An Overview of Physical Layer Security with Finite Alphabet Signaling

Downloaded from: https://research.chalmers.se, 2019-08-08 22:18 UTC

Citation for the original published paper (version of record):
An Overview of Physical Layer Security with Finite Alphabet Signaling
IEEE Communications Surveys and Tutorials
http://dx.doi.org/10.1109/COMST.2018.2880421

N.B. When citing this work, cite the original published paper.

©2018 IEEE. Personal use of this material is permitted.
However, permission to reprint/republish this material for advertising or promotional purposes
or for creating new collective works for resale or redistribution to servers or lists, or to
reuse any copyrighted component of this work in other works must be obtained from
the IEEE.
An Overview of Physical Layer Security with Finite-Alphabet Signaling

Sina Rezaei Aghdam, Member, IEEE, Alireza Nooraiepour, Student Member, IEEE, and Tolga M. Duman, Fellow, IEEE

Abstract—Providing secure communications over the physical layer with the objective of achieving secrecy without requiring a secret key has been receiving growing attention within the past decade. The vast majority of the existing studies in the area of physical layer security focus exclusively on the scenarios where the channel inputs are Gaussian distributed. However, in practice, the signals employed for transmission are drawn from discrete signal constellations such as phase shift keying and quadrature amplitude modulation. Hence, understanding the impact of the finite-alphabet input constraints and designing secure transmission schemes under this assumption is a mandatory step towards a practical implementation of physical layer security. With this motivation, this article reviews recent developments on physical layer security with finite-alphabet inputs. We explore transmit signal design algorithms for single-antenna as well as multi-antenna wiretap channels under different assumptions on the channel state information at the transmitter. Moreover, we present a review of the recent results on secure transmission with discrete signaling for various scenarios including multi-carrier transmission systems, broadcast channels with confidential messages, cognitive multiple access and relay networks. Throughout the article, we stress the important behavioral differences of discrete versus Gaussian inputs in the context of the physical layer security. We also present an overview of practical code construction over Gaussian and fading wiretap channels, and discuss some open problems and directions for future research.

I. INTRODUCTION

Wireless communication technologies have become an indispensable part of our everyday lives. As more and more data is being transmitted over wireless links, along with the reliability of the transmission, ensuring security of information transfer has become a challenging issue. Maintaining confidentiality of the transmitted data, preventing corruption of the transmitted information and verifying authenticity of communication are the most important security requirements in a wireless network. Traditionally, protecting confidentiality of data transmission is addressed via computation-based mechanisms such as encryption. On the other hand, the security provided by these methods mainly relies on the conjecture that the encryption function is difficult to invert, which is not provable from a mathematical point of view. With the ever-increasing computing power, encryption may no longer prevent information leakage to sophisticated adversaries. Moreover, implementation of key-based secure communications requires complicated protocols for key distribution and management, which is highly challenging to implement in the cases of decentralized wireless networks.

Securing transmission at the physical layer is an alternative or a complement to the cryptographic solutions. The basic idea is to exploit the inherent randomness of the channel for achieving secrecy. Dissimilar to encryption based methods, in the schemes employing physical layer security, no constraint is placed on the computational capability of the eavesdropper. Furthermore, no key distribution/management is required in its implementation.

During the past decade there has been an extensive growth of interest in studying the capabilities of physical layer for securing communication. Several surveys and overview papers are available describing the state-of-the-art on this topic. For instance, fundamentals of physical layer security are comprehensively explained in [1]-[6]. A review of recent results ranging from point-to-point communications to multiuser scenarios is given in [7]. The authors in [8] present a survey on physical layer security under channel state information (CSI) uncertainty at different nodes. A comprehensive overview on various multiple antenna techniques in physical layer security is provided in [9]. The potentials of physical layer security for safeguarding data confidentiality in the fifth generation (5G) networks is studied in [10] and two updated summaries of the recent research results on physical layer security techniques for 5G and next generation wireless networks are given in [11] and [12].

The existing literature reviews provide a comprehensive understanding of the fundamentals of physical layer security. However, less is said about the practical aspects. Adoption of physical layer security techniques in the next generation networks requires a number of important practical issues to be addressed. With this motivation, a high-level overview of these issues is provided in a recent IEEE Communications Magazine paper [13]. While [13] provides the big picture, our objective in this survey is to shed light on two important (and highly related) practical aspects, namely, secure transmission with discrete signaling and practical code construction. We present a detailed review of the related recent results, underline the important behavioral differences between Gaussian inputs and discrete signaling, and study the transmit signal design problem for securing communications in a variety of scenarios. We also provide a review of recent results on code design for secrecy, and make connections with the information theoretic results with finite inputs.

The remainder of the paper is organized as follows. In Section II, we provide a description of the fundamental
II. Fundamentals

The wiretap channel model, originally introduced by Wyner in [15], is the most basic scenario in the study of physical layer security. In this model, while the sender Alice wishes to transmit a message signal to a legitimate receiver Bob; a third party, Eve, is present with the capability of eavesdropping on Alice’s signal. The channel between Alice and Bob, and the channel between Alice and Eve are referred to as the main channel and the eavesdropper’s channel, respectively. In physical layer security, one is interested in transmitting in such a way to maximize the transmission rate over the main channel while keeping the eavesdropper ignorant about the message. This leads to the notion of secrecy capacity, which is defined as the maximum rate at which the transmitter can use the main link so as to deliver its message to the legitimate receiver in a way that the eavesdropper cannot successfully obtain any information about it. Characterization of the secrecy capacity is one of the most fundamental problems in the literature of physical layer security in that it can provide us with vital implications on how secure transmission techniques can be designed.

A. Different Notions of Secrecy

Consider a transmitter who wishes to transmit a message $M$ to a legitimate receiver while trying to keep an eavesdropper ignorant about it. $M$ is mapped to a codeword $X^n$ using a stochastic encoder with $n$ denoting the number of channel uses. Then, $X^n$ is transmitted, and $Y^n$ and $Z^n$ are received at the legitimate receiver and the eavesdropper, respectively. The eavesdropper’s level of uncertainty is quantified using the equivocation rate given by

$$R_e = \frac{1}{n} H(M|Z^n),$$

which is nothing but a measure of how unlikely it is for an eavesdropper to infer source information from its received signal. According to this definition, large equivocation rates are equivalent to higher secrecy levels.

There are different metrics for measuring the secrecy guaranteed by a specific scheme among which the information-theoretic ones are recognized as the strongest. Perfect secrecy is achievable in a system if the message $M$ and its corresponding encoder output $X^n$ are statistically independent [16]. This can be expressed as $I(M;X^n) = 0$, which means that the mutual information between the message and the encoded signal is exactly zero. This is to say, the eavesdropper is not capable of recovering the message via observation of the case and uppercase bold letters. Non-bold letters are used to denote scalar values and calligraphic letters denote sets. The expectation and probability mass (or density) function of a random variable $W$ is denoted by $E_W(\cdot)$ and $P_W(\cdot)$, respectively. $Pr(\cdot)$ is used to denote the probability of an event. $A^n$ stands for a sequence of length $n$, i.e., $A^n = \{A(1), A(2), \ldots, A(n)\}$. The notation $CN(m, R)$ denotes a circularly symmetric complex Gaussian random vector with mean vector $m$ and covariance matrix $R$. 

---

Notation: Vectors and matrices are denoted with lower-case and uppercase bold letters. Non-bold letters are used to denote scalar values and calligraphic letters denote sets. The expectation and probability mass (or density) function of a random variable $W$ is denoted by $E_W(\cdot)$ and $P_W(\cdot)$, respectively. $Pr(\cdot)$ is used to denote the probability of an event. $A^n$ stands for a sequence of length $n$, i.e., $A^n = \{A(1), A(2), \ldots, A(n)\}$. The notation $CN(m, R)$ denotes a circularly symmetric complex Gaussian random vector with mean vector $m$ and covariance matrix $R$. 

---

Fig. 1: The organizational structure of the paper.
encoder output. For this form of secrecy to be guaranteed, a secret key of the length at least equal to the length of the message should be used. However, weaker requirements can be adopted for evaluation of the secrecy in practical scenarios. One example of such notions is semantic secrecy, which is achieved if it is guaranteed that $Z^n$ does not increase the probability of guessing an arbitrary function $f$ of $M$ (with an arbitrary distribution $P_M$) by more than $\epsilon > 0$ if $\text{Adv}(M; Z^n) < \epsilon$, where advantage $\text{Adv}(M; Z^n)$ is defined as the maximum of

$$\sum_z P_{Z^n}(z^n) \max_{f, \text{supp}(f)} \Pr(f(M) = f_z | z^n) - \max_{f, \text{supp}(f)} \Pr(f(M) = f_z),$$  \hspace{1cm} (2)$$

where $\text{supp}(.)$ stands for the support of the function in its argument, over $f$ and $P_M$.

Assuming a uniform distribution for the message $M$, a system is called to operate with strong secrecy if $\lim_{n \to \infty} I(M; Z^n) = 0$, which means that if we consider codewords of $n$ symbols being transmitted by Alice as $X^n$, the information leaked by observation of the received vector $Z^n$ shall go to zero as $n$ goes to infinity. On the other hand, the condition $\lim_{n \to \infty} \frac{1}{n} I(M; Z^n) = 0$ is referred to as weak secrecy. In particular, weak secrecy requires the asymptotic rate of information leakage to be sublinear in $n$. Different notions of secrecy, which are used to measure the secrecy guaranteed by a specific scheme are summarized in Table I.

<table>
<thead>
<tr>
<th>Notion</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect Secrecy [16]</td>
<td>$I(M; X^n) = 0$</td>
</tr>
<tr>
<td>Semantic Secrecy [14]</td>
<td>$\text{Adv}(M; Z^n) &lt; \epsilon$</td>
</tr>
<tr>
<td>Strong Secrecy [17]</td>
<td>$\lim_{n \to \infty} I(M; Z^n) = 0$</td>
</tr>
<tr>
<td>Weak Secrecy [15]</td>
<td>$\lim_{n \to \infty} \frac{1}{n} I(M; Z^n) = 0$</td>
</tr>
</tbody>
</table>

B. Secrecy Capacity of Single-Antenna Wiretap Channels

The notion of secrecy capacity was originally introduced by Wyner in [15] for degraded wiretap channels. In the scenario studied in [15], the main channel and the eavesdropper’s channel are assumed to be discrete memoryless channels (DMCs) where Eve’s observation of the transmitted signal is a degraded version of the signal received by Bob. Wyner showed that the secrecy capacity for this scenario can be expressed as the difference of the mutual information between Alice’s and Bob’s signals with that of Alice and Eve, maximized over all input distributions. An extension of the notion of secrecy capacity to the case of additive white Gaussian noise (AWGN) channels is studied in [18]. It is shown that, for the case of a Gaussian wiretap channel, a nonzero secrecy capacity can be attained if Bob receives the signal through a less noisy channel.

While the scenarios considered in [15] and [18] put the eavesdropper at a disadvantage, Csiszár and Körner have studied the more general problem of secrecy for non-degraded channels in [19]. In this channel model, referred to as broadcast channel with confidential messages, the message transmitted by Alice contains a common message intended to both Bob and Eve as well as a secret message, which is intended for Bob only, and is needed to be kept secret from Eve. It is shown that the secrecy capacity over a discrete memoryless wiretap channel (DMWC) is given by

$$C_s = \max_{P_{UX}, P_U} I(U; Y) - I(U; Z),$$  \hspace{1cm} (3)$$

where $U$ is an auxiliary random variable. With a given $P_{Y|Z,X}(y,z|x)$, the secrecy capacity is achieved by maximizing the difference of the mutual information terms corresponding to the main and the eavesdropper’s channels over all joint distributions $P_{UX}(u,x)$ where $U$ satisfies the Markov chain relationship $U \to X \to YZ$.

In order to study the physical layer security in wireless communication scenarios, we need to extend the wiretap channel to a model that takes into account the fading phenomenon. Assuming a narrowband transmission, fading is modeled as multiplicative gains over the main and eavesdropper’s channels, denoted by $H_B$ and $H_E$, respectively. The first characterization of the role of fading in providing physical layer security is given in [20] where quasi-static fading channels are considered towards the legitimate receiver and the eavesdropper. It is shown via analyzing the secrecy outage probability that, in presence of fading, secure transmission is possible even in the scenarios where the eavesdropper has a better signal-to-noise ratio (SNR). The authors in [21] obtain the ergodic secrecy capacity region for the fading broadcast channel with confidential messages (BCC) by deriving the boundary achieving optimal power allocation strategies. The secrecy capacity over block fading channels is derived in [22] under the assumption that each coherence interval is long enough to allow for invoking proper random coding arguments. Given that the transmitter knows either the main CSI (MCSI) or both the main and the eavesdropper CSI (ECISI) perfectly, an achievable secrecy rate can be formulated as [22]:

$$R_s = \mathbb{E}_{H_B, H_E} \left[ I(X; Y|H_B) - I(X; Z|H_E) \right]^+, \hspace{1cm} (4)$$

where $[a]^+ = \max\{a, 0\}$. It is shown in [22] that the secrecy capacity that is achieved under the full CSI assumption serves as an upper bound on the secrecy capacity when only MCSI is known at the transmitter. This is due to the fact that with the knowledge of the instantaneous ECSI, the transmitter is capable of realizing a more efficient transmission, which provides additional gains in terms of secrecy rates. For instance, when perfect knowledge of both channels are available, the transmitter can optimize the transmit power according to the instantaneous values of the channel gains $H_B$ and $H_E$, providing a higher secrecy rate with respect to the case where the optimization of the transmit power is carried out according to the main channel only.

Under the assumption that the transmitter does not know the realizations of the channels, and it only has their statistics, assuming that each receiver knows its own channel, the secrecy
capacity can be obtained with the aid of the results for the case of DMWC [19] as

$$C_s = \max_{P_{X(u),P_U}} \left( I(U;Y,H_B) - I(U;Z,H_E) \right),$$

where $U$ satisfies the Markov chain relationship $U \rightarrow X \rightarrow (Y,H_B),(Z,H_E)$. This is obtained by treating $H_B$ and $H_E$ as channel outputs at the legitimate receiver and at the eavesdropper, respectively [23]. Using the chain rule of mutual information and noting the fact that $H_B$ and $H_E$ are independent of $X$ and $U$, the expression in (5) can be modified as

$$C_s = \max_{P_{X|U},P_U} \left( \mathbb{E}_{H_B} I(U;Y|H_B) - \mathbb{E}_{H_E} I(U;Z|H_E) \right).$$

Determining the optimal joint distribution $P_{UX}(u,x)$ and the resulting exact secrecy capacity is an open problem, however, by ignoring the channel prefixing, the performance can be quantified using an achievable ergodic secrecy rate as given by

$$R_s = \left[ \mathbb{E}_{H_B} I(X;Y|H_B) - \mathbb{E}_{H_E} I(X;Z|H_E) \right]^{\dagger}. \quad (7)$$

### C. Secrecy Capacity of Multi-Antenna Wiretap Channels

Multiple-input multiple-output (MIMO) wiretap channel is an extension of the wiretap channel to a scenario where all the nodes, namely, Alice, Bob and Eve, are equipped with multiple antennas. Under the full CSI assumption, secrecy capacity for this setup is determined independently by Khisti and Wornell [24], and Oggier and Hassibi [25]. The results in [19] are employed by Khisti and Wornell so as to develop a genie-aided upper bound on the secrecy capacity in the general case of non-degraded MIMO wiretap channels. While their characterization is through a saddlepoint, Oggier and Hassibi propose an alternative approach through a single optimization process. It is shown that the achievable secrecy rate is maximized when the channel input is Gaussian, and the secrecy capacity of a MIMO wiretap channel is given by

$$C_s = \max_{K_x \in \mathbb{C}^{N_B \times N_b}, H_B \in \mathbb{C}^{N_B \times N_1}} \log \det(I + H_B K_x H_B^{H})$$

$$- \log \det(I + H_E K_x H_E^{H}),$$

where $K_x = \mathbb{E}\{xx^H\}$, and $H_B \in \mathbb{C}^{N_B \times N_1}$ and $H_E \in \mathbb{C}^{N_E \times N_1}$ are the channel matrices corresponding to the legitimate receiver and the eavesdropper, respectively.

An alternative characterization of the secrecy capacity of the MIMO wiretap channel is presented by Liu and Shamai in [26] where they consider a more general matrix constraint on the channel input. We highlight that, while [24]–[26] prove that the optimal input is Gaussian, the optimal input covariance matrix needs to be determined using numerical optimization approaches (see, e.g., [27]–[28]). A set of equations describing the optimal channel input under partial CSI at the transmitter (CSIT) along with algorithms for obtaining the solution are given in [29].

### D. Other Secrecy Metrics

Secrecy capacity is the most widely used measure for evaluation of the secrecy performance. However, a complete characterization of the secrecy capacity region is prohibitive in some scenarios. In such cases, alternative metrics can be adopted as surveyed below.

1. **Secrecy Outage Probability:** Secrecy outage occurs when the instantaneous secrecy capacity falls below a target secrecy rate $R_s^\text{c}$. In delay-critical applications and for scenarios with quasi-static fading, encoding over multiple channel states is not possible. In these cases, secure transmission schemes can be designed such that the probability of occurrence of an outage event is minimized. Obviously, this serves as a weaker notion than secrecy capacity. This is due to the fact that, with this formulation, secrecy is not guaranteed for the entire transmission duration.

2. **Error Probability Based Metrics:** In a number of works on physical layer security, secure transmission schemes are designed based on constraining the bit error rate at the legitimate receiver and the eavesdropper to pre-specified threshold values. This leads to the definition of a performance metric called “security gap,” which quantifies the minimum required difference between the received SNR values at Bob and Eve to ensure that the bit error rate (BER) at Bob is smaller than a threshold while that at the Eve’s receiver is larger than another threshold [30]. One may note that imposing a high BER at the eavesdropper does not satisfy any of the requirements given in Table I. However, security gap serves as a practical measure of secrecy in a number of applications such as explicit and implementable code design for physical layer secrecy. Some studies (e.g., [31]) consider the difference in the signal-to-interference-plus-noise ratio (SINR) at the legitimate receiver and the eavesdropper as a secrecy metric as well. This is also a weak notion, which does not guarantee secrecy in an information theoretic sense.

### III. Gaussian Wiretap Channels with Finite-Alphabet Inputs

The secrecy capacity achieving input distribution over a Gaussian MIMO wiretap channel is proved to be Gaussian [18]. However, since Gaussian signals take on a continuum of values, their detection complexity is considerably high. In addition, noting that Gaussian signals are unbounded, they have high peak to average power ratios and they cannot be generated in practice. Accordingly, Gaussian signaling is not used in practice, and instead the transmission is carried out with the aid of discrete inputs drawn from standard constellations such as phase shift keying (PSK) or quadrature amplitude modulation (QAM). Hence, it is important to study the implications of discrete signaling in the context of physical layer security.

As a first step in studying physical layer security under the finite-alphabet input assumption, the impacts of standard constellations on the achievable secrecy rates of Gaussian wiretap channels are determined in [32] and [33]. In [32], the authors evaluate the constellation-constrained secrecy capacity and highlight an important behavioral difference between
finite-alphabet (e.g., PSK and QAM) and Gaussian inputs, that is, for a fixed noise variance at Eve, the secrecy rate curves for PSK or QAM plotted against the SNR have global maxima at finite SNR values. Investigation of the secrecy capacity of the pulse amplitude modulation (PAM) inputs over a degraded Gaussian wiretap channel in [33] leads to a similar conclusion. In addition, the authors in [33] demonstrate that when finite constellations are employed, using all the available power for the information bearing signal transmission may not be optimal. The reason behind these observations has its roots in the mutual information expressions. That is, the capacity under an average power constraint over an AWGN channel is given by

\[ C = \log_2(1 + SNR) \]  

(9)

in bits per complex dimension, and it is an unbounded function of the SNR. On the other hand, the capacity of the constellation-constrained AWGN channel can be calculated as [34]

\[ I(X; Y) = \log_2 |\mathcal{X}| - E_{X,Y}\left\{ \log_2 \sum_{x' \in \mathcal{X}} p_{Y|X}(y|x') \right\} \left( \frac{p_{Y|X}(y|x)}{p_{Y|X}(y|x')} \right), \]  

(10)

where \( \mathcal{X} \) is a constellation set and \( |\mathcal{X}| \) denotes its cardinality. Dissimilar to the capacity expression in (9), the mutual information term in (10) converges to \( \log_2 |\mathcal{X}| \) at high SNRs (see Fig. 2). This means that by increasing the transmit power, the mutual information terms \( I(X; Y) \) and \( I(X; Z) \) (calculated using (10)) will reach their saturation value, and the secrecy rate drops to zero. This is also exemplified in Fig. 3, where the secrecy rates are obtained by evaluating the difference of the mutual information terms corresponding to the main and eavesdropper’s channels, under the assumption that the SNR at the eavesdropper (denoted by \( \text{SNR}_e \)) is lower than the legitimate receiver’s SNR (denoted by \( \text{SNR}_r \)) by 1.5 dB. This also clarifies that using all the available transmit power for information transmission in the high SNR regime is not optimal [33]. Furthermore, it can be observed from Fig. 3 that higher secrecy rates are achieved over Gaussian wiretap channels by increasing the modulation orders. This is because by increasing the modulation order, roughly speaking, the distribution of the channel input can be made to resemble the optimal Gaussian signal more closely for a wider range of SNRs.

The impact of output quantization on the secrecy rate of binary-input Gaussian wiretap channels is investigated in [35]. It is shown through a theoretical analysis that when the legitimate receiver has unquantized inputs and the eavesdropper has (binary) quantized outputs, higher secrecy rates are achievable with respect to the case where both receivers have binary quantized or unquantized outputs.

To conclude this section, we note that, while it may be tempting to argue that, as in the case of transmission over an AWGN channel with no security considerations, one can design general transmission schemes using Gaussian codebooks and simply adopt the ideas to the case of inputs from standard constellations (e.g., PSK, QAM, etc.), this approach does not extend in a straightforward manner to the case of physical layer security. The optimal transmit signal design strategies under the Gaussian input assumption lead to considerable losses when applied to inputs drawn from standard constellations. For instance, it is shown in [33] that the optimal power allocation strategy with PAM inputs is in sharp contrast to that of Gaussian inputs. This motivates development of new algorithms and ideas for practical constellation-constrained channels to achieve security at the physical layer.

IV. FADING AND MULTI-ANTENNA WIRETAP CHANNELS DRIVEN BY FINITE-ALPHABET INPUTS

Fading introduces new potentials for securing communications at the physical layer. Specifically, randomness due to channel fluctuations can be exploited opportunistically by a transmitter to guarantee secrecy even in the scenarios where the eavesdropper possesses a higher SNR (on average) than the legitimate receiver. Dissimilar to the Gaussian wiretap channel where Gaussian inputs are secrecy capacity achieving, in the presence of fading, discrete signaling may achieve higher...
secrecy rates. For instance, it is shown in [36] that, when Bob’s channel gain is on average worse than that of Eve’s, QAM inputs achieve higher secrecy rates at low and moderate SNR regimes. This is because the discrete nature of QAM inputs limits the leakage at the eavesdropper whose channel is unusually good. At sufficiently high SNRs, however, the secrecy rate with QAM inputs drops to zero similar to what is observed for the case of Gaussian wiretap channels.

Taking advantage of the spatial degrees of freedom introduced by multi-antenna transmission can serve as an efficient solution to prevent secrecy rate of standard constellations from dropping to zero at high SNRs. In the remainder of this section, we focus on a system model where Alice, Bob and Eve are equipped with $N_t$, $N_B$ and $N_E$ antennas, as depicted in Fig. 4. The received signals are given by

$$y = H_B x + n_y,$$

$$z = H_E x + n_z,$$  \hspace{1cm} (11) \hspace{1cm} (12)

where $H_B$ and $H_E$ are the channel matrices corresponding to the legitimate receiver and the eavesdropper, respectively, while $n_y$ and $n_z$ denote circularly symmetric complex AWGN vectors, elements of which are independent and follow $CN(0, \sigma_n^2)$ and $CN(0, \sigma_n^2)$, respectively.

Transmit precoding and artificial noise injection serve as important approaches for enhancing secrecy [7]. Via precoding, it is possible to strengthen (or weaken) the transmitted signals in certain directions. Accordingly, it can be used as an effective tool to increase the quality difference between the signals received at the legitimate receiver and the eavesdropper. Beamforming is a special case of precoding where the transmitter is restricted to using rank-one signals. It has a lower complexity with respect to the general transmit precoding solutions, achieved at the price of some performance loss. Furthermore, injection of artificial noise [37] can degrade the reception at the eavesdropper effectively while having no (or minimal) effect on the signal received at the legitimate receiver. We also note that, to take advantage of the potentials offered by channel fading, the transmitter needs a certain level of knowledge on the CSI of both channels.

In what follows, we review the existing solutions for secure transmission over MIMO wiretap channels with finite-alphabet inputs under different CSIT assumptions.

### A. Transmit Signal Design with Full CSI

The most optimistic assumption regarding the transmitter’s CSI knowledge is availability of perfect instantaneous realizations of the channel matrices $H_B$ and $H_E$. This assumption is justifiable in the scenarios where Eve is an authorized user in the network, however, she should be kept ignorant about the confidential messages transmitted from Alice to Bob. In this case, the transmitter can take advantage of the full CSI knowledge to design a precoder, which results in the largest quality difference between the signals received at the legitimate receiver and the eavesdropper.

With the aid of the perfect CSI corresponding to both channels, a generalized singular value decomposition (GSVD) precoding solution is proposed in [38] using which the MIMO multi-antenna eavesdropper (MIMOME) channel is converted into a bank of parallel channels, and a power allocation strategy is formulated to maximize the achievable secrecy rate.

The necessary conditions for optimum transmit precoding is derived in [39], and it is demonstrated that the GSVD precoding [38] is suboptimal. Alternatively, a gradient descent optimization is proposed in which the precoder matrices are updated with steps proportional to the gradient of the instantaneous secrecy rate. Furthermore, it is shown that transmission along the null-space of the eavesdropper’s channel is the optimal secure transmission strategy at the high-SNR regime. This is because, with a precoder matrix along the null-space of the eavesdropper’s channel, her achievable rate is suppressed to zero, and when the SNR is sufficiently high, the rate at the legitimate receiver approaches the maximum value that can be achieved with the use of specific finite-alphabet inputs, which guarantees that the secrecy rate is maximized. However, it should be noted that, when the number of transmit antennas is less than or equal to the number of antennas at the eavesdropper, it is not possible to suppress the Eve’s information rate by transmitting along the null-space of $H_E$. Therefore, to maximize the secrecy rates, only part of the available power should be allocated to data transmission. Given $N_t > N_B$, the authors in [39] suggest to use the excess power for transmission of an artificial noise signal along the null-space of $H_B$ so as to prevent the secrecy rates from dropping to zero.

Fig. 5 compares the achievable secrecy rates by the schemes proposed in [38] and [39] for a wiretap channel with $N_t = 2$, $N_B = 1$ and $N_E = 1$, and fixed channels given by

$$h_B = \begin{bmatrix} 0.5128 - 0.3239j & -0.8903 - 0.0318j \end{bmatrix},$$

$$h_E = \begin{bmatrix} 0.3880 + 1.2024j & -0.9825 + 0.5914j \end{bmatrix}.$$  \hspace{1cm} (13) \hspace{1cm} (14)

The precoding approach proposed in [39] outperforms the GSVD-based precoding [38]. This improved performance is

\^This example has also been considered in [39] and [40].
MISO multi-antenna eavesdropper (MISOME) channels, the optimal precoders for Gaussian inputs. For example, over finite-alphabet input assumption is different from that of the channels, the structure of the optimal precoders under the results in [38] and [39] is that, as in the case of non-fading channel reveals the advantages offered by fading and CSI.

Furthermore, comparing the results given in Fig. 5 with those of Fig. 3 for QPSK inputs over the Gaussian wiretap channel reveals the advantages offered by fading and CSI-based precoding approaches.

An important conclusion that can be drawn from the results in [38] and [39] is that, as in the case of non-fading channels, the structure of the optimal precoders under the finite-alphabet input assumption is different from that of the optimal precoders for Gaussian inputs. For example, over MISO multi-antenna eavesdropper (MISOME) channels, the optimal transmission strategy is beamforming along the eigenvector corresponding to the largest generalized eigenvalue of \((H_0, H_e)\) [40]. However, these precoders (which correspond to signaling with rank one covariance), when applied to finite-alphabet inputs, undergo a considerable loss with respect to the precoders specifically optimized for this case (see, e.g., Fig. 7 in [39]).

As an alternative method to design a suboptimal precoder, one may also replace the secrecy capacity with practical metrics (as opposed to those ensuring secrecy in an information-theoretic sense) such as SINR-based metrics. For instance, the authors in [41] design artificial noise beamformers, which serve as constructive interference to the legitimate receiver (improving Bob’s SINR) while disrupting the reception at the eavesdropper (degrading Eve’s SINR). In Table II, we summarize the precoder design approaches with full CSI.

### B. Perfect MCSI and Statistical ECSI

While availability of perfect ECSI at the transmitter may be practical in some limited scenarios, in general, it is highly challenging to obtain it instantaneously. A more practical assumption regarding CSIT is availability of perfect MCSI along with the statistical ECSI at the transmitter. On the other hand, the precoder design in this scenario is not as effective as those carried out with the perfect knowledge of both channels. In other words, with the absence of the instantaneous ECSI, precoding is not as forceful in suppressing the reception at the eavesdropper. For instance, transmission along the null-space of the eavesdropper’s channel is not even possible in this scenario.

An example of secure transmission schemes under perfect MCSI and statistical ECSI is given in [43] where the authors consider a MISOME channel and define a practical secrecy metric (instead of an information theoretic one), which quantifies the symbol error probability of the confidential data, and show that in the absence of artificial noise, secrecy diversity (i.e., the high-SNR slope of their defined metric) vanishes. This result underlines the importance of artificial noise injection in these scenarios.

Artificial noise-aided transmit precoding strategies with the objective of maximizing the secrecy rate are proposed in [44]–[47]. In [44], the authors employ naive beamforming along with the artificial noise injection while considering single-antenna receivers. The strategy proposed in [39], on the other hand, relies on an iterative maximization of an approximation to the instantaneous secrecy rate. In both of these studies, it is shown that the optimal schemes allocate only a fraction of the total power for signal transmission at high SNRs. Hence, the remaining power is employed for artificial noise injection. It is shown in [45] that jointly optimizing the precoder matrix and portion of the power allocated to the artificial noise outperforms the solutions which rely on optimizing the precoder only. Moreover, inspired by the idea proposed in [46], a generalized artificial noise aided transmission is introduced in [47], which guarantees high secrecy rates for the scenarios with \(N_t < N_E\) for which injection of artificial noise along the null-space of the main channel is not possible. A summary of the existing solutions for the scenarios with perfect MCSI and statistical ECSI at the transmitter is given in Table III.

### C. Perfect MCSI Only

In the scenarios with a passive eavesdropper, a realistic assumption is that the transmitter does not know anything about the eavesdropper’s channel. Under this assumption, [48] and [49] propose a secure transmission strategy referred to as directional modulation, in which the amplitude and the phase of the transmit signal are adjusted by varying the length of the reflector antennas for each symbol. This scrambles the PSK symbols in all the directions other than that of the legitimate receiver. Other strategies for securing communications without the knowledge of the eavesdropper’s channel are proposed in [50] and [52]. Zhang et al. develop a Tomlinson-Harashima precoding in [50] where the transmitter allocates part of its power in order to achieve a target mean squared error (MSE) at the legitimate receiver, and the remaining power is used to transmit artificial noise to degrade the eavesdropper’s reception. In [52], a secure space-time block coding (STBC) scheme is proposed, which enables the legitimate receiver to perform separable decoding, while requiring an exhaustive maximum likelihood (ML) detection at the eavesdropper. Furthermore, the authors combine this scheme with artificial noise injection to ensure a high uncoded BER at Eve. In [53], the authors propose a practical precoded orthogonal STBC method for MISO wireless networks where space-time codewords are...
Preceded by a matrix that minimizes the error rate at the legitimate receiver and a pre-equalization step is added to take security requirements into account. Table IV summarizes the existing transmit signal design approaches under perfect MCSI and no ECSI at the transmitter.

D. Statistical MCSI and ECSI

While most of the existing physical layer security solutions rely on the assumption that the transmitter is capable of estimating at least the instantaneous main CSI, in some scenarios (e.g., for fast fading channels), it may be difficult for the transmitter to track the rapidly varying channel coefficients. For these cases, the impact of the discrete inputs on the achievable secrecy rates is analyzed in [54]. Under the assumption that both channels are doubly correlated, using the knowledge of transmit and receive correlation matrices, transmit signal design algorithms are proposed in [55] to maximize the resulting achievable secrecy rates. The results in [55] reveal that jointly optimizing the precoder matrix and artificial noise results in increased achievable secrecy rates with respect to precoding without artificial noise injection. To existing studies on MIMO wiretap channels with statistic CSI are summarized in Table V.

We now compare the secrecy rates achieved by different transmit signal design approaches under different CSIT assumptions. In particular, we overview the results for scenarios with perfect or statistical CSI at the transmitter. When perfect CSI corresponding to the main channel or the eavesdropper’s channel, a commonly adopted model in the literature employs doubly correlated channels with

\[ \mathbf{H}_B = \mathbf{H}_{tb}^{1/2} \mathbf{H}_{tb} \mathbf{H}_{tb}^{1/2}, \quad \mathbf{H}_E = \mathbf{H}_{te}^{1/2} \mathbf{H}_{te} \mathbf{H}_{te}^{1/2}, \]

where the elements of \( \mathbf{H}_{tb} \) and \( \mathbf{H}_{te} \) are i.i.d., and the transmitter is capable of acquiring the transmit and receive correlation matrices, i.e., \( \mathbf{H}_{tb} \), \( \mathbf{H}_{te} \), \( \mathbf{H}_{rb} \) and \( \mathbf{H}_{re} \), from the long-term statistics of the channel. Perfect instantaneous CSI is available at both receivers.

Fig. 6 demonstrates the achievable secrecy rates by different transmission schemes over a MIMOME wiretap channel with \( N_t = 4 \) and \( N_B = N_E = 2 \). Correlation matrices are assumed to have exponentially decaying entries as

\[ [\mathbf{H}_{tb}(\rho)]_{ij} = \rho^{i-j}, \quad i, j = 1, 2, \ldots, N_t, \]

\[ [\mathbf{H}_{te}(\rho)]_{mk} = \rho^{m-k}, \quad m, k = 1, 2, \ldots, N_B \] (or, \( N_E \)).

It is demonstrated that the precoder design using a gradient descent optimization approach [39], [47] is a promising strategy for maximizing the secrecy rates. Fig. 6 also reveals the importance of CSIT. When perfect CSI corresponding to the eavesdropper is available at the transmitter, it is possible to construct precoders, which are aligned with the null-space of the eavesdropper’s channel. By this means, the achievable rate of the eavesdropper can be suppressed to zero. Furthermore, even when only the statistical CSI of the eavesdropper along

<table>
<thead>
<tr>
<th>Authors</th>
<th>System Model</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. Bashar et al. [38]</td>
<td>MIMOME</td>
<td>Proposes a GSVD-based precoding algorithm.</td>
</tr>
<tr>
<td>Y. Wu et al. [39]</td>
<td>MIMOME</td>
<td>Derives the necessary conditions for optimality of the precoder matrix and</td>
</tr>
<tr>
<td>M. R. A. Khandaker et al. [41]</td>
<td>MISO, multiple</td>
<td>develops a gradient descent based precoder optimization algorithm.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Designs artificial noise which is constructive to the legitimate receiver and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>destructive to the eavesdroppers.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Authors</th>
<th>System Model</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. V. Nguyen et al. [43]</td>
<td>MISO</td>
<td>Defines a weak notion of secrecy based on symbol error probability and quantifies the impact of artificial noise injection with beamforming on diversity order.</td>
</tr>
<tr>
<td>S. Bashar et al. [44]</td>
<td>MISOSE</td>
<td>Employs a naive beamforming with the aid of MCSI only along with a power optimization using perfect MCSI and statistical ECSI. The excess power is used for artificial noise injection in the null-space of ( \mathbf{H}_B ).</td>
</tr>
<tr>
<td>Y. Wu et al. [39]</td>
<td>MIMOME</td>
<td>Develops an iterative algorithm for precoder design and allocate the excess power to inject artificial noise along the null-space of ( \mathbf{H}_B ).</td>
</tr>
</tbody>
</table>

Fig. 6: Secrecy rates with different transmit signal design algorithms under different CSIT assumptions.
with perfect CSI of the main user is available, high secrecy rates are still achievable with the aid of appropriately designed artificial noise injection strategies. However, in the absence of instantaneous CSI, the precoder matrices designed with the knowledge of correlation matrices are not capable of preventing the ergodic secrecy rates from dropping to zero at high SNRs. In other words, since statistical CSI does not provide sufficient degrees of freedom to take advantage of fading for improving secrecy, the behavior is similar to that in AWGN channels described in Section III.

E. Transmit Signal Design Based on Channel Reciprocity

A group of transmission strategies in the literature of physical layer security rely on the assumption of channel reciprocity, that is, the transmitter and the legitimate receiver observe the same channel, simultaneously. For systems working in time division duplex (TDD) mode, this property can be used to enable the transmitter to obtain the MCSI using pilot signals transmitted from the legitimate receiver. In [55], a secure transmission scheme is proposed over a MIMO wiretap channel under the assumption that no training signals are transmitted by Alice, and hence, no CSI is available at Bob or Eve. Alice realizes an orthogonal space-time block coded PSK transmission and uses the CSI she obtains from the reverse channel estimation to design a phase shifting precoder, which compensates the phase shift by the channel. Therefore, Bob can recover the transmitted messages with no need to CSI. On the other hand, Eve who observes an independent channel can only recover the information with non-coherent detection, ensuring positive secrecy rates over MIMO wiretap channels.

In the scenarios with channel reciprocity, when channel estimation is carried out both in the forward (from Alice to Bob) and the reverse (from Bob to Alice) directions, the channel between Alice and Bob can serve as a source of common randomness, which can further improve secrecy. This common randomness is used in [57] for generation of a random phase sequence, which is then used to manipulate the transmitted symbols (by simply multiplying them with the obtained random phase terms). Since the eavesdropper is unable to find out these random phase values, an enhanced secrecy is attained. In [58], the authors propose a secure Alamouti scheme using channel reciprocity. In their proposed approach, after performing a two-way probing, the legitimate users exchange a secret key based on their received signal strength indicator, and this key is used for rotation of the PSK constellation from one codeword to the next. A cross-layer analysis of secure MIMO transmission is given in [59]. It is shown that when using the main channel as a (not necessarily secret) key between the legitimate users, complexity of the eavesdropper’s decoding of the legitimate receiver’s message is as hard as solving standard lattice problems, and it grows exponentially in the number of transmit antennas.

F. Performance-Complexity Trade-off in Transmit Signal Design

The iterative algorithms proposed in [39] and [47] for optimizing the precoder matrix are shown to offer promising results in terms of the achievable secrecy rates. However, their implementation complexity may be too high for some practical applications. There are two main reasons for this high computational complexity. First, both the mutual information and its gradient lack closed-form expressions. Therefore, Monte Carlo simulations need to be utilized to evaluate them, which require averaging over a large number of noise realizations in order to maintain a high level of accuracy. More importantly, evaluation of the mutual information and the minimum MSE (MMSE) term involves additions over the modulation signal space, which grows exponentially with the number of transmit antennas, resulting in a prohibitively high computational complexity when the number of transmit antennas is large.

Different strategies are proposed for lowering the computational complexity associated with the transmit signal design for secrecy under the finite-alphabet input assumption. The authors in [39] and [45] derive closed-form approximations for the secrecy rate expressions using bounds on the mutual information. While the former employs the bounds given in [60], the latter approximates the mutual information using the cut-off rate expression. The authors in [61]-[62] derive

<table>
<thead>
<tr>
<th>Authors</th>
<th>System Model</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Kalani et al. [48], [49]</td>
<td>MIMOME</td>
<td>Steers the array beam such that the phase of the PSK modulated signal is scrambled in directions other than Bob’s.</td>
</tr>
<tr>
<td>L. Zhang et al. [50]</td>
<td>MIMOME</td>
<td>Proposes a nonlinear Tomlinson-Harashima precoding along with along with artificial noise injection.</td>
</tr>
<tr>
<td>S. A. A. Fakoorian et al. [52]</td>
<td>MIMOME</td>
<td>Designs a STBC which results in a high BER for Eve and a low BER for Bob.</td>
</tr>
<tr>
<td>J. H. Hamamreh et al. [53]</td>
<td>MISOSE</td>
<td>Proposes a precoded orthogonal STBC which minimizes BER at Bob and is capable of satisfying the security requirements.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Authors</th>
<th>System Model</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. Girmay et al. [54]</td>
<td>MIMOME</td>
<td>Derives a large system approximation for secrecy rates under arbitrary channel inputs and quantifies them for Gaussian and finite-alphabet inputs.</td>
</tr>
<tr>
<td>S. Rezaei Aghdam and T. M. Duman [55]</td>
<td>MIMOME</td>
<td>Proposes precoding and artificial noise injection approaches with the knowledge of correlation matrices.</td>
</tr>
</tbody>
</table>
asymptotic secrecy rates in a regime where the numbers of antennas at the transmitter and both receivers grow infinitely large with a constant ratio. Maximization of this expression is shown to yield a satisfactory performance even in the cases of small number of antennas. The authors in [64] and [65] propose per-group precoding strategies under perfect and partial CSIT assumptions, respectively. In these works, the channels are diagonalized, and the precoder matrices are designed by grouping the transmit antennas and designing the precoder matrix for each group, independently.

G. Secure Spatial Modulation

Spatial modulation and space shift keying (SSK) are relatively new MIMO transmission schemes, which rely on transmitting information via active antenna indices. Similar to the amplitude or phase modulation schemes, in spatial modulation and SSK, the channel inputs are drawn from discrete and finite sets. More specifically, in spatial modulation [64], the data bits are mapped onto two information carrying blocks: 1) an antenna index selected from the set of transmit antennas, and 2) a symbol drawn from a standard constellation. Space shift keying is a special case of spatial modulation where no amplitude of phase modulation is employed, i.e., the transmit antenna index serves as the only information carrying unit [65].

Different studies explore the advantages offered by spatial modulation over state-of-the-art MIMO schemes (see [66] for a comprehensive overview). Along with different application areas of spatial modulation and SSK, studying these transmission schemes in the context of physical layer security has been of recent interest. Secrecy capacity of spatial modulation for the scenarios with two transmit antennas is studied in [67], and the results are extended to MISOSE and MIMOME scenarios in [68] and [69], respectively.

Different secure spatial modulation based transmission schemes are also proposed with the aid of MCSI and ECSI at the transmitter, and for systems with channel reciprocity. A precoder optimization algorithm for maximizing secrecy rates with SSK transmission using full CSI is proposed in [70]. The authors in [71] provide a precoder design and artificial noise injection approach for securing spatial modulation for the scenarios with passive eavesdroppers (with no ECSI at the transmitter). The secrecy rates of spatial modulation with artificial noise injection are characterized in [72]. While the conventional artificial noise injection (similar to what is considered in [72]) requires multiple antennas to be activated, a novel secure artificial noise aided transmission is introduced in [73], which requires only two antennas to be activated at each time instance. Existing secure transmit signal design approaches for spatial modulation and SSK are summarized in Table VI.

For the scenarios where channel reciprocity holds, a transmit preprocessing technique is proposed in [74] under the assumption that the training sequences are transmitted from Bob to Alice, and no CSI is available at the receiver nodes. The authors in [75] propose a dynamic antenna index assignment based on the main channel when training is carried out at both directions, which prevents an eavesdropper from obtaining the antenna index assignment pattern.

V. Multi-Carrier Wiretap Channels with Discrete Inputs

A. OFDM Wiretap Channel

Orthogonal frequency-division multiplexing (OFDM) provides robustness against multipath channel fading and offers flexibility in resource allocation. These features have made OFDM a widely used technique in recent wireless standards including 4G and 5G wireless systems. If there are no time variations across an OFDM word (i.e., for slow fading), the system can be modeled as a set of independent parallel channels. It is proved in [77] that the secrecy capacity of \( K \) independent parallel channels (as depicted in Fig. 7) is achieved when each channel achieves its individual secrecy capacity. For parallel Gaussian wiretap channels, the optimal input is Gaussian distributed [78]. However, a more practical study of OFDM wiretap channel is possible by assuming that the data symbols belong to finite constellations.

An initial study on parallel Gaussian wiretap channels under finite-alphabet inputs is reported in [33] where the authors characterize the secrecy capacity with PAM signaling. They also provide a mercury-waterfilling interpretation of the optimal power allocation strategy. It is observed in [33] that dissimilar to standard parallel Gaussian channels (with no constraints on the input distribution), it may not be optimal to use all the available power for PAM inputs. In [79], the OFDM wiretap channel is studied with QAM inputs. After quantifying the loss in the secrecy rates with respect to the scenarios with Gaussian inputs, a bit loading strategy is proposed to minimize this loss by assigning an appropriate number of bits to each subchannel.

The sensitivity of OFDM based transmissions to frequency synchronization errors can be exploited to achieve secure transmission as well. The authors in [80] propose a secure

![Fig. 7: OFDM wiretap channel modeled as set of independent parallel channels. \( K \) : number of subcarriers (i.e., the number of parallel channels).](image-url)
OFDM transmission over reciprocal channels, which relies on carrier frequency offset that is pre-compensated (using the transmitter’s knowledge on the MCSI) in such a way that it is received without inter-carrier interference at the legitimate receiver, while the reception at the eavesdropper receiving the signal through an independent channel is considerably degraded. In [81], an eavesdropping resilient OFDM scheme is developed over reciprocal channels where Alice performs a subcarrier interleaving using her knowledge on the instantaneous MCSI. The legitimate receiver who is also capable of acquiring the MCSI can obtain the interleaving pattern and de-interleave the received signals. In contrast, the eavesdropper cannot derive the interleaving pattern initiated by the transmitter, and as a result, fails to acquire the transmitted information.

Cyclic prefix is used in OFDM systems to avoid inter carrier interference (ICI) incurred by multi-path channels. However, using cyclic prefix makes the transmission vulnerable to cyclostationarity based attacks [51]. This temporal degrees of freedom offered by the cyclic prefix can be used for artificial noise injection so as to improve secrecy over OFDM wiretap channels. In [82], a single-antenna OFDM wiretap channel is considered and an artificial noise signal is inserted into the time-domain signal (over the cyclic prefix samples). A framework for joint optimization of the power allocated to each subcarrier and artificial noise covariance matrix is also proposed. In [83], Akitaya et al. study the application of time-domain artificial noise injection in MIMOME OFDM systems. The results in [82] and [83] demonstrate that this approach is efficient for improving the secrecy rates in OFDM wiretap channels.

Efficient resource allocation for secure OFDM systems is formulated as a convex optimization problem in [84], and solved assuming that the transmitter uses artificial noise to combat an eavesdropper equipped with multiple antennas. Li et al. [85] propose scrambling of data in the time domain after the IFFT operation in order to improve physical layer security for OFDM signals. This, in fact, corresponds to sending modified constellation points (in terms of phase and amplitude) on each sub-carrier. The authors in [88] introduce channel frequency response based pre-coder and post-coder designs for enhancing security in OFDM systems. Their solution relies on extracting orthonormal matrices via singular value decomposition (SVD) of the legitimate user’s channel frequency amplitude and using it for shuffling OFDM sub-channels as well as for designing the frequency-based precoder and post-coder. In [89], physical layer security is studied for passive OFDM-based optical networks where a scrambled precoding technique is utilized to improve the security. The authors in [90] propose an efficient approach to enhance secrecy in OFDM systems that relies on employing an adaptive interleaver maximizing the overall diversity gain delivered to the legitimate receiver. Hamamreh et al. [86] propose an efficient physical layer security technique for transmission of OFDM-based waveforms by developing an optimal joint subcarrier index selection and adaptive interleaver design to enhance the security and reliability of 5G services. In order to provide QoS based security for OFDM systems, channel shortening (CS) equalizer coefficients are designed in [91] based on the legitimate receiver’s channel. Specifically, CS is used at the transmitter in such a way that the effective channel ensures no ISI at Bob, while causing ISI and performance degradation at Eve. More recently, application of deep learning for physical layer security in multi carrier systems is studied in [87] where an adversary aims at sending spurious data to the legitimate receiver.

Table VII summarizes important results on the study of OFDM wiretap channels under the finite-alphabet input assumption.

### B. Filter-Bank Multicarrier Wiretap Channel

OFDM systems face important challenges when they are adopted for more complex networks. For instance, it is very difficult to establish full synchronization for the signals corresponding to different users in the base station in an orthogonal frequency division multiple access (OFDMA) network. Furthermore, OFDM introduces significant out of band interference to other users when it is used to transmit over a set of non-contiguous frequency bands as in cognitive radio where transmission is limited to certain portions of the band and 5G wireless systems [92].

Filter bank multicarrier (FBMC) modulation is shown to be able to resolve the above problems by utilizing high quality filters for transmission, which makes it a potential candidate for future networks including 5G. With this motivation, recently, there has been some interest in the physical layer security aspects of FBMC based networks, as well. In [93], the authors

<table>
<thead>
<tr>
<th>Paper</th>
<th>System Model</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. Sumanovic et al. [67]</td>
<td>( N_t = 2, N_M = N_E = 1, ) No CSI at the transmitter</td>
<td>Characterizes secrecy rates for spatial modulation and SSK transmission.</td>
</tr>
<tr>
<td>X. Guan et al. [68]</td>
<td>MISOE</td>
<td>Evaluates the secrecy rates by deriving mutual information terms and proposes a precoding with the aid of full CSI.</td>
</tr>
<tr>
<td>S. Rezaei Aghdam and T. M. Duman [69]</td>
<td>MIMOME, No CSI at the transmitter</td>
<td>Derives expressions for achievable secrecy rates with spatial modulation and SSK schemes.</td>
</tr>
<tr>
<td>S. Rezaei Aghdam and T. M. Duman [70]</td>
<td>MIMOME, Full CSI at the transmitter</td>
<td>Develops an iterative precoder design algorithm.</td>
</tr>
<tr>
<td>F. Wu et al. [71]</td>
<td>MIMOME, ( N_r &gt; N_M, ) MCSI at the transmitter</td>
<td>Proposes a precoder and artificial noise design algorithm.</td>
</tr>
<tr>
<td>L. Wang et al. [72]</td>
<td>MIMOME, ( N_r &gt; N_M, ) MCSI at the transmitter</td>
<td>Characterizes the secrecy rates of artificial noise aided spatial modulation.</td>
</tr>
<tr>
<td>Z. Huang et al. [73]</td>
<td>MIMOME, MCSI at the transmitter</td>
<td>Proposes an artificial noise aided secure quadrature spatial modulation with a low hardware complexity.</td>
</tr>
</tbody>
</table>

**Table VI: Transmit signal design for secure spatial modulation and SSK.**
present a filter hopping FBMC approach for security. They argue that the choices of filters for each symbol provide a major degree of freedom for the FBMC waveforms, which can be exploited for security purposes. Furthermore, a mismatch in the receive filter results in violation of the orthogonality of different subcarriers, and it causes the energy of a transmit symbol at a certain subcarrier to spread into the adjacent subcarriers. This effect is shown to be significant even for a slight mismatch of the filters between the transmitter and the receiver [93], hence it can be exploited for providing physical layer security. For instance, in a practical system, Alice and Bob can exchange their filter designs by exchanging keys or using a pre-shared sequence of filters. Then, the transmitter can change the filters continuously, and make it extremely difficult for an eavesdropper to find the exact filter being used, hence securing the transmission.

VI. SPREAD SPECTRUM TECHNIQUES FOR PHYSICAL LAYER SECURITY

One of the potential methods for improving physical layer security is spread spectrum (SS) communications. In this method, a signal is spread over a frequency band such that the resulting bandwidth is much wider than that of the original signal. Historically, this method has been first used in military applications to achieve transmission schemes with low probability of intercept (LPI). Two main approaches in SS communications are direct sequence spread-spectrum (DSSS) and frequency-hopping spread-spectrum (FHSS).

FHSS involves hopping the carrier frequency of the outgoing signal in a predetermined fashion, which is known to both the transmitter and the receiver [94]. In this way, it is extremely difficult for an eavesdropper to monitor the resulting spread spectrum signal. In DSSS, the outgoing symbol stream gets multiplied by a symbol sequence called spreading code which has a much smaller symbol duration, hence increasing the bandwidth of the resulting signal. For the case where multiple users share the same channel, they utilize different spreading codes to distinguish their signals. In order to avoid an eavesdropper to detect the signal, the spread signal can be scrambled using a pseudo-noise (PN) sequence [95]. One of the main differences between DSSS and traditional cryptography methods (where the key space is very large) is that the key space is limited by the available bandwidth.

Physical layer security of code division multiple access (CDMA) systems is studied in [96] based on the use of relatively long PN scrambling sequences. In most cases, Walsh codes are used to generate channelization codes for this purpose as they are easy to generate. The authors in [96] propose a method to generate the scrambling sequences by using advanced encryption standard (AES) operation. Specifically, they design an AES-CDMA scheme secured against exhaustive key-search attacks utilizing AES with three different key sizes up to 256. We close this discussion by noting that, while valid in practice, such spread spectrum based solutions do not provide information theoretically (provable) secure communications, unlike the physical layer security algorithms being reviewed.

VII. CONSTELLATION CONSTRAINED MULTIUSER AND COOPERATIVE WIRETAP CHANNELS

The concept of physical layer security has also been extended to multi-user and cooperative networks in recent years. In addition to the numerous studies on the limits of secure communications for scenarios with more than two receivers, multiple transmitters and presence of relays/jammers/helpers (see [7] for a survey), some recent focus has also been placed on the practical constellation-constrained secure transmission in these settings.

A. Multiuser and Multi-Eve Networks

1) Broadcast Channel with Confidential Messages: The achievable rate region for a Gaussian broadcast channel with confidential messages is characterized for discrete inputs in [97]. In this scenario, the transmitter’s objective is to send a common message to two receivers as well as to deliver a confidential message to one of them. Dissimilar to [32] and [33] where standard constellations are considered, in [97], the symbols are allowed to be arbitrary, and the achievable secrecy rate region is enlarged by optimizing the symbol positions and the probability distribution of the symbols. The shrinkage of the rate region due to employing PAM inputs is quantified with respect to the optimal Gaussian inputs, and a number

<table>
<thead>
<tr>
<th>Paper</th>
<th>Model</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. R. D. Rodrigues et al. [33]</td>
<td>Parallel Gaussian Wiretap Channels</td>
<td>Characterizes secrecy rates with PAM inputs and derive optimal power allocation strategy.</td>
</tr>
<tr>
<td>F. Renna et al. [79]</td>
<td>SISOSE</td>
<td>Evaluates the secrecy rates for OFDM systems with QAM input constellations and provides a bit-loading strategy.</td>
</tr>
<tr>
<td>M. Yusuf et al. [80]</td>
<td>SISOSE with reciprocal main channel</td>
<td>Suggests inducing carrier frequency offset which is pre-compensated at the legitimate receiver.</td>
</tr>
<tr>
<td>H. Li et al. [81]</td>
<td>SISOSE with reciprocal main channel</td>
<td>Proposes subcarrier interleaving based on the MCSI which prevents eavesdropper from de-interleaving the received signals.</td>
</tr>
<tr>
<td>H. Qin et al. [82]</td>
<td>SISOSE</td>
<td>Introduces a time-domain artificial noise injection approach and proposes a joint optimization of data power and artificial noise injection covariance matrix.</td>
</tr>
<tr>
<td>M. Yusuf et al. [90]</td>
<td>SISOSE with reciprocal main channel</td>
<td>Employs an interleaving whose pattern is adapted to the main channel.</td>
</tr>
<tr>
<td>J. M. Hamanureh et al. [86]</td>
<td>SISOSE with reciprocal main channel</td>
<td>Develops a joint optimal subcarrier index selection and adaptive interleaving design to enhance the security and reliability</td>
</tr>
</tbody>
</table>
of important behavioral differences between these cases are highlighted.

2) Multiple Access Channels: The authors in [98] focus their attention on multiple access wiretap channels (as depicted in Fig. 9) under the assumption that the users employ the Alamouti STBC scheme. Specifically, they consider a case where multiple transmitters communicate with one legitimate receiver in the presence of an eavesdropper. Under the assumption that the transmitter has the knowledge of the legitimate user’s channel and does not know the eavesdropper’s channel gain, an artificial noise injection is proposed along the null-space of the main channel, which is shown to provide considerable improvements in the secrecy sum-rate.

3) Cognitive Radio Networks: The linear precoder design for the cognitive multiple-access wiretap and the cognitive multi-antenna wiretap channels with finite-alphabet inputs are studied in [99] and [100], respectively. In [99], the authors consider a setup in which two secondary-user transmitters are communicating with a secondary-user receiver in the presence of an eavesdropper and under interference threshold constraints at the primary users. A two-layer precoding algorithm is proposed using statistical CSI at the transmitters to maximize the ergodic secrecy sum rate. In [100], on the other hand, the authors study a scenario where a multi-antenna secondary-user transmitter communicates with a multi-antenna secondary-user receiver, and the communication is wiretapped by a multi-antenna eavesdropper (Fig. 10). In this work, the precoder matrix is optimized using an iterative algorithm (assuming statistical CSI of all the channels) so that the secrecy rate of the secondary user is maximized while power leakage to the primary user sharing the same frequency spectrum is controlled.

4) Multi-Eavesdropper Scenarios: The authors in [101] study a setup comprised of a source, a legitimate receiver and multiple multi-antenna eavesdroppers. They first obtain an expression for an achievable secrecy rate, and then, under the full CSI assumption, they optimize the transmit power and the beamforming vector. The problem of power allocation for secrecy over MIMO wiretap channels with multiple eavesdroppers is studied in [102]. Under the assumption that the transmitter has perfect MCSI and statistical ECSI, the proposed power allocation strategy gives non-zero secrecy rates at high transmit powers. This is important because in the absence of this power control strategy, secrecy rate decreases with increasing transmit power, and it drops to zero when the eavesdroppers’ SNRs become higher than a certain value.

B. Relay Channels and Cooperative Communications

Cooperation serves as an efficient method for enhancing secrecy in wireless networks. In this context, along with the studies conducted on secrecy capacity in various cooperative scenarios, a number of recent studies focus on finite-alphabet inputs. In this context, the secrecy rates achieved with decode-and-forward (DF) relay beamforming under the finite-alphabet input constraints are characterized in [103]. It is shown that by optimizing the source power and also the relay weights using the global CSI, it is possible to prevent the secrecy rates from dropping to zero. The authors in [104] consider a cooperative jamming network in presence of multiple eavesdroppers, and develop a secure transmission scheme with finite-alphabet inputs, which relies on a joint optimization of the artificial noise (injected in the null-space of the relay to destination links), and power allocation among the source and the relays. The study in [105] considers secure transmission over a MIMOME setup with the aid of a multi-antenna helper, which transmits a jamming signal to degrade the signal received by the eavesdropper. Similar to [104], a framework is provided for joint optimization of the precoder matrix and power allocation between the transmitter and the helper. Furthermore, low-complexity transmit signal design schemes are developed for both low and high SNR regimes. The authors in [106] focus on the problem of secure transmission with cooperative jamming via an untrusted relay. They propose a power efficient secure DF-based cooperative communication technique and demonstrate the efficacy of their scheme via simulations using...
Table VIII summarizes the recent research results of secure transmission over finite-input multi-user, multi-eavesdropper and cooperative channels.

VIII. PHYSICAL LAYER SECURITY FOR NEXT GENERATION WIRELESS SYSTEMS

With the advent of the 5G and beyond wireless communication systems, a tremendous growth is expected in the number of connected devices. To meet the high traffic demand, various technologies have emerged during the past years. For instance, massive MIMO [107], millimeter wave (mmWave) communication systems [108], full duplex transmission [115] and non-orthogonal multiple access (NOMA) [116] have attracted significant interest from researchers both in industry and academia.

Exploiting large antenna arrays and taking advantage of a spectrum from 30 GHz to 300 GHz away from the almost fully occupied spectral band have proven to provide various benefits, which make these technologies among the key enablers for the next generation wireless systems and beyond. Along with the different aspects of massive MIMO and mmWave communications, investigating their potential in providing physical layer security has been of recent interest (see e.g., [109]-[114]). The problem of linear precoder design for large-scale (massive) MIMO wiretap channel under the finite-alphabet input assumption is investigated in [63] where GSVD-based precoders are shown to achieve the maximal secrecy rate at high SNR values. The potential of large antenna arrays to realize secure directional modulation at mmWave frequencies is demonstrated in [117] where the proposed solution relies on driving only a subset of antennas in the array and choosing the switching configuration such that the desired modulation symbols are received at the legitimate receiver while scrambling the constellation in other (undesired) directions. The authors in [118] propose an antenna selection-aided secure mmWave transmission using a switched phased array architecture in large-scale transmit antenna arrays. We note that, different from the solution proposed in [117], which requires modulation at the RF end, the scheme in [118] can use traditional baseband modulation schemes.

Offering several advantages such as high throughput, low latency and improved connectivity, NOMA is one of the key enabling technologies for next generation wireless networks. It relies on removing the orthogonality (which exists in the conventional practical multiple access techniques), and assigning a specific resource block (e.g., time or frequency) to more than one user. Investigating the security issues of NOMA has been a topic of recent interest. For instance, the secrecy performance of large-scale NOMA networks is studied in [119] where their secrecy outage probability is derived for both single-antenna and multi-antenna transmitter scenarios. Proposing eavesdropping resilient transmission schemes with NOMA (especially, under a discrete signaling assumption) is an interesting research direction, which is left for future work as also stated in [120].

Full duplex wireless communications is another promising technology, which is expected to lead to a considerable increase in spectral efficiency. The main idea behind full duplex communications is to enable the radios to transmit and receive at the same time and within the same frequency band. The research on physical layer security in the scenarios where different nodes are capable of full duplex operation is a hot research topic (for an overview, see [11, Section VII]). However, only a limited number of the existing studies focus their attention on the practical constellation-constrained scenarios. One example of such studies is reported in [121] where secure communications in presence of full duplex relays is investigated under the assumption that the transmitter employs an Alamouti-based rank-two beamforming scheme.

Waveform design for 5G has also been a topic of interest during the past years. In addition to an increased bandwidth efficiency and reduced inter-user interference, new waveforms may also provide gains in terms of secrecy. For instance, the authors in [122] propose a secure waveform design for 5G wireless systems referred to as orthogonal transform division multiplexing (OTDM), which relies on replacing Fourier transform blocks in conventional OFDM systems by new blocks allowing for diagonalization of the main channel while degrading the eavesdropper’s reception.

We conclude this section by noting that the relatively recent paradigm of Internet of things (IoT) is expected to offer connectivity to billions of devices and it will be a centerpiece in the next generation wireless technologies including 5G. In view of the fact that, IoT systems are expected to consist of a large number of devices, the conventional cryptographic solutions, which require key distribution and management, may be too complex for implementation. Accordingly, physical layer security techniques (especially, the ones with lower complexities) can play an important role in safeguarding IoT. A recent survey paper [123] provides a thorough discussion of the potentials of physical layer security techniques and pinpoints a number of candidate schemes that can serve as an alternative or complement to the key-based solutions in securing IoT communications.

IX. PRACTICAL WIRETAP CODE CONSTRUCTIONS

The achievability of the secrecy rates described in the previous sections is proved with the aid of random coding arguments, i.e., from an information theoretic point of view. However, in practice, one needs to realize secure communications with the aid of practical codes. In this regard, explicit and implementable code design for secrecy has also been a topic of recent interest in the literature. In particular, error-correcting codes have found a new role since the emergence of physical layer security. Aside from providing reliable transmission, carefully designed codes are also capable of securing the transmitted messages against an eavesdropper. On the other hand, it is challenging to design these codes because they are required to meet two conflicting requirements: 1) they need to add redundancy in order to counter the randomness of the channel and provide reliable communication, and 2) in order to offer a secure transmission, they need to maximize the randomness perceived at the eavesdropper.

An important step towards designing codes for the wiretap channel is choosing a metric for measuring security. In contrast
to measuring reliability where probability of error is being used as a universally accepted metric, there are multiple metrics for quantifying secrecy for this scenario as listed in Table I. Computation of these metrics for explicit finite length coding schemes, however, has been limited to some specific channels\(^2\). One example is [124] in which the authors apply semantic secrecy to a finite length Reed-Muller code used over a binary erasure channel (BEC). For AWGN channels, a BER based metric called security gap (as mentioned in Section II-D2) is widely used by the research community, which is easily applicable to all practical coding schemes, however, its relation to the information theoretic notions is not established in a rigorous manner.

An analytical definition of security gap, which is nothing but the difference between the qualities (SNRs) of the main and the eavesdropper’s channels, can be given as follows. Let us fix two threshold values for the maximum desired BER over the main channel (denoted by \(P_{\text{Eve}}^{\text{max}} \leq \frac{1}{2}\)) and the minimum desired BER at the eavesdropper (denoted by \(P_{\text{Eve}}^{\text{min}} > 0\)). Then, the security gap equals to the difference between the SNR values corresponding to these two, i.e., \(\text{SNR}_B - \text{SNR}_E\) measured in dBs. For AWGN channels without feedback, this value must be greater than zero in order to result in a positive secrecy capacity (which equals to the difference between the capacities of the main and the eavesdropper’s channels [18]). Small security gaps are desirable because they provide secrecy even with a small degradation of the eavesdropper’s channel with respect to the main one.

The major shortcoming of the above BER-based analysis is that it assumes that if \(P_{\text{Eve}}^{\text{min}} \approx 0.5\), then Eve cannot extract any information about the secret message. However, if there exists a statistical correlation between the bit errors or between the Eve’s observation and the transmitted message, the amount of leaked information will not be zero. The authors in [129] employ tools from cryptography literature to eliminate such correlations. Specifically, they propose the use of substitution permutation networks (SPNs) without any secret key to satisfy the strict avalanche criterion (SAC) and the bit independence criterion (BIC). The required SPNs can be implemented efficiently [129], hence they can be applied to the message bits before the encoding stage. Therefore, one can design practical coding schemes, which result in BERs very close to 0.5 with no (or, small) correlation at Eve.

The general coding scheme for physical layer security is the randomized encoding proposed by Wyner in [15] to prove that there exists a code that achieves the secrecy capacity of the degraded wiretap channel. This method is also known as coset-coding studied further in the subsequent literature, e.g., in [125] and [126]. In this method, the all-zero message is mapped to a linear block code \(C\), with the generator matrix \(G\), and a non-zero message \(s = [s_1, s_2, \ldots, s_k]\) is mapped to the coset obtained by adding the \(n\)-tuple \(s_1 h_1 + s_2 h_2 + \ldots + s_k h_k\) to all the codewords in \(C\) where \(h_i\)’s are linearly independent \(n\)-tuples outside \(C\). Assuming \(h_i\)’s form a matrix \(H\) of size \(k \times n\), the transmitted codeword through the channel is

\[
c = [s \quad v] [H \quad G] \tag{18}\]

where \(v = [v_1, v_2, \ldots, v_r]\) denotes the random bit vector used for choosing a codeword in the corresponding coset uniformly randomly. Fig. 11 illustrates this method where the small code is denoted by \(C\) and \(c_{ij}\) denotes the \(i^{\text{th}}\) codeword in the \(j^{\text{th}}\) coset. We note that in conventional (non-stochastic) encoding, the leakage of information about the secret message to Eve, \(H(S^K[Z^n])\), is equal to the leakage of the codewords \(H(c^n[Z^n])\) [124] since there is a one-to-one mapping between the messages and the codewords. In the randomized encoding scheme, however, this is not true anymore since different codewords may represent the same message.

Application of lattice codes to the Gaussian wiretap channel is studied in [127] where the authors define an alternative security metric (called secrecy gain), which is related to the

<table>
<thead>
<tr>
<th>Paper</th>
<th>Network</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z. Mheich et al. [97]</td>
<td>Broadcast channel with confidential message</td>
<td>Characterizes secrecy rate region with PAM inputs and enhances it by optimizing symbol positions and the joint probability distribution.</td>
</tr>
<tr>
<td>T. Allen et al. [98]</td>
<td>MAC Wiretap</td>
<td>Develops a secure Alamouti transmission with artificial noise injection with perfect MCSI and statistical ECSI.</td>
</tr>
<tr>
<td>J. Jin et al. [99]</td>
<td>Cognitive MAC Wiretap</td>
<td>A two-layer precoding is proposed using statistical CSI to maximize secrecy sum-rate.</td>
</tr>
<tr>
<td>W. Zeng et al. [100]</td>
<td>Cognitive Multi-Antenna Wiretap</td>
<td>Proposes an iterative precoder optimization algorithm with statistical CSI.</td>
</tr>
<tr>
<td>K. Cao et al. [101]</td>
<td>MISO Wiretap with Multiple Multi-Antenna Eavesdroppers</td>
<td>Studies joint optimization of beamforming vector and transmitting power with full CSI.</td>
</tr>
<tr>
<td>S. Vishwakarma et al. [102]</td>
<td>MIMO Wiretap with Multiple Eavesdroppers</td>
<td>Employs a power allocation strategy with perfect MCSI and statistical ECSI.</td>
</tr>
<tr>
<td>S. Vishwakarma et al. [103]</td>
<td>DF Relaying in Presence of Multiple Eavesdroppers</td>
<td>Optimizes the source power and the relay weights with the aid of full CSI so that the secrecy rate is maximized.</td>
</tr>
<tr>
<td>K. Cao et al. [104]</td>
<td>Cooperative Jamming Network in Presence of Multiple Eavesdroppers</td>
<td>Provides a joint optimization of the artificial noise and the power allocated between the source and the relays.</td>
</tr>
<tr>
<td>K. Cao et al. [105]</td>
<td>MIMOME and a Multi-Antenna Helper</td>
<td>Proposes a joint optimization of precoder and power allocation between the source and the helper.</td>
</tr>
<tr>
<td>H. M. Furqan et al. [106]</td>
<td>Cooperative communications with untrusted relay</td>
<td>Proposes a destination-assisted jamming for improving secrecy and power efficiency.</td>
</tr>
</tbody>
</table>
theta series of lattices and shows the amount of confusion at the eavesdropper. Without introducing a decoding method, they evaluate the performance of different lattices based on the secrecy gain. The confusion at the eavesdropper in [127] is the result of using a random lattice in addition to the lattice that is responsible for transmitting the original message. As a further development, the authors in [128] prove the existence of lattices that are semantically secure over the Gaussian wiretap channel.

The authors in [137] and [138] provide a lower bound on the mutual information between the secret message and the Eve’s observation. They utilize this lower bound to quantify the equivocation for finite length codes for Gaussian wiretap channels. Specifically, they use finite length LDPC codes and demonstrate certain achievable points in the equivocation-rate region. In [136], the level of security at the physical layer is assessed from an information theoretic perspective while taking into account the constraints of practical transmissions over realistic wireless wiretap channels, i.e., by considering practical codes with finite length, discrete modulation formats and channels with fading.

Several practical coding schemes aiming directly at reducing the security gap are also proposed in the literature. Specifically, punctured LDPC codes are exploited for physical layer security in [30]. [130] demonstrates how non-systematic codes can be effective to reduce the resulting security gap, while [131] applies different techniques including scrambling, concatenation, and hybrid automatic repeat-request to LDPC and BCH codes in order to further reduce the security gap. Concatenation of polar and LDPC codes is proposed in [133]. Coset codes based on convolutional and turbo codes are constructed in [132] and [134] for physical layer security. Moreover, application of serially concatenated low density generator matrix (SCLDGM) codes to the randomized scheme is studied in [135].

Authors in [139] introduce a method which provides strong secrecy for the Gaussian wiretap channel by means of hash functions and capacity approaching channel codes of infinite length. Using this scheme, it is also possible to quantify the level of leakage obtained for a fixed block length. Finally, [140] provides the maximal secrecy rate over a wiretap channel subject to reliability and secrecy constraints at a given block length.

Table IX summarizes the existing practical coding schemes for Gaussian wiretap channel along with their main contributions. Furthermore, Fig. 12 demonstrates several representative security gaps from some of these practical coding schemes. One can see that when a BER of 0.5 is desired at Eve, security gaps as small as 1 dB can be obtained using carefully designed codes of length $10^4$ or longer.

In addition to the Gaussian wiretap channels, some efforts are also made to develop practical coding schemes for fading wiretap channels. The authors in [142] and [143] propose polar coding schemes for achieving secrecy over block fading wiretap channels. In [142], a secure coding scheme is designed with the aid of instantaneous MCSI and ECSI, while the polar coding approach proposed in [143] relies on statistical CSI of both channels. Application of algebraic lattice codes is also proposed over block fading wiretap channels. For instance, a sequence of non-random lattice codes is developed in [144] that achieve strong secrecy and semantic security over ergodic fading channels.

We emphasize that the easy-to-compute property of security gap makes it a widely used metric in majority of the aforementioned code designs, which try to avoid potential complexities associated with computing the original information-theoretic metrics for finite-length scenarios. Despite some limited efforts such as [138] in providing a lower bound for information-
A. Lessons Learned

The usual approaches for designing secure transmit signals with Gaussian inputs lead to considerable losses when applied to the signals drawn from finite constellations. This fact motivates the need for development of secure transmission schemes under the finite-alphabet input assumption.

In single-antenna transmission with constellation constrained inputs, achieving secrecy requires adoption of a proper power allocation strategy. This is because transmission with full power may allow the eavesdropper to successfully detect the transmitted symbols. Moreover, the increased dimensionality in multi-antenna transmissions can be employed to enhance secrecy with finite-alphabet inputs. Precoding serves as a promising strategy for maximizing the quality difference between the signals received at the legitimate receiver and the eavesdropper. Artificial noise injection offers considerable secrecy gains and it well-suits the nature of secure transmission with finite-alphabet inputs since the optimal power, which is allocated to the information bearing signal is only a fraction of the total power, and the excess power can be used for artificial noise injection. In many cases (especially, in the absence of instantaneous ECSI), without artificial noise injection, secrecy rates drop to zero when the transmit power increases. Therefore, injecting artificial noise is crucial for guaranteeing a good saturation behavior at high SNRs.

The optimal secure transmission schemes are not known in closed-form in (almost) any of the scenarios studied in this survey. This is due to the non-convexity of the objective function (i.e., secrecy rate). Furthermore, the transmit signal design algorithms, which rely on direct maximization of secrecy rates, are computationally complex mainly due to the fact that the mutual information expression under finite-alphabet inputs lacks closed-form. Therefore, different studies focus their attention on proposing lower-complexity alternatives to direct secrecy rate maximization. There is also an important trade-off between the secrecy rates and the computational complexity of the existing solutions in the literature of physical layer security with finite-alphabet inputs. Moreover, it is noteworthy that the transmit signal designs based on metrics such as SINR or BER, though simple and may be practical in some scenarios, do not ensure secrecy in an information theoretic sense.

The amount of CSI at the transmitter plays a central role in determining the secrecy performance of a particular transmission scheme. Transmission algorithms employing full CSI typically have a high capability in simultaneously increasing the information rates at the legitimate receiver and suppressing the reception at the eavesdropper, however, acquiring a perfect instantaneous CSI of the eavesdropper may not be practically possible in many scenarios. The achievable secrecy rates undergo considerable losses in the absence of instantaneous ECSI. However, nonzero secrecy rates are still attainable with the aid of statistical ECSI or in some cases even without it (i.e., with the aid of the MCSI only) [145].

Secure transmission design over multiuser channels, e.g., broadcast channel, is a more challenging task compared to the point-to-point scenarios. This is because the proposed solutions need to meet additional requirements, e.g., in terms of inter-user interference. Similarly, various challenges arise when moving to the scenarios with multiple eavesdroppers. For instance, secure transmission along the null-space of the eavesdropper’s channel is no longer possible in these scenarios and designing an artificial noise signal to effectively degrade the reception at the eavesdroppers is more complex than the single-eavesdropper case.

Practical coding schemes exploit different tools from coding theory and cryptography in order to ensure a certain level of security and reliability (in terms of BERs) for Eve and Bob, respectively. By combining coset coding and scrambling approaches along with code concatenation, security gaps as small as 1 dB can be obtained using finite-length explicit and implementable codes. However, finding specific (and implementable) finite-length codes, which can provide physical layer security from an information-theoretic point of view, remains as a future challenge.

B. Challenges and Open Problems

Various studies have been carried out on physical layer security with discrete signaling with the objective of bridging the gap between the information theoretic limits of secrecy and practical and implementable secure transmission schemes.
However, there still are many important open problems and future challenges to be addressed in this area.

Physical layer security with finite-alphabet inputs has been studied only for a limited set of CSI scenarios including the cases with perfect and statistical MCSI and ECSI at the transmitter. Some other important scenarios regarding the CSI uncertainty at different nodes, e.g., noisy estimation of CSI, outdated CSI or limited CSI feedback, which are well-investigated under the Gaussian input assumption [8], have not yet been considered for cases with finite-alphabet inputs. It would be interesting and of significant practical value to quantify the losses in the secrecy rates under these different imperfect CSI scenarios in the context of finite-alphabet inputs.

While the majority of artificial noise injection strategies rely on transmitting Gaussian distributed noise either by aligning it in the null-space of the main channel or by optimizing its covariance matrix, such artificial noise signals are not necessarily optimal. For instance, it has been shown in [146] that the worst case noise for AWGN channels with binary inputs has a discrete distribution. This motivates seeking interference distributions (rather than merely optimizing covariance matrices of Gaussian distributed signals), which maximize secrecy rates.

While some steps have been taken in studying physical layer security over finite-input multiuser networks (as surveyed in Section VII), many important scenarios have not yet been tackled. For example, no results have been reported for interference channels with secrecy constraints under the finite-alphabet input assumption (refer to [147] for an example with Gaussian inputs). Secrecy issues concerning the networks with simultaneous wireless information and power transfer (SWIPT) is also a topic of recent interest [148]. It is demonstrated that SWIPT networks are prone to information leakage at the energy receivers, and hence, the main focus of the existing articles is the characterization of the trade-off between the secrecy capacity/rate and the harvested power in these networks [149]-[152]. Studying this trade-off under the practical discrete signaling assumption is an important subject for future research. Furthermore, physical layer security in heterogeneous networks [153] and with full duplex transmitter or receivers (see, e.g. [154] and [155]) are very important, and extending the existing results for these scenarios to the case of finite-alphabet inputs is an interesting direction for future research.

Massive MIMO systems provide considerable gains in secrecy capacity. However, for such systems, besides the opportunities offered, many challenges arise in secure transmit signal designs with finite-alphabet inputs. For instance, most of the existing precoder design approaches (such as the ones proposed in [39] and [47]) are intractable when the number of antennas is high. Secure mmWave transmission is another promising research frontier. Communicating at the mmWave spectrum considerably changes the channel model with respect to the microwave communication systems, and introduces some new features (e.g., high sensitivity to blockages and considerable differences between the line-of-sight and non-line-of-sight receptions), which motivates development of new frameworks for designing practical secure transmission schemes.

Taking into account the hardware constraints in designing secure massive MIMO systems is another topic worth investigating. For example, secure transmission schemes can be proposed with finite-alphabet inputs for multi-antenna transmitters with a single radio frequency (RF) chain (or a few, but less than the number of antennas). Performance of wireless systems can be considerably affected by different RF impairments such as inphase/quadrature-phase (I/Q) imbalance, phase noise or distortions due to the nonlinear power amplifiers, which degrade the reception at the receivers (including the eavesdropper); hence, this behavior can be exploited to enhance secrecy. Moreover, various practical constraints arise in IoT systems (e.g., in terms of complexity and energy efficiency) that should be considered in designing secure transmission schemes. We also note that, designing testbeds and conducting tests by implementing physical layer security schemes is an essential step, which should be taken before these techniques can be adopted by actual communication systems. One example of such experiments is reported in [156] where the authors evaluate the secrecy performance of coset coding of lattice constellations with the aid of a software defined radio testbed.

Another interesting direction for future work is to further develop cross-layer design based solutions for securing communications as only a limited number of works have been reported on this topic so far. For instance, the authors in [157] demonstrate the advantages of cross-layer protocol interactions on the achievable secrecy by joint exploitation of automatic-repeat-request (as a MAC layer operation) and maximal ratio combining (as a physical layer operation). The potentials of other layers can also be explored to enhance security. For instance, employing authentication and watermarking strategies at the application layer along with the coding and signal processing at the physical layer [158] can lead to considerable secrecy gains.

The main problem associated with designing practical coding schemes for physical layer security is the following: Although the security gap is proposed and used as an alternative practical metric, the most widely accepted one (i.e., equivocation) relies on information theoretic quantities, and computing these quantities for finite-length codes over AWGN channels is a highly challenging problem, and can be an important research direction on physical layer security. As a final suggestion for future research, we recommend the problem of precoder design in the scenarios where practical wiretap codes are employed, which would be an important step towards designing implementable secure transmission schemes.

XI. Conclusions

An obligatory step towards practical implementation of physical layer security is understanding the impacts of employing standard constellations on the secrecy performance. With this motivation, this survey paper provides a summary of the current literature on physical layer security with discrete signaling and practical codes. We outline why and how transmit signal design problem is different for the case of
finite-alphabet inputs compared to Gaussian signaling, and review existing algorithms for maximizing secrecy rates in a variety of scenarios. We also provide an overview on practical code constructions for physical layer security, and point out important challenges and open problems on this topic.

REFERENCES

L. Zhou, D. Wu, B. Zheng, and M. Guizani, “Joint physical-application...

J. M. Hamamreh, M. Yusuf, T. Baykas and H. Arslan, “Cross...

J. Lu, J. Harshan, F. Oggier, “Performance of lattice coset codes on...

M. R. Abedi, N. Mokari, H. Saeedi, “How to manage resources to...


