A Critical Review of Physical Layer Security in Wireless Networking

Jiwei Li

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Master of Philosophy

of

University College London.

Department of Computer Science

University College London

September 15, 2015

I, Jiwei Li, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Wireless networking has kept evolving with additional features and increasing capacity. Meanwhile, inherent characteristics of wireless networking make it more vulnerable than wired networks. In this thesis we present an extensive and comprehensive review of physical layer security in wireless networking.

Different from cryptography, physical layer security, emerging from the information theoretic assessment of secrecy, could leverage the properties of wireless channel for security purpose, by either enabling secret communication without the need of keys, or facilitating the key agreement process. Hence we categorize existing literature into two main branches, namely keyless security and key-based security. We elaborate the evolution of this area from the early theoretic works on the wiretap channel, to its generalizations to more complicated scenarios including multiple-user, multiple-access and multiple-antenna systems, and introduce not only theoretical results but practical implementations.

We critically and systematically examine the existing knowledge by analyzing the fundamental mechanics for each approach. Hence we are able to highlight advantages and limitations of proposed techniques, as well their interrelations, and bring insights into future developments of this area.

Acknowledgements

Many thanks to my supervisor Dr. Kyle Jamieson for his support of this research and care of my general well-beings, and I genuinely appreciate all the helps from my colleagues in the system and networks research group at UCL.

My deepest gratitude goes to my parents and girlfriend for their continuous encouraging and comforts.

Contents

List of Figures

Chapter 1

Introduction

The proliferation of wireless networking brings great challenges for ensuring secure communications. The ubiquitous presence and growing capacity of wireless networks increases the danger that users' privacy could be comprised. Moreover, some inherent characteristics of wireless communication, such as *openness*, *broadcast* and *superposition* present different difficulties in ensuring security. Due to its inherent complexity, security in wireless networking is concerned in various layers in the Open Systems Interconnection Model (OSI). For instance, link layer solutions like Extensible Authentication Protocol (EAP) and application layer encryption like Wired Equivalent Privacy (WEP) and Wireless Application Protocol (WAP) have been developed and widely deployed throughout today's Wi-Fi networks. However, those protocols have been proven compromised [1, 2, 3], especially facing more dedicated adversaries. Traditionally, networking security is largely trusted to cryptosystems. However, the nature of wireless communications raises issues such as key distribution and management for symmetric schemes, and increasing computational complexity for asymmetric systems. Cryptosystems also face the risk of being broken by adversaries, in those cases where the reliability of the encryption code is not mathematically guaranteed.

In response, the design of secure physical layer attracts increasing attention, which could be deployed in parallel with higher layer methods. Physical layer security holds the potential for enabling secret communication and reliable authentication, as well as sparing computational complexity for devices with fewer resources. The study of secrecy features in physical layer communications can be dated back to the 70s, with the groundbreaking works by Wyner. However, it has been a theoretical topic for decades. In recent years, due to the development in wireless technology such as cognitive radio

and Multiple Input Multiple Output (MIMO) [4], we have witnessed an resurgence of interest in this area, with better prospects of being practically applied.

There are two major aspects of security in wireless networking, which could be referred to as *active attack* and *passive attack*. The latter is also commonly addressed as *eavesdropping*. Active attack could be an intrusion with faked authentication, or sabotaging the networking by malicious jamming. Eavesdropping, on the other hand, is aimed at breaking the secrecy of the communication and acquire private information. While active attack presents an serious issue in special scenarios such as military network, eavesdropping remains a more common threat in general use. In current literature, physical layer security is mainly focused on preventing eavesdropping, in other words, ensuring the secrecy of communication. Physical layer security against active attacks is also well discussed in the community, but often with separate terms, such as jamming and counter-jamming techniques. Methods for ensuring reliable authentication is commonly addressed as *physical layer authentication*. In the rest of this thesis, we will employ those terminologies, and physical layer security will be used to mainly represent protection from eavesdropping.

Despite the rich existing knowledge, secure and reliable wireless communication is still largely an open problem. With the widespread of wireless networking as well as developments in wireless technology, there will be more challenges to come in the future, making this topic a constant active research area. Physical layer security could contribute to overall secure communication in many ways. The fundamental idea behind physical layer security is to leverage features of wireless channel and randomness of signal noise to limit the amount of information that can be obtained by unauthorized eavesdropper. Based on theoretical models, with proper coding design and signal processing method, physical layer security could ensure secret communication without the need of established keys. On the other hand, when the use of secret keys is needed, by exploiting the inherent randomness of wireless medium, physical layer security provides an alternative way of establishing keys, which reduces the computational burden for application layer encryption. Hence, physical layer security methods could be further categorized into two sections: *keyless* schemes and *key-based* schemes, although there are also *exotic* approaches that could not fall in either category.

As mentioned, the topic was first discussed decades ago and is experiencing an re-

naissance in recent years, hence retrospective examinations as well as taxonomy of current trends are necessary to further progress towards practical applications. In [5], Shiu et al provide a brief tutorial on physical layer security for wireless networks, where the authors classify existing techniques into five categories, namely information-theoretic security, channel approaches, code approaches, power approaches and signal design approaches. Yet such classification failed to explain the field in a unified framework and highlight the connections between different approaches, as it is based on exploited features instead of design objectives and incentives. Also, the tutorial doesn't contain sufficient theoretical and applied details, as well as up-to-date developments in the area. A more recent survey of the topic is provided in [6], where Mukherjee et al employ a methodology similar to this thesis, which assesses applied designs with a unified information-theoretic foundation. However, the survey places much greater emphasis on the theoretical capabilities of systems instead of their practicalities and detailed designs. The survey provides a comprehensive coverage of a wide range of channel models, yet some of those models are impractical for security implementations.

In this thesis, we present a comprehensive and critical review of this rich field, while balancing between applied techniques and theoretical assessments. We will not only present the theoretical origin that lays the foundation of the topic, but also focus on the recent developments which lead to practical implementations. From the vast existing literature, we selectively cover those that are more relevant to practical and up-to-date wireless networking systems. Hence, works included in this thesis should meet following criteria. First, selected works should be significant in theoretical or experimental results. Second, selected works should consider typical and applicable networking models. Third, we are particularly interested in security with emerging technologies such as multiple antennas, which holds great potentials for security enhancement.

Questions that we would like to address include not only theoretical properties, but the practicality of physical layer security in modern wireless system. Practical challenges and limitations faced by such systems will be discussed, such as scalability and energy consumption. As diverse research developments are ongoing is this active field, we believe such review is needed to give clear guidance for future progress. The rest of this chapter will present an introduction of typical security threats in wireless

network, and an outline of the thesis.

1.1 Wireless Adversary Models

Attacks in wireless networks fall into two basic categories: passive attacks and active attacks. The most common type of passive attack is eavesdropping. On the other hand, active attacks significantly interfere with normal network operations because the adversary tries to alter the network data. The most common types of active attacks include denial-of-service (DoS) attacks, man-in-the-middle attacks, and information disclosure and message modification attacks.

A DoS attack is when an adversary attempts to exhaust or suspend resources available to legitimate users. In wireless, DoS generally consists of efforts to temporarily or indefinitely interrupt service of legitimate AP to its intended receiver, and involves more diverse techniques compared to wired DoS. Wireless DoS attacks can be further divided into three categories: Jamming attacks, flooding attacks, and semantic attacks. A jamming attack is when the adversary transmits artificial noise on the legitimate signal band to disrupt the communication. Flooding attacks exhaust resources by sending a large amount of messages to a protocol participant. Jamming attacks and flooding attacks have many similarities, and differences between them mainly lay in two aspects. First, the jammer is usually a illegal node not associated with the network while flooding attacker could be a recognized user in the network transmitting overwhelming traffics. Second, jamming attacks usually create noise to corrupt the signal at the receiver's end, while flooding attacks usually use overwhelming amounts of traffic to exhaust resources at AP, thus to prevent service for legitimate users.

On the other hand, semantic attacks exploit protocol weaknesses by transmitting valid protocol messages with forged message fields. One example of a semantic attack is the deauthentication attack against IEEE 802.11 networks [7]. After an 802.11 client has selected an access point to use for communication, it must first authenticate itself to the AP before further communication may commence. Moreover, part of the authentication framework is a message that allows clients and access points to explicitly request deauthentication from one another. Unfortunately, this message itself is not authenticated using any keying material. Consequently the attacker may spoof this message, either pretending to be the access point or the client, and direct it to the other

party. In response, the access point or client will exit the authenticated state and will refuse all further packets until authentication is re-established, and by repeating the attack persistently a client may be kept from transmitting or receiving data indefinitely. This example shows the insufficiency of the current 802.11 link layer authentication mechanism at defending against semantic attacks.

Another form of active attack in wireless networking is the man-in-the-middle attack, where an intruder pretends to be a legitimate node of the network and deceives the authentication system so as to usurp system resources and steal private information. It makes independent connections with the victims and relays messages between them, making them believe that they are talking directly to each other over a private connection, when in fact the entire conversation is controlled by the attacker. A man-in-the-middle attack usually involves passive attack like eavesdropping, to capture authentication sequences and secret keys, and therefore to obtain privilege to access information illegally. Hence, a man-in-the-middle attack could be regarded as a overlapping type of active and passive attack. Another example of man-in-the-middle attack is the rogue AP attack, where unauthorized APs placed by the attacker pretend to be legitimate ones.

Other types of active attacks less commonly seen in wireless network include information disclosure and message modification attacks. In information disclosure, a compromised node acts as an information leaker by deliberate disclosure of confidential information to unauthorized nodes. Information such as the amount and periodicity of the traffic between a selected pair of nodes and the changing traffic patterns can be valuable to the adversaries in many military applications. Message modification refers to an attack in which an aggressor performs additions or deletions to the network communication content.

Eavesdropping is the most common attack in wireless networking, which is to capture confidential and private message from private communication. In some scenarios, an eavesdropper might be able to intercept the transmitted signal but cannot obtain any critical information from it due to the encryption. On the other hand, traffic analysis could be used to determine the locations and identities of the communicating parties by intercepting and examining the transmitted messages. The traffic information may be useful for tracking the communication patterns of any two parties and break the en-

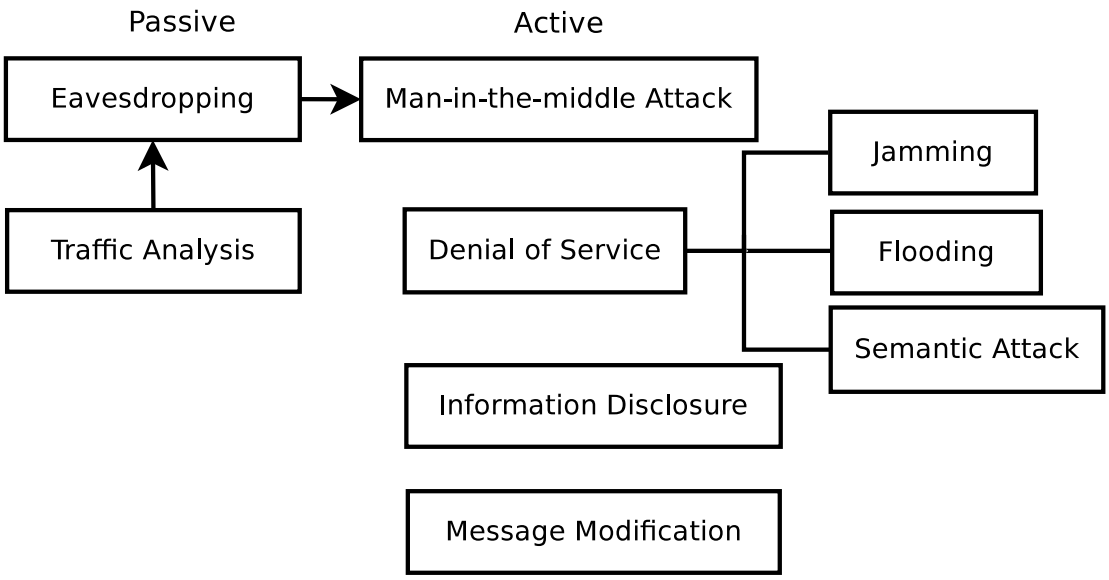

Figure 1.1: Diagram of common security attacks in wireless

cryption. Hence eavesdropping can be performed even if the messages are encrypted and the malicious users can use the information gleaned from eavesdropping for other forms of attack. Thus, eavesdropping is not popular and dangerous in itself, but also linked to other type of threats, and brings chances for active attacks. Protection from eavesdropping remains a major front of developing secure wireless communication.

A diagram of common types of attacks in wireless networking is shown in figure 1.1. In wireless networking, traffic analysis and eavesdropping are closely related, and commonly regarded as a single problem, since both of them passively capture communicating information in the networking. In this context, we use eavesdropping as a general terminology to refer to wireless passive attack, which consists of traffic analysis and narrowly defined eavesdropping. It should be noted that attacks in wireless networking happen in different stages of communication. As shown in figure 1.1, eavesdropping is the prerequisite that enables man-in-the-middle attack, and jamming is the most common form of denial of service attack in wireless networking.

As mentioned, in this thesis we focus on the mainstream physical layer security research, which tries to address the problem of confidentiality, i.e., to keep information secret from unintended receivers. Hence eavesdropping attack is the main form of threat considered in following chapters. It should be noted that security risks in today's wireless networks are much more complex than any threat model introduced above. Firstly, different aspects of secure communications, from authentication, reliability to

confidentiality and privacy, are interconnected, as an exposure in one direction may leave other aspects vulnerable. Also, an attacker is not necessarily limited to a fixed type of attacks. A *hybrid attacker* may be able to launch both active and passive attacks, which facilitate each other.

In the research of physical layer security, a *divide-and-conquer* approach is commonly employed as to focus on an individual problem with a simplified setting. Hence, physical layer security designs tackle confidentiality as a single problem, and commonly assume established authentication. It should be noted that only the initial authentication is required in such assumption, as subsequent message authentication can be implemented by sacrificing fractions of the previous message rate. On the other hand, even when such initial authentication is not available, techniques proposed for confidential communication in broadcasting channels can be used to treat any unintended receiver as eavesdropper and allow transmissions of public and private messages simultaneously. Meanwhile, most of the existing literature also assumes *pure eavesdropper*, which is limited to only passive attack.

More specifically, two basic security models are addressed in this thesis, namely secret communication and secret key generation, corresponding to keyless and key-based approaches respectively. As distinctions between the two have been previously introduced, it should also be noted that physical layer security designs commonly assume certain conditions on the communication channel, corresponding to different security models and techniques. Hence, while physical layer security captures the underlying properties of physical layer communication, which are ignored by conventional security protocols, it is at the same time restricted by those properties. In practical system design, physical layer security could therefore *complement* and *reinforce* existing conventional techniques, instead of being considered as an overall *replacement*.

1.2 Structure of the Thesis

In the next chapter, we introduce the fundamentals and recent development of information theoretic security, an important tool for accessing the secrecy capacity of wireless communication, and the common ground for physical layer security designs. Chapter 3 discusses different layers of keyless physical layer strategy, from the design of wiretap codes, to the construction of secrecy channels in multiple antenna system. Chapter 4

presents a review of key-based physical layer security.

Conclusions of this study, and discussion on future directions considering emerging topics in the area are presented in Chapter 5.

Chapter 2

Secrecy Capacity of Wireless Channels

In this chapter we present the major information theoretic analysis results on secrecy of wireless networks. The term *information theoretic security* is derived from its close relationship from information theory, as the fundamental reasoning is basically to exploit the information theoretic properties of wireless channel for security purposes. By examining the secrecy capacity and security related features of wireless channels, information theoretic security provides common ground and guidance for the design of secure wireless networking, and is an important theoretical tool for practical system design to assess its performance. This area, pioneered by Wyner in his groundbreaking article [8], lays the foundation of research in physical layer security. Wyner's work has been constantly extended to more sophisticated channels, and with different assumptions. In recent years, we witness an resurgence of information theoretic security together with vast developments of wireless technology. In this chapter we introduce the most notable works and recent developments in this area.

2.1 The Wiretap Channel Model

With the widespread deployment of wireless networks, secrecy becomes an equally important feature as capacity. The secrecy problem of a networking system is that of communicating a message through the main channel without conveying information about the message through the eavesdropper's channel. The wire-tap channel model proposed by Wyner [8] is a significant work that examines the relationship of secrecy and channel capacity in a networking system, and attracted great attention since it was published. The wiretap model is illustrated in figure 2.1, where the legal communication through the main channel is being wire-tapped by a malicious eavesdropper, and

the channel that the eavesdropper listen on is referred as the wiretap channel. **m** denotes the message being sent to the legitimate receiver, **x** as the encoded signal and **y** and **z** as the received signal at the legitimate receiver and the eavesdropper respectively, and **u** represents the message decoded at the legitimate receiver. We assume the original message contains k bits of information, which will be decoded into n bits of signal. This pair $\{k,n\}$ is a basic feature of the coding scheme.

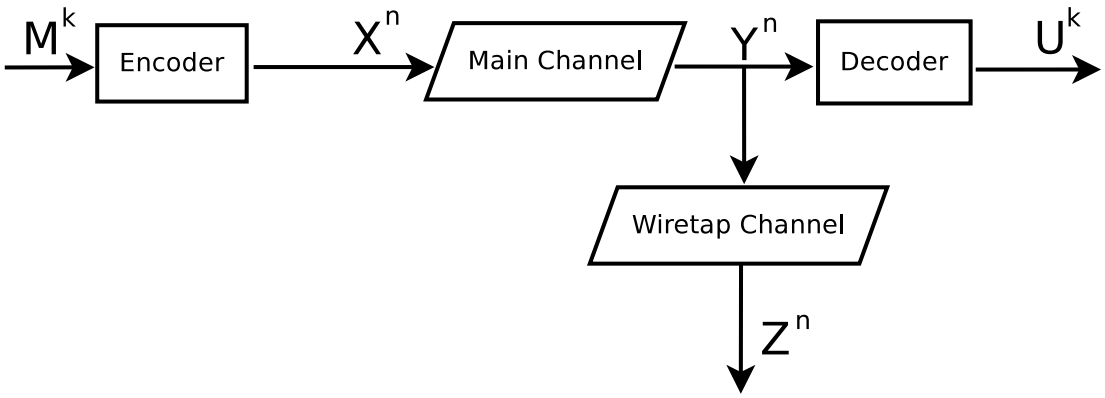

Figure 2.1: Diagram of a generic wiretap channel model

Wyner's work follows Claude Shannon's work on perfect secrecy [9]. Shannon assumed a scenario where both the main and eavesdropper's channels are noiseless. As traditionally, encryption is the major route to achieve confidential communication, Shannon adapted the use of the secret key **k** in his model. The key **k** is used to encrypt the message **y** into cryptogram **c**. The assumption on the eavesdropper is that it has unlimited computational power, knowledge of the coding scheme, and receives an identical copy of the signal received at the legitimate receiver through a noiseless channel.

Naturally, Shannon defined the communication to be perfectly secure when the encrypted signal **c** received by the eavesdropper gives no difference to the eavesdropper's perception of the message **m**. In other words, **c** and **m** are mutually independent. Following classic information theoretic paradigm, the independency between messages is measured with *mutual information*. On the other hand, the amount of information contained in messages is modeled with the uncertainty of information, measured with *information entropy*. Assuming two discrete random variables X and Y with probablity distribution $p(x)$ and $p(y)$, and joint distribution $p(x,y)$, their mutual information and

the information entropy of X can be defined as following.

$$I(X;\, Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) log(\frac{p(x,y)}{p(x)p(y)}) \tag{2.1}$$

$$H(X) = -\sum_{x \in X} p(x) log p(x) \tag{2.2}$$

Hence, in Shannon's definition of perfect secrecy, the eavesdropper has the same level of uncertainty of the information, before or after receiving the encrypted signal. Thus the mutual information between $\mathbf{m}$ and $\mathbf{z}$ is zero.

$$\lim_{n \to \infty} I(\mathbf{m};\, \mathbf{z}) = 0 \tag{2.3}$$

Following this statement, [10] proved that the secret key $\mathbf{k}$ needs to have at least the same entropy as the actual message itself, i.e., $H(\mathbf{k}) \geq H(\mathbf{m})$. In practice, an example of such encryption system is *one-time pad*, which uses a secret key with the same length as the message itself, and encrypts the messages bitwise with corresponding bits in the key. Shannon's result discouraged further research in information theoretic secrecy, since the availability of a scheme that distributes a secret key at both communicating nodes with entropy exceeding that of the message, means that the message can be shared securely between the two nodes without a key. In wiretap channel model, Wyner revisited the problem with modified assumptions. First, the noiseless assumption of Shannon was relaxed by assuming a possibly noisy main channel and an wiretap channel that is nosier than the main channel. Second, Wyner relax the perfect secrecy condition to that based on normalized leakage information, hence transfer equation 2.3 to:

$$\lim_{n \to \infty} \frac{I(\mathbf{m};\, \mathbf{z})}{n} = 0 \tag{2.4}$$

Wyner further defined the equivocation rate of such a system with those assumptions, which measures the residual ambiguity about the transmitted message at the wiretapper, or eavesdropper. It should be noted that, since Wyner defines secrecy in a normalized manner, this statement is usually considered weaker than Shannon's statement. To better identify the condition for secrecy, we could rewrite equation 2.4 to the form

of conditional entropy:

$$\lim_{n\to\infty} \frac{H(\mathbf{m};\, \mathbf{z})}{n} = \Delta \tag{2.5}$$

Δ represents the level of confusion that the eavesdropper would perceive. Naturally, we would like Δ to be as large as possible. Here, we also consider the *error-rate* of the coding scheme, which is a important feature for wireless communication. With a given pair of encoder and decoder, we define the error-rate as following:

$$P_e = \frac{1}{k} \sum_{i=1}^{k} P\{m_i \neq u_i\} \tag{2.6}$$

From a system designer's perspective, we would like to have P_e as small as possible. It is common to set a upper limit for the error rate, represented as ε, such that:

$$P_e \leq \varepsilon \tag{2.7}$$

Here, we define a pair of factors as $\{R, R_e\}$, where R represents the transmission rate and R_e represents the equivocation rate that we would like to maximize. Then two constraints about this pair can be defined as follows:

$$R - \varepsilon \leq \frac{H(\mathbf{m})}{n} \tag{2.8}$$

$$R_e - \varepsilon \leq \Delta \tag{2.9}$$

Hence, the problem of channel secrecy is refined as to characterize achievable $\{R, R_e\}$ under those constraints. Wyner's remarkable work [8] assess this problem. It should be noted that, as shown in figure 2.1, with the assumption that the wiretap channel is degraded, hence the signal X, Y, and Z forms a Markov chain, which is important for Wyner's result. Let C_m and C_w represent the capacities of the main channel and the wiretap channel respectively, there exits an rate C_s, such that when $0 \leq R \leq C_s$, $0 \leq R_e \leq H(\mathbf{m})$. The set of $\{R, R_e\}$ when $R_e = H(\mathbf{m})$ is of particular importance, as it indicates that the equivocation at the eavesdropper is the same as the entropy of the message itself. In other words, *perfect secrecy* is achieved.

Hence, as a conclusion, in the general wiretap channel model, when $C_m > C_w$, a maximum rate C_s, $0 \leq C_s \leq C_m$ is achievable to allow secure communication.

It should be noted that, as shown in figure 2.1, the received signal at the eavesdropper also goes through the main channel. The total eavesdropper's channel is then a cascade of the main channel and the separate wiretap channel. Hence in some early articles, the capacity of the wiretap channel is denoted as C_{mw}. In this thesis, we employ the recent notation, and treat the eavesdropper's channel as a whole. Hence, in the rest of the thesis, we will use C_w for this purpose.

Wyner's results are of practical significance, and have since been widely discussed and further extended. The conclusion of a possible transmission rate that ensure message privacy without additional means and limitations on the eavesdropper's computational resources, facilitate both direct secure communication without keys, and reliable key agreement. However, there are preconditions of Wyner's results which limit its practicality without further extensions.

- **Degraded wiretap channel**: The basic wiretap channel assumes implicitly that the eavesdropper's channel needs to be cascade of the main channel and an additional wiretap channel following it. Hence, the whole eavesdropper's channel is degraded compared to the main channel. To be more specific, the signals transmitted and received by the legitimate receiver and the eavesdropper should form a Markov chain such that $X \rightarrow Y \rightarrow Z$. However, in practice, such ideal condition is not usually met. In other words, the received signal at the eavesdropper $\mathbf{z}$ is not ideally independent from the transmitted signal $\mathbf{y}$.

- **Static main channel**: Wyner's results are based on static scenarios, where the main channel should be fixed and its channel information should be known to the transmitter. In reality, fading and mobile channels often complicate such assumptions, especially considering wireless channels.

- **Single access**: In the original model, only single access channel is considered. However, in modern networking, especially wireless communication, multiple access is widely used to boost the capacity of the systems. In general multi-user networking, more complicated features like interference and broadcasting may further bring challenges to the secrecy capacity.

2.2 Improvement of Wiretap Channel and Its Applicability

Following Wyner's work, there has been various articles that address its potential challenges. Those works greatly improved the adaptability of the wiretap channel and generalized the results into more sophisticated scenarios. The early enhancement of Wyner's work was focused on introducing more complicated single-antenna channels.

In [11], the authors examined the possibility and practicality of *partial secrecy* regarding Wyner's model. In other words, as *perfect secrecy* is defined as $R_e = H(\mathbf{m})$, in reality, a large enough R_e could be sufficient to keep the communication secure. The authors discussed the feasibility of a partial secrecy rate, and proposed to improve its secrecy by linear coding. More specifically, as shown in Wyner's work, the perfect secure rate C_s is between 0 and main channel capacity C_m. From a coding perspective, the transmission rate R could be related to the coding ratio k/n. By limiting R to the secrecy rate C_s, the system is in fact adding redundancy bits $n-k$ to confuse the eavesdropper. The authors argued that it is possible to leverage redundancy bits to transmit useful information, with acceptable compromises on security. If we treated the k bits information and $n-k$ bits redundancy as separate messages, then the eavesdropper will be ignorant of each message separately, but could determine one of them with the knowledge of other. This work gives useful insights on the nature of the wiretap channel model. However, it is of limited practicality, as the method works on the assumption that the transmitted message has no redundancy. If the assumption is violated, however, the proposed operation may actually help the eavesdropper to reduce equivocation.

In [12], Ozarow and Wyner further characterized this situation where the eavesdropper get arbitrary partial information of the source message with additional means. The authors categorized this situation as the type-2 wiretap channel. The trade-off between the coding rate k/n, the bits that observed by the eavesdropper $n-k$ has been examined in details. To be more specific, the equivocation Δ could be expressed as:

$$\Delta = k - 1 - \frac{2.23}{\sqrt[4]{n}} \approx k - 1 \tag{2.10}$$

Thus, with any k bits from the encoded signal, the eavesdropper is kept ignorant to

the message, and the communication remains secure. This work validated the general concept of using redundancy to improve secrecy. This results are very significant for practical wireless communication. In wireless networking protocols, such as 802.11, it is common to insert fixed bits for framework control, and it should be assumed that the eavesdropper has priori knowledge of those bits. An example could be the preamble in the 802.11 protocol.

In [13], the authors expended the wiretap channel model with Gaussian noise on both main channel and wiretap channel. Leung-Yan-Cheong and Hellman made a significant contribution to applying the wiretap channel model into practical scenarios, and clearly define the secrecy rate in Gaussian wiretap channel as following:

$$C_s = C_m - C_w \tag{2.11}$$

Hence, the secrecy capacity of Gaussian wiretap channel is clearly quantified, and it is also stated that a non-zero secrecy capacity is only possible when the wiretap channel is degraded compared to the main channel. To be more specific, the two channels are assumed to have zero mean and variances σ_M^2 and σ_E^2 respectively. Let P denote the average power constraint over the transmitted symbols, the maximum achievable secure rate of communication can be expressed (considering real signals) as:

$$C_s = \frac{1}{2}\log(1+\frac{P}{\sigma_M^2}) - \frac{1}{2}\log(1+\frac{P}{\sigma_M^2+\sigma_E^2}) \tag{2.12}$$

Hence, the secrecy capacity of the system is the difference of capacities between the main channel and the eavesdropper channel.

In [14], Csiszar and Korner presented an important piece of work of further generalizing wiretap channel model to broadcasting and non-degraded channels. The work considered an generalized scenario where the transmitter is broadcasting public messages to both the intended receiver and the eavesdropper, while also sending private message that should be kept secret from the eavesdropper. Csiszar and Korner extended the capacity pair of $\{R, R_e\}$ to a rate tuple as $\{R_1, R_e, R_0\}$, representing the achievable private message rate, equivocation rate and common message rate respectively. Constraints on the rate tuple can be expressed as:

$$0 \leq R_e \leq R_1 \tag{2.13}$$

$$R_e \leq I(\mathbf{v};\ \mathbf{y}|\mathbf{m}) - I(\mathbf{v};\ \mathbf{z}|\mathbf{m}) \tag{2.14}$$

$$R_1 + R_0 \leq I(\mathbf{v};\ \mathbf{y}|\mathbf{m}) + min(I(\mathbf{m};\ \mathbf{y}), I(\mathbf{m};\ \mathbf{z})) \tag{2.15}$$

$$0 \leq R_0 \leq min(I(\mathbf{m};\ \mathbf{y}), I(\mathbf{m};\ \mathbf{z})) \tag{2.16}$$

To revise the original wiretap channel model with this extended framework, it is straightforward to see that the capacity pair $\{R, R_e\}$ represents a special situation where there is no public message transferred. In the broadcasting scenario, the secrecy capacity is defined in a similar way as in the original model, which is the most achievable rate in tuple $\{R_1, R_e, 0\}$. Based on equations 2.13 to 2.16, the secrecy capacity C_s could be expressed as:

$$C = \max_{M \to V \to X \to YZ} (I(\mathbf{m};\ \mathbf{y}) - I(\mathbf{m};\ \mathbf{z})) \tag{2.17}$$

It should be noted that the notation $M \to V \to X \to YZ$ indicates the sequence of signals forms a Markov chain, where V represents an auxiliary random variable inserted as a prefix channel to increase the secrecy capacity by facilitating the Markov chain. Thus, the assumption of degraded wiretap channel, i.e., $X \to Y \to Z$ is waived. However, the main channel should still be of better quality than the eavesdropper's channel. Csiszer and Korner [14] greatly generalized the wiretap channel model and has since been widely discussed in following works.

Recently, the secrecy capacity of fading channels becomes an active topic of discussion. As introduced above, the key assumption of wiretap channel in static scenario is that the main channel is stronger than the eavesdropper's channel. In fading channels, however, since time diversity is provided, it is straightforward to consider leverage the fading feature to improve secrecy.

To capture the time diversity and dynamic nature of fading channels, a new metric of secrecy capacity is introduced, referred to as secrecy outage probability. Similar to

traditional rate outage probability, secrecy outage probability measures the likelihood that the instantaneous achievable secrecy rate R_i is below certain threshold ε, expressed as:

$$P_o = P\{R_i < \varepsilon\}, \varepsilon > 0 \tag{2.18}$$

Apart from fading channels, the outage probability metric is commonly used in schemes where the secrecy capacity is dynamic, which will be elaborated in the next chapter.

In [15], Barros and Rodrigues studied the secrecy capacity of slow fading channels. To be more specific, the authors examined the secrecy outage probability and rate capacity of quasi-static fading channels. The results showed that secure communication on those channels is possible even if the average SNR of the main channel is lower than that of the eavesdropper's channel. Also, the results showed the instantaneous rate capacity in fading channels could be higher than the rate capacity of non-fading channel with the same average SNRs. Those results suggested that an opportunistic approach of secure communication is feasible in fading channels, which will be discussed in details in later parts of this thesis.

Gopala et al. [16] further validated the notion that fading is actually beneficial to secrecy. The authors focused on the secrecy capacity of slow fading channels with different assumptions of knowledge of CSI at the transmitter. The results showed that CSI of the eavesdropper's channel holds less importance compared to the CSI of the main channel. When the main channel is in good enough condition, the awareness of eavesdropper' CSI is irrelevant to the secrecy capacity. The authors also discussed the impacts of power allocation and rate adaptation.

In [17], the authors investigated the achievable secrecy capacity in a scenarios where the main channel is AWGN and the eavesdropper's channel is Rayleigh fading with additive Gaussian noise. The authors assumed that only the main channel information is known at the transmitter, and the fading of the eavesdropper's channel is unknown. The authors suggested that artificial noise and power bursting can be used to achieve improve secrecy, especially when the main channel is in worse condition than the eavesdropper's channel.

In [18], Liang et al. generalized the analysis of fading wiretap channel to broadcasting scenario, where additive Gaussian noise is considered. In [19], Liang et al. further considered the scenario of compound wiretap channel, where the main channel and the eavesdropper's channel could have a variety of possible states. The work also introduced the notion of secrecy degree of freedom as a metric of accessing the secrecy in compound channel. Secrecy degree of freedom could serve as an approximation of the secrecy capacity in the high SNR scenario, and is very useful to analyse situations such as multi-access channels, which will be elaborated in the next section.

Various other extensions based on Wyner's results have been proposed. In [20, 21], the authors derived the secrecy capacity of more specific broadcasting channels, where parallel channels are considered. Situations with different allocation of subchannels are discussed.

It is clearly to see, that after its first establishment in the 70's, the topic of information theoretic security is experiencing a resurgence recently, where the focus of the discussion has been shifted to the practicality of such method in complex and diverse real scenarios. The remain challenges of keyless information theoretic security are largely due to the conditions that required to make the secrecy capacity achievable. It should be noted that there are assumptions made to validate the secrecy capacity of keyless schemes, many of which still remains compulsory with recent developments.

- **Knowledge of channel state information**: A major difficulty of keyless information theoretic security is the need of CSI for both the legitimate receiver and the eavesdropper at the transmitter, since achieving secrecy capacity is essentially a coding problem and the equivocation of the eavesdropper is ensured provided that the code scheme used for transmission is correctly tailored to the channel. Although the knowledge of the legitimate receiver at the transmitter could be expected, since both legitimate users can cooperate to characterize their channel. The knowledge of the adversary, especially considering passive attackers, is problematic. However, in a situation where the eavesdropper is also a legitimate user of the network, such requirement could be met. Methods to circumvent the requirement on CSI is still an major direction of research in information theoretic security.

- **Need of randomizer**: Fundamentally, information theoretic security relies on the equivocation of messages. In practice, such protocols require proper code design to maximize the equivocation at the eavesdropper. To enable such equivocation, stochastic encoding is needed instead of deterministic mapping. In other words, one-to-many mapping code is needed to create equivocation at the eavesdropper. In practice, the performance of stochastic encoding is dependent on the random generator, and the randomization mechanism should be considered. In [22], the trade-off between the rate of randomization for the stochastic encoding and the secrecy capacity is investigated, proving that randomization is a notable feature for practical keyless secrecy communication.

- **Authentication**: It should be noted that there is an implicit assumption made by the wiretap channel model, which is often ignored. The wiretap channel model assumes that the main channel is already authenticated to begin with. In other words, the transmitter is aware of the identity of the legitimate receiver and such identity is trustworthy. This assumption is not restrictive though, since authentication mechanisms can be implemented by upper layer mechanics. It is worthy considering when the totality of the networking security is considered.

- **Weak secrecy**: As mentioned before, the perfect secrecy defined in the wiretap channel model is a normalized version of Shannon's proposed concept [9]. In the context of information theoretic security, the latter is often termed as *strong secrecy*, while the former is referred to as *weak secrecy*. Such difference should be noted, although in practice it is often considered trivia.

2.3 Extensions to Multiple-Antenna Channels

The recent development of multi-antenna technology greatly improves the capacity of wireless networking. As a result, MIMO network has been a active topic in recent years regarding both theoretical and applied research. The advantages of MIMO can be leveraged in many fronts. It has been widely discussed in recent years that MIMO may also lead to enhanced secrecy. Hence theoretic proof and practical code design have been proposed to address the secrecy capacity of MIMO network. For instance Ekrem and Ulukus [23] generalized wiretap channel model to multiuser MIMO system. The

authors identified an relationship between the secrecy capacity region and precoding, stating the secrecy capacity can be achieved with dirty paper coding.

2.3.1 Introduction to MIMO

Here we also give a brief introduction to MIMO. It is widely understood that in MIMO system, the spectral efficiency is much higher than that of the conventional single-antenna channels. The advantage of MIMO can be leveraged in various ways. Traditionally, MIMO have been used to increase diversity to combat channel fading. Since each pair of transmit and receive antennas provides a signal path from the transmitter to the receiver, by sending signals that carry the same information through diverse paths, multiple independently faded replicas of the data symbol can be obtained at the receiver end to improve the reliability of the transmission. This mechanism is usually referred to as *diversity gain*.

Other than utilizing diverse paths to combat fading, another idea suggests that in a MIMO channel, fading can in fact be beneficial due to the increased degrees of freedom available for communication [24]. Essentially, if the path gains between individual transmit-receive antenna pairs fade independently, the channel matrix could be well conditioned so that paths of pairs of transmitting and receiving antennas could form parallel *spatial channels*. Hence it is obvious that by transmitting independent information streams in parallel through the spatial channels, the data rate, or throughput, can be increased. This effect is usually referred as *multiplexing gain* while its correspondent technology referred as *spatial multiplexing*. In single stream scenario, it is usually implemented by decomposing a high rate steam into multiple lower rate ones to be transmitted by multiple antennas. With different assumptions, diversity gain and multiplexing gain provide opportunities to leverage multiple antennas of transmitter and receiver to enhance the communication. The comparison and trade-off of those two aspects of MIMO have been well discussed [25]. In the context of multi-user MIMO, it is straightforward to see that *multiplexing gain* is of more significance than *diversity gain*, as the nature of multi-user wireless networking is to transmit concurrent multiple streams to different users. Hence, in the following of this thesis, we mainly discuss *multiplexing gain* regarding MIMO.

Another concept which is closely related to MIMO is beamforming. Applied in

various fields, beamforming is essentially a signal processing technique that utilize sensor arrays to form directional transmission or reception, which is achieved by making array signals at intended location have constructive interference while having destructive interference at other ones. Hence beamforming is traditionally applied in multiple input single output (MISO) systems. However, beamforming has also strong relevance to MIMO since beamforming provides the foundation of precoding, which is essentially is beamforming with multiple streams in its narrowest definition. It should be noted that, to allow beamforming, CSI should be known at the transmitter.

In single user MIMO, when CSI is both provided to transmitter and receiver, precoding and spatial multiplexing could combine to optimize the output, i.e., signal power is maximized at its intended user. In multi-user MIMO, precoding performs a much more important role as the fundamental function that, at each intended receiver, aggregate the intended steam constructively and let other signals cancel out each other. Hence, beamforming techniques are substantial and well discussed in the context of distributed MIMO [26]. Precoding is also an important feature regarding security since failed precoding means that interfering streams are not canceled out and streams could be leaked to unintended user. Precoding also determines the signal that an eavesdropper could hear.

Essentially, precoding is to condition the transmitting streams to let signal add constructively at intended users and destructively at others. The conditioning is based on the CSI known to the transmitter. In multi-user MIMO, precoding is generally utilized for the down-link. Since the introduction of multi-user MIMO [27], there has been extensive works on its theoretical capacity [28] and practical precoding algorithms to achieve high capacity with low complexity for implementation.

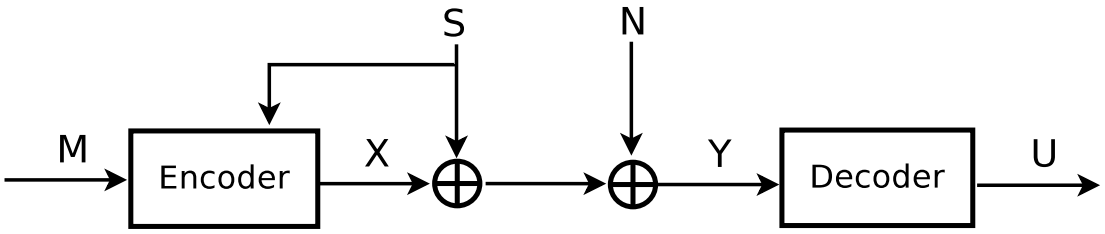

Figure 2.2: Diagram of a generic DPC model

Under assumption of known CSI, Dirty Paper Coding (DPC) [29] is known to be the optimal scheme for capacity and power allocation. While DPC is an important theo-

retical tool for realizing perfect interference pre-cancellation with no power penalty, its practical implementation is notoriously difficult [30], often considered infeasible. The generalized DPC model is shown in figure 2.2 with following mathematical expression:

$$Y = X + S + N \tag{2.19}$$

where M and U are the message to be sent and message decoded at receiver respectively, X denotes the transmitted signal, Y denotes the received signal, S is arbitrary interference known at the transmitter, and N is a statistically independent Gaussian random variable representing the channel noise with variance P_N, and the power constraint on the transmitted signal is P_X. Assuming S is known to the transmitter, to eliminate the interference, a naive approach could be pre-canceling the interference, and the transmission equation will be following:

$$Y = X' + S + N = X - S + S + N = X + N \tag{2.20}$$

However, with this approach, the transmit power will be $E[X']^2 = E[X]^2 + E[S]^2$. Hence a large portion of transmit power will be used to cancel interference, which results in reduced transmission rate. When the interference is arbitrarily far away from the symbol, the energy penalty will be unbearable. In DPC, Costa proved when interference is known to the transmitter, the capacity of such a system is

$$C = \frac{1}{2} log(1 + \frac{P_X}{P_N}) \tag{2.21}$$

which is the same as there is no interference. This result greatly encourages the research on precoding since it provides theoretical proof of the existence of precoding scheme that enables pre-cancellation of known interference without power and rate penalty.

However since Costa precoding is notoriously difficult to implement in practice, various alternative precoding algorithms have been proposed, among which Tomlinson-Harashima Preoding (THP) [31, 32] is a well-known low-complexity symbol-by-symbol precoding scheme proposed and widely studied in the context of precoding for Inter-Symbol Interference channels. At high Signal-Noise-Ratio (SNR), THP achieves

the same rate as DPC minus a fixed power penalty called the shaping loss, which is an asymptotic power loss. Although like DPC, THP is also non-linear, THP is far more practical than DPC and is a popular suboptimal approximation to DPC for distributed MIMO.

The essential of THP is simple. The idea is to replicate the constellation along the entire length of the real line to get an infinite extended constellation. Given an information symbol m, the precoding scheme maps the symbol to a representation p which is a duplicate symbol of m that is closest to the interference s. We then transmit the pre-canceling symbol $x = p - s$. Hence, we define a mapping function $q_m(s) = p$. The received signal will be:

$$y = (q_m(s) - s) + s + n = q_m(s) + n \tag{2.22}$$

Where n is the noise. Unlike the naive approach, the power penalty of the mapping does not grow unbounded with s but is fixed based on the constellation size. When the constellation is large, the power penalty is negligible.

An even simpler class of alternative precoding schemes to non-linear precoding (DPC and THP) is referred as linear precoding, among which, the simplest and best known is Zero Forcing Beamforming (ZFBF), consisting of a column-normalized version of the right pseudo-inverse of the channel matrix.

Let us consider a simple scenario where 2 single-antenna APs transmit to 2 single-antenna clients. Let h_{ij}, $i, j \in \{1,2\}$ denote the channel response from AP j to client i, $x_j(t)$ and $y_j(t)$ represents signal sent to and received at antenna j respectively, and $\mathbf{H} = [h_{ij}]$, $i, j \in \{1,2\}$ denote the channel response matrix. In a typical ZFBF system, the precoding is based on the pseudo-inverse of the channel matrix $\mathbf{H}$ to construct a diagonal matrix at the receiver's end. Assuming a simple scenario where the channel matrix $\mathbf{H}$ is invertible, such precoding can be expressed as following.

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \mathbf{H} \begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix} = \mathbf{H}\mathbf{H}^{-1} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \tag{2.23}$$

Different from symbol-by-symbol precoding like THP, ZFBF is dependent on the accurate channel measurements and CSI knowledge at the transmitter. ZFBF provides

better performance than THP when the system works in low SNR (typically not the case for WLAN broadband access, which has high spectral efficiency and operates in high SNR) or the number of clients is significantly larger than the number of AP antennas, and also ZFBF has the flexibility of user selection to approach the full capacity. In the context of multi-user MIMO, ZFBF has been a common implementation and is considered to be an easier yet sufficient alternative to THP.

ZFBF is a simple and straightforward instance of multi-user MIMO precoding, where each stream arrives at different receiving antenna. As introduced, the objective of multi-user MIMO precoding is to enable diversity gain and concurrent transmission, and ZFBF represents an intuitive way of MIMO precoding as beamforming each stream to individual antenna. In practice, sometimes making each stream arriving at different antennas, i.e., spreading stream signal over multiple receiving antenna, is more desirable. By letting each stream be received by a single antenna, when a receiving antenna fails, a whole stream would be lost, hence reducing the reliability of communication. Also, ZFBF may not be the optimal strategy to maximize the overall capacity.

Singular value decomposition (SVD) is a popular multi-user MIMO precoding, that unlike ZFBF, does not beamform each stream to single antenna. Yet, multiple streams could still be differentiated at the receiver by linear transformation. The construction of SVD is based on the $n * m$ channel matrix $\mathbf{H}$. With given $\mathbf{H}$, its SVD can be expressed as following.

$$\mathbf{H} = \mathbf{U}\mathbf{A}\mathbf{V}^H \tag{2.24}$$

Where $\mathbf{U}$ and $\mathbf{V}^H$ represent a $n * n$ unitary matrix and $m * m$ transpose unitary matrix respectively. $\mathbf{A}$ is a $n * m$ rectangular matrix whose diagonal elements are nonnegative real numbers and whose non-diagonal elements are zero. The diagonal elements λ_i of $\mathbf{A}$ are called the *singular values* of $\mathbf{H}$, and the vectors of $\mathbf{U}$ and $\mathbf{V}^H$ are called the *left singular vectors* and *right singular vectors* respectively. We take the conjugate transpose of $\mathbf{H}$ in the following form.

$$\mathbf{H}\mathbf{H}^H = \mathbf{U}\mathbf{A}\mathbf{V}^H\mathbf{V}\mathbf{A}^H\mathbf{U}^H \tag{2.25}$$

Since $\mathbf{U}$ and $\mathbf{V}$ are both unitary, and $\mathbf{A}$ is a diagonal matrix with nonnegative real

diagonal elements, we then have:

$$\mathbf{H}\mathbf{H}^H\mathbf{U}^H = \lambda_i^2\mathbf{U}^H \tag{2.26}$$

Hence, it is straightforward to see that the squared elements λ_i^2 are the eigenvalues of matrix $\mathbf{H}\mathbf{H}^H$ corresponding to eigenvectors which are vectors of $\mathbf{U}^H$. To leverage the property of SVD in multi-user MIMO precoding, SVD can be interpreted as two coordinate transformations. The input signal at the transmitter can be expressed in terms of a coordinate system defined by $\mathbf{V}$ and the output at the receiver side is expressed in terms of a coordinate system defined by $\mathbf{U}$. To be more specific, we precode the input signal with $\mathbf{V}$ and transform the output signal with conjugate transpose of $\mathbf{U}$, hence rewrite the transmission equations as following.

$$\begin{aligned} \mathbf{s} &= \mathbf{V}\mathbf{x} \\ \mathbf{y} &= \mathbf{U}^H\mathbf{s} \end{aligned} \tag{2.27}$$

Then, the resulting input output relationship of such scheme can be as following, for simplicity, assuming the channel is noiseless.

$$\mathbf{y} = \mathbf{U}^H\mathbf{H}\mathbf{V}\mathbf{x} = \mathbf{U}^H\mathbf{U}\mathbf{A}\mathbf{V}^H\mathbf{V}\mathbf{x} = \mathbf{A}\mathbf{x} \tag{2.28}$$

Hence, the resulting equivalent channel enables diagonal channel matrix and concurrent transmission of multiple streams same as ZFBF. However, SVD doesn't force each stream to be received on a single antenna, but instead spreading energy of each stream over multiple receiving antennas. Hence, SVD provides more robust and reliable precoding than ZFBF, improving the overall channel capacity. Essentially, SVD can be regarded as an extension of ZFBF with equivalent channel construction.

An derivation of SVD has been proposed as *geometry mean decomposition* (GMD) [33]. GMD is a similar decomposition technique like SVD, with the difference that the constructed equivalent channel matrix $\mathbf{A}$ is not a full diagonal matrix but a upper triangular matrix, whose diagonal elements are identical. Since the resulting matrix is not fully diagonalized, successive interference cancellation (SIC) or other techniques should be used in addition at the receiver to decode data. However,

GMD holds advantages compared to SVD as having identical diagonal elements for its equivalent channel enables the same modulation being applied to all the data streams. Hence it eliminates the trouble of assigning different modulation for different stream of data.

It should be noted that by not fully diagonalizing the equivalent channel, GMD sacrifices transmission energies that reduces communication performance. The key factor of consideration for choosing between SVD and GMD it the availability of modulation selection for different streams. As it is shown in [34], modulation selection is the essential reason of deferred performance between SVD and GMD. When modulation selection is available, SVD generally outperforms GMD, and vice versa on the contrary. Decomposition based multiple stream beamforming lays the foundation of multi-user MIMO, and its security features are examined in following sections.

2.3.2 Secrecy Gains in MIMO

As introduced, the advantages of MIMO mainly owe to its *diversity gain* and *multiplexing gain*. Those benefits not only lead to increased channel capacity, but also bring chances for enhanced secrecy. Let us describe a simple model of MIMO wiretap scenario, which contains three terminals, i.e., the transmitter, the legitimate receiver, and the eavesdropper. The terminals are equipped with n_t, n_r, n_e as amount of antennas respectively. The signals transmission is such system can be expressed as:

$$\begin{aligned} \mathbf{y} &= \mathbf{H}_r\mathbf{x} + \mathbf{n}_r \\ \mathbf{z} &= \mathbf{H}_e\mathbf{x} + \mathbf{n}_e \end{aligned} \tag{2.29}$$

Where $\mathbf{n}_r$ and $\mathbf{n}_e$ represent Gaussian white noise of the main channel and the eavesdropper's channel respectively, and $\mathbf{H}_r$ and $\mathbf{H}_e$ represent the MIMO transmission matrices.

To examine the advantage of MIMO, it should be noted that the diversity and multiplexing gain are closely related to channel conditions. The analysis of the trade-off of diversity and multiplexing gains in MIMO is first established by Zheng and Tse [25]. Both gains have been characterized and quantified in an asymptotic manner, and are functions of the channel SNR.

Similarly, in [35], the trade-off of diversity and multiplexing gains under secrecy

constraints has been discussed. Close to the conventional term, the secrecy multiplexing gain is defined as:

$$\lim_{SNR\to\infty} \frac{R(SNR)}{\log SNR} \approx r \tag{2.30}$$

Where R represents the achievable rate at the transmitter. It should be noted that the secrecy multiplexing gain r is the asymptotic value capturing the high SNR behavior of the channel.

Similarly the secrecy diversity gain of a MIMO channel is defined. As introduced, the secrecy diversity gain is also commonly referred to as the secrecy degree of freedom, which is useful the access multi-user network. The gain can be expressed as:

$$\lim_{SNR\to\infty} \frac{\log P_e(SNR)}{\log SNR} \approx -d \tag{2.31}$$

Where P_e represents the error probability. Hence, the secrecy diversity gain captures the asymptotic decrease rate of error probability with increasing SNR, under secrecy constraints. The notion of secrecy diversity and multiplexing tradeoff (DMT) is proposed [35] to access the relationship between those two metrics.

Recently, apart from accessing the asymptotic gains of MIMO secrecy, research works [36, 37, 38, 39, 40, 41, 42] have been proposed to analyse the absolute secrecy capacity of MIMO. For instance, Shafiee et al [36] started from a simple scenario of 2-2-1 MIMO wiretap channel, where both the transmitter and receiver are equipped with two antennas, and the eavesdropper has one antenna. In the assessment of MIMO secrecy, the power constraints are usually considered. According to the signals transmission expression in equation 2.29, the signal X should follow average power constraint P as:

$$\frac{1}{n}\sum_{i=1}^{n} E[x_i^* x_i] \leq P \tag{2.32}$$

The secrecy capacity C is then defined as the maximum achievable rate that can be correctly transmitted with while the eavesdropper is essentially no better informed about the transmitted information after observing the received signal than it was before. The work [36] considered a generalized case where the eavesdropper's channel is not degraded. Under the power constraint expressed in equation 2.32, the maximum

achievable rate could be expressed as:

$$C = \max_{0 \leq \mathbf{Q} \leq \mathbf{P}} (log\ det(\mathbf{I}_r + \mathbf{H}_r^H \mathbf{Q} \mathbf{H}_r) - log(\mathbf{I}_e + \mathbf{H}_e^H \mathbf{Q} \mathbf{H}_e)) \tag{2.33}$$

Where $\mathbf{I}$ represents the identity matrix with the order the transmitted signal. Hence, to achieve secrecy capacity in this situation becomes an equivalent generalized eigenvalue problem. Here it should be noted that, in early stage wiretap channel model, a degraded eavesdropper's channel is assumed. To be more specific, the transmitted signal and received signals at the legitimate receiver and the eavesdropper should form a Markov chain, i.e., $X \rightarrow Y \rightarrow Z$. Although there are discussions about loosing the assumption with auxiliary prefix signal, the degraded assumption is still an important factor in the derivation of secrecy capacity. However, as MIMO networking in its general form belongs to the class of non-degraded wiretap channels, the problem of characterizing the secrecy capacity of MIMO wiretap channel has long been a open problem.

In [37], the authors considered another special case of 2-1-2 MIMO wiretap channel, similar to the 2-2-1 model, where the transmitter and eavesdropper are equipped with two antennas, and the legitimate receiver has one antenna. The special cases of 2-2-1 and 2-1-2 wiretap channel can be further generalized into multi-input (transmitter), multi-output (receiver), single-eavesdropper (MIMOSE) model, and multi-input (transmitter), single-output (receiver), multi-eavesdropper (MISOME) mode, which are closely related to general MIMO wiretap channel, and fit in research of multi-antenna secrecy communication.

In [41], the secrecy capacity of general MISOME wiretap channel is derived. It should be noted, as introduced, the secrecy capacity of multiple antenna system is closely related to power and SNR constraints. Also, the transmitting power holds positive correlation with the resulting SNR. Hence they could be treated as connected factors for asymptotic analysis. The secrecy of general MISOME wiretap channel can be expressed as:

$$C = log \lambda_{MAX} (\mathbf{I}_r + \mathbf{h}_r^H \mathbf{Q} \mathbf{h}_r, \mathbf{I}_e + \mathbf{H}_e^H \mathbf{Q} \mathbf{H}_e) \tag{2.34}$$

Where λ_{MAX} represents the largest generalized eigenvalue of its argument matrices. The result could be further accessed asymptotically with high or low SNR/power scenario.

The general Multiple Input Multiple Output Multiple Eavesdropper (MIMOME) has been discussed in [38, 39, 40, 42]. In [39, 40, 42], general forms of MIMOMO secrecy capacity are developed independently. The expressions carry a similar form as shown in equation 2.33. It is straightforward to see that such characterization will involve optimization problem as to search the optimal covariance $\mathbf{Q}$ that satisfies the power constraint. A closed-form characterization of MIMO secrecy capacity is more desirable to examine the secrecy capacity of MIMO.

In [40], Liu and Shamai reexamined the MIMO wiretap channel and proposed an alternative proof of the MIMO secrecy capacity. Specifically, The authors suggested that MIMO secrecy capacity can be achieved with the parallel Gaussian wiretap channel model, which can be seen as a special case of the vector Gaussian wiretap channel with diagonal noise covariance matrices. The optimal transmission strategy could be to transmit independently over the subchannels for which the legitimate receiver has better conditions than the eavesdropper.

Bustin et al [38] further exploited the fundamental relationship between minimum mean-squared error (MMSE) and mutual information to provide a closed-form expression for the optimal input covariance $\mathbf{Q}$ that achieves the MIMO wiretap channel secrecy capacity, again under an input power-covariance constraint. A closed-form expression of MIMO secrecy capacity is then presented. More precisely, it is shown that with certain power constraint $\mathbf{P}$, the MIMO secrecy capacity can be given by following.

$$C(P) = \sum_{i=1} \alpha_i \tag{2.35}$$

Where α represents all the generalized eigenvalues of the following matrix pencil that is greater than one.

$$(\mathbf{P}^{1/2}\mathbf{H}_r^H\mathbf{H}_r\mathbf{P}^{1/2} + \mathbf{I}_r, \mathbf{P}^{1/2}\mathbf{H}_e^H\mathbf{H}_e\mathbf{P}^{1/2} + \mathbf{I}_e) \tag{2.36}$$

Yet, although equations 2.35 and 2.36 give closed-form characterization of the MIMO secrecy capacity itself, the model can not give the optimal transmission power matrix $\mathbf{Q}$, which is desirable to design practical secure communication. It should be also noted that the above mentioned characterizations of MIMO secrecy capacity are

commonly based on the matrix power constraint **P** which limits the per-antenna power and transmit correlation structure on the transmitter side. On the other hand, the average power constraint is much less restrictive, and provides considerable additional flexibility in increasing the secrecy rate of the MIMO wiretap channel. Using average power constraint has been proposed in [43]. The average power constraint P can be expressed as following.

$$\frac{1}{n}\sum_{i=1}^{n}\|x_i\|^2 \leq P \tag{2.37}$$

With such average constraint, MIMO secrecy capacity is then given by searching optimal matrix power constraint **P** based on the average power constraint, expressed as following.

$$C(P) = \max_{\mathbf{P}\geq 0, Tr(\mathbf{P})\leq P} C(\mathbf{P}) \tag{2.38}$$

Yet, it should be noted that the average power constraint is less practical than matrix power constraint. In reality, the transmission power is constrained that any instantaneous power allocation could not break the limit. A system satisfying power constraint on average may break the physical power limitation at certain point. Overall, the problem of characterizing MIMO secrecy capacity is actually to find optimal transmission matrix satisfying power constraint. For practicality, closed-form and low complexity algorithms for searching such problem becomes a fundamental topic of physical layer security in MIMO, which is elaborated in the next chapter.

2.4 Secret Key Agreement with Common Randomness

There is an alternative intuition about accessing secrecy property of the wireless networking. The major branch initialized by Wyner is mainly focus on the code design and signal processing perspective, as how to create enough equivocation at the eavesdropper, such that private messages will not be compromised. The scope of those research is to develop secure wireless communication without the establishment of secret keys, which differs from the traditional thought on networking security.

Another trend in the information theoretic security comes closer to the traditional key-based security means. Maurer [44] first proposed the idea of developing secret keys between the transmitter and the legitimate receiver via public (insecure) channel.

The question that Maurer tried to address, which inspired many following discussion, is that how to leverage information theoretic properties of the channel to establish secret keys via public channels with presence of eavesdropper. Maurer argued that Wyner's degraded model as being too unrealistic, and instead proposed a secret key agreement scheme that could be implemented over a noiseless but authenticated and publicly observable two-way channel.

The key ideas of Maurer's strategy are *information reconciliation* and *privacy amplification*. The information reconciliation phase is aimed at generating an identical random sequence between the transmitter and the legitimate receiver by exploiting a public discussion channel. The privacy amplification, on the other hand, extracts a secret key from the initial random sequence agreed to by two terminals in the preceding information reconciliation stage. In other words, after public discussion based on correlated randomness in the information reconciliation phase, privacy amplification reduces an initial piece of random nature into a functional secret key. The idea of privacy amplification was first discussed by Bennett et al [45].

To be more specific, in [45], the legitimate users are supposed to have two different channels at their disposal, which are an imperfect private channel and a perfect public channel. The private channel is imperfect regarding both reliability and security, as there is possibility of transmission error and information leakage. On the other hand, The public channel holds no transmission error, but is open to the eavesdropper. Bennett et al [45] proposed a protocol based on the public channel that reconciles the initial random sequence shared through the private channel. Two possible problems regarding the initial sequence were raised. First, the sequence could have been corrupted by tampering and channel noise of the private channel. Second, part of the sequence may be wiretapped by the eavesdropper. Hence, the protocol is to repair corrupted sequence, and tailor the original sequence into a shorter version, upon which the eavesdropper obtains no information.

Maurer further extended the scope of Bennett's work by considering the initial sequence and privacy amplification as a whole, and both procedures are processed with the public channel. It should be noted that, a pure passive eavesdropper is assumed in Maurer's scheme, in other words, the eavesdropper could not insert fraudulent messages nor modify messages on this public channel. Meanwhile, the public channel is

assumed to be two-way feedback channel, and both legitimate users can transmit secrecy message.

Maurer's work is based on previous discussion on the wiretap channel, and also requires randomization to create equivocation at the eavesdropper. Yet different from the original wiretap channel model, both the transmitter and receiver can send random sequences to each other, and reconciles those sequences to a shared message, which will be used to construct key. In order to protect the secrecy of the shared message, yet ensuring the mutual understanding between the legitimate terminals, the sequences send by each terminals should be stochastically constructed and correlated to each other, hence creating two different but correlated streams. This shared randomization framework is referred to as *common randomness*, and is the essential feature of physical layer key agreement. Common randomness have been proved highly beneficial for communication security [46]. The resulting shared sequence derived from common randomness is often addressed as *common message*, or *secret seed*, in the context of physical layer security.

Maurer [44] proposed a simple system where the legitimate users transmit random bits sequentially, and the random variable generated by both user follow the same distribution. To be specific, let us consider a generalized procedure in such systems, where the legitimate users have known random variable x and y respectively, and communicate message $\mathbf{c}$ to each other over the public channel one by one. The eavesdropper then makes observation z of the communication. For simplicity, we don't consider full-duplex radio [47] in this case, and assume the transmitter sends messages at odd steps and the legitimate receiver sends messages at even steps. Both legitimate users try to generate secret key $\mathbf{k}$ and $\mathbf{k'}$ based on what they receive. It is straightforward to see that the objective of the protocol is to maximize $H(\mathbf{k})$ and $H(\mathbf{k}|\mathbf{z})$, meanwhile minimizing $H(\mathbf{k}|\mathbf{k'})$. Also, the system is required that:

$$\begin{aligned} P(\mathbf{k} \neq \mathbf{k'}) &\leq \varepsilon \\ I(\mathbf{k};\, \mathbf{cz}) &\leq \gamma \end{aligned} \tag{2.39}$$

Where, ε and γ represent thresholds of the reliability and secrecy of the established key. Moreover, the general constraint of the equivocation of the established key is derived as:

$$H(\mathbf{k}) \leq I(\mathbf{x};\, \mathbf{y}|\mathbf{z}) + H(\mathbf{k}|\mathbf{k'}) + I(\mathbf{k};\, \mathbf{cz}) \tag{2.40}$$

An important factor for the performance of Maurer's proposed protocol is the *secret key rate*, which quantifies the rate at which legitimate users can agree upon the established key sequence by exchanging messages over the public channel. Similar to its counterparts in the original wiretap channel model, the *secret key rate* is defined as a maximum achievable rate R that satisfies:

$$R - \varepsilon \leq \frac{H(\mathbf{k})}{n} \tag{2.41}$$

It is straightforward to see that the assessment of secret key rate is similar to that of the secrecy capacity, where the secrecy of the shared key $\mathbf{k}$ is considered instead of the message $\mathbf{m}$ itself. The key difference is that, in the original derivation of the wiretap channel model, the message $\mathbf{m}$ is transmitted only from the transmitter to the legitimate receiver in a one-way manner, while the $\mathbf{k}$ is supposed to be established via a feedback channel through two-way communication. This difference leads to the distinct behaviors of each metric. The use of two-way feedback channel makes secrecy possible without the assumption of a degraded channel.

We denote the secret key rate in such system as C_k, representing the maximum rate at which the legitimate users can agree on the secret key K ensuring the eavesdropper obtains information under the threshold. Then the upper and lower bounds of C_k can be expressed as:

$$\begin{aligned} C_k &\leq \min(I(\mathbf{x};\, \mathbf{y}), I(\mathbf{x};\, \mathbf{y}|\mathbf{z})) \\ C_k &\geq \max(I(\mathbf{y};\, \mathbf{x}) - I(\mathbf{z};\, \mathbf{x}), I(\mathbf{x};\, \mathbf{y}) - I(\mathbf{z};\, \mathbf{y})) \end{aligned} \tag{2.42}$$

Maurer's work is of great significance in the research of information theoretic security, as it extended the wiretap channel into key-based system with two-way communication. The most notable conclusions for this work is following:

- **Common randomness**: fundamentally, Maurer's work is an extension of the wiretap channel model with two-way channel, and shares the same underpinning mechanics that requires stochastic coding and randomization. However, different from conventional wiretap channel, the legitimate users are allowed for two-way

communication over the main channel, enabling them to share an agreed random sequence. Referred to as common randomness, this framework defers from the unitary randomness used in conventional wiretap channel model, which only relies on the transmitter. Common randomness is the major distinction of key-based information theoretic security, and has been proved beneficial.

- **Interaction is better than one-way transmission**: The conventional information theoretic security founded by Wyner began with a scenario where the eavesdropper can only receive the signal through an cascaded degraded channel which reduces the capacity of the eavesdropper's channel. Extended to more general and sophisticated scenarios, keyless secrecy communication holds great potential, but is also challenged with practicality. On the other hand, Maurer proposed the benefits of interaction, assuming both legitimate users being active and employing two-way channel rather than only one-way transmission, secrecy can be achieved with loosened conditions.

- **Channel errors can be good**: Contradicting to the traditional thought of channel conditions, Maurer's work suggested that noisy channels, which bring probability of error, need not be first converted into error-free channels from a security perspective. The design strategy of cryptographic protocol based on error-free channels lead to Shannon's analysis of perfect secrecy [9], which could only be achieved with an impractically huge shared secret key. Instead, cryptographic coding and error-control coding should be combined, resulting in a system achieving virtually perfect secrecy, by leveraging the common randomness/noisiness of the imperfect channel.

As mentioned, key-based information security requires both legitimate terminals to send correlated random sequences in order to achieve common randomness. In practice, two different models can be used to describe the procedure of obtaining common randomness.

- **Source-type model**: the two terminals observe different but correlated variations of same source of randomness, and reconciles their observation and reach key agreement through two-way communication.

- **Channel-type model**: the transmitter is in control of the randomization source, and transmits random encoded message directly to the receiver. The receiver observes an the output signal which is subject to channel errors, but correlated to the input signal. The input and output signals are then reconciled through channel feedback.

It is straightforward to see that the channel-type model is very similar to keyless security, with the notable difference that here the receiver is able to feedback its observation to the transmitter.

Maurer's work [44] on key-based wiretap channel model represents another major branch of information theoretic security. similar idea have been developed independently by Ahlswede and Csiszar [46]. In [48], generalized privacy amplification by public channel was discussed, regarding both simple wiretap channel and broadcast channels.

However, there were still limitations of the information theoretic secret key establishment, which reduces the practicality of such scheme.

- **Against active eavesdropper**: As mentioned, the original information theoretic key agreement assumes a scenario with a pure passive eavesdropper, which could not modify or corrupt the communication between the legitimate users via the public channel. Although in the context of information theoretic security, passive eavesdropper is the major concern, such assumption greatly limits the practicality of information theoretic secret key. Since such schemes are based on shared keys, which are generated upon the information exchange between legitimate users itself, risks facing active eavesdropper main cause all following communications to be compromised.

- **Need of randomization**: Similar to the conventional keyless schemes, key-based information theoretic security also requires both the legitimate users to have randomization abilities for noisy main channel. For instance, in Maurer's work [44] both legitimate users are supposed to be able to extend their known random variables $\mathbf{y}$ and $\mathbf{y}$, respectively, by random bits that are statistically independent of $\mathbf{y}$, $\mathbf{y}$ and $\mathbf{z}$.

- **Dependency on channel variations**: key-based information theoretic security leverages the variations and noise of imperfect channels. Hence, the performance of such schemes, especially considering the secret key rate, is largely dependent on the level of 'imperfectness' of the channel. Hence, dynamics and errors of channel may be beneficial to the key agreement. On the other hand, the secret key capacity may not be sufficient in more static conditions.

To address the above issues, various following works have been proposed to extend Maurer's results [44]. In [49], the generation of secret key by two legitimate terminals with an additional helper terminal is considered. Each legitimate terminal observes a different component of a discrete memoryless multiple source and The helper aids the users by transmitting information to them over a noiseless public channel subject to a rate constraint. The secret key capacity in such system have been derived. Also, Csiszar and Narayan also showed that randomization is not necessary for the achievability of the secret key capacity. This conclusion greatly improve the adaptability of information theoretic key agreement.

In [50, 51, 52], Maurer and Wolf extended the original analysis of physical layer secret key agreement to the situation where an active eavesdropper is presented. The results showed that secure secret key agreement facing active adversary is possible in certain conditions on the proportion of the random sequence affected by the eavesdropper. Under certain thresholds, the information reconciliation will have same capacity as with passive adversary, otherwise such key agreement protocols wouldn't be feasible. In [53], Yakovlev et al further extends Maurer and Wolf's results into a variety of key agreement protocols, and both asymptotic and nonasymptotic cases are discussed. It is shown that in some cases the nonasymptotic key rate is not so far from the asymptotic one whenever the lengths of the transmitted strings are of the order of thousands of bits.

Comparable to the conventional information theoretic security, key agreement schemes have also been extended to more sophisticated channels. In [54], the secret key capacity of the multiple-access channel was examined. It is shown that the public channel, even in forward direction only, in combination with the multiple-access channel can provide more security in contrast to the case of the broadcast channel.

In [55], Salimi et al discussed a scenario where the transmitter tries to establish secret keys with two legitimate receivers through an interference channel, and there is

also an unlimited noisy feedback channel from the receivers to the transmitter. Inner bound of the secret key capacity in such system is derived.

In the recent work of Csiszar and Narayan [56], the authors further discussed the physical layer secret key establishment in modern set, where arbitrary number of terminals are considered. The conventional public and private key schemes are compared with physical layer key establishment, denoted as wiretap secret key in the work, regarding the key generation rate.

Not surprisingly, information theoretic key agreement has been extended to multiple antenna systems as well. As introduced before, key-based schemes holds the same problem model as the keyless ones, with the difference that feedback channel is enabled, and secrecy capacity is used to establish secret key, instead of exchanging messages directly. In [57], Renna et al analyzed the fundamental limits of secret-key agreement over MIMO quasi-static fading, or slow fading channels. In the low-power and high-power regimes, closed-form expressions for secret key capacity was established. The work also proved that the optimal strategy for low power cases signaling does not require any knowledge of the eavesdropper's channel at the transmitter and that waterfilling on the main channel achieves secret-key capacity in the high-power regime. This result highlighted a fundamental difference between information theoretic secret key agreement problem and keyless coding protocols. As mentioned, for keyless schemes, the secrecy capacity of the MIMO Gaussian wiretap channel depends strongly on the eavesdropper channel conditions. The secret key capacity of MIMO wiretap channel with Gaussian inputs are derived as:

$$C_k = \max(log\ det(\mathbf{I}_r + \mathbf{H}_x^H \mathbf{Q} \mathbf{H}_x) - log\ det(\mathbf{I}_e + \mathbf{H}_e^H \mathbf{Q} \mathbf{H}_e)) \tag{2.43}$$

Where $\mathbf{H}_x^H \mathbf{H}_x = \mathbf{H}_r^H \mathbf{H}_r + \mathbf{H}_e^H \mathbf{H}_e$ represents a compound equivalent channel.

As mentioned, similar to conventional information theoretic security based on randomization at the transmitter, key-based security requires randomization source available at both sides of the main channel. Then design of information reconciliation and privacy amplification is then a similar problem to the original wiretap channel. However, the reciprocal nature of wireless communication brings an alternative way of obtaining common randomness, which leads to a notable breakthrough in key-based

physical layer security.

Since wireless channel is reciprocal and spatially distinct, the legitimate users could perceive a channel independent from the eavesdropper's channel. In practical wireless networking, where multi-path, fading and mobility are presented, the channel conditions of the main channel is unpredictable and unique to the adversary, and can be regarded as common randomness. An important evolution of key-based physical layer security is obtaining common randomness from the wireless channel itself, which can greatly simplify the system, and waive the requirement of available randomization at terminals.

Key-based information theoretic security is not entirely different from but closely related to keyless security. Keyless security can in fact be deployed in parallel with key-based security and facilitate it. Compare to keyless security, key-based security has better flexibility and less requirements on channel conditions. The details of practical key-based security designs are elaborated in Chapter 4.

Chapter 3

Keyless Physical Layer Security

The information theoretic analysis of wiretap channel, first proposed by Wyner [8], showed that there exist codes, often referred to as wiretap codes, that allow both reliable and confidential communication between legitimate users in the presence of an eavesdropper. In this chapter, we investigate the existing knowledge of how to construct the actual wiretap codes to enable such keyless secure communication, where different channel conditions and models are considered, such as fading and MIMO respectively. Apart from wiretap code design, recent developments in keyless security also propose other techniques that enhance transmission secrecy. For instance, the exploitation of user cooperation has been recognized as an important method to improve the robustness of secure transmissions. Meanwhile, in multi-antenna systems, transmission antenna selection can be used to facilitate communication security. Those diverse techniques are also covered in this chapter present a comprehensive picture of the research on keyless communication security strategies.

3.1 Wiretap Code Design to Achieve Secrecy

Conventional information theoretic security, established with the wiretap channel model, is to design randomized code scheme where the information is hidden in the additional noise seen by the eavesdropper. In a stochastic code protocol, each message is mapped to many codewords, thus inducing maximal equivocation at the eavesdropper.

3.1.1 Coset Codes and Nested Codes

It took a decade after the publication of the original wiretap channel for an explicit wiretap coding scheme [12] to be proposed. In this work, Ozarow and Wyner considered a situation where the main channel is a binary noiseless channel and the wiretapper channel is a binary bounded erasure channel, which was referred to as wiretap channel type-2. A binary erasure channel (BEC) is a simple channel model which is widely used in information theory analysis, and its generalization in the packet erasure channel has caused much interest since the introduction of the internet, because the internet mainly considers packets instead of bits, and can cause packet loss (erasure), hence exemplifying packet erasure channel. BEC also has been widely discussed in the context of information theoretic security as a common model for code design. In a BEC, the transmitter send a bit to the receiver, which has certain probability P_e to be erased, which is illustrated in figure 3.1.

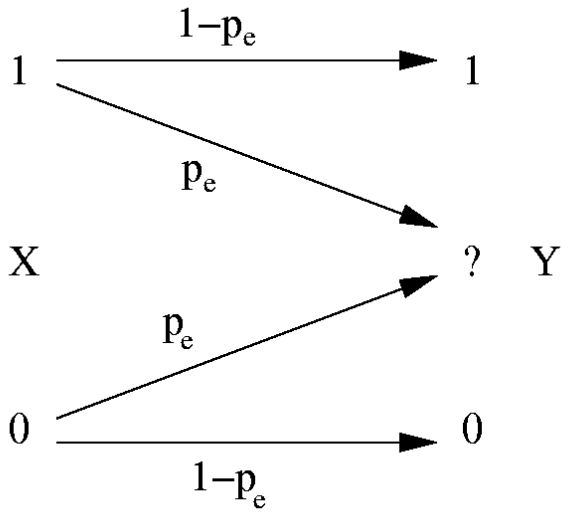

Figure 3.1: Diagram of a binary erasure channel

The proposed code scheme for type-2 wiretap channel is based on coset coding, which is first proposed in [58]. The authors have made the important observation that the signal constellation should be regarded as a finite set of points taken from an infinite lattice, and that the partitioning of the constellation into subsets corresponds to the partitioning of that lattice into a sublattice and its cosets. This lattice/coset language is both illuminating and powerful, and widely cited ever since.

Here we briefly introduce the concept of group, lattice and coset. A group is an algebraic structure consisting of a set of elements together with an operation $*$ that combines any two elements to form a third element. The operation also needs to satisfy the group axioms, which are namely closure, associativity, identity and invertibility. Closure is met when the operation on two elements in the group results another element in the same group. Associativity means when multiple operations $*$, the order in which the operations are performed does not matter as long as the sequence of the operands is not changed. Identity refers to an existing element in a group that leaves other elements unchanged when combined with them by the operation $*$. Invertibility means that for every element in the group, there exists an element that the combination of the two is the identity element.

Although the group theory can be complicated, its application in coding is straightforward. In digital communication, the group of interest in the simple additive binary vector space donated as $\{0,1\}^n$, meaning the operation $*$ for the group is bitwise addition. We define the lattice Λ as a discrete subgroup of $\{0,1\}^n$, where discrete means that there exists an $\varepsilon > 0$, that for any two distinct elements x and y in Λ, $\|x-y\| > \varepsilon$.

Using the lattice Λ as a subgroup, if there exists an element g in the group $\{0,1\}^n$, that $g*\Lambda$ or $\Lambda*g$ is still a subgroup, then those combinations are the left coset or right coset of Λ respectively. In our case, the left and right cosets coincide and can be treated as the same.

With those notions in mind, let us revise the problem of a coding scheme with coding ratio $\{n,k\}$, i.e., messages containing k bits of information is encoded into n bits signal. In the perspective of coset coding, we can regard the coding scheme as a process of partitioning the group $\{0,1\}^n$ into 2^{n-k} disjoint subsets, each of size 2^k. To be more specific, lets assume a coding scheme $\boldsymbol{C}$ with a parity check matrix $\mathbf{H}$. The transmitter simply transmits coded signal $\mathbf{x}^n$ that satisfies:

$$\mathbf{H}\mathbf{x}^n = \mathbf{m}^k \tag{3.1}$$

Where $\mathbf{m}^k$ is the intended message. To enable equivocation, the transmitter chooses uniformly from all the 2^{n-k} cosets. In other words, one of the 2^{n-k} solutions to equation 3.1 is chosen with the randomization source at the transmitter. On decoding, the

legitimate receiver simply decode the message by multiplying the parity check matrix $\mathbf{H}$ with estimated $\mathbf{x}$, assuming a noiseless main channel. On the other hand, the randomization of coset $\mathbf{x}^n$ creates equivocation at the eavesdropper, and with the assumption that the eavesdropper's channel is degraded and noisier, a positive secrecy capacity can be achieved.

The coset coding scheme for type-2 wiretap channel is an example of practical stochastic code design for information theoretic security. Coset coding is also referred to as *syndrome coding* in the context of coding theory, where considering equation 3.1, message $\mathbf{m}^k$ is called a syndrome of transmission vector $\mathbf{x}^n$. Works have been proposed to extend such schemes to more sophisticated channels. In the context of wiretap channel code design, coset code and lattice code are often interchangeable terms and hence are treated as a single framework in this thesis.

Revising the above coset coding scheme, the codeword $\mathbf{x}^n$ is randomly chosen from 2^{n-k} possible solutions of the encoding equation 3.1. The process of random choosing from 2^{n-k} possible codeword is equivalent to an auxiliary message $\mathbf{m}^0$ with $n-k$ bits. Hence we can regard the coset coding as a *nested code* structure, where the transmitter encode the message $\mathbf{m}$ together with a randomized auxiliary message $\mathbf{m}^0$. The codebook is then determined by both messages. In encoding, a subblock of the codebook is first selected based on message $\mathbf{m}$, then the codeword $\mathbf{x}^n$ is chosen based on $\mathbf{m}^0$. In other words, in order to leverage randomization, the design of wiretap code needs to employ a nested code structure, where the codebook is composed of subcodebooks, each of which can be used for encoding alone.

Coset coding can be regarded as a simple implementation of a nested code. Nested codes can be constructed recursively, meaning codes with nested structure can be further inserted in more complicated structure, hence providing further functionality including improved secrecy and error-control. As mentioned, simple coset coding shown in equation 3.1 assumes an error-free noiseless main channel. With further nested code structure, such assumption can be waived by enabling error-control on the main channel, hence improving the practicality of wiretap code. In [59], He and Yener proposed to use nested lattice code to achieve large secrecy rate in Gaussian wiretap channel. In [60], Liu et al discussed the strategy of using nested lattice code for the original setting of type-2 wiretap channel. In [61], Belfiore and Oggier proposed a nested lattice

code design for the Rayleigh fading wiretap channel.

The nested code structure and the randomization among multiple codewords are two key features for wiretap codes design, and have been widely used and discussed in the area of information theoretic security. As introduced in the type-2 wiretap channel, the model for such design assumes a noisier eavesdropper's channel compared to the main channel. For instance, for type-2 wiretap channel, an important performance metric for wiretap code is the *minimal erased bits*, which is defined as the minimal number of erased bits at the eavesdropper needed to ensure the secrecy of the message. Such performance is dependent on the individual code design.

3.1.2 Low-Density Rarity-Check Codes and Puncturing

Although the development of dedicated wiretap code leads to many notable results, and have been experiencing an resurgence in recent years, the explicit construction of wiretap code is still largely a topic in its early stage. It is naturally to consider the conventional coding methods and access their secrecy performances. Particularly, error-correction codes have been widely proposed for secrecy purposes, and the possibility of integrating them with dedicated wiretap codes was discussed.

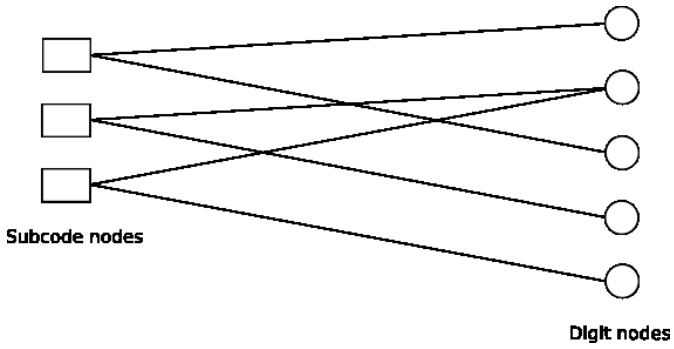

Figure 3.2: Illustration of example Tanner graph

A notable development in using conventional error-correction codes for information theoretic security is Low-Density Rarity-Check (LDPC) codes. Application of LDPC to wiretap channel was first proposed by Thangaraj et al [62]. LDPC codes are graph-based linear block codes first proposed by Gallager in his thesis [63], which can be characterized by a sparse parity-check matrix H that can be equivalently described by a bipartite graph known as a Tanner graph, shown in figure 3.2. The graph is formed by associating each column with a variable node (digit node), each row with a check

node (subcode node) , and connecting the ith check node to the jth variable node by an edge if and only if the entry h_{ij} in H is equal to one. The number of edges incoming on a node is called its degree, and the degree distribution pair (λ, ρ) of H can be defined as following:

$$\begin{aligned} \lambda(x) &= \sum_{i \geq 1} \lambda_i x^{i-1} \\ \rho(x) &= \sum_{i \geq 1} \rho_i x^{i-1} \end{aligned} \tag{3.2}$$

Where λ_i and ρ_i represent the probabilities that a randomly chosen edge in the Tanner graph of the code is incident on a variable and check node of degree i, respectively. LDPC can also be characterized by parity check matrix containing mostly 0's and few 1's, which is where its name low-density comes from, i.e., the parity check matrix in consideration should be sparse. Tanner graph gives intuitive representation of LDPC coding scheme. To be more specific, we can regard that each variable node represents an coordinate in the codewords, and each line represents a codewords. On the other hand, the check nodes represents parity check constraints, and for lines connected to the same check node, their variables must sum, modulo two, to zero. It is straightforward to see that for a parity check matrix in a coding scheme with coding ratio $\{n,k\}$, there should be n variable nodes and k check nodes in the Tanner graph.

To construct LDPC codes, we usually consider deriving the codeword generation matrix $\mathbf{G}$ from the parity check matrix $\mathbf{H}$. Common methods of doing including transforming the parity check matrix $\mathbf{H}$ with Gaussian elimination such that the matrix becomes an combination of an identity matrix and a submatrix $\mathbf{P}$, expressed as following:

$$\mathbf{H} = [\mathbf{P}^T | \mathbf{I}_{n-k}] \tag{3.3}$$

Then, the generation matrix $\mathbf{G}$ can be derived with the submatrix $\mathbf{P}$ as:

$$\mathbf{G} = [\mathbf{I}^k | \mathbf{P}] \tag{3.4}$$

The message $\mathbf{x}$ will be encoded into message $\mathbf{m}$ by multiplying the generation matrix as:

$$\mathbf{x} = \mathbf{G}\mathbf{m} \tag{3.5}$$

The above illustrates a general framework of LDPC encoding. To speed up the

decoding process to achieve linear time, many specific construction methods of LDPC codes have been proposed. For instance, Repeat Accumulate code [64] has been widely discussed a subclass of LDPC codes with fast encoder structure and graph representation.

The algorithm used to decode LDPC codes was discovered independently several times and naturally comes under different names. The most common ones are the belief propagation algorithm, which is an iterative the message passing algorithm, which is closely related to the distribution pair (λ, ρ) and graph representation. In an iterative message passing algorithm, each round messages are passed from message nodes to check nodes, then from check nodes back to message nodes. The messages from message nodes to check nodes are computed based on the observed value of the message node and some of the messages passed from the neighboring check nodes to that message node. In belief propagation, messages passed along the edges in this algorithm are probabilities, or beliefs.

To be more specific, considering the Tanner graph shown in figure 3.2, each round, message passed from a variable node to a check node is the probability that the variable node has a certain value given the observed value of that message node, and all the values communicated back to variable in previous rounds from check nodes. On the other hand, the message passed from check nodes to variable nodes are the probabilities that a variable has a certain value given all the messages passed to the constant variable in the previous rounds, assuming messages from the other variable nodes are correct. The algorithm will finish when messages received and sent at the variable nodes coincide.

Baed on the distribution pair (λ, ρ), an error threshold $\varepsilon(\lambda, \rho)$ can be defined, that if the error probability of the channel is under the threshold, all error can be corrected. The threshold of LDPC codes are often referred to as the *reliability threshold* of such code scheme. On the other hand, as introduced, in the setting of wiretap code, imperfect channel conditions and error probability are beneficial for communication secrecy, and the minimum error probability that required for creating equivocation at the eavesdropper is called the *security threshold*. In the context of wiretap code design, the difference between the reliability threshold and the security threshold are often termed as the *security gap*.

To enable further protection, LDPC can be used in a nested structure. In [62] for

instance, Thangaraj et al proposed to use dual of LDPC codes. To be more precise, consider a (n,k) LDPC code $\boldsymbol{C}$ with $(n-k,n)$ parity check matrix $\mathbf{H}$, we can derive $\boldsymbol{C}^d$ as the dual of $\boldsymbol{C}$, which has $\mathbf{H}$ as its generator matrix. The random coset of $\boldsymbol{C}^d$ can then be used as the nested code.

Another extension to LDPC codes for secrecy is the two-edge LDPC codes proposed in [65]. The proposed scheme follows the same intuition as to create nested code structure. Two-edge LDPC codes are characterized with parity check matrix that can be partitioned into two parts, expressed as following:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} \tag{3.6}$$

Where $\mathbf{H}$ is a $(n-k)*n$ matrix, and $\mathbf{H}_1$ is a $(n-k_1)*n$ matrix, $k_1 > k$. We define the code design $\boldsymbol{C}_1$ as the scheme represented by the parity check matrix $\mathbf{H}_1$. Thus, it is straightforward to see that the coding scheme $\boldsymbol{C}$ represented by matrix $\mathbf{H}$ is a subcode of $\boldsymbol{C}_1$, and the cosets of $\boldsymbol{C}$ in $\boldsymbol{C}_1$ forms a partition. To transmit $k_1 - k$ bits of message $\mathbf{m}$, the signal $\mathbf{x}$ is randomly chosen such that:

$$\mathbf{x}\mathbf{H}^T = \begin{bmatrix} 0 \\ \mathbf{m} \end{bmatrix} \tag{3.7}$$

It should be noted that simple coset coding introduced previously can be regarded as an special case when $\boldsymbol{C}_1$ is then entire space of GF(2), or $\{0,1\}^n$. It is straightforward to see that two-edge LDPC provides a further nested structure compared to simple coset coding. To be more precise, two-edge LDPC code partitions the parity check matrix into two parts, one of them provides regular error check on the codeword $\mathbf{x}$, and the other is a simple coset coding with random message $\mathbf{m}^0$. Since the LDPC code of single-edge distributions may have degree two variable, the resulting partitioned matrix $\mathbf{H}_1$ may have degree one variable node, hence there is no error-check functionality on the main channel. Hence, in such code designs, the main channel is supposed to be noiseless.

Another solution of adapting LDPC codes for secrecy communication is puncturing LDPC. Puncturing is the process of removing some of the parity bits after encoding with an error-correction code. This has the same effect as encoding with an error-

correction code with a higher rate, or less redundancy. In puncturing code, on a subset of the bits, the nonpunctured bits, is transmitted. The purpose of puncturing is similar to that of dual LDPC, which is to create nested structure to achieve equivocation at the eavesdropper.

In [66], the authors proposed random puncturing on LDPC code, where each bit is independently punctured with a given probability, and the puncturing operation can be seen as passing the outputs of an LDPC encoder through a virtual erasure channel prior to transmission. Hence, when the random puncturing pattern is public, i.e., known to both legitimate users and the eavesdropper, and the eavesdropper's channel is worse than the main channel, by finding largest admissible puncturing probability that ensures reliable communication for the legitimate users, the eavesdropper will essentially operate above its threshold, and receive more erasures than those he can correct. On the other hand, when the puncturing pattern is shared secret between legitimate users, and unknown to the eavesdropper, a positive secrecy capacity can be reached even when the main channel is worse than the eavesdropper's channel.

In [67], similar scheme is proposed, where both the legitimate receiver and the eavesdropper are supposed to user belief propagation decoders, and the message is hided by puncturing. The results showed that the proposed design is encodable in linear time and can be effectively applied as an addition to existing secrecy code designs.

3.1.3 Polar Codes

Among various error-correction codes, polar codes are another important and popular method for communication secrecy other than LDPC codes. Since proposed by Arikan [68], polar codes has draw increasing attention from the physical layer security community as a potential code scheme for wiretap code design. The idea of polar codes comes from a phenomenon called polarization. By channel polarization, the bit-channels will be divided into noiseless bit channels or pure noisy bit-channels. In polar codes design, only those noiseless bit channels are selected to transmit information.

In polar codes, the signal bits n is supposed to be an power of two so that $n = 2^m$, $m \geq 0$. The construction of code is to making the generation matrix $\mathbf{G}_n$ as the Kronecker power of standard matrix $\mathbf{G}_2$, which can be expressed as following:

$$\mathbf{G}_n = \mathbf{G}_2^{\otimes m}$$
$$\mathbf{G}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{3.8}$$

The Kronecker power can be computed with a procedure that at each iteration, replacing the elements of the current matrix with its multiplication with the base matrix. It should be noted that such generation matrix $\mathbf{G}_n$ is a square matrix, hence with such matrix, the message is only transformed but not coded in a conventional perspective. To leverage such constructed matrix for coding, it is naturally to extend the message $\mathbf{m}$ with size $l < n$ to n bits by inserting 0 into the message vector. As results, the corresponding lines in the generation matrix of those inserted 0 will be ignored. The question is then how to select positions to insert the 0.

Polar codes proposed by Arikan [68] provided an interesting solution to address this problem by introducing *successive decoding*. Successive decoding is to decode the received signal bit by bit successively from the beginning, and at each iteration, all previous bits are considered known, the bit will be decoded with all received bits $\mathbf{y}^n$ together with previous bits. Such successive decoding framework leads to the establishment of equivalent bit channel, as at each turn, we can assume an channel where the input is the current bit and outputs are $\mathbf{y}^n$ and the previous bits. The channel of the ith bit can be expressed as:

$$W_i : m_i \rightarrow \mathbf{y}^n \mathbf{m}^{n-i} \tag{3.9}$$

The key observations of polar codes is that in such schemes, as the block length n increase, the bit channels W^n will polarize. In other words, some of the bit channels become almost noise free while others become almost completely noisy. Also, such polarization will always go towards the same direction regardless of the message $\mathbf{m}$ itself. The reason of such behavior of bit channels is because with the generation matrix $\mathbf{G}_n$, the error probability of the underlying channel will accumulate at certain bit channels, and cancel at others, causing difference in channel capacities.

The intuition of leveraging such polarization phenomenon to construct code is straightforward. Since the bad bit channels are completely noisy, their values can be set to 0 or a arbitrary fixed value. On the other hand, the real message bits will be

transmitted over the noiseless channels for reliable communication. Naturally, in a (n,k) code scheme, there will be k good bit channels and $n-k$ bad bit channels. It should be noted that the polarization of bit channels should be shared by the legitimate users prior to the communication.

To apply polar code to secrecy communication is straightforward. It should be assumed that the eavesdropper's channel is degraded with respect to the main channel, i.e., the eavesdropper's channel is a cascade of the main channel and a separate noisier channel. Hence, considering the equivalent bit channels, the good channels for the eavesdropper, denoted as W_e, will be a subset of the good channels for the legitimate receiver, denoted as W_r, which can be further decomposed into two parts, W_e and W_{r-e}, while the latter represents the bit channels that are only decodable by the legitimate receiver.

The bit channels that are bad for the eavesdropper can be represented as W_b. Hence, the whole n bits of the transmitted signal can be decomposed into three parts, namely W_{r-e}, W_e and W_b. The construction of randomization and nested structure on this setting is straightforward. We transmitted the real message $\mathbf{m}$ over W_{r-e}, and generate random message $\mathbf{m}^0$ and transmit over W_e. Hence, the codebook will be partitioned in cosets of the real message $\mathbf{m}$ with randomized $\mathbf{m}^0$.

The application of polar codes to wiretap code design have been proposed in [69, 70]. In [69], Mahdavifar and Vardy proposed to apply polar codes for secrecy to symmetric binary discrete memoryless channels, but also discussed the probability of extending such schemes to non-binary channels, as it is proved that channels with an input alphabet of prime size are polarized in the same manner [71]. The authors also discussed the difference in strong security and weak security regarding polar codes wiretap code, and proved that the former requires noiseless main channel.

In [70], Koyluoglu and Gamal applied polar code to multiple access fading channels with a degraded eavesdropper, which is used to establish secret key agreement.

The recent development of polar codes greatly influences the wiretap code design for secrecy, since it holds significant advantages over other error-correction codes.

- **Low encoding and decoding complexity**: polar codes are known to be a capacity achieving code scheme with low encoding and decoding complexity. In fact, when the polarization is known, i.e., the set of (W_{r-e}, W_e, W_b) is presented, the

construction of wiretap codes can be very simple, as to transmit the real message and the randomized message on corresponding bits.

- **Weak security without noiseless main channel**: as introduced, polar codes can achieve weak security without the main channel to be noiseless. In fact, the noisiness of the main channel is beneficial in that situation since it leads to polarization. Although, is has been reported that noiseless main channel is needed for achieving strong security, in practice, weak security is widely considered sufficient for secrecy communication.

However, there are limitations of polar codes for communication secrecy.

- **Degraded eavesdropper's channel**: a key requirement for the polar codes to work in wiretap channel is the degraded eavesdropper's channel. To be more specific, the good bit channels of the eavesdropper, W_e, should be a subset of that of the legitimate receiver, W_r.

- **Agreement on polarization**: to successfully decode the polar codes, the legitimate users should have priori agreement on the polarization of the main channel, i.e., the pair of (W_r, W_b) should be agreed upon by both terminals. The polarization perceived by the terminal is naturally dependent on the measurement of channel errors at that terminal. The agreement on polarization requires the legitimate users to have exact and identical channel error measurement.

Considering the wiretap codes in general, with both dedicated methods or error-correction codes, there exists limitations and requirements that are shared by many of the methods.

- **Better main channel conditions**: in general, wiretap codes require the main channel to be of better conditions compared to the eavesdropper's channel. In fact, many methods require the main channel to be noiseless. For polar codes, additional assumption is made that the eavesdropper's channel is degraded.

- **Dependency on block length**: the performance of wiretap codes is often dependent on the block length n, as larger n can be beneficial to secrecy. It is desirable for both simple coset code and LDPC code to have larger block length regarding

both coding efficiency and security. For polar codes, a large enough block length n is needed to have sharp polarization, as for a small number of bits, the good and bad bit channels will not be clearly divided.

3.2 Cooperation and Artificial Noise

As mentioned above, the feasibility of conventional physical layer security approaches based on single antenna systems is hampered by channel conditions. To obtain a positive secrecy capacity, the main channel is often required to be of better conditions, or noiseless. On the other hand, a noisier, or degraded eavesdropper's channel is required for practical secrecy transmission schemes. In response, cooperation based physical layer security have been widely proposed and attracting increasing attention. Cooperation based schemes further improve the practicality of physical layer security, and extend the topic into system design, instead of pure coding scheme development.

The strategy of cooperation based physical layer security is first proposed in the context of relay networking. In the case that the relay is presented and can act as a 'helper' for the legitimate users, it is natural to think of a cooperation strategy that leverages such features. The basic idea behind cooperation based schemes is to create additional noise, or dummy messages for further confusing the eavesdropper and making favorable conditions for the main channel. In other words, cooperation based physical layer security deliberately jams the eavesdropper to improve secrecy. In conventional single stream communication, such a strategy results in using meaningless dummy messages to jam. Here we define *artificial noise* as a specific case of jamming, where the jamming signal doesn't carry any useful information, but is pure noise. Various cooperation based jamming strategies and systems have been proposed and are examined in this section. It should be noted that cooperation doesn't necessarily rely on external helper, but can be the communication nodes themselves.

3.2.1 Cooperative Relay Networking

The design of cooperation based physical layer security emerges from the discussion of security in relay networking. Regarding security, there two kinds of relay nodes should be considered.

- **Untrusted relay**: in some cases, nodes are used to relay messages, which at

the same time should be kept confidential to the relays. Hence the relay can be regarded as an eavesdropper and at the same time a helper.

- **Trustworthy relay**: relay are considered trustworthy when the messages need not to be kept secret from those nodes.

The situation with an untrusted relay is first discussed by Oohama [72]. The author evaluated a general relay wiretap channel model, where at least part of the message should be confidential to the relay. The work was focus on the coding problem associated with such scenario and both stochastic and deterministic coding schemes are considered. The results proved that stochastic coding is beneficial for secrecy and leads to larger achievable rate.

In [73], He and Yaner further considered multi-access untrusted relay wiretap channel with two variations. The first is a scenario where the relay and the legitimate transmitter communicate with the receiver via a multiple access channel, while the relay and the legitimate communicate with each other through an channel orthogonal to the receiver's channel. The second, the transmitter broadcasts to both the relay and the receiver, while the relay and the receiver share an orthogonal channel. The results showed that only the second model with orthogonal channel between the relay and the receiver can establish a positive secrecy capacity. Meanwhile, in the second model, when the untrusted node can act as a helper to relay information instead of a pure eavesdropper, higher secrecy capacity can be achieved.

For relay networking, it is important to examine the different relaying strategies and why such relay should be deployed. In wireless networking, relays are mainly used for two purposes.

- **Interconnection**: a naturally utility of relay networking is to expand the reach of the network. In situations the source (transmitter) and destination (receiver) cannot communicate to each other directly because the distance between them is greater than the transmission range, relays are needed as intermediate node to establish the connection. Apart from extending the transmission range, having intermediate relay nodes brings other benefits. For instance, with multiple relays on the same level, the transmission can be speeded up by choosing the optimal path and continue even when some of the paths fail. Such scheme represents

the narrowly defined relay networking. A simple interconnection system with one source, one relay and one destination is often addressed as *two-hop relay network*, while the link from source to relay is called the first hop, and the link from relay to destination is called the second hop.

- **Cooperative diversity**: relays can be used to improve channel capacities when the source and destination can communicate directly but suffer from poor channel conditions. From this perspective, the source and destination need not to be out of range for each other. The purpose is to allow the destination decode the information by combining multiple messages, the one received by itself and those relayed to it. By introducing channel diversity, such relaying structure leads to better channel capacity and transmission reliability.

In the context of secret cooperative relay networking, both above aspects have been discussed respectively [74, 75]. In fact, physical layer security provides a new angle to access relay networking by considering its security capacity. Meanwhile, various relay strategies have been proposed for conventional relay networking, which should be analyzed from a security perspective. Conventional relay strategies can be categorized into three classes.

- **Amplify and forward (AF)**: A simple strategy is to let the relay to amplify the received signal from the source before forwarding it to the destination.

- **Decode and forward (DF)**: the relay can act more actively and decode its received signal. In the case of noisy channel and error-correction, the relay can help reducing channel errors by decoding before forwarding.

- **Compress and forward (CF)**: in this case, the relay will compress its receive signal from the source and then forward compressed signal to the destination, in order to achieve higher transmission rate.

Apart from those conventional strategies, in the context of physical layer security, another strategy designed specifically for security is proposed by Koyluoglu, Koksal and Gamal [76].

- **Randomize and forward (RF)**: as introduced, a fundamental feature of information theoretic security, especially considering the associated coding scheme,

is to leverage randomization. RF scheme makes the relays to add independent randomization at each hop, hence improving the communication secrecy.

The term of randomize and forward is proposed in [76] through the discussion on the scaling behavior of wireless networking with security constraints. However, similar concept was proposed earlier [77] as noise forwarding, where AWGN channel is considered. The proposed system fits in a scenario where multiple relays are presented at each hop, which are also spatially distributed separately. Hence, the system will choose relay nodes in a way that no eavesdroppers exist in appropriately constructed secrecy zones around them. Meanwhile, independent randomization is added at each hop, further improving the secrecy.

In [78], secrecy outage probability is used to examine the security in relay networking instead of the secret rate. The authors focus on scenarios with untrusted relays and AF strategy. The results prove, with cooperative diversity, AF provides secrecy gain compared to non-relaying. The work also suggests in multiple antenna scenario, antenna selection will be beneficial for security. As it turns out, antenna selection is an important technique for secrecy in multiple antenna system, and will be elaborated in the next section.

In [75], Ekrem and Ulukus considered a system similar to the second model proposed in [73], where the transmitter broadcasts messages to two nodes who have orthogonal channels between themselves. The receiving nodes are assumed untrusted but not malicious to each other, and should be kept ignorant to other's messages. The proposed system employs a compress and forward strategy and showed that in such scenario, positive secrecy capacities can be achieved for both receiver. The authors also discussed using jamming when the intended receiver has weaker channel condition compared to the helping user.

In [79], an AF based cooperative protocol is proposed. The proposed scheme considers the cooperative diversity aspect of relay networking, where all the relays are with the transmission range of the source, and the multiple relaying signal accumulate at the destination. The system design objective is to find optimal power weight, i.e., amplification factors among multiple relays under certain power constraint. Achieved secrecy capacity has been scaled with increasing number of relays, and DF and AF based methods are compared. The results showed that secrecy capacities of cooperative protocols

as functions of the number of relays are monotonically increasing and concave. Also, DF is reported to have better performance than AF, but with higher complexity.

In [80], the secure relay communication in cellar networking is considered, using DF and RF strategies. The results show that RF outperforms DF in different situations, stating the former as a general better design for security purpose. The authors also discussed the impact of relay placement, by proposing a heuristic relay placement strategy, as to minimize the secrecy outage probability of the intended receiver (destination).

As multiple relaying network is a common model for cooperative relay network, the management of relay nodes, apart from their forwarding strategy, becomes an important issue. A topic that related to relay placement but having more popular is *relay selection*, especially considering large scale network. As shown in [76], such technique plays a vital role in the performance of the secure system. The notion of relay selection is firstly proposed in [74], which comes from the observation that in general, a relay located closer to the main destination provides a higher secrecy enhancement. The authors also assumes a cooperative diversity scenario where all relays are within the range of the source. Various selection strategies have been proposed.

- **Conventional selection**: this scheme represents the normal relay selection method without the security constraint, i.e., considering no eavesdropper. The selection is simply based on the instantaneous quality of relay links. In some literature, such scheme is also called *Max-Min selection*, as to simply maximize the channel capacity and minimize channel error with chosen forward strategy.

- **Optimal selection**: when the conditions of the eavesdropper's channel is known at the source and relay, the optimal relay selection can be enabled. The selection is based on the classical information theoretic analysis of secrecy capacity, which measures the secrecy rate regarding the eavesdropper and each relay.

- **Suboptimal selection**: as the knowledge of the eavesdropper's channel information can be hardly assumed in practice and instantaneous channel estimation is difficult, a more practical implementation of relay selection is to avoid the need of instantaneous measurements of the eavesdropper's channel. To do so, the scheme leverages empirical observations or expectations of the eavesdropper's

channel instead of its absolute values. A simple example of suboptimal selection criterion s can be expressed as:

$$s = \arg\max_{s\in\mathbf{S}}\{\frac{\gamma_{s,D}}{E\{\gamma_{s,E}^2\}}\} \tag{3.10}$$

Where $\mathbf{S}$ denotes the set of all possible relay nodes, and $\gamma_{s,D}$ and $E\{\gamma_{s,E}^2\}$ represent the SNR of the link to destination and expectation of the eavesdropper's link respectively.

The relay selection is not only for choosing proper transmission path, but also deciding the relaying weights with multiple relays. In [81], Sun et al examined the secrecy performance of untrusted relay selection, where the selected relay is also regarded as an eavesdropper and the message should be kept confidential. The proposed selection criterion can be expressed as following.

$$s = \arg\max_{s\in\mathbf{S}}\{\frac{\alpha\gamma_{s,D}\gamma_{s,T}}{1+\alpha\gamma_{s,T}+(2-\alpha)\gamma_{s,D}}\} \tag{3.11}$$

Where α represents the transmission power factor with certain power constraint P, such that the transmitted power will be αP.

Relay selection is of practical importance for multiple relay system. In [82], the optimal relay selection assuming DF strategy and trustworthy relay is proposed. The channels for both the relay and the eavesdropper are assumed to be Rayleigh fading with same variance σ_n^2, the system employs a compound selection strategy that takes into account both channels and their statistics, showed as following.

$$s = \arg\max_{s\in\mathbf{S}}\{\frac{\min(\gamma_{s,D}^2,\gamma_{s,T}^2)P+2\sigma_n^2}{\gamma_{s,E}^2+2\sigma_n^2}\} \tag{3.12}$$

It is straightforward to see that the proposed selection method is tailored specifically for DF strategy, as $\min(\gamma_{s,D}^2,\gamma_{s,T}^2)$ represents the capacity constraint of DF protocol. The results shows that such relay selection outperforms simple Max-Min selection. Unsurprisingly, both selection methods get better performance with increasing number of relays.

An opportunistic relay selection protocol is proposed in [83]. An interconnection

system is considered where there is no direct link from the source to the destination. The authors proposed to use simple instantaneous channel conditions from multiple relays and select the relay in an opportunistic manner.

In [84], two-hop half-duplex relay networks with a buffer-aided relay is considered. The proposed protocol adapts reception and transmission time slots based on only the channel quality of the relay, while the condition of the eavesdropper's channel is assumed unknown. The use of relay buffer is simple. When the first-hop channel is significantly better than the second-hop, the relay should store the decoded data in its buffer. The relay will release its buffer and transmit the message to the destination when certain channel performance with security constraint is met. In other words, the second-hop channel should reach good capacity, at the same time, guarantee enough secrecy capacity before receiving the message. A threshold pair (α, β) is defined, while α represents the relative performance of the two channels, and β denotes the stand-alone second hop channel, expressed as following.

$$\begin{aligned} \gamma_{s,D} &\geq \alpha \gamma_{s,T} \\ \gamma_{s,D} &\geq \beta \end{aligned} \tag{3.13}$$

Based on the thresholds pair, two suboptimal selection policies are proposed respectively. The first is dependent on the relative channel conditions of the two hops as well as a fixed secrecy rate threshold R, such that $\gamma_{s,D} \geq \max(\alpha\gamma_{s,T}, 2^R - 1)$. The second policy is simply to choose relays that the second hop satisfies the threshold β.

3.2.2 Cooperative Jamming and Blinding

The cooperation based physical layer security was first discussed in the context of relay networking, where intermediate nodes are used to facilitate the networking by extending transmission range or introducing channel diversity. Conventional relay stations are used to convey messages, and their associated security properties have been analyzed. An novel branch of this research is to consider an relay that doesn't not necessarily transmit or forward any information, but to act purely as a helper for security. In some literature, the distinction of such schemes is not fully recognized, and such protocols are often addressed in the context of relay network. Here, we separate such designs from cooperative relay networking, which is fundamentally different in many aspects.

Since, in such schemes, where helping nodes are not to send meaningful information, but focus on degrading the signals that perceived by the eavesdropper, we can address such design as *cooperative jamming*.

Cooperative jamming can be performed by different parts of the network, which is not necessary to be intermediate or external nodes, but the transmitter and receiver themselves. To more specific, we define cooperative jamming as cooperation based security protocols that the cooperative nodes contribute to the secrecy of communication without handling and modifying the message of concern. In such cases, the helper transmit pure noises, or in other words, jamming signals.

Cooperative jamming is related to *randomize and forward* in cooperative relay network. The main difference is that the latter will intervene the coding scheme and send random codebook to confuse the eavesdropper, while the former just send jamming signals. The use of artificial noise for secrecy gain was first discussed in [85]. Negi and Goel considered two different settings. In the first model, no helping nodes are presented but the transmitter is equipped with multiple antennas. In the second model, transmitter is single-antenna but with help of a cooperative node. In the first scenario, which is basically a MIMO or MISO setting, the transmitter can leverage beamforming that causes additional interferences only at the eavesdropper.

In the second scenario of the proposed system of Negi and Goel, an intermediate helper is presented. The design is divided into several stages. First, the legitimate transmitter and receiver will both transmit independent artificial noise signals to the helper and the eavesdropper. The helper nodes and the eavesdropper receive different weighted versions of these two signals. In the second stage, the helper nodes replay a weighted version of the received signal. At the same time, the transmitter will send the secrecy message together with another constructed noise, which will cancel the relay noise signal at the intended receiver, but not the eavesdropper. The basic idea is to construct linear combinations of noises through those procedures that will cause only interferences at the eavesdropper.

The work of Negi and Goel [85] is one of the earliest works on applying artificial noise for secrecy. However, the two stage process and noise construction are not convenient and straightforward enough. Following works of this front have been proposed to further simplify such designs and integrate with conventional wireless networking.

In [86], Tekin and Yener considered a multiple transmitter and single receiver scenario. The authors first proposed to use a beamforming-like technique to create favorable main channel conditions by multiple transmitter. On the other hand, extra transmitters can act as helpers to send jamming signals, in order to improve the secrecy capacity of the chosen transmitter.

The helper doesn't necessarily be an external node of the communication. Lai, Gamal and Poor [87] are arguably the first to discuss using the receiver (destination) to send the jamming signal. Such design is sometimes referred to as *destination based jamming (DBJ)*. The authors proposed a noise construction scheme that based on the feedback of the destination without dedicated helpers. Interestingly, the results suggests that in the modulo-additive discrete memoryless channel with a full-duplex destination, the secrecy capacity with noisy feedback equals the capacity of the main channel in the absence of the wiretapper. A simple jamming protocol is proposed, where no noise is constructed on the transmitter side, and the destination only sends random noise to confuse the eavesdropper.

The more practical scenario without a full-duplex destination is also considered. Lai, Gamal and Poor [87] proposed an channel equivalent model based on time division. To be more specific, if the destination transmits noise instead of receiving signal at time i, the destination then lose the opportunity to receive the ith symbol, which is equivalent to the erasure channel shown in figure 3.1, assuming binary symbols are used. Then the security strategy becomes a question of finding a balance between confusing the wiretapper and degrading the main channel. The authors introduce a fixed probability t, which is known to both the transmitter and the eavesdropper. At any time slot, the destination transmits noise with probability t, otherwise receives the signal. The results shows that by simply setting t to 0.5, positive secrecy capacity can be achieved. Although such simple instance of t is suboptimal, and the optimal solution is not closed-form. The idea of equivalent channel with half-duplex jamming destination becomes an important notion in cooperative jamming.

A more recent development in destination based jamming is presented in iJam [88], where cooperative jamming is used to enhance physical layer key agreement, and the transmitter send two copies of the secret key back-to-back. Upon receiving the message, the receiver jams one of the two copies randomly. Since the two copies are

sent back-to-back, the receiver can differentiate the two copies, while the eavesdropper being confused about which copy is the jammed one. The advantage of iJam is that its compatibility with existing wireless system as it doesn't require full-duplex destination and third party. However, iJam also introduces larger overhead by sending duplicate message. Hence, such design is only feasible for occasional communication like secret key establishment.

Although the intuition of iJam is simple, the practical system design is not trivial. There are several issues about putting the framework into practice. First, the jammed sample should be ensured indistinguishable from the clean one. Alterations on the signal may change its characteristics, which is detectable by sample analysis. iJam addresses such issue by exploiting statistical properties of the transmission scheme. For OFDM particularly, since the OFDM time samples approximate random Gaussian variables, the authors proposed to use Gaussian distributed noise, which is validated by experiments.

Another practical issue is the *near-far problem*. If the eavesdropper is too close to the transmitter, the jamming power at the eavesdropper will be far lower than the power of the transmitter's signal, which may allow the eavesdropper to decode the sender's signal despite jamming. To circumvent this issue, iJam employs a two-way mechanism where both legitimate users send and jam successively, and the eavesdropper's will be jammed by at least one of them. However, such scheme further increases the overhead. Moreover, iJam is not 'secure' from an classic information theoretic security perspective, since not enough equivocation of message is created, and the eavesdropper gets at least one correct copy of the message.

A fundamental limitation of conventional cooperative jamming schemes is due to the openness of wireless medium. To confuse the eavesdropper, the jamming signal should be sent simultaneously with the intended signal. How to effectively jam the eavesdropper while keeping the legitimate communication reliable becomes an vital problem of cooperative jamming. Proposed solutions can be categorized into three aspects.

- **Two-way jamming**: as proposed in [85, 88], a common method to create eavesdropper-only jam is to construct noise with two-way communication, where legitimate nodes are required to send jamming signals. Constructed noise is then

used to null the noise at the legitimate receiver but not the eavesdropper. Such nulling is either performed by a third party [85], or the destination itself [88]. The obvious shortcoming of such scheme is the introduced transmission overhead, which will increase with reduced number of external helpers.

- **Stochastic jamming**: The framework proposed by Lai Gamal and Poor [87] represents another direction of create targeting noise. The idea is to let the destination switch between transmitting and receiving symbol-wise, hence creating an equivalence of erasure channel. Conventional information theoretic security methods can then be applied to such equivalent channel. However, the frequent switch of transmitting and receiving mode is hard to implement in practice, and may introduce problems including hardware limitation, error probability and energy overhead.

- **Full-duplex radio**: as proposed in [85], an ideal system for cooperative jamming is full-duplex radio, with which the destination itself can act as a cooperative jammer at the same time without penalties. Full-duplex radio has been an major challenge for wireless communication in general for a long time, and single access full-duplex radio have been recently reported achieved [47]. Full-duplex radio holds great potential for improving wireless security. Yet, full-duplex radio is still in its early stage and not widely available in contemporary infrastructure, although its future application is promising.

3.3 MIMO Jamming and Antenna Selection

The prior section has introduced techniques considering conventional single antenna radio systems. The spread of multiple antenna technology like Multiple Input Multiple Output (MIMO) and Multiple Input Single Output (MISO) greatly changes the problem model of cooperative jamming. Multiple antenna system enable directional transmission by beamforming, and other signal processing and precoding techniques like interference alignment and cancellation with single station, and transmit antenna selection. Also, the *multiplexing gain* of MIMO as introduced in the last chapter enable multiple streams to be transmitted simultaneously on the same channel. The multiple stream transmission enabled by MIMO can be leveraged to enhance security as well,

since different streams can be used to create interference as well. Hence, jamming signal does not necessarily need to be artificial noise signal, but meaning full messages as well, as long as they interfere with each other at the eavesdropper. Those capabilities leads to novel designs in wireless physical layer security. As a result, various strategies specifically designed for *multiple antenna systems* have been proposed.

3.3.1 Artificial Noise with Multiple Antennas

In a conventional signal antenna system, cooperative jamming transmits a constructed Artificial Noise (AN) signal to create inferior conditions at the eavesdropper. As introduced, practical single-antenna jamming protocols often introduce transmission and require dedicated helper nodes. Multiple antenna system allows the intended communication and jamming to be operated on same stations, hence avoiding the need of external helper and simplify the jamming procedure. We first consider AN based multiple antenna jamming, where the system spares part of its power to transmit meaningless jamming signal.

In [89], Goel and Negi present an early and remarkable design, where two different scenarios are considered, namely a multiple antenna transmitter, or a multiple antenna AF relay. In the first scenario where multiple antenna transmitter is presented, a beamforming based protocol is used to null the jamming signals at the receiver, but not the eavesdropper. The general concept of MIMO and its precoding techniques has been introduced in Section 2.3. Here we introduce multiple antenna cooperative jamming specifically, which doesn't necessarily require multiple antenna receivers. The construction of beamforming noise $\mathbf{n}$ is based on the channel matrix of the receiver $\mathbf{H}_r$, such that $\mathbf{H}_r\mathbf{n} = 0$. Hence, with transmitted signal $\mathbf{x} = \mathbf{x}_0 + \mathbf{n}$, the received signals at the receiver $\mathbf{y}$ and the eavesdropper $\mathbf{z}$ can be expressed as following.

$$\begin{aligned} \mathbf{y} &= \mathbf{H}_r\mathbf{x} + \mathbf{n}_r = \mathbf{H}_r\mathbf{x}_0 + \mathbf{n}_r \\ \mathbf{z} &= \mathbf{H}_e\mathbf{x}_0 + \mathbf{H}_e\mathbf{n} + \mathbf{n}_e \end{aligned} \tag{3.14}$$

Where $\mathbf{H}_e$ denotes the channel matrix of the eavesdropper, and $\mathbf{n}_e$ and $\mathbf{n}_r$ represent channel noise of the eavesdropper and the receiver respectively, $\mathbf{x}_0$ denotes the encoded message.

When the transmitter is not equipped with multiple antennas, the authors proposed

to use external multiple antenna helper instead. The source itself transmits the message combined with a jamming signal, which is known to the helper. Then the third party transmits an anti-jamming signal that cancels the jamming signal at the receiver but not at the adversary, with similar beamforming construction as shown in equation 3.14. The results show that multiple antenna cooperative jamming can create positive secrecy capacity regardless of the eavesdropper's location.

From equation 3.14, it is straightforward to see that the jamming effect on the eavesdropper is due to the signal component $\mathbf{nH}_e$, whose value can be arbitrary in a simple jamming scheme. A vital factor of the jamming performance is the eavesdropper's channel matrix $\mathbf{H}_e$. More dedicated and efficient jamming signal $\mathbf{n}$ can be constructed with priori knowledge of $\mathbf{H}_e$. Hence, known CSI of the eavesdropper can help the transmitter, or helper, to optimize jamming signal and greatly enhance secrecy.

However, it is impractical to assume usual knowledge of eavesdropper's CSI, especially for pure active attackers. Techniques that compensate the unavailability of eavesdropper's channel information. Swindlehurst [90] presented an early thought of this problem. The author proposed a simple compromising strategy. It is known that optimal secrecy capacity can not be achieved without priori knowledge of eavesdropper's CSI. Hence, instead of trying to maximize the secrecy rate and eavesdropper's equivocation, Swindlehurst proposed a strategy based on a *quality of service* (QoS) perspective. The system will set a QoS level based on signal-to-interference-plus-noise-ratio (SINR). During the communication phase, the transmitter will first calculate the power allocation p needed to meet the fixed SINR target based on channel information of the receiver. Then, the transmitter uses all of its remaining power to create orthogonal noise signal. Hence, the strategy is to maximizing jamming effect while satisfying minimum main channel requirement, which is more feasible in practical scenarios than using secrecy capacity as the performance metric, since when the eavesdropper's CSI is unknown, the actual secrecy capacity and its optimization can not be accurately measured.

In [91], Mukherjee and Swindlehurst further elaborated the QoS based strategy, by further considering the impact of channel estimation errors, or imperfect CSI. In cooperative jamming schemes, even those do not assume knowledge of the eavesdropper's channel, the channel estimation itself is assumed to be reliable. However, in

practice, channel estimation is imperfect. The authors discussed the impact of such imperfectness on jamming performance. To be more specific, the authors considered two practical scenarios of channel state estimation.

- **Estimation feedback**: The receiver conducts channel estimation upon receiving signals from the transmitter and send its estimation back to the transmitter. Such procedure is often carried out in frequency-division system where the transmission and feedback are on different channels. In other words, separate feedback channel is presented.

- **Pair estimations**: the legitimate nodes obtain their own estimations of the main channel separately, which will commonly be implemented in a time-division system, where the legitimate nodes send training symbols and conduct channel estimation successively.

Apart from the main channel estimation, Mukherjee and Swindlehurst also considered imperfect CSI of the eavesdropper, and suggested that when the benefits of known eavesdropper's CSI will lose if the knowledge is imprecise. Hence, when there may be estimation errors on the eavesdropper's channel, especially when the eavesdropper's channel is less noisy, eavesdropper's CSI should better be ignored than leveraged. The authors conducted a detailed second-order perturbation analysis in order to precisely quantify the effects of inaccurate CSI of the main channel. Unsurprisingly, the result showed that estimation feedback outperforms pair estimations, since the former allow more reliable agreement on the channel estimation than the latter. The results also proved that the secrecy rate increases with the target SINR at the receiver until certain threshold, before declining with increasing target SINR. Hence, it is suggested that a balance between the target SINR and the secrecy rate exists, and the optimal target SINR should be found.

A practical system design with unknown eavesdropper's channel is presented in STROBE [92], where a beamforming based blinding (jamming) scheme is proposed. STROBE utilizes multiple antennas access point (AP) in 802.11 network, and does not require cooperation from external node. As introduced in the last Chapter, ZFBF is a simple precoding algorithm in distributed MIMO to allow parallel transmission. In STROBE, ZFBF is used to transmit the message together with blinding signals. Blind-

ing signals will cancel out each other at the legitimate user as in distributed MIMO, while creating interference at the eavesdropper.

The main problem that STROBE is aimed to address is how to precode the jamming signal without unknown eavesdropper's channel. Unlike the QoS based strategy that just maximizes jamming power, STROBE introduces Gram-Schmidt orthonormalization which is used to create orthogonal CSI vectors from that of the legitimate user to construct the transmission matrix needed for ZFBF. The Gram-Schmidt orthonormalization, or simply Gram-Schmidt process, is a technique to construct orthogonal vector set from a non-orthogonal finite set that spans the same subspace. To be more specific, considering equation 3.14, where $\mathbf{H}_r$ is known. STROBE uses Gram-Schmidt process to create additional orthogonal vectors of $\mathbf{H}_r$. Combining the constructed vectors with $\mathbf{H}_r$, a new channel matrix $\mathbf{H}_c$ is generated.

STROBE assumes the number of antennas on AP exceeds the number of transmitting streams, hence the extra antennas, or dimensions of transmission, can be used to send jamming signal which is corresponding to the constructed orthogonal vectors. Hence STROBE presents a simple way of constructing pseudo beamforming matrix for the eavesdropper, which is constructed as orthogonal, when the channel information of eavesdropper is unknown.

STROBE validates its performance in indoor environment, with the presence of a nomadic eavesdropper which can move around the space of interests. The results show that the performance of STROBE is robust regardless of the eavesdropper's proximity to the transmitter and the receiver. However, in outdoor scenario, where there are less multipath effects, STROBE is reported to be inefficient. As stated by Jakes uniform scattering model [93], in an multipath rich environment, locations with separations of few wavelengths face uncorrelated channels. The performance lost of STROBE is likely to be caused by the fact that in a less multipath environment, the eavesdropper's channel will have more similarities and correlations with the main channel. Hence orthogonal vectors will fail to simulate the real eavesdropper's channel.

Fundamentally, multiple antenna jamming is based on beamforming, which creates dedicated noise that only affects the eavesdropper and be canceled at the legitimate receiver. Given the various works and techniques proposed in the area, fundamental challenges and limitations of such protocols remain in two major

fronts.

- **Dependency on channel estimation**: optimal beamforming based jamming requires knowledge of the eavesdropper's channel, which is impractical to be assumed aware at the legitimate nodes. The walk around of jamming strategy without knowledge of the eavesdropper's channel is an important issue for cooperative jamming. Another aspect not as widely discussed in the impact of imperfect channel estimation on the main channel, which is proposed in [91]. Both absent and inaccurate CSI used in jamming will result in the jamming signal to be leaked at the legitimate receiver, causing degradation on the main channel and insufficient jamming at the eavesdropper.

- **Power consumption**: the fundamental drawback of cooperative jamming is the power consumption. Sending jamming signal will consume transmission power which can have been used to improve the channel capacity and receiver's SNR. The cost of jamming will escalate with the unawareness of the eavesdropper's channel information, as the QoS based strategy used in such situation will maximize the jamming power, causing significant power consumption.

3.3.2 MIMO Precoding for Secrecy

The power consumption of AN based physical layer security greatly affect its practicality and usage, and will not be necessarily compensated with optimal jamming performance. Another branch of thoughts in multiple antenna physical layer security in wireless networking, is to leverage the *multiplexing gain* of MIMO for secrecy enhancement. The basic idea of such scheme is to design precoding such that the multiple streams can transmit reliably to the intended receiver, but interfere each other at the eavesdropper. Such scheme require different technique compared to conventional multiple stream precoding, as the latter lets unintended streams cancel each other, while the former requires all streams to be canceled at a different station and provides desirable precoding for the receiver at the same time.

To understand multiple stream precoding with secrecy constraint, let us first consider a simple single stream model which is Multiple Input Single Output Multiple Eavesdropper (MISOME), where the number of antenna at the legitimate receiver n_r

is 1, and multiple antennas are available at the transmitter and eavesdropper, such that $n_t, n_e > 1$. The objective of secrecy communication in such system is equivalent of creating directional beamforming to the legitimate receiver, and preventing power leakage at the eavesdropper. With known channel information of the legitimate receiver and the eavesdropper, such problem can be modeled as an generalized eigenvalue calculation. To be more specific, with $\mathbf{h}_r$ and $\mathbf{H}_e$ representing channel estimation of the legitimate receiver and the eavesdropper from the transmitter respectively, the problem can be modeled as equation 2.34, proven in [41]. The model can be expressed alternatively as following.

$$(\mathbf{I}_r + \mathbf{h}_r^H \mathbf{Q} \mathbf{h}_r)\mathbf{w} = \lambda(\mathbf{I}_e + \mathbf{H}_e^H \mathbf{Q} \mathbf{H}_e)\mathbf{w} \tag{3.15}$$

Where $\mathbf{Q}$ donates the transmission covariance matrix that satisfies the power constraint, such that $\mathbf{Q} \leq \mathbf{P}$, and $\mathbf{w}$ and λ represent the generalized eigenvector and eigenvalue respectively. It is straightforward to see that the generalized eigenvector and eigenvalue $\mathbf{w}$ and λ are the eigenvector and eigenvalue of equivalent matrix $\mathbf{H}_e^{-1}\mathbf{h}_r$. The problem of finding optimal beamforming construction is then transformed to find the largest generalized eigenvalue λ, which defines the secrecy capacity of such system as shown in equation 2.34.

Similarly, the secrecy capacity of Multiple Input Multiple Output Multiple Eavesdropper (MIMOME) system, where $n_t, n_r, n_e > 1$, can be characterized with similar forms as shown in equations 2.33, 2.35 and 2.36. An alternative expression of the MIMOME secrecy capacity as a generalized eigenvalue equation can be expressed as following.

$$(\mathbf{I}_r + \mathbf{H}_r^H \mathbf{Q} \mathbf{H}_r)\mathbf{w} = \lambda(\mathbf{I}_e + \mathbf{H}_e^H \mathbf{Q} \mathbf{H}_e)\mathbf{w} \tag{3.16}$$

MIMO secrecy precoding is essentially normal multiple stream precoding with secrecy constraints. In normal MIMO precoding, multiple streams are transmitter concurrently by techniques like zero forcing beamforming. MIMO secrecy precoding provides the same multiplexing gain as normal MIMO precoding, but additionally nulls all streams at the eavesdropper. Hence, the multiple streams themselves are leveraged as jamming signal to blind the eavesdropper. The objective of such scheme is to find the transmit structure that satisfies secrecy and power constraints. It should be noted that MIMOME precoding is a specific case of MIMO secrecy precoding where the eaves-

dropper is also equipped with multiple antennas. However, MIMOME is the major model for research MIMO secrecy capacity, as single-antenna eavesdropper is easy to deal with.

The closed-form solution of MIMO secrecy precoding with average power constraint has been proposed in [94, 95]. Such solution is based on Karush-Kuhn-Tucker (KKT) conditions and assume the transmit covariance matrix $\mathbf{Q}$ to be full rank. However, average power constraint is less useful in practical scenario as it does not fully capture the physical power limits of wireless device.

An important technique for MIMO secrecy precoding with matrix power constraint is *generalized singular value decomposition* (GSVD). GSVD is reported efficient in high SNR regime, where the asymptotically optimal solution can be found [42, 94]. As mentioned, the problem of MIMO precoding with secrecy constraints can be regarded as a generalized eigenvalue problem. GSVD is an extension to SVD introduced in the last chapter, which is associated with eigenvalue modeling. Essentially, GSVD is an generalization of SVD into matrices pairs.

Khisti and Wornell [42] first proposed the idea of GSVD in MIMO secrecy precoding. We define a simple equivalent channel matrix $\mathbf{H}$, expressed as following.

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_r \\ \mathbf{H}_e \end{bmatrix} \tag{3.17}$$

Similar to SVD, the GVSD of above matrices pair transforms the pairs by following decompositions, which decomposes each channel matrix into two unitary matrices $\mathbf{U}_r$ or $\mathbf{U}_e$ and $\mathbf{V}^H$, and non-negative diagonal matrix $\mathbf{A}_r$ and $\mathbf{A}_e$.

$$\begin{aligned} \mathbf{H}_r &= \mathbf{U}_r \mathbf{A}_r \mathbf{V}^H \\ \mathbf{H}_e &= \mathbf{U}_e \mathbf{A}_e \mathbf{V}^H \end{aligned} \tag{3.18}$$

Hence, GSVD provides decomposition to the matrices pair with the same right singular vectors matrix $\mathbf{V}$. Thus, GSVD transforms the coordinate systems at the transmitter, the receiver and the eavesdropper. The input signal at the transmitter can be expressed in terms of a coordinate system defined by $\mathbf{V}$ and the outputs at the receiver and the eavesdropper are expressed in terms of a coordinate system defined by $\mathbf{U}_r$ and $\mathbf{U}_e$ respectively.

Similar to SVD, we precode the transmission signal with the right singular vectors matrix $\mathbf{V}$, expressed as following.

$$\mathbf{s} = \mathbf{V}\mathbf{x} \tag{3.19}$$

Then the received signals at the legitimate receiver and the eavesdropper are transformed with their left singular value matrices $\mathbf{U}_r$ and $\mathbf{U}_e$ respectively. The resulting transformation is following, assuming noiseless channels for simplicity.

$$\begin{aligned} \mathbf{y} &= \mathbf{U}_r^H \mathbf{H}_r \mathbf{V}\mathbf{x} = \mathbf{A}_r \mathbf{x} \\ \mathbf{z} &= \mathbf{U}_e^H \mathbf{H}_e \mathbf{V}\mathbf{x} = \mathbf{A}_e \mathbf{x} \end{aligned} \tag{3.20}$$

It is straightforward to see that hence GSVD simultaneously diagonalizes both the main channel and the eavesdropper's channel, transforming the system into an equivalent parallel channel model. To be specific, GSVD creates a set of parallel independent subchannels between the transmitter and the receivers (the legitimate one and the eavesdropper), and then independent Gaussian codebooks can be use across these subchannels. As such, a capacity approaching scheme can be used only on those subchannels for which conditions at the legitimate receiver is larger than the eavesdropper. Hence, a positive secret capacity can be achieved in such scheme. Of course, proper wiretap codes should be used on those subchannels.

Khisti and Wornell [42] proved that by using GSVD, which diagonalizes both the legitimate channel and the eavesdropper's channel, the transmission system can be modeled as equivalent parallel independent subchannels. Following the guidance of information theoretic security, positive secrecy capacity can be achieved by picking subchannels that bring favorable conditions at the legitimate receiver. Yet, similar with single-antenna wiretap code, although GSVD provides a viable framework to achieve secrecy capacity of multi-user MIMO, the challenge then lays on how to design practical communication strategy to enable such capacity.

MIMO secrecy precoding is deployed in parallel with wiretap codes, which are previously introduced in this chapter. The main front of design practical multi-user MIMO with secrecy constraint is *power allocation*. Different from *power allocation* in conventional MIMO networking, secrecy MIMO requires unique power allocation

methods. Khisti and Wornell [42] proposed simple uniform power allocation is sufficient to attain the secrecy capacity of a MIMO wiretap channel in the high SNR regime.

In [96], Fakoorian and Swindlehurst further discussed the optimal power allocation strategy for GSVD based MIMO secrecy precoding. The authors argued that power allocation has significant impact on the secrecy performance of MIMO instead of only beamforming matrix, and presented a optimal power allocation strategy, represented by power allocation matrix **P**. As mentioned, the power allocation will only allocate power on subchannels where the legitimate receiver has better conditions than the eavesdropper. The power allocation on qualified subchannels can be expressed as following.

$$P_i = \max(0, \frac{-1+\sqrt{1-4r_ie_i+4(r_i-e_i)r_ie_i/(\mu\alpha_i)}}{2r_ie_i}, r_i > e_i \tag{3.21}$$

Where P_i, r_i and e_i represent the ith diagonal element in the power allocation matrix **P**, and matrices $\mathbf{U}_r^H\mathbf{U}_r$, $\mathbf{U}_e^H\mathbf{U}_e$ respectively. α_i denotes the ith diagonal element of matrix $\mathbf{V}^H\mathbf{V}$. $\mu > 0$ is a chosen Lagrange parameter to satisfy the system's power constraint. For $r_i < e_i$, the corresponding power allocation element P_i is assigned 0.

Fakoorian and Swindlehurst[96] proved that with proper power allocation, GSVD based MIMO secrecy procoding can approach the theoretic secrecy capacity of MIMO. The authors also suggested that conventional power allocation algorithm used in multi-user MIMO procoding is insufficient for secret communication.

In [97], Reboredo el al further discussed the optimal construction of the right singular vector matrix **V**, i.e., the transmitting filter itself. The authors proposed an optimal solution based on KKT conditions, and the secrecy capacity is measured with mean-square-error (MSE).

Essentially, MIMO secrecy procoding is to design transmitting and receiving filters and power allocation algorithm, so that the eavesdropper's channel and the legitimate channel are jointly diagonalized. Given such procoding scheme, conventional information theoretic security methods can then be applied to the system. Hence MIMO secrecy precoding constructs effective channels at the destination and the eavesdropper to enable adoption of wiretap codes for secrecy communication. The performance of MIMO secrecy precoding is subject to limitations of

common multi-user MIMO and wiretap codes. Also, MIMO secrecy precoding has its own prerequisites.

- **Full knowledge of the eavesdropper's channel**: it is obvious that MIMO secrecy precoding requires full availability of the eavesdropper's CSI at the transmitter, which is a fundamental limitation of this strategy. Dedicated eavesdropper does not reveal themselves to the legitimate users and their CSI is hardly detectable. However, the assumption of known eavesdropper's CSI at the transmitter is not entirely impractical. In practice, wireless networking is often concerned with non-dedicated eavesdropper, or partial attacker, meaning the eavesdropper is the same time a legitimate user in the networking. Assuming in a multi-user MIMO network, the transmitter sends private messages to each user, which is not permitted to elicit information from other users' messages. Hence in such system, every receiver can be regarded as legitimate user and at the same time, a potential eavesdropper as well. To ensure each user's privacy in such system is non-trivial, and MIMO secrecy provides a desirable and practical solution to such scenario without additional power assumption.

- **Dependency on channel estimation**: Similar to MIMO artificial noise, the performance of MIMO secrecy precoding is also dependent on the accuracy of channel estimation and resulting CSI. As introduced, the derivation of GSVD, or general linear precoding with secrecy constraints assume perfect CSI. In practice, wireless channels are often not known perfectly by the procoder, hence affecting the precoding performance. Although the problem of imperfect CSI can be mitigated to an extend by modeling the estimation error, and designing transmission strategy that guarantee minimum performance [98], it can not be fully eliminated, and hence affect the robustness of networking security.

- **Limitation on the eavesdropper's resources**: Most of existing techniques on MIMO secure precoding assume an eavesdropper with limited resources. Specifically, the eavesdropper are commonly assumed to have smaller amount of antennas compared to the eavesdropper, hence the increased degree of freedom will allow the eavesdropper to break the blinding effect from precoding. Similar limitation is presented with MIMO artificial noise. Although, it has been proposed

that secrecy against an eavesdropper with more antennas than the transmitter is possible [99], such solution requires additional power assumption for increasing noise. Hence, secrecy against resourceful eavesdropper remains a major challenge to MIMO security.

3.3.3 Transmit Antenna Selection

As multiple antenna system becomes widely applied, drawing increasing attention from the research community as well as physical layer security, many conventional features of MIMO have been discussed regarding secrecy. *Transmit antenna selection*, naturally enabled with multiple antenna system, has been recently proposed to be adopted in physical layer security.

Originally, antenna selection is to reduce the cost of the MIMO system while maintaining part of the high performance, since the power, hardware and computational burdens of practical multiple antenna implementation are considerable. In conventional MIMO networking, two aspects of antenna selection are considered, namely transmit antenna selection and receive antenna selection. The problem of such design is how to correctly and adaptively selection the best antenna subset among all the available antennas.

In the context of physical layer security, a specific model is considered, which is mainly focused on transmit antenna selection. We assume a MISO system, where the transmitter is equipped with multiple antennas and the legitimate receive is single-antenna. It is straightforward to see that in such scenario, multiple transmission antennas is not strictly necessary, as single antenna point to point transmission is inherently sufficient. In practice, MISO transmission is often used to leverage the multiplexing gain of multiple antennas system, which increase the channel condition and gains at the receiver. Regarding physical layer security, an emerging question is then how to leverage the flexibility of MISO system for security purposes.

In [100], Alves et al discussed a scenario where only the transmitter has multiple antennas, which can be referred to as multiple-input, single-output, single-eavesdropping (MISOSE). The authors proposed a simple strategy to transmit antenna selection. In the first stage, the legitimate nodes communicate through the public insecure channel about the selection antenna index, which is leaked to the eavesdropper.

However, the optimal selection is only meant for the legitimate receiver, and does not achieve optimum for the eavesdropper. Hence the resulting antenna selection can not bring gains at the eavesdropper the same way it do to the legitimate user. and creates favorable conditions on the main channel.

In [101], the authors further extended such strategy into the more common MISOME scenario. A same intuition is applied that the eavesdropper and the legitimate user do not benefit from transmit antenna selection in the same manner. However, with an eavesdropper more sophisticated and resourceful than the legitimate user, the design of such strategy is non-trivial. The authors proved that a positive secrecy capacity can be achieved measured by the secrecy outage probability. Also, the results suggested that higher secrecy capacity can be achieved when the number of available transmit antennas increases.

It should be noted that MISO networking do not involve multiple stream transmission. Unlike the most mentioned models of multiple antenna systems in this thesis, MISO leverages the *diversity gain* instead of *multiplexing gain*, which lets the same message arrive from diverse paths, improving the gains of the final signal. Transmit antenna selection system is much concerned by how the signals from different paths are combined at the receiver, which are commonly referred to as *diversity combing*. In the context of physical layer security, two common combing methods are considered.

- **Maximum-radio combining** (MRC): also known as ratio-squared combining, MRC assigns each channel (path) with a combing factor that is proportional to the SNR of its signal and inversely proportional to its noise level. Hence, signals from strong channels are further amplified, while weak signals are attenuated.

- **Selection combining** (SC): a simple yet effective combining strategy is to pick the strongest signal from all received signals. SC benefits form simplified procedure and reduced cost, as it doesn't require signal addition and phase shifting.

In [102], full MIMOME system is considered where the transmitter is not aware of the eavesdropper's channel, and the eavesdropper and the legitimate user can use either MRC or SC. Secrecy capacities based on the secrecy outage probability are derived with different combinations of combining methods at the eavesdropper and the legitimate user. The proposed strategy employs single transmit antenna selection, which

picks only the strongest transmit antenna. The results suggest that, for all the cases, higher secrecy capacity can be achieved with increasing number of antennas at the transmitter and the receiver, and the secrecy capacity is independent from the number of antennas at the eavesdropper.

Fundamentally, transmit antenna selection creates favorable conditions on the main channel by dynamically allocating channel resources with receiver feedback. Transmit antenna selection based security strategy exploits the channel diversity and differentiation enabled by multiple antenna. Contrary to conventional thought of security in multiple antenna systems, which is mainly focused on the aspect of *multiplexing gain* and how to leverage multiple streams to blind the eavesdropper, transmit antenna selection represents another branch of secure multiple antenna system which relies on the *diversity gain*.

The limitation of transmit antenna selection, however, comes from the diversity gain itself. As introduced, in practice, multiplexing is the main benefit of multiple antenna system. Current transmission of multiple streams greatly improve the capacity of wireless networking, especially for multi-user networking. By leveraging channel diversity, transmit antenna selection based strategy sacrifices the desirable feature of spatial multiplexing, limiting its applicability into practical multiple antenna system. Yet, the alternative thought of diversity based multiple antenna security is worth attention, as inherent limitations of multiplexing systems can be eliminated. For instance, it has been reported that the resources available to the eavesdropper, i.e., its equipped antennas, are irrelevant for transmit antenna selection strategies [101, 102].

Chapter 4

Key-Based Physical Layer Security

Traditional thoughts of communication security commonly involves an establishment of secret key. In information theoretic security, the study of perfect secrecy by Shannon [9] was also based on secret key encryption. Since Wyner's work [8] on the wiretap channel, large amount of works have been focused on designing physical layer security without keys. Following the theoretical assessments of secrecy capacity of different channels, practical solutions have been developed to achieve such capacity, from wiretap code design, to construction of secrecy channels in cooperative networking and MIMO networking.

Yet, it is not sufficient to say that keyless physical layer security is superior than key-based schemes, as in reality, the establishment of secret keys is often desirable. Keyless strategies place prerequisites on channel states of the legitimate receiver and the eavesdropper, and requires specific code design and channel construction to achieve positive secrecy capacity. Those requirements commonly come with additional power, hardware and computational costs. On the other hand, key-based strategies loosen the requirements on channel conditions and simplify the design. Key-based schemes are also more coherent with existing security measures based on encryption.

Key-based strategies are not entirely different from keyless strategies. In fact, they can be deployed in parallel with existing techniques to enhance security. Many keyless security methods can be used to facilitate and bootstrap key-based schemes. As introduced, the branch of key-based information theoretic security was pioneered by Maurer's work [44] on secret key agreement over public channel. The proposed protocol includes two main components, *information reconciliation* and *privacy amplification*.

Maurer assumes a two-way main channel, with which the legitimate transmitter

and receiver can communicate. An important notion for Maurer's design is *common randomness*, which are correlated random information available at both parties. As a result, the legitimate users establish a shared random sequence, also known as *secret seed* or *common message*, by exchanging information over the public channel, which is wiretapped by the eavesdropper, and subject to channel errors. In the first stage of information reconciliation, mismatches and errors on the random sequence from the two sides should be corrected, in order to construct a fully agreed sequence. However, since the random sequence has been wiretapped, it is not fully secure yet. In the proceeding privacy amplification phase, the legitimate users follow certain protocol to extract secret key from the random sequence and reduce leaked information to the eavesdropper.

It should be noted that the design of the those two procedures should not be considered separately. The process of information reconciliation and privacy amplification is closely related to conventional keyless wiretap channel, which shares the same objective of reducing information leakage with public message. In fact, stochastic encoding proposed by Ozarow and Wyner in type-2 wirtep channel [12], which pioneers the design of wiretap codes, can be seen as an example of privacy amplification.

Two models are commonly considered regarding the realization of common randomness in key-based physical layer security, namely *source-type model* and *channel-type model*. In source-type model, two legitimate parties observe different but correlated random variations from a same source. On the other hand, in channel-type model, transmitter simply initializes the process by sending random sequence to the receiver, which then feedbacks its observation of the transmitted sequence.

The secrecy capacity of key-based physical layer security has been examined with similar techniques used for keyless strategies. Here, the secrecy capacity is measured by *secret key rate*, which is similar to conventional *secret rate* and measures the rate of secure sharing of keys instead of message exchanging. The main benefit from key-based physical layer security is enhanced secret key sharing and distribution, which is highly desirable in modern wireless networking.

Moreover, in wireless networking, where radio channels are reciprocal and location sensitive, the fact that the response of legitimate channel could be highly uncorrelated to that of the eavesdropper's channel, brings the possibility of obtaining common

randomness from the wireless medium itself, hence further reducing computational cost and security risks. Key-based physical layer security could also be integrated with cooperative jamming and multiple antenna techniques, which could enhance the secret key rate.

Compared to keyless strategy, key-based physical layer security brings more flexibility to the system, with better compatibility and less requirements. Also, key-based physical layer is useful against not only passive attacks (eavesdropping), but active attacks as well, since the established secret key can be used for authentication of following communications. As mentioned, key-based physical layer security usually requires both legitimate nodes on the transmitting and receiving side to be active and responsible for communication security. Such requirement is compatible with modern wireless networking, where the establishment of a link usually involves two-way communication.

It should be noted that not all key-based physical layer security designs strictly follow the notion of *information reconciliation* and *privacy amplification*. Key-based security is long established and well studied in various contexts, and diverse schemes have been proposed for the physical layer. However, key-based information theoretic security provide a common framework and benchmark for designing practical systems.

4.1 Information Reconciliation and Privacy Amplification

Key-based physical layer security share many similarities with keyless strategies, and can be regarded as an extension of the latter with the use of common randomness, instead of one side randomization. The procedure of information reconciliation and privacy amplification can be regarded as a single framework of concealing secrecy message into exchanged sequence between legitimate terminals. The design of such protocols could involve conventional error-control coding, and same methods used in keyless security protocols like wiretap codes, introduced in the last chapter.

To be specific, the simple stochastic coding scheme proposed in type-2 wiretap channel [12] is indeed an example of privacy amplification. Also, it is implicitly assumed in key-based system that the legitimate terminals have more information on the

common randomness than the eavesdropper, in other words, the eavesdropper only has partial information of the shared message compared to legitimate terminals. In following procedures, the legitimate parties reconcile their shared message, and generate fully secure key by discarding the partial information known at the eavesdropper. Hence, an information advantage of the initial random message at the legitimate parties should be ensured before following process, which is often referred to as *advantage distillation*.

In channel-type model, the common randomness is initialized by the transmitter and send to the receiver. Hence, the advantage distillation in channel-type model is closely related to keyless security, and various keyless methods from wiretap codes to cooperative jamming, could be leveraged to at least partially blind the eavesdropper compared to the legitimate user. On the other hand, in channel-type model, where both legitimate nodes and the eavesdropper receive random message from an external source, advantage distillation is more difficult to realize, and a common thought on this problem is to leverage priori secure authentication. In his pioneering work [44], Maurer considered an example of source-type model, where all terminals receive broadcasting random variables from a satellite with independent noise. To guarantee advantage distillation, Maurer proposed an simple exchanging protocol where the legitimate parties send their observation to each other successively.

As mentioned, the information reconciliation stage of key-based security is to correct unmatching components in the shared random sequence (common message) and reach agreement between two parties, which is similar to error-control in classic communication model. However, in conventional error-control framework, the transmitter (the source) have error-free knowledge of the message, and the error-control process is only concerned with the receiver.

Different from this classical setting, key-based system requires both the legitimate parties to be active and reach secrecy agreement by mutual communication. Thus, both parties are assumed to have imperfect observation of the common message. To be more specific, assuming the legitimate transmitter and receiver observe the common message as $\mathbf{s}_t$ and $\mathbf{s}_r$ respectively, then $\mathbf{s}_t$ and $\mathbf{s}_r$ can be regarded as two different noise variances within the same codeword. In other words, the messages obtained at each side are different but highly correlated information.

Such characteristics of information reconciliation in key-based strategy can be

modeled as an approximate distributed source coding system, illustrated in figure 4.1. It should be noted, though, that in the classical model of distributed source coding, the correlated streams are jointly decoded at a same decoder. However, in key-based physical layer security, the information reconciliation is conducted at both legitimate parties respectively. Also, as a type of source coding, distributed source coding is mainly concerned with data compression and coding rate, while in the context of physical layer security, the objective is reconciliation of the messages.

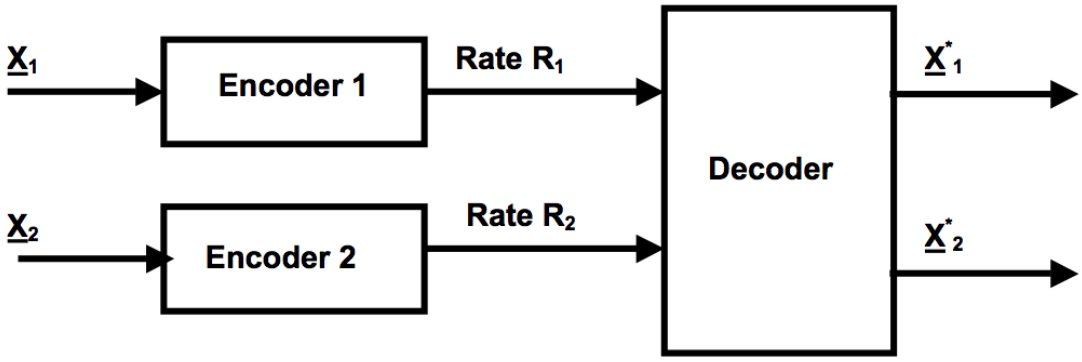

Figure 4.1: Block diagram of two stream distributed source coding

In information reconciliation, the legitimate parties act as distributed sources and receivers at the same time. Each of them sends correlated messages to each other, and reconcile their received messages with the knowledge of their own message. Hence, in key-based physical layer security, the information reconciliation process represents a special case where two decoder are involved, and one of the two correlated sources is directly available at each of the encoder. Such scenario is often addressed as *source coding with side information* in the context of coding theory.

An important result of distributed source coding, which has also been widely discussed in the context of key-based physical layer security, is *Slepian-Wolf coding*. Also known as the SlepianWolf bound, Slepian-Wolf coding considers lossless compression in distributed source coding, where the source outputs can be constructed from the compression version with arbitrary small error probability. An important aspect of the Slepian-Wolf theorem is that, compared to an encoder that assumes all data streams are independent, the separate encoders can achieve better compression rates by exploiting correlation between streams in a distributed source coding setting. The Slepian-wolf bound gives specific achievable rates for correlated streams. In an information recon-

ciliation setting with messages $\mathbf{s}_t$ and $\mathbf{s}_r$, the coding rate at the two legitimate parties, donated as R_t and R_r, can be characterized as following.

$$\begin{aligned} R_t &\geq H(\mathbf{s}_t|\mathbf{s}_r) \\ R_r &\geq H(\mathbf{s}_r|\mathbf{s}_t) \\ R_t + R_r &\geq H(\mathbf{s}_t, \mathbf{s}_r) \end{aligned} \tag{4.1}$$

It should be noted that Slepian-Wolf coding is not a specific coding scheme but rather a framework and analysis of distributed source coding model. In practice, Slepian-Wolf coding share similarities with error-control code in the single source setting. As mentioned, key-based physical layer security is closely related to keyless strategies, and many techniques proposed for keyless strategies, especially wiretap codes, also facilitate key-based methods.

Particularly, LDPC code, elaborated in the last chapter, has been considered a suitable implementation of Slepian-Wolf coding [103] in source coding with side information, and is reported to be able to approach the Slepian-Wolf bound. To be specific, in source coding with side information, each legitimate terminal has full access of one of the correlated streams, and is to estimate the other stream with this priori knowledge. It is straightforward to see that in such scenario, for each legitimate terminal, it is not necessary to send the whole message, but only side information that will allow the opposite party to successfully decode.

The practical coding design for such system is closely related to wiretap code for keyless security, elaborated in the last chapter. The main difference here is that the legitimate users share common randomness, and obtain their own realization of it. Recalling the wiretap code and nested code structure introduced in the last chapter, where the intended message $\mathbf{m}$ is encoded together with an auxiliary random message $\mathbf{m}^0$, here the legitimate terminals are equivalent to recover the message $\mathbf{m}$ with two different auxiliary random message $\mathbf{m}^0$ and $\mathbf{m}^1$.

A practical and intuitive approach of Slepian-Wolf coding is through *syndrome coding*, which is to compress the source message to its syndrome, which could be used as the side information. Assuming a parity check matrix $\mathbf{H}$, the syndrome $\mathbf{s}$ of the message $\mathbf{m}$ can be expressed as following.

$$\mathbf{s} = \mathbf{Hm} \tag{4.2}$$

It should be noted that the syndrome coding scheme shown in equation 4.2 is similar to coset coding shown in equation 3.1, with the difference that the message and resulting signal lie on opposite positions. In fact, such phenomenon demonstrates the difference between encoding and data compression. Yet both processes are related with communication secrecy, and bring the effect of confusion at the eavesdropper, since both information redundancy and information insufficiency can contribute to secrecy.

It is elaborated in the last chapter that LDPC code is a useful technique for constructing wiretap code. Similarly, the use of LDPC code for information reconciliation has been discussed in [104, 105], where both proposed schemes use dual LDPC code to implement Slepian-Wolf coding in information reconciliation. In the context of data compression and Slepian-Wolf coding, it is common to regard the compression process as an equivalent channel. Here we describe the equivalent channel coding with sparse parity check matrix $\mathbf{H}_s$ and generation matrix$\mathbf{G}_s$. And the actual channel coding for data transmission is with parity check matrix $\mathbf{H}_m$ and generation matrix$\mathbf{G}_m$, the construction of transmitted signal $\mathbf{x}$ with message $\mathbf{m}$ with dual LDPC code can be expressed as following.

$$\mathbf{s} = \mathbf{G}_m\mathbf{H}_s\mathbf{m} \tag{4.3}$$

For its efficiency and low complexity, polar codes have also been proposed to be applied to physical layer key agreement [106, 107]. It should be noted that polar code is essentially very different from traditional linear error-correction code such as LDPC. As elaborated in the last chapter, polar code leverages the polarization phenomenon and successive decoding. The physical channel can then be polarized into equivalent bit channels, which are either very good or bad. Since the main channel and the eavesdropper's channel are polarized differently, the messages are only send through bit channels that are only good for the legitimate receiver. Hence the intuition of applying polar code for key-based physical layer security departs from the case for LDPC, which is based on syndrome coding.

In [106], Renes et al assumed a communication system similar to the original

setting of Bennet [45], where two separate channels exist, a noiseless and insecure public channel and a secure but noisy private channel. The authors argued that a simple one-way transmission using polar code, i.e., the common sequence and secret key will be transmitted directly to the receiver, which follows the paradigm of conventional keyless wiretap code. Chen et al [107] further discussed a scenario where a priori common message, or secret seed, has already been obtained by the legitimate parties, and focus on the design of privacy amplification using polar code.

Constructing information reconciliation with stochastic code requires the legitimate parties to have randomization abilities, which could be computationally expensive. A simpler protocol is proposed in [108], which is referred to as *cascade-based information reconciliation*. Cascade method simplify the design by fully leveraging public discussion between legitimate stations, which randomly permute their respective streams, divide them into blocks of appropriate size, and calculate the parity check matrix of each block. This parity matrix as well as permutation information is communicated to the other party on a public channel. The other party uses this information and repeats the same permutation and parity calculation steps on their stream. For such blocks whose parity does not match, a binary search is performed to find a small number of bits that can be changed to make the block match the parity information. These steps are repeated until the probability of success reaches a desired level. However, cascade is an exhaustive protocol, and its worst case complexity, dependent on main channel conditions, could be very high, hence not being able to provide bounded performance as stochastic coding.

In key-based physical layer security, Information reconciliation allows two legitimate parties with correlated random variables to agree on a common message. Privacy amplification then allows two parties sharing a partially secret message about which an attacker has partial information, to distill a shorter but almost completely secret key by communicating over the insecure channel. As introduced, in classical key-based security setting, information reconciliation is modeled as source coding with side information. Those two components are closely related and can be regarded as a single process.

Particularly, the construction of side information used in the initial reconciliation stage has great impact on the following privacy amplification, since the side informa-

tion is also wiretapped by the eavesdropper. Intuitively, the side information may also cause information leakage and should be carefully designed. Such intuition is validated in [109], where Cachin and Maurer analysis the impact of side information on privacy amplification, and showed that each bit of side-information reduces the size of the key that can be safely distilled by at most two bits.

Privacy amplification is essentially an information compression process, which extracts a secret key from the shared message and leave the eavesdropper only negligible information about the compressed data. Assuming a binary common message $\mathbf{m}$ of n bits, which is compressed to a secret key with k bits, it should be noted that such scheme takes a similar form as wiretap code decoding with a (n,k) coding scheme. In practice, an important and popular technique for privacy amplification is universal hashing, introduced by Cater and Wegman [110].

Similar with wiretap code, universal hashing is also based on randomization. Universal hashing was originally proposed to address collisions in deterministic hash functions, where the hash function is constructed randomly from a family of hash functions. assuming the hash function h is randomly chosen from the class of functions H, which map a set S of n elements in space U to b bits index such that the resulting dictionary is of size $m = 2^b$, for $x \neq y$; $x, y \in U$. The class H is universal when following condition is met.

$$P[h(x) = h(y)] \leq 1/m \tag{4.4}$$

In practice, assuming binary messages, universal hashing can be constructed with matrix. Assuming two messages $\mathbf{x}$ and $\mathbf{y}$ differ on ith coordinate, then the corresponding ith column in the hashing matrix is randomized. Since each column contains b bits, the chance of collision of $\mathbf{x}$ and $\mathbf{y}$ will be exactly $1/m$. It should be noted that, in this case, the universal hashing takes similar form as coset coding in wiretap code.

The research on constructing information reconciliation and privacy amplification protocols is influenced by keyless wiretap codes, and leads to the design of practical key-based security system. In [111], Bloch et al presented a system design covering the whole procedure of physical layer key agreement and message protection with established key. The authors assumed a quasi-static fading channel for both legitimate users

and the eavesdropper. In the protocol design, the authors further decompose the information reconciliation stage into two steps, as first sending random message and then reconciliation processing. The transmitter then send random sequence to the receiver opportunistically.

To be more precise, the authors assume perfect CSI of the main channel and partial CSI of the eavesdropper's channel are available at the transmitter, which enable it to estimate the instantaneous secrecy capacity of the system. Since wireless channel is fluctuating, the transmission of secret seed could be opportunistically conducted only when the secrecy capacity is sufficient, hence achieving positive secrecy capacity even when the eavesdropper's channel is better than main channel on average. The authors proposed an optimized nested LDPC code for information reconciliation (error correction) and privacy amplification.

As introduced, due to the inherent common foundation of information theory, many error-control methods can be used in physical layer security design. Most of such applications involve code designs with error-correction code such as LDPC code. In [112], Abdallah el al proposed an interesting system that leverages error-detection with automatic repeat request (ARQ). The proposed system is built on 802.11 network, and only require randomization at the transmitter side.

Similar to opportunistic transmission proposed in [111], Abdallah et al [112] proposed an different technique to leverage channel fluctuation, which could be summarized as greedy transmission with feedback. To be more precise, in [112], the transmitter divide the secret key into several blocks and stochastically encode them with nested LDPC code. The proposed system leverages ARQ framework for key establishment, and each block is corresponding to an ARQ-epoch. The transmitter starts with the first stochastically encoded block. Upon receiving, the receiver conducts standard error detection. If no errors occurred, the transmitter continues to the next block, otherwise both legitimate parties drop the block, and retransmit the block but with another random codeword. It should be noted that, with stochastic coding, the transmitter could repeat corrupted blocks without retransmission of same signals. Positive secret key rate can be achieved when the eavesdropper's channel has better average SNR than the main channel. Meanwhile, greedy transmission will not require knowledge of the eavesdropper's channel. However, it is straightforward to see that such system requires

established authentication.

With various proposed techniques, the main challenge in key-based physical layer security is the obtainment of common randomness. To be more precise, the requirement of common randomness can be further decomposed into two parts. First, referred to as advantage distillation, the two legitimate users should share correlated information, on which they have better knowledge than the eavesdropper. Second, the correlated information should be constructed randomly instead of deterministically, hence reducing the risk of being compromised.

In channel-type model, the random message is initialized by one of the legitimate terminal, which follows a similar paradigm as keyless security, and experiences similar difficulties and limitations. In source-tye model, an external source of randomness is provided, of which the practical implementation is no trivia. Practical key-based security system commonly leverages two-way communication and repetition to overcome harsh channel conditions, which introduce communication overhead and reduce efficiency.

Randomization is also a key feature for key-based physical layer security. In classical key-based physical layer security, both legitimate terminals have imperfect knowledge of the common randomness, are required to have random coding ability. In practice, random coding, especially the decoding part, could be computationally expensive, and can not be assumed universally available. Hence, more efficient and convenient ways of obtain common randomness are highly desirable for key-based physical layer security.

4.2 Common Randomness of Wireless Channel

In classic key-based physical layer security, the realization of common randomness is described with two models. In channel-type model, the random message is simply initialized by one of the legitimate party. In source-type model, an external yet shared source of randomness is provided. Considering the latter, early works commonly assumed an third party broadcasting station received by both legitimate terminals, such as a broadcasting satellite. Although satellite signal is commonly available, it is difficult to ensure that the legitimate parties receive it in better conditions than the eavesdropper, which is highly beneficial for security.

Yet, the developments of key-based physical layer security lead to the understanding of inherent security advantage in wireless networking regarding the key-based framework. An important breakthrough in key-based wireless physical layer security is to characterize the wireless channel as a source of common randomness.

Common randomness of wireless channel is based on some inherent features of wireless communication. First is the reciprocity of radio wave propagation, that is, the multipath properties of the radio channel (gains, phase shifts, and delays) at any point in time are identical on both directions of a link. Hence, by utilizing reciprocity, a common message can be shared between the nodes without being exchanged through public or private channels. Second, wireless channel faces temporal variations due to movement of either end of the link and objects in the channel path. Third, the location discrimination of wireless channel proved in [113] states radio channel is spatial distinct, and an eavesdropper at a third location more than a few wavelengths from either endpoint will measure a different, uncorrelated radio channel.

Recalling the classical key-based security framework introduced in the last section, it is straightforward to see that common randomness of wireless channel perfectly suits in the source-type model, and brings very desirable features. The reciprocity and location sensitiveness of wireless channel ensure that the legitimate parties share correlated information which the eavesdropper has only inferior knowledge on. On the other hand, the temporal variations and fading subject to environmental factors give the wireless channel good randomization properties. Hence, the wireless channel provides an ideal source of common randomness for key-based security, and allows more efficient design of information reconciliation and privacy amplification.

Therefore, in practical wireless networking, where multipath, fading and station mobility are present, the main channel obtains unpredictable and distinct characters that can only be perceived by the legitimate parties, enabling them to establish key from the communication itself, which greatly simplify the system design and improve its efficiency. In this thesis, we call such secret key designs *physical layer key generation*.

The idea of leveraging common randomness of wireless channel for security is arguably first discussed by Hassan el al [114], where the authors proposed to leverage short term reciprocity of wireless channel with user mobility, and extract secret keys from channel impulse response (CIR). The idea was further elaborated in [115], where

Koorapaty el al proposed a simple secure protocol which is carried in two steps. In the first step, the phase response of the channel is probed, and in the second step the information bearing signal is modified to precompensate for the phase shifts of the channel. Hence, the measured phase response can be regarded as a secret key, the precompensation is a simple encryption method.

The authors also discussed the impact of user mobility, and stated that the transmission rate should be higher than the frequency of user's movements, in order to ensure channel reciprocity. Hence, due to the decorrelation time of the wireless channel, there is a tradeoff between transmission rate and achievable secrecy. The proposed scheme in [115] is more useful for low data rate applications with no direct line of sight (LOS) path available to the adversary. For an adversary with LOS, the phase independence argument is weaker and the security of the proposed method is compromised.

It should be noted that, since physical layer key generation follows source-type model, the communication terminals in such scenario don't have full control of the source of randomness, making the system less configurable. For instance, the information quantity, or length of the generated random message and resulted key is dependent on the wireless channel. In practice, apart from the secrecy of the key with regard to the eavesdropper, it is often desirable to have a large enough key $\mathbf{k}$, which reduces the risk of being compromised by *brute force*.

In conventional secret key protocols, whether it is physical layer or conventional encryption, the length of key is not of much significance as how to distribute such key. However, in key generation based on channel randomness, it becomes an important vector of the system performance. Since information entropy is the common metric for information quantity, Hence we introduce an unique performance metric of physical layer key generation, the *key entropy*, donated by $H(\mathbf{k})$.

With such framework, some unique challenges for physical layer key generation have emerged. Although the wireless channel provides an suitable source of common randomness, how to quantize and transform such randomness into feasible messages and keys are non-trivial. Also, performance of such designs is highly dependent on channel conditions, the problem of selecting suitable channel parameters and quantization methods to improve the robustness of the system is also very important.

4.2.1 Key Generation with Channel Estimations

The reciprocity property of wireless channels is universal, and various features can be leveraged to generate secret keys. Following the early works on physical layer key generation [114, 115], more sophisticated methods have been proposed. Many of such designs involve using channel state information, obtained with channel estimation in practice. For instance, the channel impulse response and phase response proposed by [114] and [115] respectively fall into this category.

In [116], the authors discussed key generation with multipath channels. As mentioned, from a classical view of physical layer security system, whether keyless or key-based, *randomization* is a key factor that governs the resulting performance. In scenarios with single path fading channels with user mobility, as proposed in [115], the obtained sequence of channel response can be regarded as Gaussian independent random variable, suitable for key construction. However, such simple assumption is highly suboptimal in the multipath case, where there is more dependence among channel response samples.

The challenge in such scenario is similar to information reconciliation in the classical setting, where terminals correct errors respectively, yet reconciling their selected bits. In the proposed system, the legitimate terminals should detect and remove dependent components among measured samples, yet ensuring the correlation between samples on each side. In response, the authors proposed a post processing scheme based on discretizing constituent multipath components. The selection on candidate components are based on two criterion.

- **Consistency**: the resulting components should represent an identical path that both legitimate terminals go through, hence ensuring the consistency of agreement between two sides.

- **Independency**: the resulting components should be closely independent to other components in the sample, bringing sufficient randomization for key construction.

Ye et al [116] proposed to address the information reconciliation for multipath key generation with orthogonal greedy decomposition algorithm (OGA), which at each iteration, selects a single discrete component that provides the best possible outcome

according to the above criterion. In following steps, the contribution of selected component is then computed and subtracted from the observed sample. This process is repeated until reaching a predefined threshold. The simulation results showed that with such post processing on multipath channel components, the resulting secrecy capacity, measured by secret key rate, achieves the same capacity as single path channel.

In [117], Sayeed and Perrig discussed the key entropy with multipath channels. As mentioned, the information quantity of the key itself is of particular importance to key-based system with channel randomness, since the length of key could not be instructed with terminal-based randomization, but dependent on the channel. Sayeed and Prrig proposed an interesting intuition that multipath components could help to generate larger key as compared to single path channel. To be specific, the proposed system directly quantizes the phases of channel coefficients, assuming a multipath rich environment and sufficient spatial distance between legitimate receiver and eavesdropper.

In [118], Ye et al further discussed physical layer key generation with richly scattering environment, where the short term fading channel response will be time-varying, location-sensitive, unpredictable. Assuming the fading process to be suitable for Rayleigh and Rician fading models, the authors proposed a key generation scheme based on level crossing algorithm. Considering the problem of dependency among channel response sample as suggested in [116], the authors proposed to use low-pass filter on the samples by subtracting windowed moving average, resulting in independent small scale fading variations. Such variation is supposed to follow symmetrical distribution with zero mean. The channel response samples are then quantized based on their departure from the mean. The level crossing algorithm is based on detection of successive samples that depart to the same direction, referred to as *excursion*.

To be more specific, assuming the filtered channel component is an random variable following a symmetrical distribution with zero mean and standard deviation σ, two thresholds are defined with a scaling parameter α and the standard deviation. In the following quantization process, a sample x of the channel component is then quantized with function Q based on the two thresholds, expressed as following.

$$Q(x) = \begin{cases} 1 & x > q_+ = \alpha * \sigma \\ 0 & x < q_- = -\alpha * \sigma \end{cases} \tag{4.5}$$

The intuition behind such quantization is that elements departing from the mean of the distribution hold less probability, hence bringing more randomness and entropy to the system. The algorithm then search successive 1 or 0 in the quantized sequence with minimum bits m. Each such continuous repeating sequence is called an excursion. In the following information reconciliation based on the quantization, one of the legitimate party reports his detected L excursions to the other legitimate party by sending an index $\mathbf{I}^L$ for the L excursions. The index is used to identify the position of the excursion, and can be constructed by picking the starting, ending, or central point of the excursion. Upon receiving the index $\mathbf{I}^L$, the opposite party checking his own quantized sequence. An index element is considered agreed if the corresponding bits on the receiver's quantized sequence represent an repeating sequence of at least $m-1$ bits. The receiver then report his agreed index $\mathbf{I}_r^L$ back to the transmitter. Finally, the two legitimate parties generate a shared message by picking the bits corresponding to index $\mathbf{I}_r^L$. The exchanging of excursion index is wiretapped by the eavesdropper. However, since the quantization excursion is only correlated between the legitimate parties, the eavesdropper can not decipher the shared message. It should be noted that the confusion at the eavesdropper is dependent on the amount of excursions L.

The level crossing algorithm is specifically suitable for physical layer key generation, which assumes symmetrical distributed channel fading samples. For general channel response distributions, the authors proposed a modified quantization method which converts the input channel samples into uniformly distributed quantized samples using empirical distribution function. The resulting quantized samples is not binary as in level crossing, but continuous within range $[0,1)$. Assuming the raw channel samples $\mathbf{x}^n$ at the transmitter, for the ith element x_i, the algorithm computes the number of previous samples that are less or equal to x_i, donated by $K(x_i)$, the empirical distribution function can then be expressed as.

$$U(x_i) = \frac{K(x_i)}{n} \tag{4.6}$$

Based on well established knowledge in probability theory, the Glivenko-Cantelli Theorem, such function converges to corresponding cumulative distribution function. Based on such quantization, the authors proposed to use information reconciliation

based on LDPC code, introduced in the last chapter. As LDPC coding scheme, the decoding could be much more demanding than encoding, which involves belief-propagation algorithm. On the other hand, the quantization performed in advance may introduce information loss, further complicate the problem. The authors addressed such problem by constructing per-bit log likelihood ratio with over-quantization. In the decoding stage, the log likelihood ratio together with received side information and the over-quantized bits, hence facilitating the process.

It is straightforward to see that, considering practical wireless networking, where channel noises and channel estimation errors are present, physical layer key generation also follows the notion of information reconciliation and privacy amplification. Meanwhile, since the common randomness used in physical layer key generation is obtained for the wireless channel, such designs face an additional challenge not experienced in terminal-based secret key system, which is *imperfect randomization*. While in terminal-based system, we could simply assume ideal random generator available at the terminal, in key generation based on channel randomness, the possible dependency among samples could not be ignored. Thus, the information reconciliation process in physical layer key generation, which involves in mitigating imperfect randomization, could be more complicated than conventional key-based schemes. For instance, the estimation samples may need to be filtered to eliminate dependent components.

Another notable prerequisite for information reconciliation in physical layer key generation is *quantization*. As introduced, the complex channel estimations should be first quantized before further processing. The quantization method is closely related to the following information reconciliation. Those quantization techniques could be characterized with two categories.

- **lossy quantization**: certain information or samples from the raw channel estimations will be dropped during the quantization. For instance, in binary level crossing, samples too close to the mean will be ignored. The purpose of lossy quantization is to increase the uncertainty and randomness of the shared message, hence increasing the key entropy.

- **lossless quantization**: the quantization process drops no, or only negligible number of samples from the raw estimations, hence fully leveraging the channel ca-

pacity and hardware resources, leading to potentially better secret key rate.

The tradeoff between secrecy and key capacity regarding lossy and lossless quantization affects following key generation procedures. It should be noted in key-based security, information loss is inevitable since privacy amplification is based on data compression. However, on what stage such compression should be performed, and how to design efficient algorithm to generate secure yet high-rate keys are fundamental for such designs. On the other hand, depending on the corresponding quantization, two types of information reconciliation are commonly considered in physical layer key generation systems,

- **Binary reconciliation**: the reconciliation is based on binary information. Binary reconciliation is often involved in systems assuming binary channels, or in quantization-based key generation, where raw channel estimations will be quantized to binary messages.

- **Continuous reconciliation**: the reconciled messages are continuous random variables. Such scheme is involved in physical layer generation using direct channel estimations.

In [119], Mathur et al presented an practical key generation system based on 802.11 network. The proposed system employs binary quantization and leverages level crossing for information reconciliation. Based on such framework, the authors also considered facing against active attackers, which could insert fake level crossing index into the legitimate communication. The proposed system employs an authentication protocol based on excursion index. Since the main channel is unique to the legitimate parties, the eavesdropper could only randomly insert fake index, which is highly uncorrelated to the authentic index. In response, upon receiving, legitimate terminals will first access the proportion of matching bits in the index. Index with low proportion of matching will be rejected. An active eavesdropper could also insert fake information in the privacy amplification stage. In response, the authors proposed to divide the reconciled message into two parts, one of which will be used as authentication signal. The remaining bits are then used as secret key. It should be noted that the proposed framework requires the main channel to be error-free, or has lower error probability than the eavesdropper's channel.

Physical layer key generation with channel estimation could simplify the system design, allowing more convenient algorithm such as level crossing to be deployed. Since the channel parameters provide a competent source of randomness, the computational costs and system requirements introduced by conventional key-based security could be mitigated to a certain extent. However, such designs also bring new challenges to be addressed.

- **Dependency on channel estimation**: channel estimation in practical wireless networking is non-trivial, facing challenges regarding accuracy, speed and reliability. The design of high performance channel estimation is still being explored. Particularly, key generation with channel estimation usually assumes instantaneous channel estimation, which is more demanding than long-term estimation. The resulting key generation performance faces corresponding limits, and is subject to estimation errors.

- **Imperfect randomization**: performance of key-based physical layer security is subject to the wireless channel. Particularly, schemes based on channel estimation put more demands on channel conditions, since such designs leverage channel parameters which are sensitive and subject to environmental factors. To extract more random and unpredictable components from the raw samples, signal processing should be performed to discard undesirable factors like shadow fading.

- **Difficulty on quantization**: quantization of complex channel parameters such as impulse response could be difficult and involve sophisticated signal processing. Meanwhile, such quantization also introduces information loss, especially considering lossy quantization.

4.2.2 Key Generation with Signal Strength

Received signal strength indicator (RSSI) is a popular metric in wireless communication, which also could be used as common randomness for key generation. RSSI defers from normal channel parameters since it is also subject to hardware device, and has its own measurement other than channel state estimation. A major advantage of RSSI is the fact that most of the current off-the-shelf wireless devices, without any modi-

fication, can measure it on a per-frame basis. Also, different from complex channel parameters, RSSI is a quantitative index itself, making it more straightforward to be further accessed.

RSSI is not a direct measurement of the channel. In order to obtain common randomness from the wireless channel, RSSI needs to be further transformed. For instance, a common thought in RSSI based physical layer key generation is to access the fluctuation of RSSI, instead of its absolute or average values commonly used in other systems. For physical layer key generation, raw RSSI is stable and predictable, while variations of RSSI are more unique to individual user. Also, raw RSSI is not reciprocal, and could not fully demonstrate the underneath channel properties since it is also affected by transmission factors including the output power of the transmitter and the sensitivity of the receiver. On the other hand, the fluctuation of RSSI is known unique to individual links, and hard to model, especially in multipath rich environment such as indoor spaces. With two-way communication commonly assumed in key-based security, other techniques could be used to mitigate non-channel factors in RSSI measurement, including normalization and user calibration.

In [120], Aono et al proposed to use steerable parasitic array radiator (ESPAR) antenna to create beneficial fluctuation on the wireless channel, and measure the resulting RSSI for key generation. ESPAR antenna is a type of smart antenna enabled for variable directional transmission. The proposed system employs a successive transmission protocol, where an access point equipped with ESPAR antenna transmits probing signals with different beam-forming patterns. Upon receiving successful acknowledgement from the single-antenna receiver, the access point switch to receive mode to receive feedback message, before continuing to transmission with another beam forming configuration. It should be noted that with certain configuration, the receiver may not be able to receive the message, and the access point has to retransmit. The measured RSSI at each side is normalized to have identical patterns.

In following procedure, the proposed system employs lossy quantization. Any measurements that are too close to the median value are dropped, while only wildcard measurements are used to extract key. Based on the remaining measurements, the median value is recalculated and the measurements are binary quantized with regard to the median.

Techniques proposed in channel estimation based schemes can also be leveraged in RSSI based designs. In [119], the authors discussed the feasibility of applying level crossing on RSSI measurement, which is available for off-the-shelf 802.11 devices. To discard long-term shadow fading and hardware factors, the raw RSSI measurement are filtered and normalized by subtracting out a moving average of the trace. The resulting normalized samples then represent the relative fluctuations of RSSI, and are quantized with level crossing algorithm expressed in 4.5.

In [121], a lossless binary quantization protocol is presented. Different from common binary lossy quantization which involves positive and negative thresholds, the proposed system is focusing on measuring deep fade. Deep fade is strong destructive interference experienced in wireless communication, which may result in severe drop in the channel SNR. The protocol defines a deep fading threshold, and quantized RSSI measurements that are below the threshold as 1, otherwise the sample is quantized as 0. It straightforward to see that compared to excursion based quantization, deep fade quantization provides lower message entropy and requires deep fade environment, on the other hand, it may provide higher secret key rate.

An notable quantization technique is *adaptive secret bit generation* (ASBG) proposed by Jana et al [122], where the authors proposed a hybrid scheme built on foundations of existing protocols. ASBG is essentially a lossy quantization method similar to level crossing. However, ASBG divides the consecutive RSSI measurements into blocks of appropriate size, which is a configurable parameter. For each block, they independently calculate two thresholds t_+ and t_-. It should be noted that without filtering, the resulting distribution for each block is not necessarily zero mean. Assuming the average measurement for a block μ, the thresholds can be defined as following.

$$\begin{aligned} t_+ &= \mu + \alpha * \sigma \\ t_- &= \mu - \alpha * \sigma \end{aligned} \tag{4.7}$$

In the following quantization, similar to level crossing, ASBG drops those measurements lying between those thresholds. However, for information reconciliation, ASBG makes the legitimate parties exchange index of dropped samples instead of accepted ones. Also, unlike level crossing, which normalizes the samples by subtracting a moving average, ASBG divides the measurements into blocks and computes the thresh-

olds for each block, making the protocol more adaptive to slow shifts of RSSI caused by shadow fading.

Recalling the upper and lower bound quantization in level crossing, shown in equation 4.5, only samples at the indexed positions are quantized. Each index element represent a sequence of repeating bits, and only one bit of each sequence is picked. Hence such quantization is based on occurrence of matched fading patterns. It should be noted that within each such sequence, the raw measurements of each bit may vary, the information of such discrepancy is lost during the quantization. In order to further improve efficiency, ASBG proposed *multiple bits quantization*. When excursions of RSSI measured are agreed by the two legitimate parties, each sample in the excursions will be quantized to multiple bits symbol, instead of binary bits.

ABSG conducts multiple bits quantization by following steps. Upon collecting measurements, the legitimate parties first determine the range of measurements R from the minimum and the maximum measured RSSI values, and identify a possible quantization size n, that $n \leq log_2 R$. The legitimate parties then divide the range into $m = 2n$ equal sized intervals, and choose an n bit symbol for each interval with arbitrary deterministic code. Finally, for each RSS measurement, extract n bits quantization according to the interval in which the measurement lies. Such multiple bits quantization has been reported to be able to significantly increase the secret key rate, since it further characterizes the short-term fadings experienced in the main channel.

An essential factor for RSSI-based secret key extraction is the need of channel variations. Since the secret key extraction leverages not the absolute value, but the variation of RSSI to create secrecy, the mobility of client has huge impacts on the secrecy of the extracted key. Also, the multipath structure in indoor scenario, especially whether there is a line of sight path between the legitimate nodes, affect the variation of RSSI. An experimental study on those factors is presented in ASBG. In stationary scenario, the RSSI variation is mainly caused not by actual changes in the channel, but the noise caused by hardwares of transmitting and receiving devices, hence stationary user not only bring lower entropy, but also higher disagreement rate. On the other hand, mobile user introduces more channel shifting, enabling high entropy secret key with low disagreement rate. ASBG also conducts an interesting experiment in a stationary scenario where the adversary could intentionally block the LOS path between the legit-

imate nodes, causing predictable patterns of variations on RSSI. For existing extraction schemes, such a adversary brings great danger that the key could be compromised. The same effect can be caused by a mobile client with predictable moving patterns known to the adversary. Hence, not only mobility, but also the randomness of the mobility is essential for RSSI based secret key extraction. This observation is coherent with other works [121] in RSSI based key generation, which commonly suggests that user mobility is highly desirable.

It should be noted that the half-duplex nature of the majority of existing wireless infrastructure could not allow the two legitimate parties to measure the channel at the same time on the same channel. Considering time-division system, the time slot of channel measurement, whether channel estimation and signal strength, should be kept within the coherence time to validate the reciprocity. Existing schemes commonly order the measurements at each side by sequence number instead of absolute timestamp, assuming the channel within coherence time is invariable. In [123], the authors proposed to use timestamps, and compensate the mismatch with interpolation. In the following quantization, the authors proposed to to rank the interpolated measurements in order to eliminate hardware factors.

To be more precise, the ranking algorithm convert continuous channel measurements and interpolations to quantized index by computing their ranking in the chosen sequence, then replacing the values with ranking orders. The proposed system leverages an modified multiple bits quantization protocol, where the quantization intervals are based on cumulative distribution functions of the collected samples. The cumulative distribution is divided into 2^n equal size intervals based on probability, instead of the range of values.

Signal strength based key generation becomes a popular technique for physical layer security since the measurement of RSSI is commonly available on various wireless networking including 802.11, bluetooth and cellular networking. The shortcomings of such design is mainly due to the fact that RSSI is not a direct estimate of the channel and affected by non-channel factors. Hene additional means of eliminating asymmetrical RSSI measurements should be considered. Also, RSSI is much more stable compared to channel estimations, especially in static environment. Thus, RSSI based schemes require more dynamic environ-

ment or user mobility to generate sufficient results. On the other hand, the RSSI measurement on consumer devices is not ideally sufficient on speed. State of the art physical layer key generation security achieves secret key rate in the order of bit per second, or bit per measurement/estimation, which may results a key generation time larger than the coherence time. Hence, the design of effective physical layer key generation is still an open problem.

4.3 Key Agreement with Cooperation and Multiple Antennas

Developments in keyless physical layer security can be leveraged to facilitate key-based security. Apart from the common information theoretic tools and stochastic code designs, practical enhancement including cooperation and multiple antenna systems are useful for key-based schemes as well. In [124], Pierrot and Bloch considered key-based security with destination based cooperative jamming where both legitimate parties could send additional messages to confuse the eavesdropper. The authors considered conventional scheme based on randomization at the terminals and stochastic coding. The secret key capacity of such system is derived. Assuming Gaussian two-way wiretap channel, the simulation results show that destination based jamming brings improvement on the secret key rate.

In [125], Shimizu et al proposed secret key agreement based on a two-way two-hop relaying system where the legitimate parties communicate with each other via a trusted relay, and generate secret keys with the help of the relaying signal. The proposed scheme considers amplify and forward reply with different settings, namely simple AF, signal combining AF, multiple access AF and AF with artificial noise. In signal combining AF and multiple access, the relay combines its simultaneously or successively received signals from two parties before transmitting, instead of relaying one way transmission directly. The legitimate parties could therefore extract intended messages with knowledge of their own transmitted message, which the eavesdropper could not, hence further improving the secrecy. Meanwhile, the authors proposed an AF with artificial noise strategy based on combining signals, assuming the relay and eavesdropper are equipped with multiple antennas. Standard beam forming is utilized to create

deconstructive interference at the eavesdropper. The following key generation is based on fading coefficients obtained from channel estimations.

Another two-hop relaying key generation system is proposed in [126]. The proposed system employs successive channel probing between the legitimate parties and the relay. To be more precise, the fading timeslot is divided into three parts. In the first two slots, the legitimate parties transmit training signals $\mathbf{x}$ and $\mathbf{y}$ to the relay respectively. In the final slot, the relay broadcast its training signal $\mathbf{r}$. Such process allows the three parties to obtain channel estimations of the two hop channels. In the following key generation stage, both legitimate parties reach agreements of secret keys K_1 and K_2 with the relay respectively. The relay then broadcasts bit wise summation $K_1 \oplus K_2$. Upon receiving, the legitimate parties will decompose two keys, and choose the one with bigger entropy as the common secret key. Particularly, the proposed scheme considers an active eavesdropper, which could insert jam signal in each of the training phase. The authors modeled the optimal active attack strategy for the eavesdropper based on jamming power allocation. The results show that the optimal power allocation outperforms simple average power allocation, yet a positive secret key rate is achievable even facing the optimal attack strategy.

The role of multiple antennas in improving secrecy key rate has also been discussed . In [127], Wallace and Sharma considered key-based security for indoor MIMO networking on 2.51 GHz to 2.59GHz regarding both LOS and non-LOS environments. A measurement campaign is presented to analyze factors that my affect the performance of MIMO physical layer key generation. The results show that non-LOS environment is beneficial for such systems, as LOS channels exhibit about half of the available key rate compared to non-LOS, due to the more predictable channel patterns. Assuming non-LOS environment, the authors also considered the impact of the number of antennas. Similar to existing literature, the measurement assumes the transmitter and the eavesdropper has same amount of antennas N_1, while the receiver is equipped with N_2 antennas. The results suggest that larger number of antennas could bring higher secret key rate. Two important factors for key-based security have been identified through simulations, namely sufficient multipath and spatial separation of the eavesdropper.

As introduced, an unconventional key-based security with destination based jamming in OFDM is presented in iJam[88]. The proposed system makes the transmitter to

send repeated messages for key establishment. Upon receiving such transmission two copies of the repeated messages, the receiver randomly decides to jam the samples of one of the copies. The jamming signal is constructed following Gaussian distribution, hence preventing the eavesdropper from identifying jammed samples. In order to make the system robust against the dependency on eavesdropper's location, the proposed system leverages two-way jamming, where both the legitimate parties will send and jam successively. The protocol is reported to significantly improve the secret key rate from the order of bits per second to kilobits per second. However, it should be noted that the proposed scheme differs from conventional jamming strategy, and is not fully secure from an information theoretic perspective, as the eavesdropper has large chance to break the secret key by brute force.

The various practical enhancement proposed in keyless physical layer security hold great potentials to be applied in key-based systems. The integration of those fields is still in its early stage, and further investigation in this area will likely bring promising results. Meanwhile, techniques like cooperative jamming also brings the chance for designing unconventional systems that do not follow the framework of information theoretic security, making key-based designs more flexible.

Chapter 5

Conclusion

Through decades of development, physical layer security has evolved from theoretic modeling to practical system design, and has been adapted to new networking technologies. As wireless networking becomes increasingly popular and capable, communication privacy and security become increasingly important issues, which are challenged by some inherent features of wireless communication such as superposition and openness. The openness of wireless medium makes it difficult to prevent unintended recipients, and superposition increases the risk that the intended message might be altered or jammed.

Naturally, designing secure physical layer wireless networking draw great attention in the research community, and recently we have witnessed active discussion and notable proposals in this area due to both theoretical and technical breakthroughs.

Realistically, security in wireless networks can be complicated and composed of various components. Physical layer security, leveraging inherent properties of the wireless channel, can be used to enhance and complement conventional security measures by introducing information-theoretic secrecy. For example, the advent of wireless networks has fostered the development of mobile ad-hoc networks comprised of many devices with heterogeneous capabilities; the wide range of computing power available in the devices makes it difficult to deploy a public-key infrastructure. Physical layer security could be a means to refresh the secret keys used in upper layers without deploying a traditional key management infrastructure. Also, radio access in today's wireless networks can be chaotic, as roaming and frequent hand-off procedures of mobile devices cause regular signalling exchanges between terminals for updating their associations. The heterogeneity and scalability of modern wireless networks lead to simplified de-

sign of early negotiation procedure without sufficient protection. In WLAN networks, identifiers or hardware address of terminals are directly used, and some of the protocol management frames are not encrypted, leaving the transmission vulnerable to traffic analysis and privacy leakage. In Zigbee, where a master key is used, key attack such as KillerBee [128] is able to extract the key through traffic analysis. With physical layer security techniques reviewed in this thesis, existing network security can be enhanced by facilitating key management, or introducing additional protection during vulnerable periods such as the early negotiation stages. Moreover, physical layer security can be integrated with conventional security protocols with low cost since it mainly operates at the radio interface, and recent developments in the field have greatly improve its technology readiness, as latest works are commonly capable for prototype demonstration.

Yet, physical layer security still has its limitations, as its performance is restricted by physical channel conditions. To further improve practicality and achieve guaranteed or predictable capacity remain major challenges in the field, especially considering the diverse nature of wireless networks. In this section, we introduce some emerging topics in physical layer security and present our conclusions of the state of the art, and thoughts about possible future research directions in this area.

5.1 New Topics in Physical layer Security

Physical layer security can involve interactions between various agents (transmitters, receivers, helpers, and attackers). In practice, how those agents are constrained with computational, hardware and energy limits, which are closely related to the security performance and optimal strategy. Inter-disciplinary analyses based on *game theory* and microeconomics have been reported useful to fully understand such problem.

Game theory has been widely applied in the analysis of multiple agents systems. The central principle of game theory is to model agents as rational entities whose sole focus is to maximize their individual gains or payoff functions by different strategies. A non-cooperative strategy assumes agents reject coordination with one another terminals, while in a cooperative strategy players may choose to cooperate to achieve mutual benefit. The dynamics of agent interactions will reach a point where the state of the system becomes stable, and no player has an incentive to deviate are known as Nash equilibria.

A formulation of zero-sum game model in physical layer security in presented in [129]. The model uses SINR as payoff functions, assumes the legitimate parties have only partial knowledge of the channel. It is reported that without knowledge of the eavesdropper's channel, an optimal strategy can be uniformly distribute signal energy over the subchannels.

In [130], Mukherjee and Swindlehurst presented a game model for MIMO wiretap channel with active eavesdropper. The authors considered a MIMO security system based on artificial noise, where the transmitter spares part of its power on meaningless jamming signals. On the other hand, the active eavesdropper can choose between passive eavesdropping and active jamming. The tradeoffs of the transmitter and the eavesdropper are modelled with the secrecy capacity as payoff functions. As a results, assuming different channel knowledge at the agents, equilibria with corresponding strategies are derived.

As mentioned, physical layer security is very concerned with coding design, especially stochastic coding. Emerging technologies in wireless networking require novel coding techniques, which should be accessed from a security perspective. An example of such new techniques is the design of resilient codes for distributed data and cloud storage systems, known as *distributed storage coding*. The objective of such systems is to accurately reconstruct the data which is divided and distributed over various storage nodes. Since storage nodes are assumed imperfect and subject to failure, thus fault-tolerance with bandwidth constraints is the primary code design criterion. Meanwhile, it is also critical to protect the information from being reconstructed by eavesdropper. The attack on such system usually occure during node repairing, when a node experiences failure and needs to download subsets of data stored at a set of surviving nodes, hence reconstructing the lost block using the downloaded data. A modeling of such problem is presented in [131]. The authors derived an upper bound of the amount of data that can be stored securely and reliably, known as the secrecy capacity of such system, based on the storage capacity of each node, and the total amount of downloaded data.

As introduced, existing literature in information theoretic security and physical layer security mainly consider passive attack, or eavesdropping, which is popular in wireless networking and can lead to active attacks. Meanwhile, passive attack and

active attack are closely related in practice. Many physical layer security protocols assume pure passive attacker, and require priori established authentication, limiting their practicality. Hence, the design of *physical layer authentication* becomes increasingly imperative.

Physical layer authentication can be seen as a sub-area in *non-cryptographic authentication*. Existing schemes can be divided into two categories, namely hardware fingerprinting and channel/location-based fingerprinting. Hardware fingerprinting. Hardware-based fingerprinting is based on the reflection of unique features of the hardware on the transmitted waveforms, such as characteristics of radio transceiver, clock and integrated circuits. However, such features are usually difficult to measure in practice. On the other hand, channel/location-based fingerprinting is more commonly deployed, which involves typical channel measurement like channel state information and RSSI. Meanwhile, multiple antenna technology gives new possibilities to such designs. A location fingerprinting scheme based on angle of arrival (AoA) detection enabled by multiple antennas has been proposed in [132]. Linking physical layer authentication with physical layer secrecy requires examining existing techniques from an information theoretic perspective, which is presented in [133].

5.2 Conclusions and Future Directions

In this thesis, we provide an extensive and comprehensive review on physical layer security in wireless networking. The problem of designing secure networking itself is dynamic and brings new challenges as novel technologies emerge. Existing techniques in this area have been discussed and proposed in different contexts and with various terminologies. Meanwhile, as we have shown, the many designs for physical layer security are not unconnected, but instead closely related to each other. Hence we believe such a review is necessary not only to access and categorize existing literature based on their characteristics and limitations, but also highlight their relationships.

Research on physical layer security is both theoretical and practical. Emerging from information theoretic security pioneered by Wyner, physical layer security is focused on achieving secrecy capacity indicated by theoretical results in practical networking. The essence of information theoretic security is the stochastic code designs for wiretap channel, where the communication of two legitimate parties are challenges by an eavesdropper. Proved in theory, under certain conditions regarding the legitimate channel and the eavesdropper's channel, secret communication can be achieved without encryption. In practical wireless networking, techniques like relaying, cooperative terminals and multiple antennas can contributed to create desirable channel conditions, making such approaches more feasible.

On the other hand, the establishment of a shared secret key is often desirable in wireless networks, and physical layer security can also facilitate such a functionality. Conventional secrecy communication techniques enable confidential information exchange between two or more legitimate parties, enhancing key distribution. Additional methods can then be leveraged to reconcile information on the two sides, and extract secret keys. Such methods commonly share the same principle of randomization as keyless secrecy communication. Moreover, the reciprocity and discriminability of wireless medium enables secret keys to be extracted from the wireless channel itself.

The foundation of such systems is an information theoretic assessment of information leakage and secrecy. Hence the security is defined as the level of equivocation created at the eavesdropper. Apart from means in wired network, physical layer security in wireless networking can also exploit the independence of wireless channels and common fading and noise, or leverage superposition and construct dedicated interfer-

ence towards the attackers.

Wireless communication is subject to environmental factors, and involves diverse signal processing methods and transmission protocols. On the other hand, physical layer security in wireless networking places requirements on the wireless channel, hardware as well as the underlying signal. To build practical security systems for the diverse settings of wireless networking is non-trivial. During recent years, a notable resurgence has been witnessed in this area, and promising new results should be expected. The scope of future directions in this area is extensive, here we highlight several notable examples.

- **Experimental validation**: physical layer security in wireless networking is build on various assumptions of channel models. Real time experimental study emulating differing channel conditions and adversaries is needed to fully understand the capacity and limits of physical layer security in practical networks.

- **Scalability**: it should be noted that most of existing literature in physical layer security consider rather simplified communication models. Classical wiretap channel considers a simple three-terminal situation, while recent schemes include relaying and multiple access scenarios. Very little discussion has been presented about security in large scale networks. Modern wireless networking is impressive in its scalability, and large scale commercial networks are operated on a daily base. The security challenges in such large scale networks considering user interactions and capacity constraints should be better understood.

- **Compatibility with emerging technologies**: as mentioned, the performance of physical layer security is closely related to the underlying signal. Since wireless communication has kept evolving, and new network standards are being developed, it can be expected that new network scenarios and corresponding security schemes will continue to be developed as well, which require more fundamental understanding of physical layer security, and designing flexible methods independent upon different types of network.

- **Cross-layer designs**: physical layer security can be deployed in parallel with conventional encryption based security. In practice, upper layer security proto-

cols like WEP and WPA are widely deployed, and the implementation of physical layer security in a real system is likely to be part of a layered-approach. The design of protocols that combine traditional cryptographic techniques with physical layer techniques is an rich but unexplored resource for further study. A key part of such research is the definition of proper metrics that can assess the performance of such hybrid schemes.

- **Security with full-duplex radio**: the development in full-duplex radio has recently achieved remarkable breakthroughs [47], which will significantly change the current understanding of wireless networking. Full-duplex radio also brings opportunities for physical layer security. For instance, key-based security in half-duplex network can not have ideally symmetrical channel estimations. With full-duplex, the efficiency of key agreement will be greatly improved. The feasibility of full-duplex radio and its security applications is an notable direction to follow.

Bibliography

[1] Benjamin Bertka. 802 . 11w Security : DoS Attacks and Vulnerability Controls. In *Proceedings of IEEE INFOCOM*, 2012.

[2] Martin Eian and Stig F. Mjolsnes. The modeling and comparison of wireless network denial of service attacks. In *Proceedings of the 3rd ACM SOSP Workshop on Networking, Systems, and Applications on Mobile Handhelds - MobiHeld '11*, pages 1–6, New York, NY, USA, 2011.

[3] Erik Tews and Martin Beck. Practical attacks against WEP and WPA. In *Proceedings of the second ACM conference on Wireless network security - WiSec '09*, pages 79–86, New York, NY, USA, 2009.

[4] Shyamnath Gollakota, Samuel David Perli, and Dina Katabi. Interference Alignment and Cancellation. In *ACM SIGCOMM Computer Communication Review*, volume 39, pages 159–170, August 2009.

[5] Yi-Sheng Shiu, Shin Yu Chang, Hsiao-Chun Wu, S.C.-H Huang, and Hsiao-Hwa Chen. Physical Layer Security in Wireless Networks: A Tutorial. *IEEE Wireless Communications Magazine*, 18(2):66–74, 2011.

[6] Amitav Mukherjee, S. Ali A. Fakoorian, Jing Huang, and A Lee Swindlehurst. Principles of Physical Layer Security in Multiuser Wireless Networks : A Survey. 2014.

[7] John Bellardo and Stefan Savage. 802.11 Denial-of-Service Attacks: Real Vulnerabilities and Practical Solutions. In *Proceedings of the 12th USENIX Security Symposium*, Berkeley, US, 2001.

[8] A.D. Wyner. The Wire-Tap Channel. *The Bell System Technical Journal*, 1975.

[9] C E Shannon. Communication theory of secrecy systems. 1945. *Bell System Technical Journal*, 28:656–715, 1949.

[10] P.R. Geffe. Secrecy systems approximating perfect and ideal secrecy. *Proceedings of the IEEE*, 53(9):1229–1230, 1965.

[11] A. B. Carleial and M. E. Hellman. A note on Wyner's wiretap channel (Corresp.). *IEEE Transactions on Information Theory*, 23:387–390, 1977.

[12] L. H. Ozarow and a. D. Wyner. Wire-Tap Channel II. In *Proceedings of Eurocrpt. Workshop on Advances in Cryptology*, pages 33–51, 1985.

[13] S. Leung-Yan-Cheong and Martin E Hellman. The Gaussian Wire-Tap Channel. *IEEE Transactions on Information Theory*, 24(4), 1978.

[14] I. Csiszar and J. Korner. Broadcast Channels with Confidential Messages. *IEEE Transactions on Information Theory*, IT-24(3):339–348, 1978.

[15] J. Barros and M.R.D. Rodrigues. Secrecy Capacity of Wireless Channels. *2006 IEEE International Symposium on Information Theory*, pages 356–360, 2006.

[16] Praveen Kumar Gopala, Lifeng Lai, and Hesham El Gamal. On the secrecy capacity of fading channels. *IEEE Transactions on Information Theory*, 54(10):4687–4698, 2008.

[17] Zang Li, Roy Yates, and Wade Trappe. Secret Communication with a Fading Eavesdropper Channel. In *Proceedings of IEEE ISIT*, pages 1296–1300, 2007.

[18] Yingbin Liang Yingbin Liang, H.V. Poor, and S. Shamai. Secure Communication Over Fading Channels. *IEEE Transactions on Information Theory*, 54(6):2470–2492, 2008.

[19] Yingbin Liang, Gerhard Kramer, H. Vincent Poor, and Shlomo Shamai. Compound wiretap channels. *Eurasip Journal on Wireless Communications and Networking*, 2009.

[20] Ersen Ekrem and Sennur Ulukus. Secrecy Capacity of a Class of Broadcast Channels with an Eavesdropper. *EURASIP Journal on Wireless Communications and Networking*, 2009.

[21] Ashish Khisti, Aslan Tchamkerten, and Gregory W. Wornell. Secure Broadcasting Over Fading Channels. *IEEE Transactions on Information Theory*, 54(6):2453–2469, June 2008.

[22] Shun Watanabe and Yasutada Oohama. Broadcast channels with confidential messages by randomness constrained stochastic encoder. *IEEE International Symposium on Information Theory - Proceedings*, pages 61–65, 2012.

[23] Ersen Ekrem and Sennur Ulukus. The Secrecy Capacity Region of the Gaussian MIMO Multi-Receiver Wiretap Channel. *IEEE Transactions on Information Theory*, 57(4):2083–2114, 2011.

[24] Gerard J. Foschini. Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multi-Element Antennas. *Bell Labs Technical Journal*, 1(2):41–59, 2002.

[25] Lizhong Zheng and D.N.C. Tse. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *IEEE Transactions on Information Theory*, 49(5):1073–1096, May 2003.

[26] Horia Vlad Balan, Ryan Rogalin, Antonios Michaloliakos, Konstantinos Psounis, and Giuseppe Caire. Achieving High Data Rates in a Distributed MIMO System. In *Proceedings of the 18th annual international conference on Mobile computing and networking - Mobicom '12*, pages 41–52, New York, NY, USA, 2012.

[27] Giuseppe Caire and Shlomo Shamai. On the Achievable Throughput of a Multiantenna Gaussian Broadcast Channel. *IEEE Transactions on Information Theory*, 49(7):1691–1706, 2003.

[28] Sriram Vishwanath and Nihar Jindal. Duality , Achievable Rates , and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels. *IEEE Transactions on Information Theory*, 49(10):2658–2668, 2003.

[29] Max H.M. Costa. Writing on Dirty Paper. *IEEE Transactions on Information Theory*, IT-29(3):439–441, 1983.

[30] U. Erez and S. ten Brink. A close-to-capacity dirty paper coding scheme. *IEEE Transactions on Information Theory*, 51(10):3417–3432, October 2005.

[31] Hiroshi Harashima and Hiroshi Miyakawa. Matched-Transmission Technique for Channels With Intersymbol Inteiference. *IEEE Transactions on Communications*, com-20(4):774–780, 1972.

[32] M Tomlinson. New Automatic Equaliser Employing Modulo Arithmetic. *Electronics Letters*, 7(5-6):138–139, 1971.

[33] Yi Jiang, Jian Li, and William W. Hager. Joint Transceiver Design for MIMO Communications Using Geometric Mean Decomposition. *IEEE Transactions on Signal Processing*, 53(10):3791–3803, 2005.

[34] Chau Yuen, Sumei Sun, and Jian Kang Zhang. Comparative Study of SVD and QRS in Closed-loop Beamforming Systems. *Proceedings - IEEE Military Communications Conference MILCOM*, pages 2–5, 2007.

[35] Melda Yuksel and Elza Erkip. Diversity-multiplexing tradeoff for the multiple-antenna wire-tap channel. *IEEE Transactions on Wireless Communications*, 10(3):762–771, 2011.

[36] Shabnam Shafiee, Nan Liu, and Sennur Ulukus. Secrecy capacity of the 2-2-1 Gaussian MIMO wire-tap channel. *2008 3rd International Symposium on Communications, Control, and Signal Processing, ISCCSP 2008*, pages 207–212, 2008.

[37] Ashish Khisti, Gregory Wornell, Ami Wiesel, and Yonina Eldar. On the Gaussian MIMO wiretap channel. *IEEE International Symposium on Information Theory - Proceedings*, (2):2471–2475, 2007.

[38] Ronit Bustin, Ruoheng Liu, H.Vincent Poor, and Shlomo Shamai (Shitz). An MMSE Approach to the Secrecy Capacity of the MIMO Gaussian Wiretap Channel. In *EURASIP Journal on Wireless Communications and Networking*, pages 2602–2606, 2009.

[39] Frédérique Oggier and Babak Hassibi. The secrecy capacity of the MIMO wiretap channel. *IEEE Transactions on Information Theory*, 57(8):4961–4972, 2011.

[40] Tie Liu and Shlomo Shamai. A note on the secrecy capacity of the multiple-antenna wiretap channel. *IEEE Transactions on Information Theory*, 55(6):2547–2553, 2009.

[41] Ashish Khisti and Gregory Wornell. Secure Transmission with Multiple Antennas : The MISOME Channel. *IEEE Transactions on Information Theory*, 56(7):3088–3104, 2010.

[42] Ashish Khisti and Gregory W. Wornell. Secure Transmission With Multiple Antennas-Part II: The MIMOME Wiretap Channel. *IEEE Transactions on Information Theory*, 56(11):5515–5532, November 2010.

[43] Ruoheng Liu, Tie Liu, H. Vincent Poor, and Shlomo Shamai Shitz. Multiple-Input Multiple-Output Gaussian Broadcast ChannelsWith Confidential Messages. *IEEE Transactions on Information Theory*, 59(9):1346–1359, 2013.

[44] Ueli M. Maurer. Secret key agreement by public discussion from common information. *IEEE Transactions on Information Theory*, 39(3):733–742, May 1993.

[45] Charles H. Bennett, Gilles Brassard, and Jean-Marc Robert. Privacy Amplification by Public Discussion. *SIAM Journal on Computing*, 17(2):210–229, 1988.

[46] R Ahlswede and I Csiszár. Common randomness in information theory and cryprography-Part I: Secrect Sharing. *IEEE Trans. Info. Theory*, 39(4):1121–1131, 1993.

[47] Dinesh Bharadia, Emily McMilin, and Sachin Katti. Full Duplex Radios. *ACM SIGCOMM Computer Communication Review*, 43(4):375–386, 2013.

[48] Charles H Bennett, Gilles Brassard, Claude Crkpeau, Ueli M Maurer, and Senior Member. Generalized privacy amplification. *Information Theory, IEEE Transactions on*, 41(6):1915–1923, 1995.

[49] Imre Csiszár and Prakash Narayan. Common randomness and secret key generation with a helper. *IEEE Transactions on Information Theory*, 46(2):344–366, 2000.

[50] U. Maurer and S. Wolf. Secret-Key Agreement Over Unauthenticated Public Channels-Part I: Definitions and a Completeness Result. *IEEE Transactions on Information Theory*, 49(4):822–831, 2003.

[51] Ueli Maurer and Stefan Wolf. Secret-Key Agreement Over Unauthenticated Public Channels-Part II: The Simulatability Condition. *IEEE Transactions on Information Theory*, 49(4):832–838, 2003.

[52] U. Maurer and S. Wolf. Secret-Key Agreement Over Unauthenticated Public Channels-Part III: Privacy Amplification. *IEEE Transactions on Information Theory*, 49(4):822–831, 2003.

[53] Viktor Yakovlev, Valery Korzhik, Guillermo Morales-Luna, and Mihail Bakaev. Key Distribution Protocols Based on Extractors Under the Condition of Noisy Channels in the Presence of an Active Adversary. 54(6):2535–2549, 2010.

[54] Somayeh Salimi, Mahmoud Salmasizadeh, Mohammad Reza Aref, and Jovan Dj. Key Agreement Over Multiple Access Channel. *IEEE Trans. Inf. Forensics Security*, 6(3):775–790, 2011.

[55] Somayeh Salimi, Eduard A Jorswieck, and Mikael Skoglund. Secret Key Agreement over an Interference Channel Using Noiseless Feedback. In *Proceedings of IEEE ISIT*, 2013.

[56] Imre Csiszár and Prakash Narayan. Secrecy capacities for multiple terminals. *IEEE Transactions on Information Theory*, 50(12):3047–3061, 2004.

[57] Francesco Renna, Matthieu R. Bloch, and Nicola Laurenti. Semi-blind key-agreement over MIMO fading channels. *IEEE Transactions on Communications*, 61:620–627, 2013.

[58] A.R. Calderbank and N. Sloane. New trellis codes based on lattices and cosets. *IEEE Transactions on Information Theory*, 33(2):177–195, 1987.

[59] Xiang He and Aylin Yener. Providing secrecy with structured codes: Two-user Gaussian channels. *IEEE Transactions on Information Theory*, 60(4):2121–2138, 2009.

[60] Ruoheng Liu, Yingbin Liang, H. Vincent Poor, and Predrag Spasojević. Secure nested codes for type II wiretap channels. *2007 IEEE Information Theory Workshop*, pages 337–342, 2007.

[61] Jean Claude Belfiore and Frédérique Oggier. Lattice code design for the Rayleigh fading wiretap channel. *IEEE International Conference on Communications*, 2011.

[62] Andrew Thangaraj, Souvik Dihidar, a. R. Calderbank, Steven W. McLaughlin, and Jean Marc Merolla. Applications of LDPC codes to the wiretap channel. *IEEE Transactions on Information Theory*, 53(8):2933–2945, 2007.

[63] Robert G Gallager. *Low Density Parity-Check Codes*. PhD thesis, 1963.

[64] Aliazam Abbasfar, Dariush Divsalar, and Kung Yao. Accumulate-repeat-accumulate codes. *IEEE Transactions on Communications*, 55(4):692–702, 2007.

[65] Vishwambhar Rathi, Mattias Andersson, Ragnar Thobaben, Joerg Kliewer, and Mikael Skoglund. Performance analysis and design of two edge-type LDPC codes for the BEC wiretap channel. *IEEE Transactions on Information Theory*, 59:1048–1064, 2013.

[66] Joao Almeida and Joao Barros. Random Puncturing for Secrecy. In *Asilomar Conference on Signals, Systems and Computers*, 2013.

[67] Demijan Klinc, Jeongseok Ha, Steven W. McLaughlin, João Barros, and Byung-jae Kwak. LDPC Codes for the GaussianWiretap Channel. *IEEE Transactions on Information Forensics and Security*, 6(3):551–564, 2011.

[68] Erdal Arikan. Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *IEEE Transactions on Information Theory*, 55:3051–3073, 2009.

[69] Hessam Mahdavifar and Alexander Vardy. Achieving the secrecy capacity of wiretap channels using polar codes. *IEEE Transactions on Information Theory*, 57(10):6428–6443, 2011.

[70] O. Ozan Koyluoglu and Hesham El Gamal. Polar Coding for Secure Transmission and Key Agreement. *IEEE Transactions on Information Forensics and Security*, 7(5):1472–1483, 2012.

[71] Eren Sasoglu, Emre Telatar, and Erdal Arikan. Polarization for arbitrary discrete memoryless channels. *2009 IEEE Information Theory Workshop*, pages 144–148, 2009.

[72] Yasutada Oohama. Capacity Theorems for Relay Channels with Confidential Messages. *IEEE International Symposium on Information Theory - Proceedings*, pages 926–930, 2007.

[73] Xiang He and Aylin Yener. Cooperation with an untrusted relay: A secrecy perspective. *IEEE Transactions on Information Theory*, 56(8):3807–3827, 2010.

[74] Gaojie Chen, Vincent Dwyer, Ioannis Krikidis, John S. Thompson, Steve McLaughlin, and Jonathon Chambers. Relay Selection for Secure Cooperative Networks with Jamming. *IEEE Transactions on Wireless Communications*, 8(10):5003–5011, 2009.

[75] Ersen Ekrem and Sennur Ulukus. Secrecy in cooperative relay broadcast channels. *IEEE Transactions on Information Theory*, 57:137–155, 2011.

[76] O. Ozan Koyluoglu, Can Emre Koksal, and Hesham El Gamal. On secrecy capacity scaling in wireless networks. *IEEE Transactions on Information Theory*, 58(5):3000–3015, 2012.

[77] Lifeng Lai and Hesham El Gamal. The RelayEavesdropper Channel: Cooperation for Secrecy. *IEEE Transactions on Information Theory*, 54(9):4005–4019, September 2008.

[78] Jing Huang, Amitav Mukherjee, and A. Lee Swindlehurst. Secure communication via an untrusted non-regenerative relay in fading channels. *IEEE Transactions on Signal Processing*, 61(10):2536–2550, 2013.

[79] Lun Dong, Zhu Han, Athina P. Petropulu, and H. Vincent Poor. Amplify-and-forward based cooperation for secure wireless communications. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 2613–2616, 2009.

[80] Jianhua Mo, Meixia Tao, and Yuan Liu. Relay placement for physical layer security: A secure connection perspective. *IEEE Communications Letters*, 16:878–881, 2012.

[81] Li Sun, Taiyi Zhang, Yubo Li, and Hao Niu. Performance study of two-hop amplify-and-forward systems with untrustworthy relay nodes. *IEEE Transactions on Vehicular Technology*, 61(8):3801–3807, 2012.

[82] Yulong Zou, Xianbin Wang, and Weiming Shen. Intercept probability analysis of cooperative wireless networks with best relay selection in the presence of eavesdropping attack. *IEEE International Conference on Communications*, pages 2183–2187, 2013.

[83] Yulong Zou, Xianbin Wang, Weiming Shen, and Lajos Hanzo. Security Versus Reliability Analysis of Opportunistic Relaying. *IEEE Transactions on Vehicular Technology*, 63(6):2653–2661, 2014.

[84] Jing Huang and A. Lee Swindlehurst. Wireless Physical Layer Security Enhancement with Buffer-Aided Relaying. In *Asilomar Conference on Signals, Systems and Computers*, 2013.

[85] Rohit. Negi and Satashu. Goel. Secret communication using artificial noise. *IEEE 62nd Vehicular Technology Conference (VTC)*, 3:1906–1910, 2005.

[86] Ender Tekin and Aylin Yener. Achievable Rates for the General Gaussian Multiple Access Wire-Tap Channel with Collective Secrecy. In *Annual Allerton Conference on Communication, Control, and Computing*, 2006.

[87] Lifeng Lai, Hesham El Gamal, and H Vincent Poor. The Wiretap Channel With Feedback: Encryption Over the Channel. *IEEE Transactions on Information Theory*, 54(11):5059–5067, 2008.

[88] Shyamnath Gollakota and Dina Katabi. Physical layer wireless security made fast and channel independent. In *2011 Proceedings IEEE INFOCOM*, pages 1125–1133, April 2011.

[89] Satashu Goel and Rohit Negi. Guaranteeing Secrecy using Artificial Noise. *IEEE Transactions on Wireless Communications*, 7(6):2180–2189, 2008.

[90] A. Lee Swindlehurst. FIXED SINR SOLUTIONS FOR THE MIMO WIRETAP CHANNEL. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2437–2440, 2009.

[91] Amitav Mukherjee and A. Lee Swindlehurst. Robust Beamforming for Security in MIMOWiretap Channels With Imperfect CSI. *IEEE Transactions on Signal Processing*, 59(1):351–361, 2011.

[92] Narendra Anand and Edward W. Knightly. STROBE: Actively securing wireless communications using Zero-Forcing Beamforming. In *Proceedings of IEEE INFOCOM*, pages 720–728, March 2012.

[93] P Dent, G E Bottomley, and T Croft. Jakes Fading Model Revisited. *Electronics Letters*, 29(13), 1993.

[94] Sergey Loyka and Charalambos D Charalambous. On Optimal Signaling over Secure MIMO Channels. In *IEEE International Symposium on Information Theory - Proceedings*, pages 443–447, 2012.

[95] S. Ali A Fakoorian and A. Lee Swindlehurst. Full rank solutions for the MIMO gaussian wiretap channel with an average power constraint. *IEEE Transactions on Signal Processing*, 61(10):2620–2631, 2013.

[96] S. Ali. A. Fakoorian and A. Lee Swindlehurst. Optimal Power Allocation for GSVD-Based Beamforming in the MIMO Wiretap Channel. In *IEEE International Symposium on Information Theory - Proceedings*, 2010.

[97] Hugo Reboredo, Vinay Prabhu, Miguel R D Rodrigues, and João Xavier. Filter design with secrecy constraints: The multiple-input multiple-output Gaussian wiretap channel with zero forcing receive filters. *IEEE Transactions on Signal Processing*, 61(15):3799–3814, 2013.

[98] Amitav Mukherjee and A. Lee Swindlehurst. Robust Beamforming for Security in MIMO Wiretap Channels with Imperfect CSI. *IEEE Transactions on Signal Processing*, 59(1):351–361, 2011.

[99] Shuiyin Liu, Yi Hong, and Emanuele Viterbo. Practical Secrecy using Artificial Noise. *IEEE Communications Letters*, 17(7):1483–1486, 2013.

[100] Hirley Alves, Richard Demo Souza, and Merouane Debbah. Enhanced physical layer security through transmit antenna selection. *2011 IEEE GLOBECOM Workshops*, pages 879–883, 2011.

[101] Hirley Alves, Richard Demo Souza, Mérouane Debbah, and Mehdi Bennis. Performance of transmit antenna selection physical layer security schemes. *IEEE Signal Processing Letters*, 19(6):372–375, 2012.

[102] Nan Yang, Phee Lep Yeoh, Maged Elkashlan, Robert Schober, and Iain B. Collings. Transmit antenna selection for security enhancement in MIMO wiretap channels. *IEEE Transactions on Communications*, 61(1):144–154, 2013.

[103] Jun Muramatsu, Tomohiko Uyematsu, and Tadashi Wadayama. Low-Density Parity-Check Matrices for Coding of Correlated Sources. *IEEE Transactions on Information Theory*, 51(10):3645–3654, 2005.

[104] Jalal Etesami, Werner Henkel, and I I S Ystem M Odel. LDPC Code Construction for Wireless Physical-Layer Key Reconciliation. In *IEEE International Conference on Communications in China: Communications Theory and Security*, pages 208–213, 2012.

[105] Oana Graur, Nazia Islam, Alexandra Filip, and Werner Henkel. Quantization Aspects in LDPC Key Reconciliation for Physical Layer Security. In *10th International ITG Conference on Systems, Communications and Coding*, pages 1–6, 2015.

[106] Joseph M. Renes, Renato Renner, and David Sutter. Efficient one-way secret-key agreement and private channel coding via polarization. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8269:194–213, 2013.

[107] Rémi a. Chou, Matthieu R. Bloch, and Emmanuel Abbe. Polar coding for secret-key generation. *2013 IEEE Information Theory Workshop*, 2013.

[108] Gilles Brassard and Louis Salvail. Secret-Key Reconciliation by Public Discussion. *Lecture Notes in Computer Science*, pages 410–423, 1994.

[109] Christian Cachin and Ueli M. Maurer. Linking Information Reconciliation and Privacy Amplification. *Journal of Cryptology*, 10:97–110, 1997.

[110] J. Lawrence Carter and Mark N. Wegman. Universal Classes of Hash Functions. *Journal of Computer and System Sciences*, 18:143–154, 1979.

[111] Matthieu Bloch, João Barros, Miguel R. D. Rodrigues, and Steven W. Mclaughlin. Wireless Information-Theoretic Security. *IEEE Transactions on Information Theory*, 54(6):2515–2534, 2008.

[112] Yara Abdallah, Mohamed Abdel Latif, Moustafa Youssef, Ahmed Sultan, and Hesham El Gamal. Keys through ARQ: Theory and practice. *IEEE Transactions on Information Forensics and Security*, 6:737–751, 2011.

[113] G.D. Durgin. *Space-Time Wireless Channels*. Prentice Hall PTR, January 2002.

[114] Amer A. Hassan, Wayne E. Stark, John E. Hershey, and Sandeep Chennakeshu. Cryptographic key agreement for mobile radio. *IEEE Digital Signal Processing Magazine*, 6(1996):207–212, 1996.

[115] H. Koorapaty, A. A. Hassan, and S. Chennakeshu. Secure information transmission for mobile radio. *IEEE Communications Letters*, 4(2):52–55, 2000.

[116] Chunxuan Ye, Alex Reznik, Gregory Sternberg, and Yogendra Shah. On the Secrecy Capabilities of ITU Channels. In *IEEE VTS 66th Vehicular Technology Conference*, pages 2030–2034, 2007.

[117] Akbar Sayeed and Adrian Perrig. Secure wireless communications: Secret keys through multipath. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 3013–3016, 2008.

[118] Chunxuan Ye, Suhas Mathur, Alex Reznik, Yogendra Shah, Wade Trappe, and Narayan B. Mandayam. Information-Theoretically secret key generation for fading wireless channels. *IEEE Transactions on Information Forensics and Security*, 5(2):240–254, 2010.

[119] Suhas Mathur, Wade Trappe, Narayan Mandayam, Chunxuan Ye, and Alex Reznik. Radio-telepathy : Extracting a Secret Key from an Unauthenticated Wireless Channel. In *ACM MobiCom*, pages 128–139, 2008.

[120] T. Aono, K. Higuchi, T. Ohira, B. Komiyama, and H. Sasaoka. Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels. *IEEE Transactions on Antennas and Propagation*, 53(11):3776–3784, November 2005.

[121] Babak Azimi-Sadjadi, Aggelos Kiayias, Alejandra Mercado, and Bulent Yener. Robust Key Generation from Signal Envelopes in Wireless Networks. In *Proceedings of the 14th ACM conference on Computer and communications security*, pages 401–410, 2007.

[122] Suman Jana, Sriram Nandha Premnath, Mike Clark, Sneha K. Kasera, Neal Patwari, and Srikanth V. Krishnamurthy. On the Effectiveness of Secret Key Extraction from Wireless Signal Strength in Real Environments. *Proceedings of the 15th annual international conference on Mobile computing and networking - MobiCom '09*, pages 321–333, 2009.

[123] Jessica Croft, Neal Patwari, and Sneha K. Kasera. Robust uncorrelated bit extraction methodologies for wireless sensors. In *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks - IPSN '10*, pages 70–82, New York, NY, USA, 2010.

[124] Alexandre J Pierrot and Matthieu R Bloch. Strongly Secure Communications

Over the Two-Way Wiretap Channel. *IEEE Transactions on Information Forensics and Security*, 6(3):595–605, 2011.

[125] Takayuki Shimizu, Hisato Iwai, and Hideichi Sasaoka. Physical-layer secret key agreement in two-way wireless relaying systems. *IEEE Transactions on Information Forensics and Security*, 6(3):650–660, 2011.

[126] Heng Zhou, Lauren Huie, and Lifeng Lai. Key Generation Through Two-Way Relay Channels under Active Attacks. In *Asilomar Conference on Signals, Systems and Computers*, 2013.

[127] Jon W. Wallace and Rajesh K. Sharma. Automatic secret keys from reciprocal MIMO wireless channels: Measurement and analysis. *IEEE Transactions on Information Forensics and Security*, 5(3):381–392, 2010.

[128] KillerBee. `http://code.google.com/p/killerbee/`.

[129] Andrey Garnaev and Wade Trappe. An eavesdropping game with SINR as an objective function. *Securecomm*, 19:142–162, 2009.

[130] Amitav Mukherjee and A. Lee Swindlehurst. Jamming games in the MIMO wiretap channel with an active eavesdropper. *IEEE Transactions on Signal Processing*, 61(1):82–91, 2013.

[131] Sameer Pawar, Salim El Rouayheb, and Kannan Ramchandran. Securing dynamic distributed storage systems against eavesdropping and adversarial attacks. *IEEE Transactions on Information Theory*, 57(10):6734–6753, 2011.

[132] Jie Xiong and Kyle Jamieson. SecureArray : Improving WiFi Security with Fine-Grained Physical-Layer Information. In *ACM MobiCom*, pages 441–452, 2013.

[133] Ueli M. Maurer. Authentication theory and hypothesis testing. *IEEE Transactions on Information Theory*, 46(4):1350–1356, 2000.