Wireless Channel Modeling Based on Extreme Values Theory for Ultra-Reliable Low Latency Communications

by

Niloofar Mehrnia

A Dissertation Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Doctor of Philosophy

in

Electrical and Electronics Engineering

KOÇ ÜNİVERSİTESİ

October, 2022

Wireless Channel Modeling Based on Extreme Values Theory for Ultra-Reliable Low Latency Communications

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a doctoral dissertation by

Niloofar Mehrnia

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Committee Members:

Prof. Sinem Çöleri

Assoc. Prof. Ertuğrul Başar

Prof. Fikri Karaesmen

Assoc. Prof. Mehmet Kemal Özdemir

Prof. Hakan Ali Çırpan

Date:

To my brothers, Milad and Miad, who have my endless pure love.

ABSTRACT

Wireless Channel Modeling Based on Extreme Values Theory for Ultra-Reliable Low Latency Communications
Niloofar Mehrnia
Doctor of Philosophy in Electrical and Electronics Engineering
October, 2022

Ultra-reliable low latency communication (URLLC) is one of the most important features of the fifth generation (5G) networks with the aim of supporting mission critical applications, such as remote control of robots, remote surgery, autonomous vehicles and vehicular teleoperation applications. At URLLC, the packet error rate (PER) is guaranteed to be as low as 10^{-9}-10^{-5} to address the strict reliability constraint, while the acceptable latency is on the order of a few milliseconds. A key building block in the design of ultra-reliable communication systems is a wireless channel model that captures the statistics of rare events occurring due to significant fading. Extreme value theory (EVT) is a powerful framework that characterizes the probabilistic distribution of infrequent extreme events or equivalently the tail distribution. In this thesis, we propose a novel methodology based on EVT to statistically model the behavior of extreme events in a wireless channel for ultra-reliable communication.

In the first part of the thesis, we initially include techniques based on EVT for fitting the lower tail distribution of the received power to the generalized Pareto distribution (GPD), determining the optimum threshold over which the tail statistics are derived, ascertaining the optimum stopping condition on the number of samples required to estimate the tail statistics by using GPD, and finally, assessing the validity of the derived Pareto model.

Second, we model the tail distribution of non-stationary channel based on EVT by including techniques for splitting the channel data sequence into multiple groups concerning the environmental factors causing non-stationarity, and fitting the lower tail distribution of the received power in each group to the GPD. The proposed approach also consists of optimally determining the time-varying threshold over

which the tail statistics are derived as a function of time, and assessing the validity of the derived Pareto model.

Third, we propose EVT-based framework dealing with relatively low number of data samples to estimate the optimal transmission rate and validate it by assessing the outage probability so that reliability constraints are met with a given confidence for ultra-reliable communications.

Fourth, we propose a novel channel modeling methodology based on multivariate EVT (MEVT) for a system using spatial diversity in multiple input multiple output (MIMO)-URLLC to derive the lower tail statistics of the received signal power in multiple dimensions while efficiently dealing with a massive amount of corresponding data. Accordingly, we adopt EVT to determine the optimum threshold over which the tail statistics are derived by using the UGPD model, validate the final model by using probability plots, and utilize MEVT to model the tail of the joint probability distribution by using the EVT-based logistic distribution and Poisson point process approaches.

In the second part of the thesis, we address the requirement of ultra-reliability at the upper communication level by proposing a new algorithm based on the prior collection of data to eliminate power-consuming beam-tracking techniques while ensuring high received power with minimum diversity level.

ÖZETÇE

Ultra Güvenilir Düşük Gecikmeli Haberleşme için Uç Değer Teorisine Dayalı Kablosuz Kanal Modelleme
Niloofar Mehrnia
Elektrik ve Elektronik Mühendisliği, Doktora
Ekim 2022

Ultra güvenilir düşük gecikmeli haberleşme (URLLC), robotların uzaktan kontrolü, uzaktan cerrahi, otonom araçlar, ve araç teleoperasyon gibi kritik görev uygulamalarını desteklemek amacıyla beşinci nesil (5G) ağların en önemli özelliklerinden biridir. URLLC'de, katı güvenilirlik kısıtlamasını ele almak için paket hata oranının (PER) 10^{-9}-10^{-5} kadar düşük olması garanti edilirken, kabul edilebilir gecikme birkaç milisaniye mertebesindedir. Ultra güvenilir haberleşme sistemlerinin tasarımında önemli bir yapı taşı, önemli ölçüde sönümleme nedeniyle meydana gelen nadir olayların istatistiklerini yakalayan bir kablosuz kanal modelidir. Uç değer teorisi (EVT), sık olmayan aşırı olayların olasılık dağılımını veya eşdeğer olarak kuyruk dağılımını karakterize eden güçlü bir çerçevedir. Bu tezde, ultra güvenilir iletişim için kablosuz bir kanaldaki ekstrem olayların davranışını istatistiksel olarak modellemek için EVT'ye dayalı yeni bir metodoloji öneriyoruz.

Tezin ilk bölümünde, başlangıçta, alınan gücün alt kuyruk dağılımını genelleştirilmiş Pareto dağılımına (GPD) uydurmak, kuyruk istatistiklerinin türetildiği optimum eşiği belirlemek, GPD'yi kullanarak kuyruk istatistiklerini tahmin etmek için gereken örnek sayısında optimum durdurma koşulunu belirlemek, ve son olarak, türetilmiş Pareto modelinin geçerliliğinin değerlendirilmesi için EVT'ye dayalı teknikleri dahil ediyoruz.

İkinci olarak, kanal veri dizisini, durağanlığa neden olan çevresel faktörlerle ilgili olarak birden çok gruba bölmek için EVT teknikleri dahil ederek ve alınan gücün alt kuyruk dağılımını her gruba sığdırarak GPD'ye dayalı durağan olmayan kanalın kuyruk dağılımını modelliyoruz. önerilen yaklaşım aynı zamanda kuyruk istatistiklerinin zamanın bir fonksiyonu olarak türetildiği zamanla değişen eşiğin optimal olarak belirlenmesinden ve türetilmiş Pareto modelinin geçerliliğinin değerlendiril-

mesinden oluşur.

Üçüncü olarak, optimum iletim hızını tahmin etmek ve kesinti olasılığını değerlendirerek doğrulamak için nispeten düşük sayıda veri örneğiyle ilgilenen EVT tabanlı çerçeve öneriyoruz, böylece güvenilirlik kısıtlamaları ultra güvenilir iletişim için belirli bir güvenle karşılanır.

Dördüncüsü, alınan sinyal gücünün alt kuyruk istatistiklerini birden çok boyutta elde etmek için çoklu giriş çoklu çıkış (MIMO)-URLLC'de uzamsal çeşitlilik kullanan bir sistem için çok değişkenli EVT'ye (MEVT) dayalı yeni bir kanal modelleme metodolojisi öneriyoruz. Buna göre, UGPD modelini kullanarak kuyruk istatistiklerinin türetildiği optimum eşiği belirlemek için EVT'yi kullanıyoruz, olasılık grafiklerini kullanarak nihai modeli doğruluyoruz ve EVT tabanlı lojistik dağıtım ve Poisson noktası süreç yaklaşımları kullanarak ortak olasılık dağılımının kuyruğunu modellemek için MEVT'yi kullanıyoruz.

Tezin ikinci bölümünde, güç tüketen ışın izleme tekniklerini ortadan kaldırmak için önceden veri toplamaya dayalı yeni bir algoritma önererek üst iletişim seviyesinde ultra güvenilirlik gereksinimini ele alıyoruz. Burada minimum çeşitlilik seviyesiyle yüksek alınan güç garanti ediliyor.

ACKNOWLEDGMENTS

I would like to express the deepest appreciation to my dear adviser, Dr. Sinem Çöleri, who has been my constant source of inspiration. I was always so impressed by her motivation, enthusiasm, and immense knowledge. Without her never-ending supports, this dissertation would not have been made possible.

A debt of gratitude is also owed to Dr. Mehmet Kemal Özdemir for pointing me toward the Communication Engineering, and providing me with the guides in my MS.c. journey.

I would also like to thank my dissertation committee, Dr. Ertuğrul Başar, Dr. Fikri Karaesmen, Dr. Mehmet Kemal Özdemir, and Dr. Hakan Ali Çırpan for their encouragements and insightful comments.

Last but not least, I appreciate my family and friends for their support and encouragement during my Ph.D. journey, always being my accompany on sad and happy days, and making my life bright with their positive vibes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

5G	fifth generation
ACF	autocorrelation function
AD	Anderson-Darling
ADF	Augmented Dickey-Fuller
AIC	Akaike information criterion
APTR	acknowledge packet transmission rate
AR	autoregressive
AR	average reliability
ARIMA	auto-regressive integrated moving average
AWGN	additive white Gaussian noise
BER	bit error rate
BEVT	bi-variate EVT
BGPD	bi-variate GPD
BIC	Bayesian information criterion
CDF	cumulative distribution function
CI	confidence interval
CIR	channel impulse response
CRN	cognitive radio network
CSI	channel state information
EDF	empirical distribution function
ESR	ergodic secrecy rate
EVT	extreme value theory

FL	federated learning
GARCH	generalized auto-regressive conditional heteroskedasticity
GEV	generalized extreme value
GoF	goodness-of-fit
GPD	generalized Pareto distribution
HD	high definition
HPBW	half power beam-width
i.i.d	independent and identically distributed
i.n.i.d.	independent and non-identically distributed
KDE	kernel density estimation
LoS	line of sight
MA	moving average
MDBA	minimum diversity beam-forming
MEF	mean excess function
MEVT	multivariate EVT
mg	minimum gap
MIMO	multiple input multiple output
MLE	maximum likelihood estimator
mmWaves	millimeterwaves
MRC	maximum-ratio combining
MRL	mean residual life
MSSD	Minimum Sample Size Determination
PACF	partial ACF
PCR	probably correct reliability
PDF	probability density function
PER	packet error rate

PP	probability-probability
PWL	piecewise linear model
QoS	quality of service
QQ	quantile-quantile
RMS	root mean square
RMSE	root mean square error
Rx	receiver
SBR	shooting and bouncing ray
SIR	signal-to-interference ratio
SNR	signal-to-noise-ratio
ToD	tele-operation driving
Tx	transmitter
UEVT	uni-variate EVT
UGPD	uni-variate GPD
URLLC	ultra-reliable low latency communication
V2V	vehicle-to-vehicle
VNA	vector network analyzer
VR	virtual reality
VUE	vehicular user
WI	Wireless InSite

Chapter 1

INTRODUCTION

Ultra-reliable low latency communication (URLLC) as one of the most important features of the fifth generation (5G) and beyond networks requires ultra-reliability of 10^{-9}-10^{-5} packet error rate (PER) with ultra-low latency on the order of milliseconds, to enable mission-critical applications, such as remote control of robots, self-driving cars, and remote surgeries. Establishing a URLLC system that addresses the strict requirements of ultra-reliability necessitates fundamental breakthroughs in the statistical modeling of the wireless channel in the ultra-reliable region, incorporating novel techniques to analyze the lower tail of the distributions encompassing extremely low probabilities, handling and optimizing a large amount of data to model the extreme events occurring infrequently, and establishing transmission strategies. Existing studies on channel modeling for URLLC either provide an extrapolation-based framework that extends the applicability of the available practical channel models in the ultra-reliability relevant regime, or propose performance measures for the ultra-reliability of the channel. However, none of these frameworks propose a channel modeling methodology to derive and verify ultra-reliability statistics. Extreme value theory (EVT) is a unique statistical discipline that develops techniques and models for describing extreme events occurring rarely based on the implementation of mathematical limits as finite-level approximation.

EVT has been used for a long time in a wide variety of fields, such as hydrology to quantify risks of extreme floods, rainfalls, and waves, or financial engineering to estimate the unexpected losses due to the extreme events, but has been recently utilized in communication engineering at the data link and network layers to model the tail statistics of queue length and delay, or derive closed-form asymptotic expressions for the throughput, bit error rate (BER), and PER over different fading channels. Nevertheless, EVT has never been incorporated into wireless channel modeling to estimate the tail statistics and address the constraint in the ultra-reliability region, which requires the development of methodologies for the determination of the optimum threshold below/above which the samples correspond to the extreme events, or to acquire a large number of samples to capture the extreme events occurring rarely. On the other hand, the studies on data link and network layer use the extrapolation of existing average statistics-based channel models in the ultra-reliable region, which have not yet been proven to be accurate experimentally. Therefore, a proper channel modeling approach based on EVT that incorporates the statistics of the channel tail is the key to achieve URLLC not only at the physical layer but also at the upper layers, including data link and network layers. In general, the EVT applications can be categorized into two folds: In the first category, the asymptotic distribution of the maxima/minima of a long finite sequence of independent and identically distributed (i.i.d.) random variables is modeled by the generalized extreme value (GEV) distribution, while in the second category, the distribution of the values exceeding a given threshold is modeled by the GPD. Nevertheless, in URLLC, the focus is on the second category of EVT applications, as all values exceeding a given threshold are considered as extreme events and need to be incorporated into the tail statistical analysis.

In this thesis, we propose a novel channel modeling methodology for URLLC based on EVT to derive the lower tail statistics in a consistent manner while efficiently dealing with a massive amount of corresponding data, for the first time in the literature. Additionally, we propose an EVT-based wireless channel modeling methodology for non-stationary channel data by determining the measurable external factors that cause non-stationarity and then, modeling the parameters of fitting distributions as a change-point function of time. Moreover, a novel methodology based on EVT is proposed in this thesis to determine the optimum transmission rate that addresses the constraints of an ultra-reliable communication. A framework is also proposed to estimate the statistics of the EVT-based transmission rate by incorporating the confidence intervals of the estimated Pareto parameters for different sample numbers to meet the targeted error probability for the ultra-reliable communication. Finally, a multivariate EVT (MEVT)-based channel modeling methodology is proposed for estimating the tail of the joint distribution of the multi-channel data by characterizing the multivariate extremes of a multiple input multiple output (MIMO) system.

1.1 Organization

The rest of the thesis is organized as follows:

- **Chapter 2** expresses a new methodology to model the communication channel at URLLC by adapting EVT to first, determine the optimum threshold over which the tail statistics are derived for stationary processes based on the assumption that all values exceeding the threshold correspond to the extreme events happening rarely, second, determine the stopping conditions for ascertaining the minimum amount of data required to model the tail characteristics, and finally, assess the validity of the final model by using the probability plots. It is demonstrated that the proposed channel modeling methodology achieves a significantly better fit to the empirical data in the lower tail than the conventional extrapolation-based approach, especially in the ultra-reliability region. Furthermore, the conventional approach has been demonstrated to be highly dependent on the portion of the data considered in the extrapolation, which can seriously affect the reliability performance.[1]

- **Chapter 3** proposes a novel channel modeling methodology based on EVT for the statistical characterization of the lower tail distribution of a non-stationary channel data by first, splitting the wireless channel data sequence into smaller groups, each of which corresponds to the same value of the external factor causing non-stationarity and then, fitting the generalized Pareto distribution to the lower tail of the received power distribution in each group. It is demonstrated that the proposed methodology for modeling a non-stationary channel achieves significantly better fit to the empirical data than the modeling approach for a stationary channel.[2]

- **Chapter 4** presents a novel EVT-based framework for the estimation and validation of the optimal transmission rate for ultra-reliable communications through determination of the optimum transmission rate by using the estimated GPD parameters, and then, assessment of the system reliability by means of the outage probability metric. This chapter also demonstrates the superior performance of the proposed methodology in determining the maximum transmission rate compared to the traditional extrapolation-based approaches.[3]

- **Chapter 5** proposes a novel framework based on EVT for the estimation of the optimal transmission rate along with the corresponding confidence interval

---

[1]N. Mehrnia and S. Coleri, "Wireless Channel Modeling Based on Extreme Value Theory for Ultra-Reliable Communications,"in IEEE Transactions on Wireless Communications, vol. 21, no. 2, pp. 1064-1076, Feb. 2022, doi: 10.1109/TWC.2021.3101422.

[2]N. Mehrnia and S. Coleri, "Non-Stationary Wireless Channel Modeling Approach Based on Extreme Value Theory for Ultra-Reliable Communications,"in IEEE Transactions on Vehicular Technology, vol. 70, no. 8, pp. 8264-8268, Aug. 2021, doi: 10.1109/TVT.2021.3091378.

[3]N. Mehrnia and S. Coleri, "Extreme Value Theory Based Rate Selection for Ultra-Reliable Communications,"in IEEE Transactions on Vehicular Technology, vol. 71, no. 6, pp. 6727-6731, June 2022, doi: 10.1109/TVT.2022.3158193.

with a sufficiently low number of samples for ultra-reliable communications by incorporating the estimates of the generalized Pareto distribution (GPD) parameters fitted to the channel tail into the rate selection. We also show that proper estimation of the transmission rate based on the proposed framework requires a lower number of samples compared to the traditional extrapolation-based approaches.[4]

- **Chapter 6** proposes a novel channel modeling methodology based on MEVT for a system using spatial diversity in MIMO-URLLC to derive the lower tail statistics of the received signal power in multiple dimensions. In this chapter, MEVT is adapted to first fit GPD to the tail distribution of the received power samples exceeding a given threshold in each sequence and derive the scale and shape parameters of GPD while the thresholds are determined optimally. Then, Fréchet transformation is applied to each data sequence, so that the marginal distributions of the resulting variables are Fréchet. Finally, upon modeling the tail of the joint distribution as a function of estimated GPD parameters based on two approaches: Logistic distribution and Poisson point process, the validity of the proposed models is assessed by incorporating the mean constraint on the probability measure function of Pichanks coordinates. Additionally, we demonstrate the superiority of the proposed methodology compared to the conventional extrapolation-based methods in providing the best fit to the multivariate extremes.[5]

- **Chapter 7** presents a power efficient beamforming algorithm based on the

[4]N. Mehrnia and S. Coleri, "Incorporation of Confidence Interval into Rate Selection Based on the Extreme Value Theory for Ultra-Reliable Communications,"2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit), 2022, pp. 118-123, doi: 10.1109/EuCNC/6GSummit54941.2022.9815831.

[5]N. Mehrnia and S. Coleri, "Multivariate Extreme Value Theory Based Channel Modeling for Ultra-Reliable Communications,"2022, submitted to IEEE Transactions on Wireless Communications.

prior collection of data to eliminate power consuming beam-tracking techniques while minimum diversity level ensures high received power and ultra reliability at the upper layers of communication. Extensive simulated results for different scenarios illustrate that the proposed scheme guarantees low latency ultra-reliability while maximum throughput is achieved with minimum diversity level to consume less energy.[6]

- **Chapter 8** concludes the study and predicts the possible future works.

[6]N. Mehrnia and S. C. Ergen, "Power Efficient Beam-Forming Algorithm for Ultra-Reliable Low Latency Millimeter-Wave Communications,"2019 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), 2019, pp. 1-5, doi: 10.1109/BlackSeaCom.2019.8812783.

Chapter 2

WIRELESS CHANNEL MODELING BASED ON EXTREME VALUE THEORY FOR ULTRA-RELIABLE COMMUNICATIONS

The content of this chapter has been published in IEEE Transactions on Wireless Communications (TWC).

2.1 Overview

A key building block in the design of ultra-reliable communication systems is a wireless channel model that captures the statistics of rare events occurring due to the significant fading. In this chapter, we propose a novel methodology based on extreme value theory (EVT) to statistically model the behavior of extreme events in a wireless channel for ultra-reliable communication. This methodology includes techniques for fitting the lower tail distribution of the received power to the generalized Pareto distribution (GPD), determining the optimum threshold over which the tail statistics are derived, ascertaining the optimum stopping condition on the number of samples required to estimate the tail statistics by using GPD, and finally, assessing the validity of the derived Pareto model. Based on the data collected within the engine compartment of Fiat Linea under various engine vibrations and driving scenarios, we demonstrate that the proposed methodology provides the best fit to the collected data, significantly outperforming the conventional extrapolation-based methods. Moreover, the usage of the EVT in the proposed method decreases the required number of samples for estimating the tail statistics significantly.

2.2 Introduction

Ultra-reliable and low latency communication (URLLC) is an essential part of beyond fifth generation (5G) networks with the potential to enable mission-critical applications, such as remote control of robots, self-driving cars, and remote surgeries [Popovski et al., 2018], [Gustafson et al., 2014], [Popovski et al., 2019], [Eggers et al., 2019]. URLLC is defined as a communication with targeted packet error rate in the range of 10^{-9}-10^{-5}, and acceptable latency on the order of milliseconds or less [Eggers et al., 2019], [Nielsen et al., 2017]. Addressing the strict requirements of URLLC necessitates fundamental breakthroughs in the statistical modeling of the wireless channel in the ultra-reliable region, incorporating novel techniques to analyze the lower tail of the distributions encompassing extremely low probabilities, and handling and optimizing a large amount of data to model the extreme events occurring infrequently. A proper channel modeling approach is the key to achieve URLLC not only at the physical layer but also at the upper layers, including data link and network layers.

Previous studies on the channel modeling for URLLC focus on either providing a unified framework by extrapolating a wide range of practically important channel models to extend their applicability to the ultra-reliability regime of operation [Eggers et al., 2019], or proposing new channel parameters incorporating extreme reliability requirements into the communication [Swamy et al., 2018], [Swamy et al., 2019], [Angjelichinoski et al., 2019]. In [Eggers et al., 2019], a simple power-law expression is proposed for estimating the tail of the cumulative distribution function (CDF) of the received power in block fading channels in the regime of extremely rare events by extrapolating the commonly used practical fading models. On the other hand, the authors in [Swamy et al., 2018] and [Swamy et al., 2019] propose a new channel parameter by challenging the definition of coherence time, during which channels are considered to be static for an average performance of traditional cellular and WiFi networks. As an alternative, a more nuanced notion of coherence time, considering an ultra-reliability regime, is defined as the time over which a channel is predictable to a given reliability. In [Angjelichinoski et al., 2019], two substitute performance measures have been provided for the reliability of the channel, considering that the exact knowledge of the channel is not available at the ultra-reliability level: averaged reliability, for dynamic environments in which the channel changes frequently; and probably correct reliability, for the static environments in which the most recent channel estimate can be used by the system over the future transmissions. However, still no wireless channel modeling framework has been proposed to derive and verify these ultra-reliability statistics. Deriving an appropriate wireless channel model is immensely important as model uncertainty and/or mismatch degrades the communication performance by several orders of magnitude, which is not acceptable at URLLC level [Angjelichinoski et al., 2019]. One might think that deriving the tail statistics is equivalent to collecting a large amount of data and fitting these data to probabilistic distributions. However, these derived distributions may not capture rare events due to the limited amount of collected data or may not be valid in another setting with different environmental conditions, such as different temperatures and vibrations. Extreme value theory (EVT) is a unique statistical discipline that develops techniques and models for describing rare events based on the implementation of mathematical limits as finite-level approximation [Eggers et al., 2019], [Kotz and Nadarajah, 2000], [Coles et al., 2001].

EVT has been utilized at data link and network layers to model the tail statistics of queue length and delay [Samarakoon et al., 2018], [Liu and Bennis, 2018], [Abdel-Aziz et al., 2020], [Mahmood and Jäntti, 2016], [Song and Li, 2006], [Al-Badarneh et al., 2018a], in a limited context at physical layer for wireless channel modeling, to provide a fit to the whole distribution of large-scale or small-scale fading [Molina-Garcia et al., 2008], [Dahman et al., 2010], [Reyes-Guerrero and Mariscal, 2014], [Sawada et al., 2015], [Li et al., 2018], and to analyze the asymptotic behavior of the ergodic secrecy rate (ESR)/ergodic multicast rate [Xu et al., 2016], [Subhash et al., 2020]. EVT has been applied in [Samarakoon et al., 2018] and [Liu and Bennis, 2018] to characterize the statistics of the large queue lengths and accordingly propose a queue-aware resource allocation approach for enabling ultra reliable and low latency vehicular communications. The authors in [Abdel-Aziz et al., 2020] adopt EVT to characterize the tail, and the excess value of the vehicles' queueing systems and arrival rates to further incorporate its results in the transmission power minimization problem. Moreover, EVT is utilized in [Mahmood and Jäntti, 2016], [Song and Li, 2006], [Al-Badarneh et al., 2018a] to derive closed-form asymptotic expressions for the throughput, bit error rate (BER), and packet error rate (PER) over different fading channels. In [Mahmood and Jäntti, 2016], EVT is used to find a limiting distribution for the PER in the additive white Gaussian noise (AWGN) channel with the goal of indicating the superiority of EVT-based approach to accurately approximate the average PER for the uncoded schemes. In [Song and Li, 2006], the authors apply EVT to find the limiting distribution of the maximum throughput and then, alleviate the complexity of the average throughput analysis for the channels in multi-user diversity. Likewise, in [Al-Badarneh et al., 2018a], EVT is used to analyze the effective throughput, average throughput and average bit error probability of the k-th best link with the highest throughput in the selection diversity schemes over various fading channels, such as Weibull, Gamma, $\alpha - \mu$ and Gamma-Gamma. Nevertheless, these data link and network layer studies use the existing average statistics-based channel models, and therefore, their extrapolation in the ultra-reliable region, the accuracy of which has not yet been analyzed experimentally. On the other hand, the authors in [Molina-Garcia et al., 2008], [Dahman et al., 2010], [Sawada et al., 2015] fit extreme value distribution (EVD) to the whole path-loss or power distribution, concluding that EVD models wireless channel fading better than the well-known models, such as Rayleigh and Rician. Additionally, in [Reyes-Guerrero and Mariscal, 2014] and [Li et al., 2018], generalized extreme value (GEV) distribution is used to model the small scale fading in maritime communication, and root-mean-square (RMS) delay spread for vehicle-to-vehicle (V2V) communication, respectively. Moreover, EVT is applied in [Xu et al., 2016] to fit Gumbel distribution to the lower bound of the ESR, which is then incorporated as the statistical constraints within privacy-preserving problem and achieves higher secrecy rate for downlink transmission in multiuser relay systems. Also, in [Subhash et al., 2020] and [Al-Badarneh et al., 2018b], EVT is used to fit a distribution to the minimum of signal-to-interference ratio (SIR) random variables and then determine the optimum value of the transmit power of a secondary network in a Cognitive Radio Network (CRN) subject to Quality of Service (QoS) constraints. However, EVT has never been incorporated into wireless channel modeling to estimate the tail statistics nor to use the theorems for the consistency of the distributions, stopping conditions on determining the sufficient number of samples required in EVT, and validation procedures to address the reliability constraint at the URLLC levels.

The goal of this chapter is to propose a novel channel modeling methodology for URLLC based on EVT to derive the lower tail statistics of the received signal power in a consistent manner while efficiently dealing with a massive amount of corresponding data, for the first time in the literature. We use received signal power at constant transmit power as channel data as it is equivalent to the squared amplitude of the channel state information [Eggers et al., 2019]. The modeling approach adapts EVT to (i) determine the optimum threshold over which the tail statistics are derived based on the assumption that all values exceeding the threshold correspond to the extreme events happening rarely, (ii) determine the stopping conditions for ascertaining the minimum amount of data required to model the tail characteristics, and (iii) assess the validity of the final model by using probability plots. The original contributions of this chapter are listed as follows:

- We propose a comprehensive channel modeling approach for URLLC based on EVT, for the first time in the literature. The methodology includes techniques for fitting the tail distribution to the generalized Pareto distribution (GPD), determining the optimum threshold over which consistent tail statistics are derived, specifying the optimum stopping conditions on the sufficient number of measurement samples, and assessing the model validity.

- We derive the parameters of the tail distribution of the received power by fitting the GPD to the independent and identically distributed (i.i.d.) samples extracted from two methods: Auto-Regressive Integrated Moving Average (ARIMA)-Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) and declustering method. Both ARIMA-GARCH and declustering methods remove the dependency among the observations and provide i.i.d. samples to the EVT input.

- We propose a novel methodology for determining the optimum threshold over which the tail statistics are derived, for the first time in the wireless communication literature. The existing studies in the URLLC regime of operation assume that the value of the threshold in EVT is known [Samarakoon et al., 2018], [Liu and Bennis, 2018], [Abdel-Aziz et al., 2020], [Mahmood and Jäntti, 2016], [Song and Li, 2006], [Al-Badarneh et al., 2018a], while the choice of optimum threshold is of paramount importance as it separates non-extreme observations from the extreme ones. We apply two complementary methods adopted from EVT to improve the accuracy of the threshold estimation: mean residual life method that is applied prior to the estimation of the tail distribution by using GPD, and parameter stability method that is applied after deriving the tail statistics by using GPD.

- We propose a novel methodology for determining the stopping conditions on the number of measured samples, sufficiently large to model the channel tail statistics in a consistent manner based on the adaption of the framework in [Cai and Hames, 2010], for the first time in the literature. The adaptation is made to the minimum number of samples required to estimate the distribution of values exceeding a given threshold, instead of the probability distribution of the maxima/minima in a time series in [Cai and Hames, 2010], and using GPD, rather than GEV distribution in [Cai and Hames, 2010].

- We propose a validation procedure for the accuracy assessment of the channel tail model derived by EVT using the probability plots, for the first time in the literature. We utilize the probability-probability (PP) plot and quantile-quantile (QQ) plot to assess the goodness-of-fit (GoF) of the Pareto model that characterizes the channel tail distribution.

- We demonstrate the superiority of the proposed methodology for deriving the tail statistics in terms of the modeling accuracy, compared to the conventional method based on the extrapolation of the average statistics to the ultra-reliable region, over the data collected within the engine compartment of Fiat Linea under various engine vibrations and driving scenarios.

The rest of the chapter is organized as follows. Section 2.3 describes the basics of EVT together with the theorems and corollaries used in the development of the proposed channel modeling approach, and the methods for removing the dependency in the observation samples and providing the i.i.d. input to the EVT. Section 2.4 presents the proposed channel modeling framework based on EVT for characterizing the channel tail distribution in the ultra-reliable region. Section 2.5 provides the channel measurement setup, and the performance evaluation of the proposed algorithm in determining the optimum threshold and minimum number of required samples for the derivation of the channel tail statistics, and compared to the conventional methods in terms of the estimation accuracy. Finally, concluding remarks and future works are given in Section 2.6.

2.3 Background

2.3.1 Extreme Value Theory

EVT is a powerful tool for characterizing the probabilistic distribution of extreme events occurring with low probability. EVT has been used for a long time in a wide variety of fields, such as hydrology to quantify risks of extreme floods, rainfalls and waves [Kotz and Nadarajah, 2000], or financial engineering to estimate the unexpected losses due to the extreme events [Finkenstadt and Rootzén, 2003]. However, it has only been recently employed in communication engineering, mainly in the analysis of the extreme values in network traffic, worst-case delay, queue lengths, throughput and BER/PER to address the strict delay and reliability requirements of URLLC [Samarakoon et al., 2018], [Bennis et al., 2018].

In general, one can divide EVT applications into two categories [Finkenstadt and Rootzén, 2003], [Galambos et al., 1994], [Kotz and Nadarajah, 2000], [Coles et al., 2001], [Bensalah et al., 2000], [Makarov, 2007], [Caires, 2011], [Mata, 2000]. In the first category, the asymptotic distribution of the maxima/minima of a long finite sequence of i.i.d. random variables is modeled by the generalized extreme value (GEV) distribution, while in the second category, the distribution of the values exceeding a given threshold is modeled by the GPD [Bensalah et al., 2000], [Makarov, 2007]. In URLLC, the main focus is on the second category of EVT applications, as all values exceeding a given threshold are considered as extreme events and need to be incorporated into the tail statistic analysis. In the following, we briefly introduce EVT and its major results used to develop the methodologies in the upcoming sections.

Theorem 1. *Let $X = [X_1, X_2, ..., X_n]$ be a sequence of independent random variables with a CDF denoted by F. Then, for low enough threshold u, the probabilistic distribution of the values exceeding u, i.e., $Pr\{u - X < y | X < u\}$, is approximated by the GPD with CDF given by*

$$H_u(y) = 1 - \left[1 + \frac{\xi y}{\tilde{\sigma}_u}\right]^{-1/\xi}. \tag{2.1}$$

defined on $\{y : y > 0\}$, where ξ and $\tilde{\sigma}_u = \sigma + \xi(u - \mu)$ are shape and scale parameters of the GPD, respectively [de Zea Bermudez and Kotz, 2010]. Here, μ and σ are the location and scale parameters of GEV distribution fitted to the CDF of $M_n = max\{X_1, ..., X_n\}$, respectively [Finkenstadt and Rootzén, 2003], [Coles et al., 2001].

Proof. For large sample size n, the CDF of M_n is given by

$$Pr\{M_n \leq z\} \approx F^n(z),$$

where

$$F^n(z) = \exp\Big\{ - \Big[1 + \xi\Big(\frac{z - \mu}{\sigma}\Big)\Big]^{-1/\xi}\Big\} \tag{2.2}$$

is in the form of GEV distribution, and μ, ξ, and σ are the location, shape, and scale parameters of the GEV distribution, respectively [Coles et al., 2001, Theorem 3.3]. By taking the logarithm of both sides of (2.2), and using the Taylor series expansion, we obtain

$$1 - F(z) = n^{-1}\Big[1 + \xi\Big(\frac{z - \mu}{\sigma}\Big)\Big]^{-1/\xi}. \tag{2.3}$$

The CDF of the excesses over threshold z is then given by

$$\begin{aligned} Pr\{X - z < y | X > z\} &= 1 - \frac{1 - F(z+y)}{1 - F(z)} \\ &= 1 - \left[1 + \frac{\xi y}{\sigma + \xi(z - \mu)}\right]^{-1/\xi}. \end{aligned} \tag{2.4}$$

Assuming that $\tilde{X} = -X$, the lower tail is transformed to the upper tail and then, GPD is fitted to $\{u - \tilde{X} | \tilde{X} < u\}$. With this transformation, the CDF of the excesses below threshold u is expressed as $1 - \left[1 + \frac{\xi y}{\tilde{\sigma}}\right]^{-1/\xi}$, where $\tilde{\sigma} = \sigma + \xi(u - \mu)$.

Theorem 1 implies that the distribution of the i.i.d. sequence of random variables exceeding a given threshold can be approximated by GPD. The parameters of the GPD are estimated by using the maximum likelihood estimator (MLE) [Finkenstadt and Rootzén, 2003], [Coles et al., 2001], [Van Zyl, 2015]. The consequences of Theorem 1 are given in the following corollaries. Corollaries 1.1 and 1.2 will be used in Section 2.4.3 to determine the optimum threshold for estimating the channel tail distribution by using GPD, whereas Corollary 1.3 will be utilized in Section 2.4.4 to evaluate the minimum number of samples, sufficiently large to estimate the channel tail distribution by GPD.

Corollary 1.1. *Let $X = [X_1, X_2, ..., X_n]$ be a sequence of i.i.d. random variables, where GPD(ξ,$\tilde{\sigma}_{u_0}$) models the probabilistic distribution of exceedances below threshold u_0. Then, for all thresholds $u < u_0$, the CDF of values below threshold u is modeled by GPD(ξ,$\tilde{\sigma}_u$), where $\tilde{\sigma}_u = \tilde{\sigma}_{u_0} + \xi(u - u_0)$, and the estimates of shape parameter (ξ) and modified scale parameter defined as $\sigma^* = \tilde{\sigma}_u - \xi u$ are constant against u. However, due to the sampling variability, the estimated ξ and σ^* cannot be exactly constant, but linear with respect to u.*

Proof. By Theorem 1, if a generalized Pareto distribution is a valid model for the excesses below threshold u_0, then the excesses of a threshold u, $u < u_0$, should also follow a generalized Pareto distribution. Accordingly, if exceedances below threshold u_0 have GPD(ξ,$\tilde{\sigma}_{u_0}$), with $\tilde{\sigma}_{u_0} = \sigma + \xi(u_0 - \mu)$, then, the exceedances below threshold $u < u_0$ are distributed by GPD(ξ,$\tilde{\sigma}_u$), where $\tilde{\sigma}_u = \sigma + \xi(u - \mu)$. Therefore, $\tilde{\sigma}_u = \tilde{\sigma}_{u_0} + \xi(u - u_0)$. Though the shape parameter is threshold invariant, the scale parameter changes with u unless $\xi = 0$. To make $\tilde{\sigma}_u$ constant with respect to u, we define the modified scale parameter as

$$\sigma^* = \tilde{\sigma}_u - \xi u, \tag{2.5}$$

where σ^* depends only on $\tilde{\sigma}_{u_0}$ and ξ, so, is u independent. It should be noted that the estimated ξ and σ^* cannot be exactly constant since different thresholds are associated with different number of samples included in the tail analysis. The MLEs of the estimated shape parameters change linearly with the number of samples included in the tail [Galambos et al., 1994], [Coles et al., 2001], therefore, the shape parameters should be linear with respect to $u < u_0$.

Corollary 1.2. *Let $X = [X_1, X_2, ..., X_n]$ be a sequence of i.i.d. random variables, where GPD(ξ,$\tilde{\sigma}_{u_0}$) models the probabilistic distribution of exceedances below threshold u_0. Then, for all thresholds $u < u_0$, the mean excess function (MEF), also known as the mean residual life function, defined by $e(u) = E(u - X|X < u)$, is a linear function of u.*

Proof. Let Y be a random variable with a probabilistic distribution GPD(ξ,σ). Then, the expectation of Y is given by $E(Y) = \sigma/(1 - \xi)$, where $E(.)$ denotes the expectation function. By Theorem 1, and Corollary 1.1, if GPD(ξ,$\tilde{\sigma}_{u_0}$) models the probabilistic distribution of exceedances below threshold u_0, then, for any threshold u, $u < u_0$, the CDF of values below threshold u is modeled by GPD(ξ,$\tilde{\sigma}_{u_0}+\xi(u-u_0)$)

with the mean excess function

$$e(u) = E(u - X|X < u) = \frac{\tilde{\sigma}_{u_0} + \xi(u - u_0)}{1 - \xi}, \tag{2.6}$$

which is a linear function of u.

According to Corollary 1.2, if the mean excesses are plotted against the threshold u, and the GPD assumption is satisfied at threshold u_0, then, for all thresholds $u < u_0$, the plot should be a straight line with slope and intercept equal to $\frac{\xi}{1-\xi}$ and $\frac{\tilde{\sigma}_{u_0} - \xi u_0}{1-\xi}$, respectively.

Corollary 1.3. *Let $X = [X_1, X_2, ..., X_n]$ be a sequence of i.i.d. random variables, where GPD(ξ,$\tilde{\sigma}_u$) models the probabilistic distribution of exceedances below threshold u. Then, the return level $r_m < u$, defined as the extreme quantile that is expected to be exceeded with probability $1/m$ on average, is given by*

$$r_m = u + \frac{\tilde{\sigma}_u}{\xi}[(m\zeta_u)^{\xi} - 1], \tag{2.7}$$

where $\zeta_u = Pr\{X < u\}$, and r_m is normally distributed if the sample size n is sufficiently large to estimate the tail distribution of values exceeding threshold u by using GPD(ξ,$\tilde{\sigma}_u$).

Proof. Assume $r = u - y$ in Eqn. (2.1). Then,

$$Pr\{X < r|X < u\} = \Big[1 + \frac{\xi(u - r)}{\tilde{\sigma}_u}\Big]^{-1/\xi}. \tag{2.8}$$

Since the return level, r_m is the extreme quantile that is expected to be exceeded on average once in m observations,

$$\frac{1}{m} = \zeta_u\Big[1 + \frac{\xi(u - r_m)}{\tilde{\sigma}_u}\Big]^{-1/\xi}, \tag{2.9}$$

where $\zeta_u = Pr\{X < u\}$ and m is large enough to ensure that $r_m < u$. The return level r_m can then be derived by rearranging the terms in Eqn. (2.9).

The behavior of the MLEs of the Pareto parameters, including the return level, depends on the value of the shape parameter ξ such that for $\xi < -1$, the MLEs are unlikely to be obtained; for $-1 < \xi < -0.5$, the MLEs can be obtained but are not associated with any standard asymptotic property; and if $\xi > -0.5$, the MLEs have the usual asymptomatic distribution. The case $\xi < -0.5$ happens when there are a few extreme events incorporated into the tail analysis, resulting in a distribution with a very short bounded tail. In order to have a shape parameter greater than -0.5, more extreme events should be captured by increasing the sample size. Therefore, if the sample size n is sufficiently large to estimate the tail distribution by using GPD accurately, not only the MLEs of the return levels have an asymptotic property, but also they are normally distributed [Cai and Hames, 2010], [Finkenstadt and Rootzén, 2003], [Coles et al., 2001], [Knapp, 2009], [Smith, 1985]. Otherwise, MLE is unstable and may provide unrealistic estimates for the model parameters.

2.3.2 Dependency Removal Methods

The input of EVT is necessarily a sequence of i.i.d. random variables. However, the threshold exceedances in the observation sequence are dependent samples inherently, as they occur in groups, i.e., one extremely low power value is likely to be followed by another. To obtain the i.i.d. samples from the sequence of dependent channel data, we either utilize declustering method introduced in [Coles et al., 2001], or extract the i.i.d. residuals from ARIMA-GARCH model [Nystrom and Skoglund, 2002a], [Nystrom and Skoglund, 2002b], [Glosten et al., 1993], [Ruiz et al., 2007]. In the declustering method, we model the tail distribution of the original observation sequence by splitting the observations into multiple clusters separated in time with at least a certain sample gap, each cluster consisting of a group of successive dependent observations. Then, upon extracting the minimum of each cluster, we apply EVT to the i.i.d. cluster minima. This method is based on the fact that the extreme events are close to independent at times that are far enough apart [Coles et al., 2001]. On the other hand, in the ARIMA-GARCH model, we apply EVT to the filtered residuals. The filtered residuals represent the randomness of the original sequence that cannot be determined by channel prediction using ARIMA-GARCH. However, in the case that the channel prediction is not the concern, we apply EVT to the original channel data by extracting the i.i.d. samples through the declustering approach.

Declustering Model:

Declustering is a widely adopted method based on the filtering of the dependent observations to obtain a set of threshold excesses that are approximately independent [Coles et al., 2001]. In declustering, we specify a threshold u and search for the first observation that falls below u to start the first cluster. The cluster is formed by assigning consecutive values below threshold u to the same cluster. The cluster is deemed to continue until we observe a sample that is above u. Upon obtaining an observation above u, we let the cluster to conditionally continue storing r more successive values. If any observation within these r samples is found to be below u, we continue with the same cluster. Otherwise, we terminate the cluster, and the next sample below u then initiates the next cluster. In the declustering method, we let the cluster to store r more values not exceeding the threshold u to overcome the effect of the random noise and deficiency of this method in dependency to the initial choice of threshold [Coles et al., 2001]. The optimum values of u and r need to be specified such that the minimum values of the clusters satisfy i.i.d. assumptions and meet the EVT requirements stated in Corollary 1.1, so that we can model the tail distribution of the received power by using the generalized Pareto distribution.

To obtain the optimum u and r values, we fit the generalized Pareto distribution to the cluster minima for different u and r values , and estimate the corresponding shape and modified scale parameters. Then, based on Corollary 1.1, we look for the largest u value below which the estimated parameters have a linear relation to the changes in u. The linearity relation is evaluated by using the statistical measure, named R squared, denoted by R^2. This measure is defined as the proportion of variations of one variable explained by the other variable in a linear regression model, taking values between zero and one [Gelman et al., 2019]. Moreover, there is a trade-off in the choice of r value: Too low r creates dependency among the cluster minima due to the consecutive clusters being too close to each other, whereas too high r results in the elimination of some valuable samples that could be extreme in a separated cluster. Therefore, we look for the minimum r value for which the extracted minima are i.i.d., and the estimated Pareto parameters associated to the higher r values do not differ significantly.

ARIMA-GARCH Model:

The hybrid ARIMA-GARCH model is used to predict the wireless channel by modeling the mean and variance dependency among the samples. First, ARIMA model is used to model the dependency of the variation in the conditional mean of the time series data, assuming that the variance is constant. Then GARCH model captures the variation in the conditional variance of the data sequence over time. The residuals of ARIMA-GARCH filtering then form an i.i.d. sequence, which represents the remaining randomness from the channel prediction, and is used in the analysis of the channel tail statistics by using EVT.

ARIMA(p, d, q) is commonly used to model the conditional mean of a sequence, where p and q stand for the number of autoregressive (AR) and moving average (MA) terms, respectively, and d is the degree of differencing [Ruiz et al., 2007]. If the time series data is not stationary, ARIMA parameters estimation starts by transforming the non-stationary process to a stationary one. If the non-stationarity is caused by the non-constant variance, power transformation is applied by taking the square root, cube root, or logarithm of the time series. If non-stationarity is due to the non-constant mean, differencing is applied by creating a new series whose value at time t is the difference between the samples at time t and time $t+d$, i.e., $x(t)-x(t+d)$ [Kwiatkowski et al., 1992]. The output of power transformation/differencing is a stationary process denoted by ARIMA$(p,0,q)$ and given by

$$r_t = c + \sum_{i=1}^{p} \theta_i r_{t-i} + \sum_{j=1}^{q} \beta_j \epsilon_{t-j} + \epsilon_t, \tag{2.10}$$

where r_t is the process mean at time t conditioned on the past realizations, c is a constant, θ_i is the i^{th} AR coefficient for $i \in \{1,...,p\}$, β_j is the j^{th} MA coefficient for $j \in \{1,...,q\}$, and ϵ_t is the residual at time t. MA and AR coefficients of ARIMA$(p,0,q)$ can be determined by analyzing the estimated autocorrelation function (ACF) and partial ACF (PACF), respectively [Demir and Coleri Ergen, 2016], [Guidolin and Pedio, 2018], [Zhou et al., 2005].

GARCH model characterizes the conditional variance of the predicted ARIMA-$(p,0,q)$ residuals, ϵ_t [Nystrom and Skoglund, 2002a], [Nystrom and Skoglund, 2002b], [Chen et al., 2011]. For most practical purposes, the simplest form of the GARCH model, GARCH$(1,1)$, is used, which is given by

$$\sigma_t^2 = k + \gamma\sigma_{t-1}^2 + \phi\epsilon_{t-1}^2 + \psi sgn(-\epsilon_{t-1})\epsilon_{t-1}^2, \tag{2.11}$$

where k is a constant, γ and ϕ are the coefficients of conditional variance and squared residual at time $t-1$, respectively, and ψ is the coefficient of squared residual at time $t-1$ incorporating the impact of the sign of the residual at time $t-1$. Note that $sgn(-x) = 1$ if $x < 0$; 0 if $x = 0$; and -1 if $x > 0$. The output of ARIMA-

GARCH filtering is z_t, which is a sequence of i.i.d. residuals with zero mean and the normalized variance one [Nystrom and Skoglund, 2002a], [Nystrom and Skoglund, 2002b], [Glosten et al., 1993], [Ruiz et al., 2007].

The ACF of the residuals, as well as the ACF of the squared residuals, determine the suitability of the fitted ARIMA and GARCH models, respectively. If in both aforementioned ACFs, the significant correlations at lag 0 are followed by the correlations approximately equal to zero, one may conclude that the parameters of the hybrid ARIMA-GARCH model are determined properly, and the extracted residuals are independent and identically distributed. Otherwise, the estimated ARIMA-GARCH parameters should be revised to remove the additional correlation in the ACFs of the residuals and squared residuals [Finkenstadt and Rootzén, 2003], [Ruiz et al., 2007].

2.4 Proposed Channel Modeling Framework

We propose a novel EVT-based channel modeling methodology with the goal of estimating the lower tail statistics of the communication channel in the ultra-reliable regime with very high accuracy and minimum amount of data. The methodology consists of the following steps: The sequence of measured received power samples is converted into a sequence of i.i.d samples by removing their dependency via either declustering or ARIMA-GARCH filtering. Upon applying EVT to the resulting sequence of i.i.d. samples, the parameters of the Pareto distribution associated with different thresholds are estimated by using the MLE.

The optimum threshold is then determined by applying mean residual life and parameter stability methods. Next, the stopping conditions are specified to determine the minimum number of samples required for the tail estimation by using GPD. If the number of the collected samples is below this minimum number, the algorithm continues by collecting more data and restarting the process. Otherwise, the valid-

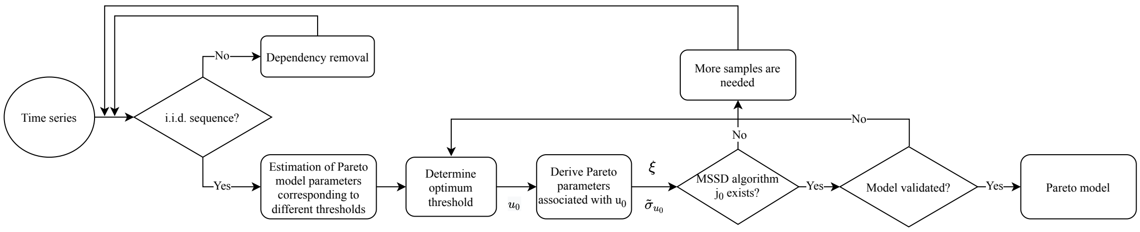


Figure 2.1: Flowchart of the proposed channel modeling framework.

ity of the model corresponding to the optimum threshold is assessed by using the probability plots. The proposed algorithm is depicted in Fig. 2.1 and explained in detail next.

2.4.1 Dependency Removal

The EVT input needs to be a sequence of i.i.d. random variables, thus, at the first step, the dependency behavior of the input process should be investigated. The proposed algorithm starts by calculating the autocorrelation of the input time series sequence. If the autocorrelation at all lags except lag 0 is effectively 0, i.e., below a predetermined small threshold value, we skip this step. Otherwise, the dependency removal methods given in Section 2.3.2 are employed: Declustering method is used in estimating the channel tail distribution of the original observed data, while ARIMA-GARCH filtering is used when the channel tail estimation of the residuals extracted from the channel prediction is of interest.

2.4.2 Estimation of Pareto Model Parameters

The parameters of the Pareto distribution, ξ and $\tilde{\sigma}_u$, are calculated by using MLE for different threshold values u. The threshold value inherently determines the percentage of the data considered as extreme events, and thus, included in the tail statistics. The lower and upper boundaries of the threshold values are chosen as the minimum and average of the i.i.d. input sequence, respectively.

2.4.3 Determination of Optimum Threshold

The suitability of the Pareto model fitted to the tail distribution is significantly affected by the threshold value: Too high a threshold leads to bias due to the immense portion of the data in the tail, while too low a threshold leads to high variance due to a few excesses with which the model is estimated [Caires, 2011]. To determine the optimum threshold, mean residual life and parameter stability methods are employed prior to and after the estimation of the Pareto model parameters, respectively. These two methods are used to validate and complement each other for consistency and robustness in case one of them is not able to determine the optimum threshold.

Mean Residual Life Method:

The mean residual life plot consists of the pairs of threshold and the expected value of the samples exceeding the corresponding threshold, as given by

$$\Big\{\big(u, E(u - X|X < u)\big) : u > X_{min}\Big\}, \tag{2.12}$$

where $X = [X_1, X_2, ..., X_n]$ is the i.i.d. input sequence and X_{min} is the smallest value in the sequence of X [Galambos et al., 1994], [Coles et al., 2001], [Mata, 2000]. Referring to Corollary 1.2, the best threshold, u_0, is the highest threshold below which, the mean excess function, i.e., $E(u-X|X < u)$, is a linear function of u, based on the R^2 measure: Larger value of R^2 indicates a better linear fit to the empirical data. Therefore, the mean residual life plot should be a straight line for thresholds $u < u_0$, with slope and intercept equal to $\frac{\xi}{1-\xi}$ and $\frac{\tilde{\sigma}_{u_0} - \xi u_0}{1-\xi}$, respectively. However, if the mean residual life plot behaves almost linearly for all possible thresholds, the decision of optimum threshold is determined by using a parallel approach entitled parameter stability method.

Parameter Stability Method:

Parameter stability is a complementary method to the mean residual life plot in the determination of the optimum threshold. In this method, first, the scale and shape parameters of the GPD fitted to the tail distributions corresponding to different thresholds are extracted. Then, based on Corollary 1.1, the optimum threshold, u_0, is the highest threshold below which the estimated scale and shape parameters are linear function of u, according to the R^2 measure.

2.4.4 Minimum Sample Size Determination (MSSD) Algorithm

MSSD is developed with the goal of determining the stopping conditions for the collection of enough samples used in modeling the tail statistics. Although the increase in sample size provides a more accurate estimation of the tail statistics, collecting a large number of samples might be expensive, time-consuming, and even impossible in many experiments [Cai and Hames, 2010]. The MSSD algorithm is a smart enumeration based algorithm, mainly inspired by the methodology given in [Cai and Hames, 2010], which has been proposed to determine the minimum sample size required to model the distribution of maxima/minima in a time sequence by using GEV distribution. We adapt this method for the determination of the minimum number of samples required to estimate the distribution of the values exceeding a given threshold in a time series by using GPD.

MSSD Algorithm is based on the normality of the return levels given in Corollary 1.3: If the sample size n is sufficiently large to estimate the tail distribution of values exceeding threshold u by using GPD(ξ, $\tilde{\sigma}_u$), then, the return level, r_m, is normally distributed. However, in most of the experiments, it is not practical to check the normality assumption of the return levels as there exists only one or a few time series. To tackle this problem, standard bootstrap method with replacement is applied to the observations with the goal of generating more sets for assessing the distribution of the return levels [Cai and Hames, 2010]. MSSD Algorithm consists of two main parts:

1. *Two-step bootstrapping process:* Bootstrapping is carried out in two steps to achieve a better bootstrapping performance in terms of the estimation accuracy, as stated in [Cai and Hames, 2010], [Knapp, 2009]. In the first step, the bootstrap is performed on the original observations to generate $M-1$ data sets, each with sample size n, same as the size of the original data set. In the second step, sample sets of different sizes are extracted from each of the M data sets in order to obtain the return levels associated with different sample sizes, performing $K-1$ bootstrapping on these data sets. Therefore, at the end of the two-step bootstrapping, $M \times K$ data sets with different lengths are generated.

2. *Normality assessment of the return levels:* Upon extracting the shape and scale parameters of the GPD fitted to the tail distribution of each sample set, we calculate the corresponding return levels for assessing their normality by using Anderson-Darling (AD) normality test. AD test is used instead of Shapiro-Wilk test in [Cai and Hames, 2010] since AD test is the most powerful empirical distribution function (EDF) test with higher focus on the tail of the distribution, compared to the other existing normality tests, and more appropriate for assessing the normality of the data with sample sizes greater than 50 [Razali et al., 2011]. After applying AD test, we extract the p-values, where p-value is defined as the index measuring the strength of the evidence against the null hypothesis H_0 that the data follows the normal distribution. Then, we compare these p-values with the critical value α, which is defined as the probability of rejecting the null hypothesis when the null hypothesis is true, i.e., probability of making a wrong decision, to (i) check whether

there are enough samples for modeling the tail distribution, (ii) determine the minimum number of samples required to estimate the tail distribution, and (iii) approximate the gain attained by increasing the sample size more than its minimum required value.

MSSD Algorithm, given in Algorithm 1, is described in detail as follows. The

---

Algorithm 1 **Minimum Sample Size Determination (MSSD) Algorithm**

Input: $\boldsymbol{X} = [X_1, X_2, ..., X_n]$, α, M, K, and n_0;
Output: j_0, G_j, $j \in \{j_0, ..., n\}$;

$\boldsymbol{x}_1 = (x_{11}, ..., x_{1n}) = \boldsymbol{X}$;
for m=1:M **do**
    if $m \neq 1$ **then**
        obtain bootstrap samples of $\boldsymbol{x}_1$, as $\boldsymbol{x}_m = (x_{m1}, ..., x_{mn})$;
    end if
    for $j = n_0 : n$ **do**
        $\boldsymbol{y}^1_{mj} = (x_{m1}, ..., x_{mj})$;
        for $k = 1 : K$ **do**
            if $k \neq 1$ **then**
                obtain bootstrap samples of $\boldsymbol{y}^1_{mj}$, as $\boldsymbol{y}^k_{mj}$;
            end if
            calculate r^k_{mj} using parameters of GPD fit for $\boldsymbol{y}^k_{mj}$;
        end for
        $\boldsymbol{d}_{mj} = (r^1_{mj}, ..., r^K_{mj})$;
        store p-value of AD normality test on $\boldsymbol{d}_{mj}$ in p_{mj};
    end for
end for
for $j = n_0 : n$ **do**
    calculate the mean of p-value, $\bar{p}_j$;
    calculate the standard deviation of p-value, s_j;
end for
if $\exists j_0 \in [n_0, n]$ **such that** $\bar{p}_j - t^* s_j > \alpha$ for all $j \in [j_0, n]$, **and** $\nexists i < j_0$ **such that** $\bar{p}_j - t^* s_j > \alpha$ for all $j \in [i, n]$ **then**
    for $j = (j_0 + 1) : n$ **do**
        $G_j = s_j - s_{j_0}$;
    end for
return j_0, $\{G_{j_0+1}, ..., G_n\}$;
else
return no feasible channel tail estimation exists;
end if

---

inputs of the algorithm are the observation sample set $\boldsymbol{X}$; the critical value α, which is used as a threshold for assessing the p-values of the AD normality test; the number of new generated data sets in the first and second steps of the bootstrapping process denoted by M and K, respectively; and the minimum sample size n_0 for which the estimated shape parameter ξ of the GPD fitted to the values exceeding the optimum threshold u_0 is greater than -0.5, based on Corollary 1.3. The algorithm starts by storing the observation samples in a new vector $\boldsymbol{x}_1$ for more convenient notation (Line 1). Then, the first step of bootstrapping with replacement is performed on the original data set to generate $M-1$ sample sets, each of size n, denoted by $\boldsymbol{x}_m$, $m \in \{2, ..., M\}$ (Lines $2-4$). Next, to perform the second step of bootstrapping, upon extracting the first j samples from each data set $\boldsymbol{x}_m$, and assigning them to a new vector denoted by $\boldsymbol{y}_{mj}^1$, for $j \in \{n_0, ..., n\}$ and $m \in \{1, ..., M\}$ (Lines $6-7$), $K-1$ bootstrapping with replacement is performed on $\boldsymbol{y}_{mj}^1$, storing each newly generated data set in the corresponding $\boldsymbol{y}_{mj}^k$, where $k \in \{2, ..., K\}$ (Line 10). Then, the GPD fit is applied on the values below the optimum predefined threshold of the data set $\boldsymbol{y}_{mj}^k$, and the corresponding scale and shape parameters are used in the calculation of the corresponding return levels r_{mj}^k from Eqn. (2.7) (Line 12). The return levels associated to the same sample size are placed into the vector $\boldsymbol{d}_{mj}$, for $j \in \{n_0, ..., n\}$ and $m \in \{1, ..., M\}$ (Line 14). Next, the normality assumption of the return levels is checked by applying the AD normality test on $\boldsymbol{d}_{mj}$, storing the p-values of the test in p_{mj} (Line 15). To enhance the performance of the bootstrapping process and decrease the test error in the estimation of the p-value, the average of M p-values corresponding to the same sample size j, denoted by $\bar{p}_j$, is calculated by using $\bar{p}_j = (1/M) \sum_{m=1}^{M} p_{mj}$, for $j \in \{n_0, ..., n\}$ (Line 19). However, the expected p-value is not exactly $\bar{p}_j$, and typically takes a value within the range of $[\bar{p}_j - t^* s_j, \bar{p}_j + t^* s_j]$, where t^* is the critical value obtained from a t-distribution with $M-1$ degrees of freedom, and s_j is the standard deviation of the p-value corresponding to the same sample size j, calculated by $s_j = \sqrt{(1/M)\sum_{m=1}^{M}(p_{mj} - \bar{p}_j)^2}$ (Line 20). Based on Corollary 1.3, if there exists a sample size j_0 such that the return level is normally distributed for all sample sizes greater than j_0, i.e., p-value, or more explicitly its lower bound denoted by $\bar{p}_j - t^* s_j$, is greater than α for $j > j_0$, and there exists no sample size i less than j_0, such that $\bar{p}_j - t^* s_j > \alpha$ for $j > i$, then, j_0 is the minimum plausible sample size with which we can estimate the tail distribution by using GPD (Line 22). Assuming a small quantity, such as 0.01 or 0.05, for the critical value α, we are $100(1-\alpha)\%$ confident that j_0 is the minimum number of samples required to obtain a properly normally distributed return level, and therefore, estimate the tail distribution by using GPD. Then, the algorithm determines the gain attained by increasing the sample size beyond the sufficient number j_0 (Lines $23-24$). The gain is defined as the difference between the standard deviations s_j and s_{j_0} corresponding to the sample sizes j and j_0, respectively, where $j \in \{j_0, ..., n\}$. In other words, the gain determines how much we can expand the range of values for the expected p-value, i.e., $[\bar{p}_j - t^* s_j, \bar{p}_j + t^* s_j]$, while still keeping the p-value above the critical value α, i.e., the distribution of the return levels normal. The algorithm terminates by returning j_0, if it exists, and the obtained gains for the sample sizes greater than j_0 (Line 26). Otherwise, the algorithm terminates stating that no feasible channel tail estimation exists since the existing samples are not sufficient to characterize the tail statistic (Lines $27-28$). Then, the sample size needs to increase by at least $0.1 \times n_0$ samples. The complexity of the MSSD algorithm is $O(nMK)$, since it bootstraps the sequence with sample size n for M iterations with complexity $O(nM)$, and then, by extracting a portion of samples from the newly generated sequences, MSSD algorithm bootstraps them for K iterations.

2.4.5 Model Validation

The model validity is assessed by using the probability plots in extreme value analysis. The probability plots are graphical techniques used to investigate whether the output of the proposed model fits well to the actual values. Two kinds of probability plots are used for validation: Probability/Probability (PP) plot and Quantile/Quantile (QQ) plot.

In the PP plot, the CDF of the received power is plotted versus the modeled CDF by GPD. **PP plot** includes the pairs

$$\{(i/k+1), H_u(y_i); i=1,...,k\}, \quad H_u(y_i) = 1 - \big(1 + \frac{\xi y_i}{\tilde{\sigma}_u}\big)^{-1/\xi}, \tag{2.13}$$

where y_i is the difference between the threshold u and X_i value exceeding the threshold such that $y_1 \leq ... \leq y_k$; k is the total number of excess values; H_u is the estimated Pareto model fitted to the CDF of the threshold excesses for threshold u with the associated shape and scale parameters ξ and $\tilde{\sigma}_u$, respectively [Coles et al., 2001]. For any one of the y_i, exactly i out of the k observations have a value less than or equal to y_i, so the empirical probability of an observation being less than or equal to y_i is $i/k+1$. Note that 1 in the denominator is a slight adjustment that is usually made to avoid reaching the CDF exactly equal to 1 [Coles et al., 2001].

In the QQ plot, empirical and modeled quantiles are plotted against each other. **QQ plot** consists of the pairs

$$\{X_i, (H_u^{-1}(i/(k+1)), i=1,...,k\}, \;\; H_u^{-1}(z_i) = u + \frac{\tilde{\sigma}_u}{\xi}\big[1 - z_i^{-\xi}\big], \tag{2.14}$$

where u denotes the threshold; ξ and $\tilde{\sigma}_u$ are the associated shape and scale parameters of the GPD fitted to the values below threshold u, respectively; z_i is the probability whose associated quantile is of interest; X_i is the i^{th} input sample exceeding the threshold such that $X_1 \leq ... \leq X_k$; and k is the number of values exceeding u. For any one of the X_i, exactly i out of the k observations have a value less than or equal to X_i.

If the generalized Pareto distribution appropriately estimates the extreme values exceeding a threshold u, then, nearly all data points in both probability plots lie on the unit diagonal, i.e., the 45° line [Finkenstadt and Rootzén, 2003], [Coles et al., 2001].

2.5 Performance Evaluation

The goal of this section is to evaluate the performance of the proposed channel modeling algorithm in determining the optimum threshold and optimum stopping conditions on the number of samples required to estimate the channel tail, and compare its performance with the traditional extrapolation-based methods in the estimation of the channel tail statistics. The traditional extrapolation-based approach first estimates the distribution of the existing channel data for the reliability order of 10^{-3}-10^{0} PER [Bas and Coleri Ergen, 2012], [Demir et al., 2013], and then, extrapolates the fitted distribution toward the ultra reliable region of 10^{-9}-10^{-6} PER [Eggers et al., 2019].

We have collected channel measurement data within the engine compartment of Fiat Linea under various engine and driving scenarios at 60 GHz by using a Vector Network Analyzer (VNA) (R & S® ZVA67). The transmitter and receiver are attached to the VNA ports through the R & S® ZV-Z196 port cables with 610 mm length and maximum 4.8 dB transmission loss. Both transmitter and receiver antennas are horn antennas operating between 50-75 GHz with a nominal 24 dBi gain and 12° vertical beamwidth. The antennas are connected to the waveguide operating at the frequency span of 50-65 GHz, with insertion loss 0.5 dB and impedance 50 Ω. The locations of the transmitter and receiver antennas are selected out of the possible locations for the wireless sensors located within the engine compartment, namely locations 5 and 13 in [Demir et al., 2013, Fig. 1], and shown in Fig. 2.2.

Figure 2.2: Transmitter and receiver antennas in the engine compartment.

The data were collected while the car is moving on the asphalt and stone roads at Koc University campus. The driving scenarios for the car include pushing the gas paddle while the car is static, moving up a ramp, and driving on the flat road. About 10^6 successive samples are captured for 30 minutes with time resolution of 2 ms. We use MATLAB® for the implementation of the proposed algorithm as well as the traditional extrapolation-based approach [Eggers et al., 2019].

In the following, first, we provide the numerical results in the determination of the optimum threshold over which the tail statistics are derived based on two approaches, ARIMA-GARCH and declustering, and then validate the tail model corresponding to the optimum threshold using probability plots in Section 2.5.1. Next, in Section 2.5.2, we present the performance evaluation of the proposed MSSD algorithm in determining the minimum number of samples required to estimate the channel tail statistics. Finally, we compare the performance of the proposed methodology to that of the traditional extrapolation-based technique in the estimation of the channel tail statistics in Section 2.5.3.

2.5.1 Optimum Threshold Determination

ARIMA-GARCH Approach:

ARIMA-GARCH filtering is used to predict the channel by its mean and variance conditioned on the past observations, and remove the dependency among the samples. Upon applying EVT on the i.i.d. residuals of ARIMA-GARCH, the best threshold is determined by using the mean residual life plot and the parameters stability methods. The probability plots are then shown to validate the GPD model fitted to the channel tail distribution. Please note that due to the stationarity of the process, the d parameter of ARIMA-GARCH is 0, and the best performance of ARIMA-GARCH at which the correlations of the residuals are almost 0 for any time lag greater than 0 is achieved when p and q are 2 [Demir and Coleri Ergen, 2016], [Guidolin and Pedio, 2018].

Fig. 2.3 shows the mean excess of the filtered residuals at different thresholds for ARIMA-GARCH filtering based on the mean residual life method. Mean excesses quantify the expected values of the residuals exceeding a given threshold, where the threshold varies from -3.85 dBm to -0.47 dBm. The lower and upper boundaries of the threshold are chosen such that, 0.3% and 30% of the extremely low-value residuals are in the tail, respectively. Since the mean excess increases linearly as threshold increases for all threshold values, i.e., $R^2 > 0.95$ for all thresholds, it is not possible to accurately distinguish the optimum threshold value based on Corollary 1.2, requiring the use of the parameter stability method.

Fig. 2.4 shows the shape and modified scale parameters of the fitted GPD model at different thresholds on the i.i.d. input sequence obtained by ARIMA-GARCH filtering for the parameter stability method. As opposed to Fig. 2.3, Fig. 2.4 can be used to determine the optimum value below which both shape and modified scale parameters change linearly with respect to the threshold. The optimum threshold, $u_0 = -1.6$ dBm, is in fact, the maximum threshold value for which the corresponding

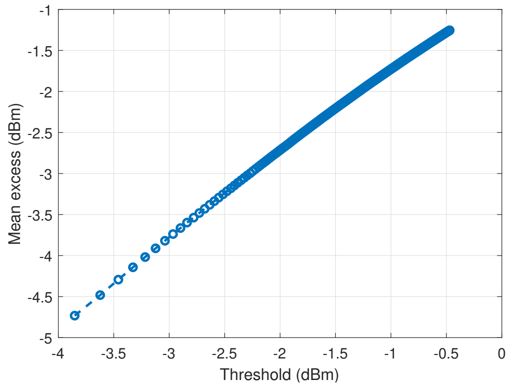


Figure 2.3: Mean excesses at different thresholds for ARIMA-GARCH filtering.

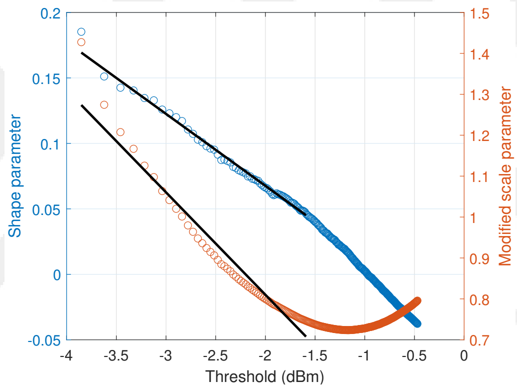
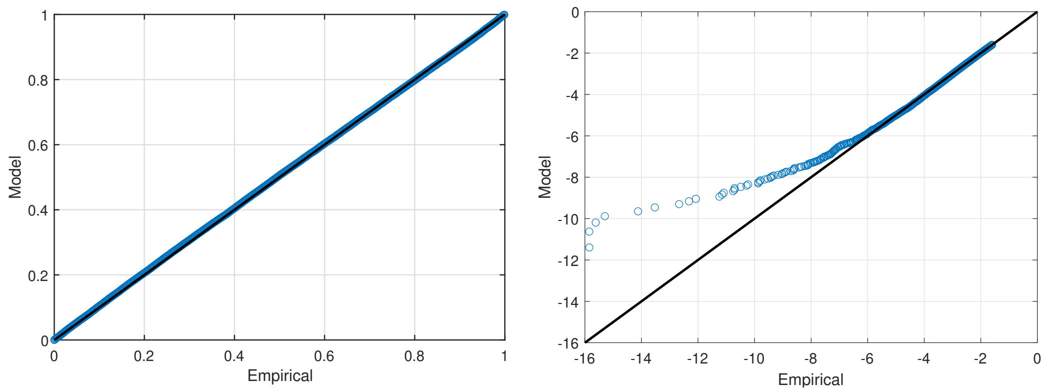


Figure 2.4: Shape and modified scale parameters of GPD at different thresholds for ARIMA-GARCH filtering. The black lines are the fitted linear regression models.

regression models fitted to the estimated parameters at $u < u_0$ results in R^2 value greater than 0.95. The black lines illustrate the linear regressions fitted to the parameters at the optimum threshold and the thresholds below that. σ^* and ξ corresponding to the optimum threshold -1.6 dBm are 0.746 and 0.046, respectively.

Fig. 2.5 shows the probability plots, including PP plot and QQ plot, for ARIMA-GARCH approach. This figure is used to validate the performance of the fitted Pareto model associated to the optimum threshold derived from Fig. 2.4. The black line is the diagonal line for diagnosing whether the values obtained by the Pareto model are in a good agreement with those of the empirical model. The black line has slope 1 and intersections 0 and -16 in the PP plot and QQ plot, respectively.



(a) (b)

Figure 2.5: Probability plots for ARIMA-GARCH filtering: (a) PP plot, and (b) QQ plot. The black lines are diagonal lines assessing the goodness of fit.

The linearity of the plots is reasonable, but not perfect, especially at the left side of the QQ plot. Fig. 2.5a shows that the CDF values obtained by the Pareto model are identical to the corresponding empirical values. However, Fig. 2.5b shows that the Pareto model provides better performance at high quantile values than the quantile values lower than -7 dBm, mainly due to the deviation of the i.i.d. residual distribution from the normal distribution. In spite of the divergence of the modeled from the empirical in Fig. 2.5b, since Fig. 2.5a illustrates a perfect agreement between the empirical and modeled values, one can still conclude that the fitted Pareto distribution models the empirical i.i.d. residuals obtained by ARIMA-GARCH filtering accurately [Finkenstadt and Rootzén, 2003].

Declustering Approach:

Declustering is used to remove the dependency among the samples and obtain i.i.d. observation for the EVT input. First, the optimum values for the threshold u and the minimum gap r between the samples are determined by using mean residual life plot and parameter stability method. Then, the probability plots are used to validate the accuracy of the GPD model fitted to the channel tail distribution.

Fig. 2.6 shows the mean excess of the observations at different thresholds and minimum gaps between the clusters for the declustering approach based on the mean residual life method. The expected values of the samples exceeding a given threshold are plotted against the threshold, where the threshold varies from -25 dBm to -10 dBm. When there is no minimum gap restriction between the clusters, i.e., $r = 0$, the mean excess increases linearly as threshold increases for all threshold values, similar to Fig. 2.3. However, by choosing $r > 0$, the mean residual life plot is linear in threshold u, for $u < -19$ dBm, where -19 dBm is the largest threshold value for which $R^2 > 0.95$ for all $r > 0$. Therefore, according to the mean residual life plot for the declustering method, the optimum threshold and the minimum r values are -19 dBm and 1, respectively.

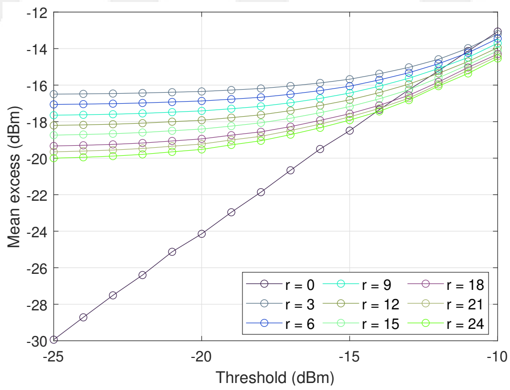


Figure 2.6: Mean excesses at different thresholds and minimum gaps between the clusters for declustering approach.

Fig. 2.7 shows the shape and modified scale parameters of the fitted GPD model on the i.i.d. sequence obtained by the declustering approach at different u and r values for the parameter stability method. By choosing r value greater than 15, the estimated Pareto parameters are linear in u for $u < -19$ dBm. Although in Fig. 2.6, it is possible to attain linearity assumption of the mean excesses for $u < -19$ dBm with r values less than 15, the linearity is observed only for $r > 15$ at $u < -19$ dBm in Fig. 2.7. Therefore, r value should be at least 15. Additionally, -19 dBm is the minimum threshold for which the R^2 value of the regression model fitted to the estimated parameters is greater than 0.95. As a result, considering Figs. 2.6 and 2.7 together, the optimum values of u and r are 16 and -19 dBm, respectively. Also, for these optimum values, the corresponding shape and modified scale parameters are 0.108 and 2.06, respectively.

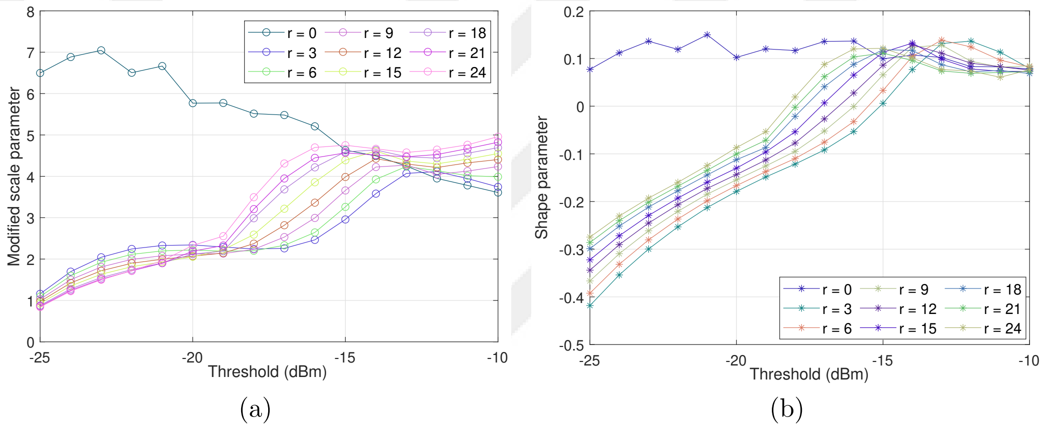


Figure 2.7: Shape and modified scale parameters of GPD at different thresholds and minimum gaps between the clusters for declustering method: (a) Modified scale parameter, and (b) Shape parameter.

Fig. 2.8 shows the probability plots for the declustering approach. This figure is used to validate the performance of the Pareto model corresponding to the threshold value evoked from Figs. 2.6 and 2.7. The black line is the diagonal line for diagnosing the goodness of fit of the GPD to the empirical values. The black lines have both slope 1, but intersections 0 and -55 for the PP plot and QQ plot, respectively. Both PP and QQ plots reveal the robust performance of the generalized Pareto model applied on the output of the declustering approach. Similar to Fig. 2.5a, Fig. 2.8a shows that the values modeled by GPD for the declustering approach are almost identical to that of the empirical values. However, as opposed to Fig. 2.5b, Fig. 2.8b shows that the QQ plot of the declustering approach fits to the diagonal line for all the quantile values. Therefore, the Pareto model obtained by the declustering approach models the empirical values with a much better performance than that of the ARIMA-GARCH approach.

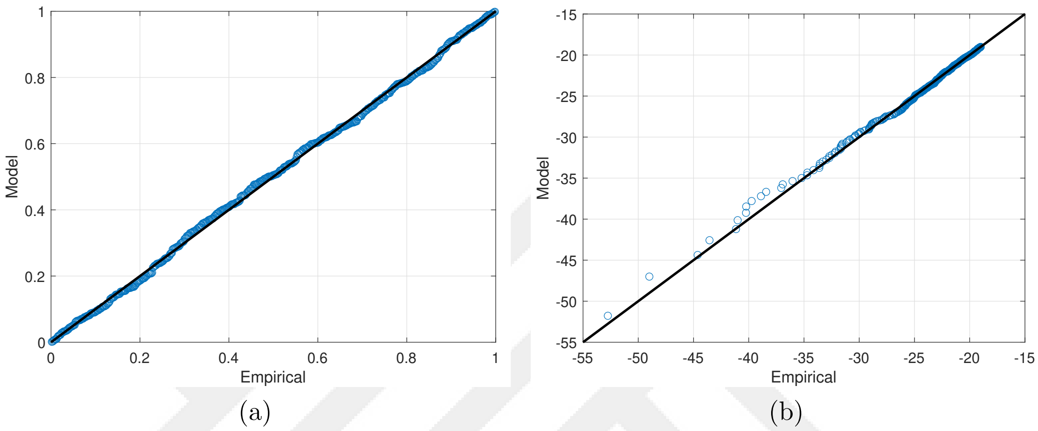


Figure 2.8: Probability plots for declustering method: (a) PP plot, and (b) QQ plot. The black lines are diagonal lines assessing the goodness of fit.

2.5.2 Minimum Required Number of Samples for Channel Tail Estimation

Fig. 2.9 shows the p-value of the AD normality test, as explained in detail in Section 2.4.4, for different number of samples. This figure is used to determine the minimum number of samples required to model the channel tail statistics. The horizontal line is at the critical value $\alpha = 0.05$, and the vertical line is at sample size equal to 2×10^5. For the sample size greater than 2×10^5, all the p-values are above the horizontal line, indicating that the sufficient number of samples required to estimate the tail statistics is 2×10^5. The gain we attain by increasing the sample size from 2×10^5 to 9×10^5, is only 0.08 in terms of the standard deviation of the p-values. As a result, although it is required to capture 10^6 samples or more for modeling the events occurring once in million for URLLC, MSSD allows the estimation of the tail statistics with lower number of samples.

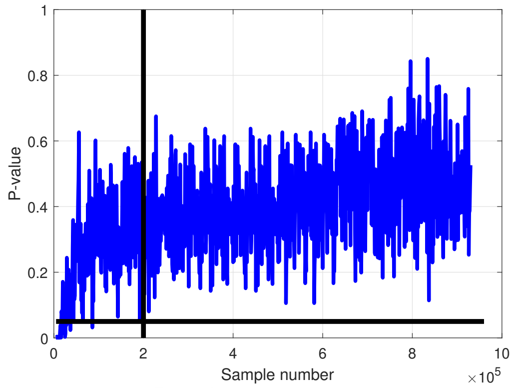


Figure 2.9: p-value of AD test for different sample numbers. The vertical and horizontal black lines indicate the minimum required number of samples and the critical value α, respectively.

2.5.3 Comparison with Conventional Tail Estimation Method

Fig. 2.10 shows the CDF of the normalized power for the empirical model, conventional extrapolation-based method, and the proposed Pareto model. To apply our proposed methodology, we use Pareto distribution to estimate the upper and lower tails of the channel, denoted by *Pareto Lower Tail* and *Pareto Upper Tail*, respectively, while the kernel-smoothing function in MATLAB® is used to estimate the

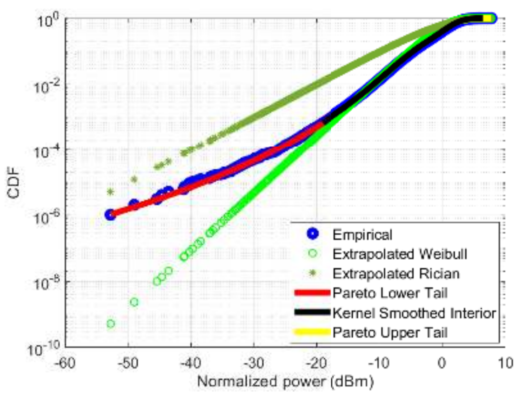


Figure 2.10: CDF of the normalized power for empirical, extrapolation-based, and Pareto model.

CDF of the interior values between the lower and upper tail probabilities, denoted by *Kernel Smoothed Interior* in the figure. The optimum threshold for the lower tail has been determined in Section 2.5.1 as -19 dBm by considering 10^6 samples. It is worth mentioning that, for the sake of simplicity and without loss of generality, we purposefully ignore threshold modeling for the upper tail as it is not the aim of this study, thus, the same sample portion in the upper tail is used as that of the lower tail. To obtain the extrapolated Weibull curve, we extract the observations corresponding to the reliability order of 10^{-3}-10^{0} PER, i.e. the observations whose probability of occurrences are between $1/10^3$ and $1/10^0$. Upon applying different distributions to these samples, we select the best-fitting distribution as the Weibull distribution according to the Akaike information criterion/Bayesian information criterion (AIC/BIC) metric, and then compute the extrapolated tail distribution. In addition, we extract first 10^3 observations regardless of their probability of occurrences, and then apply different distributions to obtain the best fitted one. In contrast to the aforementioned Weibull case, this time, the Rician distribution is found to be the best fitting distribution to the first 10^3 observations according to the AIC/BIC metric. Then, to obtain the extrapolated Rician curve in Fig. 2.10, we compute the corresponding extrapolated tail. This figure reveals that the traditional extrapolation-based approach strongly depends on the portion of the data used for the modeling, resulting in the over-estimation or under-estimation of the empirical results. The proposed method has been demonstrated to perform significantly better than the conventional extrapolation-based approach, especially in the ultra-reliability region. The proposed model outperforms the extrapolated Rician and Weibull models by 0.01 and 1.94×10^{-4} in terms of RMSE, respectively, which is a significant improvement for the ultra-reliability region, where the reliability orders are in the range of 10^{-9}-10^{-5}.

2.6 Conclusions

In this chapter, we introduced a novel framework based on the extreme value theory with the goal of estimating the channel tail statistics for URLLC. The proposed methodology adopts EVT to determine the optimum threshold over which the tail statistics are derived, the stopping condition to specify the minimum number of samples required to model the tail characteristics, and the assessment to validate the final model by using probability plots. The proposed algorithm for determining the optimal stopping condition provides a significantly lower number of required samples for estimating the channel tail distribution. Moreover, the proposed channel modeling methodology achieves a significantly better fit to the empirical data in the lower tail than the conventional extrapolation-based approach, especially in the ultra-reliability region. Furthermore, the conventional approach has been demonstrated to be highly dependent on the portion of the data considered in the extrapolation, which can seriously affect the reliability performance.

Chapter 3

NON-STATIONARY WIRELESS CHANNEL MODELING APPROACH BASED ON EXTREME VALUE THEORY FOR ULTRA-RELIABLE COMMUNICATIONS

The content of this chapter has been published in IEEE Transactions of Vehicular Technology (TVT)-Correspondence.

3.1 Overview

A proper channel modeling methodology that characterizes the statistics of extreme events is key in the design of a system at an ultra-reliable regime of operation. The strict constraint of ultra-reliability corresponds to the packet error rate (PER) in the range of 10^{-9}-10^{-5} within the acceptable latency on the order of milliseconds. Extreme value theory (EVT) is a robust framework for modeling the statistical behavior of extreme events in the channel data. In this chapter, we propose a methodology based on EVT to model the extreme events of a non-stationary wireless channel for the ultra-reliable regime of operation. This methodology includes techniques for splitting the channel data sequence into multiple groups concerning the environmental factors causing non-stationarity, and fitting the lower tail distribution of the received power in each group to the generalized Pareto distribution (GPD). The proposed approach also consists of optimally determining the time-varying threshold over which the tail statistics are derived as a function of time, and assessing the validity of the derived Pareto model. Finally, the proposed approach chooses the best model with minimum complexity that represents the time variation behavior of the non-stationary channel data sequence. Based on the data collected within the engine compartment of Fiat Linea under various engine vibrations and driving scenarios, we demonstrate the capability of the proposed methodology in providing the best fit to the extremes of the non-stationary data. The proposed approach significantly outperforms the channel modeling approach using the stationary channel assumption in characterizing the extreme events.

3.2 Introduction

Ultra-reliable and low latency communication (URLLC) is one of the key features of beyond fifth generation (5G) networks, with the potential to support a vast set of mission-critical applications in vehicular communications, remote surgeries, and virtual reality [Popovski et al., 2018], [Eggers et al., 2019]. Under the constraints of a URLLC service, the reliability is defined as achieving the packet error rate (PER) in the range of 10^{-9}-10^{-5} with the latency on the order of milliseconds. An accurate channel modeling quantifying the tail statistics produced by extreme events is needed to satisfy the requirements of the URLLC systems.

Existing studies on channel modeling for URLLC either provide an extrapolation-based framework that extends the applicability of the available practical channel models in the ultra-reliability relevant regime [Eggers et al., 2019], or propose performance measures for the ultra-reliability of the channel [Swamy et al., 2018], [Swamy et al., 2019], [Angjelichinoski et al., 2019]. However, none of these frameworks propose a channel modeling methodology to derive and verify ultra-reliability statistics. In [Mehrnia and Coleri, ions], we propose a novel channel modeling methodology for URLLC based on Extreme value theory (EVT) to statistically derive the lower tail of the received power by characterizing the probabilistic distribution of extreme events happening rarely. However, the channel was assumed stationary, whereas, in reality, the wireless channel statistics vary over time due to the dynamic environment, requiring the usage of techniques for the estimation of time-varying parameters of the non-stationary channel data sequence.

EVT is a powerful framework characterizing the probabilistic distribution of infrequent extreme events or equivalently the tail distribution. EVT has been recently utilized at the data link and network layers to model the tail statistics of queue length and delay [Samarakoon et al., 2018], [Liu and Bennis, 2018] or derive closed-form asymptotic expressions for the throughput, bit error rate (BER), and PER over different fading channels [Mahmood and Jäntti, 2016], [Song and Li, 2006], [Al-Badarneh et al., 2018a]. However, these data link and network layer studies use the extrapolation of existing average statistics-based channel models in the ultra-reliable region, which have not yet been proven to be accurate experimentally. EVT has also been utilized at physical layer for wireless channel modeling to provide a fit to the whole distribution of large/small-scale fading [Molina-Garcia et al., 2008], [Dahman et al., 2010], [Reyes-Guerrero and Mariscal, 2014], [Sawada et al., 2015], [Li et al., 2018]. Nevertheless, EVT has never been incorporated into wireless channel modeling to estimate the tail statistics and address the constraint in the ultra-reliability region, which requires the development of methodologies for the determination of the optimum threshold below/above which the samples correspond to the extreme events, or to acquire a large number of samples to capture the extreme events occurring rarely. [Mehrnia and Coleri, ions] is the only existing study focusing on modeling the statistics of the channel model for URLLC by using the concept of EVT, however, stationary channel assumption may not be realistic for channels with time-varying characteristics.

The goal of this chapter is to propose a novel channel modeling methodology based on EVT for the statistical characterization of the lower tail distribution of the samples in a non-stationary sequence of the channel data. We use received signal power at constant transmit power as channel data, since the received signal power is equivalent to using the squared amplitude of the channel state information [Eggers et al., 2019]. This chapter extends the EVT based wireless channel modeling methodology in 2 for non-stationary channel data by determining the measurable external factors that cause non-stationarity and modeling the parameters of fitting distributions as a change-point function of time. First, the wireless channel data sequence is split into smaller groups, each of which corresponds to the same value of the external factor causing non-stationarity. Then, the generalized Pareto distribution is fitted to the lower tail of the received power distribution in each group. The original contributions of this chapter are listed as follows:

- We provide a novel methodology for determining the measurable external factors that cause non-stationarity in the channel data sequence, and then, splitting the data into smaller groups, each of which contains the samples collected under the same factor.

- We provide a novel approach for modeling the extremes of each group exceeding a given threshold by fitting the generalized Pareto distribution (GPD) to the distribution of the channel lower tail, while the threshold is determined optimally. Thereupon, we model the shape and scale parameters of GPD as a change-point function of time.

- We demonstrate the superiority of the proposed channel modeling methodology for deriving the tail statistics of a non-stationary channel sequence in terms of the deviance statistic, compared to the case in which the channel is assumed stationary.

The rest of this chapter is organized as follows: Section 3.3 describes the basics of EVT for stationary and non-stationary sequences. Section 3.4 presents the channel modeling methodology for characterizing the extremes of a non-stationary sequence.

Section 3.5 provides the performance evaluation of the proposed algorithm on the data collected within the engine compartment of Fiat Linea under various engine vibrations and driving scenarios. Finally, Section 3.6 concludes the chapter.

3.3 Background

3.3.1 Extreme Value Theory for Stationary Sequences

EVT provides a robust framework for analyzing the statistics of extreme events happening rarely through modeling the probabilistic distribution of the values exceeding a given threshold by using the GPD. According to the EVT, if $\{x_1, ..., x_N\}$ is an independent and identically distributed stationary sequence, where x_i denotes the i^{th} received power for $i \in \{1, ..., N\}$, then, the tail distribution of the power sequence, i.e., the probabilistic distribution of the power values exceeding a given threshold u, can be expressed as

$$G_u(y) = 1 - \left[1 + \frac{\xi y}{\tilde{\sigma}_u}\right]^{-1/\xi},$$

where y is a non-negative value denoting the exceedance below threshold u, i.e., $(y = u - X)$, where X denotes any x_i below threshold u; $G_u(y)$ is in the form of the GPD; and ξ and $\tilde{\sigma}_u$ are shape and scale parameters of the GPD, respectively [Mehrnia and Coleri, ions, Theorem 1], [de Zea Bermudez and Kotz, 2010].

3.3.2 Extreme Value Theory for Non-Stationary Sequences

In a non-stationary sequence, the threshold, shape, and scale parameters of the GPD model are time-varying. Therefore, the distribution of the values exceeding a given threshold is modeled by using a general Pareto model with time-varying parameters given by

$$\left(u_{(t)} - X_t \middle| X_t < u_{(t)}\right) \sim GPD\left(\tilde{\sigma}_{(u,t)}, \xi_{(t)}\right), \tag{3.1}$$

where $u_{(t)}$, $\tilde{\sigma}_{(u,t)}$, and $\xi_{(t)}$ are the time-varying threshold, scale and shape parameters of the generalized Pareto distribution, respectively.

In general, the non-stationarity is caused by a time-varying factor, which can be considered constant over short enough intervals of time. This allows us to break the non-stationary sequence into smaller stationary ones, and then, fit GPD to the extremes of each group and model its time-varying parameters by using the change-point expression as

$$\theta_{(t)} = \begin{cases} \theta_1, & t \leq t_1 \\ \theta_2, & t_1 < t \leq t_2 \\ . & . \\ \theta_M, & t_{M-1} < t \end{cases} \tag{3.2}$$

where $\theta_{(t)}$ is the time varying $u_{(t)}$, $\tilde{\sigma}_{(u,t)}$, or $\xi_{(t)}$; and θ_m is the determined constant u, $\tilde{\sigma}_{(u)}$, or ξ for group $m \in \{1, ..., M\}$. For different values of the factor causing non-stationarity, different threshold, scale and shape parameters of the GPD are derived.

3.4 Methodology

The objective of the proposed methodology is to model a non-stationary channel at an ultra-reliable regime of operation based on EVT. The main features of the suggested algorithm are as follows:

1. We perform the Augmented Dickey-Fuller (ADF) test on the data to check if

it is non-stationary. The test is based on the fact that the mean or variance of a stationary time series does not change over time [Fuller, 2009].

2. If the ADF test logical result is 0, we do not have enough evidence to reject the null hypothesis of stationary assumption. Otherwise, we reject the null hypothesis and make the decision of observing non-stationarity in the time-series.

3. If the channel data sequence is not stationary, we determine the external factors under which the parameters of the GPD are changing and categorize the sequence into M different groups accordingly.

4. We model the tail distribution of the received power in each group by using GPD.

5. We select the simplest GPD model with enough accuracy by using the deviance statistic D.

The proposed approach starts by applying ADF test on the channel data to check the stationarity. If the test logical result is 1, the channel data sequence is non-stationary and then, the samples are categorized into M groups based on the factors that cause non-stationarity. Otherwise, the channel data is stationary and $M = 1$. Then, since the input of EVT is necessarily a sequence of i.i.d. samples, we remove the time-dependency between the samples of each group. Afterward, we apply EVT and obtain the optimum threshold under which the received powers are considered extremely low values. Hereupon, we fit GPD with the associated shape and scale parameters to the values bellow the threshold to model the tail of the receive power channel data. Finally, we obtain the maximum log-likelihoods of the GPD models under different factors to select the model with the lowest complexity, i.e., minimum number of different thresholds, scale and shape parameters, based on the deviance statistic. The proposed algorithm is depicted in Fig. 3.1 and explained in detail next.

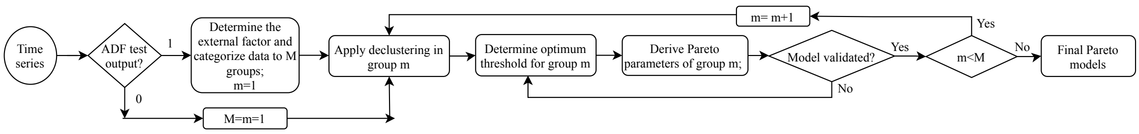


Figure 3.1: Flowchart of the proposed channel modeling framework for a non-stationary process.

3.4.1 Modeling of Tail Distribution by Using GPD

To model the tail distribution of the received power in each group, first, the sequence of measured received power samples is converted into a sequence of i.i.d samples by removing their dependency. The threshold exceedances in the observation sequence are inherently time dependent since one extremely low power is likely to be followed by another. On the other hand, since the input of EVT needs to be a sequence of i.i.d random variables, we use declustering approach to remove time dependencies [Mehrnia and Coleri, 2021c], [Coles et al., 2001]. Then, EVT is applied to the sequence of i.i.d. samples for estimating the parameters of the Pareto distribution corresponding to the different thresholds by using the maximum likelihood estimator (MLE). The optimum threshold is determined by applying two complementary methods, mean residual life (MRL) and parameter stability methods, in EVT. Next, the validity of the Pareto model corresponding to the optimum threshold is assessed by using probability plots.

The time-dependency of the observed samples is removed via the declustering approach, where the samples in each individual group are divided into multiple clusters such that each cluster includes consecutive dependent observations and the clusters are separated by a certain sample gap to ensure the independency between clusters [Mehrnia and Coleri, 2021c]. In this method, first, we assume a threshold

u and look for the first sample x_i below this threshold to initiate the first cluster. Upon observing the first $x_i < u$, x_i and all its consecutive samples below u will be assigned to the first cluster. Once a sample over u is detected, we let the cluster to continue for r more successive values and then, terminate the cluster, if no value below u is observed. The next cluster starts with the next value below the threshold u. In the second step of the declustering method, we extract the minimum value of each cluster, apply EVT to the cluster minima, and model their tail distribution by using GPD.

Optimum threshold determination is performed by using MRL and parameter stability methods. MRL method states that if we determine the mean value of samples exceeding a given threshold u and then, plot mean excesses, i.e., $E(u - X|X < u)$, against the threshold, the optimum threshold is the highest threshold below which, mean excess is a linear function of u. Though the MRL method is applied to the data sequence prior to the estimation of the Pareto model parameters with less complexity than the parameter stability method, it is sometimes difficult to obtain the optimum threshold explicitly, requiring the usage of the complementary parameter stability method. Parameter stability method states that if we fit GPD to the values exceeding a given threshold for a variety of the thresholds and extract the corresponding Pareto parameters and plot parameters against the threshold u, the optimum threshold is the highest threshold below which the estimated shape and modified scale parameters are linear with respect to u, where modified scale parameter is defined as $\sigma^* = \tilde{\sigma}_u - \xi u$.

The validity of the GPD model is assessed by using probability plots. Probability plots, consisting of Probability/Probability (PP) plot and Quantile/Quantile (QQ) plot, are graphical techniques used to assess the validity of the models fitted to the empirical values. In the PP plot, we compare the empirical probability of occurrence for an extreme value with the corresponding probability obtained by the GPD, while in the QQ plot, we compare the empirical extreme quantile with the corresponding probability obtained by the inverse of GPD. If the generalized Pareto distribution appropriately models the extreme values exceeding threshold u, then, both PP and QQ plots fit the unit diagonal line, i.e., the 45° line [Coles et al., 2001], [Finkenstadt and Rootzén, 2003]. PP plot consists of the pairs

$$\{(\frac{i}{k+1}), G_u(y_i)\}; \quad G_u(y_i) = 1 - \left(1 + \frac{\xi y_i}{\tilde{\sigma}_u}\right)^{-1/\xi}, \tag{3.3}$$

where y_i, $i \in \{1, ..., k\}$, is the absolute value of the difference between the threshold u and x_i sample exceeding the threshold; k is the number of values exceeding threshold u; $G_u(.)$ is the Pareto model fitted to the tail distribution; and ξ and $\tilde{\sigma}_u$ are the associated shape and scale parameters, respectively [Coles et al., 2001]. On the other hand, QQ plot consists of the pairs

$$\{x_i, (G_u^{-1}(\frac{i}{k+1})\}; \quad G_u^{-1}(z_i) = u + \frac{\tilde{\sigma}_u}{\xi}\left[1 - z_i^{-\xi}\right], \tag{3.4}$$

where x_i, $i \in \{1, ..., k\}$, is the i^{th} extreme quantile exceeding threshold u; $G_u^{-1}(.)$ is the inverse of the Pareto model fitted to the tail distribution; z_i is the probability of occurrence associated with the extreme quantile x_i; and k, ξ and $\tilde{\sigma}_u$ are same as those defined in (3.3).

3.4.2 Modeling Time-Varying Parameters of GPD

To model the extreme values in a non-stationary sequence, upon determining the external factors that affect Pareto parameters, we determine an optimum threshold and the corresponding Pareto model parameters for each group. The time-varying parameters are modeled according to the change-point expression given in Eqn. (3.2).

3.4.3 Final Model Selection

With the possibility of modeling any combination of the GPD parameters as functions of time, there is a catalog of models to choose from, and selecting an appropriate model becomes an important issue. The basic principle is parsimony, obtaining the simplest model that explains as much of the variation in the data as possible. In this regard, we obtain the maximum log-likelihood of the GPD associated with each external factor in Section 3.4.2 as

$$l_{GPD} = -\Big[\sum_{t=1}^{N} \log \tilde{\sigma}_{(u,t)} + \big(1 + \frac{1}{\xi_{(t)}}\big) \log \big(1 + \frac{\xi_{(t)} y_{(t)}}{\tilde{\sigma}_{(u,t)}}\Big], \tag{3.5}$$

where N is the maximum number of the samples; $\tilde{\sigma}_{(u,t)}$ and $\xi_{(t)}$ are the time varying scale and shape parameters at time t, respectively; and $y_{(t)} = u_{(t)} - x_{(t)}$ is the absolute value of the difference between the threshold $u_{(t)}$ and the sample $x_{(t)}$ exceeding $u_{(t)}$ at time t. Finally, we apply the deviance statistic test D,

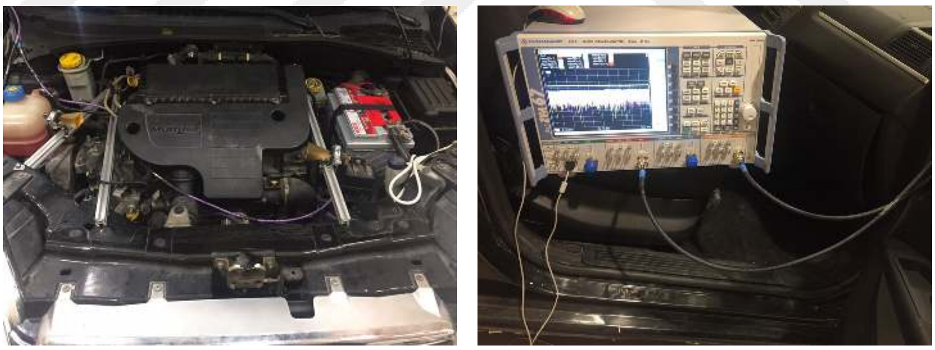
$$D = 2\{l_{GPD_1} - l_{GPD_0}\}, \tag{3.6}$$

where l_{GPD_1} and l_{GPD_0} are the maximum log-likelihoods of GPD_1 and GPD_0 models, respectively. The primarily model GPD_0 is rejected by the test at the α-level of significance if $D > c_\alpha$, where c_α is the $(1-\alpha)$ quantile of the χ_k^2 distribution for a small α value between 0.01 and 0.1; and k in χ_k^2 is the dimensionality difference, i.e., difference between the number of scale and shape parameters in each model.

3.5 Performance Evaluation

The goal of this section is to evaluate the performance of the proposed methodology in modeling the non-stationary channel for URLLC, and choosing the best GPD fitted to the non-stationary sequence according to the deviance statistic test.

The measured channel data were collected within the engine compartment of Fiat Linea at 60 GHz under different driving scenarios and road conditions, including the static car, driving on a smooth road, and ramp road. The antennas within the engine compartment are located such that the effect of the engine vibration is observed in the received power, as shown in Fig. 3.2a. A Vector Network Analyzer (VNA) (R & S® ZVA67) depicted in Fig. 3.2b is connected to the transmitter and receiver via the R & S® ZV-Z196 port cables with maximum 4.8 dB transmission loss. The horn transmitter and receiver antennas with a nominal 24 dBi gain and 12° vertical beam-width operate at 50-75 GHz. We have captured about 10^6 successive samples for 30 minutes with a time resolution of 2 ms. We use MATLAB® for the implementation of the proposed algorithm.



(a) (b)

Figure 3.2: Measurement setup (a) Transmitter (Tx) and receiver (Rx) antennas in the engine compartment of Fiat Linea, (b) VNA connection to the antennas.

Next, we first specify the parameters affecting the non-stationarity and then, categorize the sequence into different groups, accordingly, in Section 3.5.1. In Section 3.5.2, we determine the optimum threshold in each group and obtain the corresponding GPD shape and scale parameters by using the change-point expression, as provided in Sections 3.4.1 and 3.4.2. Upon obtaining the maximum log-likelihood of

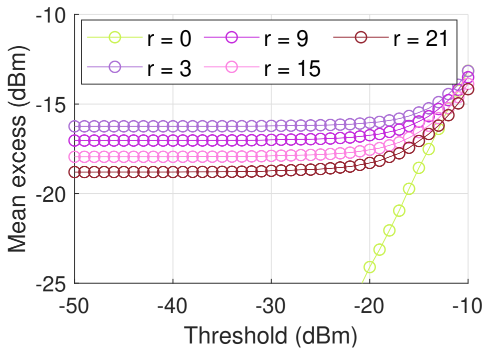
GPD from (3.5) under both non-stationary and stationary channel assumptions, we choose the best model according to the deviance statistic test D in Section 3.5.3.

3.5.1 External Factors Causing Non-Stationarity

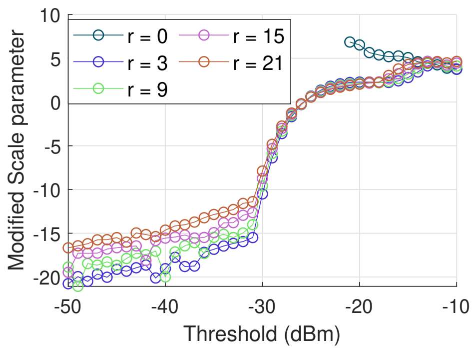
We have identified the dependence of the parameters that cause non-stationarity on both the driving scenario and the quality of the road assessed by the driver of the vehicle. Therefore, for the collected channel data, we define the parameter taking three discrete values corresponding to the following three groups: static car, smooth road, and ramp road.

3.5.2 Modeling Time-Varying Parameters of GPD Within the Engine Compartment

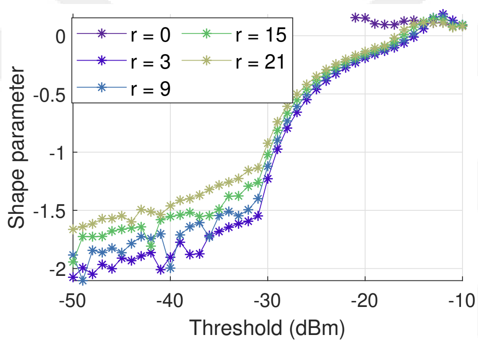
Fig. 3.3 shows the mean excess, shape and modified scale parameters of the GPD fitted to the filtered i.i.d. received power samples at different thresholds and minimum gaps between the clusters, as well as the probability plots for the group of ramp road. Fig. 3.3a illustrates the MRL plot, i.e., the mean of samples exceeding a given threshold, where the threshold varies between -50 dBm and -10 dBm. When the minimum gap between the clusters r is 0, the MRL plot linearly increases as threshold increases at all threshold values, thus, the optimum threshold cannot be recognized. However, by choosing $r > 0$, the MRL changes linearly in u only for $u < -20$ dBm, denoting that -20 dBm is the largest threshold value for which R^2 is greater than 0.95 for all $r > 0$. On the other hand, Figs. 3.3b and 3.3c show the shape and modified scale parameters of the fitted GPD model for the parameter stability method. These figures illustrate that by choosing $r > 15$, the estimated parameters of GPD are linear in u for $u < -32$ dBm, where -32 dBm is the largest threshold value for which $R^2 > 0.95$ for all $r > 15$. As a result, considering the intersection of the results obtained by Figs. 3.3a, 3.3b, and 3.3c, the optimum val-



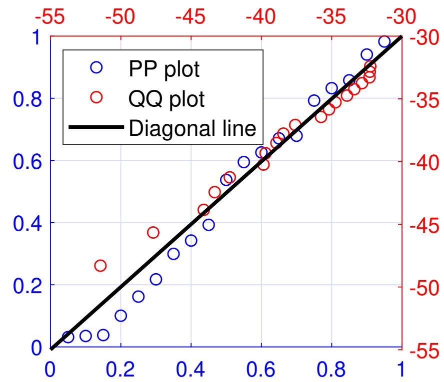
(a)



(b)



(c)



(d)

Figure 3.3: Mean excesses, shape, and modified scale parameters of GPD at different thresholds and minimum gaps between the clusters for the ramp road group: (a) Mean excess, (b) Modified scale parameter, (c) Shape parameter, and (d) Probability plots where the black line is a diagonal line, i.e., $y = x$, assessing the goodness of fit; x and y axes are empirical and modeled values.

ues of u and r are -32 dBm and 16, respectively. Additionally, Fig. 3.3d shows the probability plots where the black line is the diagonal line to graphically determine the goodness of fit of the GPD. Both PP and QQ plots illustrate that the generalized Pareto model applied to the received power properly follows the empirical results.

We summarize the specifications of Pareto models for other groups in Table 3.1, where $u_{(t)}$, $\xi_{(t)}$, $\tilde{\sigma}_{(u,t)}$, and l_{GPD} denote the optimum threshold in dBm, shape and scale parameters corresponding to $u_{(t)}$, and the log likelihood of GPD, respectively.

Table 3.1: Estimated time-varying parameters of GPD for different groups.

Group	$u_{(t)}$	$\xi_{(t)}$	$\tilde{\sigma}_{(u,t)}$	l_{GPD}
Static car	-6	0.126	0.63	-27.55
Smooth road	-24	-0.457	9.21	-55.25
Ramp road	-32	-0.284	8.08	-53.30

3.5.3 Final Channel Model for the Engine Compartment

According to Section 3.4.3, the deviance statistic between the GPD models under non-stationary and stationary assumptions $D = 2(-136.1 - (-1343)) >> 13.28$, where -136.1 is the summation of log-likelihoods of *Static car*, *Smooth road*, and *Ramp road* in Table 3.1; -1343 is the log-likelihood of the GPD fitted to the whole data sequence, assuming that the channel is stationary [Mehrnia and Coleri, 2021c]; 13.28 is the 0.99 quantile of the χ_4^2, while 4 in χ_4^2 is the difference between the complexities of GPD models under stationary and non-stationary channels which are 2 and 6, respectively. To minimize the probability of wrong decision, we assume α as small as possible, i.e., 0.01, in (3.6). The D statistic result implies that the change-point trend of the threshold, scale and shape parameters of GPD explains a substantial amount of the variation in the channel data sequence while outperforming the GPD model under stationarity assumption which returns a large log-likelihood of -1343.

3.6 Conclusions

In this chapter, we introduce a novel framework based on the extreme value theory with the goal of modeling the extremes of a non-stationary channel for URLLC. The proposed methodology for modeling a non-stationary channel achieves significantly better fit to the empirical data than the modeling approach for a stationary channel. In the future, we are planning to extend the proposed framework for the EVT analysis of the non-stationary processes to include a more extensive set of parameters that affect stationarity, such as vehicle speed, road material whether it is asphalt or stone, the temperature in the engine compartment, and traffic condition. Also, the extension of this chapter for deriving the mathematical expressions of the BER and PER due to the outage based on the proposed model is subject to future work.

Chapter 4

EXTREME VALUE THEORY BASED RATE SELECTION FOR ULTRA-RELIABLE COMMUNICATIONS

The content of this chapter has been published in IEEE Transactions of Vehicular Technology (TVT)-Correspondence.

4.1 Overview

Ultra-reliable low latency communication (URLLC) requires the packet error rate to be on the order of 10^{-9}-10^{-5}. Determining the appropriate transmission rate to satisfy the ultra-reliability constraint requires deriving the statistics of the channel in the ultra-reliable region and then incorporating these statistics into the rate selection. In this chapter, we propose a framework for determining the rate selection for ultra-reliable communications based on the extreme value theory (EVT). We first model the wireless channel at URLLC by estimating the parameters of the generalized Pareto distribution (GPD) best fitting to the tail distribution of the received powers, i.e., the power values below a certain threshold. Then, we determine the maximum transmission rate by incorporating the Pareto distribution into the rate selection function. Finally, we validate the selected rate by computing the resulting error probability. Based on the data collected within the engine compartment of Fiat Linea, we demonstrate the superior performance of the proposed methodology in determining the maximum transmission rate compared to the traditional extrapolation-based approaches.

4.2 Introduction

Ultra-reliable low latency communication (URLLC) is one of the most important features of the fifth generation (5G) networks with the aim of supporting mission critical applications, such as remote control of robots, remote surgery, autonomous vehicles and vehicular tele-operation applications [Popovski et al., 2018], [Eggers et al., 2019], [Angjelichinoski et al., 2019]. At URLLC, the packet error rate (PER) is guaranteed to be as low as 10^{-9}-10^{-5} to address the strict reliability constraint, while the acceptable latency is on the order of a few milliseconds [Popovski et al., 2018], [Gustafson et al., 2014], [Popovski et al., 2019], [Eggers et al., 2019], [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b]. Establishing a URLLC system necessitates the statistical modeling of the wireless channel tail quantifying the statistics of extreme events [Angjelichinoski et al., 2019], [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b], and the transmission strategies in the ultra-reliable region [Angjelichinoski et al., 2019].

Previous studies on the statistical modeling of the wireless channel for ultra-reliable communications extrapolate a wide range of practically important channel models to the ultra-reliability region [Eggers et al., 2019], [Angjelichinoski et al., 2019]. A simple power-law expression is proposed to estimate the tail of the cumulative distribution function (CDF) of the received power by extrapolating the commonly used practical channel models. Only recently, we have demonstrated that these extrapolated distributions are not accurate in the ultra-reliable region and can result in several orders of magnitude difference in the estimated packet error probabilities [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b]. We have proposed the usage of fading statistics from extreme value theory (EVT) and developed a novel methodology to derive these statistics efficiently with minimum amount of data. Nevertheless, none of these studies determine the maximum transmission rate in URLLC or assess the system reliability by considering ultra-reliable channel statistics.

Calculation of the reliability and rate selection for ultra-reliable communications has been addressed in the context of proposing new channel parameters and new performance measures. The definition of coherence time/distance has been modified as the time or distance over which a channel is predictable with a given reliability in [Swamy et al., 2018], [Swamy et al., 2019]. New performance measures are defined as average reliability for dynamic environments where the channel changes frequently; and probably correct reliability for the static environments where the channel statistics remain constant for a large enough time interval [Angjelichinoski et al., 2019]. However, these reliability measures have been derived by using the extrapolation of the traditional average statistic based channel models, which may not be accurate in the ultra-reliable region. Deriving outage probability for assessing the reliability of a wireless channel model is immensely important as uncertainty may degrade the communication performance by several orders of magnitude [Angjelichinoski et al., 2019]. It is worth noting that the asymptotic idea of outage probability, which is widely employed in wireless communication systems, is an accurate performance criteria even in the finite blocklength context [Zhou et al., 2019].

The goal of this chapter is to propose a novel EVT-based framework for the estimation and validation of the optimal transmission rate for ultra-reliable communications. The original contributions of this chapter are listed as follows

- We propose a novel framework based on the modeling of the tail distribution of the channel by using the GPD, determination of the optimum transmission rate by using the estimated values of the GPD parameters, and then, assessment of the system reliability by means of the outage probability metric.

- We formulate the rate selection function of GPD by incorporating the estimated Pareto parameters into the optimum transmission rate that guarantees the ultra-reliability constraints.

- We assess the reliability of the system by calculating the outage probability based on the rate selection function for the GPD.
- We demonstrate the superiority of the proposed methodology in terms of the reliability assessment, compared to the conventional method based on the extrapolation of the average statistics to the ultra-reliable region, over the data collected within the engine compartment of Fiat Linea under various engine vibrations and driving scenarios.

The rest of this chapter is organized as follows. Section 4.3 describes the system model and assumptions considered throughout the chapter. Section 4.4 presents the EVT-based framework for determining and validating the rate selection for the tail distribution of the channel by using GPD. Section 4.5 provides the channel measurement setup, and the performance evaluation in determining the optimum rate and outage probability. Finally, concluding remarks and future works are given in Section 4.6.

4.3 System Model

In this chapter, we consider an ultra-reliable communication system experiencing significant fading, resulting in extremely low received power values. We assume that the transmitter (Tx) sends a packet to a receiver (Rx) at rate R over an unknown channel. The receiver estimates the parameters of the GPD fitted to the channel tail distribution by applying EVT to the received power values. Then, the transmitter determines the transmission rate based on these estimated parameter values.

We assume that the channel is stationary, i.e., the distribution function of the received powers does not change over time, and the parameters specifying the distribution class are fixed over time. In the case of non-stationarity based on the results of Augmented Dickey-Fuller (ADF) test, the external factors varying the parameters of the GPD are determined such that the sequence is divided into M groups, in which the channel can be considered stationary, as detailed in [Mehrnia and Coleri, 2021b]. Also, the transmit power is fixed and known in advance. In such a case, using the received signal power is equivalent to using the squared amplitude of the channel state information [Angjelichinoski et al., 2019], [Mehrnia and Coleri, 2021c].

To study the outage probability and define the maximum transmission rate, we assume that the dominant source of error in block fading channels is the link outage, where we neglect other sources such as environmental noise. This model has been shown to be very suitable in the transmission of short packets in URLLC scenarios [Eggers et al., 2019], [Angjelichinoski et al., 2019], [Durisi et al., 2016]. Prior to transmission, at the training phase and after collecting the stationary channel samples by Rx, the channel sequence is converted into n independent and identically distributed (i.i.d.) sample sequence by using declustering method [Mehrnia and Coleri, 2021c], [Coles et al., 2001], denoted by $X^n = \{x_1, ..., x_n\}$, where x_i is the i^{th} i.i.d. sample, $i \in \{1, ..., n\}$. Let F be the CDF of the training samples X^n. The training phase uses either the dedicated pilot signals and training sequences or, the history of previous data transmissions and the associated feedback. The link outage is defined as

$$R(X^n) > \log_2(1 + Z), \tag{4.1}$$

where $R(X^n)$ denotes the optimum transmission rate estimated based on the n training received power samples, and Z is the received power from the test samples. Note that in (4.1), unit bandwidth is assumed; hence, $R(.)$ represents the transmission rate per unit bandwidth, i.e., spectral efficiency, in bits/sec. Then, according to

(4.1), the outage probability at transmission rate $R(X^n)$ is defined as

$$p_F(R(X^n)) = P\big[R(X^n) > \log_2(1+Z)\big]. \tag{4.2}$$

The goal of ultra-reliable communication is to choose the maximal rate that meets a predetermined reliability constraint, such that

$$p_F(R(X^n)) \leq \epsilon. \tag{4.3}$$

Guaranteeing constraint (4.3), the transmitter determines the maximum rate as a function of F as follows [Angjelichinoski et al., 2019]:

$$\begin{aligned} R_\epsilon(F) &= sup\big\{R(X^n) \geq 0 : p_F(R(X^n)) \leq \epsilon\big\} \\ &= \log_2\left(1 + F^{-1}(\epsilon)\right), \end{aligned} \tag{4.4}$$

where $R_\epsilon(F)$ denotes ϵ-outage capacity, and $F^{-1}(\epsilon)$ is the ϵ-quantile of F.

4.4 Rate Selection Framework

In our EVT-based framework, the rate selection function of GPD is formulated by incorporating the estimated Pareto parameters, i.e., scale and shape parameters, into the derivation of the optimum transmission rate guaranteeing ultra-reliability. First, received power samples are collected in the training phase. Afterwards, the training samples are converted to n i.i.d. samples by using the declustering method [Mehrnia and Coleri, 2021c], [Coles et al., 2001], as an input to the EVT process. Then, GPD is fitted to the lower tail of the i.i.d. sequence. Modeling the lower tail is followed by formulating the rate selection function of GPD. Finally, the channel reliability is assessed by comparing the estimated outage probability obtained in the test phase with the targeted PER. The proposed algorithm is depicted in Fig. 4.1

and explained in detail next.

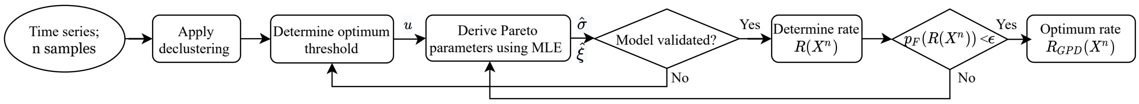


Figure 4.1: Flowchart of the proposed rate selection framework.

4.4.1 Channel Estimation

EVT provides a robust framework for analyzing the statistics of extreme events happening rarely through modeling the probabilistic distribution of the values exceeding a given threshold by using the GPD [Mehrnia and Coleri, 2021c], [Coles et al., 2001]. Assume that $X^n = \{x_1, ..., x_n\}$ is an i.i.d. stationary sequence of received powers x_i for $i \in \{1, ..., n\}$. Accordingly, the probabilistic distribution of the power values exceeding a given threshold u can be expressed as

$$F_u(y) = 1 - \left[1 + \frac{\xi y}{\tilde{\sigma}_u}\right]^{-1/\xi}, \tag{4.5}$$

where y is a non-negative value denoting the exceedance below threshold u, i.e., ($y = u - X$, X is any x_i below threshold u); and ξ and $\tilde{\sigma}_u = \sigma + \xi(u - \mu)$ are shape and scale parameters of the GPD, respectively, and μ and σ are the location and scale parameters of the generalized extreme value (GEV) distribution fitted to the CDF of $M_n = max\{x_1, ..., x_n\}$, respectively [Mehrnia and Coleri, 2021c, Theorem 1], [de Zea Bermudez and Kotz, 2010].

To model the tail distribution of the received power sequence, the measured samples are converted into an i.i.d. sequence by removing their dependency using the declustering approach [Mehrnia and Coleri, 2021c]. Then, EVT is applied to the i.i.d. samples for optimum threshold determination and the estimation of Pareto distribution parameters by using the MLE. The optimum threshold is determined by utilizing two complementary methods, mean residual life (MRL) and parameter stability methods [Mehrnia and Coleri, 2021c], [Coles et al., 2001]. The MLE estimates of the Pareto distribution parameters are formulated as follows.

Theorem 2. *Let GPD(σ,ξ) be the Pareto model fitted to the training samples $X^n = \{x_1, x_2, ..., x_n\}$. Then, the MLE of σ and ξ can be obtained as*

$$\begin{aligned} &\hat{\xi} = \tfrac{1}{k}\big[\textstyle\sum_{i=1}^{k} \log\big(1+\hat{\theta}y_i\big)\big]; \\ &\hat{\sigma} = \tfrac{\hat{\xi}}{\hat{\theta}} \end{aligned} \tag{4.6}$$

where k is the number of samples in the tail, i.e., exceeding the optimum threshold; $y_i = u - x_i$ for all $x_i < u$, where $i \in \{1, ..., n\}$; and $\hat{\theta}$ is the root of the following equation:

$$\Psi_k(\hat{\theta}) = \Big[\frac{1}{k}\sum_{i=1}^{k}(1+\hat{\theta}y_i)^{-1}(\frac{1}{k}\sum_{i=1}^{k}\log(1+\hat{\theta}y_i)+1)\Big]-1. \tag{4.7}$$

Proof. The log-likelihood function corresponding to the Pareto distribution fitted to the tail of i.i.d. sequence $X^n = \{x_1, x_2, ..., x_n\}$ is defined as $l_Y(\hat{\sigma}, \hat{\xi}) = \sum_{i=1}^{k} l_{y_i}(\hat{\sigma}, \hat{\xi})$, where $Y = \{y_1, y_2, ..., y_k\}$; $\hat{\sigma}$ and $\hat{\xi}$ are the estimated scale and shape parameters, respectively; and $l_{y_i}(\hat{\sigma}, \hat{\xi})$ is expressed as [Coles et al., 2001]

$$l_{y_i}(\hat{\sigma}, \hat{\xi}) = \log(\frac{1}{\hat{\sigma}}) - (1+\frac{1}{\hat{\xi}})\log(1+\frac{\hat{\xi}}{\hat{\sigma}}y_i). \tag{4.8}$$

By taking the partial derivative of (4.8) with respect to $\hat{\sigma}$ and $\hat{\xi}$, and making them equal to 0, we obtain the estimation of scale and shape parameters of GPD, respectively. Equating the partial derivatives to 0, we come up with the following equations:

$$\hat{\xi} = \frac{1}{k}\sum_{i=1}^{k}\log(1+\frac{\hat{\xi}}{\hat{\sigma}}y_i), \tag{4.9a}$$

$$\frac{1}{1+\hat{\xi}} = \frac{1}{k}\sum_{i=1}^{k}\frac{1}{1+\frac{\hat{\xi}}{\hat{\sigma}}y_i}. \tag{4.9b}$$

Substituting (4.9a) into (4.9b), we end up with a new equality as $\Psi_k(\hat{\theta})$ function expressed in (4.7). Then, the MLE of θ, i.e., $\hat{\theta}$, is the root of (4.7) and accordingly, the estimations of ξ and σ are obtained as shown in (4.6).

The validity of the estimated Pareto model is then assessed by using probability plots including probability/probability (PP) plot and quantile/quantile (QQ) plot. We refer the readers to [Mehrnia and Coleri, 2021c] and [Mehrnia and Coleri, 2021b] for more details on the channel modeling methodology of the extreme events at URLLC.

4.4.2 Rate Selection

Rate selection requires the estimation of the GPD parameters and calculation of the rate as a function of these parameters. Therefore, in order to determine the ϵ-outage in (4.4), we estimate $F^{-1}(\varepsilon_n)$ as

$$F_u^{-1}(\varepsilon_n) = u + \frac{\sigma_u}{\xi}\big[1-\varepsilon_n^{-\xi}\big], \tag{4.10}$$

where $F_u^{-1}(.)$ denotes the inverse of $F_u(.)$; u is the optimum threshold for GPD; and ε_n quantile is the probability whose associated quantile is of interest [Mehrnia and Coleri, 2021c]. The shape and scale parameters of GPD in (4.10) are the MLE of the GPD parameters fitted to the tail of the training samples $X^n = \{x_1, x_2, ..., x_n\}$.

Theorem 3. *Let $\hat{\sigma}$ and $\hat{\xi}$ be the MLE of the scale and shape parameters of the*

GPD(σ,ξ) fitted to the tail distribution of the training samples $X^n = \{x_1, x_2, ..., x_n\}$, *respectively. Then, the maximum transmission rate is defined as*

$$R_{GPD}(X^n) = \log_2\big(1 + u + \frac{\hat{\sigma}}{\hat{\xi}}\big[1 - \varepsilon_n^{-\hat{\xi}}\big]\big), \tag{4.11}$$

where u *is the optimum threshold for the channel tail estimation; and* ε_n *is the outage probability.*

Proof. The transmission rate for a specific training sample $X^n = \{x_1, x_2, ..., x_n\}$ is defined as

$$R(X^n) = \log_2\big(1 + \hat{F}^{-1}(\varepsilon_n)\big), \tag{4.12}$$

where the $\hat{F}^{-1}(\varepsilon_n)$ is the estimate of ε_n-quantile of F, for any distribution F. The goal is to find ε_n such that $R(X^n)$ is maximized and (4.3) is satisfied. Substituting (4.10) in (4.12), the maximum transmission rate guaranteeing a certain reliability, i.e., error probability ϵ, is defined as (4.11).

4.4.3 Validation of Selected Rate

The average outage probability, i.e., the average of error probability due to the outage of the GPD fitted to the extreme values exceeding a given threshold, is obtained by substituting (4.11) in outage probability equation (4.2), and taking the expectation as follows:

$$\begin{aligned} p_F(R_{GPD}(X^n)) &= E\Big[P\big[\log_2(1 + u + \frac{\hat{\sigma}}{\hat{\xi}}(1 - \varepsilon_n^{-\hat{\xi}}) > \log_2(1 + Z)\big]\Big] \\ &= E\Big[P\big(-\frac{\hat{\sigma}}{\hat{\xi}}(1 - \varepsilon_n^{-\hat{\xi}}) < u - Z\big)\Big] \\ &= E\Big[\big(1 - \frac{\xi}{\sigma}(\frac{\hat{\sigma}}{\hat{\xi}}(1 - \varepsilon_n^{-\hat{\xi}}))\big)^{-1/\xi}\Big], \end{aligned} \tag{4.13}$$

where $E[.]$ denotes the expectation function; $\hat{\xi}$ and $\hat{\sigma}$ are the MLE of estimated shape and scale parameters fitted to the training samples, respectively; and ξ and σ are the shape and scale parameters of the GPD, given that $F_u(.)$ is perfectly known and GPD is fitted to the sample sequence including the data in training and test phases. By referring to the reliability constraint expressed in (4.3), the outage probability obtained from (4.13) should be less than or equal to the ϵ. Therefore, the maximum allowed error probability, ε_n, can be calculated with respect to the targeted PER ϵ, as follows:

$$\varepsilon_n = \left[1 - \frac{\hat{\xi}}{\hat{\sigma}}\left(\frac{\sigma}{\xi}\left(1 - \epsilon^{-\xi}\right)\right)\right]^{-\frac{1}{\hat{\xi}}}. \tag{4.14}$$

If $R_{GPD}(X^n)$ is a valid transmission rate, the corresponding outage probability $p_F(R_{GPD}(X^n))$ is expected not to exceed the targeted reliability ϵ, according to constraint (4.3).

The complexity of the proposed rate selection framework is $O(n)$, similar to the traditional extrapolation approach [Angjelichinoski et al., 2019], [Swamy et al., 2018], [Swamy et al., 2019], since the optimum transmission rate is determined based on n training samples.

4.5 Numerical Results

The goal of this section is to evaluate the performance of the proposed methodology in determining the maximum transmission rate for URLLC, evaluating the system reliability by using the outage probability metric, comparing it to the traditional extrapolation-based methods for the Rayleigh fading in the estimation of the transmission rate, and illustrating the impact of channel mismatch on the reliability performance of the system. It is worth mentioning that throughout this section, the Rayleigh distribution refers to the channel model under the Rayleigh fading.

To obtain the extrapolated Rayleigh curve, we fit the Rayleigh distribution to the first 10^3 samples from the received power sequence and upon estimating the Rayleigh parameter, we compute the transmission rate as expressed in [Angjelichinoski et al., 2019, Equation 23]. Then, we extrapolate the fitted Rayleigh distribution toward the ultra reliable region and compute the corresponding outage probability as expressed in [Angjelichinoski et al., 2019, Equation 26], yet assuming that the transmission rate is the Rayleigh rate computed based on [Angjelichinoski et al., 2019, Equation 23].

To represent the impact of channel mismatch on the system reliability, we fit the Rayleigh distribution to the received power values and then, compute the maximum rate as expressed in [Angjelichinoski et al., 2019, Equation 23]. Then, we assume that the channel is no longer Rayleigh, i.e., F in constraint (4.2) is different than Rayleigh (the true distribution F is estimated by the GPD at different thresholds as it refers to the distribution of the channel tail); yet Tx assumes that the channel is Rayleigh and hence, setting the rate as in [Angjelichinoski et al., 2019, Equation 23] with ε_n obtained as expressed in [Angjelichinoski et al., 2019, Equation 27]. Due to mismatch, the estimated transmission rate w.r.t. the Rayleigh distribution does not converge to $R(X^n)$ computed w.r.t. the true distribution F. Additionally, the outage probability is computed such that the channel tail distribution is estimated by the GPD while the system operating at the Rayleigh transmission rate obtained by [Angjelichinoski et al., 2019, Equation 23].

The measured channel data were collected within the engine compartment of Fiat Linea at 60 GHz. The locations of the transmitter and receiver antennas are selected out of the possible locations for the wireless sensors located within the engine compartment, namely locations 5 and 13 in [Demir et al., 2013, Fig. 1], such that the effect of the engine vibration is observed in the received power, as shown in Fig. 3.2. A Vector Network Analyzer (VNA) (R & S® ZVA67) is connected to the transmitter and receiver via the R & S® ZV-Z196 port cables with maximum 4.8

dB transmission loss. The horn transmitter and receiver antennas with a nominal 24 dBi gain and 12° vertical beam-width operate at 50-75 GHz. We have captured about 10^6 successive samples for 30 minutes with a time resolution of 2 ms. We use MATLAB® for the implementation of the proposed algorithm. Due to the existence of a non-stationarity trend among the samples, the measured samples are categorized into two stationary groups according to the engine vibration. Group one includes bunches of 10^3 successive samples all above -12 dBm. If any sample within 10^3 samples is less than -12 dBm, the set of successive samples is assigned to group two.

4.5.1 Optimum Transmission Rate

To determine the maximum transmission rate at URLLC, we plot the normalized $R_{GPD}(X^n)$, denoted by w, for different sample numbers. The normalization is performed with respect to the optimal throughput, given that $F_u(.)$ is perfectly known, i.e., the last $R_{GPD}(X^n)$ estimated based on the whole set of data samples, not only the training ones.

Fig. 4.2 shows the normalized rate w for GPD fitted to the filtered i.i.d. samples of groups 1 and 2 at different sample numbers, threshold values, and error rates $\epsilon \in \{10^{-3}, 10^{-4}, 10^{-5}\}$. Fig. 4.2a shows that the normalized rate selection function w converges faster to 1, meaning that the transmission rate converges to its optimum value, at lower targeted PER $\epsilon = 10^{-3}$. Comparing Fig. 4.2a and Fig. 4.2b at each targeted PER ϵ, the normalized rate in Fig. 4.2b approaches to 1 significantly faster than that shown in Fig. 4.2a. This is expected since by relaxing the threshold from -10 dBm to -5 dBm, determining GPD and its corresponding transmission rate requires the collection of less samples in the training phase. Additionally, by increasing the threshold, at lower sample numbers, greater portion of received power with higher range and variety is included in the tail and therefore, the GPD parameters and the transmission rate estimated at lower sample number remain valid for higher sample numbers. Fig. 4.2c also demonstrates that the convergence of w function to 1 is slow for targeted PER $\epsilon = 10^{-5}$. The fluctuations in Fig. 4.2c is due to the low number of collected samples for estimating the GPD and its corresponding rate function.

Figs. 4.2a, 4.2b, and 4.2c also illustrate the normalized rate w for GPD compared to the extrapolation based method for groups 1 and 2 of data and for different thresholds of -5 dBm and -10 dBm at group 1, and -25 dBm at group 2. We observe that the normalized rate under the GPD assumption converges faster or simultaneously to the optimum rate, compared to that of the extrapolated approach for both groups (referring to Figs. 4.2a, 4.2b, and 4.2c), and under different thresholds (referring to Figs. 4.2a and 4.2b). The normalized rate of extrapolated approach in Fig. 4.2 corresponds to all ϵ values, as the normalized rate under the extrapolated Rayleigh rate-selection function appears to be (almost) independent from the targeted error probability ϵ, implying that the rate-selection function obtained for a given ϵ is valid for the others as well, as also stated in [Angjelichinoski et al., 2019].

Figs. 4.2a, 4.2b, and 4.2c besides show the impact of mismatch between the estimated channel and the true one. The proposed method in Fig. 4.2 performs significantly better than the case where the true channel distribution is actually modeled by the GPD while Rayleigh distribution was assumed. As also stated in [Angjelichinoski et al., 2019], the convergence of the rate is heavily affected by the model mismatch, degrading the normalized rate to almost 0 for all ϵ values. However, the effect of model mismatch can be reduced by increasing the sample number to the higher values.

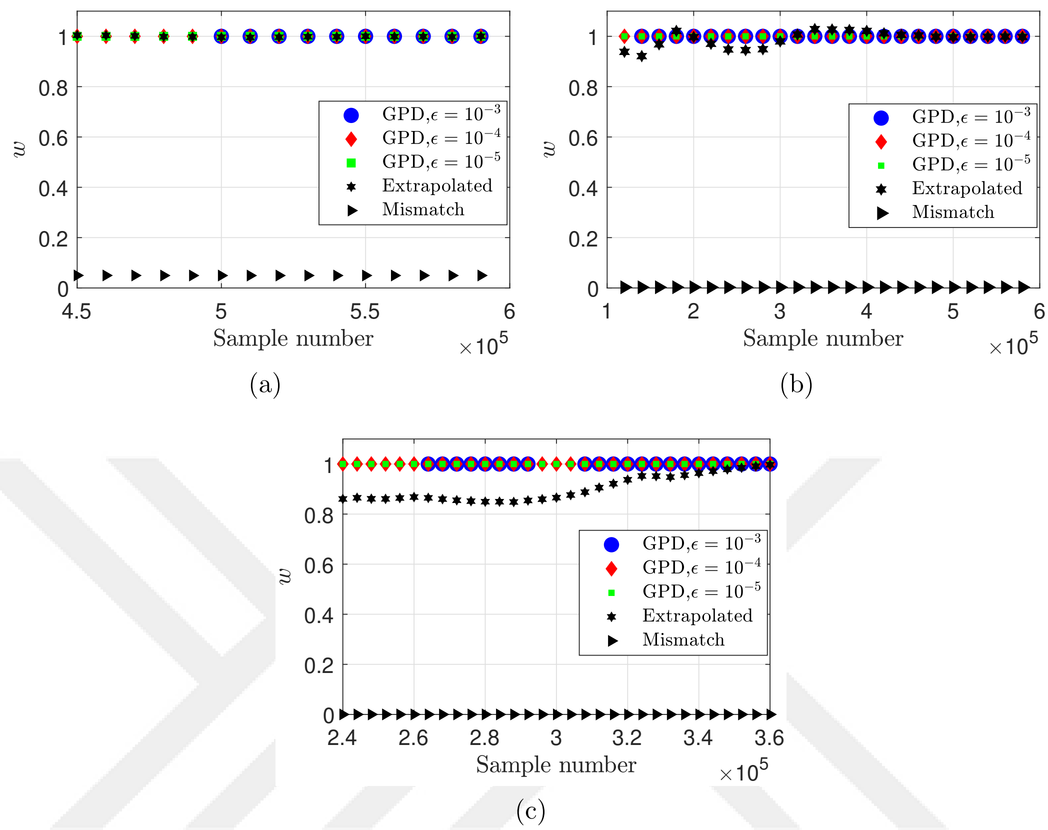


Figure 4.2: Normalized transmission rate for: (a) group 1 at $u = -10$ dBm, (b) group 1 at $u = -5$ dBm, and (c) group 2 at $u = -25$ dBm. The single *Extrapolated* and *Mismatch* plots correspond to all ϵ values. *Extrapolated* and *Mismatch* plots are both based on the Rayleigh assumptions.

4.5.2 Reliability Assessment

Fig. 4.3 illustrates the reliability performance of the fitted GPD based on the calculated outage probability for a range of thresholds for 2 groups of data. In this figure, the outage probability in (4.13) is plotted for fixed error probability $\epsilon \in \{10^{-3}, 10^{-4}, 10^{-5}\}$ for large enough n to ensure convergence [Angjelichinoski et al., 2019]. This demonstrates that at both groups and for all error probabilities, the reliability constraints in (4.3) are never violated.

Figs. 4.3a and 4.3b also show the reliability performance of the proposed Pareto model compared to the conventional extrapolation-based method by comparing their

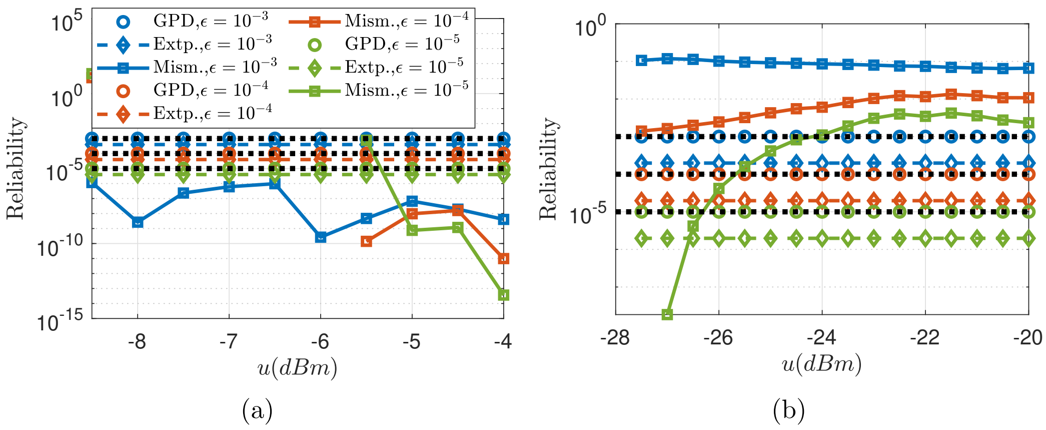


Figure 4.3: Reliability measure of the proposed model and the traditional extrapolation approach for: (a) group 1, and (b) group 2; The dashed-black lines are the reference lines at targeted error rate $\epsilon \in \{10^{-3}, 10^{-4}, 10^{-5}\}$. *Extrapolated* and *Mismatch* plots are both based on the Rayleigh assumptions.

outage probabilities for $\epsilon \in \{10^{-3}, 10^{-4}, 10^{-5}\}$. Since in the traditional approach, the outage probability is independent of any GPD thresholds, we plot a single outage probability obtained by the extrapolated method at all the threshold values. In Fig. 4.3a, the reliability of the extrapolated approach at each ϵ value is almost identical to that of the proposed approach and equal to the desired error rate. However, in Fig. 4.3b, the achieved reliability obtained by the extrapolated approach differs from the proposed GPD-based approach, as well as the targeted ϵ. It is due to the fact that in Group 2, the system suffers from more extreme values with lower received power. In such a situation, the channel tail needs to be estimated by the GPD and not the extrapolation.

Figs. 4.3a and 4.3b illustrate the effect of mismatch on the reliability of the system through comparing their outage probabilities for $\epsilon \in \{10^{-3}, 10^{-4}, 10^{-5}\}$. The proposed method in Fig. 4.3a has been demonstrated to perform significantly better than the case where the true channel distribution differs from the estimated Rayleigh distribution. The effect of mismatch is more clear for the low ϵ values and at higher thresholds of the GPD. However, the proposed method in Fig. 4.3b performs much better for all ϵ values and thresholds. On the other hand, while with the proposed GPD approach, the reliability is always equivalent to the required error rate ϵ, the reliability of the mismatch case with Rayleigh rate assumption deviates from the desired ϵ almost at all thresholds.

4.6 Conclusions

In this chapter, we proposed an EVT-based rate selection framework for ultra-reliable communications. First, we represent the channel by GPD distribution and estimate its parameters. Then, we determine the maximum transmission rate of the estimated channel and validate the selected rate by assessing the resulting error probability. The achieved reliability from the proposed EVT-based approach outperforms the traditional methods based on the extrapolation of average statistic channel models. Additionally, the GPD threshold has significant impact on the minimum required number of samples to achieve a certain reliability order.

Chapter 5

INCORPORATION OF CONFIDENCE INTERVAL INTO RATE SELECTION BASED ON THE EXTREME VALUE THEORY FOR ULTRA-RELIABLE COMMUNICATIONS

The content of this chapter has been presented in the European Conference on Networks and Communications (EuCNC) and the 6G Summit (2022 EuCNC and 6G Summit).

5.1 Overview

Proper determination of the transmission rate in ultra-reliable low latency communication (URLLC) needs to incorporate a confidence interval (CI) for the estimated parameters due to the large amount of data required for their accurate estimation. In this chapter, we propose a framework based on the extreme value theory (EVT) for determining the transmission rate along with its corresponding CI for an ultra-reliable communication system. This framework consists of characterizing the statistics of extreme events by fitting the generalized Pareto distribution (GPD) to the channel tail, deriving the GPD parameters and their associated CIs, and obtaining the transmission rate within a confidence interval. Based on the data collected within the engine compartment of Fiat Linea, we demonstrate the accuracy of the estimated rate obtained through the EVT-based framework considering the confidence interval for the GPD parameters. Additionally, we show that proper estimation of the transmission rate based on the proposed framework requires a lower number of samples compared to the traditional extrapolation-based approaches.

5.2 Introduction

Ultra-reliable low latency communication (URLLC) is a substantial new feature brought by the fifth generation (5G), with a potential to support a vast set of machine type and mission-critical applications such as reliable remote control of robots, tele-operation driving (ToD) applications, or coordination among vehicles [Popovski et al., 2018], [Mehrnia and Coleri, 2021c], [Bennis et al., 2018]. A system operating at URLLC is expected to address the strict reliability requirement of packet error rate (PER) at 10^{-9}-10^{-5} level, while the tolerable latency is on the order of milliseconds or less [Popovski et al., 2018], [Gustafson et al., 2014], [Popovski et al., 2019], [Eggers et al., 2019], [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b]. Fulfillment of the reliability requirements at URLLC necessitates a statistical model of the channel tail distribution [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b], as well as the transmission strategies guaranteeing ultra-reliability concerns [Angjelichinoski et al., 2019]. An accurate estimation of the wireless channel parameters in the ultra-reliable regime requires a massive amount of data. However, analyzing the confidence intervals in the estimated parameters as a function of the amount of historical data allows a significant reduction in the minimum sample numbers required for the estimation of wireless channel tail distribution at URLLC.

Current studies on the statistical modeling of the wireless channel for ultra-reliable communications either extrapolate a wide range of commonly used practical channel models toward the ultra-reliability region [Eggers et al., 2019], [Angjelichinoski et al., 2019], or utilize extreme value theory (EVT) to characterize the tail distribution of the channel data and derive the statistics of extreme events efficiently with a minimum amount of data [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b]. However, none of these studies assess the reliability of the system through the estimation of the transmission rate.

Measuring the system reliability and the transmission rate for ultra-reliable communications has been addressed in the context of new performance measures of the reliability based on the transmission rate [Angjelichinoski et al., 2019], and proposing a rate selection framework in the absence of knowledge about the channel state information (CSI) [Pase et al., 2021]. In [Angjelichinoski et al., 2019], two performance measures have been proposed with the goal of selecting a transmission rate guaranteeing ultra-reliability: average reliability (AR) for dynamic environments where the channel changes frequently; and probably correct reliability (PCR) for the static environments where the channel statistics can be considered constant for an extensive time interval. Moreover, [Pase et al., 2021] evaluates if the training times are affected by the cumulative distribution function (CDF) of the fading channel when a perfect CSI is not available, through estimating an optimal transmission rate for a channel and then, measuring the reliability of the system accordingly. Nevertheless, such approximations characterize the tail statistic based on the extrapolation of the average-statistics channel models, and therefore, might be deficient to fulfill the reliability targets of ultra-reliability [Mehrnia and Coleri, 2021c]. Only recently, we have proposed a novel EVT-based framework for the estimation and validation of the optimal transmission rate for ultra-reliable communications [Mehrnia and Coleri, 2021a]. However, the possible changes in the transmission rate due to the estimation error have been neglected in the computation.

Reliability of the system is strongly affected by the accuracy of the estimated channel parameters due to their effect on the transmission rate. Therefore, confidence interval (CI) is derived as a metric to guarantee the reliability constraint despite the limited amount of training data. In [Zhang et al., 2021], the authors propose a non-parametric statistical learning algorithm that utilizes kernel density estimation (KDE) for estimating the probability density function (PDF) of the channel distribution and then, upon selecting a suitable transmission channel rate, evaluate the outage probability within a confidence level to maintain the URLLC reliability. However, the proposed empirical non-parametric approach is based on the extrapolation of the average statistics fading channels. The authors in [Samarakoon et al., 2020] proposed a data-driven learning framework to characterize the statistics of a non-blocking connectivity duration within a confidence interval, in the absence of the statistical knowledge about the dynamic wireless channel, with the goal of addressing the ultra-reliability constraint in URLLC. However, the machine learning data driven approaches require an extensive large amount of training samples, much more than $(10 \times \epsilon^{-1})$ to meet the targeted error probability ϵ.

The goal of this chapter is to propose a novel framework based on EVT for the estimation of the optimal transmission rate along with the corresponding confidence interval with a sufficiently low number of samples for ultra-reliable communications. According to the proposed framework, upon incorporating the estimates of the generalized Pareto distribution (GPD) parameters fitted to the channel tail into the rate selection, the upper and lower bounds of the GPD parameters are employed to acquire the confidence interval for the transmission rate. The original contributions of this chapter are listed as follows:

- We propose a methodology based on EVT for ultra-reliable communications, where the channel tail distribution is modeled by using GPD, the upper and lower bounds for the estimated Pareto parameters are derived and then incorporated into the estimation of the confidence interval for the transmission rate.

- We determine the confidence intervals of the estimated GPD parameters by incorporating t-distribution and multiple iterations for the parameter estimation using the maximum likelihood estimator (MLE).

- We formulate the confidence interval of the transmission rate in conjunction with the confidence interval of the estimated GPD parameters.

- Based on the data collected within the engine compartment of Fiat Linea, we apply our methodology at different numbers of training samples to ensure that the targeted error probability of URLLC is met even when limited data is available.

The rest of the chapter is organized as follows. Section 5.3 describes the system model and assumptions considered throughout the chapter. Section 5.4 presents the EVT-based framework for determining the rate selection and its affiliated confidence interval in connection with the accuracy of the GPD parameters fitted to the tail distribution of the channel. Section 5.5 provides the channel measurement setup, and the performance evaluation in determining the confidence interval for the estimated transmission rate. Finally, concluding remarks and future works are given in Section 5.6.

5.3 System Model

We consider a single transmitter (Tx)-receiver (Rx) pair communicating over a stationary channel, i.e., the parameters of the channel distribution are fixed over a vast period of time. If the channel is non-stationarity according to the Augmented Dickey-Fuller (ADF) test results, the external factors causing time variation in the parameters of the GPD are determined such that the sequence is divided into M groups, each of which can be considered stationary, as explained in details in [Mehrnia and Coleri, 2021b]. The transmit power is assumed to be fixed and known in advance. Therefore, estimating the received signal power is equivalent to estimating the squared amplitude of the channel state information [Angjelichinoski et al., 2019], [Mehrnia and Coleri, 2021c].

We assume that Tx sends a packet to an Rx at rate R over an unknown channel. Applying an EVT-based channel modeling methodology, the statistics of the channel tail distribution are estimated by fitting the generalized Pareto distribution to the received power values exceeding a given threshold. Afterward, the confidence intervals of the estimated parameters corresponding to different probabilities of wrong decisions are derived for different numbers of samples. Then, the statistics of the transmission rate are estimated according to the EVT by incorporating the confidence intervals of the estimated Pareto parameters for different sample numbers to meet the targeted error probability at an ultra-reliable communication.

5.4 EVT-based Rate Selection Framework

We propose an EVT-based framework to model the channel tail distribution of the received powers by using GPD, determine the transmission rate guaranteeing ultra-reliability, and estimate the confidence interval for the estimated Pareto parameters and the transmission rate. The framework consists of the following steps: The received power sequence is converted into independent and identically distributed (i.i.d.) samples since EVT requires i.i.d. sequence as an input. Then, GPD is fitted to the tail of i.i.d. received powers to model the channel tail distribution. Upon estimating the Pareto parameters by using MLE, the optimum transmission rate for an ultra-reliable communication is determined. Finally, the confidence intervals of the estimated GPD parameters and the transmission rate are computed at different sample numbers. By incorporating the confidence interval into our proposed framework, we expect to decrease the number of samples required to determine the optimum rate that meets ultra-reliable constraints in URLLC.

5.4.1 Channel Estimation

EVT provides a robust framework for analyzing the statistics of extreme events happening rarely through modeling the probabilistic distribution of the values exceeding a given threshold by using the GPD [Mehrnia and Coleri, 2021c], [Coles et al., 2001]. To determine the statistics of the channel tail distribution, prior to the transmission and at the training phase, Rx collects stationary channel measurements and then, convert them into the sequence of n i.i.d. samples, denoted by $X^n = \{x_1, ..., x_n\}$, with the cumulative distribution function (CDF) F, where x_i is the i^{th} sample, $i \in \{1, ..., n\}$. Upon converting the measured samples into an i.i.d. sequence by removing their dependency using the declustering approach [Mehrnia and Coleri, 2021c], the Generalized Pareto Distribution (GPD) is fitted to the received power values exceeding a given threshold to model the channel tail distribution as

$$F_u(y) = 1 - \left[1 + \frac{\xi y}{\tilde{\sigma}_u}\right]^{-1/\xi}, \tag{5.1}$$

where y is a non-negative value denoting the exceedance below threshold u, i.e., $y = u - X|X = \{x_i < u\}$; ξ and $\tilde{\sigma}_u = \sigma + \xi(u - \mu)$ are shape and scale parameters of the GPD, respectively; and μ and σ are the location and scale parameters of the generalized extreme value (GEV) distribution fitted to the CDF of $M_n = max\{x_1, ..., x_n\}$, respectively [Mehrnia and Coleri, 2021c, Theorem 1], [de Zea Bermudez and Kotz, 2010].

The optimum threshold u is determined by utilizing two complementary methods, mean residual life (MRL) and parameter stability methods [Mehrnia and Coleri, 2021c], [Coles et al., 2001], while the estimation of Pareto distribution parameters is performed by using the Maximum Likelihood Estimation (MLE). The validity of the estimated Pareto model is then assessed by using probability plots including probability/probability (PP) plot and quantile/quantile (QQ) plot.

5.4.2 Rate Selection

The optimum transmission rate, assuming the unit bandwidth, is determined such that:

$$R_{GPD}(X^n) = \log_2\big(1 + u - \frac{\hat{\sigma}}{\hat{\xi}}\big[1 - \varepsilon_n^{-\hat{\xi}}\big]\big), \tag{5.2}$$

where $R_{GPD}(X^n)$ denotes the transmission rate determined based on n training received power samples; u is the optimum threshold for the channel tail estimation; $\hat{\sigma}$ and $\hat{\xi}$ are the MLE of the scale and shape parameters of the GPD(σ,ξ) fitted to the tail distribution of the training samples $X^n = \{x_1, x_2, ..., x_n\}$, respectively; and ε_n is the outage probability calculated with respect to the targeted PER ϵ, as follows:

$$\varepsilon_n = \big[1 - \frac{\hat{\xi}}{\hat{\sigma}}\big(\frac{\sigma}{\xi}\big(1 - \epsilon^{-\xi}\big)\big)\big]^{-\frac{1}{\hat{\xi}}}, \tag{5.3}$$

where ξ and σ are the shape and scale parameters of the GPD, given that $F_u(.)$ is perfectly known and GPD is fitted to the sample sequence including the data in both training and test phases [Mehrnia and Coleri, 2021a]. Then, the lower and upper bounds of the estimated Pareto parameters, as well as the estimated transmission rate are calculated for different values of probability of wrong decision α and at different sample numbers. The small α refers to a critical value based on which, we are $100(1-\alpha)\%$ confident that the estimated parameter is within the derived confidence interval.

5.4.3 Confidence Interval

CI for the Estimated Pareto Parameters:

The estimated scale and shape parameters of GPD are not exactly $\hat{\sigma}$, $\hat{\xi}$ and typically take values within the confidence interval. Suppose that we have estimated the value of the GPD parameter $\theta(i)$ for a given realization i, where $\theta(i)$ is either scale (σ) or shape (ξ) parameter. Our goal is to determine $(1-\alpha)$ confidence interval, $[l(i), u(i)]$, such that

$$Pr\big(E(\theta_i) \in \big[l(i), u(i)\big]\big) \geq 1 - \alpha,$$

for a small value of $\alpha \in (0, 1)$ [Chen et al., 2014], where

$$\begin{aligned} l(i) &= \bar{\theta}_i - t^* s_i, \\ u(i) &= \bar{\theta}_i + t^* s_i \end{aligned} \tag{5.4}$$

are the lower and upper bounds of the CI, respectively. In (5.4), t^* is the critical value obtained from a t-distribution with $M - 1$ degrees of freedom, M denotes the iteration number to estimate the parameter; and s_i is the standard deviation of the θ_i, calculated by $s_i = \sqrt{(1/M)\sum_{j=1}^{M}(\theta_j - \bar{\theta}_i)^2}$ while $\bar{\theta}_i$ is the expected value of parameter θ derived by taking an average of the estimated values over M iterations. Assuming a small quantity, such as 0.01 or 0.05 for the critical value α, we are $100(1-\alpha)\%$ confident that the estimated parameter θ is within the range $[l(i), u(i)]$.

CI for the Transmission Rate:

According to (5.2), the transmission rate of a channel operating at URLLC is a function of the GPD parameters. Since changes in the estimated Pareto parameters alter the transmission rate, the upper and lower bounds of the GPD parameters yield corresponding upper and lower bounds for the transmission rate. Referring to (5.2), the upper/lower bound of scale parameter $\hat{\sigma}$ along with the lower/upper bound of the shape parameter $\hat{\xi}$ provide the upper/lower bound of the transmission rate $R_{GPD}(X^n)$. Someone might say that $R_{GPD}(X^n)$ is also a function of ε. However, by assuming large number of training samples on the order of $1/\epsilon$, $\varepsilon_n \approx \epsilon$. It is due to the fact that the very small difference between the actual GPD parameters and the estimated ones are negligible and therefore, $(\frac{\hat{\xi}}{\hat{\sigma}})(\frac{\sigma}{\xi})$ and $(-\xi)(-\frac{1}{\hat{\xi}})$ in the power of (5.3), are ≈ 1. As a result, the only parameters affecting the upper and lower bounds of the transmission rate in (5.2), are the lower and upper bounds of the estimated GPD parameters fitted to n training samples, i.e., $\hat{\xi}$ and $\hat{\sigma}$. Accordingly, the lower/upper bound of the transmission rate is calculated as follows:

$$R_{GPD}(X^n)_{l/u} = \log_2\big(1 + u - \frac{\hat{\sigma}_{l/u}}{\hat{\xi}_{u/l}}\big[1 - \varepsilon_n^{-\hat{\xi}_{u/l}}\big]\big), \tag{5.5}$$

where subscription l/u denotes the lower/upper bound of the confidence interval for the parameters already defined in (5.2).

5.5 Numerical Results

The goal of this section is to evaluate the performance of the proposed framework in modeling the tail distribution of the channel at URLLC, estimating the transmission rate of the channel, and determining the confidence interval of the rate through the computation of the confidence interval for the parameters of GPD fitted to the tail distribution of the channel data, for different probabilities of wrong decision, α values, and different sample numbers.

The channel data were collected within the engine compartment of Fiat Linea at 60 GHz. The implemented antennas at the engine compartment are located such that the effect of the engine vibration is observed in the received power, as shown in Fig. 3.2a. A Vector Network Analyzer (VNA) (R & S® ZVA67) is connected to the transmitter and receiver via the R & S® ZV-Z196 port cables with maximum 4.8 dB transmission loss, as shown in Fig. 3.2b. The horn transmitter and receiver antennas with a nominal 24 dBi gain and 12° vertical beam-width operate at 50-75

GHz. About 10^6 successive samples have been captured for 30 minutes with a time resolution of 2 ms.

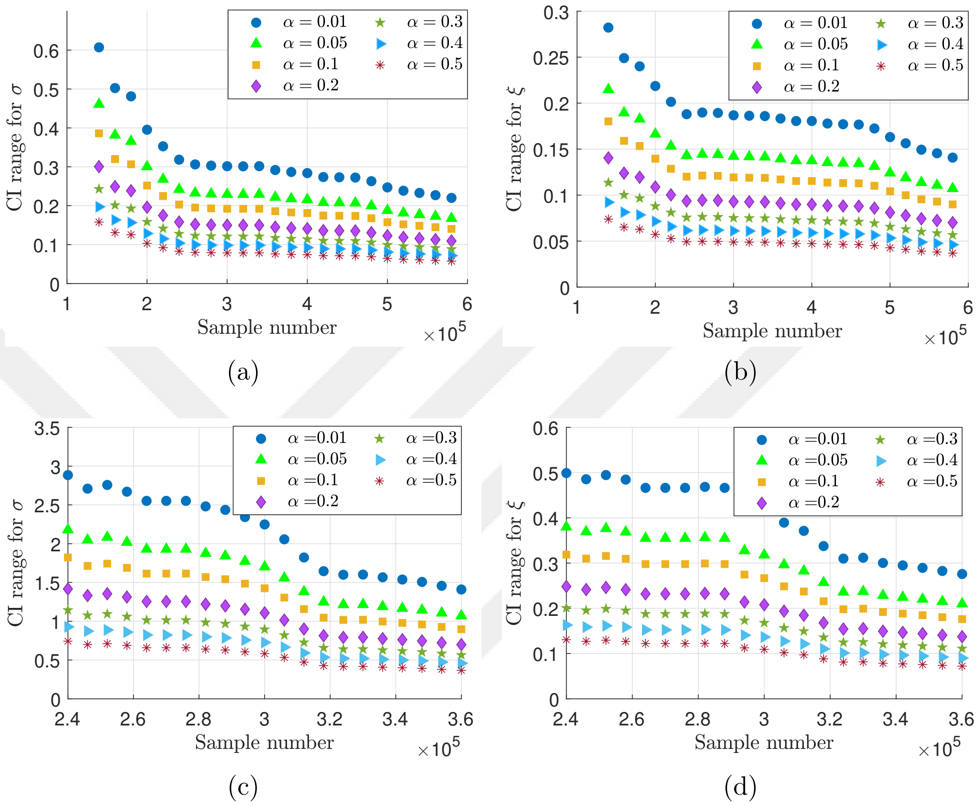


Figure 5.1: CI range for the estimated Pareto parameters considering different α values for groups 1 and 2 at different sample numbers: (a) CI range of the scale parameter for group 1, (b) CI range of the shape parameter for group 1, (c) CI range of the scale parameter for group 2, and (d) CI range of the shape parameter for group 2. The CI range refers to the difference between the upper and lower bounds of the CI.

MATLAB® is used for the implementation of the proposed framework. Due to the existence of a non-stationarity trend among the samples, the measured samples are categorized into two stationary groups according to the engine vibration. Group 1 includes bunches of 10^3 successive samples all above -12 dBm. If any sample within 10^3 samples is less than -12 dBm, the set of successive samples is assigned to group 2. It should be noted that the optimum threshold for group 1 and group 2

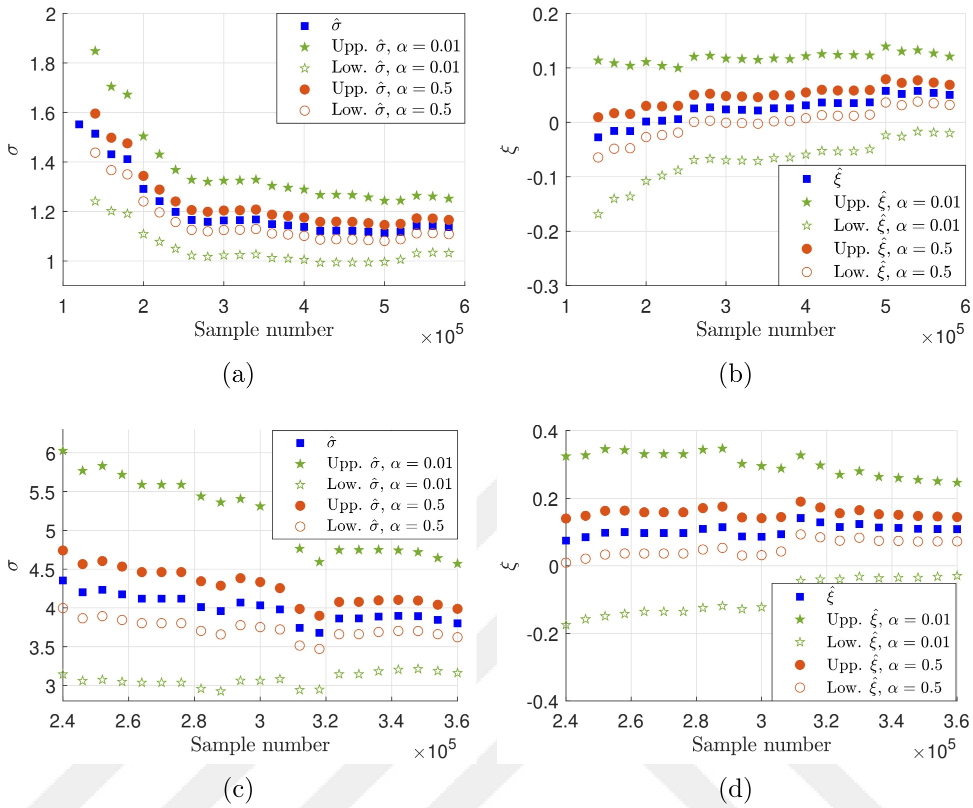


Figure 5.2: The estimated Pareto parameters along with their CI considering $\alpha = 0.01, 0.5$ for groups 1 and 2 at different sample numbers: (a) Scale parameter and the corresponding CI for group 1, (b) Shape parameter and the corresponding CI for group 1, (c) Scale parameter and the corresponding CI for group 2, and (d) Shape parameter and the corresponding CI for group 2. Blue plot is the estimated parameter; circle/pentagram corresponds to $\alpha = 0.5/0.01$; filled/empty plot corresponds to upper/lower bound of CI.

are estimated as -5 dBm and -20 dBm, respectively, and the targeted packet error rate, ϵ in (5.3), is 10^{-5}.

5.5.1 CI for the GPD Parameters

Fig. 5.1 illustrates the confidence interval range for the estimated GPD parameters considering different α values at different sample numbers for groups 1 and 2. The range of confidence interval refers to the absolute value of the difference between the upper and lower bounds of the CI interval. As the α value increases from 0.01 to 0.5, i.e., allowing more error in the estimation of the GPD parameters, the CI range decreases significantly. Additionally, by increasing the number of training samples, the CI range decreases remarkably and, the estimation accuracy for the GPD parameters increases. Moreover, since group 2 suffers from more severe extreme values, the minimum number of training samples required to estimate the GPD parameters fitted to the tail distribution is higher that that of group 1.

Fig. 5.2 shows the estimated GPD parameters along with the lower and upper bounds of CIs for $\alpha = \{0.01, 0.5\}$ and different sample numbers in both groups 1 and 2. Please note that Fig. 5.1 depicts the CI range while Fig. 5.2 illustrates the CI itself for the estimated GPD parameters. Above a certain number of samples, the estimated GPD parameters become almost constant and invariant with respect to the number of samples. Furthermore, the MLE with fewer samples yields a high level of uncertainty in the estimated scale (σ) and shape (ξ) parameters, but the uncertainty decreases as the sample number increases. This highlights the trade-off between the uncertainty of the GPD parameters and the cost of collecting data. Unlike the traditional non-parametric rate selection technique, which requires substantially more channel training than $1/\epsilon$ stated in [Angjelichinoski et al., 2019], our EVT-based framework appropriately estimates the rate within a confidence interval with a sample size of roughly $1/\epsilon$.

5.5.2 CI for the Transmission Rate

Fig. 5.3 illustrates the CI range for the transmission rate considering different α values at different sample numbers for groups 1 and 2. Similar to what we observed in Fig. 5.1, increasing the α value reduces the CI range. In addition, increasing the sample size during the training phase dramatically decreases the CI range, while converging to a fixed value beyond a certain large enough sample number. Furthermore, the minimum sample numbers necessary to fit the GPD to the tail distribution of channel data are higher in group 2 than in group 1.

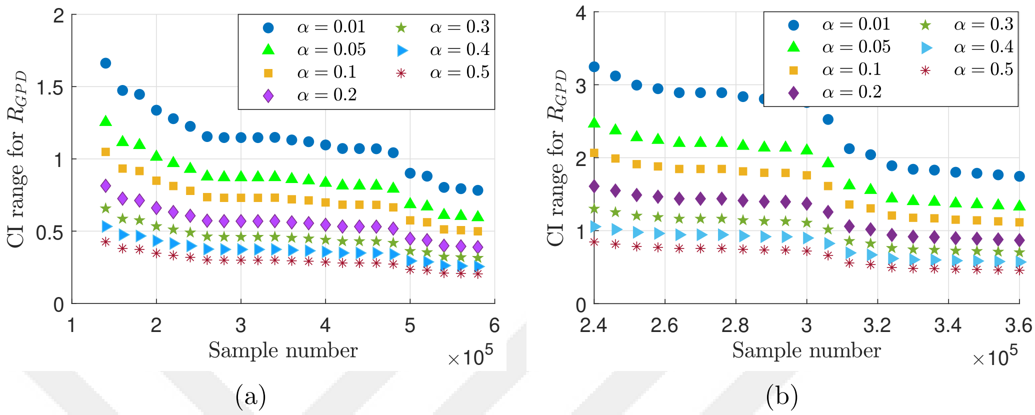


Figure 5.3: CI range for the estimated transmission rate considering different α values and different sample numbers for: (a) group 1, and (b) group 2.

Fig. 5.4 depicts the estimated transmission rate with the corresponding CI for $\alpha = \{0.01, 0.5\}$ and different sample numbers for groups 1 and 2. Same as the difference between Figs. 5.1 and 5.2, Fig. 5.3 illustrates the CI range while Fig. 5.4 shows the CI itself for the estimated transmission rate. The estimated transmission rate is constant with respect to the sample sizes. However, the confidence interval fluctuates with sample size, especially for lower α values. By incorporating a substantial large sample number in the training phase in both groups 1 and 2, the lower and upper boundaries of the confidence interval for the transmission rate become constant, and so the estimated transmission rate stays within an approximately fixed interval.

5.6 Conclusions

In this chapter, we present a novel EVT-based methodology for calculating the transmission rate of a system operating in the ultra-reliable regime within its ap-

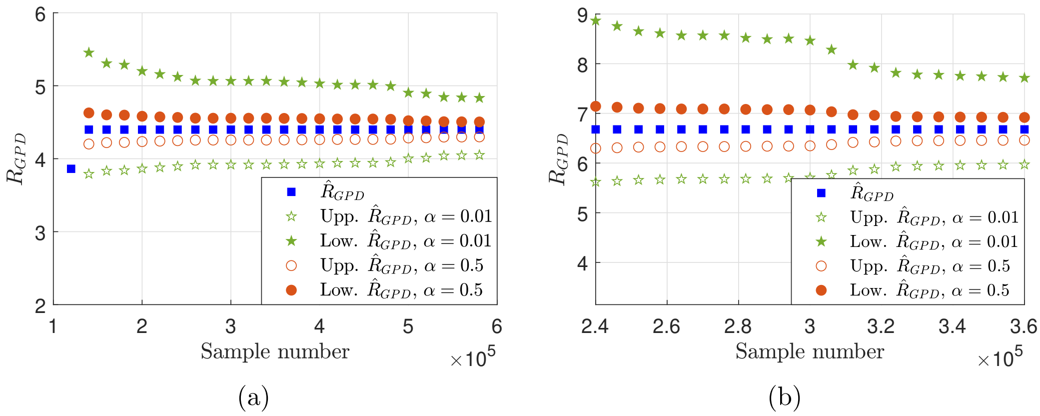


Figure 5.4: Estimated transmission rate with the corresponding CI at $\alpha = 0.5, 0.01$ and different sample numbers for: (a) group 1, and (b) group 2. Blue plot is the estimated parameter; circle/pentagram corresponds to $\alpha = 0.5/0.01$; filled/empty plot corresponds to upper/lower bound of CI.

propriate confidence interval. First, we estimate the generalized Pareto distribution (GPD) parameters fitted to the channel tail and then, we use these estimates along with their confidence intervals, to compute the transmission rate and its accompanying confidence interval. The incorporation of CI results into rate selection estimate reduces sample complexity by reducing the number of samples necessary for rate estimation in the training phase to attain a specific level of reliability. Furthermore, calculating the GPD parameters with few samples results in a wider CI in the rate estimation, indicating considerable uncertainty, but the uncertainty decreases as the sample number increases. This emphasize the trade-off between the GPD parameters uncertainty, controlling the uncertainty in the rate selection, and the cost of data collection.

Chapter 6

MULTIVARIATE EXTREME VALUE THEORY BASED CHANNEL MODELING FOR ULTRA-RELIABLE COMMUNICATIONS

The content of this chapter has been submitted for possible publication to IEEE Transactions on Wireless Communications (TWC).

6.1 Overview

Ultra-reliable low latency communication (URLLC) as one of the most important features of the fifth generation (5G) and beyond networks requires ultra-reliability of 10^{-9}-10^{-5} packet error rate (PER) with ultra-low latency on the order of milliseconds. Satisfying the ultra-reliability constraint requires deriving the statistics of the channel in the ultra-reliable region by modeling the extreme events happening rarely. Extreme value theory (EVT) has been previously adopted in channel modeling strategies to characterize the lower tail of the received powers in URLLC systems. In this chapter, we propose a multivariate EVT (MEVT)-based channel modeling methodology for the tail of the joint probability distribution of the multi-channel data by characterizing the multivariate extremes of a multiple input multiple output (MIMO) system with spatial diversity operating in URLLC. Spatial diversity-based schemes are considered to meet the high reliability requirement of URLLC when communicating over fading channels by using multiple transmission antennas. The proposed approach first derives the lower tail statistics of the received signal power for each channel based on the uni-variate EVT (UEVT) by using the generalized

Pareto distribution (GPD). Then, upon estimating the parameters of the GPDs of each individual channel, the tail of the joint distribution is modeled as a function of estimated GPD parameters by means of two approaches: logistic distribution and Poisson point process approaches. Finally, the validity of the proposed models is assessed by incorporating the mean constraint on the probability measure function of the Pichanks pseudo-polar coordinates. Based on the data collected within the engine compartment of Fiat Linea under various engine vibrations and driving scenarios, we demonstrate the capability of the proposed methodology in providing the best fit to the multivariate extremes, significantly outperforming the conventional extrapolation-based methods.

6.2 Introduction

Ultra-reliable low latency communication (URLLC) is one the most important features of the fifth generation (5G) and beyond networks with the goal of supporting mission critical applications including remote surgery, virtual reality (VR), remote control of robots, intra/inter vehicular communication, and tele-operated driving [Popovski et al., 2018], [Gustafson et al., 2014], [Popovski et al., 2019], [Eggers et al., 2019]. Strict requirements of a system operating in a URLLC necessitates to achieve ultra reliability of 10^{-9}-10^{-5} packet error rate (PER) while the acceptable latency is on the order of milliseconds or less [Popovski et al., 2018], [Gustafson et al., 2014], [Popovski et al., 2019], [Eggers et al., 2019], [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b]. Fulfillment of the reliability requirements at URLLC necessitates fundamental breakthrough in the derivation of the channel lower tail distribution by using novel statistical tools, such as extreme value theory (EVT) [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b], power law expression based on the extrapolation of the channel data [Eggers et al., 2019], [Angjelichinoski et al., 2019], and data-driven learning framework [Samarakoon et al., 2020].

Different forms of diversity techniques including time diversity [Serror et al., 2015a], frequency diversity [Pocovi et al., 2015a], space diversity [Pocovi et al., 2015b], and interface diversity [Nielsen et al., 2018a] can be applied in URLLC to reduce the required signal-to-noise-ratio (SNR) for achieving a certain reliability. Among the aforementioned diversity techniques, spatial diversity is the most viable solution to achieve high reliability over fading channels by utilizing multiple transmission antennas [Khosravirad et al., 2020]. Consequently, towards achieving ultra reliability in spatial diversity, the key is to obtain a proper multi-channel modelling approach that characterizes the statistics of multivariate extreme events for the systems with multiple input multiple output (MIMO) connections [Coles et al., 2001].

Current studies on the statistical modeling of the wireless channel operating in the ultra-reliable communication region are categorized into four folds: *a*) Extrapolation of a wide range of commonly used average statistics based wireless channel models such as Rayleigh and Rician towards the ultra-reliability region [Eggers et al., 2019], [Angjelichinoski et al., 2019]; *b*) Recommendation of the usage of new channel parameters by incorporating extreme reliability requirements into the communication [Swamy et al., 2018], [Swamy et al., 2019]; *c*) Usage of a non-parametric statistical learning algorithm to estimate the probability density function (PDF) of the channel distribution [Zhang et al., 2021]; and *d*) Application of the EVT to characterize the statistics of the channel tail distribution by deriving the statistics of extreme events happening rarely [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b], [Mehrnia and Coleri, 2021a], [Mehrnia and Coleri, 2022]. In [Eggers et al., 2019], a simple power-law expression is proposed to extrapolate the cumulative distribution function (CDF) of the commonly used average statistics channels to the ultra reliable regime of operation. In [Angjelichinoski et al., 2019], two performance measures, average reliability (AR) for dynamic environments and probably correct reliability (PCR) for the static environments, have been proposed based on the extrapolation-based channels in [Eggers et al., 2019] with the goal of selecting a transmission rate guaranteeing ultra-reliability. However, these extrapolated distributions cannot accurately estimate the lower tail distribution, and therefore, results in several orders of magnitude difference in the estimated PER [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b]. In [Swamy et al., 2018], [Swamy et al., 2019], the authors have modified the definitions of coherence time and distance to the time and distance over which a channel is predictable with a given reliability, respectively. Nevertheless, none of these frameworks derives ultra-reliability statistics by proposing a channel modeling methodology. In [Zhang et al., 2021], the authors propose a non-parametric statistical learning algorithm that utilizes the kernel density estimation (KDE) to estimate the PDF of channel distribution that meets the requirements of ultra-reliability. However, non-parametric estimations are strongly affected by amount of the data in the training set. Only recently in [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b], we have proposed a novel channel modeling methodology based on EVT to derive the statistics of the lower tail distribution while efficiently dealing with massive amount of data. EVT is a unique and robust framework to develop techniques for modelling the statistics of rare events based on the implementation of mathematical limits as finite-level approximation [Eggers et al., 2019], [Coles et al., 2001], [Kotz and Nadarajah, 2000]. In [Mehrnia and Coleri, 2021c], we have used EVT techniques to determine the optimum threshold below which the received power samples are considered extreme events and included in the tail distribution, then, model the channel tail distribution by using the generalized Pareto distribution (GPD), and finally, assess the validity of the proposed EVT-based model by using the probability plots. Additionally, in [Mehrnia and Coleri, 2021b], we have extended our work, [Mehrnia and Coleri, 2021c], to model the tail of the non-stationary channel distribution by first determining an external factor causing non-stationarity and accordingly, dividing the channel data into multiple stationary sequences where applying the EVT techniques initially presented in [Mehrnia and Coleri, 2021c] is feasible to model the parameter of the fitting distribution as a change-point function of time. However, none of these studies consider the usage of diversities in an ultra-reliable communication system to study the inter-relationships of extreme events.

Ultra-reliability can be achieved through different forms of diversity techniques including time diversity [Serror et al., 2015b], frequency diversity [Boyd et al., 2019], interface diversity [Nielsen et al., 2018a], or spatial diversity [Eggers et al., 2019], [Swamy et al., 2019], [Wallace and Jensen, 2002], [Zheng et al., 2014], [Almers et al., 2007]. In [Serror et al., 2015b], a combination of the stochastic geometry and the queuing analysis is proposed to derive the URLLC achievable ratio through the concept of effective bandwidth which transforms the constraints of delay and reliability into the constraint of achievable transmission rate. The authors in [Boyd et al., 2019] explore the diversity aspects of random access schemes where users transmit over multiple orthogonal Rayleigh fading subchannels and are treated as diversity branches by a maximum-ratio combining (MRC) receiver. The authors in [Nielsen et al., 2018a] consider a trade-off between transmission latency and reliability by allocating coded fragments of the encoded payload message to different interfaces, according to their bit-rate, latency and reliability properties. [Wallace and Jensen, 2002] proposes a statistical model by incorporating the time of arrival (TOA), angle of arrival (AOA), and angle of departure (AOD) of each multi-path component to estimate the joint magnitude and phase PDFs of measured data for realistic model parameters. A survey presented in [Zheng et al., 2014] investigates the main properties of massive MIMO channels based on different configurations of an antenna array, which directly affect the channel models and system performance. On the other hand, [Almers et al., 2007] provides a survey of the most important concepts in propagation channel modeling for MIMO systems by classifying the channels into physical models focusing on double-directional propagation and analytical models considering the channel impulse response and antenna properties. Nevertheless, an important building block of an ultra-reliable wireless system is a model of the wireless channel at the physical layer that captures the statistics of rare events and large fading dips [Eggers et al., 2019], [Swamy et al., 2019]. The authors in [Swamy et al., 2019] use a standard spatially independent quasi-static Rayleigh fading model of wireless channels to examine channel dynamics in the URLLC context. In addition, the authors in [Eggers et al., 2019] apply the power law results to analyze the performance of receiver diversity schemes and obtain a new simplified expression for maximum ratio combining applicable for an ultra-reliable communication. However, the channels have been estimated based on the extrapolation of the average statistics channel models. Also, the channels are assumed independent and non-identically distributed (i.n.i.d.), which is not a realistic assumption in reality. Therefore, there is a lack of study focusing on the channel modeling strategies in the ultra reliable regime of operation by discovering the inter-relation of extreme events. Multivariate extreme value theory (MEVT) is a robust statistical discipline that develops techniques to model the relation of rare events based on the multidimensional limiting relations [Eggers et al., 2019], [Coles et al., 2001], [Kotz and Nadarajah, 2000].

EVT has been used at the data link and network layers to model the tail statistics of queue length and delay [Samarakoon et al., 2018], [Liu and Bennis, 2018], [Abdel-Aziz et al., 2019], and derive closed-form asymptotic expressions for the throughput, bit error rate (BER), and PER over different fading channels [Mahmood and Jäntti, 2016], [Song and Li, 2006], [Al-Badarneh et al., 2018a], [Zhu et al., 2021]. Moreover, EVT has been used to derive the distribution of the maximum end-to-end SNR in an opportunistic relay selection-based cooperative relaying network consisting of a large number of independent and non-identical relay links [Subhash and Kalyani, 2021]. Additionally, EVT and federated learning (FL) approaches have been combined in

[Samarakoon et al., 2019] to propose a Lyapunov-based distributed transmit power and resource allocation procedure for vehicular users (VUEs) while the statistics of the queue lengths exceeding a high threshold are characterized by using the generalized Pareto distribution (GPD). However, these upper layer studies assume the average statistic-based channel models such as Rayleigh, Rician, or Nakagami fading, therefore, their usage may not be suitable in a system operating at URLLC [Eggers et al., 2019], [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b]. On the other hand, [Zhang et al., 2021] proposes a statistical learning algorithm to estimate the channel PDF, and accordingly, selects a proper transmission channel rate. However, the proposed data-driven non-parametric approach requires a massive number of training samples, about $10 \times \epsilon^{-1}$, to meet the targeted error probability ϵ. Until relatively recently, in [Mehrnia and Coleri, 2021a], [Mehrnia and Coleri, 2022], we have proposed a novel EVT-based framework dealing with relatively low number of data samples to estimate the optimal transmission rate and validate it by assessing the outage probability so that reliability constraints are met with a given confidence for ultra-reliable communications. Nevertheless, we have not considered spatial diversity by incorporating MEVT to address the reliability constraints in more than one dimension and assess the dependency between the extreme events.

The goal of this chapter is to propose a novel channel modeling methodology based on MEVT for a system using spatial diversity in MIMO-URLLC to derive the lower tail statistics of the received signal power in multiple dimensions while efficiently dealing with a massive amount of corresponding data. We focus on two-dimensional, or bi-variate case to highlight the main concepts and issues of MEVT without increasing the complexity of notation that is required for a full multivariate perspective. The modeling approach adapts MEVT to (i) fit GPD to the tail distribution of the received power samples exceeding a given threshold in each sequence and derive the scale and shape parameters of GPD while the thresholds are determined optimally; (ii) apply Fréchet transformation to each data sequence, so that the marginal distributions of the resulting variables are Fréchet; (iii) determine the dependency factor between the Fréchet transformed sequences; (iv) apply logistic distribution approach and Poisson point process approach to fit bi-variate GPD (BGPD) to the tail of the joint distribution of the Fréchet transformed sequences, considering the dependency factor; and (v) assess the validity of the fitted BGPD model by checking the mean constraint on the probability measure function. The original contributions of this chapter are listed as follows:

- We propose a comprehensive channel modeling methodology for a system operating at URLLC based on MEVT, for the first time in the literature. The methodology consists of techniques for determining the optimum threshold for each channel data sequence below which the tail statistics are derived, fitting Uni-variate GPD (UGPD) to each channel sequence, assessing the validity of the UGPD models, fitting BGPD to the tail of the joint probability distribution through the logistic distribution-based and Poisson point process-based approaches, and assessing the validity of these two proposed models by incorporating the probability measure function of the Pichanks coordinates.

- In the logistic distribution-based approach, for each channel data sequence obtained from different receivers, we separately determine the optimum threshold below which the samples are considered extremes. Then, upon applying the Fréchet transformation by incorporating the UGPD parameters and the corresponding thresholds fitted to each channel data sequence, we propose novel techniques based on MEVT to fit BGPD to the tail of the joint probability distribution by first determining the dependency factor among the Fréchet transformed sequences, and then, deriving a closed form expression for the BGPD model. Then, we determine the probability measure function of the

Pickands coordinates corresponding to the fitted BGPD model.

- In the Poisson point process-based approach, upon applying the Fréchet transformation, we propose techniques to represent the Fréchet transformed data by their pseudo-polar Pickands pairs, radial and angular components, and accordingly, determine the probability measure function of the Pickands angular component and the corresponding BGPD model.

- We assess the validity of the proposed BGPD models fitted to the joint distribution of the channel tail by checking if the mean constraint on the corresponding probability measure function of the Pickands coordinates obtained from BGPD approaches is satisfied.

- Based on the data collected within the engine compartment of Fiat Linea using one transmitter and two receivers under various engine vibrations and driving scenarios, we demonstrate the superiority of the proposed methodology for deriving the tail statistics of multivariate extremes compared to the conventional extrapolation-based models of the average statistics channels to the ultra-reliable region, in terms of the modeling accuracy.

The rest of this chapter is organized as follows. Section 6.3 describes the system model and assumptions considered throughout the chapter. Section 6.4 describes the basics of uni-variate and multivariate EVT together with the theorems used in the development of the proposed multivariate channel modeling approach. Section 6.5 presents the proposed channel modeling framework based on MEVT for characterizing the joint multi-channel tail distribution in the ultra-reliable region. Section 6.6 provides the channel measurement setup, and the performance evaluation of the proposed algorithm in determining the optimum threshold, fitting BGPD to the joint probability distribution of the extremes, and comparing the proposed methodology to the conventional methods in terms of the estimation accuracy. Finally, concluding remarks and future works are given in Section 6.7.

6.3 System Model

We consider multiple transmitter (Tx)-receiver (Rx) pairs communicating over a stationary channel, i.e., the parameters of the channel distribution are fixed over a vast period of time. If the channel is non-stationary according to the Augmented Dickey-Fuller (ADF) test results [Mehrnia and Coleri, 2021b], the external factors causing time variation in the parameters of the GPD are determined such that the sequence is divided into M groups, each of which can be considered stationary, as explained in detail in [Mehrnia and Coleri, 2021b]. The transmit power is assumed to be fixed and known in advance. Therefore, estimating the received signal power is equivalent to estimating the squared amplitude of the channel state information [Mehrnia and Coleri, 2021c], [Angjelichinoski et al., 2019].

To model the multivariate extremes of received powers, the sequences of measured received power samples at Tx-Rx pairs are converted into a sequence of i.i.d samples by removing their dependency via declustering approach [Mehrnia and Coleri, 2021c]. Upon applying EVT to the resulting sequence of i.i.d. samples in each pair, the optimum threshold is determined for each sample sequence. Determining an optimum threshold is of paramount importance as it determines the amount of samples that are included in the channel tail and considered as an extreme value. Then, the parameters of the Pareto distribution associated with the optimum threshold are estimated by using the maximum likelihood estimator (MLE). Finally, a multivariate version of GPD, i.e., a family with which to approximate a joint distribution on regions where we observe joint extreme values is obtained by using the MEVT.

6.4 Background

EVT is a robust framework that models the probabilistic distribution of extreme events occurring rarely. EVT has been applied in different fields, including hydrology to quantify risks of extreme floods, rainfalls and waves [Kotz and Nadarajah, 2000], and finance to estimate losses due to the extreme events [Finkenstadt and Rootzén, 2003]. However, it has been recently used in the telecommunication field to analyze the behavior of extreme values either in network traffic, worst-case delay, queue lengths, throughput and BER/PER of URLLC [Mehrnia and Coleri, 2021a], [Samarakoon et al., 2018], [Bennis et al., 2018], or in channel modeling to statistically derive the lower tail of the received power for URLLC system [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b]. In the following, Section 6.4.1 presents the uni-variate technique for modeling the extremes of a single process. Section 6.4.2 gives the general modeling technique of multivariate extremes to describe the extreme value inter-relationships of two or more processes. Sections 6.4.2 and 6.4.2 present bi-variate extreme value techniques based on the logistic family distribution and Poisson point process approach, respectively. Finally, Section 6.4.2 provides techniques to assess the validity of the proposed bi-variate extreme value modeling techniques.

6.4.1 Uni-variate Extreme Value Theory

Uni-variate EVT (UEVT) in general focus on the representation and modeling techniques for the extremes of a single process. UEVT is used for modeling extreme events in two main ways. The first concerns models for block maxima by using the generalized extreme value (GEV) distribution, given by

$$F(z) = \exp\Big\{ - \Big[1 + \xi\Big(\frac{z-\mu}{\sigma}\Big)\Big]^{-1/\xi} \Big\} \tag{6.1}$$

where μ, ξ, and σ are the location, shape, and scale parameters of the GEV distribution, respectively [Coles et al., 2001, Theorem 3.3]. The second uses EVT to characterize the tail of a distribution, i.e., the extremes of a sequence, by modeling the probabilistic distribution of values exceeding a given threshold through the generalized Pareto distribution (GPD).

Assume that $\{x_1, ..., x_N\}$ is an i.i.d. stationary sequence, where x_i denotes the i^{th} received power for $i \in \{1, ..., N\}$, and N is the total number of received power samples. Then, according to the EVT, the tail distribution of the power sequence, i.e., the probabilistic distribution of the power values exceeding a given threshold u, can be expressed as

$$G(l) = 1 - \left[1 + \frac{\xi l}{\tilde{\sigma}}\right]^{-1/\xi}, \tag{6.2}$$

where l is a non-negative value denoting the exceedance below threshold u, i.e., $(l = u - X)$, X denotes any x_i below threshold u; $G(l)$ expresses the GPD; and ξ and $\tilde{\sigma} = \sigma + \xi(u - \mu)$ are shape and scale parameters of the GPD, respectively. It should be noted that, μ and σ are the location and scale parameters of the GEV distribution fitted to the CDF of $m_N = min\{x_1, ..., x_N\}$, respectively [Mehrnia and Coleri, 2021c, Theorem 1], [de Zea Bermudez and Kotz, 2010].

6.4.2 Bi-variate Extreme Value Theory

When studying the extremes of two or more processes, each individual process can be modeled by using uni-variate techniques. However, the possible dependency between the extreme events requires the investigation of their joint behavior. Bi-variate EVT (BEVT) allows us to estimate the probability of exceeding thresholds simultaneously, and analyze the inter-dependence of two variables in the extreme value region [Joe et al., 1992].

In the following, we provide Theorems and Corollaries that are required to develop our channel modeling methodology. Theorem 4 incorporates BEVT applications to investigate the general form of BGPD models that are valid to estimate the bi-variate tail distribution. Then, based on Theorem 4, we obtain the logistic family distribution-based BGPD and Poisson point process-based BGPD models in Theorems 5 and 6, respectively. Also, we define the Fréchet transformation in Definition 1 as the input of BEVT in Theorem 4 is required to be Fréchet distributed. Additionally, the definition of Pseudo-polar Pickands coordinate transformation is provided in Definition 2 to be used in the Poisson point process approach in Theorem 6. Moreover, the constraints on the Pickands coordinates are defined in Definition 3 which will be used to estimate the probability density function and the probability measure function of the BGPD model in Theorem 5. Finally, Theorems 7 and 8 express the requirements of a valid BGPD model based on the 0.5 mean constraint on the probability measure functions of the Pickands coordinates corresponding to the BGPD models obtained in Theorems 5 and 6, respectively.

Definition 1 (Fréchet transformation). *Suppose $(x_1, y_1), ..., (x_n, y_n)$ are independent realizations of a random variable (X, Y) with joint distribution $F(x, y)$. For optimum thresholds u_x and u_y, each marginal distribution of F has an approximation of the form (6.2), with the parameters $(\tilde{\sigma}_x, \xi_x)$ and $(\tilde{\sigma}_y, \xi_y)$ for X and Y, respectively. Let's apply Fréchet transformation to variables X and Y to induce new variables $\tilde{X}$ and $\tilde{Y}$, given by*

$$\tilde{X} = -\Big(\log\Big\{1 - \zeta_x\Big[1 + \frac{\xi_x(u_x - X)}{\tilde{\sigma}_x}\Big]^{-1/\xi_x}\Big\}\Big)^{-1}, \qquad X < u_x, \tag{6.3}$$

and

$$\tilde{Y} = -\Big(\log\Big\{1 - \zeta_y\Big[1 + \frac{\xi_y(u_y - Y)}{\tilde{\sigma}_y}\Big]^{-1/\xi_y}\Big\}\Big)^{-1}, \qquad Y < u_y. \tag{6.4}$$

where $\zeta_x = Pr(X < u_x)$ *and* $\zeta_y = Pr(Y < u_y)$. *Then, the joint distribution function of* $\tilde{x}$ *and* $\tilde{y}$, $\tilde{F}$, *has margins that are approximately standard Fréchet distribution, i.e.,* $\tilde{F}(\tilde{x}) = exp(-1/\tilde{x}), \tilde{x} > 0$, $\tilde{F}(\tilde{y}) = exp(-1/\tilde{y}), \tilde{y} > 0$, *where* $\tilde{x}$ *and* $\tilde{y}$ *are any realizations from* $\tilde{X}$ *and* $\tilde{Y}$, *respectively [Coles et al., 2001].*

Theorem 4. *Let us define* $M_n^* = (M_{\tilde{x},n}^*, M_{\tilde{y},n}^*)$, *where* $M_{\tilde{x},n}^* = \max_{i=1,\ldots,n}\{\tilde{X}_i\}/n$, *and* $M_{\tilde{y},n}^* = \max_{i=1,\ldots,n}\{\tilde{Y}_i\}/n$, *and* $(\tilde{X}_i, \tilde{Y}_i)$ *are independent vectors with standard Fréchet marginal distributions as expressed in (6.3) and (6.4), respectively, and* n *is the number of realizations in the tail of* X *or* Y. *Then,*

$$Pr\{M_{\tilde{x},n}^* \le \tilde{x}, M_{\tilde{y},n}^* \le \tilde{y}\} \xrightarrow{d} G(\tilde{x}, \tilde{y}), \tag{6.5}$$

where $\xrightarrow{d}$ *denotes limit of distribution, and* G *is a bi-variate non-degenerate distribution function with the form*

$$G(\tilde{x}, \tilde{y}) = \exp\{-V(\tilde{x}, \tilde{y})\}, \qquad \tilde{x} > 0, \tilde{y} > 0 \tag{6.6}$$

$$V(\tilde{x}, \tilde{y}) = 2\int_0^1 \max(\frac{\omega}{\tilde{x}}, \frac{1-\omega}{\tilde{y}})H(d\omega), \tag{6.7}$$

and H *is a distribution function on* $[0, 1]$ *satisfying the mean constraint*

$$\int_0^1 \omega H(d\omega) = 1/2. \tag{6.8}$$

Any family of distributions that arise as limits in (6.5) can be considered the class of bi-variate extreme value distributions [Coles et al., 2001], [Zheng et al., 2018].

Theorem 4 implies that the class of bi-variate extreme value distributions is in one-to-one correspondence with the set of distribution functions H on $[0, 1]$ satisfying

(6.8). If the probability measure H is differentiable with probability density h, integral (6.7) is simplified to

$$V(\tilde{x}, \tilde{y}) = 2 \int_0^1 \max(\frac{\omega}{\tilde{x}}, \frac{1-\omega}{\tilde{y}}) h(\omega) d\omega. \tag{6.9}$$

This assumes that the thresholds u_x and u_y are small enough to justify the limit (6.5) as an approximation [Coles et al., 2001].

Theorem 4 also implies that for generating a class of BGPD models expressed in (6.6), it is required to obtain a parametric family for H over $[0, 1]$ whose mean is 0.5 for every value of the parameter. However, in practice, it is hard to find parametric families whose mean is parameter-free and for which, the integral (6.9) is tractable. Two approaches that can be utilized to model the bi-variate extreme value distribution are the logistic family and Poisson point process, which will be discussed in the following.

BGPD Based on Logistic Family:

Theorem 5. *Let $(\tilde{X}, \tilde{Y})$ be independent vectors with standard Fréchet marginal distributions as expressed in (6.3) and (6.4). The logistic family $G(\tilde{x}, \tilde{y})$ is a standard class expressing BGPD of $\tilde{X}$ and $\tilde{Y}$ as:*

$$G_l(\tilde{x}, \tilde{y}) = \exp\big\{ -\big(\tilde{x}^{-1/\alpha} + \tilde{y}^{-1/\alpha}\big)^{\alpha}\big\}. \tag{6.10}$$

where $\alpha \in (0, 1)$ denotes the dependency factor between variables $\tilde{x}$ and $\tilde{y}$ with the constraint $\tilde{x}, \tilde{y} > 0$ [Coles et al., 2001], [Zheng et al., 2018].

Proof. Let the parametric family of probability measure $H(\omega)$ have the probability density function $h(\omega)$ given by [Kotz and Nadarajah, 2000]

$$h(\omega) = \frac{1}{2}(\alpha^{-1} - 1)\{\omega(1-\omega)\}^{-1-1/\alpha}\{\omega^{-1/\alpha} + (1-\omega)^{-1/\alpha}\}^{\alpha-2}, \tag{6.11}$$

on $0 < \omega < 1$, where the $\frac{1}{2}$ mean constraint in (6.8) is automatically satisfied due to the symmetry around $\omega = \frac{1}{2}$. Upon calculating the integral in (6.7), $V(\tilde{x}, \tilde{y})$ is given by

$$V(\tilde{x}, \tilde{y}) = \big(\tilde{x}^{-1/\alpha} + \tilde{y}^{-1/\alpha}\big)^{\alpha}, \tag{6.12}$$

where $\tilde{x}$ and $\tilde{y}$ are expressed as (6.3) and (6.4), respectively, for large enough thresholds u_x and u_y. Consequently, the bi-variate extremes can be modeled as $G_l(\tilde{x}, \tilde{y})$ [Coles et al., 2001], [Kotz and Nadarajah, 2000].

The variables $\tilde{x}$ and $\tilde{y}$ of the logistic distribution family in (6.10) are exchangeable and have correlation $\rho = 1 - \alpha^2$ [Kotz and Nadarajah, 2000]. The α value very close to 1 denotes strong dependence between the variables x and y, even at moderately extreme levels.

BGPD Based on Poisson Point Process:

Definition 2 (**Pickands coordinates**). *Let $(\tilde{X}, \tilde{Y})$ be independent vectors with standard Fréchet marginal distributions as expressed in (6.3) and (6.4). Let's define the transformation $T_P(\tilde{x}, \tilde{y})$ by*

$$T_P(\tilde{x}, \tilde{y}) := \big(\frac{-\tilde{x}}{n} + \frac{-\tilde{y}}{n}, \frac{-\tilde{x}/n}{-\tilde{x}/n - \tilde{y}/n}\big) =: (\omega, r).$$

where $n \in N$, is the length of vector $\tilde{X}$ or $\tilde{Y}$. $T_P(\tilde{x}, \tilde{y})$ is called transformation with respect to the Pickands coordinates (ω, r), where ω is pseudo-polar angular component measuring angle in $[0, 1]$ scale, and r is pseudo-polar radial component measuring the distance from the origin [Coles et al., 2001]. This transformation is one-to-one with the inverse

$$T_P^{-1}(\omega, r) =: r\big(\omega, 1 - \omega\big) =: (\tilde{x}, \tilde{y}).$$

The pseudo-polar Pickands mapping has the same geometrical interpretation as standard polar coordinates with the difference that polar coordinates use the Euclidian norm for the angular and radial component, while the Pickands coordinates use the sum norm [Michel, 2006], [Falk et al., 2010].

Theorem 6. *Assume that $(\tilde{X}, \tilde{Y})$ be a vector of independent bi-variate observations with standard Fréchet margins satisfying (6.5). Let N_n denote a sequence of point processes defined as*

$$N_n = \{(n^{-1}\tilde{x}_1, n^{-1}\tilde{y}_1), ..., (n^{-1}\tilde{x}_n, n^{-1}\tilde{y}_n)\}. \tag{6.13}$$

where $\tilde{x}_i$ and $\tilde{y}_i$ are the i^{th} realization of $\tilde{X}$ and $\tilde{Y}$, respectively. Then, N_n converges to N, i.e., $N_n \xrightarrow{d} N$, where N is a non-homogeneous Poisson process on space A defined by

$$A = \{(0, \infty) \times (0, \infty) \backslash (0, \tilde{x}) \times (0, \tilde{y})\},$$

denoting space $\{(0, \infty) \times (0, \infty)\}$ excluding sub-space $\{(0, \tilde{x}) \times (0, \tilde{y})\}$. Therefore, according to Poisson point process limit, the bi-variate extremes are modeled as

$$G_{pp}(\tilde{x}, \tilde{y}) \rightarrow Pr\{N(A) = 0\} = \exp\{-\Lambda(A)\}, \tag{6.14}$$

where the intensity measure of Poisson point process, $\Lambda(A)$, given by [Coles et al., 2001], [Falk et al., 2010]

$$\Lambda(A) = 2 \int_{\omega=0}^{1} \max\big(\frac{\omega}{-\tilde{x}/n}, \frac{1-\omega}{-\tilde{y}/n}\big)\, H_{pp}(d\omega), \tag{6.15}$$

has the same form as function $V(\tilde{x}, \tilde{y})$ in (6.7), and the probability measure function $H_{pp}(d\omega) = \int_0^1 h_{pp}(\omega)\, d\omega$ with the probability density function $h_{pp}(\omega)$ of the angular *component ω based on Definition 2 [Coles et al., 2001], [Joe et al., 1992], [Falk et al., 2010].*

Proof. Let r and ω denote the Pickands coordinates of $\tilde{x}$ and $\tilde{y}$ given by

$$r = \frac{-\tilde{x}}{n} + \frac{-\tilde{y}}{n}, \qquad \textit{and} \qquad \omega = \frac{-\tilde{x}/n}{r}, \tag{6.16}$$

where $\tilde{x}$ and $\tilde{y}$ have standard marginal Fréchet distributions. Then, the intensity function of N on space A in Theorem 6 is [Coles et al., 2001], [Joe et al., 1992], [Falk et al., 2010]

$$\lambda(r, w) = 2\frac{dr}{r^2}\, H_{pp}(d\omega). \tag{6.17}$$

Additionally, let N_n be the point process defined as (6.13) with intensity function $\lambda(r, \omega)$ in (6.17). Then, $\Lambda(A)$ in Theorem 6 is given by

$$\Lambda(A) = \int_A 2\frac{dr}{r^2}\, H_{pp}(d\omega) \; = \int_{\omega=0}^{1}\int_{r_{min}}^{\infty} 2\frac{dr}{r^2}\, H_{pp}(d\omega),$$

where r_{min} is the minimum of $\frac{-\tilde{x}/n}{\omega}$ and $\frac{-\tilde{y}/n}{1-\omega}$, depending on how r is defined based on $\omega = \frac{-\tilde{x}/n}{r}$ or $1 - \omega = \frac{-\tilde{y}/n}{r}$. Therefore,

$$\begin{aligned}\Lambda(A) &= \int_{\omega=0}^{1}\int_{r=\min\{\frac{-\tilde{x}/n}{\omega}, \frac{-\tilde{y}/n}{1-\omega}\}}^{\infty} 2\,\frac{dr}{r^2}\, H_{pp}(d\omega) \\ &= 2\int_{\omega=0}^{1} \max\big(\frac{\omega}{-\tilde{x}/n}, \frac{1-\omega}{-\tilde{y}/n}\big)\, H_{pp}(d\omega).\end{aligned} \tag{6.18}$$

Remark 1. *The Pickands probability measure $H_{pp}(\omega)$ has the following property [Falk et al., 2010], [Capéraà and Fougères, 2000], [Cooley et al., 2010], [Nadarajah,* *2000]:*

$$\int_{[0,1]} \omega \, H_{pp}(d\omega) = \int_{[0,1]} (1-\omega) \, H_{pp}(d\omega). \tag{6.19}$$

The fact expressed in Remark 1 will be used in Theorem 8 to assess the validity of the Poisson point process-based BGPD.

BGPD Model Assessment:

Any distribution function H defined on the space $[0,1]$ in (6.7) that satisfies the mean constraint in (6.8), gives rise to a valid limit in (6.5) [Coles et al., 2001]. Therefore, if $G(\tilde{x},\tilde{y})$ is a valid model to estimate the tail of bi-variate extremes, its corresponding probability measure function $H(\omega)$ should satisfy the 0.5 mean constraint according to Theorem 4.

Definition 3 (**Pickands constraints**). *Let $(\tilde{X},\tilde{Y})$ be independent vectors with standard Fréchet marginal distributions as expressed in (6.3) and (6.4), and corresponding Pickands coordinates (r,ω). Then, for a radial cut-off point $r_0 < 0$ close to 0, we have the following Pickands constraints [Michel, 2006]:*

1. *Conditional on $r > r_0$, the radial and angular components of the Pickands coordinates (ω, r), are independent.*

2. *Conditional on $r > r_0$, the radial component of the Pickands coordinates is uniformly distributed. Therefore,*

$$Pr(r \geq R | r > r_0) = \frac{R}{r_0}, \qquad r_0 \leq R \leq 0. \tag{6.20}$$

Theorem 7. *If the BGPD model based on the logistic distribution family $G_l(\tilde{x},\tilde{y})$, given by (6.10), is a valid model to estimate the tail of the bi-variate extremes, the* *mean constraint defined in (6.8) should be satisfied on* $H_l(\omega) = \int h_l(\omega)\, d\omega$*, where*

$$h_l(\omega) = \frac{\phi(\omega)}{\mu}, \tag{6.21}$$

$\mu = \int \phi(\omega)\mathrm{d}\omega > 0$*, and* $\phi(\omega)$ *is the Pickands density function given by*

$$\phi(\omega) = |r|\Big(\frac{\partial^2}{\partial\tilde{x}\partial\tilde{y}} G_l\Big)\big(T_P^{-1}(\omega, r)\big), \qquad r > r_0, \tag{6.22}$$

where G_l *is the BGPD expressed as (6.10) and* r_0 *is the optimum cut-off point obtained based on the constraints on the Pickands coordinates in Definition 3.*

Proof. Assume $(\tilde{X}, \tilde{Y})$ follow a BGPD expressed as $G_l(\tilde{x}, \tilde{y})$ in (6.10) whose standard Pickands coordinates are given by $r := -(n^{-1}\tilde{x} + n^{-1}\tilde{y})$ and $\omega := (-n^{-1}\tilde{x}/r)$. Let us define the Pickands density as [Michel, 2006]

$$\phi(\omega, r) = |r|\Big(\frac{\partial^2}{\partial\tilde{x}\partial\tilde{y}} G_l\Big)\big(T_P^{-1}(\omega, r)\big), \qquad r > r_0.$$

Then, since ω is also a function of r, $\phi(\omega, r)$ depends only on ω and therefore, $\phi(\omega) = \phi(\omega, r)$ [Falk et al., 2010]. Also, conditional on $r > r_0$, the Pickands angular component ω has the probability density [Michel, 2006]

$$h_l(\omega) = \frac{\phi(\omega)}{\mu},$$

where $\phi(\omega)$ is the Pickands density function depending on the logistic-based $G_l(\tilde{x}, \tilde{y})$ as shown in (6.22), and $\mu = \int \phi(\omega)\mathrm{d}\omega > 0$ [Michel, 2006]. Remark that $\phi(\omega)$ is the density of a *probability* measure only after the division by μ. Otherwise, it is the density of a measure, which assigns the mass μ to the defined space of r. Additionally, the optimal cut-off point r_0 is determined such that the constraints on the Pickands coordinates (r, ω) in Definition 3 are satisfied.

Theorem 8. *If the BGPD model $G_{pp}(\tilde{x}, \tilde{y})$ based on the Poisson point process approach given by (6.14) is a valid model to estimate the tail of the bi-variate extremes, the mean constraint defined in (6.8) is always satisfied for the $H_{pp}(\omega)$ function in (6.15).*

Proof. Let r and ω denote the Pickands coordinates of $\tilde{x}$ and $\tilde{y}$ based on Definition 2. Then, referring to (6.19), we have

$$2 \int_{[0,1]} \omega \, H_{pp}(d\omega) = \int_{[0,1]} H_{pp}(d\omega),$$

where $\int_{[0,1]} H_{pp}(d\omega) = 1$, and therefore, the 0.5 mean constraint defined in (6.8) is satisfied for the probability measure function $H_{pp}(\omega)$.

6.5 Proposed Channel Modeling Methodology

We propose a novel BEVT-based channel modeling methodology with the goal of estimating the joint lower tail statistics of multiple channels in the ultra-reliable regime. The methodology consists of the following steps: The sequences of measured received power samples are converted into sequences of i.i.d samples by removing their dependency via declustering, where the samples are divided into multiple clusters each of which includes consecutive dependent observations and the clusters are separated by a certain sample gap to ensure the independency between clusters [Mehrnia and Coleri, 2021c]. Upon applying EVT to the resulting sequence of i.i.d. samples, the optimum thresholds are determined. Since the bi-variate analysis is restricted to the time intervals in which both thresholds are exceeded, we keep only the exceedances occur in the same time interval. Otherwise, we ignore the exceedance in any sequence. Thereafter, the parameters of the UGPD associated with the optimum thresholds are estimated by using the Maximum Likelihood (ML) estimator. Then, we assess the validity of the fitted UGPD model to the tail distribution by using the probability plots, including PP plot and QQ plot. Afterwards, we model the inter-relationship of bi-variate extremes based on Theorem 4 by applying two methods, logistic distribution and Poisson point process. In the logistic distribution approach, upon determining the dependency factor α between the Fréchet transformed variables in Definition 1, we model the tail of the joint PDF based on Theorem 5 and verify the model according to Theorem 7 by using the doubled-transformed data to the Pickands coordinate based on Definition 2. In Poisson point process approach, based on Theorem 6, we transform the data sequence in two steps: First, based on the Fréchet transformation in Definition 1, and second, based on Pickands coordinates in Definition 2. Then, we determine the probability measure of the Pickands angular coordinate $H_{pp}(\omega)$, and the intensity measure of Poisson point process $\Lambda(A)$. Finally, we asses the validity of the Poisson point process-based bi-variate model based on Theorem 8. The proposed algorithm for the bi-variate extremes is depicted in Fig. 6.1 and explained in detail next.

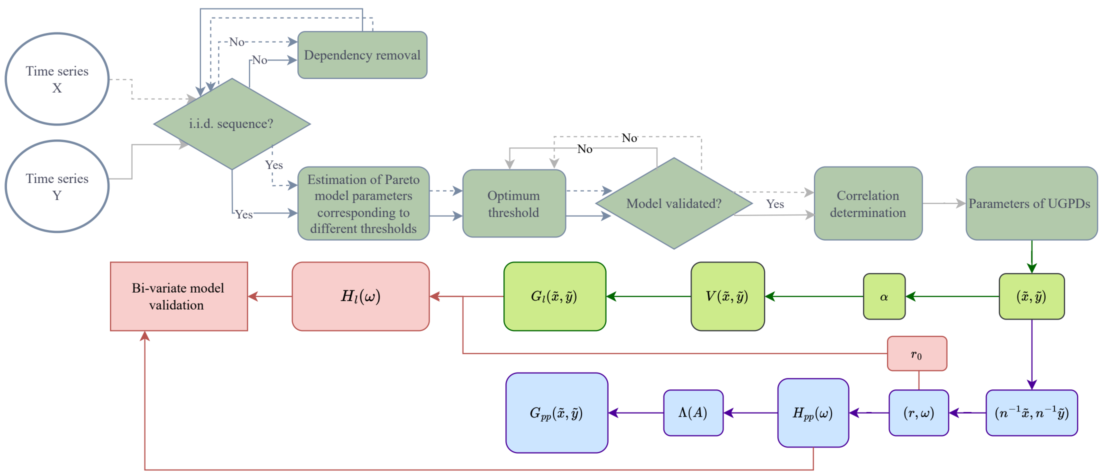


Figure 6.1: Flowchart of the proposed multi-channel modeling framework.

6.5.1 Modeling the Uni-variate Extremes

Declustering Approach:

In the declustering approach, first, we assume a low enough threshold u and start the first cluster by the first sample x_i below this threshold. Then, all the successive samples below threshold u followed by the first sample of the initial cluster are assigned to the first cluster. Right after observing a sample over u, we let the cluster to continue for mg more consecutive values and then, if no value below u is detected, close the cluster. The next cluster starts with the next value below the threshold u. Upon determination of the clusters of samples, we extract the minimum value of each cluster, apply EVT to the i.i.d. cluster minima, and model their tail distribution by using GPD. The declustering approach is based on the fact that the minimum values of the clusters are far enough to be considered independent and identically distributed [Mehrnia and Coleri, 2021c], [Coles et al., 2001].

Optimum Threshold Determination:

Optimum threshold of each uni-variate case is determined by using the mean residual life (MRL) and parameter stability methods. Based on the MRL method, u is the optimum threshold if u is the highest threshold below which the expected value of samples exceeding u is a linear function of u, i.e., $E(u - X|X < u)$ is linear against u. The remarkable advantage of the MRL method is the simplicity as it can be applied to the data sequence prior to the estimation of the UGPD parameters. However, due to the lower precision of MRL method compared to parameter stability method, it is sometimes difficult to obtain the optimum threshold explicitly. Therefore, MRL method is usually utilized as a complementary method or in the case that the optimum threshold determination does not require high precision. Parameter stability method states that the optimum threshold u is the highest threshold below which the estimated shape and modified scale parameters of the UGPD fitted to the tail distribution associated with a variety of thresholds is a linear function of the threshold. It is worth noting that the modified scale parameter is defined as $\sigma^* = \tilde{\sigma}_u - \xi u$. Additionally, the linearity relation is assessed by using the R-squared statistical measure, denoted by R^2 [Mehrnia and Coleri, 2021c], [Mehrnia and Coleri, 2021b].

Model Validity Assessment:

The validity of the GPD model is assessed by using probability plots, i.e., PP and QQ plots. These plots are graphical techniques used to assess the validity of the models fitted to the empirical values. In the PP plot, we compare the empirical probability of the occurrence for an extreme value with the corresponding probability obtained by the GPD, while in the QQ plot, we compare the empirical extreme quantile with the corresponding value obtained by the inverse of GPD [Mehrnia and Coleri, 2021c]. If the GPD appropriately models the extreme values exceeding threshold u, then, both PP and QQ plots fit the unit diagonal line, i.e., the $45°$ line [Coles et al., 2001], [Finkenstadt and Rootzén, 2003].

6.5.2 *Evaluation of the Correlation Coefficient*

The first step in the bi-variate extreme analysis is to assess the amount of dependence in the tails [Joe et al., 1992]. However, since the bi-variate analysis is restricted to those time intervals in which both thresholds u_x and u_y are exceeded, before determining the correlation among the tail samples, we are required to revise the tail samples by removing those exceedances happening in one sequence but not in the other one, at the same time interval [Coles et al., 2001]. Accordingly, we consider a time interval consisting of M samples and check if u_x and u_y are exceeded simultaneously in sequences X and Y, respectively. If so, we capture the minima of clusters; otherwise, we ignore the minima of each cluster [Coles et al., 2001], [Joe et al., 1992]. Then, EVT is applied on the obtained exceedances to determine the UGPD parameters.

Upon determining the extremes of each sample sequence, if the extremes of the two sequences are independent, no bi-variate modelling like (6.5) is required. Otherwise, the existence of correlation among the tail samples suggests the requirement of investigating the inter-relationship of the extreme values of receivers. On the other hand, correlation among the received powers in the total samples is used to assess the feasibility of the spatial-diversity. If the correlation coefficient among the total samples is a value between 0.1 and 0.5, the spatial diversity is suggested [Shi et al., 2012]. Otherwise, the employment of spatial-diversity is pointless due to the concurrent fading at different links. If the spatial diversity is feasible and the correlation among the tail samples is high enough, we are required to analyze the bi-variate tail characteristics for a system operating in URLLC.

6.5.3 Modeling the Multivariate Extremes Based on the Logistic Distribution Approach

Data Conversion Based on the Fréchet Transformation:

Applying the Fréchet transformation by using the Definition 1 on the obtained tail sequences X and Y from Section 6.5.1, we induce variables $\tilde{X}$ and $\tilde{Y}$ whose marginal distribution functions have Fréchet distribution for $x < u_x$ and $y < u_y$, approximately. This transformation is required as the input of the bi-variate extreme modelling approach is expected to have the Fréchet distribution [Coles et al., 2001], [Kotz and Nadarajah, 2000].

Determining the Dependency Factor α:

To determine the dependency parameter α of the logistic model in (6.12), we either estimate the correlation coefficient ρ between the variables $\tilde{x}$ and $\tilde{y}$ and then, calculate α as $\sqrt{1-\rho}$, or directly estimate it by MLE, where the log-likelihood function is defined as

$$l(\tilde{x}, \tilde{y}|\alpha) = \sum_{i=1}^{n} \ln\, v_\alpha(\tilde{x}_i, \tilde{y}_i), \tag{6.23}$$

where,

$$v_\alpha(\tilde{x}, \tilde{y}) = \frac{\partial^2}{\partial\tilde{x}\partial\tilde{y}} V(\tilde{x}, \tilde{y}), \tag{6.24}$$

and $V(\tilde{x}, \tilde{y})$ is the logistic distribution function defined in (6.12).

BGPD Model Based on the Logistic Distribution:

The bi-logistic family of distributions formulated in (6.12) is used to model the bi-variate extremes based on Theorem 4. Upon obtaining the dependency parameter α as well as the Fréchet transformed variables $\tilde{x}$ and $\tilde{y}$, we build function (6.12) and then, by taking the exponential of its negative function, we determine the BGPD as expressed in (6.10). If this BGPD model is reliable to characterize the tail of the bi-variate distribution, the mean constraint on the probability measure function defined in Theorem 7 should be satisfied.

6.5.4 Modeling the Multivariate Extremes Based on the Poisson Point Process Approach

Pseudo-polar Pickands Coordinates Transformation:

In Poisson point process-based approach for modeling the bi-variate extremes, it is required to first transform the data from Cartesian to the polar coordinate: $(\tilde{x}, \tilde{y}) \to (r, \omega)$. To this end, we apply pseudo-polar Pickands transformation based on Definition 2 that induces two new variables, radial component r and angular component ω.

Determining the Probability Measure Function of the Radial and Angular Components:

The CDF of the probability measure function of the angular component $H_{pp}(\omega)$ is determined based on Theorem 6. This probability measure will be used in the next step to determine the intensity of Poisson point process $\Lambda(A)$ and later, to assess the validity of the proposed Poisson point process-based BGPD, according to Theorem 8.

BGPD Model Based on the Poisson Point Process:

Upon determining the angular component probability measure function $H_{pp}(\omega)$, we compute the density function of the Poisson point process $\Lambda(A)$ for the defined space A based on Theorem 6. $\Lambda(A)$ is actually equivalent to $V(\tilde{x}, \tilde{y})$ in the logistic distribution based-BGPD model. Afterwards, the BGPD model based on the Poisson point process approach is determined as $\exp(-\Lambda(A))$.

6.5.5 BGPD Model Validation

Model Validation for Logistic-Based BGPD:

According to Theorem 7, if the logistic-based BGPD is a valid model to estimate the tail of the bi-variate extremes, the probability measure function $H_l(\omega)$, conditional on $\{\omega : r > r_0\}$, needs to satisfy the 0.5 mean constraint, i.e., $\int_\omega \omega\, H_l(d\omega) = 0.5$, where ω is the Pickands angular component of $\tilde{x}$ and $\tilde{y}$.

To determine r_0 based on Definition 3, a plot of r versus ω for $r < r_0$, referred to as the r-ω plot, can be utilized to address the first constraint in Definition 3 by checking the dependency between r and ω. The dependency of r and ω is assessed based on the correlation results in which, r and ω are independent if their correlation is less than the critical value 0.05. Additionally, to address the second constraint in

Definition 3, the distribution of the radial components $r_0 < r < 0$ is expected to be fitted to the uniform distribution, where $Pr(r > R|r > r_0) = \frac{R}{r_0}$.

Model Validation for Poisson Point Process-based BGPD:

According to Theorem 8, the Poisson point process-based BGPD is by default a valid model to estimate the tail of the bi-variate extremes as its intensity function $H_{pp}(\omega)$ satisfies the mean constraint. Therefore, the constraint $\int_{\omega} \omega \, H_{pp}(d\omega) = 0.5$ is insured in Poisson point process-based BGPD while determining $H_{pp}(\omega)$.

6.6 Numerical Results

The goal of this section is to evaluate the performance of the proposed channel modeling algorithm in estimating the BGPD model fitted to the joint distribution of the channel data sequences based on two approaches: logistic distribution approach and Poisson point process approach, and also comparing their performances with the traditional extrapolation-based approaches to estimate the statistics of the channel tail for a system operating in the spatial diversity MIMO-URLLC. In the traditional extrapolation-based approach, upon estimating the distribution of the existing channel data for the reliability order of 10^{-3}-10^{0} PER [Bas and Coleri Ergen, 2012], [Demir et al., 2013] for individual channel data sequences, we determine their joint probability distribution, and then, extrapolate it towards the ultra reliable region 10^{-9}-10^{-5} PER [Eggers et al., 2019].

We have collected channel measurement data within the engine compartment of Fiat Linea under various engine and driving scenarios at 60 GHz by using a Vector Network Analyzer (VNA) (R & S® ZVA67). The VNA is connected to the transmitter and receivers through the R & S® ZV-Z196 and PE361 port cables, respectively, with 610 mm length as shown in Fig 6.2. The transmitter is an omnidirectional antenna operating from 58 GHz to 63GHz with 0 dBi nominal gain. The receivers are horn antennas operating between 50-75 GHz with a nominal 24 dBi gain, and 11° and 9.5° horizontal and vertical half power beamwidth, respectively. The antennas are connected to the coax cables through the waveguide to coax adaptor operating at the frequency span of $50 - 65$ GHz, with insertion loss 0.5 dB and impedance 50 Ω.

The data were collected while driving the car at Koc University campus and emulating different scenarios including start/stop the car, moving up/down on a ramp, and driving on the flat road. More than 10^6 successive samples are captured from each receiver antenna for about 1.5 hour with time resolution of 3 ms. We use MATLAB® for analyzing the data and the implementation of the proposed algorithm as well as the traditional extrapolation-based approach [Eggers et al., 2019]. Meanwhile, the estimation error of the proposed methodology and the extrapolated-based approach have been reported by means of the Root Mean Square Error (RMSE) metric.

In the following, first in Section 6.6.1, we provide the numerical results in the determination of the optimum threshold over which the tail statistics are derived for two uni-variate channel data sequences obtained from receivers Rx1 and Rx2, and then, validate the tail model corresponding to the optimum threshold by means of the probability plots. Upon determining the correlation among the tail samples in Section 6.6.2, we apply the logistic distribution approach to model the tail distribution of the bi-variate extremes in Section 6.6.3. Next, in Section 6.6.4, we estimate the tail distribution of the bi-variate extremes by applying the Poisson point process approach. Afterwards, in Section 6.6.5, we assess the validity of the determined BGPD models based on logistic distribution and Poisson point process approaches. Finally, we compare the performance of both proposed methodologies, logistic distribution-based and Poisson point process-based approaches, to that of the traditional extrapolation-based technique in the estimation of the bi-variate channel tail statistics in Section 6.6.6.

Figure 6.2: Transmitter and receiver antennas in the engine compartment.

6.6.1 UGPD Model

Towards determining the optimum threshold of the GPD fitted to the tail distribution of samples in each group, declustering approach is used to remove the dependency among the samples and obtain i.i.d. observation for the EVT input. In this regard, MRL plot and parameter stability method are utilized to determine the thresholds u_x and u_y, and the minimum gaps(mgs) between the samples mg_1 and mg_2, for receiver Rx1 and Rx2, respectively. We skip the illustration of these results and only present the key information as the subject of the uni-variate channel tail modeling has been discussed in our paper [Mehrnia and Coleri, 2021c] in detail. According to the results, for the uni-variate channel data sequence obtained from receiver Rx1, $u_x = -15$ dBm is the optimum threshold below which, the $R^2 > 0.95$ for all $\mathrm{mg}_1 > 1$. Additionally, $u_y = -30$ dBm is the optimum threshold below which the linearity condition is observed for all $\mathrm{mg}_2 > 1$ for the uni-variate channel data of receiver Rx2. Upon determining the optimum thresholds (u_x and u_y) and minimum gaps mg_1 and mg_2 between the clusters, the probability plots, including PP plot and QQ plot, are used to validate the accuracy of the UGPD models fitted to the channel tail distribution of the i.i.d. samples obtained from receivers Rx1 and Rx2, corresponding to the optimum thresholds u_x and u_y, respectively.

6.6.2 Determination of Correlation

To determine the correlation among the tail samples, first, we have considered the time intervals consisting of 1000 samples, and then, at each interval, obtained those minima of Rx1 and Rx2 sequences that simultaneously exceed $u_x = -15$ dBm and $u_y = -30$ dBm, respectively. The parameters of UGPD fitted to these new exceedances are $(\xi_x, \tilde{\sigma}_x) = (-0.1469, 4.0367)$, and $(\xi_y, \tilde{\sigma}_y) = (-0.4245, 8.5886)$ for receivers Rx1 and Rx2, respectively. The correlation among the total samples obtained from Rx1 and Rx2 is about 0.2766, which confirms the applicability of spatial diversity as it is less than 0.5. Additionally, the 0.3957 correlation coefficient among the tail samples indicates the existence of correlation in the tail samples of Rx1 and Rx2, and therefore, confirms the necessity of studying the inter-relationships of the extreme values.

6.6.3 BGPD Based on the Logistic Distribution Approach

Fig. 6.3 illustrates the CDF of the transformed data by using the Fréchet transformation for the received power tail samples obtained from receivers Rx1 and Rx2. Figs. 6.3a and 6.3b show that after the transformations of the Rx1 and Rx2 tail samples, the Fréchet distribution is fitted well to the CDF of the empirical data. This confirms that $\tilde{x}$ and $\tilde{y}$ are the reliable transformed variables for building BGPD model.

Fig. 6.4 shows the proposed BGPD model determined based on the logistic distribution approach, $G_l(\tilde{x}, \tilde{y})$. Empirical joint CDF of the tail samples has also been depicted for better compatibility assessment of the logistic-based BGPD. The BGPD model obtained by the logistic distribution approach can estimate the empirical joint CDF with RMSE 0.8655. However, the exact pattern of the empirical CDF has not been preserved by logistic-based BGPD model.

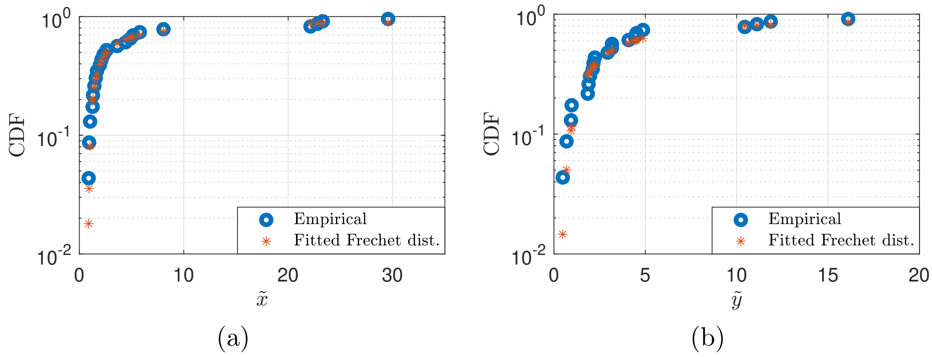


Figure 6.3: CDF of the received power tail samples transformed based on the Fréchet transformation: (a) for the transformed samples of receiver Rx1, $x \to \tilde{x}$, and (b) for the transformed samples of receiver Rx2, $y \to \tilde{y}$.

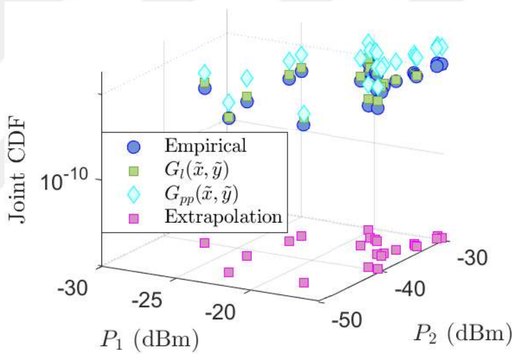


Figure 6.4: Joint bi-variate CDF of the normalized tail power of empirical, extrapolation-based, logistic-based BGPD, and Poisson point process-based BGPD.

6.6.4 BGPD Based on the Point Process Approach

Fig. 6.4 also depicts the proposed BGPD model determined based on the Poisson point process approach. It can be seen that the estimated joint CDF based on the Poisson point process approach is in a good agreement with the corresponding empirical CDF with RMSE of 0.8686. Although the Poisson point process-based approach for modeling the bi-variate tail distribution seems to have the same performance as for the logistic distribution-based approach according the RMSE results, it preserves the shape of the empirical joint CDF very well. This is mainly due to the higher complexity of the additional transformation step as well as integral computation for determining the density of Poisson point process $\Lambda(A)$.

6.6.5 Model Validity Assessment

Figs. 6.5 and 6.6 are utilized to determine the optimum cut-off point r_0 for which the Pickands constraints are satisfied and $H_l(\omega)$ is defined on $r > r_0$. Fig. 6.5 illustrates the r-ω plot of Poisson point process approach for determining the optimum cut-off point r_0. In this figure, the radial component of the Pickands coordinate is plotted versus the corresponding angular component, where r and ω are the pseudo-polar Pickands coordinates of $\tilde{x}$ and $\tilde{y}$. Please note that $\{(r,\omega)|r < r_0\}$ correspond to independent pairs and $\{(r,\omega)|r > r_0\}$ are matched to dependent pairs of r and ω. According to this plot, $r_0 \approx -0.47$ is the minimum radial component above which the correlation coefficient of Pickands radial and angular components is less than the critical value 0.05, resulting in the independent r and ω pairs. Therefore, $r_0 \approx -0.47$ is the cut-off radial component that satisfies the first Pickands constraint.

Fig. 6.6 shows the distribution of the radial component of the Pickands coordinate for $r > r_0$ where r is the pseudo-polar Pickands radial coordinate of $\tilde{x}$ and $\tilde{y}$, and $r_0 \approx -0.47$ is the optimum radial cut-off point captured from Fig. 6.5. It is observed that the distribution of the radial components $r > r_0$, i.e., those radial components independent from their angular component pairs, fits to the uniform distribution, which means that $P(r > R) = \frac{R}{r_0}$, and results in satisfying the second constraint on Pickands coordinates.

We assess the validity of the proposed BGPD models shown in Fig. 6.4 based on the CDF results of Pickands probability density functions $h_l(\omega)$ and $h_{pp}(\omega)$ obtained from the logistic-based for $r > r_0$, and Poisson point process-based BGPD approaches, respectively. Accordingly, it is observed that the CDF of the estimated density function of the angular Pickands coordinate ω, for $h_l(\omega)$, the 0.5 mean

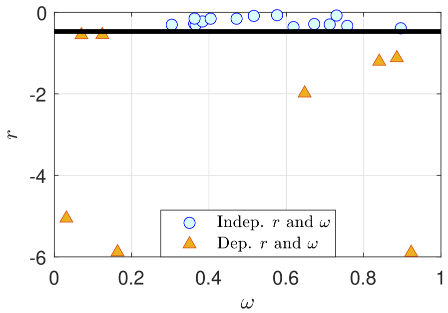


Figure 6.5: r-ω plot. The Horizontal black line corresponds to $r_0 \approx -0.47$.

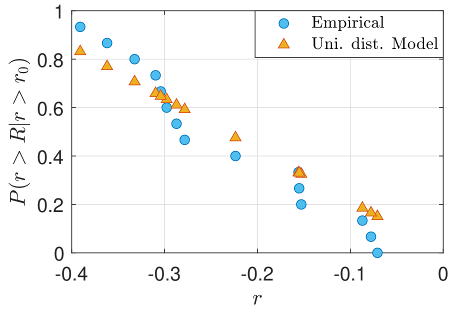


Figure 6.6: The distribution of the radial component of the Pickands coordinate for $r > r_0$.

constraint in (6.8) is satisfied as $\int_\omega \omega H_l(d\omega) = 0.5001$, indicating the validity of the the BGPD model based on the logistic distribution approach. On the other hand, referring to the Pickands density function $h_{pp}(\omega)$ obtained from the Poisson point process-based BGPD approach, the 0.5 mean constraint is satisfied as $\int_\omega \omega H_{pp}(d\omega) = 0.5064$ for the derived $H_{pp}(\omega)$.

6.6.6 *Comparison with Conventional Tail Estimation Method*

Fig. 6.4 additionally illustrates the CDF of the normalized power values for the traditional extrapolation-based approach compared to the empirical and the proposed BGPD models based on the logistic distribution and Poisson point process approaches. To derive the traditional extrapolation-based model, we first extract the observations from Rx1 and Rx2 corresponding to the reliability order of 10^{-3}-10^{0} PER. Then, upon fitting a variety of distributions to the samples in each observation sequence, the best-fitting distribution is determined as the Gaussian distribution according to the Akaike information criterion/Bayesian information criterion (AIC/BIC) metric. It is worth noting that Gaussian and Rician distribution had almost the same AIC/BIC for both sequences. However, for the sake of simplicity and without loss of generality, we opt to continue with the Gaussian distribution in our analysis as its joint PDF calculation is more straightforward. Then, we determine the joint bi-variate CDF of the Gaussian distribution for variables X and Y, corresponding to the received power values obtained from Rx1 and Rx2, respectively. Upon determining the joint CDF of X and Y, we extrapolate it towards the ultra-reliable region based on the tail approximation approach in [Eggers et al., 2019]. According to the results illustrated in Fig. 6.4, the proposed BGPD models outperform the extrapolated Gaussian model remarkably for a system operating in URLLC, where the reliability orders are in the range of 10^{-9}-10^{-5}. The RMSE of the extrapolated bi-variate Gaussian model is 7.8177×10^{09}, indicating the superiority of the proposed model with a significant improvement in the estimation of the tail distribution for ultra-reliability region, where the reliability orders are in the range of 10^{-9}-10^{-5}.

6.7 Conclusions

In this chapter, we introduced a novel framework based on the extreme value theory with the goal of estimating the multivariate channel tail statistics for a MIMO-URLLC. The proposed methodology adopted EVT to determine the optimum threshold over which the tail statistics are derived by using the UGPD model, validate the final model by using probability plots, and utilize MEVT to model the tail of the joint probability distribution by using the EVT-based logistic distribution and Poisson point process approaches. The proposed multi-dimensional channel modeling methodology achieved a significantly better fit to the empirical data in the lower tail than the conventional extrapolation-based approach. In the future, we are planning to extend this work by first determining the optimum transmission rate for a system operating in MIMO-URLLC, and incorporating a confidence interval for the estimated BGPD parameters and the corresponding optimum transmission rate to enable the usage of MEVT in real-time communication.

Chapter 7

POWER EFFICIENT BEAM-FORMING ALGORITHM FOR ULTRA-RELIABLE LOW LATENCY MILLIMETER-WAVE COMMUNICATIONS

The content of this chapter has been presented in International Black Sea Conference on Communications and Networking (IEEE BlackSeaCom2019).

7.1 Overview

Achieving ultra low-latency and ultra-high reliable communications at millimeter-waves (mmWaves) is challenging due to the sensitivity of this frequency band to blockage. Different space diversity schemes based on either equal gain or beam-tracking algorithms, are proposed to cope with the significant propagation losses introduced by the blockage at mmWaves. However, equal gain algorithms suffer from the low throughput as they allocate considerable gain to the non-optimal links. On the other hand, beam-tracking algorithms suffer from the link outage as the beam-forming is redone after dropping a large amount of data. In this chapter, we propose a new algorithm based on the prior collection of data to eliminate power consuming beam-tracking techniques while minimum diversity level ensures high received power, ultra-reliability, and low latency.

7.2 Introduction

The fifth generation (5G) mobile network exploits the millimeter wave (mmWave) bands due to its enormous amount of spectrum to greatly increase communication capacity. On the other hand, 5G applications ranging from machine-type communications to mission-critical communications such as autonomous driving and virtual reality require ultra reliability and low latency [Bennis et al., 2018]. Although, small wave length at mmWaves, allows us to place multiple antennas in a tiny area and steer the beam in the desired direction to combat significant propagation loss at mmWave [Niu et al., 2015], using directional antennas brings challenges in achieving ultra reliability due to the blockage [Niu et al., 2015], [Gao et al., 2014]. Diversity is a promising technique based on the transmission data in multiple dimensions to achieve high reliability and low latency at the cost of extra energy consumption.

Different forms of diversity techniques that can be applied to data transmission include time diversity [Serror et al., 2015a], frequency diversity [Pocovi et al., 2015a], space diversity [Pocovi et al., 2015b], and interface diversity [Nielsen et al., 2018a]. The gains of these diversity techniques are either adapted based on the channel properties or chosen equally. Although, in the equal gain diversity, there is no need to track the channel continuously, the method results in sub-optimal throughput due to the allocation of large transmission power to non-optimal propagation paths. On the contrary, adaptively changing the diversity gains requires real-time channel-tracking which is costly from the standpoint of power consumption.

By assuming perfect beam-tracking at upper layers, mmWave channel measurements focus on deriving statistical parameters for channel impulse response (CIR), path loss, root mean square (RMS) delay spread, and outage probability [Hemadeh et al., 2018], [Rappaport et al., 2015], [Rappaport et al., 2017], [Sun et al., 2016], [Maccartney et al., 2015], [MacCartney et al., 2015], [Rappaport et al., 2013], [Sun et al., 2015], [Samimi and Rappaport, 2016a], [MacCartney et al., 2017]. These studies either use the best direction with maximum received power [Samimi and Rappaport, 2016a], [MacCartney et al., 2017], [MacCartney et al., 2015], [Maccartney et al., 2015], [Sun et al., 2015], [Palacios et al., 2017], [Rappaport et al., 2015] or build a diversity scheme by combining N best directions with maximum powers [Slezak et al., 2017], [Samimi and Rappaport, 2016b], [Rappaport et al., 2013], [Kwon and Widmer, 2018] to obtain the channel parameters. However, to have a power efficient system and consume less transmitted power, diversity level should be as minimum as possible[Gao et al., 2014], [Xiao, 2013].

The constant beam-tracking can not achieve ultra high reliability and low latency simultaneously. The beam-tracking algorithms need to send an acknowledge packet continuously to check whether the beam is blocked or not [Gao et al., 2014], [Xiao, 2013]. The acknowledge packet transmission rate (APTR) determines a trade-off between power and delay efficiency such that considering very high APTR, makes the system power consuming. On the other hand, with a low APTR, transmitter (Tx) and receiver (Rx) detect that the link is lost only after dropping a large number of packets [Park and Pan, 2012], and this is not tolerable for some applications with delay constraint.

The goal of this chapter is to propose a novel beam-forming methodology based on the prior collection of channel measurements under the blockage scenarios based on piecewise linear model (PWL) introduced in [Peter et al., 2012] to eliminate the need for continuous beam-tracking and achieve a certain level of reliability. The original contributions of this chapter are listed as follows:

- We collect the data for the scenarios based on PWL for different transmitter antenna rotations and different half power beam-widths (HPBWs). This data collection is performed only once before the system begins to work, to eliminate the costly continuous beam-tracking when the system is operating.

- We determine a diversity scheme containing the minimum number of best beams with the goal of achieving a certain level of reliability. For this diversity scheme, the gains are determined optimally through a linear optimization problem to attain maximum received power.

- We demonstrate the superior performance of the proposed algorithm compared to the existing equal gain and online beam-tracking diversity schemes under different scenarios via extensive simulations.

The rest of the chapter is organized as follows: Section 7.3 describes the system model and assumptions considered throughout the chapter. Section 7.4 presents the beam-forming methodology for choosing minimum number of best beams with the goal of achieving maximum average received power for a certain level of reliability. Simulation results are presented in Section 7.5. Finally Section 7.6 concludes the chapter.

7.3 System Model

The system contains transmitter-receiver pairs requiring an ultra-reliable communication at mmWaves. We assume that the position of Tx and Rx are fixed while the environment is changing over the time. To better illustrate such systems, we bring up a very common example of an office room with fixed transmitter and receiver where the human shadowing or objects in the room introduce fading to the signal or disconnect the communication for a while. Some other instances might be an intra-vehicular communication considering fixed antennas in a moving car, where the people in the car may fade the communication or the temperature increment of the engine makes the environment time-varying for the transceivers located therein. This system model can be found also in the medical cases where the transceiver nodes inside the body are placed in the fixed positions and circulation system makes the communication vulnerable to fading.

The system applies beam-forming, which essentially either equally or optimally weighs and sums signals from different directions, such that the power of received signal is maximized while never dropping under a specific threshold [Xu et al., 2018]. Space diversity is utilized to keep the received power over the threshold for achieving a high reliable communication. Over an ultra-reliable communication, the transmitted packet is guaranteed to be received at the destination with about 10-5 to 10-9 packet error rate, uncorrupted and in the order they were sent [Nielsen et al., 2018b].

The system adjusts the parameters of the directional antenna, HPBW and pointing angle of Tx to Rx, to compensate the propagation loss introduced by the signal fading. Selecting the best combination of HPBW and pointing angle helps us to obtain more power at the receiver while the threshold is concerned . As the line of sight (LoS) component accounts for a sizable percentage of received power, its existence is crucial to the operation of a reliable mmWave communication [Kutty and Sen, 2016]. However, as this link might be blocked, we need to search for some supportive links to compensate the fading effect when the LoS path is temporary or permanently blocked. Therefore other HPBWs and pointing angle pairs between Tx and Rx should be considered to find a link with the 2^{nd}, 3^{rd} or even higher order maximum power. Hereupon, we guarantee a high reliable communication with least probability of packet error.

The system contains a transmitter's antenna array having N_t antenna elements. following the approach introduced in [Park and Pan, 2012], it is assumed that the l^{th} reflector is located in direction (ϕ_{tl}, θ_{tl}) from the transmitter. As expressed in [Park and Pan, 2012], the steering vector associated with direction (ϕ_{tl}, θ_{tl}) is assumed $h_l = [e^{j2\pi f_0 \tau_1(\phi_{tl},\theta_{tl})}, ..., e^{j2\pi f_0 \tau_{N_t}(\phi_{tl},\theta_{tl})}]^T$, where f_0 is the carrier frequency of the signal. Also, the signal arrives from the i^{th} antenna element and l^{th} reflector with $\tau_i(\phi_{tl}, \theta_{tl})$ seconds delay.Note that ^{T} is the transpose operator. The $N_t \times N_t$ channel matrix over the l^{th} propagation path C_l can be expressed as

$$C_l = \lambda_l h_l^T, \tag{7.1}$$

where λ is the channel gain over the l^{th} propagation path. To estimate the channel between the transmitter and the receiver over multiple propagation paths, first we need to find the N strongest paths by sweeping the predefined sectors and then refine the channel for each path [Gao et al., 2014], [Xiao, 2013], [Park and Pan, 2012]. The channel matrix $\boldsymbol{C}$ between the transmitter and the receiver over N chosen propagation paths is expressed as

$$\boldsymbol{C} = \boldsymbol{\Lambda} \boldsymbol{H}^T, \tag{7.2}$$

where $\boldsymbol{\Lambda}$ is $N \times N$ diagonal matrix with $[diag\{\lambda_1\ \lambda_2\ ...\ \lambda_N\}]_{ii} = \lambda_i$, and $\boldsymbol{H}$ is $N_t \times N$ matrix defined as $\boldsymbol{H} = [h_1\ h_2\ ...\ h_N]$ [Park and Pan, 2012].

7.4 Methodology

The objective of the proposed methodology is to achieve an ultra-reliable communication with minimum number of diversity level without costly real-time beam-tracking procedure. The main features of the suggested algorithm are as follows:

1. Before the system starts operation, under PWL scenarios in the environment, beam training is performed only once to find all possible links from Tx to Rx.

2. The links with maximum powers will be weighted optimally to maximize the power over a certain threshold.

3. Since the beam-tracking is done only once to find out the best measurement setup that guarantees achieving a reliable communication, the power consuming online beam-tracking is eliminated in our algorithm.

In the proposed methodology, Algorithm 2 is executed for PWL model based on the environment in which the measurement setup is going to operate. Minimum Diversity Beam-forming (MDBA) algorithm includes two phases: in Phase-I, transmit and receive antenna arrays adjust pointing angle and HPBW to check all possible combinations to calculate the $N_t \times 1$ weight vector, w_t such that the N beams are steered toward the N strongest propagation paths as

$$[h_1 \, h_2 \, ... \, h_N]^T w_t = [\alpha 1 \, \alpha 2 \, ... \, \alpha N]^T, \tag{7.3}$$

where α_l is a constant coefficient controlling the gain of the beam steered toward the l^{th} path. The transmit weight vector, w_t can be expressed as

$$w_t = (\boldsymbol{H}^T)^{-1} \boldsymbol{\alpha}, \tag{7.4}$$

where $\boldsymbol{\alpha} = [\alpha 1 \, \alpha 2 \, ... \, \alpha N]^T$. Consequently, the received signal, r, can be calculated as follows

$$\begin{aligned} r &= \boldsymbol{C} w_t s + n = \boldsymbol{\Lambda} \boldsymbol{H}^T (\boldsymbol{H}^T)^{-1} \boldsymbol{\alpha} + n \\ &= \boldsymbol{\Lambda} \boldsymbol{\alpha} s + n = \sum_{l=1}^{N} \alpha_l \lambda_l s + n, \end{aligned} \tag{7.5}$$

where s and n are the transmitted signal and additive white Gaussian noise (AWGN), respectively.

In Phase-II of MDBA, the diversity level (i.e. N) and the gains of chosen links (i.e. α_l) are determined to satisfy the condition of reliable communication in which the received power is above a certain threshold while maximum power is utilized. The first link with maximum power is passed through the MDBA algorithm (Line 12) to check whether the power does not exceed the threshold evermore. In the case where the first link violates this condition, two links or more with maximum powers are combined according to the optimization problem presented in (7.6), not only to satisfy the threshold condition, but also to achieve maximum received power.

The optimal gain assignment of the best links is therefore formulated as follows:

Algorithm 2 Minimum Diversity Beam-forming (MDBA) Algorithm

PHASE-I: Best HPBW and Rotation Angle Combination
for All HPBWs **do**
  for All Rotation Angles **do**
    α = link coefficient for the HPBW and rotation angle combination;
    Sort the powers corresponding to the link coefficients in descending order;
  end for
end for
PHASE-II: Minimum Diversity Selection
$i = 1$;
n = Maximum number of links;
while $i \leq n$ **do**
  if $\exists\ \alpha_i$ Such That $r \geq$ Threshold **then**
    Optimize α_i;
    break;
  end if
  $i = i + 1$;
end while

Maximize

$$r = \sum_{l=1}^{N} \alpha_l \lambda_l s + n, \tag{7.6a}$$

subject to

$$\sum_{i=1}^{i=N} \alpha_i = 1, \tag{7.6b}$$

variables

$$a_i \geq 0,\ i \in Z, \tag{7.6c}$$

where r is the received power of each link obtained through MDBA algorithm, N is the maximum number of links, and i is the index of sorted power. Also, α represents the unit-less link gain. The objective of the optimization problem is maximizing the received power. The sum of link gains should be one, as stated in (7.6b).

It is notable that there exist a dependency between rotation angle and HPBW such that if the HPBW is θ, the transmitter antenna should be rotated by θ increment step. This dependency between rotation angle and HPBW is for avoiding the overlap or correlation between the links. Additionally, as a rule of thumb to find the best rotation angle, maximum power is received when the incidence angle is almost same as the reflection angle, i.e. the reflection coefficient is -1.

7.5 Performance Evaluation

The goal of this section is to evaluate the performance of the proposed MDBA algorithm compared to the existing equal gain combining [Park and Pan, 2012] and beam-tracking algorithms [Xiao, 2013]. In beam-tracking algorithm, we assume that the transceiver communicates over the link with maximum received power, which is usually the LoS one. Once the link is being blocked, the communication will be switched to the link with second largest power. To make the comparison, the corresponding channel parameters are obtained by means of exploiting the channel modeling capability of a powerful electromagnetic modeling tool, named "Wireless InSite "(WI). At the first step, a 3D environment with the specific properties of objects such as wall and ceiling is created in WI. In the next step, we choose Shooting and Bouncing Ray (SBR) method which traces the rays from a transmitter and then propagates the rays through a defined environment until they reach a receiver. The rays, which are specularly reflected, have at most 30 interactions. These interactions include transmission, reflection, and diffraction around the objects [Wireless InSite, 2019]. Then the raw data is imported to MATLAB® for more analysis. The summary of the simulation parameters is shown in Table 7.1.

Table 7.1: Summary of simulation parameters

Parameter	**Value**
Carrier Frequency	60 GHz
Bandwidth	1 GHz
Transmitter Power	16 dBm
Tx, Rx Height	0.8 m
Office Dimension	6 m × 6 m × 3 m
Tx Antenna Type	Directional
Rx Antenna Type	Omnidirectional
Tx Antenna Gain	22
Rx Antenna Gain	2

7.5.1 Simulation Scenario

Based on our system model, we have emulated an office room with objects in the existence of maximum 2 people. Tx and Rx are located exactly in front of each other at the same height with 4 m separation distance. Rx is on the desk, 0.8 m away from the ground. In the Tx position, a multi-array transmitter is placed which rotates to find the best beams with highest power level.

Fig.7.1 shows 3 different scenarios, which have been considered as PWL model scenarios in an office room: 1) A person walks parallel to the Tx-Rx line, 2) A person crosses the Tx-Rx line at the middle point, and 3) Two people taking different paths from door to the chairs in the office.

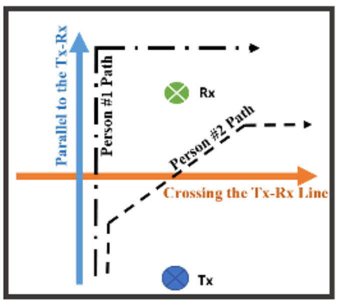


Figure 7.1: Different scenarios under the consideration.

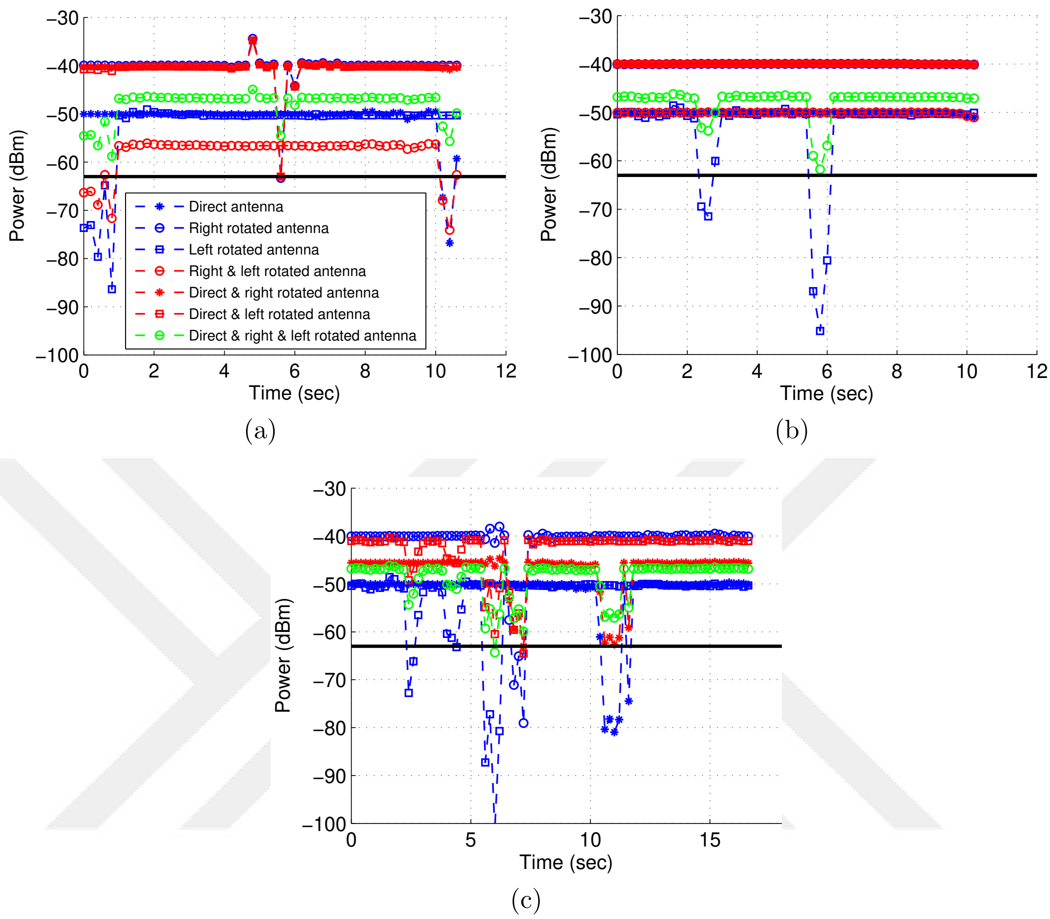


Figure 7.2: Received power vs. time based on equal gain combining; (a) one person crosses the direct link, (b) one person walks parallel to the direct link, and (c) two people walk in the office.

To choose the proper threshold for an office environment, by referring to the IEEE 802.11ad standard [Committee et al., 1999] single carrier modulation and coding scheme MCS6 can deliver 1.5 Gbit/s which is enough for high definition (HD) video streaming [Semkin et al., 2014]. To support this data rate or more, the received signal power has to be −63 dBm or higher. Therefore our requirement to have a reliable communication is to receive at least −63 dBm power at the Rx point all over the time.

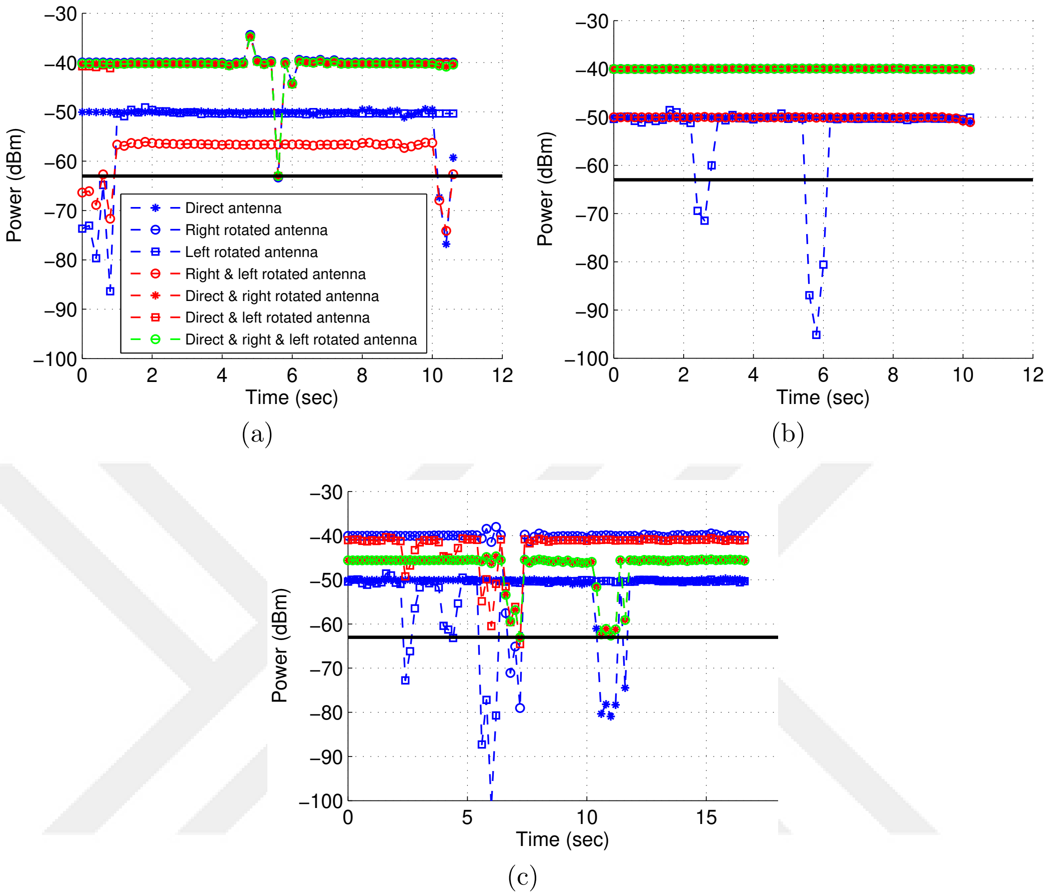


Figure 7.3: Received power vs. time based on MDBA Algorithm; (a) one person crosses the direct link, (b) one person walks parallel to the direct link, and (c) two people walk in the office.

7.5.2 Algorithm Performance

Fig. 7.2 shows the received power based on equal gain algorithm for three scenarios. Only the received power obtained through the combination of 3 links, i.e. direct, right and left rotated links, is above the threshold for all scenarios. Hence, in equal gain combining, the minimum diversity level satisfying ultra-reliability is 3. It is shown that this scheme lowers the average received power in the case of no blockage due to allocation of large transmission power to a non-optimal link.

Fig. 7.3 shows the received power based on MDBA algorithm for three scenarios.

Among all links, the combination of direct and right rotated link guarantees ultra-reliability in all scenarios under the consideration. Moreover, the performance of this link is almost same as the combination of direct, right and left rotated links. Therefore, increasing the diversity level does not improve the communication quality significantly where optimum gains are assigned to the links in the diversity scheme. The optimum coefficients for direct and right rotated links are 0.4465 and 0.5535, respectively. It is worth mentioning that the combination of direct and right rotated links has the best performance compared to the other, since most of the activities in the emulated office room occurred in the left side, and therefore more stable environment exists in the right side.

Fig. 7.4 shows the performance comparison of MDBA algorithm and the existing algorithms, i.e. equal gain and online beam-tracking [Xiao, 2013], [Park and Pan, 2012]. Fig. 7.4 validates the superiority of the proposed scheme over the other two algorithms. Furthermore Fig. 7.4 reveals that although the online beam-tracking algorithm offers more throughput, the scheme suffers from the link outage, i.e. exceeding the threshold. On the other hand, not only the average throughput of MDBA algorithm is higher than that of equal gain algorithm, but also the signal level achieved by MDBA algorithm does not drop as frequently as that of equal gain algorithm. A a result, the proposed scheme does not experience any link outage, and its average received power is only 3.6 dBm less than the averaged power obtained by online beam-tracking algorithm.

7.6 Conclusions

In this chapter, we have proposed a new algorithm to eliminate the continuous beam-tracking and achieve a certain level of reliability while minimum level of diversity is utilized. We have developed a simulation platform to evaluate the performance of the proposed scheme compared to equal gain and online beam-tracking algorithms. We

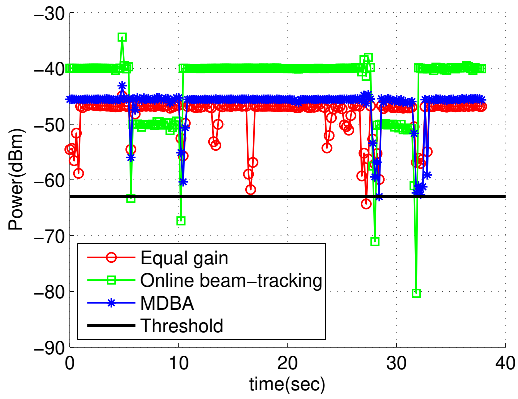


Figure 7.4: Performance comparison of equal gain, online beam-tracking, and MDBA algorithms.

have shown that equal gain scheme lowers throughput in the case of no blockage, as well as online beam-tracking suffers from the link outage although it is a power consuming scheme. Extensive simulated results for different scenarios illustrate that the proposed scheme guarantees low latency ultra-reliability while maximum throughput is achieved with minimum diversity level to consume less energy.

Chapter 8

CONCLUSIONS

In this thesis, we propose a novel framework based on EVT to model the tail distribution of stationary and non-stationary channels by incorporating the parameters of GPD, estimate the optimum transmission rate satisfying the ultra-reliability constraints, incorporate a confidence interval for the estimated GPD parameters and the corresponding optimum transmission rate to enable the usage of EVT in real-time communication, and estimating the multivariate channel tail statistics for a MIMO-URLLC.

First, we propose a channel modeling methodology that adopts EVT to determine the optimum threshold over which the tail statistics are derived, the stopping condition to specify the minimum number of samples required to model the tail characteristics, and the assessment to validate the final model by using probability plots. We have shown that the proposed channel modeling methodology achieves a significantly better fit to the empirical data in the lower tail than the conventional extrapolation-based approach, especially in the ultra-reliability region. Then, upon utilizing EVT-based techniques for modeling the extremes of a non-stationary channel for URLLC, we have demonstrated that the proposed methodology for modeling a non-stationary channel achieves significantly better fit to the empirical data than the modeling approach for a stationary channel.

We determine the maximum transmission rate of the estimated channel and validate the selected rate by assessing the resulting error probability. The achieved reliability from the proposed EVT-based approach has been shown to outperform the traditional methods based on the extrapolation of average statistic channel models. Next, we use the estimated GPD parameters along with their confidence intervals, to compute the transmission rate and its corresponding confidence interval. The incorporation of CI results into rate selection estimate reduces sample complexity by reducing the number of samples required for rate estimation in the training phase to attain a specific level of reliability.

Considering a system with spatial diversity operating in URLLC, we adopt EVT to determine the optimum threshold over which the tail statistics are derived by using the UGPD model, validate the final model by using probability plots, and utilize MEVT to model the tail of the joint probability distribution by using the EVT-based logistic distribution and Poisson point process approaches. We have shown that the proposed multi-dimensional channel modeling methodology based on MEVT achieves a significantly better fit to the empirical data in the lower tail than the conventional extrapolation-based approach.

Finally, we propose a new algorithm to eliminate the continuous beam-tracking and achieve a certain level of reliability while minimum level of diversity is utilized. We demonstrate the superior performance of the proposed beam-forming algorithm by showing that the equal gain scheme lowers throughput in the case of no blockage, as well as online beam-tracking suffers from the link outage although it is a power consuming scheme.

In future, we plan to extend this work by first determining the optimum transmission rate for a system operating in MIMO-URLLC, and incorporating a confidence interval for the estimated BGPD parameters and the corresponding optimum transmission rate to enable the usage of MEVT in real-time communication.

BIBLIOGRAPHY

[Abdel-Aziz et al., 2020] Abdel-Aziz, M. K., Samarakoon, S., Liu, C., Bennis, M., and Saad, W. (2020). Optimized age of information tail for ultra-reliable low-latency communications in vehicular networks. *IEEE Transactions on Communications*, 68(3):1911–1924.

[Abdel-Aziz et al., 2019] Abdel-Aziz, M. K., Samarakoon, S., Liu, C.-F., Bennis, M., and Saad, W. (2019). Optimized age of information tail for ultra-reliable low-latency communications in vehicular networks. *IEEE Transactions on Communications*, 68(3):1911–1924.

[Al-Badarneh et al., 2018a] Al-Badarneh, Y. H., Georghiades, C. N., and Alouini, M.-S. (2018a). Asymptotic performance analysis of the k th best link selection over wireless fading channels: An extreme value theory approach. *IEEE Transactions on Vehicular Technology*, 67(7):6652–6657.

[Al-Badarneh et al., 2018b] Al-Badarneh, Y. H., Georghiades, C. N., Radaydeh, R. M., and Alouini, M.-S. (2018b). On the secrecy performance of generalized user selection for interference-limited multiuser wireless networks. *IEEE Transactions on Vehicular Technology*, 67(12):12442–12446.

[Almers et al., 2007] Almers, P., Bonek, E., Burr, A., Czink, N., Debbah, M., Degli-Esposti, V., Hofstetter, H., Kyösti, P., Laurenson, D., Matz, G., et al. (2007). Survey of channel and radio propagation models for wireless mimo systems. *EURASIP Journal on Wireless Communications and Networking*, 2007:1–19.

[Angjelichinoski et al., 2019] Angjelichinoski, M., Trillingsgaard, K. F., and Popovski, P. (2019). A statistical learning approach to ultra-reliable low latency communication. *IEEE Transactions on Communications*, 67(7):5153–5166.

[Bas and Coleri Ergen, 2012] Bas, C. U. and Coleri Ergen, S. (2012). Ultra-wideband channel model for intra-vehicular wireless sensor networks beneath the chassis: From statistical model to simulations. *IEEE Transactions on Vehicular Technology*, 62(1):14–25.

[Bennis et al., 2018] Bennis, M., Debbah, M., and Poor, H. V. (2018). Ultra reliable and low-latency wireless communication: Tail, risk, and scale. *Proceedings of the IEEE*, 106(10):1834–1853.

[Bennis et al., 2018] Bennis, M., Debbah, M., and Poor, H. V. (2018). Ultrareliable and low-latency wireless communication: Tail, risk, and scale. *Proceedings of the IEEE*, 106(10):1834–1853.

[Bensalah et al., 2000] Bensalah, Y. et al. (2000). *Steps in applying extreme value theory to finance: a review.* Citeseer.

[Boyd et al., 2019] Boyd, C., Kotaba, R., Tirkkonen, O., and Popovski, P. (2019). Non-orthogonal contention-based access for urllc devices with frequency diversity. In *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE.

[Cai and Hames, 2010] Cai, Y. and Hames, D. (2010). Minimum sample size determination for generalized extreme value distribution. *Communications in Statistics—Simulation and Computation®*, 40(1):87–98.

[Caires, 2011] Caires, S. (2011). Extreme value analysis: wave data. *JCOMM Technical Report*, 57.

[Capéraà and Fougères, 2000] Capéraà, P. and Fougères, A.-L. (2000). Estimation of a bivariate extreme value distribution. *Extremes*, 3(4):311–329.

[Chen et al., 2011] Chen, C., Hu, J., Meng, Q., and Zhang, Y. (2011). Short-time traffic flow prediction with arima-garch model. In *2011 IEEE Intelligent Vehicles Symposium (IV)*, pages 607–612. IEEE.

[Chen et al., 2014] Chen, Y., Goldsmith, A. J., and Eldar, Y. C. (2014). Backing off from infinity: Performance bounds via concentration of spectral measure for random mimo channels. *IEEE Transactions on Information Theory*, 61(1):366–387.

[Coles et al., 2001] Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*, volume 208. Springer.

[Committee et al., 1999] Committee, I. C. S. L. M. S. et al. (1999). Wireless lan medium access control (mac) and physical layer (phy) specifications. *ANSI/IEEE Std. 802.11-1999*.

[Cooley et al., 2010] Cooley, D., Davis, R. A., and Naveau, P. (2010). The pairwise beta distribution: A flexible parametric multivariate model for extremes. *Journal of Multivariate Analysis*, 101(9):2103–2117.

[Dahman et al., 2010] Dahman, G. S., Bultitude, R. J., and Hafez, R. H. (2010). Identifying and modelling multipath clusters in propagation measurement data. In *2010 IEEE 72nd Vehicular Technology Conference-Fall*, pages 1–5. IEEE.

[de Zea Bermudez and Kotz, 2010] de Zea Bermudez, P. and Kotz, S. (2010). Parameter estimation of the generalized pareto distribution—part II. *Journal of Statistical Planning and Inference*, 140(6):1374–1388.

[Demir et al., 2013] Demir, U., Bas, C. U., and Coleri Ergen, S. (2013). Engine compartment UWB channel model for intra-vehicular wireless sensor networks. *IEEE Transactions on Vehicular Technology*, 63(6):2497–2505.

[Demir and Coleri Ergen, 2016] Demir, U. and Coleri Ergen, S. (2016). Arima-based time variation model for beneath the chassis UWB channel. *EURASIP Journal on Wireless Communications and Networking*, 2016(1):178.

[Durisi et al., 2016] Durisi, G., Koch, T., and Popovski, P. (2016). Toward massive, ultrareliable, and low-latency wireless communication with short packets. *Proceedings of the IEEE*, 104(9):1711–1726.

[Eggers et al., 2019] Eggers, P. C. F., Angjelichinoski, M., and Popovski, P. (2019). Wireless channel modeling perspectives for ultra-reliable communications. *IEEE Transactions on Wireless Communications*, 18(4):2229–2243.

[Falk et al., 2010] Falk, M., Hüsler, J., and Reiss, R.-D. (2010). *Laws of small numbers: extremes and rare events*. Springer Science & Business Media.

[Finkenstadt and Rootzén, 2003] Finkenstadt, B. and Rootzén, H. (2003). *Extreme values in finance, telecommunications, and the environment*. CRC Press.

[Fuller, 2009] Fuller, W. A. (2009). *Introduction to statistical time series*, volume 428. John Wiley & Sons.

[Galambos et al., 1994] Galambos, J., Lechner, J., Simiu, E., and HAGWOOD, C. (1994). Extreme-value theory and applications-proceedings of the conference on extreme-value theory and applications. 2. gaithersburg, md, may-1993-preface.

[Gao et al., 2014] Gao, B., Xiao, Z., Zhang, C., Su, L., Jin, D., and Zeng, L. (2014). Double-link beam tracking against human blockage and device mobility for 60-

GHz WLAN. In *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 323–328.

[Gelman et al., 2019] Gelman, A., Goodrich, B., Gabry, J., and Vehtari, A. (2019). R-squared for Bayesian regression models. *The American Statistician*, 73(3):307–309.

[Glosten et al., 1993] Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801.

[Guidolin and Pedio, 2018] Guidolin, M. and Pedio, M. (2018). *Essentials of time series for financial applications*. Academic Press.

[Gustafson et al., 2014] Gustafson, C., Haneda, K., Wyne, S., and Tufvesson, F. (2014). On mm-wave multipath clustering and channel modeling. *IEEE Transactions on Antennas and Propagation*, 62(3):1445–1455.

[Hemadeh et al., 2018] Hemadeh, I. A., Satyanarayana, K., El-Hajjar, M., and Hanzo, L. (2018). Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget. *IEEE Communications Surveys Tutorials*, 20(2):870–913.

[Joe et al., 1992] Joe, H., Smith, R. L., and Weissman, I. (1992). Bivariate threshold methods for extremes. *Journal of the royal statistical society: series B (methodological)*, 54(1):171–183.

[Khosravirad et al., 2020] Khosravirad, S. R., Viswanathan, H., and Yu, W. (2020). Exploiting diversity for ultra-reliable and low-latency wireless control. *IEEE Transactions on Wireless Communications*, 20(1):316–331.

[Knapp, 2009] Knapp, K. J. (2009). *Cyber Security and Global Information Assurance: Threat Analysis and Response Solutions: Threat Analysis and Response Solutions*. IGI Global.

[Kotz and Nadarajah, 2000] Kotz, S. and Nadarajah, S. (2000). *Extreme value distributions: theory and applications*. World Scientific.

[Kutty and Sen, 2016] Kutty, S. and Sen, D. (2016). Beamforming for millimeter wave communications: An inclusive survey. *IEEE Communications Surveys Tutorials*, 18(2):949–973.

[Kwiatkowski et al., 1992] Kwiatkowski, D., Phillips, P. C., Schmidt, P., Shin, Y., et al. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of econometrics*, 54(1-3):159–178.

[Kwon and Widmer, 2018] Kwon, S. and Widmer, J. (2018). Multi-beam power allocation for mmwave communications under random blockage. In *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, pages 1–5.

[Li et al., 2018] Li, C., Yang, K., Yu, J., Li, F., Shui, Y., Chang, F., and Chen, W. (2018). V2V radio channel performance based on measurements in ramp scenarios at 5.9 GHz. *IEEE Access*, 6:7503–7514.

[Liu and Bennis, 2018] Liu, C.-F. and Bennis, M. (2018). Ultra-reliable and low-latency vehicular transmission: An extreme value theory approach. *IEEE Communications Letters*, 22(6):1292–1295.

[MacCartney et al., 2017] MacCartney, G. R., Rappaport, T. S., and Ghosh, A. (2017). Base station diversity propagation measurements at 73 ghz millimeter-wave for 5g coordinated multipoint (comp) analysis. In *2017 IEEE Globecom Workshops (GC Wkshps)*, pages 1–7.

[Maccartney et al., 2015] Maccartney, G. R., Rappaport, T. S., Sun, S., and Deng, S. (2015). Indoor office wideband millimeter-wave propagation measurements and channel models at 28 and 73 ghz for ultra-dense 5g wireless networks. *IEEE Access*, 3:2388–2424.

[MacCartney et al., 2015] MacCartney, G. R., Samimi, M. K., and Rappaport, T. S. (2015). Exploiting directionality for millimeter-wave wireless system improvement. In *2015 IEEE International Conference on Communications (ICC)*, pages 2416–2422.

[Mahmood and Jäntti, 2016] Mahmood, A. and Jäntti, R. (2016). Packet error rate analysis of uncoded schemes in block-fading channels using extreme value theory. *IEEE Communications Letters*, 21(1):208–211.

[Makarov, 2007] Makarov, M. (2007). Applications of exact extreme value theorem. *Journal of Operational Risk*, 2(1):115–120.

[Mata, 2000] Mata, A. (2000). Parameter uncertainty for extreme value distributions. In *GIRO Convention Papers*, pages 151–173.

[Mehrnia and Coleri, ions] Mehrnia, N. and Coleri, S. (2020, submitted to IEEE Transactions on Wireless Communications). Wireless channel modeling based on extreme value theory for ultra-reliable communications. *arXiv: Signal Processing*.

[Mehrnia and Coleri, 2021a] Mehrnia, N. and Coleri, S. (2021a). Extreme value theory based rate selection for ultra-reliable communications. *Submitted to IEEE Transactions on Vehicular Technology*, pages 1–1.

[Mehrnia and Coleri, 2021b] Mehrnia, N. and Coleri, S. (2021b). Non-stationary wireless channel modeling approach based on extreme value theory for ultra-reliable communications. *IEEE Transactions on Vehicular Technology*, 70(8):8264–8268.

[Mehrnia and Coleri, 2021c] Mehrnia, N. and Coleri, S. (2021c). Wireless channel modeling based on extreme value theory for ultra-reliable communications. *IEEE Transactions on Wireless Communications*, 21(2):1064–1076.

[Mehrnia and Coleri, 2022] Mehrnia, N. and Coleri, S. (2022). Incorporation of confidence interval into rate selection based on the extreme value theory for ultra-reliable communications. In *European Conference on Networks and Communications (EuCNC) and the 6G Summit*, pages 1–6. IEEE.

[Michel, 2006] Michel, R. (2006). *Simulation and estimation in multivariate generalized Pareto models.* PhD thesis, Universität Würzburg.

[Molina-Garcia et al., 2008] Molina-Garcia, M., Fernandez-Duran, A., and Alonso, J. I. (2008). Application of extreme value distribution to model propagation fading in indoor mobile radio environments. In *2008 IEEE Radio and Wireless Symposium*, pages 97–100. IEEE.

[Nadarajah, 2000] Nadarajah, S. (2000). Approximations for bivariate extreme values. *Extremes*, 3(1):87–98.

[Nielsen et al., 2017] Nielsen, J. J., Liu, R., and Popovski, P. (2017). Ultra-reliable low latency communication using interface diversity. *IEEE Transactions on Communications*, 66(3):1322–1334.

[Nielsen et al., 2018a] Nielsen, J. J., Liu, R., and Popovski, P. (2018a). Ultra-reliable low latency communication using interface diversity. *IEEE Transactions on Communications*, 66(3):1322–1334.

[Nielsen et al., 2018b] Nielsen, J. J., Liu, R., and Popovski, P. (2018b). Ultra-reliable low latency communication using interface diversity. *IEEE Transactions on Communications*, 66(3):1322–1334.

[Niu et al., 2015] Niu, Y., Li, Y., Jin, D., Su, L., and Vasilakos, A. V. (2015). A survey of millimeter wave communications (mmwave) for 5G: opportunities and challenges. *Wireless Networks*, 21(8):2657–2676.

[Nystrom and Skoglund, 2002a] Nystrom, K. and Skoglund, J. (2002a). A framework for scenario-based risk management. *Preprint, Swedbank.*

[Nystrom and Skoglund, 2002b] Nystrom, K. and Skoglund, J. (2002b). Univariate extreme value theory, garch and measures of risk. *Preprint, Swedbank.*

[Palacios et al., 2017] Palacios, J., Donno, D. D., and Widmer, J. (2017). Tracking mm-wave channel dynamics: Fast beam training strategies under mobility. In *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, pages 1–9.

[Park and Pan, 2012] Park, M. and Pan, H. K. (2012). A spatial diversity technique for IEEE 802.11ad WLAN in 60 GHz band. *IEEE Communications Letters*, 16(8):1260–1262.

[Pase et al., 2021] Pase, F., Giordani, M., and Zorzi, M. (2021). On the convergence time of federated learning over wireless networks under imperfect csi. *arXiv preprint arXiv:2104.00331*.

[Peter et al., 2012] Peter, M., Wisotzki, M., Raceala-Motoc, M., Keusgen, W., Felbecker, R., Jacob, M., Priebe, S., and Kürner, T. (2012). Analyzing human body shadowing at 60 ghz: Systematic wideband mimo measurements and modeling approaches. In *2012 6th European Conference on Antennas and Propagation (EUCAP)*, pages 468–472.

[Pocovi et al., 2015a] Pocovi, G., Soret, B., Lauridsen, M., Pedersen, K. I., and Mogensen, P. (2015a). Signal quality outage analysis for ultra-reliable communications in cellular networks. In *2015 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6.

[Pocovi et al., 2015b] Pocovi, G., Soret, B., Lauridsen, M., Pedersen, K. I., and Mogensen, P. (2015b). Signal quality outage analysis for ultra-reliable communications in cellular networks. In *2015 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6.

[Popovski et al., 2018] Popovski, P., Nielsen, J. J., Stefanovic, C., De Carvalho, E., Strom, E., Trillingsgaard, K. F., Bana, A.-S., Kim, D. M., Kotaba, R., Park, J., et al. (2018). Wireless access for ultra-reliable low-latency communication: Principles and building blocks. *IEEE Network*, 32(2):16–23.

[Popovski et al., 2019] Popovski, P., Stefanović, Č., Nielsen, J. J., De Carvalho, E., Angjelichinoski, M., Trillingsgaard, K. F., and Bana, A.-S. (2019). Wireless access in ultra-reliable low-latency communication (URLLC). *IEEE Transactions on Communications*, 67(8):5783–5801.

[Rappaport et al., 2015] Rappaport, T. S., MacCartney, G. R., Samimi, M. K., and Sun, S. (2015). Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design. *IEEE Transactions on Communications*, 63(9):3029–3056.

[Rappaport et al., 2013] Rappaport, T. S., Sun, S., Mayzus, R., Zhao, H., Azar, Y., Wang, K., Wong, G. N., Schulz, J. K., Samimi, M., and Gutierrez, F. (2013). Millimeter wave mobile communications for 5g cellular: It will work! *IEEE Access*, 1:335–349.

[Rappaport et al., 2017] Rappaport, T. S., Xing, Y., Jr., G. R. M., Molisch, A. F., Mellios, E., and Zhang, J. (2017). Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models. *CoRR*, abs/1708.02557.

[Razali et al., 2011] Razali, N. M., Wah, Y. B., et al. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, 2(1):21–33.

[Reyes-Guerrero and Mariscal, 2014] Reyes-Guerrero, J. and Mariscal, L. (2014). 5.8 GHz propagation of low-height wireless links in sea port scenario. *Electronics Letters*, 50(9):710–712.

[Ruiz et al., 2007] Ruiz, E., Pellegrini, S., Espasa, A., et al. (2007). The relationship between ARIMA-GARCH and unobserved component models with GARCH disturbances. *Job Market Paper*, 1(ws072706).

[Samarakoon et al., 2018] Samarakoon, S., Bennis, M., Saad, W., and Debbah, M. (2018). Federated learning for ultra-reliable low-latency V2V communications. In *2018 IEEE Global Communications Conference (GLOBECOM)*, pages 1–7. IEEE.

[Samarakoon et al., 2019] Samarakoon, S., Bennis, M., Saad, W., and Debbah, M. (2019). Distributed federated learning for ultra-reliable low-latency vehicular communications. *IEEE Transactions on Communications*, 68(2):1146–1159.

[Samarakoon et al., 2020] Samarakoon, S., Bennis, M., Saad, W., and Debbah, M.

(2020). Predictive ultra-reliable communication: A survival analysis perspective. *IEEE Communications Letters*, 25(4):1221–1225.

[Samimi and Rappaport, 2016a] Samimi, M. K. and Rappaport, T. S. (2016a). 3-d millimeter-wave statistical channel model for 5g wireless system design. *IEEE Transactions on Microwave Theory and Techniques*, 64(7):2207–2225.

[Samimi and Rappaport, 2016b] Samimi, M. K. and Rappaport, T. S. (2016b). Local multipath model parameters for generating 5g millimeter-wave 3gpp-like channel impulse response. In *2016 10th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5.

[Sawada et al., 2015] Sawada, H., Mizutani, K., Ishizu, K., Matsumura, T., Tran, H.-N., Murakami, H., Kojima, F., and Harada, H. (2015). Path loss and throughput estimation models for an IEEE 802.11 af prototype. In *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE.

[Semkin et al., 2014] Semkin, V., Jacob, M., Kürner, T., Bisognin, A., Ferrero, F., Luxey, C., and Räisänen, A. V. (2014). Estimation of optimum antenna configurations supported by realistic propagation models at 60 GHz. In *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*, pages 3434–3438.

[Serror et al., 2015a] Serror, M., Dombrowski, C., Wehrle, K., and Gross, J. (2015a). Channel coding versus cooperative arq: Reducing outage probability in ultra-low latency wireless communications. In *2015 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6.

[Serror et al., 2015b] Serror, M., Dombrowski, C., Wehrle, K., and Gross, J. (2015b). Channel coding versus cooperative arq: Reducing outage probability in ultra-low latency wireless communications. In *2015 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6. IEEE.

[Shi et al., 2012] Shi, J., Anzai, D., and Wang, J. (2012). Channel modeling and performance analysis of diversity reception for implant uwb wireless link. *IEICE transactions on communications*, 95(10):3197–3205.

[Slezak et al., 2017] Slezak, C., Dhananjay, A., and Rangan, S. (2017). 60 ghz blockage study using phased arrays. In *2017 51st Asilomar Conference on Signals, Systems, and Computers*, pages 1655–1659.

[Smith, 1985] Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1):67–90.

[Song and Li, 2006] Song, G. and Li, Y. (2006). Asymptotic throughput analysis for channel-aware scheduling. *IEEE Transactions on Communications*, 54(10):1827–1834.

[Subhash and Kalyani, 2021] Subhash, A. and Kalyani, S. (2021). Cooperative relaying in a swipt network: Asymptotic analysis using extreme value theory for non-identically distributed rvs. *IEEE Transactions on Communications*, 69(7):4360–4372.

[Subhash et al., 2020] Subhash, A., Srinivasan, M., Kalyani, S., and Hanzo, L. (2020). Transmit power policy and ergodic multicast rate analysis of cognitive radio networks in generalized fading. *IEEE Transactions on Communications*, 68(6):3311–3325.

[Sun et al., 2015] Sun, S., MacCartney, G. R., Samimi, M. K., and Rappaport, T. S. (2015). Synthesizing omnidirectional antenna patterns, received power and path loss from directional antennas for 5g millimeter-wave communications. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–7.

[Sun et al., 2016] Sun, S., Rappaport, T. S., Thomas, T. A., Ghosh, A., Nguyen, H. C., Kovács, I. Z., Rodriguez, I., Koymen, O., and Partyka, A. (2016). Investigation of prediction accuracy, sensitivity, and parameter stability of large-scale propagation path loss models for 5g wireless communications. *IEEE Transactions on Vehicular Technology*, 65(5):2843–2860.

[Swamy et al., 2018] Swamy, V. N., Rigge, P., Ranade, G., Nikolic, B., and Sahai, A. (2018). Predicting wireless channels for ultra-reliable low-latency communications. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 2609–2613. IEEE.

[Swamy et al., 2019] Swamy, V. N., Rigge, P., Ranade, G., Nikolić, B., and Sahai, A. (2019). Wireless channel dynamics and robustness for ultra-reliable low-latency communications. *IEEE Journal on Selected Areas in Communications*, 37(4):705–720.

[Van Zyl, 2015] Van Zyl, J. M. (2015). Estimation of the shape parameter of a generalized pareto distribution based on a transformation to pareto distributed variables. *Journal of Statistical Theory and Practice*, 9(1):171–183.

[Wallace and Jensen, 2002] Wallace, J. W. and Jensen, M. A. (2002). Modeling the indoor mimo wireless channel. *IEEE Transactions on Antennas and Propagation*, 50(5):591–599.

[Wireless InSite, 2019] Wireless InSite (2019). Remcome, Wireless InSite 2.6.3. `http://www.remcom.com/wireless-insite`.

[Xiao, 2013] Xiao, Z. (2013). Suboptimal spatial diversity scheme for 60 GHz millimeter-wave WLAN. *IEEE Communications Letters*, 17(9):1790–1793.

[Xu et al., 2016] Xu, H., Sun, L., Ren, P., Du, Q., Wang, Y., and Gao, Z. (2016). Cooperative physical-layer approach for downlink privacy preserving in multiuser relay networks. In *2016 IEEE Global Communications Conference (GLOBECOM)*, pages 1–7.

[Xu et al., 2018] Xu, Q., Jiang, C., Han, Y., Wang, B., and Liu, K. J. R. (2018). Waveforming: An overview with beamforming. *IEEE Communications Surveys Tutorials*, 20(1):132–149.

[Zhang et al., 2021] Zhang, W., Derakhshani, M., and Lambotharan, S. (2021). Non-parametric statistical learning for urllc transmission rate control. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE.

[Zheng et al., 2014] Zheng, K., Ou, S., and Yin, X. (2014). Massive mimo channel models: A survey. *International Journal of Antennas and Propagation*, 2014.

[Zheng et al., 2018] Zheng, L., Ismail, K., Sayed, T., and Fatema, T. (2018). Bivariate extreme value modeling for road safety estimation. *Accident Analysis & Prevention*, 120:83–91.

[Zhou et al., 2005] Zhou, B., He, D., Sun, Z., and Ng, W. H. (2005). Network traffic modeling and prediction with arima/garch. In *Proc. of HET-NETs Conference*, pages 1–10.

[Zhou et al., 2019] Zhou, L., Wolf, A., and Motani, M. (2019). On lossy multi-connectivity: Finite blocklength performance and second-order asymptotics. *IEEE Journal on Selected Areas in Communications*, 37(4):735–748.

[Zhu et al., 2021] Zhu, Y., Hu, Y., Yang, T., Yang, T., Vogt, J., and Schmeink, A. (2021). Reliability-optimal offloading in low-latency edge computing networks: Analytical and reinforcement learning based designs. *IEEE Transactions on Vehicular Technology*, 70(6):6058–6072.