Security by Similarity: Information-theoretic Secrecy and Decentralized Detection Algorithms for Wireless Networks

vorgelegt von
Jafar Mohammadi, M.Sc.
geb. in Bookan

von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften
- Dr.-Ing. -
genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Giuseppe Caire
Gutachter: Prof. Sławomir Stańczak
Gutachterin: Prof. Negar Kiyavash (UIUC)
Gutachter: Prof. Tobias Oechtering (KTH)

Tag der wissenschaftlichen Aussprache: 12 August 2016

Berlin, 2017

Abstract

The broadcasting attribute of the wireless communication channel poses severe security vulnerabilities since any adversary node in the range of reception is able to listen as well as to disrupt the ongoing communication. This dissertation considers the problem of security from two different perspectives: The information-theoretic approach at the physical-layer to avert eavesdropping and centralized and decentralized detection algorithms to detect the presence of a jammer.

In the first part, we apply the available tools in information-theoretic security to study the achievable secrecy rate region of the following channel models: the discrete memoryless broadcast channel with two confidential messages for each of the two users, the discrete memoryless interference channel with two confidential messages, and the discrete memoryless untrusted relay channel with decode-and-forward (DF) relays (DMURC). For the first two scenarios, we extend the existing results on the *weak secrecy* to the *strong secrecy* criterion. Owing to its definition, the strong secrecy has a practical interpretation, whereas weak secrecy provides asymptotic results. In the third scenario, we apply a precoder and a post-decoder in addition to the common channel coding required to provide secrecy, to deal with the dishonest nature of the relays. An achievable secrecy rate region for DMURC is established.

We revisit the untrusted relay scenario with some different assumptions: with Gaussian channel and both DF and amplify-and-forward (AF) relaying. The objective, however, is to develop a secure scheme that allows for weaker secrecy than Shannon's *perfect secrecy* but with a higher transmission rate. We propose a new measure for secrecy, referred to as α-secrecy ($0 \leq \alpha \leq 1$), where $\alpha = 1$ represents the *perfect secrecy*, whereas $\alpha = 0$ denotes no secrecy at all. We suggest using a cross-layer design where at the physical layer the information-theoretic security is applied and some higher layer scheme, e.g. *secure network coding*, allows to obtain any point in the transmission rate and α-secrecy trade-off. The analysis of performance achievable by the scheme leads to a mixed-integer optimization problem. Moreover, we study the problem asymptotically to gain some insights.

Active security attacks are the main focus of the second part of this thesis. We study the problem of joint jamming detection and spectrum sensing. The problem arises in the cognitive radio scenario, where the secondary users (SU)s sensing the channel to detect the primary users (PU)s. We model the problem as a multiple-hypothesis testing problem where the observations from the SUs experience spatial channel correlation. The optimal test statistic is derived according to Neyman-Pearson lemma. We derive the exact probability distribution of the test statistic in terms of Meijer's G-functions. We consider a practical application where the SUs' power efficiency is analyzed. This leads to a non-convex optimization problem that addresses the power efficiency and detection performance trade-off. The detection performance in the exact form is too complex to be handled analytically, therefore, we consider the asymptotic regime. Two heuristics are proposed to tackle the optimization problem; one based on the projected gradient method and the other one is motivated by the KKT conditions.

All of the algorithms mentioned above are designed for a centralized system. Taking into account the susceptibility of centralized systems to the Byzantine attacks, we study the decentralization of eigenvalue-based detection algorithms in the last chapter of this thesis. We use generic consensus algorithm as a building block in decentralizing an eigenvalue estimator algorithm. The algorithm estimates all eigenvalues of a covariance matrix, which is available only partly to each node. We further modify the algorithm for two special practical scenarios, including the channel probing for admission control problem.

Zusammenfassung

Die Broadcast-Eigenschaft des drahtlosen Funkknals stellt eine empfindliche Schwachstelle der Kommunikationssicherheit dar, da jeder schädliche Knoten in Reichweite die übertragene Information mithören, oder die Übertragung stören kann. Die vorliegende Dissertation befasst sich mit dem Problem der Sicherheit von Kommunikation aus zweierlei Perspektiven: der informationstheoretischen Sicherheit auf der physikalischen Schicht zur Abwehr von Lauschangriffen, sowie zentralisierten und dezentralizierten Detektionsalgorithmen zur Erkennung von Störsendern.

Im ersten Teil werden Methoden der informationstheoretischen Sicherheit angewandt, um erreichbare Sicherheitsratenregionen für den diskreten gedächtnislosen Zweinutzer-Broadcast-Kanal mit zwei vertraulichen Nachrichten je Nutzer, den diskreten gedächtnislosen Interferenzkanal mit zwei vertraulichen Nachrichten, sowie den diskreten gedächtnislosen Kanal mit nicht vertrauenswürdigen decode-and-forward (DF) Relais (DMURC) zu ermitteln. Für die ersten beiden Fälle werden bestehende Resultate basierend auf dem Konzept *schwacher informationstheoretischer Sicherheit* hin zum Konzept *strenger informationstheoretischer Sicherheit* erweitert. Aus der Definition *strenger informationstheoretischer Sicherheit* ergibt sich eine direkte praxisnahe Möglichkeit zur Interpretation, während das Konzept *schwacher informationstheoretischer Sicherheit* für asymptotische Einsichten genutzt werden kann. Im dritten Fall wird die Vor- und Nachkodierung gemeinsam mit der üblicherweise zur Wahrung der Vertraulichkeit verwendeten Kanalkodierung eingesetzt, um dem nicht vertaulichen Charakter der Relais beizukommen. Für den DMURC wird dabei die erreichbare Sicherheitsratenregion bewiesen.

Danach wird die Kommunikation über nicht vertrauenswürdige Relais unter der Annahme eines Gauß-Kanals und DF sowie amplify-and-forward (AF) Relais analysiert. Hierbei wird zwar verglichen mit der perfekten Sicherheit nach Shannon nur ein schwächeres Konzept informationstheoretischer Sicherheit entwickelt, dafür werden im Gegenzug höhere Übertragungsraten ermöglicht. Dazu wird ein neues Maß, die sog. α-Sicherheit ($0 \leq \alpha \leq 1$), eingeführt, wobei $\alpha = 1$ den Fall perfekter informationstheoretischer Sicherheit, und $\alpha = 0$ den Fall keiner informationstheoretischen Sicherheit beschreibt. Mit Hilfe eines schichtübergreifenden Entwurfs aus informationstheoretischer Sicherheit auf der physikalischen Schicht, sowie einer weiteren Methode auf höheren Schichten (bsp. sichere Netzwerkcodierung), können alle Punkte des Abtauschs zwischen Übertragungsrate und α-Sicherheit erreicht werden. Zur Bewertung der dadurch erzielbaren Güte wird ein gemischt-ganzzahliges Optimierungsproblem aufgestellt, wobei dessen asymptotische Analyse weitere Einsichten liefert.

Aktive Angriffe auf die Systemsicherheit stehen im Fokus des zweiten Teils dieser Arbeit. Dabei wird das Problem der gemeinsamen Störererkennung und spektralen Messung analysiert, welches im Bereich kognitiver Funksysteme auftritt, bei denen Sekundärnutzer (SU) die Aktivität von Primärnutzern (PU) detektieren müssen. Das Problem wird als Hypothesentest mit räumlicher Kanalkorrelation in den Messungen der Sekundärnutzer modelliert, wobei die optimale Teststatistik im Neyman-Pearon Sinn abgeleitet wird. Die Wahrscheinlichkeitsverteilung der Teststatistik wird in analytischer Form mit Hilfe der Meijer'schen G-Funktion hergeleitet. Als Anwendung wird die Analyse der Energieeffizienz von Sekundärnutzern herangezogen, welche zu einem nicht-konvexen Optimierungsproblem als Abtausch zwischen Energieeffizienz und Detektionsgüte führt. Da die Detektionsgüte für die analytische Betrachtung zu unhandlich erscheint wird sie asymptotisch betrachtet. Das Optimierungsproblem wird heuristisch durch ein projiziertes Gradientenverfahren sowie eine KKT-basierte Methode gelöst.

Die genannten Algorithmen sind üblicherweise für den Einsatz in zentralisierten System gedacht. Aufgrund der Anfälligkeit solcher Systeme für Angriffe durch byzantinische Fehler werden im letzten Kapitel dieser Arbeit dezentralisierte und Eigenwert-basierte Detektionsalgorithmen untersucht. Dabei werden generische Konsensalgorithmen als Baustein zur dezentralisierten Berechnung aller Eigenwerte einer gegebenen Kovarianzmatrix, die nur zum Teil an den Knoten bekannt ist, herangezogen. Durch weitere Modifizierungen wird der Algorithmus an praktische Einsatzfälle, wie die Kanalmessung für das Zutrittskontrollproblem, angepasst.

Acknowledgements

I was the recipient of support from many talented people during my doctoral studies. First and foremost, I would like to express my deepest gratitude to my supervisor Prof. Sławomir Stańczak for his attentive supervision, scientific support, and for providing the opportunity to do my PhD in his group. I appreciate his constant encouragement toward accurate and pure research and extraordinary passion for science which has influenced my attitude significantly. My gratitude extends to Dr. Igor Bjelankovic for his scientific support and invaluable discussions on information theory. I am also very grateful to Prof. Negar Kiyavash and Prof. Tobias Oechtering for devoting their time to be referees of this thesis.

Moreover, I would like to thank all my colleagues in both Heinrich-Hertz-Institut Fraunhofer and the Technische Universität Berlin, for providing a stimulating research environment, friendship, and comfort. In particular, I would like to thank Michał Kaliszan, Steffen Limmer and Meng Zheng for the intensive scientific discussions, co-authoring papers with me and being a friend aside from a colleague. I would like to thank Renato Cavalcante, Jan Schreck, Federico Penna, and Mario Goldenbaum for the many fruitful discussions. My special thanks also go to Saeid Haghighatshoar for the numerous philosophic and scientific conversations during the coffee breaks, for introducing me some fantastic reads, and for being a great friend with a positive attitude toward life.

I would not have been what I am now without the endless love, support, and patience of my parents. My fathers' constant encouragement to pursue science, honor, and honesty has been my source of will and motive. The unconditional love that my mother granted me, leaves me speechless to describe.

I was granted a three years scholarship from Helmholtz Research School on Security Technologies (HRSST) in 2011 to pursue my doctoral studies. Furthermore, I was the recipient of three months scholarship from Heinrich Hertz Institute Fraunhofer to finalize this dissertation.

Contents

List of Figures

1. Introduction

1.1. Motivation

The advent of mobile communication and the subsequent high penetration rate of mobile internet connectivity via smartphones in the world realize the idea of the global village more than ever. Instantly sharing ideas, photos, videos, and even (recently) live broadcasting of videos have become a reality for every person, despite being once considered far-fetched. The colossal amount of users on the Internet, sharing data and communicating with each other demands a strong measure of data privacy and security. Mobile communication systems use the wireless channel as the medium of communication, which poses exceptional opportunities and challenges regarding security and privacy issues. The innate omnidirectional propagation of electromagnetic waves, as referred to as the *broadcast* property of the wireless channels, jeopardizes the security of the communication system since the messages can be overheard by a potential eavesdropper. Jamming attacks, on the other hand, can exploit the openness of the wireless channel by transmitting high power electromagnetic waves to disrupt the communication. Eavesdropping without any attempt to sabotage the communication system is categorized as a passive attack, whereas, jamming a communication system is regarded as an active attack.

Against this background, the wireless channel allows us to conceive countermeasures by exploiting the intrinsic randomness, the diversity from multiple antennas, the cooperative communications, and the relaying. The inherent randomness in the wireless channel state information allows designing encryption keys and proper secure channel codes. Furthermore, the use of multiple antennas enables vector communication by steering the wave propagation in a specific direction, which can nullify the signal strength at the eavesdropper. We can benefit from cooperative communications for detecting an active attack, whereas, the redundancy from relaying can enhance secrecy against active and passive attacks. In the following, we briefly introduce some tools and research works available to analyze particular types of passive and active attacks and develop countermeasures to them.

Information-theoretic security commenced with Shannon's work in [Sha49] as a technique to combat passive attacks, in particular, eavesdropping. Despite being an inspiring breakthrough, the work of Shannon [Sha49] does not consider the intrinsic characteristics of wireless communications. Shannon concluded that *perfect secrecy* is only achievable when the size of the key at least as equal as the size of the message. Wyner, however, considers noise in his seminal paper "The wiretap channel" [Wyn75]. Under the assumption that the eavesdropper's channel is a degraded version of the legitimate user's channel, Wyner achieves a positive secrecy rate. The work of [Wyn75] was followed up by Csiszár and Körner [CK78], who extended the results to a non-degraded broadcast channel with only one confidential message and a common message. Since then many other works have embraced the realization of information theoretic security, e.g. [LMSY08, LP09, XCC09, HY13, HY10b, BP15, HY10a, CBA15].

How to deal with the jamming attacks has been researched vastly in different directions, including but not limited to: *channel surfing* based on spread spectrum frequency hopping [XWZ, PL12], jammer localization [LLXC12] for re-routing [MZL+12], and jamming detection in [LLK09, ZHSA05, SDČ10], in which they perform detection based on the parameters at the physical layer as well as at the higher layers. In this thesis, we mainly focus on jamming detection techniques based on the parameters at the physical layer. The problem of physical layer jamming detection is fundamentally not much different from the problem of spectrum sensing for cognitive radio [MGKP09, LH09]. The spectrum sensing entails detecting the primary users' (PU) communication by some secondary users (SU)s [PHH+12]. The SUs observe the channel at the physical layer to infer the presence of the PUs' communication [LH09, LH10]. The practical solutions for the jamming detection consider the cooperation of multiple transceivers as well as their energy restrictions. Since the transceivers are usually distributed over a given area, a decentralized implementation of such solutions is of great importance (we refer to Part II of this thesis for further references).

1.2. Thesis outline

In **Part I** of this dissertation, we study the countermeasures that safeguard the communication against passive attacks using tools from information-theoretic security. **Part II** is dedicated to detection methods dealing with active attacks, in particular, jamming detection. We further elaborate on practical concerns (such as decentralization and energy efficiency) of a jamming detection algorithm.

We begin with a brief recap on information-theoretic tools for security in **Chapter 2**. We introduce *typical sequences* and some useful lemmas, which are used in proving achievable secure rate regions in the Chapter 3. We further introduce two of the most important lemmas in the information theory; the so-called *covering lemma* and *packing* *lemma*. Since we use the Kullback-Leibler divergence later in Chapter 3, we provide the formal definition and a useful lemma, which will be used later. After providing the basic tools, we provide the definition of the well-studied wiretap channel. The purpose is to familiarize the reader with the simplest scenario in information-theoretic security. We further provide an achievable secrecy rate and a concise achievability proof. Chapter 2 is in large parts based on the content of [GK12].

Chapter 3 presents the main results of **Part I** on information-theoretic secrecy. The emphasis is placed on developing results to *strong secrecy* for the following two scenarios: The discrete memoryless broadcast channel with two receivers, each with one confidential message, is studied first in Section 3.3. Then the discrete memoryless interference channel with two senders communicating a confidential message to one of the two receivers is studied in Section 3.4. The *weak secrecy* requires upper-bounding the information that is leaked to the adversary by an arbitrarily small number while it is averaged over the message block length size. On the other hand, the *strong secrecy* demands a stronger condition: the information leakage must be upper-bounded by an arbitrarily small number without being averaged by the message block length. We fulfill this requirement by upper-bounding not only the information leakage, but also an extra non-negative term known as 'non-stealth' [HK14]. We conclude that the strong secrecy achievable rate region matches that of the weak secrecy, which is proposed by [LMSY08].

In addition to these scenarios, in Section 3.5 we consider a discrete memoryless untrusted relay channel, in which a sender transmits a confidential message via the untrusted decode-and-forward relays to a destination (the diamond network). We assume that there are no direct channels available between the sender and the destination, as well as, among the relays. Since the relays are decode-and-forward, we require a scheme to protect the messages at each relay. Basically in Section 3.5, we use the results developed in Section 3.3 with a precoder at the sender and an extra post-decoder at the destination to deal with this challenge. The results that are developed for the untrusted relay scenario, however, do not satisfy the strong secrecy criterion, and thus fall into weakly secure criterion. The content of this chapter has been published in [1, 4].

Chapter 4 considers, similar to Section 3.5, the untrusted relay model, however, from a different point of view; a cross-layer secrecy design. We consider two-hop Gaussian relay channels, where the relays are untrusted. Two different types of relays have been considered: the decode-and-forward (DF) and the amplify-and-forward (AF). The goal in this chapter is to provide a weaker form of security than that of Shannon's perfect secrecy (even weaker than the *weak secrecy*), with the help of some higher-layer schemes. This is motivated by the applications where perfect secrecy is not the highest priority, whereas, the transmission rate is. Note that, perfect secrecy comes at the expense of transmission rate

[Wyn75]. A new measure of security, referred to as the α-secrecy is proposed in Section 4.3, where $0 \leq \alpha \leq 1$. According to this scheme 1-secrecy is equivalent to Shannon's perfect secrecy (either weak or strong), whereas, 0-secrecy represents the state of no physical layer security.

Thanks to some additional higher layer schemes, e.g. secure network coding, we propose a framework in Section 4.3 to achieve any α-secure system. This framework is then applied to the untrusted DF relay channels in Section 4.4 and the untrusted AF relay channels in Section 4.5. The optimal achievable rate for a given α requires solving a mixed-integer optimization problem, which is dealt with in Section 4.5 after some relaxation. We further analyze the asymptotic behavior of the mixed-integer problem along with the secrecy outage probability in Section 4.6. With this analysis, we conclude **Part I** of this dissertation. The content of this chapter has been published in [7, 9].

The focus of **Part II**, i.e. Chapter 5 and Chapter 6, is to study detection algorithms to combat a particular active security attack: jamming.

Chapter 5 studies the joint problem of jamming detection and spectrum sensing in a cognitive radio scenario. This problem is motivated by vulnerabilities of cognitive radio systems to jamming. We assume a set of SU nodes sense the spectrum to both detect the presence of the PUs and a jammer. Modeling the scenario as a multiple hypothesis testing problem, we analyze the probability of detection of the optimal detector in the sense of Neyman-Pearson theorem in Section 5.3. The probability of detection is derived in the following forms: one exact form in terms of a series and a closed-form version. Moreover, we evaluate the asymptotic probability of detection, as it results in a simpler form to handle. In all of the above analyzes, we assumed the data collected by the SUs are spatially correlated but temporally independent. In practice, power constraints on SUs limit their versatility to adopt any detection algorithm, since they might run out of energy. Building upon the asymptotic probability of detection developed in Section 5.3 we study the trade-off between transmission power and detection performance in Section 5.4. The problem studies the optimal node placement for power efficient detection in a wireless network scenario. The resulting optimization problem is non-convex; thus, we propose suboptimal numerical approaches to tackle it. We use both a projected gradient method and a heuristic method developed based on the Karush-Kuhn-Tucker (KKT) conditions. The content of this chapter has been published in [MPS13, MSZ15].

The detection algorithms mentioned so far are based on the premise that we have a safe and secure fusion center that collects the data and performs the detection algorithms on them. However, a central processing of the data might not be secure against Byzantine attacks or directed jamming attacks on the fusion center. These arguments and many others motivates the study of a decentralized detection algorithm, which is the topic of

Chapter 6. We focus our attention in Chapter 6 on decentralization of some eigenvalue-based detection algorithms. We assume a number of sensors that cooperatively detect a phenomena, e.g. jamming. Each sensor locally stores a vector containing the observations of the channel over time. We need to compute the eigenvalues of the covariance matrix of the observations from all of the sensors. In particular, we develop an algorithm to calculate all of the eigenvalues of the covariance matrix in a decentralized fashion, i.e. each sensor performs the calculations using only parts of the matrix. The proposed algorithm is based on a generalization of the power iteration method. The decentralized version of this algorithm is developed using the generic *average consensus* algorithm. The consensus algorithms are a family of decentralized iterative methods that accept a vector as input, where each element of the vector is available at a sensor and estimates the average of the vector as the output at each sensor. Using an average consensus, we can compute 'summation' over the values that are available at each sensor locally.

For special application, where the matrix is structured, we propose using a fast converging consensus algorithm. A generic consensus algorithm, does not take the special constellations that the nodes can form into account. In Section 6.4 we exploit a particular graph shape to obtain a fast converging consensus algorithm, therefore estimating the eigenvalues in a faster way. We further demonstrate the applications of these techniques in *channel probing* for power admission control in Section 6.5. Channel probing is a process that tests the admissibility of a new link joining some established links. The test reduces to estimating the spectral radius of a matrix which is not available at every link. The matrix is combination of the channel matrices of the existing nodes and the new nodes. Using the result developed in Section 6.3, we determine the admissibility of multiple new users at once and with few iteration. The content of this chapter has been published in [MSCE11, 2].

1.3. Additional results

In the following we present some results that are not a part of this dissertation.

- In [3], we study the design of a joint sender precoding matrix and the amplify-and- forward relay precoding matrix optimization. We consider a Gaussian two-hop multiple-input-multiple-output (MIMO) relay network. The goal is to find a pair of matrices in order to minimize the power consumption and at the same time meet pre-selected quality of service constraints that are defined as the mean square error of each data stream. The resulting optimization problem is difficult to solve in the closed form. Therefore, we apply some tools from the *majorization* theory, which defines relationships between ordered vectors, to simplify the optimization problem. Thanks to some inequalities and lemmas in majorization theory we simplify the original matrix-valued optimization problem into a scalar-valued one. We then propose a lowerbound and an upperbound of the original problem, both in convex forms. Owing to the practical reasons, we focus on solving the upperbound. Using the equations imposed by the KKT conditions we propose an iterative fast converging algorithm similar to a multi-level water-filling algorithm. We analyze the convergence and the optimality of the solutions mathematically. Numerical examples corroborate the proposed studies and also demonstrate the tightness of both bounds to the original problem.

- In cooperation with Steffen Limmer in [5] we contemplate distributed computation over multiple access channel (CoMAC). One of the missing techniques in CoMAC research is how to convert an arbitrary function intro pre-and post-processing functions that are feasible to be computed via the CoMAC scheme. In this work we propose a scheme which has the potential to be considered as a solution for this problem. In particular we introduce a novel algorithmic solution for the approximation of a given continuous multivariate function by a nomographic function that is composed of a one-dimensional continuous and monotone outer function and a sum of univariate continuous inner functions. The direct application of this approach is in the CoMAC scenario, where the multivariate function is to be estimated by exploiting the superposition property of the wireless channel. We show that a suitable approximation can be obtained by solving a cone-constrained Rayleigh-Quotient optimization problem. The proposed approach is based on a combination of a dimension-wise function decomposition known as Analysis of Variance (ANOVA) and optimization over a class of monotone polynomials. An example is given to show that the proposed algorithm can be applied to solve problems in distributed function computation over multiple-access channels.

- The results in [10] were co-authored with Meng Zheng. A key to efficient mobile communication networks is an efficient utilization of scarce resources, which includes mechanisms for resource allocation and interference management. In practice, time sharing methods such as Time Division Multiple Access (TDMA) are often exploited to create orthogonal connections. TDMA schemes, however, are inflexible and not suitable for decentralized implementation. In this context, approaches based on concurrent transmissions in the same frequency spectrum admitting the interference offer a promising and attractive alternative. Yet, their performances deteriorate dramatically if the underlying interference coefficients are relatively large. Assuming a Gaussian Interference Channel (GIC), this paper characterizes the so-called

TDMA region which is defined as a set of all interference coefficient matrices for which TDMA outperforms concurrent transmissions. We use the dimensional lifting method to prove the convexity and monotonicity of the TDMA region, which provides some insights into the scheduling design under uncertainty with respect to the interference coefficients. Finally, the analysis is used to develop strategies for resource allocation in TDMA-based wireless networks.

1.4. Copyright notice

This dissertation has already been published, in part, as journal articles, conference proceedings, and workshop proceedings as listed in the Publication list on pages 135-136. These parts, which are, up to minor modifications, identical with the corresponding scientific publication are copyrighted by the IEEE (2010-2016).

Part I.

Information-Theoretic Security

2. Preliminaries on information-theoretic secrecy

In this chapter, we review some preliminary results in information theory. We use these results to study the wiretap channel [Wyn75]. Studying the wiretap channel paves the way to better understanding the results in Chapter 3. A complete review and proofs of the results in this chapter can be found in [GK12].

Traditionally, communication standards treat the problems of communication reliability and communication security separately. More precisely, the reliability is handled by the physical layer, whereas the security is guaranteed through higher layers including the application layer. At the physical layer channel coding techniques ensure a reliable communication; on the other hand, at the application layer encryption algorithms administer the security of a communication system. Information-theoretic security alters this paradigm by providing security at the physical layer with the help of proper channel coding. Information-theoretic security, however, does not relinquish the encryption algorithms. As the encryption and the information-theoretic security are applied in different layers, they can be used alongside each other.

We give a brief introduction to a generic encryption system in the following; a more comprehensive one can be found in [BB11,MVOV96]. Let Alice, Bob and Eve be three parties of a communication system, where Alice intends to send a confidential message to Bob while keeping it secret from Eve. In cryptographic terminology, Alice's message is called *plaintext*, and the encrypted version is *ciphertext*. To securely send a message (plaintext), Alice feeds the plaintext into an encryption algorithm. The encryption algorithm requires a key as well as the plaintext to produce the ciphertext. At the receiving side, Bob has a decryption algorithm, which requires a key to decrypt the ciphertext. There are different notions of secrecy in the literature. The so-called *computational security* is measured by the number of computations required for an eavesdropper to decipher a ciphertext. For instance, the computational security of a typical encryption algorithm is compared to the complexity of prime factorization of large numbers [BB11].

Shannon in [Sha49] addressed a stronger notion of secrecy known as the *perfect secrecy* in the context of information-theoretic secrecy. Information-theoretic secrecy utilizes chan-

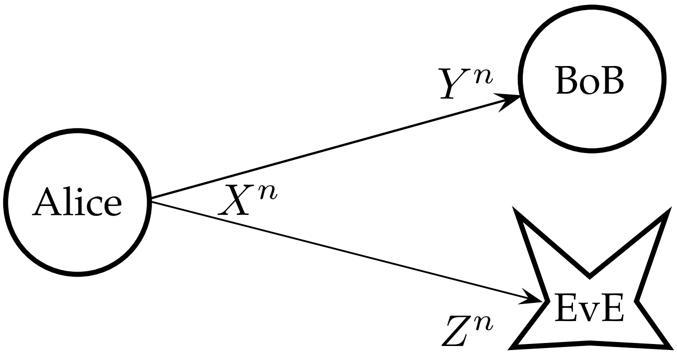

Figure 2.1.: An illustration of the three party communication scenario starring Alice, Bob, and Eve representing the legitimate sender, the legitimate receiver, and the eavesdropper.

nel coding not only to overcome the channel's impairments at the physical layer but also to tackle the security problem. This results in a fundamental trade-off between the communication rate, reliability and security. Furthermore, information-theoretic secrecy can provide security without a secret key by, for instance, taking advantage of the intrinsic randomness of the wireless communication channel. The secrecy is measured by the amount of information leakage to the eavesdropper. Intuitively, Shannon's perfect secrecy implies that Eve's observation of the codewords is not helpful to draw any conclusion about the content of the message; thus, the best Eve can do, is to guess the message.

In the first sections the notation and the very basic form of the main lemmas, namely the packing lemma and the covering lemma, are introduced. To understand the lemmas, we also mention the definition of typical sequences and the joint typicality lemma. We use these lemmas to establish the wiretap's channel capacity [Wyn75]. The coding scheme applied in this section is the building block of the more complicated coding schemes which are introduced in Chapter 3.

2.1. Notation

Throughout this dissertation, we use the following notational conventions unless otherwise is specified. We denote the sets of natural, real and complex numbers as $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$, respectively. We further let $\mathbb{R}_+$ represent the non-negative real number set and $\mathbb{R}_{++}$ the positive real number set. Sets of N-fold Cartesian products are denoted by the superscript $(\cdot)^N$. The 'equality by definition' is denoted by ":=", whereas, $\forall$, $\exists$, and $|\cdot|$ denote 'for all', 'there exists', and the absolute value of a real number[1], respectively.

---

[1] As described later it also denotes the cardinality of a set.

We denote matrices by capital bold letters, e.g. $\mathbf{X}$, and their Hermitian with $\mathbf{X}^H$. We define the operator Diag : $\mathbb{C}^K \to \mathbb{C}^{K\times K}$. We denote by $\text{Diag}[x_1, \dots, x_K]$ a $K \times K$ diagonal matrix, with $[x_1, \dots, x_K]$ as its diagonal elements. The operator diag : $\mathbb{C}^{K\times K} \to \mathbb{C}^K$ is used to denote the diagonal elements of $\mathbf{X}$ as $\text{diag}[\mathbf{X}]$. Furthermore, by $\det[\mathbf{X}]$ and $\text{trace}[\mathbf{X}]$ we denote the determinant and the trace of matrix $\mathbf{X}$, respectively. We denote the Frobenius norm of $\mathbf{X}$ by $\|\mathbf{X}\|_F$. By $[\mathbf{X}]_{i,j}$ we denote the (i,j)-th element of $\mathbf{X}$. The largest (absolute) eigenvalue of a matrix is denoted by $\lambda_{max}[\mathbf{X}]$ and the smallest by $\lambda_{min}[\mathbf{X}]$. Rank of a square matrix $\mathbf{X}$ is indicated by $\text{rank}[\mathbf{A}]$.

We depict the random variables (RV) with the capital case, e.g. X, and their realizations with small letters. Finite sets are represented by calligraphic letters, e.g. $\mathcal{X}$, and their cardinality with $|\mathcal{X}|$. Sequences are denoted by the bold small case, e.g. $\mathbf{x}$, which stands for $x^n := (x_1, \dots, x_n)$. The set of all probability distributions defined on the finite set $\mathcal{X}$ is indicated with $\mathcal{P}(\mathcal{X})$. The probability mass function (pmf) of X is given by $P_X(x)$ or in short by $P(x)$, and the n-product of $P_X(x)$ is denoted by $P_X^n(\mathbf{x}) := \prod_{i=1}^n P_X(x_i)$. The probability of an event E is $\Pr(E)$. The expected value of a function $f(X)$, where X is a RV, is $\text{E}_X[f(X)] := \sum_x P_X(x)f(x)$. Further, we define $\text{Var}[X] := \text{E}[(X - \text{E}[X])^2]$ as the variance of X.

For the RVs X and Y we denote their joint pmf as $P_{X,Y}(x,y)$ and their conditional pmf as $P_{X|Y}(x|y)$. Based on the joint and conditional pmfs of X and Y we have: the information entropy $H(X) = -\text{E}_X[\log P(X)]$, the conditional entropy $H(X|Y) = -\text{E}_{X,Y}[\log P(X|Y)]$, and the mutual information $I(X;Y) = H(X) - H(X|Y)$.

2.2. Basic concepts

The following content is considered to be the only tools needed to understand the results of this dissertation.

2.2.1. Typical sequences

Definition 1 (Typical sequences)**.** Let $\mathcal{X}$ be a finite alphabet. For $X \sim P_X(x)$ on $\mathcal{X}$ and some $\epsilon \in (0,1)$, the set of all ϵ-typical n-sequences $\mathbf{x} \in \mathcal{X}^n$, or simply the set of typical sequences, is defined as [GK12, Chapter 2.4],

$$\mathcal{T}_\epsilon^n(P_X) = \big\{\mathbf{x} :\ |N_X(a \mid \mathbf{x}) - P_X(a)| \le \epsilon P_X(a),\ \forall a \in \mathcal{X}\big\},$$

where $N_X(a \mid \mathbf{x}) := \frac{|\{i\,:\,x_i = a\}|}{n}$ is the empirical distribution of $\mathbf{x}$ on $\mathcal{X}$.

When the underlying probability distribution is fixed we drop the P_X and proceed with $\mathcal{T}_\epsilon^n$ for simplicity. This definition can be extended to joint and conditional probability distributions as well.

Definition 2. Let $\mathcal{X}, \mathcal{Y}$ be a pair of finite sets. For $(X, Y) \sim P_{XY}(x, y)$ on $\mathcal{X} \times \mathcal{Y}$ and some $\epsilon \in (0, 1)$, the set of jointly ϵ-typical n-sequences $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ is defined as [GK12, Chapter 2.4]

$$\mathcal{T}_\epsilon^n(P_{XY}) = \big\{(\mathbf{x}, \mathbf{y}) : \ | N_{XY}(a, b \mid \mathbf{x}, \mathbf{y}) - P(a, b)| \leq \epsilon P(a, b), \text{ for all } (a, b) \in \mathcal{X} \times \mathcal{Y}\big\},$$

where $N_{XY}(a, b \mid \mathbf{x}, \mathbf{y}) := \frac{|\{i \, : \, (x_i, y_i) = (a, b)\}|}{n}$ is the joint empirical distribution of $(\mathbf{x}, \mathbf{y})$ on $\mathcal{X} \times \mathcal{Y}$.

The following lemma, known as the Joint Typicality Lemma, is a prerequisite for understanding the packing lemma which follows right after.

Lemma 1 (Joint typicality lemma). *For an arbitrary joint distribution of* $(X, Y, Z) \sim P_{XYZ}(x, y, z)$ *and* $0 < \epsilon' < \epsilon$, *there exists* $\delta(\epsilon) > 0$ *that tends to zero as* $\epsilon \to 0$ *such that the following statements hold:*

1. *If* $(\mathbf{a}, \mathbf{b})$ *is a pair of arbitrary sequences and* $C^n \sim \prod_{i=1}^n P_{Z|X}(c_i|a_i)$ *then*

$$\Pr\{(\mathbf{a}, \mathbf{b}, C^n) \in \mathcal{T}_\epsilon^n(P_{XYZ})\} \leq 2^{-n(I(Y;Z|X) - \delta(\epsilon))}$$

2. *If* $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}_{\epsilon'}^n$ *and* $C^n \sim \prod_{i=1}^n P_{Z|X}(c_i|x_i)$, *then for* n *sufficiently large,*

$$\Pr\{(\mathbf{x}, \mathbf{y}, C^n) \in \mathcal{T}_\epsilon^n(P_{XYZ})\} \geq 2^{-n(I(Y;Z|X) + \delta(\epsilon))}$$

The proof is given in [GK12, Chapter 2.5].

2.2.2. Packing and covering lemmas

Packing lemma is the key tool to bound the probability of decoding error. The bound is then used to prove achievability results.

Lemma 2 (Packing lemma). *Let* $(U, X, Y) \sim P_{UXY}(u, x, y)$ *and let* $(A^n, B^n) \sim P_{AB}(a^n, b^n)$ *be a pair of arbitrarily distributed random sequences, not necessarily distributed according to* $\prod_{i=1}^n P_{UY}(a_i, b_i)$. *Let* $X^n(m)$, $m \in \mathcal{A}$ *where* $|\mathcal{A}| \leq 2^{nR}$, $R \in \mathbb{R}_+$, *be random sequences each distributed according to* $\prod_{i=1}^n P_{X|U}(x_i|a_i)$. *Further assume that* $X^n(m)$, $m \in \mathcal{A}$ *is pairwise conditionally independent of* B^n *given* A^n, *but is arbitrarily dependent on other*

$X^n(m)$ sequences. Then, there exists $\delta(\epsilon)$ that tends to zero as $\epsilon \to 0$ such that

$$\lim_{n\to\infty} \Pr\big\{(A^n, X^n(m), B^n) \in \mathcal{T}_\epsilon^n \textit{ for some } m \in \mathcal{A}\big\} = 0$$

if $R < I(X;Y|U) - \delta(\epsilon)$.

The proof is given in [GK12, Chapter 3.2].

Lemma 3 (Covering lemma)**.** *Let $(U,X,Y) \sim P(u,x,y)$ and $0 < \epsilon' < \epsilon$. Let $(U^n, X^n) \sim P(u^n, x^n)$ be a pair of random sequences with $\lim_{n\to\infty} \Pr\{(U^n, X^n) \in \mathcal{T}_{\epsilon'}^n(U,X)\} = 1$, and let $Y^n(m), m \in \mathcal{A}$ where, $|\mathcal{A}| \geq 2^{nR}$, $R \in \mathbb{R}_+$ be random sequences, conditionally independent of each other and of X^n given U^n, each distributed according to $\prod_{i=1}^n P_{Y|U}(y_i|u_i)$. Then there exists $\delta(\epsilon)$ that tends to zero as $\epsilon \to 0$ such that:*

$$\lim_{n\to\infty} \Pr\big\{(U^n, X^n, Y^n(m)) \notin \mathcal{T}_\epsilon^n \textit{ for all } m \in \mathcal{A}\big\} = 0$$

if $R > I(X;Y|U) + \delta(\epsilon)$.

The proof is given in [GK12, Chapter 3.7].

2.2.3. Kullback-Leibler divergence

The Kullback-Leibler divergence of two probability distributions P and Q defined on a set $\mathcal{A}$ is given by,

$$\mathrm{D}(P\|Q) := \begin{cases} \sum_{a\in\mathcal{A}} P(a) \log \frac{P(a)}{Q(a)} & \text{if } P \ll Q \\ +\infty & \text{if } P \not\ll Q \end{cases} \tag{2.1}$$

where, $P \ll Q$ means: $Q(a) = 0$ implies $P(a) = 0$ for $a \in \mathcal{A}$, while $P \not\ll Q$ is used if the implication does not hold.

2.2.4. Set of Π-type sequences

Definition 3. Let $\mathcal{P}_\mathcal{X}$ be the set of all probability distributions defined on a finite alphabet $\mathcal{X}$. For a distribution $\Pi \in \mathcal{P}_\mathcal{X}$, the set of all Π-type sequences $\mathbf{x} \in \mathcal{X}^n$ is

$$\mathcal{K}_\Pi := \big\{\mathbf{x} \,:\, \Pi(a) = N_X(a|\mathbf{x}) \;\forall a \in \mathcal{X}\big\}.$$

The following lemma is beneficial in the achievability proof in Chapter 3.

Lemma 4. *For all* Π*-type sequences of* $\mathbf{x} \in \mathcal{X}$ *and the probability distribution* $Q_X \in \mathcal{P}_{\mathcal{X}}$ *we have*

$$Q_X^n(\mathbf{x}) = 2^{-n(D(\Pi\|Q_X)+H(\Pi))} \qquad \textit{if} \quad \mathbf{x} \in \mathcal{K}_\Pi.$$

The proof is provided in [CK82][Chapter 2].

2.3. The wiretap channel

The main objective of this section is to concisely review a coding scheme that achieves positive secrecy rate in a wiretap channel. This will help to obtain insights in the results of the next chapter. A comprehensive proof of the converse and of the secrecy leakage is offered in [Wyn75, GK12].

The term 'wiretap channel' was coined by Wyner in [Wyn75], and the presented results here are due to Wyner [Wyn75] as well as Csiszár and Körner [CK78]. The wiretap channel is a three party communication channel with a sender (Alice) and two receivers; one of which is the legitimate (Bob) and the other is an eavesdropper (Eve), as depicted in Figure 2.2. More precisely we consider a discrete memoryless (DM) multi-terminal channel with an input alphabet $\mathcal{X}$ for the sender and output alphabets $\mathcal{Y}$ and $\mathcal{Z}$ at Bob's receiver and Eve's receiver, respectively. The channel is assumed to be memoryless in the sense that

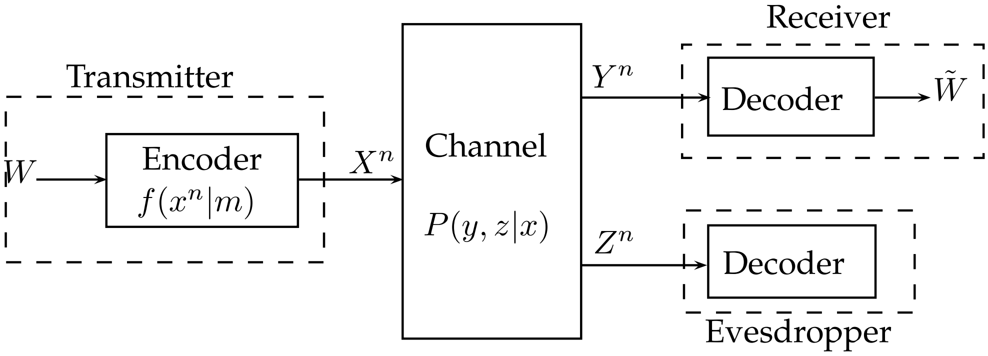

Figure 2.2.: The schematic of the network for the Wiretap channel

$$P(\mathbf{y},\mathbf{z}|\mathbf{x}) = \prod_{i=1}^{n} P(y_i, z_i \mid x_i), \tag{2.2}$$

for each, $\mathbf{x} \in \mathcal{X}^n, \mathbf{y} \in \mathcal{Y}^n$, and $\mathbf{z} \in \mathcal{Z}^n$. We consider a stochastic encoder. Intuitively, the stochastic encoder can be seen as a source of randomness that is needed to confuse the eavesdropper.

Definition 4. The stochastic encoder is a mapping $f_w : \mathcal{M} \to \mathcal{P}(\mathcal{X}^n)$ such that,

$$\sum_{\mathbf{x} \in \mathcal{X}^n} f_w(\mathbf{x} \mid m) = 1, \quad \forall\, m \in \mathcal{M} \tag{2.3}$$

where $\mathcal{M} := \{1, \ldots, M\}$ is the message set used by Alice and $\mathcal{P}(\mathcal{X}^n)$ denotes the set of all probability distributions on $\mathcal{X}^n$.

We assume that the messages are uniformly distributed over $\mathcal{M}$.

Definition 5. The decoder of Bob is a mapping $g_w : \mathcal{Y}^n \to \mathcal{M}$ from the set of channel outputs $\mathcal{Y}^n$ to the message set.

Before presenting the coding schemes we require to introduce the performance metrics. The probability of error is

$$P_{e^{(n)}} := \frac{1}{M} \sum_{m \in \mathcal{M}} \Pr[g_w(\mathbf{y}) \neq m \mid m \text{ is sent}].$$

A code $(n, M, P_{e^{(n)}})$ is consist of an encoder f_w, a decoder g_w, and a message set $\mathcal{M} := \{1, \ldots, M\}$.

Definition 6 (Weak secrecy)**.** [MW00] The system is weakly secure if for all $\epsilon > 0$ there exists a natural number $N(\epsilon)$ such that for all $n > N(\epsilon)$ we have

$$\frac{1}{n} I(W; Z^n) \leq \epsilon. \tag{2.4}$$

The mutual information in (2.4) is defined according to the following probability distribution

$$P(m; \mathbf{z}) = \frac{1}{M} \sum_{\mathbf{x} \in \mathcal{X}^n} f_w(\mathbf{x} \mid m) P(\mathbf{z} \mid \mathbf{x}).$$

Definition 7 (Strong secrecy)**.** [Mau94, MW00] The system is strongly secure if for all $\epsilon > 0$ there exists a natural number $N(\epsilon)$ such that for all $n > N(\epsilon)$ we have

$$I(W; Z^n) \leq \epsilon. \tag{2.5}$$

Definition 8. The rate R is achievable if for all $\epsilon > 0$, $\epsilon' > 0$ and $\delta > 0$ there exists a code $(n, M, P_{e^{(n)}})$, such that each of the followings holds for sufficiently large n

$$\begin{aligned} P_{e^{(n)}} &< \epsilon' \\ \frac{1}{n} I(W; Z^n) &\leq \epsilon \\ \frac{1}{n} \log M &< R - \delta. \end{aligned}$$

Definition 9. The secrecy capacity is defined as the supremum of all of the achievable secrecy rates R.

Theorem 1. *The secrecy capacity of the discrete memoryless wiretap channel is*

$$C_s = \max_{P(u,x)} (I(U,Y) - I(U,Z))$$

where $\mathcal{U}$ *is an auxiliary set with* $|\mathcal{U}| \leq |\mathcal{X}|$.

Proof. The achievability proof is presented in the following [GK12].

The coding scheme

For a fixed $P_U(u)$ we generate independently and randomly 2^{nR_t} sequences from $P_U^n(u) = \prod_{i=1}^n P_U(u_i)$. We cluster them into 2^{nR} bins each with $2^{n(R_t-R)}$ sequences $u^n(l)$, $l \in [(m-1)2^{n(R_t-R)} + 1 : m2^{n(R_t-R)}]$. Each bin serves as a subcodebook $\mathcal{C}(m)$ where $m \in [1 : 2^{nR}]$ is the message. In order to send a message m, from the corresponding bin $\mathcal{C}(m)$ we choose a sequence $u^n(l)$ at random. For simplicity we select $l = 1$ [2]. Then we generate the codeword according to $X^n(m) \sim \prod_{i=1}^n P_{X|U}(x_i|u_i(1))$ and transmit it. At the decoder side, we find a message such that $(u^n(l), y^n) \in \mathcal{T}_\epsilon^{(n)}$ for some $u^n(l) \in \mathcal{C}(\hat{m})$.

Reliable decoding

Two sorts of error might arise. The first one $\mathcal{E}_1 := \{(u^n(l), y^n) \notin \mathcal{T}_\epsilon^{(n)}\}$. Based on the definition of the typicality set, and the law of large numbers $\Pr(\mathcal{E}_1) \to 0$ as $n \to \infty$. The second type of error is $\mathcal{E}_2 := \{(u^n(l), y^n) \in \mathcal{T}_\epsilon^{(n)},\ u^n(l) \notin \mathcal{C}(m)\}$. Here we use the Packing Lemma since the codewords satisfy the terms of the packing lemma on the mutual independence. Therefore, $\Pr(\mathcal{E}_2) \to 0$ as $n \to \infty$ if $R_t < I(U; Y) - \delta(\epsilon)$.

Equivocation rate

Let's $m = 1$ be the transmitted message and the randomly chosen index in $\mathcal{C}(1)$ is also $l = 1$. We assume that the eavesdropper knows $\mathcal{C}$ in advance. To prove secrecy we bound

---

[2] without loss of generality (w. l. g.)

the average equivocation rate, which is the information leakage averaged over the entire codebooks,

$$\frac{1}{n} I(M; Z^n). \tag{2.6}$$

This has been done in [GK12][22.1.1] with extensive detail. It turns out that

$$\limsup_{n\to\infty} \frac{1}{n} I(M; Z^n) \leq \delta(\epsilon) \tag{2.7}$$

holds for $\delta(\epsilon) > 0$ and $R_t - R < I(U; Z)$. This concludes the proof of achievability since there exist a code that if $R < I(U; Y) - I(U; Z) - \delta(\epsilon)$ then $\Pr(\mathcal{E}_2)$ goes to zero and the secrecy criterion is fulfilled. ☐

3. Achievable strong secrecy rate region of broadcast and interference channel

In this chapter we use information-theoretic tools, which are reviewed in the previous chapter, to further develop some results on secrecy achievable rate regions of three scenarios: a discrete memoryless broadcast channel with two confidential messages (DMBCC), a discrete memoryless interference channel with two confidential messages (DMICC), a discrete memoryless untrusted relay channels (DMURC). Our main concern here is privacy rather than security. Privacy problems consider the leakage of one user's data to another user, whereas in security problems we consider leakage of information to an unknown adversary called the eavesdropper.

The DMBCC consists of two receivers and a sender, which has one private message for each receiver. The sender has to transmits the private message via a discrete memoryless broadcast channel reliably to the intended receiver, while keeping it secret from the unauthorized receiver. As for the second model, the DMICC, we assume two senders and two receivers communicating in a discrete memoryless interference channel where each sender has a message for one intended receiver, while treating the other receiver as an eavesdropper. The DMICC has applications in cognitive radio scenario, where both reliability and security of two communication parties become prominent. Finally, the DMURC, is consist of a sender and a receiver communicating through two decode and forward (DF) relays. We further assume there is no direct link between the sender and the receiver. The relays are untrusted, in a sense that they might be hijacked by an adversary, however, they commit to their task of relaying. From information-theoretic point of view the DMURC can be decomposed to a DM broadcast channel that is cascaded to a DM multiple access channel (MAC). This problem is challenging since the relays possess all the information the sender transmits to the receiver. The receiver must be able to reliably decode the message, whereas the relays must be unable to draw any conclusion about the message.

We consider a stronger notion of secrecy, the so called *effective secrecy*. Effective secrecy essentially includes stealth and strong secrecy together. Stealth or covert communication is roughly the situation where the eavesdropper can not distinguish if a meaningful communication is undergoing or not. The formal definition relates to hypothesis testing as well. This topic, however, is not the scope of this dissertation and is a by-product of the effective secrecy.

3.1. Background and contributions

3.1.1. Related work

Broadcast channels are single-input-multi-output channels where a transmitter sends data to many users. The broadcast channel's capacity is unknown in general form, however, for some special cases there are well defined capacity regions. For instance in the case of two users, if one transmitter-user channel is degraded with respect to the other transmitter-user channel then superposition coding obtains capacity region thoroughly [Gal74]. In terms of security, researchers consider different models for broadcast scenario, which are analysed mostly based on the seminal works of Wyner [Wyn75] and Csiszár and Körner [CK78]. The authors in [EU12] consider multiple receivers each with different degradation channels and an external eavesdropper. In this scenario the goal is to hide information from the eavesdropper only; Other users can decode each others messages. On the other hand, the authors in [LMSY08, LP09] assume two users with two confidential messages each intended for one user while kept secret from the other user; In other words, user privacy is considered. This is an extension of Csiszár and Körner's model to the case with two confidential messages and no common message. An inner and outer bound on the achievable secrecy region is derived in [LMSY08]. A rather complete model with confidential messages and one common message has been considered in [XCC09].

The interference channel is a rather more complicated scenario than the broadcast channel to investigate information-theoretically. Han and Kobayashi in [HK81] provided the closest result to the capacity of the interference channel. The authors in [ETW08] further developed more tractable Han-Kobayashi type outer-bounds. As for secrecy rate region, the authors in [LSBP$^+$07, HY09a, LMSY08] have published some results on the matter. The authors in [LSBP$^+$07] considered a two transmitter two receiver scenario with two common messages from each transmitter to both of the receivers and one confidential message at one of the transmitters. Liu et al. extends the model to two confidential messages (each transmitter has a confidential message for one receiver) criterion with weak secrecy. The strong secrecy using structural codes has been studied by [HY09a], however with only one confidential message intended for a receiver.

The authors in [HY13, HY10b, HY09b, HY10a, KMS13] have considered the untrusted relay problem. Untrusted relay is a term used to described a relay channel where the relay obeys its tasks while potentially eavesdropping the data. A variety of relaying strategies, i.e. amplify-and-forward (AF) [HMS13], decode-and-forward (DF), compute and forward (CF) [VKT15,RSEJ15] and, more recently module-and-forward (MF) [ZFPP15], have been investigated by researchers. Under the direct link availability condition, the authors in [HY10b, HY10a] have demonstrated that the untrusted AF and the compress-and-forward relaying strategies obtain positive secrecy rate. On the other hand with no direct link available, i.e. all of the data from the transmitter has to go through the relay to reach the destination, has been studied by the authors in [HY09b]. In order to increase the equivocation rate at the relay, thus providing secrecy, the destination produces pseudo noise. In most of the research works aforementioned, either the legitimate receiver (destination) or an external node (e.g. 'friendly' jammer) broadcast some 'pseudo noise' to confuse the relay in the receiving phase; and thus increase the equivocation rate at the relay. Having received the superposition of the pseudo noise and the message, the relay then transmits the superposition to the destination. The destination is able to decode the actual message by knowing the pseudo noise sequence. This method exploit the interference and its natural superposition characteristics of the wireless channel to hide the message. A bidirectional untrusted relay is considered in [ZFPP15, VKT15], where the problem is handled with lattice codes.

The secrecy criterion used in [LMSY08], as well as in [Wyn75, CK78], is the *normalized* mutual information ($\frac{1}{n}I(X;Z) < \epsilon$) between the message and the received signal at the unintended user. Strong secrecy, however, demands the *unnormalized* mutual information ($I(X;Z) < \epsilon$) to be bounded by a small number for a given code length n. Strong secrecy, first introduced by Maurer and Wolf [Mau94], provides a practical meaning to the secrecy level of a system whereas weak secrecy has not yet been linked to any practical meaning. Strong secrecy has been studied by many researchers thus far [Mau94,BL13,Csi96,HK14], of which we focus on the work of Hou and Kramer in [HK14].

3.1.2. Results and contributions

The main contributions of this chapter are provided in the following.

1. We characterize an achievable rate region with strong secrecy for broadcast channels with two confidential messages for each user. Our results are an extension of those in [LMSY08] for weak secrecy to strong secrecy. We utilize the results in [HK13, HK14] to inaugurate an upperbound on $\mathrm{D}(P_{X,Z}\|P_X Q_Z)$. This result provides a bound that is even stronger than the notion of strong secrecy, since it includes the stealth communication as well. To be more specific we consider, $\mathrm{D}(P_{X,Z}\|P_X Q_Z) = I(X;Z) + D(P_Z\|Q_Z)$, in which all of the terms are non-negative based on the definition. Therefore, bounding $\mathrm{D}(P_{X,Z}\|P_X Q_Z)$ addresses the strong

secrecy and the stealth communication whereas bounding only $I(X;Z)$ points to the strong secrecy only. We use the encoding suggested by [LMSY08], which imposes a joint distribution on the channel inputs corresponding to each user, however, to address the stealth communication the channel input of the users must be independent from each other. This results in some complexity, which is handled in this chapter. It turns out that the secrecy rate region proposed by [LMSY08] holds for strong secrecy criterion as well.

2. We upgrade the weak secrecy criterion in [LMSY08] to the strong secrecy one for IC with two confidential messages. The analysis is quite similar to that of BC, however, owing to the disjoint independent encoders the analysis does not require considering joint and independent distributions for the input channels; and thus it is simpler to study. It turns out, the achievable rate region with the strongly secure criterion is equivalent to that proposed by [LMSY08] for the weakly secure criterion, i.e. strong secrecy at no cost.

3. The untrusted DF relay channel without an external helper, the destinations help or the direct link between the sender and the destination, has not been studied yet. Building upon the results developed for BC with confidential messages, we characterize the achievable secrecy rate region for the DF relay channel with two untrusted relays. We propose to use a precoder at the sender to cope with the untrusted characteristic of the relays. We further extend this to multiple relays and propose a construction method to design the precoder.

3.2. Problem statement and system model

3.2.1. The broadcast channel with confidential messages

The broadcast channel consists of a transmitter with two receivers each with one private message as depicted in Figure 3.1. We assume the channel is discrete memoryless, i.e.

$$P(\mathbf{y}_1, \mathbf{y}_2|\mathbf{x}) = \prod_{i=1}^{n} P(y_{1i}, y_{2i}|x_i), \tag{3.1}$$

where, $\mathbf{x} \in \mathcal{X}^n, \mathbf{y}_1 \in \mathcal{Y}_1^n$, and $\mathbf{y}_2 \in \mathcal{Y}_2^n$ and $\mathcal{X}$ and $\mathcal{Y}_1, \mathcal{Y}_2$ are the input and output channel alphabets, respectively, which are assumed to be finite sets.

The joint probability distribution governing the discrete memoryless broadcast channel (DMBCC), for a given $P_X \in \mathcal{P}(\mathcal{X})$, is written as

$$P_{XY_1Y_2}(x, y_1, y_2) = P_X(x)P(y_1, y_2|x). \tag{3.2}$$

Definition 10. The stochastic encoder at the transmitter is given by,

$$f : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{P}(\mathcal{X}^n), \tag{3.3}$$

and further, we have,

$$\sum_{\mathbf{x}\in\mathcal{X}^n} f(\mathbf{x}|m_1, m_2) = 1, \ \ \forall(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2 \tag{3.4}$$

where, $m_1 \in \mathcal{M}_1 := \{1, \dots, M_1\}$ and $m_2 \in \mathcal{M}_2 := \{1, \dots, M_2\}$ are the message sets for receiver 1 and receiver 2, respectively.

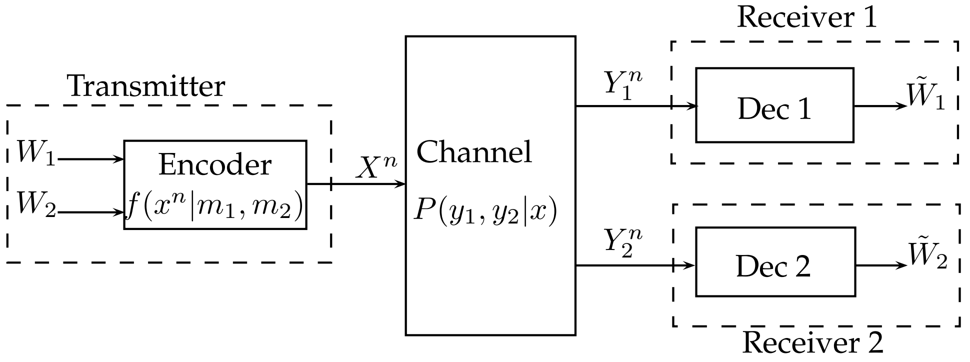

Figure 3.1.: The schematic of the network for Broadcast channel with confidential messages

Definition 11. The decoder at the tth node, $t \in \{1, 2\}$, is defined as

$$g_t : \mathcal{Y}_t^n \to \mathcal{M}_t.$$

Definition 12 (Strong Secrecy [Mau94])**.** For any given $\epsilon > 0$ there is a $N(\epsilon)$ such that for all $n \geq N(\epsilon)$,

$$\max_{m_2\in\mathcal{M}_2} I(W_1; Y_2^n|W_2 = m_2) \leq \epsilon, \tag{3.5}$$

$$\max_{m_1\in\mathcal{M}_1} I(W_2; Y_1^n|W_1 = m_1) \leq \epsilon \tag{3.6}$$

where, the random variables W_1, W_2 are distributed uniformly over $\mathcal{M}_1$, $\mathcal{M}_2$,, respectively. Besides, the random variables Y_1^n, and Y_2^n are drawn from $\mathcal{Y}_1^n$, and $\mathcal{Y}_2^n$, respectively. The mutual information quantities are evaluated with respect to the following distribution

$$P_{W_1W_2Y_1^nY_2^n}(m_1,m_2,\mathbf{y}_1,\mathbf{y}_2) = \frac{1}{M_1}\frac{1}{M_2}\sum_{\mathbf{x}\in\mathcal{X}^n} f(\mathbf{x}|m_1,m_2)P(\mathbf{y}_1,\mathbf{y}_2|\mathbf{x}),$$

with the stochastic encoder given in definition (10).

Each message m_1 (m_2) has to be decoded reliably by receiver 1 (2) while kept secret from receiver 2 (1). The probability of error at each node t is

$$P_{e\,t}^n = \frac{1}{M_1M_2}\sum_{(m_1,m_2)\in\mathcal{M}_1\times\mathcal{M}_2} \Pr[g_t(Y_t^n) \neq m_t|(m_1,m_2) \text{ is sent}], \quad t\in\{1,2\}.$$

Definition 13. A rate pair (R_1,R_2) is said to be achievable, if for every $\epsilon' > 0$, $\epsilon > 0$, $\delta > 0$, there exists a code $(M_1,M_2,n,P_{(e^n,\,1)},P_{(e^n,\,2)})$ for sufficiently large n ,in addition to (3.5) and (3.6), we have

$$P_{e\,t}^n \leq \epsilon' \quad \text{and} \quad \frac{1}{n}\log M_t \geq R_t - \delta, \quad t\in\{1,2\}.$$

3.2.2. The interference channel with confidential messages

Figure 3.2 illustrates two transmitters and two receivers communicating through an interference channel. The channel input and output alphabets are chosen from the finite sets $\mathcal{X}_1, \mathcal{X}_2$ and $\mathcal{Y}_1, \mathcal{Y}_2$, respectively. For given probability distributions $P_{X_t} \in \mathcal{P}(\mathcal{X}_t),\ t\in\{1,2\}$ and channel $P(y_1,y_2|x_1,x_2),\ Y_t\in\mathcal{Y}_t, X_t\in\mathcal{X}_t, t\in\{1,2\}$ the joint probability distribution governing the discrete memoryless interference channel (DMIC) is:

$$P_{X_1X_2Y_1Y_2}(x_1,x_2,y_1,y_2) = P(y_1,y_2|x_1,x_2)P_{X_1}(x_1)P_{X_2}(x_2).$$

By "memoryless" we mean:

$$P(\mathbf{y}_1,\mathbf{y}_2|\mathbf{x}_1,\mathbf{x}_2) = \prod_{i=1}^{n} P(y_{1i},y_{2i}|x_{1i},x_{2i}),$$

where, $\mathbf{x}_t\in\mathcal{X}_t^n$, $\mathbf{y}_t\in\mathcal{Y}_t^n$, and $t\in\{1,2\}$.

Definition 14. The stochastic encoder at the tth transmitter is defined as,

$$f_t : \mathcal{M}_t \to \mathcal{P}(\mathcal{X}_t^n),$$

and further, we have,

$$\sum_{\mathbf{x}_t \in \mathcal{X}_t^n} f_t(\mathbf{x}_t|m_t) = 1, \;\; \forall m_t \in \mathcal{M}_t, \; t \in \{1,2\},$$

where, $\mathcal{M}_1 := \{1,\ldots,M_1\}$ and $\mathcal{M}_2 := \{1,\ldots,M_2\}$ are the message sets for transmitters 1 and 2, respectively.

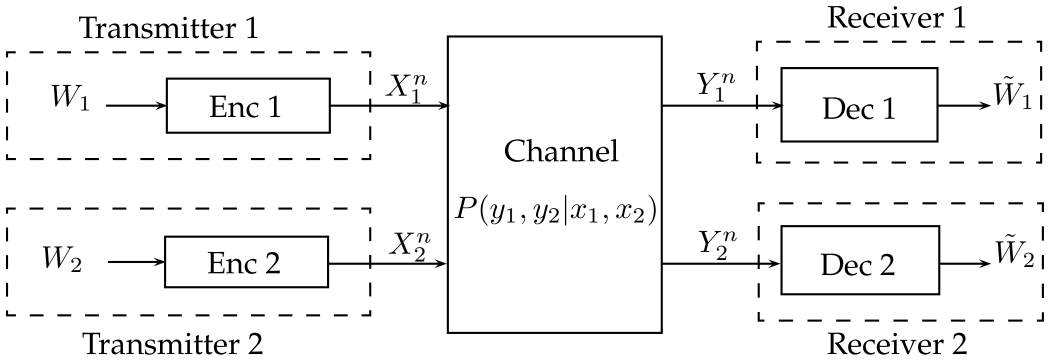

Figure 3.2.: The schematic of the interference channel with input/output distributions.

Definition 15. The decoder at the tth receiver, $t \in \{1,2\}$, is a function, given by

$$l_t : \mathcal{Y}_t^n \to \mathcal{M}_t.$$

The strong secrecy criterion imposes the following, if for every $\epsilon > 0$ there is a $N(\epsilon)$ such that for all $n \geq N(\epsilon)$,

$$I(W_1;Y_2^n) \leq \epsilon \;\text{ and }\; I(W_2;Y_1^n) \leq \epsilon \tag{3.7}$$

where W_1 and W_2 are RVs distributed uniformly over $\mathcal{M}_1$ and $\mathcal{M}_2$ and further, $Y_1^n \in \mathcal{Y}_1^n$ and $Y_2^n \in \mathcal{Y}_2^n$ are RV's corresponding to the output of the channel. The mutual information values are computed with respect to the following distribution

$$P_{W_1W_2Y_1^nY_2^n}(m_1,m_2,\mathbf{y}_1,\mathbf{y}_2) = \frac{1}{M_1}\frac{1}{M_2}\sum_{\mathbf{x}_1\in\mathcal{X}^n,\mathbf{x}_2\in\mathcal{X}^n} f_1(\mathbf{x}_1|m_1)f_2(\mathbf{x}_2|m_2)P(\mathbf{y}_1,\mathbf{y}_2|\mathbf{x}_1,\mathbf{x}_2).$$

We do not condition on the decodability of the legitimate message, i.e. $I(W_1;Y_2|W_2)$, since, the codewords are produced independently from each other and there is no cooperation assumed between the transmitters.

The probability of error at each node t is

$$P^n_{e\,t} = \frac{1}{M_1 M_2} \sum_{m_1,m_2} \Pr[l_t(Y^n_t) \neq m_t | (m_1, m_2) \text{ is sent}], \quad t \in \{1, 2\}.$$

Definition 16. A rate pair (R_1, R_2) is achievable if for any $\delta, \epsilon', \epsilon > 0$ there exists a code $(M_1, M_2, P_{(e^n,\,1)}\ P_{(e^n,\,2)}, n)$ such that (3.7) holds, and further,

$$P^n_{e\,t} \leq \epsilon' \quad \frac{1}{n} \log M_t \geq R_t - \delta \quad t \in \{1, 2\},$$

holds for n sufficiently large.

The transmitter t sends a confidential message m_t intended to the receiver t, while keeping it secret from the other receiver $r \in \{1, 2\}, r \neq t$.

3.2.3. The decode-and-forward untrusted relay channel

Consider a sender (S) that communicates through a number of decode and forward (DF) relays to a destination (D) as in Figure 3.3. We assume there is no direct channel between S and D. Although the relays obey to DF relaying paradigm, they are untrusted in a sense that they might eavesdrop on the content of the data. Further, there are no channels between the relays, thus, the relays do not cooperate to enhance the relaying task or to advance eavesdropping capabilities. According to the model in Figure 3.4, all of the data has to go through the untrusted relays. We further restrict ourselves to avoid using any side-channel information. For simplicity, we assume there are only two relays ($K = 2$). The finite sets $\mathcal{X}, \hat{\mathcal{Y}}_1, \hat{\mathcal{Y}}_2$ denote the input alphabets and the finite sets $\mathcal{Y}_1, \mathcal{Y}_2\ \mathcal{Z}$ represent the output alphabets. The discrete memoryless untrusted relay channel (DMURC) consists of a DMBCC-phase, a DF-phase, and a discrete memoryless multiple access channel (DMMAC)-phase. The DMURC is memoryless, thus

$$\text{DMBCC-phase, } P(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}) = \prod_{i=1}^{n} P(y_{1i}, y_{2i} | x_i), \text{ and}$$
$$\text{DMMAC-phase, } P(\mathbf{z} | \hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2) = \prod_{i=1}^{n} P(z_i | \hat{y}_{1i}, \hat{y}_{2i}).$$

where, $\mathbf{x} \in \mathcal{X}^n,\ \mathbf{y}_t \in \mathcal{Y}^n_t,\ \hat{\mathbf{y}}_t \in \hat{\mathcal{Y}}^n_t,\ \mathbf{z} \in \mathcal{Z}^n,$ and $t \in \{1, 2\}$.

Definition 17. The encoding process at S consists of

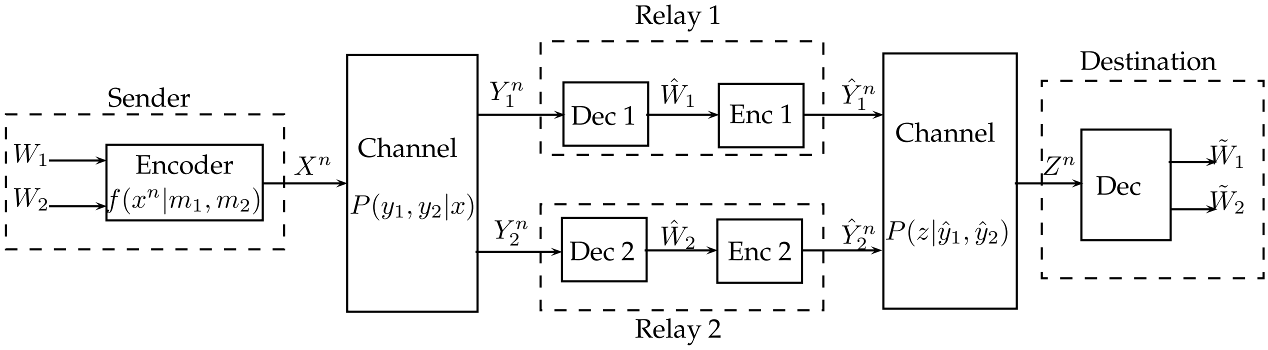

Figure 3.3.: The decode-and-forward untrusted relay channel

1. Two precoders $\theta_t : \mathcal{M}_t \times \mathcal{U}_t \to \hat{\mathcal{M}}_t,\ t \in \{1,2\}$ with the message sets $\mathcal{M}_t := \{1,\dots,M_t\},\ t \in \{1,2\}$ and the auxiliary finite sets $\mathcal{U}_t, t \in \{1,2\}$, and $\hat{\mathcal{M}}_t := \theta_t(\mathcal{M}_t,\mathcal{U}_t)$.
2. A stochastic encoder,

$$\phi : \hat{\mathcal{M}}_1 \times \hat{\mathcal{M}}_2 \to \mathcal{P}(\mathcal{X}^n).$$

3. To send the message pair $(m_1,m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$ the sender selects $(u_1,u_2) \in \mathcal{U}_1 \times \mathcal{U}_2$, and uses $\phi(\theta_1(m_1,u_1),\theta_2(m_2,u_2))$ as the channel input. We further have

$$\sum_{\mathbf{x}\in\mathcal{X}^n} \phi(\mathbf{x}|\theta_1(m_1,u_1),\theta_2(m_2,u_2)) = 1,\ \ \forall (m_1,m_2) \in \mathcal{M}_1 \times \mathcal{M}_2.$$

Remark 1. The two auxiliary sets, $\mathcal{U}_1$ and $\mathcal{U}_2$, assist securing the messages chosen from $\mathcal{M}_1$ and $\mathcal{M}_2$ at the relays as in the following. Since every message is decoded at the DF relays, to provide secrecy we require some sort of randomness (key) encoded with the messages. A common approach [HY10b, HY10a] is to have an external node jam the relays with a pseudo noise that is known to the destination, to confuse the relays. While this method proves effective in securing the message it requires elaborated coordination among the nodes, therefore, we present a simpler way. We introduce some randomness at the sender, which is modelled here as the auxiliary sets $\mathcal{U}_1$ and $\mathcal{U}_2$. We later illustrate how the destination can decode the actual messages. The encoder ϕ is a *stochastic encoder*, which is proven to be effective in achieving positive secrecy rates [CK78].

Definition 18. The decoders at tth relay $t \in \{1,2\}$ is a map $\psi_t : \mathcal{Y}_t^n \to \hat{\mathcal{M}}_t$, where, $\hat{\mathcal{M}}_t := \{1,\dots,\hat{M}_t\}$ are given by Definition 17.

Definition 19. We further define the encoder at the tth relay as a mapping $\varpi_t : \hat{\mathcal{M}}_t \to \hat{\mathcal{Y}}_t^n$.

Definition 20. The decoder at D is a mapping $\upsilon : \mathcal{Z}^n \to \mathcal{M}_1 \times \mathcal{M}_2$.

Definition 21 (Secrecy)**.** For all of the encoding and the decoding processes we impose the following secrecy conditions: For every $\epsilon > 0$ there is a non-negative integer $n \geq N(\epsilon)$ such that

$$I(W_t; \hat{W}_t) < \epsilon, \;\; t \in \{1,2\}, \tag{3.8}$$

$$I(\hat{W}_2; Y_1^n | \hat{W}_1) < \epsilon \tag{3.9}$$

$$I(\hat{W}_1; Y_2^n | \hat{W}_2) < \epsilon \tag{3.10}$$

where $W_t, \hat{W}_t, \;\; t \in \{1,2\}$, are uniformly distributed RVs with values taken from $\mathcal{M}_t, \hat{\mathcal{M}}_t, \;\; t \in \{1,2\}$ respectively; and $Y_t^n, \;\; t \in \{1,2\}$ are RVs at the output of the tth relay.

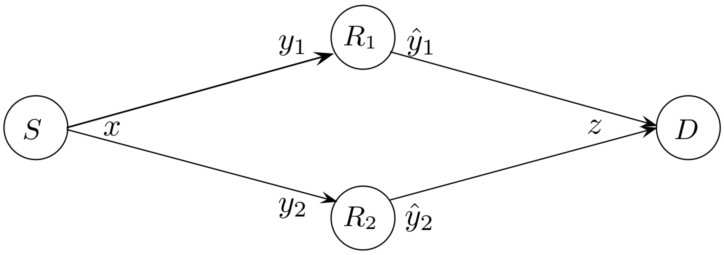

Figure 3.4.: The graph of our network

Remark 2. The criterion $I(W_t; \hat{W}_t) < \epsilon$ ensures that the tth relay can not recover m_t from $\hat{m}_t, \;\; t \in \{1,2\}$, however, it is weaker than the strong secrecy. On the other hand, (3.9) and (3.10) are strongly secure criteria guaranteeing relay 1 (2) can not decode $\hat{m}_2$ ($\hat{m}_1$).

The probability of error at D is given by

$$P_{e\,t}^n = \frac{1}{M_1 M_2} \sum_{m_1, m_2} \Pr[\rho_t(\upsilon(Z^n)) \neq m_t | (m_1, m_2) \text{ is sent}], \;\; t \in \{1,2\}.$$

Definition 22. A rate pair (R_1, R_2) is achievable if for any $\delta, \epsilon', \epsilon > 0$ there exists a code $(M_1, M_2, P_{(e^n,\,1)}\; P_{(e^n,\,2)}, n)$ such that (3.8), (3.9), and (3.10) hold, and further,

$$P_{e\,t}^n \leq \epsilon', \qquad \frac{1}{n} \log M_t \geq R_t - \delta, \qquad t \in \{1,2\},$$

holds for n sufficiently large.

3.3. Strong secrecy for broadcast channels

In this section we present secure rate region with strong secrecy criterion for a DMBCC. The following proposition presents this section's result:

Theorem 2. *The rate pair (R_1, R_2), is achievable for a DMBCC with confidential messages and strong secrecy criterion, if:*

$$\begin{aligned} 0 \leq & R_1,\ 0 \leq R_2 \\ R_1 < & I(V_1; Y_1|U) - I(V_1, X; Y_2|V_2, U) \\ R_2 < & I(V_2; Y_2|U) - I(V_2, X; Y_1|V_1, U) \\ R_1 + R_2 < & I(V_1; Y_1|U) - I(V_1, X; Y_2|V_2, U) + I(V_2; Y_2|U) - I(V_2, X; Y_1|V_1, U) - I(V_1; V_2|U) \end{aligned}$$

for some probability distribution $P_U(u)P_{V_1,V_2|U}(v_1, v_2|u)P_{X|V_1,V_2}(x|v_1, v_2)P(y_1, y_2|x)$, with $P(y_1, y_2|x)$ given as the channel and $P_{V_1V_2|U}$ a probability distribution on a finite set $\mathcal{V}_1 \times \mathcal{V}_2$.

Proof. We prove the achievablity in the following subsections. The achievablity proof borrows techniques inspired by [HK13, LMSY08, Mar79, GP80, HK14].

3.3.1. Coding scheme

We perform double binning encoding on the messages $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$ and send m_t as confidential messages to the tth receiver $t \in \{1, 2\}$.

For $t = 1, 2$ and $R_t, R_t' \in \mathbb{R}_{\geq 0}$ we set $M_t := 2^{nR_t}$, $J_t := 2^{nR_t'}$ and $\mathcal{M}_t := \{1, \ldots, M_t\}$, $\mathcal{S}_t := \{1, \ldots, J_t\}$. Let $R_{co,t} \in \mathbb{R}_{\geq 0}$, $t = 1, 2$, be such that

$$R_{co,t} \geq R_t + R_t' =: R_{t+} \tag{3.11}$$

holds for $t = 1, 2$. We fix a probability distribution $P_{V_1V_2}$ on a finite set $\mathcal{V}_1 \times \mathcal{V}_2$ and a function $f : \mathcal{V}_1 \times \mathcal{V}_2 \to \mathcal{X}$. Define $X := f(V_1, V_2)$. Finally, let $0 < \epsilon < \epsilon'$.

Codebook generation

For $t = 1, 2$ and any $(m_t, s_t) \in \mathcal{M}_t \times \mathcal{S}_t$ we consider random variables $V_t^n(m_t, s_t, k_t)$, $k_t \in \{1, \ldots 2^{n(R_{co,t} - R_{t+})}\}$, which are assumed to be independent and distributed according to $P_{V_t}^n(\mathbf{v}_t) = \prod_{i=1}^n P_{V_t}(v_{t,i})$ for $\mathbf{v}_t \in \mathcal{V}_t^n$.
The realizations of $V_t^n(m_t, s_t, k_t)$ are denoted by $\mathbf{v}_t(m_t, s_t, k_t)$.

For $t = 1, 2$ and $(m_t, s_t) \in \mathcal{M}_t \times \mathcal{S}_t$ find an index pair (k_1, k_2) such that

$$(\mathbf{v}_1(m_1, s_1, k_1), \mathbf{v}_2(m_2, s_2, k_2)) \in \mathcal{T}_\epsilon^n(P_{V_1V_2}),$$

holds. If no such (k_1, k_2) exists then choose $(1, 1)$. Generate code word $\mathbf{x}(m_1, s_1, m_2, s_2)$ as

$$\mathbf{x}(m_1, s_1, m_2, s_2) := f^n(\mathbf{v}_1(m_1, s_1, k_1), \mathbf{v}_2(m_2, s_2, k_2)) \tag{3.12}$$

where f^n denotes the component-wise application of function f.

Encoding

To send the pairs $(m_t, s_t) \in \mathcal{M}_t \times \mathcal{S}_t$, $t = 1, 2$, transmit $\mathbf{x}(m_1, s_1, m_2, s_2)$.

Decoding at the receivers

The decoder t declares that (m_t, s_t) is sent if it is unique pair such that

$$(\mathbf{v}_t(m_t, s_t, k_t), \mathbf{y}_t) \in \mathcal{T}_{\epsilon'}^n(P_{V_tY_t})$$

for some k_t. Otherwise an error is declared.

3.3.2. Error Analysis

The first error arises when the encoder can not find k_1 and k_2 that are jointly typical,

$$E_1 \triangleq \{(\mathbf{v}_1(m_1, s_1, k_1), \mathbf{v}_2(m_2, s_2, k_2)) \notin \mathcal{T}_\epsilon^n(P_{V_tY_t}) \forall k_1, k_2\}.$$

It follows from the mutual covering lemma [GK12] that

$$(R_{co,1} - R_{1+}) + (R_{co,2} - R_{2+}) > I(V_1; V_2) \tag{3.13}$$

implies

$$\Pr(E_1) \leq C \cdot 2^{-nc} \tag{3.14}$$

with positive constants $c, C \in \mathbb{R}$.
In the following we assume without loss of generality that $(\mathbf{v}_1(1, 1, k_1), \mathbf{v}_2(1, 1, k_2))$ is sent.

At the receiver t, the error events are

$$\begin{aligned}
E_{2t} &\triangleq \{(\mathbf{v}_t(1,1,k_t),\mathbf{y}_t) \notin \mathcal{T}^n_{\epsilon'}(P_{V_tY_t})\},\\
E_{3t} &\triangleq \bigcup_{\substack{(m_t,s_t)\neq(1,1)\\ k_t}} \{(\mathbf{v}_t(m_1,s_t,k_t),\mathbf{y}_t) \in \mathcal{T}^n_{\epsilon'}(P_{V_tY_t})\}.
\end{aligned}$$

By the basic properties of typical sequences we have

$$\Pr(E_{2t}) \leq C_1 \cdot 2^{-nc_1} \tag{3.15}$$

with some positive constants $c_1, C_1 \in \mathbb{R}$.
The packing lemma [GK12] shows that

$$R_{co,t} < I(V_t;Y_t) \tag{3.16}$$

for $t = 1,2$ implies

$$\Pr(E_{3t}) \leq C_2 \cdot 2^{-nc_2} \tag{3.17}$$

with positive constants $c_2, C_2 \in \mathbb{R}$. Thus the probability of error of receiver $t = 1,2$ can be bounded as

$$\Pr(E_t) \leq \Pr(E_1) + \Pr(E_{2t}) + \Pr(E_{3t}) \leq C_3 \cdot 2^{-nc_3} \tag{3.18}$$

with suitable positive constants C_3, c_3.
From (3.13), (3.16), and (3.11) we obtain,

$$R_1 + R_1' < I(V_1;Y_1) \tag{3.19}$$

and

$$R_2 + R_2' < I(V_2;Y_2) \tag{3.20}$$

as well as

$$R_1 + R_1' + R_2 + R_2' < I(V_1;Y_1) + I(V_2;Y_2) - I(V_1;V_2). \tag{3.21}$$

The bounds on R_1' and R_2' are derived in the following subsection.

3.3.3. Strong secrecy criterion

Remark 3. For the secrecy analysis we drop the random variable U for the sake of notational simplicity. It can be introduced, if desired, via standard arguments at the end of the proof [CK78, HK14].

Before we further continue, we define a deterministic function $\phi_{\mathcal{T}}$ on pairs $\{\mathbf{v}_1(m_1,s_1,k_1)\}_{k_1=1}^{2^{nR_{co,1}}}$ $\times\{\mathbf{v}_2(m_2,s_2,k_2)\}_{k_2=1}^{2^{nR_{co,2}}}$ that returns a pair (k_1,k_2) such that

$$(\mathbf{v}_1(m_1,s_1,k_1),\mathbf{v}_2(m_2,s_2,k_2))\in\mathcal{T}_\epsilon^n(P_{V_1,V_2}).$$

In the case that there are many such pairs, we choose one arbitrarily. However, if there are no such pairs, the function $\phi_{\mathcal{T}}$ returns (1,1). From now on, we denote the codeword pairs simply $(\mathbf{v}_1(m_1,s_1),\mathbf{v}_2(m_2,s_2))$ based on condition $\phi_{\mathcal{T}}\neq(1,1)$.

We extend the framework in [HK13, HK14] to find a condition that bounds

$$\mathrm{D}(P_{W_1Y_2^n|V_2}||P_{W_1}Q^n_{Y_2|V_2})\leq\xi \tag{3.22}$$

where the left hand side is known as the *effective secrecy*, $\xi>0$ is arbitrarily small, $P_{W_1,Y_2|V_2}$ is the joint distribution of W_1 and Y_2 given $\mathbf{v}_2$, which is expressed as

$$P(\mathbf{y}_2|m_1,\mathbf{v}_2):=\sum_{s_1=1}^{J_1}\frac{1}{J_1}P^n(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2(m_2,s_2)),$$

and further

$$P(\mathbf{y}_2|\mathbf{v}_2):=\sum_{m_1=1}^{M_1}\sum_{s_1=1}^{J_1}\frac{1}{M_1.J_1}P^n(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2(m_2,s_2)).$$

Further, $Q_{Y_2|V_2}$ is distribution of Y_2 given $\mathbf{v}_2$ while no meaningful message is transmitted for W_1,

$$Q^n_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)=\sum_{\mathbf{v}_1}P^n_{V_1}(\mathbf{v}_1)P^n_{Y_2|V_1,V_2}(\mathbf{y}_2|\mathbf{v}_1,\mathbf{v}_2) \tag{3.23}$$

and finally, P_{W_1} is marginal distribution of W_1. The following is useful.

Lemma 5. *We can further expand (3.22) as*

$$\mathrm{D}(P_{W_1,Y_2|\mathbf{v}_2}||Q^n_{Y_2|\mathbf{v}_2}P_{W_1})=I(W_1;Y_2|\mathbf{v}_2)+\mathrm{D}(P_{Y_2|\mathbf{v}_2}||Q^n_{Y_2|\mathbf{v}_2}), \tag{3.24}$$

The proof is given in the Appendix 3.7.1.

Remark 4. The expansion in (3.24), reveals that (3.22) addresses not only the strong secrecy notation by bounding $I(W_1;Y_2|\mathbf{v}_2)$, but also bounds $\mathrm{D}(P_{Y_2|V_2}||Q^n_{Y_2|V_2})$ which corresponds to *stealth* of the communication.

Remark 5. The difference between $P_{Y_2|V_2}$ and $Q^n_{Y_2|V_2}$ is in the absence of a message (a meaningful message) W_1. This in terms of the coding scheme it can be elaborated as in the following. For the case of $Q^n_{Y_2|V_2}$, the encoder only has secret message for the receiver 2, thus $\mathbf{v}_2$ produced according to a message m_2. On the contrast, the codeword $\mathbf{v}_1$ is chosen completely arbitrarily, since we do not communicate any meaningful message to the first receiver.

Based on chain rule for informational divergence we have,

$$\mathrm{D}(P_{W_1Y_2^n|V_2}||P_{W_1}Q^n_{Y_2|V_2}) = \mathrm{D}(P_{W_1}||P_{W_1}) + \mathrm{D}(P_{Y_2^n|W_1V_2}||Q^n_{Y_2|V_2}|P_{W_1}), \tag{3.25}$$

where, according to the definition of $\mathrm{D}(\cdot\|\cdot)$ in (2.1), the first term above equals to zero, thus we proceed with $\mathrm{D}(P_{Y_2|V_2,W_1}||Q^n_{Y_2|V_2}|P_{W_1})$. The following lemma provides an upperbound on $\mathrm{D}(P\|Q)$, which is useful for the rest of the proof.

Lemma 6. *For probability distributions P and Q, defined on a finite set A, with $P \ll Q$, we have,*

$$\mathrm{D}(P\|Q) \le \log\frac{1}{\pi_Q}, \tag{3.26}$$

where, $\pi_Q = \min\{Q(a) : Q(a) > 0\}$.

The proof is given in the Appendix 3.7.6.

Splitting the expectation with respect to (V_1, V_2) of $\mathrm{D}(P_{Y_2|V_2,W_1}||Q^n_{Y_2|V_2}|P_{W_1})$ into two cases of $\phi_\tau = (1,1)$ and $\phi_\tau = (1,1)$, we get,

$$\begin{aligned}\mathrm{E}[\mathrm{D}(P_{Y_2^n|V_2W_1}||Q^n_{Y_2|V_2}|P_{W_1})] &= \mathrm{E}\big[\mathbb{1}(\phi_\tau = (1,1))\mathrm{D}(P_{Y_2^n|V_2W_1}||Q^n_{Y_2|V_2}|P_{W_1})\\ &\quad + \mathrm{E}\big[\mathbb{1}(\phi_\tau \neq (1,1))\mathrm{D}(P_{Y_2^n|V_2W_1}||Q^n_{Y_2|V_2}|P_{W_1}),\end{aligned} \tag{3.27}$$

where, $\mathbb{1}(\cdot)$ is an indicator function. The first term of right hand side (3.27) yields,

$$\begin{aligned}\mathrm{E}\big[\mathbb{1}(\phi_\tau = (1,1))\mathrm{D}(P_{Y_2^n|V_2W_1}||Q^n_{Y_2|V_2}|P_{W_1})\big] &\le \Pr\big(\phi_\tau = (1,1)\big)\log\frac{1}{(\pi_{Q^n_{Y_2|V_2}})^n} &(3.28)\\ &= \Pr\big(\phi_\tau = (1,1)\big)(-n)\log(\pi_{Q^n_{Y_2|V_2}})\\ &\le 2^{-n\alpha} &(3.29)\end{aligned}$$

for all sufficiently large n, where, we apply Lemma 6 in (3.28). The last inequality, (3.29), comes from the mutual covering lemma, where $\alpha > 0$ is a constant independent of n. With (3.25) and (3.29), therefore, we have:

$$\mathrm{E}[\mathrm{D}(P_{Y_2^n|V_2W_1}||Q^n_{Y_2|V_2}|P_{W_1})] \leq \mathrm{E}\big[\mathbb{1}(\phi_\mathcal{T} \neq (1,1))\mathrm{D}(P_{Y_2^n|V_2W_1}||Q^n_{Y_2|V_2}|P_{W_1}) + 2^{-n\alpha}. \quad (3.30)$$

Hence, in the following we focus on bounding the first term in the right hand side of (3.30).

Lemma 7. *Taking an expectation over Y_2^n, V_1, and S_1 yields,*

$$\mathrm{E}[\mathrm{D}(P_{Y_2^n|W_1,\mathbf{v}_2}||Q^n_{Y_2|\mathbf{v}_2}|P_{W_1})] \leq \mathrm{E}\left[\log\left(\frac{P(Y_2^n|V_1^n,\mathbf{v}_2)}{J_1 Q^n_{Y_2|V_2}(Y_2^n|\mathbf{v}_2)} + 1\right)\right] \quad (3.31)$$

The proof is provided in Appendix 3.7.2. Before proceeding with the bounds on the informational divergence, we introduce the following lemma.

Lemma 8. *Let $P_{V_1V_2Y_2}(v_1,v_2,y_2) = P_{V_1V_2}(v_1,v_2)P_{Y_2|V_1V_2}(y_2|v_1,v_2)$ be a probability distribution on $\mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{Y}_2$. For $\epsilon \in (0,\frac{1}{2})$ and distribution*

$$Q(y_2|v_2) := \sum_{v_1} P_{V_1}(v_1)P_{Y_2|V_1V_2}(y_2|v_1,v_2) \quad (3.32)$$

it holds for $(\mathbf{v}_2,\mathbf{y}_2) \in \mathcal{T}^n_\epsilon(P_{V_2Y_2})$ that

$$Q^n(\mathbf{y}_2|\mathbf{v}_2) \geq 2^{-n(H(Y_2|V_2)+I(V_1;V_2)+\delta(\epsilon))} \quad (3.33)$$

with $\delta(\epsilon) > 0$ and $\lim_{\epsilon\to 0}\delta(\epsilon) = 0$.

The proof is given in the Appendix 3.7.3. The equation (3.31) can be split up as follows: (Recall that the expectation is taken over those codewords with $\phi_\mathcal{T} \neq (1,1)$)

$$\sum_{\substack{\mathbf{v}_1\\(\mathbf{v}_1,\mathbf{v}_2)\in\mathcal{T}_{\epsilon_0}}} P^n_{V_1}(\mathbf{v}_1)\sum_{Y_2^n} P^n_{Y_2|V_1,V_2}(\mathbf{y}_2|\mathbf{v}_1,\mathbf{v}_2)\log\left(\frac{P^n_{Y_2|V_2,V_1}(\mathbf{y}_2|\mathbf{v}_2,\mathbf{v}_1)}{J_1 Q^n_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)} + 1\right) = e_1 + e_2 \quad (3.34)$$

where,

$$e_1 := \sum_{\substack{\mathbf{v}_1\\(\mathbf{v}_1,\mathbf{v}_2)\in\mathcal{T}_{\epsilon_0}}} P^n_{V_1}(\mathbf{v}_1) \sum_{\substack{Y_2^n\\(\mathbf{y}_2,\mathbf{v}_2,\mathbf{v}_1)\notin\mathcal{T}_\epsilon}} P^n_{Y_2|V_1,V_2}(\mathbf{y}_2|\mathbf{v}_1,\mathbf{v}_2)\log\left(\frac{P^n(\mathbf{y}_2|\mathbf{v}_2,\mathbf{v}_1)}{J_1 Q^n(\mathbf{y}_2|\mathbf{v}_2)} + 1\right) \quad (3.35)$$

$$e_2 := \sum_{\substack{\mathbf{v}_1\\(\mathbf{v}_1,\mathbf{v}_2)\in\mathcal{T}_{\epsilon_0}}} P^n_{V_1}(\mathbf{v}_1) \sum_{\substack{\mathbf{y}_2\\(\mathbf{y}_2,\mathbf{v}_2,\mathbf{v}_1)\in\mathcal{T}_\epsilon}} P^n_{Y_2|V_1,V_2}(\mathbf{y}_2|\mathbf{v}_1,\mathbf{v}_2)\log\left(\frac{P^n(\mathbf{y}_2|\mathbf{v}_2,\mathbf{v}_1)}{J_1 Q^n(\mathbf{y}_2|\mathbf{v}_2)} + 1\right) \quad (3.36)$$

We upperbound e_1 in (3.35) as

$$e_1 \leq \sum_{\substack{\mathbf{v}_1 \\ (\mathbf{v}_1,\mathbf{v}_2)\in\mathcal{T}_{\epsilon_0}}} P^n_{V_1}(\mathbf{v}_1) \sum_{\substack{\mathbf{y}_2 \\ (\mathbf{y}_2,\mathbf{v}_2,\mathbf{v}_1)\notin\mathcal{T}_{\epsilon}}} P^n_{Y_2|V_1,V_2}(\mathbf{y}_2|\mathbf{v}_1,\mathbf{v}_2) \log\left((\frac{1}{\pi_{Y_2|V_2}})^n + 1\right) \tag{3.37}$$

$$\leq 2|\mathcal{Y}_2||\mathcal{V}_1||\mathcal{V}_2| 2^{-2n\epsilon^2\pi^2_{(Y_2|V_2,V_2)}}(-n)\log(\pi_{Y_2|V_1}). \tag{3.38}$$

where $\pi_{Y_2|V_2}$ and $\pi_{Y_2|V_2V_1}$ are derived from (3.26). The (3.38) implies that as $n \to \infty$, $e_1 \to 0$.

On the other hand, to upperbound e_2 in (3.36) we need to use the Lemma 8, since, $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{y}_2) \in \mathcal{T}_\epsilon$ defined on P_{Y_2,V_2,V_1}. Therefore we have,

$$e_2 \leq \sum_{\substack{\mathbf{v}_1 \\ (\mathbf{v}_1,\mathbf{v}_2)\in\mathcal{T}_{\epsilon_0}}} P^n_{V_1}(\mathbf{v}_1) \log\left(\frac{2^{-n(H(Y_2|V_1,V_2)-\delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2|V_2)+I(V_1;V_2)+\delta_2(\epsilon))}} + 1\right)$$

$$\leq \sum_{\substack{\mathbf{v}_1 \\ (\mathbf{v}_1,\mathbf{v}_2)\in\mathcal{T}_{\epsilon_0}}} 2^{-n(H(V_1)+I(V_1;V_2)-\delta_3(\epsilon))} \log\left(\frac{2^{-n(H(Y_2|V_1,V_2)-\delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2|V_2)+I(V_1;V_2)+\delta_2(\epsilon))}} + 1\right) \tag{3.39}$$

$$\leq 2^{n(H(V_1)+\delta_4(\epsilon))} 2^{-n(H(V_1)+I(V_1;V_2)-\delta_3(\epsilon))} \log\left(\frac{2^{-n(H(Y_2|V_1,V_2)-\delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2|V_2)+I(V_1;V_2)+\delta_2(\epsilon))}} + 1\right) \tag{3.40}$$

$$\leq 2^{-n(I(V_1;V_2)-\delta_3(\epsilon)-\delta_4(\epsilon))} \log\left(\frac{2^{-n(H(Y_2|V_1,V_2)-\delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2|V_2)+I(V_1;V_2)+\delta_2(\epsilon))}} + 1\right) \tag{3.41}$$

$$\leq 2^{-n(I(V_1;V_2)-\delta_3(\epsilon)-\delta_4(\epsilon))} \log\left(2^{-n(R'_1+H(Y_2|V_1,V_2)-H(Y_2|V_2)-I(V_1;V_2)-\delta_1(\epsilon)-\delta_2(\epsilon))} + 1\right) \tag{3.42}$$

$$\leq \log(e) 2^{-n(R'_1-I(Y_2;V_1|V_2)-\delta_5(\epsilon))}, \tag{3.43}$$

where, $\delta_5(\epsilon) := \delta_1(\epsilon) + \delta_2(\epsilon) + \delta_3(\epsilon) + \delta_4(\epsilon)$. We have to impose $R'_1 > I(Y_2; V_1|V_2) + \delta_5(\epsilon)$ in order to assure $e_2 \to 0$ with $n \to \infty$. This condition, thus, evokes strong secrecy.

Combining $R'_1 > I(Y_2; V_1|V_2)$, and symmetrically $R'_2 > I(Y_1; V_2|V_1)$, (3.19) and (3.20), we obtain the rate region as in the theorem 2.

□

Remark 6. The typicality set $\mathcal{T}_\epsilon$ above is considered with respect to the true distribution $P_{V_1V_2Y_2}$ (as opposed to the stealth distribution $Q^n_{Y_2|V_2}$), therefore we need to use Lemma 8 to upperbound the probability value of Q. The difference in the two distribution comes from the fact the our encoding structure, which dictates the distribution $P_{V_1V_2Y_2}$, imposes $P(\mathbf{v}_1, \mathbf{v}_2) \neq P(\mathbf{v}_1)P(\mathbf{v}_2)$. The stealth distribution $Q^n_{Y_2|V_2}$, however, needs $P(\mathbf{v}_1, \mathbf{v}_2) = P(\mathbf{v}_1)P(\mathbf{v}_2)$ by definition.

3.4. Strong secrecy for interference channel

The main result of this section is summarized in the following theorem.

Theorem 3. *A rate pair (R_1, R_2) is achievable for the interference channel under the strong secrecy criterion if*

$$\begin{aligned}&0 \leq R_1,\ 0 \leq R_2\\&R_1 < I(V_1;Y_1|U) - I(V_1;Y_2|V_2,U)\\&R_2 < I(V_2;Y_2|U) - I(V_2;Y_1|V_1,U)\end{aligned}$$

with some probability distribution

$$P_U(u)P_{V_1|U}(v_1|u)P_{V_2|U}(v_2|u)P_{X_1|V_1}(x_1|v_1)P_{X_2|V_2}(x_2|v_2)P(y_1,y_2|x_1,x_2),$$

where the channel $P(y_1,y_2|x_1,x_2)$ is given.

Proof. We present the achievability proof in the following subsections. The proof is inspired by techniques used in [HK13, LMSY08, Mar79, GP80, HK14].

3.4.1. Code generation, encoding and decoding

Codebook

For a fixed distribution $P(u)$, we randomly generate a sequence $\mathbf{u}$ according to $P(\mathbf{u}) = \prod_{i=1}^n P(u_i)$, which is assumed to be known by all transmitters and receivers. Each transmitter t, given the random sequence $\mathbf{u}$, forges $2^{R_t+R'_t}$ independent sequences from probability distributions $P(\mathbf{v}_t|\mathbf{u}) = \prod_{i=1}^n P(v_{t,i}|u_i)$, where $P_{v_t|u}$ is fixed. We divide each set of sequences $\mathbf{v}_t$ into 2^{R_t} bins each with $2^{R'_t}$ sequences. We refer to each sequence with $\mathbf{v}_t(m_t, s_t)$, where $m_t \in \{1, \ldots, (M_t = 2^{nR_t})\}$ and $s_t \in \{1, \ldots, (J_t = 2^{nR'_t})\}$.

Encoding

To send a message $m_t \in \mathcal{M}_t$, we select randomly and independently from the bin $\mathbf{v}_t(m_t, i)$ $i \in \{1, \ldots, J_t\}$ an index $i = s_t$, thus, $\mathbf{v}_t(m_t, s_t)$. Then for each codeword $\mathbf{v}_t(m_t, s_t)$ the stochastic encoder sends a channel input based on $P(\mathbf{x}_t|\mathbf{v}_t) = \prod_{i=1}^n P(x_{t,i}|v_{t,i})$.

Decoding

The decoder in receiver t observes $\mathbf{y}_t$. If there is a unique m_t and some s_t such that

$$(\mathbf{u}, \mathbf{y}_t, \mathbf{v}_t(m_t, s_t)) \in \mathcal{T}_\epsilon^n(P_{Y_t,V_t,U}), \tag{3.44}$$

we have decoded the message m_t, otherwise, declare an error.

Error analysis

We send w. l. g. $\mathbf{v}_1(1,1)$. There are two error events, which are specified as:

$$E_{\mathrm{IF1},t} := \{(\mathbf{u}, \mathbf{y}_t, \mathbf{v}_t(1,1)) \notin \mathcal{T}_\epsilon^n, \ |\mathbf{v}_1(1,1) \text{ sent}\} \tag{3.45}$$

$$E_{\mathrm{IF2},t} := \bigcup_{\substack{m_t \neq 1 \\ s_t}} \{(\mathbf{u}, \mathbf{y}_t, \mathbf{v}_t(m_t, s_t)) \in \mathcal{T}_\epsilon^n, \ |\mathbf{v}_1(1,1) \text{ sent}\}. \tag{3.46}$$

The union bound on total error events results in,

$$P_{(e^n,\, t)} \leq \Pr(E_{\mathrm{IF1},t}) + \sum_{m_t \neq 1, s_t} \Pr(E_{\mathrm{IF2},t}). \tag{3.47}$$

The first term $\Pr(E_{\mathrm{IF1},t}) < \epsilon$ for large n, from typicality [GK12]. From Packing Lemma the second term can be small up to a given $\epsilon > 0$ for n sufficiently large if

$$R_1 + R_1' < I(V_1, Y_1|U) \text{ and} \tag{3.48}$$

$$R_2 + R_2' < I(V_2, Y_2|U). \tag{3.49}$$

3.4.2. Strong secrecy criterion

In order to ensure the strong secrecy, we are required to show that for $\epsilon > 0$ and a sufficiently large n,

$$I(W_1; Y_2) < \epsilon \ \text{ and } \ I(W_2; Y_1) < \epsilon,$$

hold under some conditions on the rates.

Remark 7. For the secrecy analysis we drop the random variable U for the sake of notational simplicity. It can be introduced, if desired via standard arguments at the end of the proof [CK78, HK14].

We extend the framework in [HK13, HK14] to find a condition that bounds

$$\mathrm{D}(P_{W_1,Y_2} || P_{W_1} Q_{Y_2}^n) \leq \xi \tag{3.50}$$

where the left hand side is known as the *effective secrecy*, $\xi > 0$ is arbitrarily small, P_{W_1,Y_2} is the joint distribution of W_1 and Y_2, which is expressed as

$$P(\mathbf{y}_2|m_1) = \sum_{s_1=1}^{J_1} \frac{1}{J_1} P^n(\mathbf{y}_2|\mathbf{v}_1(m_1, s_1)),$$

and further

$$P(\mathbf{y}_2) = \sum_{m_1=1}^{M_1} \sum_{s_1=1}^{J_1} \frac{1}{M_1.J_1} P^n(\mathbf{y}_2|\mathbf{v}_1(m_1, s_1)).$$

Further, $Q^n_{Y_2}$ is distribution of Y_2 while no meaningful message is transmitted for W_1,

$$Q^n_{Y_2}(\mathbf{y}_2) = \sum_{\mathbf{v}_1} P^n_{V_1}(\mathbf{v}_1) P^n_{Y_2|V_1}(\mathbf{y}_2|\mathbf{v}_1) \tag{3.51}$$

and finally, P_{W_1} is marginal distribution of W_1. We can further expand (3.50) as

$$\begin{aligned} \mathrm{D}(P_{W_1,Y_2}||P_{W_1}Q^n_{Y_2}) &= \sum_{m_1,\mathbf{y}_2} P_{W_1,Y_2}(m_1;\mathbf{y}_2) \log\left(\frac{P_{W_1,Y_2}(m_1;\mathbf{y}_2)}{Q^n_{Y_2}(\mathbf{y}_2)P_{W_1}(m_1)}\right) \\ &= I(W_1;Y_2) + \mathrm{D}(P_{Y_2}||Q^n_{Y_2}). \end{aligned} \tag{3.52}$$

Remark 8. The expansion in (3.52), reveals that (3.50) addresses not only the strong secrecy notation by bounding $I(W_1;Y_2)$, but also bounds $\mathrm{D}(P_{Y_2}||Q^n_{Y_2})$ which corresponds to *stealth* of the communication.

Remark 9. The difference between P_{Y_2} and $Q^n_{Y_2}$ is in the absence of (a meaningful) message W_1.

Based on chain rule for informational divergence we have,

$$\mathrm{D}(P_{W_1,Y_2}||P_{W_1}Q^n_{Y_2}) = \mathrm{D}(P_{W_1}||P_{W_1}) + \mathrm{D}(P_{Y_2,W_1}||Q^n_{Y_2}|P_{W_1}), \tag{3.53}$$

where, according to the definition of $\mathrm{D}(\cdot\|\cdot)$ in (2.1), the first term above equals to zero, thus we proceed with $\mathrm{D}(P_{Y_2,W_1}||Q^n_{Y_2}|P_{W_1})$.

Lemma 9. *Taking an expectation over Y_2^n, V_1, and S_1 yields:*

$$\mathrm{E}[\mathrm{D}(P_{Y_2|W_1}||Q^n_{Y_2|\mathbf{v}_2}|P_{W_1})] \le \mathrm{E}\bigg[\log\left(\frac{P^n_{Y_2|V_1}(Y_2^n|V_1)}{J_1 Q^n_{Y_2}(Y_2^n)} + 1\right)\bigg], \tag{3.54}$$

The sketch of the proof is provided in the Appendix 3.7.4.

The equation (3.54) can be split up as follows:

$$\sum_{\mathbf{y}_2,\mathbf{v}_1} Q^n_{Y_2,V_1}(\mathbf{y}_2,\mathbf{v}_1)\log\left(\frac{P^n_{Y_2|V_1}(\mathbf{y}_2|\mathbf{v}_1)}{J_1 Q^n_{Y_2}(\mathbf{y}_2)}+1\right)=e_1+e_2$$

where,

$$e_1 := \sum_{(\mathbf{y}_2,\mathbf{v}_1)\notin\mathcal{T}_\epsilon} Q^n_{Y_2,V_1}(\mathbf{y}_2,\mathbf{v}_1)\log\left(\frac{P^n(\mathbf{y}_2|\mathbf{v}_1)}{J_1 Q^n(\mathbf{y}_2)}+1\right) \tag{3.55}$$

$$e_2 := \sum_{(\mathbf{y}_2,\mathbf{v}_1)\in\mathcal{T}_\epsilon} Q^n_{Y_2,V_1}(\mathbf{y}_2,\mathbf{v}_1)\log\left(\frac{P^n(\mathbf{y}_2|\mathbf{v}_1)}{J_1 Q^n(\mathbf{y}_2)}+1\right) \tag{3.56}$$

We upperbound e_1 in (3.55) as

$$\begin{aligned} e_1 &\leq \sum_{(\mathbf{y}_2,\mathbf{v}_1)\notin\mathcal{T}_\epsilon} Q^n_{Y_2,V_1}(\mathbf{y}_2,\mathbf{v}_1)\log\left((\frac{1}{\pi_{Y_2}})^n+1\right) \\ &\leq 2|\mathcal{Y}_2||\mathcal{V}_1|2^{-2n\epsilon^2\pi^2_{(Y_2)}}(-n)\log(\pi_{Y_2|V_1}). \end{aligned} \tag{3.57}$$

where π_{Y_2} and $\pi_{Y_2|V_1}$ are derived from (3.26). The (3.57) implies that as $n\to\infty$, $e_1\to 0$.

On the other hand, to upperbound e_2 in (3.56) we need to use the Lemma 8, since, $(\mathbf{v}_1,\mathbf{y}_2)\in\mathcal{T}_\epsilon$ defined on P_{Y_2,V_1}. Therefore we have,

$$\begin{aligned} e_2 &\leq \sum_{(\mathbf{y}_2,\mathbf{v}_1)\in\mathcal{T}_\epsilon} Q^n_{Y_2,V_1}(\mathbf{y}_2,\mathbf{v}_1)\log\left(\frac{2^{-n(H(Y_2|V_1)-\delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2)+\delta_2(\epsilon))}}+1\right) \\ &\leq \log\left(\frac{2^{-n(H(Y_2|V_1)-\delta_1(\epsilon))}}{J_1 2^{-n(H(Y_2)+\delta_2(\epsilon))}}+1\right) \\ &\leq \log\left(2^{-n(R_1'+H(Y_2|V_1)-H(Y_2)-\delta_1(\epsilon)-\delta_2(\epsilon))}+1\right) \\ &\leq \log(e)2^{-n(R_1'-I(Y_2;V_1)-\delta_3(\epsilon))}, \end{aligned}$$

where, $\delta_3(\epsilon):=\delta_1(\epsilon)+\delta_2(\epsilon)$. We have to impose $R_1'>I(Y_2;V_1)+\delta_3(\epsilon)$ in order to assure $e_2\to 0$ with $n\to\infty$. This condition, thus, evokes strong secrecy.

Combining $R_1'>I(Y_2;V_1)$, and symmetrically $R_2'>I(Y_1;V_2)$, (3.48) and (3.49), we obtain the rate region as in the theorem 3.

□

3.5. Applications to untrusted two-hop relay channel

We consider the transmission of messages in DMURC as two phases; the first phase is a DMBCC, which is studied in Section 3.3 and the second phase is a DMMAC, which is given in the following. The achievable rate region of DMMAC is given as union of all the rate pairs $(R_{1_{\text{MAC}}}, R_{2_{\text{MAC}}})$ such that [GK12],

$$\begin{aligned} R_{1_{\text{MAC}}} &\le I(\hat{Y}_1; Z|\hat{Y}_2, Q), \qquad (3.58)\\ R_{2_{\text{MAC}}} &\le I(\hat{Y}_2; Z|\hat{Y}_1, Q),\\ R_{1_{\text{MAC}}} + R_{2_{\text{MAC}}} &\le I(\hat{Y}_1, \hat{Y}_2; Z|Q), \end{aligned}$$

where, $P(\hat{y}_1, \hat{y}_2, q) := P(q)P(\hat{y}_1|q)P(\hat{y}_2|q)$, and $q \in \mathcal{Q}, \ \ |\mathcal{Q}| \le 2$.

We present an achievable rate region for a two-hop decode-and-forward untrusted relay channel in the following.

Theorem 4. *The R_{DMURC} is achievable for a DMURC with confidential messages if:*

$$\begin{aligned} 0 \le R_t \le \min\Big\{ & I(V_t; Y_t|U) - I(V_t; V_k|U) - I(V_t; Y_k|V_k, U),\\ & I(\hat{Y}_t; Z|\hat{Y}_k, Q), \ \frac{1}{2} I(\hat{Y}_t, \hat{Y}_k; Z|Q)\Big\},\\ & t, k \in \{1, 2\}, \ \ k \ne t, \end{aligned}$$

for some probability distributions $P_U(u)P_{V_1,V_2|U}(v_1, v_2|u)P_{X|V_1,V_2}(x|v_1, v_2)$ with $P(y_1, y_2|x)$ given as the DMBCC channel, and $P_Q(q)P_{\hat{Y}_1|Q}(\hat{y}_1|q)P_{\hat{Y}_2|Q}(\hat{y}_2|q)$ with $P(z|\hat{y}_1, \hat{y}_2)$ given as the DMMAC channel. where, U, V_t, and Q are auxiliary random variables.

Proof. The achievable rate region is

$$R_{\text{DMURC}} \subseteq R_{\text{BCC}} \cap R_{\text{MAC}}, \qquad (3.59)$$

where R_{BCC} is the achievable rate region for DMBCC in Section 3.3, and R_{MAC} is the underlying multiple access channel between the relays and the D.

In designing the encoder ϕ we take $\mathcal{U}_1 = \mathcal{U}_2 = \{1, \ldots, \min\{M_1, M_2\}\}$. The choice of the mapping part is non-trivial since it has to fulfil the conditions (3.8), (3.9) and (3.10). We propose the following

$$\theta_1(m_1, m_2) = \hat{m}_1 = m_1 \oplus m_2 \ \text{ and } \ \theta_2(m_1, m_2) = \hat{m}_2 = m_1 \ominus m_2. \qquad (3.60)$$

where, $\oplus$ and $\ominus$ are modulo plus and modulo minus operations defined in ring of size $q_F := \min\{M_1, M_2\}$. Therefore we have $\hat{m}_1, \hat{m}_2 \in \hat{\mathcal{M}}$, where $\hat{\mathcal{M}} := \{1, \ldots, q_F\}$.

Codebook Generation

We use the codebook generation similar to the one proposed in Section 3.3, with the message sets $\hat{\mathcal{M}}_1$ and $\hat{\mathcal{M}}_2$ defined according to the mapping in (3.60). [1]

Encoding at the Sender

We apply similar encoding as in the Section 3.3 with small modifications. Note that since $|\hat{\mathcal{M}}_1| = |\hat{\mathcal{M}}_2| = q_F$ the messages sets are imposed to be the same size.

Decoding at the Relays

We apply the identical decoding in the Section 3.3.

Encoding at the Relays

We use the encoding from [GK12].

Decoding at the Destination

We apply MAC simultaneous decoding in [GK12].

3.5.1. Decodability at D

In order to obtain (m_1, m_2), after successfully decoding $(\hat{m}_1, \hat{m}_2)$ at the destination, we define a mapping function as

$$\rho(\hat{m}_1, \hat{m}_2) = (m_1, m_2) \tag{3.61}$$

such that (m_1, m_2) is a unique solution of (3.60). The calculation procedure includes the following steps:

$$\rho_1 : 2m_1 = (\hat{m}_1 \oplus \hat{m}_2), \quad \rho_2 : 2m_2 = (\hat{m}_1 \ominus \hat{m}_2). \tag{3.62}$$

The following lemma ensures that ρ_1 and ρ_2 in (3.62) are capable of recovering the transmitted messages uniquely.

Lemma 10. *The mapping 3.61 is well defined, if* $\gcd(2, q_F) = 1$*, i.e.* q_F *is odd.*

The proof is given in Lemma A.1.9 in [vT93].

---

[1] Instead of $\mathcal{M}_1$ and $\mathcal{M}_1$, which are used in Section 3.3

Remark 10. For a general $\theta_1,\ \theta_2$ and ρ_t, where the divisors can be any number in the ring, the safe choice is q_F to be a prime number. However, for large q_F we do not know the prime number's density. [2] In (3.60) this problem do not appear since q_F can be any odd number that $q_F \geq 2$.

3.5.2. Security at the relays

On term of the security is that relay 1 is incapable of decoding relay 2's message, which is guaranteed with strong secrecy measures based on Section 3.3.

As far as the second secrecy criterion, in (3.8), is concerned, the following lemmas shed light on the properties of the mapping in (3.60) in more details.

Lemma 11. *Following the secrecy criterion, and the condition (3.8),* $M_1 = M_2 = q_F$ *must hold.*

This lemma is proved in the Appendix 3.7.5.

Lemma 12. *For the random variables* $W_t \in \mathcal{M}_t$ *and* $\hat{W}_t \in \hat{\mathcal{M}}_t$*, which correspond to the messages* m_t *and the precoded messages* $\hat{m}_t$ *respectively, the following holds in (3.60)*

$$I(W_j; \hat{W}_i) = 0 \tag{3.63}$$

$$H(\hat{W}_i) = H(W_j), \quad i, j \in \{1, 2\} \tag{3.64}$$

Proof. We prove that the $\hat{W}_i$ is independent of W_j. We first consider $\hat{m}_1 \triangleq m_1 \oplus m_2$.

$$\begin{aligned} p_{\hat{W}_1, W_1}(\hat{m}_1, m_1) &= \Pr(\hat{W}_1 = \hat{m}_1, W_1 = m_1) \\ &= \Pr(W_2 = \hat{m}_1 \ominus m_1, W_1 = m_1) \\ &= \Pr(W_2 = \hat{m}_1 \ominus m_1)\Pr(W_1 = m_1) \end{aligned} \tag{3.65}$$

Now we have show that $p_{\hat{W}_1}(\hat{m}_1) = p_{W_2}(\hat{m}_1 \ominus m_1)$, for every choice of m_1. This follows directly from the fact that W_2 is uniformly distributed over the set, $\Pr(W_2 = \hat{m}_1 \ominus m_1) =$

---

[2] Imagine for our message alphabet size of $M_t = 2^{nR}$ we have to choose a prime q_F which is $q_F > M_t$

$\frac{1}{q_F}$, which yields,

$$\begin{aligned}\Pr(\hat{W}_1 = \hat{m}_1) &= \Pr(W_1 + W_2 = \hat{m}_1)\\ &= \sum_{l=1}^{|W_2|} \Pr(W_2 = \hat{m}_1 \ominus n_l | W_1 = n_l)\Pr(W_1 = n_l)\\ &= \sum_{l=1}^{q_F} \Pr(W_2 = \hat{m}_1 \ominus n_l)\Pr(W_1 = n_l)\\ &= \Pr(W_2 = \hat{m}_1 \ominus n_l).\end{aligned}$$

where, we use the assumption that $p_{W_1}(n_l)$ and $p_{W_2}(n_l)$ are uniformly distributed over the same sets (Lemma 11). For $\hat{m}_2$ and other combinations of the lemma it is a similar proof according to Lemma 11.

Consequently, from independence we infer that $H(m_j|\hat{m}_i) = H(m_j)$. The second part (3.64) follows immediately from above by considering that $\hat{m}_j$ is also uniformly distributed over the same size of alphabet q_F as a result of modulo operations.

□

Remark 11. The Lemma 11 can be interpreted as follows. To secure a set of messages $\mathcal{M}_t$ with a set of key, the size of the key alphabet must be equal to the size of the sender to have *strong* secrecy. In our case the key set is also a messages set itself. This is a corollary of Shannon's perfect secrecy. The only difference is that the key's here are also messages. Therefore, we do not spend resources on sending random keys.

Lemma 13. *The proposed mapping scheme, θ_1, θ_2 and ρ_t, does not provide perfect secrecy for both (W_1, W_2), since*[3]

$$H((W_1, W_2)|\hat{W}_j) = H(W_1) \leq H((W_1, W_2)) \;\; j \in \{1, 2\} \tag{3.66}$$

Proof. Since the mapping is one to one, and uniform distribution of $\hat{W}_j$s are guaranteed by Lemma 12, the inequality is straightforward to show. As for the first part we can write:

$$\begin{aligned}\Pr\big((W_1, W_2) = (m_1, m_2)|\hat{W}_j = \hat{m}_j\big) =& \frac{\Pr\big((W_1, W_2) = (m_1, m_2)\big)}{\Pr(\hat{W}_j = \hat{m}_j)} && (3.67)\\ =& \frac{\Pr(W_1 = m_1)\Pr(W_2 = m_2)}{\Pr(\hat{W}_j = \hat{m}_j)} && (3.68)\\ =& \Pr(W_1 = m_1). && (3.69)\end{aligned}$$

---

[3] In order to accomplish the perfect secrecy we have to use random keys as the auxiliary sets $\mathcal{U}_1$ and $\mathcal{U}_2$. Then, however, we require make the random key available at the destination somehow.

Thus, we deduce the lemma. Note that we used $M_1 = M_2$. ☐

Lemma 11 implies that $R_t = \min\{R_1, R_2\}$. Moreover, since the relays do not have large buffers, the rates has to be $\min\{\hat{R}_t, R_t\}$. Combining them with theorem 2 and (3.58) we conclude the proof.

☐

3.6. Conclusions

We have analysed the achievable secrecy rate regions of three scenarios: a discrete memoryless broadcast channel with two confidential messages (DMBCC), a discrete memoryless interference channel with two confidential messages (DMICC), a discrete memoryless untrusted relay channels (DMURC).

We have considered a stronger notion of secrecy, the so called *effective secrecy*, which includes stealth and strong secrecy together. An achievable rate region with strong secrecy for DMBCC and DMICC has been developed. We used the results in [HK13, HK14] to impose an upperbound on $\mathrm{D}(P_{X,Z} \| P_X Q_Z) = I(X; Z) + D(P_Z \| Q_Z)$, in which all of the terms are non-negative by definition.We have used a random binning channel coding suggested in [LMSY08], which imposes a joint distribution on the channel inputs. The term denoting non-stealth, $D(P_Z \| Q_Z)$, however, requires the independent relationship between the channel inputs. This conflict results in some complexity, which has been handled in this chapter. This problem, however, arises only for the DMBCC scenario. We have showed that the achievable secrecy rate region with strong secrecy and stealth criteria are exactly the same as the ones with weak secrecy criterion, which is proposed in [LMSY08].

We have studied the DMURC scenario, where there are two possibilities for information leakage. The first case is when transmitting a message to a specific relay the other relay(s) can eavesdrop. The second case occurs after the target relay received the message. Unprotected messages can be eavesdropped before being forwarded to the destination. To protect against the first case we used the strongly secure channel coding proposed for DMBCC. In order to deal with the second case we have developed a pre-coder and a post-decoder. The pre-coder protects the messages at the sender by combining them with other independent messages such that they are information-theoretically protected and also reliably recovered at the destination. We use the uniform random messages in set as keys to protect the messages in the other set. The proposed scheme is, however, not strongly secure overall.

Future work

The following research directions based on the results in this chapter are proposed.

- We propose to apply the same approaches that are developed for DMBCC and DMICC to achieve the strong secrecy region of some other discrete memoryless channels.
- Extend the results to Gaussian channels instead of discrete memoryless channels.
- We suggest to investigate if the method of Output Statistic of Random Binning (OSRB) leads to similar results.
- Relays with capacity limited channel between them is worth investigation.

3.7. Appendix

3.7.1. Derivation of Lemma 5

$$\begin{aligned}
\mathrm{D}(P_{W_1,Y_2|V_2}||P_{W_1}Q_{Y_2|V_2}) &= \sum_{m_1,\mathbf{y}_2} P_{W_1,Y_2|V_2}(m_1;\mathbf{y}_2|\mathbf{v}_2)\log\left(\frac{P_{W_1,Y_2|V_2}(m_1;\mathbf{y}_2|\mathbf{v}_2)}{Q_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)P_{W_1}(m_1)}\right)\\
&= \sum_{m_1,\mathbf{y}_2} P_{W_1,Y_2|V_2}(m_1;\mathbf{y}_2|\mathbf{v}_2)\log\left(\frac{P_{W_1,Y_2|V_2}(m_1;\mathbf{y}_2|\mathbf{v}_2)}{Q_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)P_{W_1}(m_1)}\cdot\frac{P_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}{P_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\right)\\
&= \sum_{m_1,\mathbf{y}_2} P_{W_1,Y_2|V_2}(m_1;\mathbf{y}_2|\mathbf{v}_2)\log\left(\frac{P_{W_1,Y_2|V_2}(m_1;\mathbf{y}_2|\mathbf{v}_2)}{P_{W_1}(m_1)P_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\cdot\frac{P_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}{Q_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\right)\\
&\overset{(\alpha)}{=} \sum_{m_1,\mathbf{y}_2} P_{W_1,Y_2|V_2}(m_1;\mathbf{y}_2|\mathbf{v}_2)\log\left(\frac{P_{W_1,Y_2|V_2}(m_1;\mathbf{y}_2|\mathbf{v}_2)}{P_{W_1}(m_1)P_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\right)\\
&\quad +\sum_{\mathbf{y}_2} P_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)\log\left(\frac{P_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}{Q_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\right)\\
&= I(W_1;Y_2|\mathbf{v}_2)+\mathrm{D}(P_{Y_2|V_2}||Q_{Y_2|V_2}), \qquad (3.70)
\end{aligned}$$

in (α) we substitute the marginal distribution of $P_{W_1,Y_2|V_2}(m_1;\mathbf{y}_2|\mathbf{v}_2)$.

3.7.2. Proof of Lemma 7

We take the expectation over $V_1^n(m_1,s_1)$, $m_1=1,\dots,M_1$, $s_1=1,\dots,J_1$, for a given realization $\mathbf{v}_2:=\mathbf{v}_2(m_2,s_2)$

$$
\begin{aligned}
&\mathrm{E}[\mathbb{1}(\phi_\tau \neq (1,1))\mathrm{D}(P_{Y_2^n|W_1\mathbf{v}_2}||Q^n_{Y_2|\mathbf{v}_2}|P_{W_1})] = \\
&\qquad \mathrm{E}\Bigg[\mathbb{1}(\phi_\tau \neq (1,1)) \sum_{\mathbf{y}_2} \sum_{m_1,s_1}^{M_1,J_1} \frac{1}{M_1.J_1} P(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2) \times \\
&\qquad\qquad \log \frac{\sum_{s'=1}^{J_1} P(\mathbf{y}_2|\mathbf{v}_1(m_1,s'),\mathbf{v}_2)}{J_1 Q^n_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\Bigg] \\
&\qquad = \sum_{\mathbf{v}_1(1,1)\in\mathcal{T}_\epsilon(P_{V_1^n|\mathbf{v}_2})} \cdots \sum_{\mathbf{v}_1(M_1,J_1)\in\mathcal{T}_\epsilon(P_{V_1^n|\mathbf{v}_2})} \prod_{e,i}^{M_1,J_1} P_{V_1}(\mathbf{v}_1(e,i)) \sum_{\mathbf{y}_2} \\
&\qquad\qquad \sum_{\substack{m_1=1\\ s_1=1}}^{M_1,J_1} \frac{1}{M_1.J_1} P(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2) \times \\
&\qquad\qquad \left[\log \frac{\sum_{s'=1}^{J_1} P(\mathbf{y}_2|\mathbf{v}_1(m_1,s'),\mathbf{v}_2)}{J_1 Q^n_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\right]
\end{aligned}
$$

we can rearrange the above and write the residual terms as,

$$
\begin{aligned}
&\mathrm{E}[\mathbb{1}(\phi_\tau \neq (1,1))\mathrm{D}(P_{Y_2^n|W_1\mathbf{v}_2}||Q^n_{Y_2|\mathbf{v}_2}|P_{W_1})] = \\
&\qquad \sum_{\substack{m_1=1\\ s_1=1}}^{M_1,J_1} \sum_{\mathbf{y}_2} \sum_{\mathbf{v}_1(m_1,s_1)\in\mathcal{T}_\epsilon(P_{V_1^n|\mathbf{v}_2})} \frac{1}{M_1.J_1} P_{V_1}(\mathbf{v}_1(m_1,s_1)) P(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2) \\
&\qquad\qquad \times \sum_{\mathbf{v}_1(1,1),\ldots,\mathbf{v}_1(M_1,J_1)\in\mathcal{T}_\epsilon(V_1^n|\mathbf{v}_2)}^{b_{(m_1,s_1)}} \prod_{\substack{l\neq m_1\\ k\neq s_1}}^{M_1,J_1} P_{V_1}(\mathbf{v}_1(l,k)) \\
&\qquad\qquad \times \left[\log \frac{\sum_{s'=1}^{J_1} P(\mathbf{y}_2|\mathbf{v}_1(m_1,s'),\mathbf{v}_2)}{J_1 Q^n_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\right].
\end{aligned}
$$

where, $\sum_{\mathbf{v}_1(1,1),\ldots,\mathbf{v}_1(M_1,J_1)\in\mathcal{T}_\epsilon(P_{V_1^n|\mathbf{v}_2})}^{b_{(m_1,s_1)}}$ denote the series of sums of

$$
\sum_{\mathbf{v}_1(1,1)\in\mathcal{T}_\epsilon(P_{V_1^n|\mathbf{v}_2})} \cdots \sum_{\mathbf{v}_1(M_1,J_1)\in\mathcal{T}_\epsilon(V_1^n|\mathbf{v}_2)}
$$

without $\sum_{\mathbf{v}_1(m_1,s_1)\in\mathcal{T}_\epsilon(V_1^n|\mathbf{v}_2)}$. In the next step we use the Jensen's inequality and concavity of log to obtain,

$$\begin{aligned}
&\mathrm{E}[\mathbb{1}(\phi_\mathcal{T}\neq(1,1))\mathrm{D}(P_{Y_2^n|W_1\mathbf{v}_2}||Q^n_{Y_2|V_2}|P_{W_1})]\leq \qquad (3.71)\\
&\quad\sum_{\substack{m_1=1\\s_1=1}}^{M_1,J_1}\sum_{\mathbf{y}_2}\sum_{\mathbf{v}_1(m_1,s_1)\in\mathcal{T}_\epsilon(P_{V_1^n|\mathbf{v}_2})}\frac{1}{M_1.J_1}P_{V_1}(\mathbf{v}_1(m_1,s_1))P(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2)\\
&\quad\times\Bigg[\log\Bigg(\frac{P(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2)}{J_1Q^n_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\\
&\qquad+\sum_{s'\neq s_1}^{J_1}\sum_{\mathbf{v}_1(e,s')\in\mathcal{T}_\epsilon(V_1^n|\mathbf{v}_2)}\frac{P_{V_1}(\mathbf{v}_1(s',e))P(\mathbf{y}_2|\mathbf{v}_1(s',e),\mathbf{v}_2)}{J_1Q^n_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}\Bigg)\Bigg]\\
&\quad=\sum_{\substack{m_1=1\\s_1=1}}^{M_1,J_1}\sum_{\mathbf{y}_2}\sum_{\mathbf{v}_1(m_1,s_1)\in\mathcal{T}_\epsilon(P_{V_1^n|\mathbf{v}_2})}\frac{1}{M_1.J_1}P_{V_1}(\mathbf{v}_1(m_1,s_1))P(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2)\\
&\quad\times\left[\log\left(\frac{P(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2)}{J_1Q^n_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}+\frac{J_1-1}{J_1}\right)\right] \qquad (3.72)\\
&\quad\leq\sum_{\substack{m_1=1\\s_1=1}}^{M_1,J_1}\sum_{\mathbf{y}_2}\sum_{\mathbf{v}_1(m_1,s_1)\in\mathcal{T}_\epsilon(P_{V_1^n|\mathbf{v}_2})}\frac{1}{M_1.J_1}P_{V_1}(\mathbf{v}_1(m_1,s_1))P(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2)\\
&\quad\times\left[\log\left(\frac{P(\mathbf{y}_2|\mathbf{v}_1(m_1,s_1),\mathbf{v}_2)}{J_1Q^n_{Y_2|V_2}(\mathbf{y}_2|\mathbf{v}_2)}+1\right)\right]\\
&\quad=\mathrm{E}\left[\mathbb{1}(\mathcal{T}_\epsilon(V_1^n|\mathbf{v}_2))\log\left(\frac{P(Y_2^n|V_1^n,\mathbf{v}_2)}{J_1Q^n_{Y_2|V_2}(Y_2^n|\mathbf{v}_2)}+1\right)\right],
\end{aligned}$$

where in (3.72) we used the definition (3.23). To summarize we have so far derived the following inequality for every pair (m_2,s_2)

$$\mathrm{E}[\mathbb{1}(\phi_\mathcal{T}\neq(1,1))\mathrm{D}(P_{Y_2^n|W_1,\mathbf{v}_2}||Q^n_{Y_2|\mathbf{v}_2}|P_{W_1})]\leq\mathrm{E}\left[\mathbb{1}(\mathcal{T}_\epsilon(V_1^n|\mathbf{v}_2))\log\left(\frac{P(Y_2^n|V_1^n,\mathbf{v}_2)}{J_1Q^n_{Y_2|V_2}(Y_2^n|\mathbf{v}_2)}+1\right)\right] \qquad (3.73)$$

3.7.3. Proof of Lemma 8

According to Lemma 2.6 in [CK82] we have for all $(\mathbf{v}_2,\mathbf{y}_2)\in\mathcal{V}_2^n\times\mathcal{Y}_2^n$,

$$Q^n(\mathbf{y}_2|\mathbf{v}_2)=2^{-n(H(P^e_{V_2,Y_2})-H(P^e_{V_2})+\mathrm{D}(P^e_{V_2,Y_2}\|P^e_{V_2}Q))} \qquad (3.74)$$

where,

1. $P^e_{V_2,Y_2}$ denotes the empirical distribution or type of the sequence pair $(\mathbf{v}_2, \mathbf{y}_2)$
2. $P^e_{V_2}$ is the empirical distribution generated by the sequence $\mathbf{v}_2$
3. the distribution $P^e_{V_2}Q$ is the joint distribution computed with respect to $P^e_{V_2}$ and Q

Uniform continuity of the entropy (Lemma 2.7 in [CK82]), $(\mathbf{v}_2, \mathbf{y}_2) \in \mathcal{T}^n_\epsilon(P_{V_2Y_2})$, and the fact that $\epsilon \in (0, \frac{1}{2})$, yield

$$|H(P^e_{V_2,Y_2}) - H(P^e_{V_2}) - H(Y_2|V_2)| \leq \delta_1(\epsilon), \tag{3.75}$$

where, $\delta_1(\epsilon) > 0$ with $\lim_{\epsilon\to 0} \delta_1(\epsilon) = 0$. Moreover,

$$|\mathrm{D}(P^e_{V_2,Y_2} \| P^e_{V_2}Q) - \mathrm{D}(P_{V_2,Y_2} \| P_{V_2}Q)| \leq \delta_2(\epsilon), \tag{3.76}$$

with $\delta_2(\epsilon) > 0$ and $\lim_{\epsilon\to 0} \delta_2(\epsilon) = 0$.
Inequality (3.76) can be seen as follows: In a first step we show that $P^e_{V_2,Y_2} \ll P^e_{V_2}Q$. To this end we assume

$$P^e_{V_2}Q(v_2, y_2) = P^e_{V_2}(v_2)Q(y_2|v_2) = 0.$$

Then either $P^e_{V_2}(v_2) = 0$ and then $P^e_{V_2Y_2}(v_2, y_2) = 0$ automatically. Or $Q(y_2|v_2) = 0$ and then

$$\begin{aligned} 0 =& Q(y_2|v_2) \\ =& \sum_{v_1} P_{V_1}(v_1) P_{Y_2|V_1V_2}(y_2|v_1, v_2) \\ \geq& \sum_{v_1} P_{V_1V_2}(v_1, v_2) P_{Y_2|V_1V_2}(y_2|v_1, v_2) \\ =& \sum_{v_1} P_{V_1V_2Y_2}(v_1, v_2, y_2) \\ =& P_{V_2Y_2}(v_2, y_2), \end{aligned} \tag{3.77}$$

which, again, implies $P^e_{V_2Y_2}(v_2,y_2)=0$ by definition of typical sequences, and thus $P^e_{V_2Y_2}\ll P^e_{V_2}Q$. Since $P^e_{V_2Y_2}\ll P^e_{V_2}Q$ we have

$$\begin{aligned}\mathrm{D}(P^e_{V_2Y_2}\|P^e_{V_2}Q)& \\ &=\sum_{v_2,y_2}P^e_{V_2Y_2}(v_2,y_2)\log P^e_{V_2Y_2}(v_2,y_2)\\ &\quad-\sum_{v_2,y_2}P^e_{V_2Y_2}(v_2,y_2)\log P^e_{V_2}(v_2)\\ &\quad-\sum_{v_2,y_2}P^e_{V_2Y_2}(v_2,y_2)\log Q(y_2|v_2)\end{aligned}\tag{3.78}$$

and applying again the uniform continuity of entropy along with the properties of typical sequences we obtain (3.76).

To finish the proof we need to show $\mathrm{D}(P_{V_2Y_2}\|P_{V_2}Q)\leq I(V_1,V_2)$ which can be derived as follows,

$$\begin{aligned}&\mathrm{D}(P_{V_2Y_2}\|P_{V_2}Q)\\ &=\sum_{a,b}P_{V_2Y_2}(a,b)\log\frac{P_{V_2Y_2}(a,b)}{P_{V_2}(a)\sum_c P_{V_1}(c)P_{Y_2|V_1V_2}(b|c,a)}\\ &=\sum_{c,a,b}P_{V_1V_2Y_2}(c,a,b)\log\frac{\sum_c P_{V_1V_2Y_2}(c,a,b)}{\sum_c P_{V_2}(a)P_{V_1}(c)P_{Y_2|V_1V_2}(b|c,a)}\\ &\leq\sum_{c,a,b}P_{V_1V_2Y_2}(c,a,b)\log\frac{P_{V_1V_2Y_2}(c,a,b)}{P_{V_2}(a)P_{V_1}(c)P_{Y_2|V_1V_2}(b|c,a)} &&(3.79)\\ &=\sum_{c,a,b}P_{V_1V_2Y_2}(c,a,b)\log\frac{P_{V_1V_2}(c,a)P_{Y_2|V_1V_2}(b|c,a)}{P_{V_2}(a)P_{V_1}(c)P_{Y_2|V_1V_2}(b|c,a)}\\ &=\sum_{c,a,b}P_{V_1V_2Y_2}(c,a,b)\log\frac{P_{V_1V_2}(c,a)}{P_{V_2}(a)P_{V_1}(c)}\\ &=\sum_{c,a}P_{V_1V_2}(c,a)\log\frac{P_{V_1V_2}(c,a)}{P_{V_2}(a)P_{V_1}(c)}\\ &=I(V_1,V_2), &&(3.80)\end{aligned}$$

where (3.79) is by the Log-Sum-Inequality (Lemma 3.1 in [CK82]).

Combining (3.75), (3.76), (3.80), and (3.74) leads to the claim of the lemma.

3.7.4. Proof of Lemma 9

$$
\begin{aligned}
\mathrm{E}[\mathrm{D}(P_{Y_2|W_1}||Q_{Y_2|V_2}|P_{W_1})] &= \mathrm{E}\Big[\log \frac{\sum_{j=1}^{J_1} P(Y_2^n|V_1(W_1,j))}{J_1 Q_{Y_2|V_2}(Y_2^n)}\Big] \\
&= \sum_{m_1,s_1} \frac{1}{J_1 M_1} \mathrm{E}\Big[\log \frac{\sum_{j=1}^{J_1} P(Y_2^n|V_1(m_1,j))}{J_1 Q_{Y_2}(Y_2^n)}\Big| W_1 = m_1, S_1 = s_1\Big] \\
&\leq \sum_{m_1,s_1} \frac{1}{J_1 M_1} \mathrm{E}\Big[\log \frac{P(Y_2^n|V_1(m_1,s_1))}{J_1 Q_{Y_2}(Y_2^n)} + \frac{J_1-1}{J_1}\Big| W_1 = m_1, S_1 = s_1\Big] \qquad (3.81) \\
&\leq \sum_{m_1,s_1} \frac{1}{J_1 M_1} \mathrm{E}\Big[\log \frac{P(Y_2^n|V_1(m_1,s_1))}{J_1 Q_{Y_2}(Y_2^n)} + 1\Big| W_1 = m_1, S_1 = s_1\Big] \\
&\leq \mathrm{E}\Big[\log\Big(\frac{P_{Y_2|V_1}(Y_2^n|V_1)}{J_1 Q^n_{Y_2}(Y_2^n)} + 1\Big)\Big], \qquad (3.82)
\end{aligned}
$$

where, in (3.81) we used Jensen's inequality applied to the part of the expectation over $V_1(m_1, i)$ for $i \neq s_1$. The RV $S_1 \in \{1, \ldots, J_1\}$ is distributed uniformly.

3.7.5. Proof of Lemma 11

We prove that for the conditions (3.8) and (3.10), the cardinality of set of two messages groups must be equal. From Fano's inequality we have,

$$
\begin{aligned}
n\epsilon_n &\geq H(W_1|W_2, Y_1) \\
&= H(W_1; W_2, Y_1) - H(W_2, Y_1) \\
&= H(W_1, Y_1) + H(W_2|W_1, Y_1) - H(W_2, Y_1) \\
&\geq H(W_1, Y_1) - H(W_2, Y_1) \\
&\geq H(W_1) + H(Y_1) - I(W_1, Y_1) - H(W_2, Y_1) \\
&= H(W_1) - H(W_2|Y_1) - I(W_1, Y_1)
\end{aligned}
$$

We substitute $H(W_2|Y_1) = H(W_2)$, as well as, $I(W_1, Y_1) \leq \epsilon_1$ which further yields,

$$H(W_1) \leq H(W_2) + n\epsilon_n + \epsilon_1.$$

We have $H(W_1) = nR_1$ and $H(W_2) = nR_2$, thus for arbitrary small ϵ we have $R_1 \leq R_2 + \epsilon_n$. Based symmetry we have $R_2 \leq R_1 + \epsilon_n$ at the same time.

3.7.6. Proof of Lemma 6

By definition the KL divergence can be written as:

$$\mathrm{D}(P||Q) = \sum_{x} P(x)\log(\frac{P(x)}{Q(x)}) \leq \sum_{x} P(x)\log(\frac{1}{\pi_Q}) \leq \log(\frac{1}{\pi_Q}), \tag{3.83}$$

where $\pi_Q = \min\{Q(a) : Q(a) > 0\}$.

4. Secrecy for diamond network with untrusted relays

In the previous chapter, we have studied information-theoretic secrecy for different scenarios, including the discrete memoryless two-hop untrusted relay channel (DMURC). We devote this chapter chiefly to the Gaussian untrusted relay channel, however, from a cross-layer security perspective. Cross-layer design for security takes into account restrictions, conditions, and techniques of different communication layers and use them to provide security. This chapter's intention is not achieving perfect secrecy (in the Shannon sense) but a weaker secrecy with a larger achievable rate region. We present a framework for designing secrecy schemes for both decode-and-forward (DF) and amplify-and-forward (AF) two-hop untrusted relay channels that ensure some pre-defined level of secrecy.

We introduce the notion of cross-layer security, where only parts of transmitted information are protected at the physical layer (in an information-theoretical sense) against potential eavesdropping at relay nodes. On top of the information-theoretical (or physical-layer) security, we assume that there is an additional upper-layer scheme to provide the desired level of secrecy, but these additional security measures are not in the scope of this chapter (see also Section 4.3).

4.1. Background and contributions

Shamir in [Sha79] introduced the idea of secret sharing, which was well received by researchers in cryptography and computer science. The idea was to distribute a secret among different parties such that each party would not be able to reconstruct the whole secret, without having other parties information. Similar ideas have been developed for secure network coding by Cai et al. in [CY02], in which a combination of messages (e.g. linear combinations) are transmitted as the coded messages to increase both reliability and secrecy throughout the network.

4.1.1. Related work

The work of [LMB07] introduces a security scheme for network coding, which is referred to as algebraic security. In [LMB07] the encoded messages are a linear combination of original messages so that they protect each other (Shannon key length). This approach is more elaborated in an information-theoretic sense in Section 4.3. More fundamental works in secure network coding are presented by Cai et al. in [CY02, CY11]. In [CY11], the authors have studied network coding ideas of security utilizing information-theoretic security techniques. Besides the network coding community, untrusted relays draw some attention in information theory community, for example, [JKK12, HMS12] and [HY10a]. A more comprehensive overview of secure untrusted relay schemes is provided in Chapter 3.

The idea behind the cross-layer security is that only a certain proportion of the entire data stream is information-theoretically secure with respect to each entity (in our case relay). Building upon that, an additional upper-layer scheme is expected to provide the required notion of secrecy. This can be achieved via schemes like network coding [LMB07] or secrecy sharing [LK10], and in effect no parts of the entire data stream are available to any of the relays in plain text. This way, the proposed scheme can achieve different points on the trade-off between perfect secrecy and the performance expressed in terms of the transmission rate.

Finding the secrecy capacity region in closed form is challenging and still an open problem, particularity for our setup. Therefore, some authors study the asymptotic behaviour of achievable rates as certain system parameters tend to infinity [Khi04, JL06, ANDH08]. Khisti [Khi04] shows the scaling behaviour of multicast channel while the number of users increases. Developing the results of [Khi04], the authors in [JL06] considered other transmission techniques for multicast case such as beamforming, spatially white, and orthogonal transmission.

4.1.2. Contributions

We study both DF and AF relaying strategies. In the first case, we prove simple bounds describing how much secrecy can be provided at the physical layer, and propose a framework for designing protocols with the required properties. As for the AF strategy, we propose a scheme providing perfect information-theoretic secrecy and analyse the problem of finding the optimal transmit beamformer. Based on these results, we present two schemes with partial secrecy. Building upon the works of Khisti and Jindal et al. in [Khi04, JL06] we investigate the asymptotic behaviour of the achievable secrecy rate of the broadcast channel with different sets of legitimate relays and untrusted relays. This provides insights on the cross-layer design problem. We list detailed contributions of this chapter in the following.

- We introduce a framework for designing cross-layer security schemes. We further introduce a notion of secrecy as referred to as α-secrecy.
- We provide solutions that achieve α-security for the AF and DF untrusted relay Gaussian channels. The design problem reduces to a convex problem using a Charnes-Cooper transform, whereas, in a more general case it is an untraceable mixed-integer optimization problem.
- To provide further insights into the mixed-integer problem, we study its asymptotic behaviour in terms of increasing number of antennas and increasing number of relays. We further consider the secrecy outage probability for some simple scenarios.

4.2. System model

We consider the downlink of a relay network consisting of $K+2$ nodes: a base station (sender) node S, a destination node D and $K > 1$ relay nodes, depicted in Figure 4.1. The set of all relay nodes is denoted by $\mathcal{K} := \{1, 2, \ldots, K\}$. The sender node S is equipped with N_T transmit antennas, whereas, each relay and the destination D have a single antenna element. The channels from S to the relays are denoted by $\mathbf{h}_k \in \mathbb{C}^{N_T},\ k \in \mathcal{K}$ and $e_k \in \mathbb{C},\ k \in \mathcal{K}$ are used to represent the channels from the relay nodes D. Furthermore, it is assumed that the direct link between the sender and the destination as well as the links among the relay nodes cannot be used or do not exist. All channels are assumed to be randomly chosen constants that are perfectly known to the sender. The entries of vectors $\mathbf{h}_k$ and channels e_k are i.i.d. random variables drawn from the circular-symmetric Gaussian distribution $\mathcal{CN}(0,1)$. We further assume that the adversary potentially has access to the signals received and forwarded at some relays, but cannot change relays' operation, e.g. report false channel state information or receive transmissions from other relays.

Taking the half-duplex constraint into account, it is assumed that the transmission from the sender S to the relays (Phase I) and the transmission from the relays to destination D (Phase II) are always separated (using, e.g. TDMA). In Phase I the signal received by relay k is given by:

$$y_k = \mathbf{h}_k^H \mathbf{v} d + n_k, \quad k \in \mathcal{K}, \tag{4.1}$$

where $d \in \mathbb{C}$ denotes the scalar transmit signal, $\mathbf{v} \in \mathbb{C}^{N_T}$ is the transmit beamforming vector with $||\mathbf{v}||_2 \leq 1$ and $n_k \sim \mathcal{CN}(0, \sigma_n^2), k \in \mathcal{K}$ is the additive Gaussian noise, which is assumed to have equal variance at all relays. In Phase II, with the signal transmitted by relay k denoted by x_k, the destination receives:

$$y_D = e_k x + n_D, \tag{4.2}$$

where, $n_D \sim \mathcal{CN}(0, \sigma_n^2)$ is the Gaussian noise at the destination. We further define $\mathbf{e} := [e_1, \dots, e_K]$ and $\mathbf{x} := [x_1, \dots, x_K]$. The noise variables $n_k,\ k \in \mathcal{K}$ and are assumed to be i.i.d. distributed with variance $\sigma_n^2 = 1$, without loss of generality. For simplicity, we assume unit transmit powers at all nodes. Individual power constraints at the relays are assumed[1].

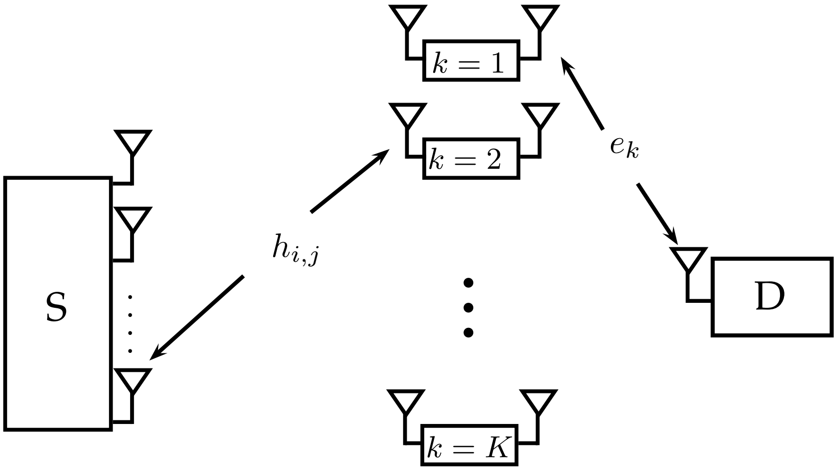

Figure 4.1.: Two-hop wireless Gaussian relay network.

4.3. Cross-layer security

In Section 3.5 we have shown that applying a precoding scheme to the independent sets of messages can provide (weak) secrecy against untrusted relays. Building upon this background, in this chapter, we propose a cross layer scheme that yields security. We assume L independent streams of data denoted by $\mathbf{d}'^L := (d'_1, \dots, d'_L) \in \mathcal{F}'^L$, where $\mathcal{F}'$ is a finite ring, to be transmitted via the K relays. We assume there exists a similar scheme to that of Section 3.5 that protects the data against the relays. In other words, we assume there exist L functions $f_l : \mathcal{F}'^L \to \mathcal{F},\ \forall l \in \{1, \dots, L\}$, where $\mathcal{F}$ is a finite ring. We use $d_l := f_l(d'_l)$ as the transmit streams. We assume that an inverse mapping at the destination exists and the f_ls fulfil security conditions[2].

---

[1] If 'fairness' in terms of total transmit power is required, sum power constraints could be assumed.

[2] For an example of such a mapping we refer to Section 3.5.

The following conditions hold while transmitting each stream:

- The stream d_l is transmitted with a rate R_l where, $R^{SD} := \sum_{l=1}^{L} R_l$ is the total transmission rate from S to D.
- We transmit the stream d_l via the relays in set $\mathcal{S}_l$ while keeping it secret from $\hat{\mathcal{S}}_l$, with $\mathcal{S}_l \cap \hat{\mathcal{S}}_l = \emptyset$ and $\mathcal{S}_l \cup \hat{\mathcal{S}}_l = \mathcal{K}$.[3]

The set of all streams that regard the relay k as an eavesdropper is represented by

$$\mathcal{M}(k) := \big\{ l : k \text{ is an eavesdropper for stream } d_l \big\}. \tag{4.3}$$

From (4.3), the total rate of the streams that are kept secret from the relay k is given by, $R^{\mathrm{u}}(k) = \sum_{l \in \mathcal{M}(k)} R_l$.

Definition 23 (α-secrecy)**.** Consider a transmission scheme between S and D, comprising L physical-layer streams $d_1, d_2, \ldots, d_L$. If

$$R^{\mathrm{u}}(k) \geq \alpha R^{SD} \quad \forall k \in \mathcal{K}, \tag{4.4}$$

for a fixed $\alpha \in [0, 1]$, the transmission is said to be α-secret for the largest α.

Remark 12. According to Definition 23, Shannon's perfect secrecy corresponds to 1-secrecy, in which non of the relays are able to decode any of the streams d_l, $\forall l \in \{1, \ldots, L\}$. On the other hand, 0-secrecy implies that at least one relay receives all the streams $\mathbf{d}'^{L}$s, thus, all of the streams are exposed to a potential eavesdropper.

4.4. Decode-and-forward

Achieving 1-secrecy is not possible with DF relays since by definition DF relays have to be able to decode a message before forwarding it. However, the α in α-secrecy can be made arbitrarily close to 1 by using an appropriate number of relay nodes and transmit antennas at the base station, as stated by the following proposition.

Proposition 1. *Assume that $N_T \geq K$. Then, α-secrecy with $\alpha = \alpha_{max} = 1 - \frac{1}{K}$ is achievable with a Decode-and-Forward scheme, while no DF scheme can achieve $\alpha > \alpha_{max}$.*

Proof. Achievability is proven by using Zero-Forcing (ZF) beamforming in Phase I of the transmission (from the BS to the relays), with $L = K$ different data streams and TDMA

---

[3] The range of the cardinalities of the sets $\mathcal{S}_l$ and $\hat{\mathcal{S}}_l$ in general determined by f, however, we do not specify any rule at this point.

multiplexing among the streams. In each stream d_l, all relays $k \neq l$ are treated as eavesdroppers, and the mutual information between the base station and these eavesdroppers is 0 by the application of ZF. Thus, using appropriate time slot durations (among the streams and for both phases of each stream), rates $R_1 = R_2 = \ldots = R_K = \frac{1}{K}R^{SD}$ with $R^{\mathrm{u}}(k) = 1 - \frac{1}{K}, k \in \mathcal{K}$ can be obtained.

On the other hand, $\alpha > \alpha_{\max} = 1 - \frac{1}{K}$ would imply that each relay can decode less than $\frac{1}{K}R^{SD}$, thus, there is at least one data stream R_i which cannot be decoded by any relay. This results in a contradiction which completes the proof. □

When $N_T < K$, in general we cannot achieve α-secrecy with $\alpha = 1 - \frac{1}{K}$, however, in the following Proposition we formulated some simple bounds.

Proposition 2. *Assume that $N_T < K$. Using DF-relaying, the largest achievable secrecy factor α is lower bounded by $\frac{N_T-1}{K}$ and upper bounded by $1 - \frac{1}{K}$.*

Sketch of the proof. According to the definition, we conclude that to achieve the best α, we must have $R_l = R \;\; \forall l$. This infers the following:

$$\begin{aligned} R^{\mathrm{u}}(k) &\geq \alpha R^{SD} \\ \sum_{l \in \mathcal{M}(k)} R_l &\geq \alpha \sum_{l=1}^{K} R_l \\ |\mathcal{M}(k)| &\geq \alpha K \\ \frac{|\mathcal{M}(k)|}{K} &\geq \alpha, \end{aligned}$$

To maximize the value for α we need to maximize $|\mathcal{M}(k)|$, i. e. the number of the streams that consider each relay eavesdropper. Using ZF results in nullifying only $N_T - 1$ relays for each stream, from which the secrecy factor $\alpha = \frac{N_T-1}{K}$ follows. As for the upper bound, the result from the proof of Proposition 1 holds. □

The general optimization problem for DF transmission with α-secrecy constraints could be formulated as in the following. Given network topology, channels, and the desired optimization goal (maximize R^{SD} given α or maximize α given rate requirements), find optimal values of the number of data streams L as well as the corresponding transmit beamforming vectors, subsets of forwarding nodes and power allocations in each stream. Such a general problem seems analytically intractable; we, therefore, propose a possible design for creating DF-based α-secret protocols and consider two specific ones of them in more detail.

Consider L data streams, multiplexed using time-sharing, such that the time assigned to data stream l is equal to τ_l and $\sum_{l=1}^{L} \tau_l = 1$. Furthermore, each stream is again subdivided into Phase I (transmission from the sender S to the relays) and Phase II (transmission from the relays to the destination D), with time durations $\tau_l^{(1)}$ and $\tau_l^{(2)}$, respectively, such that $\tau_l^{(1)} + \tau_l^{(2)} = \tau_l$.

Now, consider a given stream d_l. We define the set $\mathcal{S}_l \subset \mathcal{K}$ as the intended receivers of the data stream which are supposed to decode it, and the set $\hat{\mathcal{S}}_l = \mathcal{K} \setminus \mathcal{S}_l$ as the set of the relay nodes which are treated as eavesdroppers. Let the transmit beamformer in Phase I be denoted as $\mathbf{v}_l$. Thus, finding the optimal transmit beamformer corresponds to the multicast problem with multiple eavesdroppers [MKSS12, LM11]:

$$R_l^{(1)} = \tau_l^{(1)} \log \left(\max_{\mathbf{v}_l} \min_{\substack{i \in \mathcal{S}_l \\ j \in \hat{\mathcal{S}}_l}} \frac{1 + \mathbf{h}_i^H \mathbf{v}_l \mathbf{v}_l^H \mathbf{h}_i}{1 + \mathbf{g}_j^H \mathbf{v}_l \mathbf{v}_l^H \mathbf{g}_j} \right)^+ \tag{4.5}$$

where, we define $\mathbf{g}_j, \ \ j \in \{1, \ldots, |\hat{\mathcal{S}}_l|\}$ as the channels of the relays that belong to $\hat{\mathcal{S}}_l$ and $(a)^+ := \max(a, 0)$. Note that the problem can be solved efficiently via a convex relaxation and semidefinite programming. In Phase II since all relays $k \in \mathcal{S}_l$ are supposed to have decoded the transmitted data, the signal transmitted by relay k is $x_k = r_{kl} d_l, k \in \mathcal{S}_l$. The achievable rate is therefore equal to:

$$R_l^{(2)} = \tau_l^{(2)} \max_{\mathbf{r}_l} \log(1 + \mathbf{e}_l^H \mathbf{r}_l \mathbf{r}_l^H \mathbf{e}_l), \tag{4.6}$$

with $\mathbf{r}_l := (r_{kl})_{k \in \mathcal{S}_l}$ and $\mathbf{e}_l := (e_k)_{k \in \mathcal{S}_l}$, and $||\mathbf{r}_l||_\infty \leq 1$ under individual power constraints. Note that the above optimizations in (4.5) and (4.6) can be done independently for each data stream d_l, provided that sets $\mathcal{S}_l$ and $\hat{\mathcal{S}}_l$ are fixed. By appropriately choosing $\tau_l^{(i)}, l \in \{1, 2, \ldots, L\}, i \in \{1, 2\}$, we can obtain $R_l^{(1)} = R_l^{(2)} = R_l$ and $R_1 = R_2 = \ldots = R_L$. Furthermore, we require that sets $\mathcal{S}_l$ and $\hat{\mathcal{S}}_l$ are chosen in such a way that $|\mathcal{M}(1)| = |\mathcal{M}(2)| = \ldots = |\mathcal{M}(K)|$. Basing on the proposed design, we consider two protocols.

4.4.1. Partially-secure decode-and-forward I (PSDF1)

We assume $L = 2$ streams, and the set of all relays is assumed to be partitioned into two subsets of similar size. In general, finding a good partition is a complex problem, for simplicity we assume the following rule:

$$\begin{aligned} &\mathcal{S}_1 = \{1, \ldots, \lfloor K/2 \rfloor\}, \ \hat{\mathcal{S}}_1 = \{\lfloor K/2 \rfloor + 1, \ldots, K\} \\ &\mathcal{S}_2 = \hat{\mathcal{S}}_1, \ \hat{\mathcal{S}}_2 = \mathcal{S}_1 \end{aligned}$$

From the above, it follows that $\mathcal{M}(k) = \{l\}, k \in \hat{\mathcal{S}}_l, l = 1, 2..$

Observation 1. *Protocol PSDF1 is 0.5-secret.*

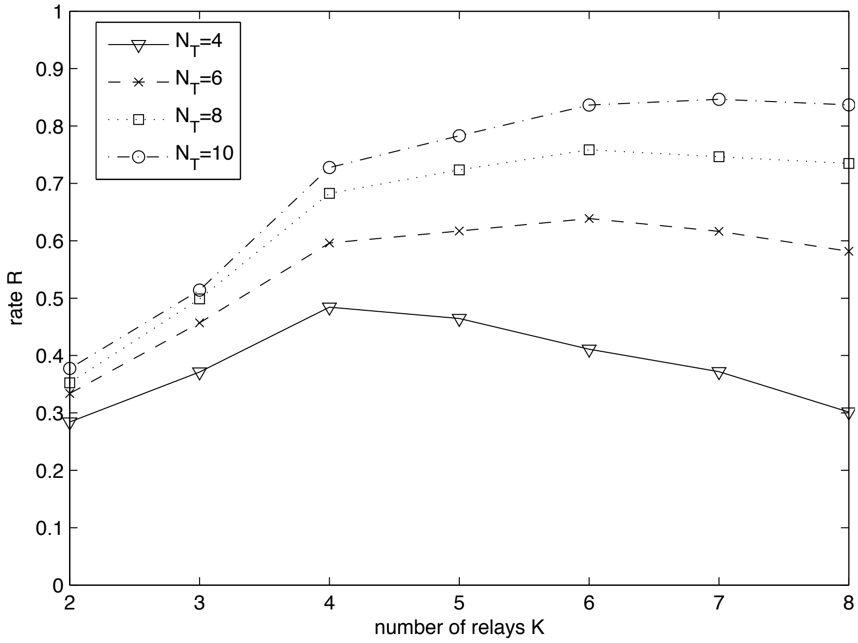

Figure 4.2.: Transmission rate as a function of the number of relays K for Scheme PSDF1 for achieving 0.5-secrecy.

In Figure 4.2, the rate achievable by protocol PSDF1 with 0.5-secrecy is presented as the function of the number or relays in the network. We can observe that, for a given number of transmit antennas N_T at the sender the transmission rate initially increases and assumes its maximum value for the number of relays K approximately equal to the number of antennas N_T, and subsequently it falls. This interesting behaviour is caused by the fact that the number of resolvable beams is essentially equal to the number of transmit antennas; if we have more relays than antennas then it is difficult to direct the energy towards the intended receivers and away from the eavesdroppers at the same time.

4.4.2. Partially-secure decode-and-forward II (PSDF2)

In scheme PSDF2, we have $L = K$ so that the number of streams is equal to the number of relays. An additional parameter $n_{\text{ev}} < K$ specifies the number of eavesdroppers per stream. The sets $\mathcal{S}_l$ and $\hat{\mathcal{S}}_l$ are constructed using the following "circular shift" rule such that $|\mathcal{S}_l| = K - n_{\text{ev}}$ and $|\hat{\mathcal{S}}_l| = n_{\text{ev}}$:

$$\mathcal{S}_l = \{1, \ldots, l - n_{\text{ev}} - 1, l, \ldots, \min(K + l - n_{\text{ev}} - 1, K)\}$$
$$\hat{\mathcal{S}}_l = \{\max(1, l - n_{\text{ev}}), \ldots, l - 1, K + l - n_{\text{ev}}, \ldots, K\}$$

With this construction, we have $|\mathcal{M}(k)| = n_{\text{ev}}, k \in \mathcal{K}$.

Observation 2. *Protocol PSDF2 is* $\frac{n_{ev}}{K}$*-secret.*

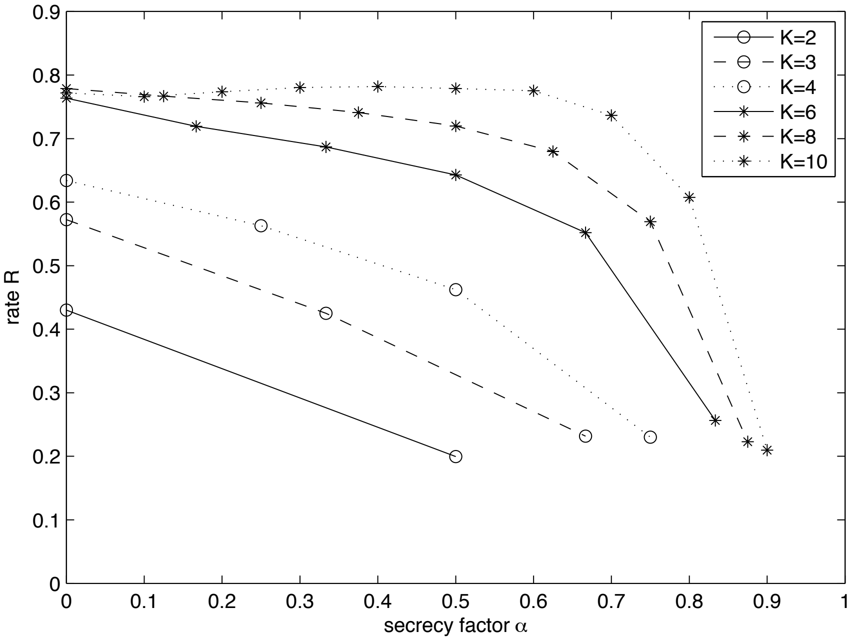

Figure 4.3.: Transmission rate as a function of the secrecy factor α for Scheme PSDF2.

In Figure 4.3 we illustrate the performance of PSDF2, with the number of antennas N_T equal to the number of relays K. It can be seen that although, as expected, the transmission rate decreases when the secrecy requirement rises, α-secrecy with α values up to approximately 0.7-0.8 can be achieved with virtually no performance loss, compared to the of transmission insecure schemes – especially if the number of transmit antennas is sufficiently high. Combining such transmission with an appropriate upper-layer security protocol, this shows that secrecy can be achieved almost “for free”.

4.5. Amplify-and-forward

We introduce three schemes for an AF setup to obtain different α-secrecy levels. The first scheme provides 1-secrecy and includes the optimum transmit beamforming problem.

Further, It serves as the basis for two other approaches to design an α-secret ($\alpha < 1$) system.

4.5.1. Secure amplify-and-forward (SAF)

Assume a single data stream d, such that the signal received at the relays is expressed in the vector notation as:

$$\mathbf{y} = \mathbf{H}\mathbf{v}d + \mathbf{n}, \tag{4.7}$$

with $\mathbf{y} := [y_1, y_2, \dots, y_K]$. $\mathbf{H} := [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K]^H$ and $\mathbf{n} := [n_1, n_2, \dots, n_K]$ as the noise vector.

Each relay is subsequently supposed to forward the received signal by multiplying it with factor f_k. Now, the signal received by the destination D is given as

$$y_D = \mathbf{e}^H\mathbf{F}\mathbf{y} + n_D = \mathbf{e}^H\mathbf{F}\mathbf{H}\mathbf{v}d + \mathbf{e}^H\mathbf{F}\mathbf{n} + n_D, \tag{4.8}$$

where we define the diagonal forwarding matrix $\mathbf{F} := \mathrm{Diag}[f_1, f_2, \dots, f_K]$. Note that no data communication between relays is required as each relay k forwards only its received signal y_k. The forwarding factors f_k should be in general determined according to the channel states $\mathbf{h}_k, k \in \mathcal{K}$ and $e_k, k \in \mathcal{K}$, which may require exchanging this values between the relays or with some designated entity. Joint data processing by the relays, however, is not considered. According to (4.8), the signal-to-noise ratio (SNR) at the destination is equal to

$$\mathrm{SNR}_D = \frac{|\mathbf{e}^H\mathbf{F}\mathbf{H}\mathbf{v}|^2}{1 + \|\mathbf{e}^H\mathbf{F}\|_2^2}, \tag{4.9}$$

in which, without loss of generality, unit noise power and unit transmit power is considered. On the other hand, if a relay k attempts to decode the data stream individually, the achieved SNR is given by

$$\mathrm{SNR}_k = |\mathbf{h}_k^H\mathbf{v}|^2. \tag{4.10}$$

Using the above, we conclude that the following rate can be achieved with perfect secrecy (i.e. we have 1-secrecy):

$$R_{\mathrm{SAF}} = \max_{\mathbf{v},\mathbf{F}}[\log(1 + \mathrm{SNR}_D) - \log(1 + \max_{k\in\mathcal{K}} \mathrm{SNR}_k)]^+ \tag{4.11}$$

Optimum transmit beamformer and relay forwarding factors

After substituting (4.10) and (4.9) into (4.11) and dropping the log function, the optimal beamformer and relay forwarding matrix are given by the solution of

$$\begin{aligned} \max_{\mathbf{v},\mathbf{F}} \min_{k} \quad & \frac{1+||\mathbf{F}^H\mathbf{e}||_2^2+|\mathbf{e}^H\mathbf{F}\mathbf{H}\mathbf{v}|^2}{(1+||\mathbf{F}^H\mathbf{e}||_2^2)(1+|\mathbf{h}_k^H\mathbf{v}|^2)} \\ \text{s.t.} \quad & ||\mathbf{v}||_2 \leq 1, \\ & |f_k|^2 \leq \frac{1}{|\mathbf{h}_k\mathbf{v}|^2+1}, k \in \mathcal{K}, \end{aligned} \tag{4.12}$$

where the constraints on f_k result from individual power constraints at the relays.

The problem above is difficult to solve jointly for $\mathbf{F}$ and $\mathbf{v}$. Therefore, we propose to solve the problem in an iterative fashion where in each iteration we solve the problem with either $\mathbf{F}$ or $\mathbf{v}$ fixed.

First, assume that the forwarding matrix $\mathbf{F}$ is fixed. With further simplification and variable change we obtain

$$\begin{aligned} \max_{\mathbf{v}} \min_{k} \quad & \frac{1+\zeta\mathbf{e}^H\mathbf{F}\mathbf{H}\mathbf{v}\mathbf{v}^H\mathbf{H}^H\mathbf{F}^H\mathbf{e}}{1+|\mathbf{h}_k^H\mathbf{v}|^2} \\ \text{s.t.} \quad & \mathbf{v}^H\mathbf{v} \leq 1, \end{aligned} \tag{4.13}$$

in which $\zeta = 1/(1+\mathbf{e}^H\mathbf{F}\mathbf{F}^H\mathbf{e})$.

In order to solve it numerically we can reformulate it using Charnes-Cooper transformation [CC62, LM11].

Proposition 3. *Optimization problem (4.13) is equivalent to:*

$$\begin{aligned} \min_{\mathbf{W},\vartheta,\mu} \quad & \vartheta \\ s.t. \quad & \mu+\mathbf{h}_k^H\mathbf{W}\mathbf{h}_k \leq \vartheta, \qquad \forall k \in \mathcal{K} \\ & \mu+\zeta\mathbf{e}^H\mathbf{F}\mathbf{H}\mathbf{W}\mathbf{H}^H\mathbf{F}^H\mathbf{e}=1, \\ & \text{trace}[\mathbf{W}] \leq \mu, \\ & \mu > 0, \mathbf{W} \succeq 0, \end{aligned} \tag{4.14}$$

where $\mathbf{W} := \mu\mathbf{v}\mathbf{v}^H$ *and* $\mu > 0$.

Based on the results in [LM11] the rank-1 constraint for $\mathbf{W}$ is fulfilled at the optimum point. The proof is similar to the one provided by [LM11]. Note that the above problem is convex and can be solved efficiently using semi-definite programming.

We proceed by considering the problem (4.12) with fixed $\mathbf{v}$. The problem (4.12) with $\mathbf{v}$ as a given fixed parameter is not trivial. We propose to use maximum transmit power at each relay and phases to be aligned at the destination, such that

$$f_k = e^{j\phi_k} / \sqrt{|\mathbf{h}_k^H \mathbf{v}|^2 + 1} \tag{4.15}$$

where $\phi_k = -\arg(e_k^* \mathbf{h}_k^H \mathbf{v})$. Numerical experiments have shown that in most cases iterating (4.14) and (4.15) converges after one to three iterations, even if no specific monotonicity has been observed.

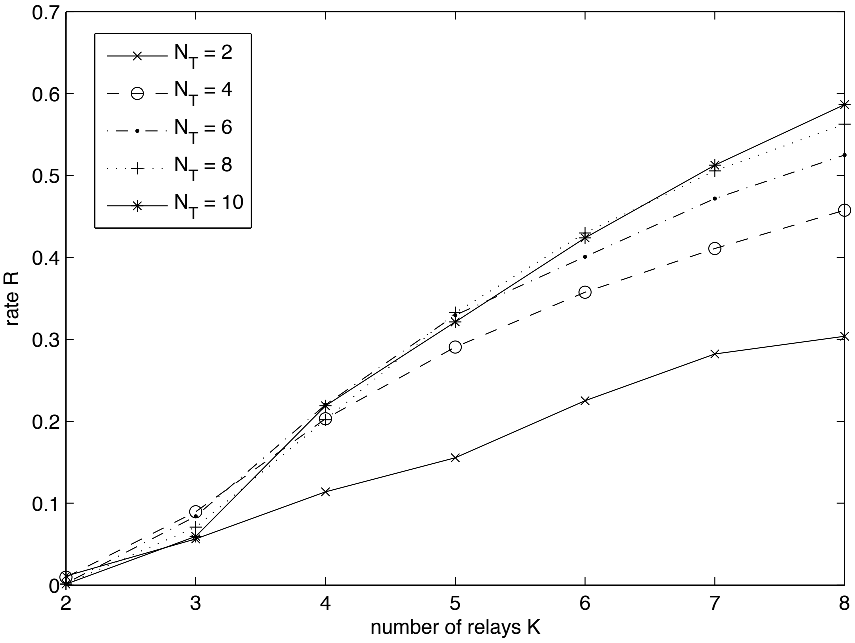

Figure 4.4.: Transmission rate as a function of the number of relays K for scheme SAF.

In Figure 4.4 we show that the rate achieved by Scheme SAF (which guarantees 1-secrecy), increases with the number of relays. In order to achieve the best possible performance for a given number of relays K, the number of transmit antennas N_T should be at least at the order of K.

4.5.2. Partially-secure amplify-and-forward I (PSAF1)

In this approach we transmit two streams from sender S to destination D in a TDMA fashion: stream d_1 at rate R_{SAF} and stream d_2 at rate R_c. R_{SAF} is an achievable information-theoretically secure (1-secret) rate (4.11), and $R_c = \max_{\mathbf{v},\mathbf{F}} \log(1 + \text{SNR}_D)$ is a non-secure achievable transmission rate (0-secret scheme). The overall rate is

$$R_{\text{PSAF1}} = \tau_1 R_c + \tau_2 R_{\text{SAF}} \geq R_{\text{SAF}}$$

in which $\tau_1 + \tau_2 = 1$. Therefore, by appropriately choosing the time sharing factors τ_1 and τ_2, we can obtain a higher transmission rate than with Scheme SAF, at the expense of a decreased secrecy coefficient α. The secrecy parameter α becomes

$$\alpha = \frac{\tau_1 R_c}{\tau_1 R_c + \tau_2 R_{\text{SAF}}}.$$

Note that under the assumption that $\tau_1 R_c = \tau_2 R_{\text{SAF}}$, the specific design for an upper-layer security scheme proposed in Section 4.3 can be used on top of Scheme PSAF1 to obtain perfect secrecy.

4.5.3. Partially-secure amplify-and-forward II (PSAF2)

Another method to achieve α-secret communication using multiple relays consists again in using L data streams, $d_1, d_2, \ldots, d_L$. In each stream l, we define only the subset of users $\hat{\mathcal{S}}_l \subset \mathcal{K}$ as the eavesdroppers of the stream. Apart from that, the transmission as in scheme SAF is performed. By carefully constructing the sets $\hat{\mathcal{S}}_l$, we can obtain the desired levels of security. The general principle adopted in this family of protocols follows the approach considered when designing the DF schemes in Section 4.4.

Since in stream l the nodes $k \in \hat{\mathcal{S}}_l$ are considered to be eavesdroppers, the beamformer optimization problem to compute transmit vector $\mathbf{v}_l$ (assuming fixed $\mathbf{F}$) becomes

$$\begin{aligned} &\max_{\mathbf{v}_l} \min_{k \in \hat{\mathcal{S}}_l} \frac{1 + \mathbf{e}^H \mathbf{F}\mathbf{F}^H \mathbf{e} + |\mathbf{e}^H \mathbf{F}\mathbf{H}\mathbf{v}_l|^2}{(1 + \mathbf{e}^H \mathbf{F}\mathbf{F}^H \mathbf{e})(1 + |\mathbf{h}_k^H \mathbf{v}_l|^2)} \\ &\text{s.t.} \quad \mathbf{v}_l^H \mathbf{v}_l \leq 1. \end{aligned} \tag{4.16}$$

Remark 13. Note that in general the optimization involves selecting optimal subsets $\hat{\mathcal{S}}_l$.The problem is a mixed integer programming, therefore some closed-form solutions are unlikely to exist. However, a solution of the problem for given subsets $\hat{\mathcal{S}}_l$ can be reduced to solving (4.14).

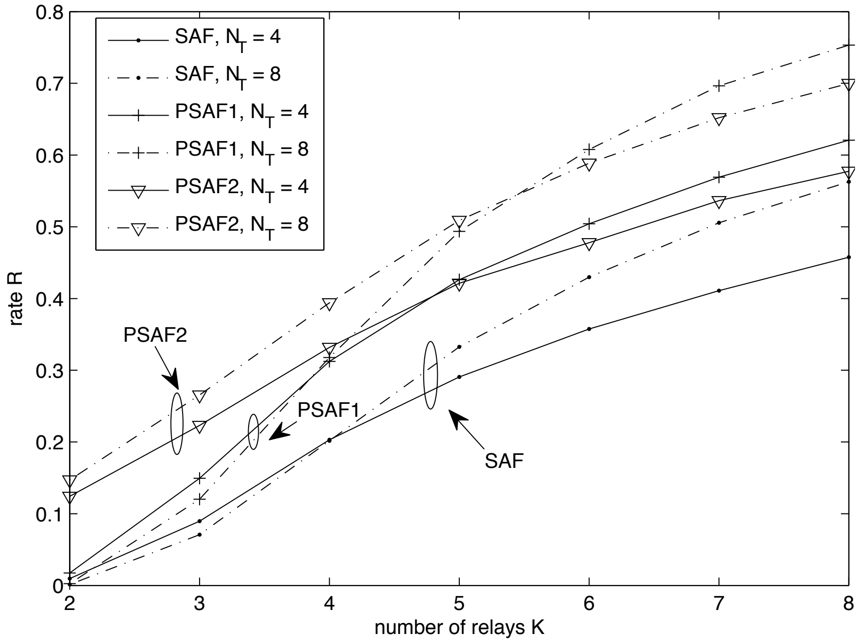

Figure 4.5.: Transmission rate as a function of the number of relays K for schemes PSAF1 and PSAF2 with 0.5-secrecy, compared to SAF with 1-secrecy.

In Figure 4.5, we present the performance results obtained for Schemes PSAF1 and PSAF2, both configured to achieve 0.5-secrecy. As for the scheme PSAF2, we used $L = 2$ streams and a simple subset construction rule from Scheme PSDF1.

4.6. Asymptotic analysis

As discussed in previous sections, obtaining the optimal solution for (4.5) involves mixed integer programming, thus, it is difficult. Instead we study the asymptotic behaviour of (4.5) to gain some insights on the number of relays that are regarded as eavesdropper, i.e. $|\hat{\mathcal{S}}_l|$ and the number of relays that are considered to be legitimate receiver of the message, i.e. $|\mathcal{S}_l|$ and the number of transmit antenna N_T. In more detail, we analyse the following cases:

- Increasing $|\mathcal{S}_l|$ while keeping N_T and $|\hat{\mathcal{S}}_l|$ constant[4].
- Increasing $|\hat{\mathcal{S}}_l|$ for some fixed N_T and $|\mathcal{S}_l|$.
- Increasing number of antennas N_T while keeping $|\mathcal{S}_l|$ and $|\hat{\mathcal{S}}_l|$ fixed.

---

[4] Note that in this section we do not fix the total number of the relays K.

We further present some basic results on outage probability for the following cases: For $N_T = 1$ with multiple legitimate users ($|\mathcal{S}_l| > 1$) and eavesdroppers ($|\hat{\mathcal{S}}_l| > 1$) and for $N_T > 1$ when $|\mathcal{S}_l| = 1$ and $|\hat{\mathcal{S}}_l| = 1$. The achievable secrecy rate is given by [KW10]:

$$R_{\text{sc}} = \log \left(\max_{\substack{\text{trace}[\mathbf{\Sigma}] \leq 1 \\ \mathbf{\Sigma} \succeq 0}} \min_{\substack{\mathbf{h}_i \in \mathcal{S}_l \\ \mathbf{g}_j \in \hat{\mathcal{S}}_l}} \frac{1 + \mathbf{h}_i^H \mathbf{\Sigma} \mathbf{h}_i}{1 + \mathbf{g}_j^H \mathbf{\Sigma} \mathbf{g}_j} \right) \tag{4.17}$$

where $\mathbf{\Sigma} \in \mathbb{C}^{N_T \times N_T}$ is the transmit covariance matrix and by $\mathbf{g}_j$ we refer to the relay channels that belong to the eavesdropper set $\hat{\mathcal{S}}_l$.

We further define $N_E := |\hat{\mathcal{S}}_l|$ and $N_L := |\mathcal{S}_l|$.

Remark 14. According to [SU07], at the optimum point we either transmit with full power (trace[$\mathbf{\Sigma}$] = 1) or do not transmit at all (trace[$\mathbf{\Sigma}$] = 0, $R_{\text{sc}} = 0$). The achievable secrecy rate, therefore, can be equivalently represented as

$$R_{\text{sc}} = \max(0, \tilde{R}_{\text{sc}})$$

with

$$\tilde{R}_{\text{sc}} \triangleq \log \left(\max_{\substack{\text{trace}[\mathbf{\Sigma}] = 1 \\ \mathbf{\Sigma} \succeq 0}} \min_{i,j} \frac{1 + \mathbf{h}_i^H \mathbf{\Sigma} \mathbf{h}_i}{1 + \mathbf{g}_j^H \mathbf{\Sigma} \mathbf{g}_j} \right). \tag{4.18}$$

Our goal is to provide some insights into the asymptotic behaviour of $\tilde{R}_{\text{sc}}$ knowing that when the number of users N_L and/or number of eavesdroppers N_E is large, $\tilde{R}_{\text{sc}}$ is expected to be negative which corresponds to the optimal secrecy rate R_{sc} equal to 0. However, the scaling of $\tilde{R}_{\text{sc}}$ allows us to analyse the order at which R_{sc} decreases to 0. Finally, we discuss the application of the presented results to find the bounds on the probability of obtaining a positive secrecy rate, i.e. $\Pr(R_{\text{sc}} > 0)$.

4.6.1. Bounds on the achievable secrecy rate

In this section we develop some bounds for $\tilde{R}_{\text{sc}}$ that are used for our scaling analysis.

Lower bound

For any $\mathbf{\Sigma} \succeq 0$ with trace[$\mathbf{\Sigma}$] ≤ 1, we have

$$\tilde{R}_{\text{sc}} \geq \min_{i,j} \left\{ \log(1 + \mathbf{h}_i^H \mathbf{\Sigma} \mathbf{h}_i) - \log(1 + \mathbf{g}_j^H \mathbf{\Sigma} \mathbf{g}_j) \right\}. \tag{4.19}$$

In particular, choosing $\mathbf{\Sigma} = \frac{1}{N_T}\mathbf{I}$ yields

$$\tilde{R}_{\text{sc}} \geq \log(1 + \frac{1}{N_T} \min_i ||\mathbf{h}_i||_2^2) - \log(1 + \frac{1}{N_T} \max_j ||\mathbf{g}_j||_2^2). \tag{4.20}$$

This bound basically uses the definition of $\max_{\mathbf{\Sigma}}$ in (4.18).

Upper bounds

An examination of the closed-form solution given by [SU07] for the case of a single user and a single eavesdropper together with the general inequality $\max_x \min_y f(x,y) \leq \min_y \max_x f(x,y)$, shows that

$$\begin{aligned} \tilde{R}_{\text{sc}} &\leq \log(\min_{i,j} \lambda_{\max}[\mathbf{B}_j^{-1/2}\mathbf{A}_i\mathbf{B}_j^{-1/2}]) \\ &\leq \log(\min_{i,j} \lambda_{\max}[\mathbf{B}_j^{-1}]\lambda_{\max}[\mathbf{A}_i]), \end{aligned} \tag{4.21}$$

where $\mathbf{A}_i \triangleq \mathbf{I} + \mathbf{h}_i\mathbf{h}_i^H$ and $\mathbf{B}_j \triangleq \mathbf{I} + \mathbf{g}_j\mathbf{g}_j^H$. We substitute the eigenvalue decomposition (EVD) $\mathbf{h}_i\mathbf{h}_i^H := \mathbf{U}_{h_i}\mathbf{\Lambda}_{h_i}\mathbf{U}_{h_i}^H$, which yields:

$$\begin{aligned} \mathbf{A}_i &= \mathbf{I} + \mathbf{U}_{h_i}\mathbf{\Lambda}_{h_i}\mathbf{U}_{h_i}^H \\ &= \mathbf{U}_{h_i}(\mathbf{I} + \mathbf{\Lambda}_{h_i})\mathbf{U}_{h_i}^H. \end{aligned} \tag{4.22}$$

Since $\mathbf{h}_i\mathbf{h}_i^H$ is a rank one matrix $\mathbf{\Lambda}_{h_i}$ has only one non-zero element, which is $\|\mathbf{h}_i\|_2^2 \geq 0$. Therefore, from (4.22) we realize that the largest eigenvalue of $\mathbf{A}_i$ is $\lambda_{\max}[\mathbf{A}_i] = 1 + \|\mathbf{h}_i\|_2^2$. Using the same procedure for $\mathbf{B}_j^{-1}$, it yields $\lambda_{\max}[\mathbf{B}_j^{-1}] = 1$. Substituting $\lambda_{\max}[\mathbf{A}_i]$ and $\lambda_{\max}[\mathbf{B}_j^{-1}]$ in (4.21), we obtain

$$\tilde{R}_{\text{sc}} \leq \log(\min_i \{1 + \|\mathbf{h}_i\|_2^2\}). \tag{4.23}$$

Another family of bounds can be obtained from

$$\max_{\mathbf{\Sigma}} \{f(\mathbf{\Sigma}) - g(\mathbf{\Sigma})\} \leq \max_{\mathbf{\Sigma}} f(\mathbf{\Sigma}) - \min_{\mathbf{\Sigma}} g(\mathbf{\Sigma}).$$

This yields

$$\tilde{R}_{\text{sc}} \leq R_{\text{mc}} - R_{\text{ev}} \tag{4.24}$$

where we defined $R_{\text{mc}} \triangleq \max_{\mathbf{\Sigma}} \min_i \log(1 + \mathbf{h}_i^H\mathbf{\Sigma}\mathbf{h}_i)$ and $R_{\text{ev}} \triangleq \min_{\mathbf{\Sigma}} \max_j \log(1 + \mathbf{g}_j^H\mathbf{\Sigma}\mathbf{g}_j)$ with $\mathbf{\Sigma} \succeq 0$ and $\text{trace}[\mathbf{\Sigma}] = 1$. R_{mc} can be interpreted as the multicast rate in a system without eavesdroppers. For R_{mc} we clearly have [JL06]

$$R_{\text{mc}} \leq \log(1 + \frac{1}{N_L}\lambda_{max}[\mathbf{H}\mathbf{H}^H]) \tag{4.25}$$

in which $\mathbf{H} = [\mathbf{h}_1\mathbf{h}_2 \dots \mathbf{h}_{N_L}] \in \mathbb{C}^{N_T \times N_L}$.

As far as R_{ev} is concerned we obtain the following lower bound

$$R_{\text{ev}} \geq \log(1 + \frac{1}{N_E}\lambda_{min}[\mathbf{G}\mathbf{G}^H]), \tag{4.26}$$

where $\mathbf{G} = [\mathbf{g}_1\mathbf{g}_2 \dots \mathbf{g}_{N_E}] \in \mathbb{C}^{N_T \times N_E}$. The proof is provided in Appendix 4.8.1.

4.6.2. Scaling of achievable secrecy rate

The objective of this section is to provide insights on scaling behavior of the expected value of $\tilde{R}_{\text{sc}}$ using the bounds introduced in the previous section. Since

$$\mathrm{E}[R_{\text{sc}}] \geq \mathrm{E}[\tilde{R}_{\text{sc}}] \quad \text{and} \quad \mathrm{E}[R_{\text{sc}}] \geq 0,$$

we have $\mathrm{E}[R_{\text{sc}}] \geq \max(\mathrm{E}[\tilde{R}_{\text{sc}}], 0)$. Therefore, $\max(\mathrm{E}[\tilde{R}_{\text{sc}}], 0)$ is a lower bound for $\mathrm{E}[R_{\text{sc}}]$.

In the following, we consider three cases. In each of these, one parameter, e.g. number of users, eavesdroppers, or antennas, is increasing while the others remain constant.

Increasing legitimate users for a constant number of eavesdroppers and antennas

Proposition 4. *Let the number of eavesdroppers and antennas be arbitrary but fixed. Then, as* $N_L \to \infty$, $\mathrm{E}[\tilde{R}_{sc}]$ *can be lower-bounded by* $c_1(\frac{1}{N_L})^{\frac{1}{N_T}} - c_2$, *for some appropriate constants* $c_1, c_2 > 0$.

Remark 15. According the Proposition 4 $\mathrm{E}[\tilde{R}_{\text{sc}}]$ scales at least as $\Omega(1)$.

The standard approach to prove the result is to apply the extreme value theory to (4.17) directly. The alternative approach is to exploit the lower bound (4.19) and the upper bound (4.23).

Before starting with the proof of Proposition 4, we present the following lemma that is used in the proof.

Lemma 14. $\mathrm{E}[\log(1 + \frac{1}{N_T}\max_j ||\mathbf{g}_j||_2^2)]$ *is upper bounded by* $\log(1 + N_E)$.

The proof is deferred to Appendix 4.8.2.

Proof. We first consider the lower bound in (4.20). From Lemma 14, we conclude that the lower bound scales as $(N_T!)^{\frac{1}{N_T}}((\frac{1}{N_L})^{\frac{1}{N_T}}) - \log(1 + N_E)$, where the scaling of the first part is given by [ANDH08]. Now, choosing $c_1 = (N_T!)^{\frac{1}{N_T}}$ and $c_2 = \log(1+N_E)$ and considering that $\lim_{N_L\to\infty}(\frac{1}{N_L})^{\frac{1}{N_T}} = 0$ completes the proof. $\square$

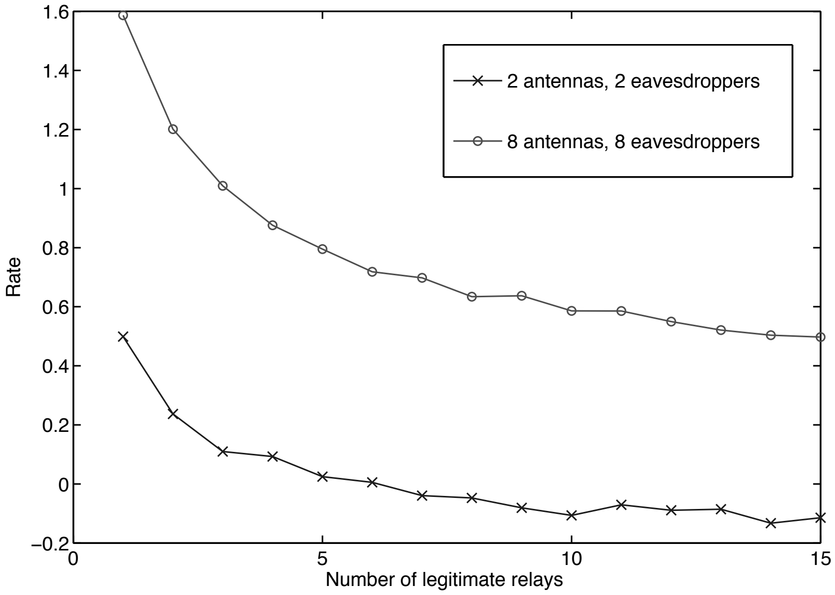

Figure 4.6.: Scaling of $\tilde{R}_{\rm sc}$ as the function of number of legitimate users for $N_T = 2$ and $N_E = 2$: average values of $\tilde{R}_{\rm sc}$ in the linear axes.

In Figure 4.6, the average rate of $\tilde{R}_{\rm sc}$ is plotted for 2 transmit antennas and 2 eavesdroppers and in Figure 4.7 for $N_T = 8$ and $N_E = 8$. In order to better illustrate scaling behaviour we leverage the fact that if $\tilde{R}_{\rm sc}$ scales as $\left(\frac{1}{N_L}\right)^{\frac{1}{N_T}} + C$, then $\log(\tilde{R}_{\rm sc} - C)$ is proportional to $-\frac{1}{N_T}\log N_L$. Therefore, a double logarithmic plot of $\log(\tilde{R}_{\rm sc} - C)$ as a function of $\log N_L$ shows asymptotically a straight line with the slope $-\frac{1}{N_T}$, which is presented in Figure 4.7.

Increasing eavesdroppers, constant legitimate users and antennas

Proposition 5. *If the number of eavesdroppers* N_E *increase and the number of legitimate users and antennas remain constant, then* $\mathrm{E}[\tilde{R}_{sc}]$ *scales as* $\Omega(\log(\log N_E)^{-1})$.

Proof. Reasoning along similar lines as in the proof of Proposition 4, we analyze the scaling rate of (4.20) as the lower bound. We proceed with applying some results in extreme value theory on the $\log(1 + \frac{1}{N_T}\max_j ||\mathbf{g}_j||_2^2)$. Theorem 8.3.2 in [ABN92] indicates

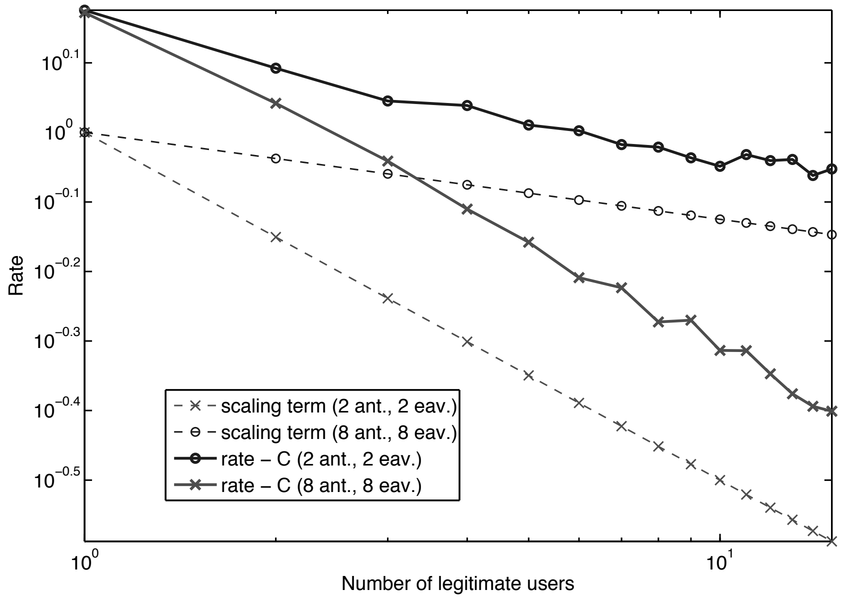

Figure 4.7.: Scaling of $\tilde{R}_{\mathrm{sc}}$ as the function of number of legitimate users for $N_T = 8$ and $N_E = 8$: $\tilde{R}_{\mathrm{sc}} - C$ in double-logarithmic axes together with the scaling terms.

that if $F(x)$ is given as the probability distribution of $||\mathbf{g}_j||_2^2$, then $\mathrm{E}[\max_{j\in\{1,\dots,N_E\}} ||\mathbf{g}_j||_2^2]$ scales according to

$$F^{-1}(1 - \frac{1}{N_E}). \tag{4.27}$$

Therefore, $\mathrm{E}[\frac{1}{N_T}\max_j ||\mathbf{g}_j||_2^2]$ approximately scales as $\log(N_E)$. We conclude that as $N_E \to \infty$, $\tilde{R}_{\mathrm{sc}}$ can be lower-bounded by $C - \log(1 + \log N_E)$ for some constant C, which results in a scaling law of $\Omega(-\log\log N_E)$. □

The average values of $\tilde{R}_{\mathrm{sc}}$ together with the corresponding scaling laws for this study case are presented in Figure 4.8.

Increasing number of transmit antennas, constant legitimate users and eavesdroppers

Proposition 6. $\mathrm{E}[\tilde{R}_{sc}]$ *scales as* $O(\log N_T)$ *while the number of antennas* N_T *increases and* N_E *and* N_L *are constants.*

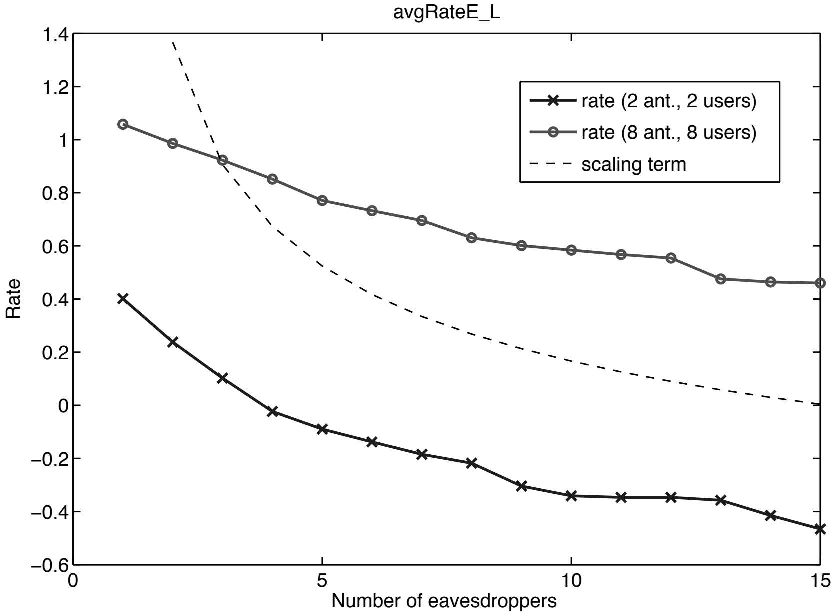

Figure 4.8.: Scaling of $\tilde{R}_{\mathrm{sc}}$ as a function of the number of eavesdroppers, together with the scaling term.

Proof. We apply the upper bound introduced in (4.23). The minimization in (4.23) is on a limited number of random variables (since N is constant), therefore, the scaling of

$$\min_{i}\{1+\|\mathbf{h}_i\|_2^2\}$$

is asymptotically equal to

$$\{1+\|\mathbf{h}\|_2^2\}$$

for any $\mathbf{h} \in \{\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_{N_L}\}$.

When N_T grows larger, based on law of large numbers, we have $\lim_{N_T\to\infty} \frac{\|\mathbf{h}\|_2^2}{N_T} = 1$. Thus, we infer the scaling rate of $\lambda_{\max}[\mathbf{A}]$ is (N_T).

Combining the results we conclude the proof with the achievable secrecy rate scaling at $O(\log(N_T))$. □

The average values of $\tilde{R}_{\mathrm{sc}}$ along with the corresponding scaling law for this study case is depicted in Figure 4.9.

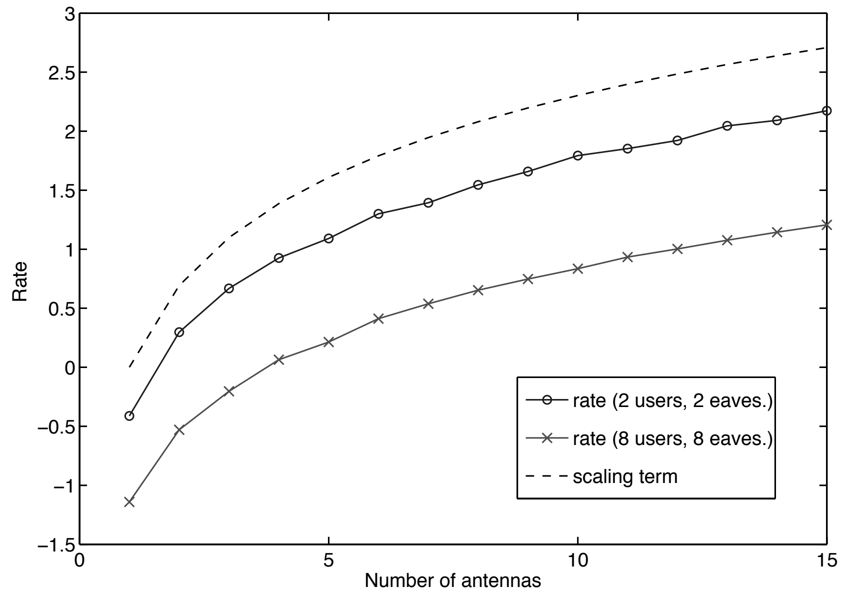

Figure 4.9.: Scaling of $\tilde{R}_{\text{sc}}$ as the function of the number of antennas.

4.6.3. Outage probability

As discussed in the previous sections, under unfavourable channel conditions there may be no positive rate achievable in the system, an event which is known as outage. Note that $R_{\text{sc}} \geq 0$ is a random variable which has a mixed type distribution with $\Pr(R_{\text{sc}} = 0) > 0$ while $\Pr(R_{\text{sc}} = r) = 0,\ \forall r > 0$. The performance measure of interest is the outage probability defined as $\Pr(R_{\text{sc}} = 0)$. As a complementary measure we sometimes use the probability of obtaining a positive rate $\Pr(R_{\text{sc}} > 0) = 1 - \Pr(R_{\text{sc}} = 0)$.

First, we present outage probability for two simple special cases.

Single transmit antenna

In the case of $N_T = 1$, we have

$$\begin{aligned}\Pr(R_{\text{sc}} = 0) &= 1 - \Pr(R_{\text{sc}} > 0)\\ &= 1 - \Pr\big(\min_{i,j} \frac{1 + |h_i|^2}{1 + |g_j|^2} > 1\big)\\ &= 1 - \Pr(\min_i |h_i|^2 > \max_j |g_j|^2).\end{aligned}$$

Single legitimate user, single eavesdropper

Since in this special case the minimum operator does not appear, we can make use of the closed-form solution from [SU07]. This leads directly to

$$\begin{aligned}\Pr(R_{\mathrm{sc}} = 0) &= \Pr(\lambda_{\max}[\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2}] \leq 1)\\ &= \Pr(\lambda_{\max}[\mathbf{B}^{-1}\mathbf{A}] \leq 1)\end{aligned}$$

with $\mathbf{A} = \mathbf{I} + \mathbf{h}\mathbf{h}^H$ and $\mathbf{B} = \mathbf{I} + \mathbf{g}\mathbf{g}^H$. However, directly finding the spectral radius is difficult.

In the following, we provide some bounds on the probability of positive rate in the general case. Since there holds $\Pr(R_{\mathrm{sc}} = 0) = \Pr(\tilde{R}_{\mathrm{sc}} \leq 0)$, we can use the bounds on the $\tilde{R}_{\mathrm{sc}}$ introduced in Section 4.6.1. In particular, for any lower bound $L(\tilde{R}_{\mathrm{sc}})$ and upper bound $U(\tilde{R}_{\mathrm{sc}})$ such that $L(\tilde{R}_{\mathrm{sc}}) \leq \tilde{R}_{\mathrm{sc}} \leq U(\tilde{R}_{\mathrm{sc}})$, we have

$$\Pr(U(\tilde{R}_{\mathrm{sc}}) \leq 0) \leq \Pr(\tilde{R}_{\mathrm{sc}} \leq 0) \leq \Pr(L(\tilde{R}_{\mathrm{sc}}) \leq 0).$$

Therefore, using (4.20) we obtain the following upper bound

$$\begin{aligned}\Pr(R_{\mathrm{sc}} = 0) &= 1 - \Pr(\tilde{R}_{\mathrm{sc}} > 0)\\ &\leq 1 - \Pr(\min_i ||\mathbf{h}_i||^2 > \max_j ||\mathbf{g}_j||^2) \qquad (4.28)\end{aligned}$$

where the result follows analogously as in the single antenna special case. As for the lower bound, considering (4.24) with (4.25) and (4.26) there holds

$$\begin{aligned}\Pr(R_{\mathrm{sc}} = 0) &= \Pr(\tilde{R}_{\mathrm{sc}} \leq 0)\\ &\geq \Pr\Big(\frac{1 + \frac{1}{N_L}\lambda_{max}[\mathbf{H}\mathbf{H}^H]}{1 + \frac{1}{N_E}\lambda_{min}[\mathbf{G}\mathbf{G}^H]} \leq 1\Big)\\ &= \Pr\big(\frac{1}{N_L}\lambda_{max}[\mathbf{H}\mathbf{H}^H] \leq \frac{1}{N_E}\lambda_{min}[\mathbf{G}\mathbf{G}^H]\big).\end{aligned}$$

Although exact computation of extreme eigenvalues is challenging, the presented bounds can be applied to characterizing conditions on outage probability in the large system limit, when $N_L, N_E, N_T \to \infty$. The outage probability and the upper bound based on (4.28) are plotted in Figure 4.10. The bounds are independent of the number of transmit antennas and, in consequence, are relatively tight only for low N_T.

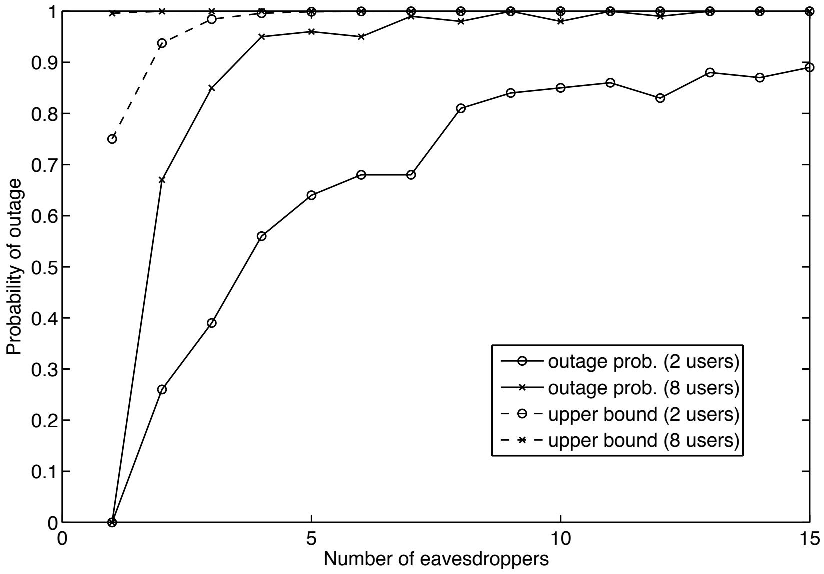

Figure 4.10.: Outage probability $\Pr(R_{\mathrm{sc}} = 0)$ vs. the number of eavesdroppers, together with corresponding upper bounds.

4.7. Conclusions

A framework for designing cross-layer security schemes has been developed. We proposed a new measure of security, referred to as α-secrecy. The idea is to use information-theoretic secrecy in conjunction with a higher layer scheme, e.g. secure network coding, to provide security. The resulting secrecy scheme does not necessarily fulfil the Shannon's perfect secrecy, but it provides higher transmission rates. We have studied two main scenarios with DF relays and AF relays.

In the case of DF relays, we proposed two schemes that can achieve any arbitrary $0 \leq \alpha < 1$. For the case of AF relays, we proposed an optimization problem that achieves $\alpha = 1$. A partially secure scheme is suggested that can be tuned to obtain a given α. The initial problem for AF relays results in a non-convex optimization problem, which is tackled using Charnes-Cooper transformation. For the case of partially secure scheme, we use beamforming to send each pre-coded message to some certain relays while assuming the rest of the relays as eavesdroppers. The pre-coding uses a simple higher layer scheme. The resulting optimization problem is of mixed-integer type, thus, difficult to solve.

To gain insight into the optimization problem, we study its asymptotic behaviour. We considered three cases, namely: Increasing number of legitimate users while the number of antennas and eavesdroppers remain constant, Increasing number of eavesdroppers while the number of antennas and legitimate users remain constant, and increasing number of antennas while the total number of users remain constant. We have provided some lowerbound and upperbound on the original problem since it is too complicated to study the original problem directly.

Future work

A cross-layer design that combines the provided schemes is worth investigating in an information-theoretic way. In Section 3.5 an initial attempt has been proposed, however, more detailed studies are required. The definition of α-secrecy, although, provides some flexibility in designing secure systems, does not provide a practical meaning in the sense of information theoretic secrecy.

4.8. Appendix

4.8.1. Proof of inequality (4.26)

The goal is to show that

$$\min_{\mathbf{\Sigma}} \max_{j} (1 + \mathbf{g}_j^H \mathbf{\Sigma} \mathbf{g}_j) \geq (1 + \frac{1}{N_E} \lambda_{min}[\mathbf{G}\mathbf{G}^H)]$$

where $\mathbf{G}$ is defined in Section III. First note

$$\min_{\mathbf{\Sigma}} \max_{j} (1 + \mathbf{g}_j^H \mathbf{\Sigma} \mathbf{g}_j) \geq \min_{\mathbf{\Sigma}} \frac{1}{N_E} \sum_{j} (1 + \mathbf{g}_j^H \mathbf{\Sigma} \mathbf{g}_j) \tag{4.29}$$

Now we prove the bound by solving the problem

$$\min_{\substack{\text{trace}[\mathbf{\Sigma}]=1 \\ \mathbf{\Sigma}\succeq 0}} \frac{1}{N_E} \sum_{j} (1 + \mathbf{g}_j^H \mathbf{\Sigma} \mathbf{g}_j).$$

This can be reformulated as

$$\begin{aligned} & \min_{\substack{\text{trace}[\mathbf{\Sigma}]=1 \\ \mathbf{\Sigma}\succeq 0}} 1 + \frac{1}{N_E} \text{trace}[\sum_{j} \mathbf{g}_j^H \mathbf{\Sigma} \mathbf{g}_j] \\ = & \min_{\substack{\text{trace}[\mathbf{\Sigma}]=1 \\ \mathbf{\Sigma}\succeq 0}} 1 + \frac{1}{N_E} \text{trace}[\mathbf{G}^H \mathbf{\Sigma} \mathbf{G}] \qquad (4.30) \\ = & \min_{\substack{\text{trace}[\mathbf{\Sigma}]=1 \\ \mathbf{\Sigma}\succeq 0}} 1 + \frac{1}{N_E} \text{trace}[\mathbf{\Sigma} \mathbf{G} \mathbf{G}^H]. \qquad (4.31) \end{aligned}$$

We rewrite the problem as

$$\min_{\substack{\mathrm{trace}[\mathbf{\Sigma}]=1\\ \mathbf{\Sigma}\succeq 0}} 1+\frac{1}{N_E}\mathrm{trace}[\mathbf{\Sigma}^{'}\mathbf{\Lambda}] \tag{4.32}$$

where $\mathbf{U}_G\mathbf{\Lambda}\mathbf{U}_G^H$ and $\mathbf{\Sigma}^{'} \triangleq \mathbf{U}_G^H\mathbf{\Sigma}\mathbf{U}_G$ is the eigenvalue decomposition (EVD) of $\mathbf{G}\mathbf{G}^H$. We use $\mathrm{trace}[\mathbf{AB}] = \mathrm{trace}[\mathbf{BA}]$ to achieve (4.32) from (4.30). The problem can be simplified to a scalar one,

$$\begin{aligned} &1+\frac{1}{N_E}\min_{\sum_i \sigma_{ii}^{'}=1}\sum_i \lambda_i\sigma_{ii}^{'} \\ &\geq 1+\frac{1}{N_E}\min_i \lambda_i \sum_i \sigma_{ii}^{'} \\ &= 1+\frac{1}{N_E}\min_i \lambda_i \end{aligned}$$

where λ_is and $\sigma_{ii}^{'}$s are diagonal elements of Λ and $\mathbf{\Sigma}^{'}$, respectively. Since this holds for any $\mathbf{\Sigma}$, the proof is complete.

4.8.2. Proof of Lemma 14

Using Jensen's inequality, we have

$$\mathrm{E}[\log(1+\frac{1}{N_T}\max_j ||\mathbf{g}_j||_2^2)] \leq \log(1+\frac{1}{N_T}\mathrm{E}[\max_j ||\mathbf{g}_j||_2^2]). \tag{4.33}$$

In the following we determine an upper bound for $\mathrm{E}[\max_j ||\mathbf{g}_j||_2^2]$.

The variables $||\mathbf{g}_j||_2^2$ are i.i.d. chi-square random variables with N_T degrees of freedom and let $f_{\chi^2,L}(t)$ and $F_{\chi^2,L}(t)$ denote their probability density function (pdf) and the cumulative distribution function (cdf), respectively. Now the expected value of $||\mathbf{g}_j||_2^2$ is given by

$$\mathrm{E}[||\mathbf{g}_j||_2^2] = \int t f_{\chi^2,L}(t)dt = N_T, \quad \forall j \in \{1,\dots N_E\}.$$

Consider $\tau \triangleq \max_j ||\mathbf{g}_j||_2^2$ and note that its cdf is given by $F_\tau(t) = F_{\chi^2,L}^{N_E}(t)$ and its pdf can be expressed as $f_\tau(f) = N_E F_{\chi^2,L}^{N_E-1}(t) f_{\chi^2,L}(t)$. Thus, the expected value of τ is also bounded because

$$\begin{aligned} \mathrm{E}[\tau] = \int t f_\tau(t) dt &= N_E \int t F_{\chi^2,L}^{N_E-1}(t) f_{\chi^2,L}(t) dt \\ &\leq N_E \int t f_{\chi^2,L}(t) dt = N_E N_T, \end{aligned} \tag{4.34}$$

where the inequality follows from $F_{\chi^2,L}(t) \leq 1$. We substitute (4.34) in (4.33) and obtain the upper bound $\log(1 + N_E)$.

Part II.

Anomaly Detection Algorithms; Implementation and Development

5. Anomaly detection algorithms for wireless sensor networks

A wireless network's security can be seized by means of two different types of attacks: passive attacks and active attacks. So far, in Part I of this dissertation, we have considered the passive attacks, where an adversary eavesdroppers the data. In Part II, on the other hand, we focus on developing some techniques that deal with active security threats in a wireless network. In this chapter, we consider the problem of jamming detection, whereas in the next chapter we deal with distributed implementation of some detection algorithms.

Combating an active attack in the wireless sensor networks leads to many important challenges, however, the first challenge is to be able to reliably detect the presence of an intruder. A proper anomaly detection algorithm for a wireless sensor network has to address the following problems in addition to detecting an anomaly. The first one is implementing the algorithms on geographically distributed nodes such that they are robust and scalable to the network's size, which is the topic of the next chapter. The second issue is power constraint at each node, which is studied in this chapter. This problem arises in many practical scenarios where the sensors are operating on batteries.

This chapter chiefly focuses on the development of the multi-objective detection algorithms. As for the power restriction, we study the fundamental trade-off between detection performance and power consumption. The results in this chapter are presented in the context of a cognitive radio (CR) scenario, because of their vulnerabilities to jamming attacks.

5.1. Background and contribution

Cognitive radio systems are designed to tackle the resource dearth in wireless communication systems by presenting a new approach to spectrum efficiency. Potential efficiency gain stem from the dynamic spectrum use as opposed to traditional fixed spectrum assignments. Traditionally the spectrum is divided into licensed sections and unlicensed ones. The unlicensed spectrum is overwhelmingly occupied by different users, which caused a huge spectrum shortage. The licensed spectrum, which is solely used by the primary users (PU), is not fully occupied by the PUs at all times and all frequencies. In order to use the spectrum more efficiently, a set of users, called the secondary users (SU), is supposed to exploit the empty time-frequency spots in the licensed spectrum. The empty time-frequency spots appear random in both time and frequency. Although we might be able to learn the pattern of appearance of empty spots and predict them, to avoid any interference to the PUs, the SUs perform spectrum sensing. Spectrum sensing is a detection technique in which the SUs execute constantly to detect the presence of the PUs. Spectrum sensing legalizes the SUs to scan the spectrum, consequently, to listen to the PU [Hay05], [SHT04]. Even though this feature of CR systems enables the possibility of utilizing the spectrum in a more efficient manner, it may also introduce security vulnerabilities.

5.1.1. Related work

Researchers present different approaches for spectrum sensing. For instance, energy detection, matched-filter, and cyclostationarity-based detection are presented respectively in [2], [CTB06], and [ALLP12]. Among those, energy detection is a long-established method since its performance is independent of the PU signal structure. In order to increase the accuracy and robustness of the energy detector, the authors in [GS05] propose a cooperative energy detection. They apply hard decision rule, an “OR” fusion rule [GS05], whereas in general the soft decision rules outperform the hard decision rules [VKP05]. The common assumption when applying energy detector is that the observation signals are independent, however, for applications related to the wireless channel that data is aggregated via multiple correlated nodes, this assumption is not accurate enough. In the case of spatially correlated signals, the authors in [LZZQ11] show that the weighted energy detector is optimal for the case that one hypothesis has correlated observations while the other hypothesis has independent signals (null hypothesis). Moreover, [LZZQ11] does not present the probability distribution of the weighted energy detector test statistic. Thus, no analytical results on the performance are established.

As far as security is concerned, one may note that certain peculiarities of CR networks make them quite vulnerable not only to the conventional threats targeted to wireless communications but also to the specific type of attacks that are tailored to CR networks [ATV^{+}12]. One of the most common attacks that could target CR networks is jamming. Jamming can be performed in different forms, for instance, constant jamming, random jamming, reactive jamming, and so on, which can target both PU and the SUs [ATV^{+}12]. There are also more intelligently designed jammers that deceive the SUs by mimicking the signal characteristics of the PU to force the SUs to vacate the spectrum [HS14]. This is known as primary user emulation attacks (PUEA) in [HS14, XS14].

5.1.2. Results and contributions

In this chapter, we consider two of the main issues of the SUs, namely spectrum sensing and jamming detection, jointly. In order to model the two problems simultaneously, we adopt multiple hypothesis testing with four cases. Building upon the work of [LZZQ11], we further derive the optimal detector when both of the hypotheses in each test are consists of correlated signals. Furthermore, we derive the exact and the asymptotic probability of detection of the optimal detector. We apply the generalized likelihood ratio test (GLRT) for practical scenarios where either the power of the jammer is unknown or the noise covariance matrix is unknown. Although there are similarities in the assumptions of this work and the research on PUEA, our work does not fall into that category owing to its different approach and objective.

We summarize the contributions of this chapter as in the following.

- We address the problem of joint detection of the jammer and the PU in the context of a multiple hypothesis testing problem and analyse the performance.
- We derive the exact and the asymptotic probability distributions of the test statistic. For the former, we present the cumulative distribution function (CDF) in terms of Meijer's G function.
- We present a correlated GLRT solution for the practical scenarios of unknown jammer's power and unknown noise's power.
- We study the problem of node placement in the context of spectrum sensing performance with limited power consumption.
- The node placement results in a non-convex optimization problem. Two heuristic methods are proposed to tackle the problem.

5.2. Problem statement

We consider a pair of PU transceivers communicating in a licensed frequency band and K SUs performing cooperative spectrum sensing. Furthermore, there is a jammer which sends jamming signals in random intervals to interrupt the PU and the SUs as in Figure 5.1. The main goal of the SUs is to determine the presence of the PU and discriminate it from the jammer. We assume the SUs send the data to a fusion center (FC) for the decision-making process. The SUs are desired to distinguish the following spectrum states out of the observation signal: the channel is idle, only the PU is present, only the jammer is present, and both the PU and the jammer are present. The sensed signal at the ith SU

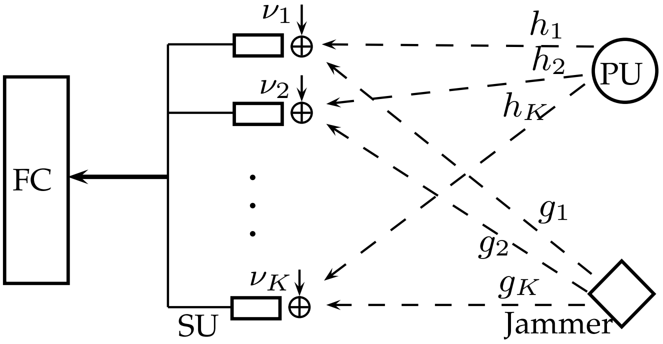

Figure 5.1.: The system setup with K SUs, a PU, a Jammer, and a FC to make the decision.

node, $i \in \{1, \ldots, K\}$, under four hypotheses at the nth time instant, $n \in \{1, \ldots, N\}$, is

$$\begin{aligned}
&H_o : x_i(n) = \nu_i(n) \\
&H_1 : x_i(n) = \sqrt{w_1} h_i(n) s(n) + \nu_i(n) \\
&H_2 : x_i(n) = \sqrt{w_2} g_i(n) u(n) + \nu_i(n) \\
&H_3 : x_i(n) = \sqrt{w_1} h_i(n) s(n) + \sqrt{w_2} g_i(n) u(n) + \nu_i(n)
\end{aligned}$$

where, $s(n)$ and $u(n)$ are complex signals with average power 1 from the PU and the jammer, respectively. The channel coefficients between the PU and the ith SU node and between the jammer and the ith SU node are $h_i(n)$ and $g_i(n)$, respectively. Further, the signal power received from the PU transmitter and the jammer is $0 < w_1$ and $0 < w_2$, respectively. Further assumptions on the channels and the noise are provided as follows:

1. The channel coefficients $h_i(n)$ and $g_i(n)$ are assumed to be Rayleigh-distributed with $\mathrm{E}[|h_i(n)|^2] = \mathrm{E}[|g_i(n)|^2] = 1$, $\forall i \in \{1, \ldots, K\}$, w. l. g. [1]

2. We assume the channels are spatially correlated but independent temporally. The channel correlation coefficient matrix for $\mathbf{g} := [g_1, \ldots, g_K]$ and $\mathbf{h} := [h_1, \ldots, h_K]$ is defined as $\mathbf{\Sigma} := \mathrm{E}[\mathbf{g}\mathbf{g}^H] = \mathrm{E}[\mathbf{h}\mathbf{h}^H]$. Further, the (i, j)th element of $\mathbf{\Sigma}$, is approximated by

$$[\mathbf{\Sigma}]_{i,j} = \begin{cases} e^{-\rho d_{i,j}} & i \neq j \\ 1 & i = j \end{cases} \tag{5.1}$$

---

[1] A more general case of $\mathrm{E}[|h_i(n)|^2] \neq \mathrm{E}[|g_i(n)|^2]$ with $\mathrm{E}[|h_i(n)|^2] = \mathrm{E}[|h_j(n)|^2]$ and $\mathrm{E}[|g_i(n)|^2] = \mathrm{E}[|g_j(n)|^2]$, $\forall i, j$ can be considered with no extra elaboration.

where $d_{i,j}$ is the Euclidean distance between nodes i and j which can be derived from the node position matrix $\mathbf{Y} \triangleq [\mathbf{y}_1, \ldots, \mathbf{y}_K] \in \mathbb{R}^{2\times K}$ as $d_{i,j} = \|\mathbf{y}_i - \mathbf{y}_j\|$ and ρ is the correlation constant depending on the environment [Gud91].

3. Since it is reasonable to assume that the locations of the jammer and the PU are separated large enough[2], it follows that the g_is and h_is are mutually independent.

4. The noise at the time n and the node i is depicted by $\nu_i(n)$, which is assumed to be i.i.d. circularly symmetric Gaussian with $\mathcal{CN}(0, \sigma_\nu^2)$. Further, we assume that $s(n)$, $u(n)$, and $\nu_i(n)$ are, w. l. g., independent.

The distribution of the observation signal under H_l hypothesis becomes,

$$\mathbf{x} \sim \mathcal{CN}(0, w_l\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I}), \qquad \mathbf{x} \in H_l \tag{5.2}$$

where $\mathbf{x} := [x_1, \ldots, x_K]^T$ and for $H_l,\ l \in \{0,1,2,3\}$ we have $w_0 := 0$ and $w_3 := w_1 + w_2$. We assume $w_1 \neq w_2$ throughout this paper since distinguishing H_1 from H_2, with $w_1 = w_2$ and having the above conditions, using energy detector is absurd [3].

5.3. Optimal a posteriori detection

In the interest of optimally selecting the correct hypothesis, we find a hypothesis that maximizes

$$P(\mathbf{x}|H_l)\Pr(H_l), \qquad l \in \{0,1,2,3\}, \tag{5.3}$$

where the $\Pr(H_l)$ is *a priori* probability of H_l [Kay93]. In order to evaluate $\arg\max_l P(\mathbf{x}|H_l)\Pr(H_l)$ we compare every two hypotheses using:

$$\frac{P(\mathbf{x}|H_l)\Pr(H_l)}{P(\mathbf{x}|H_j)\Pr(H_j)} \underset{H_j}{\overset{H_l}{\gtrless}} 1, \ \ 0 \leq l, j \leq 3, \ l \neq j \tag{5.4}$$

Then we perform the binary search to find the maximum. We compare H_0 and H_1 and in parallel H_2 and H_3 using (5.4). Then we compare the “winners” of each of the previous tests to determine the overall selected hypothesis.

---

[2] More than half of the wavelength of the carrier frequency [TV05].

[3] One has to exploit other aspects of the jammer’s signal to be able to detect efficiently in such a condition [HS14].

In this section, we study the performance of the detection tests in (5.4). Based on Section 5.3, we need to make decisions between every H_l and H_j, $l, j \in \{0, 1, 2, 3\}$. In the following, we consider three cases, namely, the average power of the jammer and the noise are known, the average power of the jammer is unknown while the noise power is given, and the noise power is unknown while the jammer power is given.

5.3.1. Known jammer's power and noise's power

In this case, based on the Neyman-Pearson theorem, one can apply log likelihood ratio (LLR) to achieve optimal detector. The following propositions present the optimal detector's performance for two hypotheses H_l and H_j where $l, j \in \{0, 1, 2, 3\}$.

Proposition 7. *Assume that we are to decide between any two hypotheses H_l and H_j in (5.2). The Neyman-Pearson optimal test statistic is given by,*

$$T_{l,j} = \sum_{n=1}^{N} \sum_{i=1}^{K} \lambda_i |z_i(n)|^2 \underset{H_j}{\overset{H_l}{\gtrless}} \vartheta_{l,j}, \tag{5.5}$$

where $\mathbf{S}_{l,j} = \mathbf{V}\text{diag}[\lambda_1, \ldots, \lambda_K]\mathbf{V}^H$ *with,*

$$\mathbf{S}_{l,j} := (w_j \boldsymbol{\Sigma} + \sigma_\nu^2 \mathbf{I})^{-1} - (w_l \boldsymbol{\Sigma} + \sigma_\nu^2 \mathbf{I})^{-1}, \tag{5.6}$$

$\mathbf{z} := \mathbf{V}^H \mathbf{x}$ *and* $\text{diag}[\cdot]$ *returns a diagonal matrix with its argument as the diagonal elements. Moreover,* $\vartheta_{l,j} := -\ln \frac{\Pr(H_l)}{\Pr(H_j)} - N(\ln \det[w_j \boldsymbol{\Sigma} + \sigma_\nu^2 \mathbf{I})] + N(\ln \det[w_l \boldsymbol{\Sigma} + \sigma_\nu^2 \mathbf{I})]$*, is the decision threshold.*

The proof is provided in the Appendix 5.7.1.

Note that for the case of correlated observations, conventional energy detector is not only suboptimal in the sense of Neyman-Pearson theorem, but also, the derivation of the probability distribution of the test statistic involves handling sum of correlated random variables [KLWH09].

To calculate a sensing performance measure, e.g. the probability of detection, we need to find the distribution of (5.5). The following propositions yield the two exact versions, and an asymptotic version of $T_{l,j}$ distribution.

Proposition 8. *The CDF of (5.5) conditioned on H_l and tested against H_j, is given by*

$$\Pr(T_{l,j} \leq \vartheta_{l,j} | H_l) = C \sum_{q=0}^{\infty} \int_{t=0}^{\vartheta_{l,j}} \frac{\delta_q t^{(NK+q-1)} e^{\frac{-t}{\lambda_1 \theta_1}}}{(\lambda_1 \theta_1)^{NK+q} \Gamma(NK+q)} dt, \tag{5.7}$$

where $C := \prod_{i=1}^{K} \left(\frac{\lambda_1\theta_1}{\lambda_i\theta_i}\right)^N$, $\lambda_1\theta_1 = \min_i\{\lambda_i\theta_i\}$, *and* θ_i *is the ith largest eigenvalue of* $(w_l\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I})$. *Further, we obtain* δ_q *using the following recursive formula,*

$$\begin{aligned}
&\delta_0 = 1\\
&\delta_{q+1} = \frac{N}{q+1}\sum_{k=1}^{q+1}\left[\sum_{i=1}^{K}(1-\frac{\lambda_1\theta_1}{\lambda_i\theta_i})^k\right]\delta_{q+1-k},\\
&\qquad\qquad\qquad k = 0,1,2,\dots
\end{aligned}$$

The proof is provided in the Appendix 5.7.2.

In practice, we need to truncate the series in (5.7) depending on the required precision [Mos85]. To present a more compact and closed-form formula, we can apply the recent results in [AYAK12].

Corollary 1. *The exact CDF of (5.5) conditioned on* H_l *and tested against* H_j, *is given by*

$$\Pr(T_{l,j} \leq \vartheta_{l,j}|H_l) = \prod_{i=1}^{K}(\frac{1}{2\lambda_i\theta_i})^N \ \mathrm{G}_{1+NK,1+NK}^{1+NK,0}\left[e^{-\vartheta_{l,j}}\middle|\begin{matrix}\Psi_K^{(1)},1\\ \Psi_K^{(2)},0\end{matrix}\right]$$

in which $\Psi_K^{(1)}$, $\Psi_K^{(2)} \in \mathbb{R}_{1\times NK}$ *are defined as in the following,*

$$\begin{aligned}
\Psi_K^{(1)} :=& \{\overbrace{(1+\frac{1}{2\lambda_1\theta_1}),\ \dots,(1+\frac{1}{2\lambda_1\theta_1})}^{N-times},\ \dots\\
&\qquad\qquad \dots,\overbrace{(1+\frac{1}{2\lambda_K\theta_K}),\ \dots,(1+\frac{1}{2\lambda_K\theta_K})}^{N-times}\}
\end{aligned} \tag{5.8}$$

$$\Psi_K^{(2)} := \underbrace{\{\overbrace{(\frac{1}{2\lambda_1\theta_1}),\ \dots,(\frac{1}{2\lambda_1\theta_1})}^{N-times},\ \dots,\overbrace{(\frac{1}{2\lambda_K\theta_K}),\ \dots,(\frac{1}{2\lambda_K\theta_K})}^{N-times}\}}_{K-times}$$

and $\mathrm{G}_{1+NK,1+NK}^{1+NK,0}$ is the Meijer's G function which is given in [AYAK12][Corollary 3] as,

$$\mathrm{G}_{p,q}^{m,n}\left[z\middle|\begin{matrix}\alpha_1,\ \dots,\alpha_p\\ \beta_1,\ \dots,\beta_q\end{matrix}\right] := \frac{1}{2\pi i}\oint_C \frac{\prod_{j=1}^{n}\Gamma(1-\alpha_j+s)\prod_{j=1}^{m}\Gamma(\beta_j-s)}{\prod_{j=n+1}^{p}\Gamma(\alpha_j-s)\prod_{j=m+1}^{q}\Gamma(1-\beta_j+s)}z^s ds$$

interested readers can refer to [Ism80] for more details.

The proof is provided in the Appendix 5.7.3.

Considering (5.8), we obtain the exact probability of detection conditioned on H_l while tested against H_j as

$$\Pr(T_{l,j} > \vartheta_{l,j}|H_l) = 1 - \prod_{i=1}^{K}(\frac{1}{2\lambda_i\theta_i})^N \, \mathrm{G}_{1+NK,1+NK}^{1+NK,0}\left[e^{-\vartheta_{i,j}}\middle|\begin{matrix}\Psi_{NK}^{(1)},1\\ \Psi_{NK}^{(2)},0\end{matrix}\right] \tag{5.9}$$

Calculating the exact forms might be computationally involved. Moreover, the asymptotic behaviour of the distribution of (5.5) delivers useful insights on studying optimum sensing strategies. Hence, in the following we present the distribution of (5.5) when $N \to \infty$.

Proposition 9. *The asymptotic probability distribution of (5.5) when $N \to \infty$, conditioned on H_l tested against H_j, is*

$$T_{l,j} \sim \mathbb{N}\left(N\sum_{i=1}^{K}\theta_i\lambda_i, N\sum_{i=1}^{K}(\theta_i\lambda_i)^2\right), \tag{5.10}$$

where θ_is are eigenvalues of $w_l\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I}$.

The proof is provided in the Appendix 5.7.4. Since we do not have the values for the *a priori* probabilities, to determine $\vartheta_{l,j}$ we proceed as in the following. First, we define the probability of false alarm. Then for a fixed probability of false alarm we obtain the corresponding $\vartheta_{l,j}$, which can be used to compute the probability of detection of different hypotheses. The probability of false alarm of H_l tested against H_j, using the Proposition 9, is given by

$$\Pr(T_{l,j} > \vartheta_{l,j}|H_j) = Q\left(\frac{\vartheta_{l,j} - N\sum_{i=1}^{K}\tilde{\theta}_i\lambda_i}{\sqrt{N\sum_{i=1}^{K}(\tilde{\theta}_i\lambda_i)^2}}\right), \tag{5.11}$$

where, $\tilde{\theta}_i$s are the eigenvalues of $w_j\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I}$ and $Q(x) = 1/\sqrt{2}\int_x^\infty \exp(-t^2/2)dt$ is the tail probability of the normal distribution.

5.3.2. Unknown jammer's power

One of the most widely used methods to deal with an unknown parameter is GLRT. Basically, in GLRT one obtains an estimate of the unknown parameter by applying maximum likelihood (ML) estimate.

In typical scenarios, we have no information of the jammer; therefore, we need to estimate the covariance matrix of the observed signal under H_2 as well as H_3. In our channel model, however, we only need to know the average power of the jammer w_2, because the covariance matrix depends on the sensors' locations.

Assuming other parameters are known, we estimate w_2 in the following.

Proposition 10. *Assuming the H_2 in (5.2), the ML estimate of w_2 is given by*

$$\hat{w_2} = \frac{\sum_{i=1}^{K}\sum_{n=1}^{N}|z_i(n)|^2 - NK\sigma_\nu^2}{N\sum_{i=1}^{K}\tau_i}, \tag{5.12}$$

where, τ_is are defined the eigenvalues of $\mathbf{\Sigma} := \mathbf{V}\mathrm{Diag}[\tau_1, \ldots, \tau_K]\mathbf{V}^H$.

The proof is presented in Appendix 5.7.5.

Next, we need to substitute (5.12) into (5.27). Therefore, the LRT for H_2 becomes:

$$\begin{aligned} &-N(\ln\det[w_j\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I})] + N(\ln\det[\hat{w_2}\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I})] \\ &\quad -\sum_{n=1}^{N}\mathbf{x}^H(n)\left((w_j\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I})^{-1} - (\hat{w_2}\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I})^{-1}\right)\mathbf{x}(n) \gtrless 0, \end{aligned}$$

which is difficult to further analyse.

5.3.3. Unknown noise power

In the case of unknown noise, similar to the previous section, we use ML estimation. Based on our assumption, we know that noise covariance matrix is diagonal with equal diagonal elements. Hence, the matrix can be described with a single parameter.

Proposition 11. *The ML estimate of σ_ν^2 for H_j*

$$\begin{aligned} \hat{\sigma}_\nu^2 = \arg\max_{\sigma_\nu^2} - NK\ln\pi - N\ln\det\left(w_j\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I}\right) \\ - N\mathrm{trace}[w_j\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I}]^{-1}\mathbf{R} \end{aligned} \tag{5.13}$$

in which $\mathbf{R} := \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)\mathbf{x}^H(n)$ is the sample covariance matrix of the received signal. The ML estimated σ_ν^2 is

$$\hat{\sigma}_\nu^2 = \frac{1}{NK}\sum_{i=1}^{K}\sum_{n=1}^{N}|z_i(n)|^2 - \frac{1}{K}\sum_{i=1}^{K}w_j\tau_i \tag{5.14}$$

The proof is similar to that of Proposition 10 given in Appendix 5.7.5.

We need to substitute (5.14) in LRT to achieve the test statistic. This in turn is given in the following,

$$\begin{aligned}&-N(\ln\det[w_j\mathbf{\Sigma}+\hat{\sigma}_\nu^2\mathbf{I}])+N(\ln\det[w_l\mathbf{\Sigma}+\hat{\sigma}_\nu^2\mathbf{I}])\\&-\sum_{n=1}^{N}\mathbf{x}^H(n)\left((w_j\mathbf{\Sigma}+\hat{\sigma}_\nu^2\mathbf{I})^{-1}-(w_l\mathbf{\Sigma}+\hat{\sigma}_\nu^2\mathbf{I})^{-1}\right)\mathbf{x}(n)\gtrless 0,\end{aligned}$$

which is difficult to analyse.

5.4. Application to power constraint detection

In the following we apply the some of the results in Section 5.3 to an energy constraint detection problem. The details are presented in the following.

5.4.1. The power-detection performance trade-off

We are interested in the problem of SUs placement for detecting the presence of the PU with power constraints at the SUs and the absence of the jammer. The detection problem simplifies to the following binary hypothesis problem,

$$\begin{aligned}&H_o: x_i(n)=\nu_i(n)\\&H_1: x_i(n)=\sqrt{w_1}h_i(n)s(n)+\nu_i(n)\end{aligned}$$

The SUs need some power to transmit $x_i(n)$ to a fusion center via the wireless channel. The fusion center is located at $\mathbf{0}$ and labelled as node 1.

According to Kim et al. in [KLWH09], the spatial correlation between the SUs degrades the performance of detection. Therefore, the optimal node placement must minimize the spatial correlation of the SUs. According to our channel correlation model, this translates to placing the SUs as far as possible from each other. On the other hand, the power consumption increases with the increase of distances of the SUs to the fusion center. This is because of the pathloss characteristics of the wireless channel. More precisely, we model this trade-off as the following optimization problem.

$$\begin{aligned}\text{Problem I}:\quad &\max_{\mathbf{Y}}\quad \mathrm{P}_D\\&\text{s.t.}\quad \sum_i u_i\le\omega\\&\qquad\;\; u_i\le\hat{\omega}\quad\forall i,\end{aligned}\tag{5.15}$$

where, $\mathrm{P}_D := \mathrm{P}_D(\alpha)$ is the probability of detection for a given probability of false alarm $\mathrm{P}_F = \alpha$, which, for N large enough, is given by proposition (9)

$$\mathrm{P}_D(\vartheta) = Q\left(\frac{\vartheta - N\sum_{i=1}^{K}\lambda_i\theta_i}{\sqrt{N\sum_{i=1}^{K}\lambda_i^2\theta_i^2}}\right), \tag{5.16}$$

where and $\vartheta_{0,1} > 0$ is the decision threshold. Evaluating the probability of false alarm P_F in the case of Gaussian noise is straightforward and is given by

$$\mathrm{P}_F := \mathrm{P}_F(\vartheta_{0,1}) = Q\left(\frac{\vartheta_{0,1} - NK\sigma_\nu^2}{\sqrt{NK\sigma_\nu^4}}\right). \tag{5.17}$$

In the spectrum sensing literature, the constraint is given in the sense of the probability of false alarm since the level of noise is usually known. The performance of the test is then compared by using the probability of detection. According to the same logic, and since we need to have s statistical measure that considers the spatial correlation, we use the probability of false alarm to obtain $\vartheta_{0,1}$ and apply it as a threshold for the probability of detection. Therefore, P_D for $\mathrm{P}_F = \alpha \in [0,1]$ yields ($\mathrm{P}_D := \mathrm{P}_D(\vartheta^*)$ where ϑ^* is defined to be $\alpha = \mathrm{P}_F(\vartheta^*)$), which is given as,

$$\mathrm{P}_D = Q\left(\frac{Q^{-1}(\alpha)\sqrt{NK\sigma_\nu^4} + NK\sigma_\nu^2 - N\sum_{i=1}^{K}\lambda_i\theta_i}{\sqrt{N\sum_{i=1}^{K}\lambda_i^2\theta_i^2}}\right), \tag{5.18}$$

where, $Q^{-1}(\cdot)$ is the inverse of $Q(\cdot)$.

5.4.2. Optimization problem

In this section, we propose some heuristic approaches to the optimization problem in (5.15) with objective function given by (5.18). We assume a communication link between the ith SU and the fusion center is established only if the signal to noise ratio (SNR) at the receiver is larger than or equal to η. Assuming pathloss is associated with the squared distance of ith SU and the fusion center ($d_{i,1}^2$), we have

$$u_i = \kappa d_{i,1}^2 \sigma_\nu^2 \eta \tag{5.19}$$

where κ is the pathloss coefficient [TV05].

Problem I can be simplified according to the following proposition:

Proposition 12. *The problem in (5.15) is equivalent to*

$$\begin{aligned}\text{Problem II}: \quad & \min_{\mathbf{Y}} \quad c_1 G(\mathbf{Y}) \\ & \text{s.t.} \quad \sum_i d_{i,1}^2 \leq \omega c_2 \\ & \qquad\; d_{1,i}^2 \leq \hat{\omega} c_2 \quad \forall i,\end{aligned} \tag{5.20}$$

where $c_1 := \big(\frac{w_1\sqrt{N}}{Q^{-1}(\alpha)\sqrt{NK}\sigma_\nu^4 + NK\sigma_\nu^4 - w_1 NK}\big)^2$, $c_2 := (\sigma_\nu^2 \eta \kappa)^{-1}$, *and* $G(\mathbf{Y})$ *is given by*

$$G(\mathbf{Y}) := \begin{cases} \sum_{i,j} e^{-2\rho d_{i,j}} & \alpha > Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^4} - 1)) \\ -\sum_{i,j} e^{-2\rho d_{i,j}} & \alpha < Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^4} - 1)) \end{cases} \tag{5.21}$$

Proof. The proof is provided in the Appendix 5.7.6. □

Remark 16. When $\alpha < Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^4} - 1))$ then the argument in (5.18) is non-negative and since for $x \geq 0$ we have $Q(x) < 0.5$, this corresponds to a rather trivial case of $\mathrm{P}_D < 0.5$.

Remark 17. The Problem II is non-convex, but the constraints are convex. For a given setup (i.e. α, K, N, w_1, and σ_ν^2), $G(\mathbf{Y})$ has one objective function which remains the same for all $\mathbf{Y}$. In the following, we propose two suboptimal methods to deal with this problem.

5.4.3. Projected gradient method (PGM)

Since the constraints are convex, we can execute PGM. Moreover, based on the results in [CM87], because the objective function in (5.20) within each domain is continuously differentiable, the PGM converges. In the following, we first investigate the projection and then we proceed with the algorithm. The projection to domain of our constraints is given by,

$$J(\mathbf{Y}) = \arg\min\{\|\mathbf{Y} - \mathbf{W}\| : \ \|\mathbf{W}\|^2 \leq \omega c_2, \ \ \|\mathbf{w}_i\|^2 \leq \hat{\omega} c_2 \ \ \forall i\},$$

where, $\|\cdot\|$ represents an inner product induced vector norm (here norm-2), and $\mathbf{w}_i$ depicts the ith column of $\mathbf{W}$.

To calculate the projection into the intersection of the hypersphere ($\|\mathbf{W}\|^2 \leq \omega c_2$) and the hyperplane ($\|\mathbf{w}_i\|^2 \leq \hat{\omega} c_2, \ \forall i$) we require an iterative method that each time projects $\mathbf{Y}$ into one of them. However, in practical scenarios, the more non-trivial dominant constraint normally is the hypersphere. Therefore, we simplify the projection as,

$$J(\mathbf{Y}) = \sqrt{\omega c_2}\mathbf{Y}/\|\mathbf{Y}\|. \tag{5.22}$$

Based on the aforementioned discussion the detailed algorithm for solving Problem II is given Alg. 1. If the $\nabla G(\mathbf{Y})$ is Lipschitz continuous, then the $0 < \{\beta_m\}_{m=1}^{\infty} < \infty$ can be calculated using the method suggested by [Dun81].

Algorithm 1 Projected Gradient method (PGM)

```
Input: Random initial positions $\tilde{\mathbf{Y}}(:)$ and the constants, i.e. $K, N, c_1, c_2, \omega, \hat{\omega}, \epsilon$, and $\{\beta_m\}_{m=1}^{\infty}$.
Output: Suboptimal positions $\mathbf{Y}$, which is a feasible stationary point of the problem.
    repeat
        $\mathbf{Y}_{m+1} \leftarrow \mathbf{Y}_m - \beta_m \nabla G(\mathbf{Y}_m)$
        if $\|\mathbf{Y}\|^2 > \omega c_2$ then
            $\mathbf{Y}_{m+1} \leftarrow \sqrt{\omega c_2}\mathbf{Y}_{m+1}/\|\mathbf{Y}\|_{m+1}$
        end if
    until $\|\mathbf{Y}_{m+1} - \mathbf{Y}_m\| \leq \epsilon$
```

5.4.4. Fixed constraint method (FCM)

In the following we provide the intuitive motivation for FCM, then we provide the detailed algorithm in Alg. 2.

As in PGM, we assume in practical and non-trivial scenarios the total power constraint dominants the individual power constraints. Furthermore, the optimal solution satisfies the total power constraint with equality. The proof by contradiction is straightforward. Motivated by Karush-Kuhn-Tucker (KKT) conditions for (5.20), we conclude that the optimal constellation is geometrically symmetric around the fusion center. Building upon these arguments, we infer that a circle or multiple concentric circles centred at the fusion center are among the possible solutions. The radii of these circles are dictated by the constraints of the problem.

In our algorithm first, we assume the nodes are uniformly scattered on one circle, $r = \sqrt{\omega c_2/K}$, which satisfies the main constraint with equality. Then in order to increase the intra-node distances, we move some of the nodes toward the center and some outward the center of the circle such that the constraint is kept with equality. This way we produce two concentric circles in which the nodes consume the same power as one circle but provide higher P_D due to the increased intra-node distances. The algorithm for F is given in the following.

```
Algorithm 2 Fixed Constraint Method (FCM)
Input: The constants, i.e. $K, N, c_1, c_2, \omega, \hat{\omega}, \epsilon$, and $\{\beta_m\}_{m=1}^{\infty}$.
Output: Suboptimal positions $\mathbf{Y}(:)$.
$r_0 \triangleq \sqrt{\omega c_2/K}$
$\theta_i = \frac{2\pi(i-1)}{K-1} \quad i \in \{2, \dots, K\}$
$\begin{cases} r_i \leftarrow r_0 - t & i \in odd \\ r_i \leftarrow \sqrt{2r_0^2 - (r_0 - t)^2} & i \in even \end{cases} \quad (5.23)$
solve
    $\min_{0 \le t \le r_0} G(\mathbf{Y}(t)), \quad (5.24)$
    where, $d_{i,j} = \sqrt{r_i(t)^2 + r_j(t)^2 - 2r_i(t)r_j(t)\cos(\theta_i - \theta_j)}$. Problem (5.24) can be handled using a search algorithm.
$\forall i$ :
    $\mathbf{y}_i = \begin{bmatrix} r_i(t)\sin(\theta_i) \\ r_i(t)\cos(\theta_i) \end{bmatrix} \quad (5.25)$
```

5.5. Numerical examples

We provide numerical experiments to validate the results in this section. For testing the results in Section 5.3, we have the following assumptions. There are $K = 8$ SUs spread out in a square area of $10m \times 10m$ and we collect $N = 10$ independent time samples unless otherwise is specified. The values for ρ are 10 (almost independent) and 0.01 (highly correlated).

We first compare the exact CDF of $T_{l,j}$, given by (5.8), with the asymptotic CDF, presented by (5.10), to study the accuracy of the results. This is depicted in Figure 5.2. One can observe that with the increase in N and K the asymptotic results tend to be more accurate.

For the case of known power of the jammer and noise, we depict the receiver operating characteristic (ROC) of $\Pr(T_{1,0} \geq \vartheta_{1,0}|H_1)$ and $\Pr(T_{2,1} \geq \vartheta_{2,1}|H_2)$ in Figure 5.3. We notice that in the fully correlated case the weighted energy detector (WED) outperforms the conventional energy detector.

Unlike the case of $\Pr(T_{2,1} \geq \vartheta_{2,1}|H_2)$, in the case of $\Pr(T_{1,0} \geq \vartheta_{1,0}|H_1)$, the performance of WED boosts when correlation increases. One notes in the case of $\Pr(T_{2,1} \geq \vartheta_{2,1}|H_2)$, both of H_1 and H_2 have correlated observations, whereas in the case of $\Pr(T_{1,0} \geq \vartheta_{1,0}|H_1)$,

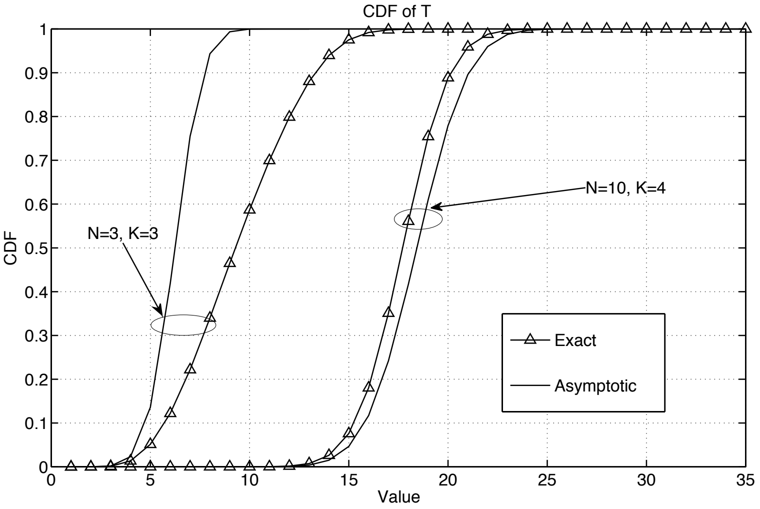

Figure 5.2.: The CDF of T versus its asymptotic CDF for the cases $N = 3, K = 3$ and $N = 10, K = 4$.

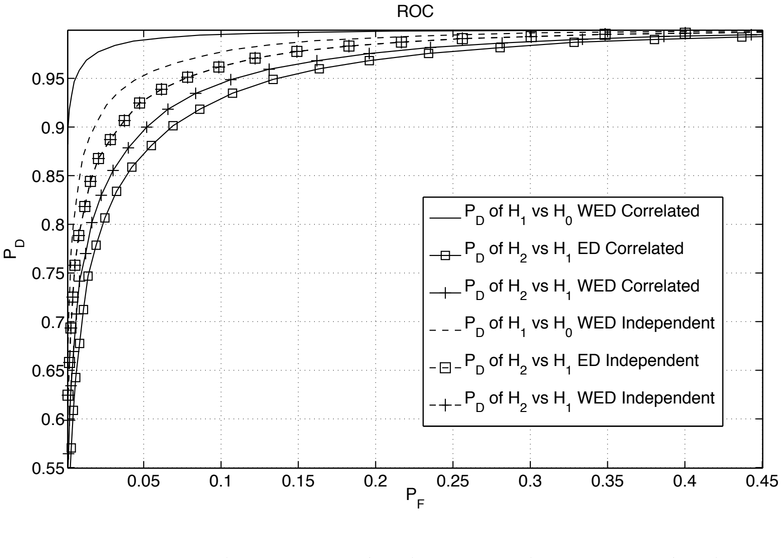

Figure 5.3.: ROC curve for $\Pr(T_{1,0} \geq \vartheta_{1,0}|H_1)$ and $\Pr(T_{2,1} \geq \vartheta_{2,1}|H_2)$ when we have full information.

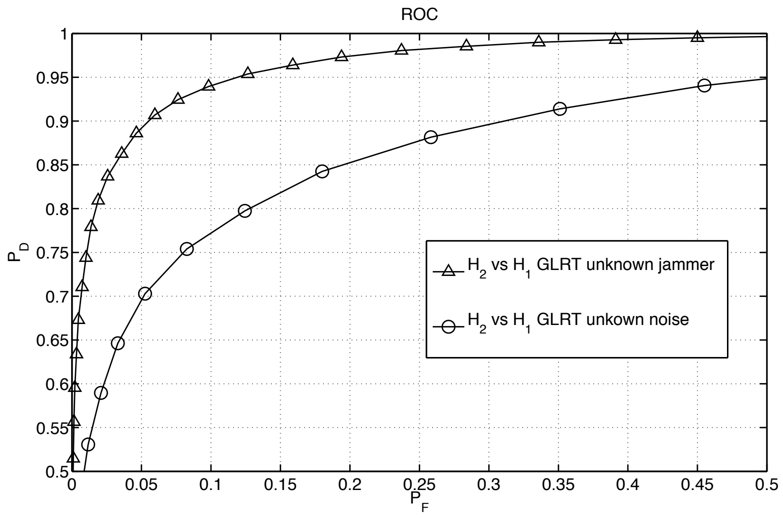

Figure 5.4.: ROC curve for $\Pr(T_{2,1} \geq \vartheta_{2,1}|H_2)$ when jammer power and noise power are unknown.

H_0 has independent observation. The performance of the case with unknown w_2 is better than the case with unknown noise power, as it is illustrated in Figure 5.4. The reason for performance degradation in GLRT in comparison to LRT is the error in ML estimation of the parameters. Since the ML estimation of noise appears in both hypotheses H_1 and H_2 the performance of the detector suffers more in comparison to the case that we incorporate the ML estimate of w_2, which appears only in H_2.

As for the results of the section 5.4 We first investigate the accuracy of our results on the probability of detection. Then we compare the performance of Alg.1 and Alg. 2 in solving (5.20). The analytical and empirical results of CDF of (5.10) is evaluated by Monte Carlo experiment of 6000 trials under random positioning of $K = 8$ nodes, $N \in \{40, 80, 120\}$ time samples, and $\rho = 0.01$. The outcome is given in Figure 5.5. We can observe the effect of using central limit theorem estimation. As a line of comparison, we have also depicted the distribution in the case of no signal ($\mathcal{H}_0$). In Figure 5.6 we compare the receiver operating characteristic (ROC) curve of ED and WED with different ρs.

In Figure 5.7 we illustrate the performances of PGM and FCM in terms of P_D versus K for the case of $\alpha > Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^2}))$. The experiment is repeated for different values of ρ. The performances of the PGM and the FCM are equal. It's noteworthy to mention that the outcome of the PGM and FCM are usually different as shown in Figure5.9 . As it is mentioned in [LZZQ11], with the increase of ρ (i. e. increase of correlation) the P_D increases as well because of the use of weighted energy detector. In order to observe the performance of PGM and FCM under the rather trivial condition of $\alpha < Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^4} - 1))$

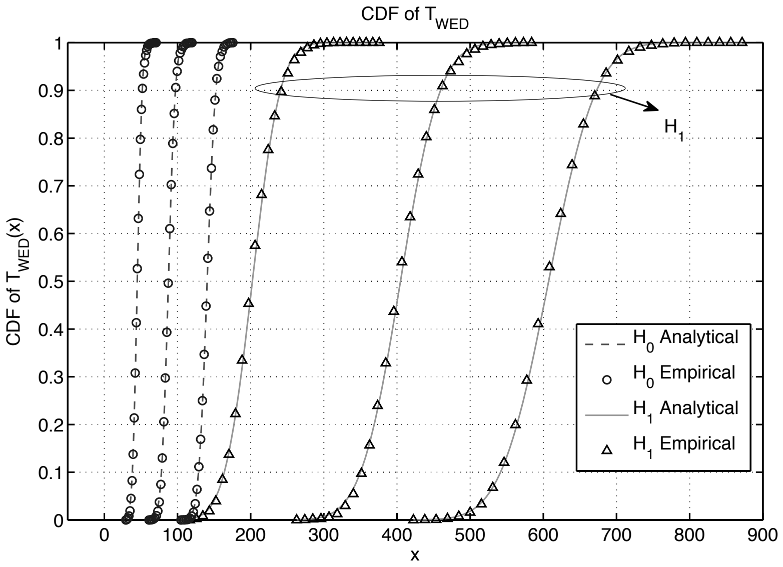

Figure 5.5.: Comparison of empirical and analytical cumulative density function of T_{WED} for $N \in \{40, 80, 120\}$ samples under $\mathcal{H}_0$ and $\mathcal{H}_1$.

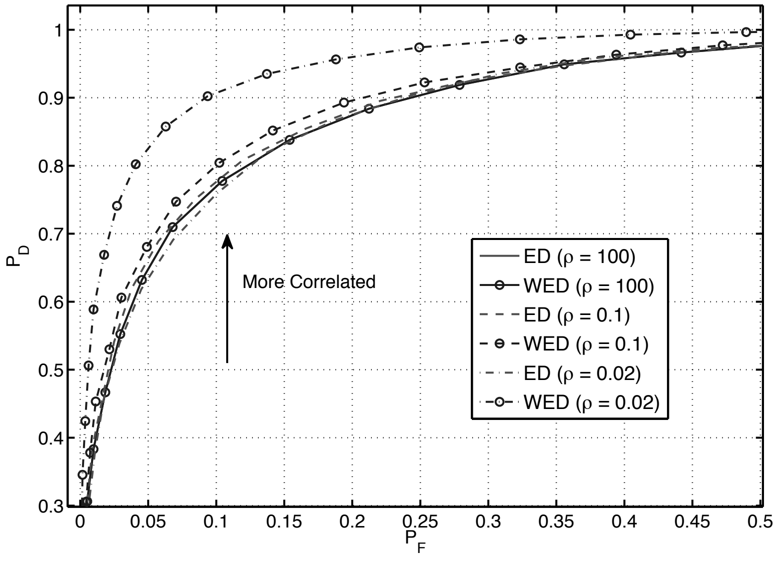

Figure 5.6.: ROC curve for the conventional energy detector (ED) and the weighted energy detector (WED) for $N = 2000, K = 8, \rho \in \{100, .1, .02, \}$.

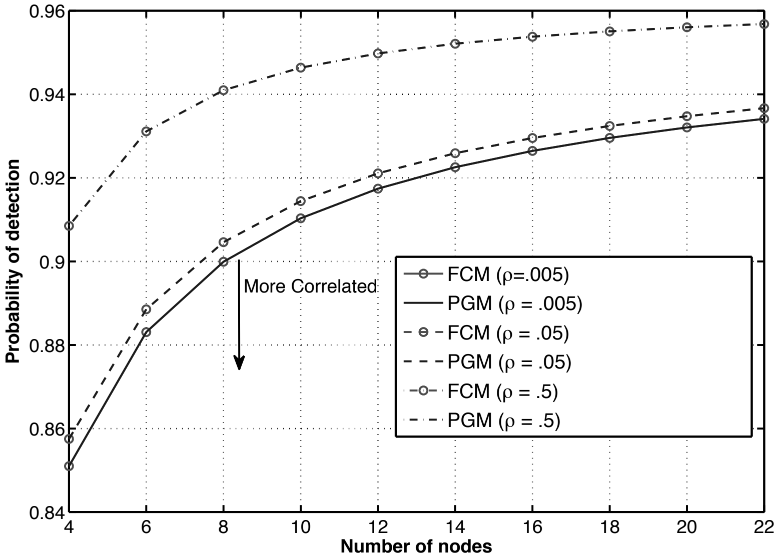

Figure 5.7.: Performance of PGM and FCM as a function of K with different correlation constants ρ when $\alpha > Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^4} - 1))$.

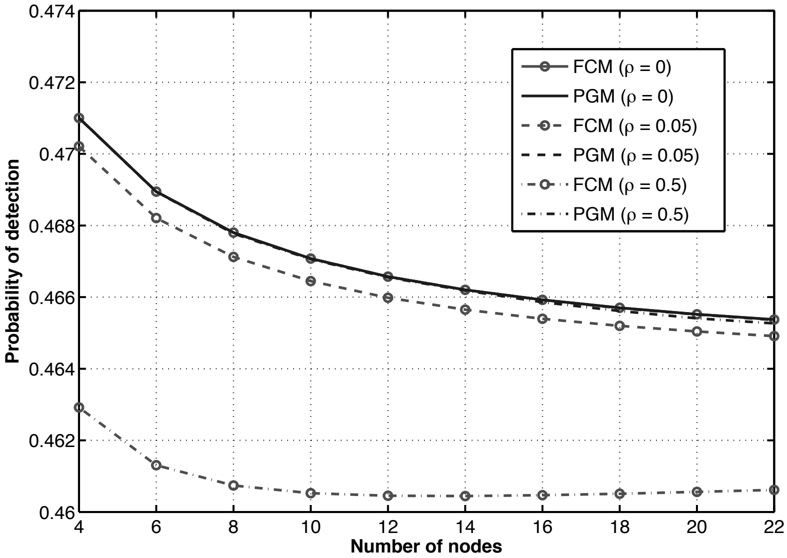

Figure 5.8.: Performance of PGM and FCM as a function of K with different correlation constants ρ when $\alpha < Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^4} - 1))$.

we performed an experiment illustrated in Figure 5.8.

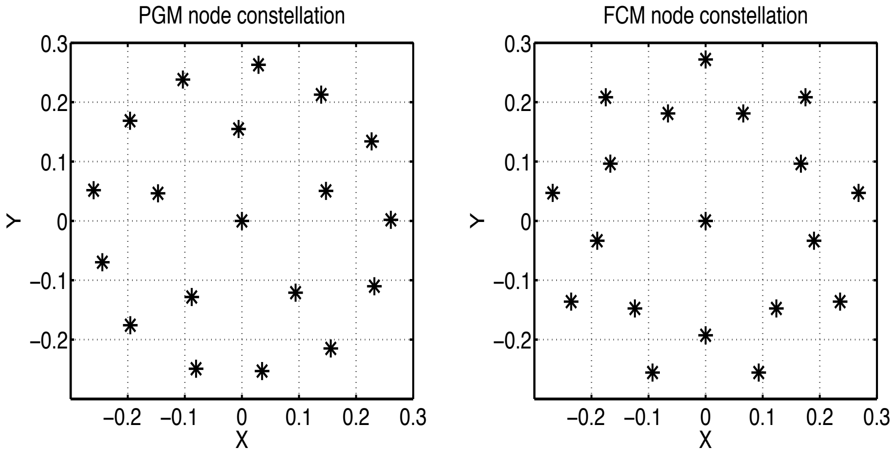

Figure 5.9.: Node constellation results proposed by PGM and FCM for $K = 19$ when $\alpha > Q(\sqrt{NK}(w_1/\sigma_\nu - 1))$.

5.6. Conclusion

We have modelled the joint problem of spectrum sensing and jammer detection using multiple hypothesis testing. For the case of spatially correlated observations we have derived the optimal test, which results in a weighted energy detector. Furthermore, The probability of detection of the optimal detector is derived in two exact versions for a given probability of false alarm. One of the exact versions is based on a series, where in practice we need to truncate. The other exact form is in terms of Meijer's G function. We further developed an asymptotic expression of the probability of detection, which results in a simpler formula. Numerical simulations showed that the asymptotic results are very accurate for a moderate size of samples.

For the case of unknown variables, we used ML to estimate the unknown variables. Using GLRT, we study the practical cases of, namely: unknown noise power and unknown jammer power. By numerical simulations, we have illustrated the accuracy of ML in estimating the unknown variables.

We have introduced the trade-off between power consumption and detection probability as a practical scenario. An optimization problem has been proposed to find the optimal locations of the nodes to achieve the maximum probability of detection while restricted by the transmission power. The optimization problem is non-convex; thus, we proposed two heuristic methods to solve it. The first method is based on the projected gradient algorithm, whereas the second one is based on the KKT conditions. Both of the methods provide the same probability of detection for a given power limitation even though with different node locations.

Future work

Node placement, which is the relaxed version of node selection problem, is of particular significance in the context of combating jamming in a wireless sensor network with energy restriction. To combat jamming, which mainly drains the nodes batteries, we can consider a node selection problem that takes into account the jamming signal strength at each location as well.

5.7. Appendix

5.7.1. Proof of Proposition 7

Based on our assumptions in (5.2) we have,

$$\begin{aligned} P(\mathbf{x}|H_l) = \prod_{n=1}^{N} & \frac{1}{\pi^K \det[w_l \mathbf{\Sigma} + \sigma_\nu^2 \mathbf{I}]} \times \\ & \exp(-\mathbf{x}(n)^H (w_l \mathbf{\Sigma} + \sigma_\nu^2 \mathbf{I})^{-1} \mathbf{x}(n)). \end{aligned} \tag{5.26}$$

We substitute (5.26) into (5.4) and take the log to achieve,

$$\begin{aligned} & N(\ln\det[w_j \mathbf{\Sigma} + \sigma_\nu^2 \mathbf{I}]) - N(\ln\det[w_l \mathbf{\Sigma} + \sigma_\nu^2 \mathbf{I}]) \\ & + \sum_{n=1}^{N} \mathbf{x}^H(n) \mathbf{S}_{l,j} \mathbf{x}(n) + \ln \frac{\Pr(H_l)}{\Pr(H_j)} \gtrless 0, \end{aligned} \tag{5.27}$$

where, $\mathbf{S}_{l,j}$ is Hermitian and defined in (5.6). We proceed by keeping the data dependent parts in the LHS as,

$$\sum_{n=1}^{N} \mathbf{x}(n)^H \mathbf{S}_{l,j} \mathbf{x}(n) \underset{H_j}{\overset{H_l}{\gtrless}} \vartheta_{l,j}, \tag{5.28}$$

in which, $\vartheta_{l,j}$ is the threshold that is defined in the proposition. We change the variable as $\mathbf{z} := \mathbf{V}^H \mathbf{x}$, where $\mathbf{V}$ is a unitary matrix given by the eigenvalue decomposition of $\mathbf{S}_{l,j} = \mathbf{V}\mathrm{Diag}[\lambda_1, \ldots, \lambda_K]\mathbf{V}^H$. Therefore, the test statistic becomes the weighted energy detector and is given as,

$$\sum_{n=1}^{N} \mathbf{z}(n)^H \mathrm{Diag}[\lambda_1, \ldots, \lambda_K] \mathbf{z}(n) \underset{H_j}{\overset{H_l}{\gtrless}} \vartheta_{l,j}, \tag{5.29}$$

by which we can infer (5.5) directly. ■

5.7.2. Proof of Proposition 8

We first find the distribution of $t_i = \sum_{n=1}^{N} \lambda_i |z_i(n)|^2$, then we use it to find $T_{l,j}$.

Based on our assumption across time the observation symbols are independent. Therefore, assuming $x_i(n) \in H_l$, $\forall i$ then the distribution of $z_i(n)$ is $\mathcal{CN}(0, \theta_i)$. It follows then, the $\lambda_i|z_i(n)|^2$ is distributed according to $\Gamma(1, \frac{1}{2\theta_i\lambda_i})$. Thus, based on the rule of summation of i.i.d. gamma random variables, t_i's distribution is given as $\Gamma(N, \frac{1}{2\theta_i\lambda_i})$.

In order to find the distribution of $T_{l,j} = \sum_{i=1}^{K} t_i$ we need to have the distribution of the sum of independent but non-identical gamma random variables. This in turn, is given by the work in [Mos85]. Therefore, we apply the results in [Mos85] with our parameters and obtain (5.7). ■

5.7.3. Proof of Corollary 1

Proceeding along similar lines as in Appendix 5.7.2, we find the probability of $T = \sum_{i=1}^{K} t_i$, where $t_i \sim \Gamma(N, \frac{1}{2\theta_i\lambda_i})$. To this end, we directly apply Corollary 3 in [AYAK12]. ■

5.7.4. Proof of Proposition 9

Since x_is are Gaussian the distribution of $\mathbf{z}$ is Gaussian as well, although with different correlation coefficient matrix. The correlation matrix of $\mathbf{z}$, conditioned on H_l is given by

$$\mathbf{\Sigma}_z = \mathrm{diag}[\theta_1, \ldots, \theta_K],$$

where θ_is are defined as $w_l\mathbf{\Sigma} + \sigma_\nu^2\mathbf{I} := \mathbf{V}\mathrm{diag}[\theta_1, \ldots, \theta_K]\mathbf{V}^H$. The test statistic is weighted sum of independent non-identical gamma random variables. According to [AAK01] the moment generating function of $T_n := \sum_{i=1}^{K} \lambda_i|z_i(n)|^2$ is given by

$$M_z(s) = \prod_{i=1}^{K}(1 - s\theta_i\lambda_i)^{-1}. \tag{5.30}$$

From (5.30) one can achieve the mean and the variance of T_n as,

$$\mathrm{E}[T_n] = \sum_{i=1}^{K} \theta_i\lambda_i, \qquad \mathrm{Var}[T_n] = \sum_{i=1}^{K}(\theta_i\lambda_i)^2. \tag{5.31}$$

Based on the assumption that temporal observations, i.e. T_n, are i.i.d. and N is sufficiently large, we apply the central limit theorem to approximate the distribution of T as,

$$T \sim \mathbb{N}\left(N\sum_{i=1}^{K} \theta_i\lambda_i, N\sum_{i=1}^{K}(\theta_i\lambda_i)^2\right).$$

■

5.7.5. Proof of Proposition 10

The ML problem becomes,

$$\begin{aligned}\hat{w}_2 =& \arg\max_{w_2 \geq 0}\ f(w_2)\\ f(w_2) :=& -NK\ln\pi - N\sum_{i=1}^{K}\ln\left(w_2\tau_i + \sigma_\nu^2\right)\\ &-\sum_{n=1}^{N}\mathbf{x}^H(n)\mathbf{V}(w_2\mathrm{Diag}[\tau_1,\ldots,\tau_K] + \sigma_\nu^2\mathbf{I})^{-1}\mathbf{V}^H\mathbf{x}(n).\end{aligned}$$

Let us define $f'(w_2)$ and $f''(w_2)$ to be the first and second derivatives of $f(w_2)$, respectively. We denote the unique solution of $f''(w_2) = 0$ as w_2^*. A simple analysis reveals:

$$\begin{cases} f''(w_2) < 0 & w_2 \in [0, w_2^*] \\ f''(w_2) > 0 & w_2 \in (w_2^*, \infty] \end{cases} \tag{5.32}$$

Thus, $f(w_2)$ is concave for $w_2 \in [0, w_2^*]$ and then it becomes convex for $w_2 \in (w_2^*, \infty]$. Therefore, we have a unique local maximum on $w_2 \in [0, w_2^*]$ which we indicate it with $\hat{w}_2$. In the following we obtain this local maximum by solving,

$$\begin{aligned} f'(w_2) =& N\sum_{i=1}^{K}\tau_i(w_2\tau_i + \sigma_\nu^2)^{-1}\\ &-\sum_{i=1}^{K}\sum_{n=1}^{N}\tau_i(w_2\tau_i + \sigma_\nu^2)^{-2}|z_i(n)|^2 =\\ &\sum_{i=1}^{K}\frac{N\tau_i(w_2\tau_i + \sigma_\nu^2) - \tau_i\sum_{n=1}^{N}|z_i(n)|^2}{(w_2\tau_i + \sigma_\nu^2)^2} = 0 \\ &= NK\sigma_\nu^2 + Nw_2\sum_{i=1}^{K}\tau_i - \sum_{i=1}^{K}\sum_{n=1}^{N}|z_i(n)|^2 = 0.\end{aligned} \tag{5.33}$$

The equation in (5.33) yields $\hat{\hat{w}}_2 = \infty$ and

$$\hat{w}_2 = \frac{\sum_{i=1}^{K}\sum_{n=1}^{N}|z_i(n)|^2 - NK\sigma_\nu^2}{N\sum_{i=1}^{K}\tau_i}. \tag{5.34}$$

($\hat{\hat{w}}_2 = \infty$) results in a minimum for f, thus we only have one maximum which is (5.34). In the following we prove that it is the global maximum.

For $w_2 \in [\hat{w}_2, \infty]$ we note that $f'(w_2) \leq 0$, hence, $f(w_2)$ is non-increasing. This implies that $\hat{w}_2$ is the global maximum. In the trivial case of $(\sum_{i=1}^{K}\sum_{n=1}^{N}|z_i(n)|^2 \leq NK\sigma_\nu^2)$ due to the non-negativity constraint, we select $\hat{w}_2 = 0$. Therefore we conclude (5.12). ■

5.7.6. Proof of Proposition 12

The two problems are equivalent when every solution of Problem I is a solution of Problem II and vice versa. Substituting (5.18) and (5.19) in (5.15) yields,

$$\begin{aligned} \max_{\mathbf{Y}} \quad & Q\left(\frac{Q^{-1}(\alpha)\sqrt{NK\sigma_\nu^4} + NK\sigma_\nu^2 - N\sum_{i=1}^{K}\lambda_i\theta_i}{\sqrt{N\sum_{i=1}^{K}\lambda_i^2\theta_i^2}}\right) \\ \text{s.t.} \quad & \sum_i d_{i,1}^2\sigma_\nu^2\eta\kappa \leq \omega \\ & d_{i,1}^2\sigma_\nu^2\kappa\eta \leq \hat{\omega} \quad \forall i. \end{aligned} \tag{5.35}$$

Since $\sigma_\nu^2\eta\kappa > 0$, the constraints of both of the problems are obviously equivalent, thus we have

$$\begin{aligned} \max_{\mathbf{Y}} \quad & Q\left(\frac{Q^{-1}(\alpha)\sqrt{NK\sigma_\nu^4} + NK\sigma_\nu^2 - N\sum_{i=1}^{K}\lambda_i\theta_i}{\sqrt{N\sum_{i=1}^{K}\lambda_i^2\theta_i^2}}\right) \\ \text{s.t.} \quad & \sum_i d_{i,1}^2 \leq \omega c_2 \\ & d_{1,i}^2 \leq \hat{\omega} c_2 \quad \forall i, \end{aligned}$$

where $c_2 := \left(\sigma_\nu^2\eta\kappa\right)^{-1}$. We need to prove that for any point in (5.35) there is a point in 12. Because $Q(x)$, $x \in \mathbb{R}$ is a strictly decreasing function, we can minimize the argument of $Q(\cdot)$ in (5.35) instead. If we rule out the dull condition of $Q^{-1}(\alpha)\sqrt{NK\sigma_\nu^4} + NK\sigma_\nu^2 = N\sum_{i=1}^{K}\lambda_i\theta_i$ we can write (5.35) as a maximization of the inverse of its argument:

$$\begin{aligned} \max_{\mathbf{Y}} \quad & \frac{\sqrt{N\sum_{i=1}^{K}(\lambda_i\theta_i)^2}}{Q^{-1}(\alpha)\sqrt{NK\sigma_\nu^4} + NK\sigma_\nu^2 - N\sum_{i=1}^{K}\lambda_i\theta_i} \\ \text{s.t.} \quad & \sum_i d_{i,1}^2 \leq \omega c_2 \\ & d_{1,i}^2 \leq \hat{\omega} c_2 \quad \forall i, \end{aligned} \tag{5.36}$$

The following lemma is needed to continue with the proof.

Lemma 15. *For the correlation matrix* $\mathbf{\Sigma} := \mathbf{U}\mathrm{Diag}[\gamma_1, \ldots, \gamma_K]\mathbf{U}^H$ *defined in (5.1) we* *have*

$$\sum_{i=1}^{K} \gamma_i^2 = \mathrm{trace}[\mathbf{\Sigma}^T \mathbf{\Sigma}] = \|\mathbf{\Sigma}\|_F^2 = \sum_{i,j} [\mathbf{\Sigma}]_{i,j}^2 = \sum_{i,j} e^{-2\rho d_{i,j}}, \tag{5.37}$$

and further,

$$\sum_{i=1}^{K} \gamma_i = \mathrm{trace}[\mathbf{\Sigma}] = K. \tag{5.38}$$

Substituting (5.38) in the denominator in (5.36) yields:

$$\begin{aligned} \max_{\mathbf{Y}} \quad & \frac{\sqrt{N \sum_{i=1}^{K} (w_1 \gamma_i / \sigma_\nu^2)^2}}{(Q^{-1}(\alpha)\sqrt{NK\sigma_\nu^4} + NK\sigma_\nu^2 - w_1 NK/\sigma_\nu^2)} \\ \text{s.t.} \quad & \sum_i d_{i,1}^2 \leq \omega c_2 \\ & d_{1,i}^2 \leq \hat{\omega} c_2 \quad \forall i, \\ \max_{\mathbf{Y}} \quad & \Big(\frac{w_1 \sqrt{N}}{Q^{-1}(\alpha)\sqrt{NK}\sigma_\nu^4 + NK\sigma_\nu^4 - w_1 NK}\Big) \sqrt{\sum_{i=1}^{K} \gamma_i^2} \\ \text{s.t.} \quad & \sum_i d_{i,1}^2 \leq \omega c_2 \\ & d_{1,i}^2 \leq \hat{\omega} c_2 \quad \forall i, \end{aligned} \tag{5.39}$$

Note that $Q^{-1}(\alpha)\sqrt{NK}\sigma_\nu^4 + NK\sigma_\nu^4 - w_1 NK < 0$ gives $\alpha > Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^4} - 1))$. We square the objective function in (5.39) and substitute (5.37) to obtain:

$$\begin{aligned} \max_{\mathbf{Y}} \quad & - c_1 G(\mathbf{Y}) \\ \text{s.t.} \quad & \sum_i d_{i,1}^2 \leq \omega c_2 \\ & d_{1,i}^2 \leq \hat{\omega} c_2 \quad \forall i, \end{aligned} \tag{5.40}$$

where, $c_1 := \big(\frac{w_1\sqrt{N}}{Q^{-1}(\alpha)\sqrt{NK}\sigma_\nu^4 + NK\sigma_\nu^4 - w_1 NK}\big)^2$ and

$$G(\mathbf{Y}) := \begin{cases} \sum_{i,j} e^{-2\rho d_{i,j}} & \alpha > Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^4} - 1)) \\ -\sum_{i,j} e^{-2\rho d_{i,j}} & \alpha < Q(\sqrt{NK}(\frac{w_1}{\sigma_\nu^4} - 1)) \end{cases} \tag{5.41}$$

This concludes the proof.

6. Decentralized computation algorithms for wireless networks

We devote this chapter to developing decentralized computation techniques for the algorithms that deal with active security threats. In particular, this chapter studies the decentralization of the eigenvalue-based detection algorithms. From the implementation point of view, the core problem of the eigenvalue-based detection algorithms is computing the eigenvalues of the covariance matrix of some observation data. We address the general problem concisely as follows. We are given multiple communication nodes with some processing power, which are referred to as the agents. Each agent has a row (column) of an observation matrix, which contains the measurements regarding some malicious activity, e.g. a jammer. We are interested in finding the eigenvalues of the covariance matrix of the observation data without the need to rebuild the whole matrix at every node. The applications of this problem include, but not limited to, channel probing for power admission control [MSCE11], spectrum sensing [PS15a], and jamming detection.

In order to compute the eigenvalues, a straightforward approach is to collect all the data and reconstruct the covariance matrix of the observations at a fusion center followed by a centralized computation of the eigenvalues. In most cases, however, this is not a desirable method according to the following arguments. Relying on a single fusion center is not robust because of single node failure. Furthermore, a centralized architecture is more susceptible to the Byzantine attacks. On the other hand, in a decentralized structure a node failure or a Byzantine attack does not disrupt the whole network's functionality. Moreover, most of the decentralized algorithms believed to be scalable to the size of the network easily, whereas a centralized regime faces a lot of problems when the network size grows larger. Therefore, there is a need for a decentralized computation-communication scheme that is able to estimate the eigenvalues without reconstructing the covariance matrix at each node.

In this chapter, we first develop an algorithm for estimating all of the eigenvalues of a given covariance matrix in the decentralized fashion by applying the generic average consensus algorithm. Then motivated by an application, we propose applying a special kind of consensus algorithm, which uses computation over multiple access channels (Co-

MAC) technique. We further apply the results to the channel probing for power admission control.

6.1. Background and contribution

6.1.1. Related work

Eigenvalue-based detection methods have been studied extensively in [NPG11,ZMcLM09]. Nadler at el. have considered two eigenvalue-based detection method, namely, the Roy's largest root test and generalized likelihood ratio test in [NPG11]. The authors in [ZMcLM09] analyse the test of largest eigenvalue divided by smallest eigenvalue of the covariance matrix.

Decentralized algorithms for eigenvalue computation has attracted many researchers' attention recently, for instance in [PS15a,PS15b,LSM11]. Penna and Stanczak in [PS15a] proposed a decentralized largest eigenvalue computation based on the *power method.* To compute the other eigenvalues they propose using the *Lanczos* method. The building block for the decentralization algorithm is the generic *average consensus* method, which inspired the results of this chapter partly. Later in [PS15b] they provide some convergence analysis of the two algorithms. The analysis suggests that the *power method* is robust against numerical errors, whereas the *Lanczos* method is prone to error, and gives some 'spurious' eigenvalues, which needs extra post-processing.

The admissibility of the new links in the context of decentralized channel admission control for cognitive radio is studied by [BCM95,XSC01,ZC01]. In [BCM95], admissibility of a new link is verified by observing the variations in the SINR of the new link when its respective transmitter changes the power in a controlled fashion. The main limitation of this approach is that it can experience convergence problems if the variations of the SINR of the new link are small. This problem has been addressed in [XSC01], which uses a technique where a new user probes the channel by transmitting at a constant power level. Then, to detect admissibility (i.e., the existence of a feasible power assignment), the new user measures its received interference after the power admission control algorithm of the existing users converges. A more general approach that allows multiple users to join the system at the same time has been proposed in [ZC01]. However, this approach requires a complex coordination scheme among users trying to join the system. To date, simple channel probing schemes, where multiple users are allowed to join the system at the same time (as required by many cognitive radio systems), remain an open problem.

6.1.2. Contributions

This chapter exploits the *power method* (PM) [GL96] to find other eigenvalues of the covariance matrix. Furthermore, we take advantage of the scalability and robustness of average consensus to implement our algorithm in a decentralized fashion. Some analyses on the convergence of the algorithm are provided; however, due to the complexity of the algorithm a complete analytical proof of convergence is missing. The results are tailored for two particular applications:

- A special scenario which induces a block diagonal structure on the covariance matrix is considered. We exploit the eigenvalue properties of a block diagonal matrix to further accelerate the convergence speed. Furthermore, we propose to utilize a particular CoMAC protocol that enables the computation of certain functions over the air and reaches consensus in only two iterations [ZGSY12, GS13, NG07, LMS15].
- The developed eigenvalue estimation technique is applied to the channel probing problem. The proposed method verifies admissibility of secondary users in cognitive radio systems (the new links to be added to the system) with a scheme that estimates whether the spectral radius of the channel matrix is above or below a given threshold (which is enough to detect admissibility). However, this estimation is done indirectly; secondary users produces two sequences that converge to the spectral radius of a matrix that has a close connection to the channel gain matrix.

6.2. Problem statement and preliminaries

We consider K sensors forming a connected graph, in which each edge represents a communication link. All links use a common wireless channel. The sensors sense some frequency spectrum to detect the presence of primary users in a given sensing time, T. Each sensor has a measurement vector $\mathbf{y} \in \mathbb{C}^N$ over N realizations of the wireless channel. The observation matrix is defined to be

$$\mathbf{Y} \triangleq \Big[\mathbf{y}(1), \ldots, \mathbf{y}(N)\Big] = \begin{bmatrix} \hat{\mathbf{y}}(1)^T \\ \vdots \\ \hat{\mathbf{y}}(K)^T \end{bmatrix} \in \mathbb{C}^{K \times N}, \tag{6.1}$$

where $\mathbf{y}$ and $\hat{\mathbf{y}}^T$ are used to denote columns and rows of $\mathbf{Y}$, respectively. The sample covariance matrix is given by

$$\mathbf{R} \triangleq \frac{1}{N}\mathbf{Y}\mathbf{Y}^H, \tag{6.2}$$

which is positive semi-definite by construction. We intend to compute the eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_K \geq 0$ of $\mathbf{R} \in \mathbb{C}^{K \times K}$, in which we further assume that all eigenvalues of $\mathbf{R}$ are distinct. Using PM, we can compute λ_1 and its corresponding eigenvector $\mathbf{v}_1$. However, some eigenvalue based inference methods require the knowledge of all eigenvalues of $\mathbf{R}$ [MPS13, MSZ15]. So our objective is to show how the PM can be used to provide estimates of all the eigenvalues. Furthermore, we show how to implement the proposed scheme in a distributed manner, which boils down to transforming the computations into local and global steps. We intend to benefit from the concept of consensus algorithms to deal with the global computations. In the following, we provide a brief introduction to the average consensus algorithm to be used later.

6.2.1. Average consensus algorithm

An average consensus algorithm can be defined on a network of nodes that form an undirected connected graph. It is a distributed algorithm that takes scalars or vectors as input from each node and delivers the overall average values of the input available at every node. Various methods have been proposed [ZGSY12, NG07, OSFM07, LSMGC13]. In this work, however, we are not interested in a particular average consensus algorithm. Instead, we assume a generic algorithm that achieves consensus without errors. An exception is Section 6.4, where we exploit physical properties of the wireless channel to perform averaging.

In this chapter, the average consensus algorithm is assumed to be an operator $\mathrm{CA}[\cdot]\ \mathbb{C}^K \to \mathbb{C}^K$. The input is the vector with entries representing different measurements at each sensor and the output is a vector with all entries equal to the average of the input vector. This can be expressed more precisely as

$$\mathbf{y} = \mathrm{CA}[\mathbf{x}] = \frac{1}{K}\left(\mathbf{1}^T\mathbf{x}\right)\mathbf{1}. \tag{6.3}$$

One can generalized this notation to the case where each sensor's input is a vector of size m. Therefore the input is a matrix of $\mathbb{C}^{K \times m}$. This is presented as

$$\mathbf{Y} = \mathrm{CA}_m[\mathbf{X}] = \frac{1}{K}\left(\mathbf{1}\mathbf{1}^T\right)\mathbf{X}. \tag{6.4}$$

6.3. Techniques for eigenvalue computation

In this section we introduce a generalization of the *power method* that is based on a matrix deflation technique (see e.g. [Saa11, Ch.4]). The generalized power method is capable of computing all eigenvalues of a given Hermitian matrix with distinct eigenvalues $\lambda_1 >$

$\lambda_2 > \cdots > \lambda_K \geq 0$. The method is shown to be amenable to distributed implementation provided that a network-wide average consensus can be achieved efficiently.

6.3.1. Generalized power method (GPM)

Motivated by the appealing characteristics of the PM such as robustness [GL96], we intend to develop an iterative algorithm based on the PM that is able to compute all eigenvalues of some eigenvalue matrix.

We are interested in estimating the eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_K \geq 0$ of a hermitian, positive semi-definite matrix $\mathbf{R} \in \mathbb{C}^{K \times K}$. Applying PM to $\mathbf{R}$ yields a robust estimate of λ_1 and the corresponding eigenvector $\mathbf{v}_1$. Now suppose that we have another matrix $\mathbf{R}_2$ with the eigenvalues $\lambda_2 > \cdots > \lambda_K \geq \lambda_{K+1} = 0$ and the corresponding eigenvectors identical to those of $\mathbf{R}$ except for λ_1 and $\mathbf{v}_1$. Then, we can again use the PM to obtain λ_2. In the following, we explain further details on the construction of $\mathbf{R}_2$ from $\mathbf{R}$.

An application of the PM to $\mathbf{R}$ provides the eigenvector $\mathbf{v}_1$ corresponding to the largest eigenvalue λ_1. If we subtract the rank-1 matrix $\lambda_1 \mathbf{v}_1 \mathbf{v}_1^H$ from $\mathbf{R}$ we can achieve a second matrix $\mathbf{R}_2$,

$$\mathbf{R}_2 = \mathbf{R} - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^H,$$

which has the desired features with $\operatorname{rank}(\mathbf{R}_2) = \operatorname{rank}(\mathbf{R}) - 1$. This procedure can be repeated to obtain the rest of the eigenvalues. Algorithm 3 presents this method in more details.

Algorithm 3 Generalized Power method (GPM)

```
Input: Initial vector $\mathbf{q}_1^0$ and hermitian PSD matrix $\mathbf{R}$
Output: Eigenvalues $\lambda_k\ \forall k$ and eigenvectors $\mathbf{v}_k\ k \in \{1..K\}$ of $\mathbf{R}$
$\mathbf{R}_1 = \mathbf{R},\ k = 0$
repeat
    $k \leftarrow k + 1$
    repeat
        $\mathbf{q}_k^{i+1} \leftarrow \mathbf{R}_k \mathbf{q}_k^i$
        $\lambda_k^i \leftarrow \frac{\mathbf{q}_k^{i^H} \mathbf{R}_k \mathbf{q}_k^i}{\|\mathbf{q}_k^i\|^2}$
        $i \leftarrow i + 1$
    until $|\lambda_k^i - \lambda_k^{i-1}| \leq \epsilon$
    $\mathbf{v}_k \leftarrow \mathbf{q}_k^i,\ \lambda_k \leftarrow \lambda_k^i$
    $\mathbf{R}_{k+1} \leftarrow \mathbf{R}_k - \lambda_k \mathbf{v}_k \mathbf{v}_k^H$
    $\mathbf{q}_{k+1} \leftarrow \mathbf{q}_k - \mathbf{v}_k(\mathbf{q}_k^H \mathbf{v}_k)$
    $\mathbf{q}_{k+1}^0 \leftarrow \mathbf{q}_{k+1} / \|\mathbf{q}_{k+1}\|$
until $k = K$
```

In the Appendix 6.7.1, a note on the proof of convergence of GPM is presented.

To see the convergence of the GPM Algorithm, we observe the descent of absolute error of each 4 estimated eigenvalues in Figure 6.1. We let each inner loop, i.e. PM, run for 50 iterations before moving to the next eigenvalue.

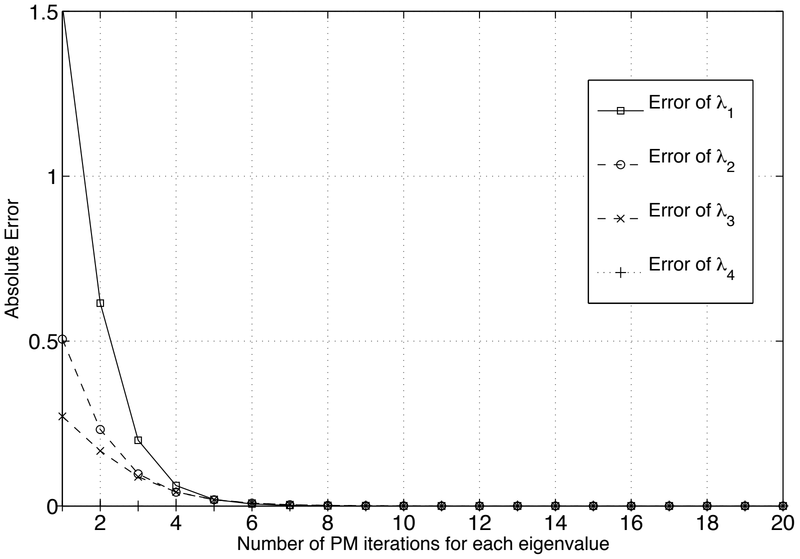

Figure 6.1.: Convergence of GPM in terms of absolute error of estimation for first four eigenvalues, $\lambda_1, \lambda_2, \lambda_3$, and λ_4 with $K = 16$.

6.3.2. Decentralized generalized power method (DGPM)

We can break down the eigenvalue computation into the steps which are possible to perform locally and into the global steps, which require data exchange between sensors. For the global steps we use a generic consensus algorithm. Building upon the results in [PS12], we use the CA operator to perform GPM in a decentralized manner. We need to evaluate the steps 5,6,10,11, and 12 using CA in Algorithm 3, which are described respectively in the following. To compute step 5 we recall that in (6.1), $\hat{\mathbf{y}}(k),\ k \in \{1, \ldots, K\}$ is available at each sensor. Thus,

$$\mathbf{R}\mathbf{q} = \frac{1}{N}\mathbf{Y}\mathbf{Y}^H\mathbf{q} = \frac{1}{N}\mathbf{Y}\begin{bmatrix}\mathbf{y}(1)^H\mathbf{q}\\ \vdots \\ \mathbf{y}(N)^H\mathbf{q}\end{bmatrix}, \tag{6.5}$$

where we dropped the indices for simplicity. We can perform $\mathbf{y}(n)^H\mathbf{q},\ n \in \{1, \ldots, N\}$ using the following command:

$$K\mathrm{CA}_1\Big[\mathrm{diag}\big[\mathrm{Diag}[\mathbf{y}(n)]\mathrm{Diag}[\mathbf{q}]\big]\Big]. \tag{6.6}$$

To compute $\mathbf{Y}^H\mathbf{q}$ we need to repeat (6.6) for every $n \in \{1, \ldots, N\}$, which can be simplified using the vector notation of consensus as

$$\mathbf{Y}^H\mathbf{q} = K\mathrm{CA}_N\left[\mathbf{Y}^H\mathrm{Diag}[\mathbf{q}]\right].$$

Now the vector $\mathbf{Y}^H\mathbf{q}$ is available at every node and since the kth row of $\mathbf{Y}$ is the local measurement of the sensor k, the rest of the operations can be done locally. This yields

$$\mathbf{Rq} = \frac{K}{N}\mathrm{diag}\Big[\mathbf{Y}\mathrm{CA}_N\big[\mathbf{Y}^H\mathrm{Diag}[\mathbf{q}]\big]\Big]. \tag{6.7}$$

At this point we can write each step in terms of CA_1 or CA_N. Step 10 in Algorithm 3 requires reconstruction of the matrix at each step, which is redundant. Therefore, we combine it with step 5. To obtain step 6 we reuse the available $\mathbf{q}_k^i$ and $\mathbf{q}_k^{i+1}$ from the previous step and then apply the consensus operator as,

$$\lambda_k^i = \frac{\mathrm{CA}_1\Big[\mathrm{diag}\big[\mathrm{Diag}[\mathbf{q}_k^i]\mathrm{Diag}[\mathbf{q}_k^{i+1}]\big]\Big]}{\mathrm{CA}_1\Big[\mathrm{diag}\big[\mathrm{Diag}[\mathbf{q}_k^i]\mathrm{Diag}[\mathbf{q}_k^i]\big]\Big]}. \tag{6.8}$$

As for steps 11 and 12 since we have $\mathbf{v}_k$ locally available at every node after power iteration, we can calculate them as in (6.7) and (6.8).

Reposing on these discussions the decentralized generalized power method is given in Algorithm 4. In each step of using CA we loose a large amount of network resources until eventually the nodes reach a consensus.

If we let the consensus algorithm converge up to a machine precision error, then DGPM converges exactly like GPM. As far as the numerical analysis of DGPM is concerned, we have considered the consensus algorithm proposed by [AYSS09]. However, for the sake of interest, we truncate the consensus algorithm in an average error of 10^{-4} to plot its convergence speed. Figure 6.2 illustrates this experiment. One can observe the effect of small error in each round of consensus algorithm on the convergence of decentralized GPM. When the value of error is more than a certain limit, the convergence for some smaller eigenvalues might not occur at all.

Algorithm 4 Decentralized Generalized Power method (DGPM)

```
Input: Initial vector $\mathbf{q}_1^0$ and $\mathbf{Y} \in \mathbb{C}^{N\times K}$
Output: Eigenvalues $\lambda_k\ \forall k$ and eigenvectors $\mathbf{v}_k\ k \in \{1..K\}$ of $\mathbf{R} = \frac{1}{N}\mathbf{Y}\mathbf{Y}^H$
$\mathbf{R}_1 = \mathbf{R},\ k = 0,\ i = 0$
repeat
    $k \leftarrow k + 1$
    repeat
        The $k$th eigenvector estimated as:
            $$\mathbf{q}_k^{i+1} \leftarrow \frac{K}{N}\mathrm{diag}\Big[\mathbf{Y}\mathrm{CA}_N\big[\mathbf{Y}^H\mathrm{Diag}[\mathbf{q}_k^i]\big]\Big] - \sum_{j=1}^{k-1}\lambda_j K\mathrm{CA}_1\Big[\mathrm{diag}\big[\mathrm{Diag}[\mathbf{v}_j]\mathrm{Diag}[\mathbf{q}_k^i]\big]\Big]\mathbf{v}_j$$
        The $k$th eigenvalue estimation after $i$ iterations:
            $$\lambda_k^i \leftarrow \frac{\mathrm{CA}_1\Big[\mathrm{diag}\big[\mathrm{Diag}[\mathbf{q}_k^i]\mathrm{Diag}[\mathbf{q}_k^{i+1}]\big]\Big]}{\mathrm{CA}_1\Big[\mathrm{diag}\big[\mathrm{Diag}[\mathbf{q}_k^i]\mathrm{Diag}[\mathbf{q}_k^i]\big]\Big]}$$
        $i \leftarrow i + 1$
    until $|\lambda_k^i - \lambda_k^{i-1}| \le \epsilon$
    $\mathbf{v}_k \leftarrow \mathbf{q}_k^i,\ \lambda_k \leftarrow \lambda_k^i$
    The Gram-Schmidt update
        $$\mathbf{q}_{k+1}^0 \leftarrow \mathbf{q}_k^0 - K\mathrm{CA}_1\Big[\mathrm{diag}\big[\mathrm{Diag}[\mathbf{q}_k^0]\mathrm{Diag}[\mathbf{v}_k]\big]\Big]\mathbf{v}_k$$
    Normalization
        $$\mathbf{q}_{k+1}^0 \leftarrow \mathbf{q}_{k+1}^0/\mathrm{CA}_1\Big[\mathrm{diag}\big[\mathrm{Diag}[\mathbf{q}_k^0]\mathrm{Diag}[\mathbf{q}_k^0]\big]\Big]$$
until $k = K$
```

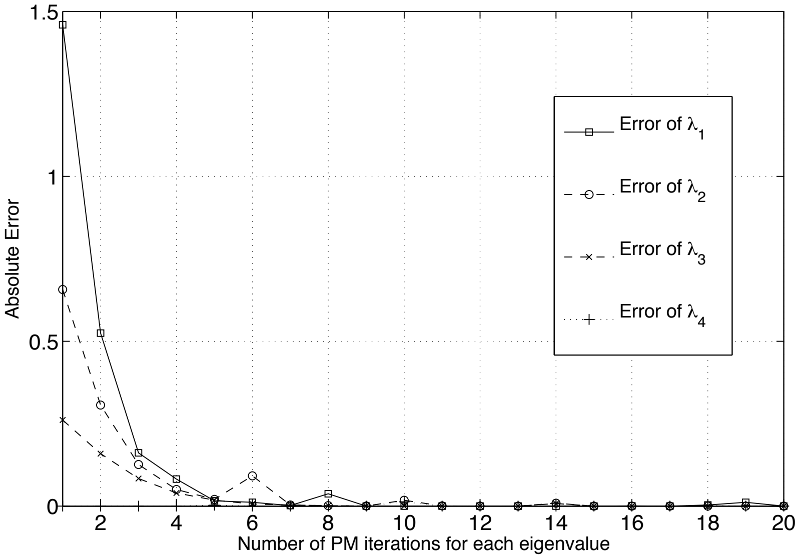

Figure 6.2.: Convergence of DGPM in terms of absolute error of estimation for first four eigenvalues, $\lambda_1, \lambda_2, \lambda_3$, and λ_4 with $K = 16$.

6.3.3. Convergence analysis

Unlike PM, which has straightforward bounds on the estimation error, the convergence properties of the truncated GPM is very involved to achieve. It depends on many variables including the ratio of consecutive eigenvalues and the truncation error of each step. In this section, we analyse the convergence of GPM for some special cases.

Throughout this section, the following assumptions are made. For the first $(k-2)$th[1] of PM rounds we let the algorithm run m iterations, where $m \to \infty$, such that the error of estimation is zero. For the $(k-1)$th round of PM we let the algorithm run $m < \infty$ iterations. Therefore we have truncation error in estimation of λ_{k-1} and $\mathbf{v}_{k-1}$. We are interested in studying the convergence conditions of the next PM round, i.e. kth, PM round which obtains an estimate of λ_k. We begin with,

$$\begin{aligned}\hat{\mathbf{R}}_k &= \mathbf{R}_{k-1} - \hat{\lambda}_{k-1}\hat{\mathbf{v}}_{k-1}\hat{\mathbf{v}}_{k-1}^H \\ &= \sum_{j=k-1}^{K} \lambda_j \mathbf{v}_j \mathbf{v}_j^H - \hat{\lambda}_{k-1}\hat{\mathbf{v}}_{k-1}\hat{\mathbf{v}}_{k-1}^H, \qquad (6.9)\end{aligned}$$

in which, $\hat{\lambda}_{k-1}$ and $\hat{\mathbf{v}}_{k-1}$ denote the estimated λ_{k-1} and $\mathbf{v}_{k-1}$ that are obtained from truncation of PM iterations at m with $\mathbf{R}_{k-1}$ as the input.

---

[1] where $k - 2 < K$.

In the following, we study a simple case of error that influence the convergence of PM with input $\hat{\mathbf{R}}_k$.

The Estimated Subspace is Aligned

We assume that the error from the estimated subspace $\hat{\mathbf{v}}_{k-1}$ is negligible, thus $\hat{\mathbf{v}}_{k-1} = \mathbf{v}_{k-1}$. Substituting $\mathbf{v}_{k-1}$ in (6.9) we obtain,

$$\begin{aligned}\hat{\mathbf{R}}_k &= \sum_{j=k-1}^{K} \lambda_j \mathbf{v}_j \mathbf{v}_j^H - \hat{\lambda}_{k-1} \mathbf{v}_{k-1} \mathbf{v}_{k-1}^H \\ &= \sum_{j=k}^{K} \lambda_j \mathbf{v}_j \mathbf{v}_j^H + (\lambda_{k-1} - \hat{\lambda}_{k-1}) \mathbf{v}_{k-1} \mathbf{v}_{k-1}^H. \end{aligned} \tag{6.10}$$

The necessary conditions for $\hat{\mathbf{R}}_k$ to converge to λ_k using the PM algorithm is $\lambda_k > |\lambda_{k-1} - \hat{\lambda}_{k-1}|$. A straightforward generalization of the result in Theorem 8.2.1 in [GL96] yields:

$$|\lambda_{k-1} - \hat{\lambda}_{k-1}| \leq c_{k-1} (\frac{\lambda_k}{\lambda_{k-1}})^{2m},$$

where, $c_{k-1} := |\lambda_{k-1} - \lambda_K|(K - k + 2)$. Thus, the necessary condition for convergence becomes

$$c_{k-1} (\frac{\lambda_k}{\lambda_{k-1}})^{m} < \lambda_k. \tag{6.11}$$

Remark 18. Despite the oversimplification, the (6.11) provides an intuition on the convergence behaviour. We observe that if we truncate PM iterations at a large m the plausibility of convergence increases. On the other hand, for small m, i.e. large truncation error, the convergence of the algorithm depends strongly on the ratio between consecutive eigenvalues.

6.3.4. Iterative method of estimating the largest eigenvalue

In the following proposition, we propose an upperbound and a lowerbound that are converging to the largest eigenvalue monotonically decreasing and monotonically increasing with m, respectively. These bounds will be used later as the core of eigenvalue estimation of a quick channel probing algorithm.

Proposition 13. *For any* $m \geq 1$, $\mathbf{p} \succ 0$ *and* $\mathbf{R} \succ 0$ *we have*

$$L_m = \min_i \frac{(\mathbf{R}^m \mathbf{p})_i}{(\mathbf{R}^{m-1}\mathbf{p})_i} \leq \rho(\mathbf{R}) \leq \max_i \frac{(\mathbf{R}^m \mathbf{p})_i}{(\mathbf{R}^{m-1}\mathbf{p})_i} = U_m.$$

where $(\cdot)_i$ *denotes the* i*th element of its argument. In addition, for any* $\mathbf{p} \succ 0$ *and* $\mathbf{R} \succ 0$, L_m *is monotonically increasing, and* U_m *is monotonically decreasing. Furthermore,* $\lim_{m\to\infty} U_m = \lim_{m\to\infty} L_m = \rho(\mathbf{R})$.

Proof. The proof is shown in Appendix 6.7.2. ☐

6.4. Application in structured networks

In some applications, especially when measuring a spatially correlated phenomenon, we might deploy the sensors in the form of a tree with leaves being densely connected clusters of sensor nodes. The motivation for this scenario can be, for instance, the physical limitations imposed by obstacles such as buildings or maybe due to some a priori information. This sort of applications imposes a structure on $\mathbf{R}$. To achieve better performance in such a condition we revisit the proposed algorithm, considering the structured $\mathbf{R}$.

In more details, the set of all clusters $\mathcal{C} := \{\mathcal{C}_1, ...\mathcal{C}_{|\mathcal{C}|}\}$ covers all the nodes and we further assume $\forall i, j \;\; \mathcal{C}_i \cap \mathcal{C}_j = \emptyset$ where $\mathcal{C}_j$ is the jth cluster. Note that the clusters are fixed. Two nodes belong to two different clusters might be able to communicate since the graph is connected. In each cluster $\mathcal{C}_l$ we define a cluster head 0_l, which can be any node that has a direct connection to all of the nodes in the cluster. The cluster head assumes to only be able to reach the nodes in the cluster. As for most of the physical phenomena, this spatial correlation greatly decays with the Euclidean distance between the sensors. We assume sensors belong to different clusters receive almost uncorrelated observation data. There is, at least, one node in each cluster that has a communication link with at least one node from another cluster (implied by the connectivity of the graph).

For the aforementioned scenario, the structure of the resulting sample covariance matrix $\mathbf{R}$ is the summation of a block diagonal matrix $\mathbf{D} \in \mathbb{C}^{K\times K}$ and a structured error matrix with very small elements $\mathbf{E} \in \mathbb{C}^{K\times K}$, where

$$[\mathbf{D}]_{i,j} := \begin{cases} \frac{1}{N}\mathbf{y}_i^H \mathbf{y}_j & i, j \in \mathcal{C}_l \\ 0 & i, j \notin \mathcal{C}_l \end{cases} \tag{6.12}$$

and

$$[\mathbf{E}]_{i,j} := \begin{cases} 0 & i,j \in \mathcal{C}_l \\ \frac{1}{N}\mathbf{y}_i^H \mathbf{y}_j & i,j \notin \mathcal{C}_l \end{cases} \tag{6.13}$$

where, $[\mathbf{D}]_{i,j}$ and $[\mathbf{E}]_{i,j}$ are the i,jth element of matrices $\mathbf{D}, \mathbf{E}$. This model is motived as in the following. For any two sensors $i,j \in \mathcal{C}_l$, the i,jth element of $\mathbf{R}$ is given by $\frac{1}{N}\mathbf{y}_i^H \mathbf{y}_j$. The spatial correlation model and the structure of the graph suggests that for any node $k \notin \mathcal{C}_l$ we have $\mathbf{y}_i^H \mathbf{y}_k \ll \mathbf{y}_i^H \mathbf{y}_j$.

In this work, we focus on the case that $\mathbf{E} \approx \mathbf{0}$. One can consider the effect of non-zero $\mathbf{E}$ and study it by means of matrix perturbation tools.

We compute the eigenvalues of each block separately as they constitute the eigenvalues of the whole matrix. Each block corresponds to a bunch of nodes (cluster $\mathcal{C}_l, \ \ l \in \{1...|\mathcal{C}|\}$) that are located close to each other, hence, have considerable data correlation. For calculating eigenvalues of the sample covariance matrix of the nodes in $\mathcal{C}_l$ we are not required to communicate with other clusters. This allows for faster convergence in consensus steps of the proposed algorithm since we have fewer nodes in the consensus algorithm. However, we can further excel the algorithm's speed by using the recent results presented as computation over MAC (CoMAC) [ZGSY12, NG07].

To be standalone work, we summarize the steps of CoMaC based consensus presented by the authors in [ZGSY12].

1. The cluster head 0_l wakes up according to an order. Then he wakes up all the nodes in $\mathcal{C}_l$.

2. Nodes send their values to be summed to 0_l simultaneously.

3. The received signal at 0_l is given by

$$y_{0_l}(t) = \sum_{(k \in \mathcal{C}_l \backslash \{0_l\})} x_k(t), \tag{6.14}$$

   where, $x_k(t)$ is the observation data of node k at time t.

4. Then 0_l broadcasts

$$z(t) = \frac{1}{|\mathcal{C}_l|}\left(y_{0_l}(t) + x_{0_l}(t)\right). \tag{6.15}$$

After every cluster achieved the eigenvalues of their corresponding block, they need to use a protocol to share it among other clusters. This can be performed by using an ordinary consensus type algorithm among all the nodes. In the following we summarize the proposed algorithm for structured graphs:

Algorithm 5 Fast Decentralized Eigenvalue Computation (FDEC)

```
Input: Initial vector $\mathbf{q}_1^0$, clusters $\mathcal{C}_l,\ l \in \{1, \dots, Z\}$ and $\mathbf{Y} \in \mathbb{C}^{N \times K}$
Output: Eigenvalues $\lambda_k\ \forall k$ and eigenvectors $\mathbf{v}_k\ k \in \{1..K\}$ of $\mathbf{R} = \frac{1}{N}\mathbf{Y}\mathbf{Y}^H$
$l = 1$, $\mathbf{R}_1 = \mathbf{R}$, $k = 0$, $i = 0$
repeat
    in $l$th cluster, $0_l$ wakes the nodes up
    run DGPM using CoMAC consensus.
    move to the next cluster $l \leftarrow l + 1$
until $l = Z$
Use a generic CA$(\cdot)$ to share the outputs the $l$s cluster among other clusters.
```

Remark 19. In the FDEC algorithm, each step of CA$(\cdot)$ is done in only two communication shots between nodes and the cluster head, namely a multiple access phase and a broadcast phase. This is considerably faster than any traditional consensus algorithm that the number of required iterations for convergence, grows with the number nodes [ZGSY12]. However, the conditions for this gains are idealized due to oversimplification of the wireless channel.

Remark 20. The comparison between FDEC and DGPM does not shed light on eigenvalue computation since the difference between them relates to the different consensus methods. The method proposed in FDEC uses 2 rounds of iteration to reach a consensus, whereas in our simulations the consensus method from [AYSS09] needs at least 50 iterations. For a detailed comparison on cluster based consensus using CoMAC and traditional average consensus we refer the reader to the works [ZGSY12, LSMGC13]. We highlight that the clear advantage of FDEC because of using CoMAC and special graph topology does not allow for a fair comparison.

6.5. Application in channel probing

We consider a system with M established links,[2] which are referred to as primary links (PLs). In addition, we assume the existence of K pairs of inactive transmitters/receivers, which we call secondary links (SLs), that want to establish new links using the same channel of the PLs. The objective is to verify whether the secondary links can be established without decreasing the SINR of primary and secondary links below a given value.

---

[2] Throughout this chapter, as in [ZC01], the terms *link* and *transmitter/receiver pair* are used interchangeably.

In more detail, the channel gain between the ith transmitter and the jth receiver (primary users) is denoted by $c_{i,j} \geq 0$, thus, the SINR measured by the primary link $i \in \{1, ..., M\}$ can be expressed as:

$$\gamma_i = \frac{c_{i,i}p_i}{n_i + \sum_{j=1,j\neq i}^{M} c_{i,j}p_j} = \frac{p_i}{v_i + \sum_{j=1,j\neq i}^{M} g_{i,j}p_j},$$

where p_i $(1 \leq i \leq M)$ is the power at which transmitter i operates. For brevity, we denote by $g_{i,j} = \frac{c_{i,j}}{c_{i,i}}$ and $v_i = \frac{n_i}{c_{i,i}}$ the normalized channel gain and noise, respectively. The following definition will be helpful to state the main results in this study.

Definition 24. A system with M users communicating in the same frequency and at the same time is called feasible if and only if there exists a set of (non-negative) powers $\mathbf{p}^* = [p_1^*, p_2^*, ..., p_M^*]^T$ such that $\gamma_i \geq \gamma^* > 0$, where γ^* is the desired SINR threshold common to all users [FM93], [Bam98].

In this study we assume that the PLs are using the classic power admission control method proposed in [FM93, Bam98, RZ] to compute a feasible power assignment that guarantees $\gamma_i \geq \gamma^*$. (Note: more general schemes can be considered.) More precisely, each PL updates its power iteratively according to

$$p_i(m+1) = \frac{\gamma^* p_i(m)}{\gamma_i}, \tag{6.16}$$

where, m is the iteration index. If the system is feasible, by using the iteration in (6.16), the users' transmitted powers converge to the following feasible power allocation:

$$\mathbf{p}_M^* = \gamma^*(\mathbf{I} - \gamma^*\mathbf{G})^{-1}\mathbf{v}, \tag{6.17}$$

where $\mathbf{p}_M^* = [p_1^*, p_2^*, ..., p_M^*]^T$ is the feasible transmit power vector of the PLs, and $\mathbf{G} = [g_{i,j}]$ and $\mathbf{v} = [v_i]$ are the normalized channel gain matrix and noise vector, respectively. The following proposition is crucial to the results that follows.

Proposition 14 ([FM93, Bam98, RZ])**.** *The users are feasible in the sense of Definition 24 if and only if the channel matrix* $\mathbf{G}$ *satisfies* $\rho(\mathbf{G}) < \frac{1}{\gamma^*}$*, where* $\rho(\cdot)$ *denotes the spectral radius of a matrix.*

Now we assume that K SLs try to join the system. The system including both PLs and SLs has a channel gain matrix of dimension $(K+M) \times (K+M)$. We use $\mathbf{Z} \in \mathbb{R}_{++}^{(K+M)\times(K+M)}$ to denote this matrix, which is given by

$$\mathbf{Z} = \begin{bmatrix} \mathbf{G} & \mathbf{G}_{ps} \\ \mathbf{G}_{sp} & \mathbf{H} \end{bmatrix},$$

where $\mathbf{G} \in \mathbb{R}_{++}^{M\times M}$, $\mathbf{G}_{ps} \in \mathbb{R}_{++}^{M\times K}$, $\mathbf{G}_{sp} \in \mathbb{R}_{++}^{K\times M}$, and $\mathbf{H} \in \mathbb{R}_{++}^{K\times K}$.

In light of Proposition 14, the spectral radius of $\mathbf{Z}$ provides a necessary and sufficient condition to verify the admissibility of the SLs (i.e., the system composed of primary and secondary links are feasible in the sense of Definition 24). In the following, we establish a biconditional relationship between the spectral radius of $\mathbf{Z}$, $\rho(\mathbf{Z})$ and the spectral radius of a smaller matrix taking advantage of the fact that $\rho(\mathbf{G}) < \frac{1}{\gamma^*}$.

In particular, suppose that the system including both SLs and PLs is feasible; then, by using (6.17), the optimum power allocation is

$$\mathbf{p}_{\text{All users}} = \gamma^* \begin{bmatrix} \mathbf{I} - \gamma^*\mathbf{G} & -\gamma^*\mathbf{G}_{ps} \\ -\gamma^*\mathbf{G}_{sp} & \mathbf{I} - \gamma^*\mathbf{H} \end{bmatrix}^{-1} \mathbf{v}.$$

Clearly, the power vector $\mathbf{p}_{\text{All users}}$ is positive for *any* vector $\mathbf{v} > 0$ if and only if the inverse matrix is positive. If one proceeds to evaluate the inversion of the matrix above, the inverse matrix has all positive elements if and only if [ZC01]

$$(\mathbf{I} - \gamma^*\mathbf{H} - \gamma^{*2}\mathbf{G}_{sp}(\mathbf{I} - \gamma^*\mathbf{G})^{-1}\mathbf{G}_{ps})^{-1} \succ 0\,.$$

where, $'\succ'$ denotes element-wise inequality. By the Neumann series, this is equivalent to

$$\rho(\mathbf{B}) < \frac{1}{\gamma^*}, \tag{6.18}$$

where

$$\mathbf{B} \triangleq (\mathbf{H} + \gamma^*\mathbf{G}_{sp}(\mathbf{I} - \gamma^*\mathbf{G})^{-1}\mathbf{G}_{ps})\,. \tag{6.19}$$

Consequently, the whole system, including PLs and SLs, is feasible if and only if (6.18) holds, which is the relation that the proposed approach tries to verify.

6.5.1. Bounding the spectral radius

In this section, we are going to establish a sequence of lower and upper bounds on the spectral radius of the matrix $\mathbf{B}$ defined by (6.19). These bounds provide a basis for the development of a novel channel probing algorithm presented in the next section. First, in the following lemma, we prove a linear relationship between the interference and the SLs' transmit power.

Lemma 16. *At time k, let the SLs fix their transmitted powers to $\mathbf{q}^{(k)} = [p_{M+1}, \ldots, p_{M+K}]^T$ and let $\mathbf{r}^{(k)} = [r_{K+1}, \ldots, r_{K+M}]^T$, where $r_i = v_i + \sum_{j=1, j\neq i}^{M+K} g_{i,j}p_j$, be the interference measured at the ith link after the primary users respond to the power increase of SLs with the power admission control in (6.16) (i.e., we assume that PLs are using the power allocation* *given by the point to which (6.16) converges when the transmitters of SLs fix their powers). Then the following holds:*

$$\mathbf{r}^{(k)} - \mathbf{r}^{(0)} = \mathbf{B}\mathbf{q}^{(k)}, \tag{6.20}$$

where we further assume that $\mathbf{q}^{(0)} = \mathbf{0}$*. (In particular,* $\mathbf{r}^{(0)}$ *represents the interference power measured at the SL receivers when the SLs are silent.)*

Proof. With the above settings, by using the results in [ZC01], we can show that the transmit power level of the PLs converges to

$$\mathbf{p}_N^{(k)} = (\mathbf{I} - \gamma^* \mathbf{G})^{-1}(\mathbf{v} + \mathbf{G}_{ps}\mathbf{q}^{(k)}),$$

where $\mathbf{G}_{ps} = [g_{ij}^{ps}]$ is the normalized channel gain between the secondary users' transmitters and primary users' receivers. We can verify that the interference at the SLs' receivers is given by

$$\mathbf{r}^{(k)} = \mathbf{r}^{(0)} + (\mathbf{H} + \gamma^* \mathbf{G}_{sp}(\mathbf{I} - \gamma^* \mathbf{G})^{-1}\mathbf{G}_{ps})\mathbf{q}^{(k)},$$

and we conclude that the linear relationship between $\mathbf{q}^{(k)}$ and $\mathbf{r}^{(k)} - \mathbf{r}^{(0)}$ is given by

$$\mathbf{r}^{(k)} - \mathbf{r}^{(0)} = \mathbf{B}\mathbf{q}^{(k)}, \tag{6.21}$$

which concludes the proof. □

The linear relationship described in the above lemma is illustrated in Figure 6.3. Applying Lemma 16 and the Proposition 13, which introduces iteratively improvable bounds for $\rho(\mathbf{B})$, we design an algorithm to determine the feasibility of the system in the next section.

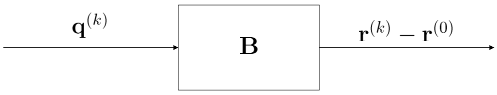

Figure 6.3.: Our system model for channel probing; a system with SL transmit powers shown as inputs and interference at the SLs receivers shown as outputs.

6.5.2. An iterative channel probing method

Now we introduce a channel probing algorithm to determine the feasibility of the system by observing the system's behaviour in response to channel probing. Based on Lemma 16, if an external user (e.g., the secondary users) increases its transmit power (and with it the interference level to the PLs), we know that the PLs compensate the loss in SINR by increasing their transmit powers. As a result, each receiver of the SLs experiences an increase in its measured interference. In our approach, starting at $k = 0$, each node of SL measures the received interference power $r_i^{(0)}$ and save this value for later use. Then, each transmitter of the SLs sends a signal at fixed power level $p_i^{(1)}$, and their corresponding receivers observe the interference power $r_i^{(1)}$ for new power allocation of the PLs. At this point, SLs compute L_1 and U_1 given in Proposition 13 to bound the spectral radius of $\mathbf{B}$ from above and below. Admissibility of the system can thus be verified if either the lower or the upper bound is above or below the value $1/\gamma^*$ (see the discussion after (6.19)). If admissibility cannot be determined with the given bounds, SLs repeat the above procedure with the new fixed power allocation given by $\mathbf{q}^{(k+1)} = \mathbf{r}^{(k)} - \mathbf{r}^{(0)} = \mathbf{B}\mathbf{q}^{(k)}$, and they again compute L_m and U_m. In doing so, we have a monotonically increasing/decreasing sequence of lower/upper bounds of $\mathbf{B}$ given by L_m and U_m in Proposition 13. One of these sequences crosses the value $1/\gamma^*$ provided that $\rho(\mathbf{Z}) \neq 1/\gamma^*$, in which moment admissibility of the secondary users can be determined. This approach is summarized below.

Note that, in the above algorithm, one of the main challenges is to implement step 5 in a truly decentralized fashion. To do so, we can use cluster-based gossip consensus algorithms [ZGSY12], which are algorithms that can efficiently compute averages of values reported by users in a network. More precisely, step 5 can be computed (by using averages) with the following approximation of the max function:[3]

$$\max\{x_1, \cdots, x_n\} \approx \sqrt[\frac{1}{p}]{\|x_1\|^p + \|x_2\|^p + ... + \|x_n\|^p}$$

for some sufficiently large $p \geq 1$. In other words, SLs reach consensus on $\|x_1\|^p + \|x_2\|^p + ... + \|x_n\|^p$, from which moment a good approximation of $\max\{x_1, \cdots, x_n\}$ is readily obtained. We refer the reader to [ZGSY12] for details on consensus algorithms.

The simulations use CDMA systems similar to those in [ZC01]. In particular, we consider $M = 18$ primary links transmitting, and $K = 18$ SLs probing the channel. The transmitters are distributed uniformly in a square area of 10km by 10km. The receivers of each link are distributed at random within a circle around their respective transmitters. In the setup, the maximum radius of the circle ranges from 100m to 500m. The channel gain from the ith receiver to the jth transmitter is given by $c_{ij} = \frac{1}{d_{ij}^3}$, where d_{ij} is the Euclidean distance. The receiver noise power is $10^{-15}W$, which is assumed to be equal for all sensors. The required SINR is $\gamma^* = 16$dB. We assume that all required signalling between receivers and transmitters is done in separate channels.

In the first experiment, we compare the performance of our algorithm with the scheme in [ZC01]. Both algorithms perturb the system by transmitting a probing signal and wait until the PLs converge. The convergence in both cases is asymptotical and the

---

[3] Min function can be computed by finding the maximum of inverted values.

Algorithm 6 Iterative Distributed Channel Probing

```
Input: Initial powers $p_i^{(1)}$, $i \in \{M+1...M+K\}$, $k = 1, T = 0$
Output: Feasibility/admissibility of the system
1: Each SU receiver measures the received interference $r_i^{(0)}$ while being silent.
2: repeat
3:     Each SU transmits with power $p_i^{(k)}$, and the corresponding receivers measure the resulting interference.
4:     After the power level of the PLs stabilizes, SU computes
           $\frac{(\mathbf{B}^k\mathbf{q})_i}{(\mathbf{B}^{k-1}\mathbf{q})_i}$
       using
           $\mathbf{B}^k\mathbf{q} = \mathbf{r}^{(k)} - \mathbf{r}^{(0)}$
       .
5:     Each SU computes
           $L_k = \min_i \frac{(\mathbf{B}^k\mathbf{q})_i}{(\mathbf{B}^{k-1}\mathbf{q})_i}$   and   $U_k = \max_i \frac{(\mathbf{B}^k\mathbf{q})_i}{(\mathbf{B}^{k-1}\mathbf{q})_i}$   (6.22)
6:     if $L_k \geq \frac{1}{\gamma^*}$ then
7:         $T = 1$ (the SUs are infeasible and leave the system)
8:     else if $U_k < \frac{1}{\gamma^*}$ then
9:         $T = 1$ (the SUs are feasible and proceed with an admission procedure)
10:    else
11:        Continue with the next step.
12:    end if
13:    Update the probing power to $p_i^{(k+1)} = r_i^{(k)} - r_i^{(0)}$ and $k = k + 1$ go to step (2)
14: until $T = 1$ or another termination condition is satisfied
```

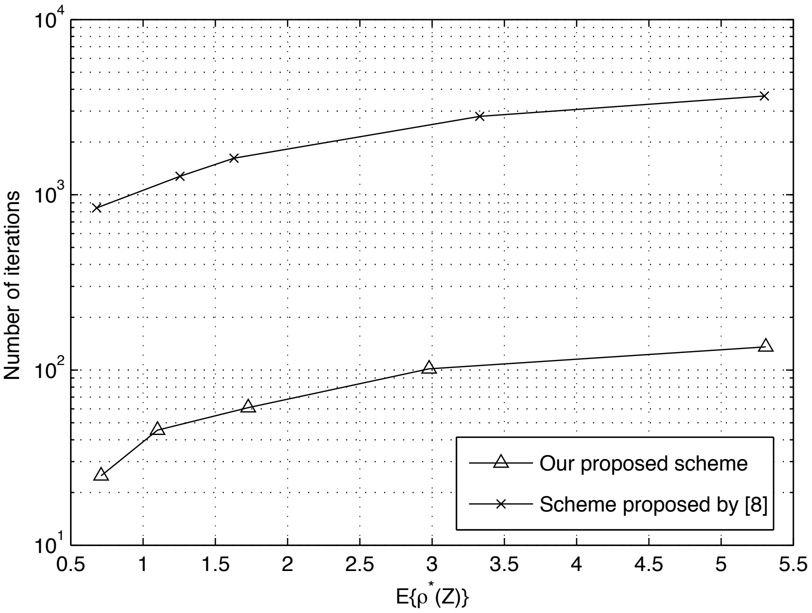

Figure 6.4.: Average number of iterations that PLs require to converge in order to SLs determine admissibility as a function of the expected value of $\gamma^* \rho(\mathbf{Z})$.

algorithm that PLs use is given in [FM93,Bam98,RZ]. We investigate the average number of iterations required by the PLs to stabilize their SIR within a range ($\epsilon = 10^{-4}$) in each scheme. This criterion is more realistic than the overall algorithm's number of iterations because the most time-consuming step in both algorithms is the convergence of PLs. As it is illustrated in Figure 6.4 our proposed algorithm needs significantly less number of PLs iterations than the other scheme. This can be explained as follows. In each iteration of [ZC01], at least, one SL is admitted. Since admitting a user causes a huge impact on the system, the PLs require a large number of iterations to adapt their powers. However, our algorithm only perturbs the PLs with small probing powers.

The second experiment focuses on the probability of reaching erroneous conclusions regarding the admissibility of secondary users when PLs use a fixed number of iterations in (6.16). The results in Figure 6.5 shows that, for the implementation of the proposed algorithm, few iterations of (6.16) are necessary to keep the probability of erroneous conclusions at low levels.

6.6. Conclusions

We have revisited the generalized power method and further developed a decentralized version, which is referred to as DGMP. We have used the generic consensus algorithm as

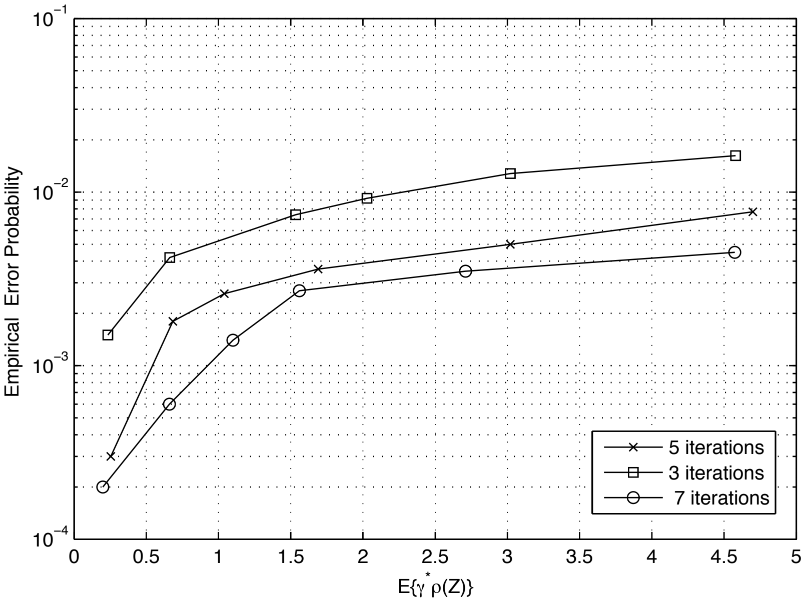

Figure 6.5.: Empirical probability of error as a function of the expected values of $\gamma^*\rho(\mathbf{Z})$ and the number of iterations used in the scheme in (6.16).

the building block for the decentralization. Furthermore, we have elaborated some analysis on the convergence of GPM, however, providing more rigorous analysis is challenging for this algorithm. We have modified this algorithm for two special applications.

We have motivated a scenario in which the covariance matrix is semi-block-diagonal. This scenario arises when the sensors are spread out forming a tree of clusters. In this scenario, we can expedite the convergence by taking advantage of CoMaC approach.

As far as the second application is concerned, we have developed a low-complexity distributed channel probing scheme. This algorithm generates two sequences that converge to a value that indicates the spectral radius of the channel. This value is the largest eigenvalue of a matrix, which determines the admissibility of the new nodes. One of the sequences provides a monotonically decreasing upper bound, and the other provides a monotonically increasing lower bound. The scheme allows multiple SLs to probe the channel and to verify their admissibility by transmitting a low-power probing tone. Simulation results show that in practice few iterations are required to detect admissibility of the secondary users in scenarios with a wide range of congestion levels of the channel.

Future work

Studying joint jamming detection and channel probing seems to be an interesting direction for some investigation. The response of the primary users to channel probing is known. This knowledge can be used to further prevent malicious users from exploiting the cognitive radios vulnerabilities. Therefore, achieving the channel probing and jamming detection jointly.

6.7. Appendix

6.7.1. Proof of convergence of ideal GPM

The proof follows from the same token given in [GL96] since for every eigenvalue λ_k we run the conventional power iteration. In order to satisfy the assumptions in the original proof by [GL96], we need the following lemma to hold:

Lemma 17. *The matrices* $\mathbf{R}_k$ *are all Hermitian, therefore diagonalizable.*

It is straightforward if we write,

$$\mathbf{R}_k = \mathbf{R} - \sum_{i=1}^{k-1} \lambda_i \mathbf{v}_i \mathbf{v}_i^H. \tag{6.23}$$

Since $\mathbf{R}$ is Hermitian positive semi-definite, $\mathbf{R}_k$s are holding the same properties as well.

6.7.2. Proof of Proposition 13

The proof is given in two parts; first we prove that

$$\lim_{m\to\infty} \frac{(\mathbf{R}^m \mathbf{p})_i}{(\mathbf{R}^{m-1}\mathbf{p})_i} = \rho(\mathbf{R}) \ \ \forall i. \tag{6.24}$$

We can reformulate (6.24) as

$$\begin{aligned} \lim_{m\to\infty} \frac{(\mathbf{R}^m \mathbf{p})_i}{(\mathbf{R}^{m-1}\mathbf{p})_i} &= \lim_{m\to\infty} \frac{(\frac{\rho(\mathbf{R})^m}{\rho(\mathbf{R})^m}\mathbf{R}^m \mathbf{p})_i}{(\frac{\rho(\mathbf{R})^{m-1}}{\rho(\mathbf{R})^{m-1}}\mathbf{R}^{m-1}\mathbf{p})_i} && (6.25) \\ &= \lim_{m\to\infty} \frac{\rho(\mathbf{R})((\rho(\mathbf{R})^{-1}\mathbf{R})^m \mathbf{p})_i}{((\rho(\mathbf{R})^{-1}\mathbf{R})^{m-1}\mathbf{p})_i} && (6.26) \end{aligned}$$

Define $a_m \triangleq ((\rho(\mathbf{R})^{-1}\mathbf{R})^m\mathbf{p})_i$. By the Perron theorem reproduced in Appendix 6.7.3, it follows that

$$\lim_{m\to\infty} a_m = (\mathbf{x}\mathbf{y}^T\mathbf{p})_i. \tag{6.27}$$

Thus, by considering (6.27), the limit in (6.26) yields

$$\begin{aligned}\lim_{m\to\infty} \frac{(\mathbf{R}^m\mathbf{p})_i}{(\mathbf{R}^{m-1}\mathbf{p})_i} &= \rho(\mathbf{R})\frac{\lim_{m\to\infty} a_m}{\lim_{m\to\infty} a_{m-1}}\\ &= \rho(\mathbf{R})\frac{(\mathbf{x}\mathbf{y}^T\mathbf{p})_i}{(\mathbf{x}\mathbf{y}^T\mathbf{p})_i}\\ &= \rho(\mathbf{R}),\end{aligned}$$

which proves (6.24).

Now we prove the monotonicity of L_m and U_m. We start with the lower bound. The objective is to prove that

$$\min_i \frac{(\mathbf{R}^m\mathbf{p})_i}{(\mathbf{R}^{m-1}\mathbf{p})_i} \leq \min_i \frac{(\mathbf{R}^{m+1}\mathbf{p})_i}{(\mathbf{R}^m\mathbf{p})_i}\ . \tag{6.28}$$

Define $q \triangleq \min_i \frac{(\mathbf{R}^m\mathbf{p})_i}{(\mathbf{R}^{m-1}\mathbf{p})_i} = \min_i \sum_k b_{i,k}\frac{(\mathbf{R}^{m-1}\mathbf{p})_k}{(\mathbf{R}^{m-1}\mathbf{p})_i}$. We can manipulate the RHS of (6.28) as follows:

$$\begin{aligned}\min_i \frac{(\mathbf{R}^{m+1}\mathbf{p})_i}{(\mathbf{R}^m\mathbf{p})_i} &= \min_i \sum_j b_{i,j}\frac{(\mathbf{R}^m\mathbf{p})_j}{(\mathbf{R}^m\mathbf{p})_i}\\ &= \min_i \sum_j b_{i,j}\frac{\sum_k b_{j,k}(\mathbf{R}^{m-1}\mathbf{p})_k}{\sum_k b_{i,k}(\mathbf{R}^{m-1}\mathbf{p})_k}.\end{aligned}$$

Hence,

$$\begin{aligned}\min_i \sum_j b_{i,j}\frac{\sum_k b_{j,k}\frac{(\mathbf{R}^{m-1}\mathbf{p})_k}{(\mathbf{R}^{m-1}\mathbf{p})_j}(\mathbf{R}^{m-1}\mathbf{p})_j}{\sum_k b_{i,k}(\mathbf{R}^{m-1}\mathbf{p})_k} &\geq \min_i \sum_j b_{i,j}\frac{q\ (\mathbf{R}^{m-1}\mathbf{p})_j}{\sum_k b_{i,k}(\mathbf{R}^{m-1}\mathbf{p})_k}\\ &= \min_i q\frac{\sum_j b_{i,j}(\mathbf{R}^{m-1}\mathbf{p})_j}{\sum_k b_{i,k}(\mathbf{R}^{m-1}\mathbf{p})_k} = q,\end{aligned}$$

where the first inequality follows from the definition of q.

The proof for the upper bound is similar to that of the lower bound. In more detail, we want to prove that

$$\max_i \frac{(\mathbf{R}^m \mathbf{p})_i}{(\mathbf{R}^{m-1} \mathbf{p})_i} \geq \max_i \frac{(\mathbf{R}^{m+1} \mathbf{p})_i}{(\mathbf{R}^m \mathbf{p})_i} \ .$$

Define $w \triangleq \max_i \frac{(\mathbf{R}^m \mathbf{p})_i}{(\mathbf{R}^{m-1} \mathbf{p})_i} = \max_i \sum_k b_{i,k} \frac{(\mathbf{R}^{m-1} \mathbf{p})_k}{(\mathbf{R}^{m-1} \mathbf{p})_i}$. Therefore,

$$\begin{aligned}\max_i \frac{(\mathbf{R}^{m+1} \mathbf{p})_i}{(\mathbf{R}^m \mathbf{p})_i} &= \max_i \sum_j b_{i,j} \frac{(\mathbf{R}^m \mathbf{p})_j}{(\mathbf{R}^m \mathbf{p})_i} \\ &= \max_i \sum_j b_{i,j} \frac{\sum_k b_{j,k} (\mathbf{R}^{m-1} \mathbf{p})_k}{\sum_k b_{i,k} (\mathbf{R}^{m-1} \mathbf{p})_k}.\end{aligned}$$

Therefore,

$$\begin{aligned}\max_i \sum_j b_{i,j} \frac{\sum_k b_{j,k} \frac{(\mathbf{R}^{m-1} \mathbf{p})_k}{(\mathbf{R}^{m-1} \mathbf{p})_j} (\mathbf{R}^{m-1} \mathbf{p})_j}{\sum_k b_{i,k} (\mathbf{R}^{m-1} \mathbf{p})_k} &\leq \max_i \sum_j b_{i,j} \frac{w \ (\mathbf{R}^{m-1} \mathbf{p})_j}{\sum_k b_{i,k} (\mathbf{R}^{m-1} \mathbf{p})_k} \\ &= \max_i w \frac{\sum_j b_{i,j} (\mathbf{R}^{m-1} \mathbf{p})_j}{\sum_k b_{i,k} (\mathbf{R}^{m-1} \mathbf{p})_k} = w,\end{aligned}$$

which concludes the proof.

6.7.3. The Perron theorem

For any primitive $\mathbf{A}$ (there exist a k such that $\mathbf{A}^k > 0$) we have,

$$\lim_{m \to \infty} [\rho(\mathbf{A})^{-1} \mathbf{A}]^m = \mathbf{x}\mathbf{y}^T,$$

where $\mathbf{x}$ and $\mathbf{y}$ are the normalized ($\mathbf{y}^T \mathbf{x} = 1$) left and right Perron vectors of $\mathbf{A}$.

Publication list

[1] J. Mohammadi, I. Bjelankovic, and S. Stanczak. Achievable secrecy rate region for discrete memoryless untrusted relay channel. *Submitted to IEEE Transaction on Information Theory*, 25(1), 2016.

[2] J. Mohammadi, S. Limmer, and S. Stanczak. A decentralized eigenvalue computation method for spectrum sensing based on average consensus. *Accepted for publication in Frequenz Journal on Wireless Communications*, 2016.

[3] J. Mohammadi, F. Gao, Y. Rong, and W. Chen. Joint source and relay design for two-hop amplify-and-forward relay networks with QoS constraints. *EURASIP Journal on Wireless Communications and Networking*, 2013(1), 2013.

[4] I. Bjelankovic, J. Mohammadi, and S. Stanczak. Strong secrecy and stealth for broadcast channels with confidential messages. In *Accepted for publication in International Symposium Information Theory*, volume 25. IEEE, 2016.

[5] S. Limmer, J. Mohammadi, and S. Stanczak. A simple algorithm for approximation by nomographic functions. In *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 453–458, Sept 2015.

[6] J. Mohammadi, S. Stanczak, and M. Zheng. Joint spectrum sensing and jamming detection with correlated channels in cognitive radio networks. In *Communication (ICC) Workshop , 2015 IEEE International Conference on*, pages 889–894. IEEE, 2015.

[7] M. Kaliszan, J. Mohammadi, and S. Stanczak. Cross-layer security in two-hop wireless gaussian relay network with untrusted relays. In *Communications (ICC), 2013 IEEE International Conference on*, pages 2199–2204, June 2013.

[8] J. Mohammadi, F. Penna, and S. Stanczak. Energy-efficient node selection for cooperative spectrum sensing with spatial correlation. In *Signals, Systems and Computers, 2013 Asilomar Conference on*, pages 1011–1015, Nov 2013.

[9] J. Mohammadi, M. Kaliszan, S. Stanczak, and J. Schreck. Secrecy capacity limits of multiple antenna multiple eavesdropper multicast. In *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, pages 1896–1900, Nov 2012.

[10] M. Zheng, J. Mohammadi, S. Stanczak, and H. Yu. Concurrent transmission versus time sharing in gaussian interference channels. In *Signal Processing Advances in Wireless Communications (SPAWC), 2012 IEEE 13th International Workshop on*, pages 314–318, June 2012.

[11] J. Mohammadi, S. Stanczak, R.L.G. Cavalcante, and J. Etesami. Iterative distributed channel probing for cognitive radios with power-controlled wireless links. In *Wireless Communication Systems (ISWCS), 2011 8th International Symposium on*, pages 567–571, Nov 2011.

Bibliography

[AAK01] M.-S. Alouini, A. Abdi, and Mostafa Kaveh. Sum of gamma variates and performance of wireless communication systems over nakagami-fading channels. *Vehicular Technology, IEEE Transactions on*, 50(6):1471–1480, 2001.

[ABN92] Barry C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *A First Course in Order Statistics*. Wiley-Interscience, 1992.

[ALLP12] E. Axell, G. Leus, E.G. Larsson, and H.V. Poor. Spectrum sensing for cognitive radio : State-of-the-art and recent advances. *Signal Processing Magazine, IEEE*, 29(3):101–116, May 2012.

[ANDH08] T.Y. Al-Naffouri, A.F. Dana, and B. Hassibi. Scaling laws of multiple antenna group-broadcast channels. *Wireless Communications, IEEE Transactions on*, 7(12):5030 –5038, december 2008.

[ATV$^+$12] A. Attar, H. Tang, A.V. Vasilakos, F.R. Yu, and V.C.M. Leung. A survey of security challenges in cognitive radio networks: Solutions and future research directions. *Proceedings of the IEEE*, 100(12):3172–3186, Dec 2012.

[AYAK12] I. S. Ansari, F. Yilmaz, M.-S. Alouini, and O. Kucur. New results on the sum of Gamma random variates with application to the performance of wireless communication systems over Nakagami-m fading channels. *ArXiv:1202. 2576 [Cs. IT]*, pages 1–10, 2012.

[AYSS09] T.C. Aysal, M.E. Yildiz, A.D. Sarwate, and A. Scaglione. Broadcast gossip algorithms for consensus. *Signal Processing, IEEE Transactions on*, 57(7):2748–2761, July 2009.

[Bam98] N. Bambos. Toward power-sensitive network architectures in wireless communications: Concepts, issues, and design aspects. *IEEE Personal Commun. Mag.*, 5:50–59, June 1998.

[BB11] Matthieu Bloch and Joao Barros. *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[BCM95] N. Bambos, S. C. Chen, and D. Mitra. Channel probing for distributed access control in wireless communication networks. In *Global Telecommunications Conference, 1995. GLOBECOM '95., IEEE*, volume 1, pages 322–326 vol.1, Nov 1995.

[BL13] M.R. Bloch and J.N. Laneman. Strong secrecy from channel resolvability. *Information Theory, IEEE Transactions on*, 59(12):8077–8098, Dec 2013.

[BP15] M. Benammar and P. Piantanida. Secrecy capacity region of some classes of wiretap broadcast channels. *IEEE Transactions on Information Theory*, 61(10):5564–5582, Oct 2015.

[CBA15] Remi A Chou, Matthieu R Bloch, and Emmanuel Abbe. Polar coding for secret-key generation. *Information Theory, IEEE Transactions on*, 61(11):6213–6237, 2015.

[CC62] A. Charnes and W. W. Cooper. Programming with linear fractional functionals. *Naval Research Logistics*, 9(3-4):181–186, 1962.

[CK78] I. Csiszar and J. Körner. Broadcast channels with confidential messages. *IEEE Trans. Inform. Theory*, 24(3):339–348, May 1978.

[CK82] Imre Csiszar and Janos Korner. *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Academic Press, Inc., Orlando, FL, USA, 1982.

[CM87] P. H. Calamai and J. J. Moré. Projected gradient methods for linearly constrained problems. *Mathematical Programming*, 39:93–116, 1987.

[Csi96] I. Csiszar. Almost independence and secrecy capacity. *Probl. Peredachi Inf.*, 32:1:48–57, 1996.

[CTB06] D. Cabric, A. Tkachenko, and R.W. Brodersen. Spectrum sensing measurements of pilot, energy, and collaborative detection. In *Proc. Military Communications Conference, MILCOM06. IEEE*, pages 1–7, Oct 2006.

[CY02] Ning Cai and R.W. Yeung. Secure network coding. In *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, pages 323–, 2002.

[CY11] Ning Cai and R.W. Yeung. Secure network coding on a wiretap network. *Information Theory, IEEE Transactions on*, 57(1):424–435, Jan 2011.

[Dun81] J. C. Dunn. Global and Asymptotic Convergence Rate Estimates for a Class of Projected Gradient Processes. *SIAM Journal on Control and Optimization*, 19(3):368–400, May 1981.

[ETW08] R.H. Etkin, D.N.C. Tse, and Hua Wang. Gaussian interference channel capacity to within one bit. *Information Theory, IEEE Transactions on*, 54(12):5534–5562, Dec 2008.

[EU12] Ersen Ekrem and Sennur Ulukus. Degraded compound multi-receiver wiretap channels. *Information Theory, IEEE Transactions on*, 58(9):5681–5698, Sept 2012.

[FM93] G.J. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Trans. Veh. Technol.*, 42(4):641–646, Nov. 1993.

[Gal74] R. G. Gallager. Coding and capacity for degraded broadcast channels. *Problemy Peridachi Informatsi*, 10(3):3–14, 1974.

[GK12] Abbas El Gamal and Young-Han Kim. *Network Information Theory.* Cambridge University Press, New York, NY, USA, 2012.

[GL96] G. H. Golub and C. F. Van Loan. *Matrix Computations.* John Hopkins University Press, New York, 1996.

[GP80] S. I. Gel'fand and M. S. Pinsker. Coding for channel with random parameters. *Problems of Control Theory*, 9(1):19–31, 1980.

[GS05] A. Ghasemi and E.S. Sousa. Collaborative spectrum sensing for opportunistic access in fading environments. In *Proc. DySPAN, IEEE International Symposium on*, pages 131–136, Nov 2005.

[GS13] M. Goldenbaum and S. Stanczak. Robust analog function computation via wireless multiple-access channels. *Communications, IEEE Transactions on*, 61(9):3863–3877, September 2013.

[Gud91] M. Gudmundson. Correlation model for shadow fading in mobile radio systems. *Electronics Letters*, 27(23):2145–2146, 1991.

[Hay05] Simon Haykin. Cognitive radio: brain-empowered wireless communications. *Selected Areas in Communications, IEEE Journal on*, 23(2):201–220, Feb 2005.

[HK81] T.S. Han and K. Kobayashi. A new achievable rate region for the interference channel. *IEEE Trans. Inform. Theory*, 27:49–60, Jan. 1981.

[HK13] Jie Hou and G. Kramer. Informational divergence approximations to product distributions. In *Information Theory (CWIT), 2013 13th Canadian Workshop on*, pages 76–81, June 2013.

[HK14] Jie Hou and G. Kramer. Effective secrecy: Reliability, confusion and stealth. In *Information Theory (ISIT), 2014 IEEE International Symposium on*, pages 601–605, June 2014.

[HMS12] Jing Huang, A. Mukherjee, and A.L. Swindlehurst. Secrecy analysis of unauthenticated amplify-and-forward relaying with antenna selection. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 2481 –2484, march 2012.

[HMS13] Jing Huang, A. Mukherjee, and A.L. Swindlehurst. Secure communication via an untrusted non-regenerative relay in fading channels. *Signal Processing, IEEE Transactions on*, 61(10):2536–2550, May 2013.

[HS14] Maryam Haghighat and Seyed Mohammad Sajad Sadough. Cooperative spectrum sensing for cognitive radio networks in the presence of smart malicious users. *AEU-International Journal of Electronics and Communications*, 68(6):520–527, 2014.

[HY09a] Xiang He and Aylin Yener. Interference channels with strong secrecy. In *Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing*, Allerton'09, pages 811–818, Piscataway, NJ, USA, 2009. IEEE Press.

[HY09b] Xiang He and Aylin Yener. Two-hop secure communication using an untrusted relay. *EURASIP J. Wireless Commun. Netw., Special Issue in Wireless Physical Layer Security*, page 13, 2009.

[HY10a] X. He and A. Yener. Cooperation with an untrusted relay: A secrecy perspective. *Information Theory, IEEE Transactions on*, 56(8):3807 –3827, August 2010.

[HY10b] Xiang He and Aylin Yener. Cooperative jamming: The tale of friendly interference for secrecy. In *Securing Wireless Communications at the Physical Layer*, pages 65–88. Springer, 2010.

[HY13] Xiang He and A. Yener. End-to-end secure multi-hop communication with untrusted relays. *Wireless Communications, IEEE Transactions on*, 12(1):1–11, January 2013.

[Ism80] Mourad E. H. Ismail. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 8(1):143–145, 1980.

[JKK12] C. Jeong, I.-M. Kim, and D. I. Kim. Joint secure beamforming design at the source and the relay for an amplify-and-forward MIMO untrusted relay system. *Signal Processing, IEEE Transactions on*, 60(1):310 –325, jan 2012.

[JL06] N. Jindal and Z.-Q. Luo. Capacity limits of multiple antenna multicast. In *Information Theory, 2006 IEEE International Symposium on*, pages 1841 –1845, July 2006.

[Kay93] Steven M. Kay. *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory (v. 1).* Prentice Hall, 1 edition, April 1993.

[Khi04] A. Khisti. Coding techniques for multicastingl. Master's thesis, Massachusetts Institute of Technology, 2004.

[KLWH09] Sungtae Kim, Jemin Lee, Hano Wang, and Daesik Hong. Sensing performance of energy detector with correlated multiple antennas. *Signal Processing Letters, IEEE*, 16(8):671–674, 2009.

[KMS13] M. Kaliszan, J. Mohammadi, and S. Stanczak. Cross-layer security in two-hop wireless gaussian relay network with untrusted relays. In *Communications (ICC), 2013 IEEE International Conference on*, pages 2199–2204, June 2013.

[KW10] A. Khisti and G. W. Wornell. Secure transmission with multiple antennas i: The MISOME wiretap channel. *IEEE Transactions on Information Theory*, 56(7):3088–3104, July 2010.

[LH09] Husheng Li and Zhu Han. Dogfight in spectrum: jamming and anti-jamming in multichannel cognitive radio systems. In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–6. IEEE, 2009.

[LH10] Husheng Li and Zhu Han. Catch me if you can: an abnormality detection approach for collaborative spectrum sensing in cognitive radio networks. *Wireless Communications, IEEE Transactions on*, 9(11):3554–3565, 2010.

[LK10] W. Luh and D. Kundur. Distributed secret sharing over the gaussian interference wiretap channel. In Ruoheng Liu and Wade Trappe, editors, *Securing Wireless Communications at the Physical Layer*, pages 39–40. Springer US, 2010.

[LLK09] Loukas Lazos, Sisi Liu, and Marwan Krunz. Mitigating control-channel jamming attacks in multi-channel ad hoc networks. In *Proceedings of the Second ACM Conference on Wireless Network Security*, WiSec '09, pages 169–180, New York, NY, USA, 2009. ACM.

[LLXC12] Z. Liu, H. Liu, W. Xu, and Y. Chen. Exploiting jamming-caused neighbor changes for jammer localization. *IEEE Transactions on Parallel and Distributed Systems*, 23(3):547–555, March 2012.

[LM11] Q. Li and W.-K. Ma. Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming. *Signal Processing, IEEE Transactions on*, 59(8):3799 –3812, aug. 2011.

[LMB07] L. Lima, M. Medard, and J. Barros. Random linear network coding: A free cipher? In *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, pages 546 –550, june 2007.

[LMS15] S. Limmer, J. Mohammadi, and S. Stanczak. A simple algorithm for approximation by nomographic functions. In *53rd Annual Allerton Conference on Communication, Control, and Computing*, Urbana, IL, USA, Sept. 29 – Oct. 2 2015.

[LMSY08] Ruoheng Liu, I. Maric, P. Spasojevic, and R.D. Yates. Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions. *Information Theory, IEEE Transactions on*, 54(6):2493–2507, June 2008.

[LP09] Ruoheng Liu and H.V. Poor. Secrecy capacity region of a multiple-antenna gaussian broadcast channel with confidential messages. *Information Theory, IEEE Transactions on*, 55(3):1235–1249, March 2009.

[LSBP$^+$07] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai (Shitz), and S. Verdu. Cognitive interference channels with confidential messages. In *in Proc. 45th Annu. Allerton Conf. Communication, Control and Computing, Monticello, IL*, pages 1–6, Sep. 2007.

[LSM11] L. Li, A. Scaglione, and J. H. Manton. Distributed principal subspace estimation in wireless sensor networks. *IEEE Journal of Selected Topics in Signal Processing*, 5(4):725–738, Aug 2011.

[LSMGC13] S. Limmer, S. Stanczak, M-Goldenbaum, and R.L.G. Cavalcante. Exploiting interference for efficient distributed learning in cluster-based wireless sensor networks. In *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP) - Network Theory Symposium*, Austin, Texas, USA, 2013. invited.

[LZZQ11] L.P. Luo, P. Zhang, G.C. Zhang, and J.Y. Qin. Spectrum sensing for cognitive radio networks with correlated multiple antennas. *Electronics Letters*, 47(23):1297–1298, 2011.

[Mar79] K. Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inform. Theory*, 25(3):306–311, May 1979.

[Mau94] UeliM. Maurer. The strong secret key rate of discrete random triples. *Communications and Cryptography, -Two Sides of One Tapestry, Kluwer Academie Publishers*, 276:271–285, 1994.

[MGKP09] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, and G. Pantziou. A survey on jamming attacks and countermeasures in wsns. *IEEE Communications Surveys Tutorials*, 11(4):42–56, Fourth 2009.

[MKSS12] J. Mohammadi, M. Kaliszan, S. Stanczak, and J. Schreck. Secrecy capacity limits of multiple antenna multiple eavesdropper multicast. In *Signals, Systems, & Computers (ASILOMAR), 2012 46th Asilomar Conference on*, pages 1896–1900, Nov 2012.

[Mos85] P. G. Moschopoulos. The distribution of the sum of independent Gamma random variables. *Ann. Inst. Statist. Math. (Part A)*, 37(6):541–544, 1985.

[MPS13] J. Mohammadi, F. Penna, and S. Stanczak. Energy-efficient node selection for cooperative spectrum sensing with spatial correlation. In *Signals, Systems and Computers, 2013 Asilomar Conference on*, pages 1011–1015, Nov 2013.

[MSCE11] J. Mohammadi, S. Stanczak, R.L.G. Cavalcante, and J. Etesami. Iterative distributed channel probing for cognitive radios with power-controlled wireless links. In *Wireless Communication Systems (ISWCS), 2011 8th International Symposium on*, pages 567–571, Nov 2011.

[MSZ15] J. Mohammadi, S. Stanczak, and M. Zheng. Joint spectrum sensing and jamming detection with correlated channels in cognitive radio networks. In *Communication (ICC) Workshop , 2015 IEEE International Conference on*, pages 889–894. IEEE, 2015.

[MVOV96] Alfred J Menezes, Paul C Van Oorschot, and Scott A Vanstone. *Handbook of applied cryptography.* CRC press, 1996.

[MW00] Ueli Maurer and Stefan Wolf. Information-theoretic key agreement: From weak to strong secrecy for free. In *Proceedings of the 19th International Conference on Theory and Application of Cryptographic Techniques*, EUROCRYPT'00, pages 351–368, Berlin, Heidelberg, 2000. Springer-Verlag.

[MZL$^+$12] Hossen Mustafa, Xin Zhang, Zhenhua Liu, Wenyuan Xu, and Adrian Perrig. Jamming-resilient multipath routing. *Dependable and Secure Computing, IEEE Transactions on*, 9(6):852–864, 2012.

[NG07] B. Nazer and M. Gastpar. Computation over Multiple-Access Channels. *IEEE Trans. on Inf. Theory*, 53, no. 10, Oct. 2007.

[NPG11] B. Nadler, F. Penna, and R. Garello. Performance of eigenvalue-based signal detectors with known and unknown noise level. In *Communications (ICC), 2011 IEEE International Conference on*, pages 1–5, June 2011.

[OSFM07] Reza Olfati-Saber, Alex Fax, and Richard M Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

[PHH$^+$12] Sazia Parvin, Farookh Khadeer Hussain, Omar Khadeer Hussain, Song Han, Biming Tian, and Elizabeth Chang. Cognitive radio network security: A survey. *Journal of Network and Computer Applications*, 35(6):1691–1708, 2012.

[PL12] Alejandro Proano and Loukas Lazos. Packet-hiding methods for preventing selective jamming attacks. *Dependable and Secure Computing, IEEE Transactions on*, 9(1):101–114, 2012.

[PS12] F. Penna and S. Stanczak. Decentralized largest eigenvlaue test for multi-Sensor signal detection . In *Proc. IEEE Global Communications Conference (GLOBECOM)*, Anaheim, CA, USA, December 2012.

[PS15a] F. Penna and S. Stanczak. Decentralized eigenvalue algorithms for distributed signal detection in wireless networks. *Signal Processing, IEEE Transactions on*, 63(2):427–440, Jan 2015.

[PS15b] F. Penna and S. Stanczak. Decentralized eigenvalue algorithms for distributed signal detection in wireless networks. *IEEE Transactions on Signal Processing*, 63(2):427–440, Jan 2015.

[RSEJ15] J. Richter, C. Scheunert, S. Engelmann, and E.A. Jorswieck. Weak secrecy in the multiway untrusted relay channel with compute-and-forward. *Information Forensics and Security, IEEE Transactions on*, 10(6):1262–1273, June 2015.

[RZ] Zvi Rosberg and Jens Zander. Toward a framework for power control in cellular systems. *Wireless Networks*, 4(4):215–222.

[Saa11] Yousef Saad. *Numerical Methods for Large Eigenvalue Problems*. Society for Industrial and Applied Mathematics, 2011.

[SDČ10] Mario Strasser, Boris Danev, and Srdjan Čapkun. Detection of reactive jamming in sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 7(2):16, 2010.

[Sha49] C.E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28:656–715, 1949.

[Sha79] Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, November 1979.

[SHT04] Anant Sahai, Niels Hoven, and Rahul Tandra. Some Fundamental Limits on Cognitive Radio. In *the 42sd Allerton Conference on Communication, Control and Computing*, 2004.

[SU07] S. Shafiee and S. Ulukus. Achievable rates in gaussian MISO channels with secrecy constraints. In *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, pages 2466 –2470, june 2007.

[TV05] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, June 2005.

[VKP05] E. Visotsky, S. Kuffner, and R. Peterson. On collaborative detection of TV transmissions in support of dynamic spectrum sharing. In *DySPAN, IEEE International Symposium on*, pages 338–345, Nov 2005.

[VKT15] S. Vatedka, N. Kashyap, and A. Thangaraj. Secure compute-and-forward in a bidirectional relay. *Information Theory, IEEE Transactions on*, 61(5):2531–2556, May 2015.

[vT93] H.C.A van Tilborg. *Coding theory, a first course.* Chartwell Bratt Studentlitteratur, Lund, Sweden, 1993.

[Wyn75] A.D. Wyner. The wire-tap channel. *Bell System Technical Journal*, 54:1355–1387, Oct. 1975.

[XCC09] Jin Xu, Yi Cao, and Biao Chen. Capacity bounds for broadcast channels with confidential messages. *Information Theory, IEEE Transactions on*, 55(10):4529–4542, Oct 2009.

[XS14] Chunsheng Xin and M. Song. Detection of PUE attacks in cognitive radio networks based on signal activity pattern. *Mobile Computing, IEEE Transactions on*, 13(5):1022–1034, May 2014.

[XSC01] Mingbo Xiao, N. B. Shroff, and E. K. P. Chong. Distributed admission control for power-controlled cellular wireless systems. *IEEE/ACM Transactions on Networking*, 9(6):790–800, Dec 2001.

[XWZ] Wenyuan Xu, Timothy Wood, and Yanyong Zhang. Channel surfing and spatial retreats: defenses against wireless denial of service. In *in Proceedings of the 2004 ACM workshop on Wireless security, 2004*, pages 80–89.

[ZC01] Chenxi Zhu and M. S. Corson. A distributed channel probing scheme for wireless networks. In *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 1, pages 403–411 vol.1, 2001.

[ZFPP15] S. Zhang, L. Fan, M. Peng, and H.V. Poor. Near-optimal modulo-and-forward scheme for the untrusted relay channel. *eprint arXiv:1503.08928*, Mar 2015.

[ZGSY12] M. Zheng, M. Goldenbaum, S. Stanczak, and H. Yu. Fast average consensus in clustered wireless sensor networks by superposition gossiping. In *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, pages 1982–1987, April 2012.

[ZHSA05] G. Zhou, T. He, J. A. Stankovic, and T. Abdelzaher. Rid: radio interference detection in wireless sensor networks. In *INFOCOM 2005. 24th Annual Joint*

Conference of the IEEE Computer and Communications Societies. Proceedings IEEE, volume 2, pages 891–901 vol. 2, March 2005.

[ZMcLM09] Yonghong Zeng, Senior Member, Ying chang Liang, and Senior Member. Eigenvalue based spectrum sensing algorithms for cognitive radio. *IEEE Trans. on Communications*, pages 1784–1793, 2009.