Physical-Layer Security for Visible-Light Communication Systems

by

Ayman Mostafa

B.Sc., Alexandria University, 2006
M.A.Sc., McMaster University, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2017

Abstract

Visible-light communication (VLC) is an enabling technology that exploits the lighting infrastructure to provide ubiquitous indoor broadband coverage via high-speed short-range wireless communication links. On the other hand, physical-layer security has the potential to supplement conventional encryption methods with an additional secrecy measure that is provably unbreakable regardless of the computational power of the eavesdropper.

The lack of wave-guiding transmission media in VLC channels makes the communication link inherently susceptible to eavesdropping by unauthorized users existing in areas illuminated by the data transmitters. In this thesis, we study transmission techniques that enhance the secrecy of VLC links within the framework of physical-layer security.

Due to linearity limitations of typical light-emitting diodes (LEDs), the VLC channel is more accurately modelled with amplitude constraints on the channel input, rather than the conventional average power constraint. Such amplitude constraints render the prevalent Gaussian input distribution infeasible for VLC channels, making it difficult to obtain closed-form secrecy capacity expressions. Thus, we begin with deriving lower bounds on the secrecy capacity of the Gaussian wiretap channel subject to amplitude constraints.

We then consider the design of optimal beamformers for secrecy rate maximization in the multiple-input single-output (MISO) wiretap channel under amplitude constraints. We show that the design problem is nonconvex and difficult to solve, however it can be recast as a solvable quasiconvex line search problem. We also consider the design of robust beamformers for worst-case secrecy rate maximization when channel uncertainty is taken into account.

Finally, we study the design of linear precoders for the two-user MISO broadcast channel with confidential messages. We consider not only amplitude constraints, but also total and per-antenna average power constraints. We formulate the design problem as a nonconvex weighted secrecy sum rate maximization problem, and provide an efficient search algorithm to obtain a solution for such a nonconvex problem. We extend our approach to handle uncertainty in channel information.

The design techniques developed throughout the thesis provide valuable tools for tackling real-world problems in which channel uncertainty is almost always inevitable and amplitude constraints are often necessary to accurately model hardware limitations.

Preface

This thesis is based on research work performed under the supervision of Professor Lutz Lampe. For all the chapters, as well as the corresponding publications, I conducted the literature surveys, formulated the problems, proposed the solutions, performed the analyses, implemented the simulations, and wrote the manuscripts. Professor Lampe helped by guiding the direction of research, validating the analyses, and providing feedback to improve the manuscripts.

The content of **Chapter 2** has been published in the following papers:

- A. Mostafa and L. Lampe, “Physical-Layer Security for MISO Visible Light Communication Channels,” in *IEEE Journal on Selected Areas in Communications - Special Issue on Optical Wireless Communications*, vol. 33, no. 9, pp. 1806–1818, Sept. 2015.

- A. Mostafa and L. Lampe, “Physical-Layer Security for Indoor Visible Light Communications,” in *Proceedings of 2014 IEEE International Conference on Communications (ICC)*, Sydney, NSW, Australia, pp. 3342–3347, Jun. 2014.

- A. Mostafa and L. Lampe, “Securing Visible Light Communications via Friendly Jamming,” in *Proceedings of 2014 IEEE Globecom Workshops (GC Wkshps)*, Austin, TX, USA, pp. 524–529, Dec. 2014.

- A. Mostafa and L. Lampe, “Enhancing the Security of VLC Links: Physical-Layer Approaches,” in *Proceedings of 2015 IEEE Summer Topicals Meeting Series (SUM)*, Invited Presentation, Nassau, Bahamas, pp. 39–40, Jul. 2015.

The content of **Chapter 3** has been published in the following papers:

- A. Mostafa and L. Lampe, “Optimal and Robust Beamforming for Secure Transmission in MISO Visible-Light Communication Links,” in *IEEE Transactions on Signal Processing*, vol. 64, no. 24, pp. 6501–6516, Dec. 2016.

- A. Mostafa and L. Lampe, “Pattern Synthesis of Massive LED Arrays for Secure Visible Light Communication Links,” in *Proceedings of 2015 IEEE International Conference on Communication Workshop (ICCW 2015)*, London, UK, pp. 1350–1355, Jun. 2015.

The content of **Chapter 4** has been submitted for publication:

- A. Mostafa and L. Lampe, “On Linear Precoding for the Two-User MISO Broadcast Channel with Confidential Messages and Per-Antenna Constraints,” submitted in Jan. 2017.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

5G	5th Generation
BC-CM	Broadcast Channel with Confidential Messages
CDMA	Code-Division Multiple Access
DAC	Digital-to-Analog Converter
DC	Direct Current
DD	Direct Detection
FoV	Field-of-View
GaN	Gallium Nitride
GEV	Generalized Eigenvalue
i.i.d.	Independent and identically-distributed
IEEE	Institute of Electrical and Electronics Engineers
IM	Intensity Modulation
IoT	Internet of Things
LD	Laser Diode
LED	Light-Emitting Diode
LoS	Line-of-Sight
MC	Multi-Carrier
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MTC	Machine-Type Communication

NLoS	Non-Line-of-Sight
OOK	On-Off Keying
PAM	Pulse-Amplitude Modulation
PD	Photodiode
PDF	Probability Density Function
PLC	Power-Line Communication
QAM	Quadrature Amplitude Modulation
RF	Radio Frequency
S-DPC	Secret Dirty-Paper Coding
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
VLC	Visible-Light Communication
VPPM	Variable Pulse-Position Modulation
w.r.t.	With respect to
ZF	Zero-Forcing

Notation

Alice	The transmitter
Bob	The (intended) receiver
Eve	The eavesdropper
$\mathbb{R}^N$	The set of N-dimensional real-valued numbers
$\mathbb{R}_+^N$	The set of N-dimensional nonnegative real-valued numbers
$\mathbf{0}$	The all-zero column vector
$\mathbf{1}_N$	The all-one column vector with length N
$\mathbf{I}_N$	The N-dimensional identity matrix
$(\cdot)^{\mathrm{T}}$	Transpose
$[x]^+$	$\max\{x, 0\}$
$\|x\|$	Absolute value of x
$\\|\mathbf{x}\\|_p$	L_p-norm of the vector $\mathbf{x}$, $p \geq 1$
$\\|\mathbf{x}\\|_\infty$	Chebyshev or l_∞-norm of the vector $\mathbf{x}$
$\\|\mathbf{X}\\|_{\mathrm{F}}$	Frobenius norm of the matrix $\mathbf{X}$
$\mathrm{Tr}(\mathbf{X})$	Trace of the square matrix $\mathbf{X}$
$\mathbf{Diag}(x_1, \ldots, x_N)$	The diagonal matrix with diagonal elements $x_1, \ldots, x_N$
$\otimes$	Kronecker product
$I(\cdot)$	Indicator function
$\lceil \cdot \rceil$	Ceiling function
$\log_b(\cdot)$	The logarithm to base b, $b \in \{2, 10\}$

$\ln(\cdot)$	The natural logarithm (i.e., the logarithm to base e)
$p(x)$	Probability density function for the random variable X
$\mathbb{E}\{X\}$	Expected value of the random variable X
$\mathbb{var}\{X\}$	Variance of the random variable X
$\mathbb{h}(X)$	Differential entropy of the random variable X
$\mathbb{D}(p(\cdot)\\|q(\cdot))$	Relative entropy between the distributions P and Q
$\mathbb{I}(X;Y)$	Mutual information between the random variables X and Y
$\mathcal{N}(0,\sigma^2)$	The Gaussian distribution with zero mean and variance σ^2
$\mathcal{U}[-a,a]$	The uniform distribution over the interval $[-a,a]$
$\mathcal{Q}(\cdot)$	The $\mathcal{Q}$-function
$\nabla_{\text{sub}} f(\mathbf{x})$	A subgradient of the function f at $\mathbf{x}$
$\{\cdot\}_{\mathrm{B}}$	A quantity relevant to Bob, the intended receiver
$\{\cdot\}_{\mathrm{E}}$	A quantity relevant to Eve, the eavesdropper

Acknowledgments

It has been a true pleasure and an enjoyable learning experience to pursue my doctoral degree under the supervision of Professor Lutz Lampe. Over the course of the past four years, I found his door always open, and I was extremely fortunate to work with such a distinguished researcher and amazing supervisor. I am sincerely grateful for his guidance, patience, encouragement, and support in so many ways.

I thank the members of my examining committee at UBC, Professors David Michelson, Julian Cheng (UBC Okanagan), Vikram Krishnamurthy, Sathish Gopalakrishnan, and Brian Marcus (Mathematics), for their time and insightful comments. Special thanks go to Professor Ashish Khisti at the University of Toronto for serving as the external examiner, and also for participating in the final exam.

I thank the Natural Sciences and Engineering Research Council of Canada (NSERC). This thesis would not have been possible without their funding and support.

I thank my colleagues in the Communications Lab at Kaiser 4090 for the wonderful times and interesting discussions we had over the years. Their companionship has significantly enriched my experience at UBC.

I also take the opportunity to express my sincere gratitude to my Master's supervisor at McMaster University, Professor Steve Hranilovic. He taught me how to be a researcher, and with whom I wrote my very first research paper.

I am deeply indebted to my wife, Shereen, for her continuous love and support. She shared with me the difficult journey of graduate studies right from the beginning, and was always there whenever I needed her encouragement. I also thank my sons, Mostafa and Omar, who always cheered for me, thinking I am a smart person and a great dad. Their excitement when I go back home was my biggest rewarding moment everyday.

Last, but definitely not least, I am forever indebted to my Parents. Whatever I have accomplished is because of them, and in fact my biggest motivation was to please them. Words would fall short to express how much I am grateful for their care, sacrifice, encouragement, endless love, and unconditional support.

Ayman Mostafa

April 19, 2017
Vancouver, BC

Dedication

To my Parents, my wife Shereen, and my sons Mostafa and Omar

Chapter 1

Introduction

1.1 Background

1.1.1 Visible-Light Communication

Visible-light communication (VLC) is an enabling technology that exploits illumination devices, mostly high-brightness light-emitting diodes (LEDs), to establish high-speed short-range wireless communication links [1, 2, 3, 4, 5, 6, 7]. In typical VLC systems, information is relayed by the means of modulating the output intensity of the LEDs, whereas at the receiver side, the data signal is recovered using simple photodiodes (PDs). The use of laser diodes (LDs), rather than LEDs, can result in higher data rates [8], however LDs are not popular for illumination purposes as they are more expensive and have the potential to cause eye or skin injuries. Data rates can also be improved by using imaging receivers [8, 9], however this comes with increased complexity and cost.

VLC systems take advantage of the license-free light spectrum and immunity to radio frequency (RF) interference. In addition, VLC transmitters can exploit the existing lighting infrastructure where legacy incandescent and fluorescent lamps are being replaced with LED-based luminaires that have longer lifespan, smaller size, lower power consumption, higher energy-conversion efficiency, and improved color rendering without using toxic chemicals [2, 5, 10]. Thus, the integration between power-line communication (PLC) and VLC systems has the potential to provide ubiquitous indoor broadband coverage with seamless handover [11, 12]. Furthermore, since typical lighting systems utilize multiple luminaires that are sufficiently separated to provide uniform illumination, VLC systems can readily benefit from multiple-antenna techniques to achieve higher data rates [9] and enhance the reliability [13] and security [14] of VLC networks. Moreover, due to line-of-sight (LoS) propagation and confinement of light waves by opaque surfaces, VLC links cause limited inter-cell interference. Such advantages qualify VLC systems for realizing small-size cells, termed as “LiFi attocells” [15], in fifth generation (5G) networks featuring cells with coverage ranges on the order of a few meters [16].

The IEEE 802.15.7 standard [17], released in 2011, was a big step towards the commercialization and widespread deployment of VLC networks [18, 19]. It defines three physical layer modes, with the second and third modes, PHY II and PHY III, respectively, supporting data rates up to 96 Mbit/sec [20, 21]. In fact, much higher data rates have already been demonstrated in laboratory conditions. A prototype VLC system utilizing high-power LEDs to achieve bidirectional real-time transmission with a total rate of 500 Mbit/sec over a 2 m distance was implemented in [22]. In [23], the authors demonstrated a 16-user multi-carrier code-division multiple access (MC-CDMA) VLC system that achieves 750 Mbit/sec sum rate over a 1.5 m distance using off-the-shelf LEDs. Furthermore, the use of μLEDs with smaller size (e.g., on the order of 50 μm) and lower junction capacitance allows higher modulation bandwidth. The authors in [24] utilized a single gallium nitride (GaN) μLED to establish a 3 Gbit/sec VLC link over a 5 cm distance.

1.1.2 Physical-Layer Security

With the unprecedented increase in traffic volumes over wireless networks, data privacy and secrecy are becoming a major concern for users, as well as for network administrators. Conventional security schemes are typically implemented at upper layers of the network stack via access control, password protection, and end-to-end encryption. Such schemes are deemed secure as long as the computational power of potential eavesdroppers remains below certain limits. For example, the eavesdroppers do not have sufficient computational power to perform an exhaustive search for the password, or determine the prime factors of a large integer (to obtain the secret key and decrypt the encrypted message), within a reasonable amount of time. During the past few years, however, physical-layer security has emerged as a promising technique that can complement conventional encryption methods with an additional secrecy layer that is provably unbreakable regardless of the computational power of the eavesdroppers [25, 26, 27, 28, 29, 30, 31]. Moreover, physical-layer security has the potential to provide lightweight standalone secrecy solutions in communication systems functioning under severe hardware or energy constraints such as machine-type communication (MTC) devices in the Internet of Things (IoT) [32].

Physical-layer security refers to transmission schemes that exploit dissimilarities among the channels of different receivers in order to hide information from unauthorized receivers, without reliance on upper-layer encryption techniques. The underlying idea behind such a secrecy scheme is to sacrifice a portion of the communication rate, that otherwise would be used for useful data transmission, in order to confuse potential eavesdroppers and diminish their capability to infer information at any positive rate, via carefully-designed signaling and coding schemes.

The innovative idea of quantifying secrecy via information-theoretic measures can be traced back to Shannon [33] who proposed *equivocation*[1] as a quantitative measure of the secrecy level of encrypted messages [26, Section 3.1]. Almost three decades later, the foundations of information-theoretic security were laid down by Wyner in his seminal paper [34] that studied the problem of secret communication over the *degraded* broadcast channel. In that paper, Wyner introduced the so-called *wiretap channel* model to describe the scenario in which the transmitter has one secret message intended for one receiver, while the other receiver, whose channel is degraded, acts as an eavesdropper. Wyner also introduced the notion of *secrecy capacity* as a performance measure that specifies the maximum communication rate that guarantees reliable reception of the secret message by the intended receiver and entire hiddenness from the eavesdropper. The work of Wyner motivated the characterization of the secrecy capacity of the scalar, i.e., the single-input single-output (SISO), Gaussian wiretap channel[2] [36]. Wyner's model was then extended to the (nondegraded) wiretap channel [37], where the eavesdropper's channel need not be degraded. Such an extension has ultimately led to the characterization of the secrecy capacity of the multiple-input single-output (MISO) [38, 39] and multiple-input multiple-output (MIMO) Gaussian wiretap channels [40, 41, 42, 43]. Furthermore, when the eavesdropper's channel is not accurately known or entirely unknown to the transmitter, the works in [44, 45, 46] proposed the transmission of jamming signals, i.e., artificial noise, in conjunction with the information-bearing signal, in order to increase the interference seen by the eavesdropper and diminish its capability to decode the secret message.

The wiretap channel model was further extended to the two-user broadcast channel with confidential messages (BC-CM) [47]. Such a model studies the scenario in

[1] As defined in his paper [33, Section 11], equivocation is the conditional entropy of the transmitted message after knowing the received signal.

[2] Recall that the scalar Gaussian broadcast channel is a degraded channel [35, Section 15.1.3].

which the transmitter has two independent secret messages, one intended for each receiver, and each message should be kept confidential from the other receiver. The secrecy capacity regions of the two-user MISO BC-CM and the two-user MIMO BC-CM were characterized in [48] and [49], respectively.

1.2 Motivation

1.2.1 Are VLC Links Secure?

VLC links are often deemed eavesdropping-proof, however this is not necessarily true, especially in public areas or in multi-user scenarios. With the lack of optical fibers, or any sort of wave-guiding transmission media, the VLC channel has a broadcast nature. This makes VLC links inherently susceptible to eavesdropping by unintended or unauthorized users having access to areas illuminated by the data transmitters. Typical scenarios include public spaces such as classrooms, meeting rooms, libraries, shopping centers, and aircrafts, to name a few.

Accordingly, our research efforts in this thesis are directed towards enhancing the secrecy of VLC networks within the framework of physical-layer security.

Figure 1.1 depicts a typical VLC scenario in which physical-layer security is applicable. The figure shows, for example, a governmental office utilizing a VLC network. The shaded area (at the bottom of the figure) is open to the public, making sensitive information vulnerable to overhearing by potential eavesdroppers. An interesting design problem is to devise a physical-layer security scheme that maintains reliable communication among the office personnel and prevents users located in the shaded area from reliably decoding the transmitted messages.

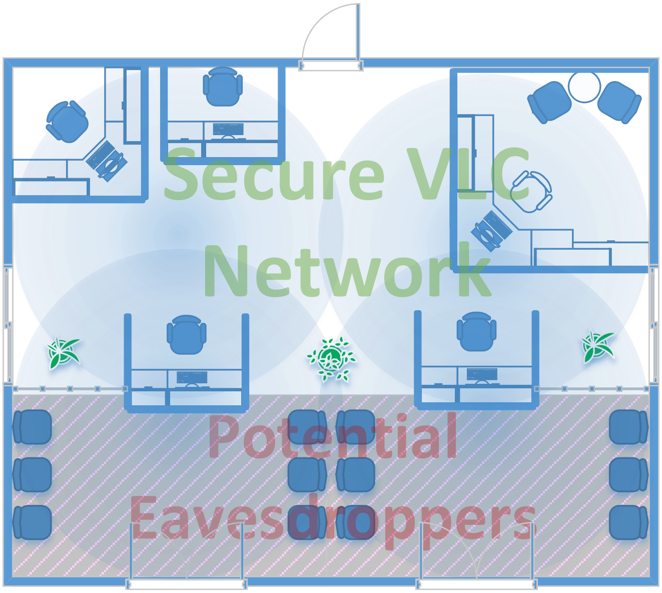

Figure 1.1: An example VLC scenario in which physical-layer security is applicable.

1.2.2 Amplitude Constraints

Typical VLC systems utilize LEDs for data transmission whereby the input current signal modulates the output intensity of the LEDs. Typical LEDs, however, have limited linear operation region beyond which electro-optical conversion becomes nonlinear (see, e.g., Figure 1.3). Such a nonlinearity can be partially compensated via predistortion of the input current signal [50]. However, predistortion can be effective only within certain operation limits beyond which the output intensity saturates, leading to clipping distortion of the transmitted signal. Thus, the modulating current signal must satisfy certain *amplitude constraints* in order to maintain linear electro-optical conversion and avoid undesirable nonlinear effects. As a consequence, intensity modulation (IM) channels are typically modelled with amplitude constraints on the channel input, rather than the conventional average power constraint [51, 52, 53].

In fact, all modern digital transmitters experience amplitude constraints because of the digital-to-analog converters (DACs) incorporated at the transmitter front-end. Clearly, these DACs have finite ranges, and thus the transmitted signals are subject to amplitude constraints. Therefore, taking amplitude constraints into account can be crucial to model hardware limitations, not only in IM systems, but in fact in all practical communication systems.

Now, an amplitude constraint on the channel input will render the prevalent Gaussian input distribution infeasible. Unfortunately, this makes amplitude constraints difficult to handle (in terms of obtaining analytic capacity expressions), and therefore they are often overlooked in favor of the more convenient average power constraint. Compared to the massive body of literature on the Gaussian wiretap channel under the average power constraint, works that considered the amplitude-constrained Gaussian wiretap channel are quite rare. Even in the absence of secrecy constraints, characterization of the capacity of amplitude-constrained Gaussian channels is quite challenging. In his seminal paper [54, Section 26], Shannon referred to the difficulty of obtaining an analytic expression for the capacity of the peak-limited, i.e., the amplitude-constrained, Gaussian channel. Instead, he derived a lower bound and an asymptotic upper bound that is valid at high peak signal-to-noise ratio (SNR). Out of his Ph.D. work [55, 51], Smith came up with the rather surprising result that the capacity-achieving input distribution for the amplitude-constrained Gaussian channel is discrete with finite support, i.e., it has a finite number of mass points. Closed-form lower and upper bounds on the capacity of the amplitude-constrained Gaussian channel were derived in [53]. For the Gaussian wiretap channel, the authors in [56] followed the approach devised in [51] and proved that the secrecy capacity-achieving distribution under the amplitude constraint is also discrete with finite support.

In this thesis, we shall characterize the performance of amplitude-constrained

Gaussian wiretap channels via closed-form secrecy rate expressions. We will also consider the design of beamformers for the MISO wiretap channel and linear precoders for the two-user MISO BC-CM when the beamformers or precoders are subject to amplitude constraints. In fact, it is fair to say that the novelty of many of the problems considered in this thesis comes from taking amplitude constraints into account.

1.2.3 Uncertain Channel Information

Compared to conventional encryption techniques, the performance of physical-layer security schemes is inherently sensitive to channel conditions, that is the secrecy performance can be severely degraded if the designed scheme is based on inaccurate channel information. In fact, this is a major drawback that may hinder any effort to deploy practical physical-layer security systems as it is almost always unrealistic, in real-world scenarios, to assume that the channel gain of the intended receiver or the eavesdropper is accurately known to the transmitter. On one hand, information regarding the intended receiver's channel may suffer from estimation errors, aside from inevitable quantization errors imposed by the finite rate of the feedback channel. On the other hand, there is probably no feedback from the eavesdropper if it is an unregistered user and shall remain silent to hide its presence. In such a case, the transmitter may resort to less reliable information sources, such as possible locations of the eavesdropper, in order to obtain an estimate of its channel gain. In all cases, channel information available to the transmitter will never be accurate, and adopting a physical-layer security scheme based on such inaccurate information may lead to a secrecy outage with catastrophic consequences.

Based on the above discussion, it becomes clear that any practical physical-layer security system must take channel uncertainty into account. In other words, we have to adopt the so-called *robust transmission schemes.* Among various possible approaches to achieve robust secure transmission, we shall consider *worst-case optimization.* In such an approach, one chooses some *uncertainty sets* that are believed to contain all possible realizations of the channel gains for the receiver and the eavesdropper. Then, the design problem is formulated to optimize the performance measure, i.e., the secrecy rate, corresponding to the *worst-case realization* of the uncertain channel gains. Now we have to face the question of how to choose *reasonable* uncertainty sets. In fact, the validity of the worst-case optimization approach depends mostly on such a choice. On one hand, unreasonably large (i.e., too conservative) uncertainty sets may render the design problem infeasible. On the other hand, small uncertainty sets can lead to an overestimate of the achievable secrecy rate and, consequently, secrecy outage may occur.

Typical works in the physical-layer security literature assume *spherical uncertainty sets* for both the receiver's and eavesdropper's channels. For example, uncertainty in the eavesdropper's channel is typically modelled by

$$\mathbf{h}_{\mathrm{E}} \in \left\{ \hat{\mathbf{h}}_{\mathrm{E}} + \mathbf{e} : \|\mathbf{e}\|_2 \leq \epsilon \right\},$$

where $\hat{\mathbf{h}}_{\mathrm{E}}$ is the transmitter's erroneous estimate of the eavesdropper's channel $\mathbf{h}_{\mathrm{E}}$, $\mathbf{e}$ is an unknown (but norm-bounded) error vector, and ϵ is some known constant that quantifies the amount of uncertainty. This error model is well accepted to take into account channel uncertainty caused by limited, i.e., finite-rate, feedback from the receiver [57, Lemma 1]. In wiretap scenarios, however, such an uncertainty model may become inapplicable if the eavesdropper is a passive receiver that remains silent to hide its presence from the transmitter, i.e., there is no feedback, $\hat{\mathbf{h}}_{\mathrm{E}}$, regarding the eavesdropper's channel. Fortunately, in indoor VLC scenarios, it is often reasonable to assume that the transmitter has some uncertain information regarding the location and/or orientation of the eavesdropper (recall, for example, the scenario in Figure 1.1, wherein potential eavesdroppers can only exist within areas of the room that are open to the public). Furthermore, the LoS path is typically dominant in VLC scenarios, and thus the channel gain can be accurately approximated by a deterministic function of the location and orientation of the receiver, as well as the emission pattern of the LEDs (see Eq. (1.4) in Section 1.3.2). This is unlike the case of RF channels wherein rich scattering environments typically give rise to significant multipath components, which are usually unpredictable.

Therefore, in this thesis we develop the idea of choosing uncertainty sets for the eavesdropper's channel based on the uncertain parameters in the LoS channel gain equation, i.e., based on uncertain information regarding the location and orientation of the eavesdropper. Then, we use such uncertainty sets, along with spherical uncertainty sets for the intended receiver's channel, in order to formulate the worst-case secrecy rate maximization problem and obtain a robust transmission scheme.

1.3 Preliminaries and Definitions

In this section, we present some of the key concepts and definitions used throughout the entire thesis. We begin with describing the VLC channel model and the modulation scheme that we adopt. We then recall the *generalized Lambert's cosine law* used to model the emission pattern of typical LEDs. We also explain how transmit beamforming can be implemented in IM channels. Furthermore, we review two fundamental constructs in physical-layer security, namely, the wiretap channel and the two-user BC-CM, and recall the relevant definitions of achievable secrecy rates and secrecy rate regions. Finally, we clarify what the term "secure transmission" precisely

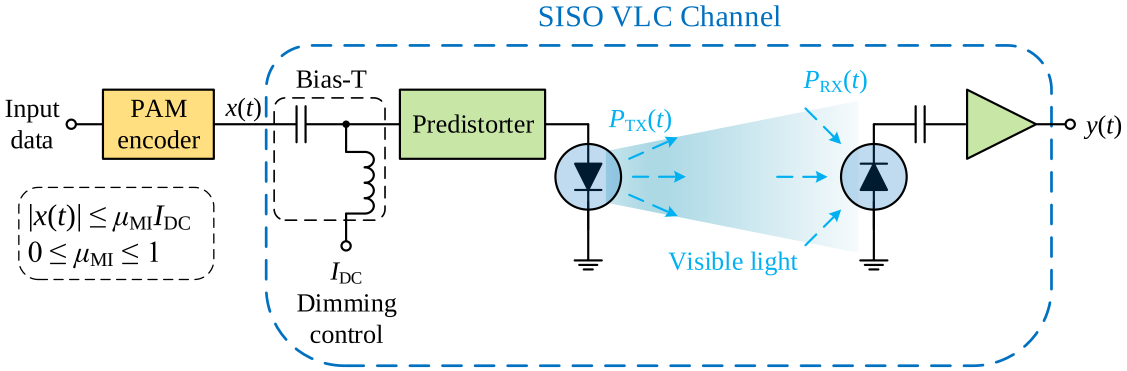

Figure 1.2: Simplified block diagram of a SISO PAM VLC system.

means in the context of physical-layer security.

1.3.1 The VLC Channel Model and Modulation Scheme

Typical VLC systems utilize illumination LEDs for data transmission. Such LEDs are incoherent light sources[3], and thus IM is the only feasible transmission scheme. As a consequence, direct detection (DD) at the receiver using simple PDs is sufficient for demodulation [5, 6, 58, 59].

In this thesis, we adopt the DC-biased pulse-amplitude modulation (PAM) scheme[4] illustrated in Figure 1.2. The transmit element is an illumination LED driven by a fixed bias current $I_{\mathrm{DC}} \in \mathbb{R}_+$ that sets the *average* radiated optical power $P_{\mathrm{opt}} = \eta I_{\mathrm{DC}}$, where η (mW/mA) is the electro-optical conversion efficiency of the LED. The current-power response of a typical LED is depicted in Figure 1.3.

The PAM scheme is described as follows. Information symbols from a single-stream data source are stochastically encoded[5] into a zero-mean current signal $x(t)$,

[3] Unlike LDs, LEDs emit photons with random phases.

[4] Note that PHY I and PHY II in the IEEE 802.15.7 standard use the OOK and (binary) VPPM schemes [20, Tables I and II]. However, restricting the transmitted signal to such *binary* schemes would not allow much room for optimization and performance enhancement, especially when secrecy constraints are taken into account.

[5] Stochastic encoding adds randomization to confuse the eavesdropper. See, e.g., [26, Chapter 3].

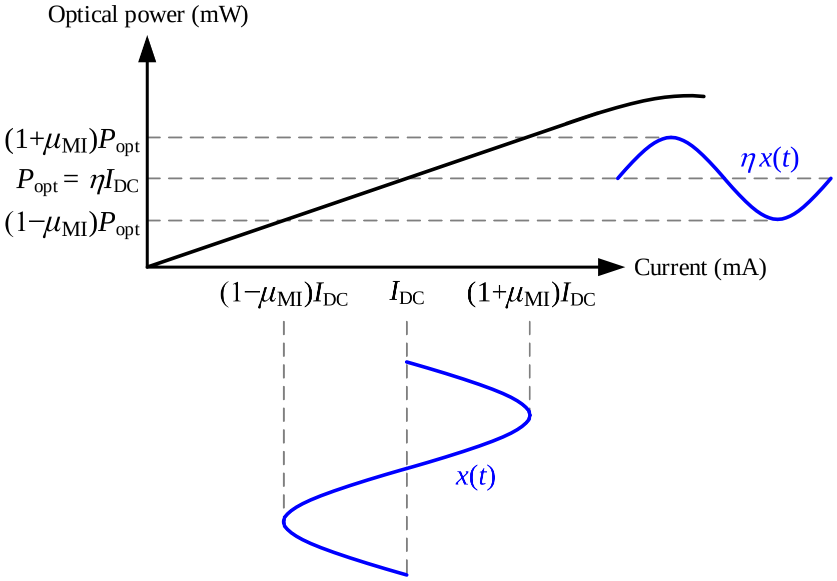

Figure 1.3: Current-power response of a typical LED.

$t = 1, 2, \ldots$. The codewords are chosen such that $\mathbb{E}\{X\} = 0$ and $|x(t)| \leq A \; \forall t$, where X is the random variable counterpart of $x(t)$, $A \triangleq \mu_{\mathrm{MI}} I_{\mathrm{DC}}$, and $\mu_{\mathrm{MI}} \in [0, 1]$ is termed as the *modulation index*. The modulation index, in turn, is chosen such that the LED maintains linear electro-optical conversion over the input current range $[I_{\mathrm{DC}} - A, I_{\mathrm{DC}} + A]$, as illustrated in Figure 1.3. If nonlinearity is severe, digital predistortion of the input current signal $x(t)$ may become necessary to linearize the LED response around the DC bias point [50]. The codewords are then superimposed on the DC bias, via a bias-T circuit, to imperceptibly modulate the output intensity of the LED. Thus, the *instantaneous* emitted optical power $P_{\mathrm{TX}}(t)$ can be expressed as

$$P_{\mathrm{TX}}(t) = \eta(I_{\mathrm{DC}} + x(t)). \tag{1.1}$$

Since $\mathbb{E}\{X\} = 0$, the data[6] signal $x(t)$ does not alter the average radiated optical

[6] With slight abuse of notation, we shall use the term "*data*" to refer to the "codewords corresponding to the secret message". However, it is essential to keep in mind that $x(t)$ is a sequence of secrecy codewords rather than uncoded data symbols.

power and, consequently, it has no effect on the illumination level.

We shall assume narrow-band transmission, that is the bandwidth of the transmitted signal is well below the modulation bandwidth (or the cutoff frequency) of the LED, and is also smaller than the inverse of the maximum excess delay of the VLC channel. In other words, we shall ignore possible low-pass filtering caused by the LED characteristics or multipath propagation. Consider, for example, a VLC system in which phosphorus-coated blue LEDs are utilized for transmission, and blue filtering is applied at the receiver. This setup allows 3-dB modulation bandwidth of about 20 MHz [60, Figure 3], and the excess delay in a medium-sized room is about 10-20 nsec [61]. Thus, a transmitted signal whose bandwidth is limited to 10 MHz, for example, should not suffer noticeable distortion from the frequency response of the LEDs or the channel.

Thus, under the assumption of narrow-band transmission, the frequency response of the VLC channel is almost flat near DC [58], and it is sufficient to characterize the optical channel by its DC gain given by the ratio of transmitted to received optical powers. From (1.1), the instantaneous received optical power is

$$\begin{aligned} P_{\mathrm{RX}}(t) &= h_{\mathrm{opt}} P_{\mathrm{TX}}(t) \\ &= h_{\mathrm{opt}} \eta (I_{\mathrm{DC}} + x(t)), \end{aligned} \tag{1.2}$$

where $h_{\mathrm{opt}} \in \mathbb{R}_+$ is the DC optical channel gain that shall be specified in the next subsection. The received optical power, in turn, is converted by a PD into a proportional photocurrent $R_{\mathrm{PD}} P_{\mathrm{RX}}(t)$, where R_{PD} (μA/mW) is the responsivity of the PD. Then, the DC term $R_{\mathrm{PD}} h_{\mathrm{opt}} \eta I_{\mathrm{DC}}$ is blocked, and the resulting signal is amplified by a transimpedance amplifier with gain T_a (mV/μA) to produce a voltage signal $y(t) \in \mathbb{R}$ that is a scaled, but noisy, version of the input signal $x(t)$. Dominant noise sources in

VLC channels are the thermal noise in the receiver electronic circuits, i.e., the amplifier noise, and the shot noise caused by ambient illumination from sunlight or other light sources. Both noise processes are well modelled as signal-independent additive white Gaussian noise [52, 59]. Thus, the discrete-time VLC channel in Figure 1.2 can be modelled by

$$y(t) = hx(t) + n(t), \qquad t = 1, 2, \ldots, \tag{1.3}$$

where $h \triangleq \eta h_{\text{opt}} R_{\text{PD}} T_a$ is the DC channel gain, and $n(t)$ denotes independent and identically-distributed (i.i.d.) zero-mean Gaussian noise samples with variance σ^2, i.e., $N \sim \mathcal{N}(0, \sigma^2)$, where N is the random variable counterpart of $n(t)$. The channel model in (1.3) is a scalar Gaussian channel whose input $x(t)$ is subject to the amplitude constraint $|x(t)| \leq A \ \forall t = 1, 2, \ldots$.

1.3.2 The Optical Channel Gain

Figure 1.4 illustrates the geometry of an LoS VLC link. The receiver is pointing towards an arbitrary direction specified by the unit vector

$$\mathbf{u} = \left[\sin\theta\cos\phi \quad \sin\theta\sin\phi \quad \cos\theta\right]^{\mathrm{T}},$$

where $\theta \in [0, \pi]$ is the zenith (or polar) angle, and $\phi \in [0, 2\pi]$ is the azimuth angle. We shall refer to $\mathbf{u}$ as the orientation vector.

We assume that the LED has an azimuth-symmetric generalized[7] Lambertian emission pattern. We also assume that the LoS path is dominant over multipath components caused by diffuse reflections from nearby surfaces[8]. Under these assumptions,

[7] In the case of (non-generalized) Lambertian emission pattern, the Lambertian order m is equal to 1, which corresponds to a half-intensity angle $\zeta^{\text{3-dB}} = 60°$.

[8] This assumption will be relaxed in Section 3.4.4 wherein non-line-of-sight (NLoS) components are taken into account.

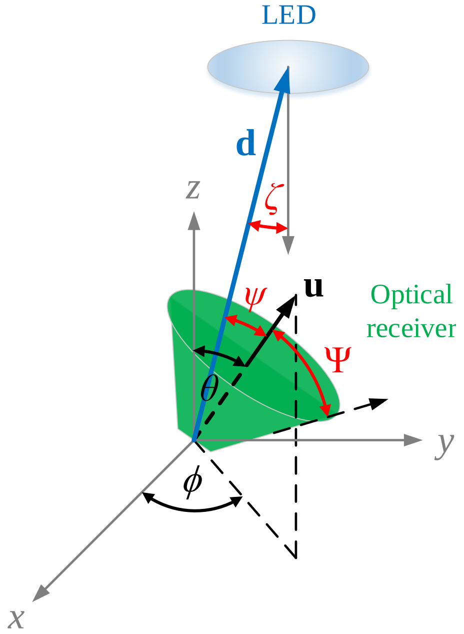

Figure 1.4: Geometry of an LoS VLC link with arbitrary receiver orientation.

the DC optical channel gain h_{opt} can be accurately approximated by [58, Eq. (10)]

$$h_{\text{opt}} = \frac{(m+1)A_{\text{PD}}}{2\pi\|\mathbf{d}\|_2^2}\,(\cos\zeta)^m\,T_s\,g_c\cos\psi\,I_\Psi(\psi) \tag{1.4a}$$

$$= \frac{(m+1)A_{\text{PD}}}{2\pi\|\mathbf{d}\|_2^{m+3}}\,d_z^m\,T_s\,g_c\,\mathbf{d}^{\text{T}}\mathbf{u}\,I_\Psi\left(\cos^{-1}\frac{\mathbf{d}^{\text{T}}\mathbf{u}}{\|\mathbf{d}\|_2}\right), \tag{1.4b}$$

where m is the Lambertian order, A_{PD} is the area of the PD, $\mathbf{d} = [d_x\ d_y\ d_z]^{\text{T}}$ is the displacement vector between the PD and the LED, ζ is the angle of irradiance from the LED (measured w.r.t. the LED axis), T_s is the gain of the optical filter, g_c is the gain of the optical concentrator within its field-of-view (FoV), ψ is the angle of incidence from the LED (measured w.r.t. the receiver axis), and $I_\Psi(\cdot)$ is an indicator function defined as

$$I_\Psi(\psi) \triangleq \begin{cases} 1 & |\psi| \le \Psi \\ 0 & |\psi| > \Psi \end{cases},$$

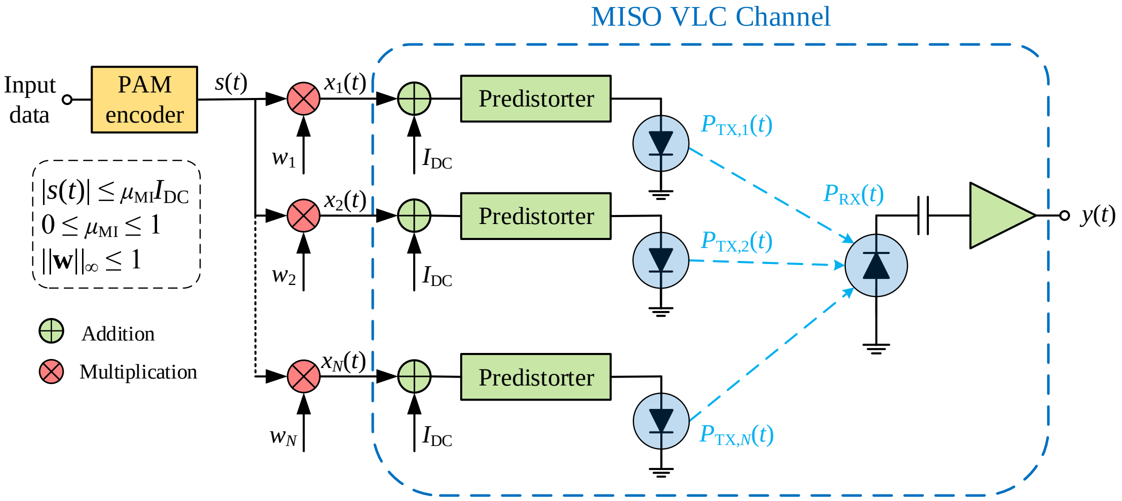

Figure 1.5: Beamforming in conjunction with PAM for the MISO VLC Channel.

where Ψ is the semi-angle FoV of the concentrator. Assuming an idealized non-imaging concentrator, the gain g_c can be approximated by [58, Eq. (8)]

$$g_c = \frac{n_r^2}{\sin^2 \Psi}, \tag{1.5}$$

where n_r is the refractive index of the concentrator material. Furthermore, the Lambertian order m is determined by

$$m = \frac{-1}{\log_2(\cos \zeta^{\text{3-dB}})}, \tag{1.6}$$

where $\zeta^{\text{3-dB}}$ is the half-intensity angle of the LED.

1.3.3 Beamforming for the MISO VLC Channel

When the transmitter has $N > 1$ LEDs that are sufficiently separated and can be modulated independently of each other using separate drivers, we end up with a MISO channel having N transmit elements. Figure 1.5 illustrates a MISO VLC system utilizing transmit beamforming along with PAM. Similar to the SISO case, information symbols are stochastically encoded into codewords S such that $\mathbb{E}\{S\} = 0$ and $|s(t)| \leq A \ \forall t$. Then, the codewords are multiplied by a fixed vector $\mathbf{w} \in \mathbb{R}^N$, $\|\mathbf{w}\|_\infty \leq 1$, termed as the *beamformer*, resulting in the modulation current vector

$$\mathbf{x}(t) = \mathbf{w}s(t). \tag{1.7}$$

Thus, after adding the DC bias to each LED, the vector of instantaneous optical powers transmitted from the LEDs can be expressed as

$$\begin{aligned} \mathbf{P}_{\mathrm{TX}}(t) &= \eta(I_{\mathrm{DC}}\mathbf{1}_N + \mathbf{x}(t)) \\ &= \eta(I_{\mathrm{DC}}\mathbf{1}_N + \mathbf{w}s(t)). \end{aligned} \tag{1.8}$$

With multiple-LED transmission, the total received optical power, $P_{\mathrm{RX}}(t)$, is the sum of optical powers collected from individual LEDs, i.e., $P_{\mathrm{RX}}(t)$ is given by

$$\begin{aligned} P_{\mathrm{RX}}(t) &= \mathbf{h}_{\mathrm{opt}}^{\mathrm{T}}\mathbf{P}_{\mathrm{TX}}(t) \\ &= \eta\mathbf{h}_{\mathrm{opt}}^{\mathrm{T}}(I_{\mathrm{DC}}\mathbf{1}_N + \mathbf{w}s(t)), \end{aligned} \tag{1.9}$$

where $\mathbf{h}_{\mathrm{opt}} \in \mathbb{R}_+^N$ is the DC optical channel gain vector. Then, after removing the DC component from the output of the PD, the received signal $y(t)$ can be expressed as

$$y(t) = \mathbf{h}^{\mathrm{T}}\mathbf{w}s(t) + n(t), \qquad t = 1, 2, \ldots, \tag{1.10}$$

where $\mathbf{h} \triangleq \eta\mathbf{h}_{\mathrm{opt}}R_{\mathrm{PD}}T_a$ is the DC channel gain vector, and $n(t)$ denotes i.i.d. Gaussian noise samples with variance σ^2. Equation (1.10) specifies a Gaussian MISO channel with transmit beamforming, and the transmitted signal vector is subject to

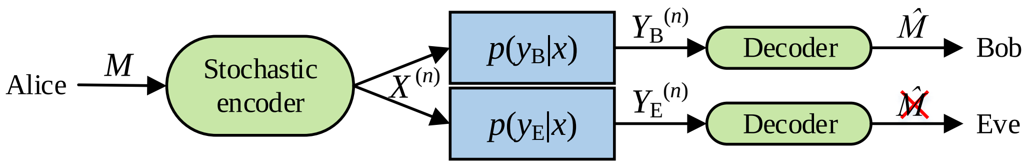

Figure 1.6: A general wiretap channel.

the amplitude constraint

$$\|\mathbf{w}\|_\infty \leq 1, \tag{1.11a}$$

$$|s(t)| \leq A \ \forall t = 1, 2, \ldots. \tag{1.11b}$$

1.3.4 The Wiretap Channel and the Secrecy Capacity

The wiretap channel is a broadcast channel model that was originally proposed by Wyner [34], and later extended by Csiszár and Körner [37], to study the following communication problem: The transmitter (Alice) aims to send a confidential message $M \in \{1, 2, \cdots, 2^{nR_s}\}$ to the receiver (Bob) and keep the message entirely secret from the eavesdropper (Eve) without using secret-key encryption. Figure 1.6 illustrates such a scenario, and the individual channels to Bob and Eve are specified by the marginal transition probability density functions (PDFs) $p(y_\text{B}|x)$ and $p(y_\text{E}|x)$, respectively.

In order to send the secret message M, Alice will stochastically encode M into a codeword $X^{(n)}$ that is transmitted over the broadcast channel in n channel uses. Thus, the information rate is $\frac{1}{n}\log_2(2^{nR_s}) = R_s$ bits/channel use. Both Bob and Eve will attempt decoding their received signals. Let $\hat{M}$ denote the message decoded by Bob, where $\hat{M} \in \{1, 2, \cdots, 2^{nR_s}\}$. Then, decoding error happens when $\hat{M} \neq M$. Let $P_e^{(n)}$

denote the average probability of decoding error at Bob, then the communication rate R_s is said to be *achievable and secure*, i.e., R_s is an *achievable secrecy rate*, if there exists a sequence of $(2^{nR_s}, n)$ codes such that

$$\lim_{n\to\infty} P_e^{(n)} = 0, \tag{1.12a}$$

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(M; Y_{\mathrm{E}}^{(n)}) = 0. \tag{1.12b}$$

The condition in (1.12a) requires the transmission rate R_s to be reliable, i.e., can be reliably decoded by Bob. On the other hand, (1.12b) is the *weak secrecy* constraint which requires the rate of information leaked to Eve to vanish [26, Section 3.3].

The *secrecy capacity* is defined as the maximum achievable secrecy rate. By definition, any achievable secrecy rate is a lower bound on the secrecy capacity[9]. Csiszár and Körner [37] have shown that the secrecy capacity of the (nondegraded) wiretap channel illustrated in Figure 1.6 is [26, Corollary 3.4]

$$C_s = \max_{p(u,x)} \left(\mathbb{I}(U; Y_{\mathrm{B}}) - \mathbb{I}(U; Y_{\mathrm{E}})\right), \tag{1.13}$$

where U is an auxiliary random variable that satisfies the Markov chain $U \to X \to (Y_{\mathrm{B}}, Y_{\mathrm{E}})$. Except for a few specials cases, the optimization problem in (1.13) is typically difficult to solve, and usually it is unclear how to choose the auxiliary variable U in an optimal way. For the special case of the degraded wiretap channel, i.e., when $X \to Y_{\mathrm{B}} \to Y_{\mathrm{E}}$ forms a Markov chain, it can be shown that the choice $U = X$ is optimal (see [26, Corollary 3.5]), and thus (1.13) simplifies to

$$C_s = \max_{p(x)} \left(\mathbb{I}(X; Y_{\mathrm{B}}) - \mathbb{I}(X; Y_{\mathrm{E}})\right). \tag{1.14}$$

[9] Therefore, we use the terms *"achievable secrecy rate"* and *"lower bound on the secrecy capacity"* interchangeably.

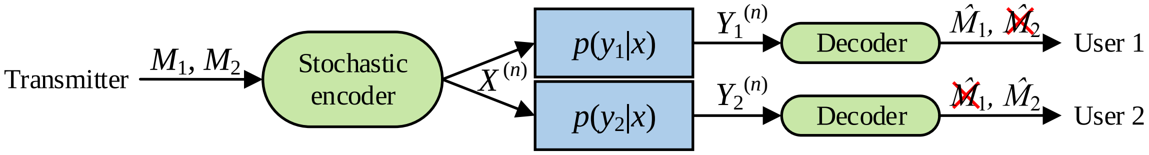

Figure 1.7: A general two-user broadcast channel with confidential messages (BC-CM).

1.3.5 The Two-User Broadcast Channel with Confidential Messages (BC-CM)

The wiretap channel model was extended by Liu *et al.* [47] to the two-user BC-CM illustrated in Figure 1.7. In such a model, the transmitter has two independent confidential messages: $M_1 \in \{1, 2, \cdots, 2^{nR_1}\}$ is intended for User 1 and should be kept secret from User 2, and $M_2 \in \{1, 2, \cdots, 2^{nR_2}\}$ is intended for User 2 and should be kept secret from User 1.

The transmitter encodes the pair (M_1, M_2) into a codeword $X^{(n)}$ that is transmitted in n channel uses. Similar to the wiretap channel, let $P_{e,1}^{(n)}$ denote the average probability of decoding error at User 1, i.e., the average probability that $\hat{M}_1 \neq M_1$, where $\hat{M}_1$ is the decoded message, and $P_{e,2}^{(n)}$ denote the average probability of decoding error at User 2. Then, the rate pair (R_1, R_2) is said to be achievable and secure if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that

$$\lim_{n\to\infty} P_{e,1}^{(n)} = 0, \qquad \lim_{n\to\infty} P_{e,2}^{(n)} = 0, \tag{1.15a}$$

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(M_1; Y_2^{(n)}) = 0, \qquad \lim_{n\to\infty} \frac{1}{n}\mathbb{I}(M_2; Y_1^{(n)}) = 0, \tag{1.15b}$$

where (1.15a) specifies the reliability requirements for both users, and (1.15b) is the mutual confidentiality constraint using the weak secrecy measure.

Compared to the wiretap channel, evaluating the secrecy performance of the two-user BC-CM is obviously more challenging as it requires the characterization of a *secrecy capacity region* rather than the secrecy capacity (which is just a scalar). Let U_1 and U_2 be auxiliary random variables such that $(U_1, U_2) \to X \to (Y_1, Y_2)$ forms a Markov chain. Then, it was shown in [47, Theorem 4] that the secrecy rate pair (R_1, R_2) satisfying

$$0 \leq R_1 \leq \mathbb{I}(U_1; Y_1) - \mathbb{I}(U_1; Y_2|U_2) - \mathbb{I}(U_1; U_2), \tag{1.16a}$$

$$0 \leq R_2 \leq \mathbb{I}(U_2; Y_2) - \mathbb{I}(U_2; Y_1|U_1) - \mathbb{I}(U_1; U_2) \tag{1.16b}$$

is achievable for the general two-user BC-CM illustrated in Figure 1.7.

1.3.6 What Does "Secure Transmission" Mean?

The term "secure transmission scheme" can be ambiguous to a reader not familiar with the terminology used in the physical-layer security literature. Thus, it may be useful to clarify what "secure transmission" literally means.

In the context of physical-layer security, transmission schemes, such as the beamformers proposed in Chapter 3 and precoders proposed in Chapter 4, are said to be "secure" when they *lead to positive secrecy rates.* Thus, a typical problem of interest is to find transmission schemes that maximize the achievable secrecy rate. Note, however, that having a positive secrecy rate is a *necessary but not sufficient condition* to achieve secure transmission. In other words, applying a transmission scheme that leads to a positive secrecy rate does not immediately render the communication link secure. Instead, it makes secure transmission possible provided that an appropriate secrecy codebook is constructed and used to encode the transmitted messages. The secrecy codebook, which is revealed to all parties, should ensure reliable reception by the receiver (like regular channel codes), and also have sufficient randomization to allow stochastic encoding and confuse the eavesdropper. In other words, the codebook should satisfy the reliability and secrecy constraints in (1.12) for the wiretap channel, or the corresponding constraints in (1.15) for the two-user BC-CM. The design of secrecy codebooks is an involved subject that is beyond the scope of this thesis. The interested reader, however, can refer to [26, Chapter 6] or [31, Section VII] and the references therein.

1.4 Contributions of the Thesis

We claim that this thesis is the first to consider enhancing the secrecy of VLC systems within the framework of physical-layer security. By taking amplitude constraints into account, we encounter a novel category of design problems in which closed-form solutions usually cease to be possible. Furthermore, by taking channel uncertainty into account, we help make physical-layer security schemes more applicable to real-world scenarios in which the assumption of perfect channel information is almost always impractical. Our contributions in the entire thesis are summarized as follows.

1. **Achievable Secrecy Rates subject to Amplitude Constraints:** With the lack of analytic expressions for the secrecy capacity of amplitude-constrained Gaussian wiretap channels, we resort to closed-form bounds. In Chapter 2, we begin with deriving lower and upper bounds on the secrecy capacity of the scalar channel under the amplitude constraint. We derive the lower bounds using the uniform input distribution in conjunction with the entropy power inequality. For the upper bound, we devise an approach to obtain upper bounds on the secrecy capacity of degraded wiretap channels, and apply the devised

approach to the scalar Gaussian wiretap channel. We then exploit the lower bound along with transmit beamforming in order to obtain an achievable secrecy rate for the MISO wiretap channel. This achievable rate will serve as the design equation, i.e., the objective function, in all the optimization problems encountered in Chapter 3 wherein the design of the beamformer is studied in detail. We also consider in Chapter 2 the scenario in which the scalar channel between the transmitter and intended receiver is aided by a friendly jammer capable of sending jamming signals using multiple transmit elements. We derive a closed-form secrecy rate expression when both the data and jamming signals are subject to amplitude constraints. Our contributions in Chapter 2 were published in [62, 14, 63, 64].

2. **Optimal and Robust Beamforming for the MISO VLC Wiretap Channel:** In Chapter 3, we focus on the MISO VLC wiretap channel. In particular, we study the design of transmit beamformers that maximize the achievable secrecy rate, subject to amplitude constraints. Such constraints render the design problem nonconvex and difficult to solve. We show, however, that this nonconvex problem can be transformed into a solvable quasiconvex line search problem. Our approach to solve the optimization problem is generic in the sense that it can handle general l_p-norm constraints on the beamforming vector, i.e., for any $p \geq 1$. We also consider the more realistic case of imperfect channel information regarding the receiver's and eavesdropper's links. We tackle the worst-case secrecy rate maximization problem, again subject to amplitude constraints. In our treatment, uncertainty in the receiver's channel is due to limited feedback, and is modelled by spherical uncertainty sets. On the other hand, there is no feedback from the eavesdropper, and the transmitter shall uti-

lize the LoS channel gain equation to map the eavesdropper's nominal location and orientation into an estimate of the channel gain. Thus, we derive channel uncertainty sets based on inaccurate information regarding the eavesdropper's location and orientation, as well as the emission pattern of the LEDs. We also consider channel mismatches caused by the uncertain NLoS components. The work in Chapter 3 was published in [65, 66].

3. **Linear Precoding for the Two-User MISO BC-CM:** In Chapter 4, we turn our focus to the more general two-user MISO BC-CM communication model. We study the design of linear precoders for secure transmission on such a channel subject to total and per-antenna[10] average power constraints, and also subject to amplitude constraints. In both cases, we tackle the design problem by formulating a weighted secrecy sum rate maximization problem. The formulated problem involves a fractional objective function, making it nonconvex and difficult to solve. Nevertheless, we show that this nonconvex problem can be transformed into an equivalent, but more tractable, problem. We propose a subgradient-based search algorithm to obtain a solution, and characterize the condition under which the obtained solution is guaranteed to be globally optimal. Furthermore, we show that our problem formulation and solution approach can be easily extended to handle the robust version of the design problem with uncertain channel information regarding both receivers. Our work in Chapter 4 was submitted for possible publication [67].

[10] In Chapter 4, we generalize the channel model by considering different types of constraints on the channel input. Accordingly, in that chapter, we use the general term "*antenna*" to denote general transmit and receive elements. In a VLC system, for example, the transmit antenna would be an LED and the receive antenna would be a PD.

1.5 Organization of the Thesis

The structure of the thesis reflects the list of contributions in the previous section, and is as follows.

In Chapter 2, we derive closed-form secrecy rate expressions for the Gaussian wiretap channel subject to amplitude constraints. Three cases are considered, namely, the scalar wiretap channel, the MISO wiretap channel, and the scalar channel aided by a friendly jammer having multiple transmit elements. We provide numerical examples from typical VLC scenarios in order to get insight into the secrecy performance of VLC wiretap channels.

In Chapter 3, we consider the design of beamformers for the MISO VLC wiretap channel. The design equation is the secrecy rate expression derived in Chapter 2, and the design parameter is the beamformer subject to amplitude constraints. Under the premise of perfect channel information, we show that the nonconvex secrecy rate maximization problem can be optimally solved using a simple line search algorithm. We then extend our approach to the design of robust beamformers that maximize the worst-case secrecy rate with imperfect channel information. In order to obtain reasonable uncertainty models for the eavesdropper's channel, we derive uncertainty sets based on the uncertain parameters in the VLC channel gain equation. We use numerical examples to compare the performance of the optimal and robust beamformers with conventional beamforming schemes, and also to illustrate the secrecy performance in typical VLC scenarios.

In Chapter 4, we consider linear precoding for the two-user MISO BC-CM subject to total and per-antenna average power constraints, and also subject to amplitude constraints. We begin with deriving closed-form secrecy rate pair expressions. Then, we provide a unified framework to tackle the design problem via weighted secrecy sum rate maximization. We also extend our approach to take channel uncertainty into account. We use numerical examples to validate the solution method and compare the performance of the proposed linear precoder with conventional precoding schemes.

Finally, in Chapter 5, we summarize our contributions and findings in the thesis, and outline some topics for future research.

Appendices A, B, and C contain proofs and derivations relevant to Chapters 2, 3, and 4, respectively.

Chapter 2

Achievable Secrecy Rates for VLC Wiretap Channels

2.1 Introduction

Intensity modulation (IM) is the only feasible transmission scheme for VLC systems that utilize LEDs. Due to linearity limitations of typical LEDs, the input current signal, i.e., the intensity-modulating signal, must satisfy certain amplitude constraints in order to maintain linear electro-optical conversion and avoid nonlinear or clipping distortion (see Figure 1.3). Therefore, IM channels are typically modelled with amplitude constraints on the channel input, rather than the conventional average power constraint [52, 53]. Consequently, a proper characterization of the secrecy performance of VLC links should involve the secrecy capacity of amplitude-constrained Gaussian wiretap channels. In [56], it was shown that the secrecy capacity of the scalar wiretap channel under the amplitude constraint is achieved by a discrete input distribution having a finite number of mass points. For sufficiently-small amplitude constraints, the symmetric binary input distribution has been shown to be optimal [56, Section IV]. For the general case, however, it is difficult to explicitly solve for the maximizing distribution, and thus the secrecy capacity can be only found via numerical methods. Since closed-form expressions are typically crucial for system design purposes, one might resort to lower bounds on the secrecy capacity.

Accordingly, in this chapter, we derive closed-form secrecy rate expressions for the wiretap channel subject to amplitude constraints. Three scenarios are considered, namely, the scalar wiretap channel, the MISO wiretap channel, and the scalar wiretap channel aided by a friendly jammer. In all scenarios, the data and jamming signals (when applicable) are subject to amplitude constraints. For the scalar channel, we use the uniform input distribution in conjunction with the entropy power inequality to obtain lower bounds on the secrecy capacity. We also devise a technique to derive an upper bound. Next, we leverage beamforming to obtain a lower bound on the secrecy capacity of the MISO channel. We characterize the secrecy performance when simple zero-forcing (ZF) beamforming is applied. Finally, we consider the scalar channel when it is aided by a friendly jammer having multiple transmit elements, but does not know the message that is being transmitted. We derive a closed-form secrecy rate expression after restricting the jamming signal such that it causes no interference to the intended receiver.

The remainder of this chapter is divided into three main sections, corresponding to the three scenarios we consider, besides the conclusions section. The scalar and MISO wiretap channels are considered in Sections 2.2 and 2.3, respectively, whereas the scalar channel aided by a friendly jammer is considered in Section 2.4. In each section, we begin with describing the problem scenario and system model, then we derive closed-form secrecy rate expressions followed by a numerical example. We conclude the chapter in Section 2.5.

2.2 The Scalar VLC Wiretap Channel

In this section, we consider the scalar VLC wiretap channel, i.e., the amplitude-constrained scalar Gaussian wiretap channel. Because of the amplitude constraint,

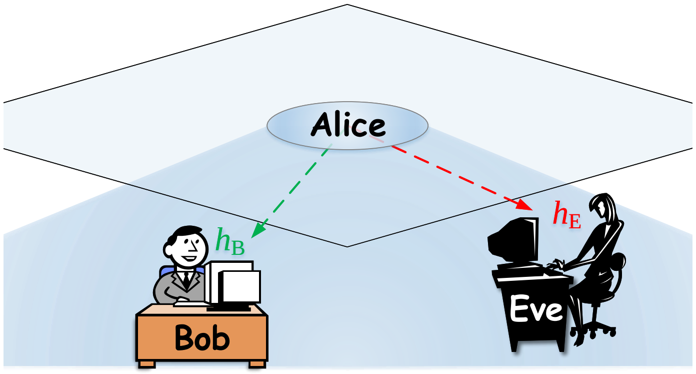

Figure 2.1: Problem scenario for the SISO case.

there is no analytic expression for the secrecy capacity, and thus we derive closed-form lower and upper bounds.

2.2.1 System Model

We consider the simple VLC scenario illustrated in Figure 2.1. The service area, or simply the room, is illuminated by a single light fixture that is also utilized by Alice for data transmission. The fixture may have one LED, or multiple LEDs modulated by the same current signal, e.g., all the LEDs are connected in series. The intended receiver (Bob) and the eavesdropper (Eve) have a single photodiode (PD), each.

Utilizing the Gaussian channel model in (1.3), the signals received by Bob and Eve, respectively, are given by

$$y_{\mathrm{B}}(t) = h_{\mathrm{B}} x(t) + n_{\mathrm{B}}(t), \tag{2.1a}$$

$$y_{\mathrm{E}}(t) = h_{\mathrm{E}} x(t) + n_{\mathrm{E}}(t), \tag{2.1b}$$

where $x(t) \in [-A, A]$ is the transmitted signal, $h_{\mathrm{B}} \in \mathbb{R}_+$ and $h_{\mathrm{E}} \in \mathbb{R}_+$ are Bob's and Eve's channel gains, respectively, and $n_{\mathrm{B}}(t)$ and $n_{\mathrm{E}}(t)$ are i.i.d. Gaussian noise samples with variances σ_{B}^2 and σ_{E}^2, respectively. For simplicity, and without loss of generality, we assume that $\sigma_{\mathrm{B}}^2 = \sigma_{\mathrm{E}}^2 = \sigma^2$. Such an assumption can be simply fulfilled by properly scaling y_{B} or y_{E}.

2.2.2 Achievable Secrecy Rates

Assuming $h_{\mathrm{B}} > h_{\mathrm{E}}$, the secrecy capacity of the scalar wiretap channel in (2.1) is [36]

$$C_s^{\mathrm{SISO}} = \max_{p(x)} \ \left(\mathbb{I}(X;Y_{\mathrm{B}}) - \mathbb{I}(X;Y_{\mathrm{E}})\right) \tag{2.2a}$$

$$\text{s.t.} \ \ |X| \leq A, \tag{2.2b}$$

where maximization is performed over all the input distributions $p(x)$ that satisfy the amplitude constraint $|X| \leq A$. Now, because of the amplitude constraint, obtaining a closed-form solution for (2.2) is a formidable task, if not unfeasible [56]. Nevertheless, it was shown that the maximization problem in (2.2) is convex [56, Eq. (9)], and the optimal distribution $p^{\star}(x)$ that maximizes the difference $\mathbb{I}(X;Y_{\mathrm{B}}) - \mathbb{I}(X;Y_{\mathrm{E}})$ is discrete with a finite number of mass points. Thus, the problem in (2.2) can be efficiently solved via numerical methods. Nevertheless, closed-form expressions are typically of great interest for system design purposes. Therefore, we provide closed-form lower bounds on the secrecy capacity of the wiretap channel in (2.1), as follows.

Proposition 2.1. *(Lower Bound on the Secrecy Capacity)*

The secrecy capacity of the scalar Gaussian wiretap channel in (2.1) *subject to the amplitude constraint* $|x(t)| \leq A \ \forall t$ *is lower-bounded as*

$$\begin{aligned} C_s^{\mathrm{SISO}} \geq \frac{1}{2}\ln\left(1 + \frac{2A^2h_{\mathrm{B}}^2}{\pi e \sigma^2}\right) - \left(1 - 2\mathcal{Q}\left(\frac{\delta + Ah_{\mathrm{E}}}{\sigma}\right)\right)\ln\frac{2(Ah_{\mathrm{E}} + \delta)}{\sqrt{2\pi\sigma^2}\left(1 - 2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)} \\ - \mathcal{Q}\left(\frac{\delta}{\sigma}\right) - \frac{\delta}{\sqrt{2\pi\sigma^2}}e^{-\frac{\delta^2}{2\sigma^2}} + \frac{1}{2}, \end{aligned} \tag{2.3}$$

where $\delta > 0$ is a free parameter, and $\mathcal{Q}(\cdot)$ is the $\mathcal{Q}$-function.

Proof: The secrecy capacity in (2.2) can be lower-bounded by the difference between the capacities of Alice-Bob and Alice-Eve channels, as follows.

$$\begin{aligned} C_s^{\mathrm{SISO}} &= \max_{p(x)} \left(\mathbb{I}(X;Y_{\mathrm{B}}) - \mathbb{I}(X;Y_{\mathrm{E}})\right) \\ &\geq \max_{p(x)} \mathbb{I}(X;Y_{\mathrm{B}}) - \max_{p(x)} \mathbb{I}(X;Y_{\mathrm{E}}) \\ &= C_{\mathrm{B}} - C_{\mathrm{E}}, \end{aligned} \tag{2.4}$$

where the inequality follows from the fact that

$$\max_{\mathbf{u}} \left(f_1(\mathbf{u}) - f_2(\mathbf{u})\right) \geq \max_{\mathbf{u}} f_1(\mathbf{u}) - \max_{\mathbf{u}} f_2(\mathbf{u})$$

for arbitrary functions f_1 and f_2. Then, C_{B} and C_{E}, respectively, can be lower- and upper-bounded as [53, Theorem 5]

$$C_{\mathrm{B}} \geq \frac{1}{2} \ln\left(1 + \frac{2A^2 h_{\mathrm{B}}^2}{\pi e \sigma^2}\right), \tag{2.5a}$$

$$C_{\mathrm{E}} \leq \left(1 - 2\mathcal{Q}\left(\frac{\delta + Ah_{\mathrm{E}}}{\sigma}\right)\right) \ln \frac{2(Ah_{\mathrm{E}} + \delta)}{\sqrt{2\pi\sigma^2}\left(1 - 2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)} + \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi\sigma^2}} e^{-\frac{\delta^2}{2\sigma^2}} - \frac{1}{2}, \tag{2.5b}$$

where $\delta > 0$ is a free parameter. Replacing C_{B} and C_{E} in (2.4) with the lower and upper bounds in (2.5a) and (2.5b), respectively, yields the lower bound in (2.3). ■

Proposition 2.2. *(Lower Bound on the Secrecy Capacity)*

The secrecy capacity of the scalar Gaussian wiretap channel in (2.1) *subject to the* *amplitude constraint* $|x(t)| \leq A \ \forall t$ *is lower-bounded as*

$$C_s^{\text{SISO}} \geq \frac{1}{2} \ln \frac{6A^2 h_{\text{B}}^2 + 3\pi e \sigma^2}{\pi e A^2 h_{\text{E}}^2 + 3\pi e \sigma^2}. \tag{2.6}$$

Proof: The secrecy capacity in (2.2) can be lower-bounded using the entropy-maximizing uniform input distribution as follows.

$$\begin{aligned}
C_s^{\text{SISO}} &\overset{\text{(a)}}{\geq} \mathbb{I}(X; Y_{\text{B}}) - \mathbb{I}(X; Y_{\text{E}}) \\
&= \mathbb{h}(Y_{\text{B}}) - \mathbb{h}(Y_{\text{B}}|X) - \mathbb{h}(Y_{\text{E}}) + \mathbb{h}(Y_{\text{E}}|X) \\
&= \mathbb{h}(Y_{\text{B}}) - \mathbb{h}(Y_{\text{E}}) \\
&= \mathbb{h}(h_{\text{B}}X + N_{\text{B}}) - \mathbb{h}(Y_{\text{E}}) \\
&\overset{\text{(b)}}{\geq} \frac{1}{2} \ln \left(e^{2\mathbb{h}(h_{\text{B}}X)} + e^{2\mathbb{h}(N_{\text{B}})} \right) - \frac{1}{2} \ln \left(2\pi e \, \mathbb{var}\{Y_{\text{E}}\} \right) \\
&\overset{\text{(c)}}{=} \frac{1}{2} \ln(4A^2 h_{\text{B}}^2 + 2\pi e \sigma^2) - \frac{1}{2} \ln \left(2\pi e \left(\frac{4A^2 h_{\text{E}}^2}{12} + \sigma^2 \right) \right) \\
&= \frac{1}{2} \ln \frac{6A^2 h_{\text{B}}^2 + 3\pi e \sigma^2}{\pi e A^2 h_{\text{E}}^2 + 3\pi e \sigma^2},
\end{aligned} \tag{2.7}$$

where (a) follows from dropping the maximization over $p(x)$, (b) from lower-bounding $\mathbb{h}(h_{\text{B}}X + N_{\text{B}})$ using the entropy power inequality [35, Theorem 17.7.3] and upper-bounding $\mathbb{h}(Y_{\text{E}})$ by the differential entropy of a Gaussian random variable having variance $\mathbb{var}\{Y_{\text{E}}\}$, and (c) from choosing $X \sim \mathcal{U}[-A, A]$, i.e., $p(x)$ is the uniform distribution over the interval $[-A, A]$, and substituting with

$$\begin{aligned}
&\mathbb{h}(h_{\text{B}}X) = \ln(2Ah_{\text{B}}), \\
&\mathbb{var}\{Y_{\text{E}}\} = \mathbb{var}\{h_{\text{E}}X\} + \mathbb{var}\{N_{\text{E}}\} = \frac{(2Ah_{\text{E}})^2}{12} + \sigma^2.
\end{aligned}$$

■

Note that the uniform distribution $p(x) = \mathcal{U}[-A, A]$ is the maximum-entropy distribution over the input range $[-A, A]$ subject to the constraint $\mathbb{E}\{X\} = 0$. Such a constraint is necessary to ensure that the average radiated optical power, and consequently the illumination level, is not altered by $\mathbb{E}\{X\}$ (recall the modulation scheme described in Section 1.3.1).

2.2.3 Upper Bound on the Secrecy Capacity

In [68, 69, 53], the authors used the *dual* channel capacity expression in [70, Theorem 8.4] to obtain upper bounds on the capacity of the Gaussian channel under amplitude constraints. Here, we follow a similar approach in order to derive an upper bound on the secrecy capacity of the scalar Gaussian wiretap channel. First, we note that, for the case $h_{\mathrm{B}} > h_{\mathrm{E}}$, the wiretap channel in (2.1) has the same secrecy performance as that of the *physically degraded* wiretap channel characterized by [26, Section 5.1]

$$y_{\mathrm{B}}(t) = x(t) + n_{\mathrm{B}}(t), \tag{2.8a}$$

$$y_{\mathrm{E}}(t) = y_{\mathrm{B}}(t) + n_{\mathrm{E}}(t), \tag{2.8b}$$

with $N_{\mathrm{B}} \sim \mathcal{N}(0, \frac{\sigma^2}{h_{\mathrm{B}}^2})$ and $N_{\mathrm{E}} \sim \mathcal{N}(0, \frac{\sigma^2}{h_{\mathrm{E}}^2} - \frac{\sigma^2}{h_{\mathrm{B}}^2})$. Note from (2.8) that $X \to Y_{\mathrm{B}} \to Y_{\mathrm{E}}$ forms a Markov chain. Next, we introduce the following theorem.

Theorem 2.1. *(Upper Bound on Conditional Mutual Information)*
Let X, Y_{B}, and Y_{E} be three random variables with a joint distribution $p(x, y_{\mathrm{B}}, y_{\mathrm{E}})$ that factors as $p(x)p(y_{\mathrm{B}}|x)p(y_{\mathrm{E}}|y_{\mathrm{B}})$, i.e. $X \to Y_{\mathrm{B}} \to Y_{\mathrm{E}}$ forms a Markov chain. Then, the

conditional mutual information $\mathbb{I}(X;Y_{\rm B}|Y_{\rm E})$ *is upper-bounded as*

$$\mathbb{I}(X;Y_{\rm B}|Y_{\rm E}) \leq \mathbb{E}_{p(x)}\{\mathbb{D}(p(y_{\rm B}|X,y_{\rm E})\|q(y_{\rm B}|y_{\rm E}))\}\,, \tag{2.9}$$

where $\mathbb{D}(\cdot\|\cdot)$ *denotes the relative entropy,* $p(y_{\rm B}|x,y_{\rm E}) = \dfrac{p(y_B|x)p(y_{\rm E}|y_{\rm B})}{p(y_{\rm E}|x)}$*, and* $q(y_{\rm B}|y_{\rm E})$ *is an arbitrary conditional distribution.*

Proof: We begin with [71, Eq. (2.4.20)]

$$\mathbb{I}(X;Y_{\rm B}|Y_{\rm E}) = \iiint p(x,y_{\rm B},y_{\rm E}) \ln \frac{p(y_{\rm B}|x,y_{\rm E})}{p(y_{\rm B}|y_{\rm E})}\, dx\, dy_{\rm B}\, dy_{\rm E}. \tag{2.10}$$

We also have [35, Eq. (2.65)]

$$\begin{aligned} \mathbb{D}(p(y_{\rm B}|y_{\rm E})\|q(y_{\rm B}|y_{\rm E})) &= \iint p(y_{\rm B},y_{\rm E}) \ln \frac{p(y_{\rm B}|y_{\rm E})}{q(y_{\rm B}|y_{\rm E})}\, dy_{\rm B}\, dy_{\rm E} \\ &= \iiint p(x,y_{\rm B},y_{\rm E}) \ln \frac{p(y_{\rm B}|y_{\rm E})}{q(y_{\rm B}|y_{\rm E})}\, dx\, dy_{\rm B}\, dy_{\rm E}. \end{aligned} \tag{2.11}$$

Adding (2.10) to (2.11) yields

$$\begin{aligned} &\mathbb{I}(X;Y_{\rm B}|Y_{\rm E}) + \mathbb{D}(p(y_{\rm B}|y_{\rm E})\|q(y_{\rm B}|y_{\rm E})) \\ &\quad = \iiint p(x,y_{\rm B},y_{\rm E}) \ln \frac{p(y_{\rm B}|x,y_{\rm E})}{q(y_{\rm B}|y_{\rm E})}\, dx\, dy_{\rm B}\, dy_{\rm E} \\ &\quad = \mathbb{E}_{p(x)}\left\{\iint p(y_{\rm B},y_{\rm E}|X) \ln \frac{p(y_{\rm B}|X,y_{\rm E})}{q(y_{\rm B}|y_{\rm E})}\, dy_{\rm B}\, dy_{\rm E}\right\} \\ &\quad = \mathbb{E}_{p(x)}\{\mathbb{D}(p(y_{\rm B}|X,y_{\rm E})\|q(y_{\rm B}|y_{\rm E}))\}\,. \end{aligned} \tag{2.12}$$

Then, the inequality in (2.9) follows since the relative entropy $\mathbb{D}(p(y_{\rm B}|y_{\rm E})\|q(y_{\rm B}|y_{\rm E}))$ is always nonnegative [35, Theorem 2.6.3]. ■

Note from (2.12) that equality holds in (2.9) when $\mathbb{D}(p(y_{\rm B}|y_{\rm E})\|q(y_{\rm B}|y_{\rm E})) = 0$, i.e., when $p(y_{\mathrm{B}}|y_{\mathrm{E}}) = q(y_{\mathrm{B}}|y_{\mathrm{E}})\ \forall y_{\mathrm{B}}, y_{\mathrm{E}}$. Note also that the inequality in (2.9) holds for any input distribution $p(x)$. Now, consider the specific distribution $p^{\star}(x)$ that achieves the secrecy capacity, i.e.,

$$p^{\star}(x) \triangleq \underset{p(x)}{\operatorname{argmax}}\ \mathbb{I}(X; Y_{\mathrm{B}}|Y_{\mathrm{E}}), \tag{2.13}$$

where maximization is over all the distributions that satisfy the constraints on the channel input X. Using $p^{\star}(x)$ in (2.9) results in the following upper bound on the secrecy capacity.

Corollary 2.1. *(Upper Bound on The Secrecy Capacity of the Degraded Wiretap Channel)*
An upper bound on the secrecy capacity of the degraded wiretap channel $X \to Y_{\mathrm{B}} \to Y_{\mathrm{E}}$ is given by

$$C_s \leq \mathbb{E}_{p^{\star}(x)}\{\mathbb{D}(p(y_{\mathrm{B}}|X, y_{\mathrm{E}}) \| q(y_{\mathrm{B}}|y_{\mathrm{E}}))\}, \tag{2.14}$$

where $p^{\star}(x)$ is as defined in (2.13)*, and $q(y_{\mathrm{B}}|y_{\mathrm{E}})$ is an arbitrary conditional distribution.*

Now, we are ready to derive an upper bound on the secrecy capacity of the wiretap channel in (2.1), as follows.

Proposition 2.3. *(Upper Bound on the Secrecy Capacity of the Scalar Gaussian Wiretap Channel)*
The secrecy capacity of the scalar Gaussian wiretap channel in (2.1) *subject to the amplitude constraint $|x(t)| \leq A\ \forall t$ is upper-bounded as*

$$C_s^{\mathrm{SISO}} \leq \frac{1}{2} \ln \frac{A^2 h_{\mathrm{B}}^2 + \sigma^2}{A^2 h_{\mathrm{E}}^2 + \sigma^2}. \tag{2.15}$$

Proof: Substituting for $\mathbb{D}(\cdot\|\cdot)$ in (2.14) yields

$$\begin{aligned} C_s &\leq \mathbb{E}_{p^\star(x)}\left\{\iint p(y_\mathrm{B}, y_\mathrm{E}|X) \ln \frac{p(y_\mathrm{B}|X, y_\mathrm{E})}{q(y_\mathrm{B}|y_\mathrm{E})}\, dy_\mathrm{B}\, dy_\mathrm{E}\right\} \\ &= \underbrace{\mathbb{E}_{p^\star(x)}\left\{\iint p(y_\mathrm{B}, y_\mathrm{E}|X) \ln p(y_\mathrm{B}|X, y_\mathrm{E})\, dy_\mathrm{B}\, dy_\mathrm{E}\right\}}_{I_1} \\ &\quad \underbrace{-\mathbb{E}_{p^\star(x)}\left\{\iint p(y_\mathrm{B}, y_\mathrm{E}|X) \ln q(y_\mathrm{B}|y_\mathrm{E})\, dy_\mathrm{B}\, dy_\mathrm{E}\right\}}_{I_2}. \end{aligned} \tag{2.16}$$

Now, we have to calculate the terms I_1 and I_2.

The first term I_1 can be written as

$$\begin{aligned} I_1 &= \mathbb{E}_{p^\star(x)}\left\{\iint p(y_\mathrm{B}, y_\mathrm{E}|X) \ln p(y_\mathrm{B}|X, y_\mathrm{E})\, dy_\mathrm{B}\, dy_\mathrm{E}\right\} \\ &= \mathbb{E}_{p^\star(x)}\{-\mathbb{h}(Y_\mathrm{B}|X = x, Y_\mathrm{E})\} \\ &= -\mathbb{h}(Y_\mathrm{B}|X, Y_\mathrm{E}). \end{aligned} \tag{2.17}$$

Recall that for a Markov chain $X \to Y_\mathrm{B} \to Y_\mathrm{E}$, we have

$$\mathbb{h}(X, Y_\mathrm{B}, Y_\mathrm{E}) = \mathbb{h}(X) + \mathbb{h}(Y_\mathrm{B}|X) + \mathbb{h}(Y_\mathrm{E}|Y_\mathrm{B}). \tag{2.18}$$

In addition, for any random variables X, Y_B, and Y_E, we have

$$\mathbb{h}(X, Y_\mathrm{B}, Y_\mathrm{E}) = \mathbb{h}(X) + \mathbb{h}(Y_\mathrm{E}|X) + \mathbb{h}(Y_\mathrm{B}|X, Y_\mathrm{E}). \tag{2.19}$$

From (2.17)–(2.19), we can see that

$$I_1 = -\left(\mathbb{h}(Y_\mathrm{B}|X) + \mathbb{h}(Y_\mathrm{E}|Y_\mathrm{B}) - \mathbb{h}(Y_\mathrm{E}|X)\right). \tag{2.20}$$

For notational convenience, define γ_{B}^2 and γ_{B}^2, respectively, as

$$\gamma_{\mathrm{B}}^2 \triangleq \frac{\sigma^2}{h_{\mathrm{B}}^2}, \qquad \gamma_{\mathrm{E}}^2 \triangleq \frac{\sigma^2}{h_{\mathrm{E}}^2}.$$

Thus, we have

$$I_1 = -\frac{1}{2}\ln\left(2\pi e \frac{\gamma_{\mathrm{B}}^2(\gamma_{\mathrm{E}}^2 - \gamma_{\mathrm{B}}^2)}{\gamma_{\mathrm{E}}^2}\right). \tag{2.21}$$

In order to calculate I_2 in (2.16), we choose the conditional distribution $q(y_{\mathrm{B}}|y_{\mathrm{E}})$ as

$$q(y_{\mathrm{B}}|y_{\mathrm{E}}) = \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(y_{\mathrm{B}} - \mu y_{\mathrm{E}})^2}{2s^2}}, \tag{2.22}$$

where μ and s^2 are constants to be determined in (2.25). Again, for a Markov chain $X \to Y_{\mathrm{B}} \to Y_{\mathrm{E}}$, we have

$$\begin{aligned} p(y_{\mathrm{B}}, y_{\mathrm{E}}|x) &= p(y_{\mathrm{B}}|x)\, p(y_{\mathrm{E}}|y_{\mathrm{B}}) \\ &= \frac{1}{\sqrt{2\pi\gamma_{\mathrm{B}}^2}} e^{-\frac{(y_{\mathrm{B}} - x)^2}{2\gamma_{\mathrm{B}}^2}} \frac{1}{\sqrt{2\pi(\gamma_{\mathrm{E}}^2 - \gamma_{\mathrm{B}}^2)}} e^{-\frac{(y_{\mathrm{E}} - y_{\mathrm{B}})^2}{2(\gamma_{\mathrm{E}}^2 - \gamma_{\mathrm{B}}^2)}}. \end{aligned} \tag{2.23}$$

Using (2.22) and (2.23), we get

$$\begin{aligned} I_2 &= -\mathbb{E}_{p^\star(x)}\left\{\iint p(y_{\mathrm{B}}, y_{\mathrm{E}}|X) \ln q(y_{\mathrm{B}}|y_{\mathrm{E}})\, dy_{\mathrm{B}}\, dy_{\mathrm{E}}\right\} \\ &= -\mathbb{E}_{p^\star(x)}\left\{\frac{1}{\sqrt{2\pi\gamma_{\mathrm{B}}^2}} \int_{-\infty}^{\infty} e^{-\frac{(y_{\mathrm{B}} - X)^2}{2\gamma_{\mathrm{B}}^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\left(\gamma_{\mathrm{E}}^2 - \gamma_{\mathrm{B}}^2\right)}} e^{-\frac{(y_{\mathrm{E}} - y_{\mathrm{B}})^2}{2(\gamma_{\mathrm{E}}^2 - \gamma_{\mathrm{B}}^2)}} \times \right. \\ &\qquad \left. \left(-\frac{1}{2}\ln(2\pi s^2) - \frac{(y_{\mathrm{B}} - \mu y_{\mathrm{E}})^2}{2s^2}\right) dy_{\mathrm{B}}\, dy_{\mathrm{E}}\right\} \\ &= \frac{1}{2}\ln(2\pi s^2) + \mathbb{E}_{p^\star(x)}\left\{\frac{1}{\sqrt{2\pi\gamma_{\mathrm{B}}^2}} \int_{-\infty}^{\infty} e^{-\frac{(y_{\mathrm{B}} - X)^2}{2\gamma_{\mathrm{B}}^2}} \frac{1}{2s^2}\left(\mu^2\left(\gamma_{\mathrm{E}}^2 - \gamma_{\mathrm{B}}^2\right) + (\mu - 1)^2 y_{\mathrm{B}}^2\right) dy_{\mathrm{B}}\right\} \\ &= \frac{1}{2}\ln(2\pi s^2) + \mathbb{E}_{p^\star(x)}\left\{\frac{1}{2s^2}\left(\mu^2\left(\gamma_{\mathrm{E}}^2 - \gamma_{\mathrm{B}}^2\right) + (\mu - 1)^2\left(X^2 + \gamma_{\mathrm{B}}^2\right)\right)\right\} \\ &\le \frac{1}{2}\ln(2\pi s^2) + \frac{1}{2s^2}\left(\mu^2\left(\gamma_{\mathrm{E}}^2 - \gamma_{\mathrm{B}}^2\right) + (\mu - 1)^2\left(A^2 + \gamma_{\mathrm{B}}^2\right)\right), \end{aligned} \tag{2.24}$$

where the inequality follows from $\mathbb{E}_{p^\star(x)}\{X^2\} \leq A^2$. In order to minimize the expression in (2.24), we differentiate w.r.t. μ and s^2. After setting the resulting partial derivatives to zero, we obtain the minimizers

$$\mu = \frac{A^2 + \gamma_{\mathrm{B}}^2}{A^2 + \gamma_{\mathrm{E}}^2} \qquad \text{and} \qquad s^2 = \frac{(A^2 + \gamma_{\mathrm{B}}^2)(\gamma_{\mathrm{E}}^2 - \gamma_{\mathrm{B}}^2)}{A^2 + \gamma_{\mathrm{E}}^2}. \tag{2.25}$$

Substituting from (2.25) back into (2.24) and adding the result to (2.21), we get

$$\begin{aligned} C_s &\leq \frac{1}{2} \ln \frac{(A^2 + \gamma_{\mathrm{B}}^2)\gamma_{\mathrm{E}}^2}{(A^2 + \gamma_{\mathrm{E}}^2)\gamma_{\mathrm{B}}^2} \\ &= \frac{1}{2} \ln \frac{h_{\mathrm{B}}^2 A^2 + \sigma^2}{h_{\mathrm{E}}^2 A^2 + \sigma^2}, \end{aligned} \tag{2.26}$$

which is the upper bound in (2.15). ■

It is worth mentioning that the upper bound in (2.26) can be simply obtained by relaxing the amplitude constraint $|X| \leq A$ into the average power constraint $\mathbb{E}\{X^2\} \leq A^2$ and noting that (2.26) is the secrecy capacity of the Gaussian channel under the average power constraint. Nevertheless, the framework we proposed via Theorem 2.1 and Corollary 2.1 can be used to derive upper bounds on the secrecy capacity of degraded wiretap channels with arbitrary conditional distributions $p(y_{\mathrm{B}}|x)$ and $p(y_{\mathrm{E}}|y_{\mathrm{B}})$, i.e., the main and degraded channels need not be Gaussian.

2.2.4 Numerical Example

Figure 2.2 depicts the bounds in (2.3), (2.6), and (2.15). Three groups of these bounds are shown using $20 \log_{10}(h_{\mathrm{B}}/h_{\mathrm{E}}) = 10, 20,$ and 30 dB. The lower bound in (2.3) is calculated using $\delta = \sigma \ln(1 + 2Ah_{\mathrm{E}}/\sigma)$ as proposed in [53]. As can be seen, both (2.3) and (2.6) along with (2.15) tightly bound the secrecy capacity at asymptotically low and high $\mathrm{SNR}_{\mathrm{B}}$, where $\mathrm{SNR}_{\mathrm{B}} \triangleq h_{\mathrm{B}}^2 A^2/\sigma^2$. Note that the lower

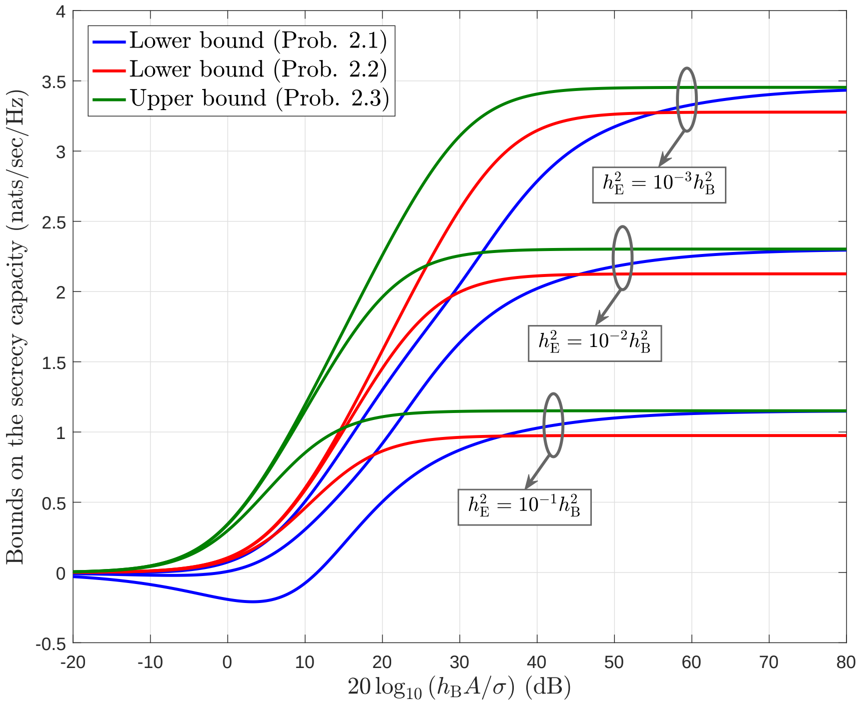

Figure 2.2: Lower and upper bounds on the secrecy capacity of the scalar Gaussian wiretap channel.

bound in (2.6) incurs a fixed gap $\ln\sqrt{\pi e/6} = 0.1765$ nats/sec/Hz at asymptotically high $\mathrm{SNR}_{\mathrm{B}}$. Nevertheless, since typical VLC links operate at SNR values below 40 dB (see, e.g., Figure 2.5), the lower bound in (2.6) is more appropriate for VLC scenarios. Furthermore, (2.6) is more analytically-tractable, and therefore it will be used to obtain secrecy rate expressions for the MISO wiretap channel.

2.3 The MISO VLC Wiretap Channel

In this section, we utilize one of the lower bounds we derived in the previous section along with beamforming to obtain a secrecy rate expression for the MISO wiretap channel subject to amplitude constraints on the channel input vector.

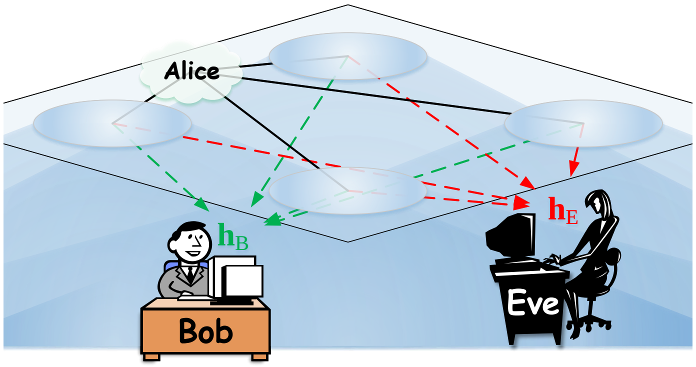

Figure 2.3: Problem scenario for the MISO case.

2.3.1 System Model

We consider the MISO scenario illustrated in Figure 2.3. The room is illuminated by N identical light fixtures utilized for data transmission. Using the vectorized version of the channel model in (1.3), the signals observed by Bob and Eve, respectively, are

$$y_{\mathrm{B}}(t) = \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{x}(t) + n_{\mathrm{B}}(t), \tag{2.27a}$$

$$y_{\mathrm{E}}(t) = \mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{x}(t) + n_{\mathrm{E}}(t), \tag{2.27b}$$

where $\mathbf{x}(t) \in \mathbb{R}^N$ is the transmitted signal vector subject to the amplitude constraint $\|\mathbf{x}(t)\|_\infty \leq A \; \forall t$, $\mathbf{h}_{\mathrm{B}} \in \mathbb{R}_+^N$ and $\mathbf{h}_{\mathrm{E}} \in \mathbb{R}_+^N$ are fixed channel gain vectors, and $n_{\mathrm{B}}(t)$ and $n_{\mathrm{E}}(t)$ are i.i.d. Gaussian noise samples with variance σ^2. Unlike the scalar wiretap channel in (2.1), the MISO wiretap channel in (2.27) is nondegraded, provided that $\mathbf{h}_{\mathrm{B}}$ and $\mathbf{h}_{\mathrm{E}}$ are linearly independent.

2.3.2 Achievable Secrecy Rates

A single-letter characterization of the secrecy capacity of the nondegraded wiretap channel in (2.27) was derived by Csiszár and Körner as [37]

$$C_s^{\text{MISO}} = \max_{p(\mathbf{u},\mathbf{x})} \left(\mathbb{I}(\boldsymbol{U};Y_{\text{B}}) - \mathbb{I}(\boldsymbol{U};Y_{\text{E}})\right), \tag{2.28}$$

where $\boldsymbol{U}$ is an auxiliary random vector that satisfies the Markov chain

$$\boldsymbol{U} \to \boldsymbol{X} \to (Y_{\text{B}}, Y_{\text{E}}).$$

Unlike the scalar case, the optimization problem in (2.28) is nonconvex, in general. Furthermore, it is unclear how to choose $\boldsymbol{U}$. For the Gaussian MISO channel under total average power constraint, it was shown in [39] that the secrecy capacity is achieved via beamforming, i.e., the choice $\boldsymbol{U} = \boldsymbol{X} = \mathbf{w}S$ is optimum, where $\mathbf{w}$ is the beamformer, i.e., a fixed vector, and S is a Gaussian random variable. Accordingly, we propose the use of beamforming to obtain a lower bound on the secrecy capacity of the MISO wiretap channel in (2.27) under the amplitude constraint, as follows.

Proposition 2.4. *(Lower Bound on the Secrecy Capacity)*
The secrecy capacity of the MISO wiretap channel in (2.27) *subject to the amplitude constraint* $\|\mathbf{x}(t)\|_\infty \leq A \; \forall t$ *is lower-bounded as*

$$C_s^{\text{MISO}} \geq \frac{1}{2} \ln \frac{6A^2(\mathbf{h}_{\text{B}}^{\text{T}}\mathbf{w})^2 + 3\pi e\sigma^2}{\pi e A^2(\mathbf{h}_{\text{E}}^{\text{T}}\mathbf{w})^2 + 3\pi e\sigma^2}, \tag{2.29}$$

where $\mathbf{w} \in \mathbb{R}^N$ *is any beamforming vector that satisfies the constraint* $\|\mathbf{w}\|_\infty \leq 1$*.*

Proof: The proof follows directly from combining beamforming and the lower bound in (2.6), as follows.

$$\begin{aligned} C_s^{\text{MISO}} &= \max_{p(\mathbf{u},\mathbf{x})} \left(\mathbb{I}(\boldsymbol{U};Y_{\text{B}}) - \mathbb{I}(\boldsymbol{U};Y_{\text{E}})\right) \\ &\overset{\text{(a)}}{\geq} \mathbb{I}(\boldsymbol{X};Y_{\text{B}}) - \mathbb{I}(\boldsymbol{X};Y_{\text{E}}) \\ &\overset{\text{(b)}}{\geq} \mathbb{I}(\mathbf{w}S;Y_{\text{B}}) - \mathbb{I}(\mathbf{w}S;Y_{\text{E}}) \\ &\overset{\text{(c)}}{\geq} \frac{1}{2}\ln\frac{6A^2(\mathbf{h}_{\text{B}}^{\text{T}}\mathbf{w})^2 + 3\pi e\sigma^2}{\pi e A^2(\mathbf{h}_{\text{E}}^{\text{T}}\mathbf{w})^2 + 3\pi e\sigma^2}, \end{aligned} \tag{2.30}$$

where (a) follows from dropping the maximization and setting $\boldsymbol{U} = \boldsymbol{X}$, (b) from choosing $\boldsymbol{X} = \mathbf{w}S$ such that $\|\mathbf{w}\|_\infty \leq 1$ and $|S| \leq A$, i.e., restricting the transmission scheme to beamforming, and (c) from choosing $p(s) = \mathcal{U}[-A, A]$ and utilizing the lower bound in (2.6). ■

Although suboptimal, beamforming is preferable as it is a linear scheme with low implementation complexity. Furthermore, beamforming reduces the vector channel into a scalar version which enables the use of scalar channel codes. Note that the secrecy rate expression in (2.30) provides a design equation for the beamformer $\mathbf{w}$.

2.3.2.1 Optimal Beamforming

The optimal beamformer $\mathbf{w}^\star$ that maximizes the secrecy rate in (2.30) is

$$\mathbf{w}^\star = \underset{\|\mathbf{w}\|_\infty \leq 1}{\text{argmax}} \ \frac{1}{2}\ln\frac{6A^2(\mathbf{h}_{\text{B}}^{\text{T}}\mathbf{w})^2 + 3\pi e\sigma^2}{\pi e A^2(\mathbf{h}_{\text{E}}^{\text{T}}\mathbf{w})^2 + 3\pi e\sigma^2}. \tag{2.31}$$

The optimization problem in (2.31) is nonconvex and difficult to solve, mainly because of the amplitude constraint $\|\mathbf{w}\|_\infty \leq 1$. In fact, we shall devote a considerable portion of Chapter 3 to solving (2.31).

2.3.2.2 Zero-Forcing Beamforming

The secrecy rate expression in (2.30) can be simplified by restricting the beamformer $\mathbf{w}$ to Eve's null space. Then, the best ZF beamformer $\mathbf{w}_{\mathrm{ZF}}$ is obtained by

$$\mathbf{w}_{\mathrm{ZF}} = \underset{\|\mathbf{w}\|_\infty \leq 1}{\operatorname{argmax}} \ \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} \tag{2.32a}$$

$$\text{s.t.} \ \mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w} = 0, \tag{2.32b}$$

which yields the ZF secrecy rate

$$R_s^{\mathrm{ZF}} = \frac{1}{2}\ln\left(1 + \frac{2A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}_{\mathrm{ZF}})^2}{\pi e \sigma^2}\right). \tag{2.33}$$

Unlike (2.31), the problem in (2.32) is a linear program, and thus can be solved with lower computational complexity. Furthermore, the ZF beamformer $\mathbf{w}_{\mathrm{ZF}}$ makes it unnecessary to use secrecy codebooks, i.e., secure transmission can be achieved with regular channel codes.

2.3.3 Numerical Example from a VLC Scenario

Here we provide some numerical results to get insight into the secrecy performance of the ZF beamformer in a typical indoor VLC scenario. The problem geometry is illustrated in Figure 2.4, and the simulation parameters are provided in Table 2.1. There exist 16 down-facing light fixtures attached to the ceiling. Each fixture encloses 4 LEDs, and each LED radiates 1 W optical power. The half-intensity angle is 60°, and the modulation index is set to 10%. Bob and Eve are located at height 0.85 m above the floor level, e.g., on desks, and their receivers have a 60° FoV (semi-angle). We use a Cartesian coordinate system (x, y) at the receivers height to specify their

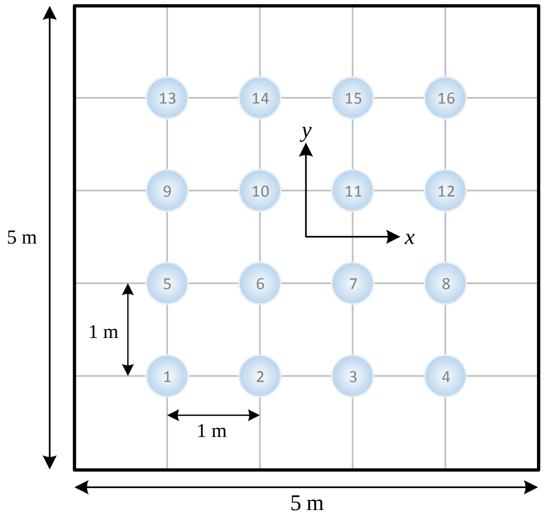

Figure 2.4: Layout of the LEDs for the MISO case.

locations. The origin $(0, 0)$ corresponds to the room center, and all distances are specified in meters. Noise power is calculated using [9, Eq. (6) and Table I] with 70 MHz receiver bandwidth, and the result is averaged over the entire room area. The average electric noise power is -98.82 dBm.

Figure 2.5 shows the spatial distribution of the SNR at the receivers height without beamforming, i.e., $\mathbf{w} = \mathbf{1}_N$. As can be seen, the SNR reaches its maximum value, 39.40 dB, at the room center, and decays to 24.97 dB at the corners.

Figure 2.6 shows the achievable communication rate R_{B}, between Alice and Bob, as a function of Bob's location, without secrecy constraints. This rate is obtained using (2.33) after replacing $\mathbf{w}_{\mathrm{ZF}}$ with $\mathbf{w} = \mathbf{1}_N$.

In Figure 2.7, Bob's location is fixed at $(-0.9, -2.0)$ and the secrecy rate (2.33) is depicted as a function of Eve's location within the entire room area. As expected, the secrecy rate significantly decreases when Eve is close to Bob. Once Eve is relatively faraway, e.g., more than about 2.5 m apart, the secrecy rate is almost independent of Eve's exact location. It is also interesting to characterize the loss in communication

Table 2.1: Simulation parameters for the MISO wiretap channel.

Problem geometry	
Room dimensions ($W \times L \times H$)	$5 \times 5 \times 3$ m^3
Light fixtures height (Alice)	3 m
Receivers height (Bob and Eve)	0.85 m
Number of light fixtures N	16
Transmitter characteristics	
Number of LEDs per fixture	4
Average optical power per LED P_{opt}	1 W
Modulation index μ_{MI}	10%
LEDs half-intensity angle $\zeta^{\text{3-dB}}$	60°
Receiver characteristics	
Receiver FoV Ψ	60°
Refractive index of the concentrator n_r	1.5
PD responsivity R_{PD}	0.54 (A/W)
PD surface area A_{PD}	1 cm^2
Average noise power σ^2	-98.82 dBm

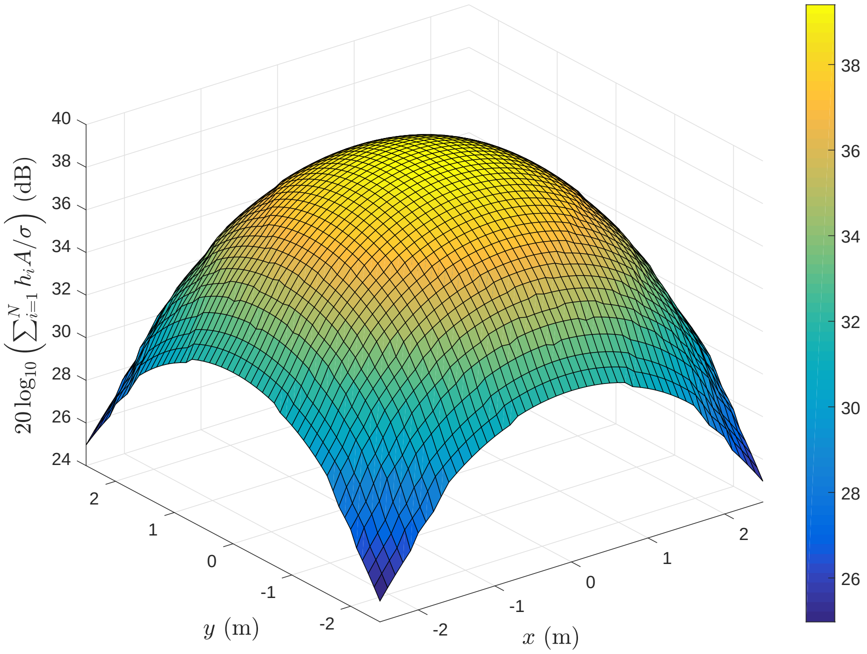

Figure 2.5: Spatial distribution of the SNR at the receivers height (0.85 m above the floor level) without beamforming.

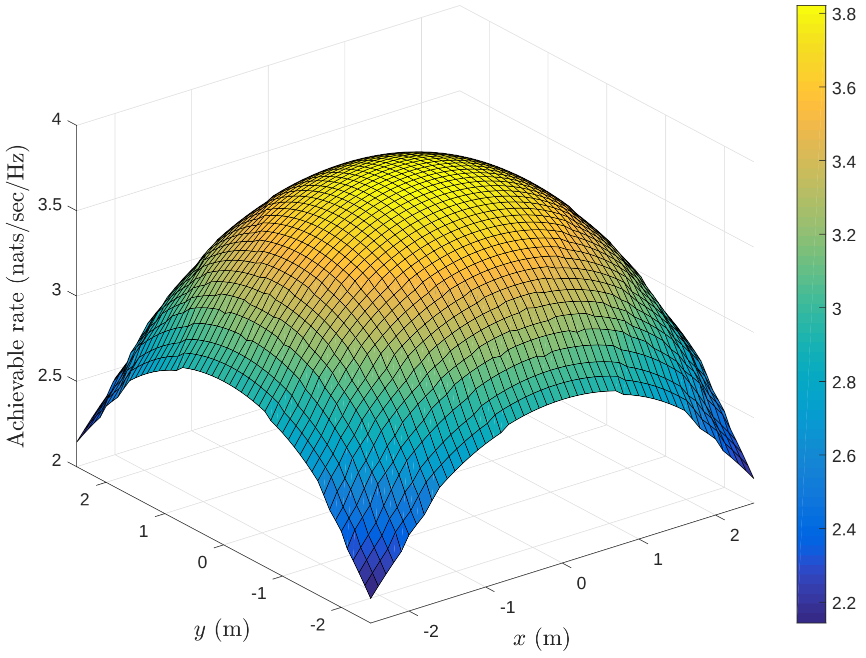

Figure 2.6: Achievable communication rate between Alice and Bob as a function of Bob's location without secrecy constraints.

rate caused by the secrecy constraint, i.e., $R_{\mathrm{B}} - R_s$, by comparing the secrecy rates in Figure 2.7 with $R_{\mathrm{B}}(-0.9, -2.0) = 3.2256$ nats/sec/Hz from Figure 2.6.

Finally, in Figure 2.8, Eve's location is fixed at $(1.6, -0.7)$ and the secrecy rate (2.33) is shown as a function of Bob's location. As can be seen, even when Bob is relatively faraway from Eve, the secrecy rate R_s still depends on Bob's location, i.e., R_s exhibits stronger dependence on $\mathbf{h}_{\mathrm{B}}$ than $\mathbf{h}_{\mathrm{E}}$.

2.4 The Scalar VLC Wiretap Channel Aided by a Friendly Jammer

In this section, we study the secrecy performance of the scalar VLC wiretap channel when it is aided by a friendly jammer having multiple transmit LEDs. A jamming

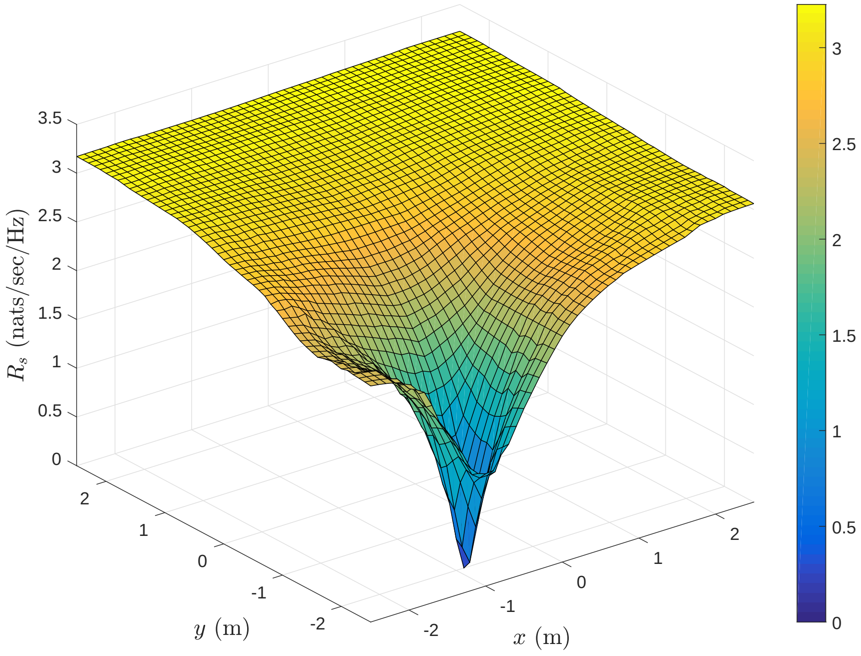

Figure 2.7: Secrecy rate obtained with the ZF beamformer (2.32) as a function of Eve's location when Bob is located at $(-0.9, -2.0)$.

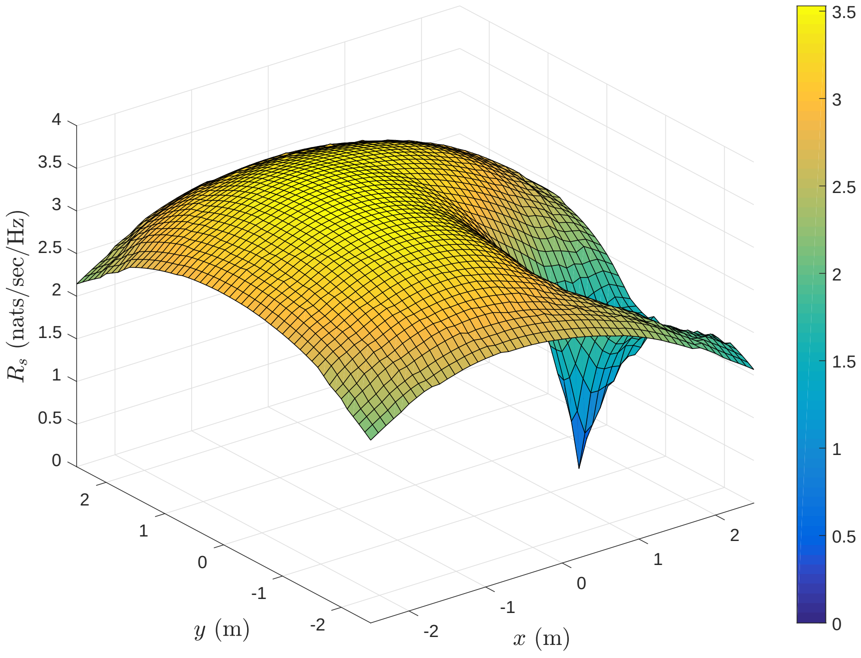

Figure 2.8: Secrecy rate obtained with the ZF beamformer (2.32) as a function of Bob's location when Eve is located at $(1.6, -0.7)$.

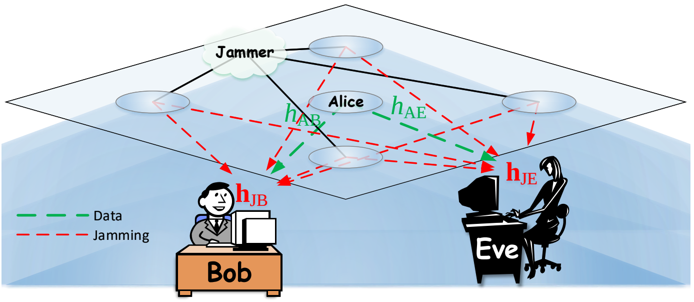

Figure 2.9: Problem scenario for the scalar channel aided by a friendly jammer.

signal is transmitted to degrade Eve's reception while causing no interference to Bob, which leads to an increase in the achievable secrecy rate between Alice and Bob. Both the data and jamming signals are subject to amplitude constraints.

After a formal description of the system model, we derive a closed-form expression for the achievable secrecy rate. Then, we provide a numerical example to illustrate the performance in a typical VLC scenario.

2.4.1 System Model

We consider the VLC scenario illustrated in Figure 2.9. The room is illuminated by $N_{\mathrm{J}} + 1$ identical light fixtures. Each fixture consists of a group of LEDs modulated by the same current signal. Alice, the transmitter, sends her data via a single fixture. On the other hand, the jammer utilizes the remaining N_{J} fixtures, but it does not know the data transmitted by Alice. Bob and Eve have a single PD, each.

Without help from the jammer, securing the connection between Alice and Bob is not possible unless Bob is closer to Alice than Eve. On the other hand, a jammer equipped with multiple transmit elements, and without having access to the transmitted data, can help secure the connection by transmitting a carefully-designed jamming signal that increases the interference seen by Eve, i.e., degrades her signal-to-interference-plus-noise ratio (SINR), while causing no interference to Bob.

Utilizing the channel model in (1.3), the signals received by Bob and Eve, respectively, are

$$y_{\mathrm{B}}(t) = h_{\mathrm{AB}}x(t) + \mathbf{h}_{\mathrm{JB}}^{\mathrm{T}}\mathbf{x}_{\mathrm{J}}(t) + n_{\mathrm{B}}(t), \tag{2.34a}$$

$$y_{\mathrm{E}}(t) = h_{\mathrm{AE}}x(t) + \mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{x}_{\mathrm{J}}(t) + n_{\mathrm{E}}(t), \tag{2.34b}$$

where $h_{\mathrm{AB}} \in \mathbb{R}_+$ and $h_{\mathrm{AE}} \in \mathbb{R}_+$ are the channel gains from Alice to Bob and Eve, respectively, $\mathbf{h}_{\mathrm{JB}} \in \mathbb{R}_+^{N_{\mathrm{J}}}$ and $\mathbf{h}_{\mathrm{JE}} \in \mathbb{R}_+^{N_{\mathrm{J}}}$ are the channel gain vectors from the jammer to Bob and Eve, respectively, $x(t) \in \mathbb{R}$ is the data signal, $\mathbf{x}_{\mathrm{J}}(t) \in \mathbb{R}^{N_{\mathrm{J}}}$ is the jamming signal, and $n_{\mathrm{B}}(t)$ and $n_{\mathrm{E}}(t)$ are i.i.d. Gaussian noise samples with variance σ^2. The data and jamming signals are subject to the amplitude constraints $|x(t)| \leq A \ \forall t$ and $\|\mathbf{x}_{\mathrm{J}}(t)\|_\infty \leq A \ \forall t$, respectively. Furthermore, both $x(t)$ and $\mathbf{x}_{\mathrm{J}}(t)$ are designed such that $\mathbb{E}\{X\} = 0$ and $\mathbb{E}\{\boldsymbol{X}_{\mathrm{J}}\} = \mathbf{0}$. Thus, neither $x(t)$ nor $\mathbf{x}_{\mathrm{J}}(t)$ has an effect on illumination. Finally, we assume that $\mathbf{h}_{\mathrm{JB}}$ and $\mathbf{h}_{\mathrm{JE}}$ are linearly independent, and all the channel gains are accurately known to all the terminals.

In order to derive a secrecy rate expression for the wiretap channel in (2.34), we have to simplify the expressions in (2.34) by imposing the following restrictions. First, the jamming signal $\mathbf{x}_{\mathrm{J}}(t)$ shall cause no interference to Bob, i.e., $\mathbf{h}_{\mathrm{JB}}^{\mathrm{T}}\mathbf{x}_{\mathrm{J}}(t) = 0 \ \forall t$. Such a restriction is not necessarily optimal as it might well be the case that allowing nonzero interference at Bob would permit higher interference at Eve and probably higher achievable secrecy rate. Second, the jammer shall adopt a beamforming strategy, i.e., the jamming signal is constructed as $\mathbf{x}_{\mathrm{J}}(t) = \mathbf{w}_{\mathrm{J}}j(t)$, where $\mathbf{w}_{\mathrm{J}} \in \mathbb{R}^{N_{\mathrm{J}}}$, $\|\mathbf{w}_{\mathrm{J}}\|_\infty \leq 1$, is the jamming beamformer, and $j(t) \in [-A, A]$ is a zero-mean random jamming symbol. Beamforming is preferred as it allows simple implementation, however it might be an inappropriate jamming strategy if there are many eavesdroppers with probably orthogonal or near-orthogonal channels. Finally, to simplify the derivation of a closed-form secrecy rate expression, we assume that both X and J have uniform distributions over the interval $[-A, A]$. After applying such restrictions, the wiretap channel in (2.34) simplifies to

$$y_{\mathrm{B}}(t) = h_{\mathrm{AB}}x(t) + n_{\mathrm{B}}(t), \tag{2.35a}$$

$$y_{\mathrm{E}}(t) = h_{\mathrm{AE}}x(t) + \mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}j(t) + n_{\mathrm{E}}(t). \tag{2.35b}$$

We are now ready to derive an achievable secrecy rate expression for (2.35), which will also be achievable for (2.34).

2.4.2 Achievable Secrecy Rate

Proposition 2.5. *(Achievable Secrecy Rate)*

An achievable secrecy rate, in (nats/sec/Hz), for the wiretap channel in (2.35) *is* $[R_s]^+$*, where* R_s *is given by*

$$\begin{aligned} R_s &= \frac{1}{2}\ln\left(1 + \frac{2A^2h_{\mathrm{AB}}^2}{\pi e\sigma^2}\right) - \begin{cases} \ln\dfrac{h_{\mathrm{AE}}}{|\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}|} + \dfrac{|\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}|}{2h_{\mathrm{AE}}} & |\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}| \le h_{\mathrm{AE}} \\ \dfrac{h_{\mathrm{AE}}}{2|\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}|} & \text{otherwise} \end{cases} \\ &= \frac{1}{2}\ln\left(1 + \frac{2A^2h_{\mathrm{AB}}^2}{\pi e\sigma^2}\right) - \min\left\{\ln\frac{h_{\mathrm{AE}}}{|\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}|} + \frac{|\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}|}{2h_{\mathrm{AE}}}\,,\ \frac{h_{\mathrm{AE}}}{2|\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}|}\right\}, \end{aligned} \tag{2.36}$$

where $\mathbf{w}_{\mathrm{J}} \in \mathbb{R}^N$ *is any jamming vector that satisfies the constraints* $\mathbf{h}_{\mathrm{JB}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}} = 0$ *and* $\|\mathbf{w}_{\mathrm{J}}\|_\infty \le 1$.

Proof: Without loss of generality, we assume in the following that $\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}$ is non-negative. If $\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}} < 0$, then $\mathbf{w}_{\mathrm{J}}$ can be replaced with $-\mathbf{w}_{\mathrm{J}}$ without violating the amplitude constraint or changing the secrecy rate results.

First, we recall our assumption in the previous subsection that $X \sim \mathcal{U}[-A, A]$

and $J \sim \mathcal{U}[-A, A]$. Thus, we have

$$\mathbb{h}(h_{\mathrm{AB}}X) = \ln(2Ah_{\mathrm{AB}}), \tag{2.37a}$$

$$\mathbb{h}(h_{\mathrm{AE}}X) = \ln(2Ah_{\mathrm{AE}}), \tag{2.37b}$$

$$\mathbb{h}(\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}J) = \ln(2A\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}). \tag{2.37c}$$

Next, let the random variable V_{E} be defined as

$$V_{\mathrm{E}} \triangleq h_{\mathrm{AE}}X + \mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}J. \tag{2.38}$$

Then, V_{E} has a trapezoidal distribution (see Appendix A), and its differential entropy is

$$\mathbb{h}(V_{\mathrm{E}}) = \min\left\{\ln(2Ah_{\mathrm{AE}}) + \frac{\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}}{2h_{\mathrm{AE}}}\ ,\ \ln(2A\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}) + \frac{h_{\mathrm{AE}}}{2\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}}\right\}. \tag{2.39}$$

Furthermore, it is clear from (2.35b) and (2.38) that $X \to V_{\mathrm{E}} \to Y_{\mathrm{E}}$ forms a Markov chain. Now, the secrecy capacity of the wiretap channel in (2.35) can be lower-bounded as follows.

$$\begin{aligned}
C_s &\geq \mathbb{I}(X; Y_{\mathrm{B}}) - \mathbb{I}(X; Y_{\mathrm{E}}) \\
&\overset{\text{(a)}}{\geq} \mathbb{I}(X; Y_{\mathrm{B}}) - \mathbb{I}(X; V_{\mathrm{E}}) \\
&= \mathbb{h}(Y_{\mathrm{B}}) - \mathbb{h}(Y_{\mathrm{B}}|X) - \mathbb{h}(V_{\mathrm{E}}) + \mathbb{h}(V_{\mathrm{E}}|X) \\
&\overset{\text{(b)}}{\geq} \frac{1}{2}\ln\left(e^{2\mathbb{h}(h_{\mathrm{AB}}X)} + e^{2\mathbb{h}(N_{\mathrm{B}})}\right) - \mathbb{h}(N_{\mathrm{B}}) - \mathbb{h}(V_{\mathrm{E}}) + \mathbb{h}(\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}J) \\
&\overset{\text{(c)}}{=} \frac{1}{2}\ln(4A^2h_{\mathrm{AB}}^2 + 2\pi e\sigma^2) - \frac{1}{2}\ln(2\pi e\sigma^2) - \mathbb{h}(V_{\mathrm{E}}) + \ln(2A\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}) \\
&\overset{\text{(d)}}{=} \frac{1}{2}\ln\left(1 + \frac{2A^2h_{\mathrm{AB}}^2}{\pi e\sigma^2}\right) - \min\left\{\ln\frac{h_{\mathrm{AE}}}{\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}} + \frac{\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}}{2h_{\mathrm{AE}}}\ ,\ \frac{h_{\mathrm{AE}}}{2\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}}\right\},
\end{aligned} \tag{2.40}$$

where (a) follows from the data-processing inequality [35, Theorem 2.8.1], (b) from lower-bounding $\mathbb{h}(Y_{\mathrm{B}})$ using the entropy power inequality [35, Theorem 17.7.3], (c) by substituting from (2.37a) and (2.37c) for $\mathbb{h}(h_{\mathrm{AB}}X)$ and $\mathbb{h}(\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}J)$, respectively, and (d) by substituting from (2.39) for $\mathbb{h}(V_{\mathrm{E}})$. ■

Figure 2.10 depicts R_s in (2.36) as a function of $\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}$ for different values of h_{AE}. Note that R_s is upper-bounded by $\frac{1}{2}\ln\left(1+\frac{2A^2h_{\mathrm{AB}}^2}{\pi e\sigma^2}\right)$, which is the achievable rate between Alice and Bob, without secrecy constraints, subject to $|x(t)| \leq A \ \forall t$ [53, Theorem 5]. Note also that R_s is a nondecreasing function of $\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}}$ for $\mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}} \geq 0$, and a nonincreasing function of h_{AE}. Thus, under the assumption that $\mathbf{h}_{\mathrm{JB}}$ and $\mathbf{h}_{\mathrm{JE}}$ are perfectly known to the jammer, the optimal jamming beamformer that maximizes R_s while causing no interference to Bob is obtained by

$$\underset{\|\mathbf{w}_{\mathrm{J}}\|_\infty \leq 1}{\text{maximize}} \quad \mathbf{h}_{\mathrm{JE}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}} \tag{2.41a}$$

$$\text{s.t.} \quad \mathbf{h}_{\mathrm{JB}}^{\mathrm{T}}\mathbf{w}_{\mathrm{J}} = 0, \tag{2.41b}$$

which is a simple linear program and can be efficiently solved.

2.4.3 Numerical Example from a VLC Scenario

In this subsection, we provide a numerical example by simulating a typical indoor VLC scenario. The problem geometry is illustrated in Figure 2.11, and the simulation parameters are provided in Table 2.2. The room has a size of $5 \times 5 \times 3$ m^3, and is illuminated by 9 identical light fixtures. Each fixture has 7 LEDs, and each LED radiates 1 W optical power. The fixture at the center is used by Alice for data transmission, while the remaining 8 fixtures are exploited for jamming. The modulation index for all the LEDs is 10%. Bob and Eve are located at height 0.85 m above the

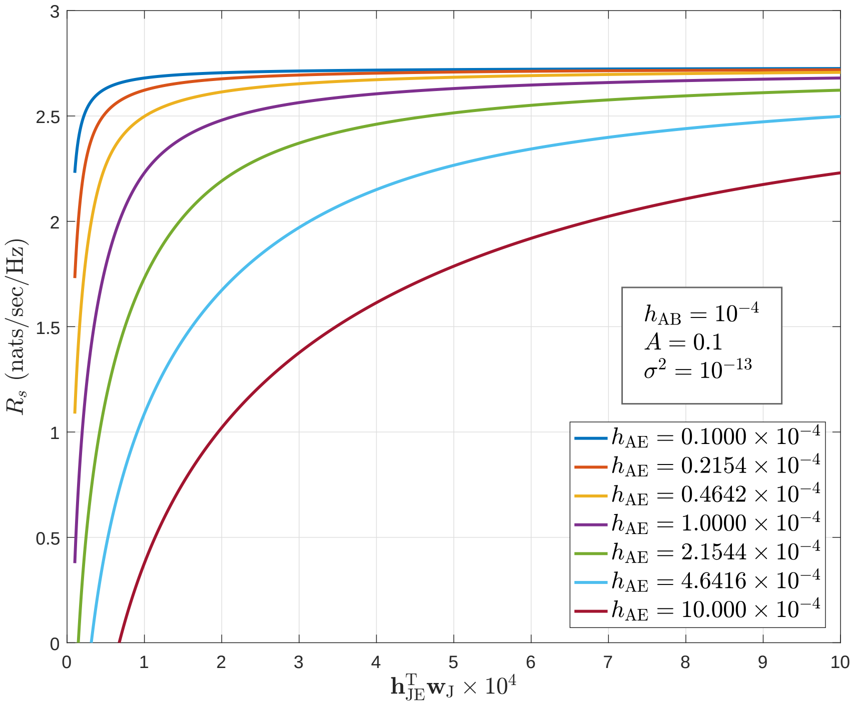

Figure 2.10: Achievable secrecy rates (2.36) for the scalar channel aided by a friendly jammer.

floor level, and their receivers have a 70° FoV and a single PD, each. We use a two-dimensional coordinate system (x, y) to identify the receivers locations. The origin $(0, 0)$ corresponds to the room center at the receivers level. Noise power is calculated using [9, Eq. (6) and Table I] with a receiver bandwidth of 70 MHz, and the result is averaged over the entire room area. The average noise power is -98.39 dBm.

In Figure 2.12, we plot the secrecy rate (2.36) as a function of Eve's location when Bob is located at $(-0.7, -0.9)$, while in Figure 2.13, we fix Eve's location at $(0.3, -1.5)$ and plot (2.36) as a function of Bob's location.

In both figures, the jamming beamformer $\mathbf{w}_{\mathrm{J}}$ is obtained with (2.41). We note that, when Eve is sufficiently close to Bob, jamming is restrained by the null space of Bob, resulting in considerably reduced secrecy rates. On the other hand, when Bob and Eve are faraway, the jammer is able to significantly degrade Eve's reception,

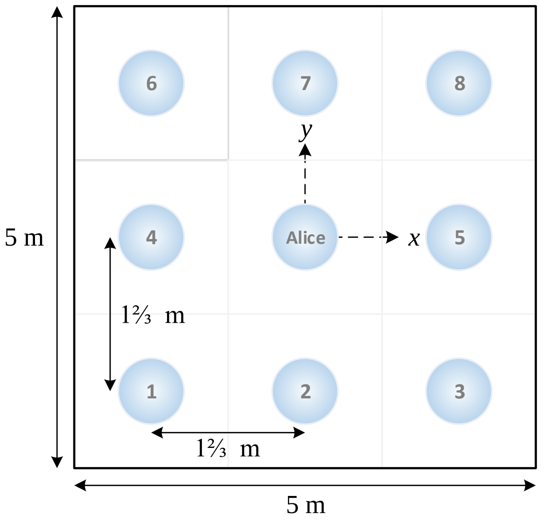

Figure 2.11: Layout of the LEDs for the scalar channel aided by a friendly jammer.

Table 2.2: Simulation parameters for the scalar channel aided by a friendly jammer.

Problem geometry	
Room dimensions ($W \times L \times H$)	$5 \times 5 \times 3$ m^3
Light fixtures height (Alice and the jammer)	3 m
Receivers height (Bob and Eve)	0.85 m
Total number of light fixtures $N_\mathrm{J} + 1$	9
Transmitter characteristics	
Number of LEDs per fixture	7
Average optical power per LED P_opt	1 W
Modulation index μ_MI	10%
LEDs half-intensity angle $\zeta^{\text{3-dB}}$	60°
Receiver characteristics	
Receiver FoV Ψ	70°
Refractive index of the concentrator n_r	1.5
PD responsivity R_PD	0.54 A/W
PD surface area A_PD	1 cm^2
Average noise power σ^2	-98.39 dBm

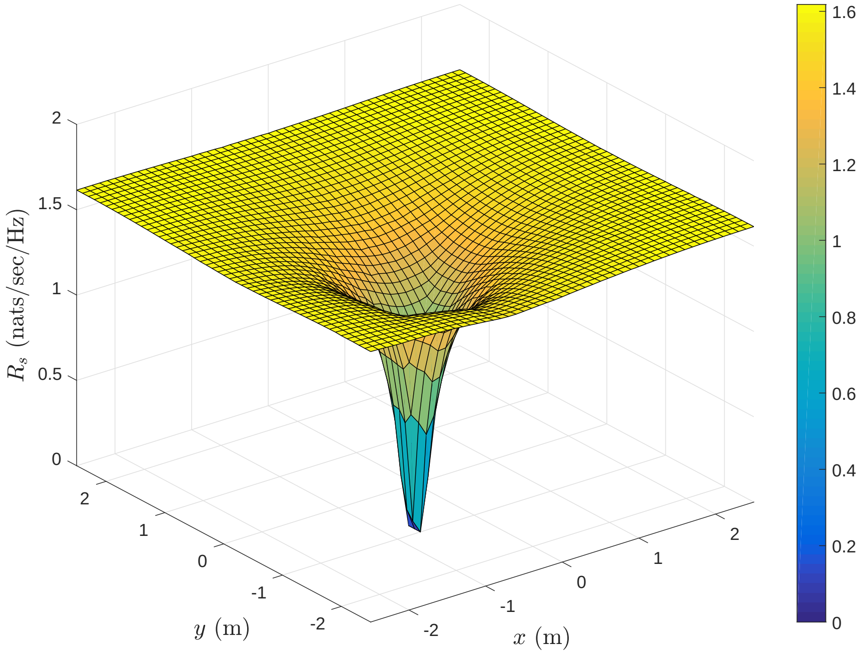

Figure 2.12: Secrecy rate obtained with the jamming beamformer (2.41) as a function of Eve's location when Bob is located at $(-0.7, -0.9)$.

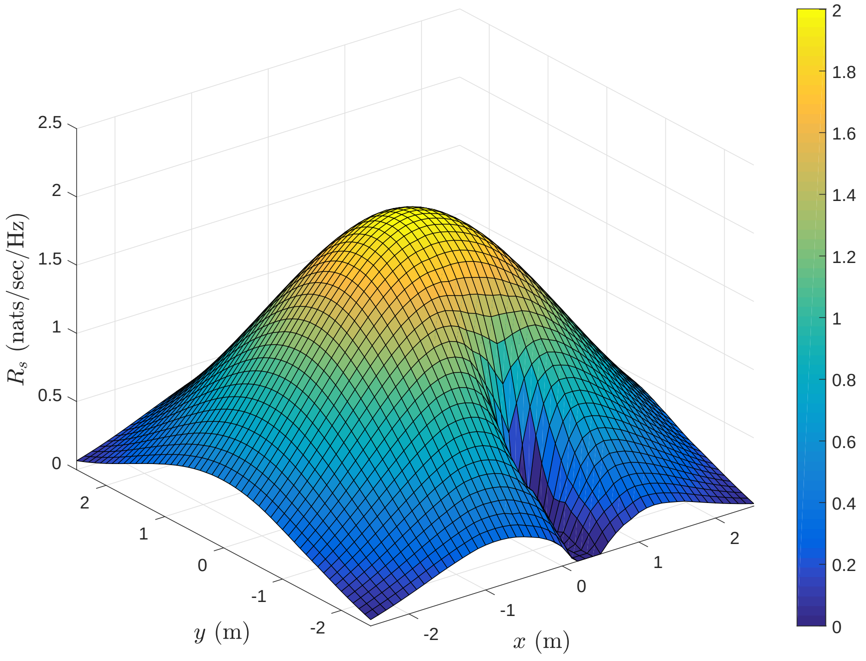

Figure 2.13: Secrecy rate obtained with the jamming beamformer (2.41) as a function of Bob's location when Eve is located at $(0.3, -1.5)$.

and the resulting secrecy rate is almost independent of Eve's channel, but is upper-bounded by the achievable rate between Alice and Bob.

2.5 Conclusions

Unlike RF channels, the VLC channel is more accurately modelled with amplitude constraints on the channel input, making it difficult to obtain analytic secrecy capacity expressions even for the simple SISO case. Therefore, we derived closed-form lower and upper bounds on the secrecy capacity of the amplitude-constrained scalar wiretap channel. Then, we utilized beamforming to obtain an achievable secrecy rate for the MISO channel. The numerical results revealed that ZF is an appropriate strategy for secure transmission in VLC scenarios, provided that the transmitter has accurate channel information. When feasible, ZF is a favorable transmission scheme as it eliminates the need to use secrecy codebooks.

We also derived a closed-form secrecy rate expression for the scalar wiretap channel when the signal received by the eavesdropper is degraded by amplitude-constrained jamming signals transmitted from a helper node. In addition, we formulated a simple linear program to optimize the jamming beamformer, assuming perfect channel information.

In the next chapter, we will focus on the MISO channel and study the design of beamformers for secrecy rate maximization subject to amplitude constraints.

Chapter 3

Optimal and Robust Beamforming for Secure MISO VLC Links

3.1 Introduction

In the previous chapter, we utilized the uniform input distribution to derive closed-form secrecy rate expressions for the scalar wiretap channel under the amplitude constraint. Then, we leveraged beamforming to obtain a closed-form secrecy rate expression for the MISO wiretap channel. In this chapter, we focus on the design of the beamformer itself. In particular, we study the design of transmit beamformers for secure downlink transmission in indoor MISO VLC links in the presence of a passive eavesdropper (Eve) attempting to overhear the message conveyed by light waves to the intended receiver (Bob). Assuming uniform input distribution, our performance measure is the secrecy rate expression (2.29) derived in the previous chapter for the amplitude-constrained MISO wiretap channel.

Under the premise of perfect channel information, we first consider the design of *optimal beamformers* that maximize the achievable secrecy rate subject to amplitude constraints. Such constraints render the optimization problem nonconvex and difficult to solve. Nevertheless, we show that this nonconvex problem can be recast as a solvable quasiconvex line search problem. We then consider the more general and more realistic case in which the transmitter (Alice) has uncertain information regarding Bob's and Eve's channels. We study the design of *robust beamformers* that maximize the *worst-case secrecy rate*, again subject to amplitude constraints. The resulting max-min optimization problem is more complex than its non-robust counterpart, but still can be reformulated as a quasiconvex line search problem. Tractability of the reformulated problem, however, depends on the geometries of the uncertainty sets. For Bob's channel, we consider uncertainty arising from quantization errors imposed by the finite rate of the feedback channel. Such uncertainty is well modelled with N-dimensional spherical sets centered at the nominal estimate available to Alice, where N is the number of transmit elements. For Eve's channel, however, we do not assume any feedback because Eve is a passive eavesdropper. Instead, we take advantage of the fact that the line-of-sight (LoS) path is typically dominant in VLC channels. Moreover, the LoS channel gain can be accurately approximated by a deterministic function of the receiver's location and orientation, along with the emission pattern of the LEDs (recall the LoS channel gain expression in (1.4)). In typical VLC scenarios, it is sensible to assume that Alice has some knowledge of Eve's location and orientation (recall, for example, the scenario in Figure 1.1). Thus, a reasonable estimate of Eve's channel can be obtained from such information. Accordingly, we derive uncertainty sets that reflect Alice's imprecise knowledge of Eve's location and orientation, as well as the emission pattern of the LEDs. We also consider possible channel mismatches caused by non-line-of-sight (NLoS) components. Such components are due to diffuse reflections from nearby surfaces, and they are not taken into account by the channel gain equation in (1.4). All the derived uncertainty sets are well structured in the sense that they lead to solvable worst-case secrecy rate maximization problems.

The secrecy performance of the Gaussian MISO wiretap channel with perfect channel information, subject to a total average power constraint, was studied in [72, 38, 39, 43]. Lower bounds on the secrecy capacity were obtained in [72] and [38]. In addition, it was shown in [38] that beamforming is the optimal transmission strategy if the channel inputs are Gaussian. These results were generalized in [39] and [43] where it was shown that Gaussian signaling, along with beamforming, is in fact optimal, and closed-form secrecy capacity expressions were derived.

The design of robust transmission schemes with imperfect channel information, based on worst-case secrecy rate maximization, was considered in [73, 74, 75, 76, 77, 78]. In [73], the authors observed similarities between the cognitive radio and wiretap channel models, and considered the design of robust beamformers in conjunction with spherical uncertainty sets for Eve's channel. The authors in [74] studied robust beamforming along with discrete uncertainty sets corresponding to inaccurate information regarding Eve's location, under the assumption of LoS propagation for RF channels. Worst-case secrecy rate maximization for the MISO channel wiretapped by multiple eavesdroppers having multiple antennas was considered in [75] using spherical uncertainty sets for the receiver's and eavesdroppers' channels. In [76], the authors considered the use of artificial noise generated by a friendly jammer and studied the design of robust data and jamming covariance matrices, under both individual and global power constraints. The work in [77] considered the design of robust transmit covariance matrices for the MIMO wiretap channel in the low SNR regime using a linearized secrecy rate expression, i.e., the secrecy rate is approximated by a linear function of the covariance matrix. A similar approach was utilized in [78] where the data and jamming covariance matrices are alternatively optimized after linearizing the nonconcave term in the secrecy rate expression based on Taylor's first-order approximation.

Compared to the previously mentioned works, our work in this chapter has the following two key differences:

1) We design the beamformer $\mathbf{w}$ subject to a per-transmit-element amplitude constraint, i.e., $\|\mathbf{w}\|_\infty \leq 1$. As mentioned in Section 1.2.2, amplitude constraints explicitly arise in VLC systems because of limitations on the linear operation region of the LEDs. Furthermore, as a side advantage, our approach to solve the design problem is in fact applicable to general l_p-norm constraints, i.e., $\|\mathbf{w}\|_p \leq 1$, for any $p \geq 1$. On the other hand, the works in [72, 38, 39, 43, 73, 74, 75, 76, 77, 78] consider a total power constraint P_{Tot} on the transmitted signal vector, that is $\|\mathbf{w}\|_2 \leq \sqrt{P_{\mathrm{Tot}}}$, or, more generally, $\mathrm{Tr}(\mathbb{E}\{\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\}) \leq P_{\mathrm{Tot}}$, where $\mathbb{E}\{\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\}$ is the transmit covariance matrix.

2) We do not assume feedback from Eve regarding her channel information. Instead, we exploit Alice's imprecise knowledge of Eve's location and orientation to obtain an estimate of Eve's channel gain. Specifically, we derive uncertainty sets for Eve's channel based on the uncertain parameters in the LoS channel gain equation in (1.4). We also consider uncertainty caused by the NLoS components. On the other hand, the works in [73, 75, 76, 77, 78] assume spherical uncertainty sets for Eve's channel, that is $\|\mathbf{h}_{\mathrm{E}} - \hat{\mathbf{h}}_{\mathrm{E}}\|_2 \leq \epsilon_{\mathbf{h}_{\mathrm{E}}}$, where $\hat{\mathbf{h}}_{\mathrm{E}}$ is Alice's erroneous estimate of $\mathbf{h}_{\mathrm{E}}$, and $\epsilon_{\mathbf{h}_{\mathrm{E}}}$ is some known constant. This model is well accepted to take into account channel uncertainty caused by limited feedback from the receiver [57, Lemma 1]. In wiretap scenarios, however, the spherical uncertainty model becomes inapplicable if Eve is a passive eavesdropper and not part of the communication network.

The remainder of this chapter is organized as follows. The system model is described in Section 3.2. In Section 3.3, we consider the design of optimal and robust beamformers under the assumptions of perfect and imperfect channel information, respectively. In Section 3.4, we derive uncertainty sets for Eve's channel based on the uncertain parameters in the LoS channel gain equation. In Section 3.5, we provide numerical examples to compare the performance of the proposed beamformers with conventional schemes, and evaluate the worst-case secrecy rate performance in a typical VLC scenario. We conclude the chapter in Section 3.6.

3.2 System Model

We consider secure downlink transmission from Alice to Bob over an indoor VLC link in the presence of a passive eavesdropper, Eve (recall the scenario in Figure 2.3). The service area is illuminated by N_{Fix} light fixtures attached to the ceiling. Each fixture encloses N_{LED} high-brightness LEDs that can be modulated independently of each other using separate drivers. Thus, the total number of LEDs is $N = N_{\text{Fix}} \times N_{\text{LED}}$.

Next, we recall the beamforming scheme described in Section 1.3.3 whereby the transmitted signal vector $\mathbf{x}(t) \in \mathbb{R}^N$ is constructed as

$$\mathbf{x}(t) = \mathbf{w}s(t), \tag{3.1}$$

where $\mathbf{w} \in \mathbb{R}^N$ is the beamformer and $s(t) \in \mathbb{R}$ is the data symbol. Due to linearity limitations of the LEDs, the transmitted signal vector $\mathbf{x}(t)$ must satisfy the amplitude constraint

$$\|\mathbf{x}(t)\|_\infty \leq A \ \forall t. \tag{3.2}$$

In order to satisfy (3.2), we let $S \sim \mathcal{U}[-A, A]$, where S is the random variable counterpart of the data symbol $s(t)$, and choose the beamformer $\mathbf{w}$ such that it satisfies the constraint

$$\|\mathbf{w}\|_\infty \leq 1. \tag{3.3}$$

Thus, utilizing the MISO channel model in (1.10), the signals received by Bob and Eve, respectively, are

$$y_{\mathrm{B}}(t) = \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}s(t) + n_{\mathrm{B}}(t), \tag{3.4a}$$

$$y_{\mathrm{E}}(t) = \mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}s(t) + n_{\mathrm{E}}(t), \tag{3.4b}$$

where $\mathbf{h}_{\mathrm{B}} \in \mathbb{R}_+^N$ and $\mathbf{h}_{\mathrm{E}} \in \mathbb{R}_+^N$ are Bob's and Eve's channel gain vectors, respectively, and $n_{\mathrm{B}}(t)$ and $n_{\mathrm{E}}(t)$ are i.i.d. Gaussian noise samples with variance σ^2.

3.3 Optimal and Robust Beamformer Design

3.3.1 Problem Formulation

Utilizing the result of Proposition 2.4, an achievable secrecy rate, in (bits/sec/Hz), for the MISO wiretap channel in (3.4) is

$$R_s = \left[\frac{1}{2}\log_2 \frac{6A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w})^2 + 3\pi e\sigma^2}{\pi e A^2(\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w})^2 + 3\pi e\sigma^2}\right]^+, \tag{3.5}$$

where the beamformer $\mathbf{w}$ is subject to the amplitude constraint $\|\mathbf{w}\|_\infty \leq 1$. A typical problem of interest is to find the optimal beamformer $\mathbf{w}^\star$ that maximizes the achievable secrecy rate, i.e.,

$$\mathbf{w}^\star = \underset{\|\mathbf{w}\|_\infty \leq 1}{\operatorname{argmax}}\ R_s. \tag{3.6}$$

In fact, our main goal in this chapter is to solve the design problem in (3.6). To this end, we have to overcome two major difficulties. Firstly, the optimization problem in (3.6) is clearly nonconvex, and the amplitude constraint $\|\mathbf{w}\|_\infty \leq 1$ makes it different from the well-known *Rayleigh quotient maximization* problem. In the next subsection, we introduce Proposition 3.1 to transform this nonconvex problem into a solvable quasiconvex line search problem. Secondly, it is unrealistic to assume that the channel gain vectors $\mathbf{h}_\mathrm{B}$ and $\mathbf{h}_\mathrm{E}$ are precisely known to Alice. Therefore, a more appropriate design approach is to devise reasonable uncertainty sets, $\mathcal{H}_\mathrm{B}$ and $\mathcal{H}_\mathrm{E}$, that enclose all possible realizations of $\mathbf{h}_\mathrm{B}$ and $\mathbf{h}_\mathrm{E}$, respectively, and solve the *robust counterpart* [79] of (3.6) to maximize the secrecy rate corresponding to the worst-case realization of $(\mathbf{h}_\mathrm{B}, \mathbf{h}_\mathrm{E}) \in \mathcal{H}_\mathrm{B} \times \mathcal{H}_\mathrm{E}$. That is to solve

$$\underset{\|\mathbf{w}\|_\infty \leq 1}{\text{maximize}}\; R_s \quad \forall (\mathbf{h}_\mathrm{B}, \mathbf{h}_\mathrm{E}) \in \mathcal{H}_\mathrm{B} \times \mathcal{H}_\mathrm{E}, \tag{3.7a}$$

or, equivalently,

$$\underset{\|\mathbf{w}\|_\infty \leq 1}{\text{maximize}}\; \min_{\substack{\mathbf{h}_\mathrm{B} \in \mathcal{H}_\mathrm{B},\\ \mathbf{h}_\mathrm{E} \in \mathcal{H}_\mathrm{E}}} R_s. \tag{3.7b}$$

We will tackle the robust design problem (3.7) in Section 3.3.3 via Proposition 3.2, whereas in Section 3.4, we shall discuss methods to model uncertainty in Eve's channel, in VLC scenarios, without feedback from Eve.

3.3.2 Optimal Beamforming with Perfect Channel Information

Our focus in this subsection is on solving the design problem in (3.6) under the premise of perfect channel information. Although the constraint on the beamformer is specified by $\|\mathbf{w}\|_\infty \leq 1$, i.e., an amplitude or l_∞-norm constraint, we shall in fact solve the problem subject to a general l_p-norm constraint, i.e., for any $p \geq 1$.

Proposition 3.1. *(Certain $\mathbf{h}_{\mathrm{B}}$ and $\mathbf{h}_{\mathrm{E}}$) Let $\|\mathbf{w}\|_p$, $p \geq 1$, denote the l_p-norm of $\mathbf{w}$, then the maximization problem*

$$\underset{\|\mathbf{w}\|_p \leq 1}{\text{maximize}} \ \frac{6A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w})^2 + 3\pi e\sigma^2}{\pi e A^2(\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w})^2 + 3\pi e\sigma^2} \tag{3.8}$$

is equivalent to the quasiconvex optimization problem (or quasiconcave maximization problem)

$$\underset{\alpha \in [\alpha_{\min}, \sqrt{6/\pi e}]}{\text{maximize}} \ \frac{6A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}_\alpha)^2 + 3\pi e\sigma^2}{\pi e \alpha^2 A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}_\alpha)^2 + 3\pi e\sigma^2}, \tag{3.9}$$

where $\alpha_{\min}$, the lower bound on α, is

$$\alpha_{\min} = \min_{\mathbf{w},\alpha} \ \alpha \tag{3.10a}$$

$$\text{s.t.} \ \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} = 1, \tag{3.10b}$$

$$|\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \alpha, \tag{3.10c}$$

and, for each $\alpha \in [\alpha_{\min}, \sqrt{6/\pi e}]$, $\mathbf{w}_\alpha$ is obtained by

$$\mathbf{w}_\alpha = \underset{\|\mathbf{w}\|_p \leq 1}{\text{argmax}} \ \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} \tag{3.11a}$$

$$\text{s.t.} \ |\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \alpha \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}. \tag{3.11b}$$

Proof: Our goal is to prove that the problem in (3.8) is equivalent to the line search problem in (3.9), and the objective function in (3.9) is quasiconcave w.r.t. the search variable α. The latter part, in particular, is not straightforward.

Using the auxiliary variable $\tau \geq 3\pi e\sigma^2$, the problem in (3.8) can be expressed as

$$\underset{\|\mathbf{w}\|_p \leq 1, \tau}{\text{maximize}} \ \frac{6A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w})^2 + 3\pi e\sigma^2}{\tau} \tag{3.12a}$$

$$\text{s.t.} \ \ \pi e A^2(\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w})^2 + 3\pi e\sigma^2 \leq \tau, \tag{3.12b}$$

or, equivalently,

$$\underset{\tau}{\text{maximize}} \ \frac{f(\tau)}{\tau}, \tag{3.13}$$

where $f(\tau)$ is defined as

$$f(\tau) \triangleq \max_{\|\mathbf{w}\|_p \leq 1} \ 6A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w})^2 + 3\pi e\sigma^2 \tag{3.14a}$$

$$\text{s.t.} \ \ |\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \sqrt{\frac{\tau - 3\pi e\sigma^2}{\pi e A^2}}. \tag{3.14b}$$

Note that the constraints in (3.12b) and (3.14b) are equivalent. In the following, we show that the objective function in (3.13) is quasiconcave w.r.t. τ by establishing the concavity of $f(\tau)$. For notational convenience, we introduce a new variable $\varepsilon \geq 0$, defined as

$$\varepsilon \triangleq \sqrt{\frac{\tau - 3\pi e\sigma^2}{\pi e A^2}}. \tag{3.15}$$

Then, we define the *perturbation function* $\varphi(\varepsilon)$ as

$$\varphi(\varepsilon) \triangleq \max_{\|\mathbf{w}\|_p \leq 1} \ \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} \tag{3.16a}$$

$$\text{s.t.} \ \ |\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \varepsilon. \tag{3.16b}$$

It is clear that $\varphi(\varepsilon)$ is nonnegative and nondecreasing for all $\varepsilon \geq 0$. Furthermore, the perturbed problem in (3.16) is convex, and thus $\varphi(\varepsilon)$ is concave [80, Section 5.6.1].

As a consequence, $\varphi(\varepsilon)$ is continuous and its *right and left derivatives*[11], $\varphi'_+(\varepsilon)$ and $\varphi'_-(\varepsilon)$, exist for all $\varepsilon > 0$. These derivatives are nonincreasing in the sense that, for any $\varepsilon_2 > \varepsilon_1 > 0$, we have [81, Theorem 1.6]

$$\varphi'_-(\varepsilon_1) \geq \varphi'_+(\varepsilon_1) \geq \varphi'_-(\varepsilon_2) \geq \varphi'_+(\varepsilon_2) \geq 0, \tag{3.17}$$

where the last inequality holds since $\varphi(\varepsilon)$ is nondecreasing. Moreover, for any $\varepsilon_0 \geq 0$ and any $\varepsilon \in \{\varepsilon : \varepsilon > 0, \varphi'_+(\varepsilon) = \varphi'_-(\varepsilon)\}$, i.e., any ε at which $\varphi(\varepsilon)$ is differentiable, we have [80, Section 3.1.3]

$$\varphi(\varepsilon_0) \leq \varphi(\varepsilon) + \varphi'(\varepsilon)(\varepsilon_0 - \varepsilon). \tag{3.18}$$

Substituting with $\varepsilon_0 = 0$ into (3.18), we get

$$\varphi(\varepsilon) \geq \varphi(0) + \varepsilon\varphi'(\varepsilon) \geq \varepsilon\varphi'(\varepsilon), \tag{3.19}$$

where the second inequality holds since $\varphi(0)$ is nonnegative. We are now ready to prove that $f(\tau) \equiv 6A^2(\varphi(\varepsilon))^2 + 3\pi e\sigma^2$ is concave w.r.t. $\tau \equiv \pi e A^2\varepsilon^2 + 3\pi e\sigma^2$. The right and left derivatives of $f(\tau)$ can be written in terms of $\varphi'_+(\varepsilon)$ and $\varphi'_-(\varepsilon)$, respectively, as

$$f'_+(\tau) = \frac{6}{\pi e}\frac{\varphi(\varepsilon)}{\varepsilon}\varphi'_+(\varepsilon), \quad f'_-(\tau) = \frac{6}{\pi e}\frac{\varphi(\varepsilon)}{\varepsilon}\varphi'_-(\varepsilon). \tag{3.20}$$

From (3.17) and (3.20), it is clear that

$$f'_-(\tau) \geq f'_+(\tau) \quad \text{for any } \tau > 3\pi e\sigma^2. \tag{3.21}$$

[11] We resort to one-sided derivatives, rather than the ordinary two-sided derivative $\varphi'(\varepsilon)$, because $\varphi(\varepsilon)$ is not necessarily smooth or differentiable over the whole interior of its domain. Particularly, there exist, in general, some $\varepsilon > 0$ at which $\varphi'_+(\varepsilon) \neq \varphi'_-(\varepsilon)$. These are the points where $\varphi'_+(\varepsilon)$ and $\varphi'_-(\varepsilon)$ have jump discontinuities. Nevertheless, since $\varphi(\varepsilon)$ is concave, there are only countably many such jumps, i.e., the set $\{\varepsilon : \varepsilon > 0, \varphi'_+(\varepsilon) \neq \varphi'_-(\varepsilon)\}$ has zero Lebesgue measure [81, Section 1.8].

Furthermore, when $\varphi(\varepsilon)$ is twice differentiable (and consequently $f(\tau)$ is twice differentiable), we have

$$f''(\tau) = \frac{3}{\pi e \varepsilon^2}\left(\left(\varphi'(\varepsilon) - \frac{\varphi(\varepsilon)}{\varepsilon}\right)\varphi'(\varepsilon) + \varphi(\varepsilon)\varphi''(\varepsilon)\right) \leq 0, \tag{3.22}$$

where the inequality holds since $\varphi'(\varepsilon) \leq \varphi(\varepsilon)/\varepsilon$, $\varphi'(\varepsilon) \geq 0$, $\varphi(\varepsilon) \geq 0$, and $\varphi''(\varepsilon) \leq 0$ (the last inequality follows from (3.17) or the second-order condition of concavity [80, Section 3.1.4]). Combining (3.21) and (3.22) yields

$$f'_-(\tau_1) \geq f'_+(\tau_1) \geq f'_-(\tau_2) \geq f'_+(\tau_2), \tag{3.23}$$

for any $\tau_2 > \tau_1 > 3\pi e\sigma^2$. Hence, $f(\tau)$ is concave [82, Theorem 24.2]. Then, it is straightforward to verify that $f(\tau)/\tau$ is quasiconcave by noting that all the β-superlevel sets $\{\tau : \tau \geq 3\pi e\sigma^2,\ f(\tau)/\tau \geq \beta\}$, for all $\beta \in \mathbb{R}$, are convex, i.e., intervals, including the empty set and infinite intervals [80, Section 3.4.1].

Next, we define the new variable $\alpha \geq 0$ as

$$\alpha \triangleq \frac{\varepsilon}{\varphi(\varepsilon)} = \sqrt{\frac{\tau - 3\pi e\sigma^2}{\pi e A^2(\varphi(\varepsilon))^2}}, \quad \varphi(\varepsilon) \neq 0. \tag{3.24}$$

Alternatively, for some given $\alpha \geq 0$, τ can be expressed in terms of α as

$$\tau = g(\alpha) \triangleq \pi e\alpha^2 A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}_\alpha)^2 + 3\pi e\sigma^2, \tag{3.25}$$

where $\mathbf{w}_\alpha$ is defined as

$$\mathbf{w}_\alpha \triangleq \underset{\|\mathbf{w}\|_p \leq 1}{\operatorname{argmax}} \;\; \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} \tag{3.26a}$$

$$\text{s.t.} \;\; |\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \alpha \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}. \tag{3.26b}$$

The problem in (3.26) is clearly equivalent to the perturbed problem in (3.16) when α and ε satisfy (3.24), or, equivalently, when α and τ satisfy (3.25). Thus, $\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}_\alpha \equiv \varphi(\varepsilon)$. Furthermore, we note from (3.26) that $\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}_\alpha$ is nondecreasing w.r.t. α (since increasing α relaxes the constraint in (3.26b)). Thus, $g(\alpha)$, as defined in (3.25), is a strictly increasing function of α. Substituting with $\tau = g(\alpha)$ back into (3.13) and changing the optimization variable into α, the problem in (3.13) can be written as

$$\underset{\alpha}{\text{maximize}} \; \frac{f(g(\alpha))}{g(\alpha)},$$

or, equivalently,

$$\underset{\alpha}{\text{maximize}} \; \frac{6A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}_\alpha)^2 + 3\pi e\sigma^2}{\pi e\alpha^2 A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}_\alpha)^2 + 3\pi e\sigma^2}, \tag{3.27}$$

where $\mathbf{w}_\alpha$ is as defined in (3.26). Since $f(\tau)/\tau$ is quasiconcave w.r.t. τ, and $\tau = g(\alpha)$ is strictly increasing w.r.t. α, we conclude that $f(g(\alpha))/g(\alpha)$ is quasiconcave w.r.t. α, and hence the problem in (3.27) is quasiconvex, i.e., a quasiconcave maximization problem.

Finally, the search interval for optimal α can be lower-bounded by the smallest feasible α, given by

$$\begin{aligned}\alpha_{\min} = \min_{\mathbf{w},\alpha} \;\; & \alpha && (3.28\text{a})\\ \text{s.t.} \;\; & \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} = 1, && (3.28\text{b})\\ & |\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \alpha, && (3.28\text{c})\end{aligned}$$

and the upper bound $\alpha_{\max} = \sqrt{6/\pi e}$ is simply obtained by noting that $f(g(\alpha)) \geq g(\alpha)$, and thus $R_{\mathrm{s}} \geq 0$, only if $\alpha \leq \sqrt{6/\pi e}$, which completes the proof. ■

Remarks:

- Proposition 3.1 has a practical interpretation. It states that the achievable secrecy rate is a quasiconcave function of the parameter α, which is the ratio of the signal level at Eve to the signal level at Bob. This is provably true for an arbitrary l_p-norm constraint on the beamformer $\mathbf{w}$.
- Setting $\alpha = 0$ in (3.11) corresponds to ZF, i.e., $\mathbf{w}_{\alpha=0}$ is the best ZF beamformer.
- If $N \geq 2$, and $\mathbf{h}_{\mathrm{B}}$ and $\mathbf{h}_{\mathrm{E}}$ are linearly independent, then $\alpha_{\min} = 0$ and ZF is feasible.
- The case of $K > 1$ *colluding eavesdroppers*, or, equivalently, a single eavesdropper having K receiving elements, can be also handled using Proposition 3.1 after replacing the inequalities in (3.10c) and (3.11b) with $\|\mathbf{H}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}\|_2 \leq \alpha$ and $\|\mathbf{H}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}\|_2 \leq \alpha\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w}$, respectively, where $\mathbf{H}_{\mathrm{E}} \triangleq [\mathbf{h}_{\mathrm{E}_1} \; \ldots \; \mathbf{h}_{\mathrm{E}_K}]$ and $\mathbf{h}_{\mathrm{E}_k}$, $k = 1, \ldots, K$, is the channel gain vector of the kth eavesdropper.

Proposition 3.1 involves two optimization problems; the outer problem (3.9) and the inner problem (3.11). The outer problem is a quasiconvex line search problem whose globally optimal solution can be found by performing a bisection search on

$\alpha \in [\alpha_{\min}, \sqrt{6/\pi e}]$, on a logarithmic scale. In the next subsection, we propose Algorithm 3.1 to solve (3.9), as well as the corresponding problem in the more general case of uncertain channel information. In each iteration of the bisection search, the inner problem (3.11) should be solved to obtain $\mathbf{w}_\alpha$ and calculate the objective function in (3.9). The inner problem is clearly convex for any $p \geq 1$, and thus it can be efficiently solved.

Using (3.9), the achievable secrecy rate, as a function of α, is

$$R_s(\alpha) = \left[\frac{1}{2}\log_2 \frac{6A^2(\mathbf{h}_\mathrm{B}^\mathrm{T}\mathbf{w}_\alpha)^2 + 3\pi e\sigma^2}{\pi e\alpha^2 A^2(\mathbf{h}_\mathrm{B}^\mathrm{T}\mathbf{w}_\alpha)^2 + 3\pi e\sigma^2}\right]^+ . \tag{3.29}$$

Let $\alpha^\star$ denote the global maximizer of (3.9), then the optimal beamformer $\mathbf{w}^\star$ is the solution of (3.11) corresponding to $\alpha = \alpha^\star$, i.e., $\mathbf{w}^\star \equiv \mathbf{w}_{\alpha^\star}$, and the maximum achievable secrecy rate is $R_s(\alpha^\star)$.

3.3.3 Robust Beamforming with Imperfect Channel Information

In this subsection, we extend Proposition 3.1 to take into account uncertainty in channel information for both Bob and Eve.

Proposition 3.2. *(Uncertain $\mathbf{h}_\mathrm{B}$ and $\mathbf{h}_\mathrm{E}$) Given a convex set $\mathcal{H}_\mathrm{B}$ and an arbitrary set $\mathcal{H}_\mathrm{E}$, the max-min problem*

$$\underset{\|\mathbf{w}\|_p \leq 1}{\text{maximize}} \; \min_{\substack{\mathbf{h}_\mathrm{B}\in\mathcal{H}_\mathrm{B},\\ \mathbf{h}_\mathrm{E}\in\mathcal{H}_\mathrm{E}}} \frac{6A^2(\mathbf{h}_\mathrm{B}^\mathrm{T}\mathbf{w})^2 + 3\pi e\sigma^2}{\pi e A^2(\mathbf{h}_\mathrm{E}^\mathrm{T}\mathbf{w})^2 + 3\pi e\sigma^2}, \tag{3.30}$$

for any $p \geq 1$, is equivalent to the quasiconvex problem

$$\underset{\alpha\in[\alpha_{\min},\sqrt{6/\pi e}]}{\text{maximize}} \ \frac{6A^2t_\alpha^2+3\pi e\sigma^2}{\pi e\alpha^2A^2t_\alpha^2+3\pi e\sigma^2}, \tag{3.31}$$

where $\alpha_{\min}$ is

$$\begin{aligned} \alpha_{\min} = \min_{\mathbf{w},\alpha} \quad & \alpha && (3.32a)\\ \text{s.t.} \quad & \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} \geq 1 \qquad \forall \mathbf{h}_{\mathrm{B}} \in \mathcal{H}_{\mathrm{B}}, && (3.32b)\\ & |\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \alpha \quad \forall \mathbf{h}_{\mathrm{E}} \in \mathcal{H}_{\mathrm{E}}, && (3.32c) \end{aligned}$$

and, for each $\alpha \in [\alpha_{\min}, \sqrt{6/\pi e}]$, t_α is obtained from

$$\begin{aligned} (\mathbf{w}_\alpha, t_\alpha) = \underset{\|\mathbf{w}\|_p\leq 1,t}{\text{argmax}} \quad & t && (3.33a)\\ \text{s.t.} \quad & \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} \geq t \qquad \forall \mathbf{h}_{\mathrm{B}} \in \mathcal{H}_{\mathrm{B}}, && (3.33b)\\ & |\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \alpha t \quad \forall \mathbf{h}_{\mathrm{E}} \in \mathcal{H}_{\mathrm{E}}. && (3.33c) \end{aligned}$$

Proof: The proof is mostly along the same line as that of Proposition 3.1. The max-min problem in (3.30) can be expressed as

$$\underset{\tau}{\text{maximize}} \ \frac{f(\tau)}{\tau}, \tag{3.34}$$

where $f(\tau)$ is defined as

$$\begin{aligned} f(\tau) \triangleq \max_{\|\mathbf{w}\|_p \leq 1} \min_{\mathbf{h}_\mathrm{B} \in \mathcal{H}_\mathrm{B}} & \; 6A^2(\mathbf{h}_\mathrm{B}^\mathrm{T}\mathbf{w})^2 + 3\pi e\sigma^2 && (3.35a) \\ \text{s.t. } & |\mathbf{h}_\mathrm{E}^\mathrm{T}\mathbf{w}| \leq \sqrt{\frac{\tau - 3\pi e\sigma^2}{\pi e A^2}} \quad \forall \mathbf{h}_\mathrm{E} \in \mathcal{H}_\mathrm{E}. && (3.35b) \end{aligned}$$

Next, we define the perturbation function $\varphi(\varepsilon)$ as

$$\begin{aligned} \varphi(\varepsilon) \triangleq \max_{\|\mathbf{w}\|_p \leq 1, t} & \; t && (3.36a) \\ \text{s.t. } & |\mathbf{h}_\mathrm{B}^\mathrm{T}\mathbf{w}| \geq t \quad \forall \mathbf{h}_\mathrm{B} \in \mathcal{H}_\mathrm{B}, && (3.36b) \\ & |\mathbf{h}_\mathrm{E}^\mathrm{T}\mathbf{w}| \leq \varepsilon \quad \forall \mathbf{h}_\mathrm{E} \in \mathcal{H}_\mathrm{E}, && (3.36c) \end{aligned}$$

where ε is defined as in (3.15). Note from (3.35) and (3.36) that $f(\tau) \equiv 6A^2(\varphi(\varepsilon))^2 + 3\pi e\sigma^2$. Note also that, unlike (3.16), the perturbed problem in (3.36) is not convex because of the constraint in (3.36b). This nonconvexity can be eliminated by imposing the additional constraint

$$\mathbf{h}_\mathrm{B}^\mathrm{T}\mathbf{w} \geq 0 \quad \forall \mathbf{h}_\mathrm{B} \in \mathcal{H}_\mathrm{B}, \tag{3.37}$$

or, equivalently, replacing (3.36b) with

$$\mathbf{h}_\mathrm{B}^\mathrm{T}\mathbf{w} \geq t \quad \forall \mathbf{h}_\mathrm{B} \in \mathcal{H}_\mathrm{B}. \tag{3.38}$$

The additional constraint, however, may render the solution suboptimal. Let $\underline{\varphi}(\varepsilon)$

be defined as

$$\underline{\varphi}(\varepsilon) \triangleq \max_{\|\mathbf{w}\|_p \leq 1, t} \quad t \tag{3.39a}$$

$$\text{s.t.} \quad \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} \geq t \qquad \forall \mathbf{h}_{\mathrm{B}} \in \mathcal{H}_{\mathrm{B}}, \tag{3.39b}$$

$$|\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \varepsilon \quad \forall \mathbf{h}_{\mathrm{E}} \in \mathcal{H}_{\mathrm{E}}. \tag{3.39c}$$

Then, $\underline{\varphi}(\varepsilon) \leq \varphi(\varepsilon)$, i.e., a nonzero gap may exist between the two optimal values. In the sequel, we show that this gap actually disappears with an additional technical assumption on $\mathcal{H}_{\mathrm{B}}$.

Lemma 3.1. *If $\mathcal{H}_{\mathrm{B}}$ is a convex set, then $\underline{\varphi}(\varepsilon) = \varphi(\varepsilon)$, i.e., the problems in* (3.36) *and* (3.39) *are equivalent.*

Proof: The proof is provided in Appendix B.1.

Following the same approach from the proof of Proposition 3.1, it can be shown that $f(\tau) \equiv 6A^2(\underline{\varphi}(\varepsilon))^2 + 3\pi e\sigma^2$ is concave w.r.t. τ, and thus $f(\tau)/\tau$ is quasiconcave. Next, we introduce the variable $\alpha \geq 0$ via the substitution

$$\tau = \pi e\alpha^2 A^2 t_\alpha^2 + 3\pi e\sigma^2, \tag{3.40}$$

where t_α is obtained from

$$(\mathbf{w}_\alpha, t_\alpha) = \underset{\|\mathbf{w}\|_p \leq 1, t}{\operatorname{argmax}} \quad t \tag{3.41a}$$

$$\text{s.t.} \quad \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} \geq t \qquad \forall \mathbf{h}_{\mathrm{B}} \in \mathcal{H}_{\mathrm{B}}, \tag{3.41b}$$

$$|\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \alpha t \quad \forall \mathbf{h}_{\mathrm{E}} \in \mathcal{H}_{\mathrm{E}}. \tag{3.41c}$$

Note from (3.39) and (3.41) that $t_\alpha \equiv \underline{\varphi}(\varepsilon)$ whenever α and τ satisfy (3.40). Substituting (3.40) back into (3.34), the latter can be rewritten as

$$\underset{\alpha}{\text{maximize}} \ \frac{6A^2 t_\alpha^2 + 3\pi e \sigma^2}{\pi e \alpha^2 A^2 t_\alpha^2 + 3\pi e \sigma^2}. \tag{3.42}$$

Similar to (3.27) in the proof of Proposition 3.1, we note from (3.40) and (3.41) that τ is strictly increasing w.r.t. α. Thus, the objective function in (3.42) is quasiconcave w.r.t. α. Finally, $\alpha_{\min}$ can be obtained by modifying the problem in (3.28) to

$$\begin{aligned} &\underset{\mathbf{w},\alpha}{\text{minimize}} && \alpha && \text{(3.43a)}\\ &\text{s.t.} && \mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} \geq 1 \qquad \forall \mathbf{h}_{\mathrm{B}} \in \mathcal{H}_{\mathrm{B}}, && \text{(3.43b)}\\ &&& |\mathbf{h}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}| \leq \alpha \quad \forall \mathbf{h}_{\mathrm{E}} \in \mathcal{H}_{\mathrm{E}}, && \text{(3.43c)} \end{aligned}$$

which completes the proof. ■

Remarks:

- $\alpha_{\min} > 0$ implies that ZF is not feasible.
- $\alpha_{\min} \geq \sqrt{6/\pi e}$ implies that the max-min problem is not feasible and the worst-case secrecy rate is zero (e.g., when $\mathcal{H}_{\mathrm{B}} \cap \mathcal{H}_{\mathrm{E}} \neq \emptyset$).

Similar to (3.9) in Proposition 3.1, the outer problem (3.31) is quasiconvex, and thus it can be efficiently solved by performing a bisection search on α. We propose Algorithm 3.1, provided in Table 3.1, to obtain a solution $\alpha^\star$ with accuracy ϵ_α (dB).

Assuming $\epsilon_\alpha = 0.2$ dB, Algorithm 3.1 shall converge in at most [80, Section 4.2.5]

$$\left\lceil \log_2 \left(20 \log_{10} \frac{\sqrt{6/\pi e}}{10^{-10}} \right) - \log_2 \epsilon_\alpha \right\rceil = 10 \text{ iterations.}$$

Table 3.1: Bisection search to solve the maximization problem in (3.31).

Algorithm 3.1 Bisection search to solve (3.31) in Proposition 3.2

```
Solve (3.32) to obtain $\alpha_{\min}$
**if** $\alpha_{\min} < 10^{-10}$, **then** $\alpha_{\min} := 10^{-10}$
Initialize $\overline{\alpha} = 20\log_{10}\sqrt{6/\pi e}$ and $\underline{\alpha} = 20\log_{10}\alpha_{\min}$
**given** the required accuracy $\epsilon_\alpha$ (dB), set the positive constant $\Delta_\alpha$ such that $0 < 20\log_{10}\Delta_\alpha < \epsilon_\alpha$
**while** $\overline{\alpha} - \underline{\alpha} \geq \epsilon_\alpha$ **do**
    $\alpha_{(\mathrm{dB})} := \dfrac{\overline{\alpha} + \underline{\alpha}}{2}$
    Solve (3.33) with $\alpha$ to obtain $t_\alpha$, where $\alpha = 10^{\frac{\alpha_{(\mathrm{dB})}}{20}}$
    Calculate the objective in (3.31), $f(\alpha) = \dfrac{6A^2t_\alpha^2 + 3\pi e\sigma^2}{\pi e\alpha^2A^2t_\alpha^2 + 3\pi e\sigma^2}$
    Solve (3.33) with $\alpha + \Delta_\alpha$ to obtain $t_{\alpha+\Delta_\alpha}$
    Calculate $f(\alpha + \Delta_\alpha) = \dfrac{6A^2t_{\alpha+\Delta_\alpha}^2 + 3\pi e\sigma^2}{\pi e(\alpha+\Delta_\alpha)^2A^2t_{\alpha+\Delta_\alpha}^2 + 3\pi e\sigma^2}$
    **if** $f(\alpha + \Delta_\alpha) - f(\alpha) > 0$, **then** $\underline{\alpha} := \alpha_{(\mathrm{dB})}$ **else** $\overline{\alpha} := \alpha_{(\mathrm{dB})}$
**end while**
**return** $\alpha^\star := \alpha$
```

Note, however, that the inner problem (3.33) should be solved twice in each iteration. Thus, although Proposition 3.2 is valid in principle for any convex set $\mathcal{H}_{\mathrm{B}}$ and an arbitrary set $\mathcal{H}_{\mathrm{E}}$, it is practically useful only when (3.33) is tractable, i.e., can be efficiently solved. The inner problem (3.33) is a *robust convex program* whose tractability depends solely on the geometries of $\mathcal{H}_{\mathrm{B}}$ and $\mathcal{H}_{\mathrm{E}}$ [79, 83, 84]. In Section 3.4, we use a spherical set $\mathcal{H}_{\mathrm{B}}$ to accommodate quantization errors caused by limited feedback from Bob. For Eve's channel, we use discrete, interval, and ellipsoidal sets to model different uncertainty sources that cause inaccurate estimates of $\mathbf{h}_{\mathrm{E}}$ in VLC scenarios. Using a spherical set $\mathcal{H}_{\mathrm{B}}$, and discrete, interval, or ellipsoidal sets $\mathcal{H}_{\mathrm{E}}$, and assuming that[12] $p \in \{1, 2, \infty\}$, the inner problem (3.33) can be expressed as a second-order cone program, which can be efficiently solved.

From (3.31), the worst-case secrecy rate, as a function of α, is

$$R_s^{\mathrm{wc}}(\alpha) = \left[\frac{1}{2}\log_2 \frac{6A^2 t_\alpha^2 + 3\pi e\sigma^2}{\pi e\alpha^2 A^2 t_\alpha^2 + 3\pi e\sigma^2}\right]^+. \tag{3.44}$$

The best worst-case secrecy rate is equal to $R_s^{\mathrm{wc}}(\alpha^\star)$, and is achieved by the robust beamformer $\mathbf{w}_{\alpha^\star}$.

3.4 Uncertainty Sets for the Eavesdropper's Channel in VLC Scenarios

Recall from Sections 1.3.1 and 1.3.2 that the LoS DC channel gain from the ith transmit LED can be accurately approximated by

[12] We need the assumption $p \in \{1, 2, \infty\}$ merely to state that the resulting problem is a second-order cone program. However, the problem is still convex and equally solvable, e.g., via the CVX toolbox [85], for any $p \geq 1$.

$$h_i = \eta R_{\mathrm{PD}} T_a \frac{(m+1)A_{\mathrm{PD}}}{2\pi \|\mathbf{d}_i\|_2^2} \left(\cos\zeta_i\right)^m T_s\, g_c \cos\psi_i\, I_\Psi(\psi_i) \tag{3.45a}$$

$$= \eta R_{\mathrm{PD}} T_a \frac{(m+1)A_{\mathrm{PD}}}{2\pi \|\mathbf{d}_i\|_2^{m+3}}\, d_z^m\, T_s\, g_c\, \mathbf{d}_i^{\mathrm{T}}\mathbf{u}\, I_\Psi\left(\cos^{-1}\frac{\mathbf{d}_i^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i\|_2}\right), \tag{3.45b}$$

where all the terms in (3.45) are defined in Sections 1.3.1 and 1.3.2. Note that we assume equal heights for all the LEDs, i.e., the vertical distance between the PD at the receiver and each LED is d_z regardless of the LED index (d_z is independent of i).

Now, our focus in this section is on deriving uncertainty sets for Eve's channel based on the uncertain parameters in (3.45). Our motivation towards this approach is the lack of feedback from Eve regarding her channel when Eve is a passive or non-cooperative receiver. In particular, we take advantage of the fact that $\mathbf{h}_{\mathrm{E}}$ can be predicted from Eve's location and orientation using (3.45) if the LoS path is dominant and the emission pattern of the LEDs is known. Such information can be mapped into an estimate of $\mathbf{h}_{\mathrm{E}}$ surrounded by a reasonable uncertainty set $\mathcal{H}_{\mathrm{E}}$. Unfortunately, the channel gain expression in (3.45) is quite complex, and mapping such uncertain parameters altogether into a useful $\mathcal{H}_{\mathrm{E}}$ that makes the inner problem (3.33) solvable is quite difficult. Thus, we begin with studying uncertainty sets corresponding to one uncertain parameter at a time. We also consider uncertainty caused by the NLoS components in $\mathbf{h}_{\mathrm{E}}$. Cases involving more than one uncertainty source will also be briefly discussed.

Throughout the entire section, we assume an amplitude constraint on $\mathbf{w}$, i.e., $\|\mathbf{w}\|_\infty \leq 1$. Furthermore, we assume a spherical uncertainty set for Bob's channel, i.e., $\mathbf{h}_{\mathrm{B}} \in \mathcal{H}_{\mathrm{B}}$,

$$\mathcal{H}_{\mathrm{B}} = \left\{\hat{\mathbf{h}}_{\mathrm{B}} + \mathbf{e}_{\mathbf{h}_{\mathrm{B}}} : \|\mathbf{e}_{\mathbf{h}_{\mathrm{B}}}\|_2 \leq \epsilon_{\mathbf{h}_{\mathrm{B}}}\right\}, \tag{3.46}$$

where the nominal vector $\hat{\mathbf{h}}_{\mathrm{B}}$ is known to Alice via limited feedback from Bob, and the bounded error term $\mathbf{e}_{\mathbf{h}_{\mathrm{B}}}$ is due to quantization errors. Substituting (3.46) back into (3.33b), the latter can be expressed as

$$\hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}}\|\mathbf{w}\|_2 \geq t. \tag{3.47}$$

3.4.1 Uncertain Eavesdropper's Location

In this subsection, we consider uncertainty caused by inaccurate information regarding Eve's location. We assume that Eve is located inside a three-dimensional rectangular region (or box) $\mathcal{B}$ with dimensions $(2l_x, 2l_y, 2l_z)$. We also assume, without loss of generality, that $\mathcal{B}$ is centered at the origin, i.e.,

$$\mathcal{B} = \left\{\mathbf{L}\boldsymbol{v} : \boldsymbol{v} \in \mathbb{R}^3, \|\boldsymbol{v}\|_\infty \leq 1\right\}, \tag{3.48}$$

where $\mathbf{L} \triangleq \mathbf{Diag}(l_x, l_y, l_z)$. Furthermore, we choose the origin (or the center of $\mathcal{B}$) as the nominal location of Eve.

Let $\boldsymbol{\delta} = [\delta_x \ \delta_y \ \delta_z]^{\mathrm{T}}$, $\boldsymbol{\delta} \in \mathcal{B}$, denote the deviation of the actual location of Eve from the origin. Using (3.45b), the channel gain h_i, $i = 1, \ldots, N$, anywhere inside $\mathcal{B}$, as a function of $\boldsymbol{\delta}$, is

$$h_i(\boldsymbol{\delta}) = \eta R_{\mathrm{PD}} T_a \frac{(m+1)A_{\mathrm{PD}}}{2\pi\|\mathbf{d}_i - \boldsymbol{\delta}\|_2^{m+3}} (d_z - \delta_z)^m \, T_s \, g_c \, (\mathbf{d}_i - \boldsymbol{\delta})^{\mathrm{T}}\mathbf{u} \, I_{\Psi_{\mathrm{E}}}\left(\cos^{-1}\frac{(\mathbf{d}_i - \boldsymbol{\delta})^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i - \boldsymbol{\delta}\|_2}\right), \tag{3.49}$$

and the set of all possible channel realizations inside $\mathcal{B}$ can be written as

$$\mathcal{H}_{\mathrm{E}}^{\mathcal{B}} = \{\mathbf{h}(\boldsymbol{\delta}) : \boldsymbol{\delta} \in \mathcal{B}\}. \tag{3.50}$$

If we substitute with $\mathcal{H}_{\mathrm{E}} = \mathcal{H}_{\mathrm{E}}^{\mathcal{B}}$ back into (3.33c), we will end up with an intractable semi-infinite optimization problem. Therefore, we shall discuss methods to approximate $\mathcal{H}_{\mathrm{E}}^{\mathcal{B}}$, based on the volume of $\mathcal{B}$, in order to make (3.33) solvable.

3.4.1.1 Small Uncertainty Region

For sufficiently-small $\mathcal{B}$, e.g., $\max\{2l_x, 2l_y, 2l_z\} \leq 0.5$ m, we can assume that the subset of LEDs seen by Eve's receiver at a particular location $\boldsymbol{\delta}$,

$$\mathcal{I}_{\boldsymbol{\delta}} = \left\{ i : I_{\Psi_{\mathrm{E}}}\left(\cos^{-1} \frac{(\mathbf{d}_i - \boldsymbol{\delta})^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i - \boldsymbol{\delta}\|_2} \right) = 1,\, i \in \{1, \ldots, N\} \right\},$$

is identical for all $\boldsymbol{\delta} \in \mathcal{B}$. In other words, the output of the indicator function in (3.49) is independent of $\boldsymbol{\delta}$ for all the LEDs and is solely determined by the nominal location of Eve. Under this assumption, the channel gain in (3.49) can be written as

$$h_i(\boldsymbol{\delta}) = c_i \, \frac{(d_z - \delta_z)^m (\mathbf{d}_i - \boldsymbol{\delta})^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i - \boldsymbol{\delta}\|_2^{m+3}}, \tag{3.51}$$

where

$$c_i \triangleq \eta R_{\mathrm{PD}} T_a \frac{(m+1)A_{\mathrm{PD}}}{2\pi} \, T_s \, g_c \, I_{\Psi_{\mathrm{E}}}\left(\cos^{-1} \frac{\mathbf{d}_i^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i\|_2} \right). \tag{3.52}$$

Furthermore, with sufficiently-small $\mathcal{B}$, $\mathbf{h}(\boldsymbol{\delta})$ can be well approximated by its first-order approximation around the center of $\mathcal{B}$, that is

$$\mathbf{h}(\boldsymbol{\delta}) \approx \bar{\mathbf{h}}(\boldsymbol{\delta}) = \mathbf{h}_{\mathbf{0}} + \mathbf{J}_{\mathbf{0}}\boldsymbol{\delta}, \tag{3.53}$$

where $\mathbf{h_0} \equiv \mathbf{h}(\mathbf{0})$, $\mathbf{J} \in \mathbb{R}^{N\times 3}$ is the *Jacobian matrix* (or matrix of partial derivatives), defined as

$$\mathbf{J} \triangleq \begin{bmatrix} \frac{\partial h_1(\boldsymbol{\delta})}{\partial \delta_x} & \frac{\partial h_1(\boldsymbol{\delta})}{\partial \delta_y} & \frac{\partial h_1(\boldsymbol{\delta})}{\partial \delta_z} \\ \vdots & \vdots & \vdots \\ \frac{\partial h_N(\boldsymbol{\delta})}{\partial \delta_x} & \frac{\partial h_N(\boldsymbol{\delta})}{\partial \delta_y} & \frac{\partial h_N(\boldsymbol{\delta})}{\partial \delta_z} \end{bmatrix}, \tag{3.54}$$

and $\mathbf{J_0} \equiv \mathbf{J}(\mathbf{0})$. The entries of $\mathbf{h_0}$ and $\mathbf{J_0}$ are provided in Appendix B.2. Using the linearized channel gain expression in (3.53), the uncertainty set in (3.50) can be approximated by

$$\bar{\mathcal{H}}_{\mathrm{E}}^{\mathcal{B}} = \left\{ \mathbf{h_0} + \mathbf{J_0}\mathbf{L}\boldsymbol{v} : \boldsymbol{v} \in \mathbb{R}^3, \|\boldsymbol{v}\|_\infty \leq 1 \right\}. \tag{3.55}$$

Substituting with $\bar{\mathcal{H}}_{\mathrm{E}}^{\mathcal{B}}$ back into (3.33c), the inner problem (3.33) can be expressed as

$$\underset{\|\mathbf{w}\|_\infty \leq 1, t}{\text{maximize}} \quad t \tag{3.56a}$$

$$\text{s.t.} \quad \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}}\|\mathbf{w}\|_2 \geq t, \tag{3.56b}$$

$$|\mathbf{h}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w} + \boldsymbol{v}^{\mathrm{T}}\mathbf{L}\mathbf{J}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w}| \leq \alpha t \quad \forall \boldsymbol{v} : \|\boldsymbol{v}\|_\infty \leq 1, \tag{3.56c}$$

or, equivalently,

$$\underset{\|\mathbf{w}\|_\infty \leq 1, t}{\text{maximize}} \quad t \tag{3.57a}$$

$$\text{s.t.} \quad \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}}\|\mathbf{w}\|_2 \geq t, \tag{3.57b}$$

$$\mathbf{h}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w} + \|\mathbf{L}\mathbf{J}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w}\|_1 \leq \alpha t, \tag{3.57c}$$

$$\mathbf{h}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w} - \|\mathbf{L}\mathbf{J}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w}\|_1 \geq -\alpha t, \tag{3.57d}$$

which is a second-order cone problem. Similarly, the problem in (3.32) can be expressed as

$$\begin{aligned}
&\underset{\mathbf{w},\alpha}{\text{minimize}} && \alpha && (3.58\text{a})\\
&\text{s.t.} && \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}}\|\mathbf{w}\|_2 \geq 1, && (3.58\text{b})\\
& && \mathbf{h}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w} + \|\mathbf{L}\mathbf{J}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w}\|_1 \leq \alpha, && (3.58\text{c})\\
& && \mathbf{h}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w} - \|\mathbf{L}\mathbf{J}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w}\|_1 \geq -\alpha. && (3.58\text{d})
\end{aligned}$$

3.4.1.2 Large Uncertainty Region

If the uncertainty region $\mathcal{B}$ is relatively large, the first-order approximation in (3.53) may become poor. Nevertheless, $\mathcal{B}$ can first be divided into K non-overlapping boxes, $\mathcal{B}_k$, $k = 1, \ldots, K$, such that $\bigcup_{k=1}^{K} \mathcal{B}_k = \mathcal{B}$. Then, the first-order approximation is applied inside each box, around its center, and (3.56) is solved with the corresponding K constraints.

Alternatively, the region $\mathcal{B}$ can be discretized using a three-dimensional fine grid $\dddot{\mathcal{B}}$, and the inner problem (3.33) is approximated by

$$\begin{aligned}
&\underset{\|\mathbf{w}\|_\infty \leq 1, t}{\text{maximize}} && t && (3.59\text{a})\\
&\text{s.t.} && \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}}\|\mathbf{w}\|_2 \geq t, && (3.59\text{b})\\
& && |\mathbf{h}^{\mathrm{T}}(\boldsymbol{\delta})\,\mathbf{w}| \leq \alpha t \quad \forall \boldsymbol{\delta} \in \dddot{\mathcal{B}}, && (3.59\text{c})
\end{aligned}$$

where the entries of $\mathbf{h}(\boldsymbol{\delta})$ are obtained with (3.49). Although discretization is a straightforward approach that leads to linear constraints, the number of constraints may grow up very quickly with large uncertainty regions.

3.4.2 Uncertain Eavesdropper's Orientation

In this subsection, we assume that Eve has the freedom to adjust the direction of her receiver, $(\theta_{\mathrm{E}}, \phi_{\mathrm{E}})$, $\theta_{\mathrm{E}} \in [\theta_{\min}, \theta_{\max}]$, $\phi_{\mathrm{E}} \in [\phi_{\min}, \phi_{\max}]$, to her advantage. In other words, the exact direction of Eve's receiver is unknown to Alice. The uncertainty set $\mathcal{U}$ containing all possible realizations of Eve's orientation vector $\mathbf{u}$ (refer to Figure 1.4) can be written as

$$\mathcal{U} = \left\{\mathbf{u} = [\sin\theta\cos\phi \quad \sin\theta\sin\phi \quad \cos\theta]^{\mathrm{T}} : \theta \in [\theta_{\min}, \theta_{\max}], \phi \in [\phi_{\min}, \phi_{\max}]\right\}, \quad (3.60)$$

and the channel gain h_i, $i = 1, \ldots, N$, as a function of $\mathbf{u}$, is given by

$$h_i(\mathbf{u}) = c_i \frac{d_z^m}{\|\mathbf{d}_i\|_2^{m+3}} \mathbf{d}_i^{\mathrm{T}} \mathbf{u}, \quad (3.61)$$

where c_i is as defined in (3.52). For notational convenience, let $\mathbf{D} \in \mathbb{R}^{N\times 3}$ be defined as

$$\mathbf{D} \triangleq d_z^m \left[\frac{c_1 \mathbf{d}_1}{\|\mathbf{d}_1\|_2^{m+3}} \quad \cdots \quad \frac{c_N \mathbf{d}_N}{\|\mathbf{d}_N\|_2^{m+3}}\right]^{\mathrm{T}}. \quad (3.62)$$

Then, $\mathbf{h}(\mathbf{u})$ can be expressed as

$$\mathbf{h}(\mathbf{u}) = \mathbf{D}\mathbf{u}. \quad (3.63)$$

Note from (3.52) and (3.62) that $\mathbf{D}$ depends on $\mathbf{u}$ via the indicator function in the definition of c_i, $i = 1, \ldots, N$. Thus, the mapping from $\mathbf{u}$ to $\mathbf{h}$ in (3.63) is not linear, in general. The set of all possible channel gains for Eve is given by

$$\mathcal{H}_{\mathrm{E}}^{\mathcal{U}} = \left\{\mathbf{D}\mathbf{u} : \mathbf{u} \in \mathcal{U}\right\}. \quad (3.64)$$

Substituting with $\mathcal{H}_{\mathrm{E}}^{\mathcal{U}}$ back into (3.33c), the inner problem (3.33) can be written as

$$\underset{\|\mathbf{w}\|_\infty \leq 1, t}{\text{maximize}} \quad t \tag{3.65a}$$

$$\text{s.t.} \quad \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}}\|\mathbf{w}\|_2 \geq t, \tag{3.65b}$$

$$\max_{\mathbf{u}\in\mathcal{U}} |\mathbf{u}^{\mathrm{T}}\mathbf{D}^{\mathrm{T}}\mathbf{w}| \leq \alpha t. \tag{3.65c}$$

In order to efficiently solve (3.65), we shall differentiate between two cases, as follows.

3.4.2.1 Small Angle Variations

In this case, we assume that Eve's freedom to adjust her receiver's orientation is limited in the sense that the subset of LEDs inside Eve's FoV at a particular direction $\mathbf{u}$,

$$\mathcal{I}_{\mathbf{u}} = \left\{ i : I_{\Psi_{\mathrm{E}}}\left(\cos^{-1}\frac{\mathbf{d}_i^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i\|_2}\right) = 1,\, i \in \{1, \ldots, N\} \right\},$$

remains unchanged for all $\mathbf{u} \in \mathcal{U}$. Perhaps the most practical case in which the above assumption may hold is when the permissible variations of the zenith angle θ_{E} is relatively small and close to zero, i.e., $\theta_{\mathrm{E}} \in [0, \theta_{\max}]$, where $\theta_{\max}$ is relatively small (e.g., $\theta_{\max} \leq 30°$). If $\mathcal{I}_{\mathbf{u}}$ is fixed for all $\mathbf{u} \in \mathcal{U}$, then $\mathbf{D}$ is independent of $\mathbf{u}$, and $\mathbf{h}$, as given in (3.63), is a linear function of $\mathbf{u}$. In this case, the left-hand side of the inequality in (3.65c) can be upper-bounded as

$$\max_{\mathbf{u}\in\mathcal{U}} |\mathbf{u}^{\mathrm{T}}\mathbf{D}^{\mathrm{T}}\mathbf{w}| \leq \max_{\|\mathbf{u}\|_2 \leq 1} \mathbf{u}^{\mathrm{T}}\mathbf{D}^{\mathrm{T}}\mathbf{w} = \|\mathbf{D}^{\mathrm{T}}\mathbf{w}\|_2. \tag{3.66}$$

Then, the problem in (3.65) is replaced by

$$\underset{\|\mathbf{w}\|_\infty \le 1, t}{\text{maximize}} \quad t \tag{3.67a}$$

$$\text{s.t.} \quad \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}} \mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}} \|\mathbf{w}\|_2 \ge t, \tag{3.67b}$$

$$\|\mathbf{D}^{\mathrm{T}} \mathbf{w}\|_2 \le \alpha t, \tag{3.67c}$$

which is a second-order cone problem.

3.4.2.2 Large Angle Variations

With arbitrary zenith and/or azimuth angle variations for Eve's receiver, $\mathbf{D}$ becomes dependent on $\mathbf{u}$, and linearity between $\mathbf{h}$ and $\mathbf{u}$ no longer holds. In this case, it becomes difficult to obtain a mathematically-convenient uncertainty set $\mathcal{H}_{\mathrm{E}}^{\mathcal{U}}$ over the continuum of θ_{E} and ϕ_{E}. Thus, we resort to sampling $\mathbf{h}(\mathbf{u})$ over $\mathcal{U}$, and the inner problem (3.33) is approximated by

$$\underset{\|\mathbf{w}\|_\infty \le 1, t}{\text{maximize}} \quad t \tag{3.68a}$$

$$\text{s.t.} \quad \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}} \mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}} \|\mathbf{w}\|_2 \ge t, \tag{3.68b}$$

$$|\mathbf{h}^{\mathrm{T}}(\theta, \phi)\, \mathbf{w}| \le \alpha t \quad \forall (\theta, \phi) \in \dddot{\Theta} \times \dddot{\Phi}, \tag{3.68c}$$

where the components of $\mathbf{h}(\theta, \phi)$ are obtained with (3.45b), and $\dddot{\Theta}$ and $\dddot{\Phi}$ are fine grids on the intervals $[\theta_{\min}, \theta_{\max}]$ and $[\phi_{\min}, \phi_{\max}]$, respectively.

3.4.3 Uncertain LEDs Half-Intensity Angle

Assuming generalized Lambertian emission, the emission pattern of the LEDs is fully determined by the Lambertian order

$$m = -1/\log_2(\cos\zeta^{\text{3-dB}}), \tag{3.69}$$

where $\zeta^{\text{3-dB}}$ is the half-intensity angle of the LEDs. This angle is typically specified by the LED manufacturer as a nominal value in the datasheet. In practice, however, the actual angle of each LED will deviate from the nominal value. In this subsection, we study channel uncertainty caused by this deviation. In particular, we assume an interval uncertainty model in which $\zeta^{\text{3-dB}} \in [\zeta^{\text{3-dB}}_{\text{min}}, \zeta^{\text{3-dB}}_{\text{max}}]$, and allow independent realizations of $\zeta^{\text{3-dB}}$ for each LED. Then, we map the interval $[\zeta^{\text{3-dB}}_{\text{min}}, \zeta^{\text{3-dB}}_{\text{max}}]$ into independent interval uncertainties for each entry of $\mathbf{h}_{\text{E}}$.

We begin with rewriting the channel gain from (3.45a) as

$$h_i(m_i) = \kappa_i(m_i+1)(\cos\zeta_i)^{m_i}, \quad i = 1,\ldots,N, \tag{3.70a}$$

where

$$m_i = -1/\log_2(\cos\zeta_i^{\text{3-dB}}), \quad \zeta_i^{\text{3-dB}} \in \left[\zeta^{\text{3-dB}}_{\text{min}}, \zeta^{\text{3-dB}}_{\text{max}}\right], \tag{3.70b}$$

and

$$\kappa_i \triangleq \eta R_{\text{PD}} T_a \frac{A_{\text{PD}}}{2\pi\|\mathbf{d}_i\|_2^2} T_s\, g_c \cos\psi_i\, I_{\Psi_{\text{E}}}(\psi_i). \tag{3.70c}$$

Next, we define m_{min} and m_{max}, respectively, as

$$m_{\text{min}} \triangleq -1/\log_2(\cos\zeta^{\text{3-dB}}_{\text{max}}), \tag{3.71a}$$

$$m_{\text{max}} \triangleq -1/\log_2(\cos\zeta^{\text{3-dB}}_{\text{min}}). \tag{3.71b}$$

Then, in order to map the interval $[m_{\min}, m_{\max}]$ into $[h_i^{\min}, h_i^{\max}]$, $i = 1, \ldots, N$, we first show that h_i is a quasiconcave function of m_i. Differentiating h_i w.r.t. m_i yields

$$h_i'(m_i) = \kappa_i (\cos\zeta_i)^{m_i} (1 + (m_i + 1)\ln(\cos\zeta_i)). \tag{3.72}$$

From (3.72), for $\kappa_i \neq 0$ and $i = 1, \ldots, N$, we note that

$$\begin{cases} h_i'(m_i) \geq 0 & \text{for } m_i \leq m_i^\star, \\ h_i'(m_i) < 0 & \text{for } m_i > m_i^\star, \end{cases}$$

where $m_i^\star \triangleq -(1 + 1/\ln(\cos\zeta_i))$. Thus, $h_i(m_i)$ is a quasiconcave function with global maximizer $m_i^\star$. Consequently, the uncertainty set $\mathcal{H}_{\mathrm{E}}^{\zeta^{\text{3-dB}}}$ corresponding to the interval $[\zeta_{\min}^{\text{3-dB}}, \zeta_{\max}^{\text{3-dB}}]$ can be written as

$$\mathcal{H}_{\mathrm{E}}^{\zeta^{\text{3-dB}}} = \left\{ [h_1 \ \ldots \ h_N]^{\mathrm{T}} : h_i \in [h_i^{\min}, h_i^{\max}], i = 1, \ldots, N \right\}, \tag{3.73a}$$

where, for $i = 1, \ldots, N$,

$$h_i^{\min} = \begin{cases} h_i(m_{\min}) & \text{if } m_i^\star > m_{\max}, \\ \min\{h_i(m_{\min}), h_i(m_{\max})\} & \text{if } m_i^\star \in [m_{\min}, m_{\max}], \\ h_i(m_{\max}) & \text{if } m_i^\star < m_{\min}, \end{cases} \tag{3.73b}$$

$$h_i^{\max} = \begin{cases} h_i(m_{\max}) & \text{if } m_i^\star > m_{\max}, \\ h_i(m_i^\star) & \text{if } m_i^\star \in [m_{\min}, m_{\max}], \\ h_i(m_{\min}) & \text{if } m_i^\star < m_{\min}. \end{cases} \tag{3.73c}$$

Define $\hat{\mathbf{h}} \in \mathbb{R}_+^N$ and $\hat{\mathbf{H}} \in \mathbb{R}_+^{N \times N}$, respectively, as

$$\hat{\mathbf{h}} \triangleq \frac{1}{2}[h_1^{\max} + h_1^{\min} \ \dots \ h_N^{\max} + h_N^{\min}]^{\mathrm{T}}, \tag{3.74a}$$

$$\hat{\mathbf{H}} \triangleq \frac{1}{2}\mathbf{Diag}(h_1^{\max} - h_1^{\min}, \dots, h_N^{\max} - h_N^{\min}). \tag{3.74b}$$

Then, $\mathcal{H}_{\mathrm{E}}^{\zeta^{\text{3-dB}}}$ can be written as

$$\mathcal{H}_{\mathrm{E}}^{\zeta^{\text{3-dB}}} = \left\{ \hat{\mathbf{h}} + \hat{\mathbf{H}}\boldsymbol{v} : \boldsymbol{v} \in \mathbb{R}^N, \|\boldsymbol{v}\|_\infty \leq 1 \right\}. \tag{3.75}$$

Similar to (3.55)–(3.57), substituting with $\mathcal{H}_{\mathrm{E}}^{\zeta^{\text{3-dB}}}$ into (3.33c), the inner problem (3.33) can be expressed as

$$\underset{\|\mathbf{w}\|_\infty \leq 1, t}{\text{maximize}} \quad t \tag{3.76a}$$

$$\text{s.t.} \quad \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}}\|\mathbf{w}\|_2 \geq t, \tag{3.76b}$$

$$\hat{\mathbf{h}}^{\mathrm{T}}\mathbf{w} + \|\hat{\mathbf{H}}\mathbf{w}\|_1 \leq \alpha t, \tag{3.76c}$$

$$\hat{\mathbf{h}}^{\mathrm{T}}\mathbf{w} - \|\hat{\mathbf{H}}\mathbf{w}\|_1 \geq -\alpha t. \tag{3.76d}$$

3.4.4 Uncertain NLoS Components

In this subsection, we consider channel uncertainty arising from the NLoS components caused by diffuse reflections from nearby surfaces. Taking into account signal contributions from both the LoS and NLoS paths, the channel gain can be written as

$$h_i = h_i^{\mathrm{LoS}} + h_i^{\mathrm{NLoS}}, \quad i = 1, \dots, N, \tag{3.77}$$

where h_i^{LoS} is the LoS component obtained with (3.45), and h_i^{NLoS} is the unknown NLoS component. We shall consider a simple multiplicative uncertainty model in which h_i^{NLoS} is an uncertain fraction, γ_i, of h_i^{LoS}, that is

$$h_i^{\mathrm{NLoS}} = \gamma_i h_i^{\mathrm{LoS}}, \quad 0 \leq \gamma_i \leq \gamma_{\max}, \, i = 1, \ldots, N, \tag{3.78}$$

where $\gamma_{\max} \triangleq \max_i \gamma_i$. The actual value of $\gamma_{\max}$ depends mostly on the problem geometry as well as the diffuse reflectivity of nearby surfaces. In practice, $\gamma_{\max}$ can be measured or predicted using numerical simulations. Simulation results reported in [59] show $\gamma_{\max}$ of about 12% (see the discussion after Figure 6 in [59]). Note, however, that the multiplicative model in (3.78) is applicable only when the LoS path between the ith LED and the PD exists, i.e., $h_i^{\mathrm{LoS}} \neq 0$, and is dominant. In other words, (3.78) does not take into account the case in which the received signal consists entirely of NLoS components, e.g., when the LoS path is blocked or outside the receiver FoV.

From (3.77) and (3.78), the set of all possible channel gain vectors can be written as

$$\mathcal{H}_{\mathrm{E}}^{\gamma_{\max}} = \left\{ [h_1 \ \ldots \ h_N]^{\mathrm{T}} : h_i \in [h_i^{\mathrm{LoS}}, (1 + \gamma_{\max}) h_i^{\mathrm{LoS}}], i = 1, \ldots, N \right\}, \tag{3.79}$$

which is similar to $\mathcal{H}_{\mathrm{E}}^{\zeta^{\text{3-dB}}}$ in (3.73a), and thus we can proceed with the same steps from the previous subsection.

3.4.5 Combined Uncertainties

So far we have derived separate sets corresponding to uncertainties in location, orientation, half-intensity angle, and NLoS components. In practice, however, these uncertainties will mostly happen in combination with each other. Thus, more inclusive sets that take into account the aggregate uncertainty are required. Unfortunately, it is difficult, in general, to derive such sets or provide a unified treatment for different combinations of uncertainties because, as we mentioned earlier, the channel gain expression in (3.45) is a complex function of the uncertainty sources. Nevertheless, one intuitive approach to circumvent such a difficulty is to sample the channel gain vector over the variables with lower dimension or smaller uncertainty size. Consider, for example, the case of uncertain location and LEDs half-intensity angle, that is

$$\begin{aligned}\mathcal{H}_{\mathrm{E}}^{\mathcal{B}\times\boldsymbol{\zeta}^{3\text{-dB}}} &= \mathcal{H}_{\mathrm{E}}^{\mathcal{B}} \times \mathcal{H}_{\mathrm{E}}^{\boldsymbol{\zeta}^{3\text{-dB}}} \\ &= \left\{\mathbf{h}(\boldsymbol{\delta}, \boldsymbol{\zeta}^{3\text{-dB}}) : \boldsymbol{\delta} \in \mathcal{B},\, \boldsymbol{\zeta}^{3\text{-dB}} \in [\zeta_{\min}^{3\text{-dB}}, \zeta_{\max}^{3\text{-dB}}]^N\right\},\end{aligned}$$

where $\boldsymbol{\zeta}^{3\text{-dB}} = [\zeta_1^{3\text{-dB}} \ \dots \ \zeta_N^{3\text{-dB}}]^{\mathrm{T}}$. If $N > 3$, i.e., the dimension of $\boldsymbol{\zeta}^{3\text{-dB}}$ is bigger than the dimension of $\boldsymbol{\delta}$, then $\mathcal{B}$ can be discretized using a three-dimensional K-point grid, $\dddot{\mathcal{B}} = \{\boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_K\}$, and the problem in (3.76) is modified to

$$\begin{aligned}\underset{\|\mathbf{w}\|_\infty \le 1, t}{\text{maximize}} \quad & t \\ \text{s.t.} \quad & \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}}\|\mathbf{w}\|_2 \ge t, \\ & \hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w} + \|\hat{\mathbf{H}}_k\mathbf{w}\|_1 \le \alpha t, \qquad k = 1, \dots, K, \\ & \hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w} - \|\hat{\mathbf{H}}_k\mathbf{w}\|_1 \ge -\alpha t, \qquad k = 1, \dots, K,\end{aligned}$$

where $\hat{\mathbf{h}}_k$ and $\hat{\mathbf{H}}_k$ are obtained as in (3.74) using the components of $\mathbf{h}^{\min}(\boldsymbol{\delta}_k)$ and $\mathbf{h}^{\max}(\boldsymbol{\delta}_k)$, for $k = 1, \dots, K$. The same idea can be applied to other combinations of uncertainty sources.

Furthermore, there exist specific cases of combined uncertainties in which discretization may not be necessary. Consider, for example, the special, but practically relevant, case of small location and angle uncertainties. With such a combination, the linear channel gain models considered in Sections 3.4.1.1 and 3.4.2.1 are both applicable, and an explicit formulation of the optimization problem can be obtained as follows. First, we rewrite the linearized channel gain expression from (3.53) as

$$\begin{aligned}\bar{\mathbf{h}}(\boldsymbol{\delta},\mathbf{u}) &= \mathbf{h_0} + \mathbf{J_0}\boldsymbol{\delta} \\ &= \mathbf{D}\mathbf{u} + \mathbf{G_0}(\mathbf{I}_3 \otimes \mathbf{u})\boldsymbol{\delta} \\ &= \mathbf{D}\mathbf{u} + \mathbf{G_0}(\boldsymbol{\delta} \otimes \mathbf{I}_3)\mathbf{u}, \end{aligned} \tag{3.80}$$

where $\mathbf{D}$ is as defined in (3.62), and the entires of $\mathbf{G_0}$, $\mathbf{G_0} \in \mathbb{R}^{N\times 9}$, can be inferred from (B.2b)–(B.2d) in Appendix B.2. Then, the inner problem (3.33) can be expressed as

$$\underset{\|\mathbf{w}\|_\infty \le 1, t}{\text{maximize}} \quad t \tag{3.81a}$$

$$\text{s.t.} \quad \hat{\mathbf{h}}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w} - \epsilon_{\mathbf{h}_{\mathrm{B}}}\|\mathbf{w}\|_2 \ge t, \tag{3.81b}$$

$$\max_{\substack{\boldsymbol{\delta}\in\mathcal{B},\\ \|\mathbf{u}\|_2\le 1}} |\mathbf{u}^{\mathrm{T}}\mathbf{D}^{\mathrm{T}}\mathbf{w} + \mathbf{u}^{\mathrm{T}}(\boldsymbol{\delta}^{\mathrm{T}} \otimes \mathbf{I}_3)\mathbf{G}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w}| \le \alpha t. \tag{3.81c}$$

The constraint in (3.81c) can be replaced by a set of second-order cone constraints, given by

$$\|\mathbf{D}^{\mathrm{T}}\mathbf{w} + (\mathbf{v}_{(k)}^{\mathrm{T}} \otimes \mathbf{I}_3)\mathbf{G}_{\mathbf{0}}^{\mathrm{T}}\mathbf{w}\|_2 \le \alpha t, \quad k = 1,\ldots,8, \tag{3.82}$$

where $\mathbf{v}_{(k)} \in \mathbb{R}^3$, $k = 1,\ldots,8$, are the vertices (or corners) of $\mathcal{B}$.

3.5 Numerical Examples

In this section, we provide numerical examples to verify the performance gains of the beamformers proposed in Section 3.3 compared to conventional beamforming schemes. We also demonstrate the design of robust beamformers in a typical VLC scenario and investigate the resulting worst-case secrecy rate performance in conjunction with the uncertainty models devised in Section 3.4.

3.5.1 Performance Comparisons

All the results presented in this subsection are obtained under the following assumptions. The number of transmit elements is $N = 4$. The entries of $\mathbf{h}_{\mathrm{B}}$ and $\mathbf{h}_{\mathrm{E}}$ are generated i.i.d. according to the uniform distribution over the interval $[0, 1]$, and the results are averaged over 1000 independent trials. The optimal and robust beamformers are obtained via Algorithm 3.1, where the outer maximization problem is solved with accuracy $\epsilon_\alpha = 0.2$ dB and the inner problem is solved using the CVX toolbox [85] along with the MOSEK solver [86].

3.5.1.1 Optimal versus Suboptimal Beamformers under Different l_p-norm Constraints

In this example, we compare the secrecy rate performance of the optimal beamformer with the generalized eigenvalue (GEV) and ZF beamformers, under the premise of perfect channel information.

Figure 3.1(a) depicts the secrecy rates (3.5) versus A/σ. These secrecy rates are obtained with $\mathbf{w}_{\alpha^\star}$, $\mathbf{w}_{\mathrm{GEV}}$, and $\mathbf{w}_{\alpha=0}$, corresponding to the optimal, GEV, and ZF beamformers, respectively, subject to the constraint $\|\mathbf{w}\|_p \leq 1$, for $p = 1, 2, \infty$. The optimal beamformer $\mathbf{w}_{\alpha^\star}$ is obtained with Proposition 3.1, and the corresponding $\alpha^\star$

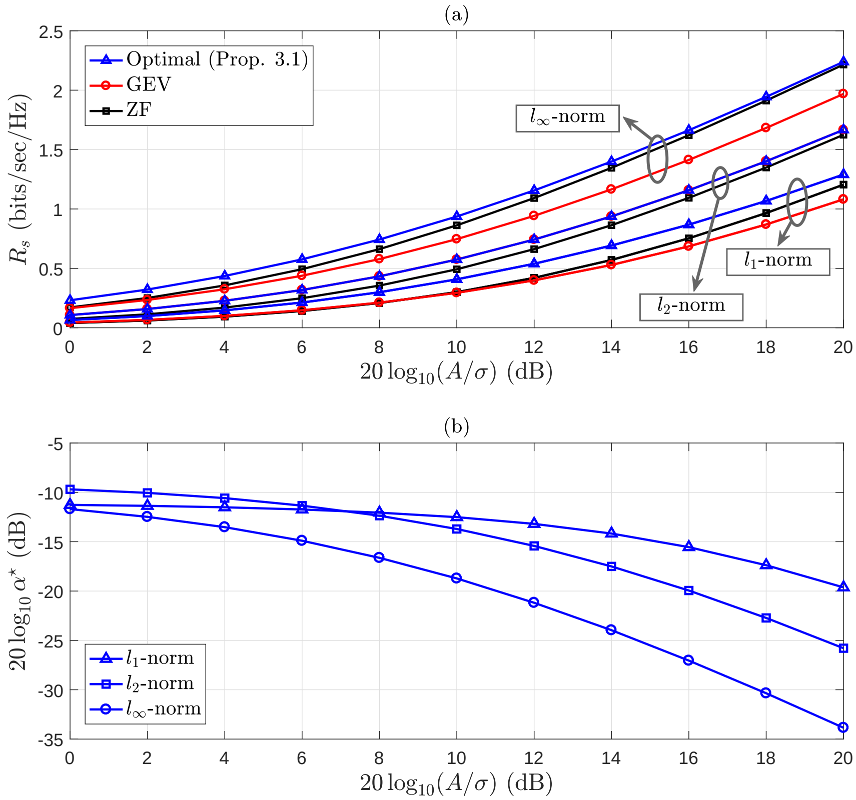

Figure 3.1: (a) Secrecy rates (3.5) obtained with the optimal, GEV, and ZF beamformers versus A/σ, subject to the constraint $\|\mathbf{w}\|_p \leq 1$, for $p = 1, 2, \infty$. (b) $\alpha^\star$ corresponding to the optimal beamformer $\mathbf{w}_{\alpha^\star}$.

is shown in Figure 3.1(b). The beamformer $\mathbf{w}_{\text{GEV}}$ is the generalized eigenvector of the matrix pair $(6A^2\mathbf{h}_\text{B}\mathbf{h}_\text{B}^\text{T} + 3\pi e\sigma^2\mathbf{I}_N, \pi e A^2\mathbf{h}_\text{E}\mathbf{h}_\text{E}^\text{T} + 3\pi e\sigma^2\mathbf{I}_N)$ corresponding to its largest generalized eigenvalue, where $\mathbf{w}_{\text{GEV}}$ is scaled such that $\|\mathbf{w}_{\text{GEV}}\|_p = 1$, for $p = 1, 2, \infty$. The ZF beamformer $\mathbf{w}_{\alpha=0}$ is obtained by solving (3.11) with $\alpha = 0$.

As expected, we note from Figure 3.1(a) that the optimal beamformer provides the best performance for all $p = 1, 2, \infty$, however at the cost of increased computational complexity. We also note that the secrecy rates of the optimal and GEV beamformers coincide when $p = 2$. This is because GEV beamforming is optimal under the l_2-norm constraint [39]. Furthermore, we note that the ZF beamformer outperforms its GEV counterpart under the l_∞-norm constraint, and it approaches the performance of the

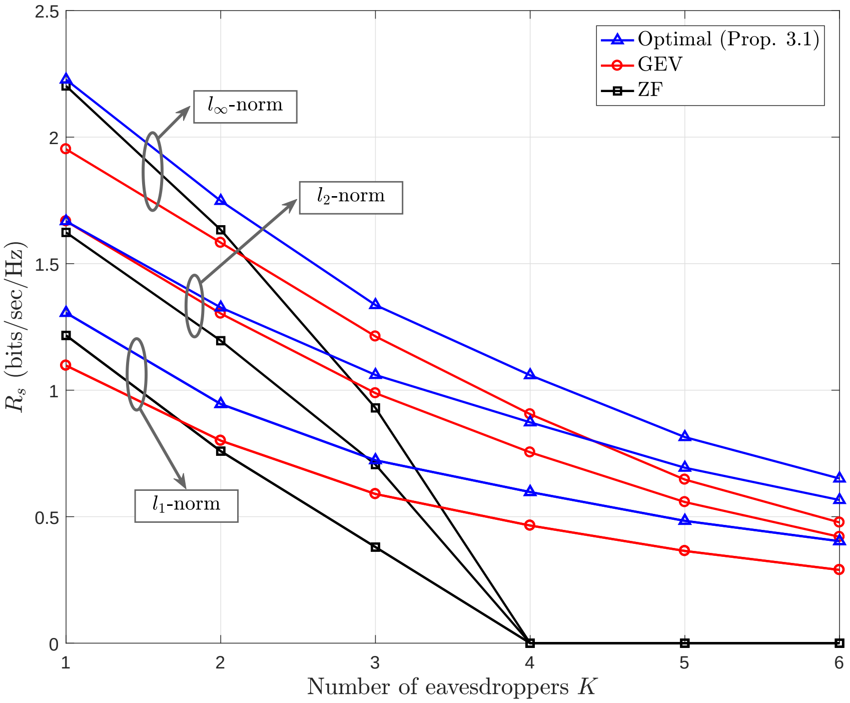

Figure 3.2: Secrecy rates (3.83) of the optimal, GEV, and ZF beamformers under different l_p-norm constraints, versus the number of eavesdroppers when $20\log_{10}(A/\sigma) = 20$ dB.

optimal beamformer as A/σ increases. Figure 3.1(b) shows that $\alpha^\star$ is nonincreasing w.r.t. A/σ for all $p = 1, 2, \infty$. Thus, the ZF beamformer is in fact asymptotically optimal at high A/σ for all p. Moreover, it can be observed that $\alpha^\star$ decreases rapidly as p increases. This reveals that the performance gap between the ZF and optimal beamformers narrows quickly with increasing p at high A/σ.

Figure 3.2 shows the secrecy rate performance versus the number of eavesdroppers K when $20\log_{10}(A/\sigma) = 20$ dB. Each eavesdropper has a single receive element, and there is no collaboration among the eavesdroppers, i.e., centralized processing of the received signals is not permitted. The secrecy rates are obtained with

$$R_s(\mathbf{w}) = \left[\frac{1}{2}\log_2 \frac{6A^2(\mathbf{h}_{\mathrm{B}}^{\mathrm{T}}\mathbf{w})^2 + 3\pi e\sigma^2}{\pi e A^2\|\mathbf{H}_{\mathrm{E}}^{\mathrm{T}}\mathbf{w}\|_\infty^2 + 3\pi e\sigma^2}\right]^+, \tag{3.83}$$

where $\mathbf{H}_{\rm E} \triangleq [\mathbf{h}_{{\rm E}_1} \ \ldots \ \mathbf{h}_{{\rm E}_K}]$. The GEV beamformer is the generalized eigenvector of the matrix pair $(6A^2\mathbf{h}_{\rm B}\mathbf{h}_{\rm B}^{\rm T} + 3\pi e\sigma^2\mathbf{I}_N, \pi e A^2\mathbf{H}_{\rm E}\mathbf{H}_{\rm E}^{\rm T} + 3\pi e\sigma^2\mathbf{I}_N)$ corresponding to the largest generalized eigenvalue. The optimal beamformer is obtained with Proposition 3.1 after replacing the constraint in (3.11b) with $\|\mathbf{H}_{\rm E}^{\rm T}\mathbf{w}\|_\infty \leq \alpha\mathbf{h}_{\rm B}^{\rm T}\mathbf{w}$, and the ZF beamformer is obtained by setting $\alpha = 0$.

We note that the GEV beamformer is optimal when $K = 1$ and $p = 2$. We also note that, as K increases, the GEV beamformer outperforms the ZF scheme even when $p \neq 2$. Obviously, ZF becomes infeasible once $K \geq N$.

3.5.1.2 Robust versus Non-Robust Schemes

In this example, we compare the worst-case secrecy rate performance of the robust beamformer with non-robust schemes. We assume that the uncertainty sets for Bob's and Eve's channels, respectively, are

$$\mathcal{H}_{\rm B} = \left\{\hat{\mathbf{h}}_{\rm B} + \mathbf{e}_{\mathbf{h}_{\rm B}} : \|\mathbf{e}_{\mathbf{h}_{\rm B}}\|_2 \leq \epsilon_{\mathbf{h}_{\rm B}}\right\}, \tag{3.84a}$$

$$\mathcal{H}_{\rm E} = \left\{\hat{\mathbf{h}}_{\rm E} + \mathbf{e}_{\mathbf{h}_{\rm E}} : \|\mathbf{e}_{\mathbf{h}_{\rm E}}\|_\infty \leq \epsilon_{\mathbf{h}_{\rm E}}\right\}. \tag{3.84b}$$

The entries of the nominal vectors $\hat{\mathbf{h}}_{\rm B}$ and $\hat{\mathbf{h}}_{\rm E}$ are generated at random, and the results are averaged over 1000 trials.

In Figure 3.3, we plot the worst-case secrecy rate

$$R_s^{\rm wc}(\mathbf{w}) = \left[\frac{1}{2}\log_2 \frac{\min\limits_{\mathbf{h}_{\rm B}\in\mathcal{H}_{\rm B}} 6A^2(\mathbf{h}_{\rm B}^{\rm T}\mathbf{w})^2 + 3\pi e\sigma^2}{\max\limits_{\mathbf{h}_{\rm E}\in\mathcal{H}_{\rm E}} \pi e A^2(\mathbf{h}_{\rm E}^{\rm T}\mathbf{w})^2 + 3\pi e\sigma^2}\right]^+ \tag{3.85}$$

versus $\epsilon_{\mathbf{h}_{\rm E}}$, for $\epsilon_{\mathbf{h}_{\rm B}} = 0, 0.2, 0.4$, and $20\log_{10}(A/\sigma) = 20$ dB. We compare the performance of the robust beamformer from Proposition 3.2 with its non-robust counterpart from Proposition 3.1, as well as the GEV and ZF beamformers. All the beamformers

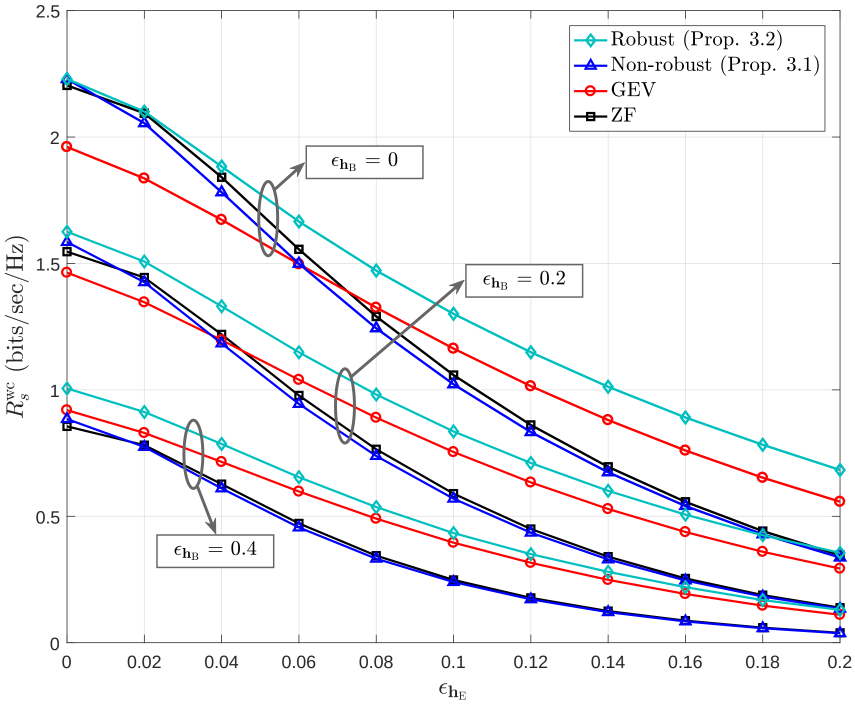

Figure 3.3: Worst-case secrecy rates (3.85) of the robust, non-robust, GEV, and ZF beamformers versus $\epsilon_{\mathbf{h}_\mathrm{E}}$ with $\epsilon_{\mathbf{h}_\mathrm{B}} = 0, 0.2, 0.4$. All the beamformers are subject to the amplitude constraint $\|\mathbf{w}\|_\infty \leq 1$, and $20\log_{10}(A/\sigma) = 20$ dB.

are subject to the amplitude constraint $\|\mathbf{w}\|_\infty \leq 1$. Substituting from (3.84a) and (3.84b) into (3.33b) and (3.33c), respectively, the inner problem (3.33) is expressed as

$$\underset{\|\mathbf{w}\|_\infty \leq 1, t}{\text{maximize}} \quad t \tag{3.86a}$$

$$\text{s.t.} \quad \hat{\mathbf{h}}_\mathrm{B}^\mathrm{T}\mathbf{w} - \epsilon_{\mathbf{h}_\mathrm{B}}\|\mathbf{w}\|_2 \geq t, \tag{3.86b}$$

$$\hat{\mathbf{h}}_\mathrm{E}^\mathrm{T}\mathbf{w} + \epsilon_{\mathbf{h}_\mathrm{E}}\|\mathbf{w}\|_1 \leq \alpha t, \tag{3.86c}$$

$$\hat{\mathbf{h}}_\mathrm{E}^\mathrm{T}\mathbf{w} - \epsilon_{\mathbf{h}_\mathrm{E}}\|\mathbf{w}\|_1 \geq -\alpha t. \tag{3.86d}$$

Then, the robust beamformer is obtained via Algorithm 3.1. On the other hand, the non-robust, GEV, and ZF beamformers are obtained using the nominal vectors $\hat{\mathbf{h}}_\mathrm{B}$

and $\hat{\mathbf{h}}_{\mathrm{E}}$.

As expected, we note from Figure 3.3 that the robust beamformer outperforms its non-robust counterparts, and the performance gain becomes more evident with increasing $\epsilon_{\mathbf{h}_{\mathrm{B}}}$ and $\epsilon_{\mathbf{h}_{\mathrm{E}}}$.

3.5.2 Worst-Case Secrecy Rate Performance in VLC Scenarios

In this subsection, we investigate the worst-case secrecy rate performance in a typical VLC scenario using the robust beamformer from Proposition 3.2 along with the uncertainty sets derived in Section 3.4.

We consider a room of size $5 \ \times 5 \ \times 3$ m^3 illuminated by 25 square-shaped light fixtures uniformly distributed over 4×4 m^2 of the ceiling area, as depicted in Figure 3.4. Each fixture occupies 10×10 cm^2 and encloses 4 high-brightness 2.5-W LEDs located at the corners of the fixture. Each LED radiates 570 mW optical power (or radiant flux). Emitted light is “warm-white” (i.e., color temperature is between 2700 and 3000 K) with luminous efficiency 284 lm/W [10, Table 3.2]. The resulting luminous flux is $0.570 \times 284 \approx 162$ lm per LED. The nominal half-intensity angle (measured from the center) is 60°, and the peak (or center) luminous intensity is 51 cd. The resulting illuminance, averaged over a horizontal 4×4 m^2 illumination grid at height 0.85 m above the floor level, is 339 Lux. For convenience, all simulation parameters are provided in Table 3.2.

All the following results are generated with Bob and Eve having PDs of area $A_{\mathrm{PD}} = 1$ cm^2 and responsivity $R_{\mathrm{PD}} = 100$ μA/mW/cm^2. The modulation index μ_{MI} is set at 10%. The noise power is assumed to be equal everywhere with $20 \log_{10} \sigma = -114$ dBm. This value is obtained with [9, Eq. (6)] using the average received DC

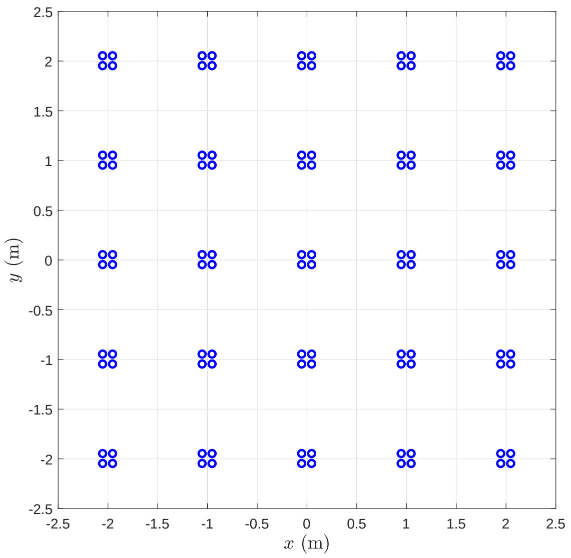

Figure 3.4: Layout of the LEDs on the ceiling. There exist 25 light fixtures. Each fixture has 10 × 10 cm^2 surface area and encloses 4 LEDs located at the corners of the fixture.

Table 3.2: Simulation parameters for the VLC scenario.

Simulation setup	
Room size	$5 \times 5 \times 3$ m^3
Number of fixtures N_{Fix}	25
Fixture size	10×10 cm^2
Number of LEDs per fixture N_{LED}	4
Total number of LEDs N	100
LED electrical and optical characteristics	
Forward voltage	3.6 V
Forward current I_{DC}	700 mA
Input electric power	2.52 W
Optical power / electric current η	813.6 μW/mA
Output optical power (or radiant flux) P_{opt}	569.52 mW
Luminous efficiency (warm-white color)	284 lm/W
Luminous flux	161.74 lm
Luminous efficacy	64.18 lm/W
Nominal half-intensity angle $\zeta^{\text{3-dB}}$	60°
Peak (or center) luminous intensity	51.48 cd
Modulation index μ_{MI}	10%
Optical receiver characteristics	
Gain of the optical filter T_s	1
Lens refractive index n_r	1.5
PD responsivity R_{PD}	100 μA/mW/cm^2
PD surface area A_{PD}	1 cm^2

optical power (averaged over the horizontal plane at height 0.85 m) with FoV $\Psi = 70^\circ$ and receiver bandwidth of 10 MHz. All location coordinates are specified in meters w.r.t. the room center at the floor level.

In all scenarios, we assume that Bob is located at $(x_\mathrm{B}, y_\mathrm{B}, z_\mathrm{B}) = (1.7173, 0.7496, 0.85)$ with orientation $(\theta_\mathrm{B}, \phi_\mathrm{B}) = (15^\circ, 240^\circ)$ and FoV $\Psi_\mathrm{B} = 70^\circ$. Furthermore, we use the spherical set in (3.46) to model uncertainty in Bob's channel, where the entries of $\hat{\mathbf{h}}_\mathrm{B}$ are obtained with (3.45), i.e., $\hat{\mathbf{h}}_\mathrm{B} = \mathbf{h}(x_\mathrm{B}, y_\mathrm{B}, z_\mathrm{B}, \theta_\mathrm{B}, \phi_\mathrm{B}, \Psi_\mathrm{B})$, and $\epsilon_{\mathbf{h}_\mathrm{B}} = 0.1\|\hat{\mathbf{h}}_\mathrm{B}\|_2$. The nominal estimate $\hat{\mathbf{h}}_\mathrm{B}$ is fixed and assumed to be known to Alice via feedback from Bob. Parameters relevant to Eve are provided in the caption of each figure. In all cases, for the sake of illustration, we plot the worst-case secrecy rate versus α using (3.44), where $20 \log_{10} \alpha = -50, -49, \ldots, 0$ dB. We also include the case of certain Eve's channel for comparison purposes. For each α, we use the CVX toolbox [85], in conjunction with the MOSEK solver [86], to solve (3.33) using the relevant uncertainty set $\mathcal{H}_\mathrm{E}$ from Section 3.4.

3.5.2.1 Uncertain Eavesdropper's Location

Figure 3.5 shows the worst-case secrecy rate performance with uncertain Eve's location. We include three groups of curves corresponding to three uncertainty regions, $\mathcal{B}$, of different volumes. All the regions are rectangular and centered at $(x, y, z) = (-1.25, 0, 0.85)$. Four curves are generated for each $\mathcal{B}$ corresponding to the combinations of two methods to approximate $\mathcal{H}_\mathrm{E}^{\mathcal{B}}$ and two methods to modulate the LEDs. We refer to the case in which the affine approximation (3.53) is used as "Linearized", and to the case in which $\mathcal{B}$ is discretized as "Discretized". For the "Linearized" case, $\mathcal{B}$ is divided into identical boxes, each of volume $2l_x \times 2l_y \times 2l_z = 0.5 \times 0.5 \times 0.25$ m^3, then (3.53) is applied to each box and $\mathbf{w}_\alpha$ is obtained with (3.57). For the "Discretized" case, $\mathcal{H}_\mathrm{E}^{\mathcal{B}}$ is approximated by sampling the channel gain in the

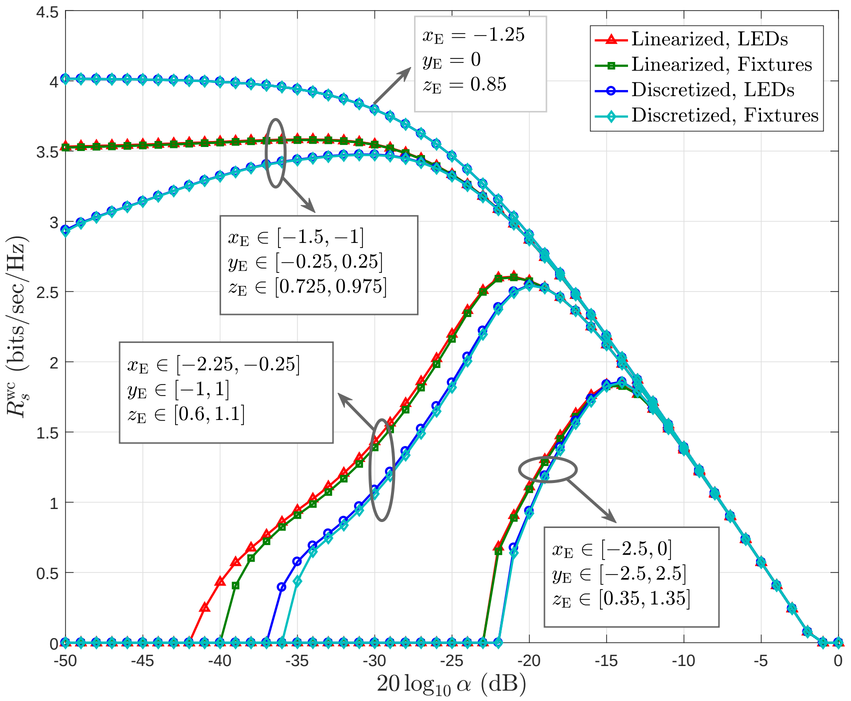

Figure 3.5: Worst-case secrecy rate (3.44) versus α with uncertain Eve's location. $\theta_{\mathrm{E}} = 0$ and $\Psi_{\mathrm{E}} = 70°$.

three-dimensional space using a $10 \times 10 \times 10$ cm^3 grid, and $\mathbf{w}_\alpha$ is obtained with (3.59). Furthermore, we refer to the case in which each LED is modulated independently as "LEDs", and to the case in which all LEDs in one fixture are modulated with the same current signal as "Fixtures".

As expected, we note from Figure 3.5 that R_s^{wc} decreases as the uncertainty about Eve's location increases. For the case of certain Eve's location, we can see that the ZF beamformer is practically optimal. In addition, Figure 3.5 reveals that independent-LED modulation does not provide much improvement, if any, compared to the more practical and less expensive "Fixture" modulation scheme. This is also expected since LEDs in the same fixture are relatively close to each other and have almost identical channel gains. Figure 3.5 also validates the affine approximation in (3.53) when l_x, l_y, and l_z are chosen properly.

3.5.2.2 Uncertain Eavesdropper's Orientation

In Figure 3.6, we plot the worst-case secrecy rate performance with uncertain Eve's orientation. We also investigate the impact of Eve's FoV on the secrecy rate. The curve "Small $\theta_{\max}$" is generated with $\mathbf{w}_{\alpha}$ obtained from (3.67), whereas all other curves are obtained with (3.68) after discretizing the intervals $\Theta = [0, \theta_{\max}]$ and $\Phi = [0, 360°]$ using $\Delta_{\theta} = \Delta_{\phi} = 4°$.

For the case $\theta_{\max} = 0$, i.e., no uncertainty about Eve's orientation, we can see that ZF is essentially optimal. We also note that the curve "Small $\theta_{\max}$" almost coincides with the curve corresponding to $\Psi_{\mathrm{E}} = 90°$ and $\theta_{\max} = 30°$. Thus, the linear channel gain model that leads to the problem in (3.67) is indeed valid for small angle variations and wide FoV. Figure 3.6 also reveals that reducing Eve's FoV has a negative impact on the secrecy rate performance, which can be explained as follows. First, we recall from (1.5) that reducing the FoV of the concentrator increases its gain (inside the FoV). Second, the limited FoV of Eve's receiver, in conjunction with her ability to adjust orientation, increases the space of her received signal as measured by the number of nonzero singular values of the matrix $\mathbf{H}_{\mathrm{E}}^{\ddot{\mathcal{U}}}$ whose columns are the elements of the discretized uncertainty set $\mathcal{H}_{\mathrm{E}}^{\ddot{\mathcal{U}}}$. Obviously, increasing the signal space available to Eve makes it more difficult for Alice to suppress Eve's signal, i.e., more of the degrees of freedom available to Alice are exploited, and thus the secrecy rate is reduced.

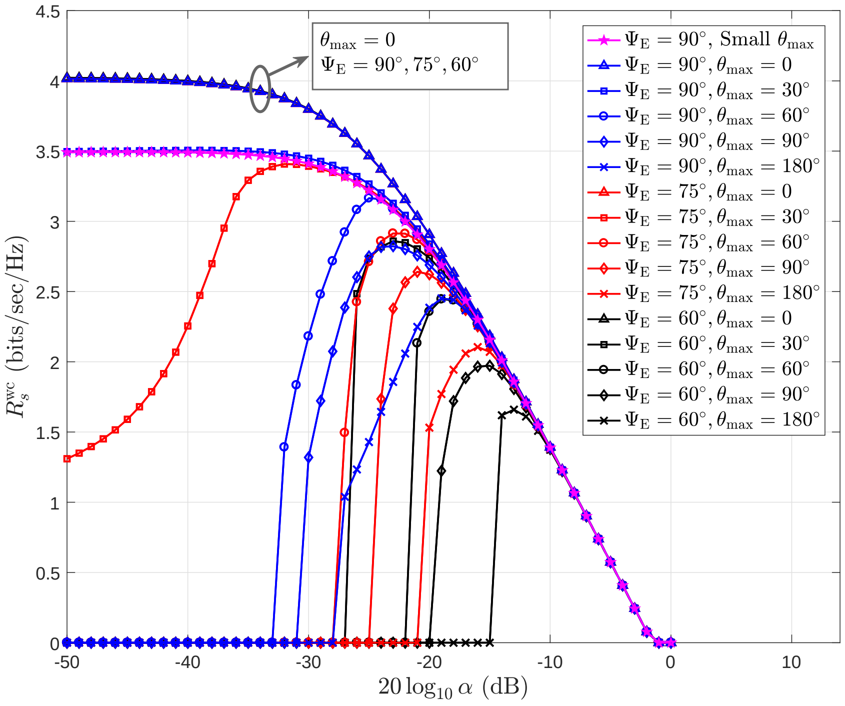

Figure 3.6: Worst-case secrecy rate (3.44) versus α with uncertain Eve's orientation. $x_{\mathrm{E}} = -1.25$, $y_{\mathrm{E}} = 0$, $z_{\mathrm{E}} = 0.85$, $\theta_{\mathrm{E}} \in [0, \theta_{\max}]$, and $\phi_{\mathrm{E}} \in [0, 360°]$.

3.5.2.3 Uncertain Eavesdropper's Location and LEDs Half-Intensity Angle

Figure 3.7 depicts the secrecy performance with uncertain Eve's location and LEDs half-intensity angle. We consider half-intensity angle uncertainties up to ±20° around the nominal value of 60°, and the location uncertainty region

$$\{(x_{\mathrm{E}}, y_{\mathrm{E}}) : x_{\mathrm{E}} \in [-2.25, -0.25], y_{\mathrm{E}} \in [-2.5, 2.5]\}$$

is discretized using $\Delta_x = \Delta_y = 20$ cm. As can be seen, even relatively small half-intensity angle deviations, e.g., ±5°, can significantly reduce the worst-case secrecy rate.

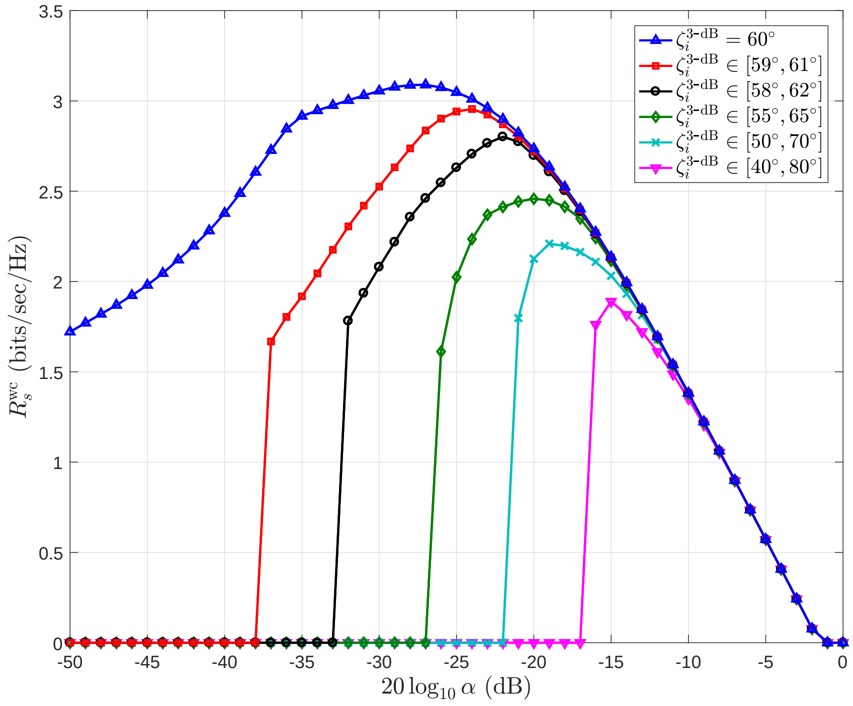

Figure 3.7: Worst-case secrecy rate (3.44) versus α with uncertain Eve's location and LEDs half-intensity angle $\zeta_i^{3\text{-dB}}$, $i = 1, \ldots, N$. $x_{\mathrm{E}} \in [-2.25, -0.25]$, $y_{\mathrm{E}} \in [-2.5, 2.5]$, $z_{\mathrm{E}} = 0.85$, $\theta_{\mathrm{E}} = 0$, and $\Psi_{\mathrm{E}} = 70^\circ$.

3.5.2.4 Uncertain Eavesdropper's Location and NLoS Components

In Figure 3.8, we show the worst-case secrecy rate performance with uncertain Eve's location and NLoS components. Similar to the results in Figure 3.7, the location uncertainty region is discretized using $\Delta_x = \Delta_y = 20$ cm. We investigate the performance when the strength of the NLoS components can go up to $\gamma_{\max} = 60\%$ of the LoS path. Note that $\gamma_{\max} = 60\%$, or even 40%, is a too pessimistic or too conservative assumption. In typical scenarios with only diffuse reflections, i.e., no large windows or mirrors, $\gamma_{\max}$ will probably be less than 20% (see, e.g., Figure 4 in [87] or the discussion after Figure 6 in [59]).

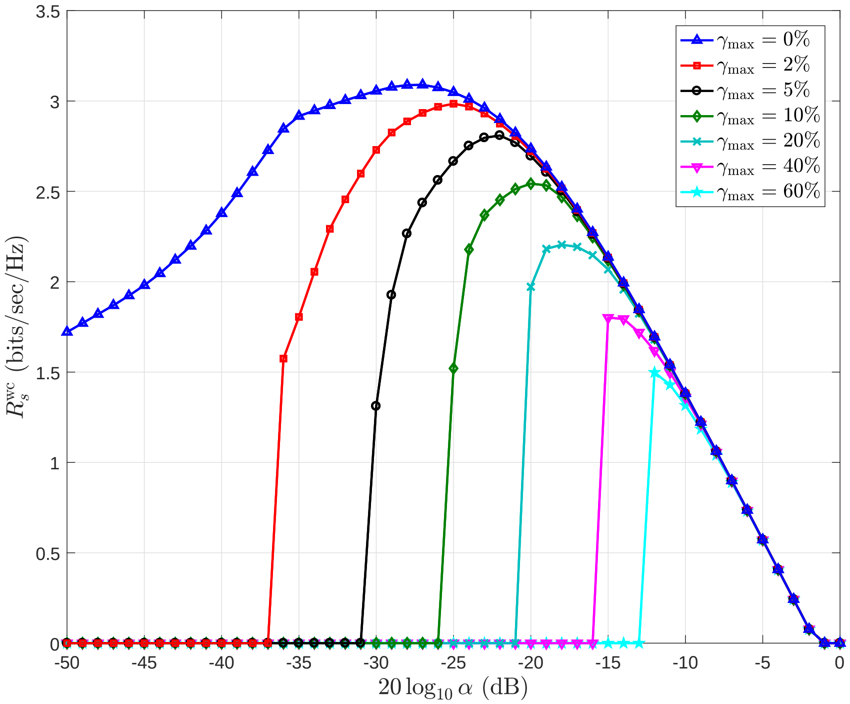

Figure 3.8: Worst-case secrecy rate (3.44) versus α with uncertain Eve's location and NLoS components. $x_{\mathrm{E}} \in [-2.25, -0.25]$, $y_{\mathrm{E}} \in [-2.5, 2.5]$, $z_{\mathrm{E}} = 0.85$, $\theta_{\mathrm{E}} = 0$, and $\Psi_{\mathrm{E}} = 70°$.

3.6 Conclusions

In this chapter, we studied the design of transmit beamformers for secrecy rate maximization in MISO wiretap channels subject to amplitude constraints. Such constraints are typically difficult to handle and, because of that, they are often overlooked in favor of the more convenient total power constraint.

We tackled the nonconvex secrecy rate maximization problem by transforming it into an equivalent quasiconvex line search problem. Our approach is conceptually simple but provably optimal for general l_p-norm constraints, and the equivalent problem itself is practically meaningful. Furthermore, our approach proved helpful in tackling the more complex robust design problem with uncertain channel information.

We used the VLC scenario as a practical example in which reasonable estimates of the eavesdropper's channel can be obtained without feedback from the (passive) eavesdropper. Numerical results show that the excess degrees of freedom provided by the large number of LEDs in typical VLC transmitters can be effectively utilized to compensate for the lack of accurate information regarding the eavesdropper's channel.

In the next chapter, we consider the more general two-user MISO BC-CM model.

Chapter 4

Linear Precoding for the Two-User MISO Broadcast Channel with Confidential Messages

4.1 Introduction

In the previous chapter, we studied the design of transmit beamformers for the MISO wiretap channel under the amplitude constraint. In such a scenario, the transmitter had one secret message for the intended receiver (Bob), while the other receiver (Eve) acted only as an eavesdropper. In this chapter, we consider the more general scenario in which the transmitter has two independent secret messages, one intended for each user, and each message should be kept confidential from the other user. Such a model is referred to as the two-user broadcast channel with confidential messages (BC-CM). Note that the two-user BC-CM reduces to the wiretap channel when the information rate to one of the users is set to zero.

Extension of the wiretap channel to the two-user BC-CM was considered in [47] wherein the authors derived inner and outer bounds on the secrecy capacity region of the discrete memoryless BC-CM. Achievability of the secrecy capacity region of the two-user MISO BC-CM was established in [48] using the so-called *secret dirty-*

paper coding (S-DPC) scheme under the total (average) power constraint. This coding scheme was extended in [49] to the MIMO BC-CM, and it was shown that the secrecy capacity region is rectangular under the matrix power (or input covariance matrix) constraint. Under the total power constraint, however, the secrecy capacity region can be only found by performing an exhaustive search over the set of all input covariance matrices that satisfy the total power constraint. Due to the complexity of S-DPC and the lack of a simple solution to the practical case of total power constraint, the authors in [88] proposed a low-complexity linear precoding scheme for the two-user MIMO BC-CM based on generalized singular value decomposition. The work in [89] also considered the secrecy rate region of the two-user MIMO BC-CM under the total power constraint via formulating a nonconvex weighted secrecy sum rate maximization problem. An iterative algorithm based on a *block successive lower-bound maximization method* was proposed to solve such a nonconvex problem.

In this chapter, we study the design of linear precoders for the two-user MISO BC-CM. Our treatment will not be limited to VLC channels in the sense that we design the precoders not only subject to amplitude constraints, but also subject to total and per-antenna power constraints. Note that, in this chapter, we use the term "*antenna*" to designate general transmit and receive elements. In a VLC system, the transmit antenna is an LED and the receive antenna is a PD.

Under amplitude constraints, the secrecy capacity region of the two-user MISO BC-CM is unknown, and thus our motivation to find achievable secrecy rate regions based on linear precoding is clear. However, we also consider linear precoding for the cases of total and per-antenna power constraints, wherein S-DPC is known to be optimal, for the following reasons:

1) Our approach to formulate and solve the problem of linear precoder design

is equally applicable to all the aforementioned constraints. In other words, total and per-antenna power constraints can be considered without additional difficulty.

2) More importantly, the case of total power constraint is the only case in which the secrecy capacity region is precisely known and has been characterized in closed form. Therefore, it sets a benchmark that can be used to quantify the performance loss incurred by using a suboptimal linear precoding scheme, and also to validate our approach to solve the precoder design problem when the total power constraint is considered.

3) Finally, for the case of per-antenna power constraint, it seems that the secrecy capacity region can be only found via an exhaustive search over the set of all input covariance matrices that satisfy the per-antenna power constraint, which entails high computational complexity. Even when it is found, the S-DPC scheme utilized to achieve the secrecy capacity region is difficult to implement. Therefore, our proposed linear precoding scheme provides a viable solution with lower implementation complexity for the case of per-antenna power constraint.

After fixing the input distribution, our goal in this chapter is to find linear precoders that achieve the boundary points of the secrecy rate region. To this end, we formulate the precoder design problem as a weighted secrecy sum rate maximization problem, subject to any of the aforementioned constraints. The resulting problem, however, has a fractional objective function, making it nonconvex and difficult to solve. To circumvent such a difficulty, we first simplify the objective function using a lower bound on the weighted secrecy sum rate. Then, we transform the problem into an equivalent one having only two optimization variables. We show that the equivalent problem is more tractable and can be solved iteratively using the subgradient method. In each iteration, we solve a convex *inner* optimization problem to update the value of the *outer* problem, and also to obtain the subgradient vector that specifies the search direction for the next iteration. We provide a condition under which the obtained solution is guaranteed to be globally optimal. Furthermore, we show that the inner problem can be easily modified to take into account channel uncertainty caused by limited feedback from both receivers. This leads us to the design of *robust linear precoder*s that guarantee a certain *worst-case secrecy rate* performance.

To the best of our knowledge, the work in this chapter is the first to consider linear precoding for the two-user MISO BC-CM, subject to per-antenna power or amplitude constraints. Furthermore, it is the first work to consider robust precoding, for the same channel, by taking channel uncertainty into account.

The remainder of this chapter is organized as follows. The system model, precoding scheme, and transmit constraints are described in detail in Section 4.2. In Section 4.3, we solve the precoder design problem under the premise of perfect channel information. In Section 4.4, we extend the design problem to its robust counterpart by considering uncertainty in channel information. In Section 4.5, we provide our numerical examples to illustrate the achievable secrecy rate regions of the proposed precoder. We conclude the chapter in Section 4.6.

4.2 System Model

In this section, we describe the channel model, the linear precoding scheme, the achievable secrecy rate regions, and the constraints imposed on the transmitted signal vector.

4.2.1 The Two-User MISO BC-CM

We study the problem of secret communication between one transmitter and two independent receivers over the Gaussian MISO broadcast channel. The transmitter has $N \geq 2$ antennas, and each receiver has a single antenna. In each communication session, the transmitter has two independent confidential messages, one intended for each receiver. The two messages are simultaneously broadcasted, and the transmitter shall ensure that each message can be reliably decoded by its intended receiver, and is kept confidential from the other one.

We assume narrow-band transmission over a quasi-static, i.e., non-fading, Gaussian broadcast channel. The transmitted and received baseband signals, as well as the channel gain vectors, are real-valued, i.e., the carrier phase is not modulated. This model is typical for intensity-modulation (IM) channels, including VLC channels, and is also applicable to RF systems utilizing amplitude modulation schemes, such as amplitude-shift keying (where the baseband data symbols are real-valued). Under these assumptions, the signals observed by the two receivers can be expressed as

$$y_1(t) = \mathbf{h}_1^{\mathrm{T}}\mathbf{x}(t) + n_1(t), \tag{4.1a}$$

$$y_2(t) = \mathbf{h}_2^{\mathrm{T}}\mathbf{x}(t) + n_2(t), \tag{4.1b}$$

where $\mathbf{x}(t) \in \mathbb{R}^N$ is the transmitted signal vector, $\mathbf{h}_1 \in \mathbb{R}^N$ and $\mathbf{h}_2 \in \mathbb{R}^N$ are the channel gain vectors, and $n_1(t)$ and $n_2(t)$ are i.i.d. Gaussian noise samples with variance σ^2. We assume that $\mathbf{h}_1$ and $\mathbf{h}_2$ are linearly independent to ensure that the MISO broadcast channel in (4.1) is nondegraded.

Let $\boldsymbol{X}$ be an input random vector that satisfies the constraints on the channel input, and Y_1 and Y_2 be the output random variables. Also let $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ be auxiliary random variables. Then, it was shown in [47, Theorem 4] (see also [48, Lemma 2]) that for any joint PDF $p(\mathbf{u}_1, \mathbf{u}_2, \mathbf{x}, y_1, y_2)$ that can be written as[13]

$$p(\mathbf{u}_1, \mathbf{u}_2)\, p(\mathbf{x}|\mathbf{u}_1, \mathbf{u}_2)\, p(y_1, y_2|\mathbf{x}),$$

the secrecy rate pair (R_1, R_2) satisfying

$$0 \leq R_1 \leq \mathbb{I}(\boldsymbol{U}_1; Y_1) - \mathbb{I}(\boldsymbol{U}_1; Y_2|\boldsymbol{U}_2) - \mathbb{I}(\boldsymbol{U}_1; \boldsymbol{U}_2), \tag{4.2a}$$

$$0 \leq R_2 \leq \mathbb{I}(\boldsymbol{U}_2; Y_2) - \mathbb{I}(\boldsymbol{U}_2; Y_1|\boldsymbol{U}_1) - \mathbb{I}(\boldsymbol{U}_1; \boldsymbol{U}_2) \tag{4.2b}$$

is achievable for the two-user MISO BC-CM in (4.1). Achievability of the rate pair in (4.2) was proved based on a *double-binning* scheme [48, Section IV]. Thus, given a joint PDF $p(\mathbf{u}_1, \mathbf{u}_2, \mathbf{x})$, the achievable secrecy rate region can be determined using (4.2). On the other hand, given a certain constraint on the channel input $\boldsymbol{X}$, it remains unclear how to choose $p(\mathbf{u}_1, \mathbf{u}_2, \mathbf{x})$ such that the secrecy rate region is maximized. For the case of total power constraint, it was shown that the secrecy capacity region of the MISO BC-CM in (4.1) can be characterized in closed form [48, Theorem 1], and the boundary points are achievable with the S-DPC scheme. This scheme, however, is difficult to implement in practice [88]. In addition, with per-antenna power constraints, there is no closed-form characterization, and apparently the secrecy capacity region can be only found via an exhaustive search over all input covariance matrices that satisfy the per-antenna power constraint. Furthermore, the S-DPC scheme proposed in [48] does not seem to be applicable to the case with amplitude constraints. This motivates us to consider the linear precoding scheme

[13] In other words, $(\boldsymbol{U}_1, \boldsymbol{U}_2) \rightarrow \boldsymbol{X} \rightarrow (Y_1, Y_2)$ forms a Markov chain.

described in the next subsection.

4.2.2 Linear Precoding

We study the secrecy performance of the two-user MISO BC-CM in (4.1) when the transmitted signal vector is constructed as

$$\mathbf{x}(t) = \mathbf{w}_1 s_1(t) + \mathbf{w}_2 s_2(t) = \mathbf{W}\mathbf{s}(t), \tag{4.3}$$

where $\mathbf{w}_1 \in \mathbb{R}^N$ and $\mathbf{w}_2 \in \mathbb{R}^N$ are fixed beamformers, $s_1(t) \in \mathbb{R}$ and $s_2(t) \in \mathbb{R}$ are independent symbols (i.e., secrecy codewords) intended for Users 1 and 2, respectively, $\mathbf{W} = [\mathbf{w}_1 \ \mathbf{w}_2]$ is termed as the *precoding matrix*, or simply the *precoder*, and $\mathbf{s}(t) = [s_1(t) \ \ s_2(t)]^{\mathrm{T}}$ is the vector of transmitted symbols. Although suboptimal, the precoding scheme in (4.3) is simple to implement. Furthermore, it will enable us to handle per-antenna power or amplitude constraints.

Substituting (4.3) back into (4.1), the signals received at both users can be written as

$$y_1(t) = \mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1 s_1(t) + \mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2 s_2(t) + n_1(t), \tag{4.4a}$$

$$y_2(t) = \mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1 s_1(t) + \mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2 s_2(t) + n_2(t). \tag{4.4b}$$

Let S_1 and S_2 denote the random variable counterparts of $s_1(t)$ and $s_2(t)$, respectively. Then, the transmission scheme in (4.3) corresponds to choosing

$$\boldsymbol{U}_1 = \mathbf{w}_1 S_1, \quad \boldsymbol{U}_2 = \mathbf{w}_2 S_2, \quad \text{and} \ \ \boldsymbol{X} = \boldsymbol{U}_1 + \boldsymbol{U}_2. \tag{4.5}$$

Substituting from (4.5) into (4.2), the achievable secrecy rate pair in (4.2) can be written as

$$0 \leq R_1 \leq \mathbb{I}(S_1; Y_1) - \mathbb{I}(S_1; Y_2|S_2), \tag{4.6a}$$

$$0 \leq R_2 \leq \mathbb{I}(S_2; Y_2) - \mathbb{I}(S_2; Y_1|S_1). \tag{4.6b}$$

Note that joint encoding is not utilized in (4.5), i.e., S_1 and S_2 are independent, and thus $\mathbb{I}(\boldsymbol{U}_1; \boldsymbol{U}_2) = \mathbb{I}(S_1; S_2) = 0$.

4.2.3 Transmit Constraints and Secrecy Rate Regions

In this subsection, we describe the transmit constraints considered in the chapter, and derive closed-form expressions for the corresponding secrecy rate pairs (R_1, R_2).

4.2.3.1 Total Average Power Constraint

The most common constraint imposed on the input of Gaussian channels is the total (or sum) average power constraint. It is mathematically convenient, and often leads to closed-form solutions. Furthermore, it provides much insight into the performance of the communication system for a given power budget. Mathematically, a total average power constraint P_{Tot} requires the transmitted codewords $\boldsymbol{X}$ to satisfy the inequality

$$\mathrm{Tr}(\mathbb{E}\{\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\}) \leq P_{\mathrm{Tot}}, \tag{4.7}$$

where $\mathbb{E}\{\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\}$ is the transmit covariance matrix. An obvious way to comply with the transmission scheme in (4.3) and satisfy the constraint in (4.7) is to choose S_1 and S_2 to be i.i.d. standard Gaussian random variables, that is

$$S_1 \sim \mathcal{N}(0, 1), \quad S_2 \sim \mathcal{N}(0, 1), \tag{4.8a}$$

and to ensure that the precoder $\mathbf{W}$ satisfies the inequality

$$\|\mathbf{W}\|_{\mathrm{F}}^2 \le P_{\mathrm{Tot}}. \tag{4.8b}$$

Note that our choice of equal variances for the distributions of S_1 and S_2 (both have unity variance) involves no loss of generality because the power allocated to each user can still be adjusted from the entries of the precoding matrix $\mathbf{W}$.

Now, for a given $\mathbf{W}$, and with Gaussian codewords $S_1, S_2 \sim \mathcal{N}(0,1)$, the mutual information terms in (4.6a) are simply calculated as

$$\mathbb{I}(S_1;Y_1) = \frac{1}{2}\log_2\left(1 + \frac{(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^2}{(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2}\right), \tag{4.9a}$$

$$\mathbb{I}(S_1;Y_2|S_2) = \frac{1}{2}\log_2\left(1 + \frac{(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2}{\sigma^2}\right), \tag{4.9b}$$

where information is measured in (bits/sec/Hz). Similar expressions can be obtained for the corresponding terms in (4.6b), and thus we end up with the achievable secrecy rate pair

$$R_1 = \frac{1}{2}\left[\log_2\left(1 + \frac{(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^2}{(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2}\right)\left(\frac{\sigma^2}{(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2}\right)\right]^+, \tag{4.10a}$$

$$R_2 = \frac{1}{2}\left[\log_2\left(1 + \frac{(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2)^2}{(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2}\right)\left(\frac{\sigma^2}{(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2}\right)\right]^+. \tag{4.10b}$$

4.2.3.2 Per-Antenna Average Power Constraint

Despite its simplicity, the total average power constraint (4.7) is often not sufficient to capture practical limitations that arise from implementation constraints. For example, the so-called *digital beamforming*[14] scheme requires a dedicated transmit RF chain for each antenna element[15]. Clearly, each of these chains has its own power budget. Thus, a more realistic approach to model power limitations at the transmitter is to impose an individual power constraint on each RF chain, or, equivalently, on each antenna element, in addition to the total power constraint. A per-antenna average power constraint $P_i, i = 1, \ldots, N$, can be expressed as

$$\mathbb{E}\{X_i^2\} \leq P_i, \quad i = 1, \ldots, N, \tag{4.11}$$

where X_i is the ith entry of $\boldsymbol{X}$. Depending on the values of P_{Tot} and $P_1, \ldots, P_N$, one of the constraints in (4.7) and (4.11) may become redundant. In particular:

i) If $\sum_{i=1}^{N} P_i \leq P_{\text{Tot}}$, the per-antenna power constraint (4.11) becomes dominant and (4.7) can be ignored.

ii) If $P_i \geq P_{\text{Tot}}$ for all $i \in \{1, \ldots, N\}$, then (4.11) is obviously redundant and the total power constraint (4.7) is sufficient.

iii) If neither of the above two cases holds, both (4.7) and (4.11) can be active, and thus they should be taken into account.

Similar to the case of dominant total average power constraint, we let the codewords S_1 and S_2 be i.i.d. standard Gaussian random variables. Then, in order to satisfy the constraint in (4.11), the entires of $\mathbf{W}$ should be chosen such that

$$w_{1i}^2 + w_{2i}^2 \leq P_i, \quad i = 1, \ldots, N, \tag{4.12}$$

[14] In fact, all the transmission schemes considered in this thesis fall into the category of digital (or baseband) beamforming.

[15] Such a constraint is relaxed in the so-called *hybrid beamforming* scheme where the number of RF chains can be smaller than the number of antennas.

where w_{1i} and w_{2i} are the ith entries of $\mathbf{w}_1$ and $\mathbf{w}_2$, respectively. Since S_1 and S_2 are Gaussian, the secrecy rate pair expressions in (4.10) remain valid for any $\mathbf{W}$ that satisfies (4.12).

4.2.3.3 Per-Antenna Amplitude Constraint

By now, we already know that amplitude constraints typically arise in the design of IM systems. In such a case, the input signal must satisfy the amplitude constraint

$$|X_i| \leq A_i, \quad i = 1, \ldots, N. \tag{4.13}$$

This constraint obviously renders the Gaussian distribution infeasible for the channel input. Nevertheless, (4.13) can be fulfilled by choosing the codewords S_1 and S_2 according to the uniform distribution over the interval $[-1, 1]$, i.e.,

$$S_1 \sim \mathcal{U}[-1, 1], \quad S_2 \sim \mathcal{U}[-1, 1], \tag{4.14a}$$

and choosing the entries of the precoder $\mathbf{W}$ such that they satisfy the constraint

$$|w_{1i}| + |w_{2i}| \leq A_i, \quad i = 1, \ldots, N. \tag{4.14b}$$

Unlike the Gaussian input distribution in (4.8a), the uniform distribution in (4.14a), along with Gaussian noise, do not immediately lead to closed-form expressions for $\mathbb{I}(S_1; Y_1) - \mathbb{I}(S_1; Y_2 | S_2)$ in (4.6a), or the similar terms in (4.6b). Nevertheless, we can lower-bound these terms to obtain closed-form expressions for the secrecy rate pair (R_1, R_2), as follows.

First, we rewrite $\mathbb{I}(S_1; Y_1) - \mathbb{I}(S_1; Y_2|S_2)$ as

$$\mathbb{h}(Y_1) - \mathbb{h}(Y_1|S_1) - \mathbb{h}(Y_2|S_2) + \mathbb{h}(Y_2|S_1, S_2). \tag{4.15}$$

Using the entropy power inequality [35, Theorem 17.7.3], the differential entropy $\mathbb{h}(Y_1)$ can be lower-bounded as

$$\begin{aligned}\mathbb{h}(Y_1) &= \mathbb{h}(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1 S_1 + \mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2 S_2 + N_1)\\ &\geq \frac{1}{2}\log_2\left(2^{2\mathbb{h}(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1 S_1)} + 2^{2\mathbb{h}(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2 S_2)} + 2^{2\mathbb{h}(N_1)}\right)\\ &= \frac{1}{2}\log_2\left(4(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^2 + 4(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + 2\pi e\sigma^2\right).\end{aligned} \tag{4.16}$$

On the other hand, the conditional differential entropy $\mathbb{h}(Y_1|S_1)$ can be upper bounded by the differential entropy of a Gaussian random variable having equal variance, that is

$$\begin{aligned}\mathbb{h}(Y_1|S_1) &= \mathbb{h}(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2 S_2 + N_1)\\ &\leq \frac{1}{2}\log_2\left(2\pi e\left(\tfrac{1}{3}(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2\right)\right).\end{aligned} \tag{4.17}$$

Similarly, we have

$$\mathbb{h}(Y_2|S_2) \leq \frac{1}{2}\log_2\left(2\pi e\left(\tfrac{1}{3}(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2\right)\right). \tag{4.18}$$

We also have

$$\mathbb{h}(Y_2|S_1, S_2) = \mathbb{h}(N_2) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right). \tag{4.19}$$

Substituting (4.16)-(4.19) back into (4.15) yields the secrecy rate expression

$$R_1 = \left[\frac{1}{2}\log_2 \frac{\left(4(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^2 + 4(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + 2\pi e\sigma^2\right)\sigma^2}{2\pi e\left(\frac{1}{3}(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2\right)\left(\frac{1}{3}(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2\right)}\right]^+. \tag{4.20a}$$

Similarly, we have

$$R_2 = \left[\frac{1}{2}\log_2 \frac{\left(4(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2)^2 + 4(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + 2\pi e\sigma^2\right)\sigma^2}{2\pi e\left(\frac{1}{3}(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2\right)\left(\frac{1}{3}(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2\right)}\right]^+. \tag{4.20b}$$

4.3 Precoder Design with Perfect Channel Information

In this section, we focus on the design of the precoder $\mathbf{W}$ under the assumption of perfect channel information. We begin with the case of total and per-antenna average power constraints. Then, we show in Section 4.3.5 that the problem formulation and solution method can be easily modified to handle per-antenna amplitude constraints.

4.3.1 Problem Formulation

By designing $\mathbf{W}$ we mean finding the set of precoding matrices that achieve the boundary of the secrecy rate region characterized by (R_1, R_2). Assuming total and per-antenna average power constraints, the design problem can be expressed by the two-objective optimization problem

$$\begin{aligned} \underset{\mathbf{W}}{\text{maximize}} \quad & (R_1, R_2) && \text{(4.21a)} \\ \text{s.t.} \quad & \|\mathbf{W}\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tot}}, && \text{(4.21b)} \\ & w_{1i}^2 + w_{2i}^2 \leq P_i, \quad i = 1, \ldots, N, && \text{(4.21c)} \end{aligned}$$

where partial ordering and maximization of the pair (R_1, R_2) are w.r.t. the nonnegative orthant $\mathbb{R}_+^2$ [80, Section 4.7.5]. In the context of multi-objective optimization, a feasible matrix $\mathbf{W}$ that achieves a secrecy rate pair on the boundary of the set of all achievable rate pairs is referred to as *Pareto optimal*, and the corresponding secrecy rate pair (R_1, R_2) is a *Pareto optimal pair*. Thus, solving (4.21) means finding Pareto optimal matrices $\mathbf{W}$.

The standard approach towards solving (4.21) is to scalarize the objective via a weighted sum [80, Section 4.7.5], that is to replace (R_1, R_2) with $\rho_1 R_1 + \rho_2 R_2$, where the weights $\rho_1 \geq 0$ and $\rho_2 \geq 0$ are free parameters. Different Pareto optimal points can be obtained by adjusting the relative weight ρ_1/ρ_2 to different values between 0 and ∞. This can be carried out by choosing[16] $\rho_1 = \rho$ and $\rho_2 = 1 - \rho$, where ρ is a free parameter taking values in the interval $[0, 1]$. Thus, for any $\rho \in [0, 1]$, we have the weighted secrecy sum rate maximization problem

$$\begin{aligned} \underset{\mathbf{W}}{\text{maximize}} \quad & R_{\mathrm{wsum}}(\rho) && \text{(4.22a)} \\ \text{s.t.} \quad & \|\mathbf{W}\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tot}}, && \text{(4.22b)} \\ & w_{1i}^2 + w_{2i}^2 \leq P_i, \quad i = 1, \ldots, N, && \text{(4.22c)} \end{aligned}$$

[16] Although constraining ρ_1 and ρ_2 to sum to 1 looks arbitrary here, we will need this restriction in the proof of Proposition 4.2, particularly to ensure that the optimization problem in (C.6) is convex.

where

$$R_{\mathrm{wsum}}(\rho) \triangleq \rho R_1 + (1-\rho) R_2 \tag{4.22d}$$

is the weighted secrecy sum rate. It is clear that solving (4.22) with $\rho = 1$ corresponds to finding the maximum achievable secrecy rate for User 1 when User 2 is treated as an eavesdropper, while $\rho = 0$ yields the maximum achievable secrecy rate for User 2.

Ideally we would like to solve (4.22) with the objective R_{wsum} calculated using the rate expressions in (4.10). Using these expressions, however, would make (4.22) very difficult to solve, except for the special cases $\rho = 0$ and $\rho = 1$. In order to make the problem tractable, we will simplify the objective of (4.22) by replacing R_{wsum} with the lower bound $\hat{R}_{\mathrm{wsum}}$, given by

$$\hat{R}_{\mathrm{wsum}}(\rho) = \rho \hat{R}_1 + (1-\rho)\hat{R}_2, \tag{4.23}$$

where, for $\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1 \neq 0$ and $\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2 \neq 0$, $\hat{R}_1$ and $\hat{R}_2$, respectively, are given by

$$\hat{R}_1 = \log_2 \frac{\left|\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1\right| \sigma}{((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2)^{\frac{1}{2}}((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2)^{\frac{1}{2}}}, \tag{4.24a}$$

$$\hat{R}_2 = \log_2 \frac{\left|\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2\right| \sigma}{((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2)^{\frac{1}{2}}((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2)^{\frac{1}{2}}}. \tag{4.24b}$$

From (4.10) and (4.24), it is clear that[17] $\hat{R}_1 < R_1$ and $\hat{R}_2 < R_2$. Thus, for any $\rho \in [0,1]$, we have the inequality $\hat{R}_{\mathrm{wsum}} < R_{\mathrm{wsum}}$. Substituting from (4.24) into (4.23), we obtain

$$\hat{R}_{\mathrm{wsum}}(\rho) = \log_2 \frac{\left|\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1\right|^{\rho} \left|\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2\right|^{1-\rho} \sigma}{((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2)^{\frac{1}{2}}((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2)^{\frac{1}{2}}}. \tag{4.25}$$

Note that $\hat{R}_{\mathrm{wsum}}$ is a tight lower bound for R_{wsum} in (4.22d) when the SNR at both

[17] The inequality $\hat{R}_1 < R_1$ results from dropping the term 1 in the logarithm in (4.9a) and dropping the operator $[\cdot]^+$ from the rate expression in (4.10a). In a similar way, it can be shown that $\hat{R}_2 < R_2$.

receivers is sufficiently high. However, unlike R_{wsum}, whose nonnegativity is ensured by the $[\cdot]^+$ operators in (4.10), $\hat{R}_{\text{wsum}}$ can be negative since $\hat{R}_1$ and/or $\hat{R}_2$ can be negative when the corresponding SNR is sufficiently low. Nonetheless, maximizing $\hat{R}_{\text{wsum}}$ is still beneficial even when its optimal value ends up to be negative because the maximization problem is only used for the design of $\mathbf{W}$. The achievable secrecy rate pair, however, is obtained by substituting $\mathbf{W}$ into (4.10), i.e., the achievable rate pair is guaranteed to be nonnegative. Now, we formulate our design problem as[18]

$$\underset{\mathbf{W}}{\text{maximize}} \quad \ln \frac{(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^{\rho}(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2)^{1-\rho}}{((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2+\sigma^2)^{\frac{1}{2}}((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2+\sigma^2)^{\frac{1}{2}}} \tag{4.26a}$$

$$\text{s.t.} \quad \|\mathbf{W}\|_{\mathrm{F}}^2 \le P_{\mathrm{Tot}}, \tag{4.26b}$$

$$w_{1i}^2 + w_{2i}^2 \le P_i, \quad i = 1, \ldots, N. \tag{4.26c}$$

Note that the formulation in (4.26) implicitly adds the two constraints $\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1 \ge 0$ and $\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2 \ge 0$. These additional constraints cause no loss in performance because an optimal $\mathbf{w}_1$ that results in negative $\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1$ can always be replaced with $-\mathbf{w}_1$ without reducing the optimal value or violating the constraints on $\mathbf{W}$. In a similar way, the implicit constraint $\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2 \ge 0$ can be justified. Note also that, unlike the expressions in (4.24), the formulation in (4.26) does not exclude the cases $\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1 = 0$ and $\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2 = 0$ as the objective function remains well defined even when optimal $\mathbf{W}$ leads to $\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1 = 0$ or $\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2 = 0$. For example, the solution $\mathbf{w}_1 = \mathbf{0}$ (which results in $\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1 = 0$) would be optimal only when[19] $\rho = 0$, resulting in $(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^{\rho} = 0^0 = 1$.

In the next subsection, we shall explain in detail our approach to solve (4.26).

[18] Using the natural logarithm in the objective of (4.26) (instead of the logarithm to base 2) will slightly simplify the notation when differentiation becomes involved.

[19] This is true because we assume that $\mathbf{h}_1$ and $\mathbf{h}_2$ are linearly independent. On the other hand, if $\mathbf{h}_1$ and $\mathbf{h}_2$ are collinear and $\|\mathbf{h}_1\|_2 \le \|\mathbf{h}_2\|_2$, then $\mathbf{w}_1 = \mathbf{0}$ would be optimal for all $\rho \in [0,1]$, i.e., User 1 cannot achieve positive secrecy rates and should always be treated as an eavesdropper because its channel $\mathbf{h}_1$ is degraded.

4.3.2 The Outer Problem

Using the auxiliary variables $\delta_1 \geq 0$ and $\delta_2 \geq 0$, the problem in (4.26) can be expressed as

$$\underset{\mathbf{W},\delta_1,\delta_2}{\text{maximize}} \quad \ln \frac{(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^{\rho}(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2)^{1-\rho}}{(\delta_1^2+\sigma^2)^{\frac{1}{2}}(\delta_2^2+\sigma^2)^{\frac{1}{2}}} \tag{4.27a}$$

$$\text{s.t.} \quad |\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1| \leq \delta_1, \quad |\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2| \leq \delta_2, \tag{4.27b}$$

$$\|\mathbf{W}\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tot}}, \tag{4.27c}$$

$$w_{1i}^2 + w_{2i}^2 \leq P_i, \quad i = 1, \ldots, N. \tag{4.27d}$$

Let $f(\delta_1, \delta_2)$ denote the optimal value of the *perturbed problem*

$$\underset{\mathbf{W}}{\text{maximize}} \quad \rho \ln(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1) + (1-\rho)\ln(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2) \tag{4.28a}$$

$$\text{s.t.} \quad |\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1| \leq \delta_1, \quad |\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2| \leq \delta_2, \tag{4.28b}$$

$$\|\mathbf{W}\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tot}}, \tag{4.28c}$$

$$w_{1i}^2 + w_{2i}^2 \leq P_i, \quad i = 1, \ldots, N. \tag{4.28d}$$

Then, the problem in (4.27) can be written as

$$\underset{\delta_1,\delta_2 \geq 0}{\text{maximize}} \quad \varphi(\delta_1, \delta_2), \tag{4.29a}$$

where

$$\varphi(\delta_1, \delta_2) \triangleq f(\delta_1, \delta_2) - \frac{1}{2}\ln\left((\delta_1^2+\sigma^2)(\delta_2^2+\sigma^2)\right). \tag{4.29b}$$

Now, we can see that solving the design problem in (4.26) entails solving (4.28) and (4.29) iteratively. For obvious reasons, we shall refer to (4.29) as the *outer* *problem*, and to (4.28) as the *inner problem*.

The inner problem is clearly convex, and thus can be efficiently solved using standard convex optimization packages. On the other hand, the outer problem is nonconvex because the objective function $\varphi(\delta_1, \delta_2)$ is not concave, in general. Nevertheless, the following two propositions reveal that $\varphi(\delta_1, \delta_2)$ has a special structure that makes the outer problem solvable, i.e., its global maximum can be efficiently obtained, when a certain condition is satisfied. Even when such a condition is not satisfied, these propositions still give us guidelines for approaching the outer problem.

Proposition 4.1. *The objective function of the outer problem* (4.29) *is concave when restricted inside the region* $\{(\delta_1, \delta_2) : 0 \leq \delta_1 \leq \sigma, 0 \leq \delta_2 \leq \sigma\}$.

Proof: The proof is fairly straightforward. The first term in (4.29b), i.e., $f(\delta_1, \delta_2)$, is concave for all $\delta_1, \delta_2 \geq 0$ because the perturbed problem in (4.28) is convex [80, Section 5.6.1]. On the other hand, the second term $-\frac{1}{2}\ln\left((\delta_1^2 + \sigma^2)(\delta_2^2 + \sigma^2)\right)$ is concave only when $0 \leq \delta_1 \leq \sigma$ and $0 \leq \delta_2 \leq \sigma$ (this can be easily verified after writing down the Hessian matrix). Thus, $\varphi(\delta_1, \delta_2)$ is concave when $\delta_1, \delta_2 \leq \sigma$. ■

Proposition 4.2. *The objective function of the outer problem* (4.29) *is quasiconcave when restricted to any line (in the nonnegative orthant* $\mathbb{R}_+^2$*) passing through the origin.*

The proof, which is provided in Appendix C.1, is based on the observation that the first term in (4.29b) is nondecreasing (w.r.t. $\mathbb{R}_+^2$), while the second term is monotonically decreasing[20]. Note that the condition in Proposition 4.2 is weaker than stating that $\varphi(\delta_1, \delta_2)$ is quasiconcave, as the latter condition would require φ to be quasiconcave when restricted to *any line* in $\mathbb{R}_+^2$.

Combining Propositions 4.1 and 4.2 immediately yields the following conclusion:

[20] See [80, Section 3.6.1] for the notion of *monotonicity w.r.t. a generalized inequality* on the nonnegative orthant.

Corollary 4.1. *For the outer problem* (4.29)*, any local maximum inside the region* $\{(\delta_1, \delta_2) : 0 \le \delta_1 \le \sigma, 0 \le \delta_2 \le \sigma\}$ *is a global maximum.*

Corollary 4.1 suggests that we start searching for the solution of (4.29) inside the region $\delta_1, \delta_2 \le \sigma$. If the search algorithm terminates at $\boldsymbol{\delta}^\star = (\delta_1^\star, \delta_2^\star)$ such that $\delta_1^\star, \delta_2^\star \le \sigma$, then $\boldsymbol{\delta}^\star$ is guaranteed to be the (globally) optimal solution of (4.29). On the other hand, if $\delta_1^\star > \sigma$ or $\delta_2^\star > \sigma$, then we will accept $\boldsymbol{\delta}^\star$ as a (possibly) suboptimal solution. It is worth to mention that the numerical results show that $\varphi(\delta_1, \delta_2)$ is a *unimodal* function with only one maximum, for all $\delta_1, \delta_2 \ge 0$, and no other stationary points. However, it is difficult, in general, to rigorously prove that a multivariable function is unimodal, beyond concavity or quasiconcavity. Therefore, we can only conjecture that $\varphi(\delta_1, \delta_2)$ is unimodal (for all $\delta_1, \delta_2 \ge 0$), and consequently any local maximum is global.

Now, we have to choose a reasonable search algorithm to solve (4.29). Since the objective function $\varphi(\delta_1, \delta_2)$ is differentiable almost everywhere[21], a natural choice for the search algorithm is the *subgradient method* in which the subgradient vectors are used as the search directions [90, 91]. Let the vector $\nabla_{\text{sub}} f(\delta_1, \delta_2) \in \mathbb{R}_+^2$ be a subgradient[22] of f at (δ_1, δ_2), where the two entries of $\nabla_{\text{sub}} f$ are both nonnegative since f is nondecreasing w.r.t. δ_1 and δ_2. Then, the corresponding subgradient of φ is given by

$$\nabla_{\text{sub}} \varphi(\delta_1, \delta_2) = \nabla_{\text{sub}} f(\delta_1, \delta_2) - \begin{bmatrix} \dfrac{\delta_1}{\delta_1^2 + \sigma^2} & \dfrac{\delta_2}{\delta_2^2 + \sigma^2} \end{bmatrix}^{\mathrm{T}}. \tag{4.30}$$

Before we proceed to the details of the search algorithm, we need to find $\nabla_{\text{sub}} f$ in order to calculate the search direction $\nabla_{\text{sub}} \varphi$ at any (δ_1, δ_2). This will be our goal in the next subsection.

[21] This is because $f(\delta_1, \delta_2)$ is not necessarily differentiable (everywhere). Nevertheless, since $f(\delta_1, \delta_2)$ is concave, it is differentiable *almost* everywhere.

[22] The term "*supergradient*" is probably more appropriate here because $f(\delta_1, \delta_2)$ is a concave function.

4.3.3 The Dual of the Inner Problem

The inner problem (4.28) is a convex problem whose constraints satisfy *Slater's condition*, and thus strong duality holds [80, Section 5.2.3]. As a consequence, the optimal value of the inner problem, i.e., $f(\delta_1, \delta_2)$, is identical to the optimal value of its (Lagrange) dual. Furthermore, the optimal Lagrange multipliers associated with the two constraints in (4.28b) provide a subgradient vector[23] for f at (δ_1, δ_2) [92, Section 8.5.6]. Therefore, our next task is to derive the dual problem for (4.28).

We begin with reformulating (4.28) as

$$\begin{aligned}
\underset{\mathbf{W}, z_1, \ldots, z_4}{\text{maximize}} \quad & \rho \ln z_1 + (1-\rho) \ln z_2 && \text{(4.31a)}\\
\text{s.t.} \quad & |z_3| \leq \delta_1, \quad |z_4| \leq \delta_2, && \text{(4.31b)}\\
& \|\mathbf{w}_1\|_2^2 + \|\mathbf{w}_2\|_2^2 \leq P_{\text{Tot}}, && \text{(4.31c)}\\
& w_{1i}^2 + w_{2i}^2 \leq P_i, \quad i = 1, \ldots, N, && \text{(4.31d)}\\
& \mathbf{h}_1^{\mathrm{T}} \mathbf{w}_1 = z_1, \quad \mathbf{h}_2^{\mathrm{T}} \mathbf{w}_2 = z_2, && \text{(4.31e)}\\
& \mathbf{h}_2^{\mathrm{T}} \mathbf{w}_1 = z_3, \quad \mathbf{h}_1^{\mathrm{T}} \mathbf{w}_2 = z_4, && \text{(4.31f)}
\end{aligned}$$

where we have introduced four new variables, $z_1, \ldots, z_4$, and four associated equality constraints (4.31e)-(4.31f). The Lagrangian associated with the reformulated problem in (4.31) is

$$\begin{aligned}
& L(\mathbf{W}, z_1, \ldots, z_4, \lambda_1, \lambda_2, \gamma, \boldsymbol{\mu}, \nu_1, \ldots, \nu_4)\\
&\quad = \rho \ln z_1 + (1-\rho) \ln z_2 - \lambda_1 \left(|z_3| - \delta_1\right) - \lambda_2 \left(|z_4| - \delta_2\right)\\
&\qquad - \gamma \left(\|\mathbf{w}_1\|_2^2 + \|\mathbf{w}_2\|_2^2 - P_{\text{Tot}}\right) - \sum_{i=1}^{N} \mu_i (w_{1i}^2 + w_{2i}^2 - P_i)\\
&\qquad - \nu_1(\mathbf{h}_1^{\mathrm{T}} \mathbf{w}_1 - z_1) - \nu_2(\mathbf{h}_2^{\mathrm{T}} \mathbf{w}_2 - z_2) - \nu_3(\mathbf{h}_2^{\mathrm{T}} \mathbf{w}_1 - z_3) - \nu_4(\mathbf{h}_1^{\mathrm{T}} \mathbf{w}_2 - z_4), \qquad (4.32)
\end{aligned}$$

[23] Note that f has more than one subgradient at the points (δ_1, δ_2) where f is non-differentiable.

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the Lagrange multipliers associated with the perturbed constraints in (4.31b), $\gamma \geq 0$ is the Lagrange multiplier associated with the total power constraint (4.31c), $\boldsymbol{\mu} = [\mu_1 \dots \mu_N]^{\mathrm{T}}$, with entries $\mu_i \geq 0, i = 1, \dots, N$, is the Lagrange multiplier vector associated with the per-antenna power constraint (4.31d), and $\nu_1, \dots, \nu_4$ are the Lagrange multipliers associated with the equality constraints in (4.31e)-(4.31f). Upon rearranging the terms in the Lagrangian (4.32), the dual function g is obtained by maximization over the primary variables $\mathbf{W}, z_1, \dots, z_4$, that is

$$\begin{aligned} g(\lambda_1, \lambda_2, \gamma, \boldsymbol{\mu}, \nu_1, \dots, \nu_4) = {} & \lambda_1 \delta_1 + \lambda_2 \delta_2 + \gamma P_{\mathrm{Tot}} + \sum_{i=1}^{N} \mu_i P_i \\ & + \sum_{i=1}^{N} \max_{w_{1i}} \left(- \left(\nu_1 h_{1i} + \nu_3 h_{2i} \right) w_{1i} - (\gamma + \mu_i) w_{1i}^2 \right) \\ & + \sum_{i=1}^{N} \max_{w_{2i}} \left(- \left(\nu_2 h_{2i} + \nu_4 h_{1i} \right) w_{2i} - (\gamma + \mu_i) w_{2i}^2 \right) \\ & + \max_{z_1} \left(\nu_1 z_1 + \rho \ln z_1 \right) + \max_{z_2} \left(\nu_2 z_2 + (1-\rho) \ln z_2 \right) \\ & + \max_{z_3} \left(\nu_3 z_3 - \lambda_1 |z_3| \right) + \max_{z_4} \left(\nu_4 z_4 - \lambda_2 |z_4| \right), \end{aligned} \tag{4.33}$$

where h_{1i} and h_{2i} are the ith entries of $\mathbf{h}_1$ and $\mathbf{h}_2$, respectively. Now, we have to solve all the maximization terms in (4.33) analytically. In fact, we have

$$\begin{aligned} & \max_{w_{1i}} \left(- \left(\nu_1 h_{1i} + \nu_3 h_{2i} \right) w_{1i} - (\gamma + \mu_i) w_{1i}^2 \right) \\ & \qquad = \frac{\left(\nu_1 h_{1i} + \nu_3 h_{2i} \right)^2}{4(\gamma + \mu_i)}, \quad \gamma + \mu_i > 0, \quad i = 1, \dots, N, \end{aligned} \tag{4.34a}$$

$$\max_{z_1} \left(\nu_1 z_1 + \rho \ln z_1 \right) = -\rho \ln \frac{-\nu_1}{\rho} - \rho, \quad \nu_1 < 0, \tag{4.34b}$$

$$\max_{z_3} \left(\nu_3 z_3 - \lambda_1 |z_3|\right) = \begin{cases} 0 & |\nu_3| \leq \lambda_1 \\ \infty & \text{otherwise} \end{cases}, \tag{4.34c}$$

where (4.34a) is a simple unconstrained quadratic concave maximization problem, (4.34b) follows from the conjugate of the negative logarithm function (see [80, Example 3.21]), and (4.34c) follows from the conjugate of the absolute value function (see [80, Example 3.26]). Note that the condition $\gamma + \mu_i > 0$ in (4.34a) is always satisfied because, for each antenna, at least one of the constraints (i.e., the total power constraint or the per-antenna power constraint) must be active. Thus, $\gamma + \mu_i$ is strictly positive for all $i = 1, \ldots, N$. Using the expressions in (4.34), the dual problem can be formulated as[24]

$$\underset{\substack{\lambda_1,\lambda_2,\gamma,\boldsymbol{\mu},\boldsymbol{\tau}_1,\\ \boldsymbol{\tau}_2,\nu_1,\ldots,\nu_4}}{\text{minimize}} \left(\begin{array}{c} \delta_1\lambda_1 + \delta_2\lambda_2 + P_{\text{Tot}}\gamma + \sum_{i=1}^{N}(P_i\mu_i + \tau_{1i} + \tau_{2i}) \\ -\rho \ln \dfrac{-\nu_1}{\rho} - (1-\rho)\ln \dfrac{-\nu_2}{1-\rho} \end{array} \right) - 1 \tag{4.35a}$$

$$\text{s.t.} \quad \nu_1, \nu_2 < 0, \quad |\nu_3| \leq \lambda_1, \quad |\nu_4| \leq \lambda_2, \tag{4.35b}$$

$$\gamma \geq 0, \quad \mu_i \geq 0, \quad \gamma + \mu_i > 0, \qquad i = 1, \ldots, N, \tag{4.35c}$$

$$\begin{bmatrix} \tau_{1i} & \nu_1 h_{1i} + \nu_3 h_{2i} \\ \nu_1 h_{1i} + \nu_3 h_{2i} & 4(\gamma + \mu_i) \end{bmatrix} \succeq 0, \qquad i = 1, \ldots, N, \tag{4.35d}$$

$$\begin{bmatrix} \tau_{2i} & \nu_2 h_{2i} + \nu_4 h_{1i} \\ \nu_2 h_{2i} + \nu_4 h_{1i} & 4(\gamma + \mu_i) \end{bmatrix} \succeq 0, \qquad i = 1, \ldots, N, \tag{4.35e}$$

where we have used *Schur complement*, in conjunction with the auxiliary variables τ_{1i} and τ_{2i}, $i = 1, \ldots, N$, to formulate the linear matrix inequality constraints in (4.35d) and (4.35e). Two special cases of the dual problem (4.35) are worth mentioning.

[24] We maintain the fixed term -1 in the objective function in (4.35a) to have its optimal value equal to the optimal value of the inner problem (4.28), i.e., equal to $f(\delta_1, \delta_2)$.

First, at the corner point $\rho = 0$, the Lagrange multipliers λ_1, ν_1, and ν_3 are set to zero, and the dual problem (4.35) simplifies to

$$\underset{\substack{\lambda_2,\gamma,\boldsymbol{\mu},\\ \boldsymbol{\tau}_2,\nu_2,\nu_4}}{\text{minimize}} \left(\delta_2\lambda_2 + P_{\text{Tot}}\gamma + \sum_{i=1}^{N}(P_i\mu_i + \tau_{2i}) - \ln(-\nu_2) \right) - 1 \tag{4.36a}$$

$$\text{s.t.} \quad \nu_2 < 0, \quad |\nu_4| \le \lambda_2, \tag{4.36b}$$

$$\gamma \ge 0, \quad \mu_i \ge 0, \quad \gamma + \mu_i > 0, \qquad i = 1, \dots, N, \tag{4.36c}$$

$$\begin{bmatrix} \tau_{2i} & \nu_2 h_{2i} + \nu_4 h_{1i} \\ \nu_2 h_{2i} + \nu_4 h_{1i} & 4(\gamma + \mu_i) \end{bmatrix} \succeq 0, \qquad i = 1, \dots, N, \tag{4.36d}$$

where we have used the convention that $0 \ln \frac{0}{0} = 0$ while simplifying the objective. A similar simplification can be obtained for the other corner point, i.e., at $\rho = 1$.

Second, for the case in which there is only a total power constraint, i.e., when the per-antenna constraint in (4.31d) does not exist or is not active, the Lagrange multiplier vector $\boldsymbol{\mu}$ is set to $\mathbf{0}$, and (4.35) simplifies to

$$\underset{\substack{\lambda_1,\lambda_2,\gamma,\tau_1,\\ \tau_2,\nu_1,\dots,\nu_4}}{\text{minimize}} \begin{pmatrix} \delta_1\lambda_1 + \delta_2\lambda_2 + P_{\text{Tot}}\gamma + \tau_1 + \tau_2 \\ -\rho \ln \dfrac{-\nu_1}{\rho} - (1-\rho)\ln \dfrac{-\nu_2}{1-\rho} \end{pmatrix} - 1 \tag{4.37a}$$

$$\text{s.t.} \quad \nu_1, \nu_2 < 0, \quad |\nu_3| \le \lambda_1, \quad |\nu_4| \le \lambda_2, \tag{4.37b}$$

$$\gamma > 0, \tag{4.37c}$$

$$\begin{bmatrix} \tau_1 & (\nu_1\mathbf{h}_1 + \nu_3\mathbf{h}_2)^{\mathrm{T}} \\ \nu_1\mathbf{h}_1 + \nu_3\mathbf{h}_2 & 4\gamma\mathbf{I}_N \end{bmatrix} \succeq 0, \tag{4.37d}$$

$$\begin{bmatrix} \tau_2 & (\nu_2\mathbf{h}_2 + \nu_4\mathbf{h}_1)^{\mathrm{T}} \\ \nu_2\mathbf{h}_2 + \nu_4\mathbf{h}_1 & 4\gamma\mathbf{I}_N \end{bmatrix} \succeq 0. \tag{4.37e}$$

The dual problem (4.35) is, of course, convex and thus can be efficiently solved to obtain $f(\delta_1, \delta_2)$, as well as $\nabla_{\text{sub}} f(\delta_1, \delta_2)$. Let $\{\lambda_1^\star, \lambda_2^\star, \gamma^\star, \boldsymbol{\mu}^\star, \boldsymbol{\tau}_1^\star, \boldsymbol{\tau}_2^\star, \nu_1^\star, \ldots, \nu_4^\star\}$ be the optimal solution of (4.35) for given δ_1 and δ_2. Then, $f(\delta_1, \delta_2)$ is equal to the optimal value of the objective, and the vector $[\lambda_1^\star \; \lambda_2^\star]^{\mathrm{T}}$ is a subgradient of f at (δ_1, δ_2). Consequently, the subgradient vector in (4.30) can be written as

$$\nabla_{\text{sub}} \varphi(\delta_1, \delta_2) = \left[\lambda_1^\star - \frac{\delta_1}{\delta_1^2 + \sigma^2} \quad \lambda_2^\star - \frac{\delta_2}{\delta_2^2 + \sigma^2} \right]^{\mathrm{T}}. \tag{4.38}$$

Having obtained $\nabla_{\text{sub}} \varphi(\delta_1, \delta_2)$, we are now ready to use the subgradient method to solve (4.29).

4.3.4 The Search Algorithm

In this subsection, we turn our focus to the search algorithm used to find a solution for the outer problem (4.29), i.e., to find $\boldsymbol{\delta}^\star = [\delta_1^\star \; \delta_2^\star]^{\mathrm{T}}$ that maximizes $\varphi(\delta_1, \delta_2)$. A typical subgradient method uses the iteration [91]

$$\boldsymbol{\delta}^{(k+1)} = \boldsymbol{\delta}^{(k)} + \alpha^{(k)} \, \nabla_{\text{sub}} \varphi(\boldsymbol{\delta}^{(k)}), \quad k = 1, 2, \ldots, \tag{4.39}$$

where $\boldsymbol{\delta}^{(k)}$ is the start point at the kth iteration (with $\boldsymbol{\delta}^{(1)}$ being the initial point), $\alpha^{(k)} > 0$ is the kth step size, and $\boldsymbol{\delta}^{(k+1)}$ is the end point after k iterations. The numerical results in Section 4.5 reveal that, when the noise variance σ^2 is equal to 1, the values of $\delta_1^\star$ and $\delta_2^\star$ can be on the order of 10^{-4} up to 10^{1}. This several orders of magnitude difference suggests that the search is better carried out on a logarithmic scale, rather than the ordinary linear scale, in order to improve the accuracy and maintain numerical stability (so convergence is achieved within a reasonable number of iterations).

Let $\boldsymbol{\delta}_{\mathrm{dB}}$ be defined as $\boldsymbol{\delta}_{\mathrm{dB}} \triangleq [20\log_{10}(\delta_1) \quad 20\log_{10}(\delta_2)]^{\mathrm{T}}$. Then, the subgradient $\nabla_{\mathrm{sub}}\varphi$ on the logarithmic scale, i.e., when differentiation is w.r.t. $20\log_{10}(\delta_1)$ and $20\log_{10}(\delta_2)$, is given by

$$\nabla_{\mathrm{sub}}\varphi(\boldsymbol{\delta}_{\mathrm{dB}}) = \frac{\ln 10}{20}\begin{bmatrix}\delta_1\left(\lambda_1^\star - \dfrac{\delta_1}{\delta_1^2+\sigma^2}\right)\\ \delta_2\left(\lambda_2^\star - \dfrac{\delta_2}{\delta_2^2+\sigma^2}\right)\end{bmatrix}. \tag{4.40}$$

Now, we proceed with the search algorithm as follows. First, we choose an initial point $\boldsymbol{\delta}_{\mathrm{dB}}^{(1)}$, such that $\delta_1^{(1)} \leq \sigma$ and $\delta_2^{(1)} \leq \sigma$. This point is iteratively updated by

$$\boldsymbol{\delta}_{\mathrm{dB}}^{(k+1)} = \boldsymbol{\delta}_{\mathrm{dB}}^{(k)} + \alpha_{\mathrm{dB}}^{\mathrm{Fix}}\,\frac{\nabla_{\mathrm{sub}}\varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(k)})}{\|\nabla_{\mathrm{sub}}\varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(k)})\|_2}, \quad k = 1, 2, \ldots, \tag{4.41}$$

where $\alpha_{\mathrm{dB}}^{\mathrm{Fix}}$ is a fixed step size in dB. That is, for each iteration, we take a step $\alpha_{\mathrm{dB}}^{\mathrm{Fix}}$ in the direction of the subgradient. This iteration shall continue until we overshoot the peak, i.e., when $\varphi(\boldsymbol{\delta})$ starts to decrease. Once the peak is encountered, we reduce the step size and use the iteration

$$\boldsymbol{\delta}_{\mathrm{dB}}^{(K+l+1)} = \boldsymbol{\delta}_{\mathrm{dB}}^{(K+l)} + \frac{\alpha_{\mathrm{dB}}^{\mathrm{Fix}}}{l}\,\frac{\nabla_{\mathrm{sub}}\varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(K+l)})}{\|\nabla_{\mathrm{sub}}\varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(K+l)})\|_2}, \quad l = 1, \ldots, L, \tag{4.42}$$

where K is the number of iterations using (4.41), i.e., with a fixed step size, and L is the maximum number of iterations with a decreasing step size. Unlike K, L is determined in advance according to the required accuracy of the solution. Therefore, the search will terminate after $K+L$ total iterations, and the solution $\boldsymbol{\delta}_{\mathrm{dB}}^\star$ is obtained with accuracy $\alpha_{\mathrm{dB}}^{\mathrm{Fix}}/L$ dB. For convenience, the algorithm is summarized in Table 4.1.

Upon solving the outer problem (4.29), we solve the inner problem (4.28) using $\boldsymbol{\delta}^\star$ to obtain the optimum precoding matrix $\mathbf{W}^\star$. Then, the secrecy rate pair (R_1, R_2)

Table 4.1: Subgradient-based search algorithm to solve the maximization problem in (4.29).

Algorithm 4.1 Subgradient-based method to solve (4.29)

```
Set the initial (fixed) step size $\alpha_{\mathrm{dB}}^{\mathrm{Fix}}$ and the maximum number of iterations with decreasing step size $L$
Set the binary switch REDUCE = false
Set the indexes $k = 0$ and $l = 1$
Choose an initial point $\boldsymbol{\delta}_{\mathrm{dB}}^{(1)}$ such that $\delta_1^{(1)} \leq \sigma$, $\delta_2^{(1)} \leq \sigma$
while $l \leq L$ do
    Solve (4.35) to obtain $f(\boldsymbol{\delta}_{\mathrm{dB}}^{(k+l)})$, $\lambda_1^{\star(k+l)}$, $\lambda_2^{\star(k+l)}$
    Calculate $\varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(k+l)})$ using (4.29b)
    Calculate $\nabla_{\mathrm{sub}}\varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(k+l)})$ using (4.40)
    if $\varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(k+l)}) \leq \varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(k+l-1)})$, then
        REDUCE = true
    end if
    if REDUCE = false, then
        Update $\boldsymbol{\delta}_{\mathrm{dB}}^{(k+l)}$ using (4.41)
        $k := k + 1$
    else
        Update $\boldsymbol{\delta}_{\mathrm{dB}}^{(k+l)}$ using (4.42)
        $l := l + 1$
    end if
end while
return $\boldsymbol{\delta}_{\mathrm{dB}}^{\star} = \operatorname{argmax}\{\varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(1)}), \ldots, \varphi(\boldsymbol{\delta}_{\mathrm{dB}}^{(k+L+1)})\}$
```

is calculated by substituting $\mathbf{W}^\star$ into (4.10). We repeat the entire procedure with different values of $\rho \in [0, 1]$ to obtain different points (R_1, R_2) on the boundary of the secrecy rate region.

4.3.5 Per-Antenna Amplitude Constraint

In this subsection, we design the precoding matrix $\mathbf{W}$ subject to the per-antenna amplitude constraint (4.14b). Fortunately, the problem formulation and solution techniques developed in the previous subsections are immediately applicable. In fact, we just need to modify the weighted secrecy sum rate expression (4.25) and the inner problem (4.28), and consequently its dual (4.35), to take the amplitude constraint into account.

Similar to (4.25), we need a weighted secrecy sum rate expression that is amenable to optimization. From (4.20), the rate expressions R_1 and R_2, respectively, can be lower-bounded by

$$\hat{R}_1 = \log_2 \frac{3\sqrt{2}\left|\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1\right|\sigma}{\sqrt{\pi e}\left((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + 3\sigma^2\right)^{\frac{1}{2}}\left((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + 3\sigma^2\right)^{\frac{1}{2}}}, \tag{4.43a}$$

$$\hat{R}_2 = \log_2 \frac{3\sqrt{2}\left|\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2\right|\sigma}{\sqrt{\pi e}\left((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + 3\sigma^2\right)^{\frac{1}{2}}\left((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + 3\sigma^2\right)^{\frac{1}{2}}}. \tag{4.43b}$$

Then, for any $\rho \in [0, 1]$, we have the weighted secrecy sum rate

$$\hat{R}_{\text{wsum}}(\rho) = \log_2 \frac{3\sqrt{2}\left|\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1\right|^{\rho}\left|\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2\right|^{1-\rho}\sigma}{\sqrt{\pi e}((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + 3\sigma^2)^{\frac{1}{2}}((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + 3\sigma^2)^{\frac{1}{2}}}. \tag{4.44}$$

Similar to (4.28), we formulate the inner problem as

$$\begin{aligned}
\underset{\mathbf{W}}{\text{maximize}} \quad & \rho \ln(\mathbf{h}_1^{\mathrm{T}} \mathbf{w}_1) + (1-\rho)\ln(\mathbf{h}_2^{\mathrm{T}} \mathbf{w}_2) && \text{(4.45a)}\\
\text{s.t.} \quad & |\mathbf{h}_2^{\mathrm{T}} \mathbf{w}_1| \le \delta_1, \quad |\mathbf{h}_1^{\mathrm{T}} \mathbf{w}_2| \le \delta_2, && \text{(4.45b)}\\
& |w_{1i}| + |w_{2i}| \le A_i, \quad i = 1, \ldots, N. && \text{(4.45c)}
\end{aligned}$$

Then, following the same procedure as in Section 4.3.3, it can be shown that the dual problem for (4.45) is

$$\begin{aligned}
\underset{\substack{\lambda_1,\lambda_2,\boldsymbol{\mu},\\ \nu_1,\ldots,\nu_4}}{\text{minimize}} \quad & \begin{pmatrix} \delta_1\lambda_1 + \delta_2\lambda_2 + \sum_{i=1}^{N}(A_i\mu_i) \\ -\rho \ln \dfrac{-\nu_1}{\rho} - (1-\rho)\ln\dfrac{-\nu_2}{1-\rho} \end{pmatrix} - 1 && \text{(4.46a)}\\
\text{s.t.} \quad & \nu_1, \nu_2 < 0, \quad |\nu_3| \le \lambda_1, \quad |\nu_4| \le \lambda_2, && \text{(4.46b)}\\
& |\nu_1 h_{1i} + \nu_3 h_{2i}| \le \mu_i, \quad i = 1, \ldots, N, && \text{(4.46c)}\\
& |\nu_2 h_{2i} + \nu_4 h_{1i}| \le \mu_i, \quad i = 1, \ldots, N, && \text{(4.46d)}
\end{aligned}$$

where the Lagrange multipliers $\lambda_1, \lambda_2, \nu_1, \ldots, \nu_4$ are defined as in (4.35), and $\boldsymbol{\mu} = [\mu_1 \ldots \mu_N]^{\mathrm{T}}$ is the Lagrange multiplier vector associated with the amplitude constraint (4.45c).

4.4 Robust Precoder Design with Imperfect Channel Information

Our solutions in Section 4.3 were based on the assumption that the channel gain vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ are precisely known to the transmitter. In this section, we capitalize on our approach and tackle the more general design problem in which the transmitter has only uncertain estimates of $\mathbf{h}_1$ and $\mathbf{h}_2$. We will see that the problem formulation is very similar to its non-robust counterpart, and thus the solution approach will also be similar. Therefore, our pace in this section will be relatively fast, and much of the details and derivations encountered in the previous section will be omitted for brevity.

4.4.1 Channel Uncertainty Model

We adopt the spherical uncertainty model (or norm-bounded error model) in which the actual channel gain vectors, $\mathbf{h}_1$ and $\mathbf{h}_2$, respectively, are modelled by

$$\mathbf{h}_1 \in \mathcal{H}_1, \quad \mathcal{H}_1 = \left\{ \hat{\mathbf{h}}_1 + \mathbf{e}_1 : \|\mathbf{e}_1\|_2 \leq \epsilon_1 \right\}, \tag{4.47a}$$

$$\mathbf{h}_2 \in \mathcal{H}_2, \quad \mathcal{H}_2 = \left\{ \hat{\mathbf{h}}_2 + \mathbf{e}_2 : \|\mathbf{e}_2\|_2 \leq \epsilon_2 \right\}, \tag{4.47b}$$

where $\mathcal{H}_1$ and $\mathcal{H}_2$ are N-dimensional spherical sets, $\hat{\mathbf{h}}_1 \in \mathbb{R}^N$ and $\hat{\mathbf{h}}_2 \in \mathbb{R}^N$ are the channel vector estimates available to the transmitter, $\mathbf{e}_1 \in \mathbb{R}^N$ and $\mathbf{e}_2 \in \mathbb{R}^N$ are unknown (but norm-bounded) error vectors, and ϵ_1 and ϵ_2 are known constants that quantify the amount of uncertainty for each channel. This error model is well accepted for representing channel uncertainty caused by quantization errors and finite-rate feedback from each receiver to the transmitter [57, Lemma 1].

Given the uncertain channel information in (4.47), our goal in this section is to design the precoding matrix $\mathbf{W}$ in order to optimize the performance in terms of the worst-case secrecy rate pair $(R_1^{\mathrm{wc}}, R_2^{\mathrm{wc}})$. That is to solve the two-objective optimization problem

$$\underset{\mathbf{W}}{\text{maximize}} \;\; (R_1^{\mathrm{wc}}, R_2^{\mathrm{wc}}) \tag{4.48}$$

subject to power or amplitude constraints, where, for fixed $\mathbf{W}$, the worst-case secrecy rates R_1^{wc} and R_2^{wc} are determined by

$$R_1^{\mathrm{wc}} = \min_{\substack{\mathbf{h}_1 \in \mathcal{H}_1, \\ \mathbf{h}_2 \in \mathcal{H}_2}} R_1, \tag{4.49a}$$

$$R_2^{\mathrm{wc}} = \min_{\substack{\mathbf{h}_1 \in \mathcal{H}_1, \\ \mathbf{h}_2 \in \mathcal{H}_2}} R_2. \tag{4.49b}$$

Similar to our approach in the previous section, we shall tackle (4.48) by solving a weighted worst-case secrecy sum rate maximization problem, as we see in the following two subsections.

4.4.2 Total and Per-Antenna Average Power Constraints

In this subsection, we solve the weighted worst-case secrecy sum rate maximization problem subject to total and per-antenna power constraints. First, we need to simplify the worst-case secrecy rate expressions in order to obtain a weighted sum rate that is amenable to optimization. Substituting from (4.10a) into (4.49a), we obtain

$$\begin{aligned}
R_1^{\mathrm{wc}} &= \left[\frac{1}{2}\log_2 \min_{\mathbf{h}_1 \in \mathcal{H}_1} \left(1 + \frac{(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^2}{(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2}\right) + \frac{1}{2}\log_2 \min_{\mathbf{h}_2 \in \mathcal{H}_2} \left(\frac{\sigma^2}{(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2}\right)\right]^+ \\
&\geq \frac{1}{2}\log_2 \left(1 + \frac{\min\limits_{\mathbf{h}_1 \in \mathcal{H}_1} (\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^2}{\max\limits_{\mathbf{h}_1 \in \mathcal{H}_1} (\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2}\right) + \frac{1}{2}\log_2 \left(\frac{\sigma^2}{\max\limits_{\mathbf{h}_2 \in \mathcal{H}_2} (\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2}\right) && (4.50a) \\
&\geq \log_2 \frac{\min\limits_{\mathbf{h}_1 \in \mathcal{H}_1} \left|\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1\right| \sigma}{\max\limits_{\mathbf{h}_1 \in \mathcal{H}_1} ((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2 + \sigma^2)^{\frac{1}{2}} \max\limits_{\mathbf{h}_2 \in \mathcal{H}_2} ((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2 + \sigma^2)^{\frac{1}{2}}}, && (4.50b)
\end{aligned}$$

where the first inequality follows from dropping the $[\cdot]^+$ operator and applying the inequality

$$\min_x \frac{f_1(x)}{f_2(x)} \geq \frac{\min\limits_x f_1(x)}{\max\limits_x f_2(x)},$$

which holds for arbitrary functions f_1 and f_2, and the second inequality follows from dropping the term 1. We shall use (4.50b) to formulate the weighted secrecy sum rate for the optimization problem, while we use the better bound in (4.50a) to calculate the worst-case secrecy rate R_1^{wc} after obtaining $\mathbf{W}$. Similarly, we have

$$\begin{aligned} R_2^{\mathrm{wc}} &\geq \frac{1}{2}\log_2\left(1+\frac{\min\limits_{\mathbf{h}_2\in\mathcal{H}_2}(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2)^2}{\max\limits_{\mathbf{h}_2\in\mathcal{H}_2}(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2+\sigma^2}\right)+\frac{1}{2}\log_2\left(\frac{\sigma^2}{\max\limits_{\mathbf{h}_1\in\mathcal{H}_1}(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2+\sigma^2}\right) && \text{(4.51a)}\\ &\geq \log_2\frac{\min\limits_{\mathbf{h}_2\in\mathcal{H}_2}\left|\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2\right|\sigma}{\max\limits_{\mathbf{h}_2\in\mathcal{H}_2}((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2+\sigma^2)^{\frac{1}{2}}\max\limits_{\mathbf{h}_1\in\mathcal{H}_1}((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2+\sigma^2)^{\frac{1}{2}}}. && \text{(4.51b)}\end{aligned}$$

Next, we combine the rate expressions in (4.50b) and (4.51b) using the weights ρ and $1-\rho$, for any $\rho\in[0,1]$, to formulate the robust design problem

$$\begin{aligned} \underset{\mathbf{W}}{\text{maximize}}\quad & \ln\frac{\min\limits_{\mathbf{h}_1\in\mathcal{H}_1}(\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1)^{\rho}\min\limits_{\mathbf{h}_2\in\mathcal{H}_2}(\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2)^{1-\rho}}{\max\limits_{\mathbf{h}_2\in\mathcal{H}_2}((\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1)^2+\sigma^2)^{\frac{1}{2}}\max\limits_{\mathbf{h}_1\in\mathcal{H}_1}((\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2)^2+\sigma^2)^{\frac{1}{2}}} && \text{(4.52a)}\\ \text{s.t.}\quad & \|\mathbf{W}\|_{\mathrm{F}}^2\leq P_{\mathrm{Tot}}, && \text{(4.52b)}\\ & w_{1i}^2+w_{2i}^2\leq P_i,\quad i=1,\dots,N. && \text{(4.52c)}\end{aligned}$$

Problem (4.52), in turn, can be expressed as

$$\begin{aligned}
\underset{\mathbf{W},z_1,z_2,\delta_1,\delta_2}{\text{maximize}} \quad & \ln \frac{z_1^{\rho}\, z_2^{1-\rho}}{(\delta_1^2+\sigma^2)^{\frac{1}{2}}(\delta_2^2+\sigma^2)^{\frac{1}{2}}} && \text{(4.53a)}\\
\text{s.t.} \quad & \mathbf{h}_1^{\mathrm{T}}\mathbf{w}_1 \geq z_1 \quad \forall \mathbf{h}_1 \in \mathcal{H}_1, && \text{(4.53b)}\\
& \mathbf{h}_2^{\mathrm{T}}\mathbf{w}_2 \geq z_2 \quad \forall \mathbf{h}_2 \in \mathcal{H}_2, && \text{(4.53c)}\\
& |\mathbf{h}_2^{\mathrm{T}}\mathbf{w}_1| \leq \delta_1 \quad \forall \mathbf{h}_2 \in \mathcal{H}_2, && \text{(4.53d)}\\
& |\mathbf{h}_1^{\mathrm{T}}\mathbf{w}_2| \leq \delta_2 \quad \forall \mathbf{h}_1 \in \mathcal{H}_1, && \text{(4.53e)}\\
& \|\mathbf{W}\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tot}}, && \text{(4.53f)}\\
& w_{1i}^2 + w_{2i}^2 \leq P_i, \quad i = 1, \ldots, N. && \text{(4.53g)}
\end{aligned}$$

Utilizing the expressions of the uncertainty sets $\mathcal{H}_1$ and $\mathcal{H}_2$ in (4.47), the inequalities in (4.53b), (4.53c), (4.53d), and (4.53e), respectively, can be replaced by

$$\begin{aligned}
& \hat{\mathbf{h}}_1^{\mathrm{T}}\mathbf{w}_1 - \epsilon_1\|\mathbf{w}_1\|_2 \geq z_1, && \text{(4.54a)}\\
& \hat{\mathbf{h}}_2^{\mathrm{T}}\mathbf{w}_2 - \epsilon_2\|\mathbf{w}_2\|_2 \geq z_2, && \text{(4.54b)}\\
& |\hat{\mathbf{h}}_2^{\mathrm{T}}\mathbf{w}_1| + \epsilon_2\|\mathbf{w}_1\|_2 \leq \delta_1, && \text{(4.54c)}\\
& |\hat{\mathbf{h}}_1^{\mathrm{T}}\mathbf{w}_2| + \epsilon_1\|\mathbf{w}_2\|_2 \leq \delta_2. && \text{(4.54d)}
\end{aligned}$$

Similar to (4.28), let $f(\delta_1, \delta_2)$ denote the optimal value of the perturbed problem

$$\begin{aligned}
\underset{\mathbf{W},z_1,z_2}{\text{maximize}} \quad & \rho \ln z_1 + (1-\rho)\ln z_2 && \text{(4.55a)}\\
\text{s.t.} \quad & (4.54\text{a}), (4.54\text{b}), (4.54\text{c}), (4.54\text{d}), (4.53\text{f}), (4.53\text{g}). && \text{(4.55b)}
\end{aligned}$$

Then, the robust design problem (4.53) can be expressed as

$$\underset{\delta_1,\delta_2\geq 0}{\text{maximize}} \;\; f(\delta_1,\delta_2) - \frac{1}{2}\ln\left((\delta_1^2+\sigma^2)(\delta_2^2+\sigma^2)\right). \tag{4.56}$$

Again, we shall refer to (4.56) as the outer problem, and to (4.55) as the inner problem. It is clear that the inner problem (4.55) is convex, and the outer problem (4.56) is essentially identical to (4.29). Thus, it can be shown that Propositions 4.1 and 4.2 hold for the objective of (4.56) as well. Consequently, (4.56) can be solved iteratively using Algorithm 4.1. In each iteration, the subgradient vector $\nabla_{\text{sub}} f(\delta_1,\delta_2)$ is obtained by solving the dual of the inner problem (4.55). Such a dual problem can be formulated as

$$\underset{\substack{\lambda_1,\lambda_2,\gamma,\boldsymbol{\mu},\\ \boldsymbol{\tau}_1,\boldsymbol{\tau}_2,\chi_1,\chi_2,\\ \boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\nu_1,\nu_2}}{\text{minimize}} \left(\begin{matrix} \delta_1\lambda_1+\delta_2\lambda_2+P_{\text{Tot}}\gamma+\sum_{i=1}^{N}(P_i\mu_i+\tau_{1i}+\tau_{2i}) \\ -\rho\ln\dfrac{\chi_1}{\rho}-(1-\rho)\ln\dfrac{\chi_2}{1-\rho} \end{matrix}\right) - 1 \tag{4.57a}$$

$$\text{s.t.}\;\; \chi_1,\chi_2>0, \quad |\nu_1|\leq\lambda_1, \quad |\nu_2|\leq\lambda_2, \tag{4.57b}$$

$$\|\chi_1\hat{\mathbf{h}}_1-\boldsymbol{\eta}_1-\nu_1\hat{\mathbf{h}}_2\|_2\leq\lambda_1\epsilon_2+\chi_1\epsilon_1, \tag{4.57c}$$

$$\|\chi_2\hat{\mathbf{h}}_2-\boldsymbol{\eta}_2-\nu_2\hat{\mathbf{h}}_1\|_2\leq\lambda_2\epsilon_1+\chi_2\epsilon_2, \tag{4.57d}$$

$$\gamma\geq 0, \quad \mu_i\geq 0, \quad \gamma+\mu_i>0, \qquad i=1,\ldots,N, \tag{4.57e}$$

$$\begin{bmatrix} \tau_{1i} & \eta_{1i} \\ \eta_{1i} & 4(\gamma+\mu_i) \end{bmatrix}\succeq 0, \qquad i=1,\ldots,N, \tag{4.57f}$$

$$\begin{bmatrix} \tau_{2i} & \eta_{2i} \\ \eta_{2i} & 4(\gamma+\mu_i) \end{bmatrix}\succeq 0, \qquad i=1,\ldots,N, \tag{4.57g}$$

where λ_1 and λ_2 are the Lagrange multipliers associated with the constraints (4.54c) and (4.54d), respectively. Derivation of the dual problem (4.57) is provided in Appendix C.2.

4.4.3 Per-Antenna Amplitude Constraint

With amplitude constraints, we use the definitions in (4.49) to obtain the worst-case counterparts of the secrecy rate expressions in (4.20). Furthermore, the inner problem (4.55) is modified to

$$\begin{aligned}
\underset{\mathbf{W},z_1,z_2}{\text{maximize}} \quad & \rho \ln z_1 + (1-\rho)\ln z_2 && \text{(4.58a)}\\
\text{s.t.} \quad & \text{(4.54a)}, \text{(4.54b)}, \text{(4.54c)}, \text{(4.54d)}, && \text{(4.58b)}\\
& |w_{1i}| + |w_{2i}| \leq A_i, \quad i = 1, \dots, N, && \text{(4.58c)}
\end{aligned}$$

and it can be shown that its dual is given by

$$\begin{aligned}
\underset{\substack{\lambda_1,\lambda_2,\boldsymbol{\mu},\chi_1,\chi_2,\\ \boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\nu_1,\nu_2}}{\text{minimize}} \quad & \begin{pmatrix} \delta_1\lambda_1 + \delta_2\lambda_2 + \sum_{i=1}^{N}(A_i\mu_i) \\ -\rho \ln \dfrac{\chi_1}{\rho} - (1-\rho)\ln \dfrac{\chi_2}{1-\rho} \end{pmatrix} - 1 && \text{(4.59a)}\\
\text{s.t.} \quad & \chi_1, \chi_2 > 0, \quad |\nu_1| \leq \lambda_1, \quad |\nu_2| \leq \lambda_2, && \text{(4.59b)}\\
& \|\chi_1\hat{\mathbf{h}}_1 - \boldsymbol{\eta}_1 - \nu_1\hat{\mathbf{h}}_2\|_2 \leq \lambda_1\epsilon_2 + \chi_1\epsilon_1, && \text{(4.59c)}\\
& \|\chi_2\hat{\mathbf{h}}_2 - \boldsymbol{\eta}_2 - \nu_2\hat{\mathbf{h}}_1\|_2 \leq \lambda_2\epsilon_1 + \chi_2\epsilon_2, && \text{(4.59d)}\\
& |\eta_{1i}| \leq \mu_i, \quad |\eta_{2i}| \leq \mu_i, \quad i = 1, \dots, N. && \text{(4.59e)}
\end{aligned}$$

Then, we proceed with the same steps from the previous subsection and use Algorithm 4.1 to obtain the precoder $\mathbf{W}$.

4.5 Numerical Examples

In this section, we provide numerical examples to demonstrate the performance of the proposed linear precoder, in terms of the achievable secrecy rate regions, subject to power or amplitude constraints. We also show the performance of the robust precoder under different channel uncertainty levels along with different constraints.

For simulation purposes, the elements of the channel gain vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ (or $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ for the robust case) are generated randomly (i.i.d. random variables) according to the standard normal distribution $\mathcal{N}(0, 1)$. To obtain the achievable secrecy rate region, we generate 21 points, i.e., secrecy rate pairs (R_1, R_2), on the boundary of the region by solving the weighted secrecy sum rate maximization problem using $\rho = 0, 0.05, 0.10, \ldots, 1.00$. The final results that we plot are obtained by averaging over 1000 realizations of the channel gain vectors. In all cases, the noise variance σ^2 is equal to 1 at both receivers.

We begin with the case of perfect channel information and total power constraint. This case is particularly important as it is the only case for which the secrecy capacity region is precisely known, and the boundary points can be calculated using a closed-form expression. This capacity region sets a benchmark that enables us to quantify the loss incurred by using a suboptimal linear precoding scheme, and also to validate the algorithm we use to obtain the linear precoder.

In Figure 4.1, we plot the secrecy capacity region obtained with the optimal S-DPC scheme [48, Theorem 1], along with the secrecy rate region of the linear precoder proposed in Section 4.3, subject to a total power constraint specified by $P_{\text{dB}} \triangleq 10 \log_{10} P_{\text{Tot}}$. We also include two other conventional linear precoding schemes, namely, the generalized eigenvalue (GEV) and the zero-forcing (ZF) schemes, for comparison purposes.

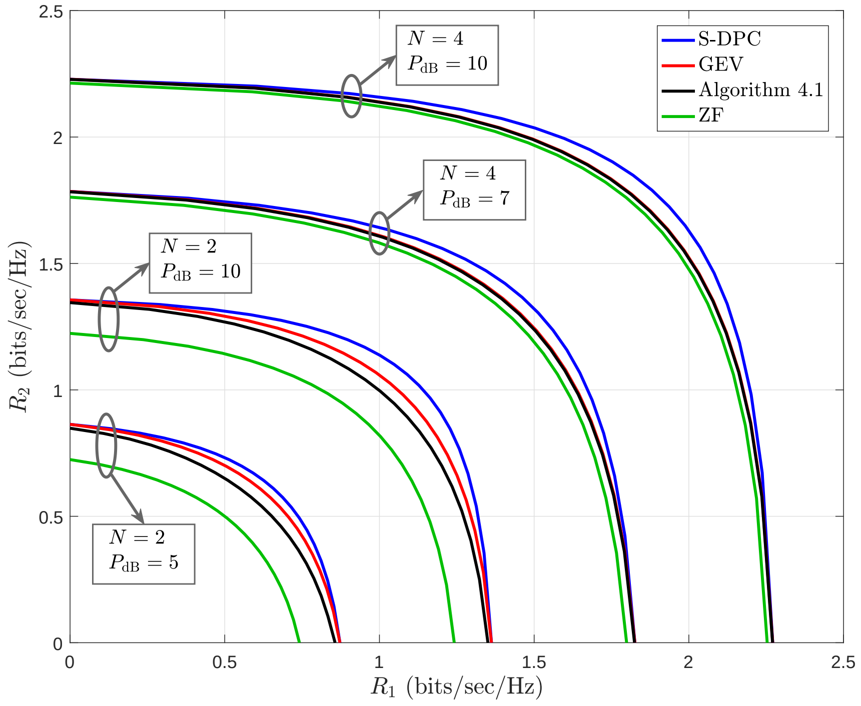

Figure 4.1: The secrecy capacity region obtained with optimal S-DPC along with the secrecy rate regions of the GEV precoder, the precoder obtained with Algorithm 4.1, and the ZF precoder, subject to the total power constraint $P_{\mathrm{dB}} = 10\log_{10} P_{\mathrm{Tot}}$. The number of antennas $N \in \{2, 4\}$.

The proposed linear precoder is obtained using Algorithm 4.1 along with the dual problem (4.37). The dual problem is solved using the CVX toolbox [85] in conjunction with the MOSEK solver [86]. For Algorithm 4.1, we use $\boldsymbol{\delta}^{(1)} = (10^{-1}, 10^{-1})$ or, equivalently, $\boldsymbol{\delta}_{\mathrm{dB}}^{(1)} = (-20 \text{ dB}, -20 \text{ dB})$, as the initial point, and start searching with a fixed step $\alpha_{\mathrm{dB}}^{\mathrm{Fix}} = 1$ dB. The maximum number of iterations after encountering a peak is $L = 10$, i.e., the final solution $\boldsymbol{\delta}_{\mathrm{dB}}^{\star}$ is obtained with accuracy $\alpha_{\mathrm{dB}}^{\mathrm{Fix}}/L = 0.1$ dB. For the GEV precoder, the beamformers $\mathbf{w}_{1,\mathrm{GEV}}$ and $\mathbf{w}_{2,\mathrm{GEV}}$ are obtained as follows. Let $\boldsymbol{v}_1$ be the generalized eigenvector of the matrix pair $(\sigma^2\mathbf{I}_N + P_{\mathrm{Tot}}\mathbf{h}_1\mathbf{h}_1^{\mathrm{T}}, \sigma^2\mathbf{I}_N + P_{\mathrm{Tot}}\mathbf{h}_2\mathbf{h}_2^{\mathrm{T}})$

corresponding to its largest generalized eigenvalue. Then,

$$\mathbf{w}_{1,\mathrm{GEV}} = \sqrt{\rho P_{\mathrm{Tot}}} \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|_2}.$$

Similarly, we have

$$\mathbf{w}_{2,\mathrm{GEV}} = \sqrt{(1-\rho) P_{\mathrm{Tot}}} \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|_2},$$

where $\boldsymbol{v}_2$ is the generalized eigenvector of the matrix pair $(\sigma^2 \mathbf{I}_N + P_{\mathrm{Tot}} \mathbf{h}_2 \mathbf{h}_2^{\mathrm{T}}, \sigma^2 \mathbf{I}_N + P_{\mathrm{Tot}} \mathbf{h}_1 \mathbf{h}_1^{\mathrm{T}})$ corresponding to its largest generalized eigenvalue. The ZF precoder is obtained by solving the inner problem (4.28), without the per-antenna power constraint (4.28d), using $\delta_1 = \delta_2 = 0$. For all three linear precoders, the achievable rate pairs (R_1, R_2) are obtained by substituting with the precoder $\mathbf{W}$ into (4.10).

Several interesting conclusions can be drawn from Figure 4.1. First, we note that the GEV precoder yields slightly better performance than our precoder from Section 4.3, especially at low power levels. This is due to the fact that we use the simplified lower bound in (4.25) as the objective function of the weighted secrecy sum rate maximization problem, rather than the more complex expression in (4.22d). This suggests that the GEV is probably a good linear precoder when there is only a total power constraint (and channel information is accurately known to the transmitter). Note, however, that there is no counterpart of the GEV scheme for cases involving per-antenna power or amplitude constraints. Particularly, unlike the cases in Section 3.5.1.1, where we simply scaled the beamformer $\mathbf{w}_{\mathrm{GEV}}$ to satisfy the l_p-norm constraint for all $p = 1, 2, \infty$, scaling the precoder $\mathbf{W}_{\mathrm{GEV}} \triangleq [\mathbf{w}_{1,\mathrm{GEV}} \ \mathbf{w}_{2,\mathrm{GEV}}]$ to satisfy the per-antenna power or amplitude constraints would significantly deteriorate the performance. Instead, $\mathbf{w}_{1,\mathrm{GEV}}$ and $\mathbf{w}_{2,\mathrm{GEV}}$ should be scaled by different factors, however it is unclear how to choose these factors in an optimal way. We also note

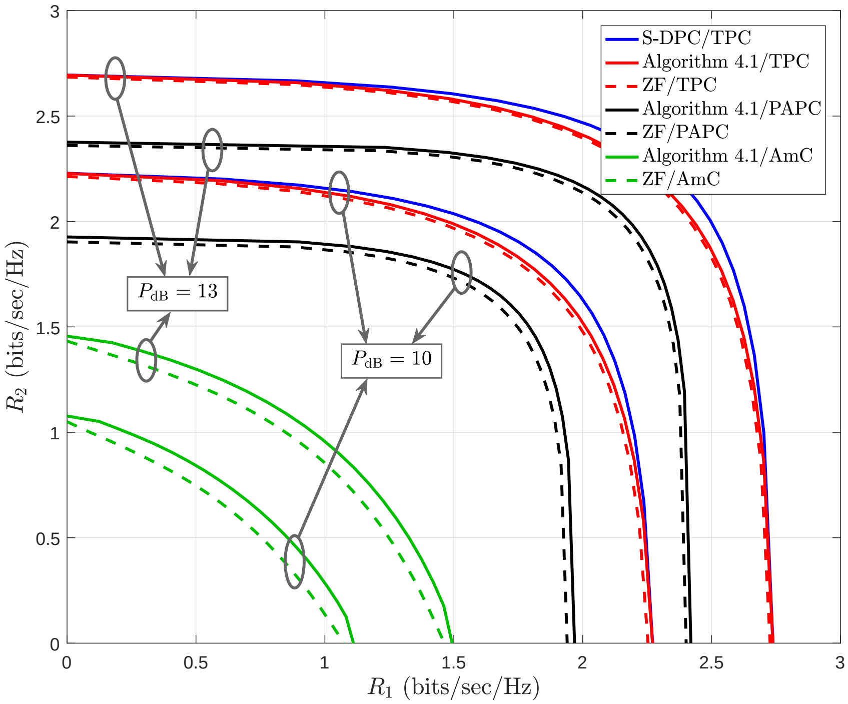

Figure 4.2: Achievable secrecy rate regions of the proposed and ZF precoders subject to total power constraint (TPC), per-antenna power constraint (PAPC), and amplitude constraint (AmC). $P_{\text{Tot}} = NP_i = NA_i^2, i = 1, \ldots, N$, $P_{\text{dB}} = 10\log_{10} P_{\text{Tot}}$, and $N = 4$. The secrecy capacity region with optimal S-DPC is included for the case of total power constraint.

from Figure 4.1 that the ZF precoder has the worst performance among all other precoders at all power levels. Performance gaps, however, significantly decrease as the number of antennas or transmit power increases.

In Figure 4.2, we show the achievable secrecy rate regions of the proposed linear precoder, subject to the total power constraint (4.8b), the per-antenna power constraint (4.12), and the per-antenna amplitude constraint (4.14b). The secrecy capacity region obtained with optimal S-DPC (for the case of total power constraint) and the secrecy rate regions of the ZF precoder (for all constraints) are also included. The power level indicated in the figure specifies the total power constraint in dB, i.e., $P_{\text{dB}} = 10\log_{10} P_{\text{Tot}}$. For comparison purposes, we choose the per-antenna power

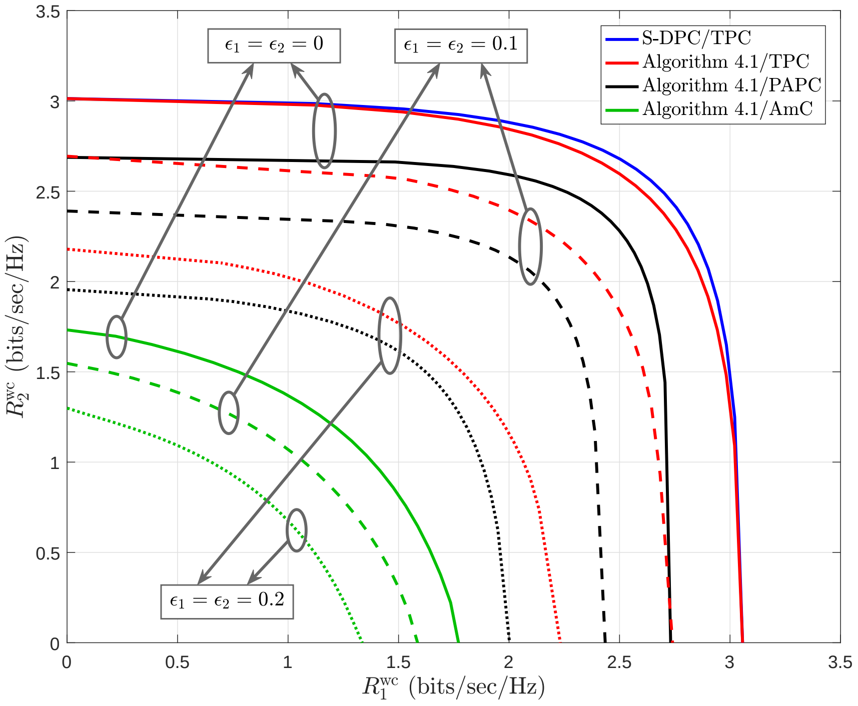

Figure 4.3: Worst-case secrecy rate regions with different channel uncertainty levels, ϵ_1 and ϵ_2, subject to total power constraint (TPC), per-antenna power constraint (PAPC), and amplitude constraint (AmC). $P_{\mathrm{Tot}} = NP_i = NA_i^2, i = 1, \ldots, N$, $P_{\mathrm{dB}} = 10 \log_{10} P_{\mathrm{Tot}} = 15$ dB, and $N = 4$.

constraint as $P_i = P_{\mathrm{Tot}}/N$, and the amplitude constraint as $A_i = \sqrt{P_{\mathrm{Tot}}/N}$, for $i = 1, \ldots, N$. Thus, the per-antenna power constraint also implies the total power constraint, and the amplitude constraint implies the total and per-antenna power constraints. The number of antennas N is set to 4. As expected, the proposed linear precoder outperforms the ZF precoder, under all constraints, however at the cost of increased computational complexity.

Finally, in Figure 4.3, we plot the worst-case secrecy rate regions obtained with the robust precoder considered in Section 4.4, subject to (4.8b), (4.12), and (4.14b), separately. Similar to the previous example, we choose $P_i = P_{\mathrm{Tot}}/N$ and $A_i = \sqrt{P_{\mathrm{Tot}}/N}$, for all $i = 1, \ldots, N$, where $10 \log_{10} P_{\mathrm{Tot}} = 15$ dB and $N = 4$. The case $\epsilon_1 = \epsilon_2 = 0$ designates perfect channel information, and is included for comparison purposes. As expected, we note from Figure 4.3 that increased uncertainty levels have negative impact on the worst-case secrecy rate region.

4.6 Conclusions

In this chapter, we considered the design of linear precoders for the two-user MISO BC-CM subject to total and per-antenna power constraints, and also subject to amplitude constraints. Per-antenna constraints are typically more difficult to handle, but they are essential for modelling hardware limitations in practical systems employing multiple transmit antennas. Although suboptimal, linear precoding is particularly attractive because of low implementation complexity. On the other hand, the optimal S-DPC scheme is difficult to implement, and can be only found via an exhaustive search when per-antenna power constraints are taken into account. Furthermore, the optimal scheme is unknown under amplitude constraints. Therefore, our proposed linear precoding scheme provides a viable solution to an open problem that has not been addressed in the published literature.

We formulated the linear precoder design problem as a weighted secrecy sum rate maximization problem that is transformed into a more tractable problem having only two optimization variables. We proposed a subgradient-based search algorithm to obtain a solution, and provided a condition under which the obtained solution is guaranteed to be optimal. Our approach is applicable to any combination of the total power, per-antenna power, and per-antenna amplitude constraints. It is also applicable to the robust design problem when channel uncertainty is taken into account.

We used the total power constraint case, in which the secrecy capacity region is precisely known, to validate our approach and compare the performance of the linear precoder with the optimal S-DPC scheme. Numerical results show negligible loss when the SNR is sufficiently high. Compared to the idealistic case of total power constraint and perfect channel information, the results show considerable reduction in the achievable secrecy rate region when per-antenna constraints and channel uncertainty are taken into account.

Chapter 5

Conclusions and Future Directions

5.1 Conclusions

Physical-layer security has the potential to complement existing encryption techniques with an additional secrecy measure that is provably unbreakable regardless of the computational power of the eavesdropper. It can also be a viable lightweight secrecy solution under severe hardware or energy constraints. While physical-layer security has been an active research area for more than a decade, it still has not got much attention from practical system designers. Perhaps the main reason for such a disregard is performance sensitivity to channel conditions. Particularly, the performance of physical-layer security schemes can be severely degraded, and secrecy outage may occur, if the design is based on inaccurate channel information.

In this thesis, we proposed the use of physical-layer security techniques to enhance the secrecy of visible-light communication (VLC) systems. We had a twofold purpose from such a proposal. First, the broadcast nature of the VLC channel makes an additional secrecy layer a sensible approach. Second, VLC links can be a reasonable platform for the deployment of physical-layer security prototypes as realistic assumptions regarding channel information can be made in typical VLC scenarios with dominant LoS path. Furthermore, by adopting robust transmission schemes that take channel uncertainty into account, performance sensitivity to channel estimation errors can be significantly alleviated. Although in this thesis we mainly focused on VLC systems functioning in indoor environments, the techniques we developed are also applicable to outdoor scenarios.

Existing physical-layer security schemes assume Gaussian input distribution and total transmit power constraint, making them inapplicable to VLC channels wherein amplitude constraints on the channel input are inherent due to linearity limitations of typical LEDs.

Accordingly, in this thesis, we studied the design of physical-layer security schemes for the Gaussian wiretap channel subject to amplitude constraints. Three major contributions have been presented:

Firstly, with the lack of closed-form secrecy capacity expressions for the amplitude-constrained Gaussian wiretap channel, we utilized the maximum-entropy uniform input distribution, along with the entropy power inequality, to establish a closed-form lower bound on the secrecy capacity. We also developed a method to derive an upper bound on the secrecy capacity of degraded wiretap channels, and applied that method to the scalar Gaussian wiretap channel. Then, we used the lower bound along with beamforming to obtain a closed-form secrecy rate expression for the MISO wiretap channel. We later used that expression as a performance metric for the beamformer design. We also derived a closed-form secrecy rate expression for the amplitude-constrained scalar wiretap channel when it is aided by a friendly jammer sending artificial noise that is also subject to amplitude constraints.

Secondly, we studied the design of beamformers for the MISO wiretap channel subject to amplitude constraints. Unlike the case of total power constraint, which is readily solvable as a Rayleigh quotient maximization problem, the design problem under the amplitude constraint is more difficult to solve. Nevertheless, we transformed such a difficult problem into a quasiconvex line search problem that is easily solved with a bisection search. In addition, we showed that our solution technique is applicable to arbitrary l_p-norm constraints on the beamformer. We also solved the worst-case secrecy rate maximization problem when channel uncertainties for the receiver and eavesdropper are taken into account.

Thirdly, we studied the design of linear precoders for the two-user MISO broadcast channel with confidential messages (BC-CM). We developed a general approach that can handle the design problem subject to any combination of the total, per-antenna power, or amplitude constraints. Although suboptimal, our linear precoding scheme entails low implementation complexity. Furthermore, it provides a viable solution to the cases of per-antenna power or amplitude constraints where there is no closed-form characterization of the secrecy capacity region. We formulated the design problem as a weighted secrecy sum rate maximization problem, then we transformed the problem into a more tractable form that can be solved with an iterative search algorithm. We used the case of total power constraint to quantify the performance loss incurred by using a suboptimal linear precoding scheme and also to validate our approach to solving the design problem. We also considered the design of robust linear precoders to maximize the worst-case secrecy rate region when channel uncertainty is taken into account.

The numerical results revealed considerable decline in the achievable secrecy rates when channel uncertainty and amplitude constraints are taken into account as compared to the idealistic case of perfect channel information and total power constraint. Therefore, the design techniques we developed throughout the thesis provide valuable tools for tackling real-world problems in which channel uncertainty is almost always inevitable and per-antenna constraints are essential for accurate modelling of hardware limitations.

Finally, it is worth mentioning that physical-layer security is a research area that has its origins from information theory. On the other hand, the design of VLC systems is mostly treated as a practical engineering problem. By combining these two research areas in one thesis, we aim to serve both communities and help narrow the gap between information theory and practical system design. For example, amplitude constraints impose an inherent practical limitation that is difficult to treat mathematically, however we derived lower bounds on the secrecy capacity, subject to these constraints, in order to circumvent such a difficulty. Furthermore, by using realistic channel gain models from VLC scenarios and linking the physical sources of channel uncertainty (e.g., location or orientation uncertainty) with the uncertainty sets used in robust optimization problems, we help make the techniques from information theory and convex optimization more approachable to practical system designers.

5.2 Future Work

Secure Transmission with Discrete Input Distribution:

In Chapter 2, we derived closed-form secrecy rate expressions for the amplitude-constrained Gaussian wiretap channel based on the (continuous) uniform input distribution. Then, we used the resulting expression for beamformer design in the MISO wiretap channel. Similarly, we used a secrecy sum rate expression based on the uniform input distribution for linear precoder design in the two-user MISO BC-CM under amplitude constraints. We note, however, that the codewords or signals transmitted in practical communication systems cannot have a continuous distribution. Instead, they must be drawn from a discrete constellation, i.e., a discrete distribution with finite support. One practical reason for such a limitation is the finite resolution of the digital-to-analog converters (DACs) incorporated at the transmitter front-end. Although we know that the optimal input distribution for the amplitude-constrained scalar wiretap channel is discrete (and we conjecture that this is also the case for the MISO wiretap channel and the two-user MISO BC-CM), there is no closed-form characterization of these optimal distributions and they can only be found via numerical techniques. Since closed-form expressions are required for beamformer or precoder design, an interesting research direction is to find secrecy capacity-approaching discrete input distributions which yield closed-form secrecy rate expressions that are amenable to optimization.

Linear Precoding for the MIMO Wiretap Channel under Amplitude Constraints:

In Chapter 3, we considered the design of beamformers for the MISO wiretap channel under amplitude constraints. Our approach in Propositions 3.1 and 3.2 took advantage of the fact that only one data stream is being transmitted (since the intended receiver has one antenna), and the signal term in the numerator of the secrecy rate expression (3.5) is a *squared linear function* of the beamformer. This ultimately led to convex formulations of the inner problems (3.11) and (3.33). If the intended receiver, however, has multiple antennas, then simultaneous transmission of multiple data streams should be considered. Consequently, equalization may become necessary or desirable at the receiver as well as the eavesdropper. The resulting secrecy rate expression, however, will become more difficult to handle, and the inner problems corresponding to (3.11) and (3.33) will no longer be convex. Thus, a natural extension to the work presented in Chapter 3 is to consider linear precoding for the MIMO wiretap channel subject to amplitude constrains.

Similar remarks can be made about the design of linear precoders for the two-user MIMO BC-CM when multiple data streams are simultaneously transmitted to each user.

Linear Precoding for the Complex-Valued MISO Wiretap Channel under Amplitude Constraints:

Throughout the entire thesis, we assumed real-valued transmitted signals and channel gain vectors. Such an assumption is applicable to intensity modulation (IM) systems, as well as RF systems utilizing only amplitude modulation, i.e., the carrier phase is not modulated. Extension to the more general case of complex-valued transmission, such as quadrature amplitude modulation (QAM), shall make the design problems considered in Chapters 3 and 4 more difficult to handle. For example, with complex-valued channel gain and beamforming vectors, the inner problem (3.11), as well as its robust counterpart (3.33), would involve maximization of the magnitude of a complex-valued quantity. Obviously, this is a nonconvex problem, and thus the techniques we developed via Propositions 3.1 and 3.2 will have to be modified in order to deal with nonconvexity of the inner problem.

Finally, the problem of deriving achievable secrecy rate expressions for the complex-valued Gaussian wiretap channel subject to amplitude constraints can also be of great interest. Among various feasible input distributions that can be utilized, the circular uniform distribution sounds like a good candidate to begin with.

Bibliography

[1] Z. Ghassemlooy, S. Arnon, M. Uysal, Z. Xu, and J. Cheng, "Emerging optical wireless communications – Advances and challenges," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 9, pp. 1738–1749, Sept. 2015.

[2] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: A survey, potential and challenges," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 4, pp. 2047–2077, 4th quarter 2015.

[3] D. Karunatilaka, F. Zafar, V. Kalavally, and R. Parthiban, "LED based indoor visible light communications: State of the art," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 3, pp. 1649–1678, 3rd quarter 2015.

[4] A. Sevincer, A. Bhattarai, M. Bilgi, M. Yuksel, and N. Pala, "LIGHTNETs: Smart LIGHTing and mobile optical wireless NETworks – A survey," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 4, pp. 1620–1641, 4th quarter 2013.

[5] Z. Ghassemlooy, W. Popoola, and S. Rajbhandari, "Visible light communications," in *Optical Wireless Communications: System and Channel Modelling with MATLAB®*. CRC Press, 2013, pp. 443–496.

[6] K. Tae-Gyu, "Visible-light communications," in *Advanced Optical Wireless Communication Systems*, S. Arnon, J. R. Barry, G. K. Karagiannidis, R. Schober, and M. Uysal, Eds. Cambridge University Press, 2012, pp. 351–368.

[7] H. Elgala, R. Mesleh, and H. Haas, "Indoor optical wireless communication: Potential and state-of-the-art," *IEEE Communications Magazine*, vol. 49, no. 9, pp. 56–62, Sept. 2011.

[8] A. T. Hussein and J. M. H. Elmirghani, "10 Gbps mobile visible light communication system employing angle diversity, imaging receivers, and relay nodes," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 7, no. 8, pp. 718–735, Aug. 2015.

[9] L. Zeng, D. O'Brien, H. Minh, G. Faulkner, K. Lee, D. Jung, Y. Oh, and E. T. Won, "High data rate multiple input multiple output (MIMO) optical wireless communications using white LED lighting," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 9, pp. 1654–1662, Dec. 2009.

[10] R. Lenk and C. Lenk, *Practical Lighting Design with LEDs.* Wiley, 2011.

[11] T. Komine and M. Nakagawa, "Integrated system of white LED visible-light communication and power-line communication," *IEEE Transactions on Consumer Electronics*, vol. 49, no. 1, pp. 71–79, Feb. 2003.

[12] H. Ma, L. Lampe, and S. Hranilovic, "Integration of indoor visible light and power line communication systems," in *2013 IEEE 17th International Symposium on Power Line Communications and Its Applications*, Mar. 2013, pp. 291–296.

[13] T. Komine, J. H. Lee, S. Haruyama, and M. Nakagawa, "Adaptive equalization system for visible light wireless communication utilizing multiple white LED lighting equipment," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 2892–2900, Jun. 2009.

[14] A. Mostafa and L. Lampe, "Physical-layer security for MISO visible light communication channels," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 9, pp. 1806–1818, Sept. 2015.

[15] H. Haas, L. Yin, Y. Wang, and C. Chen, "What is LiFi?" *Journal of Lightwave Technology*, vol. 34, no. 6, pp. 1533–1544, Mar. 2016.

[16] S. Wu, H. Wang, and C. H. Youn, "Visible light communications for 5G wireless networking systems: From fixed to mobile communications," *IEEE Network*, vol. 28, no. 6, pp. 41–45, Nov. 2014.

[17] "IEEE Standard for Local and Metropolitan Area Networks – Part 15.7: Short-Range Wireless Optical Communication Using Visible Light," *IEEE Std 802.15.7-2011*, pp. 1–309, 2011.

[18] S. Hranilovic, L. Lampe, and S. Hosur, "Visible light communications: The road to standardization and commercialization (Part 1) [guest editorial]," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 24–25, Dec. 2013.

[19] S. Hranilovic, L. Lampe, S. Hosur, and R. Roberts, "Visible light communications: The road to standardization and commercialization (Part 2) [guest editorial]," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 62–63, Jul. 2014.

[20] R. Roberts, S. Rajagopal, and S.-K. Lim, "IEEE 802.15.7 physical layer summary," in *2011 IEEE GLOBECOM Workshops (GC Wkshps)*, Dec. 2011, pp. 772–776.

[21] S. Rajagopal, R. Roberts, and S.-K. Lim, "IEEE 802.15.7 visible light communication: Modulation schemes and dimming support," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 72–82, Mar. 2012.

[22] L. Grobe, A. Paraskevopoulos, J. Hilt, D. Schulz, F. Lassak, F. Hartlieb, C. Kottke, V. Jungnickel, and K.-D. Langer, "High-speed visible light communication systems," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 60–66, Dec. 2013.

[23] C. Yang, Y. Wang, Y. Wang, X. Huang, and N. Chi, "Demonstration of high-speed multi-user multi-carrier CDMA visible light communication," *Optics Communications*, vol. 336, pp. 269 – 272, 2015.

[24] D. Tsonev, H. Chun, S. Rajbhandari, J. J. D. McKendry, S. Videv, E. Gu, M. Haji, S. Watson, A. E. Kelly, G. Faulkner, M. D. Dawson, H. Haas, and D. O'Brien, "A 3-Gb/s single-LED OFDM-based wireless VLC link using a gallium nitride μLED," *IEEE Photonics Technology Letters*, vol. 26, no. 7, pp. 637–640, Apr. 2014.

[25] Y. Liang, H. V. Poor, and S. Shamai (Shitz), *Information Theoretic Security*, ser. Foundations and Trends® in Communications and Information Theory. Now Publishers, 2008, vol. 5, no. 4-5.

[26] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, Oct. 2011.

[27] R. Liu and W. Trappe, Eds., *Securing Wireless Communications at the Physical Layer*. Springer, 2010.

[28] H. Wen, *Physical Layer Approaches for Securing Wireless Communication Systems*, ser. SpringerBriefs in Computer Science. Springer, 2013.

[29] X. Zhou, L. Song, and Y. Zhang, Eds., *Physical Layer Security in Wireless Communications*. CRC Press, 2013.

[30] Y.-W. P. Hong, P.-C. Lan, and C.-C. J. Kuo, *Signal Processing Approaches to Secure Physical Layer Communications in Multi-Antenna Wireless Systems*, ser. SpringerBriefs in Electrical and Computer Engineering. Springer, 2014.

[31] A. Mukherjee, S. Fakoorian, J. Huang, and A. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 3, pp. 1550–1573, 3rd quarter 2014.

[32] A. Mukherjee, "Physical-layer security in the Internet of Things: Sensing and communication confidentiality under resource constraints," *Proceedings of the IEEE*, vol. 103, no. 10, pp. 1747–1761, Oct. 2015.

[33] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[34] A. D. Wyner, “The wire-tap channel,” *Bell System Technical Journal*, vol. 54, pp. 1355–1387, 1975.

[35] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley, 2006.

[36] S. Leung-Yan-Cheong and M. Hellman, “The Gaussian wire-tap channel,” *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451–456, Jul. 1978.

[37] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[38] S. Shafiee and S. Ulukus, “Achievable rates in Gaussian MISO channels with secrecy constraints,” in *IEEE International Symposium on Information Theory*, Jun. 2007, pp. 2466–2470.

[39] A. Khisti and G. W. Wornell, “Secure transmission with multiple antennas I: The MISOME wiretap channel,” *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[40] S. Shafiee, N. Liu, and S. Ulukus, “Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel,” *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4033–4039, Sept. 2009.

[41] T. Liu and S. Shamai, “A note on the secrecy capacity of the multiple-antenna wiretap channel,” *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[42] A. Khisti and G. W. Wornell, “Secure transmission with multiple antennas – Part II: The MIMOME wiretap channel,” *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[43] F. Oggier and B. Hassibi, “The secrecy capacity of the MIMO wiretap channel,” *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[44] S. Goel and R. Negi, “Guaranteeing secrecy using artificial noise,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[45] X. Tang, R. Liu, P. Spasojević, and H. Poor, “Interference-assisted secret communication,” in *2008 IEEE Information Theory Workshop (ITW)*, May 2008, pp. 164–168.

[46] A. L. Swindlehurst, “Fixed SINR solutions for the MIMO wiretap channel,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2009, pp. 2437–2440.

[47] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[48] R. Liu and H. Poor, "Secrecy capacity region of a multiple-antenna Gaussian broadcast channel with confidential messages," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 1235–1249, Mar. 2009.

[49] R. Liu, T. Liu, H. V. Poor, and S. Shamai, "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4215–4227, Sept. 2010.

[50] H. Elgala, R. Mesleh, and H. Haas, "Predistortion in optical wireless transmission using OFDM," in *Ninth International Conference on Hybrid Intelligent Systems*, vol. 2, Aug. 2009, pp. 184–189.

[51] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Journal of Information and Control*, vol. 18, pp. 203–219, 1971.

[52] S. Hranilovic and F. Kschischang, "Optical intensity-modulated direct detection channels: Signal space and lattice codes," *IEEE Transactions on Information Theory*, vol. 49, no. 6, pp. 1385–1399, Jun. 2003.

[53] A. Lapidoth, S. Moser, and M. Wigger, "On the capacity of free-space optical intensity channels," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4449–4461, Oct. 2009.

[54] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[55] J. G. Smith, *On the Information Capacity of Peak and Average Power Constrained Gaussian Channels.* Ph.D. dissertation, Department of Electrical Engineering, University of California, Berkeley, California, 1969.

[56] O. Ozel, E. Ekrem, and S. Ulukus, "Gaussian wiretap channel with amplitude and variance constraints," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5553–5563, Oct. 2015.

[57] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[58] J. Kahn and J. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.

[59] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using LED lights," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 100–107, Feb. 2004.

[60] J. Grubor, S. Randel, K. D. Langer, and J. W. Walewski, "Broadband information broadcasting using LED-based interior lighting," *Journal of Lightwave Technology*, vol. 26, no. 24, pp. 3883–3892, Dec. 2008.

[61] V. Jungnickel, V. Pohl, S. Nönnig, and C. von Helmolt, "A physical model of the wireless infrared communication channel," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 631–640, Apr. 2002.

[62] A. Mostafa and L. Lampe, "Physical-layer security for indoor visible light communications," in *2014 IEEE International Conference on Communications (ICC)*, Jun. 2014, pp. 3342–3347.

[63] A. Mostafa and L. Lampe, "Securing visible light communications via friendly jamming," in *2014 Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 524–529.

[64] A. Mostafa and L. Lampe, "Enhancing the security of VLC links: Physical-layer approaches," in *2015 IEEE Summer Topicals Meeting Series (SUM)*, Jul. 2015, pp. 39–40.

[65] A. Mostafa and L. Lampe, "Pattern synthesis of massive LED arrays for secure visible light communication links," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, Jun. 2015, pp. 1350–1355.

[66] A. Mostafa and L. Lampe, "Optimal and robust beamforming for secure transmission in MISO visible-light communication links," *IEEE Transactions on Signal Processing*, vol. 64, no. 24, pp. 6501–6516, Dec. 2016.

[67] A. Mostafa and L. Lampe, "On linear precoding for the two-user MISO broadcast channel with confidential messages and per-antenna constraints," *Submitted for publication*, pp. 1–13, Jan. 2017.

[68] A. Lapidoth and S. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.

[69] S. M. Moser, *Duality-Based Bounds on Channel Capacity.* Ph.D. dissertation, Swiss Federal Institute of Technology, Zürich, 2004.

[70] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Cambridge University Press, 2011.

[71] R. G. Gallager, *Information Theory and Reliable Communication.* John Wiley & Sons, 1968.

[72] Z. Li, W. Trappe, and R. Yates, "Secret communication via multi-antenna transmission," in *41st Annual Conference on Information Sciences and Systems*, Mar. 2007, pp. 905–910.

[73] L. Zhang, Y. C. Liang, Y. Pei, and R. Zhang, "Robust beamforming design: From cognitive radio MISO channels to secrecy MISO channels," in *IEEE Global Telecommunications Conference, 2009*, Nov. 2009, pp. 1–5.

[74] W. Shi and J. Ritcey, "Robust beamforming for MISO wiretap channel by optimizing the worst-case secrecy capacity," in *Forty Fourth Asilomar Conference on Signals, Systems and Computers*, Nov. 2010, pp. 300–304.

[75] Q. Li and W. K. Ma, "Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3799–3812, Aug. 2011.

[76] J. Huang and A. L. Swindlehurst, "Robust secure transmission in MISO channels based on worst-case optimization," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1696–1707, Apr. 2012.

[77] K. Cumanan, Z. Ding, B. Sharif, G. Y. Tian, and K. K. Leung, "Secrecy rate optimizations for a MIMO secrecy channel with a multiple-antenna eavesdropper," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 4, pp. 1678–1690, May 2014.

[78] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. Y. L. Goff, "Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 1833–1847, May 2015.

[79] A. Ben-Tal and A. Nemirovski, "Robust convex optimization," *Mathematics of Operations Research*, vol. 23, no. 4, pp. 769–805, Nov. 1998.

[80] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, 2009.

[81] J. V. Tiel, *Convex Analysis: An Introductory Text.* John Wiley & Sons, 1984.

[82] R. T. Rockafellar, *Convex Analysis.* Princeton University Press, 1970.

[83] A. Ben-Tal and A. Nemirovski, "Robust solutions to uncertain linear programs," *Operations Research Letters*, vol. 25, pp. 1–13, 1999.

[84] A. Ben-Tal and A. Nemirovski, "Robust optimization – Methodology and applications," *Mathematical Programming, Series B*, vol. 92, pp. 453–480, 2002.

[85] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[86] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28).*, 2015. [Online]. Available: http://docs.mosek.com/7.1/toolbox/index.html

[87] F. J. López-Hernández, R. Pérez-Jiménez, and A. Santamaría, "Ray-tracing algorithms for fast calculation of the channel impulse response on diffuse IR wireless indoor channels," *Optical Engineering*, vol. 30, no. 10, pp. 2775–2780, Oct. 2000.

[88] S. A. A. Fakoorian and A. L. Swindlehurst, "On the optimality of linear precoding for secrecy in the MIMO broadcast channel," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1701–1713, Sept. 2013.

[89] D. Park, "Weighted sum rate maximization of MIMO broadcast and interference channels with confidential messages," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 1742–1753, Mar. 2016.

[90] N. Z. Shor, *Minimization Methods for Non-differentiable Functions.* Springer Series in Computational Mathematics. Springer, 1985.

[91] S. Boyd, L. Xiao, and A. Mutapcic, *Subgradient methods.* Notes for EE392o, Stanford University, Autumn, 2003.

[92] G. Calafiore and L. El Ghaoui, *Optimization Models.* Cambridge University Press, 2014.

[93] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis.* Springer, 2009.

Appendix A

The Trapezoidal Distribution

In this appendix, we introduce the trapezoidal distribution encountered in the proof of Proposition 2.5, and provide the associated PDF and differential entropy.

The trapezoidal distribution arises from adding two independent random variables having uniform distributions. Particularly, let a and b be two positive real numbers, and X and Y be two independent random variables having the distributions $p_X(x) = \mathcal{U}[-a, a]$ and $p_Y(y) = \mathcal{U}[-b, b]$, respectively. Then, the PDF of the sum $Z \triangleq X + Y$ is obtained by the convolution

$$\begin{aligned} p_Z(z) &= (p_X * p_Y)(z) \\ &= \begin{cases} \dfrac{z + a + b}{4ab} & -a - b \leq z \leq -|a - b|, \\ \min\left\{\dfrac{1}{2a}, \dfrac{1}{2b}\right\} & -|a - b| \leq z \leq |a - b|, \\ \dfrac{-z + a + b}{4ab} & |a - b| \leq z \leq a + b, \\ 0 & \text{otherwise}, \end{cases} \end{aligned} \tag{A.1}$$

as depicted in Figure A.1 for the case $b < a$. For obvious reasons, the distribution $p_Z(z)$ is referred to as the *trapezoidal distribution.* Note that the uniform distribution $\mathcal{U}[-a, a]$ is a special case of the trapezoidal distribution (A.1) when $b \to 0$.

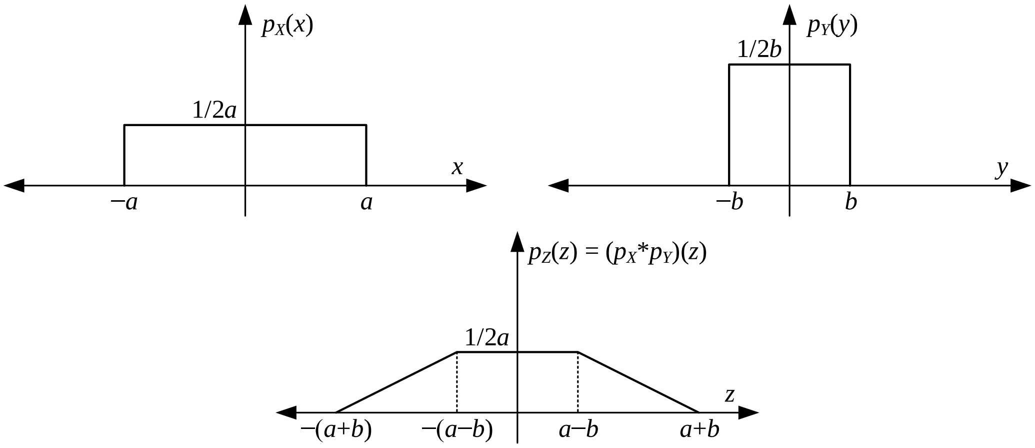

Figure A.1: The trapezoidal distribution in (A.1) with $b < a$.

The differential entropy of Z (in nats) is

$$\begin{aligned} \mathbb{h}(Z) &= -\int p_Z(z) \ln p_Z(z) dz \\ &= \begin{cases} \ln(2a) + \dfrac{b}{2a} & b \le a \\ \ln(2b) + \dfrac{a}{2b} & \text{otherwise} \end{cases} \\ &= \min\left\{ \ln(2a) + \frac{b}{2a}\ ,\ \ln(2b) + \frac{a}{2b} \right\}. \end{aligned} \tag{A.2}$$

In Figure A.2, we plot the differential entropy (A.2) as a function of b when $a = 2$. Note that $\mathbb{h}(Z)$ is differentiable for all $b > 0$, including the point $b = a$.

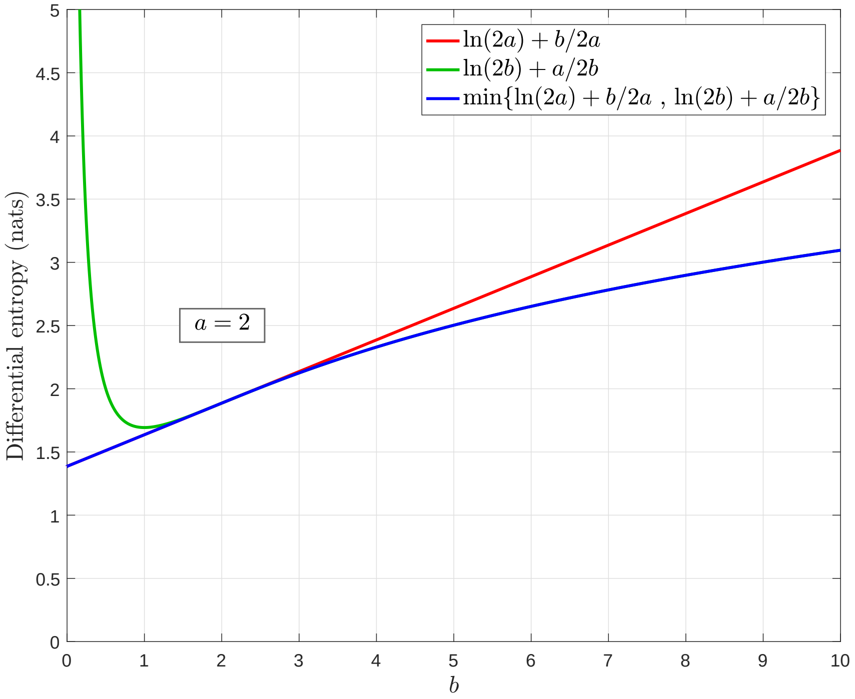

Figure A.2: The differential entropy (A.2) as a function of b when $a = 2$.

Appendix B

Proofs and Derivations for Chapter 3

B.1 Proof of Lemma 3.1

Let the pair $(\mathbf{w}^\star, t^\star)$ be an optimal solution of the nonconvex perturbed problem in (3.36), where $t^\star \equiv \varphi(\varepsilon)$. If $\mathcal{H}_\text{B}$ is convex, then the linear function $f_{\mathbf{w}^\star}(\mathbf{h}_\text{B}) \triangleq \mathbf{h}_\text{B}^\text{T}\mathbf{w}^\star$ maps $\mathcal{H}_\text{B}$ into an interval with three possible outcomes:

i) If $\mathbf{h}_\text{B}^\text{T}\mathbf{w}^\star \geq 0$ for all $\mathbf{h}_\text{B} \in \mathcal{H}_\text{B}$, then $t^\star \geq 0$.

ii) If $\mathbf{h}_\text{B}^\text{T}\mathbf{w}^\star \leq 0$ for all $\mathbf{h}_\text{B} \in \mathcal{H}_\text{B}$, then $t^\star \geq 0$. This also implies that $-\mathbf{h}_\text{B}^\text{T}\mathbf{w}^\star \geq 0$ for all $\mathbf{h}_\text{B} \in \mathcal{H}_\text{B}$. Note that if $(\mathbf{w}^\star, t^\star)$ is a solution to (3.36), then $(-\mathbf{w}^\star, t^\star)$ is also a solution.

iii) If there exist $\mathbf{h}_1, \mathbf{h}_2 \in \mathcal{H}_\text{B}$ such that $\mathbf{h}_1^\text{T}\mathbf{w}^\star > 0 > \mathbf{h}_2^\text{T}\mathbf{w}^\star$, then $t^\star = 0$.

From the above cases, we see that $\mathbf{h}_\text{B}^\text{T}\mathbf{w}^\star \geq 0$, or $\mathbf{h}_\text{B}^\text{T}\mathbf{w}^\star \leq 0$, for all $\mathbf{h}_\text{B} \in \mathcal{H}_\text{B}$, is a necessary condition to obtain nonzero $t^\star$. Thus, we lose nothing by imposing the constraint in (3.37), provided that $\mathcal{H}_\text{B}$ is a convex set. ■

B.2 Components of $\mathbf{h}_0$ and $\mathbf{J}_0$ from (3.53)

From (3.51), for $i = 1, \ldots, N$, $c_i \neq 0$, we have

$$\frac{1}{c_i} h_i(\boldsymbol{\delta}) = \frac{(d_z - \delta_z)^m (\mathbf{d}_i - \boldsymbol{\delta})^\text{T}\mathbf{u}}{\|\mathbf{d}_i - \boldsymbol{\delta}\|_2^{m+3}}, \tag{B.1a}$$

$$\frac{1}{c_i}\frac{\partial h_i(\boldsymbol{\delta})}{\partial \delta_x} = \frac{-(d_z-\delta_z)^m \mathbf{e}_1^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i-\boldsymbol{\delta}\|_2^{m+3}} + \frac{(m+3)(d_{x,i}-\delta_x)(d_z-\delta_z)^m(\mathbf{d}_i-\boldsymbol{\delta})^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i-\boldsymbol{\delta}\|_2^{m+5}}, \tag{B.1b}$$

$$\frac{1}{c_i}\frac{\partial h_i(\boldsymbol{\delta})}{\partial \delta_y} = \frac{-(d_z-\delta_z)^m \mathbf{e}_2^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i-\boldsymbol{\delta}\|_2^{m+3}} + \frac{(m+3)(d_{y,i}-\delta_y)(d_z-\delta_z)^m(\mathbf{d}_i-\boldsymbol{\delta})^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i-\boldsymbol{\delta}\|_2^{m+5}}, \tag{B.1c}$$

$$\begin{aligned}\frac{1}{c_i}\frac{\partial h_i(\boldsymbol{\delta})}{\partial \delta_z} &= \frac{-m(d_z-\delta_z)^{m-1}(\mathbf{d}_i-\boldsymbol{\delta})^{\mathrm{T}}\mathbf{u} - (d_z-\delta_z)^m \mathbf{e}_3^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i-\boldsymbol{\delta}\|_2^{m+3}} \\ &\quad + \frac{(m+3)(d_z-\delta_z)^{m+1}(\mathbf{d}_i-\boldsymbol{\delta})^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i-\boldsymbol{\delta}\|_2^{m+5}},\end{aligned} \tag{B.1d}$$

where $\mathbf{e}_j, j = 1, 2, 3$, is the jth column of the identity matrix $\mathbf{I}_3$. Substituting with $\boldsymbol{\delta} = \mathbf{0}$ back into (B.1), we obtain

$$h_i(\mathbf{0}) = c_i \frac{d_z^m \mathbf{d}_i^{\mathrm{T}}\mathbf{u}}{\|\mathbf{d}_i\|_2^{m+3}}, \tag{B.2a}$$

$$\frac{\partial h_i(\mathbf{0})}{\partial \delta_x} = c_i \left(\frac{-d_z^m \mathbf{e}_1^{\mathrm{T}}}{\|\mathbf{d}_i\|_2^{m+3}} + \frac{(m+3)d_{x,i} d_z^m \mathbf{d}_i^{\mathrm{T}}}{\|\mathbf{d}_i\|_2^{m+5}} \right)\mathbf{u}, \tag{B.2b}$$

$$\frac{\partial h_i(\mathbf{0})}{\partial \delta_y} = c_i \left(\frac{-d_z^m \mathbf{e}_2^{\mathrm{T}}}{\|\mathbf{d}_i\|_2^{m+3}} + \frac{(m+3)d_{y,i} d_z^m \mathbf{d}_i^{\mathrm{T}}}{\|\mathbf{d}_i\|_2^{m+5}} \right)\mathbf{u}, \tag{B.2c}$$

$$\frac{\partial h_i(\mathbf{0})}{\partial \delta_z} = c_i \left(\frac{-m d_z^{m-1} \mathbf{d}_i^{\mathrm{T}} - d_z^m \mathbf{e}_3^{\mathrm{T}}}{\|\mathbf{d}_i\|_2^{m+3}} + \frac{(m+3) d_z^{m+1} \mathbf{d}_i^{\mathrm{T}}}{\|\mathbf{d}_i\|_2^{m+5}} \right)\mathbf{u}. \tag{B.2d}$$

Appendix C

Proofs and Derivations for Chapter 4

C.1 Proof of Proposition 4.2

Consider the unit vector $\mathbf{u} = [u_1 \; u_2]^{\mathrm{T}}$, $u_1 \geq 0$, $u_2 \geq 0$, $\|\mathbf{u}\|_2 = 1$, and let $\varphi_{\mathbf{u}}(t)$, $t \geq 0$, denote the function φ from (4.29b) with its domain restricted to the line passing through the origin along the direction $\mathbf{u}$, i.e.,

$$\begin{aligned}\varphi_{\mathbf{u}}(t) &\triangleq \varphi(t\mathbf{u}) = \varphi(tu_1, tu_2) \\ &= f_{\mathbf{u}}(t) - \frac{1}{2}\left(\ln(u_1^2 t^2 + \sigma^2) + \ln(u_2^2 t^2 + \sigma^2)\right),\end{aligned} \tag{C.1}$$

where $f_{\mathbf{u}}(t) \triangleq f(t\mathbf{u})$. Our goal here is to prove that, for each $\mathbf{u}$, there exists one point $t^\star$ such that $\varphi_{\mathbf{u}}(t)$ is nondecreasing for $t \in [0, t^\star]$ and nonincreasing for $t \geq t^\star$, i.e., $\varphi_{\mathbf{u}}(t)$ is quasiconcave.

Since $f(\delta_1, \delta_2)$ is concave, its restriction to a line is also concave. As a consequence, $f_{\mathbf{u}}(t)$ is continuous and twice differentiable *almost* everywhere, meaning that there are only countably many points where $f''_{\mathbf{u}}(t)$ may not exist [93, Chapter 13]. In order to simplify the notation, we will first restrict ourselves to the points at which $f_{\mathbf{u}}(t)$ is twice differentiable, then we will see that extension to all $t > 0$ is straightforward. Differentiating (C.1) w.r.t. t, we obtain

$$\varphi'_{\mathbf{u}}(t) = f'_{\mathbf{u}}(t) - \left(\frac{u_1^2 t}{u_1^2 t^2 + \sigma^2} + \frac{u_2^2 t}{u_2^2 t^2 + \sigma^2}\right). \tag{C.2}$$

Further differentiation yields

$$\varphi''_{\mathbf{u}}(t) = f''_{\mathbf{u}}(t) + u_1^2 \frac{u_1^2 t^2 - \sigma^2}{(u_1^2 t^2 + \sigma^2)^2} + u_2^2 \frac{u_2^2 t^2 - \sigma^2}{(u_2^2 t^2 + \sigma^2)^2}. \tag{C.3}$$

Let $t^\star$ denote any point at which $\varphi'_{\mathbf{u}}(t) = 0$. Then, we need to show that there is only one such point. Setting $t = t^\star$ and substituting with $\varphi'_{\mathbf{u}}(t^\star) = 0$ in (C.2) yield

$$f'_{\mathbf{u}}(t^\star) = \frac{u_1^2 t^\star}{u_1^2 t^{\star 2} + \sigma^2} + \frac{u_2^2 t^\star}{u_2^2 t^{\star 2} + \sigma^2}. \tag{C.4}$$

Using (C.3) and (C.4), $\varphi''_{\mathbf{u}}(t^\star)$ can be written as

$$\begin{aligned}\varphi''_{\mathbf{u}}(t^\star) = f''_{\mathbf{u}}(t^\star) + \left(f'_{\mathbf{u}}(t^\star)\right)^2 - \frac{2u_1^2 u_2^2 t^{\star 2}}{(u_1^2 t^{\star 2} + \sigma^2)(u_2^2 t^{\star 2} + \sigma^2)} \\ - \frac{u_1^2 \sigma^2}{(u_1^2 t^{\star 2} + \sigma^2)^2} - \frac{u_2^2 \sigma^2}{(u_2^2 t^{\star 2} + \sigma^2)^2}.\end{aligned} \tag{C.5}$$

Now we will show that the sum $f''_{\mathbf{u}}(t^\star) + \left(f'_{\mathbf{u}}(t^\star)\right)^2$ is always nonpositive, and thus $\varphi''_{\mathbf{u}}(t^\star)$ is also nonpositive. To do this, we first need to show that $e^{f_{\mathbf{u}}(t)}$ is a concave function. Let $G(\delta_1, \delta_2)$ denote the optimal value of the perturbed problem

$$\underset{\mathbf{W}}{\text{maximize}} \quad (\mathbf{h}_1^{\mathrm{T}} \mathbf{w}_1)^{\rho} (\mathbf{h}_2^{\mathrm{T}} \mathbf{w}_2)^{1-\rho} \tag{C.6a}$$

$$\text{s.t.} \quad |\mathbf{h}_2^{\mathrm{T}} \mathbf{w}_1| \le \delta_1, \quad |\mathbf{h}_1^{\mathrm{T}} \mathbf{w}_2| \le \delta_2, \tag{C.6b}$$

$$\|\mathbf{W}\|_{\mathrm{F}}^2 \le P_{\mathrm{Tot}}, \tag{C.6c}$$

$$w_{1i}^2 + w_{2i}^2 \le P_i, \quad i = 1, \ldots, N. \tag{C.6d}$$

Since the objective function in (C.6a) is concave (see [80, Problem 3.16 (f)]), the perturbed problem (C.6) is convex, and thus $G(\delta_1, \delta_2)$ is a concave function. Next, we note from (4.28) and (C.6) that $G(\delta_1, \delta_2) = e^{f(\delta_1, \delta_2)}$. Thus, $G_{\mathbf{u}}(t) \triangleq G(t\mathbf{u}) = e^{f_{\mathbf{u}}(t)}$, and we have

$$G''_{\mathbf{u}}(t) = G_{\mathbf{u}}(t)\left(f''_{\mathbf{u}}(t) + (f'_{\mathbf{u}}(t))^2\right). \tag{C.7}$$

Since $G_{\mathbf{u}}(t)$ is concave, it holds that $G''_{\mathbf{u}}(t) \le 0$ [80, Section 3.1.4]. Furthermore, since $G_{\mathbf{u}}(t)$ is nonnegative, we must have

$$f''_{\mathbf{u}}(t) + (f'_{\mathbf{u}}(t))^2 \le 0. \tag{C.8}$$

Thus, $f''_{\mathbf{u}}(t^\star) + \left(f'_{\mathbf{u}}(t^\star)\right)^2 \le 0$ and, consequently, $\varphi''_{\mathbf{u}}(t^\star) \le 0$. The last inequality tells us that $\varphi'_{\mathbf{u}}(t)$ can experience zero-crossing only from positive to negative. Since this can happen only once, we conclude that there is only one point $t^\star$ such that

$$\begin{cases} \varphi'_{\mathbf{u}}(t) \ge 0 & \text{for} \;\; t \le t^\star, \\ \varphi'_{\mathbf{u}}(t) \le 0 & \text{for} \;\; t \ge t^\star. \end{cases}$$

Hence $\varphi_{\mathbf{u}}(t)$ is quasiconcave.

In order to extend the proof to include the points at which $f_{\mathbf{u}}(t)$ is not differentiable, we just need to replace the derivative of $f_{\mathbf{u}}(t)$ with any element from its *subdifferential*. Specifically, since $f_{\mathbf{u}}(t)$ is concave, it is continuous and has right and left derivatives over the whole interior of its domain (i.e., for all $t > 0$) [81, Theorem 1.6]. Such derivatives are nonincreasing in the sense that, for any $t_2 > t_1 > 0$, we have

$$f'_{\mathbf{u}}(t_1^-) \ge f'_{\mathbf{u}}(t_1^+) \ge f'_{\mathbf{u}}(t_2^-) \ge f'_{\mathbf{u}}(t_2^+). \tag{C.9}$$

Now, at the points where $f'_{\mathbf{u}}(t^+) \ne f'_{\mathbf{u}}(t^-)$, i.e., $f_{\mathbf{u}}(t)$ is non-differentiable, we will allow $f''_{\mathbf{u}}(t) \to -\infty$ and let $f'_{\mathbf{u}}(t)$ take any value in the interval $[f'_{\mathbf{u}}(t^+), f'_{\mathbf{u}}(t^-)]$, making (C.8) hold for all $t > 0$. Thus, $\varphi''_{\mathbf{u}}(t^\star)$ is always nonpositive including, possibly,

$\varphi''_{\mathbf{u}}(t^\star) \to -\infty$. In other words, $\varphi'_{\mathbf{u}}(t^+)$ (or, equivalently, $\varphi'_{\mathbf{u}}(t^-)$) can experience zero-crossing only from positive to negative, even if $\varphi'_{\mathbf{u}}(t^+)$ has jump discontinuity at the crossing point. Following the same argument for the differentiable case, we conclude that $\varphi_{\mathbf{u}}(t)$ is quasiconcave for all $t \geq 0$. ■

C.2 Derivation of the Dual Problem (4.57)

The problem in (4.55) can be reformulated as

$$\begin{aligned}
\underset{\mathbf{W},\mathbf{M},z_1,\ldots,z_4}{\text{maximize}} \quad & \rho \ln z_1 + (1-\rho)\ln z_2 && \text{(C.10a)}\\
\text{s.t.} \quad & \hat{\mathbf{h}}_1^{\mathrm{T}}\mathbf{w}_1 - \epsilon_1\|\mathbf{w}_1\|_2 \geq z_1, && \text{(C.10b)}\\
& \hat{\mathbf{h}}_2^{\mathrm{T}}\mathbf{w}_2 - \epsilon_2\|\mathbf{w}_2\|_2 \geq z_2, && \text{(C.10c)}\\
& |z_3| + \epsilon_2\|\mathbf{w}_1\|_2 \leq \delta_1, && \text{(C.10d)}\\
& |z_4| + \epsilon_1\|\mathbf{w}_2\|_2 \leq \delta_2, && \text{(C.10e)}\\
& \|\mathbf{m}_1\|_2^2 + \|\mathbf{m}_2\|_2^2 \leq P_{\text{Tot}}, && \text{(C.10f)}\\
& m_{1i}^2 + m_{2i}^2 \leq P_i, \quad i = 1,\ldots,N, && \text{(C.10g)}\\
& \mathbf{w}_1 = \mathbf{m}_1, \quad \mathbf{w}_2 = \mathbf{m}_2, && \text{(C.10h)}\\
& \hat{\mathbf{h}}_2^{\mathrm{T}}\mathbf{w}_1 = z_3, \quad \hat{\mathbf{h}}_1^{\mathrm{T}}\mathbf{w}_2 = z_4, && \text{(C.10i)}
\end{aligned}$$

where we have introduced the new variables $\mathbf{M}$, z_3, and z_4, along with the equality constraints in (C.10h)-(C.10i).

The Lagrangian associated with (C.10) is

$$
\begin{aligned}
&L(\mathbf{W},\mathbf{M},z_1,\dots,z_4,\chi_1,\chi_2,\lambda_1,\lambda_2,\gamma,\boldsymbol{\mu},\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\nu_1,\nu_2)\\
&\quad=\rho\ln z_1+(1-\rho)\ln z_2\\
&\qquad-\chi_1(-\hat{\mathbf{h}}_1^{\mathrm{T}}\mathbf{w}_1+\epsilon_1\|\mathbf{w}_1\|_2+z_1)-\chi_2(-\hat{\mathbf{h}}_2^{\mathrm{T}}\mathbf{w}_2+\epsilon_2\|\mathbf{w}_2\|_2+z_2)\\
&\qquad-\lambda_1\left(|z_3|+\epsilon_2\|\mathbf{w}_1\|_2-\delta_1\right)-\lambda_2\left(|z_4|+\epsilon_1\|\mathbf{w}_2\|_2-\delta_2\right)\\
&\qquad-\gamma\left(\|\mathbf{m}_1\|_2^2+\|\mathbf{m}_2\|_2^2-P_{\mathrm{Tot}}\right)-\sum_{i=1}^{N}\mu_i(m_{1i}^2+m_{2i}^2-P_i)\\
&\qquad-\boldsymbol{\eta}_1^{\mathrm{T}}\left(\mathbf{w}_1-\mathbf{m}_1\right)-\boldsymbol{\eta}_2^{\mathrm{T}}\left(\mathbf{w}_2-\mathbf{m}_2\right)-\nu_1(\hat{\mathbf{h}}_2^{\mathrm{T}}\mathbf{w}_1-z_3)-\nu_2(\hat{\mathbf{h}}_1^{\mathrm{T}}\mathbf{w}_2-z_4).
\end{aligned}
\tag{C.11}
$$

Rearranging the terms, and maximizing w.r.t. the primary variables $\mathbf{W},\mathbf{M},z_1,\dots,z_4$, we obtain the dual function

$$
\begin{aligned}
&g(\chi_1,\chi_2,\lambda_1,\lambda_2,\gamma,\boldsymbol{\mu},\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\nu_1,\nu_2)\\
&\quad=\lambda_1\delta_1+\lambda_2\delta_2+\gamma P_{\mathrm{Tot}}+\sum_{i=1}^{N}\mu_iP_i\\
&\qquad+\max_{\mathbf{w}_1}\Big((\chi_1\hat{\mathbf{h}}_1-\boldsymbol{\eta}_1-\nu_1\hat{\mathbf{h}}_2)^{\mathrm{T}}\mathbf{w}_1-\ (\lambda_1\epsilon_2+\chi_1\epsilon_1)\|\mathbf{w}_1\|_2\Big)\\
&\qquad+\max_{\mathbf{w}_2}\Big((\chi_2\hat{\mathbf{h}}_2-\boldsymbol{\eta}_2-\nu_2\hat{\mathbf{h}}_1)^{\mathrm{T}}\mathbf{w}_2-\ (\lambda_2\epsilon_1+\chi_2\epsilon_2)\|\mathbf{w}_2\|_2\Big)\\
&\qquad+\sum_{i=1}^{N}\max_{m_{1i}}\left(\eta_{1i}m_{1i}-(\gamma+\mu_i)m_{1i}^2\right)+\sum_{i=1}^{N}\max_{m_{2i}}\left(\eta_{2i}m_{2i}-(\gamma+\mu_i)m_{2i}^2\right)\\
&\qquad+\max_{z_1}\left(-\chi_1z_1+\rho\ln z_1\right)+\max_{z_2}\left(-\chi_2z_2+(1-\rho)\ln z_2\right)\\
&\qquad+\max_{z_3}\left(\nu_1z_3-\lambda_1|z_3|\right)+\max_{z_4}\left(\nu_2z_4-\lambda_2|z_4|\right).
\end{aligned}
\tag{C.12}
$$

The first maximization in the Lagrangian (C.12) is the conjugate of the l_2-norm function [80, Example 3.26], and is solved as

$$\max_{\mathbf{w}_1} \Big((\chi_1 \hat{\mathbf{h}}_1 - \boldsymbol{\eta}_1 - \nu_1 \hat{\mathbf{h}}_2)^{\mathrm{T}} \mathbf{w}_1 - (\lambda_1 \epsilon_2 + \chi_1 \epsilon_1) \|\mathbf{w}_1\|_2 \Big) \tag{C.13}$$

$$= \begin{cases} 0 & \|\chi_1 \hat{\mathbf{h}}_1 - \boldsymbol{\eta}_1 - \nu_1 \hat{\mathbf{h}}_2\|_2 \leq \lambda_1 \epsilon_2 + \chi_1 \epsilon_1 \\ \infty & \text{otherwise.} \end{cases} \tag{C.14}$$

Then, after solving the other maximization terms in (C.12), which are similar to those in (4.34), the dual problem (4.57) follows.