

Physical Layer Aspects of LoRa and Full-Duplex Wireless Transceivers

Présentée le 18 juin 2020

à la Faculté des sciences et techniques de l'ingénieur
Laboratoire de circuits pour télécommunications
Programme doctoral en génie électrique

pour l'obtention du grade de Docteur ès Sciences

par

Orion AFISIADIS

Acceptée sur proposition du jury

Prof. C. Dehollain, présidente du jury
Prof. A. P. Burg, Dr A. K. Balatsoukas Stimming, directeurs de thèse
Prof. L. Van der Perre, rapporteuse
Prof. G. Karagiannidis, rapporteur
Prof. M. Gastpar, rapporteur

Στους γονείς μου, Γιώργο και Κατερίνα,
και τους δασκάλους μου, Χρήστο Δημάκη και Αντώνη Παπαγιαννάκη.

Acknowledgements

First of all, I would like to thank my advisor, Prof. Andreas Burg, who believed in me and gave me the opportunity to perform my PhD studies at TCL. I deeply admire his ability to provide new paths to solve problems that seem insurmountable, as well as his ability to create an environment where people feel so comfortable to live and work in. He has been there for me whenever I asked for his guidance and he has extended his helping hand to me at all difficult times. I feel very lucky to have him as an advisor. I also would like to thank my co-advisor Prof. Alexios Balatsoukas Stimming for his great guidance and good friendship all these years. Thanks to him I am a better scientist and a better person. I owe him big part of my scientific mentality, and I feel the utmost respect for him.

I would like to thank Prof. Catherine Dehollain for acting as the president of my PhD jury, Prof. Liesbet Van der Perre and Prof. Georgios Karagiannidis for serving as external examiners, as well as Prof. Michael Gastpar for serving as an internal examiner. Their comments and feedback improved the quality of this manuscript.

Taking a step back in time, I want to thank all the people from the sixth floor of our School at Aristotle University of Thessaloniki. I thank Prof. Christos Dimakis who took me by the hand, guided me, and helped me understand how important integrity in science is. I thank Prof. Georgios Karagiannidis for all the inspiration and the clear solutions he always suggested to me. I thank Prof. Traianos Yioultsis for the uncountable discussions along with nice coffee in his office. I want to thank my first mentor and good friend Konstantis Arkoudogiannis, who breathed love and passion into my studies. I also want to thank Vasileios Kapinas, who believed in me and always pushed me to the next step in my studies. Finally, I thank my good friend the late Prof. Antonios Papagiannakis (The Tall Dude) for all the unending discussions over a glass, which have largely affected my way of thinking. I wish I could meet him and discuss with him one more evening.

I could not omit thanking the administrative assistant of TCL, Ioanna Paniara, for making everything smooth for me, so that I did not spend any time in administrative issues. Particular thanks go to all my current and former colleagues in TCL for all the technical and non-technical discussions all these years. They made and continue making TCL feel a great place to be! In alphabetical order I would like to thank: Konstantinos Alexandris, Andrew Austin, Alexios Balatsoukas Stimming, Pavle Belanovic, Andrea Bonetti, Jeremy Constantin, Matthieu Cotting, Shrikanth Ganapathy, Pascal Giard, Robert Giterman, Georgios Karakonstantis, Andreas Kristensen, Reza Ghanaatian, Sitian Li, Pascal Meinerzhagen, Christoph Müller, Nicholas Preyss, Lorenz Schmid, Adrian Schumacher, Christian Senning, Adi Teman, and Ning Xu.

Acknowledgements

An important part of the PhD studies has been related to the interaction with students that worked with me during their projects at TCL has been invaluable. For this reason, I thank alphabetically: Matthieu Cotting, Juan Martin Galicia, Juan Guirado, Guillaume Jaume, Joffrey Martinelli, Paul Mayoraz, and Joachim Tapparel.

I would like to thank all the friends that directly or indirectly became part of my PhD studies, mainly in Switzerland and Greece, but not only there. Without them my PhD studies would not have been so colorful. Special thanks to all the friends with whom we shared lunch and coffee breaks at EPFL all these years.

Finally, I would like to thank my parents, Georgios and Katerina, as well as my sisters, Katerina, Efi, and Maria, for supporting me and all my decisions throughout this long journey. Special thanks to my wife Eirini for her unrestrained love and support all these years and her patience during the stressful times.

Lausanne, June 11, 2020

Orion Afisiadis

Abstract

Wireless communications are currently faced with two main challenges. The first challenge stems from the enormous number of Internet of Things (IoT) devices that transmit very small amounts of data. The second challenge is the need for ever-increasing data rates required by users of multimedia rich services, as well as the extremely low latency required in emerging applications such as autonomous vehicles and augmented reality. In this thesis we deal with important physical layer (PHY) aspects that have not been analyzed in-depth in the existing literature, and whose study can help to address the aforementioned challenges.

Low-power wide-area networks (LPWANs) comprise a big part of the IoT. For energy efficiency reasons, most of LPWAN technologies adopt uncoordinated channel access schemes which result in collisions. This issue becomes more severe as the number of devices increases, putting the scalability of LPWANs at risk as they become interference-limited. To evaluate and support LPWAN scalability, in the first part of this thesis we perform a thorough analysis of the performance of one of the most important LPWAN technologies, namely LoRa. We analyze the LoRa performance in interference scenarios, and we derive expressions, as well as very accurate low-complexity approximations, for the error rate of LoRa for both the uncoded and coded cases, and with carrier frequency offset (CFO). We also propose and analyze the coherent demodulation of LoRa under interference, as a potential receiver improvement in collision scenarios. Finally, we build a standard-compatible LoRa PHY software-defined radio (SDR) prototype based on GNU Radio, which can be used for measurements of LoRa PHY performance.

The second part of this thesis focuses on full-duplex radios, which allow simultaneous transmission and reception in the same frequency band, and have been proposed as a possible solution to overcome the capacity bottleneck of high data-rate applications. However, full-duplex transceivers suffer from strong self-interference. Perfect self-interference cancellation is difficult to achieve due to the presence of strong non-linear signal components, which are introduced by hardware imperfections inherent in the transmitter and receiver chains. We propose the digital predistortion of the transmit signal to compensate for the cascade of the transceiver non-linearities and enhance self-interference cancellation. Unfortunately, a residual self-interference component always remains, particularly when operating at realistic transmit powers. To increase the usefulness of full-duplex technology, we examine communication schemes where using full-duplex transceivers can significantly improve the performance in terms of both throughput and latency, even under imperfect self-interference suppression. In particular, we examine the use of full-duplex technology in cognitive radios,

Abstract

and in communication links with asymmetric capacity requirements between the uplink and downlink channels.

Key words: Internet of Things (IoT), LoRa, LoRaWAN, In-Band Full-Duplex, Cognitive Radio, Wireless Communications, Software-Defined Radio (SDR).

Résumé

Les communications sans fil sont actuellement confrontées à deux grands défis. Le premier défi provient du nombre énorme d'appareils utilisant l'internet des objets (IoT) qui transmettent de très petites quantités de données. Le second défi est le besoin de débits de données toujours plus élevés pour les utilisateurs de services multimédias, tout comme la latence extrêmement faible des applications émergentes telles que les véhicules autonomes et la réalité augmentée. Dans cette thèse, nous traitons des aspects importants de la couche physique qui n'ont pas été analysés en profondeur dans la littérature existante et dont l'étude peut aider à relever les défis mentionnés ci-dessus. Les réseaux étendus de faible puissance (LPWAN) constituent une grande partie de l'IoT. Pour des raisons d'efficacité énergétique, la plupart des technologies LPWAN adoptent des schémas d'accès aux canaux non-coordonnés qui entraînent des collisions. Ce problème s'aggrave à mesure que le nombre d'appareils augmente, ce qui met en péril l'évolutivité des réseaux LPWAN à mesure qu'ils deviennent limités par les interférences. Pour évaluer et soutenir l'extensibilité du LPWAN, dans la première partie de cette thèse, nous effectuons une analyse approfondie des performances de l'une des plus importantes technologies LPWAN, à savoir LoRa. Nous analysons la performance de LoRa dans des scénarios d'interférence et nous dérivons des expressions ainsi que des approximations très précises et peu complexes du taux d'erreur de LoRa pour les cas non-codés et codés et avec décalage de la fréquence porteuse (CFO). Nous proposons et analysons également la démodulation cohérente de la LoRa sous interférence comme amélioration potentielle du récepteur dans les scénarios de collision. Enfin, nous construisons un prototype de radio logicielle (SDR) LoRa PHY compatible avec la norme, basé sur GNU radio, qui peut être utilisé pour mesurer les performances de LoRa PHY.

La deuxième partie de cette thèse se concentre sur les radios full-duplex qui permettent la transmission et la réception simultanées dans la même bande de fréquence et qui ont été proposées comme solution possible pour surmonter la limitation de la capacité des applications à haut débit de données. Cependant, les émetteurs-récepteurs en duplex intégral souffrent d'une forte auto-interférence. Il est difficile d'obtenir une annulation parfaite de l'interférence en raison de la présence de fortes composantes de signal non-linéaires introduites par des imperfections matérielles héritées de la chaîne d'émission et de réception. Nous proposons la pré-distorsion numérique du signal d'émission pour compenser la cascade des non-linéarités du trans-receveur et améliorer l'annulation de l'auto-interférence. Malheureusement, une composante d'auto-interférence résiduelle subsiste toujours, en particulier lorsque l'on fonctionne à des puissances d'émission réalistes. Pour accroître l'utilité

Résumé

de la technologie full-duplex, nous examinons des schémas de communication dans lesquels l'utilisation d'émetteurs-récepteurs full-duplex peut améliorer considérablement les performances en termes de débit et de latence, même en cas de suppression imparfaite des auto-interférences. En particulier, nous examinons l'utilisation de la technologie full-duplex dans les radios cognitives et dans les liaisons de communication avec des exigences de capacité asymétriques entre la liaison montante et la liaison descendante.

Mots clefs : Internet des Objets (IoT), LoRa, LoRaWAN, Radio Full-Duplex Intrabande, Radio Cognitive, Communication sans Fil, Radio Logiciel.

Contents

Acknowledgements	i
Abstract (English/Français)	iii
List of Figures	xi
1 Introduction	1
1.1 Low-Power Wide-Area Networks for the Internet of Things	2
1.2 In-Band Full-Duplex Wireless Communications	3
1.3 Thesis Outline and Contributions	5
1.4 Notation	8
I LoRa Physical Layer Aspects	9
2 Introduction to LoRa and LoRaWAN	11
2.1 The LoRa Modulation	12
2.1.1 Modulation	12
2.1.2 Demodulation	13
2.2 Complete LoRa Transceiver Chain	14
2.3 LoRa Packet Structure	15
2.4 LoRaWAN Multiple Access	17
2.4.1 LoRaWAN Network Architecture	17
2.4.2 LoRaWAN Classes of Devices	18
2.4.3 Adaptive Data Rate Mechanism	19
2.5 State of the Art	20
2.5.1 LoRaWAN Applications	20
2.5.2 Reverse Engineering and SDR Implementations of the LoRa PHY	21
2.5.3 Interference Models for LoRa PHY	21
2.5.4 LoRa PHY Link Performance Evaluation through Deployments	24
2.5.5 Probabilistic Models for the Performance of LoRa PHY	26
2.5.6 LoRaWAN Scalability	26
2.6 Discussion and Summary	30
3 LoRa Error Rate under Same-Technology Interference	33

Contents

3.1	Symbol Error Rate Under AWGN	34
3.1.1	Distribution of the Demodulation Decision Metric	34
3.1.2	Symbol Error Rate Expression	34
3.1.3	Symbol Error Rate Approximations	35
3.2	Symbol Error Rate Under Same-SF LoRa Interference	36
3.2.1	Distribution of the Demodulation Decision Metric	38
3.2.2	Symbol Error Rate Expression	39
3.2.3	Complexity Reduction Using Equivalent Interference Patterns	40
3.2.4	Symbol Error Rate Approximation	44
3.2.5	Results	46
3.3	Frame Error Rate Under Same-SF Interference	49
3.3.1	Frame Error Rate Expression and Approximation	50
3.3.2	Results	51
3.4	Discussion and Summary	52
4	Coherent LoRa Detection under Same-Technology Interference	55
4.1	Coherent LoRa Detection: How and Why?	56
4.1.1	Symbol Error Rate Under AWGN	57
4.2	Coherent LoRa Symbol Error Rate Under Same-SF Interference	59
4.2.1	Symbol Error Rate Under Same-SF Interference	59
4.3	Symbol Error Rate Approximation	62
4.3.1	Interference Patterns	62
4.3.2	Symbol Error Rate Approximation	63
4.4	Coherent LoRa Frame Error Rate Under Same-SF Interference	66
4.5	Results	67
4.6	Discussion and Summary	70
5	Coded LoRa Error Rate	73
5.1	Coded LoRa Frame Error Rate Under AWGN	73
5.1.1	Codeword Error Rate	74
5.1.2	Frame Error Rate	75
5.1.3	Results	77
5.2	LoRa Coded Frame Error Rate under Carrier Frequency Offset	78
5.2.1	Distribution of the Decision Metric	78
5.2.2	Uncoded Symbol Error Rate	79
5.2.3	Complexity Reduction	81
5.2.4	Frame Error Rate under AWGN and CFO	81
5.2.5	Results	81
5.3	Discussion and Summary	84
6	An Open-Source LoRa Physical Layer Prototype on GNU Radio	85

6.1	LoRa PHY Reverse Engineering Details	85
6.1.1	Gray mapping	86
6.1.2	Interleaving	86
6.1.3	Whitening	89
6.1.4	Hamming Coding	90
6.1.5	PHY Header	92
6.1.6	Payload CRC	96
6.2	LoRa Frame Synchronization	96
6.2.1	Frame Synchronization Signal Model	96
6.2.2	Synchronization and Offset Correction	97
6.3	LoRa Testbed and Measurement Results	99
6.3.1	GNU Radio LoRa PHY Implementation	99
6.3.2	Testbed description	100
6.3.3	Testbed results	101
6.4	Discussion and Summary	101
II	In-Band Full-Duplex Wireless Communications	103
7	Introduction to Full-Duplex Transceivers	105
7.1	Self-interference Cancellation in Full-Duplex Wireless Transceivers	106
7.1.1	Analog Cancellation	107
7.1.2	Digital Cancellation	108
7.2	Full-Duplex Transceiver Impairments	109
7.2.1	DAC Non-Linearities	109
7.2.2	Phase Noise	109
7.2.3	IQ Imbalance	110
7.2.4	Power Amplifier Non-Linearities	110
7.3	State of the Art	110
7.4	Summary and Discussion	113
8	Digital Predistortion of Hardware Impairments for Full-Duplex Transceivers	115
8.1	Application of Digital Predistortion to Full-Duplex Transceivers	117
8.1.1	Characterization of Hardware Impairments in Full-Duplex Transceivers	117
8.1.2	Predistortion for All Hardware Impairments	118
8.1.3	Application to Regeneration-Based Full-Duplex Transceivers	119
8.2	Experimental Validation	120
8.3	Discussion and Summary	121
9	Full-Duplex Systems with Imperfect Self-Interference Cancellation	123
9.1	Sliding Window Spectrum Sensing for Full-Duplex Cognitive Radios	124
9.1.1	Half-Duplex and Full-Duplex Cognitive Radio System Models	125
9.1.2	Latency-Throughput Analysis	126
9.1.3	Average Latency of the Slotted Full-Duplex System	128

Contents

9.1.4	Results	129
9.1.5	Full-Duplex Cognitive Radio Demonstration Setup	132
9.2	Full-Duplex Communications for Wireless Links with Asymmetric Capacity Requirements	134
9.2.1	System Models for Half-Duplex and Full-Duplex Asymmetric Links . . .	134
9.2.2	Results	137
9.3	Discussion and Summary	141
10	Summary and Discussion	143
10.1	Summary of the Thesis	143
10.2	Discussion	144
A	Appendix	147
	Bibliography	151

List of Figures

2.1	Illustration of a DFT-based LoRa demodulation chain.	14
2.2	The LoRa PHY Tx and Rx chains.	16
2.3	The structure of a LoRa packet.	16
2.4	The structure of the LoRa preamble.	16
2.5	The LoRaWAN Network Architecture. The diagram was created based on a figure owned by Semtech that can be found at https://www.semtech.com/lora/what-is-lora	18
3.1	Symbol error rate of the LoRa modulation under AWGN for all supported spreading factors $SF \in \{7, \dots, 12\}$. Results for Monte Carlo simulations and the approximation in (3.7) are shown.	36
3.2	Illustration of LoRa uplink transmission with one interfering user having an arbitrary offset τ	37
3.3	The magnitude of an example transmitted interference pattern for $SF = 7$, $s_{I_1} = 83$, $s_{I_2} = 4$, and $\tau = 48.8$	38
3.4	Vector representation of the signal of interest (green), the interference $h_I R_s$ (black), the noise (red), and the bin value Y_s at the bin of interest s	39
3.5	Symbol error rate approximation of the LoRa modulation under same-SF interference and AWGN for $SF \in \{9, 10, 11\}$ and $SIR = 3$ dB, for various values of the oversampling factor $1/\epsilon$	46
3.6	Symbol error rate of the LoRa modulation under AWGN and same-SF interference for $SF \in \{7, \dots, 12\}$ and $SIR = 3$ dB. Black dotted lines show the SER when ignoring the phase offset ω and thick transparent lines show the SER when there is only AWGN for comparison (taken from Fig. 3.1).	47
3.7	Symbol error rate of the LoRa modulation under AWGN and same-SF interference for $SF \in \{9, 10, 11\}$ and $SIR = 3$ dB. The approximations of [1] and (3.54) are shown with black dotted lines.	48
3.8	Required SNR for a target symbol error rate of $2 \cdot 10^{-5}$ as a function of the SIR for $SF = 7$ and $SF = 12$	49
3.9	Frame error rate of the LoRa modulation for a frame length $F = 10$ under AWGN and same-SF interference for $SF \in \{9, 10, 11\}$ and $SIR = 3$ dB. The approximations of [1] and (3.54) are shown with black dotted lines.	51

List of Figures

3.10	Required SNR for a target frame error rate of 10^{-1} as a function of the SIR for SF = 7 and SF = 12.	52
3.11	Required SNR for target frame error rates of 10^{-1} and 10^{-2} , for frame lengths $F = 10, 20, 30$, as a function of the SIR for SF = 7.	53
4.1	The magnitude of an example transmitted (green) and received (black) interference pattern for SF = 7, $s_{I_1} = 83$, $s_{I_2} = 4$, $\tau = 88.4$, $ h_I = 1$, and $\lambda_{\text{cfo}} = 0.4$	60
4.2	The received interference pattern for SF = 7, $s_{I_1} = 83$, $s_{I_2} = 4$, $\tau = 88.4$, $ h_I = 1$, and $\lambda_{\text{cfo}} = 0.4$: magnitude (a); and values on the complex plane (b).	61
4.3	Simple illustration of LoRa bins projected on the real axis.	63
4.4	Frame error rate of the coherent and non-coherent receiver for three different values of CFO, and a packet of length $F = 20$ LoRa symbols under AWGN and same-SF interference for SF = 7 and $P_I = 0$ dB.	68
4.5	Symbol error rate of the coherent and non-coherent LoRa receiver under AWGN and same-SF interference for SF $\in \{7, 9, 11\}$ and $P_I = -3$ dB. The approximations for the coherent and the non-coherent case are shown with black dotted lines.	69
4.6	Frame error rate of the coherent and non-coherent receiver for a packet of length $F = 20$ LoRa symbols under AWGN and same-SF interference for SF $\in \{7, 9, 11\}$ and $P_I = -3$ dB. The approximations for all cases are shown with black dotted lines.	70
4.7	Required SNR for a target frame error rate of 10^{-1} as a function of the SIR for SF = 7 for coherent and non-coherent receiver.	71
4.8	Frame error rate of the coherent receiver under AWGN and same-SF interference with perfect estimation, and for estimation error $\sigma^2 \in \{0.2, 0.3, 0.4\}$. The packet length is $F = 20$ LoRa symbols, SF=7 and $P_I = -3$ dB. The performance of the non-coherent receiver is shown in thick transparent line.	71
5.1	LoRa deinterleaving for SF = 7 and a (4, 8) Hamming code.	75
5.2	Frame error rate of the LoRa modulation under AWGN for all supported spreading factors SF $\in \{7, \dots, 12\}$	78
5.3	The magnitude of the DFT bins for SF = 7, $s = 83$, and $\lambda_{\text{cfo}} = 0.4$	80
5.4	LoRa FER under three different values of CFO for SF = 7.	82
5.5	Frame error rate of the LoRa modulation under AWGN and CFO ($\lambda = 0.2$) for all supported spreading factors SF $\in \{7, \dots, 12\}$. The thick gray lines show the FER when only AWGN is present for comparison.	83
5.6	LoRa FER under three different values of CFO for SF = 7 with and without Gray mapping.	83
6.1	Differential patterns for CR = $4/8$ and SF $\in \{7, 8, 9\}$	87
6.2	Differential pattern for CR = $4/6$ and SF=8.	88
6.3	Differential patterns for CR = $4/5$ and SF = 8.	89
6.4	Differential patterns swapped every two for CR = $4/5$ and SF = 8.	89

6.5	Interleaving blocks for three combinations of CR and SF: (a): CR = $\frac{4}{8}$ and SF = 8, (b): CR = $\frac{4}{6}$ and SF = 7, and (c): CR = $\frac{4}{5}$ and SF = 9.	90
6.6	Whitening matrices for different coding rates	91
6.7	Bits in each deinterleaved codeword	91
6.8	Exhaustive list of whitened codewords	92
6.9	Differential message between transmissions with CR = $\frac{4}{7}$ and CR = $\frac{4}{8}$ in explicit header mode.	93
6.10	Summary matrix of many different headers	94
6.11	Header received for CR= $\frac{4}{7}$, Payload length of 16 Bytes and CRC off	94
6.12	Summary matrices for three different coding rates	95
6.13	Mask and bit corresponding to the coding rate information	95
6.14	LoRa testbed in USRPs using our GNU Radio SDR platform.	99
6.15	Uncoded and coded LoRa BER for SF = 7 (testbed vs simulation).	100
7.1	Power budget for a full-duplex transceiver with multiple self-interference suppression stages. This figure was created based on [2].	106
7.2	Full duplex transceiver architectures: (a) RF circuit-based; (b) Regeneration-based from the digital baseband.	107
7.3	Block diagram of the transmitter with the hardware components that introduce non-idealities.	109
8.1	Full duplex system architectures: (a) RF cancellation circuit; (b) regeneration from the digital baseband; and (c) proposed regeneration with digital predistortion for hardware impairments.	116
8.2	Block diagram showing the application of predistortion to compensate for the cascaded non-idealities of the transmitter hardware impairments.	118
8.3	Application of the proposed predistortion method in regeneration-based full-duplex transceivers.	119
8.4	Full-duplex testbed used to obtain the experimental results.	120
8.5	Experimentally measured self-interference power spectral density with: (a) no digital predistortion; (b) and predistortion only for the power amplifier.	121
8.6	Experimentally measured self-interference power spectral density with predistortion of all significant hardware impairments.	122
9.1	Frame Structures	125
9.2	Average latency - normalized throughput comparison for the 3 systems, SNR _{PU} = 0 dB with perfect self-interference suppression, $N_s = 16$, $N = 128$	130
9.3	Maximum acceptable latency - normalized throughput comparison for the 95% and 99% quantiles for the 3 systems, primary user SNR _{PU} = 0, $N_s = 16$, $N = 128$	130
9.4	Average access latency with increasing residual self-interference, for a normalised throughput of 0.9, SNR _{PU} = 0 dB, $N_s = 16$, $N = 128$	131
9.5	Normalised throughput with increasing residual self-interference, for an average access latency of 16 samples, SNR _{PU} = 0 dB, $N_s = 16$, $N = 128$	131

List of Figures

9.6	Annotated photograph of the full-duplex cognitive radio demonstration setup.	133
9.7	Allocation of uplink and downlink channels for time-division duplexing (TDD). For asymmetric scenarios, TDD can allocate a larger proportion of the frame to the high-rate link, up to a limit defined by the standard for the minimum uplink rate.	135
9.8	Capacity comparison for HD and FD over the link distance range for the IEEE 802.11 scenario, with self-interference suppression of 100 dB.	138
9.9	Uplink transmit power comparison for HD and FD over the link distance range for the IEEE 802.11 scenario, with self-interference suppression of 90, 100 and 110 dB.	139
9.10	Capacity comparison for HD and FD over the link distance range for the LTE scenario, with self-interference suppression of 120 dB.	139
9.11	Downlink (a) and uplink (b) transmit power comparison for HD and FD over the link distance range for the LTE scenario, with self-interference cancellation of 110, 120 and 130dB.	140
9.12	Impact of different rate asymmetry on the useful range of a full-duplex system compared to LTE-TDD for different self-interference cancellation capability. . .	141

1 Introduction

One of the greatest challenges in today's wireless communications is the sustainability of networks which wirelessly connect a rapidly increasing number of devices. In a few years, almost every device that we use in our everyday life will be connected to the Internet of Things (IoT). Most of the things connected to the IoT typically need to transmit a very small amount of information, such as temperature, humidity, and pressure values or position coordinates. Devices used in low-power wide-area networks (LPWANs) need to be able to autonomously transmit at long ranges and for many years. For this reason, very low-power wireless-transmission protocols need to be applied. Most of the very low-power transmission protocols are based on uncoordinated channel access schemes, such as the ALOHA protocol. For this reason, LPWANs result in ad-hoc transmissions which interfere with each other. Therefore, the main challenge for LPWANs is that they are *interference limited*, rather than noise limited.

The second big challenge of contemporary wireless communications is the ever-increasing data rate demands of everyday users using multimedia-rich services. Furthermore, emerging applications such as autonomous vehicles and augmented reality require extremely low end-to-end latency. During the last years, LTE-Advanced and 5G have tried to accommodate these increasing demands for high-throughput and low-latency systems. In particular, 5G paves the way for small-cell networks which allow reusing the limited radio spectrum in a more effective way. In addition, 5G standards set the latency requirements to less than 1 ms. Most contemporary systems use schemes which separate the transmission and the reception in orthogonal channels. The separation is achieved either in time, utilizing time-division duplexing (TDD) or in frequency, utilizing frequency-division duplexing (FDD). These approaches introduce overhead due to the required guard intervals and guard bands, respectively. Moreover, high latency can be introduced, especially for TDD systems.

In this thesis we address physical layer aspects related to both aforementioned challenges in wireless communications. On one hand, since massive low data rate LPWANs are interference limited, we model the physical layer performance with and without interference, of one of the most popular LPWAN solutions, namely LoRaWAN. On the other hand, to partially overcome

the capacity and latency bottlenecks of the high data rate wireless networks, we examine the adoption of the in-band full-duplex technology in such networks. For both topics, we perform our analysis considering the limitations introduced by practical transceiver aspects such as hardware impairments. In addition, our analysis results in two corresponding software-defined radio (SDR) implementations, which allow the experimental performance evaluation and demonstration of the systems under consideration. In the remainder of this introductory chapter, we will discuss more on these two directions, and present an outline of the thesis.

1.1 Low-Power Wide-Area Networks for the Internet of Things

The number of devices with wireless connectivity has vastly increased in the last years, and is expected to grow even more rapidly in the years to come [3, 4]. This increase in connectivity is mainly due to devices that are used in IoT applications, and to a much lesser extent due to newly connected high data rate multimedia-rich devices. These IoT devices typically need to wirelessly transmit small amounts of information. The reason why more and more devices connect to the IoT is because of the wide range of useful applications it can provide. These applications include industrial and environmental sensing, energy and smart grid applications, smart metering, agricultural monitoring, health monitoring, localization and tracking, logistics, and smart city applications.

This broad scope of applications comes with highly diverse wireless connectivity requirements. For this reason, the IoT standards landscape is very rich [5]. On the short-range wireless connectivity side, Bluetooth Low Energy (BLE) [6], ZigBee [7], and 802.15.4 [8] are predominant standards. These technologies provide a single-hop coverage area, which can be extended by using multihop relaying of the data in a mesh network. Most of the IoT applications though require some type of direct long-range wireless connectivity. We can distinguish between two main categories for low-power wide-area networks (LPWANs), namely licensed LPWANs (e.g., cellular IoT) and unlicensed LPWANs. The main common characteristic of the two categories is the wide area of coverage, on the range of a few kilometers for urban areas and tens of kilometers for rural areas.

Cellular IoT standards, such as NB-IoT, are an evolution of existing 3GPP cellular standards. For this reason cellular IoT technologies have relatively complex physical (PHY) and medium-access control (MAC) layers. Cellular IoT standards cover many use cases and use licensed frequency bands to provide robustness and security. Since cellular standards have been designed to cover high data rate services for smaller number of devices, they face the challenge of a fundamental paradigm shift to low data rate services for a much larger number of devices.

On the other hand, unlicensed LPWAN technologies are not an evolution of any existing standard. The PHY and MAC layers of the unlicensed LPWAN technologies have been designed from scratch for the IoT, and are therefore relatively simple. Some examples of unlicensed technologies include Sigfox [9], Weightless [10], Ingenu [11], and LoRaWAN [12, 13, 14]. All these technologies use the unlicensed industrial, scientific and medical (ISM) bands. As

such, they have lower operational and infrastructure costs than the cellular IoT. For energy efficiency reasons, unlicensed LPWANs use uncoordinated MAC layer schemes, in contrast to the centralized MAC protocols of the cellular IoT.

Uncoordinated MAC protocols are either based on the ALOHA scheme or on listen-before-talk (LBT) schemes. They are usually chosen as channel access mechanisms for networks with a large and variable number of connected devices with unpredictable traffic. Slotted ALOHA divides time into predefined slots and allows every device to start transmitting packets only at the beginning of each slot. Slotted ALOHA needs some type of synchronization in time, often achieved by the periodic broadcast of beacons. However, very low power networks mostly rely on pure ALOHA, which is the lowest-power channel access solution. In pure ALOHA, every device attempts to transmit a packet whenever a new message is generated. Depending on the application, the end-node devices can ask for acknowledgments (ACKs) or not. If a packet is not positively acknowledged, a retransmission may be attempted.

The total lack of coordination in pure ALOHA channel access schemes leads to reduced network throughput due to packet collisions when the number of connected devices becomes large. Moreover, a large number of retransmission attempts can reduce the energy efficiency [3]. In order to alleviate the problems arising from the increase in the number of connected devices, LPWAN standards often introduce constraints regarding the maximum allowed transmission time for each device, which are typically referred to as *duty-cycle* restrictions.

In this thesis we consider one of the most popular unlicensed LPWAN technologies, namely LoRaWAN. LoRaWAN, and its physical layer, called LoRa, use spread-spectrum modulation to achieve long-range connectivity and resilience to noise and interference. LoRaWAN adopts a pure ALOHA access scheme for the vast majority of the devices, and thus faces the risk to become interference limited. Therefore, in this thesis, we analyze the performance of LoRa with and without interference, and we propose simple physical layer modifications to increase the resilience of LoRa under interference.

1.2 In-Band Full-Duplex Wireless Communications

In-band *full-duplex* is a relatively recent technology that promises to double the spectral efficiency compared to half-duplex by allowing simultaneous transmission and reception in the same frequency band [15, 16, 17, 18, 19]. The capacity advantage of full-duplex when compared to existing systems is particularly important, as congestion in the radio spectrum suitable for wireless communications has significantly increased the access cost, and limits the maximum transmission rates and the number of devices that can operate concurrently, and leads to larger latencies [20].

However, full-duplex nodes suffer from self-interference generated by the transmitter chain. The self-interference signal is received at orders of magnitude higher power than the desired signal. As the baseband transmit signal is known within the full-duplex transceiver, it is in

principle possible to generate an appropriate cancellation signal that effectively suppresses the self-interference, ideally to (or below) the receiver noise-floor [20]. However, the need to avoid overloading the low-noise amplifier (LNA), and the limited dynamic range of the analog-to-digital converters (ADCs), impose practical constraints, such that the cancellation signal must typically be applied immediately following the receive antenna. These limitations necessitate the generation of the cancellation signal in the RF domain [17, 2]. Perfect self-interference cancellation is difficult to achieve in practice due the presence of strong non-linear signal components, which are introduced by various hardware imperfections inherent in the transmitter and receiver chains [20]. Thus, in many reported implementations, a residual self-interference component remains, particularly when operating at realistic transmit powers [17, 2, 21]. This residual self-interference requires full-duplex systems to operate at lower transmission rates to maintain error-free operation, i.e., the link capacity may *not* be doubled compared to the half-duplex case.

Nevertheless, doubling the capacity, which necessitates perfect self-interference cancellation, is not the only reason to adopt full-duplex technology in wireless communication systems. Full-duplex systems with residual self-interference can be very useful in wireless communication systems where some of the nodes inherently include some type of processing asymmetry in their receiver. In such systems, instead of aiming to double the capacity, full-duplex can be applied to reduce the end-to-end latency, while maintaining or even increase the throughput. This approach can be particularly beneficial in latency-critical 5G and beyond-5G systems. In this thesis we examine the use of full-duplex technology in two such systems, namely cognitive radios and links with asymmetric capacity requirements.

Cognitive radio

One of the most significant challenges faced by wireless systems today is the scarcity of the available spectrum. However, recent studies have shown that, while the radio spectrum is densely allocated, it is often not heavily occupied or utilized by the licensed primary users (PUs) [22, 23]. Frequency-agile cognitive radio networks have been proposed to take advantage of this situation, by allowing unlicensed secondary users (SU) to opportunistically reuse licensed frequency bands [24]. One of the fundamental requirements of such systems is that SUs should not generate harmful interference to PUs. Consequently, SU transceivers must be capable of sensing the radio channel to determine if a PU is present [25]. Similarly to spectrum reuse, we can also consider a scenario that requires low-latency medium access for high-priority users or high-priority transmissions. In such systems, a less latency-sensitive ongoing transmission would need to stop instantly as a PU accesses the channel to transmit its high-priority message. Such scenarios are of particular importance for real-time services, such as virtual/augmented reality and autonomous vehicles. System architectures that periodically stop SU transmissions to sense the channel have been widely proposed [26, 27] to detect the start of PU transmissions. Due to the necessary sensing intervals, these conventional cognitive radio system architectures inevitably reduce the throughput of the SUs. Furthermore,

conventional cognitive radio system architectures introduce blind intervals where the SU system is unable to detect the come-back of a PU. Secondary users with simultaneous transmission and sensing capabilities, using self-interference cancellation technology, have been proposed in the last years as a solution for cognitive radio to overcome the aforementioned problems [28, 29, 30, 31]. In this thesis we perform an analysis on the sensing-throughput trade-off of full-duplex cognitive radio with residual self-interference, and we propose a sliding-window sensing method. We show that the proposed method is particularly beneficial for SU full-duplex transceivers with residual self-interference.

Asymmetric links

Asymmetries between up-link and down-link throughput requirements are common in modern wireless systems, often reflecting typical usage patterns, e.g., web-browsing and video/audio streaming [32, p. 459]. Asymmetry is also often inherent in the structure of wireless communication protocols, e.g., ACK frames that are sent back to confirm the successful decoding of the data frames. The amount of information contained in the ACK up-link frame is very small compared to the downlink frame containing the data. Current wireless communications systems that use either TDD or FDD to separate the downlink and the uplink, adapt to the link asymmetries by readjusting their time and frequency resources [33]. However, such a readjustment is difficult or even impossible to perform in highly asymmetric links, due to minimum constraints in time or frequency resources for both links. To overcome the aforementioned problem, we examine the use of full-duplex technology on asymmetric links, and we show the benefits of such an approach. Full-duplex transceivers can use transmit power allocation as a tuning control to always adapt to the link asymmetry, thus increasing the overall spectral efficiency.

1.3 Thesis Outline and Contributions

In summary, these are the main contributions of the thesis. We perform an analysis of important open physical layer aspects for both LoRa and full-duplex technologies. In particular, we include in our models hardware impairments that can act as a limiting factor to the performance. Our analysis is the basis on which we build SDR implementations of the examined systems. First, we perform a thorough analysis of the LoRa physical layer, and we derive expressions for the LoRa error rate under interference and noise. We gradually introduce models for both the uncoded and coded modes of LoRa, as well as for one of the main hardware impairments. Our models result in low-complexity approximations for the performance of LoRa for all the aforementioned cases. Our low-complexity approximations are shown to be very accurate in the useful operating regimes of LoRa, and can directly be used as inputs to network simulators in order to assess the scalability of LoRaWAN. Second, we present our fully-functional open-source SDR implementation of LoRa PHY that, to the best of our knowledge, is the only implementation in the literature that works in the low-SNR

operating regime of LoRa. Our experimental LoRa performance results are within 1 dB of the theoretical LoRa performance, which indicates the correctness of the implementation. Therefore, our SDR implementation can serve as a performance evaluation prototype of future receiver algorithms for improved LoRa performance. Third, to alleviate the limited analog self-interference cancellation of regeneration-based full-duplex transceivers, we propose a predistortion method which jointly compensates for the cascade of the transceiver's hardware impairments. The improved self-interference suppression of the proposed method is experimentally validated using a full-duplex SDR testbed. Finally, we examine the benefits of using full-duplex transceivers with imperfect self-interference cancellation capabilities in wireless communication systems, in particular in cognitive radio systems and in links with asymmetric capacity requirements. In the following we will summarize the contents of each chapter in this thesis.

Part I: LoRa Physical Layer Aspects

In Chapter 2, we present an introduction to LoRaWAN and to its physical layer. In particular, we start with the physical layer description by presenting the LoRa modulation and demodulation procedures in Section 2.1, in Section 2.2 the LoRa transceiver chain, and in Section 2.3 the LoRa packet structure. In Section 2.4 we present a brief description of the MAC layer characteristics that affect the scalability of LoRaWAN. Finally, in Section 2.5 we discuss the state of the art in LoRaWAN research. In particular, we present the numerous application fields of LoRaWAN, the related research on physical layer interference models, and the evaluation of the performance of LoRa in the physical layer. We continue with the research on the scalability of LoRaWAN and finally we present the ongoing research on LoRa SDR platforms.

Chapter 3 describes our work on the performance analysis of uncoded LoRa under interference. We note that this chapter is derived from our works of [34, 35]. In Section 3.1 we derive an expression for the LoRa symbol error rate (SER) in AWGN, and in Section 3.2 we derive an expression for the SER of LoRa under interference from another LoRa user. We also derive an accurate low-complexity approximation for the evaluation of the derived expression. In addition, we show how to reduce the complexity of calculating the SER. Finally, in Section 3.3 we derive the expression for the calculation of the LoRa FER under interference, since the frame error rate (FER) is more useful as an input in network simulators which evaluate the network throughput.

In Chapter 4 we propose the coherent detection of LoRa symbols as a potential method to improve the performance of LoRa under interference. In Section 4.1, we present the intuition behind our proposal. Following the methodology of Chapter 3 for non-coherent detection, in Section 4.2 we first present the uncoded error rate analysis when only AWGN is present, and then we derive an expression for the error rate of a coherent receiver under interference from another LoRa user. We continue in Section 4.3 with the derivation of a low-complexity approximation for the evaluation of the derived error-rate expression. Finally, we perform the FER analysis for the coherent LoRa receiver, and we compare the performance

to the standard non-coherent receiver, showing the benefits of coherent LoRa detection in interference scenarios.

In Chapter 5 we perform an analysis of the performance of the coded modes of LoRa, under AWGN and carrier frequency offset (CFO). We note that this chapter is derived from our work of [36]. We show that the channel coding, the interleaving, and the Gray mapping blocks of a LoRa transceiver are built in such a way to be resilient not only to AWGN, but also to CFO. In particular, in Section 4.1 we explain the joint impact of the channel coding, interleaving, and Gray mapping blocks of LoRa transceivers on the demodulation bit errors. Specifically, we first derive the codeword error rate (CWER) and the FER of coded LoRa transmissions under AWGN. In Section 5 we extend our analysis to derive an expression and a low-complexity approximation for the FER under the impact of residual CFO. We present results showing the positive impact of coding, interleaving, and Gray mapping on the LoRa performance under residual CFO, which is particularly important for low-cost LoRa hardware modules.

In Chapter 6 we present our open-source SDR implementation of the LoRa physical layer in GNU Radio. This chapter is derived from our work of [37]. In Section 6.1 we first present some details of our reverse engineering effort concerning the LoRa physical layer, which include small differentiations from previous incomplete approaches in the literature. Moreover, in Section 6.2, we explain the LoRa packet synchronization procedure we use in our implementation, along with the sampling time offset (STO) and CFO estimation and compensation algorithms that are necessary both for the correct packet synchronization and the correct demodulation of the payload. Finally, in Section 6.3, we describe our GNU Radio implementation. Moreover, we use the SDR implementation to build a hardware testbed that provides experimental results on the error rate of LoRa. Our experimental performance results are the first ones reported in the literature that are obtained for low SNRs, indicating the correctness of our implementation. The measured performance using our testbed is shown to be no more than 1 dB away from the theoretical LoRa performance, both for the uncoded and the coded modes.

Part II: In-Band Full-Duplex Wireless Communications

Chapter 7 serves as a brief introduction to full-duplex technology. In Section 7.1 we discuss the self-interference problem in full-duplex transceivers, and we explain why self-interference cancellation is needed both in the radio frequency (RF)/analog domain and in the digital baseband domain. Moreover, we discuss the two main full-duplex architectures that have been presented in the literature, namely the circuit-based and the regeneration-based architectures. In Section 7.2, we present the transceiver hardware impairments that render self-interference cancellation a non-trivial task. Finally, in Section 7.3 we present the state of the art in those full-duplex research fields that are connected to this thesis. In particular, we present the research in non-linear digital self-interference cancellation methods, in full-duplex cognitive radios, and in resource allocation for full-duplex systems.

In Chapter 8, we propose the digital predistortion of the transmitter and cancellation chains

for regeneration-based full-duplex transceivers, which suffer more from transceiver non-linearities. We note that this chapter is derived from our work of [38]. In Section 8.1 we perform a characterization of the cascaded chain of hardware impairments and we show how the digital predistortion of the transmitted baseband signal can effectively linearize the transmitted RF signal. In Section 8.2 we present our full-duplex hardware testbed that we use to experimentally validate the proposed method. The results show that digital predistortion helps to achieve improved levels of self-interference cancellation already from the analog stage. This is particularly beneficial to avoid over-provisioning the RF components and the analog-to-digital converter. However, perfect self-interference cancellation is not achieved, and a small amount of residual self-interference is present.

For this reason, in Chapter 9, we study the use of full-duplex transceivers with residual self-interference in wireless systems. We note that this chapter is derived from our works of [39, 40]. We discuss the fact that the impact of residual self-interference cancellation cannot be neglected. However, residual self-interference cancellation is not necessarily catastrophic for all systems. Specific wireless systems can in fact benefit from full-duplex in terms of throughput, latency and low transmit power levels, compared to their half-duplex counterparts, even under the presence of residual self-interference. In Section 9.1, we examine cognitive radio systems, and we derive the access-latency of the primary user (PU) both when the secondary user (SU) operates in half-duplex or in full-duplex mode. Moreover, we propose a sliding-window sensing technique for full-duplex cognitive radio that further reduces the PU access-latency and renders the sensing process of the SUs more resilient to residual self-interference. We also present our proof-of-concept full-duplex cognitive radio SDR implementation. In Section 9.2, we propose the use of full-duplex technology in links with asymmetric capacity requirements. We show that full-duplex systems can more easily adjust to the link asymmetry by adjusting their transmit power, and can profit both in terms of throughput and lower transmit power compared to their half-duplex counterparts.

1.4 Notation

Bold lowercase letters (e.g., \mathbf{a}) denote vectors, while bold uppercase letters (e.g., \mathbf{A}) denote the discrete Fourier transform (DFT) of \mathbf{a} , i.e., $\mathbf{A} = \text{DFT}(\mathbf{a})$. Bold calligraphic letters (e.g., \mathcal{A}) denote matrices. We define $[x]_y = x \bmod y$. We denote the normal and complex normal (with i.i.d. components) probability density functions (PDFs) with mean μ and variance σ^2 as $\mathcal{N}(\mu, \sigma^2)$ and $\mathcal{CN}(\mu, \sigma^2)$, respectively. Moreover, we denote the PDF and the cumulative density function (CDF) of the Rayleigh and Rice distributions by $f_{\text{Ra}}(y; \sigma)$, $f_{\text{Ri}}(y; \nu, \sigma)$ and $F_{\text{Ra}}(y; \sigma)$, $F_{\text{Ri}}(y; \nu, \sigma)$, respectively, where σ and ν denote the *scale* and *location* parameters.

LoRa Physical Layer Aspects

Part I

2 Introduction to LoRa and LoRaWAN

LoRaWAN is a very popular LPWAN communications protocol for the Internet of things (IoT). Its physical layer (PHY), which is called LoRa, is based on a proprietary spread spectrum modulation scheme that uses chirp modulation as its basis [41]. Along with the chirp modulation, the LoRa PHY chain includes whitening, channel coding, interleaving and Gray mapping [42]. LoRa is able to work in a wide range of operational signal-to-noise ratios (SNRs), due to the support of multiple spreading factors (SF) and code rates.

The physical layer of LoRa is the only part of LoRaWAN which is not open-source. Therefore, only part of its technical details have been made publicly available by its designers, while the rest have been revealed through reverse-engineering attempts. On the contrary, the details of the medium access control (MAC) of LoRaWAN have been clearly described from the very beginning of this technology. One of the main contributions of this thesis is to provide an in-depth analysis of the performance of the LoRa PHY, both as a basis for future PHY algorithms that can directly improve the performance of LoRa receivers, but also to serve as a solid basis for the correct evaluation and further development of LoRaWAN, which is typically performed in higher layers of abstraction. Therefore, in order to perform the aforementioned performance analysis, a clear description of the LoRa PHY fundamentals, along with their connection to the channel access scheme of LoRaWAN is necessary.

The aim of this chapter is to serve as a self-contained introduction to LoRa and LoRaWAN, by providing a description of the LoRa PHY fundamentals, as well as a brief description of some basic MAC layer characteristics that are necessary in order to understand the current performance limitations of LoRa networks. Moreover, this chapter provides an overview of the related work in order to clearly position the contributions of this thesis.

More specifically, in Section 2.1 we begin with a description of the LoRa chirp modulation and demodulation, along with a description of all the additional tasks present in the transmitter (Tx) and receiver (Rx) chains in Section 2.2. The complete structure of a LoRa packet is presented in Section 2.3. We continue in Section 2.4 with the description of the basic different channel access schemes of LoRaWAN, as well as the adaptive data rate (ADR) mechanism,

which is the main way of dynamically choosing the wireless transmission parameters to optimize the LoRaWAN throughput. Finally, in Section 2.5, we present an overview of the state-of-the-art research work that has been conducted in the field during the last five years.

2.1 The LoRa Modulation

LoRa is a low-rate, low-power, and high-range modulation that uses a chirp spread-spectrum modulation for its physical layer [41]. LoRa supports multiple spreading factors, coding rates, and packet lengths, to support a very wide range of signal-to-noise ratios (SNRs). The LoRa physical layer is proprietary [42], but reverse engineering attempts [43, 44] have led to detailed mathematical descriptions [45, 46, 35, 47]. More recently, the effects of hardware impairments, such as carrier- and sampling frequency offsets, on LoRa digital receivers have also been modeled and analyzed [48, 49, 50]. In this section, we briefly summarize the LoRa modulation and how the demodulation can be performed.

2.1.1 Modulation

LoRa is a spread-spectrum modulation that uses a bandwidth B and $N = 2^{\text{SF}}$ chips per symbol, where SF is called the *spreading factor* with $\text{SF} \in \{7, \dots, 12\}$. The continuous-time baseband-equivalent equation of a modulated LoRa symbol s is [48, 46]

$$x_s(t) = \begin{cases} e^{j2\pi\left(\frac{B}{2T_s}t^2 + \left(s\frac{B}{N} - \frac{B}{2}\right)t\right)}, & 0 \leq t < t_{\text{fold}}, \\ e^{j2\pi\left(\frac{B}{2T_s}t^2 + \left(s\frac{B}{N} - \frac{3B}{2}\right)t\right)}, & t_{\text{fold}} \leq t < T_s, \end{cases} \quad (2.1)$$

where $T_s = \frac{N}{B}$ is the symbol duration, and $t_{\text{fold}} = \frac{N-s}{B}$.

When considering the discrete-time baseband-equivalent signal, the bandwidth B is split into N frequency steps. A symbol $s \in \mathcal{S}$, where $\mathcal{S} = \{0, \dots, N-1\}$, begins at frequency $\left(\frac{sB}{N} - \frac{B}{2}\right)$. The frequency increases by $\frac{B}{N}$ at each chip until it reaches the Nyquist frequency $\frac{B}{2}$. When the Nyquist frequency is reached, there is a frequency fold to $-\frac{B}{2}$ at chip $n_{\text{fold}} = N - s$ [46]. The general discrete-time baseband equivalent equation of a LoRa symbol s is

$$x_s[n] = \begin{cases} e^{j2\pi\left(\frac{1}{2N}\left(\frac{B}{f_s}\right)^2 n^2 + \left(\frac{s}{N} - \frac{1}{2}\right)\left(\frac{B}{f_s}\right)n\right)}, & n \in \mathcal{S}_1, \\ e^{j2\pi\left(\frac{1}{2N}\left(\frac{B}{f_s}\right)^2 n^2 + \left(\frac{s}{N} - \frac{3}{2}\right)\left(\frac{B}{f_s}\right)n\right)}, & n \in \mathcal{S}_2, \end{cases} \quad (2.2)$$

where $\mathcal{S}_1 = \{0, \dots, n_{\text{fold}} - 1\}$ and $\mathcal{S}_2 = \{n_{\text{fold}}, \dots, N - 1\}$. In the practically relevant case where the sampling frequency f_s is equal to B , which we assume for the remainder of this manuscript, the discrete-time baseband equivalent description of a LoRa symbol s can be simplified to

$$x_s[n] = e^{j2\pi\left(\frac{n^2}{2N} + \left(\frac{s}{N} - \frac{1}{2}\right)n\right)}, \quad n \in \mathcal{S}. \quad (2.3)$$

The recent work of [46] showed that the continuous-time LoRa chirp in (2.1) in fact occupies a bandwidth that is slightly larger than B , so that a sampling rate $f_s = B$ introduces aliasing. Nevertheless, (2.2) and (2.3) are very good approximations of the actual samples and are analytically tractable.

After transmission over a time-invariant and frequency-flat wireless channel with complex-valued channel gain $h \in \mathbb{C}$, the received LoRa symbol is given by

$$y[n] = hx_s[n] + z[n], \quad n \in \mathcal{S}, \quad (2.4)$$

where $z[n] \sim \mathcal{CN}(0, \sigma^2)$ is complex AWGN with variance $\sigma^2 = \frac{N_0}{2N}$ and N_0 is the single-sided noise power spectral density. The channel is $h = |h|e^{j\phi}$, where $\phi = \angle h$ denotes a phase shift introduced by the transmission channel that we assume to be fixed for each transmitted packet, but generally uniformly distributed in $[0, 2\pi)$. This is essentially an AWGN channel, but with a phase shift, where we assume $|h| = 1$ without loss of generality, and since we defined the transmitted signal to have unit power, the received signal-to-noise ratio (SNR) before despreading is

$$\text{SNR} = \frac{1}{N_0}. \quad (2.5)$$

2.1.2 Demodulation

To demodulate the symbols, the inner product of the received signal with all the possible symbols $k \in \mathcal{S}$ is computed as

$$X_k = \sum_{n=0}^{N-1} y[n] x_k^*[n] \quad (2.6)$$

$$= \underbrace{|h|}_{=1} \sum_{n=0}^{N-1} e^{j2\pi(\frac{s-k}{N})n+\phi} + \sum_{n=0}^{N-1} z[n] x_k^*[n] \quad (2.7)$$

$$= \sum_{n=0}^{N-1} e^{j2\pi(\frac{s-k}{N})n+\phi} + \tilde{z}_k, \quad (2.8)$$

where $\tilde{z}_k \sim \mathcal{CN}(0, N\sigma^2)$. In a typical non-coherent LoRa receiver a symbol estimate \hat{s} is obtained as

$$\hat{s} = \underset{k \in \mathcal{S}}{\operatorname{argmax}} (|X_k|). \quad (2.9)$$

The SF-bit label of \hat{s} , which is defined in the LoRa standard, corresponds to an estimate of the SF transmitted bits.

The complexity of computing (2.9) is $O(N^2)$. The following equivalent and low-complexity method can also be used to perform the demodulation. First, a *dechirping* is performed, where the received signal is multiplied by the complex conjugate of a reference signal x_{ref} . A typical



Figure 2.1 – Illustration of a DFT-based LoRa demodulation chain.

choice for this reference signal is a pure *upchirp*, i.e., the LoRa symbol for $s = 0$

$$x_{\text{ref}}[n] = e^{j2\pi\left(\frac{n^2}{2N} - \frac{n}{2}\right)}, \quad n \in \mathcal{S}. \quad (2.10)$$

Then, the non-normalized discrete Fourier transform (DFT) is applied to the dechirped signal in order to obtain $\mathbf{Y} = \text{DFT}(\mathbf{y} \odot \mathbf{x}_{\text{ref}}^*)$, where \odot denotes the Hadamard product and $\mathbf{y} = [y[0] \ \dots \ y[N-1]]$ and $\mathbf{x}_{\text{ref}} = [x_{\text{ref}}[0] \ \dots \ x_{\text{ref}}[N-1]]$. Non-coherent demodulation can be performed by selecting the frequency bin index with the maximum magnitude

$$\hat{s} = \arg \max_{k \in \mathcal{S}} (|Y_k|). \quad (2.11)$$

Using the fast Fourier transform (FFT), the complexity of computing (2.11) is $O(N \log N)$. These demodulation steps are illustrated in Fig. 2.1.

2.2 Complete LoRa Transceiver Chain

In this section, we provide some background on the complete LoRa transceiver chain, since apart from the modulation and demodulation explained in the previous section, the transmitter and receiver chains also perform some additional processing, as shown in Fig. 2.2. In particular, the LoRa transmitter chain performs whitening, Hamming encoding, interleaving, and Gray mapping prior to the chirp modulation. The receiver performs Gray demapping, deinterleaving, Hamming decoding, and dewhitening. Most of the details of these blocks were revealed during reverse engineering attempts. The contributions of existing reverse engineering works as well as of our own reverse engineering work are explained in Section 6.1.

Whitening

The payload data to be transmitted may contain long sequences of either ones or zeros, introducing a DC-bias which results in the signal to have non-uniform power distribution over the transmission bandwidth [41]. Moreover, such long sequences of ones or zeros introduce data dependencies on the receiver side, before the decoder [41], since the same LoRa symbol may be received multiple consecutive times. For the above reasons, the LoRa transceiver chain includes a whitening and dewhitening block, which perform an XOR of the information bits with a pseudo-random sequence.

Error-Correction Coding

LoRa uses simple schemes for error detection and error correction. LoRa supports four error-correction coding (ECC) rates $CR \in \{4/5, 4/6, 4/7, 4/8\}$ [42], and a totally uncoded mode. For the three lowest rates, LoRa uses (k, n) Hamming codes with $k = 4$ and $n \in \{6, 7, 8\}$, where k is the data length and n is the codeword length. The $(4, 6)$ Hamming code is a punctured version of the standard $(4, 7)$ Hamming code, and can detect all single-bit errors as well as some double-bit errors. The standard $(4, 7)$ Hamming code can correct all single-bit errors. The $(4, 8)$ Hamming code, which is an extended version of the $(4, 7)$ Hamming code, can correct all single-bit errors, and in addition, can detect all double-bit errors or correct some double-bit errors. For $CR = 4/5$, LoRa uses an even parity-check code instead of a (punctured) $(4, 5)$ Hamming code. The generator and parity-check matrices, are described in detail in Section 6.1.

Interleaving

LoRa uses a diagonal interleaver to distribute the (up to SF) bit errors resulting from a symbol error over multiple (SF) codewords. The combination of interleaving with the ECC leads to a higher probability of correctly decoded codewords since most codewords will only contain a single bit error, as explained in detail in Chapter 5. Furthermore, if a single symbol error occurs in a deinterleaving block, there is no way that a codeword of that block will contain more than one error.

Gray mapping

LoRa uses a reverse Gray code for the mapping from bits to symbols. Thus, a symbol error that mistakes a symbol for one of its adjacent symbols (also called a ± 1 *demodulation error*) only causes *a single* bit error, which can always be corrected by the Hamming codes with $CR = 4/7$ and $CR = 4/8$. This property is particularly useful if CFO or STO can not fully be corrected, which typically leads to demodulation errors that deviate by ± 1 from the actual symbol index [42, 49, 50, 36].

2.3 LoRa Packet Structure

The structure of a LoRa packet, which has been explained in detail in [42, 41, 44], is shown in Fig. 2.3. The LoRa packet consists of a preamble, an optional PHY Header, the PHY payload, and an optional CRC of the payload.

Preamble

The first part of a LoRa packet is the preamble, which is used for packet synchronization and frequency offset estimation. The details of how the synchronization and the offset estimation

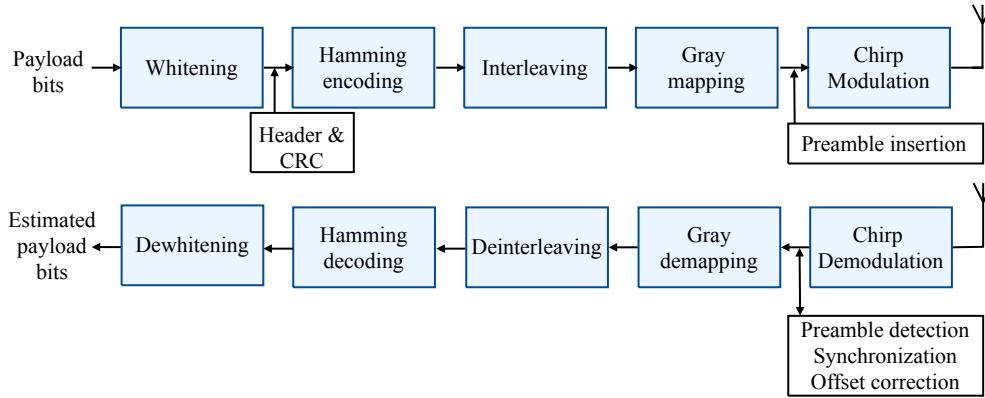


Figure 2.2 – The LoRa PHY Tx and Rx chains.

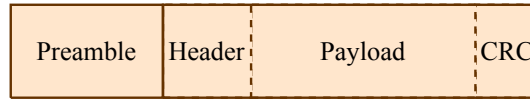


Figure 2.3 – The structure of a LoRa packet.

and compensation can be performed using the preamble will be presented in Section 6.2. The preamble consists of a programmable part of N_{pr} upchirps, and a fixed part consisting of two network identifier symbols and two-and-a-quarter downchirps, as can be seen in Fig. 2.4.

- **Upchirps:** The preamble begins with a programmable number of upchirps, i.e., symbols \mathbf{x}_0 , used to detect the existence of a LoRa packet. The default value for N_{pr} is 8, but generally the number of upchirps in the preamble can be in the range $N_{pr} \in \{6, \dots, 65532\}$ [41]. Such a wide range of preamble lengths enables the preamble detection for a very large range of SNRs.
- **Network Identifiers:** After the preamble, the packet contains two symbols that are used for frame synchronization and to distinguish between devices from different networks. In [42] it is mentioned that the network identifier symbols are modulated as $\{x, N - x\}$, where x is the network identifier, and that they should have a minimum distance of three for different networks to avoid problems caused by ± 1 demodulation errors. However, it

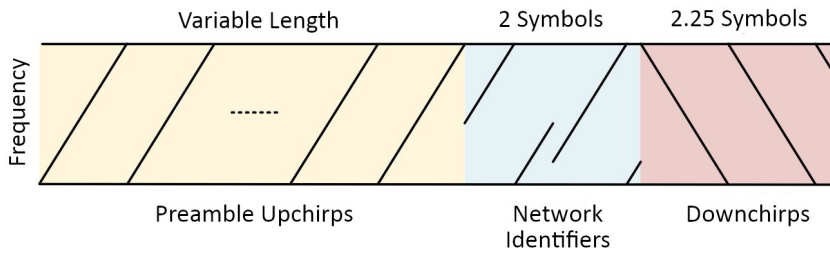


Figure 2.4 – The structure of the LoRa preamble.

is also mentioned in [49] that the network identifiers observed were actually of the form $\{x, x\}$. When the receiver demodulates two symbols with the expected network identifier values after the pure upchirps, a new frame is detected. Moreover, a packet is discarded if it contains different network identifiers than the ones expected, since it belongs to a different LoRa network.

- **Downchirps:** After the network identifiers, there are two and a quarter frequency synchronization symbols, which are downchirps, i.e., \mathbf{x}_0^* . In Section 6.2, we will explain how these two symbols can be used to partially distinguish between the CFO and the STO.

Header (Optional)

The packet continues with an optional header, which contains information about the length of the packet in bytes, the code rate, the presence of a cyclic redundancy check (CRC) for the payload, and a checksum of the header bits. If the header is not present (implicit header mode [41]), the receiver parameters need to be configured manually. The structure of the header was explained in [44] and [51], and is presented in detail in Section 6.1.

Payload and CRC

Finally, the last part of the packet is the PHY payload, which contains either data packets or MAC layer control packets. The maximum length of the payload is 255 bytes. An optional 16-bit CRC of the payload bits follows [41].

2.4 LoRaWAN Multiple Access

The MAC scheme of LoRaWAN is relatively simple and is based on an open-source protocol of the LoRa Alliance that is stacked on top of the LoRa PHY layer. Its first version has been described in detail in [52], and the updated current version is [12]. In this section we will briefly introduce the LoRaWAN network architecture, and present the basic characteristics of the LoRaWAN MAC layer, in order to understand how the LoRaWAN channel access scheme impacts the network throughput of LoRaWAN, and how the PHY layer characteristics play a role in this matter.

2.4.1 LoRaWAN Network Architecture

The LoRaWAN network architecture is depicted in Fig. 2.5. On one side there are the numerous LoRa end nodes that transmit data wirelessly to the LoRa gateways. LoRaWAN uses a star topology, i.e., the end nodes cannot transmit information directly to each other, but only through centralized gateways. The gateways are responsible for the reception of the LoRa packets and the demodulation of the data, but they do not have any further intelligence.

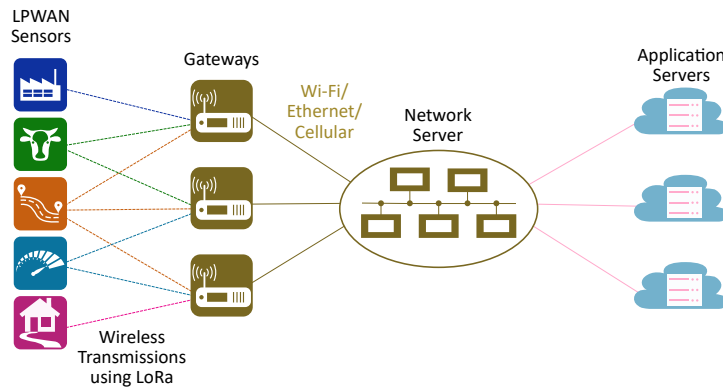


Figure 2.5 – The LoRaWAN Network Architecture. The diagram was created based on a figure owned by Semtech that can be found at <https://www.semtech.com/lora/what-is-lora>.

Instead, they forward the raw data to the network server, serving only as an interface between the wireless transmission and the network server. The network server on the other side is the real intelligence of the system: it collects receptions from multiple gateways, and therefore the same packet can arrive at the network server from multiple paths, taking advantage of the associated diversity. Moreover, the network server schedules the downlink packets to the end nodes, containing acknowledgments (ACKs) if needed (only for confirmed uplink traffic) as well as MAC commands. The application servers can be owned by third parties, e.g., The Things Network [53], and multiple applications can be connected to the same network server.

2.4.2 LoRaWAN Classes of Devices

LoRaWAN consists of three types of devices, namely Class A, Class B, and Class C. We will briefly explain the main characteristics of each class.

- **Class A:** The vast majority of end-node devices belong in this class. The features of Class A devices should also be implemented by Class B and Class C devices as they are the basic functionalities for an end node to be part of LoRaWAN [52, 12]. Class A end nodes are the ones with the lowest power consumption, since the transmissions can only be initiated by the end-nodes themselves, using the ALOHA protocol [54]. As a result, the end-node radio is turned off most of the time, except when it needs to transmit a packet. The uplink traffic can be either unconfirmed or confirmed. The LoRaWAN specifications advise to use confirmed uplink traffic only when necessary to avoid congestion, since a gateway cannot receive uplink packets from end nodes when is transmitting a packet in downlink. A downlink transmission from the gateway to the end node can only happen after an uplink transmission, during a preset reception window (Rx1) which starts $1s \pm 20\mu s$ after the end of the uplink transmission. The downlink message in the window Rx1 is transmitted using the same parameters (same frequency band, same bandwidth B and same SF) as the ones used in the uplink transmission. In case a downlink message is not received during the first transmission, the end node must turn

on a second reception window (Rx2) $2s \pm 20\mu s$ after the end of the uplink transmission, in a predefined frequency band and with a predefined bandwidth B and spreading factor SF. The predefined values for Rx2 can be modified using MAC commands, and the default values are region specific. For example, the default values for Rx2 in Europe are a carrier frequency at 869.525 MHz, a bandwidth $B = 125$ kHz, and a maximum spreading factor of $SF = 12$ [55]. The end node is obliged to open the receive windows even when it transmits an unconfirmed uplink message, since those receive windows are the only opportunity for the network server to send data to the end nodes.

- **Class B:** the devices in this class have all the functionalities of Class A devices with the additional feature that the end nodes periodically open reception windows (called *ping slots* [12]) for scheduled downlink messages. For this reason, the end nodes require a precise timer that is synchronized to the network server. To this end, periodic beacons are sent in the downlink by the gateways to synchronize all end nodes of the network. Class B devices can run on battery, but consume more power than Class A devices.
- **Class C:** the devices of this class have the highest power consumption, and need to be connected to a power source. Class C end nodes are supposed to continuously listen for downlink messages, with parameters set to the Rx2 window parameters of Class A devices, except when they are transmitting an uplink message or during the short Rx1 window that follows an uplink message [12].

2.4.3 Adaptive Data Rate Mechanism

The fundamental trade-off of LoRaWAN between the data rate and the communication range is handled by the adaptive data rate (ADR) mechanism. The ADR mechanism dynamically changes the spreading factor, the coding rate, and the transmission power of an end node to maximize both the battery life of the end node and to optimize the total network throughput. Although the ADR mechanism is part of the MAC layer, it requires an understanding of the underlying physical layer performance. For example, the ADR mechanism, based on the channel state information, i.e., the knowledge of the SNR, chooses the desired error rate using a physical layer model, such as the one presented in this thesis, and thus derives a suitable data rate and adjusts the transmission parameters for the given SNR.

The ADR mechanism runs in both the end node and the network server, and each end node can choose if it will allow the network server to govern the ADR mechanism or if the end node will govern it itself. The ADR algorithm on the end node is specified by the LoRa Alliance [12], whereas the algorithm on the network server can be defined by the network operator. The ADR algorithm on the network server can modify the SF and the transmission power, while the ADR algorithm on the end node can only increase the SF, after some failed attempts to deliver a packet.

When the network server is able to control the spreading factor and the transmission power of a device, it informs the end node by setting the downlink-ADR bit to one, during a downlink

transmission. If an end node chooses the SF and the transmit power to be controlled by the network server, it can inform the network server by setting the uplink-ADR bit to one, during an uplink transmission. The LoRaWAN specifications mention that the end nodes should choose the network server to control the ADR mechanism whenever this is possible [12]. If the network server is unable to control the ADR mechanism, e.g., due to rapid variations in the channel characteristics, it sets the downlink-ADR bit to zero. In such a case it is suggested that a mobile end node should also set the uplink-ADR bit to zero, and make use of the ADR mechanism that resides in the end node. A stationary end node is suggested to keep the uplink-ADR bit at one, waiting for the network server to control the ADR mechanism again.

2.5 State of the Art

As of March 2020 there were 744 publications in IEEE Xplore database which contain the term ‘LoRa’ or ‘LoRaWAN’ in their title. Since an exhaustive list of the publications is not particularly useful, in this section we present an overview of the high-impact research papers, first on the LoRaWAN applications to motivate our focus on LoRa, then the LoRa physical layer aspects, focusing mainly on interference, and on LoRaWAN scalability, which are directly related to the focus of this thesis.

2.5.1 LoRaWAN Applications

LoRa networks are useful in a wide range of applications where latency is not critical. LoRa has been used extensively in agriculture monitoring, among other LPWAN technologies [56, 57, 58]. In particular, it has been used for the monitoring of bee colonies [59], livestock monitoring [60], water monitoring [61], and soil monitoring systems [62]. It has also been used in wireless sensor networks [63, 64], smart city applications [65, 66, 67, 68], smart metering [69, 70, 71], and industrial monitoring [72, 73, 74, 75, 76]. Health monitoring is another field where LoRa networks have been used during the last years [77, 78]. An overview of LPWANs for health-monitoring applications can be found in [79].

Despite the fact that LoRa uses a low bandwidth of a few hundred kHz, it has recently been used in localization applications in outdoor environments. For example, [80, 81] use multiple algorithms for LoRa-based localization based on the received signal strength indicator (RSSI), and [82] uses a Gaussian-process-based RSSI method for indoor and outdoor localization. A large dataset used with a fingerprinting algorithm is presented in [83]. In [84], the authors use the time difference of arrival (TDoA) method while [85] uses the TDoA differences measured at different gateways along with an extended Kalman filter to estimate the location of an end node. In [86] TDoA is combined with angle of arrival (AoA) to reduce the number of gateways needed to obtain a location estimate. The work in [87] explores the use of LoRaWAN for mobile sensing and tracking applications, and [88] uses LoRaWAN for outdoor tracking. The authors in [89] discuss geolocation based on LoRa, using gateway timestamps. Based on the results from the aforementioned works, we note that the low-bandwidth nature of LoRa limits

the localization accuracy. Therefore, LoRa is not a great solution for high-accuracy location tracking. However, we note that more research is required on the interesting direction of geolocation and the use of LoRa along with Global Navigation Satellite System (GNSS) for asset location in a combination of device-centric and cloud-based geolocation approaches [90].

2.5.2 Reverse Engineering and SDR Implementations of the LoRa PHY

Numerous reverse engineering efforts and corresponding SDR implementations (e.g., [43, 44, 91, 92]) in the past few years have revealed many important details of the LoRa PHY. In particular, [43] was the first attempt to reverse engineer the PHY layer of LoRa. This seminal reverse engineering work gave a great stimulus to the LoRa community toward understanding the LoRa PHY structure, and also gave the first partial SDR implementation of the LoRa PHY in GNU Radio.

The work in [44] went a step further and revealed some details that were slightly overlooked in [43]. Moreover, the authors of [44] reverse-engineered parts of the LoRa packet that were not considered in [43], e.g., the header structure. However, the authors of [44] admit that their GNU Radio implementation works only for high SNR values, since they use a demodulator that is more robust to frequency offsets, but is much worse in terms of error-rate performance than the one described in Section 2.1. A more recent SDR implementation of LoRa PHY in GNU Radio is described in [92, 93], where GNU Radio simulation results for AWGN and Rayleigh fading are presented, for coherent and non-coherent demodulation and for hard and soft decoding of LoRa. The authors in [93] also mention that their SDR implementation does not work well with actual transmissions for low SNR values.

All of the existing SDR implementations lack important aspects. For example, sampling time offset (STO) and carrier frequency offset (CFO) estimation and correction. This can have a devastating effect on the demodulation of LoRa symbols. This is the main reason why the existing SDR implementations can only operate at very high SNRs. An SDR implementation of the LoRa PHY that works in the operational SNR regime of LoRa is still missing in the literature.

2.5.3 Interference Models for LoRa PHY

LoRa networks operate in the industrial, scientific and medical (ISM) band, together with many other wireless technologies. Moreover, due to the pure ALOHA MAC scheme of LoRaWAN, the number of LoRa packet collisions in a network increases rapidly with the number of connected end nodes. In this subsection, we will discuss the research work on both types of interference, which are directly related to the research of this thesis.

Cross-Technology Interference

LoRa uses the ISM band, which is free and shared with many other technologies, e.g., Sigfox [9] and IEEE 802.15.4g. We call this type of interference *cross-technology interference*. The impact of interference coming from different technologies on the performance of the LoRa modulation has received significant attention in the literature [94, 95, 96, 97, 98, 99, 100, 101].

Specifically, [94] studies the co-existence of LoRa with IEEE 802.15.4g using a controlled experimental setup with continuous interference between the two technologies in an anechoic chamber. The presented results show that LoRa is relatively resilient to interference from IEEE 802.15.4g, especially for higher spreading factors, even in the presence of high interfering power – up to 16 dB higher than the LoRa signal. For the lower spreading factors a high packet reception ratio (PRR) is shown to be possible up to 6 dB of interference power. On the contrary, IEEE 802.15.4g is shown to be less resilient to interference coming from LoRa nodes.

The work in [95] studies the co-existence of LoRa with ultra-narrowband technologies (UNB), such as Sigfox. Using a PHY layer model, as well as a network simulator for the MAC layer model, it is shown that LoRa is more susceptible to interference from UNB than the other way around. Furthermore, in networks where LoRa and Sigfox end nodes co-exist, LoRa end nodes that are farther away from the gateway suffer more from the UNB interference than end nodes that are closer to the gateway.

In [96], the interference between LoRa modulation and GFSK is examined using an experimental setup. The experiments showed that LoRa is relatively resilient to GFSK interference that has less than 6 dB higher power than LoRa.

The work in [97] evaluates the impact of interference from multiple sub-GHz ISM technologies like Sigfox, Z-wave [102], and IO Home Control [103] to LoRa, using a controlled measurement environment. The results show that Sigfox has the most detrimental effect on LoRa, and that all three technologies affect the LoRa packet reception mainly when they interfere during the preamble and the header part of the LoRa packet.

An experimental validation of the impact of ISM interference in the 863-870 MHz band on LoRa is performed in [98]. The measurements are performed in different environments (business park, shopping area, hospital, etc) in the city of Aalborg, Denmark. The results show that high probabilities of interference exist in the business and shopping areas, while LoRa does not face much cross-technology interference in the hospital, industrial, and residential areas of the city. Moreover, in [99] the authors combine the aforementioned experimental measurements with a system-level simulator to assess the impact of ISM interference on LoRaWAN coverage and capacity.

The work in [100] presents the IEEE 802.15 LPWAN interference model based on a Poisson point process. Moreover, the impact of the interference of minimum-shift keying (MSK) modulation on the BER of LoRa is evaluated with Monte-Carlo simulations in [101].

The aforementioned works show that cross-technology interference is relevant and important to be examined. However, due to the spread spectrum nature of LoRa, cross-technology interference essentially shows up – and can be modeled – as AWGN. Hence, the effect of cross-technology interference can be included in PHY error-rate models, such as the one described in this thesis, and analyzed as a reduced SNR. However, the exact reduction in SNR depending on the number of interfering users with different SFs is still an open question missing from the literature.

Same-Technology Interference

Apart from the cross-technology interference in the ISM band, a LoRa node can also experience interference from other LoRa nodes. We call this type of interference *same-technology interference*. Same-technology interference can be divided in two main categories: the first is interference from other LoRa nodes which use different spreading factors, which is called *inter-SF interference*. The second, and most severe type of same-technology interference comes from LoRa nodes transmitting with the same spreading factor, and is called *same-SF interference*. Same-technology interference is directly related to the research of this thesis. The impact of same-technology interference of both types has also received significant attention in the literature [104, 105, 106, 107, 108, 109, 110, 111, 96, 112, 1, 113].

The work of [104] examines the scalability of a LoRa network that is limited by same-SF interference. The authors use a network simulator and a simple interference model, where a packet affected by interference survives if both of the following independent assumptions hold: first, the packet has a value of received signal strength (RSS) that is above the sensitivity threshold for the chosen SF. Second, the received packet has an SIR > 6 dB. The sensitivity threshold for every SF has been measured experimentally, and the 6 dB threshold is chosen according to [42], and also after collision measurements performed by the authors. The fact that a LoRa packet can survive the impact of a lower-power interfering packet is mentioned in the literature as the *capture effect*, and has been examined in many works. Including the capture effect in the model leads to higher values of network throughput compared to the pure ALOHA model where any colliding packet is considered lost. The work in [105] uses the capture-effect model and the simulator of [104] to show the impact of directional antennas and multiple gateways in improving the packet reception rate in coexisting, interfering LoRa networks.

The work in [106] measures the impact of same-SF interference in colliding LoRa packets using a controllable setup with real transmissions. The PHY collision results are then used in a simulation model that determines the number of end nodes that can be served by one gateway. Similarly, the authors in [107] measure the impact of colliding packets with actual transmissions and then create a simple simulation model to examine the coexistence of up to 1000 LoRa end nodes. Finally, the authors in [108] measure colliding LoRa packets for multiple SIRs to experimentally evaluate the capture effect in a controllable environment. The results show that a packet under interference from another LoRa packet can survive even

with an SIR < 6 dB. The measurement results are then used in a simulation model to evaluate the capacity of LoRa networks, showing that it is higher than the capacity of a pure ALOHA network.

The work of [109] extended the simulator of [104] in order to study the impact of imperfect orthogonality between different LoRa spreading factors. The authors in [110] also examine the effect of imperfect orthogonality by examining the SIR threshold for receiving a packet correctly for all combinations of spreading factors. The SIR thresholds are derived both by simulations and by experimental results and rectify the values found in [111]. Interference is particularly detrimental when users with the same spreading factor collide since the spreading can no longer mitigate the interference. The work in [110] shows that interferers with different spreading factors have an average rejection SIR threshold of -16 dB while the SIR threshold for same-SF interference is 0 dB. A similar result is mentioned in [96], where the authors experimentally show that although the most severe degradation of the packet reception rate happens for same-SF interference, LoRa signals with different SFs can also interfere with each other. The authors of [112] perform an experimental assessment of the link-level characteristics of the LoRa system, followed by a system-level simulation to assess the capacity of a LoRaWAN network, where inter-SF interference is considered.

We conclude from the works of the previous paragraph that the assumption of perfect orthogonality between different SFs does not hold, and inter-SF interference should indeed be considered, especially in large networks. However, due to spreading gain, the impact of inter-SF interference is quite different than the impact of same-SF interference. For this reason, due to the approximate orthogonality of different spreading factors, it is widely accepted that inter-SF interference can be treated as white noise. Therefore, the inter-SF interference can be included in LoRa PHY error-rate models, such as the one we present in this thesis, by properly adjusting the SNR. Unfortunately, this simple model is not applicable for same-SF interference which requires a specific analysis that is missing in the literature. This is why we include same-SF interference in our error-rate model with a detailed analysis.

None of the aforementioned works treats interference and noise in a unified probabilistic PHY layer model, where decisions on the survival of a packet are taken with a probability that depends both on interference and noise. The first work that analyzes the coverage of LoRa under a simple unified model is [1], which provides an approximation for the BER of the LoRa modulation under AWGN and interference from a single same-SF LoRa interferer. Capacity planning for LoRa with the aforementioned interference model is addressed in [113]. We believe that the current lack of a detailed interference model in the physical layer leads to misconceptions about the actual network throughput of LoRaWAN.

2.5.4 LoRa PHY Link Performance Evaluation through Deployments

The coverage and the performance of LoRa links can also be evaluated using measurements obtained through transmissions between actual LoRa nodes. In [63], the coverage and the

reception ratio were measured for a direct LoRa link in the center of Glasgow, Scotland, where one of the transceivers was static and the other was mobile, as well as with a mobile node and multiple static gateways. Mainly outdoor, along with few indoor tests were performed, and it was shown that ACKs were received correctly for 42% of the transmissions in a range of 1.9 km. Similarly, in [66] a single link between a static gateway and a mobile node has been measured in terms of device sensitivity and packet capture rate at two districts of Bologna, Italy. Path-loss model curves were fitted to the experimental results for the two districts. Moreover, measurements between a node located either on the roof rack of a car or on a boat were reported for the city of Oulu, Finland [114], and path-loss models were fitted for the two experimental cases. An experimental coverage test is presented in [13] for the city of Padova, Italy. A gateway is installed in an elevated spot, allowing to cover a cell of 2 km, where the connections at the edges of the cell were only guaranteed with a SF = 12. The packet delivery ratio for different spreading factors and for a deployment consisting of a static gateway and an end-node at five different points in a suburb of Paris, France, were reported in [115]. Received packets in distances above 2.3 km could be achieved only with SF = 12.

In [116], an indoor deployment of a gateway and an end-node in different positions inside a building is presented. Results for the RSSI, the packet-loss ratio, and the packet error rate for different spreading factors are shown, along with results concerning the daily throughput per SF, power consumption and ACK-reception delays. In [78], an outdoor gateway is placed in the campus of the University of Oulu, Finland, and sensor nodes are placed indoors in different buildings all over the campus. The authors provide measurement results regarding the RSSI, the power consumption and the packet delivery ratio. It is shown that for SF = 12 and 14 dBm transmit power the average indoor packet delivery ratio for the entire campus is 96.7%. The work in [117] evaluates the indoor propagation performance of LoRa in a building on the campus of Glasgow Caledonian University. The experimental measurements are compared to four propagation models, namely the International Telecommunication Union (ITU) site generic model, the log-distance model, the multi-wall model, and the ray-tracing model. It is shown that the multi-wall model is the most appropriate for the examined scenario.

In the works of [118, 119], a point-to-point (P2P) topology for LoRa is examined and experimentally measured, where the terminals are placed at a low height (1.5m above the ground), as opposed to the star topology of LoRaWAN with elevated positioned gateways. The coverage and the path loss of P2P LoRa links are measured with a portable setup developed by the authors [120] in three different environments, namely a coastal environment, an urban environment, and a forest environment. Path-loss models are fitted based on the measured censored data, and the packet error rate of P2P LoRa links is determined for all three environments.

In [72, 73, 74], indoor deployments for measurements in industrial areas are presented. In particular, the results in [73] show a total coverage area of 34000m² for SF = 7 and a single gateway. Moreover, the work in [74] performs a multi-week industrial condition monitoring case study at ABB Switzerland for three different LPWAN technologies. The results show that

LoRa provides the largest communication range and packet delivery ratio in the examined indoor industrial setup among the three LPWAN technologies.

2.5.5 Probabilistic Models for the Performance of LoRa PHY

Despite the large number of experimental works and empirical models for the LoRa PHY, the number of works that provide rigorous probabilistic models for the performance of LoRa PHY is still very small, indicating that more research needs to be performed in this important topic. This fact is one of the main motivations for the work in this thesis. The first approximations for the BER of LoRa under AWGN using curve-fitting were given in [95, 121]. Moreover, approximations for the bit-error rate (BER) of the LoRa modulation when transmission takes place over additive white Gaussian noise (AWGN) and Rayleigh fading channels are given in [122], but interference is not considered. The work of [1], provides an approximation for the BER of the LoRa modulation under AWGN and same-SF interference. Closed-form expressions for the LoRa BER under AWGN, Nakagami, and Rice fading channels are given in [123]. The evaluation of LoRa BER using a Marcum function was proposed in [124]. Finally, the work in [125] shows and numerically evaluates the bit error rate expressions of coded LoRa in AWGN for both coherent and non-coherent detection. A basic contribution of this thesis is that it enhances the probabilistic modeling of the LoRa PHY performance, which is clearly under-considered in the literature.

2.5.6 LoRaWAN Scalability

In Section 2.5.4, we presented deployments of LoRa links between an end-node and a gateway, or in peer-to-peer links, which can provide important experimental measurements regarding the RSSI, the coverage for different spreading factors, the packet delivery ratio of a single device placed in different positions, etc. These deployments are very useful to experimentally evaluate the transmission characteristics and the performance of a single LoRa link in different environments, and to compare with the sensitivity values reported by the designers, and the theoretical error rates. However, LoRaWAN scalability is difficult to be assessed directly through deployments, since the assessment of LoRaWAN scalability with large deployments is difficult, costly, and the results cannot easily be generalized to any network configuration. For this reason, either mathematical LoRaWAN models have been developed or system-level simulations have been carried out to assess the scalability of LoRa networks. Although the research on LoRaWAN scalability is not performed on the physical layer, both of the two aforementioned approaches rely on some type of an error-performance PHY model.

Mathematical Models for LoRaWAN

The work in [126] presents the results produced by a mathematical model for different values of network load, and as the number of nodes goes to infinity. The model is pure-ALOHA,

i.e., all colliding packets are considered lost. The authors extend their mathematical network model in [127] to include ACKs, retransmissions, and the capture effect. Finally, the authors present an extended mathematical model for confirmed traffic in LoRa networks with one gateway in [128]. The capture-effect model from [127] is used, but the model is extended with transmission failures due to random (interference) noise in the ISM band that is not related to LoRa transmissions. The model presents an allocation scheme for SF and coding rate, so that the packet loss ratio is above the value required by the quality-of-service (QoS) needs for a given application. The total delay from the transmission of a packet to a correctly received ACK is also modeled.

A stochastic geometry framework for modeling the uplink performance of a single-gateway LoRa network under Rayleigh fading is used in [129]. For a packet to survive, the received power has to be above the SNR threshold for the chosen SF. The collision model for same-SF packets is based on the idea of the 6 dB SIR threshold (capture effect) discussed in Section 2.5.3. The authors show that same-SF interference is the main limiting factor for the scalability of LoRaWAN. The model described in [130] considers the joint effect of same-SF and inter-SF interference. In particular, the same-SF interference is both modeled considering the dominant same-SF interferer, and also considering the cumulative same-SF interference. It is shown that same-SF interference is the main limiting factor, but inter-SF interference leads also to a further 15% coverage loss. The work in [131] extends the model of [129] to include message replication for time diversity and multiple receive antennas for space diversity. Moreover, the work in [132] uses a stochastic geometry model to jointly analyze the interference in time and frequency domains, to account for overlapping packets in a two-dimensional plane.

Closed-form expressions for the collision probability and the packet loss probability were derived in [133] using the ALOHA model for same-SF collisions, and the expressions are compared to a Poisson-process model and simulations. The authors in [134] propose a Markov-chain model for the on-air activation of LoRaWAN nodes, where the delay and the energy required to join the network are examined. An investigation of the latency, collision rate, and throughput for LoRaWAN under duty-cycle restrictions is performed in [135]. In [136], the authors present a single-gateway model which accounts for the capture effect as in [127], and also accounts for the limited number of concurrently received packets at the gateway, the duty-cycle restrictions in both the uplink and the downlink, and the number of ACKs sent as a response to a confirmed uplink packet. The number of retransmissions is not accounted for in the model. Packet retransmissions and the coexistence of confirmed and unconfirmed traffic is added in the model of [137]. Inter-SF interference is not included in any of the five aforementioned works. In [138], the authors attempt to minimize the collision probability by a power- and SF-allocation scheme that introduces packet error rate fairness among nodes, to overcome the near-far effect, i.e., the situation where nodes that are further away from the gateway cannot deliver their messages successfully when the nodes closer to the gateway are transmitting.

Most of the mathematical modeling of the LoRaWAN MAC layer in the aforementioned works is performed using simplistic collision models on the physical layer. These collision models typically use either pure ALOHA or consider the capture effect with the heuristic 6 dB SIR threshold. With our research, we contribute to enable the extension of these models to include a more accurate physical layer performance description.

System-level Simulations

System-level simulators allow the evaluation of the overall network performance. In such simulators it is possible to tune many important parameters of the network, such as the number of nodes per SF, the total number of nodes in the network, the amount of transmitted data, how often each node transmits, the transmission power, etc. Such system-level simulation environments are therefore versatile and can be relatively easily tuned to provide performance results for different network configurations, including large and heavily-loaded networks. The decisions on the survival of the simulated LoRa packets are made using some type of an underlying PHY error-performance model. Such a detailed PHY model that is necessary as an input to system-level simulators is going to be presented as part of this thesis. Many system-level LoRaWAN simulators have been proposed in the literature during the last years.

One of the first network simulators written in Python was presented in [104]. The discrete-event simulator is called LoRaSim and is publicly available in [139]. The PHY model used in [104] treats interference and noise independently. A sensitivity threshold for each SF is derived from measurements coming from actual LoRa transmissions. Moreover, the capture effect is included with a simple model which decides whether colliding packets survive, depending on their relative time offset and the relative strength of the received power. The network simulator in [104] measures the data extraction rate and the network energy consumption in either single- or multi-gateway configurations. The gateways and the end-nodes are placed randomly in two-dimensional space, with predefined limits. It is shown that multiple gateways help the scalability of LoRa networks. However, the simulator considers different SFs as being entirely orthogonal, and does not simulate downlink ACK messages or any other downlink traffic. LoRaSim – or modified versions of it – have been extensively used in the literature for the evaluation of different LoRa networks [105, 140, 109, 141, 142, 143, 144, 145].

MAC-layer support was included in LoRaSim in the work of [140], but the source code is not made publicly available by the authors. The extended simulator is named LoRaWANSim, and allows for downlink ACK messages, downlink packets, retransmissions, and rate-adaptation functionalities, considering also the duty-cycle restrictions of the LoRaWAN MAC layer. The simulation results show that the network throughput is negatively affected not only by a mass-scaling of the network, but also from downlink traffic (both ACK and data), as well as from packet retransmissions which occur when a packet in a confirmed uplink is not acknowledged.

The work in [109] extends the LoRaSim simulator to include the impact of inter-SF collisions in the overall network throughput, and introduces a log-normal channel-fading model in the

simulator. The PHY model for the survival of colliding packets is based on deriving the SIR thresholds for all combinations of inter-SF collisions in MATLAB, but then setting the SIR threshold to a simple conservative value of 6 dB for the network simulations. The results show that the imperfect orthogonality of different SFs affects the total throughput of a LoRa network with end nodes transmitting using all SFs. Moreover, the impact of fading is shown to be severe only in single-gateway deployments, since multi-gateway networks profit from diversity.

LoRaSim was also extended in [146] to include bidirectional traffic. The extended simulator, called LoRaFREE [147] supports confirmed and unconfirmed uplink traffic, downlink traffic, retransmissions, as well as the modeling of inter-SF interference using the inter-SF SIR table of [110]. The network simulator is used to compare the typical LoRaWAN protocol with a proposed coordinated MAC protocol, where the end-devices collect data in a buffer for longer periods and transmit them in bursts.

Another open-source system-level LoRaWAN simulator written in Python is presented in [148]. The simulator is named LoRaEnergySim and can be found at [149]. The cross-layer simulator focuses on the energy efficiency of LoRa networks. The capture effect for the case of same-SF colliding packets is accounted for according to the collision model of [104]. No collisions are assumed for packets with different SFs. The detailed energy profile of an end node used in the simulator is based on the energy consumption reported in [150]. The simulator allows for both confirmed and unconfirmed uplink traffic and for the inclusion of the adaptive data rate scheme. The simulated metrics are the data extraction rate and the energy per payload byte.

LoRaWAN system-level simulators have also been implemented in ns-3, which is an open-source discrete-event network simulator, used for simulating Wi-Fi, LTE, IEEE 802.15.4, and other systems. Inter-technology interference can generally be simulated in ns-3. The work in [151] presents such a LoRaWAN simulator, where the PHY layer error-rate model is based on MATLAB simulations, and any interference (including same-SF interference) is treated equally as noise which degrades the SINR value of a received packet. Configurations with one, two or four gateways are possible, as well as both confirmed and unconfirmed uplink traffic. Downlink ACKs and other downlink traffic can also be included in the simulator. However, MAC commands and the ADR mechanism are not included in the simulator of [151]. The simulator measures the packet delivery ratio but not the energy consumption or the energy efficiency of the system.

Another ns-3 LoRaWAN simulator is presented in [65], where multi-gateway configurations are allowed, as well as MAC commands. Inter-SF interference is accounted for using the SINR threshold matrix of [111]. Confirmed uplink traffic as well as downlink traffic is not supported, but these features, among others, were introduced in the extended versions of [152, 153], and can be found at [154]. Furthermore, an ns-3 LoRaWAN simulator is also presented in [155], where MAC commands are included, and multi-gateway configurations can be simulated, along with downlink traffic and downlink ACKs. Interference from different configurations of SF and BW that result in similar chirp characteristics is also included. The differences between

the aforementioned ns-3 simulators are discussed in detail in [155, 55, 153].

An ns-3 module to examine the impact of using carrier sense multiple access (CSMA) was introduced in [156]. Furthermore, an ns-3 simulator with a component to evaluate the use of persistent-carrier sense multiple access (p-CSMA) in LoRaWAN was build in [157].

FLoRa is an open-source system-level LoRaWAN simulator in OMNeT++ that was presented in [158], and can be found at [159]. This implementation focuses on the evaluation of the dynamic management of the network parameters, such as modifications of the ADR mechanism. The collision model used is the one from [104]. The simulator supports bidirectional communication, the ADR mechanisms both on the node and on the network server, and the evaluated metrics are the delivery ratio and the energy consumption per successful transmission. A modified ADR mechanism on the network server is shown to be beneficial in non-static channel conditions.

Finally, a system-level LoRaWAN simulator written in Java is presented in [160]. The authors use the collision model of [106], and the simulator can be found at [161]. A simple event-driven simulator in Matlab to emulate basic functionalities of the MAC protocol is shown in [162], where the authors focus on the negative impact of downlink traffic. A short review of many of the aforementioned simulators can be found in [163].

We note that the PHY collision models used in current system-level simulators in the literature are in some cases overly simplistic. We consider that replacing the simplistic ALOHA or capture effect collision models with probabilistic PHY models for LoRa collisions, as the ones we propose in this thesis, will result in more accurate network throughput evaluation results.

2.6 Discussion and Summary

LoRa is one of the main LPWAN technologies in many different application fields and it is one of the most promising candidates for the massive IoT deployments of the future. In this chapter we presented a brief overview of LoRa and LoRaWAN. In Section 2.1, a digital baseband model for the LoRa PHY has been established that can serve as a solid basis for the analysis that will be carried out in the following chapters. Moreover, in Section 2.2 and Section 2.3, we described the complete LoRa transceiver chain and the LoRa packet structure respectively. In Section 2.4, we introduced a brief description of the basic channel-access protocol which is important to understand where the LoRaWAN limitations originate from. Finally, in Section 2.5, we presented research works related to the LoRa PHY performance, LoRaWAN scalability, and various LoRa SDR implementations.

Existing works have shown that scalability limitations are inherently present in LoRaWAN due to the low-power ALOHA-based channel access scheme that leads to frequent and unavoidable packet collisions. In order to design future LoRa networks, tools to accurately evaluate LoRaWAN scalability first need to be developed, in order to propose methods to improve this

scalability. A big effort has been made toward this research direction during the last years. Error-performance models of the LoRa PHY lie at the heart of the MAC mechanisms that can dynamically adjust (and allocate to the users) the available spreading factors, coding rates and transmission power levels, to optimize the network throughput. Moreover, error-performance models of the LoRa PHY lie at the heart of the tools that are used to evaluate the scalability of LoRaWAN. However, the vast majority of the LoRaWAN system-level performance evaluation methods and tools rely on over-simplified PHY error models that treat noise and interference independently, and are based in heuristic or empirical collision-survival thresholds. Furthermore, we saw that, regardless of the important reverse-engineering efforts for the LoRa PHY that have been carried out by the community, an SDR LoRa implementation that can work in the low-SNR regime is still lacking.

Therefore, in the next chapters, we gradually introduce a detailed model for the LoRa PHY error rate in AWGN both with and without LoRa interference. Such a probabilistic error-rate model that treats both noise and interference can help toward the goal of obtaining a complete PHY model for the accurate evaluation of LoRaWAN scalability in multi-layer analytical models or system-level simulators. Such a PHY model will also help toward improvements of the MAC layer mechanisms that adjust and allocate the transmission parameters to the end nodes. Finally, we will present our efforts toward the implementation of a fully-functional SDR LoRa prototype that also operates correctly in the low-SNR regime. Our SDR implementation can serve as a solid basis toward building testbeds for the experimental evaluation of new algorithms for the LoRa PHY and MAC layers.

3 LoRa Error Rate under Same-Technology Interference

In the previous chapter we have discussed the fundamentals of the LoRa modulation and demodulation, as well as the prior art on the analysis of LoRa networks. We have seen that the main limiting factor for the scalability of massive LoRa networks is the same-technology interference. However, the majority of the works that evaluate the throughput of large LoRa networks through system-level simulations, use simplistic physical layer models to infer frame error rates from the SNR and to decide on the survival of colliding packets. In most of the cases interference and noise are treated independently, and the threshold for the capture effect is heuristically chosen at $SIR = 6$ dB. As a result of such simplistic collision models, the system-level simulations may not estimate the actual throughput of different LoRa networks.

To be able to estimate correctly the throughput of LoRa networks, well-motivated probabilistic models for the survival of both colliding and non-colliding packets are necessary for the physical layer. In this chapter we present the first probabilistic error-rate model with both interference and noise, for the case of uncoded LoRa packets. Such a detailed error-rate analysis provides a better understanding of the nature and the impact of LoRa interference. Furthermore, it has high practical relevance, since it results in low-complexity approximations that can directly be used as an input for higher-layer simulators to provide accurate results on the scalability of LoRaWAN.

In particular, in this chapter, we first present our probabilistic interference model in Section 3.2, and we derive the expression for the symbol error rate (SER) under this new model. Since evaluating the derived expression may result in numerical stability issues, and since the complexity is very high, we also derive an approximation for the SER under the presented interference model, which is very accurate in the useful operating regime of LoRa. Furthermore, we prove two properties of LoRa-induced interference that enable a significant reduction of the complexity of calculating both the exact and the approximated SER. Finally, in Section 3.3, we derive an approximation for the frame error rate (FER), which is generally of greater practical interest for system-level simulators.

3.1 Symbol Error Rate Under AWGN

In this section, we derive the expression for the LoRa SER under additive white Gaussian noise (AWGN), which is useful for later explaining how the SER and the FER can be calculated in the presence of both AWGN *and* interference.

3.1.1 Distribution of the Demodulation Decision Metric

As explained in Section 2.1.2, the non-normalized DFT is applied to the dechirped signal, and non-coherent demodulation can be performed by selecting the frequency bin index with the maximum magnitude, as in (2.11). In the absence of noise, and with perfect synchronization, the DFT of the dechirped signal \mathbf{Y} has a single frequency bin that contains all the signal energy (i.e., a bin with magnitude N) and all remaining $N - 1$ bins have zero energy. On the other hand, when AWGN is present, all frequency bins will contain some energy. The distribution of the frequency bin values Y_k for $k \in \mathcal{S}$ is

$$Y_k \sim \begin{cases} \mathcal{CN}(0, 2\sigma^2), & k \in \mathcal{S}/s, \\ \mathcal{CN}(N(\cos\phi + j\sin\phi), 2\sigma^2), & k = s, \end{cases} \quad (3.1)$$

where s is the transmitted symbol.

Let us define $Y'_k = \frac{Y_k}{\sigma}$ for $k \in \mathcal{S}$. The normalized values Y'_k can be used in (2.11) instead of Y_k without changing the result and their distribution is

$$Y'_k \sim \begin{cases} \mathcal{CN}(0, 2), & k \in \mathcal{S}/s, \\ \mathcal{CN}\left(\frac{N\cos\phi}{\sigma} + j\frac{N\sin\phi}{\sigma}, 2\right), & k = s. \end{cases} \quad (3.2)$$

Thus, using basic properties of the complex normal distribution, we can show that the demodulation metric $|Y'_k|$ follows a Rayleigh distribution for $k \in \mathcal{S}/s$ and a Rice distribution for $k = s$, i.e.,

$$|Y'_k| \sim \begin{cases} f_{\text{Ra}}(y; 1), & k \in \mathcal{S}/s, \\ f_{\text{Ri}}\left(y; \frac{N}{\sigma}, 1\right), & k = s. \end{cases} \quad (3.3)$$

3.1.2 Symbol Error Rate Expression

A symbol error occurs if and only if any of the $|Y'_k|$ values for $k \in \mathcal{S}/s$ exceeds the value of $|Y'_s|$, or, equivalently, if and only if $|Y'_{\max}| > |Y'_s|$, where $|Y'_{\max}| = \max_{k \in \mathcal{S}/s} |Y'_k|$. Using order statistics [164] and the fact that all $|Y'_k|$ for $k \in \mathcal{S}/s$ are i.i.d., the PDF of $|Y'_{\max}|$ can be obtained as

$$f_{|Y'_{\max}|}(y) = (N - 1) f_{\text{Ra}}(y; 1) F_{\text{Ra}}(y; 1)^{(N-2)} \quad (3.4)$$

Using $f_{|Y'_{\max}|}(y)$, the conditional SER when symbol s is transmitted can be calculated as

$$P(\hat{s} \neq s|s) = \int_{y=0}^{+\infty} \int_{x=0}^y f_{\text{Ri}}(x; \nu, 1) f_{|Y'_{\max}|}(y) dx dy \quad (3.5)$$

$$= \int_{y=0}^{+\infty} F_{\text{Ri}}(y; \nu, 1) f_{|Y'_{\max}|}(y) dy, \quad (3.6)$$

with $\nu = \frac{N}{\sigma}$. The SER for all symbols s is identical and therefore (3.6) is in fact equal to the average SER and, if we assume that all symbols are equiprobable, it is also equal to the expected SER.

3.1.3 Symbol Error Rate Approximations

While the evaluation of (3.6) is in principle straightforward, in practice the values of N in the LoRa modulation are very large so that numerical problems arise. For this reason, two approximations that can be used to efficiently evaluate (3.6) were derived in [122]. Specifically, [122] used a Gaussian approximation so that $|Y'_s| \sim \mathcal{N}(\frac{N}{\sigma}, 1)$ and $|Y'_{\max}| \sim \mathcal{N}(\mu_\beta, \sigma_\beta^2)$ and where appropriate expressions are given to calculate μ_β and σ_β^2 . The work in [46] showed that [122] uses a slightly different description for the LoRa symbol than (2.3), but this difference does not affect the derived error rate approximation. By using our definition of the SNR in (2.5), the SER from [122] can be approximated as

$$P(\hat{s} \neq s) \approx Q \left(\frac{\sqrt{\text{SNR}} - \left((H_{N-1})^2 - \frac{\pi^2}{12} \right)^{1/4}}{\sqrt{H_{N-1} - \sqrt{(H_{N-1})^2 - \frac{\pi^2}{12} + 0.5}}} \right), \quad (3.7)$$

where $H_n = \sum_{k=1}^n \frac{1}{k}$ denotes the n th harmonic number and $Q(\cdot)$ denotes the Q-function. Using several additional approximations [122], the following more concise version of (3.7) is obtained

$$P(\hat{s} \neq s) \approx Q \left(\sqrt{2\text{SNR}} - \sqrt{2(\log(2)\text{SF} + \gamma_{\text{EM}})} \right), \quad (3.8)$$

where $\gamma_{\text{EM}} \approx 0.57722$ is the Euler–Mascheroni constant. We note that it is also possible to directly arrive at (3.8) using the methodology of [122] by skipping the intermediate result (3.7) and the required additional approximations. It is sufficient to observe that the distribution of the random variable $\hat{\gamma}$ defined in [122, Section II-B] converges to a Gumbel distribution with $\mu_{\hat{\gamma}} = 2\sigma^2(\log(2)\text{SF} + \gamma_{\text{EM}})$ and $\sigma_{\hat{\gamma}}^2 = 4\sigma^2 \frac{\pi^2}{6}$ for large N due to the Fisher–Tippett–Gnedenko extremal value theorem [164].

The SER of the LoRa modulation under AWGN for all supported spreading factors $\text{SF} \in \{7, \dots, 12\}$ is provided in Fig. 3.1. In this figure, we show results obtained from both Monte Carlo simulations and the approximation given in (3.7).

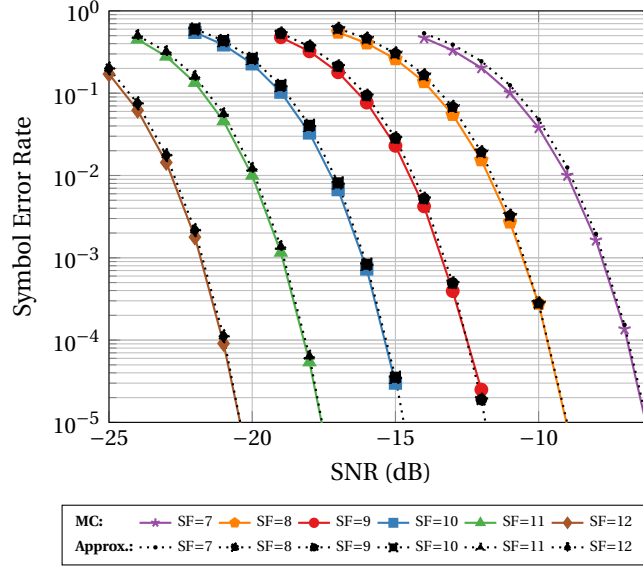


Figure 3.1 – Symbol error rate of the LoRa modulation under AWGN for all supported spreading factors $SF \in \{7, \dots, 12\}$. Results for Monte Carlo simulations and the approximation in (3.7) are shown.

3.2 Symbol Error Rate Under Same-SF LoRa Interference

In this section, we analyze the case of a gateway trying to decode the message of a user in the presence of an interfering LoRa device, as depicted in Fig. 3.2. As mentioned in Chapter 2, this scenario becomes particularly relevant in future LoRa deployments with a high density of end nodes due to the uncoordinated ALOHA-based random channel access of LoRaWAN. We assume that the LoRa gateway is perfectly synchronized to the user whose message is decoded. Various synchronization techniques for LoRa have been proposed in the literature [44, 49, 50], and will be explained in Section 6.2. It has been shown in [110] that interferers with different spreading factors have an average rejection SIR threshold of -16 dB while the SIR threshold for same-SF interference is 0 dB. As such, even though the inter-SF interference has a non-negligible effect on the error rate, it is the same-SF interference that has a dominant impact and needs to be modeled first, since it cannot be simplified to a simple SNR degradation. Therefore, in our work we limit our model to interference signals with the same spreading factor as the one employed by the user of interest. Finally, for simplicity, in our work we only consider one interfering user. If multiple interfering users collide at the same time, the strongest interfering user typically dominates the error rate. Therefore, the signal model is

$$y[n] = hx[n] + h_I x_I[n] + z[n], \quad n \in \mathcal{S}, \quad (3.9)$$

where h is the channel gain between the user of interest and the LoRa gateway, $x[n]$ is the signal of interest, h_I is the channel gain between the interferer and the LoRa gateway, $x_I[n]$ is the interfering signal, and $z[n] \sim \mathcal{N}(0, \sigma^2)$ is additive white Gaussian noise. Since we assume

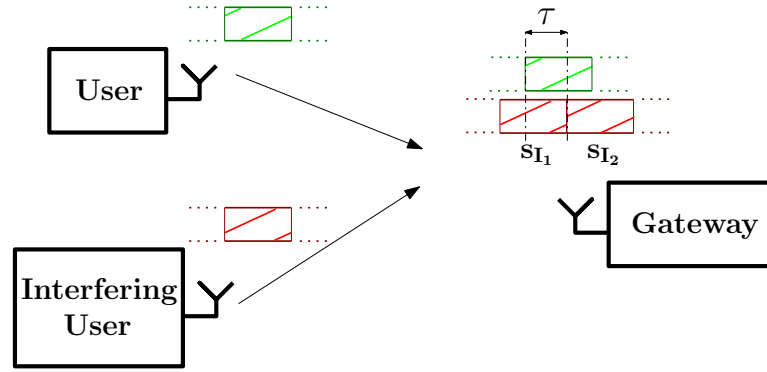


Figure 3.2 – Illustration of LoRa uplink transmission with one interfering user having an arbitrary offset τ .

that $|h| = 1$, the signal-to-interference ratio (SIR) can be defined as

$$\text{SIR} = \frac{1}{|h_I|^2} = \frac{1}{P_I}, \quad (3.10)$$

where we use P_I to denote the power of the interfering user. Since LoRa uses the non-slotted ALOHA protocol for medium access control, the interfering signal $y_I[n] = h_I x_I[n]$ is not synchronized in any way to the user of interest or the gateway. Due to the lack of synchronization, each LoRa symbol of the user of interest is generally affected by a combination of parts of two distinct interfering LoRa symbols, which we denote by s_{I_1} and s_{I_2} , as shown in Fig. 3.2.

Let τ denote the relative time-offset between the first chip of the symbol of interest s and the first chip of the interfering symbol s_{I_2} (i.e., the first chip of the interfering symbol s_{I_2} starts τ chip durations *after* the first chip of s). Due to the complete lack of synchronization between the users, we assume that τ is uniformly distributed in $[0, N)$. We note that in [1], the offset τ is constrained to integer chip durations, which is not particularly realistic since it effectively assumes that the interferer is chip-aligned with the user. Let $\mathcal{N}_{L_1} = \{0, \dots, \lceil \tau \rceil - 1\}$ and $\mathcal{N}_{L_2} = \{\lceil \tau \rceil, \dots, N - 1\}$. The discrete-time baseband equivalent equation of $x_I[n]$ can be found using (2.3) for s_{I_1} and s_{I_2} , appropriately adjusted to include the offset τ ¹

$$x_I[n] = \begin{cases} e^{j2\pi \left(\frac{(n+N-\tau)^2}{2N} + (n+N-\tau) \left(\frac{s_{I_1}}{N} - \frac{1}{2} \right) \right)}, & n \in \mathcal{N}_{L_1}, \\ e^{j2\pi \left(\frac{(n-\tau)^2}{2N} + (n-\tau) \left(\frac{s_{I_2}}{N} - \frac{1}{2} \right) \right)}, & n \in \mathcal{N}_{L_2}. \end{cases} \quad (3.11)$$

¹We note that the exact expression for $x_I[n]$ requires the use of (2.2) for representing s_{I_1} and s_{I_2} , resulting in up to four terms in (3.11). Therefore, (3.11) is a simplified version of the expression for $x_I[n]$ in the presence of non-integer time offsets τ . It corresponds to the exact expression for $x_I[n]$ only for integer values of τ . We use this simplification for tractability, and for being able to show the essence of the LoRa interference analysis, without delving into overly complex models.

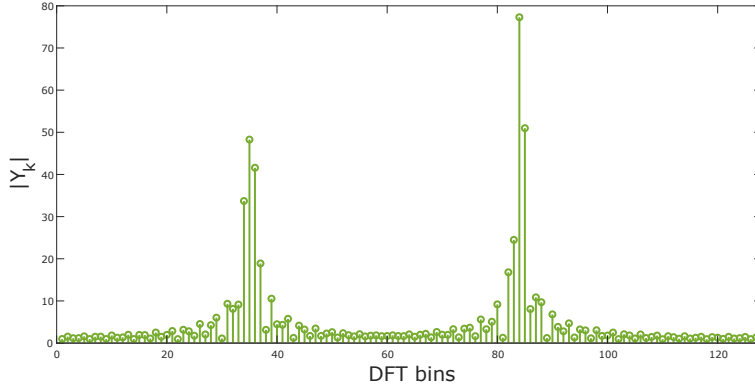


Figure 3.3 – The magnitude of an example transmitted interference pattern for $SF = 7$, $s_{I_1} = 83$, $s_{I_2} = 4$, and $\tau = 48.8$.

The demodulation of $y[n]$ at the receiver yields

$$\mathbf{Y} = \text{DFT}(\mathbf{y} \odot \mathbf{x}_{\text{ref}}^*) \quad (3.12)$$

$$= \text{DFT}(h_I \mathbf{x}_I \odot \mathbf{x}_{\text{ref}}^*) + \text{DFT}(\mathbf{z} \odot \mathbf{x}_{\text{ref}}^*). \quad (3.13)$$

We call $\text{DFT}(\mathbf{x}_I \odot \mathbf{x}_{\text{ref}}^*)$ and $\text{DFT}(h_I \mathbf{x}_I \odot \mathbf{x}_{\text{ref}}^*) = \text{DFT}(\mathbf{y}_I \odot \mathbf{x}_{\text{ref}}^*)$ the transmitted and received *interference patterns*, respectively. The received interference pattern depends on the time-domain interference signal \mathbf{y}_I , which is in turn a function of the interfering symbols s_{I_1} , s_{I_2} , the channel h_I , and the interferer time-offset τ . Fig. 3.3 presents an example of a transmitted interference pattern. In particular, in Fig. 3.3, we show the magnitude of the DFT bins for a transmitted interference pattern for the particular choice of $SF = 7$, $s_{I_1} = 83$, $s_{I_2} = 4$, and $\tau = 48.8$. We observe the two main clusters of high-power bins, which stem from the fact that the interference pattern consists of two symbols. We note that the interference pattern is the result of a continuous envelope that is shifted by a non-integer offset and ‘sampled’ at the DFT bins. As we see from (3.13), the magnitude of the DFT bins of the received interference pattern will have the same shape as Fig. 3.3, but scaled by $|h_I|$.

3.2.1 Distribution of the Demodulation Decision Metric

Let R_k denote the value of the transmitted interference pattern at frequency bin k , i.e.,

$$R_k = \text{DFT}(\mathbf{x}_I \odot \mathbf{x}_{\text{ref}}^*)[k], \quad k \in \mathcal{S}. \quad (3.14)$$

For a specific combination of a symbol s and an interference pattern \mathbf{y}_I , adding the interference to the signal of interest corresponds to changing the mean value of the distribution of Y'_k

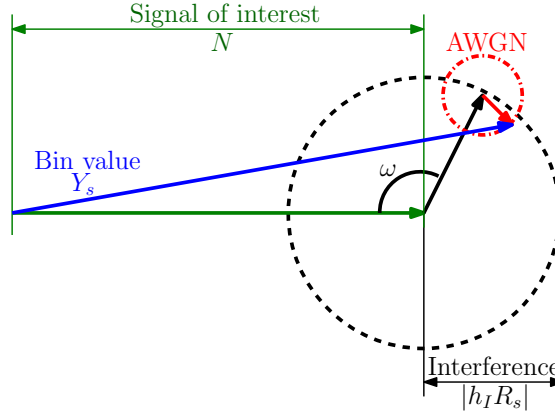


Figure 3.4 – Vector representation of the signal of interest (green), the interference $h_I R_s$ (black), the noise (red), and the bin value Y_s at the bin of interest s .

in (3.2), as follows

$$Y'_k \sim \begin{cases} \mathcal{CN}\left(\frac{|h_I R_k| \cos \theta}{\sigma} + j \frac{|h_I R_k| \sin \theta}{\sigma}, 1\right), & k \in \mathcal{S} / s, \\ \mathcal{CN}\left(\frac{N \cos \phi + |h_I R_k| \cos \theta}{\sigma} + j \frac{N \sin \phi + |h_I R_k| \sin \theta}{\sigma}, 1\right), & k = s. \end{cases} \quad (3.15)$$

where $\theta = \angle h_I$ is the phase shift introduced by the interference channel which is fixed for each packet transmission, but generally uniformly distributed in $[0, 2\pi)$. Thus, in the presence of interference, the demodulation metric $|Y'_k|$ used in (2.11) is distributed according to

$$|Y'_k| \sim \begin{cases} f_{\text{Ri}}\left(y; \frac{|h_I R_k|}{\sigma}, 1\right), & k \in \mathcal{S} / s, \\ f_{\text{Ri}}\left(y; \frac{\sqrt{N^2 + |h_I R_k|^2 + 2N|h_I R_k| \cos(\omega)}}{\sigma}, 1\right), & k = s, \end{cases} \quad (3.16)$$

where we define the phase shift between the user and the interfering user as $\omega = \phi - \theta$ for simplicity. The joint effect of the AWGN and the interference on the signal of interest is illustrated in Fig. 3.4.

3.2.2 Symbol Error Rate Expression

Similarly to (3.6), in the presence of interference, the SER for a given symbol s , conditioned on \mathbf{y}_I and ω , can be written as

$$P(\hat{s} \neq s | s, \mathbf{y}_I, \omega) = 1 - \int_{y=0}^{+\infty} f_{\text{Ri}}(y; v_s, 1) F_{|Y'_{\max}|}(y) dy, \quad (3.17)$$

where $v_s = \frac{1}{\sigma} \sqrt{N^2 + |h_I R_s|^2 + 2N|h_I R_s| \cos(\omega)}$ is the location parameter for the bin $k = s$. The CDF of the N th order statistic (i.e., the CDF of the maximum) is known to be $F_n(x) = P(X_1 < x)P(X_2 < x) \dots P(X_n < x)$. Due to the conditioning on \mathbf{y}_I and ω , each $|Y'_m|$, for $m \in \{1, \dots, N\} / s$, is independent from any other $|Y'_n|$, for $n \in \{1, \dots, N\} / \{s, m\}$. Thus, we can directly deduce that

the CDF of the maximum interfering bin is

$$F_{|Y'_{\max}|}(y) = \prod_{\substack{k=1 \\ k \neq s}}^N F_{\text{Ri}}(y; \nu_k, 1), \quad (3.18)$$

where $\nu_k = |h_I R_k|/\sigma$. By taking the expectation of $P(\hat{s} \neq s | s, \mathbf{y}_I, \omega)$ with respect to ω , we get the SER conditioned on s, \mathbf{y}_I

$$P(\hat{s} \neq s | s, \mathbf{y}_I) = \frac{1}{2\pi} \int_{\omega=0}^{2\pi} P(\hat{s} \neq s | s, \mathbf{y}_I, \omega) d\omega. \quad (3.19)$$

Recall that, by assumption, s_{I_1} and s_{I_2} are uniformly distributed in \mathcal{S} and τ is uniformly distributed in $[0, N)$. As such, the conditional SER (across multiple packets with different time offsets) $P(\hat{s} \neq s | s)$ can be computed as

$$P(\hat{s} \neq s | s) = \frac{1}{N^3} \sum_{s_{I_1}=0}^{N-1} \sum_{s_{I_2}=0}^{N-1} \int_0^N P(\hat{s} \neq s | s, \mathbf{y}_I) d\tau. \quad (3.20)$$

Finally, s is also uniformly distributed in \mathcal{S} , so that the unconditional SER becomes

$$P(\hat{s} \neq s) = \frac{1}{N} \sum_{s=0}^{N-1} P(\hat{s} \neq s | s). \quad (3.21)$$

The full expression for $P(\hat{s} \neq s)$ is then written as

$$P(\hat{s} \neq s) = 1 - \frac{1}{2\pi N^4} \sum_{s=0}^{N-1} \sum_{s_{I_1}=0}^{N-1} \sum_{s_{I_2}=0}^{N-1} \int_{\tau=0}^N \int_{\omega=0}^{2\pi} \int_{y=0}^{+\infty} f_{\text{Ri}}(y; \nu_s, 1) \prod_{\substack{k=1 \\ k \neq s}}^N F_{\text{Ri}}(y; \nu_k, 1) dy d\omega d\tau. \quad (3.22)$$

3.2.3 Complexity Reduction Using Equivalent Interference Patterns

Apart from the numerical problems that arise from the product of $(N-1)$ CDFs in (3.22), an additional practical issue is that the computational complexity of evaluating (3.22) is very high. In particular, the complexity of computing the three sums scales as N^3 and three integrals need to be numerically evaluated in order to obtain each of the N^3 summation terms. In this section we show that there exist sets of *equivalent interference patterns* that can be exploited in order to reduce the complexity of evaluating (3.22).

We first derive an explicit form for the magnitude of the transmitted interference pattern R_k , $k \in \{0, \dots, N-1\}$, which we will then use to show the existence of equivalent interference patterns. Note that the offset τ can be split into an integer part L and a non-integer part λ

$$L = \lfloor \tau \rfloor, \quad (3.23)$$

$$\lambda = \tau - \lfloor \tau \rfloor, \quad (3.24)$$

3.2. Symbol Error Rate Under Same-SF LoRa Interference

where L and λ correspond to the inter-chip and intra-chip misalignments between the user and the interferer, respectively.

Using the definition of the DFT and after some algebraic transformations, we have

$$R_k = \sum_{n=0}^{N-1} Z_{k,n}, \quad (3.25)$$

where $Z_{k,n}$ is defined as

$$Z_{k,n} = \begin{cases} T_1 e^{j\frac{2\pi}{N}n(s_{I_1}-k-\tau)} e^{-j2\pi\lambda}, & n \in \mathcal{N}_{L_1}, \\ T_2 e^{j\frac{2\pi}{N}n(s_{I_2}-k-\tau)}, & n \in \mathcal{N}_{L_2}, \end{cases} \quad (3.26)$$

in which T_1 and T_2 are terms that are independent of the summation variable n which are given by

$$T_1 = e^{j2\pi\frac{\tau^2}{2N}} e^{j2\pi\frac{\tau}{2}} e^{-j2\pi\frac{s_{I_1}\tau}{N}}, \quad (3.27)$$

$$T_2 = e^{j2\pi\frac{\tau^2}{2N}} e^{j2\pi\frac{\tau}{2}} e^{-j2\pi\frac{s_{I_2}\tau}{N}}. \quad (3.28)$$

Using the geometric series sum formula and after some relatively straightforward operations, R_k can be written as

$$R_k = A_{k,1} e^{-j\theta_{k,1}} + A_{k,2} e^{-j\theta_{k,2}}, \quad (3.29)$$

where

$$A_{k,1} = \frac{\sin\left(\frac{\pi}{N}(s_{I_1}-k-\tau)\lceil\tau\rceil\right)}{\sin\left(\frac{\pi}{N}(s_{I_1}-k-\tau)\right)}, \quad (3.30)$$

$$A_{k,2} = \frac{\sin\left(\frac{\pi}{N}(s_{I_2}-k-\tau)(N-\lceil\tau\rceil)\right)}{\sin\left(\frac{\pi}{N}(s_{I_2}-k-\tau)\right)}, \quad (3.31)$$

and

$$\theta_{k,1} = \frac{\pi}{N} \left(-\tau^2 + (\lambda-L)N + s_{I_1}(2\tau - \lceil\tau\rceil + 1) + k(\lceil\tau\rceil - 1) + \tau(\lceil\tau\rceil - 1) \right), \quad (3.32)$$

$$\theta_{k,2} = \frac{\pi}{N} \left(-\tau^2 + s_{I_2}(2\tau - \lceil\tau\rceil + 1 - N) + k(\lceil\tau\rceil - 1 + N) + \tau(\lceil\tau\rceil - 1) \right). \quad (3.33)$$

For the special case where τ is an integer and $k = [s_{I_1} - \tau]_N$ and $k = [s_{I_2} - \tau]_N$, (3.30) and (3.31), respectively, are of the indeterminate form $\frac{0}{0}$. Using L'Hôpital's rule, it can be shown that in these cases we have $A_{k,1} = \lceil\tau\rceil$, and $A_{k,2} = N - \lceil\tau\rceil$. Using Euler's formula, and for all $k \in \mathcal{S}$, the magnitude of R_k in (3.29) can be written as

$$|R_k| = \sqrt{A_{k,1}^2 + A_{k,2}^2 + 2A_{k,1}A_{k,2}\cos(\theta_{k,1} - \theta_{k,2})}. \quad (3.34)$$

We first give a definition for the equivalent interference patterns. Then, we show two equivalent interference pattern properties and we explain how they can be used in order to reduce the computational complexity of evaluating (3.22).

Definition 1. *An interference pattern \mathbf{y}_{I_1} is said to be equivalent with respect to some other interference pattern \mathbf{y}_{I_2} if it contains exactly the same set of frequency bin magnitudes $|R_k|$, $k \in \mathcal{S}$, irrespective of the order of these magnitudes within the set.*

We note that the ordering of the magnitudes $|R_k|$ only slightly affects the distribution of $|Y'_{\max}| = \max_{k \in \mathcal{S}/s} |Y'_k|$, because the cardinality of \mathcal{S} is large, $|\mathcal{S}| = 2^{\text{SF}}$. As such, the probability of $|Y'_{\max}| > |Y'_s|$ is only marginally affected, especially for larger SFs. Therefore, equivalent interference patterns result in approximately the same conditional SER $P(\hat{s} \neq s | s, \mathbf{y}_I)$, and hence it is sufficient to compute each distinct interference pattern once for the evaluation of the unconditional SER $P(\hat{s} \neq s)$ given in (3.22). Naturally, care has to be taken so that the contribution of each distinct interference pattern is weighted according to how many other equivalent interference patterns exist.

Proposition 1. *Let $\delta \in \{0, 1, \dots, N-1\}$ and $s_{I_1} \geq s_{I_2}$ without loss of generality and let τ be fixed. Moreover, let $s'_{I_1} = [s_{I_1} + \delta]_N$ and $s'_{I_2} = [s_{I_2} + \delta]_N$. Then there exist the following two sets of equivalent interference patterns*

$$\mathcal{Y}_{I_1} = \left\{ \mathbf{y}_I(s'_{I_1}, s'_{I_2}, \tau) : s'_{I_1} \geq s'_{I_2} \right\}, \quad (3.35)$$

$$\mathcal{Y}_{I_2} = \left\{ \mathbf{y}_I(s'_{I_1}, s'_{I_2}, \tau) : s'_{I_1} < s'_{I_2} \right\}, \quad (3.36)$$

where the interference patterns in \mathcal{Y}_{I_1} are generally not equivalent versions of the patterns in \mathcal{Y}_{I_2} . Furthermore, the cardinalities of the two sets are

$$|\mathcal{Y}_{I_1}| = N - (s_{I_1} - s_{I_2}), \quad (3.37)$$

$$|\mathcal{Y}_{I_2}| = (s_{I_1} - s_{I_2}). \quad (3.38)$$

In the special case where $\lambda = 0$ (i.e., when τ is an integer), all interference patterns in both \mathcal{Y}_{I_1} and \mathcal{Y}_{I_2} are equivalent.

A detailed proof is given in the Appendix.

Proposition 2. *Let $\tau \in [0, N-1)$ and let τ' be*

$$\tau' = (N-1) - \tau. \quad (3.39)$$

Then, the interference patterns $\mathbf{y}_I(s_{I_1}, s_{I_2}, \tau)$ and $\mathbf{y}_I(s_{I_1}, s_{I_2}, \tau')$ are equivalent.

A detailed proof is given in the Appendix.

The essence of Proposition 1 is that there are only two sets of distinct interference patterns for each value of $s_I = [s_{I_1} - s_{I_2}]_N$. Let $P_e(\mathcal{Y}_{I_i}) = P(\hat{s} \neq s | s, \mathbf{y}_{I_i}(s_I, \tau))$, where $\mathbf{y}_{I_i}(s_I, \tau)$ denotes any

3.2. Symbol Error Rate Under Same-SF LoRa Interference

(equivalent) element of \mathcal{Y}_{I_i} . Then, the double sum in (3.20) can be simplified to

$$P(\hat{s} \neq s|s) = \frac{1}{N^3} \sum_{s_{I_1}=0}^{N-1} \sum_{s_{I_2}=0}^{N-1} \int_0^N P(\hat{s} \neq s|s, \mathbf{y}_I) d\tau \quad (3.40)$$

$$= \frac{1}{N^2} \sum_{s_I=0}^{N-1} \int_0^N \left(\frac{1}{N} \sum_{i=1}^2 |\mathcal{Y}_{I_i}| P_e(\mathcal{Y}_{I_i}) \right) d\tau, \quad (3.41)$$

which reduces the complexity of evaluating (3.22) by a factor of $N/2$. In the special case of integer offsets τ (i.e., $\lambda = 0$ and $\tau = L$) the above integral can be simplified to a sum over all the integer offsets L . Moreover, in such a case, the sets \mathcal{Y}_{I_i} are equivalent, meaning that there is only one set of equivalent interference patterns \mathcal{Y}_I . Therefore, the above expression can be further simplified to the expression found in [1]

$$P(\hat{s} \neq s|s) = \frac{1}{N^2} \sum_{s_I=0}^{N-1} \sum_{L=0}^{N-1} P_e(\mathcal{Y}_I). \quad (3.42)$$

Proposition 2 essentially means that any interference pattern with $\tau \in ((N-1)/2, N-1]$ is equivalent with exactly one interference pattern with $\tau \in [0, (N-1)/2)$. It is important to note that so far we have not shown any property about the interference patterns in the region $\tau \in (N-1, N)$, so we consider this region separately.² If we let $\tilde{P}_e = \left(\frac{1}{N} \sum_{i=1}^2 |\mathcal{Y}_{I_i}| P_e(\mathcal{Y}_{I_i}) \right)$, the expression in (3.41) can be re-written as

$$P(\hat{s} \neq s|s) = \frac{1}{N^2} \sum_{s_I=0}^{N-1} \left(2 \int_0^{\frac{N-1}{2}} \tilde{P}_e d\tau + \int_{N-1}^N \tilde{P}_e d\tau \right), \quad (3.43)$$

which reduces the complexity of evaluating (3.22) by an additional factor of approximately 2. In the special case of integer offsets τ (i.e., $\lambda = 0$ and $\tau = L$) any interference pattern with $\tau \in \{N/2, N-1\}$ is equivalent with exactly one interference pattern with $\tau \in \{0, N/2-1\}$. Therefore, the above two integrals can be simplified to a summation over all the integer offsets L . In this integer-offset case we have

$$P(\hat{s} \neq s|s) = \frac{1}{N \left(\frac{N}{2} \right)} \sum_{s_I=0}^{N-1} \sum_{L=0}^{\frac{N}{2}-1} P_e(\mathcal{Y}_I). \quad (3.44)$$

We note that the corresponding simplification that is used in [1], corresponding to the special chip-aligned case, is different than the simplification we gave in (3.44). Specifically, in [1] the upper limit of the sum over L is $N/2$ and, consequently, the normalization is done with $\left(\frac{N}{2} + 1 \right)$ instead of $\frac{N}{2}$. However, the interference pattern resulting from $\tau = \frac{N}{2}$ is equivalent with the interference pattern resulting from $\tau = \frac{N}{2} - 1$, which has already been considered in the sum.

²It is relatively simple to verify that the region $\tau \in (N-1, N)$ contains equivalent interference patterns around the point $\tau = N - \frac{1}{2}$ following the syllogism of the proof of Property 2. However, this property only marginally reduces the computational complexity and we thus neither use it nor prove it explicitly.

3.2.4 Symbol Error Rate Approximation

Since, even with the above simplifications, the complexity of evaluating (3.22) is very high, we derive a low-complexity approximation for (3.22). Using the triangle inequality, we can simplify (3.34) to

$$|R_k| \approx |A_{k,1}| + |A_{k,2}|. \quad (3.45)$$

With this simplification, the $A_{k,1} A_{k,2} \cos(\theta_{k,1} - \theta_{k,2})$ term that leads to the existence of two sets of equivalent interference patterns \mathcal{Y}_{I_1} and \mathcal{Y}_{I_2} (cf. proof of Proposition 1) disappears. Thus, there is only a single set of interference patterns \mathcal{Y}_I and (3.41) can be simplified to

$$P(\hat{s} \neq s|s) \approx \frac{1}{N^2} \sum_{s_I=0}^{N-1} \int_0^N P_e(\mathcal{Y}_I) d\tau. \quad (3.46)$$

Moreover, we also approximate (3.43) by ignoring the second integral for $\tau \in (N-1, N)$ so that

$$P(\hat{s} \neq s|s) \approx \frac{2}{N^2} \sum_{s_I=0}^{N-1} \int_0^{\frac{N-1}{2}} P_e(\mathcal{Y}_I) d\tau, \quad (3.47)$$

We now follow the following procedure in order to derive a simple approximation for $P_e(\mathcal{Y}_I)$. First, we assume that the interference-induced SER is dominated by the maximum of $|R_k|$. Thus, we are interested in evaluating

$$|R_{k_{\max}}| = \max_k (|A_{k,1}| + |A_{k,2}|). \quad (3.48)$$

Without loss of generality, we assume that $s_{I_2} = 0$, so that $s_I = s_{I_1}$. Since $\tau \in [0, (N-1)/2]$ and due to (3.30) and (3.31) it holds that $\max_k (|A_{k,2}|) > \max_k (|A_{k,1}|)$. Based on this observation, we choose

$$k_{\max} \approx \arg \max_k (|A_{k,2}|) = \lfloor \tau \rfloor, \quad (3.49)$$

so that we can easily approximate $|R_{k_{\max}}|$ as

$$|R_{k_{\max}}| \approx |A_{\lfloor \tau \rfloor,1}| + |A_{\lfloor \tau \rfloor,2}|. \quad (3.50)$$

The probability of the event that the (maximum) interference bin $\lfloor \tau \rfloor$ coincides with the bin of the signal-of-interest s is $\frac{1}{N}$. Since in LoRa N is relatively large ($N > 2^7$), the impact of the aforementioned event on the total error probability is negligible, and therefore, for the approximation of the SER, we only consider the cases where $\lfloor \tau \rfloor \neq s$.³ The aforementioned fact, combined with (3.50) which says that all bins except s and $\lfloor \tau \rfloor$ are zero-valued, means that

³We note that wherever this simplification becomes relevant, e.g., in the infinite-SNR region, the derived approximation loses accuracy. However, we show in the results that the aforementioned simplification does not lead to inaccuracies in the performance evaluation of non-coherent receivers under interference in the operational regime that is relevant for LoRa.

3.2. Symbol Error Rate Under Same-SF LoRa Interference

the approximation of the SER does not depend on the value of s . As such, $P(\hat{s} \neq s | s) \approx P(\hat{s} \neq s)$ and calculating the expectation over s can be avoided. Only considering $\lfloor \tau \rfloor \neq s$ also has the convenient side-effect that we ignore the only case of (3.16) which contains ω , meaning that we can entirely avoid the integration over ω in the computation of $P_e(\mathcal{B}_I)$. Let $P^{(I)}(\hat{s} \neq s)$ denote the interference-dominated SER resulting from the approximation in (3.50). As explained in Section 4.2.1, $|Y'_{k_{\max}}|$ follows a Rice distribution, which can be approximated by a Gaussian distribution for large location parameters [1] so that

$$|Y'_{k_{\max}}| \sim \mathcal{N}\left(\frac{|h_I||R_{k_{\max}}|}{\sigma}, 1\right). \quad (3.51)$$

Using the Gaussian approximation, the interference-dominated SER $P^{(I)}(\hat{s} \neq s)$ can be computed as

$$P^{(I)}(\hat{s} \neq s) \approx \frac{1}{N\left(\frac{N}{2}\right)} \sum_{s_I=0}^{N-1} \int_0^{\frac{N-1}{2}} Q\left(\frac{N - |h_I||R_{k_{\max}}|}{\sqrt{2}\sigma^2}\right) d\tau, \quad (3.52)$$

where $Q(\cdot)$ denotes the Q-function and the integral can be evaluated numerically by discretizing the interval $[0, (N-1)/2]$ with a step size ϵ as

$$P^{(I)}(\hat{s} \neq s) \approx \frac{\epsilon}{N\left(\frac{N}{2}\right)} \sum_{s_I=0}^{N-1} \sum_{\tau \in \mathcal{T}} Q\left(\frac{N - |h_I||R_{k_{\max}}|}{\sqrt{2}\sigma^2}\right), \quad (3.53)$$

where $\mathcal{T} = \{0, \epsilon, 2\epsilon, \dots, \frac{N-1}{2} - \epsilon\}$.

We note that, in the low SNR (i.e., AWGN-limited) regime, the above approximation becomes inaccurate, since all bins have similar values and no single bin dominates the error rate. To fix this inaccuracy, let $P^{(N)}(\hat{s} \neq s)$ denote the SER under AWGN given in (3.6) (which can be evaluated efficiently using the approximation in (3.7)). Then, a final estimate of the SER that is more accurate also in the low SNR regime [1] can be obtained as

$$P(\hat{s} \neq s) \approx P^{(N)}(\hat{s} \neq s) + (1 - P^{(N)}(\hat{s} \neq s)) P^{(I)}(\hat{s} \neq s). \quad (3.54)$$

Finally, we note that the derivation of upper bounds for the LoRa SER, with and without interference, as well as the derivation of performance bounds for the infinite-SNR regime, is an interesting and important theoretical problem that is worth to be addressed in the future literature. Without underestimating the importance of the aforementioned approach, in this thesis we provide performance models of the LoRa PHY under practical constraints, as well as to provide accurate low-complexity approximations that accurately model the behavior of the examined systems in the operating SINR regimes of LoRa.

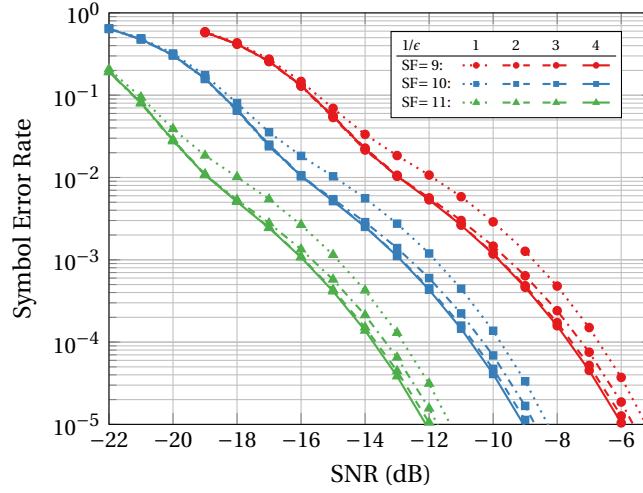


Figure 3.5 – Symbol error rate approximation of the LoRa modulation under same-SF interference and AWGN for $SF \in \{9, 10, 11\}$ and $SIR = 3$ dB, for various values of the oversampling factor $1/\epsilon$.

3.2.5 Results

In this subsection, we provide numerical results for the SER of LoRa with same-SF interference. The derived low-complexity approximations are compared to Monte Carlo simulations to check the accuracy of the approximations in the operational SNR regime of LoRa. We note that for the Monte Carlo simulations, the interferer time-offset τ is simulated by oversampling (2.3) to create s_{I_1} and s_{I_2} , concatenating the oversampled symbols, applying the appropriate offset, and downsampling to obtain $N = 2^{SF}$ samples for \mathbf{x}_I .

Choice of Discretization Step

In Fig. 3.5 we show the evaluation of (3.53) with different discretization steps ϵ (or, equivalently, oversampling factors $1/\epsilon$). We observe that very small gains in accuracy are obtained after $\epsilon = 1/3$. As such, in the remainder of this section we use $\epsilon = 1/5$ to err on the side of caution. This choice also means that the complexity of evaluating (3.53) is not significantly higher than the complexity of evaluating the corresponding approximation in [1] (which is obtained for $\epsilon = 1$). In general, the complexity of evaluating (3.53) is $1/\epsilon$ times higher than the complexity of evaluating the corresponding approximation in [1], but still very low compared to simulation-based studies. We note that for the Monte Carlo simulations we use an even more conservative $\epsilon = 1/10$, since they are used as a comparison baseline.

Symbol Error Rate

In Fig. 3.6, we show the results of a Monte Carlo simulation for the SER of a LoRa user for all possible spreading factors $SF \in \{7, \dots, 12\}$, under the effect of same-SF interference with an

3.2. Symbol Error Rate Under Same-SF LoRa Interference

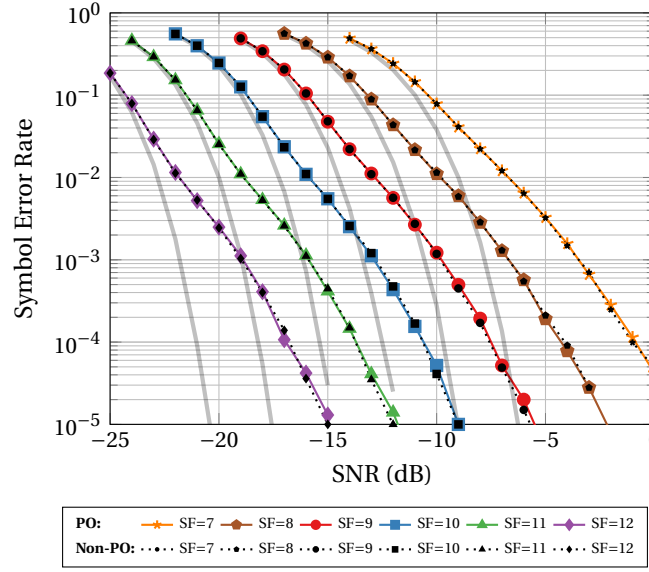


Figure 3.6 – Symbol error rate of the LoRa modulation under AWGN and same-SF interference for $SF \in \{7, \dots, 12\}$ and $SIR = 3$ dB. Black dotted lines show the SER when ignoring the phase offset ω and thick transparent lines show the SER when there is only AWGN for comparison (taken from Fig. 3.1).

SIR of 3 dB (i.e., $P_I = -3$ dB) and AWGN. The SER when there is only AWGN is also included in the figure with thick transparent lines (taken from Fig. 3.1). We can clearly observe the strong impact of the interference on the SER when comparing to the case where there is only AWGN. The black dotted lines in the figure depict the SER when the relative phase offset ω between the interferer and the user is not taken into account in the Monte Carlo simulation. It is interesting to observe that ω does not play an important role for the SER in the useful operating SNR regime for LoRa, which further justifies ignoring ω in the approximation of Section 3.2.4.

In Fig. 3.7, we show the results of a Monte Carlo simulation for the SER of a LoRa user with $SF \in \{9, 10, 11\}$ using the chip-aligned model of [1] and the non-aligned model we described in this chapter, as well as the corresponding approximations in (3.54) and [1], respectively. We observe that there is a significant difference of approximately 1 dB between the two models, and that the chip-aligned model of [1] is pessimistic in the computation of the SER. This can be intuitively explained as follows. When the offset τ is an integer, the maximum value of the interference magnitudes $A_{k,1}$ and $A_{k,2}$ with respect to the index k is always larger than when τ is not an integer. As such, considering only chip-aligned interference is a worst-case scenario. Finally, we observe that the low-complexity computation of (3.22) using (3.54) is very accurate in the useful operating SNR regime for LoRa. We note that in the infinite-SNR regime the Monte Carlo simulation curves will reach an error floor due to the impact of the interference on the bin of the transmitted symbol s with a phase ω . The error floor is different for different SIR values and different SFs and exists due to the following reason: a large interference at

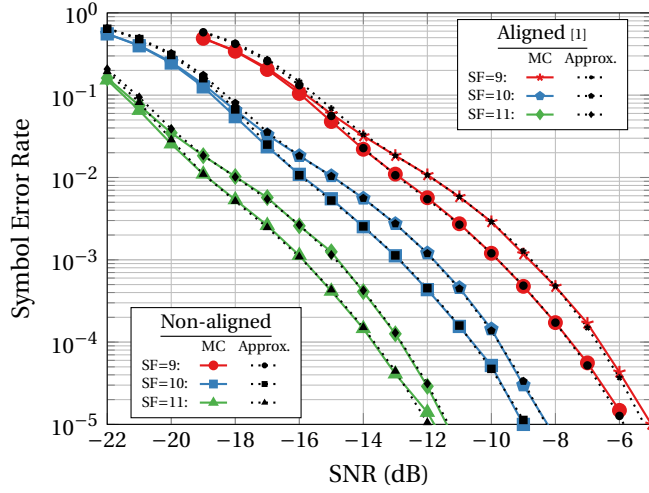


Figure 3.7 – Symbol error rate of the LoRa modulation under AWGN and same-SF interference for $SF \in \{9, 10, 11\}$ and $SIR = 3$ dB. The approximations of [1] and (3.54) are shown with black dotted lines.

bin s will significantly decrease the magnitude of bin s for some ranges of values for ω , as shown in Fig. 3.4. In such cases, one of the remaining bins that contain only interference can have a larger magnitude than bin s , even in the absence of AWGN. Thus, in some cases, a demodulation error will always happen in the infinite-SNR regime, leading to an error floor in the average SER. The approximation in (3.54) does not capture this error-floor effect, since (3.54) does not include the impact of the interference at bin s . The SER of the curves resulting from the approximation in (3.54) will continue decreasing even for infinite SNR.

In Fig. 3.8, we show the required SNR for a target SER performance of $2 \cdot 10^{-5}$, for different SIR levels and for the two extremal spreading factors, i.e., $SF = 7$ and $SF = 12$. The target SER performance of $2 \cdot 10^{-5}$ is chosen similarly to [1] and in accordance with the sensitivity thresholds for a bandwidth of $B = 125$ kHz and a coding rate of $R = 4/5$, as provided by Semtech [41]. The SNR vs SIR plot is important for every framework that considers AWGN and interference jointly and was introduced in [1]. It is obvious that as the interference power increases, there is a significant increase in the required SNR to obtain the same error-rate performance. We can observe that the chip-aligned model of [1] overestimates the required SNR increase for both the extremal spreading factors. Moreover, the overestimation is slightly more pronounced for higher levels of interference, since in the interference-limited regime, the impact of the overestimation of the chip-aligned model is higher than in the noise-limited regime.

It is important to note that the framework that treats the signal and the interference in a unified fashion, which is used both in [1, 113] and in this work, does not contradict, but rather enhances the view of the non-unified framework of [129, 104, 111]. In particular, in the non-unified framework, the SNR and the SIR are considered independently, and a LoRa

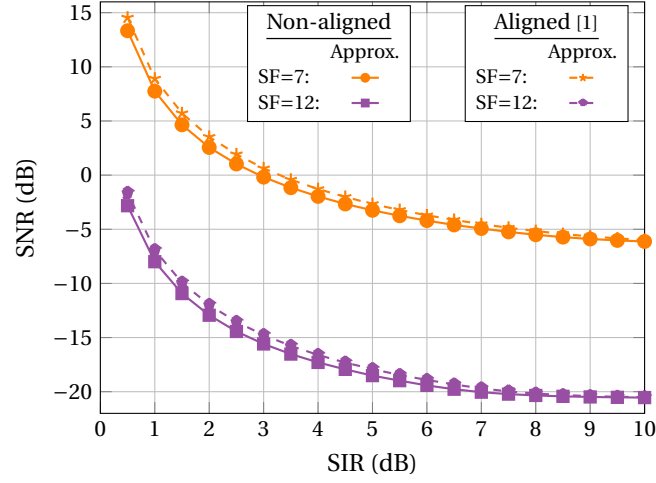


Figure 3.8 – Required SNR for a target symbol error rate of $2 \cdot 10^{-5}$ as a function of the SIR for SF = 7 and SF = 12.

message is received successfully only if both of the following two assumptions hold

1. $\text{SNR} > \text{SNR}_{\text{thr}}^{(\text{SF})}$, where $\text{SNR}_{\text{thr}}^{(\text{SF})}$ is the SF-specific SNR threshold for a given target error probability, and
2. $\text{SIR} > 6 \text{ dB}$

In the unified SINR framework, a message can potentially survive even if $\text{SIR} < 6 \text{ dB}$ if in turn the SNR is high enough. This SNR-SIR trade-off is in essence the information that Fig. 3.8 provides. Therefore, the unified SINR framework allows for a softer decision threshold on the successful reception of a LoRa message, rather than the hard 6 dB SIR threshold that is commonly used in the literature. Moreover, the availability of an error rate allows a probabilistic setup rather than a threshold for determining if a packet is actually deleted. The effect of multiple same-SF interferers can only be handled approximately under the non-unified framework. As explained in [129], the analysis for multiple same-SF interferers can be simplified by considering only the most powerful interferer in the SIR value; the larger the number of same-SF users in the network, the higher the probability that the most powerful interferer will lead to a critical SIR (e.g., $< 6 \text{ dB}$). The effect of multiple same-SF interferers on the error rate under the unified SINR framework is a field that needs to be explored further.

3.3 Frame Error Rate Under Same-SF Interference

Since network simulators, such as the ones presented in [104, 151, 155, 148], typically operate on a frame level, the frame error rate (FER) is generally of greater practical relevance than the SER. The expression for the SER derived in Section 3.2 can not be used directly to evaluate the FER because it includes an expectation over τ and an expectation over ω , while all symbols

in a frame would experience the same τ and ω . For this reason, in this section we derive an expression to approximate the FER. We note that we are considering the FER of an uncoded system. Thus, the expression we derive can only be used for the LoRa modes that use channel codes of rate $4/5$ and $4/6$, which have error-detection, but no error-correction capabilities, as well as for the uncoded mode.

3.3.1 Frame Error Rate Expression and Approximation

We first make the following simplifying assumptions. We assume that perfect frame synchronization for the user of interest is achieved even in the presence of interference. Then, we also assume that the interfering frame has the same length as the frame of interest. The latter assumption is taken only for clarity of the presentation and does not prevent our results from being easily generalizable to any interfering frame length. Due to the time offset between the frames, only part of the frame of interest is affected by interference.

Let the vector \mathbf{s} denote a frame of F LoRa symbols and let the vector $\hat{\mathbf{s}}$ denote the estimated frame at the receiver. The FER can then be defined as $P(\hat{\mathbf{s}} \neq \mathbf{s})$. Moreover, let F_I , where $F_I \in \{1, \dots, F\}$, denote the number of symbols in the frame that are affected by the interfering frame. The value of F_I depends on the relative position of the two frames. As mentioned in [110], the random relative position of the two frames plays an important role on the final FER. We consider the final FER as the expectation over all the possible relative positions of the two frames. We note that, except in the case of perfect alignment between the frame of interest and the interference, there always exists one symbol that is only partially affected by interference. For simplicity, we approximate this situation by considering the partially-affected symbol as fully-affected by interference, thus including it in F_I . The number of symbols in the frame that are affected only by AWGN is $F - F_I$. Since we consider the same length for the frame of interest and the interfering frame, the number F_I of interfered symbols in a frame is uniformly distributed in $\{1, \dots, F\}$.

Since all symbols of the interfering frame have the same offset τ , the probability, for a given F_I , that all F_I symbols under interference and all $F - F_I$ symbols under AWGN are correct is

$$P(\hat{\mathbf{s}} = \mathbf{s} | F_I, \tau) = (1 - P(\hat{\mathbf{s}} \neq \mathbf{s} | \tau))^{F_I} (1 - P^{(N)}(\hat{\mathbf{s}} \neq \mathbf{s}))^{F - F_I}, \quad (3.55)$$

where $P(\hat{\mathbf{s}} \neq \mathbf{s} | \tau)$ can be approximated as

$$P(\hat{\mathbf{s}} \neq \mathbf{s} | \tau) = \frac{1}{N} \sum_{s_I=0}^{N-1} Q\left(\frac{N - |h_I| |R_{k_{\max}}|}{\sqrt{2\sigma^2}}\right). \quad (3.56)$$

By using a similar simplification to (3.47), the conditional frame error rate $P(\hat{\mathbf{s}} \neq \mathbf{s} | F_I)$ can be approximated as

$$P(\hat{\mathbf{s}} \neq \mathbf{s} | F_I) \approx \frac{2}{N} \int_0^{\frac{N-1}{2}} (1 - P(\hat{\mathbf{s}} = \mathbf{s} | F_I, \tau)) d\tau. \quad (3.57)$$

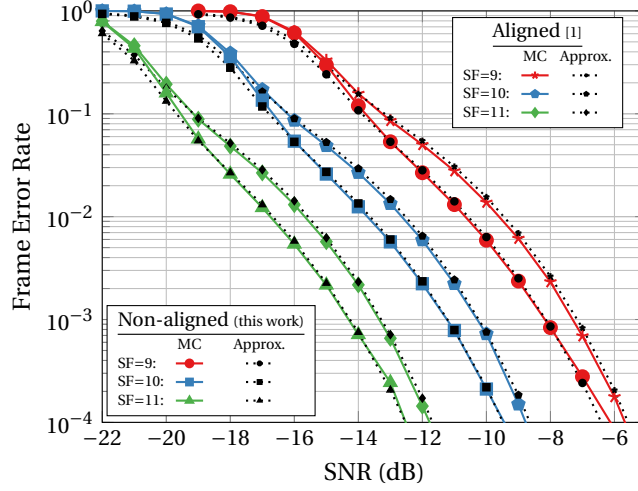


Figure 3.9 – Frame error rate of the LoRa modulation for a frame length $F = 10$ under AWGN and same-SF interference for $SF \in \{9, 10, 11\}$ and $SIR = 3$ dB. The approximations of [1] and (3.54) are shown with black dotted lines.

Finally, we take the expectation over all possible values of F_I and we obtain the final expression for the FER

$$P(\hat{\mathbf{s}} \neq \mathbf{s}) \approx \frac{1}{F} \sum_{F_I=1}^F P(\hat{\mathbf{s}} \neq \mathbf{s} | F_I). \quad (3.58)$$

3.3.2 Results

In Fig. 3.9, we show the results of a Monte Carlo simulation for the FER of a LoRa user with $SF \in \{9, 10, 11\}$ using the chip-aligned model of [1] and the model we described in this work, as well as the corresponding approximation described in Section 3.3.1. The frame contains $F = 10$ LoRa symbols, which is a valid data payload length for LoRa. We observe the same difference of approximately 1 dB between the two models. Moreover, we can see that the approximation for the FER described in Section 3.3 is very accurate. We note that the error-floor discussion of Section 3.2.5 applies also here for the FER curves. The error floor for the FER depends also on the frame length, additionally to the SIR value and the SF.

In Fig. 3.10, we show the required SNR for a target FER performance of 10^{-1} , for different SIR levels, and for the two extremal spreading factors, i.e., $SF = 7$ and $SF = 12$. Similarly to Fig. 3.8, as the interference power increases, there is an increase in the required SNR to obtain the same performance. For the chosen target FER performance of 10^{-1} we can observe that the overestimation of the chip-aligned model of [1] is clearly more pronounced for higher levels of interference.

Finally, in Fig. 3.11, we show the required SNR for two different target FERs and three different

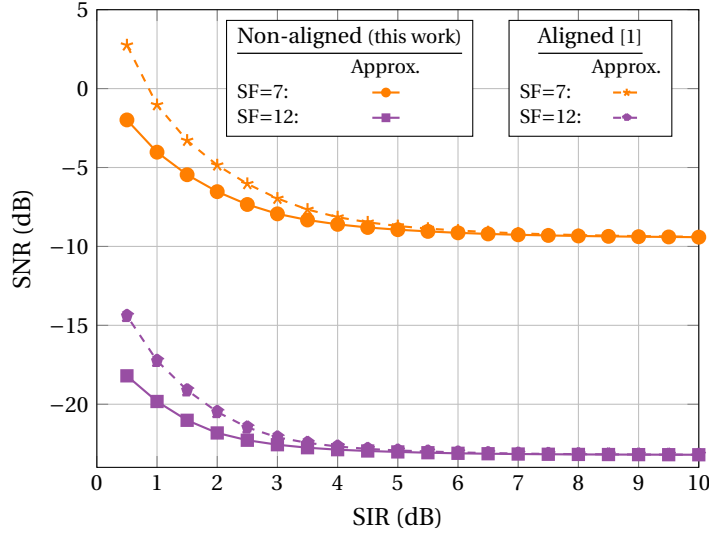


Figure 3.10 – Required SNR for a target frame error rate of 10^{-1} as a function of the SIR for SF = 7 and SF = 12.

frame lengths, for different SIR levels and for SF = 7. Specifically, we choose both a typical target FER performance of 10^{-1} and a stricter target FER of 10^{-2} [110], and we choose frames of length $F = 10$, $F = 20$, and $F = 30$ LoRa symbols. We observe that longer frames require a larger increase in the required SNR for successful reception under the same interference power. Moreover, the performance requirement plays an important role, since the increase in the required SNR values for a packet to survive under a stricter target FER performance is very pronounced.

3.4 Discussion and Summary

In this chapter, we introduced a unified probabilistic interference-and-noise model for LoRa PHY. We tried to avoid over-simplistic assumptions that could affect the performance results, e.g., any type of alignment between the LoRa interferer and the LoRa user. In particular, in Section 3.1 we showed an expression for the SER under AWGN, as well as a corresponding low-complexity approximation. In Section 3.2 we derived an expression for the SER under same-SF interference and AWGN. Moreover, we proved two properties of same-SF LoRa-induced interference that enabled us to reduce the complexity of calculating the SER by a factor that is approximately equal to the LoRa symbol length N . Furthermore, we derived a low-complexity approximation for both the SER and the FER, and finally, in Section 3.3, we derived the expression for the FER of LoRa under interference and AWGN.

The conducted analysis in this chapter gives a deeper understanding of how to model the LoRa interference as well as how to attempt to overcome the problems it introduces. By using the results from this chapter, we propose such an improvement in Chapter 4 of this

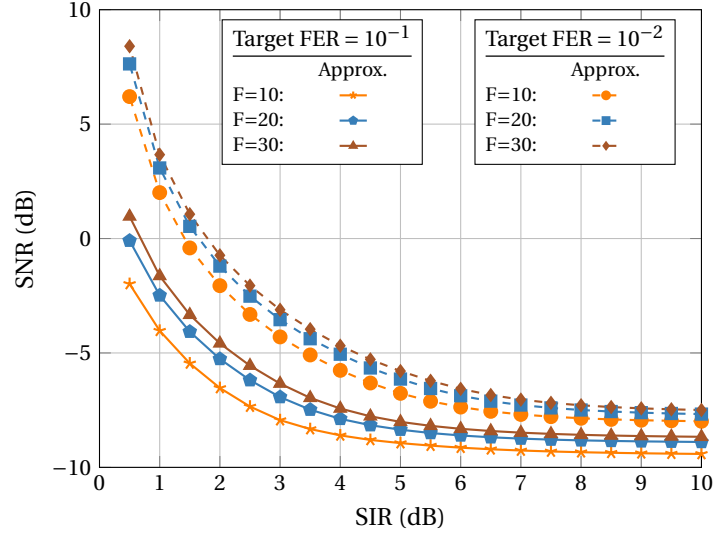


Figure 3.11 – Required SNR for target frame error rates of 10^{-1} and 10^{-2} , for frame lengths $F = 10, 20, 30$, as a function of the SIR for SF = 7.

thesis. Furthermore, the model of this chapter can be used directly as an input to system-level LoRaWAN simulators to evaluate LoRaWAN scalability in a more accurate framework than the more simplistic models currently in use. We note that in this chapter we have only modeled the uncoded modes of the LoRa PHY. We further expand this error-rate model to the coded modes of the LoRa PHY in Chapter 5 of this thesis.

4 Coherent LoRa Detection under Same-Technology Interference

In the previous chapter we have analyzed the performance of a typical non-coherent LoRa receiver under same-SF interference. We have seen that same-SF interference can result in severe performance degradation, especially for low SIR values. LoRaWAN can be designed to allocate the transmission parameters of the end-nodes, e.g., the spreading factor and the transmit power, to minimize the catastrophic collisions between end-nodes that use the same-SF. However, this task is not always easy, especially for large networks where the probability of colliding packets inevitably becomes very high. LoRaWAN would therefore directly benefit from a LoRa receiver with lower probability of demodulation errors under same-SF interference than the one described in the previous chapter.

To this end, we propose coherent LoRa detection as a way to improve the performance of LoRa receivers under same-SF interference. The aim of this chapter is to present an in-depth error-rate analysis of LoRa receivers with coherent detection, using the knowledge we have gained from performing the error-rate analysis in the previous chapter for non-coherent receivers. We show that the coherent detection of LoRa is a good way to improve the demodulation robustness of colliding LoRa packets directly in the physical layer, with only a very small increase in the receiver complexity.

In particular, in Section 4.1 we explain why the coherent detection of LoRa can be useful for colliding packets, as well as the differences between a coherent and a non-coherent LoRa receiver and how a non-coherent receiver can be designed. We start our analysis in Section 4.2 with modeling the performance of coherent LoRa receivers under same-SF interference from another LoRa user, and we derive an expression for the symbol error rate (SER). Since this expression is difficult to evaluate in practice, in Section 4.3 we also derive a corresponding low-complexity approximation that is based on Q-functions, and is accurate in the operating SNR regime of LoRa, similarly to the previous chapter. Finally, in Section 4.4, we extend our analysis to the frame error rate (FER), which is more useful for LoRa network simulators. We note that in this chapter we go one step further in the direction of building a realistic model by including the carrier frequency offset (CFO) of the interferer in our model.

4.1 Coherent LoRa Detection: How and Why?

The general discrete-time baseband-equivalent description of a LoRa symbol s can be written in two forms. In the first form, the phase of the first chip of each LoRa symbol is not clearly and unambiguously defined. This description is therefore not suited for coherent demodulation. In the second form, as shown in (2.2) and (2.3), all LoRa symbols start with the same phase. Therefore, only the second form has inter-symbol phase continuity, which has been mentioned in the LoRa patent [42] as a desirable property.

Coherent LoRa detection would necessitate a phase rotation of the received signal to compensate for the channel phase rotation, i.e., $\mathbf{U} = \mathbf{Y}e^{-j\phi}$. The phase shift ϕ introduced by the transmission channel can be estimated using the preamble LoRa symbols, which can be used as pilots. Due to the inter-symbol phase continuity property of (2.3), every transmitted preamble symbol starts with a phase of zero. Therefore, the starting phase of every received preamble symbol should be ϕ . After the preamble symbols are used to synchronize the LoRa packet on a sample level as will be explained in Chapter 6, the samples of the preamble symbols can be realigned, so that the first sample of each preamble symbol can be extracted. The estimation of ϕ should be done in the frequency domain after dechirping the preamble symbols, since the first bin of the DFT of the dechirped symbol contains all the signal energy and, thus, has a higher SNR than any sample in the time domain. The phase estimation should be performed over all preamble symbols, to average out the noise. Therefore, ϕ can be estimated as

$$\hat{\phi} = \arg \left(\sum_{i=1}^{N_{pr}} Y_1^{(i)} \right), \quad (4.1)$$

where N_{pr} is the number of preamble symbols in the LoRa packet, and $Y_1^{(i)}$ is the first frequency bin of the i -th preamble symbol.

After the rotation, coherent detection is performed by selecting the bin index from the maximum projection on the real axis [47, 92, 125]

$$\hat{s} = \arg \max_{k \in \mathcal{S}} (\Re(U_k)). \quad (4.2)$$

We will briefly present the intuition behind proposing coherent LoRa detection as a good receiver candidate for colliding LoRa packets. The standard non-coherent LoRa receiver is affected by AWGN in both real and imaginary axes, where AWGN has the same variance σ^2 on both axes. Due to the large values of $N = 2^{\text{SF}}$ in LoRa, and since every frequency bin has a different phase, we can imagine the N bins of the DFT of the dechirped received symbol to be with high probability all over $[0, 2\pi)$. Coherent LoRa detection performs first a phase correction of the received signal to compensate for the channel phase rotation. Hence, all the bins would change angles, so that the bin s would lie on top of the positive real axis. Therefore, under coherent detection, bin s does not lose any energy due to the projection on the real axis. However, all the remaining bins, containing only AWGN, lose part of the noise energy

by projecting onto the real axis. Nevertheless, since the AWGN bins are all around the circle, the chance that some bins are lying close to the positive real axis, even after the rotation, is very high. Therefore, the maximum projection of the noise in a coherent receiver, would not be that much lower than the maximum magnitude of the noise in a non-coherent receiver. This is the intuitive explanation of the small performance improvement – around 0.7 dB – of coherent LoRa detection in AWGN reported in the literature [47, 92, 125].

Let us now examine the case where both interference and noise are present. Although AWGN lies in all bins all around the circle with the same average power, the same does not hold for interference. As seen also in the previous chapter, there exist some bins with considerably higher power than the remaining bins, as illustrated in Fig. 4.2 and Fig. 4.3. The distribution of the high-power interference bins all around the circle is also random due to the interference channel realization. However, since the number of high-power bins is very small, there is a considerable probability that the high-power bins for a given transmission will lie at such angles that their projection on the real axis will be considerably smaller than their magnitude. Therefore, the coherent detection of colliding LoRa packets does not remove any energy from the demodulated signal of interest, but on average removes considerable amount of energy from the interfering signal. To verify the impact of these observations on the performance, in the following sections of this chapter, we will model and analyze the performance of coherent LoRa detection under both same-SF interference and AWGN.

4.1.1 Symbol Error Rate Under AWGN

We first show how the SER of coherent LoRa can be modeled under AWGN, as well as how the error rate under AWGN has been evaluated in the existing literature. As in the previous chapter, the AWGN discussion is useful both on its own, and also for later describing how the symbol error rate (SER) can be calculated in the presence of both interference and AWGN.

As in the previous chapter, we consider that in an ideal noiseless receiver and under perfect sample synchronization, the DFT of the dechirped signal \mathbf{Y} would result in a single frequency bin which would contain all the signal energy (i.e., a bin that has magnitude N) and all remaining $N - 1$ bins would have zero energy. In a real receiver, where AWGN is present, no bin will have zero energy. As we showed in the previous chapter, the distribution of the frequency bin values Y_k for $k \in \mathcal{S}$ is given by (3.1).

After the compensation for the channel phase by the coherent receiver, the frequency bin s , which contains all of the transmitted signal energy, lies exactly on the positive real axis. Therefore, the demodulation metric $\Re(U_s)$ of the coherent receiver for bin s does not remove any useful-signal energy, but the projections of all the remaining bins remove part of the energy of the noise. The corresponding demodulation metric then follows a normal distribution with

zero mean for $k \in \mathcal{S}/s$ and a normal distribution with mean N for $k = s$

$$\Re(U_k) \sim \begin{cases} \mathcal{N}(0, \sigma^2), & k \in \mathcal{S}/s \\ \mathcal{N}(N, \sigma^2), & k = s. \end{cases} \quad (4.3)$$

A symbol error occurs if and only if any of the $\Re(U_k)$ values for $k \in \mathcal{S}/s$ exceeds the value of $\Re(U_s)$, or, equivalently, if and only if $\Re(U_{\max}) > \Re(U_s)$, where $\Re(U_{\max}) = \max_{k \in \mathcal{S}/s} \Re(U_k)$. The probability density function (PDF) of bin s is $f_{\Re(U_s)}(y) = \mathcal{N}(N, \sigma^2)$ and the cumulative distribution function (CDF) of the maximum projection of the remaining bins, $\Re(U_{\max})$, is

$$F_{\Re(U_{\max})}(y) = \left(1 - Q\left(\frac{y}{\sigma}\right)\right)^{(N-1)}, \quad (4.4)$$

where $Q(\cdot)$ denotes the Q-function. Therefore, the probability of symbol error for a given transmitted symbol s is

$$P(\hat{s} \neq s|s) = \frac{1}{\sigma\sqrt{2\pi}} \int_{y=0}^{+\infty} \left(1 - \left(1 - Q\left(\frac{y}{\sigma}\right)\right)^{(N-1)}\right) e^{-\frac{(y-N)^2}{2\sigma^2}} dy. \quad (4.5)$$

The SER for all symbols s is identical, therefore (4.5) is also equal to the expected SER $P(\hat{s} \neq s)$.

The SER in (4.5) is similar to [125, Eq. (17)], where the authors explain how to evaluate it numerically without suffering from numerical problems. The authors in [47] give a low-complexity approximation for the coherent LoRa SER under only AWGN. The approximation from [47], which we write below using our notation, was derived using curve fitting.

$$P(\hat{s} \neq s|s) \approx Q\left(\frac{1 - \sqrt{\sigma^2}(1.161 + 0.2074 \cdot \text{SF})}{\sqrt{\sigma^2 + \sigma^2(0.2775 - 0.0153 \cdot \text{SF})}}\right). \quad (4.6)$$

The approximation in (4.6) can be evaluated with very low complexity and is very accurate in the operational SNR regime of LoRa.

We note here that using Newton's binomial identity, we have

$$1 - \left(1 - Q\left(\frac{y}{\sigma}\right)\right)^{(N-1)} = \sum_{q=1}^{N-1} (-1)^{q+1} \binom{N-1}{q} \left(Q\left(\frac{y}{\sigma}\right)\right)^q, \quad (4.7)$$

and thus we can write (4.5) as

$$P(\hat{s} \neq s|s) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{q=1}^{N-1} (-1)^{q+1} \binom{N-1}{q} \int_{y=0}^{+\infty} \left(Q\left(\frac{y}{\sigma}\right)\right)^q e^{-\frac{(y-N)^2}{2\sigma^2}} dy. \quad (4.8)$$

The SER written in the form of (4.8) contains integer powers of the Q-function, and thus it can be evaluated using the simple and tight approximation for the integer powers of the Q-function of [165].

4.2 Coherent LoRa Symbol Error Rate Under Same-SF Interference

In this section we will derive the SER expression of LoRa with coherent detection under same-SF interference as well as a corresponding practical low-complexity approximation.

4.2.1 Symbol Error Rate Under Same-SF Interference

As in the previous chapter for the non-coherent receiver, we assume that a LoRa gateway is trying to decode a user's message while at the same time a same-SF interfering packet is superimposed on the user's packet. The gateway is perfectly synchronized to the user, but the interfering signal $y_I[n] = h_I x_I[n]$ is not synchronized in any way to the user or the gateway. Moreover, we assume that any carrier frequency offset (CFO) of the synchronized user has been perfectly estimated and compensated. The gateway can perform CFO estimation and compensation only for one user, therefore, there is no CFO estimation and compensation for the interfering user. Thus, in this chapter, we also include the CFO of the interferer in order to obtain an even more accurate model than in the previous chapter.

As shown in Fig. 3.2, and as explained in the previous chapter, due to the lack of any synchronization, the interfering signal $x_I[n]$ will generally comprise parts of two distinct LoRa symbols s_{I_1} and s_{I_2} . Following the notation of Chapter 3, let τ be the relative time-offset between the first chip of the transmitted symbol of interest s and the first chip of the second interfering symbol s_{I_2} (i.e., the first chip of s_{I_2} starts τ chip durations *after* the first chip of s). We note that the offset τ can be split into an integer part $L = \lfloor \tau \rfloor$, and a non-integer part $\lambda = \tau - \lfloor \tau \rfloor$. We consider that τ has a uniform distribution in $[0, N)$, due to the complete lack of synchronization. A simplified version of the discrete-time baseband-equivalent equation of the transmitted interference signal $x_I[n]$ is given in (3.11).

Further, let f_{c_1} be the carrier frequency used during up-conversion at the transmitter of the interfering user and f_{c_2} be the carrier frequency used during down-conversion at the gateway after alignment with the carrier frequency of the desired user. The carrier frequency offset of the interfering user is the difference $\Delta f_c = f_{c_1} - f_{c_2}$, while any frequency offset of the desired user is perfectly compensated. As a result, the corresponding signal model is

$$y[n] = hx[n] + h_I c_I[n] x_I[n] + z[n], \quad n \in \mathcal{S}, \quad (4.9)$$

where h is the channel gain between the user and the LoRa gateway, $x[n]$ is the transmitted signal, h_I is the channel gain between the interferer and the gateway, $x_I[n]$ is the transmitted interfering signal, $c_I[n] = e^{j2\pi(n+(m-1)N)\frac{\Delta f_c}{f_s}}$ is the CFO term affecting the m -th symbol in the interfering packet, and $z[n] \sim \mathcal{N}(0, \sigma^2)$ is AWGN. We note that the CFO translates into an offset $\tau_{\text{cfo}} = \frac{\Delta f_c N}{f_s}$ [36], and similarly to τ , the offset τ_{cfo} can be split into an integer part $L_{\text{cfo}} = \lfloor \tau_{\text{cfo}} \rfloor$, and a non-integer part $\lambda_{\text{cfo}} = \tau_{\text{cfo}} - \lfloor \tau_{\text{cfo}} \rfloor$. Since $|h| = 1$, the signal-to-interference ratio (SIR) can be defined as $\text{SIR} = \frac{1}{P_I}$, where $P_I = |h_I|^2$ is the received power of the interfering user at the

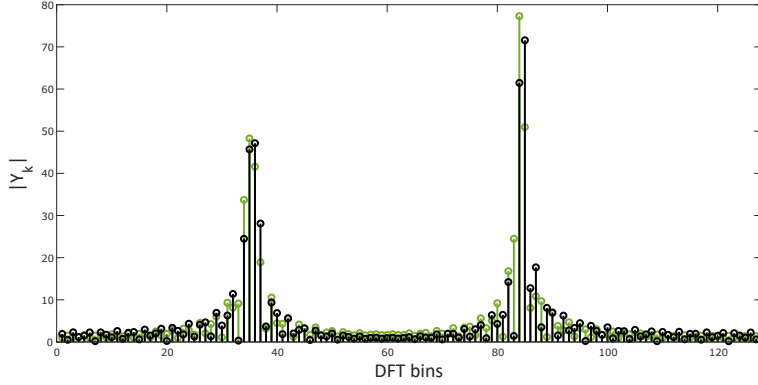


Figure 4.1 – The magnitude of an example transmitted (green) and received (black) interference pattern for $SF = 7$, $s_{I_1} = 83$, $s_{I_2} = 4$, $\tau = 88.4$, $|h_I| = 1$, and $\lambda_{\text{cfo}} = 0.4$.

gateway. The demodulation of $y[n]$ at the receiver yields

$$\mathbf{Y} = \text{DFT}(\mathbf{h}\mathbf{x} \odot \mathbf{x}_{\text{ref}}^*) + \text{DFT}(h_I \mathbf{c}_I \odot \mathbf{x}_I \odot \mathbf{x}_{\text{ref}}^*) + \text{DFT}(\mathbf{z} \odot \mathbf{x}_{\text{ref}}^*). \quad (4.10)$$

We call $\text{DFT}(\mathbf{x}_I \odot \mathbf{x}_{\text{ref}}^*)$ the *transmitted interference pattern* and $\text{DFT}(h_I \mathbf{c}_I \odot \mathbf{x}_I \odot \mathbf{x}_{\text{ref}}^*) = \text{DFT}(\mathbf{y}_I \odot \mathbf{x}_{\text{ref}}^*)$ the *received interference pattern*. The received interference pattern depends on the time-domain interference signal $\mathbf{y}_I = (h_I \mathbf{c}_I \odot \mathbf{x}_I)$, which is in turn a function of the interfering symbols s_{I_1} , s_{I_2} , the channel h_I , the interferer time-offset τ , and the CFO between the interferer and the gateway.

In Fig. 4.1, we present an example showing the difference between the received interference pattern (black plot) and the transmitted interference pattern (green plot), for the particular choice of $SF = 7$, $s_{I_1} = 83$, $s_{I_2} = 4$, $\tau = 88.4$, $|h_I| = 1$, and $\lambda_{\text{cfo}} = 0.4$. We choose $|h_I| = 1$, so that the only difference between the two patterns is due to the CFO. We observe that, in the presence of CFO, the received interference pattern is not just a scaled version of the transmitted pattern, but it has a different shape. The same received interference pattern is also shown in Fig. 4.2, both using a magnitude plot and on the complex plane. In coherent detection of LoRa, the received interference pattern has to be examined on the complex plane, since the decision is *not* performed using the magnitudes of the DFT bins, as was the case for the non-coherent LoRa detection which was explained in the previous chapter.

Distribution of the Decision Metric

Let R_k be the value of the received interference pattern at frequency bin k . For a specific combination of a symbol s and an interference pattern \mathbf{y}_I , the distribution of Y_k is

$$Y_k \sim \begin{cases} \mathcal{CN}(|R_k| \cos \theta_k + j |R_k| \sin \theta_k, 2\sigma^2), & k \in \mathcal{S} / s \\ \mathcal{CN}(N \cos \phi + |R_k| \cos \theta_k + j (N \sin \phi + |R_k| \sin \theta_k), 2\sigma^2), & k = s, \end{cases} \quad (4.11)$$

4.2. Coherent LoRa Symbol Error Rate Under Same-SF Interference

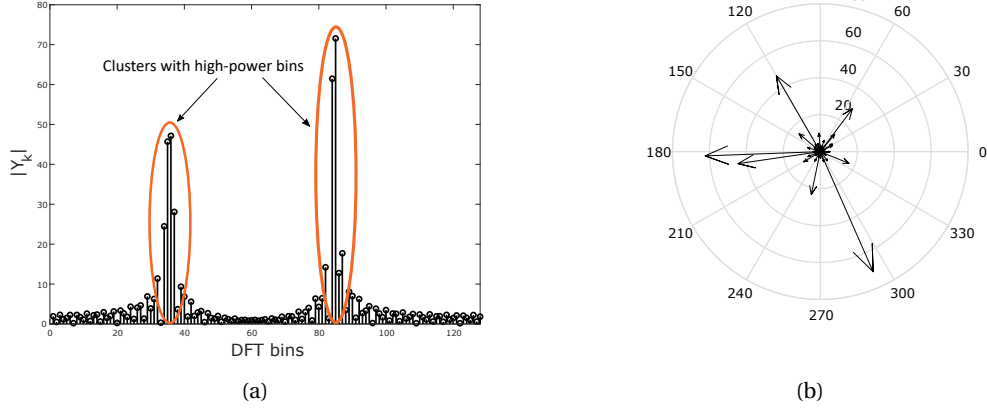


Figure 4.2 – The received interference pattern for SF = 7, $s_{I_1} = 83$, $s_{I_2} = 4$, $\tau = 88.4$, $|h_I| = 1$, and $\lambda_{\text{cfo}} = 0.4$: magnitude (a); and values on the complex plane (b).

where $\theta_k = \theta + \theta_{I_k}$ is the phase shift for bin k introduced by the interference channel and by the CFO. We note that $\theta = \angle h_I$ is fixed for all symbols in a given packet, but changes for different frame transmissions, and is generally uniformly distributed in $[0, 2\pi)$. On the contrary, θ_{I_k} are deterministic, but different for each bin k , and also, they change for different symbols in a packet. Since the received signal \mathbf{U} is rotated by $-\phi$ due to the coherent detection, we define the phase shift between the interferer and the user as $\omega_k = \omega + \theta_{I_k}$ which corresponds to relevant phase shift of the interferer, after the rotation introduced by the coherent receiver. We note that since ϕ is fixed for each transmission, but generally uniformly distributed in $[0, 2\pi)$, it holds that $\omega = \theta - \phi$, is also fixed for each transmission, but generally uniformly distributed in $[0, 2\pi)$.

Finally, the demodulation metric $\Re(U_k)$ in the presence of interference, is distributed as

$$\Re(U_k) \sim \begin{cases} \mathcal{N}(|R_k| \cos \omega_k, \sigma^2), & k \in \mathcal{S}/s \\ \mathcal{N}(N + |R_k| \cos \omega_k, \sigma^2), & k = s. \end{cases} \quad (4.12)$$

Symbol Error Rate Expression

For a given realization of a transmission, conditioning on both the phase of the user's channel ϕ and the phase of the interferer's channel θ , can be replaced by conditioning on $\omega = \theta - \phi$. Thus, in the presence of same-SF interference, the SER for a given symbol s , conditioned on s_{I_1} , s_{I_2} , the relative offset τ , the equivalent offset due to the CFO τ_{cfo} , and the phase difference ω , can be written as

$$P(\hat{s} \neq s | s, s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega) = 1 - \frac{1}{\sigma \sqrt{2\pi}} \int_{y=0}^{+\infty} e^{-\frac{(y-\mu_s)^2}{2\sigma^2}} F_{\Re(U_{\max})}(y) dy, \quad (4.13)$$

where $\mu_s = N + |R_s| \cos \omega_s$, and $F_{\Re(U_{\max})}(y) = \prod_{\substack{k=1 \\ k \neq s}}^N F_{\mathcal{N}}(y; \mu_k, \sigma^2)$, where $F_{\mathcal{N}}(y; \mu_k, \sigma^2)$ denotes the CDF of a Gaussian distribution, and $\mu_k = |R_k| \cos \omega_k$. Following a similar reasoning as in the previous chapter for the non-coherent receiver, the full expression for $P(\hat{s} \neq s)$ in the case of a coherent receiver, conditioned only on the CFO value, is given as follows

$$P(\hat{s} \neq s | \tau_{\text{cfo}}) = 1 - \frac{1}{\sigma(2\pi)^{\frac{3}{2}}} \sum_{s=0}^{N-1} \sum_{s_{I_1}=0}^{N-1} \sum_{s_{I_2}=0}^{N-1} \int_{\tau=0}^N \int_{\omega=0}^{2\pi} \int_{y=0}^{+\infty} e^{-\frac{(y-\mu_s)^2}{2\sigma^2}} \prod_{\substack{k=1 \\ k \neq s}}^N F_{\mathcal{N}}(y; \mu_k, \sigma^2) dy d\omega d\tau. \quad (4.14)$$

4.3 Symbol Error Rate Approximation

Unfortunately, the computational complexity for evaluating the expression in (4.14) is prohibitive and numerical problems may also arise. Therefore, in this section, we derive an approximation for efficiently evaluating (4.14) in the operating SNR regime of LoRa.

4.3.1 Interference Patterns

Let V_k be the value of the received interference pattern at frequency bin k after the rotation due to the coherent receiver, i.e., $V_k = e^{-j\phi} R_k$. We first derive an explicit form for the real-axis projection of V_k , i.e., $\Re(V_k)$, $k \in \mathcal{S}$. Using the definition of the DFT and after some simple algebraic transformations, we obtain

$$\Re(V_k) = |h_I| A_{k,1} \cos \theta_{k,1} + |h_I| A_{k,2} \cos \theta_{k,2}, \quad (4.15)$$

where $A_{k,1}$, $A_{k,2}$, and $\theta_{k,1}$, $\theta_{k,2}$ are

$$A_{k,1} = \frac{\sin\left(\frac{\pi}{N}(s_{I_1} - k - \tau + \tau_{\text{cfo}})\lceil\tau\rceil\right)}{\sin\left(\frac{\pi}{N}(s_{I_1} - k - \tau + \tau_{\text{cfo}})\right)}, \quad (4.16)$$

$$A_{k,2} = \frac{\sin\left(\frac{\pi}{N}(s_{I_2} - k - \tau + \tau_{\text{cfo}})(N - \lceil\tau\rceil)\right)}{\sin\left(\frac{\pi}{N}(s_{I_2} - k - \tau + \tau_{\text{cfo}})\right)}, \quad (4.17)$$

and

$$\theta_{k,1} = \frac{\pi}{N} \left(-\tau^2 - \tau N + s_{I_1}(2\tau - \lceil\tau\rceil + 1) + k(\lceil\tau\rceil - 1) + (\tau - \tau_{\text{cfo}})(\lceil\tau\rceil - 1) - 2m\tau_{\text{cfo}}N - \omega \frac{N}{\pi} \right), \quad (4.18)$$

$$\theta_{k,2} = \frac{\pi}{N} \left(-\tau^2 + \tau N + s_{I_2}(2\tau - \lceil\tau\rceil + 1 - N) + k(\lceil\tau\rceil - 1 + N) + (\tau - \tau_{\text{cfo}})(\lceil\tau\rceil - 1) - 2m\tau_{\text{cfo}}N - \omega \frac{N}{\pi} \right). \quad (4.19)$$

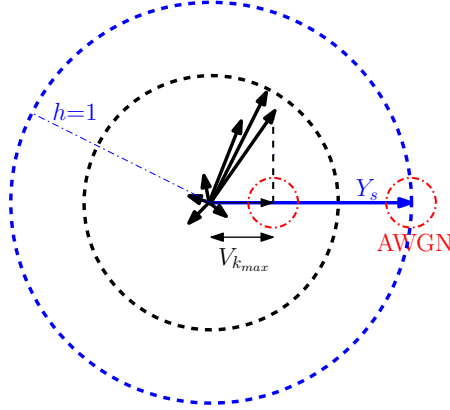


Figure 4.3 – Simple illustration of LoRa bins projected on the real axis.

4.3.2 Symbol Error Rate Approximation

We wish to derive an approximation for $P(\hat{s} \neq s | \tau_{\text{cfo}})$, with the use of Q-functions, that can be evaluated more efficiently than (4.14). First, we assume that the maximum of $\Re(V_k)$, as depicted in Fig. 4.3, dominates the interference-induced SER. Thus, we need to search over all N bins in order to evaluate the maximum projection

$$V_{k_{\max}} = \max_{k \in \mathcal{S}} (\Re(V_k)). \quad (4.20)$$

For a given transmitted symbol s , a particular realization of the user channel, and a particular realization of the interference pattern, i.e., for given s , s_{I_1} , s_{I_2} , τ , τ_{cfo} , and ω , the probability that the interfering bin has higher energy than bin s can be approximated by

$$P(\hat{s} \neq s | s, s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega) \approx Q\left(\frac{N + \Re(V_s) - V_{k_{\max}}}{\sqrt{2\sigma^2}}\right), \quad (4.21)$$

where $\Re(V_s)$ is the projection of the interference which lies on our desired symbol on bin s . Since symbol s can take any of the N possible values, the error rate after removing the conditioning on s can be written as

$$P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega) \approx \frac{1}{N} \sum_{s=0}^{N-1} Q\left(\frac{N + \Re(V_s) - V_{k_{\max}}}{\sqrt{2\sigma^2}}\right). \quad (4.22)$$

We note that, contrary to the previous chapter, the impact of the interference at the user bin s is now included in the approximation. This fact can be also considered as an extension of the SER approximation of Chapter 3, coming with a small increase in computational complexity. Including the impact of the interference at bin s allows the derived approximation to be accurate even in very high SNRs, and even for small SIR values and small SFs. Since s_{I_1} and s_{I_2} can uniformly take any value in the discrete set \mathcal{S} and τ is uniformly distributed in the

continuous set $[0, N)$, we can finally approximate the interference-induced SER as

$$P(\hat{s} \neq s | \tau_{\text{cfo}}) \approx \frac{1}{N^3} \sum_{s_{I_1}=0}^{N-1} \sum_{s_{I_2}=0}^{N-1} \int_0^{2\pi} \int_0^N P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega) d\tau d\omega. \quad (4.23)$$

Complexity Reduction

We can further reduce the complexity for evaluating the interference-induced SER in (4.23) in two ways. One is by reducing the complexity of evaluating $P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega)$ itself, and the second is by reducing the number of times $P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega)$ needs to be evaluated.

Let us first examine how we can reduce the complexity of evaluating $P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega)$. Looking at (4.16) and (4.17), we observe that the DFT bins adjacent to bins $N - \lfloor \tau - \tau_{\text{cfo}} \rfloor + s_{I_1}$ and $N - \lfloor \tau - \tau_{\text{cfo}} \rfloor + s_{I_2}$ have considerably higher energy, compared to the rest of the bins, due to the combination of the fractional part of the misalignment and the fractional part of the CFO. An example of this can be observed in Fig. 4.2a.

Let $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$, with

$$\begin{aligned} \mathcal{D}_1 &= \left\{ \left\lfloor N - \lfloor \tau - \tau_{\text{cfo}} \rfloor + s_{I_1} - \frac{K-1}{2} \right\rfloor_N, \dots, \left\lfloor N - \lfloor \tau - \tau_{\text{cfo}} \rfloor + s_{I_1} + \frac{K-1}{2} \right\rfloor_N \right\}, \\ \mathcal{D}_2 &= \left\{ \left\lfloor N - \lfloor \tau - \tau_{\text{cfo}} \rfloor + s_{I_2} - \frac{K-1}{2} \right\rfloor_N, \dots, \left\lfloor N - \lfloor \tau - \tau_{\text{cfo}} \rfloor + s_{I_2} + \frac{K-1}{2} \right\rfloor_N \right\}, \end{aligned} \quad (4.24)$$

where $K = \frac{|\mathcal{D}|}{2}$. \mathcal{D}_1 and \mathcal{D}_2 are the sets of K bins in the cluster around $N - \lfloor \tau - \tau_{\text{cfo}} \rfloor + s_{I_1}$ and of K bins in the cluster around $N - \lfloor \tau - \tau_{\text{cfo}} \rfloor + s_{I_2}$, respectively, where we denote $[x]_y = x \bmod y$. We note here that the two sets \mathcal{D}_1 and \mathcal{D}_2 can have overlapping bins. We are interested in treating the set of high-power bins $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ separately from the rest of the bins because of two reasons: first, as can be seen in Fig. 4.2b and Fig. 4.3, there is high chance that the maximum projection $V_{k_{\text{max}}}$ will occur due to one of these high-power bins that belong in set \mathcal{D} . Second, the impact of the interference at bin s is significant only when $s \in \mathcal{D}$. For all cases where $s \notin \mathcal{D}$ the impact of the interference at bin s is negligible. Increasing the cardinality of the set \mathcal{D} increases the number of cases in which the bin with the maximum projection will be included in the set \mathcal{D} . We show in the results, that choosing a very small cardinality for this set, e.g., $|\mathcal{D}| = 6$, is sufficient for accurate results, independent of the spreading factor.

We reduce the set of bins we consider from \mathcal{S} with cardinality $|\mathcal{S}| = N$, to \mathcal{D} with much smaller cardinality $|\mathcal{D}| \ll N$ which does not scale with SF. This reduction has two implications on complexity: first, the search in equation (4.20) is only conducted over \mathcal{D} , instead of \mathcal{S} .

Second, and most important, (4.22) can be written as

$$P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega) = \frac{1}{N} \left(\sum_{s \in \mathcal{D}} Q\left(\frac{N + \Re(V_s) - V_{k_{\max}}}{\sqrt{2\sigma^2}}\right) + (N - |\mathcal{D}|) Q\left(\frac{N - V_{k_{\max}}}{\sqrt{2\sigma^2}}\right) \right). \quad (4.25)$$

Using (4.25), $P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega)$ only requires the evaluation of $|\mathcal{D}| + 1$ Q-functions, instead of N Q-functions needed for (4.22), which again is specifically advantageous for large spreading factors (corresponding to large N).

Let us now make some additional observations in order to further reduce the complexity for evaluating (4.23) by reducing the number of times $P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega)$ needs to be evaluated. Every bin of a given interference pattern, for a specific choice of s_{I_1} and s_{I_2} , will be rotated on the circle of Fig. 4.3 due to ω and the interference CFO. Since the bins are already randomly rotated with $\omega \in [0, 2\pi)$, for the evaluation of the average error rate, we can avoid computing the error rate over all possible combinations of $s_{I_1} \in \mathcal{S}$ and $s_{I_2} \in \mathcal{S}$. Instead, we can compute the average error rate over all possible relative differences between the two interfering symbols. We therefore set $s_{I_2} = 0$ and consider only all possible values for $s_I = s_{I_1} - s_{I_2}$, with $s_I \in \mathcal{S}$. Such an assumption reduces the complexity of evaluating the average SER by a factor of N . Note that in Chapter 3, a similar simplification for s_{I_1} and s_{I_2} was mathematically proven to exist for the non-coherent receiver. For the coherent receiver, we therefore provide only the intuition behind the simplification explained above, but without formal proof.

After the simplifications described above, the approximation for the average interference-driven SER is written as follows

$$P(\hat{s} \neq s | \tau_{\text{cfo}}) = \frac{2\pi}{N^2} \sum_{s_I=0}^{N-1} \int_0^{2\pi} \int_0^N P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega) d\tau d\omega. \quad (4.26)$$

The integral over the offset τ can be evaluated numerically by discretizing the interval $[0, N)$ with a step size ϵ . Moreover, the integral over ω can be evaluated by discretizing the interval $[0, 2\pi)$ with a step size ρ . In the results, we show that a discretization choice of $\epsilon = \frac{1}{5}$ and $\rho = \frac{\pi}{2}$ provides a low-complexity, yet accurate, evaluation of (4.26).

In the AWGN-limited regime (i.e., at low SNR), the above approximation eventually becomes inaccurate, since all bins have similar values and no single bin projection dominates the error rate. Similarly to Chapter 3, to fix this issue, let $P^{(N)}(\hat{s} \neq s)$ denote the SER under AWGN given in (4.5), which can be efficiently evaluated using the approximation in (4.6), and let $P^{(I)}(\hat{s} \neq s | \tau_{\text{cfo}})$ be the interference-driven SER from (4.26). Then, a final estimate of the average SER that is accurate also in the low-SNR regime can be approximated by

$$P(\hat{s} \neq s | \tau_{\text{cfo}}) \approx P^{(N)}(\hat{s} \neq s) + (1 - P^{(N)}(\hat{s} \neq s)) P^{(I)}(\hat{s} \neq s | \tau_{\text{cfo}}). \quad (4.27)$$

4.4 Coherent LoRa Frame Error Rate Under Same-SF Interference

In this section we derive an expression to approximate the FER of an uncoded LoRa system with coherent detection. This expression can be used for the LoRa modes that use channel codes of rates $4/5$, and $4/6$, which have error-detection, but no error-correction capabilities, as well as for the uncoded mode. The methodology of extending the analysis to a coded LoRa system will be shown in the next chapter.

We assume perfect frame synchronization for the user, even under the impact of interference. Generally, due to the time offset between frames, only part of the user's frame is affected by interference. However, we assume that the *entire* user frame is under the impact of interference. This way, our derived expression is a straightforward basis for computing FERs for any interference scenario, i.e., any partial collision of two frames. A network simulator can easily derive the corresponding aggregate FER assuming an infinite SIR for the portion of the packet affected only by AWGN, and one SER value derived for the part affected by interference with an SIR according to the interference power of the particular collision. As mentioned in [110], the random relative position of the two frames plays a crucial role for the final FER. The averaging over all possible relative positions of the packets is inherently done by the network simulator itself.

A frame of F LoRa symbols is denoted by the vector \mathbf{s} and the estimated frame at the receiver is denoted by the vector $\hat{\mathbf{s}}$. The FER is therefore $P(\hat{\mathbf{s}} \neq \mathbf{s})$. We note that the expression for the SER derived in Section 4.2 can not be used as is for the evaluation of the FER because it includes an expectation over τ and ω , while all symbols in a frame experience the same τ and the same ω . However, the CFO of the interferer will result in a continuous rotation of the samples in the interfering packet. This effect can be approximately captured by allowing the expectation over ω to be actually included in a symbol level, as this modeling results in an effect that is intuitively similar. The frame error rate $P(\hat{\mathbf{s}} \neq \mathbf{s})$ can thus be approximated as

$$P(\hat{\mathbf{s}} \neq \mathbf{s} | \tau_{\text{cfo}}) \approx \frac{1}{N} \int_0^N \left(1 - (1 - P(\hat{s} \neq s | \tau, \tau_{\text{cfo}}))^F \right) d\tau. \quad (4.28)$$

where $P(\hat{s} \neq s | \tau, \tau_{\text{cfo}})$ can be approximated as

$$P(\hat{s} \neq s | \tau, \tau_{\text{cfo}}) \approx \frac{2\pi}{N} \sum_{s_I=0}^{N-1} \int_0^{2\pi} P(\hat{s} \neq s | s_{I_1}, s_{I_2}, \tau, \tau_{\text{cfo}}, \omega) d\omega. \quad (4.29)$$

Average FER over different overlapping portions of the packets

We also present an expression for the average FER over all possible relative positions of the colliding packets. This expression is useful in cases where the averaging over all possible relative positions of the packets is *not* inherently done by a network simulator. To this end,

as in the previous chapter, we consider that the interfering frame has the same length as the frame of interest. This assumption is taken only for clarity of the presentation and does not prevent the generalization of the results to any interfering frame length. The difference to (4.28) is that, due to the time offset between the frames, only part of the frame of interest is affected by interference.

As in Chapter 3, let F_I , where $F_I \in \{1, \dots, F\}$, denote the number of symbols in the frame that are affected by the interfering frame. The value of F_I depends on the relative position of the two frames. For the average FER, we consider the expectation over all the possible relative positions of the two frames. As explained in Chapter 3, there is a symbol in the frame that is partially affected by the interfering frame. Similarly to Chapter 3, we consider the partially-affected symbol as fully-affected by interference, thus including it in F_I . The frame error rate $P(\hat{\mathbf{s}} \neq \mathbf{s})$ is now given as

$$P(\hat{\mathbf{s}} \neq \mathbf{s} | F_I, \tau_{\text{cfo}}) \approx \frac{1}{N} \int_0^N \left(1 - (1 - P(\hat{s} \neq s | \tau, \tau_{\text{cfo}}))^{F_I} (1 - P^{(N)}(\hat{s} \neq s))^{F-F_I} \right) d\tau. \quad (4.30)$$

where $P(\hat{s} \neq s | \tau, \tau_{\text{cfo}})$ is given by (4.29), and $P^{(N)}(\hat{s} \neq s)$ is the SER under AWGN given in (4.5) (which can be evaluated efficiently using the approximation in (4.6)).

Finally, we take the expectation over all possible values of F_I and we obtain the final expression for the average FER over different overlapping portions of the packets

$$P_{\text{av}}(\hat{\mathbf{s}} \neq \mathbf{s} | \tau_{\text{cfo}}) \approx \frac{1}{F} \sum_{F_I=1}^F P(\hat{\mathbf{s}} \neq \mathbf{s} | F_I, \tau_{\text{cfo}}). \quad (4.31)$$

4.5 Results

In this section, we provide numerical results for the SER and the FER of a coherent LoRa receiver with same-SF interference. We compare the performance against the non-coherent receiver, under the same interference conditions. Moreover, we compare the derived approximations against Monte Carlo simulations to show their accuracy. We note that for the Monte Carlo simulations, as in Chapter 3, the interferer time-offset τ is simulated by oversampling (2.3) to create s_{I_1} and s_{I_2} , concatenating the oversampled symbols, applying the appropriate offset, and downsampling to obtain $N = 2^{\text{SF}}$ samples for \mathbf{x}_I .

In Fig. 4.4, we show the results of a Monte Carlo simulation for the FER of a LoRa receiver for $\text{SF} = 7$, for a packet of length $F = 20$ LoRa symbols, under the effect of same-SF interference with an SIR of 0 dB, for three values of $\lambda_{\text{cfo}} \in \{0, 0.3, 0.5\}$, and for both a coherent and a non-coherent receiver. We observe that for the coherent receiver, under the simulation method described in the previous paragraph, the CFO value of the interferer does not have a significant impact on the error rate. We note that the results that we obtained for higher SIR values show an even more negligible impact of the CFO of the interferer on the error rate. For this reason, we choose a value of $\lambda_{\text{cfo}} = 0$ for the subsequent results for the coherent receiver. We note that,

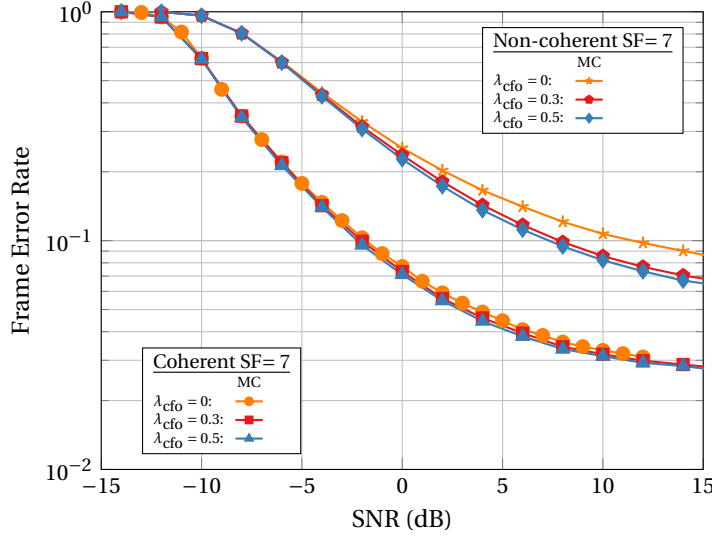


Figure 4.4 – Frame error rate of the coherent and non-coherent receiver for three different values of CFO, and a packet of length $F = 20$ LoRa symbols under AWGN and same-SF interference for $SF = 7$ and $P_I = 0$ dB.

in the results that we obtained for higher SIR values, the impact of the CFO of the interferer becomes negligible also for the non-coherent receiver, even for $\lambda_{cfo} = 0.5$. For this reason, we choose a value of $\lambda_{cfo} = 0$ for the subsequent results also for the non-coherent receiver. The inclusion of the CFO of the interferer in the model, and the discussion on its impact on the error rate, can be considered as a more realistic extension of the model we presented in Chapter 3, where the CFO of the interferer was not included in the model. However, the impact of the CFO of the interferer in low SIR values need to be studied more. In particular, it needs to be checked if the simplified version of the expression for $x_I[n]$ in (3.11) becomes relevant.

In Fig. 4.5, we show the results of a Monte Carlo simulation for the SER of a LoRa user for $SF \in \{7, 9, 11\}$ and $SIR = 3$ dB, using the coherent receiver described in this work, as well as the non-coherent receiver described in Chapter 3. The corresponding approximations for the coherent and non-coherent receivers are plotted as well. We observe that the coherent receiver has a significant performance gain of up to 2.5 dB for some SNR values compared to the non-coherent receiver. In [47, 92, 125], for the AWGN-only case, a difference of around 0.7 dB was shown between the coherent and the non-coherent receiver. It is, therefore, very interesting to see that using a coherent receiver is much more beneficial under same-SF interference than in the AWGN-only case. Finally, we observe that the low-complexity SER approximation in (4.27) with $|\mathcal{D}| = 6$ and discretization steps $\rho = \frac{\pi}{2}$ and $\epsilon = \frac{1}{5}$ is very accurate.

In Fig. 4.6, we show the results of a Monte Carlo simulation for the FER of both a coherent and a non-coherent LoRa receiver with $SF \in \{7, 9, 11\}$, as well as the corresponding approximation described in (4.28). The frame length is $F = 20$ LoRa symbols and $SIR = 3$ dB. We observe the same performance difference between the coherent and non-coherent receivers as in the

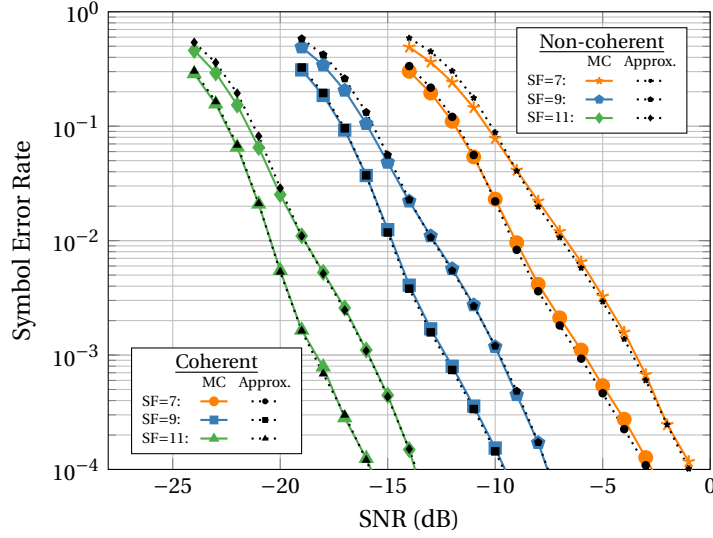


Figure 4.5 – Symbol error rate of the coherent and non-coherent LoRa receiver under AWGN and same-SF interference for $SF \in \{7, 9, 11\}$ and $P_I = -3$ dB. The approximations for the coherent and the non-coherent case are shown with black dotted lines.

SER curves. Furthermore, we see that the low-complexity approximation for the FER under same-SF interference, described in Section 4.4, is very accurate.

In Fig. 4.7, we show the required SNR for $SF = 7$ with a target FER performance of 10^{-1} , for different SIR levels. We show results for both the coherent and the non-coherent receiver. As expected, for both receivers, there is an increase in the required SNR to obtain the same FER, as the interference power increases. For the chosen target FER of 10^{-1} we can observe that the coherent receiver requires a much lower increase in the required SNR compared to the non-coherent receiver. We can clearly observe that the 0.7 dB performance difference between the coherent and non-coherent receiver reported in [47, 92, 125] for the AWGN case (i.e., for $SIR \rightarrow \infty$) is only the smallest possible performance difference. In fact, the coherent receiver is of increasingly greater importance for increasing levels of interference.

Phase estimation errors

A receiver needs to estimate the phase of the received symbols, in order to be able to be coherent. Although the initial estimation of the phase is easy to do in a LoRa packet, thanks to the preamble, the continuous phase tracking, needed due to the phase drift etc, is not going to be perfect. The phase estimation error due to the imperfect tracking can be modeled as a Gaussian process with variance σ_{tr}^2 , which depends on the phase-drift levels and the quality of the tracking algorithm.

In Fig. 4.8, we show the performance degradation of a coherent LoRa receiver for three different estimation error variances, $\sigma^2 \in \{0.2, 0.3, 0.4\}$, as well as for a perfect phase estimation. The

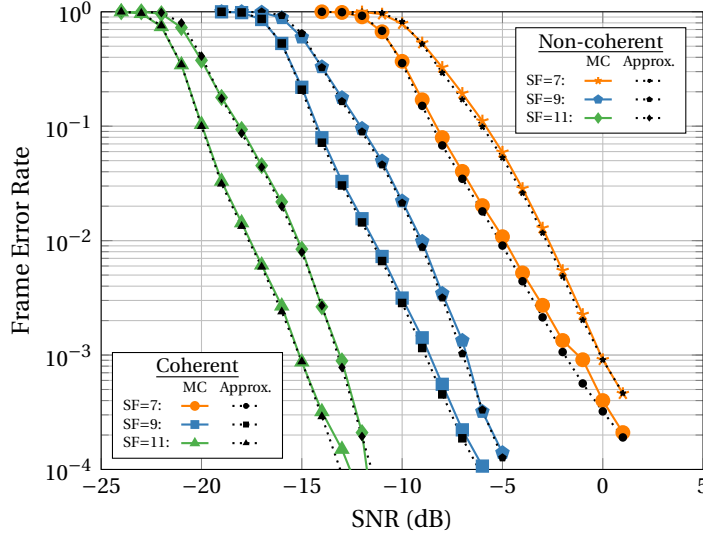


Figure 4.6 – Frame error rate of the coherent and non-coherent receiver for a packet of length $F = 20$ LoRa symbols under AWGN and same-SF interference for $SF \in \{7, 9, 11\}$ and $P_I = -3$ dB. The approximations for all cases are shown with black dotted lines.

frame length is $F = 20$ LoRa symbols, the spreading factor is $SF = 7$, and $SIR = 3$ dB. The results are compared to the non-coherent receiver, shown in a thick transparent line. We observe that the tracking phase-estimation error plays a significant role in the performance of a coherent receiver. Under small estimation errors, the coherent receiver can still be competitive to the non-coherent receiver. Therefore, for a coherent receiver to be worth applying compared to a non-coherent receiver, a careful design of the phase-tracking algorithm is needed.

4.6 Discussion and Summary

In this chapter we modeled and investigated the performance of coherent LoRa receivers under same-SF interference to validate our intuition that the adoption of coherent LoRa receivers may prove fruitful in interference scenarios. To this end, we derived the full error rate expression as well as accurate low-complexity approximations to evaluate the expression. We have seen that the benefit of employing a coherent LoRa receiver increases for higher levels of interference and can reach values close to 10 dB for $SIR = 0$ dB, making coherent LoRa receivers particularly attractive for massive LoRa networks with large probabilities of interfering packets.

In Chapters 3 and 4, we have established models to analyze the performance of both existing and modified LoRa receivers, when no error-correction coding is used. However, LoRa transceivers include options of coded transmission as well, the performance of which cannot immediately be covered by the models presented in neither of these two chapters. Therefore, in the following chapter we will establish the analysis of coded LoRa receivers.

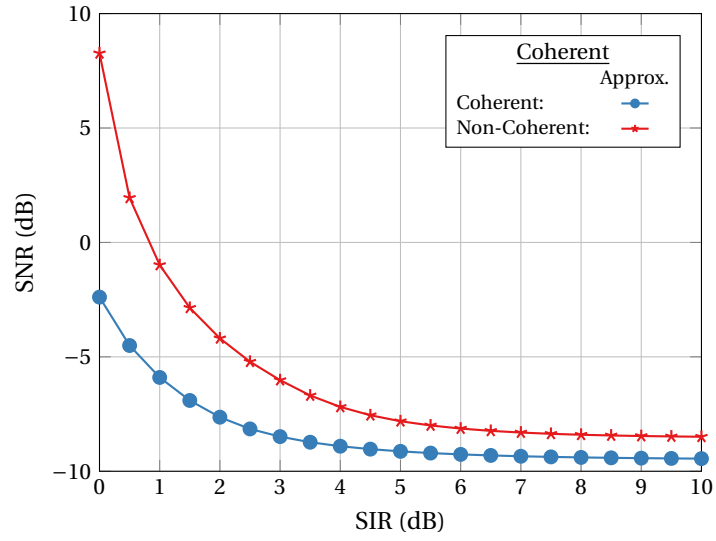


Figure 4.7 – Required SNR for a target frame error rate of 10^{-1} as a function of the SIR for SF = 7 for coherent and non-coherent receiver.

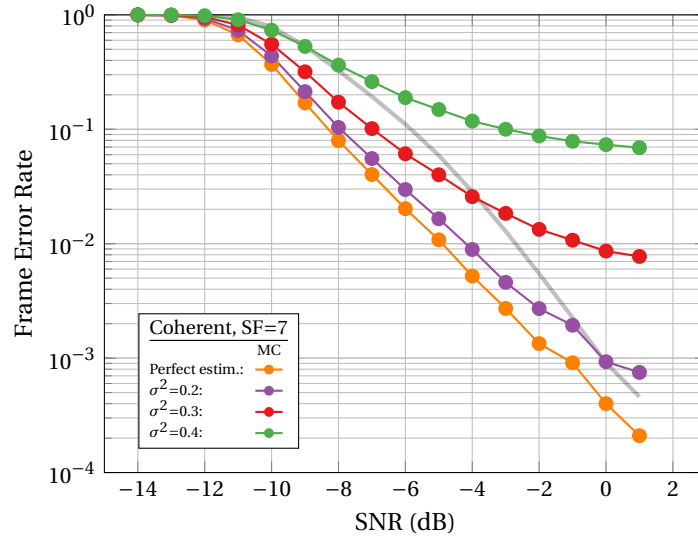


Figure 4.8 – Frame error rate of the coherent receiver under AWGN and same-SF interference with perfect estimation, and for estimation error $\sigma^2 \in \{0.2, 0.3, 0.4\}$. The packet length is $F = 20$ LoRa symbols, SF=7 and $P_I = -3$ dB. The performance of the non-coherent receiver is shown in thick transparent line.

5 Coded LoRa Error Rate

In the previous two chapters we have presented a detailed probabilistic error-rate model for both the standard non-coherent and the proposed coherent LoRa receivers. However, the FER expressions in Chapters 3 and 4 are only valid for the LoRa modes that do not use error-correction coding, since the expressions do not take into account the channel coding, interleaving, and Gray mapping present in a LoRa transceiver physical layer chain, which have a great impact on the error rate. Furthermore, hardware impairments are very important in the context of low-power radios. Chapter 3 does not model any hardware impairment, while Chapter 4 acknowledges that CFO exists for both the user of interest and the interferer, but assumes that it can be perfectly corrected for the user of interest, and only accounts for it for the interfering user. As we will show, the combination of channel coding, interleaving, and Gray mapping improves the performance in case a residual carrier frequency offset is present after its estimation and partial compensation at the LoRa receiver.

Therefore, in this chapter, we extend our error-rate model to the coded LoRa system, and we derive an approximation for the coded FER of a LoRa system in AWGN. Moreover, we derive an approximation for the coded FER of a LoRa system with residual carrier frequency offset, which is an important impairment that can significantly affect the performance of LoRa.

In particular, we begin with an analysis of the channel coding and interleaving mechanisms of the LoRa PHY in order to first derive the codeword error rate (CWER) for LoRa. We then use the CWER to derive two low-complexity approximations for the coded FER of a LoRa system under AWGN. Moreover, we derive an approximation for the coded FER of a LoRa system with residual carrier frequency offset (CFO). Finally, we corroborate the accuracy of our approximations through Monte Carlo simulations.

5.1 Coded LoRa Frame Error Rate Under AWGN

The payload comprises multiple interleaved Hamming codewords. In this section we first derive the CWER using existing results for the SER. We then use the CWER to derive the FER.

5.1.1 Codeword Error Rate

The uncoded symbol error probability P_s is defined as $P_s \triangleq P(\hat{s} \neq s)$. An approximation that can be used to efficiently evaluate the aforementioned probability under AWGN was shown in (3.7), and is repeated here for convenience

$$P(\hat{s} \neq s) \approx Q \left(\frac{\sqrt{\text{SNR}} - \left((H_{N-1})^2 - \frac{\pi^2}{12} \right)^{1/4}}{\sqrt{H_{N-1} - \sqrt{(H_{N-1})^2 - \frac{\pi^2}{12} + 0.5}}} \right). \quad (5.1)$$

The term $(H_{N-1})^2$ in the above equation stems from the fact that, in the estimation of \hat{s} , $N-1$ symbols can be mistakenly chosen instead of s . As mentioned in Chapter 3, if a symbol error happens due to AWGN only, on average half of the SF symbol bits will be erroneous, independently of whether Gray mapping is used or not. Thus, for the uncoded bit error probability P_b we have

$$P_b = 0.5 \cdot P_s. \quad (5.2)$$

The codeword error probability P_{cw} is defined as the probability that the Hamming decoder output decision $\hat{\mathbf{c}}$ does not correspond to the Hamming encoder output codeword \mathbf{c} , i.e., $P_{\text{cw}} \triangleq P(\hat{\mathbf{c}} \neq \mathbf{c})$. The Hamming distance between two vectors \mathbf{w}_1 and \mathbf{w}_2 , denoted by $d(\mathbf{w}_1, \mathbf{w}_2)$, is defined as the number of locations in which \mathbf{w}_1 and \mathbf{w}_2 differ. Thus, if \mathbf{v} is the Hamming decoder input vector, then P_{cw} for the (4, 7) and (4, 8) Hamming codes assuming that the decoder declares a failure when more than one error is detected can be equivalently defined as the probability of the event that \mathbf{v} has *at least* two erroneous bits as

$$P_{\text{cw}} = P(\{d(\mathbf{v}, \mathbf{c}) \geq 2\}). \quad (5.3)$$

Codewords in LoRa are interleaved (and deinterleaved) in blocks of SF codewords. Fig. 5.1 illustrates the LoRa deinterleaving process. Each row of the matrix before deinterleaving corresponds to the SF bits of a demodulated LoRa symbol $\hat{\mathbf{s}}_i$, $i \in \{1, \dots, n_c\}$, while each row of the matrix after deinterleaving corresponds to the n_c bits of \mathbf{v}_j , $j \in \{1, \dots, \text{SF}\}$, at the Hamming decoder input.

When a symbol error happens, e.g., $\hat{\mathbf{s}}_1 \neq \mathbf{s}_1$, on average half of the bits of row 1 of the matrix before deinterleaving will be erroneous. These bit errors are generally correlated because they come from the same symbol error, but the deinterleaver effectively removes this correlation by distributing each bit error of $\hat{\mathbf{s}}_1$ to distinct codewords, as shown in Fig. 5.1. If another symbol error occurs, e.g., $\hat{\mathbf{s}}_2 \neq \mathbf{s}_2$, the bit errors due to $\hat{\mathbf{s}}_2$ are again generally correlated, but they are independent from the bit errors of $\hat{\mathbf{s}}_1$. As a result, the bit errors after the deinterleaver in every input codeword of the Hamming decoder \mathbf{v}_j are independent and identically distributed (iid) with probability of bit error P_b given by (5.2). Thus, the CWER in (5.3) becomes

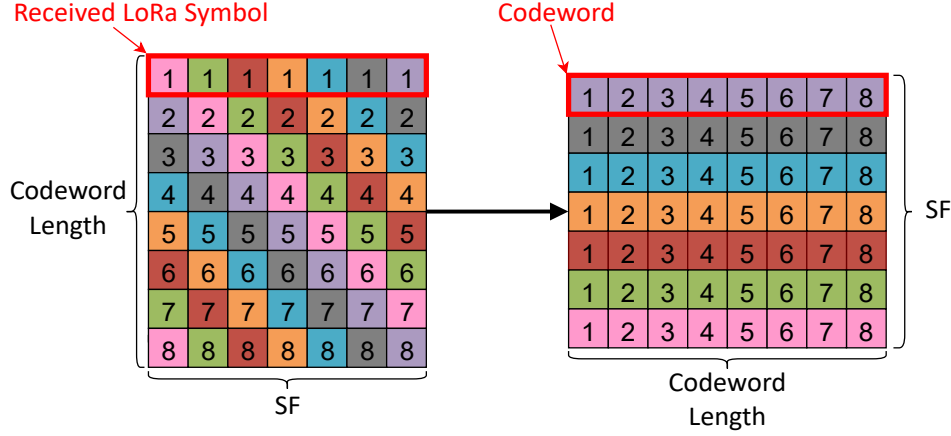


Figure 5.1 – LoRa deinterleaving for SF = 7 and a (4, 8) Hamming code.

$$P_{cw} = 1 - \left((1 - P_b)^{n_c} + \binom{n_c}{1} P_b (1 - P_b)^{n_c - 1} \right). \quad (5.4)$$

5.1.2 Frame Error Rate

We denote the number of payload symbols in a LoRa packet by N_{pl} . By construction, N_{pl} is an integer multiple of the codeword length n_c . The number of codewords in the payload is $N_{cw} = \frac{N_{pl} SF}{n_c}$, where each codeword has a CWER P_{cw} given by (5.4). Let $\mathcal{C}_{bl} \in \{0, 1\}^{SF \times n_c}$ be a matrix containing a block of SF transmitted codewords \mathbf{c}_i , $i \in \{1, \dots, SF\}$, that are input to one block of the interleaver. Let also $\mathcal{V}_{bl} \in \{0, 1\}^{SF \times n_c}$ be a matrix containing a block of SF codewords that are the output of one deinterleaving block (and also the input to the Hamming decoder). Finally, let $\hat{\mathcal{C}}_{bl} \in \{0, 1\}^{SF \times n_c}$ be a matrix containing a block of SF estimated codewords $\hat{\mathbf{c}}_i$, $i \in \{1, \dots, SF\}$, after the Hamming decoder. The probability $P(\hat{\mathcal{C}}_{bl} = \mathcal{C}_{bl})$ corresponds to the probability that all the SF decoded codewords of one deinterleaver block are correctly decoded, which is given by

$$P(\hat{\mathcal{C}}_{bl} = \mathcal{C}_{bl}) = \prod_{i=1}^{SF} P(\hat{\mathbf{c}}_i = \mathbf{c}_i | \hat{\mathbf{c}}_1 = \mathbf{c}_1, \dots, \hat{\mathbf{c}}_{i-1} = \mathbf{c}_{i-1}). \quad (5.5)$$

Let $P_{cw}^{(i)}$ denote the conditional codeword error probability given that all previous $i - 1$ codewords in the block were decoded correctly, i.e.,

$$P_{cw}^{(i)} = P(\hat{\mathbf{c}}_i \neq \mathbf{c}_i | \hat{\mathbf{c}}_1 = \mathbf{c}_1, \dots, \hat{\mathbf{c}}_{i-1} = \mathbf{c}_{i-1}). \quad (5.6)$$

Using (5.6), we can re-write (5.5) as

$$P(\hat{\mathcal{C}}_{bl} = \mathcal{C}_{bl}) = \prod_{i=1}^{SF} (1 - P_{cw}^{(i)}). \quad (5.7)$$

Let $\mathcal{C}_{pl} \in \{0, 1\}^{N_{cw} \times n_c}$ be a matrix containing all the transmitted codewords in the payload and

$\hat{\mathcal{C}}_{\text{pl}}$ be a matrix containing all the decoded payload codewords. Since the payload contains N_{pl}/n_c deinterleaving blocks which are independent from each other, using (5.7) we can obtain the overall FER as

$$P(\hat{\mathcal{C}}_{\text{pl}} \neq \mathcal{C}_{\text{pl}}) = 1 - \left(\prod_{i=1}^{\text{SF}} (1 - P_{\text{cw}}^{(i)}) \right)^{\frac{N_{\text{pl}}}{n_c}}. \quad (5.8)$$

As explained in Section 5.1.1, a correctly decoded codeword $\hat{\mathbf{c}}_i$ may come from a decoder input \mathbf{v}_i that has up to one bit error. This single non-catastrophic bit-error can be in any of the n_c positions of \mathbf{v}_i . As a result, the computation of the product in (5.8), becomes a cumbersome task, since the number of all possible non-catastrophic bit-error patterns inside a deinterleaver block is a large combinatorial quantity. For this reason, in the remainder of this section we derive two low-complexity approximations for (5.8).

Approximation 1: A simple approximation of (5.8) can be obtained by ignoring the conditioning in (5.6) to obtain

$$P(\hat{\mathcal{C}}_{\text{pl}} \neq \mathcal{C}_{\text{pl}}) \approx 1 - (1 - P_{\text{cw}})^{\frac{N_{\text{pl}} \text{SF}}{n_c}}. \quad (5.9)$$

Since dependence of decoding errors between codewords of the same block intuitively decreases the probability that the i -th codeword will be erroneous, given that all previous $i - 1$ codewords were decoded correctly, we expect that Approximation 1, which assumes independence, overestimates the FER.

Approximation 2: Let us now define a more elaborate approximation for (5.8), which has slightly higher computational complexity but gives more accurate results. To this end, in the conditional codeword error probability $P_{\text{cw}}^{(i)}$ of (5.6), we approximate the event that all previous $i - 1$ codewords in the block were decoded correctly (i.e., $\{\hat{\mathbf{c}}_1 = \mathbf{c}_1, \dots, \hat{\mathbf{c}}_{i-1} = \mathbf{c}_{i-1}\}$) with the event that all previous $i - 1$ codewords in the block did not contain any error at all (i.e., $\{\hat{\mathbf{v}}_1 = \mathbf{c}_1, \dots, \hat{\mathbf{v}}_{i-1} = \mathbf{c}_{i-1}\}$). Therefore, for all the previously decoded $i - 1$ codewords in the block, we ignore single-bit errors. The conditional codeword error rate in (5.6) can then be written as

$$P_{\text{cw}}^{(i)} \approx P(\hat{\mathbf{c}}_i \neq \mathbf{c}_i | \mathbf{v}_1 = \mathbf{c}_1, \dots, \mathbf{v}_{i-1} = \mathbf{c}_{i-1}). \quad (5.10)$$

Each conditional codeword error probability $P_{\text{cw}}^{(i)}$ in (5.10) can be interpreted as the codeword error probability of a decoder that has the additional side-information that $(i - 1)$ out of the n_c bit positions in \mathbf{v}_i are guaranteed to be error-free. As such, we expect this approximation to result in slightly lower FER values compared to (5.8). However, we need to find an expression to calculate the conditional probability of (5.10).

Let $P_s^{(i)}$ be the conditional symbol error probability

$$P_s^{(i)} = P(\hat{s} \neq s | \hat{b}_1 = b_1, \dots, \hat{b}_{i-1} = b_{i-1}), \quad (5.11)$$

where $\hat{b}_i = b_i$ denotes that the i -th bit of symbol s is estimated correctly. Then the conditional bit error probability is $P_b^{(i)} = 0.5P_s^{(i)}$, since errors are again due to AWGN. The additional information included in the conditional symbol error probability of (5.11), compared to the unconditional one, is that a potential demodulation error can only be in $2^{\text{SF}-i-1} - 1$ DFT bins instead of the $N-1$ DFT bins with possible errors in (3.7). As such, the conditional symbol error probability $P_s^{(i)}$ can be calculated by adapting the harmonic number in (3.7) as follows

$$P_s^{(i)} \approx Q \left(\frac{\sqrt{\text{SNR}} - \left((H_{\frac{N}{2^{i-1}}-1})^2 - \frac{\pi^2}{12} \right)^{1/4}}{\sqrt{H_{\frac{N}{2^{i-1}}-1} - \sqrt{(H_{\frac{N}{2^{i-1}}-1})^2 - \frac{\pi^2}{12}} + 0.5}}} \right). \quad (5.12)$$

Therefore, each $P_{\text{cw}}^{(i)}$ in (5.8) can be calculated as

$$P_{\text{cw}}^{(i)} \approx 1 - \left(\left(1 - P_b^{(i)} \right)^{n_c} + \binom{n_c}{1} P_b^{(i)} \left(1 - P_b^{(i)} \right)^{n_c-1} \right). \quad (5.13)$$

Finally, our second approximation for the FER can be obtained by replacing (5.13) in (5.8), as

$$P(\hat{\mathcal{C}}_{\text{pl}} \neq \mathcal{C}_{\text{pl}}) = 1 - \left(\prod_{i=1}^{\text{SF}} \left(1 - P_b^{(i)} \right)^{n_c} + \binom{n_c}{1} P_b^{(i)} \left(1 - P_b^{(i)} \right)^{n_c-1} \right)^{\frac{N_{\text{pl}}}{n_c}}. \quad (5.14)$$

5.1.3 Results

For consistency of the results presented throughout the chapter, we choose again a common payload length of $N_{\text{pl}} = 32$ LoRa symbols and a common codeword length $n_c = 8$, but we note that the derived approximations are also accurate for $n_c = 7$, and for any frame length.

In Fig. 5.2, we compare the results of a Monte Carlo simulation for the LoRa FER for all possible $\text{SF} \in \{7, \dots, 12\}$ with the FER results obtained by using Approximation 1 and Approximation 2 under AWGN. The payload length is chosen to be $N_{\text{pl}} = 32$ LoRa symbols. We observe that both approximations are quite accurate, especially at low FER values, but Approximation 2 is visibly better than Approximation 1. However, Approximation 1 only requires a single Q-function evaluation per SNR point, while Approximation 2 requires SF Q-function evaluations per SNR point. Moreover, we observe that Approximation 1 slightly overestimates and Approximation 2 slightly underestimates the error rate, as expected from the discussion in Section 5.1.

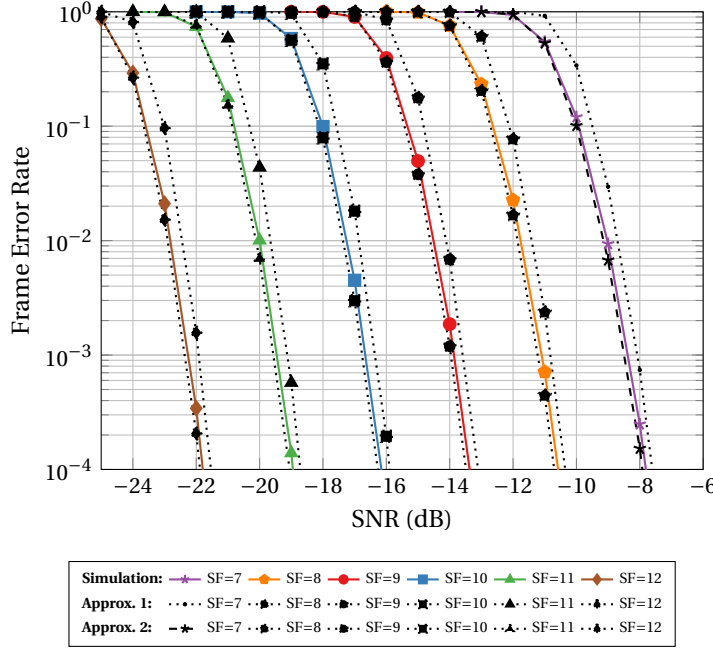


Figure 5.2 – Frame error rate of the LoRa modulation under AWGN for all supported spreading factors $SF \in \{7, \dots, 12\}$.

5.2 LoRa Coded Frame Error Rate under Carrier Frequency Offset

In this section, we first derive an expression for the coded LoRa BER under AWGN with carrier frequency offset (CFO). We then use this initial result to explain how the coded FER can be calculated in the presence of both AWGN and CFO.

5.2.1 Distribution of the Decision Metric

In order to derive the coded FER under AWGN in Section 5.1, we use the symbol and bit error probability expressions from Chapter 3. In order to derive the coded FER under CFO in this section, we first need an expression for the symbol and bit error probabilities of LoRa under residual CFO. To this end, we introduce the necessary modeling that allows us to find the uncoded symbol error probability as well as the coded bit error probability in Section 5.2.2.

Let f_{c_1} and f_{c_2} be the carrier frequencies that are used during up- and down-conversion, respectively. The carrier frequency offset is the difference $\Delta f_c = f_{c_1} - f_{c_2}$. Thus, in the presence of CFO, the signal model with unit channel gain $|h| = 1$ becomes

$$y[n] = c[n]x[n] + z[n], \quad n \in \mathcal{S}, \quad (5.15)$$

where $x[n]$ is the signal of interest, $c[n] = e^{j2\pi(n+(m-1)N)\frac{\Delta f_c}{f_s}}$ is the CFO term affecting the m -th

5.2. LoRa Coded Frame Error Rate under Carrier Frequency Offset

$$P(\hat{s} \neq s | s, \lambda) = P\left(\bigcup_{i \in \mathcal{D}} \{|Y'_i| > |Y'_s|\} | s, \lambda\right) + P\left(\bigcup_{j \in \mathcal{R}} \{|Y'_j| > |Y'_s|\} | s, \lambda\right) - P\left(\left\{\bigcup_{i \in \mathcal{D}} \{|Y'_i| > |Y'_s|\} \cap \bigcup_{j \in \mathcal{R}} \{|Y'_j| > |Y'_s|\}\right\} | s, \lambda\right) \quad (5.17)$$

symbol in the packet, and $z[n] \sim \mathcal{N}(0, \sigma^2)$ is AWGN. The demodulation of $y[n]$ yields

$$\mathbf{Y} = \text{DFT}(\mathbf{c} \odot \mathbf{x} \odot \mathbf{x}_{\text{ref}}^*) + \text{DFT}(\mathbf{z} \odot \mathbf{x}_{\text{ref}}^*), \quad (5.16)$$

where $\mathbf{c} = [c[0] \ \dots \ c[N-1]]$. We call $\text{DFT}(\mathbf{c} \odot \mathbf{x} \odot \mathbf{x}_{\text{ref}}^*)$ the *CFO pattern*. The CFO pattern depends on the symbol value s and on the CFO value Δf_c . As explained in [48, 37], and in the previous chapter, the CFO results in a frequency shift which can be modeled with an integer part L and a fractional part λ , where $L + \lambda = \frac{\Delta f_c N}{f_s}$. The fractional part λ is the fraction of the CFO relative to the distance between two adjacent DFT bins. We note that in the previous chapter, we considered that the CFO affected only the interfering signal, while we considered the demodulated user to have perfectly estimated and compensated the CFO. In this chapter, we consider *residual CFO* for the demodulated user after the estimation and correction procedures. We limit our study to the case where $L = 0$ and $-0.5 \leq \lambda \leq 0.5$, since, as will be shown in the next chapter, the integer part of the CFO can be estimated and compensated relatively easily.

Let R_k denote the value of the CFO pattern at frequency bin k , i.e., $R_k = \text{DFT}(\mathbf{c} \odot \mathbf{x} \odot \mathbf{x}_{\text{ref}}^*)[k]$, $k \in \mathcal{S}$. Using the definition of the Fourier transform and arguments that are similar to the analysis in Chapter 3 and Chapter 4, we have

$$|R_k| = \frac{\sin\left(\frac{\pi}{N}(s-k+\lambda)N\right)}{\sin\left(\frac{\pi}{N}(s-k+\lambda)\right)}. \quad (5.18)$$

From (5.18), it can be seen that, in the absence of CFO (i.e., $\lambda = 0$) we have $|R_k| = 0$, for any $k \in \mathcal{S}/s$, and $|R_s| = N$. On the other hand, if $\lambda \neq 0$, a part of the energy of bin s is spread across all $N = 2^{\text{SF}}$ bins. For a specific λ and symbol s , combining the CFO pattern with the AWGN leads to the demodulation metric $|Y_k|$ used in (2.11) with

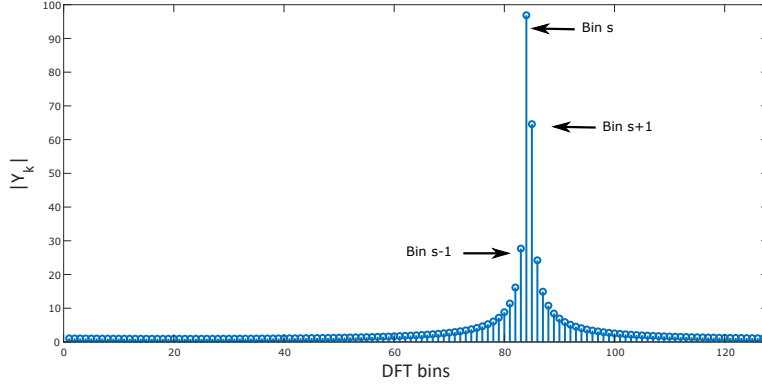
$$|Y_k| \sim f_{\text{Ri}}(y; |R_k|, \sigma^2), \quad k \in \mathcal{S}. \quad (5.19)$$

5.2.2 Uncoded Symbol Error Rate

A symbol error occurs if and only if any of the $|Y_k|$ values for $k \in \mathcal{S}/s$ exceeds the value of $|Y_s|$,

$$P(\hat{s} \neq s | s, \lambda) = P\left(\bigcup_{k \in \mathcal{S}/s} \{|Y_k| > |Y_s|\} | s, \lambda\right). \quad (5.20)$$

As can be deduced from (5.18), the DFT bins adjacent to bin s have considerably higher energy due to the fractional offset λ compared to the remaining bins, as can be seen in Fig. 5.3.


 Figure 5.3 – The magnitude of the DFT bins for $SF = 7$, $s = 83$, and $\lambda_{\text{cfo}} = 0.4$.

As a result, when CFO and AWGN are present, the adjacent bins (especially one of the two, depending on the sign and the value of λ_{cfo}) are generally more prone to causing symbol errors. However, in the case where the SNR is very low, all the bins in the set \mathcal{S}/s still have comparable symbol error probabilities. In order to model the symbol error rate for a large range of SNRs, we rewrite the symbol error probability in (5.20) in the equivalent form shown in (5.17), which separates the set $\mathcal{D} = \{s-1, s+1\}$ of the two adjacent bins, from the rest of the bins in the set $\mathcal{R} = \mathcal{S} \setminus \{s-1, s, s+1\}$. The third term of (5.17) is typically small and can be ignored.

Due to the Gray mapping, a symbol error that mistakes s for one of the symbols in \mathcal{D} only causes one bit error. On the other hand, when a symbol error happens that mistakes s for one of the symbols in \mathcal{R} , on average half of the bits are wrong. Therefore, the bit error probability can be written as

$$P(\hat{b} \neq b | s, \lambda) \approx \frac{1}{SF} P(|Y_{\max, \mathcal{D}}| > |Y_s| | s, \lambda) + \frac{1}{2} P(|Y_{\max, \mathcal{R}}| > |Y_s| | s, \lambda), \quad (5.21)$$

where $|Y_{\max, \mathcal{D}}| = \max_{i \in \mathcal{D}} |Y_i|$ and $|Y_{\max, \mathcal{R}}| = \max_{j \in \mathcal{R}} |Y_j|$. The first probability term in (5.21) is given by

$$P(|Y_{\max, \mathcal{D}}| > |Y_s| | s, \lambda) = 1 - \int_{y=0}^{+\infty} f_{\text{Ri}}(y; \nu_s, \sigma^2) F_{|Y_{\max, \mathcal{D}}|}(y) dy, \quad (5.22)$$

where $\nu_s = |R_s|$ is the location parameter for bin s . The second term in (5.21) can be obtained by replacing \mathcal{D} with \mathcal{R} in (5.22). As explained in Chapter 3, the CDF of the n -th order statistic (i.e., the CDF of the maximum) is $F_n(x) = P(X_1 < x)P(X_2 < x) \dots P(X_n < x)$. Due to the conditioning on λ , each $|Y_m|$ for $m \in \mathcal{R}$ is independent from any other $|Y_n|$ for $n \in \mathcal{D}/m$. Thus, the CDF of the maximum bin for the set \mathcal{D} is $F_{|Y_{\max, \mathcal{D}}|}(y) = \prod_{j \in \mathcal{D}} F_{\text{Ri}}(y; \nu_j, \sigma^2)$, where $\nu_j = |R_j|$. The CDF

$F_{|Y_{\max, \mathcal{S}}|}$ can be obtained accordingly. Finally, s is uniformly distributed in \mathcal{S} , so that

$$P_{b|\lambda} = P(\hat{b} \neq b|\lambda) = \frac{1}{N} \sum_{s=0}^{N-1} P(\hat{b} \neq b|s, \lambda). \quad (5.23)$$

5.2.3 Complexity Reduction

We can see from (5.18) that all CFO patterns for all values of $s \in \mathcal{S}$ contain the same set of frequency bin magnitudes $|R_k|$, $k \in \mathcal{S}$, but are circularly shifted. This circular shift does not change the distribution of $|Y_{\max}|$, thus the probability of $|Y_{\max, \mathcal{S}}| > |Y_s|$ as well as the probability of $|Y_{\max, \mathcal{S}}| > |Y_s|$ are not affected. Moreover, the errors in the adjacent bins always give one-bit errors due to the Gray mapping, independently of the circular shift. Therefore, the CFO patterns for all $s \in \mathcal{S}$ result in exactly the same bit error probability $P(\hat{b} \neq b|s, \lambda)$. This means that it is sufficient to compute the CFO pattern for any single symbol s for the evaluation of the bit error probability $P_{b|\lambda}$ given in (5.23), thus reducing the complexity of evaluating (5.23) by a factor of N .

5.2.4 Frame Error Rate under AWGN and CFO

We follow a similar approach to the AWGN case in order to approximate the coded FER under CFO. First, we approximate the CWER under CFO as

$$P_{\text{cw, cfo}} = 1 - \left((1 - P_{b|\lambda})^{n_c} + \binom{n_c}{1} P_{b|\lambda} (1 - P_{b|\lambda})^{n_c-1} \right). \quad (5.24)$$

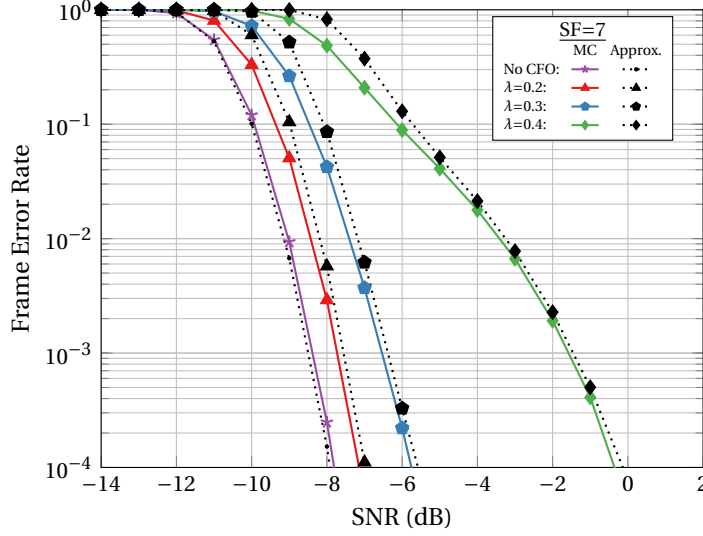
Following the same reasoning as Approximation 1 of Section 5.1.2, the FER under CFO can then be written as

$$P(\hat{\mathcal{C}}_{\text{pl, cfo}} \neq \mathcal{C}_{\text{pl, cfo}}) \approx 1 - \left(1 - P_{\text{cw, cfo}} \right)^{\frac{N_{\text{pl}} \text{SF}}{n_c}}. \quad (5.25)$$

5.2.5 Results

In Fig. 5.4, we show the FER for a coded LoRa system with $\text{SF} = 7$ for three different values of the fractional CFO ($\lambda \in \{0.2, 0.3, 0.4\}$), as well as the case without any CFO. We observe that the impact of residual CFO on the performance is significant, especially for large λ and at low error rates. Moreover, we can see that our derived approximation is quite accurate for all values of λ , when comparing to Monte Carlo simulations.

In Fig. 5.5, we compare the results of a Monte Carlo simulation for the LoRa FER for all possible $\text{SF} \in \{7, \dots, 12\}$ with the FER results obtained by using our derived approximation for $\lambda = 0.2$. The FER under only AWGN is also included in the figure with thick gray lines (taken from Fig. 5.2). We note that the choice of a common fractional offset λ for all spreading factors corresponds to different Δf_c values for each SF. We observe that, for all spreading factors, a


 Figure 5.4 – LoRa FER under three different values of CFO for $SF = 7$.

common fractional offset value leads to similar performance degradation and that our derived approximation is accurate.

Finally, in Fig. 5.6, we show the FER for a coded LoRa system with $SF = 7$ for three different values of the fractional CFO ($\lambda \in \{0.2, 0.3, 0.4\}$), with and without Gray mapping. We observe that for lower values of fractional CFO the Gray mapping does not play a significant role in the performance. However, for larger values of residual CFO, which are possible in low-cost transceivers that may avoid complex CFO correction methods, the performance improvement due to Gray mapping is significant, considering especially that it comes without any particular complexity increase in the LoRa transceiver. The fact that LoRa is relatively resilient to small values of residual CFO even without Gray mapping can be approached using the following reasoning. Adjacent demodulation errors with Gray mapping have always a weight of 1 bit error. Adjacent demodulation errors without Gray mapping, have a weight that is larger than 1, but is not necessarily very large. For example, with an exhaustive search for a typical non-Gray mapping for $SF = 7$, we found that adjacent demodulation errors have an average weight of slightly less than 2 bits. As such, the combination of interleaving and Hamming coding can be relatively resilient to small residual CFO values even in the absence of Gray mapping.

We note here that protection against ± 1 demodulation errors resulting from CFO or SFO can be obtained with the *reduced-rate* mode. In the reduced-rate mode, the interleaving and the Gray mapping are performed in blocks of $SF-2$ instead of SF codewords. The Gray mapper sets the two least significant bits of the transmitted symbol to zero. This way only $2^{(SF-2)}$ symbols are used out of the 2^{SF} possible positions. As a result, the transmitted symbols have a distance of four, allowing ± 1 demodulation errors at the receiver to be restored without ambiguity, at the cost of a reduced transmission rate [42]. The reduced rate mode is compulsorily used for the header part of the LoRa packet, which needs better protection. The reduced rate mode can also be used for the better protection of the payload part against CFO, SFO, and sampling

5.2. LoRa Coded Frame Error Rate under Carrier Frequency Offset

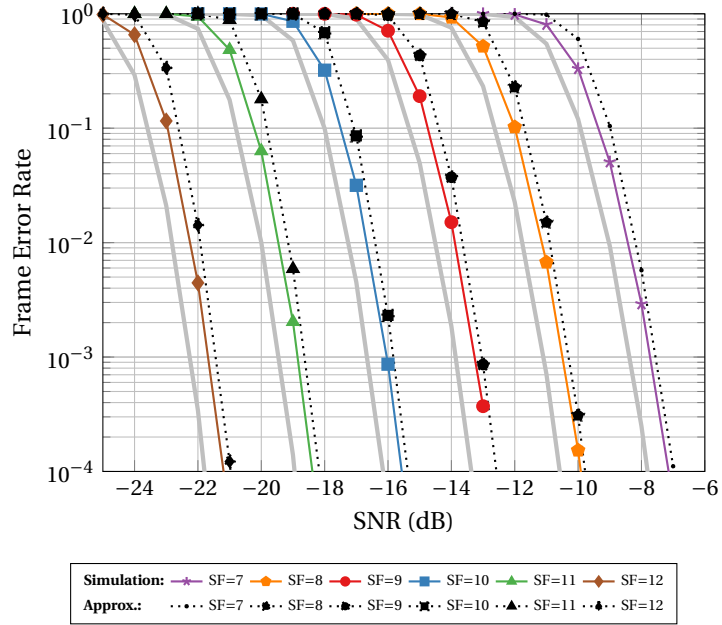


Figure 5.5 – Frame error rate of the LoRa modulation under AWGN and CFO ($\lambda = 0.2$) for all supported spreading factors $SF \in \{7, \dots, 12\}$. The thick gray lines show the FER when only AWGN is present for comparison.

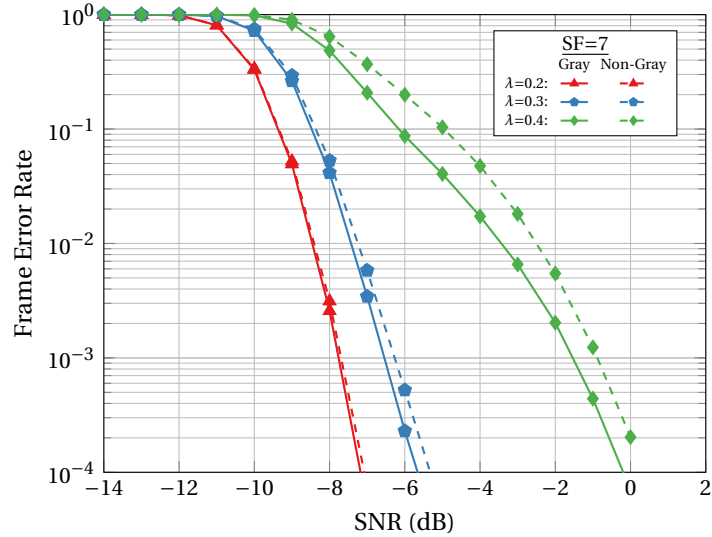


Figure 5.6 – LoRa FER under three different values of CFO for $SF = 7$ with and without Gray mapping.

jitter, especially for the larger SFs, which result in longer packets [42].

5.3 Discussion and Summary

In this chapter, we have analyzed and quantified the coded frame error rate performance of LoRa under AWGN and carrier frequency offset. Specifically, in Section 5.1 we first derived two low-complexity approximations for the coded LoRa FER under AWGN. The FER obtained by using the second approximation is shown to be within 0.2 dB of the FER obtained through Monte Carlo simulations for all LoRa spreading factors and FERs down to 10^{-5} . Moreover, in Section 5.2, we derived the FER under the impact of CFO, as well as a low-complexity approximation for the LoRa FER under CFO which is no more than 0.5 dB away from the corresponding Monte Carlo simulations.

This chapter showed that the combination of channel coding, interleaving, and Gray mapping has a great impact on the error rate when enabled, even for the simple AWGN case. Furthermore, since a residual CFO would cause consecutive LoRa symbols to be demodulated with off-by-one errors at the receiver, we analyzed the impact of these blocks on the receiver performance under such a scenario. The results showed that, thanks to the combination of Gray mapping, interleaving, and coding, LoRa is relatively robust to small values of residual CFO. The analysis performed in this chapter is also an important first step toward extending the model to the case where same-SF interference is present, in addition to residual CFO of the desired user and AWGN.

6 An Open-Source LoRa Physical Layer Prototype on GNU Radio

Since details of the LoRa PHY are not fully publicly available, many reverse engineering attempts have been carried out to reveal its structure. All these reverse engineering attempts were accompanied by software-defined radio (SDR) implementations of the PHY receiver chains. However, despite the indisputable importance of these efforts, some reverse-engineering details, as well as some details on the LoRa packet synchronization are still missing, leading thus to SDR implementations that work only at very high SNRs.

In this chapter, we describe a fully-functional GNU Radio SDR implementation of a LoRa transceiver with all the necessary receiver components to operate correctly even at very low SNRs. Moreover, our implementation fixes several issues of existing implementations (e.g., the order of channel coding and whitening). This fully-functional LoRa PHY implementation is useful for the rapid prototyping and evaluation of new algorithms that improve the performance of LoRa receivers.

In Section 6.1, we start with a description of the reverse engineering details of the LoRa PHY, focusing on the additional knowledge we have obtained compared to existing implementations. We continue in Section 6.2 with a description of the complete LoRa packet synchronization procedure, using variants of the algorithms described in [49, 50] to jointly correct the sampling time offsets and the carrier frequency offsets. Finally, we provide, to the best of our knowledge, the first fully end-to-end experimental performance results of a LoRa SDR receiver at low SNR. The testbed error rate performance is within 1 dB of the theoretical LoRa performance, as derived in the previous chapters of this thesis, even with fully decoupled transmitters and receivers. Our GNU Radio implementation of the complete LoRa transceiver chain is publicly available at [166].

6.1 LoRa PHY Reverse Engineering Details

For the reverse engineering process, we used a commercial LoRa transceiver (Adafruit Feather 32u4 [167], which uses the RFM95W LoRa modem [168]) and an NI USRP-2920 radio [169]. In

order to better understand the reverse engineering details of the LoRa PHY transceiver chain in Fig. 2.2 (except the modulation and demodulation blocks which are straightforward, and have been described already in Section 2.1) we will present them in the order in which we reverse-engineered them. Many of these steps follow the process proposed by [43, 44].

6.1.1 Gray mapping

As explained in Section 2.2, Gray mapping results in single-bit errors when a LoRa symbol is mistakenly demodulated as one of its two adjacent symbols. Such off-by-one errors can happen due to CFO as explained in Chapter 5 or also due to sampling frequency offset (SFO), and the single-bit errors they cause can be corrected by the Hamming codes with rates $CR = 4/7$ and $CR = 4/8$. In a valid Gray mapping of the $N = 2^{SF}$ binary SF-bit-long strings, any of the strings can be translated to the LoRa symbol $s = 0$. Moreover, the Gray mapping can be assigned to the decimal values of the symbols either in ascending or descending order. For these reasons, there exist $2N$ valid Gray mappings. We note here that a reverse Gray mapping is used in the Tx chain and a Gray mapping is used in the Rx chain, and not the other way around, in order for the off-by-one demodulation errors to result in single-bit errors.

6.1.2 Interleaving

A transmitted LoRa symbol can be corrupted by AWGN, resulting in multiple-bits errors in case of an erroneous demodulation. Since these errors are caused by the same symbol, they are correlated. To break this correlation between the erroneous bits the interleaving process (in particular the deinterleaver) is distributing the errors over multiple codewords. This way, the effectiveness of the error correction coding is improved, at the cost of a slightly increased latency which is caused by the necessity of receiving multiple symbols before being able to recover one codeword. Furthermore, in the case of consecutive single-bit errors caused by off-by-one demodulation errors, the fact that the interleaver is a diagonal one results also in codewords that contain with high probability only one erroneous bit, which can be handled by the Hamming decoder.

At this step of the reverse-engineering process, we want to reveal the details of the interleaving process as proposed in [43, 44]. Thus, it is convenient to remove the effect of the Hamming coding and whitening blocks. From the LoRa patent [42], we know that some variant of Hamming coding is used, therefore we know that all-zero data result in all-zero parity bits, and similarly, all-one data result in all-one parity bits. Therefore, sending either all-zero data or all-one data nulls the effect of Hamming coding. Moreover, since the whitening consists of an XOR of the data with a pseudo-random sequence, we can use the following property to null the effect of the whitening:

$$C_1 \oplus C_2 = M_1 \oplus W \oplus M_2 \oplus W = M_1 \oplus M_2 \quad (6.1)$$

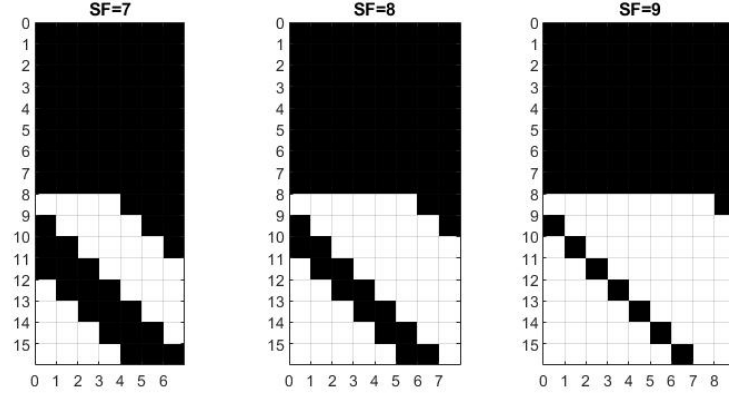


Figure 6.1 – Differential patterns for $CR = 4/8$ and $SF \in \{7, 8, 9\}$.

were C_1 and C_2 are whitened messages, M_1 and M_2 are the original non-whitened messages, and W is the whitening sequence. The result of $M_1 \oplus M_2$ will be referred to as a *differential pattern*.

Fig. 6.1 shows the differential pattern of the 16 first symbols in a packet after the preamble is removed. This differential pattern is derived by first transmitting a payload entirely composed of zeros, and then transmitting a payload entirely composed of ones. Black boxes indicate bits with value zero and white boxes indicate bits with value one. The parameters used are with a coding rate $CR = 4/8$ and spreading factors $SF \in \{7, 8, 9\}$. The objective of these differential patterns is to indicate the starting position of the payload bits inside the packet, after removing the preamble and when transmitting in implicit header mode, which means that there is no header present in the packet. By plotting the differential pattern, we expect to have a value of 1 at the bit positions that were affected by changing the payload from 0 to 1. We observe that in the first 16 symbols of the packet there exist 80 coded bits that do not depend on the payload content. This result is surprising, since we send these messages in implicit header mode, which should avoid the existence of any header, and the packet content after the preamble should only depend on the payload. We note that it is mentioned in the Arduino library used for the transmissions that 4 bytes are prepended to the payload and are namely `to`, `from`, `ID`, and `Flags`. This means that these 40 uncoded bits we observe are not part of the LoRa PHY. We will refer to these bits as payload-prepended bits in the following.

By changing the coding rate to $4/6$ we obtain the pattern shown in Fig. 6.2. We can observe that changing the coding rate impacts the number of symbols that are interleaved together in an interleaving block. Moreover, as highlighted by the red boxes, there exists a first block of 8 symbols that are interleaved using a coding rate of $4/8$, independently of the chosen coding rate. The lowest-rate code is supposed to be used for the header in the explicit-header mode [41]. The reason is because when a header is present in a LoRa packet, it is of crucial importance to receive it correctly, since the decoding of the payload depends on the information in the header. Nevertheless, we see that the low-rate that is used in the beginning of the packet in

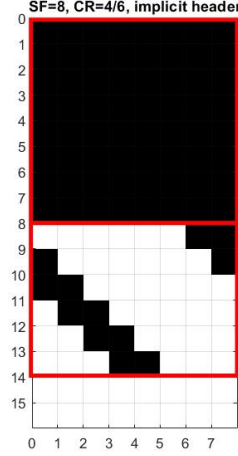


Figure 6.2 – Differential pattern for CR = 4/6 and SF=8.

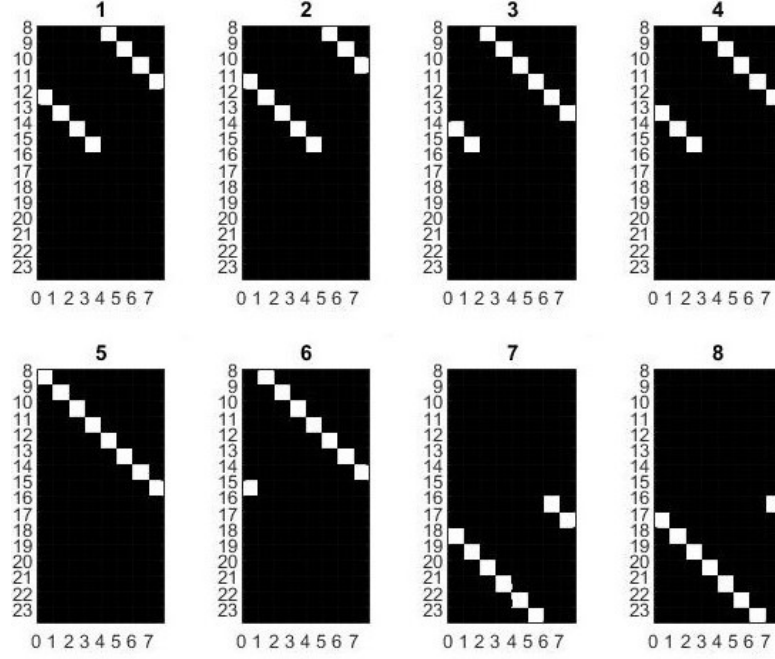
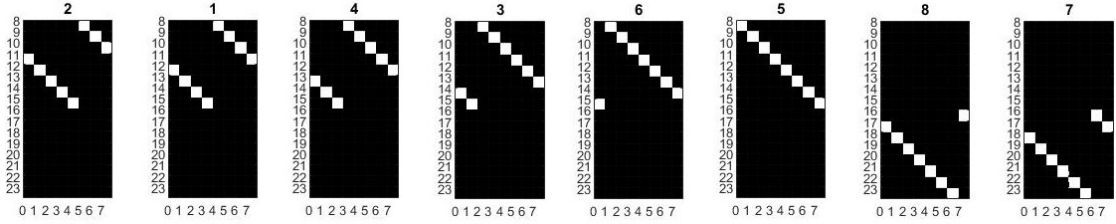
explicit header mode, is also used in the implicit header mode. We observe in Fig. 6.2 that with a coding rate of $4/6$, we get a different number of coded payload-prepended bits than with a coding rate of $4/8$ in Fig. 6.1. However, taking into account that the first 8 symbols are coded using a coding rate of $4/8$ and only the following ones are using $4/6$, we obtain $\frac{4}{8} \cdot 64 + \frac{4}{6} \cdot 12 = 40$ uncoded payload-prepended bits, as in the previous case.

The next step is to extract the interleaving method that is used to distribute the bits of each codeword to different symbols. To this end, we transmit several payloads that differ in a single nibble (i.e., 4 bits). By observing the differential patterns of these transmissions, we are able to observe the position of each codeword in the packet after the interleaving.

Fig. 6.3 presents the differential patterns obtained by inverting the first eight nibbles one at a time, using CR = $4/8$ and SF = 8. We can observe that all the bits of a codeword get distributed to different LoRa symbols. Furthermore, it appears that when the original bytes are split into two nibbles, the nibble containing the LSBs is sent first. This reordering becomes evident by looking at Fig. 6.4 which presents the same patterns as in Fig. 6.3, but swapped every two symbols. In that order the pattern appears to be shifted by one to the left each time the following payload nibble is transmitted with flipped bits.

Fig. 6.5 presents the interleaving blocks for three different combinations of coding rates and spreading factors, namely CR = $4/8$ and SF = 8, CR = $4/6$ and SF = 7, and CR = $4/5$ and SF = 9. All patterns have been obtained following the same procedure as the one explained above. We note that an interleaving block transforms a matrix of dimension $SF \times \frac{4}{CR}$ to $\frac{4}{CR} \times SF$. Each line in the matrix before interleaving corresponds to a codeword, while each line in the matrix after interleaving corresponds to a LoRa symbol. From this procedure, we can finally extract the following general interleaving formula:

$$I_{i,j} = D_{j,(i-(j+1) \bmod SF)}, \quad (6.2)$$


 Figure 6.3 – Differential patterns for $CR = 4/5$ and $SF = 8$.

 Figure 6.4 – Differential patterns swapped every two for $CR = 4/5$ and $SF = 8$.

where I is the matrix after interleaving and D the matrix before interleaving.

6.1.3 Whitening

As explained in Chapter 2, whitening helps to remove the DC-bias and to break any data dependencies at the receiver, before performing the decoding. The whitening sequence that LoRa uses can be found in [166], and was derived from the whitening matrices shown in Fig. 6.6. Recovering the whitening sequence is fairly straightforward since $M \oplus 0 = M$, meaning that when sending a long all-zero message, the received bits will correspond to the whitening sequence. To obtain the matrices in Fig. 6.6, we transmit an all-zero message, and we show the received message in every case. Black boxes indicate received bits of value 0, and white boxes indicate received bits of value 1. Special care needs only to be taken of the payload-prepended bits, since these cannot be forced to 0. We note that we obtain the same whitening matrices

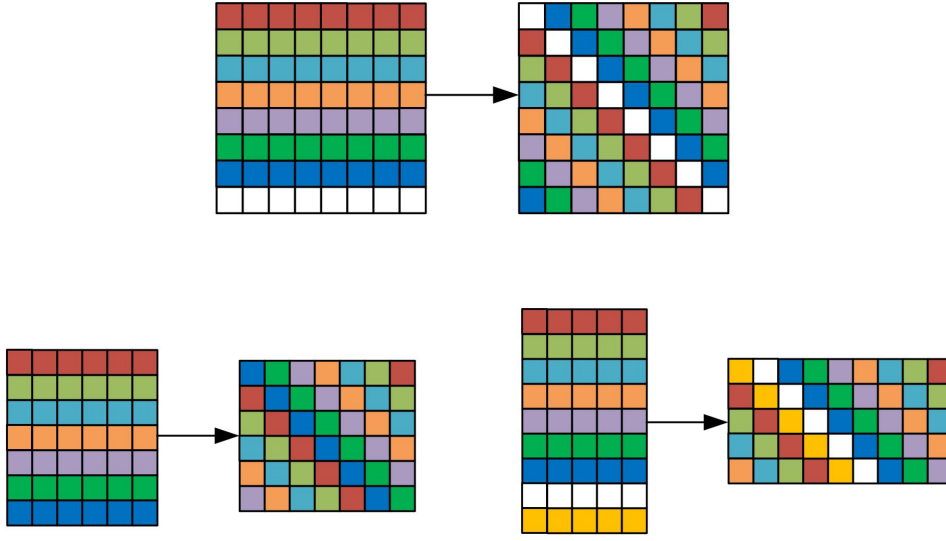


Figure 6.5 – Interleaving blocks for three combinations of CR and SF: (a): $CR = 4/8$ and $SF = 8$, (b): $CR = 4/6$ and $SF = 7$, and (c): $CR = 4/5$ and $SF = 9$.

for all possible combinations of coding rates and spreading factors, a fact which indicates also that the interleaving occurs after the whitening in the Tx chain.

The base whitening matrix shown in Fig. 6.6 is the one that corresponds to the coding rate $CR = 4/8$. For $CR = 4/7$ and $CR = 4/6$ one and two of the rightmost columns of the base whitening matrix are removed, respectively. For $CR = 4/5$ the last column in the whitening matrix is different than the corresponding fifth column in the base whitening matrix. This happens because, as explained in the sequel, for $CR = 4/5$ a single parity check is used instead of a Hamming code. Moreover, we found that for the coding rate $CR = 4/5$, the parity bit is calculated from the whitened version of the four bits, meaning that the whitening block is the first block in the transmitter chain, something that was not mentioned in the previous reverse engineering attempts of the LoRa PHY.

6.1.4 Hamming Coding

From the patent [42], we know that the error-correction scheme used by LoRa is some variation of Hamming coding. In the preceding reverse engineering steps, the bit position inside each codeword was undefined. After transmitting known data and deinterleaving them, we receive them in the order presented in Fig.6.7. We note that the parity bits are in arbitrary positions, since any other possible positioning would only shuffle the generator and parity-check matrices. We also note that d_0 denotes the least significant bit (LSB) of the data nibble.

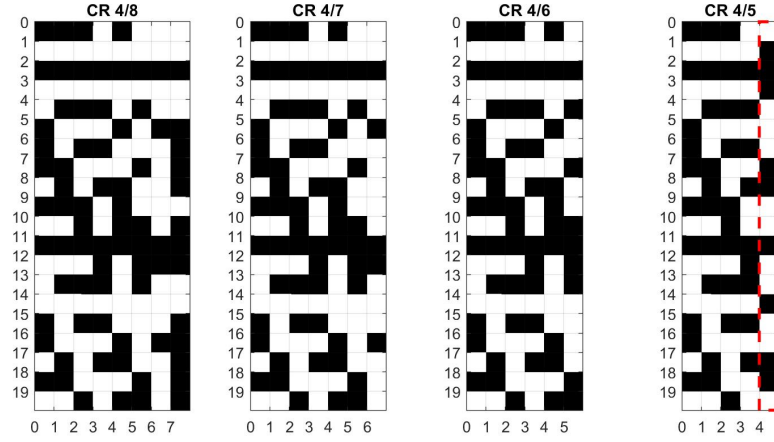


Figure 6.6 – Whitening matrices for different coding rates

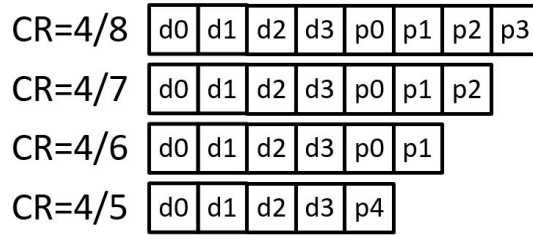


Figure 6.7 – Bits in each deinterleaved codeword

The check-sums for the parity bits are

$$p_0 = d_0 \oplus d_1 \oplus d_2 \quad (6.3)$$

$$p_1 = d_1 \oplus d_2 \oplus d_3 \quad (6.4)$$

$$p_2 = d_0 \oplus d_1 \oplus d_3 \quad (6.5)$$

$$p_3 = d_0 \oplus d_2 \oplus d_3 \quad (6.6)$$

However, for $CR = 4/5$, p_4 is a checksum of the whitened data bits. We verified this by sending all 16 different data nibbles, as shown in Fig. 6.8. We observe that the last column corresponds indeed to the parity bit of the whitened data nibble, verifying also that the whitening is performed before the coding block in the Tx chain.

Encoding: For the coding rates $CR \in \{4/6, 4/7, 4/8\}$, the encoding is performed with the following

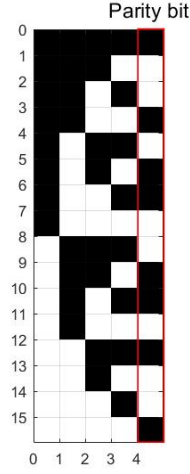


Figure 6.8 – Exhaustive list of whitened codewords

modulo-2 operation

$$\begin{bmatrix} d_3 & d_2 & d_1 & d_0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} d_0 & d_1 & d_2 & d_3 & p_0 & p_1 & p_2 & p_3 \end{bmatrix}$$

where only the first $\frac{4}{CR}$ bits are kept.

Decoding: For $CR = 4/7$ and $CR = 4/8$ we perform syndrome decoding, where the syndrome indicates the position of the erroneous bit in the codeword. Additionally, for $CR = 4/8$ we use the extra parity bit p_3 to avoid swapping a correct bit in the case of an even number of erroneous bits. If p_3 indicates that there is an odd number of erroneous bits, then we treat $CR = 4/8$ exactly the same way as the typical Hamming code with $CR = 4/7$.

6.1.5 PHY Header

In the patent [42], it is mentioned that the header should always be encoded using $CR = 4/8$ to ensure the best possible protection, since the header contains crucial information regarding the transmission parameters. Defining a fixed coding rate for the header for all LoRa transmissions is also mandatory because the coding rate is not known during the decoding of the header. The patent describes what is included in the header, namely the coding rate, the payload length, the presence of a payload CRC, a burst mode indication, a compressed mode indication, some reserved bits, a ranging bit, and a header CRC.

Fig. 6.9 presents the first 16 codewords of a differential message obtained by sending a packet two times, varying only the coding rate used between the two transmissions. This way, in

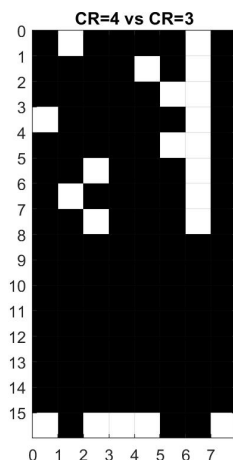


Figure 6.9 – Differential message between transmissions with $CR = 4/7$ and $CR = 4/8$ in explicit header mode.

Fig. 6.9 we can observe which part of the header changed, indicating the presence of the explicit header at that position. We can see that the explicit header is located in the first 8 symbols. We also compared the whitening sequence obtained for the implicit- and explicit-header mode, and we conclude that the header is not whitened.

In order to find the position of each element in the header, we sent packets with many different configurations, varying the coding rate, the payload length, and the presence of a payload CRC. We then combine all the received headers to create the matrix shown on Fig. 6.10. In this matrix, all the bits that kept a value of 0 in all transmissions, independently of the configuration of the aforementioned parameters, are shown in black. Similarly, all the bits that kept a value of 1 are shown in white. The bits that changed their value, depending on the configuration are shown in gray. We observe that the last two columns have a special structure. In particular, the last column always contains bits with value 0, and the second last column does not contain any position that has always a fixed value of either 0 or 1.

In Fig. 6.11 we see the received header of a transmission with $CR = 4/7$, a payload length of 16 bytes, and no payload CRC present. As expected from Fig. 6.10, the last column contains only bits with value 0. We observe that the second last column (marked in red) contains the checksum of each row.

We note that the last two columns are added to the header after interleaving, which is done in what is named the *reduced rate* mode by interleaving SF–2 codewords together, instead of SF in the typical case. Reduced rate interleaving can be done by simply replacing the term SF in Eq. 6.2 with SF–2. The advantage of the reduced rate mode for the header is to provide a greater distance between the symbols during the demodulation, helping in cases of residual CFO and SFO, among others [42].

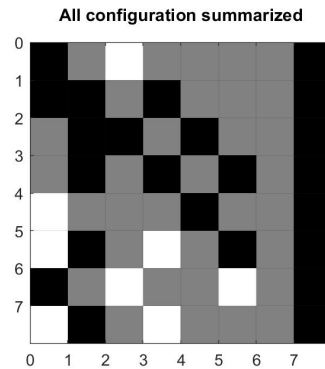


Figure 6.10 – Summary matrix of many different headers

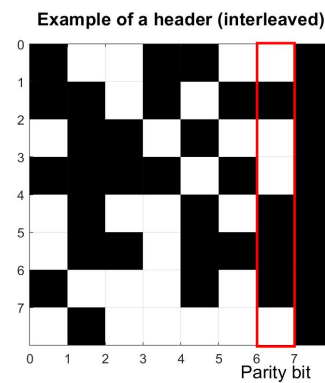


Figure 6.11 – Header received for CR=4/7, Payload length of 16 Bytes and CRC off

Coding rate: In Fig. 6.12 every matrix is created by fixing the CR value to one of the three values in $\{4/6, 4/7, 4/8\}$, and transmitting packets with all possible configurations for the rest of the parameters. These matrices do not reveal any information on their own, except if they are combined. We can combine these matrices to find the positions that are constant (either black or white) for each individual CR but changing between two CRs. These header positions should be the ones that contain the CR information. This masking method is shown in Fig. 6.13, in which the four white squares represent the position of the bits inside the header that contain the information about the CR. By applying this mask on top of the headers of figure 6.12, we get the three patterns in Fig. 6.13 presented alongside the mask. Since we have already revealed the interleaving patterns, we know that the data bits should be in the first four rows, and therefore the bit in the sixth row is a parity bit. The CR is therefore encoded in the bits of rows two, three, and four as follows

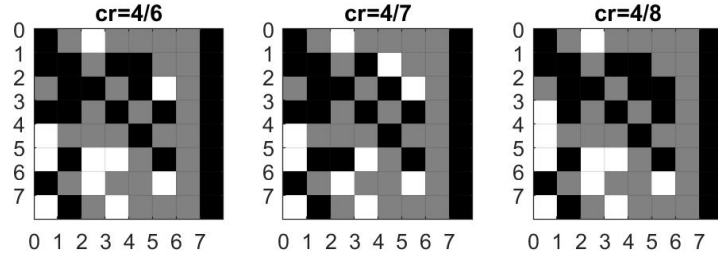


Figure 6.12 – Summary matrices for three different coding rates

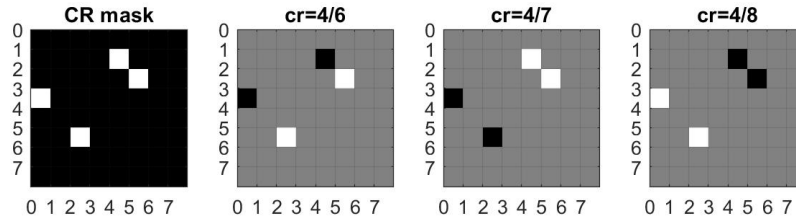


Figure 6.13 – Mask and bit corresponding to the coding rate information

- 010 for $CR = 4/6$
- 110 for $CR = 4/7$
- 001 for $CR = 4/8$

As explained in the RFM95 datasheet [168], the coding rate has to be specified by the value n , as $CR = \frac{4}{4+n}$. The aforementioned three messages correspond to the LSB-first binary representation of the values for n .

CRC presence and Payload length: Following exactly the same masking method used to find the positions of the bits that contain the CR information, we recover the position of the bit that indicates the presence or absence of a CRC, and the bits that indicate the payload length. The masking plots can be found in [51], and are not shown here for brevity.

Header CRC: In order to find the header CRC, we performed all the necessary steps to find if it is based on a polynomial. After these steps which we explained in [51], we concluded that the CRC is not based on a polynomial, therefore we extracted the CRC with an exhaustive search method, where for every transmission we set all header bits except one to 0, trying all the possible different cases.

6.1.6 Payload CRC

Following a methodology described in [51], we found that the payload CRC is based on the CCITT-16 polynomial, which is $x^{16} + x^{12} + x^5 + 1$. We also found that the 16-bits CRC is not calculated over the entire payload, but the last two bytes of the payload are excluded. Finally, the CRC of the rest of the payload is XORed with these last two bytes of the payload.

6.2 LoRa Frame Synchronization

The demodulation procedure described in Section 2.1 assumes that the receiver is perfectly synchronized to the incoming signal and that no impairments are present. However, in practice these assumptions typically do not hold and some additional processing steps are required in order to ensure correct frame synchronization which corrects the STO and CFO impairments. We note that in LoRa the STO and CFO need to be corrected already during the frame synchronization step and not only for the data part of the frame. In this section, we describe how we ensure fine-grained synchronization and correct demodulation in the presence of STO and CFO in our USRP-based GNU Radio LoRa receiver. To this end, we first write an expression for the received signal before frame synchronization, in order to later describe the appropriate frame synchronization and offset correction algorithms.

6.2.1 Frame Synchronization Signal Model

The received packet is sampled with a frequency $f_s = B$ and, in our implementation, decisions are taken in windows of N samples. The decision windows are generally not synchronized to the preamble upchirps and the N samples of a decision window will typically consist of parts of two preamble upchirps, which we denote by s_{up_1} and s_{up_2} and where $s_{\text{up}_1} = s_{\text{up}_2} = 0$. Following a similar notation as in the previous chapters, let τ_{STO} be the relative time offset between the first chip of the decision window and the first chip of s_{up_2} in the decision window. Then the discrete-time baseband-equivalent equation of the transmitted signal $x[n]$ at the receiver is

$$x[n] = \begin{cases} e^{j2\pi\left(\frac{(n+N-\tau_{\text{STO}})^2}{2N} - \frac{1}{2}(n+N-\tau_{\text{STO}})\right)}, & n \in \mathcal{N}_{L_1}, \\ e^{j2\pi\left(\frac{(n-\tau_{\text{STO}})^2}{2N} - \frac{1}{2}(n-\tau_{\text{STO}})\right)}, & n \in \mathcal{N}_{L_2}, \end{cases} \quad (6.7)$$

where $\mathcal{N}_{L_1} = \{0, \dots, \lceil \tau_{\text{STO}} \rceil - 1\}$ and $\mathcal{N}_{L_2} = \{\lceil \tau_{\text{STO}} \rceil, \dots, N - 1\}$. Moreover, let f_{c_1} denote the carrier frequency used during up-conversion at the transmitter and f_{c_2} the carrier frequency used during down-conversion at the receiver. The carrier frequency offset is the difference $\Delta f_c = f_{c_1} - f_{c_2}$. The corresponding signal model is

$$y[n] = hc[n]x[n] + z[n], \quad n \in \mathcal{S}, \quad (6.8)$$

where h is the channel gain between the transmitter and the receiver, $c[n] = e^{j2\pi(n+(m-1)N)\frac{\Delta f_c}{f_s}}$ is the CFO term affecting the m -th symbol in the packet, $x[n]$ is the transmitted signal, and $z[n] \sim \mathcal{N}(0, \sigma^2)$ is AWGN. Similarly to the models in Chapters 3 and 4, the STO τ_{STO} can be separated into an integer part $L_{\text{STO}} = \lfloor \tau_{\text{STO}} \rfloor$, and a fractional part $\lambda_{\text{STO}} = \tau_{\text{STO}} - \lfloor \tau_{\text{STO}} \rfloor$. In addition, similarly to Chapters 4 and 5, the CFO translates into an offset $\tau_{\text{CFO}} = \frac{\Delta f_c N}{f_s}$, which can also be split into an integer part $L_{\text{CFO}} = \lfloor \tau_{\text{CFO}} \rfloor$, and a fractional part $\lambda_{\text{CFO}} = \tau_{\text{CFO}} - \lfloor \tau_{\text{CFO}} \rfloor$. The STO and the CFO affect the signal in a combined manner that makes their joint estimation and correction challenging [49, 50].

6.2.2 Synchronization and Offset Correction

In this subsection, we describe the algorithms used for the synchronization of a LoRa packet, along with the algorithms to estimate and compensate the STO and CFO. We explain why these two tasks need to be performed in a combined manner.

Preamble detection

In the presence of STO and CFO and in the absence of AWGN, it can be shown that the demodulation decisions of the detection windows during the preamble upchirps are $\hat{s} = \lfloor \tau_{\text{STO}} - \tau_{\text{CFO}} \rfloor = \lfloor (L_{\text{STO}} - L_{\text{CFO}}) + (\lambda_{\text{STO}} - \lambda_{\text{CFO}}) \rfloor$. Our SDR LoRa receiver detects the presence of a preamble when $N_{\text{pr}} - 1$ consecutive symbols are demodulated with values in a range $\{s-1, s, s+1\}$. The reason for allowing this margin during preamble detection is that the fractional offsets $\lambda_{\text{STO}} - \lambda_{\text{CFO}}$ can lead to ± 1 demodulation errors in the presence of AWGN [42, 48, 49, 50]. Finally, the preamble synchronization value \hat{s}_{pr} is decided using a majority rule from the $N_{\text{pr}} - 1$ values in the range $\{s-1, s, s+1\}$. In this first part of the synchronization procedure, the receiver performs a coarse time synchronization by discarding $N - \hat{s}_{\text{pr}}$ samples from its buffer. This way the buffer now contains $\hat{N}_{\text{pr}} - 2$ symbols with a value of $\hat{s} = 0$ or $\hat{s} = \pm 1$. This coarse synchronization is necessary as a first step in order to later apply the estimation algorithms for the fractional offsets. It is important to note that after this coarse synchronization, the receiver will still be misaligned in time, since the integer part of the CFO, i.e., L_{CFO} , also affects the value of \hat{s}_{pr} and therefore results in a time misalignment. Moreover, the estimation of \hat{s}_{pr} has been performed under the effect of the combination of the fractional parts $\lambda_{\text{STO}} - \lambda_{\text{CFO}}$, which can lead to an additional time offset of ± 1 samples.

Estimation and Compensation of L_{STO} and L_{CFO}

A LoRa receiver can operate with the time misalignment resulting from L_{CFO} , since it translates into a frequency offset which finally compensates for L_{CFO} . Ideally, however, the receiver should avoid the time misalignment due to L_{CFO} because it results in inter-symbol interference [48]. Instead, L_{CFO} should be compensated by introducing a frequency shift. A simple and effective way of distinguishing the integer parts of the STO and the CFO has been de-

scribed in [49, 50], where the receiver takes advantage of the 2.25 downchirps in the preamble to separate the L_{STO} and L_{CFO} values. We implement the approach of [49], or equivalently [50, Eq. (26), (27)] in our LoRa SDR receiver for this part of the synchronization process. We re-synchronize our receiver in time, using initially only the value of L_{STO} , and we compensate the effect of L_{CFO} by introducing a frequency shift through multiplication with a complex exponential signal $e^{-j2\pi n \frac{L_{\text{CFO}}}{f_s}}$.

Estimation and Compensation of λ_{STO} and λ_{CFO}

The estimation of a fractional offset in the frequency domain is a well-studied problem [170]. In particular, interpolation between the three maximum peaks of a sinc kernel can be used in order to find the value of the fractional offset with good accuracy and low-complexity. We propose two variations of the rational combination of the three spectral lines (RCTSL) method described in [170], to estimate λ_{CFO} and λ_{STO} .

As explained in [50], λ_{CFO} has to be estimated and compensated before estimating λ_{STO} . For the estimation of λ_{CFO} we use the $N_{\text{pr}} - 2$ preamble symbols which were left in the buffer after the previous synchronization steps. We dechirp the symbols and we perform a zero-padded DFT of length $2(N_{\text{pr}} - 2)N$ (the zero-padding by a factor of two is required by the method of [170]) on an entire block of the $N_{\text{pr}} - 2$ preamble symbols left in the buffer, i.e., $\tilde{\mathbf{Y}} = \text{DFT}(\tilde{\mathbf{y}} \odot \tilde{\mathbf{x}}_0^*)$, where $\tilde{\mathbf{y}} = [\mathbf{y}_1 \dots \mathbf{y}_{N_{\text{pr}}-2} \mathbf{0}_{(N_{\text{pr}}-2)N}]$, $\tilde{\mathbf{x}}_0 = [\mathbf{x}_0 \dots \mathbf{x}_0 \mathbf{0}_{(N_{\text{pr}}-2)N}]$, and $\mathbf{0}_{(N_{\text{pr}}-2)N}$ denotes a zero vector of length $(N_{\text{pr}} - 2)N$. Let $\tilde{Y}_{k_{\text{max}}}$ be the value of the maximum bin of the DFT $\tilde{\mathbf{Y}}$. Then, we compute [170]

$$k_\alpha = \frac{N}{\pi} \frac{|\tilde{Y}_{k_{\text{max}}+1}|^2 - |\tilde{Y}_{k_{\text{max}}-1}|^2}{u(|\tilde{Y}_{k_{\text{max}}+1}|^2 - |\tilde{Y}_{k_{\text{max}}-1}|^2) + v|\tilde{Y}_{k_{\text{max}}}|^2}, \quad (6.9)$$

where $u = \frac{64N}{\pi^5 + 32\pi}$ and $v = u \frac{\pi^2}{4}$ [170]. The fractional carrier frequency offset is computed as

$$\lambda_{\text{CFO}} = \frac{k_{\text{max}} + k_\alpha}{2(N_{\text{pr}} - 2)} \mod 1. \quad (6.10)$$

Finally, the offset λ_{CFO} is corrected through a multiplication of the received signal with $e^{-j2\pi n \frac{\lambda_{\text{CFO}}}{f_s}}$.

After λ_{CFO} has been corrected, we re-use the $N_{\text{pr}} - 2$ preamble symbols that were used for the estimation of λ_{CFO} to estimate λ_{STO} as follows. For every dechirped symbol we perform a length- $2N$ DFT (again, the upsampling by a factor of two is required by the method of [170]) and we combine the DFTs of the individual symbols in the following way

$$|Y'_k|^2 = \sum_{i=1}^{N_{\text{pr}}-2} |Y_k^{(i)}|^2, \quad k \in \{0, \dots, 2N - 1\}. \quad (6.11)$$

We then find $|Y'_{k_{\text{max}}}|^2$ and we use it instead of $|\tilde{Y}_{k_{\text{max}}}|^2$ in (6.9) to calculate k_α . The fractional

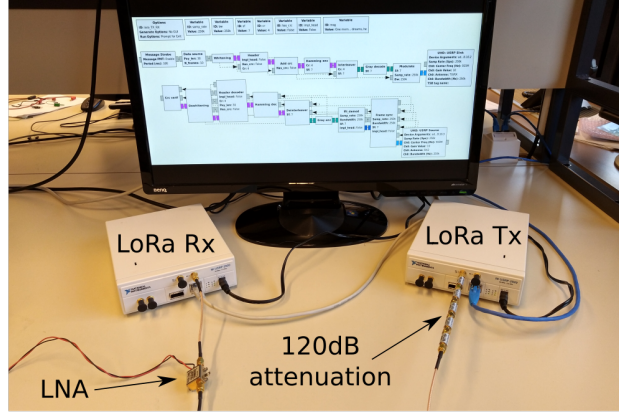


Figure 6.14 – LoRa testbed in USRPs using our GNU Radio SDR platform.

sampling time offset is

$$\lambda_{\text{STO}} = \frac{k_{\text{max}} + k_{\alpha}}{2} \mod 1. \quad (6.12)$$

Finally, λ_{STO} is compensated using time-domain interpolation.

6.3 LoRa Testbed and Measurement Results

In this section we briefly describe our GNU Radio LoRa PHY implementation. Moreover, we use this implementation in a testbed in order to experimentally assess its correctness by comparing the experimental performance results to the theoretical error rate performance of LoRa. Our open-source implementation is publicly available for the research and development of algorithms that improve the performance of LoRa receivers [166].

6.3.1 GNU Radio LoRa PHY Implementation

In the GNU Radio implementation of the LoRa Tx and Rx chains the user can choose all the parameters of the transmission, such as the spreading factor, the coding rate, the bandwidth, the presence of a header and a CRC, the message to be transmitted, etc. In the Tx chain, the implementation contains all the main blocks of the LoRa transceiver described in Section 2.2, i.e., the header- and the CRC-insertion blocks, the whitening block, the Hamming encoder block, the interleaver block, the Gray mapping block, and the modulation block. On the receiver side there is the packet synchronization block, which performs all the necessary tasks needed for the synchronization, such as the STO and CFO estimation and correction as described in Section 6.2. The demodulation block follows, along with the Gray demapping block, the deinterleaving block, the Hamming decoder block and the dewhitening block, as well as a CRC block. The implementation supports all spreading factors, both uncoded and coded modes (for all coding rates), both implicit and explicit header modes, with and without

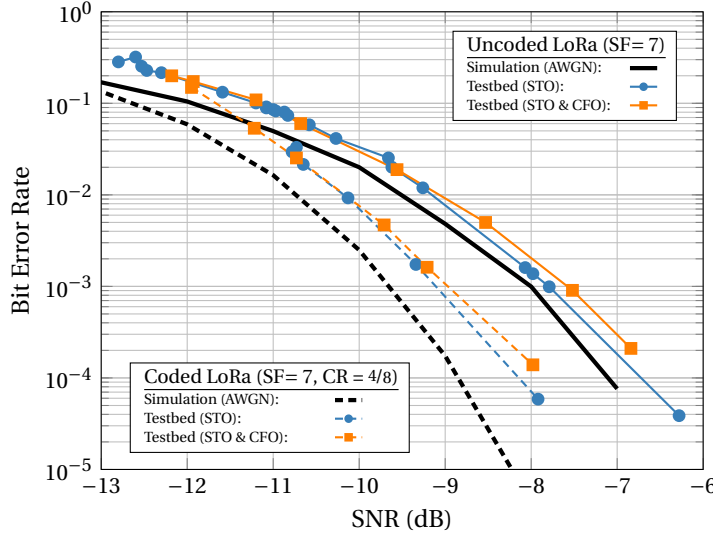


Figure 6.15 – Uncoded and coded LoRa BER for SF = 7 (testbed vs simulation).

the presence of a CRC. The reduced rate mode for the payload is not included in the current version of the GNU Radio implementation. Interoperability with a commercial LoRa module has been tested for all spreading factors and all the aforementioned modes using the Adafruit Feather 32u4, which uses the RFM95W LoRa modem.

6.3.2 Testbed description

Our testbed uses National Instruments (NI) 2920 USRP transceivers [169], but any SDR device that supports GNU Radio can be used instead. We use a bandwidth of $B = 250$ kHz because the NI 2920 USRP transceivers have a 200 kHz lower limit on the sampling frequency. In order to use the $B = 125$ kHz mode of LoRa, downsampling is required. The transmit power of the NI 2920 USRP ranges from -11 dBm to 20 dBm. The nominal carrier frequency used for our transmissions is 915 MHz. We note that, as can be seen in Fig. 6.14, in order to avoid interference from other sources in 915 MHz ISM band, we use a cable with attenuators for transmission instead of antennas, to operate in a controlled environment for test and characterization. The carrier frequency is generated within each USRP from a reference clock. This reference clock can be shared between the Tx and the Rx USRPs, leading to $L_{\text{CFO}} = \lambda_{\text{CFO}} = 0$, for initial validation experiments without CFO and to obtain the CFO ground truth for measurements of the CFO compensation.¹ We use the Adafruit Feather 32u4 RFM95W for functional verification of our GNU Radio implementation. Moreover, we use an external MiniCircuits ZX60-33LN+ low-noise amplifier [171] with a gain of approximately 19 dB to enhance the receiver gain of the USRP. Using this testbed, we have successfully transmitted LoRa packets from a USRP to a USRP, from the Adafruit Feather to a USRP, and from a USRP to the Adafruit Feather.

¹We note that, even when sharing the reference clock, a random, time-varying STO is present, since only the frequencies, not the phase, are locked.

6.3.3 Testbed results

In Fig. 6.15 we show bit error rate (BER) results for $SF = 7$, $B = 250\text{kHz}$, and for a payload of 64 bytes obtained from transmissions using our USRP testbed which runs our GNU Radio implementation of LoRa PHY for both the uncoded and coded ($CR = 4/8$) cases. The experimental curves are compared to the performance of LoRa obtained through MATLAB-based Monte Carlo simulations. Results for the cases with and without CFO are shown. Low SNR values are achieved by using multiple attenuators, reaching a total attenuation of 120 dB. We generate different SNR values by fixing the attenuation to 120 dB and sweeping the Tx gain. Even for a fixed Tx gain, we observed that the received signal power may vary between different runs. Therefore, for a given SNR we consider only the packets with received signal power no more than ± 1 dB of the mean received signal strength value for the given Tx gain. The SNR for the experimental curves is measured as the ratio between the power of the maximum DFT bin over the power contained in the rest of the bins after synchronization and STO/CFO compensation. The SNR increase due to any cross-talk between the USRPs is thus inherently captured in the measured SNR value.

For uncoded LoRa, we observe that in the case with STO, but without CFO, the experimental curve matches the simulated performance of uncoded LoRa under AWGN (as described in Chapter 3) very well. This result shows that the STO estimation and compensation algorithm described in Section 6.2 is accurate, leading to a testbed performance that is very close to the theoretical limits for AWGN. Moreover, we observe that in the case where the local oscillator is not shared, and therefore both STO and CFO are present, the performance degradation of the implementation compared to the bound given by the AWGN simulation curve is still less than 1 dB.

For coded LoRa, we observe that in the case with STO, but without CFO, the experimental curve is within 1 dB of the simulated performance of coded LoRa under AWGN (as described in Chapter 5). Additionally, we observe that the performance degradation in the presence of both STO and CFO is almost negligible, since the CFO estimation and compensation works well, and moreover, coded LoRa is relatively resistant to small residual CFO values after the correction.

6.4 Discussion and Summary

In this chapter we presented our work on the SDR implementation of LoRa PHY. We started in Section 6.1 with presenting some details of our reverse engineering work toward revealing parts of the LoRa PHY. These details are necessary for the correct implementation of LoRa PHY in SDR, but are also useful on their own, since they provide better understanding of the structure of the LoRa PHY. We continued in Section 6.2 with a description of the algorithms we use to synchronize to a received LoRa packet, in strong connection with estimation and compensation algorithms for the STO and the CFO. Finally, in Section 6.3, we presented our LoRa testbed that is able to provide experimental performance results in low SNR.

Chapter 6. An Open-Source LoRa Physical Layer Prototype on GNU Radio

The reverse engineering process and the SDR implementation of LoRa PHY presented in this chapter was performed in parallel with the theoretical works of the previous chapters, and was in strong interplay with them. The theoretical work enabled us to implement correctly the SDR transceiver, and the better knowledge of the LoRa PHY obtained during the reverse engineering and the experimental building processes further boosted our theoretical understanding. The most important contribution of the work in this chapter is that it establishes the first LoRa implementation that can be used as a basis for rapid prototyping and performance assessment of future advanced algorithms for the LoRa receiver.

In-Band Full-Duplex Wireless Communications

Part II

7 Introduction to Full-Duplex Transceivers

The increased demand in wireless traffic over the last years has led researchers to investigate wireless systems that use the spectrum more efficiently in order to support more users and a higher traffic load. In-band full-duplex technology is an example of such a novel approach that promises to double the spectral efficiency of traditional half-duplex wireless links by allowing simultaneous transmission and reception in the same frequency band [16, 17, 18]. The main challenge for this technology is the presence of a strong self-interference signal from the transmitter to the receiver of the same full-duplex device. This self-interference signal can be many orders of magnitude stronger than the received signal of interest, which comes from a transmitter that is located much farther away than the transmitter of the full-duplex device.

As the baseband transmitted signal is known within the full-duplex transceiver, it is in principle possible to generate an appropriate cancellation signal that will effectively suppress the self-interference. Ideally the self-interference should be suppressed to (or below) the receiver noise floor. However, it has been shown that the actual radio frequency (RF) self-interference signal is a highly complex non-linear function of the baseband transmitted signal [20], mainly due to non-idealities of the transceiver hardware. This function is difficult to model. As a result, an amount of residual self-interference is left in many full-duplex transceiver implementations.

In this chapter, we introduce the main principles of the in-band full-duplex wireless technology, along with the main transceiver impairments that make self-interference cancellation a cumbersome task. The aforementioned introduction to full-duplex transceivers and self-interference cancellation techniques is important both in order to propose new cancellation methods, as we do in Chapter 8, and in order to understand the limitations of the full-duplex technology. These limitations motivate our effort in Chapter 9 to propose full-duplex systems that can be useful even with only imperfect self-interference cancellation. In this chapter, we begin in Section 7.1 with a description of the analog and digital self-interference cancellation stages, as well as the different full-duplex transceiver architectures that have been proposed in the literature. In Section 7.2, we explain the different transceiver impairments that affect the self-interference signal. Finally, in Section 7.3, we present the state of the art in the research on full-duplex systems.

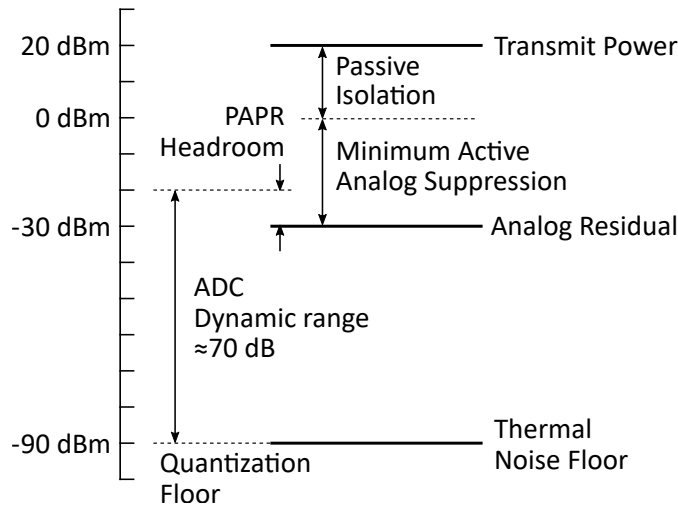


Figure 7.1 – Power budget for a full-duplex transceiver with multiple self-interference suppression stages. This figure was created based on [2].

7.1 Self-interference Cancellation in Full-Duplex Wireless Transceivers

Simultaneous transmission and reception is commonly performed either with two separate co-located antennas, one for transmission and the other for reception, or with one shared antenna and a circulator. Some amount of passive self-interference suppression is obtained by either the physical separation of the transmit and the receive antennas, or the isolation provided by the circulator [172]. Passive suppression can also be obtained using directional antennas, beamforming, polarization, or shielding. However, passive suppression alone is not sufficient for functional full-duplex transceivers, since it typically cannot reduce the self-interference signal to the noise floor. Most importantly, additional suppression is needed before the digitization of the signal in the analog to digital converter (ADC), due to the limited resolution of the ADC. This will be explained in more detail using the power budget diagram of Fig. 7.1.

The power levels in Fig. 7.1 are based on [2], and the full-duplex testbed we use, that will be presented in Chapter 8. In particular, we consider as an example a transmission with a full-duplex transceiver resulting in a measured noise floor at -90 dBm. The transmit power is set to 20 dBm. Therefore, a total self-interference suppression of 110 dB is needed for the self-interference signal to be suppressed down to the receiver noise floor. The effective number of bits in the ADC is 11.5 [173, 2], thus the dynamic range of the ADC is approximately 70 dB. The peak-to-average power ratio (PAPR) of the transmitted OFDM signal is 10 dB¹. Thus, 10 dB should be reserved as a headroom to avoid saturating the RF components. As a result, a total suppression of minimum 50 dB should be obtained in the analog domain. Since a passive suppression of 20 dB is present due to the physical separation of the transmit and

¹The worst-case PAPR is actually significantly higher, but such rare peaks are already clipped in the digital domain at the transmitter.

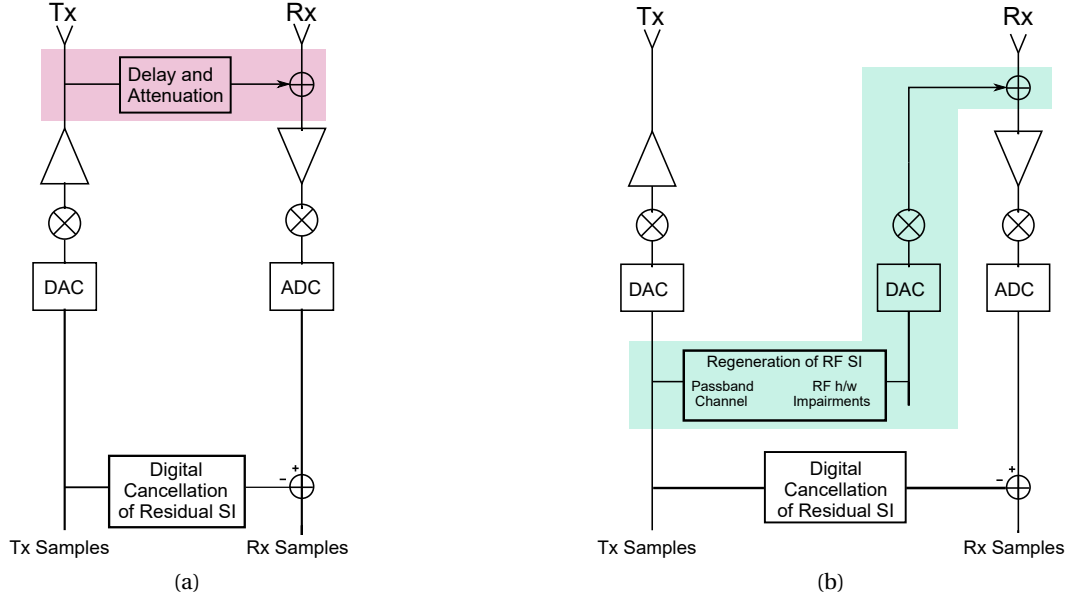


Figure 7.2 – Full duplex transceiver architectures: (a) RF circuit-based; (b) Regeneration-based from the digital baseband.

receive antennas, an additional amount of at least 30 dB of active analog suppression should be present to reduce the self-interference signal to such a power that would allow the signal of interest to be captured by the ADC.

7.1.1 Analog Cancellation

As seen in the power budget discussion of the above paragraph, to avoid saturating the receiver RF components and the ADC, a cancellation signal has to be generated and applied in the RF/analog domain [20, 174]. To this end, a number of techniques have been proposed which follow one of the two following full-duplex architectures.

The first technique uses a *circuit-based* architecture, where a replica of the transmitted signal is taken immediately before the transmit antenna, and is appropriately attenuated and phase-shifted via delay lines to produce the required RF cancellation signal, as illustrated in Fig. 7.2a. The RF cancellation signal is then subtracted from the received signal [16, 18, 175, 176]. One of the first implementations of this approach [16] achieved 40 dB of active analog cancellation, using tapped delay lines for a signal with 40 MHz bandwidth at a carrier frequency of 2.45 GHz. The main advantage of the circuit-based architecture is that all the transmitter non-idealities are inherently captured by the RF cancellation signal, since the cancellation signal is created by coupling in the transmitted signal at the very end of the transmitter chain. The main disadvantage of this architecture is that system specific RF cancellation circuits are needed to model the effect of the channel. Moreover, the scaling of this architecture to MIMO transceivers is difficult, since it requires a separate RF cancellation circuit for each of the M^2 transmitter-

receiver pairs, where M is the number of antennas [177, 2].

The second technique uses a *regeneration-based* architecture, where a separate RF chain is used to synthesize the necessary RF cancellation signal directly from the transmitted digital baseband samples [15, 17, 2], as illustrated in Fig. 7.2b. The baseband input signal to the cancellation chain is created by estimating the self-interference channel to obtain the attenuation, and delay parameters. This architecture was initially implemented on a WARP platform, and used a 40 MHz signal with transmit powers in the range 0 – 15 dBm. The reported analog self-interference cancellation values were in the range of 20 – 34 dB. Lower values of analog cancellation are generally obtained for higher transmit powers [2]. In another implementation of the same architecture, the National Instruments (NI) FlexRIO platform was used with a signal of 20 MHz bandwidth, and a transmit power of 4 dBm [178]. An active analog cancellation of 48 dB was reported. In both implementations, an additional digital cancellation step is included. The main advantage of the regeneration-based architecture is that it facilitates easier and more flexible matching of the cancellation chain to the transmitter chain for wideband signals. Specialized RF cancellation circuits are not required, and the architecture can be extended to full-duplex MIMO transceivers in a straightforward and comparatively low-complexity way. The main disadvantage of this architecture is that the transmitter hardware impairments are not inherently included in the cancellation signal. Moreover, since the cancellation signal uses an independent RF chain, additional hardware impairments are affecting the cancellation signal.

7.1.2 Digital Cancellation

Usually passive isolation and active analog cancellation are not sufficient to suppress the self-interference signal down to the noise floor. For this reason, an additional cancellation step is typically performed in the digital domain. After passive isolation and analog cancellation, the self-interference signal should be within the dynamic range of the ADC. In general, the received digital baseband signal y comprises the signal of interest s coming from a distant transmitter, a function $f(\cdot)$ of the baseband transmitted self-interference signal x and AWGN z , i.e., $y = s + f(x) + z$. The goal of the digital cancellation stage is to estimate and implement the function $f(\cdot)$ in order to remove $f(x)$ from y . In linear digital cancellation, $f(\cdot)$ is modeled as a convolution of the transmit signal with the self-interference channel. However, $f(\cdot)$ can be a complicated non-linear function due to the transceiver impairments [2]. Digital signal processing generally enables the implementation of more sophisticated algorithms to estimate and compensate for the function $f(\cdot)$ than signal processing in the analog domain. In the next section we will describe these transceiver impairments as well as the various attempts to model them that have been explored in the literature.

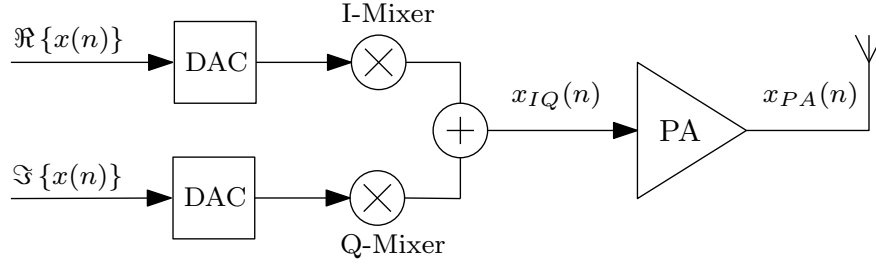


Figure 7.3 – Block diagram of the transmitter with the hardware components that introduce non-idealities.

7.2 Full-Duplex Transceiver Impairments

In this section, we will briefly present the most important transceiver hardware impairments that have been reported in the literature. In particular, these hardware impairments are digital-to-analog converter (DAC) non-linearities [2], phase noise [179, 180, 181, 182, 183], IQ imbalance [184, 185, 186, 187], and power amplifier (PA) non-linearities [187, 188, 175, 185]. In Fig. 7.3, we present a transmitter block diagram with all the important hardware components that introduce non-idealities into the self-interference signal.

7.2.1 DAC Non-Linearities

The authors in [2] observed even-order harmonics in the self-interference signal that can only occur in the analog baseband signal. The only components that can introduce non-linearities to the analog baseband signal are the two DACs for the real and imaginary components or the two ADCs. However, the authors conclude that it is mostly the DACs that introduce the observed non-linearities, since in their NI-5791R transceiver [189], the ADC [173] has a higher spurious-free dynamic range than the DAC [190]. The authors model the observed DAC non-linearities in the baseband representation of the passband signal using a Taylor series expansion.

7.2.2 Phase Noise

The baseband signal is upconverted to the carrier frequency f_c by mixing the baseband signal with a carrier signal. The carrier signal for the upconversion of the in-phase and quadrature components of the analog baseband signal is generated using a local oscillator, which is typically shared for the real and imaginary parts. The oscillator suffers from phase noise [191]. Thus, the carrier signal is not exactly $e^{j2\pi f_c t}$, but $e^{j(2\pi f_c t + \phi_{Tx}(t))}$, where $\phi_{Tx}(t)$ is the random phase noise process of the transmitter. Similarly, during downconversion the received signal is mixed with a carrier signal, which introduces a random receiver phase noise process $\phi_{Rx}(t)$. Generally, if the transmitter and the receiver use different local oscillators, the transmit and receive phase noise processes are independent. In full-duplex transceivers, the transmitter

and receiver chains are physically close, and thus, the local oscillator can be shared. It has been reported that sharing the oscillator can significantly reduce the effect of phase noise in full-duplex transceivers [181, 2]. The impact of phase noise on full-duplex transceivers has been extensively studied in the literature [179, 180, 181, 182, 183, 2].

7.2.3 IQ Imbalance

In an ideal scenario, the in-phase and quadrature components have a phase difference of $\pi/2$. IQ imbalance stems from the amplitude and phase mismatch between the in-phase and quadrature components of the upconverted and downconverted analog signal [192, 191]. In [184], the authors noticed the effect of IQ imbalance in the measurements of their full-duplex transceiver. To address this issue, widely linear digital filtering has been used for full-duplex transceivers which suffer from IQ imbalance [185]. Moreover, the joint digital cancellation of DAC non-linearities and IQ imbalance has been considered in [2].

7.2.4 Power Amplifier Non-Linearities

The power amplifier is the component that introduces the most severe non-linearities to the transmitted signal. This effect is particularly pronounced in high-power transmissions, where the power amplifier operates in its non-linear region where significant signal compression can occur. The power amplifier non-linearities are commonly modeled using a Taylor series expansion [191, 2]. Power amplifier non-linearities manifest only as odd-ordered harmonics of the complex-valued baseband signal, since the even-ordered harmonics lie out of band and are cut-off by the transmitter and receiver low-pass filters [2]. Models for power amplifier non-linearities in full-duplex systems have been extensively studied in the literature [175, 188, 193, 2, 194, 195].

7.3 State of the Art

The research scope of full-duplex technology is extremely broad. In this section we present the state-of-the-art research in field that are related to the full-duplex topics treated in this thesis. In particular, we present the state of the art in non-linear digital cancellation techniques, in full-duplex cognitive radio systems, and in resource allocation for full-duplex networks.

Non-linear Digital Cancellation Techniques

Due to the multiple transceiver non-idealities, complicated digital cancellation methods are required to suppress the residual self-interference down to the receiver noise floor. As mentioned previously, the most significant non-linearities are introduced by the transmitter power amplifier. In order to mitigate the non-linear distortions coming from the power amplifier, the digital baseband equivalent of the power amplifier transfer function should

be identified. The power amplifier stage can generally be modeled as a non-linear system with memory using the Volterra series [196, 197]. In systems where the signal bandwidth is small compared to the carrier frequency, the Volterra series can be simplified to a memory polynomial expansion [196, 198]. Many of the models for the power amplifier distortions that are used in full-duplex transceivers are based on polynomial expansions [175, 188, 193]. The parallel Hammerstein model is a special case of the Volterra series, which is commonly used to model power amplifier non-linearities. The known digital baseband transmit signal is passed through the estimated power amplifier model and subtracted from the received signal in the digital baseband. The distortion that originates from the IQ imbalance was modeled in the digital baseband using the complex conjugate of the original baseband signal [191] with memory coefficients in widely linear filters [185]. DAC non-linearities and IQ-imbalance were jointly modeled in [2], using a non-linear method to digitally cancel the combination of these two impairments. In [194] the authors used a parallel Hammerstein model that incorporates both power amplifier non-linearities and IQ imbalance. Polynomial models have a high complexity because the number of parameters that need to be estimated grows rapidly with the maximum chosen polynomial order. Since many terms of the parallel Hammerstein model are not significant, the authors in [194] use principal component analysis (PCA) to identify the most significant non-linear terms in the parallel Hammerstein model. However, the PCA operation needs to be rerun whenever the self-interference channel changes, leading thus again to high complexity. Recently, neural networks have been introduced in the literature as a promising method for non-linear digital self-interference cancellation with lower complexity than polynomial methods [195, 199, 200].

Full-duplex Cognitive Radio

In many situations the radio spectrum is densely allocated. However, it is not necessarily heavily occupied. Cognitive radio has been proposed to improve spectral efficiency by introducing opportunistic reuse of temporarily underused or unused parts of the frequency spectrum [24, 25]. Secondary users (SU) with cognitive radio capability must be able to rapidly detect the presence of primary users (PU) and switch to other unoccupied frequency bands. Half-duplex system architectures that periodically stop SU transmissions to sense the channel have been widely proposed to detect the start of PU transmissions [26, 27]. Previous research has examined the trade-off between sensing and throughput, and introduced scheduling algorithms to maximize sensing efficiency [201], as well as MAC-layer frame structures to maximize SU throughput while adequately protecting the PU [27]. The cumulative interference in large-scale cognitive radio networks has been examined and modeled in [202]. The joint impact of hardware imperfections in the performance of energy-detection-based sensing schemes has been analyzed in [203, 204]. A review on the research on hardware imperfections for half-duplex cognitive radio has been presented in [205].

An alternative approach to improve the detection-throughput trade-off, by using self-interference cancellation to enable concurrent transmission and sensing, has been proposed by a num-

ber of authors [28, 29, 30, 206, 31]. Previous research has considered the analysis of the sensing-throughput trade-off for energy detection [29], waveform detection [207], the power-throughput trade-off [31, 208, 209], cooperative spectrum sensing [210], and various adaptive algorithms for maximizing SU throughput [30, 206], in the presence of different levels of residual self-interference. In [29] and [30], the authors use energy-detection algorithms to identify primary user transmissions by assuming near-perfect cancellation of the self-interference (i.e., suppression below the thermal noise-floor).

The impact of hardware imperfections (in particular, IQ imbalance) in both cooperative and non-cooperative spectrum sensing scenarios for full-duplex cognitive radio has been analyzed in [211, 212]. The authors derive closed-form expressions for the probability of detection and the probability of false alarm for both perfect and imperfect self-interference cancellation at the SU. Moreover, in [213], the authors derive closed-form expressions for the probability of detection and the probability of false alarm in energy detection spectrum sensing for full-duplex nodes with residual self-interference in Nakagami-m fading channels.

A MAC protocol for non time-slotted full-duplex cognitive radio networks has been proposed in [214]. The authors in [215] examine the use of a distributed MAC protocol design with frame fragmentation in full-duplex cognitive radio networks. Power allocation for full-duplex relays in cognitive radio systems has been examined in [216]. A comprehensive review of the research on full-duplex cognitive radio until 2017 has been presented in [217].

In recent works of the last three years, the sensing-throughput trade-off as well as the energy efficiency of full-duplex cognitive radio have been studied in [218]. Moreover, throughput maximization under constraints on the energy efficiency and the PU protection has been formulated as a convex optimization problem in [219]. Analytical expressions for different sensing schemes (interweave, underlay, overlay, and hybrid) and a proposed adaptive scheme were derived in [220]. An adaptive scheme where SUs switch between full-duplex bi-directional transmission and sensing modes has been examined in [221]. The authors in [222] use a neural network to predict the PU activity, and switch between the two aforementioned modes. A full-duplex transceiver that uses its cognitive capabilities to detect jamming attacks has been presented in [223]. The authors in [224] examine multiple criteria and metrics under which full-duplex cognitive radios should be compared to half-duplex. Full-duplex MIMO cognitive radio has been examined in [225, 226]. The authors in [227] perform a performance analysis for energy detection spectrum sensing in full-duplex cognitive radio, considering a Rician distribution for the residual self-interference.

Resource Allocation in Full-duplex Networks

Resource allocation problems in the context of full-duplex radio have been extensively examined in the literature. In this subsection we will present the most pertinent examples of the research in this very broad field. In [228], resource allocation for half-duplex and full-duplex relay networks was examined, and dynamic resource allocation policies were proposed. Opti-

mal power allocation for full-duplex relays with residual self-interference under asymmetric fading conditions was examined in [229]. Power allocation to maximize the throughput of a full-duplex link under delay constraints was examined in [230]. Downlink and uplink power control was investigated in [231], and resource allocation in a multi-subcarrier aggregation model was examined in [232]. The authors in [233] investigate power allocation both for half-duplex and full-duplex cooperative non-orthogonal multiple access schemes. Moreover, the work in [234] examines the joint power and subcarrier allocation in full-duplex multicarrier non-orthogonal multiple access schemes. The authors of [235] and [236] study issues related to spectrum and power allocation in very dense full-duplex networks and power allocation for full-duplex relays in the absence of exact channel state information (CSI), respectively.

7.4 Summary and Discussion

In this chapter, we discussed the most important features of the in-band full-duplex technology. In particular, in Section 7.1 we presented a description of the analog and digital self-interference cancellation stages. Moreover, we presented the two main different full-duplex transceiver architectures that have been proposed in the literature. In Section 7.2, we introduced the different transceiver impairments that affect the self-interference signal, along with the state of the art in the modeling of these impairments. Finally, in Section 7.3, we discussed the state of the art in the following topics: first, the non-linear digital cancellation techniques that incorporate and subtract the effect of hardware-induced non-linearities. Second, the resource allocation in contemporary full-duplex networks, and third, one of the most important applications of full-duplex technology, namely cognitive radio. These topics summarize the state of the art for our contributions in Chapter 8 and Chapter 9.

We saw that the regeneration-based analog cancellation has advantages over the circuit-based cancellation, regarding the RF hardware complexity and the scalability. However, with the regeneration-based cancellation, it is difficult to subtract the strong non-linearities of the self-interference signal in the analog domain, leading to over-provisioning the ADC, resulting thus, in reduced receiver sensitivity. For this reason, in the next chapter, we will propose a simple scheme to improve the non-linear cancellation capabilities of the regeneration-based full-duplex architecture. Moreover, the state of the art in full-duplex research shows that the residual self-interference has a high impact in many full-duplex applications, and can not be simply neglected. Nevertheless, residual self-interference is not catastrophic to every system that incorporates full-duplex transceivers. Some types of full-duplex systems are inherently more resilient to the impact of residual self-interference. In Chapter 9 we are going to elaborate on this aspect.

8 Digital Predistortion of Hardware Impairments for Full-Duplex Transceivers

Regeneration-based full-duplex transceivers have practical advantages over transceivers that derive the self-interference cancellation signal from the transmitted RF signal in the analog/RF domain. These advantages include a reduced physical size, the ability to more rapidly adapt to changing channel conditions and better scalability for multiple-input multiple-output (MIMO) technology. However, as the cancellation signal is synthesized from the digital-baseband, regeneration-based transceivers are also more prone to the effects of transmitter and receiver hardware impairments. Straightforward analog/RF cancellation can only suppress components of the self-interference that are *linearly* proportional to delayed copies of the baseband transmitted signal. Suppression of non-linear effects requires reconstruction of these non-linearities for the analog cancellation signal. Unfortunately, this reconstruction turns out to be extremely difficult since the RF components in the cancellation chain also introduce additional (and independent) non-linear signal components. At high transmit power levels, the non-linear components in the residual (after analog suppression) can have significant power. The high power of this residual often requires the sensitivity of the receiver to be reduced, thereby decreasing the effective signal-to-noise ratio (SNR).

Given the limited analog suppression that can be practically achieved, existing regeneration-based (and circuit-based) full-duplex systems use additional digital cancellation stages to remove the non-linear components in the residual [18, 17, 178]. Unfortunately, accurate reconstruction of these components requires a very complex digital cancellation stage [2]. In particular, the self-interference signal becomes very difficult to model, as it contains the cascaded effect of multiple hardware impairments, e.g., baseband non-linearities introduced by the DACs, IQ imbalance arising from the mixers, sampling jitter, phase noise, and memory effects of the non-linear power amplifiers [2, 21, 193, 187]. While capturing the impact of all of these components requires high-order polynomials, it is important to avoid over-fitting the cancellation model.

As an alternative to the reconstruction of non-linearities, previous research in our group has examined the use of digital predistortion in full-duplex transceivers to compensate for power amplifier memory effects [21]. The basic idea behind predistortion is to distort the trans-

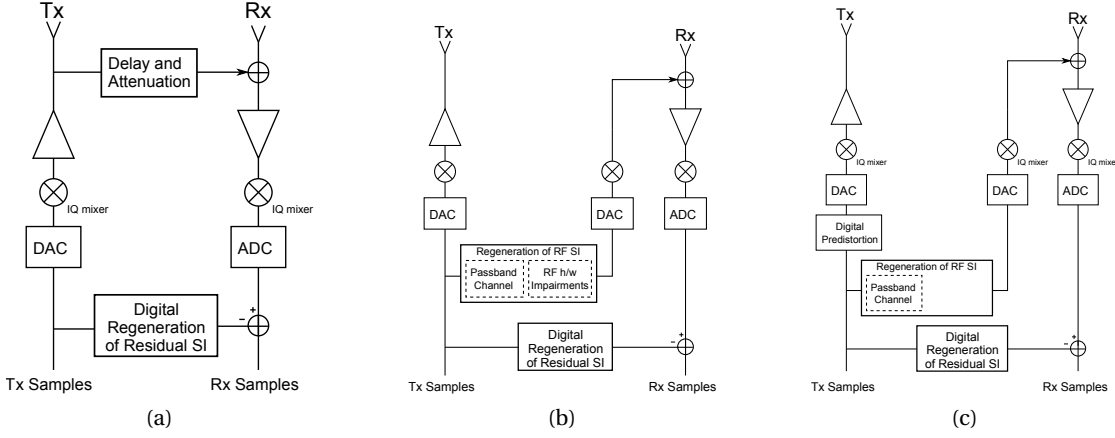


Figure 8.1 – Full duplex system architectures: (a) RF cancellation circuit; (b) regeneration from the digital baseband; and (c) proposed regeneration with digital predistortion for hardware impairments.

mit signal in the digital baseband to compensate for the non-linear distortion and memory effects introduced by the baseband and RF hardware impairments. In this case an overall 13 dB improvement in the suppression was observed (compared to the case without predistortion). However, this improvement was achieved only in the digital stage, i.e., no additional suppression in the analog stage was observed.

In this chapter, we propose to use predistortion together with signal regeneration to avoid and remove a significant amount of non-linear self-interference components in a full-duplex transceiver already in the RF domain. Since the predistorted RF self-interference signal is (ideally) only a linear copy of the baseband transmit signal, a predistorted analog cancellation stage *alone* is sufficient to remove this self-interference. In addition to the introduction of the concept of predistortion based self-interference cancellation, we also extend the power amplifier predistortion basis functions developed in [21] to incorporate the cascade of IQ imbalance, digital-to-analog converter (DAC) non-linearity, and power amplifier memory effects. We validate both our proposed predistortion-based full-duplex cancellation architecture and the extended non-linearity model on our wideband full-duplex hardware testbed.

In particular, in Section 8.1, we briefly describe the predistortion based regeneration architecture for self-interference cancellation in full-duplex radios, and we outline mathematical models for the various hardware impairments present in a typical regeneration-based full-duplex transceiver. In Section 8.2, we present our full-duplex testbed, which we use to experimentally evaluate the self-interference suppression performance of the theoretical models.

8.1 Application of Digital Predistortion to Full-Duplex Transceivers

Fig. 8.1(a) shows the block diagram of a full-duplex transceiver, where the analog-self-interference signal is generated via a dedicated RF circuit. In comparison, Fig. 8.1(b) shows a regeneration architecture, where the cancellation signal is regenerated from the digital baseband using a separate RF chain. The self-interference channel is estimated using a training frame. The response is used to compute the coefficients of the finite-impulse response (FIR) filter that produces the required cancellation signal [17, 2]. Only few hardware impairments can be included, as this process only suppresses linear signal components.

8.1.1 Characterization of Hardware Impairments in Full-Duplex Transceivers

The state of the art in modeling these hardware impairments was presented in Section 7.2 and Section 7.3 on a high level. In this section we go into more details of the digital baseband models for the various hardware impairments in the full-duplex transmitter depicted in Fig. 8.1(b). The baseband transmit signal $x[n]$ is comprised of a real and an imaginary component, which are up-converted using separate DACs and mixers. The DACs introduce baseband non-linearities, which manifest as both odd-ordered *and* even-ordered harmonics of the complex-valued baseband signal [2]. As such, the DAC non-linearities can be modeled using a Taylor series expansion

$$x_{\text{DAC}}[n] = \sum_{p=1}^{P_{\max}} a_p \Re\{x[n]\}^p + j \sum_{p=1}^{P_{\max}} b_p \Im\{x[n]\}^p, \quad (8.1)$$

where $\Re\{x[n]\}$ and $\Im\{x[n]\}$ are the real and imaginary components of the baseband transmit signal, respectively. The coefficients a_p and b_p are the coefficients for the in-phase and quadrature DACs respectively, and the expansion is truncated after P_{\max} terms. In the baseband the impact of this imbalance can be modelled as an additional complex conjugate term [191, pp. 71–79]

$$x_{\text{IQ}}[n] = \alpha x_{\text{DAC}}[n] + \beta x_{\text{DAC}}^*[n], \quad (8.2)$$

where α and β are coefficients that can be derived from the amplitude and phase difference between the in-phase and quadrature mixers [191, pp. 72–73]. In this case, for an arbitrary non-linearity with input $x[n]$, the output signal $y[n]$ can be expressed by

$$y[n] = \sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{M}} a_{km} x_{\text{IQ}}(n-m) |x_{\text{IQ}}(n-m)|^{k-1}, \quad (8.3)$$

where a_{km} are the power amplifier coefficients, and \mathcal{K} and \mathcal{M} are the sets of polynomial orders and delays, respectively [196]. For a typical RF power amplifier only odd ordered polynomial terms are included, as even ordered harmonics of the RF signal fall out of band and are removed by filters. The basis functions in (8.3) have the form $x[n] |x[n]|^{k-1}$ to ensure that the phase information in $x[n]$ is preserved.

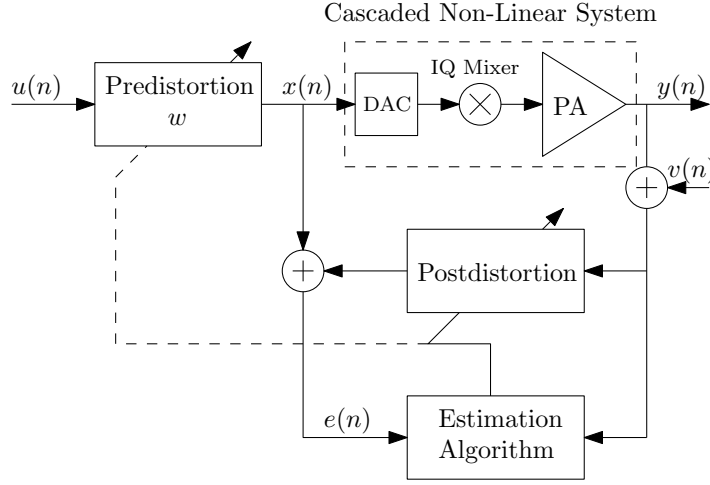


Figure 8.2 – Block diagram showing the application of predistortion to compensate for the cascaded non-idealities of the transmitter hardware impairments.

8.1.2 Predistortion for All Hardware Impairments

In the ideal case, when a baseband predistorted signal is applied to the non-linear hardware, the RF output will be free from non-linear distortion. Accordingly, for a full-duplex transceiver, predistortion of the transmitted signal would thus considerably simplify the design and implementation of the cancellation algorithms at both the analog and digital stages, as shown in Fig. 8.1(c).

Predistorting for the hardware impairments requires the inversion of (8.3). Unfortunately, explicitly inverting (8.3) is complicated. However, following the approach of [196] and [197] for signals where the bandwidth is small compared to the carrier frequency, we can use (8.3) to model the inverse of the non-linear system directly, i.e., the output is used to predict the input. This method is termed *postdistortion*. The postdistortion coefficients in (8.3) are estimated by sending a training frame and applying least-squares estimation to the training and the received signal. Following convergence, these coefficients are copied into the predistortion stage and the system is run in open-loop configuration [196], as shown in Fig. 8.2.

Cascading the hardware impairment models for the DAC non-linearities, IQ imbalance, and the power amplifier non-linearities with memory effects leads to a transmit signal

$$y[n] = \sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{M}} a_{km} [\alpha x_{\text{DAC}}(n-m) + \beta x_{\text{DAC}}^*(n-m)] \cdot \left| \alpha x_{\text{DAC}}(n-m) + \beta x_{\text{DAC}}^*(n-m) \right|^{k-1}, \quad (8.4)$$

where $x_{\text{DAC}}[n]$ includes the DAC non-linearities given by (8.1). It is difficult to directly estimate the coefficients in (8.4) using least-squares estimation. However, (8.4) includes many more degrees of freedom available to tune the model than the actual physical system. This fact increases the amount of training needed, reduces the quality of the estimates, and worsens the

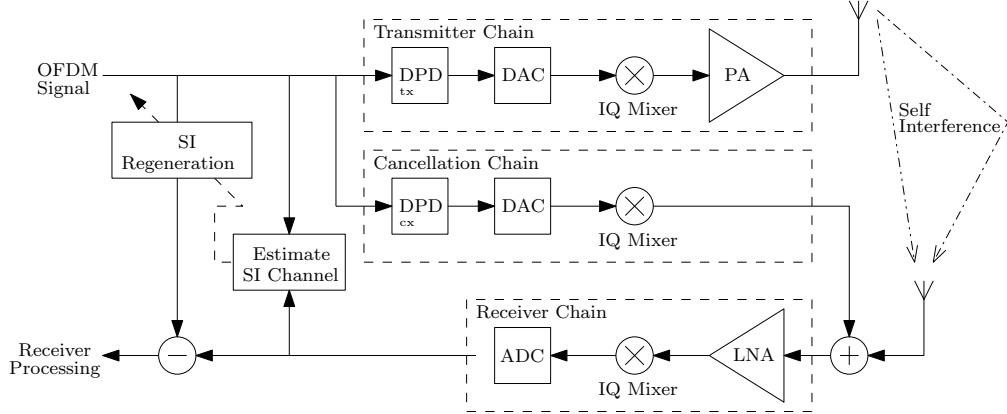


Figure 8.3 – Application of the proposed predistortion method in regeneration-based full-duplex transceivers.

condition of the problem. An attempt to overcome the above issues may lead to over-fitting. Thus, a reduced basis for the model is desirable. We observe that the expanded form in (8.4) mainly consists of real and imaginary parts of the signal that are independently delayed and raised to powers. Consequently, (8.4) can be expanded and expressed with a different basis, in which the real and imaginary parts of the signal are separated, as follows

$$y[n] = \sum_{g \in \mathcal{G}} \sum_{f \in \mathcal{F}} \sum_{h \in \mathcal{H}} \sum_{l \in \mathcal{L}} b_{gfhl} \Re(x(n-h))^f \Im(x(n-l))^g, \quad (8.5)$$

where b_{gfhl} are the coefficients, and \mathcal{F} and \mathcal{G} are the set of polynomial orders for the real and imaginary parts, and \mathcal{H} and \mathcal{L} are the corresponding sets of delay terms respectively. In this form, the required set of post-distortion coefficients, b_{gfhl} , can be computed in a straightforward manner using least-squares estimation.

8.1.3 Application to Regeneration-Based Full-Duplex Transceivers

Since the regeneration-based full-duplex architecture includes independent transmitter and cancellation chains, predistortion needs to be applied in both of them, as shown in Fig. 8.3. In particular, to compute the postdistortion coefficients for this chain, we first send a training frame in the transmitter chain and apply least-squares estimation, and we iterate. The exact same procedure is followed for the cancellation chain to obtain the postdistortion coefficients. The two sets of coefficients are then copied into the predistortion stage of the transmitter and cancellation chains. We note that the impact of the self-interference channel is inherently captured by the described predistortion method.

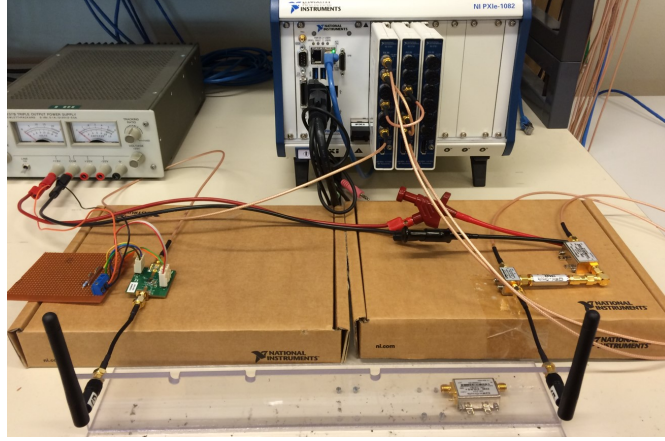


Figure 8.4 – Full-duplex testbed used to obtain the experimental results.

8.2 Experimental Validation

Our hardware testbed is based on the National Instruments PXI platform with an NI FlexRIO PXIe-1082 chassis [237] and two NI-5791 RF transceiver modules [189] operating in the 2.4 GHz ISM band [2, 21]. Baseband signal processing is performed in Matlab, and the NI-5791 modules are configured using LabVIEW. An external 30 dB gain amplifier (Skyworks SE2576L) is used to increase the average transmitted power to approximately 10 dBm. The transmit and receive antennas are rubber-duck monopoles, and are placed 25 cm apart with the same polarization, resulting in 25 dB of passive suppression. As shown in Fig. 8.1(b) and Fig. 8.1(c), the cancellation signal is injected to the received signal between the antennas and the receiver chain using an RF combiner. The testbed is shown in Fig. 8.4. The transmitted signal used to estimate the predistortion coefficients consists of a frame containing 20 2048-tone OFDM symbols, where each sub-carrier is modulated with 64-QAM representing a random bit-stream. We note that convergence is quick, and we can obtain the coefficients with just two iterations per chain. Frame synchronization symbols or pilot tones are not included. The OFDM signal has a bandwidth of 20 MHz, but the signals are sampled at 60 MHz to capture the out-of-band emissions. The carrier frequency is 2.48 GHz, and the local oscillator is shared between the transmit, receive, and cancellation chains to reduce the impact of phase noise.

Fig. 8.5a shows an experimental measurement of the power spectral density recorded on the full-duplex testbed when no digital predistortion is applied, i.e., using the architecture depicted in Fig. 8.1b. In this case the analog cancellation signal only contains linear time-delayed components. Significant out-of-band emissions are observed in the transmitted signal, and these remain in the residual self-interference after both the analog and digital suppression stages. In particular, the residual self interference after analog suppression is approximately 23 dB above the measured noise-floor. The digital suppression stage (using the model outlined in [2], which only accounts for a limited number of coupled IQ and non-linear terms) only improves the suppression by an additional 2 dB. The noise-floor is measured when transmitting an empty, i.e., all zero, frame and thus includes also the thermal noise arising

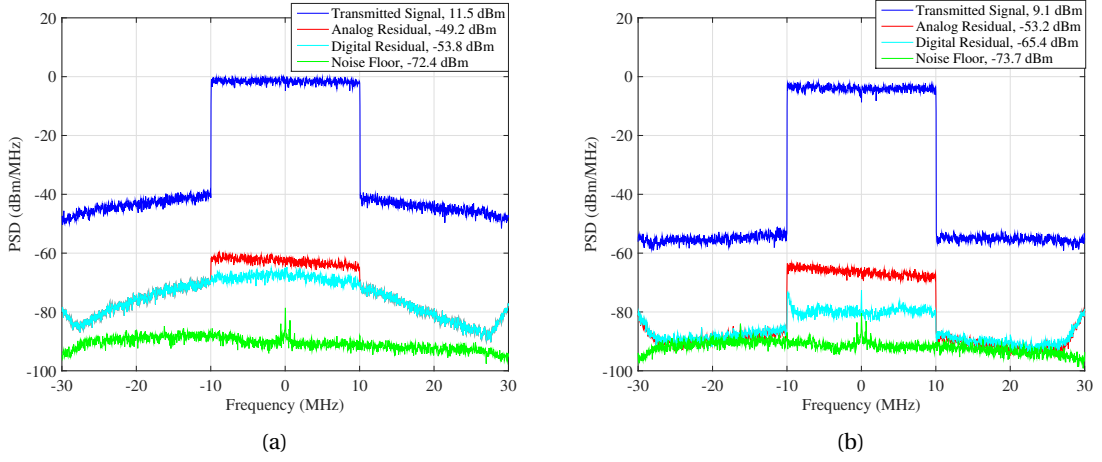


Figure 8.5 – Experimentally measured self-interference power spectral density with: (a) no digital predistortion; (b) and predistortion only for the power amplifier.

from the cancellation chain.

Fig. 8.5b shows the power spectral density when predistorting only for the power amplifier non-linearities using (8.3), with the architecture depicted in Fig. 8.1c. This result is based on the method of [21]. A comparison with Fig. 8.5a shows the out-of-band emissions in the analog residual component and the transmitted signal are significantly reduced. Unfortunately, the analog residual is still approximately 21 dB above the noise-floor. However, the digital suppression stage, with the same model applied in Fig. 8.5(a), can further reduce the self-interference by approximately 12 dB.

While predistortion for the power amplifier improves the overall suppression, this is only achieved in the digital stage. The high analog residual measured in Fig. 8.5b therefore reduces the receiver sensitivity. In contrast, as shown in Fig. 8.6, predistortion for all significant hardware impairments, using (8.5), improves the performance of the analog suppression stage by 10 dB, compared to only predistorting for the power amplifier. The analog residual is thus only approximately 10 dB above the measured noise-floor, and importantly, this suppression is achieved in the analog domain, *before* the signal reaches the receiver. In this case the digital suppression stage does not further reduce the self-interference.

8.3 Discussion and Summary

In this chapter we explored the use of digital predistortion in full-duplex transceivers to compensate for the non-linear distortions introduced by the transmitter and the cancellation chains. In particular, in Section 8.1.1, we described digital baseband models for the various transceiver impairments. In Section 8.1.2, we introduced a predistortion model for the cascade of impairments, and in Section 8.1.3 we described how it can be used in the transmitter and

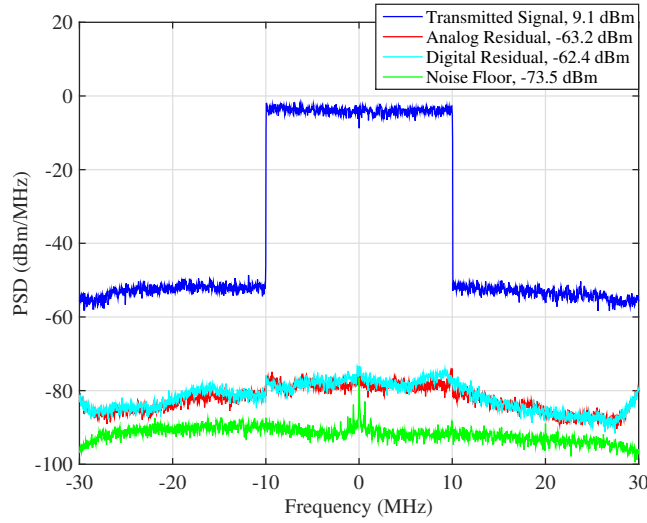


Figure 8.6 – Experimentally measured self-interference power spectral density with predistortion of all significant hardware impairments.

cancellation chains of regeneration-based full-duplex transceivers. Finally, in Section 8.2, we experimentally evaluated the performance of the proposed predistortion method using our full-duplex testbed.

Regeneration-based full-duplex transceivers are prone to the effects of hardware impairments, which introduce significant non-linear components in the self-interference signal, particularly at high transmit power. Digital predistortion for hardware impairments in the transceiver chain, e.g., mixer IQ-imbalance, DAC non-linearities, and power amplifier non-linear memory effects, effectively linearizes the self-interference signal and allows for increased suppression. Results on a hardware testbed operating at 2.48 GHz shows that predistortion for all significant hardware impairments in a full-duplex transceiver can increase the analog suppression by an additional 14 dB for a 20 MHz OFDM signal, compared to a conventional regeneration architecture with no predistortion.

9 Full-Duplex Systems with Imperfect Self-Interference Cancellation

In the previous chapter we saw that in our regeneration-based full-duplex transceiver, the measured residual self-interference was approximately 10 dB above the noise-floor, even using digital predistortion to counteract the transmitter chain non-linearities. More generally, the low cost hardware of full-duplex user nodes inevitably results in significant amount of residual self-interference, which translates into a significant SNR degradation. Therefore, using in-band full-duplex transceivers in a symmetric bi-directional link, i.e., to double the spectral efficiency, limits the scope of bi-directional full-duplex applications only to high-end full-duplex devices.

However, the scalability of contemporary wireless communications systems, along with the reduction of their end-to-end latency to less than 1 ms, as required by 5G, is a very cumbersome task. We argue that the adoption of full-duplex technology in such systems can still be helpful for the improvement of the total network throughput and the reduction of the end-to-end latency. In this chapter, we propose to use full-duplex in cognitive radio systems and bi-directional systems with highly asymmetric capacity requirements. In cognitive radio systems, the goal of a secondary user (SU) is only to *detect* the presence of a primary user (PU), and not to correctly decode its transmissions. In this case, reliable spectrum-sensing performance may be achieved at a lower SNR. Similarly, in highly asymmetric links, where the downlink requires considerably more capacity than the uplink (or vice-versa), the task of the receiver of the low-capacity link is to decode a very simple message. Thus, systems with inherent asymmetric communication requirements may benefit from the adoption of full-duplex technology, since the impact of residual self-interference can be compensated by increasing the coding rate or the spreading gain of the low-rate link.

In Section 9.1 of this chapter, we perform an analysis of the latency-throughput trade-off in full-duplex cognitive radios. We derive analytical expressions for the average access latency of half-duplex and existing full-duplex cognitive radios, and we propose a sliding-window method to reduce the access latency of full-duplex cognitive radios. We also present our proof-of-concept software-defined radio (SDR) implementation of a full-duplex cognitive radio system. In Section 9.2, we propose the use of full-duplex technology in asymmetric links,

and we show the conditions under which this approach leads to benefits for the capacity and the power consumption of the system.

9.1 Sliding Window Spectrum Sensing for Full-Duplex Cognitive Radios

Standard cognitive radio systems need to periodically stop secondary user (SU) transmissions to sense the channel in order to detect the presence of primary users (PUs) and promptly vacate the channel. This overhead reduces the throughput of the SUs, since they need to periodically stop their transmissions to sense the channel. Furthermore, while transmitting, the SU is unable to detect the start of a PU transmission until the next sensing slot (at the earliest). Decreasing the interval between successive sensing slots decreases the efficiency and throughput of the SU system, but improves its spectrum sensing capabilities [27]. Recently, full-duplex cognitive radio has been proposed in the literature as a possible solution to the aforementioned problems [28, 29, 30, 206, 31].

In existing work on cognitive radio, the PU is assumed to be protected if the SU has a probability of PU detection, P_d , that is sufficiently high (typically above 90%) [26, 206]. However, from a physical layer perspective, the *access latency* (defined as the time required by the SU to detect the PU and vacate the channel) gives a better measure of the impact of SU interference and the necessary protection. For example, a high access latency may harm PU communication, by distorting a too large portion of its training or synchronization fields, which are often found at the beginning of packets. Even though full-duplex cognitive radio systems reduce the access latency, they still take decisions after a pre-set number of samples [29] and [31]. Therefore, these schemes may still have a non-negligible access latency, as the PU may start transmitting at any time.

In this section we focus on obtaining analytical expressions for the physical access latency in various cognitive radio systems. The access latency results are presented in terms of the number of samples and can thus be scaled and applied to arbitrary hardware implementations. Furthermore, to alleviate the issue of high access latency in full-duplex cognitive radio systems, we introduce a full-duplex *sliding-window* spectrum sensing technique. Unlike existing schemes, our approach takes decisions on a sample-by-sample basis, and can detect the presence of PUs more quickly, thereby reducing the access latency as demonstrated by our simulation results.

This section is organized as follows. In Section 9.1.1 we describe the system models for the existing half-duplex and slotted full-duplex systems, along with the model of our proposed sliding-window full-duplex system. In Section 9.1.2 we analyze the throughput-latency trade-off, and in Section 9.1.3 we derive expressions for the access latency. In Section 9.1.4 we present theoretical and numerical simulation results.

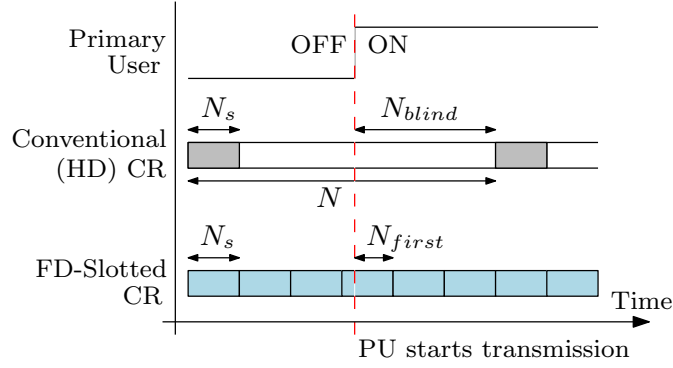


Figure 9.1 – Frame structures for half-duplex and full-duplex cognitive radio systems.

9.1.1 Half-Duplex and Full-Duplex Cognitive Radio System Models

In this section we outline the three cognitive radio system models considered in this work. For all three approaches, the presence of the PU is detected by comparing a decision metric, M , (computed from a set of N_s samples), against a threshold, ϵ . If $M > \epsilon$, the PU is assumed to be present, otherwise the channel is declared idle. A widely used decision-metric for spectrum sensing is energy detection [26], with

$$M = \frac{1}{N_s} \sum_{n=1}^{N_s} |r(n)|^2, \quad (9.1)$$

where $r(n)$ is the received signal. Other metrics to detect the presence of the PU are not considered here, however, the findings can be readily extended.

Based on the decision metric, the probabilities of detection, P_d , and false-alarm, P_f , can be found by applying appropriate hypothesis tests [26, 27]. Denoting \mathcal{H}_0 as the scenario in which the PU is inactive, and \mathcal{H}_1 where the PU is active, the probabilities of false alarm and detection are defined as

$$P_f = \Pr(M > \epsilon | \mathcal{H}_0) \quad (9.2)$$

$$P_d = \Pr(M > \epsilon | \mathcal{H}_1). \quad (9.3)$$

The probability of misdetection, P_m is defined as

$$P_m = \Pr(M \leq \epsilon | \mathcal{H}_1) = 1 - P_d. \quad (9.4)$$

Half-Duplex Cognitive Radio

Fig. 9.1 shows the SU transmission frame consisting of N samples for a conventional half-duplex cognitive radio system [27]. The frame consists of two parts: an N_s sample sensing window and an $N - N_s$ sample transmission window. The N_s samples are used to decide if

the PU is present based on the decision metric, e.g., (9.1), with transmission continuing if the PU is not detected. Decisions are thus made every N samples, i.e., once per SU frame. It is important to note that the SU is unable to detect or react to the presence of the PU while transmitting, which leads to a high access latency. Due to the sensing window, the throughput of the SU system is also reduced by a factor of $\frac{N-N_s}{N}$.

Concurrent Sensing and Transmission With Slotted Sensing

By cancelling the self-interference signal, SUs can concurrently transmit and sense the channel. Fig. 9.1 shows the frame structure of a concurrent sensing and transmission system using self-interference cancellation as proposed in [207] and [31]. The systems analysed in [207] retain the same frame structure as a conventional CR system by including a sensing-only slot at the start of each frame. In this work, we consider full-duplex systems that only include concurrent transmission and sensing slots, similar to the approach taken in [31, 208]. The SU throughput is not reduced as there is no dedicated sensing-slot. Similarly, the SU may be able to more quickly detect whether the PU starts transmitting during the frame, thereby reducing the access latency. While decisions are made more frequently (every N_s samples) than in a conventional cognitive radio system (every N samples), there still remains a blind interval of N_{first} samples. Residual self-interference also makes it more difficult to detect the presence of the PU, by effectively increasing the noise-floor. However, as the detection decisions are made more frequently compared to the half-duplex case, lower values of the P_d may still result in an acceptable access latency for the PU.

Concurrent Sensing and Transmission with Sliding Window Sensing

We propose an extension to the concurrent sensing and transmission scheme outlined in the previous paragraph by introducing a sliding-window to take decisions at every sample, i.e., the buffer does not wait to fill with N_s fresh samples before a decision is made. This approach can be implemented easily in digital hardware via a FIFO buffer. It is important to note that successive decision metrics are *not* independent as only one new sample is added (and one removed). However, the minimum access latency of this scheme is one sample, unlike the slotted full-duplex method. In our analysis it is assumed that the number of sensing samples, N_s , remains the same as in the slotted full-duplex model.

9.1.2 Latency-Throughput Analysis

Throughput of the Half-Duplex and Full-Duplex Systems

The capacity of the SU (assuming the PU is not transmitting) is given by

$$C_0 = \log_2 (1 + \text{SNR}_{\text{SU}}), \quad (9.5)$$

9.1. Sliding Window Spectrum Sensing for Full-Duplex Cognitive Radios

where SNR_{SU} is the signal-to-noise ratio of the SU measured at a receiver node. This capacity is achieved when the PU is not transmitting *and* the SU has not raised a false alarm.

Following [27, 31], we express the total throughput for the SU system with full-duplex sensing capability as

$$R_{\text{FD}} = C_0(1 - P_f). \quad (9.6)$$

For a specific SNR of the SU, the throughput is a parameter of only P_f . In practice, if a false-alarm occurs, the loss in SU throughput can be significant as an entire data frame may be lost. However, we consider that if a false alarm occurs during the transmission of a frame, the interrupted frame is retransmitted in the next attempt of the SU. This way, intermittent SU outages are avoided.

The half-duplex system has an additional throughput loss due to the sensing overhead

$$R_{\text{HD}} = \frac{N - N_s}{N} C_0(1 - P_f). \quad (9.7)$$

Clearly for the half-duplex system there exists a trade-off between PU detection latency and SU throughput that is determined by the sensing overhead. For full-duplex systems, the main consideration is the increase of P_f , due to the residual self-interference. In order for a full-duplex system with residual self-interference to maintain a throughput close to C_0 , it has to operate with lower P_d values. In the results we show that the sliding window model allows for lower access latency compared to the slotted full-duplex and half-duplex cases, while maintaining a high SU throughput.

Average Latency of the Half-Duplex System

Let D_k denote the event that the PU is detected during the k -th decision after starting transmission, and D_k^c the complementary event that the PU is not detected during the k -th decision. Let also N_k be the number of samples from when the PU starts to the k -th decision point. The average access latency for all the schemes can thus be computed from the infinite sum

$$L = N_1 P(D_1) + N_2 P(D_2 | D_1^c) + N_3 P(D_3 | D_1^c \cap D_2^c) + \dots + N_i P(D_i | D_1^c \cap D_2^c \dots \cap D_{i-1}^c) + \dots \quad (9.8)$$

For the half-duplex case, there are two possible scenarios, depending on which part of the SU activity (sensing or transmitting) the PU starts transmitting. In the first scenario the PU starts transmitting during the blind interval, N_{blind} samples before the end of the blind interval. The first decision is made after $(N_{\text{blind}} + N_s)$ samples, i.e., $N_1 = (N_{\text{blind}} + N_s)$. Let $P_d(k)$ denote the probability of detection with k samples of the PU signal and $P_m(k)$ denote the probability of misdetection for the same case; the probability of detecting the PU at the first decision is thus $P_d(N_s)$. If the PU is not detected, the number of samples for the second decision will be

$N_2 = (N_{\text{blind}} + N_s + N)$, again with probability of detection $P_d(N_s)$, and so on. Thus

$$\begin{aligned} L(N_{\text{blind}}) &= P_d(N_s) \sum_{n=0}^{\infty} (N_{\text{blind}} + N_s + nN) P_m^n(N_s) \\ &= P_d(N_s) \left[(N_{\text{blind}} + N_s) \frac{1}{P_d(N_s)} + N \frac{P_m(N_s)}{P_d^2(N_s)} \right], \end{aligned} \quad (9.9)$$

for $0 < P_d(N_s) < 1$.

N_{blind} can take each value between 1 and $N - N_s$ with probability $\frac{1}{N}$, thus the average latency under scenario 1 is

$$L_1 = \frac{1}{N} \sum_{N_{\text{blind}}=1}^{N-N_s} L(N_{\text{blind}}). \quad (9.10)$$

In the second scenario the PU starts transmitting during the sensing period, N_{first} samples before the end of the sensing period. The first decision is taken after $N_1 = N_{\text{first}}$ samples, with probability of detection $P_d(N_{\text{first}})$. If the PU is not detected, the number of samples for the second decision is $N_2 = (N_{\text{first}} + N)$, with probability of detection $P_d(N_s)$, and so on. Thus

$$\begin{aligned} L(N_{\text{first}}) &= N_{\text{first}} P_d(N_{\text{first}}) + P_d(N_s) P_m(N_{\text{first}}) \sum_{n=0}^{\infty} [N_{\text{first}} + (n+1)N] P_m^n(N_s) \\ &= N_{\text{first}} P_d(N_{\text{first}}) + P_d(N_s) P_m(N_{\text{first}}) \left[N_{\text{first}} \frac{1}{P_d(N_s)} + N \frac{1}{P_d^2(N_s)} \right] \end{aligned} \quad (9.11)$$

for $0 < P_d(N_s) < 1$.

As N_{first} can take any value between 1 and N_s , with probability $\frac{1}{N}$, the average latency under scenario 2 is

$$L_2 = \frac{1}{N} \sum_{N_{\text{first}}=1}^{N_s} L(N_{\text{first}}). \quad (9.12)$$

Therefore, the total average latency of the half-duplex system is $L = L_1 + L_2$.

9.1.3 Average Latency of the Slotted Full-Duplex System

In the slotted full-duplex scheme there is no blind interval. Let us thus consider the case where the PU turns on N_{first} samples before the end of the sensing period. The first decision is made after $N_1 = N_{\text{first}}$ samples, with probability of detection $P_d(N_{\text{first}})$. If the PU is not detected, the second decision has latency of $N_2 = (N_{\text{first}} + N_s)$, with $P_d(N_s)$, and so on. Thus the latency is

$$\begin{aligned} L(N_{\text{first}}) &= N_{\text{first}} P_d(N_{\text{first}}) + P_d(N_s) P_m(N_{\text{first}}) \sum_{n=0}^{\infty} [N_{\text{first}} + (n+1)N_s] P_m^n(N_s) \\ &= N_{\text{first}} P_d(N_{\text{first}}) + P_d(N_s) P_m(N_{\text{first}}) \left[N_{\text{first}} \frac{1}{P_d(N_s)} + N_s \frac{1}{P_d^2(N_s)} \right], \end{aligned} \quad (9.13)$$

for $0 < P_d(N_s) < 1$.

N_{first} can take any value between 1 and N_s with probability $\frac{1}{N_s}$, which leads to an average latency of

$$L = \frac{1}{N_s} \sum_{N_{\text{first}}=1}^{N_s} L(N_{\text{first}}). \quad (9.14)$$

Average Latency of the Sliding Full-Duplex System

In the sliding full-duplex model, the average latency can be derived from (9.8), however the difference in this case is that the decisions are not independent. A decision taken at a sample i is not independent of decisions taken over the previous $i + (N_s - 1)$ samples. Only decisions separated by N_s samples are independent. Accordingly, the slotted model can be regarded as a special case of the sliding model, where the only decisions kept are those separated by N_s samples. The idea of the sliding approach is that there is no real reason to discard all the decisions in between, as these will potentially reduce the access latency. It should also be noted that taking decisions every sample is feasible in contemporary implementations. However, the effect on the energy consumption, which may be a limitation for battery-powered SUs, remains to be investigated.

Since each decision is not independent of the N_s previous decisions, the conditional terms of (9.8) cannot be easily expanded using multiplicative terms as in (9.11) and (9.13). The analytical expression for the average access latency in this case would require the conditional terms to be expressed using a stochastic process model with memory. To compare the performance of our proposed sliding-window method with the existing schemes, we have used Monte-Carlo simulations which maintain the dependency between the decisions.

9.1.4 Results

In the following, we present results regarding the access latency and the throughput of half-duplex and slotted full-duplex cognitive radio systems, as well as for the proposed sliding-window full-duplex system.

In Fig. 9.2 we observe a significant improvement in the average latency of the full-duplex schemes compared to the conventional half-duplex scheme. Moreover an improvement is observed for the sliding full-duplex model compared to the slotted scheme. For these results, no residual self-interference is present, i.e., perfect self-interference suppression is assumed. The minimum average access latency of the half-duplex scheme is approximately half the length of a SU frame, as the system is unable to sense while transmitting. The slotted full-duplex system has lower access latency—as decisions are made every N_s samples—and the minimum average latency is approximately $\frac{N_s}{2}$. The sliding-window full-duplex scheme can potentially detect the PU with an average access latency of a single sample, however,

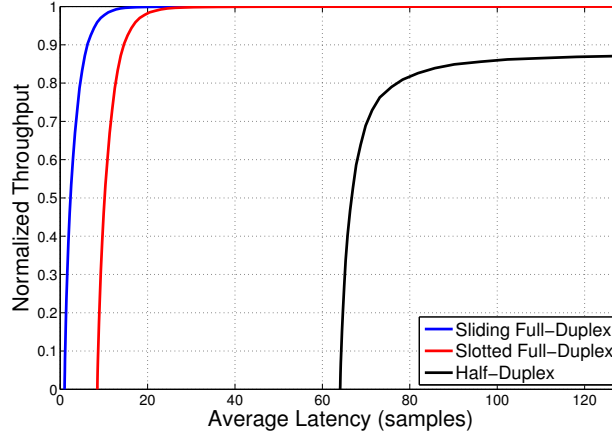


Figure 9.2 – Average latency - normalized throughput comparison for the 3 systems, $\text{SNR}_{\text{PU}} = 0$ dB with perfect self-interference suppression, $N_s = 16$, $N = 128$.

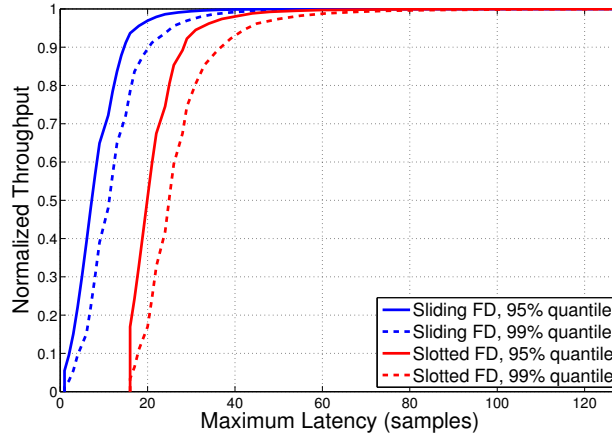


Figure 9.3 – Maximum acceptable latency - normalized throughput comparison for the 95% and 99% quantiles for the 3 systems, primary user $\text{SNR}_{\text{PU}} = 0$, $N_s = 16$, $N = 128$.

this mode of operation results in very low SU throughput (and is thus not viable in normal circumstances). Similarly, the maximum throughput is achieved for all three schemes—the half-duplex system has a reduced maximum throughput due to the sensing overhead, as given by ((9.7))—as the average latency increases.

The access latency is a random variable, which depends on the actual realizations of the signal, noise and self-interference. Thus, in order to provide useful metrics of the system performance we examine the maximum access latency that is reached with a specific confidence. Fig. 9.3 shows the 95% and 99% quantiles of the latency-throughput curves for both full-duplex schemes.

Fig. 9.4 compares the average latency for the two full-duplex models with increasing residual self-interference (measured relatively to the noise-floor). The average latency is found by

9.1. Sliding Window Spectrum Sensing for Full-Duplex Cognitive Radios

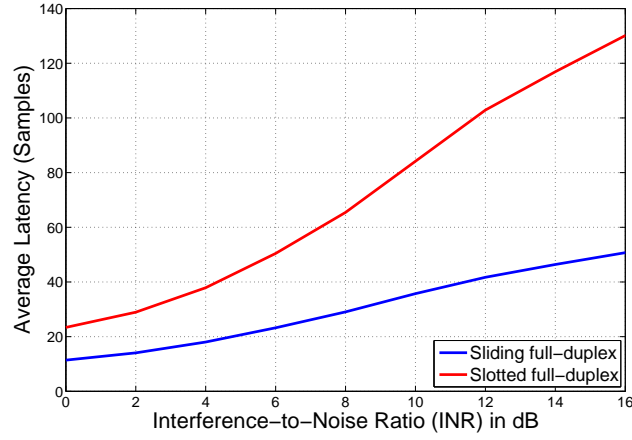


Figure 9.4 – Average access latency with increasing residual self-interference, for a normalised throughput of 0.9, $\text{SNR}_{\text{PU}} = 0$ dB, $N_s = 16$, $N = 128$.

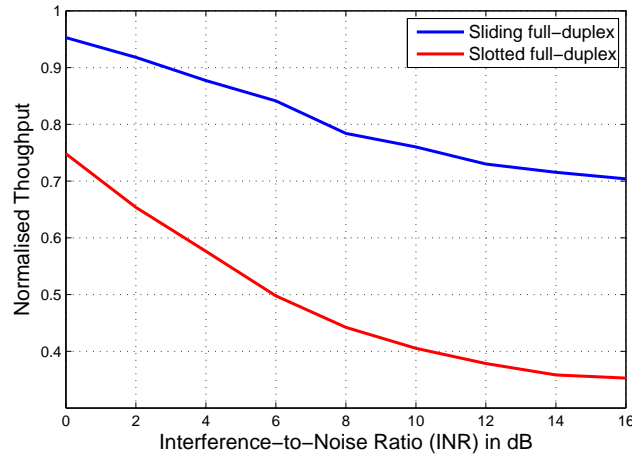


Figure 9.5 – Normalised throughput with increasing residual self-interference, for an average access latency of 16 samples, $\text{SNR}_{\text{PU}} = 0$ dB, $N_s = 16$, $N = 128$.

normalizing the throughput to 0.9, i.e., $P_f = 0.1$. The sliding scheme is observed to have lower latency for all values of residual self-interference considered. Moreover, the difference in the slopes indicates that the sliding scheme is more resilient to residual self-interference, by approximately a factor of 2.7 relative to the slotted scheme.

In Fig. 9.5 we can observe a similar result from a different perspective. In this figure we compare the normalized throughput for the two full-duplex models with increasing residual self-interference. The throughput is found for the same latency of $L = N_s$ samples for both schemes, i.e., for the same PU protection. The sliding scheme is observed to have higher throughput for all values of residual self-interference considered, again with a difference in the slopes.

9.1.5 Full-Duplex Cognitive Radio Demonstration Setup

The demonstration outlined in this section is a proof-of-concept implementation for a concurrent spectrum-sensing and transmission cognitive radio network using self-interference cancellation. An important contribution is the evaluation of spectrum-sensing algorithms for full-duplex cognitive radio under realistic conditions, i.e., in the presence of residual self-interference. For example, [29] and [30] use energy-detection algorithms to identify primary user transmissions by assuming near-perfect cancellation of the self-interference (i.e., suppression below the thermal noise-floor). In practice, perfect suppression is usually not attained, due to transceiver non-idealities, and uncertainty in the channel estimates, as explained in Chapter 7 and Chapter 8. For typical operating conditions, the residual self-interference is above the noise-floor, and this can make it difficult to demodulate weak external signals. However, for cognitive-radio, the goal is only to *detect* the presence of a primary user, and not necessarily to correctly decode the transmissions. This demonstration is a proof of concept that reliable spectrum-sensing performance can be achieved even at a lower signal-to-noise ratio due to the presence of residual self-interference. An overview of different cognitive radio SDR platforms can be found in [238].

Overview of the Demonstration

The demonstration is implemented on four National Instruments (NI) 2920 USRP transceivers [169]. The NI USRP-2920 operates at carrier frequencies of 50 MHz to 2.2 GHz and provides output power ranging from -11 dBm to 20 dBm. To prevent unwanted interference to other users, the demonstration system uses carrier frequencies in the 902–928 MHz ISM band. The USRPs are programmed over Ethernet from a laptop computer, running the LabVIEW software platform, which also performs data processing. If the spectrum-sensing algorithm detects a transmission from the PU, the cognitive radio SU quickly switches the transmitter off. The remote receiver then is able to receive the message from the PU.

In particular, USRP 3 in Fig. 9.6 is assigned the role of the PU transmitter and remote-user

9.1. Sliding Window Spectrum Sensing for Full-Duplex Cognitive Radios

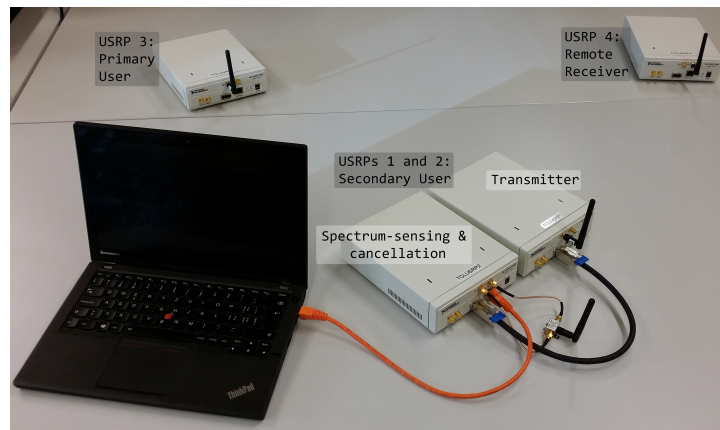


Figure 9.6 – Annotated photograph of the full-duplex cognitive radio demonstration setup.

receiver, transmitting a BPSK signal. USRPs 1 and 2 are used to form the SU full-duplex transmit node, as shown in Fig. 9.6. USRP 1 acts as the SU transmitter, modulating and sending a QPSK signal, while spectrum-sensing is performed by the SU receiver in USRP 2. USRP 3 in Fig. 9.6 is assigned the role of the remote-user receiver, which can distinguish between PU and SU transmissions.

Self-Interference Cancellation

The self-interference signal of the SU is first attenuated when propagating between the transmitting and spectrum-sensing antennas. USRP 2 also estimates the self-interference channel, and generates an appropriate RF cancellation signal. The cancellation signal is then subtracted from the incoming signal via an external combiner, reducing the power of the self-interference so that transmissions from the primary user may be detected. A digital cancellation step follows which further reduces the power of the self-interference. Analog and digital cancellation algorithms from [2] are used.

A 10 MHz reference clock is shared between transmit and cancellation USRPs. The 915 MHz carrier is regenerated within each USRP from this reference signal. In this configuration, each local oscillator has independent phase-noise, which tends to limit the self-interference cancellation that can be achieved, as explained in Section 7.2. Therefore, primary-user detection algorithms that perform well in the presence of noise and interference are required. In this demonstration we use energy-detection spectrum sensing, under real channel conditions and in the presence of varying levels of residual self-interference. The video of our fully-functional full-duplex cognitive radio demonstration can be found at [239].

9.2 Full-Duplex Communications for Wireless Links with Asymmetric Capacity Requirements

In this section, we provide another example of how full-duplex transceivers with imperfect self-interference cancellation capabilities can still be very helpful in wireless communication systems with inherent asymmetric communication requirements.

Currently, most bi-directional wireless systems separate the uplink and downlink using frequency-division duplexing (FDD) or time-division duplexing (TDD). Highly asymmetric communication links can be accommodated by allocating different amounts of time or frequency resources to the downlink and uplink streams [32, p. 459]. However, this approach introduces overhead as guard bands and guard intervals have to be used for FDD and TDD operation, respectively, in order to avoid interference between the uplink and the downlink transmission. Moreover, using highly asymmetric TDD can introduce significant latency for the uplink data, as most of the available time slots are already used for the downlink data.

While most studies on full-duplex communications focus on symmetric links and on maximizing their sum-capacity, we note that full-duplex can also be used to transmit both the low-rate uplink data and the high-rate downlink data at the same time and over the same frequency band, thus eliminating the overhead and increased latency of the asymmetric FDD and TDD approaches. It has been shown in the literature that when the data-rates are symmetric, the maximum capacity of a full-duplex link is obtained when the users and the base-station transmit at maximum power [231] even in the case where residual self-interference is present. However, in asymmetric links, transmitting at the maximum power may not be an optimal solution when the link traffic requirements are highly asymmetric. In particular, as the level of the residual self-interference depends on the transmit power, higher capacities on the downlink may be obtained by *decreasing* the uplink transmission power, depending on the degree of asymmetry. Such full-duplex links with asymmetric traffic have first been examined in the literature under the scope of resource allocation in a multi-subcarrier aggregation model [232].

This section is organized as follows. In Section 9.2.1 we describe the system model for the half-duplex asymmetric link, along with the model of the full-duplex system with residual self-interference. In Section 9.2.2, to validate the usefulness of full-duplex in asymmetric links, we present numerical evaluation results for system parameters based on the IEEE 802.11 standard along with parameters for an LTE-based system.

9.2.1 System Models for Half-Duplex and Full-Duplex Asymmetric Links

Fig. 9.7 shows an asymmetric half-duplex link employing TDD to separate the uplink and downlink, with total bandwidth W . In the depicted scenario, the proportion of the time allocated for the downlink, $\alpha_d > 0.5$, is larger than that allocated for the uplink, $\alpha_u = 1 - \alpha_d$. Consequently, the effective downlink transmission rate, R_d , (in bits/s/Hz), is higher than the uplink transmission rate, R_u . In practice, R_d and R_u are set by the requirements of the

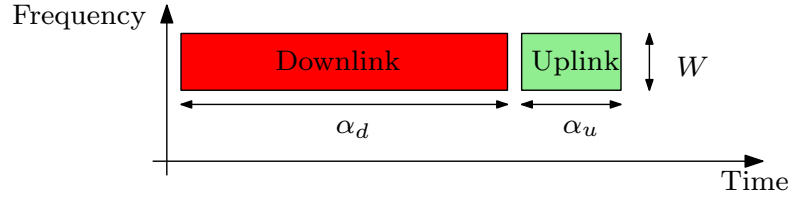


Figure 9.7 – Allocation of uplink and downlink channels for time-division duplexing (TDD). For asymmetric scenarios, TDD can allocate a larger proportion of the frame to the high-rate link, up to a limit defined by the standard for the minimum uplink rate.

users and the system adjusts α_d and α_u accordingly. However, it is important to note that in both LTE-TDD and IEEE 802.11 systems, a minimum uplink rate is provided, regardless of the actual user traffic, as α_u cannot be chosen arbitrarily small [32, 240]. For example, in LTE-TDD the maximum asymmetry between α_u and α_d is $\frac{1}{8}$, representing one 1 ms uplink sub-frame sent for eight 1 ms downlink sub-frames [32, p. 459]. It is important to note that in many cellular systems, such as LTE-FDD, spectrum allocation is standardized and is often non-contiguous. It is thus usually not possible to adjust the size of the frequency bands to reflect the traffic asymmetry [32, pp. 376–380]. In this case, asymmetry between the uplink and downlink transmission rates can significantly reduce the overall spectral efficiency as the low-rate uplink essentially squanders the excess bandwidth.

In order to make the case for asymmetric full duplex operation we first describe two models of asymmetric links, namely the corresponding half-duplex and the full-duplex models.

Half-Duplex Asymmetric Links

We first define the asymmetry ratio, r , as

$$r = \frac{R_u}{R_d} < 1, \quad (9.15)$$

where we assume the downlink has a larger throughput requirement than the uplink.

The downlink and uplink capacities in a half-duplex link are given by

$$C_{d,HD} = W \log_2 \left(1 + \frac{\delta P_{d,HD}}{N_0 W} \right) \quad (9.16)$$

$$C_{u,HD} = W \log_2 \left(1 + \frac{\delta P_{u,HD}}{N_0 W} \right), \quad (9.17)$$

where W is the bandwidth, δ is the path-loss (in linear units), $P_{d,HD}$ and $P_{u,HD}$ are the downlink and the uplink transmit powers, respectively, and N_0 is the power spectral density of the noise. As depicted in Fig. 9.7, for a system employing TDD, α_d is the portion of the frame dedicated to the downlink and α_u is the portion of the frame dedicated to the uplink.

Chapter 9. Full-Duplex Systems with Imperfect Self-Interference Cancellation

Since $R_d = \alpha_d C_{d,HD}$, and $R_u = \alpha_u C_{u,HD}$, the asymmetry ratio for half-duplex is given by

$$r_{HD} = \frac{\alpha_u C_{u,HD}}{\alpha_d C_{d,HD}}, \quad (9.18)$$

with

$$\alpha_d + \alpha_u = 1. \quad (9.19)$$

If we solve the system of linear equations (9.18) and (9.19), we can obtain values for α_d and α_u ,

$$\alpha_d = \frac{C_{u,HD}}{C_{u,HD} + r_{HD} C_{d,HD}} \quad (9.20)$$

$$\alpha_u = \frac{r_{HD} C_{d,HD}}{C_{u,HD} + r_{HD} C_{d,HD}}, \quad (9.21)$$

However, as a minimum fraction of the total frame, α_{\min} , must be allocated for the uplink according to standard and frame-format requirements, (9.20) and (9.21) are modified to

$$\alpha_u = \max\left(\alpha_{\min}, \frac{r_{HD} C_{d,HD}}{C_{u,HD} + r_{HD} C_{d,HD}}\right) \quad (9.22)$$

$$\alpha_d = 1 - \alpha_u. \quad (9.23)$$

The sum capacity of the half-duplex system is therefore

$$C_{\text{sum},HD} = \alpha_d C_{d,HD} + \alpha_u C_{u,HD}. \quad (9.24)$$

Clearly, to maximize $C_{d,HD}$ and $C_{u,HD}$, both the half duplex base-station and user-terminal transmit at the maximum allowed power, i.e., $P_{d,HD} = P_{\max}$ and $P_{u,HD} = P'_{\max}$, where P_{\max} is the maximum downlink transmit power and P'_{\max} is the maximum uplink transmit power. We use $C_{d,HD}$ and $C_{u,HD}$ obtained with P_{\max} and P'_{\max} , respectively, to compare against the full-duplex scenario.

Full-Duplex Asymmetric Links

In the full-duplex case the downlink and uplink capacities are

$$C_d = W \log_2 \left(1 + \frac{\delta P_d}{N_0 W + \beta P_u} \right) \quad (9.25)$$

$$C_u = W \log_2 \left(1 + \frac{\delta P_u}{N_0 W + \beta P_d} \right), \quad (9.26)$$

where β is the effective amount of self-interference cancellation (in linear units) of the transmitted powers P_u and P_d .

Since in a full-duplex link the downlink and the uplink are active concurrently in the same

9.2. Full-Duplex Communications for Wireless Links with Asymmetric Capacity Requirements

Table 9.1 – Parameters used in simulations for the two scenarios

Parameter	Scenario I (IEEE 802.11)	Scenario II (LTE-TDD)
Bandwidth (W)	20 MHz	20 MHz
Center frequency (f_0)	2.45 GHz	2.35 GHz
Self-interference supp.	90–110 dB	90–130 dB
Link distance	5–500 m	10–3000 m
Asymmetry ratio (r)	(2/8)	(1/9)
downlink Tx range (P_d)	-20 dBm to 23 dBm	23 dBm to 46 dBm
uplink Tx range (P_u)	-20 dBm to 23 dBm	-20 dBm to 23 dBm

band, the asymmetry ratio is

$$r_{\text{FD}} = \frac{C_u}{C_d}, \quad (9.27)$$

without any additional factors. However, generally $C_d \leq C_{d,\text{HD}}$ and $C_u \leq C_{u,\text{HD}}$, due to the residual self-interference, and only a full-duplex system with perfect self-interference suppression capabilities has $\beta = 0$, which would result in $C_d = C_{d,\text{HD}}$ and $C_u = C_{u,\text{HD}}$.

Since we are interested in the usage of full-duplex communication in asymmetric links, we require a minimum uplink rate as a fraction of the downlink data rate, i.e., $C_u \geq r_{\text{FD}} C_d$, for example for ARQ flow-control. The downlink and uplink transmit powers, namely P_d and P_u , are then chosen with the aim to maximize the downlink capacity.

Interestingly, in such an asymmetric full-duplex scenario, transmitting with maximum downlink and uplink powers will not always maximize the downlink capacity. Instead, by reducing the transmit power of the user-terminal on the uplink, the amount of self-interference of this terminal is also reduced and thus the impact on the high speed downlink can be minimized. However, the reduced power on the uplink will also result in a weaker signal at the base-station/access-point. Nevertheless, the base-station may still be able to decode the received signal, leveraging the high processing gain that is available due to the low-rate uplink. In practice, this processing gain can be introduced using channel coding or spreading applied to the uplink signal. We note that the reduced uplink power would also potentially provide an additional benefit in terms of both an increased battery life of the user terminal and a reduction in the inter-cell interference (ICI). The detrimental effects of ICI on the full-duplex performance have been extensively examined in [241, 242].

9.2.2 Results

We examine two different scenarios: scenario I resembles an IEEE 802.11 system and scenario II resembles an LTE-TDD system (in terms of parameters that are relevant to our capacity analysis). The corresponding configurations for both scenarios are presented in Table 9.1.

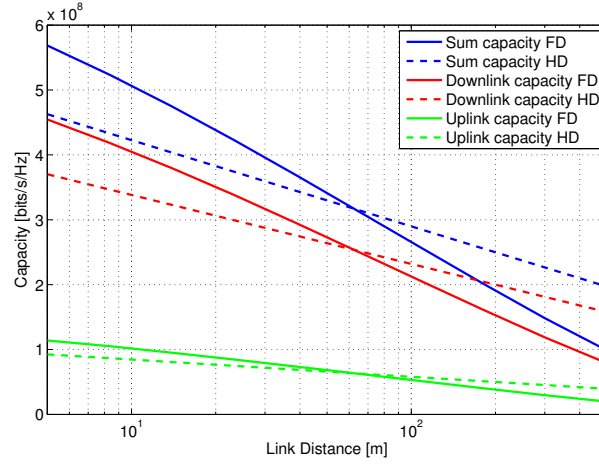


Figure 9.8 – Capacity comparison for HD and FD over the link distance range for the IEEE 802.11 scenario, with self-interference suppression of 100 dB.

Scenario I: IEEE 802.11 system

In Fig. 9.8 we observe the capacity of an asymmetric full-duplex system and the capacity of a corresponding asymmetric half-duplex system for radial distances from the base-station. The maximum transmit power is set to 23 dBm and the self-interference suppression capability of the full-duplex system is chosen to be 100 dB, since this is an achievable value in today's full duplex implementations [16, 2]. The asymmetry ratio is chosen as $r = (2/8)$, as a representative value for IEEE 802.11 systems [243]. Fig. 9.8 shows that a full-duplex system with proper power-adjustment can result in a significant potential improvement in the downlink capacity over a corresponding half-duplex system. For example, at a distance of 10 m, one can observe a 20% improvement in capacity compared to half-duplex.

In addition to the capacity improvement on the downlink, the transmit power of the user-terminal is also reduced, compared to the uplink power of a half-duplex terminal which saturates at its maximum, as shown in Fig. 9.9. As expected, the power reduction is larger for systems with higher self-interference cancellation capabilities, i.e., systems with better self-interference cancellation can transmit with lower power, for larger link distances. An interesting observation from Fig. 9.9, is that even an asymmetric full-duplex system with a relatively small cancellation capability, e.g., 90 dB, that may not look attractive under the asymmetric downlink capacity improvement criterion, may still be an interesting option due to the transmit power reduction.

Scenario II: LTE system

In a typical macro-cell scenario, the high maximum downlink transmit power of 46 dBm creates significant self-interference at the full-duplex nodes. Thus, the self-interference cancellation capabilities of the nodes need to be relatively high (e.g., 120 dB) to obtain a capacity advantage

9.2. Full-Duplex Communications for Wireless Links with Asymmetric Capacity Requirements

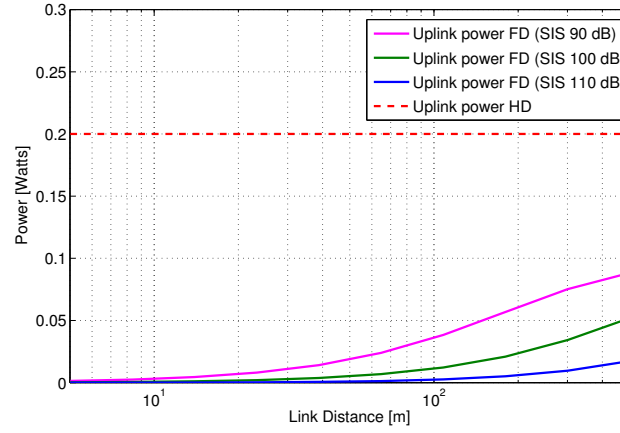


Figure 9.9 – Uplink transmit power comparison for HD and FD over the link distance range for the IEEE 802.11 scenario, with self-interference suppression of 90, 100 and 110 dB.

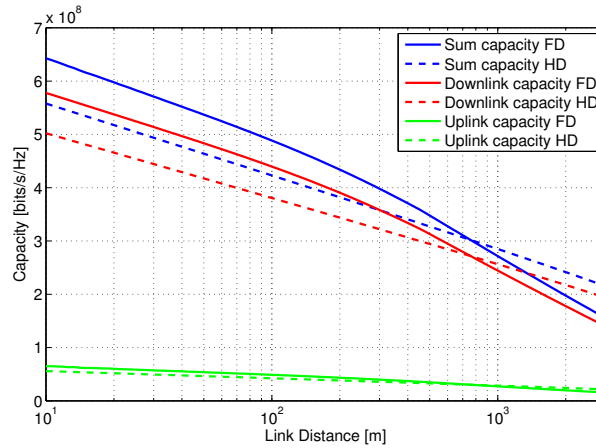


Figure 9.10 – Capacity comparison for HD and FD over the link distance range for the LTE scenario, with self-interference suppression of 120 dB.

over half-duplex operation. Nevertheless, especially in this scenario, asymmetric full-duplex links can provide interesting results as presented in Fig. 9.10. In this case the asymmetry ratio is set to $r = (1/9)$, which satisfies the α_{\min} constraint for LTE-TDD [32, pp. 459–460]. For example, at a distance of 100 m, full-duplex reveals a 16% improvement in downlink capacity compared to half-duplex. It is important to note that base-stations in small-cell networks can transmit at significantly lower powers (in the range of 10 dBm) [244], thus rendering full-duplex feasible even using nodes with much lower self-interference cancellation capabilities.

Fig. 9.11a and Fig. 9.11b present the asymmetric downlink and uplink power levels respectively, both for half-duplex and full-duplex systems, for increasing values of the link distance. We observe that for shorter distances, the power-adjustment procedure requires the base-station to transmit with maximum power, while the transmit power of the user terminal is kept as low

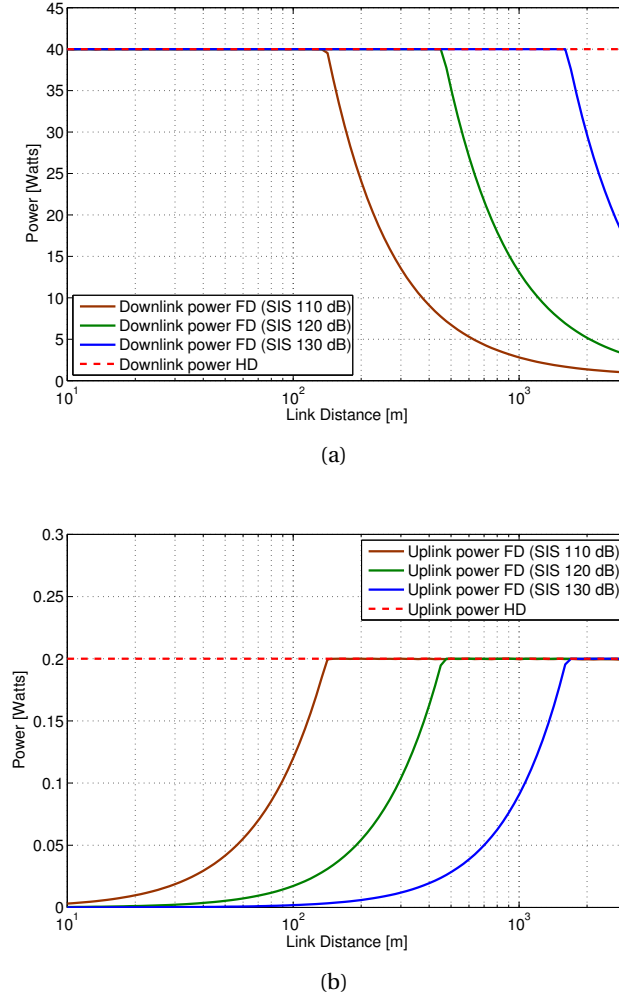


Figure 9.11 – Downlink (a) and uplink (b) transmit power comparison for HD and FD over the link distance range for the LTE scenario, with self-interference cancellation of 110, 120 and 130dB.

as possible. This observation is in line with the idea of keeping low uplink power in full-duplex systems, in order to avoid prohibitively high levels of self-interference. Moreover, we notice that for larger link distances, the user terminal saturates its transmit power at a maximum that is constrained to a value (23 dBm) which is smaller than the maximum downlink power (46 dBm) (cf. Tbl. 9.1), thus requiring the downlink power to decrease as well, to meet the asymmetry constraint. Systems with better self-interference cancellation capabilities can stay in the low uplink power region for larger link distances as the differences between the three curves in Fig. 9.11b indicate.

Fig. 9.12 shows the link distance range for which a full-duplex asymmetric LTE system outperforms a corresponding half-duplex asymmetric LTE system as the rate asymmetry is changed. We observe that for lower values of self-interference suppression, full-duplex outperforms

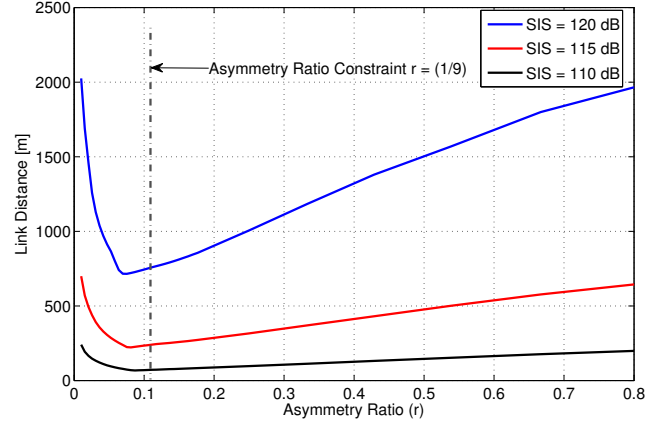


Figure 9.12 – Impact of different rate asymmetry on the useful range of a full-duplex system compared to LTE-TDD for different self-interference cancellation capability.

half-duplex for all asymmetry ratios, but only for a limited range of link distances. However, for higher values of self-interference suppression, full-duplex outperforms half-duplex in a considerably high range of link distances, especially for highly symmetric and highly asymmetric links. The great benefit of full-duplex systems for symmetric links is well-known and expected from the literature. As the asymmetry of the link increases, i.e., the asymmetry ratio r decreases, full-duplex outperforms half-duplex inside a smaller distance range. Nevertheless, full-duplex becomes particularly interesting for highly asymmetric links (i.e., for very small asymmetry ratios r). We observe that the range over which full-duplex is advantageous over half-duplex increases under highly asymmetric traffic conditions, i.e., when the uplink is sending back data with lower rate than the minimum uplink rate constraint that is set by the standard. This can happen for example when the uplink data contain only location information or ACK frames from flow control. This great improvement for highly asymmetric links, shown in Fig. 9.12, occurs as the LTE-TDD standard specifies a minimum uplink rate (for most operation modes this is $\frac{1}{8}$ which corresponds to an asymmetry ratio constraint $r = (1/9)$), regardless of the actual data asymmetry. In cases where the data asymmetry is less than this minimum provided uplink rate, full-duplex operation is more spectrally efficient and consequently can achieve a capacity improvement over half-duplex across a typical urban macro-cell for suitable values of self-interference cancellation.

9.3 Discussion and Summary

In this chapter we demonstrated that full-duplex systems can be highly beneficial in certain scenarios even with imperfect self-interference cancellation. To this end, we examined two different systems with the common characteristic of having an inherent asymmetry in their architecture. We examined cognitive radio systems, which have the inherent characteristic that the SUs do not need to decode a high-rate message, but only detect its presence. We also

examined links with asymmetric capacity requirements, where the receiver in the uplink (or the downlink) only needs to decode a low-rate signal.

In particular, in Section 9.1 we have analyzed cognitive radios from the perspective of protecting the primary users by reducing the access latency of the system. By deriving analytical formulas for the access latency of half-duplex and slotted full-duplex schemes, we have quantified the access latency problem and we have shown that it can be prohibitively high for the half-duplex model and that it is reduced by the slotted full-duplex model. In order to overcome the latency problem even more effectively, we proposed a sliding window full-duplex scheme where decisions can be taken at every sample. Our results show that there is a significant improvement in the access latency of the system by using the proposed full-duplex technique. The problem of residual self-interference that exists in every full-duplex system was also considered. The proposed sliding method was shown to be more resistant to residual self-interference than the slotted scheme. Moreover, a proof-of-concept SDR implementation of a full-duplex cognitive radio was presented. The SDR implementation showed that concurrent spectrum-sensing and transmission is possible for SUs with imperfect self-interference cancellation.

Finally, in Section 9.2, we examined systems with high levels of asymmetry between the uplink and downlink traffic. This asymmetry can lead to a significant waste in spectral resources, particularly in TDD systems where minimum uplink data rates are typically specified. In addition, the latency of the uplink traffic can be significant as it must be interleaved with the high-rate downlink data. Full-duplex operation allows the uplink to be sent simultaneously with the downlink, thereby improving spectral efficiency and reducing latency. Current full-duplex technology is unable to suppress the self-interference to (or below) the thermal noise floor. However, in asymmetric links, residual self-interference can be tolerated by increasing the spreading and/or coding rate of a low-rate uplink. It is then possible to adjust the transmit powers so that they result in maximum downlink capacity while ensuring a minimum required uplink capacity. For many configurations, the maximum capacity is in fact achieved by *reducing* the uplink transmit power. In particular, we find that, for systems based on IEEE 802.11 and LTE-TDD, full-duplex operation can outperform the corresponding half-duplex system over a wide range of typical system parameters and link asymmetry factors.

10 Summary and Discussion

In this concluding chapter we present a summary of the thesis, and discuss open research problems connected to its topics.

10.1 Summary of the Thesis

In this thesis, we have investigated various physical layer aspects of LoRa and full-duplex communication technologies. The analysis for both these technologies has been performed under a common methodology that emphasizes the understanding of the bottlenecks and limitations introduced to a wireless communication system by practical aspects on the physical layer, such as hardware impairments and interference. An additional common characteristic of our methodology has been to propose physical-layer algorithms in order to alleviate the aforementioned bottlenecks and limitations on a system. Our analysis has been accompanied by SDR implementations for both LoRa and full-duplex technologies.

Specifically, in Chapter 3 we analyzed the performance of LoRa by using a unified interference-and-noise model. We showed that this analysis gives significantly different performance results compared to the relatively simplistic physical layer models often used in the literature. Our proposed analysis can help to build an extended probabilistic performance model under interference, which can be used for a more accurate study of the scalability of LoRaWAN. Moreover, in Chapter 4 we proposed the use of coherent detection to partially mitigate the impact of same-SF interference in LoRa receivers. We showed that coherent LoRa detection can significantly improve the receiver performance, especially for high interference power. In Chapter 5 we analyzed the performance of coded LoRa in the presence of residual carrier frequency offset (CFO), which is one of the main hardware impairments that exist in low-cost LoRa end nodes. In Chapter 6 we presented our open-source GNU Radio implementation of the LoRa physical layer, which is currently the only software-defined radio (SDR) implementation that can work in the low SNR region targeted by long-range IoT systems. Our GNU Radio implementation can serve as a basis of future prototypes for the performance evaluation of improved LoRa receiver algorithms.

In the full-duplex part of the thesis, in Chapter 8 we proposed the use of predistortion to mitigate the effects of transceiver non-linearities on the self-interference cancellation performance. The improved self-interference performance was experimentally measured using a hardware full-duplex testbed. Finally, in Chapter 9 we examined two systems where full-duplex transceivers with imperfect self-interference cancellation capabilities can still be very useful. In particular, we proposed a sliding-window sensing method to reduce the access-latency of the PU in cognitive radio systems. Moreover, we proposed the use of full-duplex transceivers with residual self-interference in links with asymmetric traffic.

Some important lessons that were learned in the context of this thesis are the following. First, in order to understand the actual limitations of wireless systems, one has to always meticulously model the most relevant practical aspects that can limit the performance such as hardware impairments and destructive interference. Second, the relation between the theoretical analysis and implementations is a fundamental and crucial aspect. A physical-layer theoretical analysis of an existing technology provides the necessary knowledge to build correctly working testbeds based on SDR implementations. Then, at a second step, the implementations can be extended in order to experimentally evaluate the performance of proposed improved transceiver algorithms.

10.2 Discussion

In this section we will discuss research directions and open problems related to the topics analyzed in this thesis, which we believe to be worth examining in the future.

In Chapter 3 and Chapter 4 we analyzed the performance of uncoded LoRa under same-SF interference from one interfering user. Our error-rate model would be even more useful for network simulators if it is extended to the case where multiple same-SF interferers affect the packet of interest, both for the standard non-coherent and the coherent receivers. Moreover, although same-SF interference is the main limiting factor in LoRaWAN, the impact of inter-SF interference is also important and can be easily included in our model by considering an SNR degradation. However, the amount of the SNR decrease depending on the number of interfering users with different SFs is an open question.

Moreover, in Chapter 5, we have analyzed the performance of coded LoRa only under CFO. However, we believe that the joint effect of coding, interleaving, and Gray mapping, greatly helps the LoRa performance under other residual impairments as well, such as sampling frequency offset (SFO) and phase jitter. The impact of the aforementioned blocks under these impairments that are often present in off-the-self low-cost LoRa end-nodes is worth investigating. Moreover, coded LoRa performance should also be analyzed for the interference scenario.

The detection of LoRa packets over multipath fading channels has been very recently modeled as a time-offset estimation problem [245]. Our non-integer time-offset analysis both for the

interference in Chapter 3 and Chapter 4, and for the user of interest in Chapter 6 gives us the ability to study this important and interesting topic in much more detail. We believe that the non-integer time offset resulting from different paths will have a big impact on the correct detection in multipath environments, which is an open problem worth pursuing. The detection algorithms in multipath environments can be easily embedded in our SDR implementation and be experimentally evaluated using our testbed.

We believe that one of the most important future research directions is to analyze and implement LoRa receivers with improved receiver performance characteristics. For example, we believe that LoRa receivers that can correctly decode colliding LoRa packets are going to have a tremendous impact on the total network throughput of LoRaWAN, whose scalability is collision-limited. Multi-user receivers can be very effective, but studies on this topic have only started very recently, using relatively simple models [246, 247, 248, 249]. The interference analysis we performed in this thesis can help to study the details of such multi-user receivers in much more depth. Moreover, our current SDR implementation can be easily extended with multi-user reception capabilities for their experimental evaluation.

Finally, in the full-duplex analysis of Chapter 9, we have examined the use of low-end full-duplex transceivers with residual self-interference in links with asymmetric capacity requirements. Moreover, in the introductory LoRa chapter we saw that the uplink and downlink of many IoT systems, including LoRa, can be highly asymmetric. In an attempt to directly (and not only methodologically) combine the analysis of the two wireless communications directions examined in this thesis, the introduction of full-duplex capabilities in IoT systems can be examined. Low-cost full-duplex IoT transceivers will result in increased levels of residual self-interference, which may however be tolerated due to the highly asymmetric nature of the links. In many cases the only information that needs to be transmitted on the downlink is an ACK message from the gateway to the end node. In current IoT protocols, the single-bit ACK has to be embedded in an entire packet. As discussed in Section 2.5, confirmed traffic can have a large negative impact on the total network throughput. It would be beneficial to explore modifications to the physical and MAC layers of IoT systems, such as LoRaWAN, to allow hybrid-ARQ schemes. For example, an end node can use an adaptive-rate channel coding scheme on a packet level, on top of the current LoRa physical layer structure, to transmit as many redundancy bits as needed for the gateway to receive the message correctly. As soon as the gateway receives the uplink message, it immediately transmits an ACK. If the end node has full-duplex capabilities, it can concurrently listen for the ACK message while transmitting the redundancy bits. As soon as it receives the ACK, the end node immediately stops the uplink transmission. It would be very interesting to study if such a full-duplex IoT system with hybrid ACK would benefit both in terms of total network throughput and energy efficiency from the avoidance of retransmissions. Such a modification may also be beneficial in latency-critical applications, since it avoids all the latency introduced by duty-cycle restrictions and retransmissions.

A Appendix

Proof of Proposition 1. We will first show that \mathcal{Y}_{I_1} indeed contains $N - (s_{I_1} - s_{I_2})$ equivalent interference patterns for a given s_{I_1} and s_{I_2} , which can be obtained by setting $s'_{I_1} = [s_{I_1} + \delta]_N$, $s'_{I_2} = [s_{I_2} + \delta]_N$, and $k' = [k + \delta]_N$. Recall that for $|R_{k'}|$ we have

$$|R_{k'}| = \sqrt{(A'_{k,1})^2 + (A'_{k,2})^2 + 2A'_{k,1}A'_{k,2}\cos(\theta'_{k,1} - \theta'_{k,2})}. \quad (\text{A.1})$$

Then, our goal essentially is to show that $|R_{k'}| = |R_k|$ for all δ such that $s'_{I_1} \geq s'_{I_2}$, and for all k .

For a given s_{I_1} and s_{I_2} , the condition $s'_{I_1} \geq s'_{I_2}$ holds in the following two cases

$$\text{Case A: } s_{I_1} + \delta < N \text{ and } s_{I_2} + \delta < N, \quad (\text{A.2})$$

$$\text{Case B: } s_{I_1} + \delta \geq N \text{ and } s_{I_2} + \delta \geq N. \quad (\text{A.3})$$

As such, it is straightforward to see that $|\mathcal{Y}_{I_1}| = N - (s_{I_1} - s_{I_2})$, meaning that (3.37) holds.

We show the remainder of the proof only for case A, but it can easily be extended to case B using the same arguments. For $A'_{k,i}$, we have

$$A'_{k,i} = \frac{\sin\left(\frac{\pi}{N}([s_{I_i} + \delta]_N - [k + \delta]_N - \tau)[\tau]\right)}{\sin\left(\frac{\pi}{N}([s_{I_i} + \delta]_N - [k + \delta]_N - \tau)\right)} \quad (\text{A.4})$$

$$= \begin{cases} \frac{\sin\left(\frac{\pi}{N}(s_{I_i} - k - \tau)[\tau]\right)}{\sin\left(\frac{\pi}{N}(s_{I_i} - k - \tau)\right)}, & k + \delta < N, \\ \frac{\sin\left(\frac{\pi}{N}(s_{I_i} - k - \tau)[\tau] + \pi[\tau]\right)}{\sin\left(\frac{\pi}{N}(s_{I_i} - k - \tau) + \pi\right)}, & k + \delta \geq N. \end{cases} \quad (\text{A.5})$$

We can rewrite (A.5) as

$$A'_{k,i} = \begin{cases} +A_{k,i}, & k + \delta < N, \\ -A_{k,i}, & k + \delta \geq N \text{ and } [\tau] \text{ even}, \\ +A_{k,i}, & k + \delta \geq N \text{ and } [\tau] \text{ odd}. \end{cases} \quad (\text{A.6})$$

Appendix A. Appendix

This means that $A'_{k,1} = A_{k,1}^2$, $A'_{k,2} = A_{k,2}^2$, and $A'_{k,1} A'_{k,2} = A_{k,1} A_{k,2}$ for any k and δ . For $\cos(\theta'_{k,1} - \theta'_{k,2})$, using the assumption $s_{I_1} \geq s_{I_2}$, we have

$$\begin{aligned} \cos(\theta'_{k,1} - \theta'_{k,2}) &= \cos\left(\frac{\pi}{N} \left((\lambda - L)N + (s_{I_1} - s_{I_2})(\tau - \lceil \tau \rceil + 1) + \right. \right. \\ &\quad \left. \left. + N(s'_{I_2} - k') \right) \right). \end{aligned} \quad (\text{A.7})$$

For the term $N(s'_{I_2} - k')$, using the fact that $s_{I_2} + \delta < N$ we have the following two cases

$$N(s'_{I_2} - k') = N((s_{I_2} + \delta) - [k + \delta]_N) \quad (\text{A.8})$$

$$= \begin{cases} N(s_{I_2} - k), & k + \delta < N, \\ N(s_{I_2} - k) - N^2, & k + \delta \geq N. \end{cases} \quad (\text{A.9})$$

Combining (A.7) and (A.9) and using the fact that N is a power of two, we can show that

$$\cos(\theta'_{k,1} - \theta'_{k,2}) = \cos(\theta_{k,1} - \theta_{k,2}). \quad (\text{A.10})$$

We have thus shown the claimed result for \mathcal{Y}_{I_1} .

The corresponding proof for \mathcal{Y}_{I_2} is omitted for the sake of brevity, but it can be obtained using the same arguments. The main differences in this case are that we no longer have $A'_{k,1} A'_{k,2} = A_{k,1} A_{k,2}$ and $\cos(\theta'_{k,1} - \theta'_{k,2}) = \cos(\theta_{k,1} - \theta_{k,2})$, but it can be shown that

$$A'_{k,1} A'_{k,2} \cos(\theta'_{k,1} - \theta'_{k,2}) = A_{k,1} A_{k,2} \cos(\theta_{k,1} - \theta_{k,2} - 2\lambda\pi). \quad (\text{A.11})$$

As such, the interference patterns are identical with each other for all δ that lead to $s'_{I_1} < s'_{I_2}$, but they are different from the interference patterns that are obtained for the δ values that lead to $s'_{I_1} \geq s'_{I_2}$.

In the special case where $\lambda = 0$ (i.e., when τ is an integer), (A.11) becomes

$$A'_{k,1} A'_{k,2} \cos(\theta'_{k,1} - \theta'_{k,2}) = A_{k,1} A_{k,2} \cos(\theta_{k,1} - \theta_{k,2}), \quad (\text{A.12})$$

meaning that all interference patterns in both \mathcal{Y}_{I_1} and \mathcal{Y}_{I_2} are indeed equivalent. \square

Proof of Proposition 2. Let $k' = -k - (N - 1) + \lceil s_{I_1} + s_{I_2} \rceil_N$. In order to prove Proposition 2, we will to show that

$$|R_{k'}| = |R_k|, \quad k = 0, \dots, N - 1. \quad (\text{A.13})$$

Recall that for $|R_{k'}|$ we have

$$|R_{k'}| = \sqrt{(A'_{k,1})^2 + (A'_{k,2})^2 + 2A'_{k,1} A'_{k,2} \cos(\theta'_1 - \theta'_2)}. \quad (\text{A.14})$$

For $A'_{k,1}$ we have

$$A'_{k,1} = \frac{\cos\left(\frac{\pi}{N}(s'_{I_1} - k' - \tau')(\lceil \tau \rceil')\right)}{\cos\left(\frac{\pi}{N}(s'_{I_1} - k' - \tau')\right)} \quad (\text{A.15})$$

$$= \begin{cases} \frac{\cos\left(\frac{\pi}{N}(-s_{I_2} + k + \tau)(N - \lceil \tau \rceil)\right)}{\cos\left(\frac{\pi}{N}(-s_{I_2} + k + \tau)\right)}, & s_{I_1} + s_{I_2} < N, \\ \frac{\cos\left(\frac{\pi}{N}(-s_{I_2} + k + \tau)(N - \lceil \tau \rceil) + \pi(N - \lceil \tau \rceil)\right)}{\cos\left(\frac{\pi}{N}(-s_{I_2} + k + \tau) + \pi\right)}, & s_{I_1} + s_{I_2} \geq N, \end{cases} \quad (\text{A.16})$$

where $\lceil \tau \rceil' = N - \lceil \tau \rceil$. Note that $\lceil \tau \rceil' \neq \lceil \tau' \rceil$ because the term $\lceil \tau \rceil'$ comes from the cardinality of \mathcal{N}_{L_2} , which is $N - \lceil \tau \rceil$, and not from the time shift τ' . We can rewrite (A.16) as

$$A'_{k,1} = \begin{cases} +A_{k,2}, & s_{I_1} + s_{I_2} < N, \\ -A_{k,2}, & s_{I_1} + s_{I_2} \geq N \text{ and } \lceil \tau \rceil \text{ even}, \\ +A_{k,2}, & s_{I_1} + s_{I_2} \geq N \text{ and } \lceil \tau \rceil \text{ odd}. \end{cases} \quad (\text{A.17})$$

The equivalent expression for $A'_{k,2}$ can be obtained by exchanging the subscripts 1 and 2 in (A.17). This means that $A'^2_{k,1} = A^2_{k,2}$, $A'^2_{k,2} = A^2_{k,1}$, and $A'_{k,1}A'_{k,2} = A_{k,2}A_{k,1}$. Finally, for $\cos(\theta'_{k,1} - \theta'_{k,2})$, using the fact that (3.39) implies $L' = N - 2 - L$ and $\lambda' = 1 - \lambda$, we have

$$\begin{aligned} \cos(\theta'_{k,1} - \theta'_{k,2}) &= \cos\left(\frac{\pi}{N}((\lambda' - L')N + \right. \\ &\quad \left. + (s'_{I_1} - s'_{I_2})(2\tau' - \lceil \tau \rceil' + 1) + (s'_{I_2} - k')N)\right) \end{aligned} \quad (\text{A.18})$$

$$\begin{aligned} &= \cos\left(\frac{\pi}{N}((- \lambda + L)N + (s_{I_2} - \lceil s_{I_1} + s_{I_2} \rceil_N + k)N + \right. \\ &\quad \left. + (s_{I_1} - s_{I_2})(-2\tau + \lceil \tau \rceil - 1 + N))\right). \end{aligned} \quad (\text{A.19})$$

By taking two cases for $\lceil s_{I_1} + s_{I_2} \rceil_N$, it is straightforward to show that in both cases

$$\cos(\theta'_{k,1} - \theta'_{k,2}) = \cos(\theta_{k,1} - \theta_{k,2}). \quad (\text{A.20})$$

As such, $|R_{k'}| = |R_k|$, $k = 0, \dots, N - 1$, indeed holds. \square

Bibliography

- [1] T. Elshabrawy and J. Robert, "Analysis of BER and coverage performance of LoRa modulation under same spreading factor interference," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2018.
- [2] A. Balatsoukas-Stimming, A. Austin, P. Belanovic, and A. Burg, "Baseband and RF hardware impairments in full-duplex wireless systems: experimental characterisation and suppression," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 142, May 2015.
- [3] D. Zucchetto and A. Zanella, "Uncoordinated access schemes for the IoT: Approaches, regulations, and performance," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 48–54, Sep. 2017.
- [4] A. Boulogeorgos, P. Diamantoulakis, and G. Karagiannidis, "Low power wide area networks (LPWANS) for Internet of Things (IoT) applications: Research challenges and future trends," *ArXiv e-prints*, Nov. 2016, <https://arxiv.org/abs/1611.07449>.
- [5] A. Burg, A. Chattopadhyay, and K. Lam, "Wireless communication and security issues for cyber-physical systems and the Internet-of-Things," *Proceedings of the IEEE*, vol. 106, no. 1, pp. 38–60, Jan. 2018.
- [6] [Online]. Available: https://www.bluetooth.com/wp-content/uploads/2019/03/1901_Feature_Overview_Brief_FINAL.pdf
- [7] [Online]. Available: <https://zigbeealliance.org/wp-content/uploads/2019/11/docs-05-3474-21-0csg-zigbee-specification.pdf>
- [8] "IEEE standard for low-rate wireless networks," *IEEE Std 802.15.4-2015 (Revision of IEEE Std 802.15.4-2011)*, pp. 1–709, Apr. 2016.
- [9] [Online]. Available: https://storage.sbg.cloud.ovh.net/v1/AUTH_669d7dfced0b44518cb186841d7cbd75/prod_medias/b2be6c79-4841-4811-b9ee-61060512ecf8.pdf
- [10] [Online]. Available: <http://www.weightless.org/about/weightless-specification>

Bibliography

- [11] [Online]. Available: <https://www.ingenu.com/technology/>
- [12] “LoRaWAN specification v1.1,” <https://loro-alliance.org/resource-hub/lorawanr-specification-v11>.
- [13] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, “Long-range communications in unlicensed bands: the rising stars in the IoT and smart city scenarios,” *IEEE Wireless Communications*, vol. 23, no. 5, pp. 60–67, Oct. 2016.
- [14] U. Raza, P. Kulkarni, and M. Sooriyabandara, “Low power wide area networks: An overview,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 855–873, Secondquarter 2017.
- [15] M. Duarte and A. Sabharwal, “Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results,” in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*, Nov. 2010, pp. 1558–1562.
- [16] M. Jain, J.-I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, “Practical, real-time, full duplex wireless,” in *Proc. 17th International Conference on Mobile Computing and Networking*. ACM, Sep. 2011, pp. 301–312.
- [17] M. Duarte, C. Dick, and A. Sabharwal, “Experiment-driven characterization of full-duplex wireless systems,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Nov. 2012.
- [18] T. Riihonen and R. Wichman, “Analog and digital self-interference cancellation in full-duplex MIMO-OFDM transceivers with limited resolution in A/D conversion,” in *Asilomar Conference on Signals, Systems and Computers*, Nov. 2012, pp. 45–49.
- [19] L. Song, R. Wichman, Y. Li, and Z. Han, *Full-Duplex Communications and Networks*. Cambridge University Press, 2017.
- [20] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, “In-band full-duplex wireless: Challenges and opportunities,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Jun. 2014.
- [21] A. C. M. Austin, A. Balatsoukas-Stimming, and A. Burg, “Digital predistortion of power amplifier non-linearities for full-duplex transceivers,” in *Proc IEEE Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2016.
- [22] Spectrum Policy Task Force, “Report of the spectrum efficiency working group,” Federal Communications Commission, Tech. Rep., Nov. 2002.
- [23] R. I. C. Chiang, G. B. Rowe, and K. W. Sowerby, “A quantitative analysis of spectral occupancy measurements for cognitive radio,” in *Proc. IEEE Vehicular Technology Conf. (Spring)*, Apr. 2007, pp. 3016–3020.

- [24] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
- [25] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [26] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, Mar. 2009.
- [27] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.
- [28] E. Ahmed, A. Eltawil, and A. Sabharwal, "Simultaneous transmit and sense for cognitive radios using full-duplex: A first study," in *Proc. IEEE Antennas Propag. Int. Symp.*, Jul. 2012.
- [29] W. Afifi and M. Krunz, "Exploiting self-interference suppression for improved spectrum awareness/efficiency in cognitive radio systems," in *IEEE INFOCOM*, Apr. 2013, pp. 1258–1266.
- [30] T. Riihonen and R. Wichman, "Energy detection in full-duplex cognitive radios under residual self-interference," in *Proc. Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Jun. 2014, pp. 57–60.
- [31] Y. Liao, T. Wang, L. Song, and Z. Han, "Listen-and-talk: Full-duplex cognitive radio networks," in *Global Communications Conference (GLOBECOM), 2014 IEEE*, Dec. 2014, pp. 3068–3073.
- [32] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, 2nd ed. John Wiley & Sons, 2011.
- [33] A. Chiumento, C. Desset, S. Pollin, L. Van der Perre, and R. Lauwereins, "Impact of CSI feedback strategies on LTE downlink and reinforcement learning solutions for optimal allocation," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 550–562, Jan. 2017.
- [34] O. Afisiadis, M. Cotting, A. Burg, and A. Balatsoukas-Stimming, "LoRa symbol error rate under non-aligned interference," in *Asilomar Conference on Signals, Systems, and Computers*, Nov. 2019.
- [35] O. Afisiadis, M. Cotting, A. Burg, and A. Balatsoukas-Stimming, "On the error rate of the LoRa modulation with interference," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1292–1304, Feb. 2020.
- [36] O. Afisiadis, A. Burg, and A. Balatsoukas-Stimming, "Coded LoRa frame error rate analysis," in *IEEE International Conference on Communications (ICC) (accepted)*, Jun. 2020.

Bibliography

- [37] J. Tapparel, O. Afisiadis, A. Balatsoukas-Stimming, P. Mayoraz, and A. Burg, "An open-source LoRa physical layer prototype on GNU Radio," *ArXiv e-prints*, Nov. 2020, <https://arxiv.org/abs/2002.08208>.
- [38] A. C. M. Austin, O. Afisiadis, and A. Burg, "Digital predistortion of hardware impairments for full-duplex transceivers," in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Nov. 2017, pp. 878–882.
- [39] O. Afisiadis, A. C. M. Austin, A. Balatsoukas-Stimming, and A. Burg, "Sliding window spectrum sensing for full-duplex cognitive radios with low access-latency," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016, pp. 1–5.
- [40] —, "Full-duplex communications for wireless links with asymmetric capacity requirements," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*, Oct. 2017, pp. 1682–1686.
- [41] Semtech Corporation, "SX1272/73–860 MHz to 1020 MHz low power long range transceiver." [Online]. Available: <https://www.semtech.com/uploads/documents/sx1272.pdf>
- [42] O. B. Seller and N. Sornin, "Low power long range transmitter," US Patent 9,252,834, Feb., 2016.
- [43] M. Knight and B. Seeber, "Decoding LoRa: Realizing a modern LPWAN with SDR," in *GNU Radio Conference*, Sep. 2016.
- [44] P. Robyns, P. Quax, W. Lamotte, and W. Thenaers, "A multi-channel software decoder for the LoRa modulation scheme," in *International Conference on Internet of Things, Big Data and Security (IoTBDs)*, Mar. 2018, pp. 41–51.
- [45] L. Vangelista, "Frequency shift chirp modulation: The LoRa modulation," *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1818–1821, Dec. 2017.
- [46] M. Chiani and A. Elzanaty, "On the LoRa modulation for IoT: Waveform properties and spectral analysis," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8463–8470, Oct. 2019.
- [47] T. Elshabrawy, P. Edward, M. Ashour, and J. Robert, "On the different mathematical realizations for the digital synthesis of LoRa-based modulation," in *European Wireless Conf.*, May 2019, pp. 1–6.
- [48] R. Ghanaatian, O. Afisiadis, M. Cotting, and A. Burg, "LoRa digital receiver analysis and implementation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019.
- [49] C. Bernier, F. Dehmas, and N. Deparis, "Low complexity LoRa frame synchronization for ultra-low power software-defined radios," *IEEE Transactions on Communications*, Feb. 2020 (early access).

-
- [50] M. Xhonneux, D. Bol, and J. Louveaux, "A low-complexity synchronization scheme for LoRa end nodes," *ArXiv e-prints*, Dec. 2019, <https://arxiv.org/abs/1912.11344>.
 - [51] J. Tapparel, "Complete reverse engineering of LoRa PHY," <https://tcl.epfl.ch/resources-and-sw/lora-phy/>, EPFL, Tech. Rep., 2019.
 - [52] "LoRaWAN specification v1.0," <https://lora-alliance.org/resource-hub/lorawanr-specification-v10>.
 - [53] "The Things Network." [Online]. Available: <https://www.thethingsnetwork.org/>
 - [54] A. Tanenbaum, *Computer Networks*, 5th ed. London, UK: Pearson Studium, 2012.
 - [55] J. Haxhibeqiri, E. De Poorter, I. Moerman, and J. Hoebeke, "A survey of LoRaWAN for IoT: From technology to application," *Sensors*, vol. 18, no. 11, Nov. 2018.
 - [56] H. M. Jawad, R. Nordin, S. K. Gharghan, A. M. Jawad, and M. Ismail, "Energy-efficient wireless sensor networks for precision agriculture: A review," *Sensors*, vol. 17, no. 8, Aug. 2017.
 - [57] B. Citoni, F. Fioranelli, M. A. Imran, and Q. H. Abbasi, "Internet of Things and LoRaWAN-enabled future smart farming," *IEEE Internet of Things Magazine*, vol. 2, no. 4, pp. 14–19, Dec. 2019.
 - [58] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E. M. Aggoune, "Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk," *IEEE Access*, vol. 7, pp. 129 551–129 583, Aug. 2019.
 - [59] S. Gil-Lebrero, F. J. Quiles-Latorre, M. Ortiz-López, V. Sánchez-Ruiz, V. Gámiz-López, and J. J. Luna-Rodríguez, "Honey bee colonies remote monitoring system," *Sensors*, vol. 17, no. 1, 2017.
 - [60] S. Benaissa, D. Plets, E. Tanghe, J. Troghe, L. Martens, L. Vandaele, L. Verloock, F. Tuytens, B. Sonck, and W. Joseph, "Internet of animals: characterisation of LoRa sub-GHz off-body wireless channel in dairy barns," *Electronics Letters*, vol. 53, no. 18, pp. 1281–1283, Aug. 2017.
 - [61] D. Sartori and D. Brunelli, "A smart sensor for precision agriculture powered by microbial fuel cells," in *2016 IEEE Sensors Applications Symposium (SAS)*, Apr. 2016, pp. 1–6.
 - [62] D. Ilie-Ablachim, G. C. Pătru, I. Florea, and D. Rosner, "Monitoring device for culture substrate growth parameters for precision agriculture: Acronym: MoniSen," in *2016 15th RoEduNet Conference: Networking in Education and Research*, Sep. 2016, pp. 1–7.
 - [63] A. J. Wixted, P. Kinnaird, H. Larijani, A. Tait, A. Ahmadiania, and N. Strachan, "Evaluation of LoRa and LoRaWAN for wireless sensor networks," in *IEEE SENSORS*, Oct. 2016, pp. 1–3.

Bibliography

- [64] H. Lee and K. Ke, "Monitoring of large-area IoT sensors using a LoRa wireless mesh network system: Design and evaluation," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 9, pp. 2177–2187, Sep. 2018.
- [65] D. Magrin, M. Centenaro, and L. Vangelista, "Performance evaluation of LoRa networks in a smart city scenario," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–7.
- [66] G. Pasolini, C. Buratti, L. Feltrin, F. Zabini, C. De Castro, R. Verdone, and O. Andrisano, "Smart city pilot projects using LoRa and IEEE802.15.4 technologies," *Sensors*, vol. 18, no. 4, Apr. 2018.
- [67] P. Jörke, S. Böcker, F. Liedmann, and C. Wietfeld, "Urban channel models for smart city IoT-networks based on empirical measurements of LoRa-links at 433 and 868 MHz," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–6.
- [68] G. PremSankar, B. Ghaddar, M. Slabicki, and M. Di Francesco, "Optimal configuration of LoRa networks in smart cities," *IEEE Transactions on Industrial Informatics*, Feb. 2020 (early access).
- [69] M. de Castro Tomé, P. H. J. Nardelli, and H. Alves, "Long-range low-power wireless networks and sampling strategies in electricity metering," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1629–1637, Feb. 2019.
- [70] N. Varsier and J. Schwoerer, "Capacity limits of LoRaWAN technology for smart metering applications," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6.
- [71] M. Rizzi, P. Ferrari, A. Flammini, and E. Sisinni, "Evaluation of the IoT LoRaWAN solution for distributed measurement applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 12, pp. 3340–3349, Dec. 2017.
- [72] M. Rizzi, P. Ferrari, A. Flammini, E. Sisinni, and M. Gidlund, "Using LoRa for industrial wireless networks," in *2017 IEEE 13th International Workshop on Factory Communication Systems (WFCS)*, May 2017, pp. 1–4.
- [73] J. Haxhibeqiri, A. Karaagac, F. Van den Abeele, W. Joseph, I. Moerman, and J. Hoebeke, "LoRa indoor coverage and performance in an industrial environment: Case study," in *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, Sep. 2017, pp. 1–8.
- [74] P. Sommer, Y. Maret, and D. Dzung, "Low-power wide-area networks for industrial sensing applications," in *2018 IEEE International Conference on Industrial Internet (ICII)*, Oct. 2018, pp. 23–32.

- [75] D. Zorbas, K. Abdelfadeel, P. Kotzanikolaou, and D. Pesch, "TS-LoRa: Time-slotted LoRaWAN for the industrial Internet of Things," *Computer Communications*, vol. 153, pp. 1 – 10, Mar. 2020.
- [76] A. Seferagić, J. Famaey, E. De Poorter, and J. Hoebeke, "Survey on wireless technology trade-offs for the industrial Internet of Things," *Sensors*, vol. 20, no. 2, Jan. 2020.
- [77] P. A. Catherwood, D. Steele, M. Little, S. Mccomb, and J. McLaughlin, "A community-based IoT personalized wireless healthcare solution trial," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 6, pp. 1–13, May 2018.
- [78] J. Petäjärvi, K. Mikhaylov, R. Yasmin, M. Hämäläinen, and J. Iinatti, "Evaluation of LoRa LPWAN Technology for Indoor Remote Health and Wellbeing Monitoring," *International Journal of Wireless Information Networks*, vol. 24, no. 2, pp. 153–165, Jun. 2017.
- [79] S. B. Baker, W. Xiang, and I. Atkinson, "Internet of Things for smart healthcare: Technologies, challenges, and opportunities," *IEEE Access*, vol. 5, pp. 26 521–26 544, Nov. 2017.
- [80] K. Lam, C. Cheung, and W. Lee, "LoRa-based localization systems for noisy outdoor environment," in *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct. 2017, pp. 278–284.
- [81] —, "RSSI-based lora localization systems for large-scale indoor and outdoor environments," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 11 778–11 791, Dec. 2019.
- [82] Z. He, Y. Li, L. Pei, and K. O’Keefe, "Enhanced gaussian process-based localization using a low power wide area network," *IEEE Communications Letters*, vol. 23, no. 1, pp. 164–167, Jan. 2019.
- [83] M. Aernouts, R. Berkvens, K. Van Vlaenderen, and M. Weyn, "Sigfox and LoRaWAN Datasets for Fingerprint Localization in Large Urban and Rural Areas," *Data*, vol. 3, Jun. 2018.
- [84] N. Podevijn, D. Plets, J. Trough, L. Martens, P. Suanet, K. Hendrikse, and W. Joseph, "TDoA-Based Outdoor Positioning with Tracking Algorithm in a Public LoRa Network," May 2018.
- [85] W. Bakkali, M. Kieffer, M. Lalam, and T. Lestable, "Kalman filter-based localization for internet of things LoRaWAN end points," in *IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–6.
- [86] M. Aernouts, N. BniLam, R. Berkvens, and M. Weyn, "Simulating a combination of TDoA and AoA localization for LoRaWAN," in *Advances on P2P, Parallel, Grid, Cloud and Internet Computing*. Springer International Publishing, Oct. 2020, pp. 756–765.

Bibliography

- [87] D. F. Carvalho, A. Depari, P. Ferrari, A. Flammini, S. Rinaldi, and E. Sisinni, "On the feasibility of mobile sensing and tracking applications based on LPWAN," in *2018 IEEE Sensors Applications Symposium (SAS)*, Mar. 2018, pp. 1–6.
- [88] W. R. Da Silva, L. Oliveira, N. Kumar, R. A. L. Rabêlo, C. N. M. Marins, and J. J. P. C. Rodrigues, "An Internet of Things tracking system approach based on LoRa protocol," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2018, pp. 1–7.
- [89] B. C. Fargas and M. N. Petersen, "GPS-free geolocation using LoRa in low-power WANs," in *2017 Global Internet of Things Summit (GloTS)*, Jun. 2017, pp. 1–6.
- [90] Semtech Corporation, "LR1110–low power Wi-Fi/ GNSS scanner + LoRa transceiver." [Online]. Available: <https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/2R000000Q2YY/7hyal8gUewWREN8DSEk5R3Ee8OpqEuVOdsHpiAHb3jo>
- [91] MyriadRF. [Online]. Available: <https://github.com/myriadrf/LoRa-SDR/>
- [92] A. Marquet, N. Montavont, and G. Z. Papadopoulos, "Investigating theoretical performance and demodulation techniques for LoRa," in *IEEE Int. Symp. on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, June 2019, pp. 1–6.
- [93] A. Marquet, N. Montavont, and G. Z. Papadopoulos, "Towards an SDR implementation of LoRa: Reverse-engineering, demodulation strategies and assessment over Rayleigh channel," *Computer Communications*, vol. 153, pp. 595 – 605, Mar. 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0140366419314665>
- [94] C. Orfanidis, L. M. Feeney, M. Jacobsson, and P. Gunningberg, "Investigating interference between LoRa and IEEE 802.15.4g networks," in *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct. 2017, pp. 1–8.
- [95] B. Reynders, W. Meert, and S. Pollin, "Range and coexistence analysis of long range unlicensed communication," in *International Conference on Telecommunications (ICT)*, May 2016, pp. 1–6.
- [96] K. Mikhaylov, J. Petäjäjärvi, and J. Janhunen, "On LoRaWAN scalability: Empirical evaluation of susceptibility to inter-network interference," in *2017 European Conference on Networks and Communications (EuCNC)*, Jun. 2017.
- [97] J. Haxhibeqiri, A. Shahid, M. Saelens, J. Bauwens, B. Jooris, E. De Poorter, and J. Hoebeke, "Sub-gigahertz inter-technology interference. How harmful is it for LoRa?" in *2018 IEEE International Smart Cities Conference (ISC2)*, Sep. 2018.
- [98] M. Lauridsen, B. Vejlgård, I. Z. Kovacs, H. Nguyen, and P. Mogensen, "Interference measurements in the european 868 mhz ISM band with focus on LoRa and SigFox," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2017.

-
- [99] B. Vejlgaard, M. Lauridsen, H. Nguyen, I. Z. Kovacs, P. Mogensen, and M. Sorensen, "Interference impact on coverage and capacity for low power wide area IoT networks," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2017.
- [100] J. Robert, S. Rauh, H. Lieske, and A. Heuberger, "IEEE 802.15 low power wide area network (LPWAN) PHY interference model," in *2018 IEEE International Conference on Communications (ICC)*, May 2018.
- [101] T. Elshabrawy and J. Robert, "The impact of ISM interference on LoRa BER performance," in *2018 IEEE Global Conference on Internet of Things (GCIoT)*, Dec. 2018.
- [102] ITU-T, "SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS. Short range narrow-band digital radiocommunication transceivers – PHY, MAC, SAR and LLC layer specifications." [Online]. Available: <https://www.itu.int/rec/T-REC-G.9959-201501-I/en>
- [103] "IO Home Control." [Online]. Available: <http://www.io-homecontrol.com/index.php/en/>
- [104] M. C. Bor, U. Roedig, T. Voigt, and J. M. Alonso, "Do LoRa low-power wide-area networks scale?" in *ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, ser. MSWiM '16. New York, NY, USA: ACM, Nov. 2016, pp. 59–67.
- [105] T. Voigt, M. C. Bor, U. Roedig, and J. M. Alonso, "Mitigating inter-network interference in LoRa networks," *CoRR*, vol. abs/1611.00688, Nov. 2016. [Online]. Available: <http://arxiv.org/abs/1611.00688>
- [106] J. Haxhibeqiri, F. Van den Abeele, I. Moerman, and J. Hoebeke, "LoRa scalability: A simulation model based on interference measurements," *Sensors*, vol. 17, no. 6, Jun. 2017.
- [107] P. Ferrari, A. Flammini, M. Rizzi, E. Sisinni, and M. Gidlund, "On the evaluation of LoRaWAN virtual channels orthogonality for dense distributed systems," in *2017 IEEE International Workshop on Measurement and Networking (MN)*, Sep. 2017, pp. 1–6.
- [108] R. Fernandes, R. Oliveira, M. Luís, and S. Sargento, "On the real capacity of LoRa networks: the impact of non-destructive communications," *IEEE Comm. Letters*, vol. 23, no. 12, pp. 2437–2441, Dec. 2019.
- [109] D. Croce, M. Gucciardo, I. Tinnirello, D. Garlisi, and S. Mangione, "Impact of spreading factor imperfect orthogonality in LoRa communications," in *Digital Communication. Towards a Smart and Secure Future Internet*, A. Piva, I. Tinnirello, and S. Morosi, Eds. Cham: Springer International Publishing, Sep. 2017, pp. 165–179.
- [110] D. Croce, M. Gucciardo, S. Mangione, G. Santaromita, and I. Tinnirello, "Impact of LoRa imperfect orthogonality: Analysis of link-level performance," *IEEE Communications Letters*, vol. 22, no. 4, pp. 796–799, Apr. 2018.

Bibliography

- [111] C. Goursaud and J.-M. Gorce, "Dedicated networks for IoT : PHY / MAC state of the art and challenges," *EAI Endorsed Transactions on Internet of Things*, Oct. 2015.
- [112] L. Feltrin, C. Buratti, E. Vinciarelli, R. De Bonis, and R. Verdone, "LoRaWAN: Evaluation of link- and system-level performance," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2249–2258, Jun. 2018.
- [113] T. Elshabrawy and J. Robert, "Capacity planning of LoRa networks with joint noise-limited and interference-limited coverage considerations," *IEEE Sensors Journal*, pp. 1–1, Feb. 2019.
- [114] J. Petajajarvi, K. Mikhaylov, A. Roivainen, T. Hanninen, and M. Pettissalo, "On the coverage of LPWANs: range evaluation and channel attenuation model for LoRa technology," in *2015 14th International Conference on ITS Telecommunications (ITST)*, Dec. 2015, pp. 55–59.
- [115] A. Augustin, J. Yi, T. Clausen, and W. M. Townsley, "A study of LoRa: Long range and low power networks for the Internet of Things," *Sensors*, vol. 16, no. 9, Sep. 2016.
- [116] P. Neumann, J. Montavont, and T. Noël, "Indoor deployment of low-power wide area networks (LPWAN): A LoRaWAN case study," in *2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct. 2016, pp. 1–8.
- [117] S. Hosseinzadeh, H. Larijani, K. Curtis, A. Wixted, and A. Amini, "Empirical propagation performance evaluation of LoRa for indoor environment," in *2017 IEEE 15th International Conference on Industrial Informatics (INDIN)*, Jul. 2017, pp. 26–31.
- [118] G. Callebaut, G. Leenders, C. Buyle, S. Crul, and L. V. der Perre, "LoRa physical layer evaluation for point-to-point links and coverage measurements in diverse environments," Sep. 2019.
- [119] G. Callebaut and L. Van der Perre, "Characterization of LoRa point-to-point path loss: Measurement campaigns and modeling considering censored data," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1910–1918, Mar. 2020.
- [120] G. Callebaut, "DRAMCO/LoRaPointToPointModules: Init release," May 2019. [Online]. Available: <https://doi.org/10.5281/zenodo.2660247>
- [121] B. Reynders and S. Pollin, "Chirp spread spectrum as a modulation technique for long range communication," in *2016 Symposium on Communications and Vehicular Technologies (SCVT)*, Nov. 2016.
- [122] T. Elshabrawy and J. Robert, "Closed-form approximation of LoRa modulation BER performance," *IEEE Communications Letters*, vol. 22, no. 9, pp. 1778–1781, Sep. 2018.

-
- [123] C. Ferreira Dias, E. Rodrigues de Lima, and G. Fraidenraich, "Bit error rate closed-form expressions for LoRa systems under Nakagami and Rice fading channels," *Sensors*, vol. 19, Oct. 2019.
- [124] J. Courjault, B. Vrigneau, M. Gautier, and O. Berder, "Accurate LoRa performance evaluation using marcum function," in *2019 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2019.
- [125] G. Baruffa, L. Rugini, V. Mecarelli, L. Germani, and F. Frescura, "Coded LoRa performance in wireless channels," in *IEEE Int. Symp. on Personal, Indoor, and Mobile Radio Commun.*, Sep. 2019, pp. 1–6.
- [126] D. Bankov, E. Khorov, and A. Lyakhov, "On the limits of LoRaWAN channel access," in *2016 International Conference on Engineering and Telecommunication (EnT)*, Nov. 2016, pp. 10–14.
- [127] —, "Mathematical model of LoRaWAN channel access with capture effect," in *IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–5.
- [128] D. Bankov, E. Khorov, and A. Lyakhov, "LoRaWAN modeling and MCS allocation to satisfy heterogeneous QoS requirements," *Sensors*, vol. 19, no. 19, Sep. 2019.
- [129] O. Georgiou and U. Raza, "Low power wide area network analysis: Can LoRa scale?" *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp. 162–165, Apr. 2017.
- [130] A. Mahmood, E. Sisinni, L. Guntupalli, R. Rondón, S. A. Hassan, and M. Gidlund, "Scalability analysis of a LoRa network under imperfect orthogonality," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1425–1436, Mar. 2019.
- [131] A. Hoeller, R. D. Souza, O. L. Alcaraz López, H. Alves, M. de Noronha Neto, and G. Brante, "Analysis and performance optimization of LoRa networks with time and antenna diversity," *IEEE Access*, vol. 6, pp. 32 820–32 829, Jun. 2018.
- [132] Z. Li, S. Zozor, J. Brossier, N. Varsier, and Q. Lampin, "2D time-frequency interference modelling using stochastic geometry for performance evaluation in low-power wide-area networks," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–7.
- [133] G. Ferre, "Collision and packet loss analysis in a LoRaWAN network," in *2017 25th European Signal Processing Conference (EUSIPCO)*, Aug. 2017, pp. 2586–2590.
- [134] J. Toussaint, N. El Rachkidy, and A. Guitton, "Performance analysis of the on-the-air activation in LoRaWAN," in *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Oct. 2016, pp. 1–7.

Bibliography

- [135] R. B. Sørensen, D. M. Kim, J. J. Nielsen, and P. Popovski, "Analysis of latency and MAC-layer performance for class A LoRaWAN," *IEEE Wireless Communications Letters*, vol. 6, no. 5, pp. 566–569, Oct. 2017.
- [136] M. Capuzzo, D. Magrin, and A. Zanella, "Mathematical modeling of LoRa WAN performance with bi-directional traffic," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2018, pp. 206–212.
- [137] D. Magrin, M. Capuzzo, A. Zanella, and M. Zorzi, "A complete LoRaWAN model for single-gateway scenarios," Dec. 2019.
- [138] B. Reynders, W. Meert, and S. Pollin, "Power and spreading factor control in low power wide area networks," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6.
- [139] [Online]. Available: <https://www.lancaster.ac.uk/scc/sites/lora/lorasim.html>
- [140] A. Pop, U. Raza, P. Kulkarni, and M. Sooriyabandara, "Does bidirectional traffic do more harm than good in LoRaWAN based LPWA networks?" in *IEEE Global Communications Conference (GLOBECOM)*, Dec. 2017, pp. 1–6.
- [141] S. Sugianto, A. A. Anhar, R. Harwahu, and R. F. Sari, "Simulation of mobile LoRa gateway for smart electricity meter," in *2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, Oct. 2018, pp. 292–297.
- [142] M. O. Farooq and D. Pesch, "Evaluation of multi-gateway LoRaWAN with different data traffic models," in *2018 IEEE 43rd Conference on Local Computer Networks (LCN)*, Oct. 2018, pp. 279–282.
- [143] S. Spinsante, L. Gioacchini, and L. Scalise, "A novel experimental-based tool for the design of LoRa networks," in *2019 II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4.0 IoT)*, Jun. 2019, pp. 317–322.
- [144] S. Borkotoky, C. Bettstetter, U. Schilcher, and C. Raffelsberger, "Allocation of repetition redundancy in LoRa," in *European Wireless 2019; 25th European Wireless Conference*, May 2019, pp. 1–6.
- [145] C. M. S. Ferreira, R. A. R. Oliveira, and J. S. Silva, "Low-energy smart cities network with LoRa and Bluetooth," in *2019 7th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud)*, Apr. 2019, pp. 24–29.
- [146] K. Q. Abdelfadeel, D. Zorbas, V. Cionca, and D. Pesch, "FREE —fine-grained scheduling for reliable and energy-efficient data collection in LoRaWAN," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 669–683, Jan. 2020.
- [147] [Online]. Available: <https://github.com/kqorany/FREE>

-
- [148] G. Callebaut, G. Ottoy, and L. Van der Perre, "Cross-layer framework and optimization for efficient use of the energy budget of IoT nodes," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2019, pp. 1–6.
- [149] [Online]. Available: <https://github.com/GillesC/LoRaEnergySim>
- [150] G. Ottoy, G. Leenders, and G. Callebaut, "LoRaWAN EFM32." [Online]. Available: https://github.com/DRAMCO/LoRaWAN_EFM32
- [151] F. Van den Abeele, J. Haxhibeqiri, I. Moerman, and J. Hoebeke, "Scalability analysis of large-scale LoRaWAN networks in ns-3," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 2186–2198, Dec. 2017.
- [152] M. Capuzzo, D. Magrin, and A. Zanella, "Confirmed traffic in LoRaWAN: Pitfalls and countermeasures," in *2018 17th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, Jun. 2018, pp. 1–7.
- [153] D. Magrin, M. Capuzzo, and A. Zanella, "A thorough study of LoRaWAN performance under different parameter settings," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 116–127, Jan. 2020.
- [154] [Online]. Available: <https://github.com/signetlabdei/lorawan>
- [155] B. Reynders, Q. Wang, and S. Pollin, "A LoRaWAN module for ns-3: Implementation and evaluation," in *Workshop on ns-3*, ser. WNS3 '18. New York, NY, USA: ACM, Jun. 2018, pp. 61–68.
- [156] T. To and A. Duda, "Simulation of LoRa in NS-3: Improving LoRa performance with CSMA," in *2018 IEEE International Conference on Communications (ICC)*, May 2018, pp. 1–7.
- [157] N. Kouvelas, V. Rao, and R. R. V. Prasad, "Employing p-CSMA on a LoRa network simulator," May 2018.
- [158] M. Slabicki, G. PremSankar, and M. Di Francesco, "Adaptive configuration of LoRa networks for dense IoT deployments," in *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*, Apr. 2018, pp. 1–9.
- [159] [Online]. Available: <https://flora.aalto.fi/>
- [160] A. M. Yousuf, E. M. Rochester, B. Ousat, and M. Ghaderi, "Throughput, coverage and scalability of LoRa LPWAN for Internet of Things," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, Jun. 2018, pp. 1–10.
- [161] [Online]. Available: <http://things.cs.ucalgary.ca/lorasim.zip>
- [162] M. Centenaro, L. Vangelista, and R. Kohno, "On the impact of downlink feedback on LoRa performance," in *IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–6.

Bibliography

- [163] J. M. Marais, A. M. Abu-Mahfouz, and G. P. Hancke, "A review of LoRaWAN simulators: Design requirements and limitations," in *2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC)*, Nov. 2019, pp. 1–6.
- [164] H. A. David and H. N. Nagaraja, *Order Statistics*. Wiley, 2003.
- [165] G. K. Karagiannidis and A. S. Lioumpas, "An improved approximation for the Gaussian Q-function," *IEEE Communications Letters*, vol. 11, no. 8, pp. 644–646, Aug. 2007.
- [166] "LoRa-PHY repository," <https://tcl.epfl.ch/resources-and-sw/lora-phy/>.
- [167] "Adafruit Feather 32u4 RFM95." [Online]. Available: <https://www.adafruit.com/product/3078>
- [168] HOPERF, "Low power long range transceiver module." [Online]. Available: <https://www.hoperf.com/modules/lora/RFM95.html>
- [169] "NI USRP-292x/293x," <http://www.ni.com/datasheet/pdf/en/ds-355>.
- [170] C. Yang and G. Wei, "A noniterative frequency estimator with rational combination of three spectrum lines," *IEEE Trans. Signal Proc.*, vol. 59, no. 10, pp. 5065–5070, Oct 2011.
- [171] "MiniCircuits ZX60-33LN+ LNA." [Online]. Available: <https://www.minicircuits.com/pdfs/ZX60-33LN+.pdf>
- [172] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, Jan. 2014.
- [173] "Texas Instruments ADS4246 datasheet," <http://www.ti.com/lit/ds/symlink/ads4246.pdf>.
- [174] Y.-S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 5992–6010, Dec. 2013.
- [175] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. ACM Conference on SIGCOMM*, Aug. 2013, pp. 375–386.
- [176] S. Ramakrishnan, L. Calderin, A. Niknejad, and B. Nikolic, "An FD/FDD transceiver with RX band thermal, quantization, and phase noise rejection and >64dB TX signal cancellation," in *IEEE Radio Frequency Integrated Circuits Symposium*, Jun. 2017, pp. 352–355.
- [177] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, Seattle, WA, Apr. 2014, pp. 359–372.

-
- [178] A. Balatsoukas-Stimming, P. Belanovic, K. Alexandris, and A. Burg, "On self-interference suppression methods for low-complexity full-duplex MIMO," in *Asilomar Conference on Signals, Systems and Computers*, Nov. 2013, pp. 992–997.
- [179] T. Riihonen, P. Mathecken, and R. Wichman, "Effect of oscillator phase noise and processing delay in full-duplex OFDM repeaters," in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2012.
- [180] S. Shao, X. Quan, Y. Shen, and Y. Tang, "Effect of phase noise on digital self-interference cancellation in wireless full duplex," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 2759–2763.
- [181] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancellation in wireless full-duplex," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4494–4510, Nov. 2013.
- [182] E. Ahmed and A. M. Eltawil, "On phase noise suppression in full-duplex systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1237–1251, 2015.
- [183] V. Syrjala, M. Valkama, L. Anttila, T. Riihonen, and D. Korpi, "Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex OFDM radio transceivers," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 2977–2990, June 2014.
- [184] Y. Hua, P. Liang, Y. Ma, A. C. Cirik, and Q. Gao, "A method for broadband full-duplex MIMO radio," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 793–796, 2012.
- [185] D. Korpi, L. Anttila, V. Syrjala, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1674–1687, Sept. 2014.
- [186] D. Korpi, T. Riihonen, V. Syrjala, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.
- [187] S. Li and R. Murch, "An investigation into baseband techniques for single-channel full-duplex wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4794–4806, Jul. 2014.
- [188] L. Anttila, D. Korpi, V. Syrjälä, and M. Valkama, "Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers," in *2013 Asilomar Conference on Signals, Systems and Computers*, Nov. 2013.
- [189] "National Instruments 5791R user manual and specifications," <http://www.ni.com/pdf/manuals/373845d.pdf>.
- [190] "Texas Instruments DAC3482 datasheet," <http://www.ti.com/lit/ds/symlink/dac3482.pdf>.

Bibliography

- [191] L. Smaïni, *RF Analog Impairments Modeling for Communication Systems Simulation: Application to OFDM-based Transceivers*. Chichester: Wiley, 2012.
- [192] J. Tubbax, L. Van der Perre, S. Donnay, M. Engels, M. Moonen, and H. De Man, "Joint compensation of IQ imbalance, frequency offset and phase noise in OFDM receivers," *European Transactions on Telecommunications*, vol. 15, no. 3, pp. 283–292, 2004.
- [193] L. Anttila, D. Korpi, E. Antonio-Rodriguez, R. Wichman, and M. Valkama, "Modeling and efficient cancellation of nonlinear self-interference in MIMO full-duplex transceivers," in *Globecom Workshops*, Dec. 2014, pp. 777–783.
- [194] D. Korpi, L. Anttila, and M. Valkama, "Nonlinear self-interference cancellation in MIMO full-duplex transceivers under crosstalk," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, Dec. 2017.
- [195] A. Balatsoukas-Stimming, "Non-linear digital self-interference cancellation for in-band full-duplex radios using neural networks," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2018, pp. 1–5.
- [196] D. Morgan, M. Zhengxiang, K. Jaehyeong, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.
- [197] C. Eun and E. Powers, "A new Volterra predistorter based on the indirect learning architecture," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 223–227, Jan. 1997.
- [198] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electron. Lett.*, vol. 37, no. 23, pp. 1417–1418, Nov. 2001.
- [199] H. Guo, J. Xu, S. Zhu, and S. Wu, "Realtime software defined self-interference cancellation based on machine learning for in-band full duplex wireless communications," in *2018 International Conference on Computing, Networking and Communications (ICNC)*, Mar. 2018, pp. 779–783.
- [200] A. T. Kristensen, A. Burg, and A. Balatsoukas-Stimming, "Advanced machine learning techniques for self-interference cancellation in full-duplex radios," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, Nov. 2019, pp. 1149–1153.
- [201] W.-Y. Lee and I. F. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3845–3857, Oct. 2008.
- [202] M. Timmers, S. Pollin, A. Dejonghe, A. Bahai, L. Van der Perre, and F. Catthoor, "Accumulative interference modeling for cognitive radios with distributed channel access," in *2008 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2008)*, May 2008, pp. 1–7.

-
- [203] A. A. Boulogeorgos, N. D. Chatzidiamantis, G. K. Karagiannidis, and L. Georgiadis, "Energy detection under RF impairments for cognitive radio," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, Jun. 2015, pp. 955–960.
- [204] A. A. Boulogeorgos, N. D. Chatzidiamantis, and G. K. Karagiannidis, "Energy detection spectrum sensing under RF imperfections," *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2754–2766, Jul. 2016.
- [205] S. K. Sharma, T. E. Bogale, S. Chatzinotas, B. Ottersten, L. B. Le, and X. Wang, "Cognitive radio techniques under practical imperfections: A survey," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 1858–1884, Fourthquarter 2015.
- [206] W. Afifi and M. Krunz, "Adaptive transmission-reception-sensing strategy for cognitive radios with full-duplex capabilities," in *IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)*, Apr. 2014, pp. 149–160.
- [207] —, "Incorporating self-interference suppression for full-duplex operation in opportunistic spectrum access systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2180–2191, Apr. 2015.
- [208] Y. Liao, L. Song, Z. Han, and Y. Li, "Full duplex cognitive radio: a new design paradigm for enhancing spectrum usage," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 138–145, May 2015.
- [209] Y. Liao, T. Wang, L. Song, and Z. Han, "Listen-and-Talk: Protocol design and analysis for full-duplex cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 656–667, Jan. 2017.
- [210] Y. Liao, T. Wang, L. Song, and B. Jiao, "Cooperative spectrum sensing for full-duplex cognitive radio networks," in *2014 IEEE International Conference on Communication Systems*, Nov. 2014, pp. 56–60.
- [211] A. A. Boulogeorgos, H. A. B. Salameh, and G. K. Karagiannidis, "On the effects of I/Q imbalance on sensing performance in full-duplex cognitive radios," in *2016 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, Apr. 2016, pp. 361–366.
- [212] —, "Spectrum sensing in full-duplex cognitive radio networks under hardware imperfections," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 3, pp. 2072–2084, Mar. 2017.
- [213] A. A. Boulogeorgos and G. K. Karagiannidis, "Energy detection in full-duplex systems with residual RF impairments over fading channels," *IEEE Wireless Communications Letters*, vol. 7, no. 2, pp. 246–249, Apr. 2018.
- [214] W. Cheng, X. Zhang, and H. Zhang, "Full-duplex spectrum-sensing and MAC-protocol for multichannel nontime-slotted cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 5, pp. 820–831, May 2015.

Bibliography

- [215] L. T. Tan and L. B. Le, "Distributed MAC protocol design for full-duplex cognitive radio networks," in *2015 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2015, pp. 1–6.
- [216] H. Kim, S. Lim, H. Wang, and D. Hong, "Optimal power allocation and outage analysis for cognitive full duplex relay systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3754–3765, Oct. 2012.
- [217] M. Amjad, F. Akhtar, M. H. Rehmani, M. Reisslein, and T. Umer, "Full-duplex communication in cognitive radio networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 19, no. 4, pp. 2158–2191, Fourthquarter 2017.
- [218] A. Bayat and S. Aïssa, "Full-duplex cognitive radio with asynchronous energy-efficient sensing," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1066–1080, Feb. 2018.
- [219] H. Mokhtarzadeh, A. Taherpour, A. Taherpour, and S. Gazor, "Throughput maximization in energy limited full-duplex cognitive radio networks," *IEEE Transactions on Communications*, vol. 67, no. 8, pp. 5287–5296, Aug. 2019.
- [220] D. Li, J. Cheng, and V. C. M. Leung, "Adaptive spectrum sharing for half-duplex and full-duplex cognitive radios: From the energy efficiency perspective," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5067–5080, Nov. 2018.
- [221] M. R. Mortada, A. Nasser, A. Mansour, and K. C. Ya, "Novel sensing mechanism for full-duplex secondary users in cognitive radio," in *2019 27th European Signal Processing Conference (EUSIPCO)*, Sep. 2019, pp. 1–5.
- [222] Y. Zhang, J. Hou, V. Towhidlou, and M. R. Shikh-Bahaei, "A neural network prediction-based adaptive mode selection scheme in full-duplex cognitive networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 540–553, Sep. 2019.
- [223] M. K. Hanawal, D. N. Nguyen, and M. Krunz, "Cognitive networks with in-band full-duplex radios: Jamming attacks and countermeasures," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 296–309, Mar. 2020.
- [224] D. Li, D. Zhang, and J. Cheng, "Degrees of freedom for half-duplex and full-duplex cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2571–2584, Mar. 2019.
- [225] A. C. Cirik, R. Wang, Y. Rong, and Y. Hua, "MSE-based transceiver designs for full-duplex MIMO cognitive radios," *IEEE Transactions on Communications*, vol. 63, no. 6, pp. 2056–2070, Jun. 2015.
- [226] A. C. Cirik, S. Biswas, O. Taghizadeh, and T. Ratnarajah, "Robust transceiver design in full-duplex MIMO cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 2, pp. 1313–1330, Feb. 2018.

- [227] H. Shehata and T. Khattab, "Energy detection spectrum sensing in full-duplex cognitive radio: The practical case of Rician RSI," *IEEE Transactions on Communications*, vol. 67, no. 9, pp. 6544–6555, Sep. 2019.
- [228] W. Cheng, X. Zhang, and H. Zhang, "Full/half duplex based resource allocations for statistical quality of service provisioning in wireless relay networks," in *2012 Proceedings IEEE INFOCOM*, Mar. 2012, pp. 864–872.
- [229] P. C. Sofotasios, M. K. Fikadu, S. Muhaidat, Q. Cui, G. K. Karagiannidis, and M. Valkama, "Full-duplex regenerative relaying and energy-efficiency optimization over generalized asymmetric fading channels," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3232–3251, May 2017.
- [230] W. Cheng, X. Zhang, and H. Zhang, "Qos driven power allocation over full-duplex wireless links," in *2012 IEEE International Conference on Communications (ICC)*, Jun. 2012, pp. 5286–5290.
- [231] R. Zhang, M. Ma, D. Li, and B. Jiao, "Investigation on DL and UL power control in full-duplex systems," in *IEEE International Conference on Communications (ICC)*, June 2015, pp. 1903–1907.
- [232] M. Liu, Y. Mao, S. Leng, X. Huang, and Q. Zhao, "Traffic-aware resource allocation for full duplex wireless networks," in *IEEE/CIC International Conference on Communications in China (ICCC)*, 2016. IEEE, Jul. 2016, pp. 1–6.
- [233] G. Liu, X. Chen, Z. Ding, Z. Ma, and F. R. Yu, "Hybrid half-duplex/full-duplex cooperative non-orthogonal multiple access with transmit power adaptation," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 506–519, Jan. 2018.
- [234] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1077–1091, Mar. 2017.
- [235] G. Zhang, H. Zhang, Z. Han, and G. K. Karagiannidis, "Spectrum allocation and power control in full-duplex ultra-dense heterogeneous networks," *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4365–4380, Jun. 2019.
- [236] N. Nomikos, T. Charalambous, D. Vouyioukas, R. Wichman, and G. K. Karagiannidis, "Power adaptation in buffer-aided full-duplex relay networks with statistical CSI," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7846–7850, Aug. 2018.
- [237] "National Instruments PXIe-1082 user manual," <http://www.ni.com/pdf/manuals/372752b.pdf>.
- [238] O. Gustafsson, K. Amiri, D. Andersson, A. Blad, C. Bonnet, J. R. Cavallaro, J. Declerck, A. Dejonghe, P. Eliardsson, M. Glasse, A. Hayar, L. Hollevoet, C. Hunter, M. Joshi, F. Kaltenberger, R. Knopp, K. Le, Z. Miljanic, P. Murphy, F. Naessens, N. Nikaein, D. Nussbaum, R. Pacalet, P. Raghavan, A. Sabharwal, O. Sarode, P. Spasojevic, Y. Sun, H. M.

Bibliography

- Tullberg, T. V. Aa, L. Van der Perre, M. Wetterwald, and M. Wu, "Architectures for cognitive radio testbeds and demonstrators — an overview," in *2010 Proceedings of the Fifth International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Jun. 2010, pp. 1–6.
- [239] "Full-duplex cognitive radio demo - Telecommunications Circuits Laboratory." [Online]. Available: <https://www.youtube.com/watch?v=ENnFAYJsDXo>
- [240] "IEEE standard for information technology—telecommunications and information exchange between systems local and metropolitan area network—specific requirements. Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," IEEE Computer Society, Tech. Rep., 2012.
- [241] M. G. Sarret, M. Fleischer, G. Berardinelli, N. H. Mahmood, P. Mogensen, and H. Heinz, "On the potential of full duplex performance in 5G ultra-dense small cell networks," in *IEEE European Signal Processing Conference (EUSIPCO)*, preprint, Aug. 2016.
- [242] G. Berardinelli, D. A. Wassie, N. H. Mahmood, M. G. Sarret, T. B. Sørensen, and P. Mogensen, "Evaluating full duplex potential in dense small cells from channel measurements," in *IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016.
- [243] D. Murray, T. Koziniec, and M. Dixon, "Solving Ack inefficiencies in 802.11 networks," in *IEEE International Conference on Internet Multimedia Services Architecture and Applications*, Dec. 2009, pp. 1–6.
- [244] M. G. Sarret, G. Berardinelli, N. H. Mahmood, and P. Mogensen, "Can full duplex boost throughput and delay of 5G ultra-dense small cell networks?" in *IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016.
- [245] Y. Guo and Z. Liu, "Time-delay-estimation-liked detection algorithm for LoRa signals over multipath channels," *IEEE Wireless Communications Letters*, Mar. 2020.
- [246] U. Noreen, L. Clavier, and A. Bounceur, "LoRa-like CSS-based PHY layer, capture effect and serial interference cancellation," in *European Wireless 2018; 24th European Wireless Conference*, May 2018, pp. 1–6.
- [247] N. E. Rachkidy, A. Guitton, and M. Kaneko, "Decoding superposed LoRa signals," in *2018 IEEE 43rd Conference on Local Computer Networks (LCN)*, Oct. 2018, pp. 184–190.
- [248] N. El Rachkidy, A. Guitton, and M. Kaneko, "Collision resolution protocol for delay and energy efficient LoRa networks," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 2, pp. 535–551, Jun. 2019.
- [249] B. Laporte-Fauret, M. A. Ben Temim, G. Ferre, D. Dallet, B. Minger, and L. Fuché, "An enhanced LoRa-like receiver for the simultaneous reception of two interfering signals," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2019, pp. 1–6.

ORION AFISIADIS

CONTACT DETAILS

Telecommunications Circuits Laboratory
École Polytechnique Fédérale de Lausanne
EPFL-STI-IEL-TCL, ELG 036, Station 11
1015 Lausanne, Switzerland

E-mail: orion.afisiadis@epfl.ch
Phone: +41 21 693 7854
Date of birth: October 27th, 1986
Last update: June 11, 2020

EDUCATION

- **PhD** in Electrical Engineering 2015 – today
École polytechnique fédérale de Lausanne, Lausanne, Switzerland
- **Diploma** in Electrical and Computer Engineering 2015
Aristotle University of Thessaloniki, Thessaloniki, Greece

TEACHING EXPERIENCE

- **Teaching Assistant**
École polytechnique fédérale de Lausanne
 - *EE-442 Wireless Receivers: Algorithms and Architectures* (MSc) 2013–2020
 - *EE-543 Advanced Wireless Receivers: Algorithms and Architectures* (MSc) 2014–2020
 - *EE-390(c) TP en technologies de l'information* (MSc) 2016–2020

STUDENT SUPERVISION

- **MSc, BSc, and Intern Supervision**
 - Mr. Paul Mayoraz, BSc student at EPFL 2020
Bachelor End Project: *Correction algorithms of fractional carrier frequency offset in LoRa receivers*
 - Mr. Joachim Tapparel, MSc student at EPFL 2020
Internship: *GNU Radio Implementation of LoRa PHY*
 - Mr. Joachim Tapparel, MSc student at EPFL 2019
Project: *USRP implementation of LoRa Physical Layer for Multiple Spreading Factors and Coding Rates*
 - Mr. Juan Guirado, BSc exchange student at EPFL 2018
Bachelor End Project: *Hybrid ARQ Scheme Using In-Band Full-Duplex LoRa*
 - Mr. Juan Antonio Martin Galicia, MSc student at EPFL 2018
Project: *Reverse engineering LoRa header and CRC*
 - Mr. Joffrey Martinelli, MSc student at EPFL 2017
Project: *USRP implementation of LoRa LPWAN physical layer*
 - Mr. Guillaume Jaume & Mr. Matthieu Cotting, MSc students at EPFL 2016
Project: *LoRa PHY implementation and analysis*

PROFESSIONAL ACTIVITIES

- **Journal Reviewer:** IEEE Trans. on Wireless Communications, IEEE Trans. on Vehicular Technology, IEEE Wireless Communications Letters, IEEE Communications Letters
- **Professional Memberships:** IEEE Member (Communications, Signal Processing, and Vehicular Societies).

PUBLICATIONS

Google Scholar profile: <https://scholar.google.com/citations?user=GcfWchIAAAAJ&hl=en&oi=ao>

Total citations: 124

h-index: 3

In Review

1. J. Tapparel, **O. Afisiadis**, P. Mayoraz, A. Balatsoukas-Stimming, and A. Burg, “An open-source LoRa physical layer prototype on GNU Radio,” in *IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2020

Refereed Journal Publications

1. **O. Afisiadis**, M. Cotting, A. Burg, and A. Balatsoukas-Stimming, "On the error rate of the LoRa modulation with interference," *IEEE Transactions on Wireless Communications*, Feb. 2020

Refereed Conference Publications

1. **O. Afisiadis**, A. Burg, and A. Balatsoukas-Stimming, "Coded LoRa frame error rate analysis," in *IEEE International Conference on Communications (ICC)*, Jun. 2020
2. **O. Afisiadis**, M. Cotting, A. Burg, and A. Balatsoukas-Stimming, "LoRa symbol error rate under non-aligned interference," in *Asilomar Conference on Signals, Systems, and Computers*, Nov. 2019
3. R. Ghanaatian, **O. Afisiadis**, M. Cotting, and A. Burg, "LoRa digital receiver analysis and implementation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019
4. **O. Afisiadis**, A. C. M. Austin, A. Balatsoukas-Stimming, and A. Burg, "Analysis of full-duplex wireless links with asymmetric capacity requirements," in *Asilomar Conference on Signals, Systems, and Computers*, Nov. 2017
5. A. C. M. Austin, **O. Afisiadis**, A. Balatsoukas-Stimming, and A. Burg, "Digital predistortion of hardware impairments for full-duplex transceivers," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Nov 2017
6. **O. Afisiadis**, A. C. M. Austin, A. Balatsoukas-Stimming, and A. Burg, "Sliding window spectrum sensing for full-duplex cognitive radios with low access-latency," in *IEEE Vehicular Technology Conference*, May 2016
7. **O. Afisiadis**, A. Balatsoukas-Stimming, and A. Burg, "A low-complexity improved successive cancellation decoder for polar codes," in *Asilomar Conference on Signals, Systems, and Computers*, Nov. 2014

Invited Talks

1. "Lora physical layer modeling and performance analysis: Theory and testbed prototype," TU/e CWTe colloquium (invited by Alexios Balatsoukas-Stimming), Feb 2020

Demos

1. A. C. M. Austin, **O. Afisiadis**, A. Balatsoukas-Stimming, and A. Burg, "Concurrent spectrum sensing and transmission for cognitive radio using self-interference cancellation," in *ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC)*, Jun. 2015