ with limit μ_0. Let $f(x) = |x|$, which is bounded below (i.e. $f(x) \geq 0$), continuous and, therefore, also lower semicontinuous. By the Portmanteau theorem for weak convergence (see Theorem 3, page 42),

$$\begin{aligned}\mathbb{E}_{\mu_0}[|X|] &= \int_{-\infty}^{\infty} f(x)\mu_0(dx) \\ &\leq \liminf_{n\to\infty} \int_{-\infty}^{\infty} f(x)\mu_n(dx) \leq c.\end{aligned} \tag{3.31}$$

This means that $\mu_0 \in \Lambda(c)$. As our choice of convergent sequence was arbitrary, it follows that $\Lambda(c)$ is closed. As such, $\Lambda(c)$ is compact.

The existence of a probability measure μ is studied. We present the extreme value theorem given by

Theorem 7 (The extreme value theorem)**.** *Suppose a function $f(x)$ is continuous on a compact interval $[a, b]$. Then $f(x)$ attains both a maximum and minimum, that is, there are points x_{max} and x_{min} in $[a, b]$, so that for every other $x \in [a, b]$, $f(x_{min}) \leq f(x) \leq f(x_{max})$.*

Using Theorem 4 (page 49), $I(X;Y)$ is continuous on $\Lambda(c)$. As such, by the extreme value theorem the capacity achieving probability measure μ exists. □

We now turn to the problem of showing that the capacity achieving probability measure is unique.

Theorem 8. *The capacity achieving probability measure μ^* on $\Lambda(c)$ is unique.*

Proof. By [LMG14], the mutual information is concave. Then the input-output mutual information $I(X;Y)$ of the $AS\alpha SN$ channel is concave in $\Lambda(c)$.

To prove strict concavity, we use the same approach as [LMG14] and show that if μ_0, μ_1 both achieve the maximum, then μ_0, μ_1 are identical. This is implied if

$$\mathbb{E}_{\mu_0}[p_N(y-x)] = p_Y(y;\mu_0) = p_Y(y;\mu_1) = \mathbb{E}_{\mu_1}[p_N(y-x)], \forall y. \tag{3.32}$$

Let $\phi_N(t)$ be the characteristic function of $N \sim S_\alpha(\gamma, 0, 0)$ and $\phi_\mu(t)$ be the characteristic function of the probability measure μ. As N and μ are independent, the characteristic function can be written as $\phi_{N+\mu}(t) = \phi_N(t)\phi_\mu(t)$. Moreover, the Lévy continuity theorem is given by

Theorem 9 (Lévy's Continuity Theorem on $\mathbb{R}$)**.** *Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $\mathbb{R}$, with characteristic functions $(\phi_n)_{n\in\mathbb{N}}$. If $\mu_n \Rightarrow \mu$, then ϕ_n converges pointwise to ϕ (the characteristic function of μ), that is, $\lim_{n\to\infty} \phi_n = \phi$. Conversely if ϕ_n converges pointwise to a function ϕ which is continuous at 0, then ϕ is the characteristic function of a probability measure μ, and $\mu_n \Rightarrow \mu$.*

Then, by applying the above theorem, $p_Y(y;\mu_0) = p_Y(y;\mu_1)$ is equivalent to $\phi_N(t)\phi_{\mu_0}(t) = \phi_N(t)\phi_{\mu_1}(t)$. Since the characteristic function $\phi_N(t)$ is non-zero for all t (see (3.24)), this implies $\mu_0 = \mu_1$; completing the proof. □

We remark that in [FAF16], Fahs and Abou-Faycal have recently established a general method for proving that the support of the optimal input is compact and finitely supported for additive noise channels. In particular, in [FAF16, Theorem 9], they show that this result applies to moment constraints of the form $\mathbb{E}[|X|^r] \leq c$, $0 < r < \alpha$. Therefore, the optimal input exists, unique, compact and finitely supported.

Lower Bound

We now turn to obtain a lower bound of the capacity obtained from (3.25). We get inspired by the Gaussian case. The stability property is needed to derive the explicit form of the capacity. Our idea is to use as an input an α-stable distribution, showing the same α exponent. In that case, the resulting output is also α-stable and we are able to obtain an analytical expression of the achievable rate, which, indeed, is a lower bound of the capacity.

We compute the mutual information when the input distribution is the same as the noise, up to the scale and location parameters; in particular, the characteristic exponent $\alpha \in (1, 2]$ is the same for both the noise and the input distribution. This yields a closed-form expression, as we show in the following theorem.

Theorem 10. *The capacity of the $AS\alpha SN$ channel with $N \sim S_\alpha(\gamma, 0, 0)$ with $1 < \alpha < 2$ is lower bounded by*

$$C \geq \frac{1}{\alpha} \log_2 \left(1 + M_\alpha \left(\frac{c}{\gamma_N} \right)^\alpha \right), \tag{3.33}$$

where

$$M_\alpha = \left(\frac{\pi}{2\Gamma\left(1 - \frac{1}{\alpha}\right)} \right)^\alpha, \tag{3.34}$$

Proof. Let $X \sim S_\alpha(\gamma_X, 0, 0)$ and $\gamma_X \in \mathbb{R}_{>0}$. Consider the random variable $U \sim S_\alpha(1, 0, 0)$. By the scaling and translation properties of α-stable random variables (Corollary 1, page 23, and Property 3, page 24), we can write

$$\begin{aligned} X &\overset{d}{=} \gamma_X U \\ N &\overset{d}{=} \gamma_N U. \end{aligned} \tag{3.35}$$

Moreover, by Property 1,

$$Y = X + N \sim S_\alpha((\gamma_X^\alpha + \gamma_N^\alpha)^{\frac{1}{\alpha}}, 0, 0) \tag{3.36}$$

and hence

$$Y \overset{d}{=} \gamma_Y U, \tag{3.37}$$

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

The mutual information is then given by

$$\begin{aligned} I(X;Y) &= h(Y) - h(Y|X) \\ &= h(\gamma_Y U) - h(\gamma_N U) \\ &= h(U) + \log_2(\gamma_Y) - h(U) - \log_2(\gamma_N) \\ &= \log_2\left(\frac{(\gamma_X^\alpha + \gamma_N^\alpha)^{\frac{1}{\alpha}}}{\gamma_N}\right) \\ &= \log_2\left(\left(\frac{\gamma_X^\alpha + \gamma_N^\alpha}{\gamma_N^\alpha}\right)^{\frac{1}{\alpha}}\right) \\ &= \frac{1}{\alpha}\log_2\left(1 + \frac{\gamma_X^\alpha}{\gamma_N^\alpha}\right) \\ &= \frac{1}{\alpha}\log_2\left(1 + \left(\frac{\gamma_X}{\gamma_N}\right)^\alpha\right) \\ &= \frac{1}{\alpha}\log_2\left(1 + \left(\frac{c\pi}{2\gamma_N\Gamma\left(1 - \frac{1}{\alpha}\right)}\right)^\alpha\right), \end{aligned} \tag{3.38}$$

where we used Property 5 (page 24) and the constraint $\mathbb{E}[|X|] \leq c$. □

Remark 3. *Observe that by applying Property 5 (page 24) to γ_N in our lower bound in Theorem 10 yields (3.23).*

Remark 4. *Differently from the Gaussian case, the stable distribution does not maximize the entropy with a fractional lower or absolute moment constraint, so that the lower bound differs from the capacity.*

Upper Bounds

Denote $W(\cdot|x)$ as the channel law corresponding to the random variable $Z = x + N$, $x \in \mathbb{R}$, which is absolutely continuous since the noise, N, is absolutely continuous. Let $R(\cdot)$ be any absolutely continuous probability

measure on $\mathbb{R}$, with corresponding probability density function p_R. Since the alphabet of the input and output is $\mathbb{R}$, which is separable[1], we can apply Theorem 5.1 in [LM03]. This provides a means of obtaining an upper bound on the capacity by choosing any absolutely continuous probability measure $R(\cdot)$, which is given by

$$C \leq \int_{-\infty}^{\infty} D(W(\cdot|x)||R(\cdot))d\mu^*, \tag{3.39}$$

where $D(\cdot||\cdot)$ is the Kullback-Leibler divergence (see Definition 20, page 40) and μ^* is the optimal input measure. By the results in Section 3.2.2, μ^* exists and is unique; however, there is no explicit characterization of μ^* beyond its existence and uniqueness.

Since $W(\cdot|x)$ and R admit absolutely continuous probability density functions w.r.t the Lebesgue measure, we can write

$$C \leq \mathbb{E}_{\mu^*}\left[\int_{-\infty}^{\infty} p_Z(y)\log_2\left(\frac{p_Z(y)}{p_R(y)}\right)dy\right]. \tag{3.40}$$

As such, the key problem is to choose an appropriate measure R on the output. We consider two choices, in order to find tight bounds, which each lead to a tractable upper bound on the capacity. As we will show in Section 3.2.3, the first bound is tighter when c is small, while the second bound is tighter when c is large. The distributions that we use to obtain tractable upper bounds are given by:

(i) The *Laplace distribution*, with probability density function

$$p_{R_S}(y) = \frac{\lambda}{2}\exp(-\lambda|x|), \tag{3.41}$$

where $\lambda > 0$ is a free parameter to be chosen.

(ii) The *polynomial distribution*, with probability density function

$$p_{R_P}(x) = \begin{cases} \frac{c_{x_0}}{|x|}, & |x| \leq x_0 \\ \frac{c_{x_0}}{1+|x|}, & |x| > x_0, \end{cases} \tag{3.42}$$

where $x_0 > 0$ and c_{x_0} is chosen to normalize the density function.

Our choices of the two distributions are informed by known properties of symmetric α-stable random variables. In particular, we seek to exploit the properties detailed in Section 2.4.1; namely, the fractional moment and asymptotic probability density expressions. These properties provide a means of obtaining closed-form upper bounds, as we show in Theorems 11 and 12.

Our first upper bound is obtained from the choice of the Laplace distribution.

[1] A topological space is separable if it contains a countable, dense subset. Since $\mathbb{R}$ contains $\mathbb{Q}$, it follows that $\mathbb{R}$ is separable since every point in $\mathbb{R}$ is a limit point of $\mathbb{Q}$.

Theorem 11. *The solution to (3.25) is upper bounded by*

$$C \leq \log_2 \left(\frac{2\Gamma\left(\frac{1}{\alpha}\right)}{\lambda \gamma_N \alpha \pi} \right) + (\log 2)^{-1} \lambda \left(\frac{2\gamma_N \Gamma\left(1 - \frac{1}{\alpha}\right)}{\pi} + c \right). \tag{3.43}$$

Proof. By (3.39), the capacity is upper bounded by

$$C \leq (\log 2)^{-1} \mathbb{E}_{\mu^*} \left[\int_{-\infty}^{\infty} p_Z(y) \log \left(\frac{p_Z(y)}{p_{R_S}(y)} \right) dy \right]. \tag{3.44}$$

Using Property 4 (page 24) yields

$$\begin{aligned} C &\leq (\log 2)^{-1} \log \left(\frac{2\Gamma\left(\frac{1}{\alpha}\right)}{\lambda \gamma_N \alpha \pi} \right) \\ &\quad + (\log 2)^{-1} \lambda \mathbb{E}_{\mu^*} \left[\int_{-\infty}^{\infty} p_Z(y) |y| dy \right] \\ &= \log_2 \left(\frac{2\Gamma\left(\frac{1}{\alpha}\right)}{\lambda \gamma_N \alpha \pi} \right) \\ &\quad + (\log 2)^{-1} \lambda \mathbb{E}_{\mu^*} \left[\int_{-\infty}^{\infty} p_N(y) |y + X| dy \right]. \end{aligned} \tag{3.45}$$

Applying the triangle inequality (i.e., for any two real numbers x and y, $|x + y| \leq |x| + |y|$), then gives

$$C \leq \log_2 \left(\frac{2\Gamma\left(\frac{1}{\alpha}\right)}{\lambda \gamma_N \alpha \pi} \right) + (\log 2)^{-1} \lambda \left(\mathbb{E}[|N|] + \mathbb{E}_{\mu^*}[|X|] \right). \tag{3.46}$$

Finally, from from Property 5 (page 24)

$$C \leq \log_2 \left(\frac{2\Gamma\left(\frac{1}{\alpha}\right)}{\lambda \gamma_N \alpha \pi} \right) + (\log 2)^{-1} \lambda \left(\frac{2\gamma_N \Gamma\left(1 - \frac{1}{\alpha}\right)}{\pi} + c \right). \tag{3.47}$$

□

We now turn to proving a second approximate upper bound, obtained by using the polynomial distribution on the output. As we show in the following sections, this result forms a tighter approximation to the Blahut-Arimoto numerical capacity calculation than our first upper bound for large values of c.

Theorem 12. *Let $x_0 > 1$. The capacity satisfies the following approximate upper bound*

$$C \lessapprox \log\left(\frac{C_\alpha x_0^{-\alpha-1}}{c_{x_0}}\left(1+\mathbb{E}_N[|N|]+c\right) + \frac{\Gamma(1/\alpha)}{\alpha\pi\gamma c_{x_0}}\left[\mathbb{E}_N[|N|]+c\right]\right), \tag{3.48}$$

where $\lessapprox$ denotes asymptotically less than.

Proof. By (3.39), the capacity is upper bounded by

$$\begin{aligned} C &\leq \mathbb{E}_{\mu^*}\left[\int_{-\infty}^{\infty} p_Z(y)\log\left(\frac{p_Z(y)}{p_{R_P}(y)}\right)dy\right] \\ &\leq \log\left(\mathbb{E}_{\mu^*}\left[\int_{-\infty}^{\infty} p_Z(y)\frac{p_Z(y)}{p_{R_P}(y)}dy\right]\right), \end{aligned} \tag{3.49}$$

where we applied Jensen's inequality twice. Now, let

$$p_{R_P}(x) = \begin{cases} \frac{c_{x_0}}{1+|x|}, & |x| \leq x_0 \\ \frac{c_{x_0}}{|x|}, & |x| > x_0, \end{cases} \tag{3.50}$$

where c_{x_0} is chosen so that $p_{R_P}(x)$ is normalized to one. Recall that $x_0 > 1$.

Now,

$$\begin{aligned} C &\leq \log\left(\mathbb{E}_{\mu^*}\left[\int_{|z|>x_0} p_N(z)\frac{p_N(z)}{p_{R_P}(z+x)}dz\right]\right. \\ &\quad \left.+\mathbb{E}_{\mu^*}\left[\int_{|z|\leq x_0} p_N(z)\frac{p_N(z)}{p_{R_P}(z+x)}dz\right]\right) \\ &\lessapprox \log\left(\mathbb{E}_{\mu^*}\left[\int_{|z|>x_0} p_N(z)\frac{C_\alpha|z|^{-\alpha-1}}{p_{R_P}(z+x)}dz\right]\right. \\ &\quad \left.+\mathbb{E}_{\mu^*}\left[\int_{|z|\leq x_0} p_N(z)\frac{\Gamma(1/\alpha)}{\alpha\pi\gamma p_{R_P}(z+x)}dz\right]\right), \end{aligned} \tag{3.51}$$

where we applied the asymptotic tail representation in Property 6 (page 24), from which C_α arises.

Replacing $p_{R_P}(z+x)$ in the second integral by its expression in (3.42) and applying the triangle inequality becomes

$$\begin{aligned} C &\lessapprox \log\left(\mathbb{E}_{\mu^*}\left[\int_{|z|>x_0} \frac{p_N(z)C_\alpha|z|^{-\alpha-1}}{p_{R_P}(z+x)}dz\right]\right. \\ &\quad \left.+\mathbb{E}_{\mu^*}\left[\int_{|z|\leq x_0} \frac{p_N(z)\Gamma(1/\alpha)(|z|+|x|)}{\alpha\pi\gamma c_{x_0}}dz\right]\right) \\ &\lessapprox \log\left(\mathbb{E}_{\mu^*}\left[\int_{|z|>x_0} \frac{p_N(z)C_\alpha|z|^{-\alpha-1}}{p_{R_P}(z+x)}dz\right]\right. \\ &\quad \left.+\frac{\Gamma(1/\alpha)}{\alpha\pi\gamma c_{x_0}}\left[\mathbb{E}_N[|N|]+\mathbb{E}_{\mu^*}[|X|]\right]\right). \end{aligned} \tag{3.52}$$

We note that

$$\begin{aligned}
&\mathbb{E}_{\mu^*}\left[\int_{|z|>x_0}\frac{p_N(z)C_\alpha|z|^{-\alpha-1}}{p_{R_P}(z+x)}dz\right]\\
&=\mathbb{E}_{\mu^*}\left[\int_{|z|>x_0}\frac{p_N(z)C_\alpha(1+|z+x|)|z|^{-\alpha-1}}{c_{x_0}}dz\right]\\
&\leq\mathbb{E}_{\mu^*}\left[\int_{|z|>x_0}\frac{p_N(z)C_\alpha(1+|z+x|)x_0^{-\alpha-1}}{c_{x_0}}dz\right]\\
&\leq\frac{C_\alpha x_0^{-\alpha-1}}{c_{x_0}}\left(1+\mathbb{E}_N[|N|]+\mathbb{E}_{\mu^*}[|X|]\right), \qquad (3.53)
\end{aligned}$$

where the triangle inequality was again applied in order to produce the expected values.

Putting it all together, we have

$$\begin{aligned}
C\lessapprox\log\Bigg(&\frac{C_\alpha x_0^{-\alpha-1}}{c_{x_0}}\left(1+\mathbb{E}_N[|N|]+\mathbb{E}_{\mu^*}[|X|]\right)\\
&+\frac{\Gamma(1/\alpha)}{\alpha\pi\gamma c_{x_0}}\left[\mathbb{E}_N[|N|]+\mathbb{E}_{\mu^*}[|X|]\right]\Bigg). \qquad (3.54)
\end{aligned}$$

□

Note that when $x_0\to\infty$, our approximate bound is in fact an upper bound. We remark that this is an asymptotic approximation and other values may not hold. This is due to the fact that we used the asymptotic tail representation of α-stable probability density functions from Property 6 (page 24). Despite this, Fofack and Nolan [FN99] have numerically shown that polynomial tails are a very good approximation of the tail as can be seen in Fig. 2.9.

We also remark that another variation on the Lapidoth-Moser bound [LM03] yields an alternative method to upper bound the capacity, which corresponds to the dual problem [BV04] for the capacity. In particular, the capacity can be upper bounded by

$$C\leq\min_{\gamma\geq0}\max_{x\in\mathbb{R}}[D(W(\cdot|x)||R(\cdot))+\gamma(c-\mathbb{E}[|X|])]. \qquad (3.55)$$

This bound has been investigated by Katz and Shamai [KS04] in the context of non-coherent and partially coherent AWGN channels. However, in our case the optimization problem is challenging due to the fact that the channel law is not available in closed-form. As such, the bound cannot be obtained in closed-form and numerical methods are required. In the following section, we numerically study the $AS\alpha SN$ channel capacity via the Arimoto-Blahut algorithm, which has the advantage over the Katz and Shamai approach that the algorithm converges to the capacity as the channel law approximation converges to the stable channel law.

3.2.3 Numerical Analysis

In this section, we study numerical properties of the capacity optimization problem in (3.25). In particular, we consider a numerical approximation of the capacity using a variation of the Blahut-Arimoto algorithm. In contrast with [WKZ11], our numerical study considers the constraint $\mathbb{E}[|X|] \leq c$, rather than an average power constraint. Our key result is a set of guidelines for the choice of support size of the noise and input signal distributions in order to ensure that the error in the capacity obtained via the Blahut-Arimoto algorithm is less than approximately 0.01 bits.

Numerical Capacity Approximation Algorithm

The Blahut-Arimoto algorithm provides a means of numerically approximating the capacity of a discrete memoryless channel in the case of an input with discrete and bounded support. To approximate the capacity in (3.25), the variant of the algorithm for the capacity with constraints is required [Bla72, Section IV].

In this section, we provide details of our variation on the Blahut-Arimoto algorithm. These details are important as our discussion in Section 3.2.4 is only guaranteed to apply for our particular algorithm.

In the $AS\alpha SN$ channel, the support of the noise density is not discrete and bounded. As such, it is necessary to approximate the noise with a random variable with discrete and bounded support. More precisely, we consider the channel

$$Y = X_{X_{\max},h_X} + N_{N_{\max},h_N}, \tag{3.56}$$

where the random variable $X_{X_{\max},h_X}$ has support $S_X = h_X\mathbb{Z} \cap [-X_{\max}, X_{\max}]$ and $N_{N_{\max},h_N}$ has support $S_N = h_N\mathbb{Z} \cap [-N_{\max}, N_{\max}]$, where h_N and h_X are the step sizes of the supports. The probability mass function of $N_{N_{\max},h_N}$ is obtained by discretizing the absolutely continuous density of the α-stable distributed noise, N, and the channel law is denoted by $Q(\cdot|x)$.

The corresponding optimization problem for the capacity of the channel in (3.57) with constraint $\mathbb{E}[|X_{X_{\max},h_X}|] \leq c$ is then

$$\begin{aligned} \underset{\mathbf{q}\in\mathcal{Q}}{\text{maximize}} \quad & \sum_{i=1}^{m}\sum_{j=1}^{n} q(x_i)Q(y_j|x_i)\log_2\left(\frac{Q(y_j|x_i)q(x_j)}{p(y_i)q(x_j)}\right) \\ \text{subject to} \quad & \sum_{i=1}^{m}|x_i|q(x_i) \leq c, \end{aligned} \tag{3.57}$$

where $m = |S_X|$, $n = |S_N|$, and $\mathcal{Q}$ is the set of probability mass functions on S_X.

To obtain the solution, C_{approx}, to (3.57), the key observation due to Blahut and Arimoto [Bla72; Ari72] is that

$$C_{approx} = \max_{\mathbf{q}:\sum_{i=1}^{m}|x_i|q(x_i)\leq c} \max_{\Phi} J(\mathbf{q}, \mathbf{\Phi}), \tag{3.58}$$

where

$$J(\mathbf{q}, \Phi) = \sum_{i=1}^{m}\sum_{j=1}^{n} q(x_i)Q(y_j|x_i)\log_2\left(\frac{\Phi(x_i|y_j)}{q(x_i)}\right) \tag{3.59}$$

and Φ is an arbitrary $m \times n$ transition probability matrix. The approximate capacity is then obtained by alternating between the maximization problems, which leads to Algorithm 1[2].

Algorithm 1 Computation of the approximate capacity in (3.57).

Initialize:
(1) Set $r^{(0)}(x) = \frac{1}{|S_X|}$.
(2) Set $C_0 = 0,\ C_{-1} = -2\epsilon$.
while $C_n - C_{n-1} > \epsilon$ **do**
Compute:
(1) $C_{n-1} = C_n$.
(2) $Q^{(n)}(x|y) = \frac{r^{(n-1)}(x)P(y|x)}{\sum_{x=1}^{m} r^{(n-1)}(x)P(y|x)}$.
(3) $C_n = \sum_{x=1}^{m}\sum_{y=1}^{n} r^{(n-1)}(x)P(y|x)$
$\log_2\left(\frac{Q^{(n)}(x|y)}{r^{(n-1)}(x)}\right)$.
Solve for ν such that

$$\sum_{x=1}^{m}\left(1 - \frac{|x|}{c}\right)e^{\nu|x|}\prod_{y=1}^{n} Q^{(n)}(x|y)^{P(y|x)} = 0. \tag{3.60}$$

Compute:

$$r^{(n)}(x) = \frac{e^{\nu|x|}\prod_{y=1}^{n} Q^{(n)}(x|y)^{P(y|x)}}{\sum_{x'=1}^{m} e^{\nu|x'|}\prod_{y=1}^{n} Q^{(n)}(x'|y)^{P(y|x')}}. \tag{3.61}$$

end while
return C_n.

3.2.4 Effect of the Support Size

To compute the capacity with the Blahut-Arimoto algorithm accurately requires a good choice of the step size (h_X, h_N) and support sizes ($X_{\max}$, $N_{\max}$) in Algorithm 1. Moreover, an important question is whether the output of the Blahut-Arimoto algorithm converges to the capacity as the approximate channel law converges to the stable channel law. To see that this holds, consider the following bound on the error which is justified in

[2]The Dekker algorithm [Bre73] is used to solve for ν in (3.60). An implementation of this code is available at https://github.com/maurokenny/BlahutArimoto_brent

[EPK17]

$$|C^* - C_{approx}| \leq M|\|p_N - p_{N,approx}\|_{TV} + |o(\|p_N - p_{N,approx}\|_{TV})|, \tag{3.62}$$

where $M < \infty$ and

$$\|p_N - p_{N,approx}\|_{TV} = \frac{1}{2}\int_{\mathbb{R}} |p_N(x) - p_{N,approx}(x)|dx, \tag{3.63}$$

where $\|\cdot\|_{TV}$ is the total variation defined as

Definition 27. *Let the measures P and Q having a common measurable space $(X, \mathcal{M})$, then the total variation is given by*

$$\sup_{E\in\mathcal{M}} |P(E) - Q(E)|.$$

As such, the Blahut-Arimoto algorithm approximation converges to the capacity.

It is also possible to establish the rate of convergence. Suppose that there is no discretization and the truncation level is T. Then, by [EPK17] the error is of the order $O(T^{-\alpha})$.

To address the choice of h_X and h_N, Fig. 3.1 shows the support size required to obtain the capacity for different choices of constraint value c such that the error is approximately 0.01 with a step size of $h_X = h_N = 0.01$. For reference, we used a support size of $N_{\max} = 220$. Observe that the support size is increasing as the constraint increases.

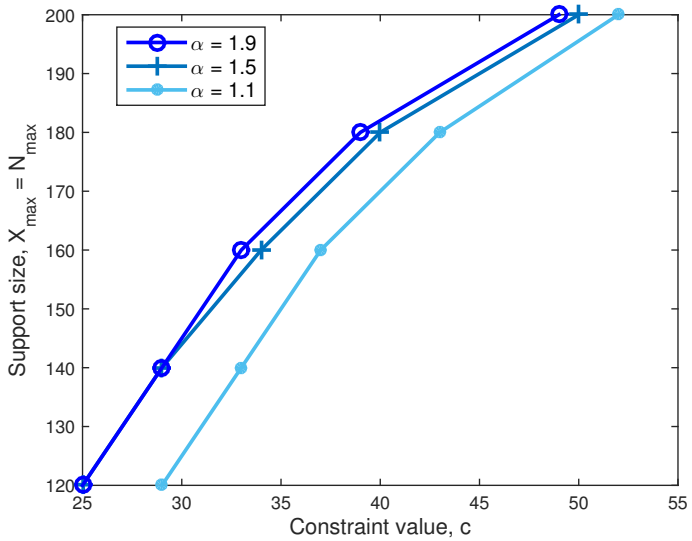


FIGURE 3.1: Plot of the support size required to ensure an error in capacity of approximately 0.01 bits for each value of the constraint, c, with $\gamma_N = 1$ and step size $h_X = h_N = 0.01$.

3.2.5 Behavior of the Bounds and Numerical Approximation

In this section, we compare our bounds with the numerical approximation obtained using the Blahut-Arimoto algorithm. We also investigate the behavior of the approximate capacity via the Blahut-Arimoto algorithm as the ratio $\frac{c}{\gamma_N}$ varies. We then evaluate the tightness of our bounds and study the effect of noise parameters.

On the Ratio $\frac{c}{\gamma_N}$

In the Gaussian channel with a power constraint, the capacity is determined by the signal-to-noise ratio (SNR). In the case of the $AS\alpha SN$ channel, the analogous quantity is the ratio $\frac{c}{\gamma_N}$. As such, we can ask whether this ratio plays a similar role to the SNR. To do this, we compare in Fig. 3.2 the numerical approximation of the capacity when c varies with γ_N fixed, or when γ_N varies with c fixed.

Observe that the curves agree when α is fixed, which shows that for the choices of α in Fig. 3.2, $\frac{c}{\gamma_N}$ determines the capacity analogously to the role of the SNR in the Gaussian channel with a power constraint. This observation is consistent with our lower bound, which is also determined by the ratio $\frac{c}{\gamma_N}$; however, this is not the case for our upper bounds. We also note that for different α values, this ratio is not sufficient and comparing different $AS\alpha SN$ channels is not straightforward.

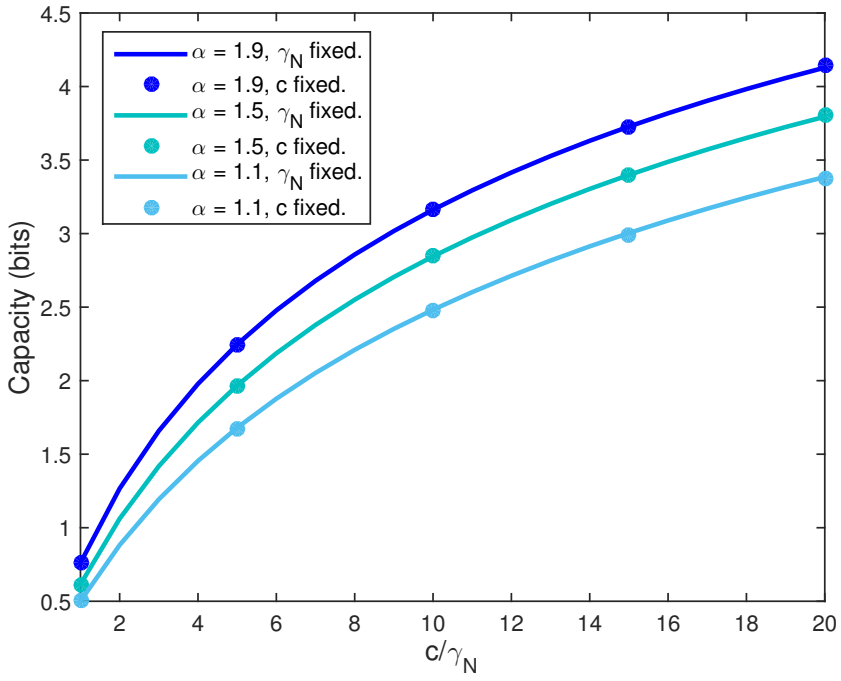


FIGURE 3.2: Plot of the capacity with $\gamma_N = 1$ or $c = 1$ using $\beta_N = 0$, $\delta_N = 0$.

Effect of Noise Parameters on the Bounds

We now compare our bounds and the numerical capacity approximation using Algorithm 1 for $\alpha = 1.9$ and $\alpha = 1.1$ in Fig. 3.3 and 3.4. Since we observed from Fig. 3.2 that only the ratio $\frac{c}{\gamma_N}$ needs to be varied —not c and γ_N separately— we only vary c and fix $\gamma_N = \sqrt{0.5}$. Regarding our first upper bound, we assume $\lambda = 0.1$ in Theorem 11 (page 55), obtained by a numerical search to optimize the tightness of the bound for a range of choices of $\mathbb{E}[|X|]$. We note that our second upper bound in Theorem 12 (page 55) is asymptotic, that holds as the parameter $x_0 \to \infty$, due to the fact that the probability density function of a symmetric α-stable random variable only has asymptotically polynomial tails (see Property 4, page 24). As polynomial tails are a very good approximation of the tail, we consider through an empirical approach the approximation of the bound, with $x_0 = 10$ in (3.42).

Fig. 3.3 compares the bounds and the numerical approximation with $\alpha = 1.9$. Observe that the lower bound and the numerical approximation are in very good agreement. Moreover, the gap between the lower and the asymptotic upper bound from Theorem 12 (page 55) is about 1 bit. We also observe that the upper bound based on the Laplace distribution in Theorem 11 (page 55) is tighter than the upper bound from Theorem 12 for a range of c between 5 and 10.

Fig. 3.4 compares the bounds and numerical approximation with $\alpha = 1.1$. In this case, the lower bound and the numerical approximation are within 2 bits for sufficiently large c. The gap between the asymptotic upper bound from Theorem 12 and the numerical approximation is also within 2 bits.

Comparing the two figures, observe that increasing α leads to an increase in the capacity. This is consistent with the results in [WKZ11], where the second moment constraint $\mathbb{E}[X^2] \leq P$ was considered.

3.2.6 Further Properties of the Lower Bounds

So far, we have focused on characterizing the capacity arising from the optimization problem in (3.25). In particular, we showed that our lower bound in Theorem 10 (page 52) is a good approximation compared with the numerical approximation in Section 3.2.3 for sufficiently large α.

In this section, we discuss the bend point property of the lower bound, which is the behavior of the lower bound at medium c and the parametrization of the input distribution. The tractability of the lower bound means that it is an attractive performance metric in settings based on the $AS\alpha SN$ channel, and can play a role similar to the power constrained capacity in settings based on the Gaussian channel.

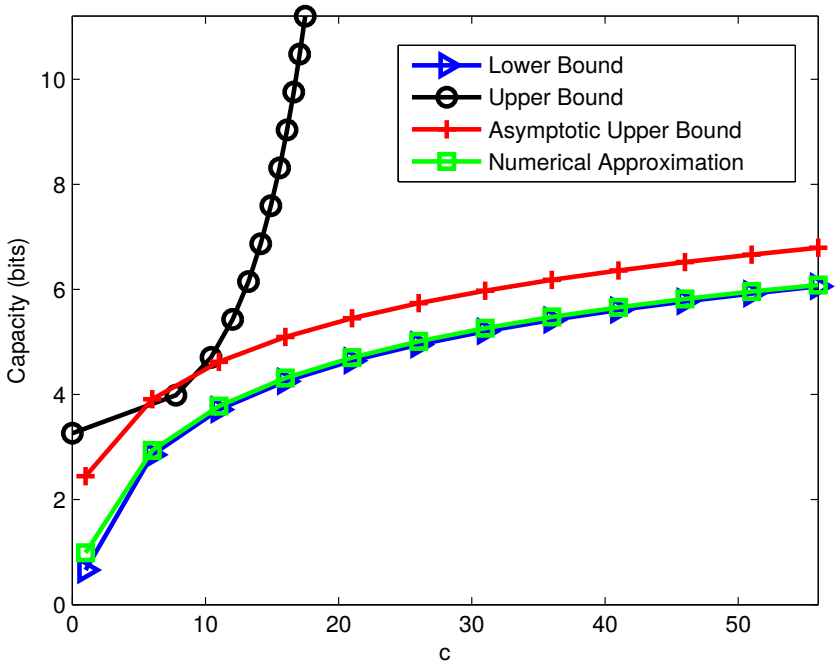


FIGURE 3.3: Comparison of capacity bounds and approximations with $\alpha = 1.9$.

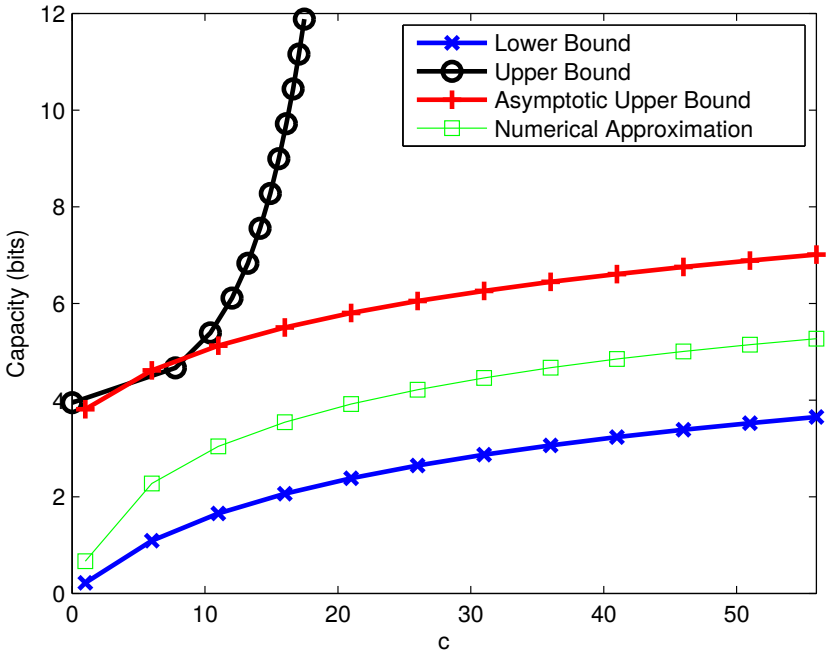


FIGURE 3.4: Comparison of capacity bounds and approximations with $\alpha = 1.1$.

Medium c Behavior

We now consider the behavior of the lower bound for medium c. As observed in [Ega15] in the context of Gaussian noise channels, a fundamental qualitative feature of the capacity curve in the medium SNR region is the *bend point*. The bend point provides a means of quantifying the transition from low to high SNR. This is defined for the $AS\alpha SN$ channel as follows.

Definition 28 (Bend point)**.** *Consider the capacity lower bound in Theorem 10 (page 52), given by*

$$C_{LB} = \frac{1}{\alpha}\log_2\left(1 + M_\alpha\left(\frac{c}{\gamma_N}\right)^\alpha\right), \tag{3.64}$$

where M_α is given by (3.34).

The bend point, c_{bend}, is then the $c_{dB} = 10\log_{10} c$ such that the second derivative of (3.64) is maximized.

For the $AS\alpha SN$ channel, the bend point corresponds to the point on the capacity lower bound curve where the rate of change of the slope is maximized. As such, it can be viewed as the transition between high and low c as the rate of change of the slope tends to zero as $c_{dB} \to -\infty$, reaches its maximum value at the bend point, and then tends to zero as $c_{dB} \to \infty$.

An important observation in [Ega15] is that the bend point is intimately related to the intersection of high and low SNR asymptotes in the capacity of power constrained Gaussian channels. We now investigate the bend point in the context of the $AS\alpha SN$ channel.

Theorem 13. *The bend point is given by*

$$c_{bend} = \frac{10}{\alpha}\log_{10}\left(\frac{\gamma_N^\alpha}{M_\alpha}\right). \tag{3.65}$$

Proof. The third derivative of the lower bound in (3.33) in Theorem 10 (page 52) is given by

$$C'''_{LB} = \frac{M_\alpha\left(\frac{\alpha}{10}\log 10\right)^3}{\alpha\gamma_N^\alpha\log 2}\left[\frac{10^{\alpha c_{db}/10}\left(1 - \frac{M_\alpha}{\gamma_N^\alpha}10^{\alpha c_{dB}/10}\right)}{\left(1 + \frac{M_\alpha}{\gamma_N^\alpha}10^{\alpha c_{dB}/10}\right)^3}\right], \tag{3.66}$$

which satisfies $C'''_{LB} = 0$ when $c_{dB} = \frac{10}{\alpha}\log_{10}\left(\frac{\gamma_N^\alpha}{M_\alpha}\right)$. Note also that the second derivative of (3.33) is given by

$$C''_{LB} = \frac{M_\alpha\left(\frac{\alpha}{10}\log 10\right)^2}{\alpha\gamma_N^\alpha\log 2}\left[\frac{10^{\alpha c_{dB}/10}}{\left(1 + \frac{M_\alpha}{\gamma_N^\alpha}10^{\alpha c_{dB}/10}\right)^2}\right], \tag{3.67}$$

and is symmetric around $c_{dB} = \frac{10}{\alpha}\log_{10}\left(\frac{\gamma_N^\alpha}{M_\alpha}\right)$ and decreasing for $c_{dB} > \frac{10}{\alpha}\log_{10}\left(\frac{\gamma_N^\alpha}{M_\alpha}\right)$, which proves the theorem. □

In Fig. 3.5, the second derivative of our lower bound C_{LB} is plotted for varying α and γ_N with $\beta = 0$ and $\delta_N = 0$ fixed, where one may seem the symmetry around the maximum point c_{bend}.

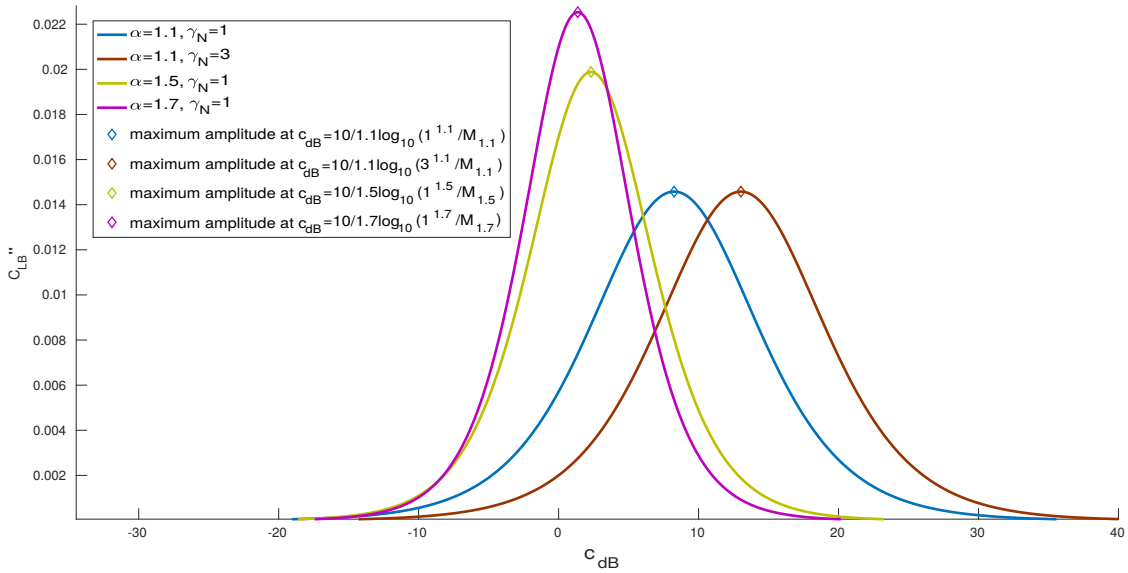


FIGURE 3.5: Plot of the second derivative of the capacity lower bound, C''_{LB} for varying α and γ_N with $\beta = 0$ and $\delta_N = 0$. The dot on each curve is the maximum point on each curve.

Now, define the asymptote (as $c_{dB} \to \infty$) of the lower bound as

$$C_{asymp} = \frac{1}{\alpha}\log_2\left(10^{\alpha c_{dB}/10}\right) + \frac{1}{\alpha}\log_2\left(\frac{M_\alpha}{\gamma_N^\alpha}\right). \tag{3.68}$$

Observe that $C_{asymp} = 0$ when $c_{dB} = \frac{10}{\alpha}\log_{10}\left(\frac{\gamma_N^\alpha}{M_\alpha}\right)$, which agrees with the bend point c_{bend}, from Theorem 13. This means that as for the power constrained Gaussian channel, the intercept asymptote of the capacity lower bound for the $AS\alpha SN$ channel agrees with the bend point; however, unlike the power constrained Gaussian channel, the bend point does not always occur at $c_{dB} = 0$.

Fig. 3.6 plots the capacity lower bound for varying α. Observe that the bend point c_{bend} is reduced as α increases. This suggests that using the asymptotic approximation is more accurate for Gaussian channels than for the $AS\alpha SN$ channel at lower values of c. As asymptotic approximations are widely used, it may mean that approximations that are valid in the Gaussian case are less accurate for other values of α.

Parametrization of the Input Distribution

A interesting feature of the lower bound is that the influence of α depends on whether it is written in terms of γ_X or $\mathbb{E}[|X|]$. To see this, first write the

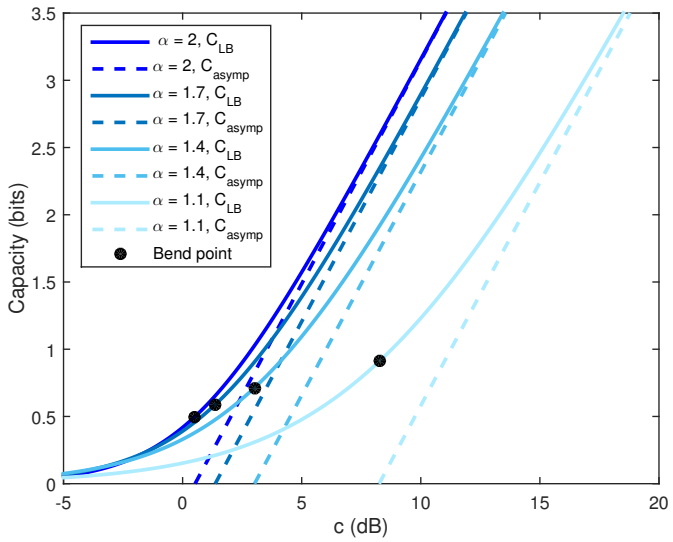


FIGURE 3.6: Plot of our capacity lower bound, C_{LB} for varying α, with $\gamma_N = 1$, $\beta = 0$ and $\delta_N = 0$. The dot on each curve is the corresponding bend point.

lower bound as

$$C_{LB,\gamma_X} = \frac{1}{\alpha}\log_2\left(1 + \frac{\gamma_X^\alpha}{\gamma_N^\alpha}\right), \tag{3.69}$$

with γ_X and γ_N fixed. In this case, the lower bound (3.69) *increases* as α is *reduced* (Fig. 3.7a). Now write the lower bound, equivalently, as

$$C_{LB,\mathbb{E}[|X|]} = \frac{1}{\alpha}\log_2\left(1 + M_\alpha\left(\frac{c}{\gamma_N}\right)^\alpha\right), \tag{3.70}$$

with $\mathbb{E}[|X|] = c$ and γ_N fixed. In this case, the lower bound (3.70) *increases* as α *increases* (Fig. 3.7b). We note that the behavior of $C_{LB,\mathbb{E}[|X|]}$ as α varies is also consistent with the effect of α with the constraint $\mathbb{E}[X^2] \leq P$ numerically studied in [WKZ11].

This behavior arises because the relationship between γ_X and $\mathbb{E}[|X|]$ itself depends on α, as detailed in Property 5 (page 24). The consequence is that the role of α is dependent on how the input signal is written (i.e., whether it is in terms of γ_X or $\mathbb{E}[|X|]$), and must be carefully considered if the bound is applied in a physical setting.

Another remark is that the comparison between different α, meaning different impulsiveness, is tricky. A criterion equivalent to the SNR that would also include impulsiveness index and sufficient to qualify the link quality in different impulsiveness conditions is needed. Further research in that direction is required.

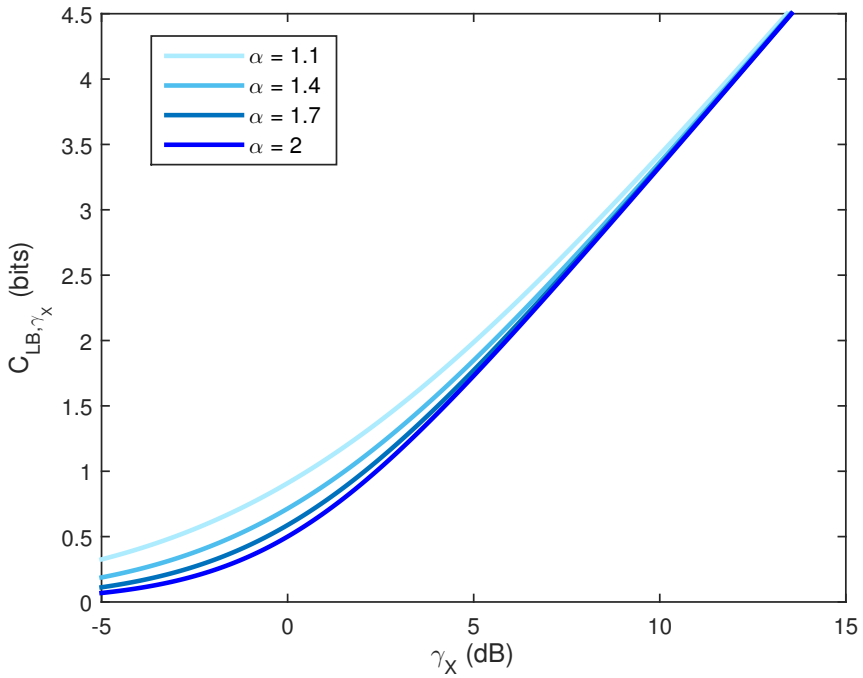
3.3 Conclusion

Impulsive noise plays a key role in many communication systems, ranging from wireless to molecular. Firstly, the capacity in many scenarios were outlined. The emblematic classical expression of the theory is Shannon's formula was revisited to be used as comparative for further impulsive scenarios. For this purpose, many information measure tools were presented, which will also be important along next chapters.

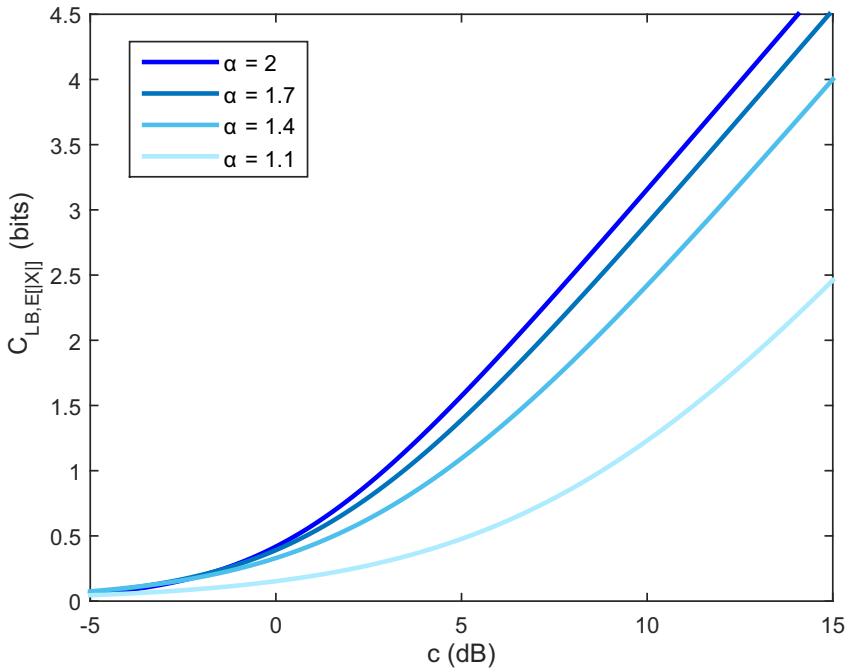
In particular, impulsive noise modeled with the symmetric α-stable distributions were studied in this chapter. We have derived lower and upper bounds for the $AS\alpha SN$ channel, with $\alpha \in (1, 2]$ and the existence and uniqueness of the optimal input distribution were proved. We have also investigated a numerical approximation via the Blahut-Arimoto algorithm, which requires discretization and truncation. A study was made in order to adjust the steps and supports sizes for the algorithm. Particularly, we show that the lower bound is a good approximation of the capacity obtained from the Blahut-Arimoto algorithm for α near 2.

We investigate the behavior of the quantity c/γ_N, in order to discover if it leads to a similar role as the SNR in Gaussian channel. In fact, it has been showed that our lower bound is consistent, although the comparison between many impulsiveness scenarios is difficult. Moreover, two additional studies were presented concerning the derived lower bound. First, the bend point case, a medium to characterize the behavior between low and high rates, was outlined. As a result, bend points may be shifted by a calculable amount in comparison to the Gaussian case. Second, the importance of the input parametrization was discussed.

There are several avenues for future work. For instance, this opens the question of the behavior and design of algorithms for parallel and MIMO additive α-stable noise channels. For the former, an approach will be presented in section 4.4.



(A) Capacity lower bound in (3.69) for varying α, with $\gamma_N = 1$ and $\beta = 0$.



(B) Capacity lower bound in (3.70) for varying α, with $\gamma_N = 1$ and $\beta = 0$.

FIGURE 3.7: Behavior of our lower bound for varying α.

Chapter 4

Capacity of Additive Isotropic α-Stable Noise Channels

In this chapter, we extend the capacity achievable bound for two dimensions, in which the channel is isotropic. Existence and uniqueness of the optimal input are proved. The achievable rate is then initially considered in the parallel channels case with fractional moments. Furthermore, the parametrization α is studied in the achievable rates using a perturbation approach from the Gaussian case, in order to understand its effect.

4.1 Position of the problem

PREVIOUS chapter has focused on the channel

$$Y = X + N, \tag{4.1}$$

where N is a real-valued symmetric α-stable random variable. In this chapter, we study the capacity of a generalization of (4.1), where the noise is an isotropic complex α-stable random variable. The additive isotropic α-stable noise ($AI\alpha SN$) channel naturally arises in the context of the baseband in wireless cellular communication networks with base stations distributed according to a homogeneous Poisson point process. For the $AI\alpha SN$ channel, we can still rely on the same system model as (4.1) but with complex valued signals. The details will be further defined in Section 4.2.

In order to study the $AI\alpha SN$ channel, we provide a real-valued vector channel representation. Unlike the Gaussian case ($\alpha = 2$) , the real and imaginary parts of isotropic α-stable random variables ($\alpha < 2$) are not independent. As such, it cannot be reduced to two parallel real-valued scalar channels and must be treated instead as a real-valued vector channel.

For the $AI\alpha SN$ channel we prove two key results:

1. We show that the optimal input for the $AI\alpha SN$ channel subject to a constraint $\mathbb{E}[|\mathbf{X}|^r] = (\mathbb{E}[|X_1|^r], \mathbb{E}[|X_2|^r])^T \preceq \mathbf{c}$, $r < \alpha$ exists and is unique, where $\preceq$ denotes a componentwise inequality, that is, $\mathbb{E}[|X_i|^r] \leq c_i$ for $i = 1, 2$.

2. We derive a lower bound on the capacity subject to $\mathbb{E}[|\mathbf{X}|^r] \preceq \mathbf{c}$, $r < \alpha$, given by

$$C \geq \frac{1}{\alpha} \log \left(1 + \frac{\left(\sqrt{2} \left(\frac{\min\{c_1, c_2\}}{C(r,\alpha)} \right)^{1/r} \right)^{\alpha}}{\sigma_{\mathbf{N}}^{\alpha}} \right) \text{ nats}, \tag{4.2}$$

where $\sigma_{\mathbf{N}}$ is a parameter characterizing the noise (defined in Section 4.2).

4.2 Additive Isotropic α-Stable Noise Channels

Firstly, it is important to establish the vector channel representation of the $AI\alpha SN$ channel. Theorem 1 (page 28) implies that we can write the $AI\alpha SN$ channel as the vector channel in $\mathbb{R}^2$

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}, \tag{4.3}$$

where $\mathbf{N} = (N_1, N_2)^T$ is the sub-Gaussian random vector induced by the isotropic α-stable noise N.

Next, we now turn to studying the capacity of this channel.

4.3 Capacity of $AI\alpha SN$ Channels

4.3.1 Capacity Optimization Problem

Let $\mathcal{B}(\mathbb{R}^2)$ be the Borel σ-algebra on $\mathbb{R}^2$ and $\mathcal{P}$ denote the collection of Borel probability measures on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ equipped with the topology of weak convergence. We define the capacity of the $AI\alpha SN$ channel as the solution to the following optimization problem,

$$\begin{aligned} &\underset{\mu \in \mathcal{P}}{\text{maximize}} \quad I(\mathbf{X}; \mathbf{Y}) \\ &\text{subject to} \quad \mathbb{E}_{\mu}[|\mathbf{X}|^r] \preceq \mathbf{c}, \end{aligned} \tag{4.4}$$

where $\mathbb{E}[|\mathbf{X}|^r] = (\mathbb{E}[|X_1|^r], \mathbb{E}[|X_2|^r])^T$, $I(\mathbf{X}; \mathbf{Y})$ is the mutual information of the channel (4.3), μ is the probability measure of X, and we restrict r to satisfy $r < \alpha$. Note that $\mathbb{E}[|\mathbf{X}|^r] \preceq \mathbf{c} = [c_1, c_2]^T$ if and only if $\mathbb{E}[|X_1|^r] \leq c_1$ and $\mathbb{E}[|X_2|^r] \leq c_2$.

4.3.2 Existence and Uniqueness

Denote $\Lambda(\mathbf{c})$ as the set of probability measures μ with support $\mathbb{R}^2$, such that $\mathbb{E}[|\mathbf{X}|^r] \preceq \mathbf{c}$. We first prove that the capacity achieving probability measure μ^* exists. To do this, as in the real case, we need to show that $\Lambda(\mathbf{c})$ is compact in $\mathcal{P}$. In turn, if $I(\mathbf{X}; \mathbf{Y})$ is continuous on $\Lambda(\mathbf{c})$, the extreme value

theorem then implies that the mutual information achieves its supremum on $\Lambda(\mathbf{c})$.

The first step is then to show that $I(\mathbf{X};\mathbf{Y})$ is continuous on $\Lambda(\mathbf{c})$.

Theorem 14. *$I(\mathbf{X};\mathbf{Y})$ is continuous on $\Lambda(\mathbf{c})$.*

Proof. Suppose that $\mu_k \Rightarrow \mu$, which means that the sequence of random vectors $\{\mathbf{X}_k\}$ in $\mathbb{R}^2$ converges weakly to a random vector $\mathbf{X}$ with measure μ. The mutual information of $\mathbf{X}_k$ is given by

$$I(\mathbf{X}_k;\mathbf{X}_k+\mathbf{N}) = h(\mathbf{X}_k+\mathbf{N}) - h(\mathbf{N}). \tag{4.5}$$

Since $h(\mathbf{N}) = h(N_1,N_2) \leq h(N_1)+h(N_2)$ and N_1, N_2 are symmetric α-stable random variables, it follows that $|h(\mathbf{N})| < \infty$.

We now show that $I(\mathbf{X};\mathbf{Y}) < \infty$. Define

$$q(\mathbf{x}) = \frac{1}{4\int_{\mathbb{R}^2_{\geq 0}} e^{-x_1^r - x_2^r} dx_1 dx_2} e^{-|x_1|^r - |x_2|^r}. \tag{4.6}$$

We have

$$I(\mathbf{X};\mathbf{Y}) = -\int_{\mathbb{R}^2} p_{\mathbf{X}+\mathbf{N}}(\mathbf{x}) \log p_{\mathbf{X}+\mathbf{N}}(\mathbf{x}) d\mathbf{x} - h(\mathbf{N}) \tag{4.7}$$

The Kullback-Leibler divergence $D(\cdot||\cdot)$ (see (3.3)) is applied to write the mutual information as

$$I(\mathbf{X};\mathbf{Y}) = -D(p_{\mathbf{X}+\mathbf{N}}||q) + \mathbb{E}[\log q(\mathbf{X}+\mathbf{N})] - h(\mathbf{N}). \tag{4.8}$$

Recall that Proposition 11 (page 40) guarantees the $D(p_{\mathbf{X}+\mathbf{N}}||q) \geq 0$, then

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &\leq \mathbb{E}[\log q(\mathbf{X}+\mathbf{N})] - h(\mathbf{N}) \\ &= -\mathbb{E}[|X_1+N_1|^r + |X_2+N_2|^r] \\ &\quad - \mathbb{E}\left[\log\left(\int_{\mathbb{R}^2_{\geq 0}} 4e^{-x_1^r - x_2^r} dx_1 dx_2\right)\right] - h(\mathbf{N}) \\ &\leq \left|-\mathbb{E}[|X_1+N_1|^r + |X_2+N_2|^r]\right. \\ &\quad \left. - \log\left(\int_{\mathbb{R}^2_{\geq 0}} 4e^{-x_1^r - x_2^r} dx_1 dx_2\right)\right| - h(\mathbf{N}), \end{aligned} \tag{4.9}$$

using the fact that the expected value of a constant is a constant itself.

Applying the triangle inequality and $\log(x) \leq x - 1$, for $x > 0$, we have

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &\leq \mathbb{E}[|X_1 + N_1|^r + |X_2 + N_2|^r] \\ &\quad + \int_{\mathbb{R}^2_{\geq 0}} 4e^{-x_1^r - x_2^r} dx_1 dx_2 - 1 - h(\mathbf{N}) \\ &\leq \mathbb{E}[|X_1 + N_1|^r + |X_2 + N_2|^r] \\ &\quad + \int_{\mathbb{R}^2_{\geq 0}} 4e^{-x_1^r - x_2^r} dx_1 dx_2 - h(\mathbf{N}) \\ &\leq 2^r(\mathbb{E}[|X_1|^r + |N_1|^r + |X_2|^r + |N_2|^r]) \\ &\quad + \int_{\mathbb{R}^2_{\geq 0}} 4e^{-x_1^r - x_2^r} dx_1 dx_2 - h(\mathbf{N}) \\ &< \infty, \end{aligned} \tag{4.10}$$

recalling that $r < \alpha$.

To conclude the proof, let $\mathbf{Y}_k = \mathbf{X}_k + \mathbf{N}$. Then,

$$p_{\mathbf{Y}_k}(\mathbf{x}) = \int p_{\mathbf{N}}(\mathbf{x} - \mathbf{y}) \mu_k(d\mathbf{y}). \tag{4.11}$$

Now consider

$$\lim_{k\to\infty} h(\mathbf{X}_k + \mathbf{N}) = -\lim_{k\to\infty} \int p_{\mathbf{Y}_k}(\mathbf{x}) \log p_{\mathbf{Y}_k}(\mathbf{x}) d\mathbf{x}. \tag{4.12}$$

Since $I(\mathbf{X}_k;\mathbf{Y}) < \infty$ and $|h(\mathbf{N})| < \infty$, it follows that $|h(\mathbf{X}_k + \mathbf{N})| < \infty$. An argument based on splitting the integral in (4.12) into positive and negative parts, justifies swapping the limit and the integral. Using the fact that $p_{\mathbf{N}}$ is bounded and continuous and the application definition of weak convergence in $\mathbb{R}^2$ yields the desired result, similarly to the real case. □

We now turn to showing that $\Lambda(\mathbf{c})$ is compact in $\mathcal{P}$.

Theorem 15. *The set of probability measures $\Lambda(\mathbf{c})$ is compact in the topology of weak convergence. Moreover, the capacity achieving probability measure μ^* exists.*

Proof. Using a similar approach as in the real case, $\Lambda(\mathbf{c})$ is compact if it is tight and closed by Prokhorov's theorem [Bil99; Sha11]. To see that $\Lambda(\mathbf{c})$ is tight, observe that for any $\epsilon > 0$, there exists an $\mathbf{a}_\epsilon = [a_{1,\epsilon}, a_{2,\epsilon}] \succ \mathbf{0}$ such that for all $\mu \in \Lambda(\mathbf{c})$,

$$\begin{aligned} \Pr(|X_1|^r > a_{1,\epsilon}, |X_2|^r > a_{2,\epsilon}) &\leq \min_{i=1,2} \frac{\mathbb{E}[|X_i|^r]}{a_{i,\epsilon}} \\ &\leq \min_{i=1,2} \frac{c_i}{a_{i,\epsilon}} < \epsilon, \end{aligned} \tag{4.13}$$

where the first inequality follows from the generalized Markov inequality [Mar84, Proposition 2.1]. Now, choose $\mathcal{K}_\epsilon = [-a_{1,\epsilon}, a_{1,\epsilon}] \times [-a_{2,\epsilon}, a_{2,\epsilon}]$, then $\mathcal{K}_\epsilon$ is compact and $\mu(\mathcal{K}_\epsilon) \geq 1 - \epsilon$ for all $\mu \in \Lambda(\mathbf{c})$. As such, $\Lambda(\mathbf{c})$ is tight.

To show that $\Lambda(\mathbf{c})$ is closed, let $\{\mu_n\}_{n=1}^{\infty}$ be a convergent sequence in $\Lambda(\mathbf{c})$ with limit μ_0. Consider the vector valued function $f(\mathbf{x}) = |\mathbf{x}|^r = [|x_1|^r, |x_2|^r]^T$, which is continuous and bounded below. By the Portmanteau theorem for weak convergence [Bil99],

$$\begin{aligned}\mathbb{E}_{\mu_0}[|\mathbf{X}|^r] &= \int f(\mathbf{x})\mu_0(d\mathbf{x}) \\ &\preceq \liminf_{n\to\infty} \int f(\mathbf{x})\mu_n(d\mathbf{x}) \preceq \mathbf{c}. \end{aligned} \tag{4.14}$$

This means that $\mu_0 \in \Lambda(\mathbf{c})$. As our choice of convergent sequence was arbitrary, it follows that $\Lambda(\mathbf{c})$ is closed. As such, $\Lambda(\mathbf{c})$ is compact.

To prove existence of μ^*, by Theorem 14 (page 71) $I(\mathbf{X};\mathbf{Y})$ is continuous on $\Lambda(\mathbf{c})$. As such, by the extreme value theorem, the capacity achieving probability measure μ^* exists. □

Next, we prove that the optimal input distribution is unique.

Theorem 16. *The capacity achieving probability measure μ^* on $\Lambda(\mathbf{c})$ is unique.*

Proof. By [LMG14, Theorem 12], the mutual information is concave. To prove strict concavity, we need to show that if μ_0 and μ_1 both achieve the maximum, then μ_0, μ_1 are identical. By the Lévy continuity theorem, this holds if both probability measures correspond to the same characteristic function. As $\mathbf{X}$ and $\mathbf{N}$ are independent, we have $\phi_{\mathbf{N}}(\boldsymbol{\theta})\phi_{\mu_1}(\boldsymbol{\theta}) = \phi_{\mathbf{N}}(\boldsymbol{\theta})\phi_{\mu_2}(\boldsymbol{\theta})$. Since $\phi_{\mathbf{N}}(\boldsymbol{\theta})$ is non-zero for all $\boldsymbol{\theta}$, the uniqueness of μ^* then follows from the strict concavity of the mutual information and the fact that $\Lambda(\mathbf{c})$ is convex. □

4.3.3 Capacity Lower Bound

We now turn to deriving a lower bound on the capacity defined by (4.4). Our result is given as follows.

Theorem 17. *The capacity of the $AI\alpha SN$ channel defined by (4.4) with noise parameter $\sigma_{\mathbf{N}}$ (see Property 10, page 29) is lower bounded by*

$$C \geq \frac{1}{\alpha}\log\left(1 + \frac{\left(\sqrt{2}\left(\frac{\min\{c_1,c_2\}}{C(r,\alpha)}\right)^{1/r}\right)^{\alpha}}{\sigma_{\mathbf{N}}^{\alpha}}\right) \text{ nats}, \tag{4.15}$$

and $C(r,\alpha)$ is given by (5.10).

Proof. Consider the random vector $\mathbf{N}$ induced by the isotropic α-stable noise N. Recall from Property 10 (page 29) that $\mathbf{N}$ has characteristic function $\phi_{\mathbf{N}}(\boldsymbol{\theta}) = e^{-2^{-\alpha/2}\sigma_{\mathbf{N}}^{\alpha}|\boldsymbol{\theta}|^{\alpha}}$.

To obtain a lower bound on the capacity, suppose that $\mathbf{X}$ is also a random vector induced by an isotropic α-stable random variable with $\sigma_{\mathbf{X}}$. The characteristic function of $\mathbf{X}$ is then given by $\phi_{\mathbf{X}}(\boldsymbol{\theta}) = e^{-2^{-\alpha/2}\sigma_{\mathbf{X}}^{\alpha}|\boldsymbol{\theta}|^{\alpha}}$.

By Property 9 (page 29), the distributions of $\mathbf{N}$ and $\mathbf{X}$ can be written in terms of another random vector $\mathbf{U}$ induced by a complex isotropic α-stable random variable with $\sigma_{\mathbf{U}}$. In particular, there exist matrices $\mathbf{V_N}, \mathbf{V_X}$ satisfying the condition in Property 9 such that

$$\mathbf{X} \stackrel{d}{=} \sigma_{\mathbf{X}} \mathbf{V_X} \mathbf{U}, \quad \mathbf{N} \stackrel{d}{=} \sigma_{\mathbf{N}} \mathbf{V_N} \mathbf{U}. \tag{4.16}$$

Since both $\mathbf{X}$ and $\mathbf{N}$ are α-stable random vectors, the random vector $\mathbf{Y} = \mathbf{X} + \mathbf{N}$ is also, due to the *stability property* in Property 1 (page 23) and can be written in terms of $\mathbf{U}$. In particular, $\mathbf{Y}$ has characteristic function

$$\phi_{\mathbf{Y}}(\boldsymbol{\theta}) = e^{-2^{-\alpha/2}(\sigma_{\mathbf{X}}^{\alpha} + \sigma_{\mathbf{N}}^{\alpha})|\boldsymbol{\theta}|^{\alpha}}. \tag{4.17}$$

This implies that there exists a matrix $\mathbf{V_Y}$ satisfying $\mathbf{V_Y}\mathbf{V_Y}^T = c\mathbf{I}$ for some $c \geq 0$ such that

$$\mathbf{Y} \stackrel{d}{=} (\sigma_{\mathbf{X}}^{\alpha} + \sigma_{\mathbf{N}}^{\alpha})^{1/\alpha} \mathbf{V_Y} \mathbf{U}. \tag{4.18}$$

Let $\sigma_{\mathbf{Y}} = (\sigma_{\mathbf{X}}^{\alpha} + \sigma_{\mathbf{N}}^{\alpha})^{1/\alpha}$. The mutual information is then given by

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &= h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \\ &= h(\sigma_{\mathbf{Y}}\mathbf{V_Y}\mathbf{U}) - h(\sigma_{\mathbf{N}}\mathbf{V_N}\mathbf{U}) \\ &= \log \sigma_{\mathbf{Y}} + h(\mathbf{U}) - \log \sigma_{\mathbf{N}} - h(\mathbf{U}) \\ &= \frac{1}{\alpha} \log \left(1 + \frac{\sigma_{\mathbf{X}}^{\alpha}}{\sigma_{\mathbf{N}}^{\alpha}}\right). \end{aligned} \tag{4.19}$$

All that remains is to write $\sigma_{\mathbf{X}}$ in terms of $\mathbb{E}[|X_1|^r] = \mathbb{E}[|X_2|^r]$. Using [Zol81], we have

$$\mathbb{E}[|X_1|^r] = C(r,\alpha) \left(\frac{\sigma_{\mathbf{X}}}{\sqrt{2}}\right)^r, \tag{4.20}$$

where

$$C(r,\alpha) = \frac{2^{r+1}\Gamma\left(\frac{r+1}{2}\right)\Gamma(-r/\alpha)}{\alpha\sqrt{\pi}\Gamma(-r/2)}. \tag{4.21}$$

The result then follows by substituting (4.20) into (4.19) □

4.4 Parallel Channels

A natural extension of the $AI\alpha SN$ channel is to the case of parallel channels. It is clear that our lower bound bears strong similarities with the capacity of Gaussian noise channels with a power constraint. To study the consequences of this observation, we apply our lower bound to compute

the achievable rate of n parallel $AI\alpha SN$ channels subject to the constraint

$$\sum_{i=1}^{n} \mathbb{E}[|\mathbf{X}|^r] \circ \mathbb{E}[|\mathbf{X}|^r] \preceq \mathbf{c} \circ \mathbf{c}, \ r < \alpha, \tag{4.22}$$

where $\circ$ is the Hadamard product. With this purpose, we consider the scenario where there are n parallel $AI\alpha SN$ channels subject to a sum fractional moment constraint. In order to study this parallel channel setting, we make use of our lower bound to compute achievable rates (in nats). More precisely, we consider the optimization problem

$$\begin{aligned} \underset{\sigma_{\mathbf{X},k},\ k=1,2,\ldots,n}{\text{maximize}} \quad & \sum_{k=1}^{n} \frac{1}{\alpha} \log\left(1 + \frac{\sigma_{\mathbf{X},k}^{\alpha}}{\sigma_{\mathbf{N},k}^{\alpha}}\right) \\ \text{subject to} \quad & \sum_{k=1}^{n} \sigma_{\mathbf{X},k}^2 \leq \sigma_{\mathbf{X},\max}^2 \\ & \sigma_{\mathbf{X},k} \geq 0, \ k = 1, 2, \ldots, n, \end{aligned} \tag{4.23}$$

where $\sigma_{\mathbf{X}}$ can be mapped to $\mathbb{E}[|\mathbf{X}|^r]$ via (4.20).

The key results in this section are a study of the convexity properties of this optimization problem and a comparison with the Gaussian case ($\alpha = 2$). In particular, we derive a bound that allows for the analytical comparison of the achievable rate for a given $\alpha < 2$ and the Gaussian case ($\alpha = 2$). We observe that in the case that $\alpha = 2$, the optimization problem reduces to the standard waterfilling solution in the case of Gaussian channels with a power constraint.

4.4.1 Convexity Properties

Observe that we can rewrite (5.28) as

$$\begin{aligned} \underset{\rho}{\text{maximize}} \quad & \sum_{k=1}^{n} \frac{1}{\alpha} \log\left(1 + \frac{\rho_k^{\alpha/2}}{\sigma_{\mathbf{N},k}^{\alpha}}\right) \\ \text{subject to} \quad & \sum_{k=1}^{n} \rho_k \leq \sigma_{\mathbf{X},\max}^2 \\ & \rho_k \geq 0, \ k = 1, 2, \ldots, n. \end{aligned} \tag{4.24}$$

The optimization problem (4.24) is convex. To see this, observe that $g(\rho_i) = 1 + \frac{\rho_i^{\alpha/2}}{\sigma_{\mathbf{N},k}^{\alpha}}$ are concave since $\alpha \leq 2$. Using the fact that $\log(\cdot)$ is concave and non-decreasing, it follows by [BV04, Eq. (3.10)] that

$$R = \sum_{k=1}^{n} \frac{1}{\alpha} \log\left(1 + \frac{\rho_k^{\alpha/2}}{\sigma_{\mathbf{N},\mathbf{k}}^{\alpha}}\right) \tag{4.25}$$

is a sum of concave functions, which implies the problem in (4.24) is convex.

The convexity of (5.28) implies that the problem can be solved efficiently by standard solvers (e.g., CVX). Moreover, for $\alpha = 2$, the problem reduces to the standard waterfilling problem that arises in the case of Gaussian noise with a power constraint.

4.4.2 The Effect of α

In this section, we are concerned with the variation of α producing modifications in the achievable rates in the case the noise has this parameter. Let $R^*(\alpha)$ denote the achievable rate arising from the solution of (4.24), which is continuous. We study the effect of varying the stability parameter as it is perturbed away from the Gaussian case ($\alpha = 2$) via the distance $|R^*(\alpha) - R^*(2)|$.

We begin by applying Taylor's theorem [KC02; Fis12] to the value function $R^*(\alpha)$, which is differentiable, yielding

$$R^*(\alpha) = R^*(2) + (D_\alpha R^*)(2)(\alpha - 2) + o(|\alpha - 2|), \quad \alpha \to 2, \tag{4.26}$$

where $D_\alpha R^*(2)$ is the derivative of the rate R^* in the direction of α evaluated at the point 2. To produce the distance between the rates, (4.26) is rearranged as following

$$R^*(\alpha) - R^*(2) = (D_\alpha R^*)(2)(\alpha - 2) + o(|\alpha - 2|), \quad \alpha \to 2. \tag{4.27}$$

Applying the triangle inequality, we produce

$$|R^*(\alpha) - R^*(2)| \leq |(D_\alpha R^*)(2)||\alpha - 2| + |o(|\alpha - 2|)|, \quad \alpha \to 2, \tag{4.28}$$

which is equivalent to the following

$$|R^*(\alpha) - R^*(2)| \leq |(D_\alpha R^*)(2)||2 - \alpha| + |o(|2 - \alpha|)|, \quad \alpha \to 2, \tag{4.29}$$

due to the fact that $|\alpha - 2| = \sqrt{(\alpha^2 - 4\alpha + 2^2)} = |2 - \alpha|$. The challenge in evaluating (4.29) lies in evaluating the directional derivative. To proceed, we can adapt a lemma from [Dan67, pg. 23], which provides an expression for the directional derivative, given by

Lemma 3. *Let the real valued function $f(\mathbf{x}, y)$ be twice differentiable on a compact convex subset $\mathcal{X}$ of $\mathbb{R}^{n+1}$, strictly concave in $\mathbf{x}$. Let $\mathbf{x}^*$ be the optimal value of f on $\mathcal{X}$ and denote $\psi(y) = f(\mathbf{x}^*, y)$. Then, the directional derivative of $\psi(y)$ in the direction y is given by*

$$\psi'(y) = f_y(\mathbf{x}^*(y), y) \tag{4.30}$$

Proof. By the implicit function theorem, we can solve for $\mathbf{x}^*$, which yields the function as a continuously differentiable function of y, denoted by

$\mathbf{x}^* = \mathbf{x}^*(y)$. Hence, $\psi(y) = f(\mathbf{x}^*(y), y)$, so that $\psi(y)$ is a continuously differentiable function of y. The derivative is then given by

$$\begin{aligned}\psi'(y) &= f_y(\mathbf{x}^*(y), y) + (\nabla_{\mathbf{x}} f(\mathbf{x}^*(y), y))^T \frac{d\mathbf{x}^*(y)}{dy} \\ &= f_y(\mathbf{x}^*(y), y),\end{aligned} \tag{4.31}$$

as required. □

We now evaluate (4.29). In particular, for the case $\sigma_{\mathbf{N},k} = 1$, $k = 1, 2, \ldots, n$ we have the approximate bound

$$\begin{aligned}&|R^*(\alpha) - R^*(2)| \lesssim |2-\alpha| \\ &\times \left| -\frac{1}{4}\sum_{k=1}^{n} \log\left(1+\rho_k^*\right) + \frac{1}{4}\sum_{k=1}^{n} \frac{\rho_k^* \log \rho_k^*}{1+\rho_k^*} \right|.\end{aligned} \tag{4.32}$$

Fig. 4.1 illustrates the effect of α as it is varied away from the Gaussian case ($\alpha = 2$), using our approximate bound and numerical solution of (4.24). Observe that for values of α near 2 our approximate bound is in good agreement with the numerical result. Note that the difference for larger $|2-\alpha|$ is in part due to the $o(|2-\alpha|)$ term in (4.29), which means that the approximation is not a strict upper bound for large $|2-\alpha|$.

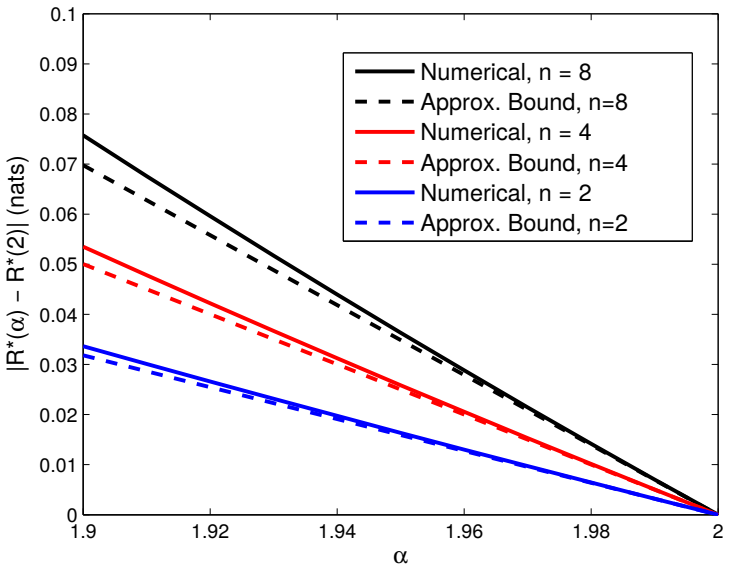


FIGURE 4.1: Plot of the rate-loss using the approximate bound (4.32) and the error obtained by solving (4.24) for varying α with $n = 2, 4$, $\rho_{\max} = 1$, and $\sigma^{\alpha}_{\mathbf{N},k} = 1$, $k = 1, 2, \ldots, n$.

4.5 Chapter conclusion

The Gaussian channels constrained by a second order moment are an exception in characterizing the capacity of continuous channels due to the limited success to characterize the capacity in the domain of continuous channels. A key question is therefore what rates are achievable in the presence of isotropic α-stable interference. We have studied the capacity of the $AI\alpha SN$ channel. In particular, we derived a tractable lower bound on the capacity, as well as existence and uniqueness of the optimal input distribution. We then applied our lower bound and the effect of α on the achievable rate was demonstrated in order to contribute with the understanding of the capacity regarding the parametrization.

We applied our lower bound to the case of parallel $AI\alpha SN$ channels, and demonstrated the effect of α on the achievable rate. The tractability of our lower bound suggests that it may play a useful role in the analysis and design of more complicated systems with α-stable noise, as will be study in the next chapters.

Chapter 5

Physical Layer Design with Dynamic Interference

Practical approaches for dynamic interference are the main interest in this chapter. Firstly, we study the effect of fading in our achievable rate. Next, the constraints are discussed for practical scenarios and some input distributions are considered. The amount of devices in an area is studied through the area spectral analysis. Finally, the parallel channels are revisited, but considering power allocation.

5.1 The Effect of Fading

FADING plays an important role in many Gaussian channels. In the case of the additive α-stable noise, we obtain the channel model given by

$$Y = gX + N, \tag{5.1}$$

where g represents the real-valued fading coefficient. For a fixed g, the capacity is lower bounded by

$$C \geq \frac{1}{\alpha} \log_2 \left(1 + |g|^\alpha M_\alpha \frac{c^\alpha}{\gamma_N^\alpha} \right), \tag{5.2}$$

which follows from Theorem 10, page 52, where M_α is defined in (3.34).

In the case of slow fading (g varies slowly, but randomly according to a fixed distribution F_g), the transmission quality is often characterized by the outage probability, which is given by

$$P_{out} = \Pr(C \leq R_0) \tag{5.3}$$

Using the lower bound in (5.2), we can give an upper bound to the outage probability:

$$P_{out} \leq \Pr \left(\frac{1}{\alpha} \log_2 \left(1 + |g|^\alpha M_\alpha \frac{c^\alpha}{\gamma_N^\alpha} \right) \leq R_0 \right). \tag{5.4}$$

A common choice for the distribution of g^2 is $F_{g^2}(x) = 1 - e^{-\lambda x}$, corresponding to Rayleigh fading. The outage probability is then bounded by

$$\begin{aligned} P_{out} &\leq \Pr\left(|g| \leq \frac{\gamma_N}{M_\alpha^{\frac{1}{\alpha}} c} \left(2^{\alpha R_0} - 1\right)^{\frac{1}{\alpha}} \right) \\ &= 1 - \exp\left[-\lambda \left(\frac{\gamma_N^2}{M_\alpha^{\frac{2}{\alpha}} c^2} \left(2^{\alpha R_0} - 1\right)^{\frac{2}{\alpha}} \right) \right] \\ &= P_{UB} \end{aligned} \tag{5.5}$$

In Fig. 5.1, we represent the outage probability upper bound, where the simulated points are plotted based on (5.4) under the Rayleigh fading, by means of 10^4 samples. In turn, the exact curves are plotted using (5.5). Observe that for small R_0 the outage probability upper bound is heavily influenced by α.

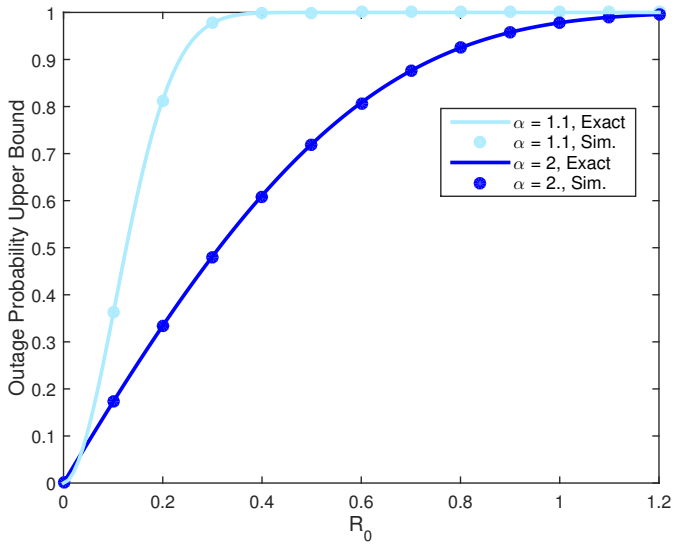


FIGURE 5.1: Plot of the outage probability upper bound (5.5) for varying R_0 and α, with $\beta = 0$, $c = \gamma_N = 1$, $\delta_N = 0$, and $\lambda = 1$.

5.2 Achievable Rates with Dynamic Interference

Rapid changes in the active transmitter set is a characteristic of wireless communication networks with very short transmissions, which arises in M2M communications. A consequence of the rapid changes in the active transmitter set is that the interference is *dynamic*. We have shown that dynamic interference is not Gaussian, as discussed in Section 2.5.1. In fact, the interference is isotropic α-stable for large scale networks with interferers located according to a PPP.

In this section, we derive the achievable rate for the access point at the origin. Unlike the power constrained Gaussian noise channel, tractable expressions are not known for the power constrained $AI\alpha SN$ channel. For this reason, it is desirable to consider alternative constraints.

To characterize the capacity of the $AI\alpha SN$ channel subject to the constraints in (5.7), recall the output y in (2.51), given by

$$y = r_d^{-\eta/2} h_d x_d + I. \tag{5.6}$$

where $h_{d,t} \sim \mathcal{CN}(0,1)$ is the Rayleigh fading coefficient and $x_{d,t}$ is the baseband emission for the typical user.

One choice of constraints is the combination of amplitude and fractional moment constraints. In particular, the input signal x_d in (5.6) is required to satisfy

$$\begin{aligned} \mathbb{E}[|\mathrm{Re}(x_d)|^r] &\leq c \\ \mathbb{E}[|\mathrm{Im}(x_d)|^r] &\leq c \\ |\mathrm{Re}(x_d)| &\leq A \\ |\mathrm{Im}(x_d)| &\leq A, \end{aligned} \tag{5.7}$$

where $0 < r < \alpha$. Note that the presence of the amplitude constraint ensures that the input has finite moments, including power.

We proceed in two steps. First, we relax the amplitude constraints and consider the capacity optimization problem given by

$$\begin{aligned} \underset{\mu \in \mathcal{P}}{\text{maximize}} \quad & I(X;y) \\ \text{subject to} \quad & \mathbb{E}[|\mathrm{Re}(X)|^r] \leq c, \\ & \mathbb{E}[|\mathrm{Im}(X)|^r] \leq c, \end{aligned} \tag{5.8}$$

where $\mathcal{P}$ is the set of probability measures on $\mathbb{C}$ and $0 < r < \alpha$. The unique solution (see Chapter 4) to (5.8) is lower bounded in the following corollary, which arrives directly from Theorem 17 (page 73).

Corollary 4. *For fixed r_d and h_d, the capacity of the additive isotropic $\frac{4}{\eta}$-stable noise channel in (5.6) subject to the fractional moment constraints in (5.8) is lower bounded by:*

$$C_L = \frac{\eta}{4} \log \left(1 + \frac{\left(\sqrt{2|r_d^{-\frac{\eta}{2}} h_d|^2} \left(\frac{c}{C(r, \frac{4}{\eta})} \right)^{\frac{1}{r}} \right)^{\frac{4}{\eta}}}{\sigma_{\mathbf{N}}^{\frac{4}{\eta}}} \right), \tag{5.9}$$

where $\Gamma(\cdot)$ is the Gamma function and

$$C\left(r,\frac{4}{\eta}\right)=\frac{2^{r+1}\Gamma\left(\frac{r+1}{2}\right)\Gamma(-\eta r/4)}{\frac{4}{\eta}\sqrt{\pi}\Gamma(-r/2)}. \tag{5.10}$$

Proof. We consider the case that x_d is an isotropic α-stable random variable satisfying the constraints in (5.8). By Theorem 17 (page 73), the mutual information of the channel $Y=x_d+I$ is given by

$$I(x_d;Y)=\frac{\eta}{4}\log\left(1+\frac{\left(\sqrt{2}\left(\frac{c}{C(r,\frac{4}{\eta})}\right)^{1/r}\right)^{\frac{4}{\eta}}}{\sigma_{\mathbf{N}}^{\frac{4}{\eta}}}\right) \tag{5.11}$$

The result then follows by observing that $r_d^{-\frac{\eta}{2}}h_dx_d$ is also an isotropic $\frac{4}{\eta}$-stable random variable with parameter $|r_d^{-\frac{\eta}{2}}h_d|\sigma_{\mathbf{N}}$ using the fact that x_d is isotropic and Property 2 (page 23). □

One can notice that the achievable rate in Corollary 4 is obtained by using input signals that are isotropic α-stable random variables, which does not satisfy the amplitude constraints in (5.7). The second step in characterizing the capacity of the $AI\alpha SN$ channel subject to (5.7) is therefore to consider a *truncated isotropic α-stable* input. This guarantees the amplitude constraints are satisfied and, as we will show, yields a mutual information in the $AI\alpha SN$ channel that is well approximated by Corollary 4 for a sufficiently large truncation level T.

The truncated isotropic α-stable random variables are defined as follows. Let X be an isotropic α-stable random variable, with real part X_r and imaginary part X_i. The truncation of X, denoted by X_T, is given by

$$X_T=\begin{cases}X, & |X_r|\le T,\ |X_i|\le T\\ \operatorname{sign}(X_r)T+iX_i, & |X_r|>T,\ |X_i|\le T\\ X_r+i\operatorname{sign}(X_i)T, & |X_r|\le T,\ |X_i|>T\\ \operatorname{sign}(X_r)T+i\operatorname{sign}(X_i)T, & |X_r|>T,\ |X_i|>T.\end{cases} \tag{5.12}$$

Using the truncated isotropic α-stable input, an achievable rate of the amplitude and fractional moment constrained $AI\alpha SN$ channel is obtained by evaluating the mutual information $I(y;X_T)$, where y is the output of the channel in (5.18). In fact, using a similar argument to that for the power constrained Gaussian noise channel presented in Definition 26, it is straightforward to show that all rates $R<I(y;X_T)$ are achievable by using a codebook consisting of 2^{nR} codewords $W^n(1),\ldots,W^n(2^{nR})$ with $W_i(w),\ i=1,2,\ldots,n,\ w=1,2,\ldots,2^{nR}$ independent truncated isotropic α-stable random variables.

Unfortunately, truncated isotropic α-stable inputs do not lead to a closed-form mutual information for the channel in (5.18). In fact, only scaling laws for the capacity have been recently derived for real-valued inputs [EPK17]. In order to characterize the achievable rates in the presence of dynamic interference, we therefore approximate $I(X_T; y)$ by the lower bound in Corollary 4.

To verify that this approximation is indeed accurate, we numerically compute the mutual information $I(X_T; y)$ and compare it with the result in Corollary 4 in Fig. 5.2 and Fig. 5.3 for $\alpha = 1.7$ and $\alpha = 1.3$, respectively. Observe that for a sufficiently large truncation level, the approximation based on Corollary 4 is in good agreement with $I(X_T; y)$. Moreover, the achievable rate is significantly larger than the case of a Gaussian input. This suggests that Gaussian signaling is not necessarily desirable in the presence of dynamic interference.

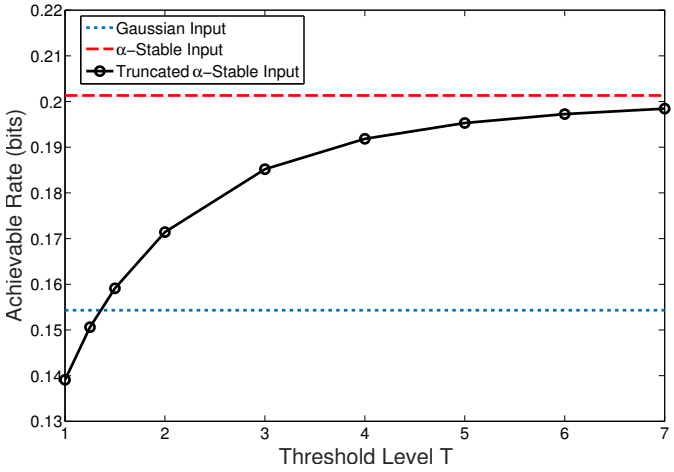


FIGURE 5.2: Achievable rates for an $AI\alpha SN$ channel with $\alpha = 1.7$, $\sigma_{\mathbf{N}} = 0.1$ and a constraint $\mathbb{E}[|X|] \leq 1$. The curves correspond to a Gaussian input, an isotropic α-stable input and a truncated isotropic α-stable input (defined in (5.26)).

5.3 Area Spectral Efficiency Analysis

In this section, we investigate the effect of device density λ on network performance. In particular, we study the area spectral efficiency, which is defined as the expected total rate per square meter. Its importance is explained due to the tradeoff between the distance of each device and its base station. In fact, the interference increases when the density is increased. Formally, let $A_1 \subset A_2 \subset \cdots$ be a sequence of discs such that $\mathrm{Area}(A_n) \to \infty$ as $n \to \infty$. The area spectral efficiency is then given by

$$\zeta = \lim_{n\to\infty} \frac{1}{\mathrm{Area}(A_n)} \mathbb{E}\left[\sum_{i\in\Phi(A_n)} R_i(A_n)\right], \tag{5.13}$$

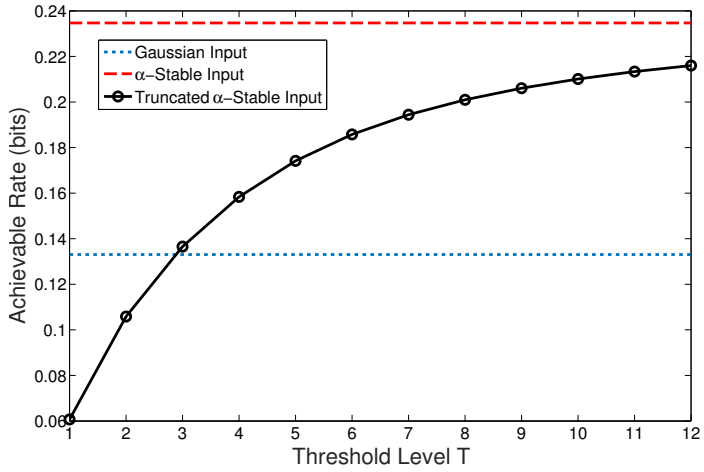


FIGURE 5.3: Achievable rates for an $AI\alpha SN$ channel with $\alpha = 1.3$, $\sigma_{\mathbf{N}} = 0.5$ and a constraint $\mathbb{E}[|X|] \leq 1$. The curves correspond to a Gaussian input, an isotropic α-stable input and a truncated isotropic α-stable input (defined in (5.26)).

where $\Phi(A_n)$ is the PPP Φ restricted to the disc A_n and $R_i(A_n)$ corresponds to the achievable rate with a truncated isotropic α-stable input and devices in $\Phi(A_n)$.

The area spectral efficiency in the large-scale network detailed in Section 2.5 is given in the following theorem.

Theorem 18. *The area spectral efficiency with device locations governed by a PPP, dynamic interference and truncated isotropic α-stable inputs is given by*

$$\zeta = \lambda \mathbb{E}_{r_d,h_d}[R_i], \tag{5.14}$$

where R_i is the achievable rate with a truncated isotropic α-stable input and devices in Φ.

Proof. See Appendix A. □

As observed in Section 5.2, $R_i = I(y_i; X_T)$ does not have a closed-form expression which makes characterizing the area spectral efficiency ζ challenging. To proceed, we exploit the approximation of $I(y_i; X_T)$ based on Corollary 4. In particular, we obtain the following approximation for the area spectral efficiency

$$\zeta \approx \lambda \mathbb{E}_{r_d,h_d}\left[\frac{\eta}{4}\log\left(1+\frac{\left(\sqrt{2|r_d h_d|^2}\left(\frac{c}{C(r,\frac{4}{\eta})}\right)^{\frac{1}{r}}\right)^{\frac{4}{\eta}}}{\sigma_{\mathbf{N}}^{\frac{4}{\eta}}}\right)\right]$$

$$= \frac{\lambda\eta}{4}\mathbb{E}_{r_d,h_d}\left[\log\left(1+\frac{\left(\sqrt{2|r_d h_d|^2}\left(\frac{c}{C(r,\frac{4}{\eta})}\right)^{\frac{1}{r}}\right)^{\frac{4}{\eta}}}{\pi\lambda C_{\frac{\eta}{4}}^{-1}\mathbb{E}[|\mathrm{Re}(h_k x_k)|^{\frac{4}{\eta}}]}\right)\right] \tag{5.15}$$

which is tight when the truncation level for the input T is sufficiently large.

The expression in (5.15) provides insight into the effect of the device density λ. In particular, consider a function of the form

$$f(\lambda) = \lambda \log\left(1+\frac{1}{\lambda}\right), \tag{5.16}$$

which captures the dependence of the spatial rate density approximation in (5.15) on the device density λ. We seek to find a stationary point such that $f'(\lambda) = 0$. Evaluating the derivative yields the condition

$$\log\left(1+\frac{1}{\lambda}\right) - \frac{1}{1+\lambda} = 0. \tag{5.17}$$

Since $\log x > 1 - \frac{1}{x}$ for $x > 1$, it follows that $\log\left(1+\frac{1}{\lambda}\right) > \frac{1}{1+\lambda}$ and hence for $\lambda > 0$, $f'(\lambda) > 0$. This implies that the area spectral efficiency ζ is an increasing function of the density λ (illustrated in Fig. 5.4). We therefore conclude that *dense networks maximize the area spectral efficiency*. We remark that dense networks are also desirable for slowly varying active interferer sets [DWLPML16]. This implies that although the optimal signaling strategy for each link is no longer Gaussian, the basic network structure is the same both for dynamic interference and interference arising from a slowly varying active interferer set.

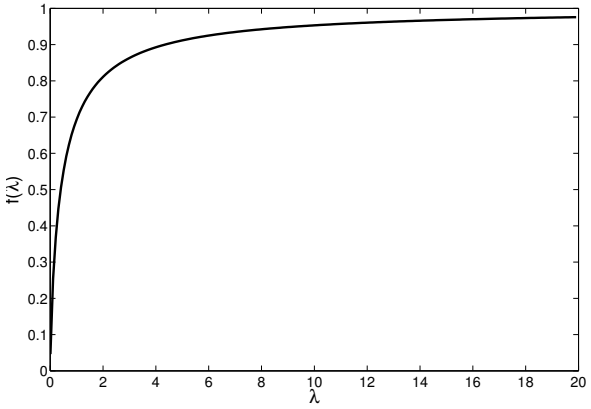


FIGURE 5.4: Plot of $f(\lambda)$ in (5.16).

5.4 Power allocation - Parallel channels

In this section, we consider the problem of power allocation in parallel symmetric α-stable noise channels subject to a power constraint. In the case of a

Gaussian input in the presence of additive Gaussian noise, the rate-optimal solution is the well-known waterfilling algorithm. However, the waterfilling algorithm is tailored to Gaussian noise channel, and for it to be used in symmetric α-stable noise channels, it is necessary to assume that the noise is Gaussian when in fact it is non-Gaussian. As such, it is highly desirable to develop alternative power control strategies that do not rely on the Gaussian noise assumption.

We adopt a two-step approach to the design of power control for the symmetric α-stable noise channel. The first step is to select the input distribution. To this end, we numerically evaluate the achievable rates of the symmetric α-stable noise channel with Gaussian and truncated symmetric α-stable inputs. Here, we show that Gaussian inputs perform comparably or outperform truncated symmetric α-stable inputs, despite the fact that the truncated symmetric α-stable inputs approximately match the input with the noise distribution.

The second step is to optimize the power control for the selected input distribution. Using the capacity lower bound in Theorem 10 (page 52), we develop a new power allocation scheme for Gaussian inputs. We show that this scheme is a convex optimization problem—readily solved using standard solvers—but differs from the waterfilling algorithm. In particular, numerical results show that our power control schemes can outperform by up to 1 bit the rate achieved by waterfilling for Gaussian inputs, where the α-stable noise is assumed to be Gaussian.

Regarding the following sections, in Section 5.4.1, we detail the parallel symmetric α-stable noise model. In Section 5.4.2, we study the effect of different input distributions. In Section 5.4.3, we develop our power control scheme. In Section 5.4.4, we explain how our algorithms can be generalized to complex noise. In Section 5.5 we conclude.

5.4.1 System Model

We are concerned with the memoryless additive symmetric α-stable noise ($AS\alpha SN$) channel

$$Y = hX + N, \tag{5.18}$$

where $h \in \mathbb{R}$ is a constant, $X \in \mathbb{R}$ is the channel input, and $N \in \mathbb{R}$ is symmetric α-stable noise.

We recall the derived closed-form expressions for achievable rates. In particular, the rate of the $AS\alpha SN$ channel with a symmetric α-stable input was derived in Theorem 10 (page 52), given by

$$R = \frac{1}{\alpha} \log \left(1 + |h|^{\alpha} \frac{\gamma_X^{\alpha}}{\gamma_N^{\alpha}} \right), \tag{5.19}$$

where γ_X is the scale parameter of the symmetric α-stable input and γ_N is the scale parameter of the symmetric α-stable noise.

The main problem we consider in this section is power control for K parallel $AS\alpha SN$ channels. In this case, the system consists of K channels defined by

$$Y_i = h_i X_i + N_i,\ i = 1, 2, \ldots, K, \tag{5.20}$$

where $h_i \in \mathbb{R}$, X_i is the real-valued input to the i-th channel and N_i is real symmetric α-stable noise, independent for each i but not necessarily identically distributed. Since each of the channels is independent, it follows from (5.19) that the sum-rate achieved using a symmetric α-stable input for each channel is given by

$$R_S = \sum_{k=1}^{K} \frac{1}{\alpha} \log\left(1 + |h_k|^\alpha \frac{\gamma_{X,k}^\alpha}{\gamma_{N,k}^\alpha}\right). \tag{5.21}$$

This result will provide a basis to construct the optimization problems underlying our power control algorithms.

5.4.2 The Input Distribution

Although the optimal input distribution for the power constrained additive Gaussian noise channel is well-known to be Gaussian, this is not the case for symmetric α-stable noise channels. In fact, the optimal input distribution is known to be discrete [FAF16], with the probability masses dependent on the power level. As such, for the purposes of power control, it is highly desirable to obtain input distributions that yield a high achievable rate with a simple parametric form.

In this section, we investigate the choice of the input distribution for the $AI\alpha SN$ channel in (5.18) subject to a power constraint. Formally, we study lower bounds of the capacity optimization problem

$$\begin{aligned} \underset{\mu \in \mathcal{P}}{\text{maximize}} \quad & I(X;Y) \\ \text{subject to} \quad & \mathbb{E}_\mu[X^2] \leq P, \end{aligned} \tag{5.22}$$

where $\mathcal{P}$ is the set of probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The optimal input distribution for (5.22) is known to be discrete and compactly supported [FAF16]; however, there are no known closed-form expressions or tight bounds for the capacity in the presence of a power constraint.

In order to investigate the choice of the input distribution, we consider the following three choices:

(i) Zero-mean Gaussian input X_G with probability density function

$$p_{X_G}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{\sigma^2}\right). \tag{5.23}$$

(ii) Truncated symmetric α-stable inputs, which are defined as follows. Let X_S be a symmetric α-stable random variable, then the truncated symmetric α-stable input X_T with truncation level T is constructed via

$$X_T = \begin{cases} X_S, & |X_S| \leq T \\ \text{sign}(X_S)T & |X_S| > T. \end{cases} \tag{5.24}$$

The power of the truncated symmetric α-stable input is given by

$$\mathbb{E}[X_T^2] = \int_{-T}^{T} x^2 p_{X_S}(x)dx + 2\int_{T}^{\infty} T^2 p_{X_S}(x)dx, \tag{5.25}$$

where p_{X_S} is the probability density function of the symmetric α-stable random variable X_S.

(iii) Truncated Gaussian inputs, which are defined as follows. Let X_G be a Gaussian random variable, then the truncated Gaussian input $X_{G,T}$ with truncation level T is constructed via

$$X_{G,T} = \begin{cases} X_G, & |X_G| \leq T \\ \text{sign}(X_G)T & |X_G| > T. \end{cases} \tag{5.26}$$

The power of the truncated Gaussian input is given by

$$\mathbb{E}[X_{G,T}^2] = \int_{-T}^{T} x^2 p_G(x)dx + 2\int_{T}^{\infty} T^2 p_G(x)dx, \tag{5.27}$$

where p_G is the probability density function of the Gaussian random variable X_G.

We have selected these distributions for the following reasons. First, the Gaussian input is a standard reference. Second, the truncated symmetric α-stable input is chosen because it approximately matches the noise distribution and also satisfies the finite power constraint. It is also an appropriate choice of input for the case where the channel is both power and amplitude constrained [Smi71], which are essential in practical systems. Third, the truncated Gaussian input is selected as it forms a natural choice of input in the case of both power and amplitude constraints, for further comparison with the truncated symmetric α-stable input. In each case, closed-form expressions for the corresponding rates are not known.

To understand how the choice of input distribution affects the achievable rate, Fig. 5.5 plots the achievable rates using a Gaussian input, a truncated symmetric α-stable input, and also a truncated Gaussian input.

In each case, the power is constrained to be $P = 3$ and the figure shows the impact on the truncation level for each input distribution. As the achievable rates of additive symmetric α-stable noise channels with the inputs detailed in Section 5.4.2 are not known, in the experiment they are estimated via Monte Carlo simulation. In particular, we use $5 \cdot 10^6$ input samples, the entropy of the output and the noise are obtained by estimating the corresponding probability density functions via the kernel method [BA97], which was performed by using a grid of 10^6 points and support $[-200, 200]$.

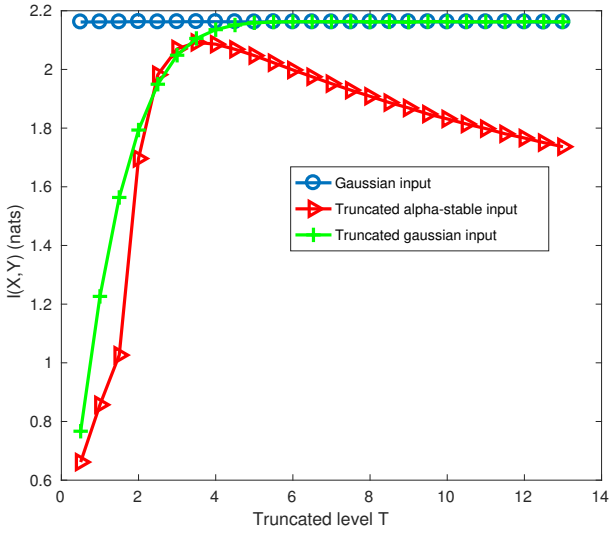


FIGURE 5.5: Comparison of achievable rates using a truncated symmetric α-stable input ($\alpha = 1.4$, $\mathbb{E}[X_T^2] = 3$), a Gaussian input and a truncated Gaussian Input ($\mathbb{E}[X_G^2] = \mathbb{E}[X_{G,T}^2] = 3$) in the presence of symmetric α-stable noise ($\alpha = 1.4$, $\gamma_N = 0.1$).

We observe in Fig. 5.5 that the Gaussian input outperforms both the truncated Gaussian and truncated symmetric α-stable inputs. Similarly, for most choices of the truncation level, the truncated Gaussian input also outperforms the truncated symmetric α-stable input. Moreover, the truncation level rapidly has no effect on the achievable rate for the truncated Gaussian input. We remark that based on extensive numerical experiments, we have observed that these trends hold for a wide range of channel parameters. This suggests that as in the Gaussian noise channel, a Gaussian input is a good choice for the symmetric α-stable noise channel. We also note that the fact that a Gaussian input performs well in the presence of a power constraint differs from the case of an absolute moment constraint, where an α-stable input performs near capacity, as presented in previews chapters.

5.4.3 Power Control Algorithm

In this section, we develop a power control algorithm for Gaussian inputs in parallel $AS\alpha SN$ channels, which is motivated by the results obtained

in Section 5.4.2. Our approach is to view the Gaussian inputs as approximations of symmetric α-stable inputs. This is possible since both of these inputs lie in the α-stable family. As such, the sum-rate in (5.21) can be used to approximate the sum-rate with each input. We verify the performance of our algorithms via numerical simulation.

The Algorithm Zero-mean Gaussian random variables are a special case of symmetric α-stable random variables corresponding to $\alpha = 2$. As such, a method to approximate a general symmetric α-stable random variable with scale parameter γ_X is via a zero-mean Gaussian random variable with variance $2\gamma_X^2$. Moreover, the achievable rate with Gaussian inputs can be approximated by (5.21). These considerations motivate the following optimization problem:

$$\begin{aligned} \underset{\gamma_{X,k},\ k=1,2,\ldots,n}{\text{maximize}} \quad & \sum_{k=1}^{n} \frac{1}{\alpha} \log\left(1 + |h_k|^{\alpha} \frac{\gamma_{X,k}^{\alpha}}{\gamma_{N,k}^{\alpha}}\right) \\ \text{subject to} \quad & \sum_{k=1}^{n} 2\gamma_{X,k}^2 \leq P_{\max} \\ & \gamma_{X,k} \geq 0,\ k = 1, 2, \ldots, n. \end{aligned} \tag{5.28}$$

Here, the parameter $\gamma_{X,k}$ is the parameter for a symmetric α-stable input. Our effective Gaussian inputs are assumed to have the same parameters $\gamma_{X,k}$ and as such, the power levels of the inputs are obtained via [ST94a]

$$P_{X,k} = 2\gamma_{X,k}^2. \tag{5.29}$$

Note that this relationship also implies that the constraint

$$\sum_{k=1}^{n} 2\gamma_{X,k}^2 \leq P_{\max} \tag{5.30}$$

in (5.28) is in fact a sum power constraint for Gaussian inputs.

To solve (5.28), we apply the transformation $\rho_k = \gamma_{X,k}^2$, which yields the problem

$$\begin{aligned} \underset{\rho_k,\ k=1,2,\ldots,n}{\text{maximize}} \quad & \sum_{k=1}^{n} \frac{1}{\alpha} \log\left(1 + |h_k|^{\alpha} \frac{\rho_k^{\alpha/2}}{\gamma_{N,k}^{\alpha}}\right) \\ \text{subject to} \quad & \sum_{k=1}^{n} 2\rho_k \leq P_{\max} \\ & \rho_k \geq 0,\ k = 1, 2, \ldots, n. \end{aligned} \tag{5.31}$$

We observe that the problem in (5.28) is convex (see section 4.4.1), which follows from the fact that the function $\rho_k^{\alpha/2}$ is concave for $0 < \alpha < 2$ and the

linearity of the constraints. As such (5.31) can be solved efficiently using standard convex optimization solvers such as CVX [GB13].

Numerical Results To evaluate the performance of our algorithm, we compare it with the waterfilling algorithm designed for Gaussian noise in the case of two parallel channels. In applying the waterfilling algorithm, we assume that the system does not know the noise is non-Gaussian. As such, the variance of the noise is estimated by observing $N_S = 5 \cdot 10^6$ samples and applying the estimator

$$\hat{\gamma}^2_{G,k} = \frac{1}{N_S - 1} \sum_{i=1}^{N_S} n^2_{i,k}, \; k = 1, 2, \tag{5.32}$$

where $n_{i,k}$ is the i-th noise sample on the k-th channel. Note that since the variance of α-stable noise is infinite, it follows that the variance estimate in (5.32) does not converge (illustrated in Fig. 5.6). Nevertheless, (5.32) provides a means of systematically choosing the noise variance parameter required for the waterfilling algorithm, corresponding to the behavior of a system that does not know the noise is non-Gaussian.

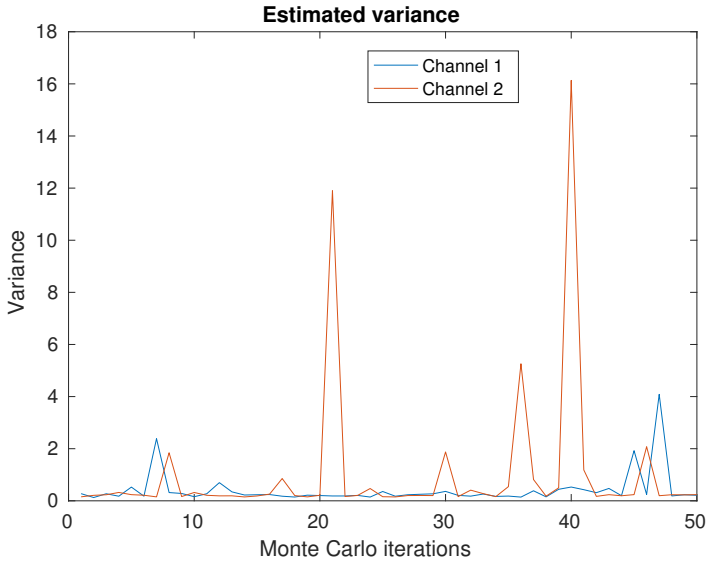


FIGURE 5.6: Estimated variance for waterfilling.

In order to provide a fair comparison with power allocation based on our proposed method, the exponent α is also estimated based on $N_S = 5 \cdot 10^6$ samples, using the characteristic function method in [McC86; Kou81]. This is to ensure that noise parameters are estimated rather than assumed known.

In the experiments, the scale parameter of the symmetric α-stable noise is $\gamma_{N,k} = 0.1, \; k = 1, 2$ and $5 \cdot 10^6$ Gaussian input samples are generated. Table 5.1 shows the estimated achievable rate for each choice of α, channel $\mathbf{h}$ and power allocation method. The rates are estimated using the same procedure as for Fig. 5.5 with 50 Monte Carlo iterations. Observe that our

TABLE 5.1: Estimated Achievable Rates

	Waterfilling (nats)	Proposed (nats)
$\alpha = 1.4$, $\mathbf{h} = [0.5; 0.5]$	1.5607	2.4291
$\alpha = 1.4$, $\mathbf{h} = [0.9; 0.7]$	2.0496	3.2443
$\alpha = 1.4$, $\mathbf{h} = [0.9; 0.85]$	2.2280	3.4203
$\alpha = 1.7$, $\mathbf{h} = [0.5; 0.5]$	2.3745	2.6966
$\alpha = 1.7$, $\mathbf{h} = [0.9; 0.7]$	3.2320	3.5605
$\alpha = 1.7$, $\mathbf{h} = [0.9; 0.85]$	3.4501	3.7454

proposed algorithm implemented in CVX [GB13] outperforms the water-filling algorithm for each choice of parameters. In particular, for $\alpha = 1.4$ and $\mathbf{h} = [0.9; 0.7]$ a gain of more than 1 nat is obtained.

Note that the proposed algorithm does not approximate the symmetric α-stable noise as Gaussian. Instead, the Gaussian input is approximated as a symmetric α-stable random variable (in order to approximate the achievable rate via (5.21)). Our experiments suggest that it is necessary to carefully account for the impulsive nature of the noise for resource allocation in symmetric α-stable noise.

5.4.4 Extensions to the Complex Case

In wireless communications, it is common to use baseband representations which induce the additive isotropic α-stable noise ($AI\alpha SN$) channel. For instance, the $AI\alpha SN$ channel arises in large-scale communication networks with fast-varying active transmitter sets as discussed in Section 2.5.

In Chapter 4, the rate of the $AI\alpha SN$ channel with an isotropic α-stable input with $\mathbb{E}[|X|^r] = c,\ r < \alpha$, was shown to satisfy

$$R_I = \frac{1}{\alpha} \log \left(1 + \frac{\left(|h| \left(\frac{c}{C(r,\alpha)} \right)^{\frac{1}{r}} \right)^{\alpha}}{\gamma_N^{\alpha}} \right), \tag{5.33}$$

Note that the form of (5.33) is similar to the rate in (5.19). This observation provides a straightforward means of extending the algorithms in Section 5.4.3 to the complex case and also can be used as base to the fading study in Section 5.1.

5.5 Chapter conclusion

We have investigated the effect of fading using the closed-form approximation for the achievable rate when an isotropic α-stable input was used. The numerical results suggest that it well approximates the achievable rate when the input signal is truncated, which corresponds to an amplitude constraint, for sufficiently large truncation levels.

Furthermore, in order to establish the effect of device density in large scale networks with dynamic interference, the area spectral efficiency was considered. The approximation suggests that dense networks maximize the area spectral efficiency. This result is consistent with analysis for networks with slowly varying active transmitter sets, which means that the basic network architecture in both settings is the same.

We then considered the problem of power control for parallel symmetric α-stable noise channels and considering a power constraint on the input. We have shown that, in this case, Gaussian inputs are a good choice, consistent with the Gaussian noise case. We then developed a new power control algorithm for Gaussian inputs tailored to symmetric α-stable noise. This algorithm significantly outperforms the rate achieved when the impulsive nature of the noise is ignored. We have also shown that our algorithm can be extended to the case of complex α-stable noise, which arise in wireless communication systems.

Conclusion

THIS thesis focused on the dynamic interference study, which is a subject that will play an essential role in the future of wireless communication networks with very short transmissions, as in machine-to-machine and heterogeneous networks.

We first show that communications in wireless networks can lead to an interference model with an impulsive behavior that will represent a significant limitation in future systems. Several models have been proposed in literature to represent this impulsive behavior. The classical Gaussian is detailed but it has been shown to poorly represent the rare events that strongly limit the performance. We then introduced two non-Gaussian models, namely, the Middleton and α-stable distributions. Though many models have been proposed, they are two of the most used. The rest of our work is based on the stable family. We describe this family and introduce some of the properties for univariate and multivariate variables that will be useful for our work. The heavier tails from α-stable models are depicted, showing that they are better suitable for impulsive behaviors in comparison to Gaussian approaches. Furthermore, the finiteness of α-stable r.v.s moments are discussed, as well as, the series representation in the form of Lepage series.

We then study the capacity of additive noise channels in impulsive environments. Information measures are introduced and the classical Gaussian noise channel capacity for continuous input constrained in power is reviewed, working as a base for our extension to impulsive scenarios. We propose achievable rates for the $AS\alpha SN$ channel, in which $\alpha \in (1, 2]$. To obtain them we use a fractional lower order moment constraint on the input distribution and make the choice of an α-stable input. The stability property allows to obtain a bound. We then derive upper bounds based on Laplace and polynomial distributions. In fact, the derived lower bound is a reasonable approximation of the capacity for α near 2. Moreover, the existence and uniqueness of the optimal input distribution are proved. Next, in order to verify the behavior of our achievable rates, a Blahut-Arimoto algorithm is proposed. The parametrization study of this algorithm was necessary due to the discretization and truncation requirements. As such, we propose guidelines to approximate the capacity. After the acquirement of new bounds, we investigate the lower bound parameters, showing that the quantity c/γ_N, where c is the maximum fractional absolute moment for the input distribution and γ_N is the noise dispersion. This parameter plays a similar role as the SNR for AWGN channels, although the comparison for

different impulsiveness is difficult. Finally, we studied the behavior of the lower bound by means of the bend point, showing similar behavior as the capacity of the Gaussian case.

The next section extends the lower bound to the complex case with $0 < \alpha < 2$. Its importance comes from its natural presence in the context of baseband in wireless cellular communication networks with interferers distributed according to a homogeneous Poisson point process. We also proved the existence and uniqueness of the optimal input distribution. We then applied the derived lower bound to parallel channels constrained by fractional order moment and we studied the convexity of this problem, which allowed its solution by standard convex solvers. Moreover, the effect of the impulsiveness controlled by the exponent α was analyzed, through a perturbation away from the Gaussian case, with the latter working as a base comparison.

Finally, in the last chapter, we apply the achievable rates obtained in Additive Isotropic α-Stable channels in practical scenarios. Firstly, the effect of slow fading modeled with a Rayleigh random variable is studied, producing an upper bound on the outage probability. We question the input distribution. Indeed an α-stable input presents an infinite mean power, which is not realistic. We first consider an amplitude constraint and, in a second time, a power constraint. In the case of amplitude constraint, we show that truncated α-stable r.v.s achieve better rates when compared with the traditional Gaussian input in the presence of dynamic interference. We also investigate the area spectral efficiency, in order to have a tool to understand how the density of devices affects the achievable rate. In fact, dense networks maximize the area spectral efficiency. Finally, when a power constraint is imposed, the Gaussian distributed input is a good choice. We propose a power allocation scheme for parallel channels that takes into account impulsiveness in the noise. Our proposed algorithm outperforms the traditional waterfilling in dynamic interference scenarios.

The work in this thesis motivates many future research directions. One interesting research topic would be to investigate extensions of this work as well as the application of our techniques to other non-Gaussian noise channels. In particular, the problems of bounding the capacity with alternative constraints, the case of $0 < \alpha \leq 1$, and asymmetric α-stable noise distributions remain open. The tractability of our lower bound and its close relationship to the capacity of Gaussian noise channel with a power constraint also suggests that it may be able to play an analogous role in applications. One interesting property that it offers is the fact that there is a parameter that accounts for the impulsiveness in the channel. This impulsiveness may play an important role and the usual traditional capacity ignores such an effect. Another important aspect to consider is an extension towards higher

dimensions. It can be important for MIMO and Massive MIMO in the additive α-stable noise channels. Such a context raises many challenges. Dealing with multi-dimensionnal stable distribution is not straightforward. For example, the dependence structure has to be re-visited, usual correlation being unable to model upper or lower tail dependence (the simultaneaous occurence of large samples on different dimensions). Also, an algorithm to estimate the capacity in a higher dimension case with a fast convergence may be necessary to assess the accuracy of bounds.

An avenue of future research is the study of dynamic interference in networks using general SCMA codebooks that have improved decoding complexity. The general class of additive vector α-stable noise channels seems a promising approach to exploit copula models in this scenario and fundamental limits of the SCMA channel remains an open question.

Appendices

Appendix A

Proof of Theorem 18

In order to compute the area spectral efficiency ζ, observe that the random variables $R_i(A_n)$ are identically distributed (but not independent) since the distances r_d are independent and identically distributed, and the locations of the devices are independently and uniformly distributed in A_n conditioned on the number of devices $N(A_n)$ in A_n [DVJ03]. By the strong law of large numbers for PPPs [Hae13], $\frac{N(A_n)}{\text{Area}(A_n)} \cong \lambda$ a.s. as $n \to \infty$. Let $\epsilon > 0$, it then follows that

$$\begin{aligned}
\zeta &= \lim_{n\to\infty} \frac{1}{\text{Area}(A_n)} \mathbb{E}\left[\sum_{i=1}^{\text{Area}(A_n)\frac{N(A_n)}{\text{Area}(A_n)}} R_i(A_n)\right] \\
&= \lim_{n\to\infty} \frac{1}{\text{Area}(A_n)} \left(\mathbb{E}\left[\sum_{i=1}^{\lfloor \text{Area}(A_n)\lambda_1 \rfloor} R_i(A_n) | \lambda_1 \in [\lambda-\epsilon, \lambda+\epsilon]\right] \right. \\
&\quad \times \Pr(\lambda_1 \in [\lambda-\epsilon, \lambda+\epsilon]) \\
&\quad + \mathbb{E}\left[\sum_{i=1}^{\lfloor \text{Area}(A_n)\lambda_1 \rfloor} R_i(A_n) | \lambda_1 \notin [\lambda-\epsilon, \lambda+\epsilon]\right] \\
&\quad \left. \times \Pr(\lambda_1 \notin [\lambda-\epsilon, \lambda+\epsilon])\right) \qquad \text{(A.1)}
\end{aligned}$$

A direct consequence of the strong law of large numbers of PPPs is that as $n \to \infty$,

$$\Pr(\lambda_1 \in [\lambda-\epsilon, \lambda+\epsilon]) \to 1. \qquad \text{(A.2)}$$

Next, for fixed large n selected A_n such that λA_n is an integer and $\epsilon > 0$ sufficiently small such that $\lambda \text{Area}(A_n)$ is the only integer in $[\lambda-\epsilon, \lambda+\epsilon]$. It then follows that

$$\begin{aligned}
\zeta &= \lim_{n\to\infty} \frac{1}{\text{Area}(A_n)} \text{Area}(A_n)\lambda \mathbb{E}[R_i(A_n)] \\
&= \lambda \lim_{n\to\infty} \mathbb{E}[R_i(A_n)]. \qquad \text{(A.3)}
\end{aligned}$$

To evaluate $\lim_{n\to\infty} \mathbb{E}[R_i(A_n)]$, let y_{i,A_n} be the received signal at the access point served by the i-th device in $\Phi(A_n)$. For fixed r_d, h_d, $R_i(A_n) = I(y_{i,A_n}; X_T)$. From the LePage series representation of the interference in

(2.52), it follows that the signal received by the access point served by the i-th device in Φ satisfies $y_i \overset{(d)}{=} r_d^{-\frac{\eta}{2}} h_d X_T + I, \text{ a.s. as } n \to \infty$.

Since the conditions in [FAF16, Theorem 1] hold, it follows that for fixed r_d, h_d we have $I(y_{i,A_n}; X_T) \to I(y_i; X_T)$ as $n \to \infty$. As $R_i(A_n)$ is positive and $R_i(A_n) \to R_i$ as $n \to \infty$, we then obtain the desired result.

Bibliography

[AALM17] O. Alhussein, I. Ahmed, J. Liang, and S. Muhaidat. "Unified Analysis of Diversity Reception in the Presence of Impulsive Noise". In: *IEEE Transactions on Vehicular Technology* 66.2 (Feb. 2017), pp. 1408–1417.

[AB07] T. Aysal and K. Barner. "Generalized Mean-Median Filtering for Robust Frequency-Selective Applications". In: 55.3 (Mar. 2007), pp. 937–948. DOI: 10.1109/TSP.2006.888882.

[Abr70] N. Abramson. "THE ALOHA SYSTEM: another alternative for computer communications". In: *Proceedings of the November 17-19, 1970, fall joint computer conference*. ACM. 1970, pp. 281–285.

[ABW10] S. Albeverio, Z. Brzeźniak, and J.-L. Wu. "Existence of global solutions and invariant measures for stochastic differential equations driven by Poisson type noise with non-Lipschitz coefficients". In: *Journal of Mathematical Analysis and Applications* 371.1 (2010), pp. 309–322.

[AFGMAA15] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash. "Internet of things: A survey on enabling technologies, protocols, and applications". In: *IEEE Communications Surveys & Tutorials* 17.4 (2015), pp. 2347–2376.

[Ahl74] R. Ahlswede. "The capacity region of a channel with two senders and two receivers". In: *The annals of probability* (1974), pp. 805–814.

[Amm14] H. M. Ammari, ed. *The art of wireless sensor networks. Vol. 1: Fundamentals*. eng. Signals and communication technology. OCLC: 931353087. Heidelberg: Springer, 2014. ISBN: 978-3-642-40009-4 978-3-642-40008-7.

[And05] J. G. Andrews. "Interference cancellation for cellular systems: a contemporary overview". In: *IEEE Wireless Communications* 12.2 (2005), pp. 19–29.

[AP10] N. Andreadou and F.-N. Pavlidou. "Modeling the Noise on the OFDM Power-Line Communications System". In: *IEEE Transactions on Power Delivery* 25.1 (Jan. 2010), pp. 150–157.

[APY16] K. Agha, G. Pujolle, and T. Yahiya. *Mobile and Wireless Networks*. Wiley, 2016. ISBN: 978-1-119-00756-2. URL: https://books.google.fr/books?id=jRzfDAAAQBAJ.

[AQM14] A. Asadi, Qing Wang, and V. Mancuso. "A Survey on Device-to-Device Communication in Cellular Networks". In: *IEEE Communications Surveys & Tutorials* 16.4 (2014), pp. 1801–1819. ISSN: 1553-877X. DOI: 10.1109/COMST.2014.2319555. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6805125 (visited on 10/07/2017).

[Ari72] S. Arimoto. "An algorithm for computing the capacity of arbitrary memoryless channels". In: *IEEE Transactions on Information Theory* 18.1 (1972), pp. 14–20.

[BA97] A. W. Bowman and A. Azzalini. *Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations*. Oxford statistical science series 18. Oxford : New York: Clarendon Press ; Oxford University Press, 1997. ISBN: 978-0-19-852396-3.

[Bak79] C. R. Baker. "Calculation of the Shannon information". In: *Journal of Mathematical Analysis and Applications* 69.1 (1979), pp. 115–123.

[Ber99] D. P. Bertsekas. *Nonlinear programming*. Athena scientific Belmont, 1999.

[Bil99] P. Billingsley. *Convergence of Probability Measures*. John Wiley and Sons, 1999.

[Bla72] R. Blahut. "Computation of channel capacity and rate-distortion functions". In: *IEEE Transactions on Information Theory* 18.4 (1972), pp. 460–473.

[BR95] G Bartle Robert. "The elements of integration and Lebesgue measure". In: *Jhon Wiley and Sons* (1995).

[Bre73] R. P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall, Englewood Cliffs, New Jersey, 1973.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[CAG08] V. Chandrasekhar, J. G. Andrews, and A. Gatherer. "Femtocell networks: a survey". In: *IEEE Communications magazine* 46.9 (2008).

[Car10] P. Cardieri. "Modeling interference in wireless ad hoc networks". In: *IEEE Communications Surveys & Tutorials* 12.4 (2010), pp. 551–572.

[Car78] A. Carleial. "Interference channels". en. In: *IEEE Transactions on Information Theory* 24.1 (Jan. 1978), pp. 60–70. ISSN: 0018-9448. DOI: 10.1109/TIT.1978.1055812. URL: http://ieeexplore.ieee.org/document/1055812/ (visited on 10/15/2017).

[CBVVJOP09] N. Czink, B. Bandemer, G. Vazquez-Vilar, L. Jalloul, C. Oestges, and A. Paulraj. "Spatial separation of multi-user MIMO channels". In: *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*. IEEE. 2009, pp. 1059–1063.

[CE12] A. Chopra and B. Evans. "Outage Probability for Diversity Combining in Interference-Limited Channels". In: *IEEE Transactions on Wireless Communications* 12.2 (2012), pp. 550–560.

[CEG87] M. H. Costa and A. El Gamal. "The capacity region of the discrete memoryless interference channel with strong interference." In: *IEEE Transactions on Information Theory* 33.5 (1987), pp. 710–711.

[CGETS09] A. Chopra, K. Gulati, B. Evans, K. Tinsley, and C. Sreerama. "Performance bounds of MIMO receivers in the presence of radio frequency interference". In: *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2009*. Apr. 2009, pp. 2817–2820.

[CL97] R. C. Cheng and W. Liu. "A continuous representation of the family of stable law distributions". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 59.1 (1997), pp. 137–145.

[CLCC11] S.-M. Cheng, S.-Y. Lien, F.-S. Chu, and K.-C. Chen. "On exploiting cognitive radio to mitigate interference in macro/femto heterogeneous networks". In: *IEEE Wireless Communications* 18.3 (2011).

[Cos83] M. Costa. "Writing on dirty paper (corresp.)" In: *IEEE transactions on information theory* 29.3 (1983), pp. 439–441.

[CS84] S. Cambanis and A. R. Soltani. "Prediction of stable processes: spectral and moving average representations". In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 66.4 (1984), pp. 593–612.

[Csi92] I. Csiszar. "Arbitrarily varying channels with general alphabets and states". In: *IEEE Transactions on Information Theory* 38.6 (1992), pp. 1725–1742.

[CT06] T. Cover and J. Thomas. *Elements of Information Theory, Second Edition*. John Wiley and Sons, Inc., 2006.

[Dan67] J. Danskin. *The Theory of Max-Min*. Springer-Verlag, 1967.

[Dig09] Digi. *Efficient data transfer over cellular networks*. White Paper. 2009.

[DVJ03] D. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer, 2003.

[DWLPML16] M. Ding, P. Wang, D. López-Pérez, G. Mao, and Z. Lin. "Performance impact of LoS and NLoS transmissions in dense cellular networks". In: *IEEE Transactions on Wireless Communications* 15.3 (2016), pp. 2365–2380.

[EAPH13] O. El Ayach, S. W. Peters, and R. W. Heath. "The practical challenges of interference alignment". In: *IEEE Wireless Communications* 20.1 (2013), pp. 35–42.

[ECD08] T. Erseghe, V. Cellini, and G. Dona. "On UWB Impulse Radio Receivers Derived by Modeling MAI as a Gaussian Mixture Process". In: *IEEE Transactions on Wireless Communications* 7.6 (June 2008), pp. 2388–2396. ISSN: 1536-1276. DOI: 10.1109/TWC.2008.070133. URL: http://ieeexplore.ieee.org/document/4543090/ (visited on 10/16/2017).

[ECFDGS17a] M. Egan, L. Clavier, M. de Freitas, L. Dorville, J.-M. Gorce, and A. Savard. "Wireless communication in dynamic interference". In: *Proc. IEEE Global Communications Conference (GLOBECOM)*. 2017.

[ECFDGS17b] M. Egan, L. Clavier, M. de Freitas, L. Dorville, J.-M. Gorce, and A. Savard. "Wireless communication in dynamic interference". In: *IEEE Globecom* (2017 - Accepted for publication).

[ECZFG18] M. Egan, L. Clavier, C. Zheng, M. Freitas, and J.-M. Gorce. "Dynamic interference in uplink SCMA for large-scale wireless networks without coordination". In: *EURASIP Journal on Wireless Communications and Networking* (2018, under review).

[EFCGPA16] M. Egan, M. de Freitas, L. Clavier, A. Goupil, G. Peters, and N. Azzaoui. "Achievable rates for additive isotropic α-stable noise channels". In: *IEEE International Symposium on Information Theory*. 2016.

[Ega15] M. Egan. "Low-high SNR transition in multiuser MIMO". In: *IET Electronics Letters* 51.3 (2015), pp. 296–298.

[EKR14] J. B. Ernst, S. C. Kremer, and J. J. Rodrigues. "A survey of QoS/QoE mechanisms in heterogeneous wireless networks". en. In: *Physical Communication* 13 (Dec. 2014), pp. 61–72. ISSN: 18744907. DOI: 10.1016/j.phycom.2014.04.009. URL: http://linkinghub.elsevier.com/retrieve/pii/S1874490714000457 (visited on 10/07/2017).

[EPK17] M. Egan, S. M. Perlaza, and V. Kungurtsev. *Capacity Sensitivity of Continuous Channels*. Tech. rep. 9012. Lyon, France: INRIA, 2017.

[FAF12] J. Fahs and I. Abou-Faycal. "On the capacity of additive white alpha-stable noise channels". In: *Proc. IEEE International Symposium on Information Theory*. 2012.

[FAF14] J. Fahs and I. Abou-Faycal. "A cauchy input achieves the capacity of a cauchy channel under a logarithmic constraint". In: *Proc. IEEE International Symposium on Information Theory*. 2014.

[FAF16] J. Fahs and I. Abou-Faycal. "Input constraints and noise density functions: a simple relation for bounded-support and discrete-capacity achieving inputs". In: *arXiv:1602.00878* (2016).

[Far08] S. Farahani. *ZigBee wireless networks and transceivers*. OCLC: ocn190777991. Amsterdam ; Boston: Newnes/Elsevier, 2008. ISBN: 978-0-7506-8393-7.

[FEC17] M. de Freitas, M. Egan, and L. Clavier. "Study of achievable Rates for Additive Symmetric α-Stable Noise Channels". In: *Gretsi* (2017).

[FECGPA17] M. de Freitas, M. Egan, L. Clavier, A. Goupil, G. W. Peters, and N. Azzaoui. "Capacity Bounds for Additive Symmetric α-Stable Noise Channels". In: *IEEE Transactions on Information Theory* (2017).

[FECSG17] M. de Freitas, M. Egan, L. Clavier, A. Savard, and J. Gorce. "Power Control in Parallel Symmetric α-Stable Noise Channels". In: *IEEE Communications Letters* (2017 - under review).

[Fel09] W. Feller. *An introduction to probability theory and its applications. Vol. 1: ...* eng. 3. ed., rev. print., [Nachdr.] Wiley series in probability and mathematical statistics. OCLC: 837266385. S.l.: Wiley, 2009. ISBN: 978-0-471-25708-0.

[FGCE15] N. Farsad, W. Guo, C.-B. Chae, and A. Eckford. "Stable distributions as noise models for molecular communication". In: *Proc. IEEE Global Communications Conference*. 2015.

[FH06] J. Fiorina and W. Hachem. "On the asymptotic distribution of the correlation receiver output for time-hopped UWB signals". In: *IEEE Transactions on Signal Processing* 54.7 (July 2006), pp. 2529–2545. ISSN: 1053-587X. DOI: 10.1109/TSP.2006.871976. URL: http://ieeexplore.ieee.org/document/1643893/ (visited on 10/16/2017).

[FI60] K. Furutsu and T. Ishida. "On the theory of amplitude distribution of impulsive random noise and its application to the atmospheric noise". In: *Journal of the radio research laboratories (Japan)* 7.32 (1960).

[Fis12] E. Fischer. *Intermediate real analysis*. Springer Science & Business Media, 2012.

[FK72] T. S. Ferguson and M. J. Klass. "A representation of independent increment processes without Gaussian components". In: *The Annals of Mathematical Statistics* 43.5 (1972), pp. 1634–1643.

[FMLFSJ17] M. L. de Freitas, W. A. Martins, E. B. de Lima Filho, and W. S. da Silva Júnior. "New Designs for Reduced-Redundancy Transceivers". In: *Circuits, Systems, and Signal Processing* 36.5 (2017), pp. 2075–2101.

[FN99] H. Fofack and J. Nolan. "Tail behavior, modes and other characteristics of stable distributions". In: *Extremes* 2.1 (1999), pp. 39–58.

[Fol13] G. B. Folland. *Real analysis: modern techniques and their applications*. John Wiley & Sons, 2013.

[Gar17] I. Gartner. *Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016*. 2017. URL: http://www.gartner.com/newsroom/id/3598917 (visited on 09/15/2017).

[GB13] M. Grant and S. Boyd. *CVX: Matlab software for disciplined convex programming, version 2.0 beta*. 2013. URL: http://cvxr.com/cvx.

[GCASR10] H. E. Ghannudi, L. Clavier, N. Azzaoui, F. Septier, and P.-a. Rolland. "α-stable interference modeling and cauchy receiver for an IR-UWB Ad Hoc network". In: *IEEE Transactions on Communications* 58.6 (June 2010), pp. 1748–1757.

ISSN: 0090-6778. DOI: 10.1109/TCOMM.2010.06.090074. URL: http://ieeexplore.ieee.org/document/5474639/ (visited on 10/16/2017).

[GDK06] N. Guney, H. Deliç, and M. Koca. "Robust detection of ultra-wideband signals in non-Gaussian noise". In: *IEEE Transactions on Microwave Theory and Techniques* 54.4 (June 2006), pp. 1724–1730.

[GEAT10] K. Gulati, B. Evans, J. Andrews, and K. Tinsley. "Statistics of co-channel interference in a field of poisson and poisson-poisson clustered interferers". In: *IEEE Transactions on Signal Processing* 58.12 (2010), pp. 6207–6222.

[GH72] A. Giordano and F. Haber. "Modeling of atmospheric noise". In: *Radio Science* 7 (1972), pp. 1011–1023.

[Gol05] A. Goldsmith. *Wireless communications*. Cambridge university press, 2005.

[Gor41] R. D. Gordon. "Values of Mills' Ratio of Area to Bounding Ordinate and of the Normal Probability Integral for Large Values of the Argument". In: *The Annals of Mathematical Statistics* 12.3 (1941), pp. 364–366. ISSN: 00034851. URL: http://www.jstor.org/stable/2235868.

[GXR09] K. Ghaboosi, Y. Xiao, and J. J. Robertson. "Overview of IEEE 802.15.1 Medium Access Control and Physical Layers". en. In: *Emerging Wireless LANs, Wireless PANs, and Wireless MANs*. Ed. by Y. Xiao and Y. Pan. Hoboken, NJ, USA: John Wiley & Sons, Inc., Apr. 2009, pp. 105–134. ISBN: 978-0-470-40368-6 978-0-471-72069-0. URL: http://doi.wiley.com/10.1002/9780470403686.ch5 (visited on 10/13/2017).

[Hae13] M. Haenggi. *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2013.

[HC17] K. Hwang and M. Chen. *Big-Data Analytics for Cloud, IoT and Cognitive Computing*. John Wiley & Sons, 2017.

[HM12] J. Hoadley and P. Maveddat. "Enabling small cell deployment with HetNet". In: *IEEE Wireless Communications* 19.2 (Apr. 2012), pp. 4–5. ISSN: 1536-1284. DOI: 10.1109/MWC.2012.6189405. URL: http://ieeexplore.ieee.org/document/6189405/ (visited on 10/07/2017).

[HTLN12a] S. Herath, N. Tran, and T. Le-Ngoc. "On optimal input distribution and capacity limit of Bernoulli-Gaussian impulsive noise channels". In: *IEEE International Conference on Communications (ICC)*. 2012.

[HTLN12b] S. Herath, N. Tran, and T. Le-Ngoc. "On optimal input distribution and capacity limit of Bernoulli-Gaussian impulsive noise channels". In: *IEEE International Conference on Communications* (June 2012), pp. 3429–3433.

[IH98] J. Ilow and D. Hatzinakos. "Analytic alpha-stable noise modeling in a poisson field of interfers or scatterers". In: *IEEE Transactions on Signal Processing* 46.6 (1998), pp. 1601–1611.

[III07] H. Ishikawa, M. Itami, and K. Itoh. "A Study on Adaptive Modulation of OFDM under Middleton's Class-A Impulsive Noise Model". In: *Digest of Technical Papers. International Conference on Consumer Electronics, 2007. ICCE 2007.* Jan. 2007, pp. 1–2.

[ILII00] IEEE Computer Society, LAN/MAN Standards Committee, Institute of Electrical and Electronics Engineers, and IEEE-SA Standards Board. *Supplement to IEEE standard for Information technology– telecommunications and information exchange between systems– local and metropolitan area networks– specific requirements: part 11 : wireless LAN medium access control (MAC) and physical layer (PHY) specifications : Higher-speed physical layer extension in the 2.4 Ghz band*. English. OCLC: 44400669. New York, N.Y., USA: Institute of Electrical and Electronics Engineers, 2000. ISBN: 978-0-7381-1811-6.

[IP16] R. J. Igual Pérez. "Platform Hardware/Software for the energy optimization in a node of wireless sensor networks". PhD thesis. Lille 1, 2016.

[JHMB05] K. Jamieson, B. Hull, A. Miu, and H. Balakrishnan. "Understanding the real-world performance of carrier sense". In: *Proceedings of the 2005 ACM SIGCOMM workshop on Experimental approaches to wireless network design and analysis*. ACM. 2005, pp. 52–57.

[JW93] A. Janicki and A. Weron. *Simulation and chaotic behavior of alpha-stable stochastic processes*. Vol. 178. CRC Press, 1993.

[KC02] D. R. Kincaid and E. W. Cheney. *Numerical analysis: mathematics of scientific computing*. Vol. 2. American Mathematical Soc., 2002.

[KG95] K. J. Kerpez and A. M. Gottlieb. "The error performance of digital subscriber lines in the presence of impulse noise". In: *IEEE Transactions on Communications* 43.5 (1995), pp. 1902–1905.

[KL51] S. Kullback and R. A. Leibler. "On information and sufficiency". In: *The annals of mathematical statistics* 22.1 (1951), pp. 79–86.

[Kou81] I. A. Koutrouvelis. "An iterative procedure for the estimation of the parameters of stable laws: An iterative procedure for the estimation". en. In: *Communications in Statistics - Simulation and Computation* 10.1 (Jan. 1981), pp. 17–28. ISSN: 0361-0918, 1532-4141.

[KS04] M. Katz and S. Shamai. "On the capacity achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels". In: *IEEE Transactions on Information Theory* 50.10 (2004), pp. 2257–2270.

[Kuc09] M. Kuczma. *An introduction to the theory of functional equations and inequalities: Cauchy's equation and Jensen's inequality*. Springer Science & Business Media, 2009.

[LeP89] R. LePage. "Appendix Multidimensional infinitely divisible variables and processes. Part I: Stable case". In: *Probability theory on vector spaces IV*. Springer, 1989, pp. 153–163.

[LM03] A. Lapidoth and S. Moser. "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels". In: *IEEE Transactions on Information Theory* 49.10 (2003), pp. 2426–2467.

[LMG14] H. Li, S. Moser, and D. Guo. "Capacity of the memoryless additive inverse Gaussian noise channel". In: *IEEE Journal on Selected Areas on Communications* 32.12 (2014), pp. 2315–2329.

[LP17] G. Last and M. Penrose. *Lectures on the poisson process*. Vol. 7. Cambridge University Press, 2017.

[LVG12] L. P. Lebedev, I. I. Vorovich, and G. M. L. Gladwell. *Functional analysis: applications in mechanics and inverse problems*. Vol. 41. Springer Science & Business Media, 2012.

[LW14] Z. Lin and H. Wang. *Weak Convergence and Its Applications*. World Scientific, 2014.

[Mar84] A. W. Marshall. "Markov's inequality for random variables taking values in a linear topological space". In: *Lecture Notes-Monograph Series* (1984), pp. 104–108.

[May79] H. Maynard. "A Radon-Nikodỳm theorem for finitely additive bounded measures". In: *Pacific Journal of Mathematics* 83.2 (1979), pp. 401–413.

[McC86] J. H. McCulloch. "Simple consistent estimators of stable distribution parameters". In: *Communications in Statistics - Simulation and Computation* 15.4 (1986), pp. 1109–1136. DOI: 10.1080/03610918608812563. eprint: http://www.tandfonline.com/doi/pdf/10.1080/03610918608812563.

[MFCJBEFC17] D. P. de Mello, M. L. de Freitas, L. C. Cordeiro, W. S. Júnior, I. V. de Bessa, B. Eddie Filho, and L. Clavier. "Verification of Magnitude and Phase Responses in Fixed-Point Digital Filters". In: *XXXV Simpósio Brasileiro de Telecomunicações e processamento de sinais - SBRT* (2017), pp. 1184–1188.

[MGC05] H Meng, Y. L. Guan, and S Chen. "Modeling and analysis of noise effects on broadband power-line communications". In: *IEEE Transactions on Power delivery* 20.2 (2005), pp. 630–637.

[Mid77a] D. Middleton. "Statistical-physical models of electromagnetic interference". In: *IEEE Transactions on Electromagnetic Compatibility* 19.3 (1977), pp. 106–127.

[Mid77b] D. Middleton. "Statistical-Physical Models of Electromagnetic Interference". In: *IEEE Transactions on Electromagnetic Compatibility* EMC-19.3 (Aug. 1977), pp. 106–127.

[Mid99] D. Middleton. "Non-gaussian noise models in signal processing for telecommunications: new methods and results for class A and class B noise models". In: *IEEE Transactions on Information Theory* 45.4 (1999), pp. 1129–1149.

[MK04] R. Metzler and J. Klafter. "The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics". In: *J. Phys. A: Math. Gen.* 37 (2004), pp. 161–208.

[MW03] J. Moller and R. P. Waagepetersen. *Statistical inference and simulation for spatial point processes*. CRC Press, 2003.

[NAHV14] B. Nikfar, T. Akbudak, and A. Han Vinck. "MIMO capacity of class A impulsive noise channel for different levels of information availability at transmitter". In: *IEEE International Symposium on Power Line Communications and Its Applications*. 2014.

[NB13] H. Nikopour and H. Baligh. "Sparse code multiple access". In: *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. IEEE. 2013, pp. 332–336.

[NGK16] K. E. Nolan, W. Guibene, and M. Y. Kelly. "An evaluation of low power wide area network technologies for the Internet of Things". In: IEEE, Sept. 2016, pp. 439–444. ISBN: 978-1-5090-0304-4. DOI: 10.1109/IWCMC.2016.7577098. URL: http://ieeexplore.ieee.org/document/7577098/ (visited on 10/07/2017).

[Nor16] A. Nordrum. *Popular Internet of Things Forecast of 50 Billion Devices by 2020 Is Outdated*. 2016. URL: https://spectrum.ieee.org/tech-talk/telecom/internet/popular-internet-of-things-forecast-of-50-billion-devices-by-2020-is-outdated (visited on 09/15/2017).

[NS95] C. L. Nikias and M. Shao. *Signal processing with alpha-stable distributions and applications*. Wiley-Interscience, 1995.

[OU11] O. Ozel and S. Ulukus. "AWGN channel under time-varying amplitude constraints with causal information at the transmitter". In: *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*. IEEE. 2011, pp. 373–377. DOI: 10.1109/ACSSC.2011.6190022.

[PFFR09] R. Pighi, M. Franceschini, G. Ferrari, and R. Raheli. "Fundamental performance limits of communications systems impaired by impulse noise". In: *IEEE Transactions on Communications* 57.1 (2009), pp. 171–182.

[PG10] K. Penson and K Górska. "Exact and explicit probability densities for one-sided Lévy stable distributions". In: *Physical review letters* 105.21 (2010), p. 210604.

[PLZB03] D. S. Pham, Y. H. Leung, A. Zoubir, and R. Brcic. "On the computational aspect of robust multiuser detection". In: *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology, 2003. ISSPIT 2003*. 2003, pp. 22–25.

[PW10a] P. Pinto and M. Win. "Communication in a Poisson field of interferers-part I: interference distribution and error probability". In: *IEEE Transactions on Wireless Communications* 9.7 (2010), pp. 2176–2186.

[PW10b] P. Pinto and M. Win. "Communication in a poisson field of interferers-part II: channel capacity and interference spectrum". In: *IEEE Transactions on Wireless Communications* 9.7 (2010), pp. 2187–2195.

[PW10c] P. C. Pinto and M. Z. Win. "Communication in a Poisson field of interferers-Part II: Channel capacity and interference spectrum". In: *IEEE Transactions on Wireless Communications* 9.7 (2010), pp. 2187–2195.

[Raa10] F. H. Raab. "Noise model for low-frequency through-the-Earth communication: NOISE MODEL FOR LF/VLF". en. In: *Radio Science* 45.6 (Dec. 2010), n/a–n/a. ISSN: 00486604. DOI: 10.1029/2010RS004378. URL: http://doi.wiley.com/10.1029/2010RS004378 (visited on 10/16/2017).

[Ros90] J. Rosinski. "On Series Representations of Infinitely Divisible Random Vectors". In: *Ann. Probab.* 18.1 (Jan. 1990), pp. 405–430. DOI: 10.1214/aop/1176990956. URL: https://doi.org/10.1214/aop/1176990956.

[RQZ05] R. Raich, H. Qian, and G. Zhou. "Optimization of SNDR for Amplitude-Limited Nonlinearities". In: *IEEE Transactions on Communications*. 2005. DOI: 10.1109/TCOMM.2005.857141.

[Sat81] H. Sato. "The capacity of the Gaussian interference channel under strong interference (corresp.)" In: *IEEE transactions on information theory* 27.6 (1981), pp. 786–788.

[SEA12] K. Srinivas, A. Eckford, and R. Adve. "Molecular communication in fluid media: the additive inverse Gaussian noise channel". In: *IEEE Transactions on Information Theory* 58.7 (2012), pp. 4678–4692.

[Sha+61] C. E. Shannon et al. "Two-way communication channels". In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the University of California. 1961.

[Sha11] A. Shapiro. "Topics in stochastic programming". In: *CORE Lecture Series, Universite Catholique de Louvain* (2011).

[Sha48] C. E. Shannon. "A mathematical theory of communication, Part I, Part II". In: *Bell Syst. Tech. J.* 27 (1948), pp. 623–656.

[Sha49] C. E. Shannon. "Communication in the presence of noise". In: *Proceedings of the IRE* 37.1 (1949), pp. 10–21.

[SKBNLH13] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi. "Non-orthogonal multiple access (NOMA) for cellular future radio access". In: *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*. IEEE. 2013, pp. 1–5.

[Smi71] J. Smith. "The information capacity of amplitude-and-variance-constrained scalar Gaussian channels". In: *Information and Control* 18.3 (1971), pp. 203–219.

[Sou92] E. Sousa. "Performance of a spread spectrum packet radio network link in a Poisson field of interferers". In: *IEEE Trans. Inform. Theory* 38.6 (Nov. 1992), pp. 1743–1754.

[ST94a] G. Samorodnitsky and M. Taqqu. *Stable Non-Gaussian Random Processes*. Chapman and Hall, 1994.

[ST94b] G. Samorodnitsky and M. S. Taqqu. *Stable non-Gaussian random processes: stochastic models with infinite variance*. Vol. 1. CRC press, 1994.

[SWH17] R. S. Sinha, Y. Wei, and S.-H. Hwang. "A survey on LPWA technology: LoRa and NB-IoT". en. In: *ICT Express* 3.1 (Mar. 2017), pp. 14–21. ISSN: 24059595. DOI: 10.1016/j.icte.2017.03.004. URL: http://linkinghub.elsevier.com/retrieve/pii/S2405959517300061 (visited on 10/07/2017).

[TMIPWMC16] V. Toldov, J. Meijers, R. Igual-Pérez, R. Wolhuter, N. Mitton, and L. Clavier. "Performance evaluation of LoRa radio solution for PREDNET wildlife animal tracking project". In: *LPWAN 2016*. Paris Roissy, France, May 2016. URL: https://hal.inria.fr/hal-01288077.

[Tol15] M. Tolstrup. *Indoor Radio Planning: A Practical Guide for 2G, 3G and 4G*. John Wiley & Sons, Inc., 2015.

[UZ99] V. V. Uchaikin and V. M. Zolotarev. *Chance and stability: stable distributions and their applications*. Walter de Gruyter, 1999.

[Vas00] S. V. Vaseghi. "NOISE AND DISTORTION". In: (2000).

[Vas84] K Vastola. "Threshold detection in narrow-band non-Gaussian noise". In: *IEEE Transactions on Communications* 32.2 (1984), pp. 134–139.

[VLNKMS17] B. Vejlgaard, M. Lauridsen, H. Nguyen, I. Kovács, P. Mogensen, and M. Sørensen. "Coverage and Capacity Analysis of Sigfox, LoRa, GPRS, and NB-IoT". In: *Proceedings of the IEEE 85th Vehicular Technology Conference, Sydney, Australia*. 2017, pp. 4–7.

[VTNH14] H. Vu, N. Tran, T. Nguyen, and S. Hariharan. "Estimating Shannon and Constrained Capacities of Bernoulli-Gaussian Impulsive Noise Channels in Rayleigh Fading". In: 62.6 (June 2014), pp. 1845–1856.

[WA12] S. Weber and J. Andrews. "Transmission capacity of wireless networks". In: *Foundations and Trends in Networking*. Vol. 5. 2-3. NOW Publishers, 2012.

[Wal96] C. Walck. *Hand-book on statistical distributions for experimentalists*. Tech. rep. 1996.

[WAYDV07] S. P. Weber, J. G. Andrews, X. Yang, and G. De Veciana. "Transmission capacity of wireless ad hoc networks with successive interference cancellation". In: *IEEE Transactions on Information Theory* 53.8 (2007), pp. 2799–2814.

[WCXDZ13] J. Wan, M. Chen, F. Xia, L. Di, and K. Zhou. "From machine-to-machine communications towards cyber-physical systems". In: *Computer Science and Information Systems* 10.3 (2013), pp. 1105–1128.

[WKZ11] J. Wang, E. Kuruoglu, and T. Zhou. "Alpha-stable channel capacity". In: *IEEE Communications Letters* 15.10 (2011), pp. 1107–1109.

[WPS09] M. Win, P. Pinto, and L. Shepp. "A Mathematical Theory of Network Interference and Its Applications". In: *Proc. IEEE* 97.2 (Feb. 2009), pp. 205–230.

[WSKTEK15] A. Walid, E. Sabir, A. Kobbane, T. Taleb, and M. El Koutbi. "On improving network capacity for downlink and uplink of two-tier LTE-FDD networks". In: IEEE, Aug. 2015, pp. 1168–1173. ISBN: 978-1-4799-5344-8. DOI: 10.1109/IWCMC.2015.7289248. URL: http://ieeexplore.ieee.org/document/7289248/ (visited on 10/07/2017).

[WST09] K. Wiklundh, P. Stenumgaard, and H. Tullberg. "Channel capacity of Middleton's class A interference channel". In: *Electronics letters* 45.24 (2009), pp. 1227–1229.

[WV12] Y. Wu and S. Verdú. "Functional properties of minimum mean-square error and mutual information". In: *IEEE Transactions on Information Theory* 58.3 (2012), pp. 1289–1301.

[YP03] X. Yang and A. Petropulu. "Co-channel interference modeling and analysis in a Poisson field of interferers in wireless communications". In: *IEEE Transactions on Signal Processing* 51.1 (2003), pp. 64–76.

[YV05] X. Yang and N. Vaidya. "On physical carrier sensing in wireless ad hoc networks". In: *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*. Vol. 4. IEEE. 2005, pp. 2525–2535.

[Zol57] V. Zolotarev. "Mellin-Stieltjes transforms in probability theory". In: *Theory Probab. Appl.* 2.4 (1957), pp. 433–460.

[Zol81] V. M. Zolotarev. "Integral transformations of distributions and estimates of parameters of multidimensional spherically symmetric stable laws". In: *Contributions to Probability*. Ed. by J. Gani and V. K. Rohatgi. 1981, pp. 283–305.

[Zol86] V. M. Zolotarev. *One-dimensional stable distributions*. Vol. 65. American Mathematical Soc., 1986.

[ZQ06] D. Zha and T. Qiu. "Underwater sources location in non-Gaussian impulsive noise environments". In: *Digital Signal Processing* 16 (2006), pp. 149–163.