Project PHYLAWS (Id 317562)
PHYsical LAYER Wireless Security

Deliverable 3.3–Coding techniques and algorithms for secrecy coding and secret key generation

Version 2.0 - 05 / 11 / 2015
## Change History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Description</th>
<th>Affected Sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>2015-9-05</td>
<td>First draft; ICL: Zheng Wang, Ling Liu and Cong Ling TPT: Jean-Claude Belfiore</td>
<td>1-8</td>
</tr>
</tbody>
</table>

## List of Contributors

<table>
<thead>
<tr>
<th>Partner</th>
<th>Contributors</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICL</td>
<td>Zheng Wang, Ling Liu and Cong Ling</td>
</tr>
<tr>
<td>TPT</td>
<td>Jean-Claude Belfiore</td>
</tr>
<tr>
<td>Thales</td>
<td>François Delaveau, Christiane-Laurie Kameni Ngassa.</td>
</tr>
</tbody>
</table>
# Table of contents

1 **Introduction** .......................................................................................................................... 5

2 **Definitions and Scoping** ....................................................................................................... 6
   2.1 Terms and concepts ................................................................................................................ 6
   2.2 Scope.................................................................................................................................... 7

3 **Wiretap Coding for Discrete Channels** .................................................................................. 8
   3.1 Basics of Coset Coding ......................................................................................................... 8
   3.2 Low Density Parity-Check (LDPC) Codes ......................................................................... 8
   3.3 Polar Codes (PC) ............................................................................................................... 10
   3.4 Modified Polar Wiretap Coding ......................................................................................... 12

4 **Wiretap coding for Gaussian Wiretap Channels (GWC)** ......................................................... 13
   4.1 Lattice Wiretap coding ....................................................................................................... 13
   4.2 Polar Lattice Wiretap Coding for mod-As GWC ............................................................... 14
      4.2.1 Mod-As GWC ............................................................................................................. 15
      4.2.2 Reliability Analysis.................................................................................................... 18
   4.3 Secrecy-good polar lattices with discrete Gaussian shaping ............................................. 19
      4.3.1 Gaussian shaping over polar lattices ........................................................................ 19
      4.3.2 Three-dimensional partition ..................................................................................... 21
      4.3.3 Strong secrecy ........................................................................................................... 23
      4.3.4 Achieving the Secrecy Capacity .............................................................................. 24

5 **Wiretap Coding for MIMO and Fading Channels** ................................................................. 26
   5.1 The MIMO Case .............................................................................................................. 26
   5.2 Block and Fast Fading Channels ...................................................................................... 29
   5.3 A MIMO example: The Alamouti Code ......................................................................... 30

6 **Secret key generation from Gaussian sources using lattice hashing** .................................... 34
   6.1 Lattice Hashing for Gaussian Sources ............................................................................ 34
   6.2 Secret Key Agreement ...................................................................................................... 35
   6.3 Uniformity ........................................................................................................................ 36
   6.4 Strong secrecy and Reliability ....................................................................................... 37
   6.5 Existence of a Sequence of Nested Lattices for Secret Key Generation ......................... 38

7 **Conclusion - future work and relations with other tasks** ...................................................... 39
   7.1 Summary ........................................................................................................................ 39
   7.2 Future works and relations to other tasks ...................................................................... 39

8 **References** ............................................................................................................................ 40
   8.1 Phylaws deliverables ....................................................................................................... 40
   8.2 Open literature ............................................................................................................... 40
List of figures

Figure 1: The parity check matrix and corresponding factor graph for LDPC code [6] (ci denotes check nodes, vi denotes variable nodes). ................................................................. 8
Figure 2: The binary erasure wiretap channel model [2] [© 2011 IEEE] ................................................................................................................................. 9
Figure 3: The construction of polar codes with block length N=8 ................................................................................................................................. 10
Figure 4: Channel capacity of bit channels ...................................................................................................................................................... 11
Figure 5: The Gaussian wiretap channel ...................................................................................................................................................... 13
Figure 6: The mod-As Gaussian wiretap channel ...................................................................................................................................................... 15
Figure 7: The multilevel lattice coding system over the mod-As Gaussian wiretap channel. ................................................................. 17
Figure 8: The relationship between W and W ...................................................................................................................................................... 20
Figure 9: Partitions of the index set [N] with shaping ...................................................................................................................................................... 22
Figure 10: Block diagram of the shaping-induced channel QN(W,S) ................................................................................................................................. 23
Figure 11: Block diagram of the new induced channel QN(W,S,R) ................................................................................................................................. 24
Figure 12: An upper bound onPC, e/CMIMO: the Alamouti code with ne = 2. .............................................................................................................. 31
Figure 13: A tighter upper bound onPC, e/CMIMO: the Alamouti code with ne = 2 .............................................................................................................. 32
Figure 14: Loose upper bounds versus tight upper bounds: the Alamouti code with ne = 2. .............................................................................................................. 33
Figure 15: Upper bounds versus simulations: the Alamouti code with ne = 2. .............................................................................................................. 33
Figure 16: Secret key generation in the presence of an eavesdropper with communication over a public channel [46]. ...................................................................................................................................................... 35
1 Introduction

Given the growing prevalence of wireless radio-communication technologies, security, privacy and reliability of the exchanged information becomes a major societal challenge for both personal and professional sphere. Moreover, the growing importance of sensing and cognitive procedures in future radio access technologies (white spectrum, cognitive networks) will occur numerous downloading and uploading procedures for geo-referenced sensing spectrum allocations, whose integrity and privacy are major industrial challenges for both operators and administrations. Secure air interface within wireless networks are thus crucial for various applications such as broadband internet, e-commerce, radio-terminal payments, bank services, machine to machine, health/hospital distant services. Most of citizens, professionals, stakeholders, services providers and economical actors are thus concerned by confidentiality lacks and by privacy improvements of the physical layer of wireless networks.

The main objective of this deliverable is to study the explicit coding techniques and algorithms for security. Instead of the work in previous deliverable [D2.3] in which we studied how much secrecy we can extract from the noise itself in the form of a secret key, we now analyze how to design an explicit coding scheme which is able to provide reliable message communication between Alice and Bob and keep it confidential to the eavesdropper Eve. Specifically, we use nested lattice structure to complete this task. Intuitively, we design two lattice codes with rates almost equal to the capacities of the main channel (between Alice and Bob) and the wiretapper's channel (between Alice and Eve), respectively. The lattice code for the wiretapper's channel is utilized to convey random bits, and the code for the main channel only makes use of the coset leader (the two lattices are nested) to send message. Furthermore, we extend our discussion to the MIMO channel and fading channel scenarios, and present more details for secrecy key generation from Gaussian source.

First, we introduce the wiretap coding scheme for discrete channels, based on the famous LDPC codes and the recently proposed polar codes. Second, we extend the coset wiretap coding scheme to continuous Gaussian channels, by constructing polar lattices. As a result, we prove that our scheme is able to achieve strong secrecy and the secrecy capacity. In the meantime, an explicit lattice shaping scheme based on discrete lattice Gaussian distribution is also presented. We will show that it achieves the optimal shaping gain. This shaping scheme is compatible to the wiretap coding structure by the versatility of polar codes, meaning that a convenient implementation is possible. Then we study in detail the wiretap coding design for the MIMO channels and fading channels, based on a similar structure of nesting lattice codes. A detailed example of the Alamouti code will be presented.

The following of the document concentrates on secret key generation from Gaussian source, using the lattice hashing technique. After a preliminary review of the state of the art, an analysis and a preliminary channel model is investigated about the meaning of secrecy in this context and the effectiveness of obtaining keys with enough secure bits, depending on the nature of the radio environment and the distance between Bob and Eve. Subsequently, a methodology is discussed about the way to conduct channel measurements or channel simulations in PHYLAWS in order to achieve the project objectives.
2 Definitions and Scoping

2.1 Terms and concepts

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beamforming</td>
<td>A signal processing technique where transmitter/receiver uses multiple antennas to send/receive the same signal. Can be used for directional signal transmission (organizing phase and amplitude of signals so that others experience constructive and others destructive interference).</td>
</tr>
<tr>
<td>CIR</td>
<td>Channel Impulse Response</td>
</tr>
<tr>
<td>DOA</td>
<td>Direction Of Arrival</td>
</tr>
<tr>
<td>GSCM</td>
<td>Geometry based Stochastic Channel Model</td>
</tr>
<tr>
<td>GSM</td>
<td>Group Special Mobile</td>
</tr>
<tr>
<td>i.i.d.</td>
<td>Independent and identically distributed</td>
</tr>
<tr>
<td>LAN</td>
<td>Local Area Network</td>
</tr>
<tr>
<td>LOS</td>
<td>Line Of Sight</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multi-input multi-output (use of multiple antennas at both the transmitter and receiver)</td>
</tr>
<tr>
<td>MISO</td>
<td>Multi-input single-output (use of multiple antennas at the transmitter)</td>
</tr>
<tr>
<td>NLOS</td>
<td>Non Line Of Sight</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal frequency division multiplex – Modulation technology where a signal is split into several narrowband channels at different frequencies.</td>
</tr>
<tr>
<td>PHYSEC</td>
<td>Physical Layer Security is generic term that will be used in this project to design all kind of protection techniques that are based on the use of the physical layer sensing and/or measurement.</td>
</tr>
<tr>
<td>RSSI</td>
<td>Received Signal Strength Indication</td>
</tr>
<tr>
<td>RT</td>
<td>Ray-Tracing</td>
</tr>
<tr>
<td>Rx</td>
<td>Receiver</td>
</tr>
<tr>
<td>SKG</td>
<td>Secret Key Generation</td>
</tr>
<tr>
<td>SIMO</td>
<td>Single-input multi-output (use of multiple antennas at the receiver)</td>
</tr>
<tr>
<td>SISO</td>
<td>Single-input single-output (use of single antennas at both the transmitter and receiver)</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise ratio</td>
</tr>
<tr>
<td>Tx</td>
<td>Transmitter</td>
</tr>
<tr>
<td>WSSUS</td>
<td>Wide-Sense Stationary and Uncorrelated Scattering</td>
</tr>
<tr>
<td>LDPC</td>
<td>Low Density Parity Check</td>
</tr>
<tr>
<td>BEC</td>
<td>Binary Erasure Channel</td>
</tr>
<tr>
<td>MP</td>
<td>Message Passing</td>
</tr>
<tr>
<td>BP</td>
<td>Belief Propagation</td>
</tr>
<tr>
<td>SC-LDPC</td>
<td>Spatially Coupled Low Density Parity Check</td>
</tr>
<tr>
<td>BMSC</td>
<td>Binary Memoryless Symmetric Channel</td>
</tr>
<tr>
<td>BER</td>
<td>Bit-Error Rate</td>
</tr>
<tr>
<td>SC</td>
<td>Successive Cancellation</td>
</tr>
<tr>
<td>GWC</td>
<td>Gaussian Wiretap Channel</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise</td>
</tr>
<tr>
<td>VNR</td>
<td>Volume-to-Noise Ratio</td>
</tr>
</tbody>
</table>
2.2 Scope

The deliverable introduces several notions relevant to information theory and the main principle that are relevant to wiretap coding at the physical layer. From a state of the current researches, several wiretap coding solutions will be highlighted, which take benefit of native PHYSEC concepts, i.e., the randomness of noise and radio channels when facing passive eavesdroppers. The deliverable also deals with practical implantation perspectives of secret key generation in existing and future radio-networks, as stand-alone added modules operating at the physical layer, or as added algorithm combined with classical solutions such as TRANSEC, NETSEC and COMSEC protections.

Secondly, the deliverable addresses radio channel aspects regarding SKG. The objective is to highlight the major channel issues regarding SKG and routes towards the development of channel models suited to the PHYLAWS needs for SKG.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF</td>
<td>Probability Distribution Function</td>
</tr>
<tr>
<td>LR</td>
<td>Likelihood Ratio</td>
</tr>
<tr>
<td>BMA</td>
<td>Binary Memoryless Asymmetric</td>
</tr>
<tr>
<td>MMSE</td>
<td>Minimum Mean Square Error</td>
</tr>
</tbody>
</table>
3 Wiretap Coding for Discrete Channels

In this section, we discuss the construction of some practical binary codes for secrecy such as the low density parity check (LDPC) code and the polar code. The design of codes for the wiretap channel turns out to be challenging, and this area of information-theoretic security is still largely in its infancy. To some extent, the major obstacles in the road to secrecy capacity are similar to those that lay in the path to channel capacity; the random coding arguments used to establish the secrecy capacity [D2.3] do not provide explicit code constructions. However, the design of wiretap codes is further impaired by the absence of a simple metric, such as a bit error rate, which could be evaluated numerically [2].

3.1 Basics of Coset Coding

The vast majority of work on PHYSEC is based on non-constructive random-coding arguments to establish the theoretic results. Such results demonstrate the existence of codes that achieve the secrecy capacity, but are of little practical usefulness. In recent years, progress has been made on the construction of practical codes for PHYSEC, to a more or less extent. The design methodology can be traced back to Wyner’s original work on coset coding [1] which suggests that several codewords should represent the same message and that the choice of which codeword to transmit should be random to confuse the eavesdropper. In what follows, we will show how the original concept of the coset coding could be implemented with binary codes to achieve strong secrecy.

![Figure 1: The parity check matrix and corresponding factor graph for LDPC code [6] (c_i denotes check nodes v_i denotes variable nots).](image)

3.2 Low Density Parity-Check (LDPC) Codes

LDPC codes are famous for their capacity-approaching performance on many communication channels. The parity check matrix and the corresponding factor graph are shown in Fig. 1. However, LDPC codes have been used to build wiretap codes only with limited success.

When the main channel is noiseless and the wiretap channel is the binary erasure channel (BEC) as in Fig. 2, LDPC codes for the BEC were presented in [3;4]. Especially in [4], the authors generalized the link between capacity-approaching codes and weak secrecy capacity. The use of capacity-achieving codes for the eavesdropper’s channel is a sufficient condition for weak secrecy. This viewpoint provided a clear construction method for coding schemes for secure communication across arbitrary wiretap channels. Then, they used this idea to construct the first secrecy-capacity-achieving LDPC codes for a wiretap channel with a noiseless main channel and a BEC under message passing decoding (MP), however in terms of weak secrecy. Later, [2] proved that the same construction can be used to guarantee strong secrecy at lower rates. A similar construction based on two-edge-type LDPC codes was proposed in [5] for the BEC wiretap channel.
The coset coding scheme using LDPC codes for the BEC wiretap channel model shown in Fig. 2 can be interpreted as follows. Prior to transmission, Alice and Bob publicly agree on a \( (n, n(1 - R)) \) binary LDPC code \( C \), where \( n \) is the blocklength of \( C \) and \( R \) is the secrecy rate. For each possible value \( m \) of the \( nR \) bits secret message \( M \), a coset of \( C \) given by \( C(m) = \{x^n \in \{0,1\}^n; x^m H^T = m\} \) is associated, where \( H \) is the parity check matrix of \( C \). To convey the message \( M \) to Bob, Alice picks a codeword in \( C(m) \) randomly and transmits it. Bob can obtain the secret message by calculating \( x^m H^T \). Suppose that code \( C \) has a generator matrix \( G = [a_1, ... a_n] \), where \( a_i \) is the \( i \)th column of \( G \). Consider that Eve observes \( u \) unerased symbols from \( X^n \), with the unerased positions given by \( \{i: z_i \neq ?\} = \{i_1, i_2, ..., i_u\} \).

Message \( m \) is secured by \( C \) in the sense that the probability \( Pr[M = m | Z = 1] = 1/2^{|R|} \) if and only if the matrix \( G_u = [a_{i_1}, a_{i_2}, ..., a_{i_u}] \) has rank \( u \). This is due to the fact that if \( G_u \) has rank \( u \), the code \( C \) has all \( 2^u \) possible \( u \)-tuples in the \( u \) unerased positions. Therefore, by linearity, each coset of \( C \) would have all \( 2^n \) possible \( u \)-tuples in the same positions as well, which means that Eve has no idea which coset Alice has chosen. To guarantee such strong secrecy, the design rate \( \alpha \) of \( C \) is used in the coset coding scheme, the conditional entropy \( H(M|Z^n) \) can be converted to the problem of constructing a dual LDPC code over a BEC with erasure probability \( 1 - \varepsilon \).

Let \( P_e^{(n)}(\varepsilon) \) denote the probability of block error for a LDPC code \( C \) with blocklength \( n \) over \( \text{BEC}(\varepsilon) \). For a parity check matrix \( H \) of \( C \), \( 1 - P_e^{(n)}(\varepsilon) \) is a lower bound on the probability that the erased columns of \( H \) form a full rank submatrix, which means that generator matrix \( G_u(C^+) \) resulted by a \( \text{BEC}(1 - \varepsilon) \) has full rank with probability larger than \( 1 - P_e^{(n)}(\varepsilon) \). If \( C^+ \) is used in the coset coding scheme, the conditional entropy \( H(M|Z^n) \) can be bounded as:

\[
H(M|Z^n) \geq H(M|Z^n, \text{rank}(G_u(C^+))) \\
\geq H(M|Z^n, G_u(C^+) \text{ is full rank}) \text{Prob}[G_u(C^+) \text{ is full rank}] \\
\geq H(M)(1 - P_e^{(n)}(\varepsilon)).
\]

Then we have

\[
I(M; Z^n) = H(M) - H(M|Z^n) \leq H(M)P_e^{(n)}(\varepsilon) \leq nRP_e^{(n)}(\varepsilon).
\]

Therefore, if code \( C \) has a BP threshold \( \varepsilon^{BP} \) such that \( P_e^{(n)}(\varepsilon) = O(\varepsilon^{BP \alpha}) \), \( (\alpha > 1) \) for \( \varepsilon < \varepsilon^{BP} \), then the dual code of \( C \) used in a coset coding scheme provides strong secrecy over a binary erasure wiretap channel with erasure probability \( \varepsilon' > 1 - \varepsilon^{BP} \). Note that weak secrecy is provided if \( \alpha > 0 \). The design rate \( R \) of \( C \) should satisfy \( R < 1 - \varepsilon^{BP} \), while the secrecy capacity \( \varepsilon' \) of the wiretap channel is larger than \( 1 - \varepsilon^{BP} \). To achieve the secrecy capacity, \( 1 - \varepsilon^{BP} \) is required to be very close to the code rate \( R \). It is known that LDPC codes achieve capacity over a BEC under the BP...
decoding, however, the parameter $\alpha$ can only be proved to be positive for the rate arbitrarily close to $1 - \varepsilon^{BP}$. To satisfy the strong secrecy requirement $\alpha > 1$, the design rate should be slightly away from $1 - \varepsilon^{BP}$.

Unfortunately, for other binary memoryless symmetric channels (BMSCs) except BECs, general LDPC codes do not have the capacity achieving property, which means that the coset coding scheme using general LDPC codes cannot achieve the secrecy capacity when the wiretapper’s channel is not a BEC. In this case, spatially coupled low density parity check (SC-LDPC) codes [26], which have been proved to be able to achieve the capacity of general BMSCs, provide us a promising approach. In [27], a coset coding scheme based on regular two edge type SC-LDPC codes is proposed over a BEC wiretap channel, where the main channel is also a BEC. It is shown that the whole rate-equivocation region of such BEC wiretap channel can be achieved by using this scheme under weak secrecy condition. Since SC-LDPC codes are universally capacity achieving, it is also conjectured that this construction is optimal for the class of wiretap channel where the main channel and wiretapper’s channel are BMSCs and the wiretapper’s channel is physically degraded with respect to the main channel.

In addition, LDPC codes has also been proposed for Gaussian wiretap channel in [6] but with a different criterion. The proposed coding scheme is asymptotically effective in the sense that it yields a bit-error rate (BER) very close to 0.5 for an eavesdropper whose SNR is lower than a threshold, even if the eavesdropper has the ability to use a MAP decoder. However, this approach does not ensure information theoretic strong secrecy.

### 3.3 Polar Codes (PC)

Polar codes [13] are the first codes with an explicit construction to provably achieve the channel capacity for symmetric binary discrete memoryless channels (B-DMCs) with low encoding and decoding complexity. The construction of polar codes with block length $N = 8$ is shown in Fig. 3. As the block length increases, the capacities of bit channels polarize either to 0 or 1. It is shown that as the block length goes to infinity. The number of the bit channels with 1 capacity is equal to the channel capacity. Therefore polar codes can achieve the channel capacity by just transmitting information bits through these perfect channels.

![Figure 3: The construction of polar codes with block length N=8](image)

In the meantime, polar coding also seems to offer a more powerful approach to design wiretap codes. Recently there has been a lot of interest in the design of wiretap codes based on polar codes. For example, [7] [8] and [9] use polar codes to build encryption schemes for the wiretap setting with binary-input symmetric channels. However, these schemes only provide weak security. In [10], it was shown that, with a minor modification of the original design, polar codes achieve strong secrecy (and also semantic security). However, they could not guarantee reliability of the main channel when it is noisy. In the following paper [11], a multi-block polar coding scheme was proposed to solve this reliability problem under the condition that the block is sufficiently large. In the meantime, a similar multi-block coding scheme was discussed in [12]. Their polar coding scheme also achieves the secrecy capacity under strong secrecy.
condition and guarantees reliably for the legitimate receiver. However, they only proved the existence of this coding scheme, and thus it might be computationally hard to find the explicit structure.

Figure 4: Channel capacity of bit channels

In the rest of this subsection, we show how polar codes can achieve the secrecy capacity of the binary wiretap channel [10, 11]. Define the sets of very reliable and very unreliable indices for a binary channel $Q$ and for $0 < \beta < 1/2$:

- $\mathcal{G}(Q) = \{i: Z(Q_N(i)) \leq 2^{-N\beta}\}$ (Reliability-good indices)
- $\mathcal{N}(Q) = \{i: I(Q_N(i)) \leq 2^{-N\beta}\}$ (Information-bad indices)

where $Z(Q_N(i))$ and $I(Q_N(i))$ represent the Bhattacharyya parameter and mutual information of the polarized bit-channel $Q_N(i)[13]$. $V$ and $W$ represent the main channel and the wiretap channel, respectively. The indices in $\mathcal{G}(V)$ and $\mathcal{N}(W)$ are the reliable and the secure indices. The index set can be partitioned into the following four sets:

$\mathcal{A} = \mathcal{G}(V) \cap \mathcal{N}(W)$
$\mathcal{B} = \mathcal{G}(V) \cap \mathcal{N}(W)^c$
$\mathcal{C} = \mathcal{G}(V)^c \cap \mathcal{N}(W)^c$
$\mathcal{D} = \mathcal{G}(V)^c \cap \mathcal{N}(W)$

Unlike the standard polar coding, the bit-channels are now partitioned into three parts: A set $\mathcal{M}$ that carries the confidential message bits, a set $\mathcal{R}$ that carries random bits, and a set $\mathcal{F}$ of frozen bits which are known to both Bob and Eve prior to transmission. We assign the bits as follows:

$\mathcal{A} = \mathcal{M}$
$\mathcal{B} \subseteq \mathcal{R}$
$\mathcal{C} \subseteq \mathcal{F}$
$\mathcal{D} \subseteq \mathcal{R}$.

According to [10], this assignment introduces a new channel which is also symmetric. The mutual information of this channel can be upper-bounded by the sum of the mutual information of bit-channels in $\mathcal{N}(W)$. The threshold of the mutual information on each bit-channel within $\mathcal{N}(W)$ is $2^{-\beta N}$. Then the mutual information between the message and the signal Eve received $I(M; Z_N)$ can be upper-bounded by $N2^{-N\beta}$. Therefore the strong secrecy is achieved if $\lim_{N \to \infty} I(M; Z_N) = 0$. Furthermore the achievable secrecy rate is equal to the secrecy capacity. This is due to the following facts of polarization theory [11]:

$$\lim_{N \to \infty} \frac{|\mathcal{G}(Q)|}{N} = C(Q),$$
$$\lim_{N \to \infty} \frac{|\mathcal{N}(Q)|}{N} = 1 - C(Q).$$
And since $W$ is degraded with respect to $V$,
\[ \lim_{N \to \infty} \frac{|\mathcal{G}(V) \cap \mathcal{N}(W)|}{N} = C(V) - C(W), \]
\[ \lim_{N \to \infty} \frac{|\mathcal{G}(V)^c \cap \mathcal{N}(W)^c|}{N} = 0. \]

This polar coding scheme is also capable of satisfying the reliability condition. According to the construction of polar codes [13], the block error probability at Bob’s end is upper-bounded by the sum of the Bhattacharyya parameters $Z$ of those subchannels that are not frozen. Let $V^{(i)}_N$ denote the $i$th subchannel of the main channel $V$, the decoding error probability $P_e^{SC}$ under the Successive Cancellation (SC) decoding is upper-bounded as
\[ P_e^{SC} \leq \sum_{i \in \mathcal{U} \cup \mathcal{D}} Z(V^{(i)}_N) = \sum_{i \in \mathcal{U}} Z(V^{(i)}_N) + \sum_{i \in \mathcal{E} \cup \mathcal{D}} Z(V^{(i)}_N). \]

The last equation is due to the fact that set $\mathcal{G}$ and $\mathcal{D}$ are disjoint. By the definition of $\mathcal{G}$, the term $\sum_{i \in \mathcal{U}} Z(V^{(i)}_N)$ is bounded by $2^{-N^p}$, which vanishes when $N$ is sufficiently large. However, the bound on the term $\sum_{i \in \mathcal{D}} Z(V^{(i)}_N)$ is difficult to derive. To overcome this problem, a modified scheme dividing the message $\mathcal{M}$ into several blocks is proposed [11]. For a specific block, $\mathcal{D}$ is still assigned with random bits but transmitted in advance in the set $\mathcal{A}$ of the previous block. By embedding $\mathcal{D}$ in $\mathcal{A}$, we induce some rate loss, but obtain an arbitrarily small term $\sum_{i \in \mathcal{D}} Z(V^{(i)}_N)$. Since the size of $\mathcal{D}$ is very small, the rate loss is negligible and finally the new scheme realizes reliability and strong security simultaneously.

### 3.4 Modified Polar Wiretap Coding

In section 3.3, the information-bad set $\mathcal{N}(Q)$ was defined according to the mutual information of the subchannels, while the reliability-good set $\mathcal{G}(Q)$ was defined according to the Bhattacharyya parameter of the subchannels. To make our wiretap coding construction more convenient, we modify the definition of $\mathcal{N}(Q)$ to be
\[ \mathcal{N}(Q) = \left\{ i : Z(Q^{(i)}_N) \geq 1 - 2^{-N^p} \right\}, \]
which is also based on the Bhattacharyya parameter of the subchannels. It can be proved that if $Z(Q^{(i)}_N) \geq 1 - 2^{-N^p}$, the mutual information of the $i$-th subchannel can be upper-bounded as $I(Q^{(i)}_N) \leq 2^{-N^p}$, for sufficiently large $N$ and $0 < \beta' < \beta < 1/2$.

Since the mutual information of subchannels in $\mathcal{N}(Q)$ can be upper-bounded in the same form, it is not difficult to understand that the strong secrecy can be achieved using the technique proposed in section 3.3. Similarly, we divide the index set into the four sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$. Then the mutual information between the message and the signal Eve received $I(M; Z^N)$ can be upper-bounded by $N2^{-N^p}$, which implies $\lim_{N \to \infty} I(M; Z^N) = 0$ as well. With regard to the secrecy rate, we can also show that the modified polar coding scheme can achieve the secrecy capacity since the modification does not change the proportion of the set $\mathcal{N}(Q)$. 
4 Wiretap coding for Gaussian Wiretap Channels (GWC)

In this section, we address how to construct polar lattices to achieve the secrecy capacity of Gaussian wiretap channels. We first start with a discussion on how to obtain the AWGN good lattice $\Lambda_b$ and secrecy good lattice $\Lambda_e$ for the mod-$\Lambda_x$ Gaussian wiretap channels. The result also holds for the polar lattices when the input distribution of each level is uniform for the genuine Gaussian wiretap channel (GWC). The setting without power constraint is similar to the Poltyrev setting in the Gaussian point-to-point channel. Then we introduce an explicit lattice shaping scheme for $\Lambda_b$ and $\Lambda_e$ simultaneously and remove the mod-$\Lambda_x$ front-end. As a result, we develop an explicit wiretap coding scheme based on polar lattices which can be proved to achieve the secrecy capacity of the GWC with no requirement on SNR.

4.1 Lattice Wiretap coding

Polar codes [13] have shown their great potential in solving the wiretap coding problem. The polar coding scheme proposed in [10], combined with the block Markov coding technique [11], was proved to achieve the strong secrecy capacity when $W$ and $V$ are both binary-input symmetric channels, and $W$ is degraded with respect to $V$. For continuous channels such as the GWC, there also has been notable progress in wiretap lattice coding. On the theoretical aspect, the existence of lattice codes achieving the secrecy capacity to within 1/2 nat under the strong secrecy as well as semantic security criterion was demonstrated in [14]. On the practical aspect, wiretap lattice codes were proposed in [15] and [16] to maximize the eavesdropper’s decoding error probability.

![Figure 5: The Gaussian wiretap channel](image)

In this work, we construct lattice codes for the GWC which is shown in Fig. 5. The confidential message $M$ drawn from the message set $M$ is encoded by the sender (Alice) into an $N$-dimensional codeword $X^{[N]}$. The outputs $Y^{[N]}$ and $Z^{[N]}$ received by the legitimate receiver (Bob) and the eavesdropper Eve are respectively given by

\[
Y^{[N]} = X^{[N]} + W_b^{[N]},
\]

\[
Z^{[N]} = X^{[N]} + W_e^{[N]},
\]

Where are $N$-dimensional Gaussian noise vectors with zero mean and variance $\sigma_b^2$, $\sigma_e^2$ respectively. The channel input $X^{[N]}$ satisfies the power constraint $P_x$, i.e.,

\[
\frac{1}{N} E\left[\|X^{[N]}\|^2\right] \leq P_x.
\]

A lattice is a discrete subgroup of $\mathbb{R}^n$ which can be described by

\[
\Lambda = \{A = Bx: x \in \mathbb{Z}^n\},
\]

where $B$ is the $n$-by-$n$ lattice generator matrix and we always assume that it has full rank.

For a vector $x \in \mathbb{R}^n$, the nearest-neighbor quantizer associated with $\Lambda$ is $Q_\Lambda(x) = \text{argmin}_{y \in \Lambda} \|y - x\|$. We define the modulo lattice operation by $x \mod \Lambda \triangleq x - Q_\Lambda(x)$. The Voronoi region of $\Lambda$, defined by $V(\Lambda) = \{x: Q_\Lambda(x) = 0\}$ specifies the nearest-neighbor decoding region. The Voronoi cell is one example of fundamental region of the lattice. A measurable set $R(\Lambda) \subset \mathbb{R}^n$ is a fundamental region of the lattice $\Lambda$ if $U_{\lambda \in \Lambda} (R(\Lambda) + \lambda) = \mathbb{R}^n$ and if $(R(\Lambda) + \lambda) \cap (R(\Lambda) + \lambda')$ has measure 0 for any $\lambda \neq \lambda'$ in $\Lambda$. The volume of a fundamental region is equal to that of the Voronoi region $V(\Lambda)$, which is given by $\text{Vol}(\Lambda) = |\det(B)|$. 
To satisfy the reliability condition for Bob, we are mostly concerned with the block error probability $P_e(\Lambda, \sigma^2)$ of lattice decoding. It is the probability that an $n$-dimensional independent and identically distributed (i.i.d.) Gaussian noise vector $x$ with zero mean and variance $\sigma^2$ per dimension falls outside the Voronoi region $V(\Lambda)$. For an $n$-dimensional lattice $\Lambda$, define the volume-to-noise ratio (VNR) by

$$Y_\Lambda(\sigma) \triangleq \frac{\text{Vol}(\Lambda)^2}{\sigma^2},$$

Then we introduce the notion of lattices which are good for the AWGN channel without power constraint. A sequence of lattices $\Lambda_0$ of increasing dimension $n$ is AWGN-good if, for any fixed $P_e(\Lambda, \sigma^2) \in (0,1)$,

$$\lim_{n \to \infty} Y_\Lambda(\sigma) = 2\pi e,$$

and if for a fixed VNR greater than $2\pi e$, $P_e(\Lambda, \sigma^2)$ goes to 0 as $n \to \infty$. It is worth mentioning here that we do not insist on exponentially vanishing error probabilities, unlike Poltyrev's original treatment of good lattices for coding over the AWGN channel. This is because a sub-exponential or polynomial decay of the error probability is often good enough.

To satisfy the secrecy condition for Eve, we are mostly concerned with the information leakage $I(M; Z^n)$ at Eve’s end. Therefore, we define the secrecy-good lattice as a lattice $\Lambda_0$ which results in fast-vanishing information leakage. Note that this definition is more general than the definition proposed in [13] which is based on the flatness factor.

The idea of Gaussian wiretap lattice coding can be explained as follows. Let $\Lambda_0$ and $\Lambda_0$ be the AWGN-good lattice and secrecy-good lattice designed for Bob and Eve accordingly. Let $\Lambda_0 \subset \Lambda_0 \subset \Lambda_0$ be a nested chain of $N$-dimensional lattices in $\mathbb{R}^N$, where $\Lambda_0$ is the shaping lattice. Note that the shaping lattice $\Lambda_0$ here is employed primarily for the convenience of designing the secrecy-good lattice and secondarily for satisfying the power constraint. Consider a one-to-one mapping: $M \to \Lambda_0/\Lambda_0$ which associates each message $m$ to a coset $\lambda_0 \in \Lambda_0/\Lambda_0$. Alice selects a lattice point $\lambda \in \Lambda_0 \cap V(\Lambda_0)$ uniformly at random and transmits $X[N] = \lambda + \lambda_0$ where $\lambda_0$ is the coset representative of $\lambda_0$ in $V(\Lambda_0)$. This scheme has been proved to achieve both reliability and semantic security in [13] by random lattice codes. We will make it explicit by constructing polar lattice codes in next section, where we assume that a shaping lattice already exists for the power constraint. As we have mentioned, the shaping lattice will be replaced by a discrete lattice Gaussian shaping in section 4.3. Before that, we firstly introduce the flatness factor and lattice Gaussian distribution.

For $c > 0$ and $c \in \mathbb{R}^n$, the Gaussian distribution of mean $c$ and variance $\sigma^2$ is defined as

$$f_{\sigma, c}(x) = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{2\sigma^2} ||x-c||^2},$$

for all $x \in \mathbb{R}^n$. For convenience, let $f_c(x) = f_{\sigma,0}(x)$.

Given lattice $\Lambda$, we define the $\Lambda$-periodic function

$$f_{\sigma, \Lambda}(x) = \sum_{\lambda \in \Lambda} f_{\sigma, \lambda}(x) = \frac{1}{(\sqrt{2\pi\sigma})^n} \sum_{\lambda \in \Lambda} e^{-\frac{1}{2\sigma^2} ||x-\lambda||^2},$$

For all $x \in \mathbb{R}^n$.

The flatness factor is defined for a lattice $\Lambda$ as

$$\epsilon_{\Lambda}(\sigma) \triangleq \max_{x \in \mathbb{R}^n} |\text{Vol}(\Lambda)f_{\sigma, \Lambda}(x) - 1|.$$

It can be interpreted as the maximum variation of $f_{\sigma, \Lambda}(x)$ from the uniform distribution over $R(\Lambda)$. The flatness factor can be calculated using the theta series

$$\epsilon_{\Lambda}(\sigma) = \max_{x \in \mathbb{R}^n} |\text{Vol}(\Lambda)f_{\sigma, \Lambda}(x) - 1|.$$

where $\theta_{\Lambda}(\tau) = \sum_{\lambda \in \Lambda} e^{-\pi||\lambda||^2}$. If the flatness factor is negligible, the discrete Gaussian distribution over a lattice preserves the capacity of the AWGN channel [17].

We define the discrete Gaussian distribution over $\Lambda$ centred at $c$ as the following discrete distribution taking values in $\lambda \in \Lambda$:

$$D_{\Lambda, \sigma, c}(\lambda) = \frac{f_{\sigma, c}(\lambda)}{\sum_{\lambda \in \Lambda} f_{\sigma, c}(\lambda)}, \quad \forall \lambda \in \Lambda.$$

Where $f_{\sigma, c}(\lambda) \triangleq \sum_{\lambda \in \Lambda} f_{\sigma, c}(\lambda)$. Again for convenience, we write $D_{\Lambda, \sigma} = D_{\Lambda, \sigma, 0}$. This distribution will be used for lattice shaping.

### 4.2 Polar Lattice Wiretap Coding for mod-$\Lambda_0$ GWC
A sublattice \( \Lambda' \subset \Lambda \) induces a partition (denoted by \( \Lambda / \Lambda' \)) of \( \Lambda \) into equivalence classes modulo \( \Lambda' \). The order of the partition is denoted by \( |\Lambda / \Lambda'| \), which is equal to the number of cosets. If \( |\Lambda / \Lambda'| = 2 \), we call this a binary partition. Let \( \Lambda / \Lambda_1 / \cdots / \Lambda_{r-1} / \Lambda' \) for \( r \geq 1 \) be an \( n \)-dimensional lattice partition chain. For each partition \( \Lambda_{l-1} / \Lambda_l \) (\( 1 \leq l \leq r \) with convention \( \Lambda_0 = \Lambda \) and \( \Lambda_r = \Lambda' \)) a code \( C_l \) over \( \Lambda_{l-1} / \Lambda_l \) selects a sequence of representatives \( a_l \) for the cosets of \( \Lambda_l \). Consequently, if each partition is binary, the code \( C_l \) is a binary code.

Polar lattices are constructed by Construction D [18] using a set of nested polar codes \( C_1 \subseteq C_2 \subseteq \cdots \subseteq C_r \). Suppose \( C_r \) has block length \( N \) and the number of information bits \( k_r \) for \( 1 \leq l \leq r \). Choose a basis \( g_1, g_2, \ldots, g_N \) from the polar generator matrix \( G_r \) such that \( g_1, g_2, \ldots, g_k \) span \( C_r \). When the dimension \( n = 1 \), the lattice \( L \) admits the form [18]

\[
L = \left\{ \sum_{i=1}^{r} 2^{l-1} \sum_{i=1}^{n-1} u_i^j g_i + 2^r z^N u_i^j \in \{0,1\} \right\}
\]

where the addition is carried out in \( \mathbb{R}^N \). The fundamental volume of a lattice obtained from this construction is given by

\[
\text{Vol}(L) = 2^{-N R_C} \cdot \text{Vol}(\Lambda)^N.
\]

where \( R_C = \frac{1}{N} \sum_{i=1}^{r} k_i \) denotes the sum rate of component codes. In this work, we limit ourselves to the binary lattice partition chain and binary polar codes for simplicity.

### 4.2.1 Mod-\( \Lambda_s \) GWC

Now we consider the construction of secrecy-good polar lattices over the mod-\( \Lambda_s \) GWC shown in Fig. 6. The difference between the mod-\( \Lambda_s \) GWC and the genuine GWC is the mod-\( \Lambda_s \) operation on the received signal of Bob and Eve. With some abuse of notation, the outputs \( Y^{[N]} \) and \( Z^{[N]} \) at Bob and Eve's ends respectively become:

\[
Y^{[N]} = X^{[N]} + W_b^{[N]} \mod \Lambda_s
\]

\[
Z^{[N]} = X^{[N]} + W_e^{[N]} \mod \Lambda_s.
\]

Let \( \Lambda_s \) and \( \Lambda_e \) be constructed from a binary partition chain \( \Lambda / \Lambda_s / \cdots / \Lambda_{r-1} / \Lambda' \), and assume \( \Lambda_s \subset \Lambda_r^{N} \subset \Lambda_e \subset \Lambda_o \). Also, denote by \( X^{[r]} \) the bits encoding \( \Lambda_r^{N} / \Lambda_s^{N} \), which include all information bits for message \( M \) as a subset. We have that \( X^{[N]} + W_e^{[N]} \mod \Lambda_s^{N} \) is a sufficient statistic for \( X^{[r]} \).

In our context, we identify \( \Lambda \) with \( \Lambda_r^{N} \) and \( \Lambda' \) with \( \Lambda_s \), respectively. Since the bits encoding \( \Lambda_r^{N} / \Lambda_s \) are uniformly distributed, the mod-\( \Lambda_s \) operation is information-lossless in the sense that

\[
I(X^{[r]}, Z^{[N]}) = I(X^{[r]}, [X^{[N]} + W_b^{[N]}] \mod \Lambda_s^{N})
\]

As far as mutual information \( I(X^{[r]}, Z^{[N]}) \) is concerned, we can use the mod-\( \Lambda_s \) operator instead of the mod-\( \Lambda_s \) operator here. Under this condition, similarly to the multilevel lattice structure introduced in [18], the mod-\( \Lambda_s \) channel can be decomposed into a series of BMS channels according to the partition chain \( \Lambda / \Lambda_s / \cdots / \Lambda_{r-1} / \Lambda' \). Therefore, the already mentioned polar coding technique for BMS channels can be employed. Moreover, the channel resulted from the lattice partition chain can be proved to be equivalent to that based on the chain rule of mutual information. Following this channel equivalence, we can construct an AWGN-good lattice \( \Lambda_s \) and a secrecy-good lattice \( \Lambda_e \), using the wiretap coding technique introduced in section 3.3 and section 3.4 at each partition level.
A mod-$\Lambda$ channel is a Gaussian channel with a modulo $\Lambda$ operator in the front end. The capacity of the mod-$\Lambda$ channel is [18]:

$$C(\Lambda, \sigma^2) = \log(\text{Vol}(\Lambda)) - h(\Lambda, \sigma^2),$$

where $h(\Lambda, \sigma^2)$ is the differential entropy of the $\Lambda$-aliased noise over $V(\Lambda)$:

$$h(\Lambda, \sigma^2) = -\int_{V(\Lambda)} f_{\sigma, \Lambda}(t) \log(f_{\sigma, \Lambda}(t)) dt.$$ 

The differential entropy is maximized to $\log(\mathbb{E}_{\Lambda, \sigma \Lambda}(\Lambda))$ by the uniform distribution over $\Lambda$ [18]. The $\Lambda^{-1}/\Lambda$ channel is defined as a mod-$\Lambda^{-1}$ channel whose input is drawn from $\Lambda^{-1} \cap V(\Lambda)$. It is known that the $\Lambda^{-1}/\Lambda$ channel is symmetric (``regular'' in the sense of Delsarte and Piret and symmetric in the sense of Gallager [18]), and the optimum input distribution is uniform [18]. Furthermore, the $\Lambda^{-1}/\Lambda$ channel is binary if $|\Lambda^{-1}/\Lambda| = 2$. The capacity of the $\Lambda^{-1}/\Lambda$ channel for Gaussian noise of variance $\sigma^2$ is given by [18]

$$C(\Lambda^{-1}/\Lambda, \sigma^2) = C(\Lambda, \sigma^2) - C(\Lambda^{-1}, \sigma^2) = h(\Lambda^{-1}, \sigma^2) - h(\Lambda, \sigma^2) + \log(\text{Vol}(\Lambda)/\text{Vol}(\Lambda^{-1}))$$

The decomposition into a set of $\Lambda^{-1}/\Lambda$ channels is used in [18] to construct AWGN-good lattices. Take the partition chain $\mathbb{Z}/2\mathbb{Z}/\ldots/2^{l-1}\mathbb{Z}$ as an example. Given uniform input $X_1^{l-1}$, let $K_l$ denote the coset indexed by $x_1^{l-1}$, i.e., $K_l = x_1 + \ldots + 2^{l-1}x_l + 2^l\mathbb{Z}$. The conditional probability distribution function (PDF) of this channel with binary input $X_l$ and output $Z_l = Z \mod \Lambda_l$ is

$$f_{Z_l|X_l} = \frac{1}{\sqrt{2\pi} \sigma_l} \sum_{a \in K_l} \exp \left( -\frac{z-a^2}{{2\sigma}_l^2} \right),$$

Since the previous input bits $x_1^{l-1}$ cause a shift on $K_l$ and will be removed by the multistage decoder at level $l$, the code can be designed according to the channel transition probability $f_{Z_l|X_l}$ with $x_1^{l-1} = 0$. Following the notation of [18], we use $V(\Lambda^{-1}/\Lambda, \sigma^2)$ and $W(\Lambda^{-1}/\Lambda, \sigma^2)$ to denote the $\Lambda^{-1}/\Lambda$ channel for Bob and Eve respectively. The $\Lambda^{-1}/\Lambda$ channel can also be used to construct secrecy-good lattices. In order to bound the information leakage of the wiretapper’s channel, we firstly express $I(X_l; Z)$ according to the chain rule of mutual information as

$$I(X_l; Z) = I(X_l; Z) + I(X_l; Z|X_1^{l-1}) + \ldots + I(X_l; Z|X_1^{l-1})$$

This equation still holds if $Z$ denotes the noisy signal after the mod-$\Lambda_l$ operation, namely, $Z = [X + W_l] \mod \Lambda_l$. We will adopt this notation in the rest of this section. We refer to the $l$-th channel associated with mutual information $I(X_l; Z|X_1^{l-1})$ as the equivalent channel denoted by $W'(X_l; Z|X_1^{l-1})$, which is defined as the channel from $X_l$ to $Z$ given the previous $X_1^{l-1}$. Then the transition probability distribution of $W'(X_l; Z|X_1^{l-1})$ is

$$f_{Z_l|X_l} = \frac{1}{|\Lambda_l/\Lambda_r|} \frac{1}{\sqrt{2\pi\sigma_r}} \sum_{a \in K_l} \exp \left( -\frac{z-2a^2}{{2\sigma}_r^2} \right), \quad z \in V(\Lambda_r)$$

From the two PDFs, we can observe that the channel output likelihood ratio (LR) of the $W'(\Lambda_l/\Lambda_r, \sigma^2)$ channel is equal to that of the $l$-th equivalent channel $W'(X_l; Z|X_1^{l-1})$. Consider a lattice $L$ constructed by a binary lattice partition chain $\Lambda/\Lambda_l/\ldots/\Lambda_1/\Lambda'$. Constructing a polar code for the $l$-th equivalent binary-input channel $W'(X_l; Z|X_1^{l-1})$ defined by the chain rule is equivalent to constructing a polar code for the $\Lambda_1/\Lambda_l$ channel $W(\Lambda_1/\Lambda_l, \sigma^2)$.
Figure 7: The multilevel lattice coding system over the mod-$\Lambda_3$ Gaussian wiretap channel.

Now it is ready to introduce the polar lattice construction for the mod-$\Lambda_3$ GWC shown in Fig. 7. A polar lattice $L$ is constructed by a series of nested polar codes $C_1(\mathcal{N}, k_1) \subseteq C_2(\mathcal{N}, k_2) \subseteq \cdots \subseteq C_r(\mathcal{N}, k_r)$ and a binary lattice partition chain $\Lambda/\Lambda_1/\cdots/\Lambda_{r-1}/\Lambda'$. The block length of polar codes is $N$. Alice splits the message $M$ into $\mathcal{M}, \mathcal{R}_2 \cdots \mathcal{R}_r$. We follow the same rule of polar wiretap coding to assign bits in the component polar codes to achieve strong secrecy. Note that $W(\Lambda_{r-1}/\Lambda_1, \sigma^{(r)}_2)$ is degraded with respect to $W(\Lambda_{r-1}/\Lambda_1, \sigma^{(r)}_1)$ for $1 \leq l \leq r$ because $\sigma^{(r)}_2 \leq \sigma^{(r)}_1$. Treating $W(\Lambda_{r-1}/\Lambda_1, \sigma^{(r)}_1)$ and $W(\Lambda_{r-1}/\Lambda_1, \sigma^{(r)}_2)$ as the main channel and wiretapper's channel at each level and using the partition rule of polar wiretap coding, we can get four sets $\mathcal{A}_l, \mathcal{B}_l, \mathcal{C}_l$ and $\mathcal{D}_l$. Similarly, we assign the bits as follows

$$A_l \leftarrow M_l, B_l \leftarrow R_l, C_l \leftarrow F_l, D_l \leftarrow R_l$$

for each level $l$ where $M_l, F_l$ and $R_l$ represent message bits, frozen bits (could be set as all zeros) and random bits at level $l$. Since the $\Lambda_{l-1}/\Lambda_1$ channel is degraded with respect to the $\Lambda_l/\Lambda_{l+1}$ channel, it is easy to obtain that $C_l \supseteq C_{l-1}$, which means $A_l \cup B_l \cup D_l \equiv A_{l+1} \cup B_{l+1} \cup D_{l+1}$. This construction is clearly a lattice construction as polar codes constructed on each level are nested.

Interestingly, the above multilevel construction yields an AWGN-good lattice $\Lambda_\mu$ and a secrecy-good lattice $\Lambda_\kappa$ simultaneously. More precisely, $\Lambda_\mu$ is constructed from a set of nested polar codes $C_1(\mathcal{N}, \mathcal{A}_1) + |\mathcal{B}_1| + |\mathcal{D}_1| \subseteq \cdots \subseteq C_r(\mathcal{N}, \mathcal{A}_r) + |\mathcal{B}_r| + |\mathcal{D}_r|$, and $\Lambda_\kappa$ is constructed from a set of nested polar codes $C_1(\mathcal{N}, |\mathcal{B}_1| + |\mathcal{D}_1|) \subseteq \cdots \subseteq C_r(\mathcal{N}, |\mathcal{B}_r| + |\mathcal{D}_r|)$ and with the same lattice partition chain. More details about the AWGN-goodness of $\Lambda_\mu$ are given in the next section. It is clear that $\Lambda_\kappa \subseteq \Lambda_{\mu_\kappa}$. Thus, our proposed coding scheme instantiates the coset coding scheme introduced in [14], where the confidential message is mapped to the coset $\tilde{\lambda}_m \in \Lambda_\mu/\Lambda_\kappa$.

By using the results of polar wiretap coding, we have

$$I(M; Z^{[N]}_l) \leq N 2^{-N \beta'}$$

where $Z_l^{[N]} = Z^{[N]} \text{ mod } \Lambda_l$. In other words, the employed polar code for the channel $W'(\Lambda_{l-1}/\Lambda_1, \sigma^{(l)}_2)$ can guarantee that the mutual information between the input message and the output is upper bounded by $N 2^{-N \beta'}$. According to the channel equivalence, this polar code can also guarantee the same upper bound on the mutual information between the input message and the output of the channel $\tilde{W}'(X_l; Z|X_{l-1})$ as shown in the following inequality ($X_l$ is independent of the previous $X_{l-1}$):

$$I(M; Z^{[N]}_{1:l-1}) \leq N 2^{-N \beta'}$$

Recall $Z^{[N]}$ is the signal received by Eve after the mod-$\Lambda_1$ operation. From the chain rule of mutual information,

$$I(M; Z^{[N]}) = \sum_{l=1}^{r} I(M; Z^{[N]}_{1:l-1})$$
where the last inequality holds because \( I(M_i; Z^{[N]}| X^{[N]}_{1:l}) = I(M_i; Z^{[N]}, U^{[N]}_{1:l}) \) and adding more variables will not decrease the mutual information. Therefore strong secrecy is achieved since \( \lim_{N \to \infty} I(M; Z^{[N]}) = 0 \).

Note that the above analysis actually implies semantic security, i.e., \( \lim_{N \to \infty} I(M; Z^{[N]}) = 0 \) holds for arbitrarily distributed \( M \). This is because of the symmetric nature of the \( \Lambda_{l-1}/\Lambda_1 \) channel [18]. Since the message \( M \) is drawn from \( R(\Lambda_e) \) and the random bits are drawn from \( \Lambda_e \cap R(\Lambda_e) \), the mod-\( \Lambda_e \) mapping is information lossless and its output is a sufficient statistic for \( M \). In this sense, the channel between the confidential message and the Eavesdropper's signal can be viewed as a \( \Lambda_b/\Lambda_e \) channel. Since the \( \Lambda_b/\Lambda_e \) channel is symmetric, the maximum mutual information is achieved by the uniform input. Consequently, the mutual information corresponding to other input distributions can also be upper bounded by \( rN2^{-N\beta'} \).

We then have the theorem for the mod-\( \Lambda_e \) Gaussian wiretap channel.

Consider a polar lattice \( L \) constructed according to “Construction D” with the binary lattice partition chain \( \Lambda/\Lambda_1/\ldots \Lambda_{r-1}/\Lambda' \). and \( r \) binary nested polar codes with block length \( N \). Scale \( \Lambda \) and \( r \) to satisfy the following conditions:

1. \( h(\Lambda, \sigma_e^2) \to \log(\text{Vol}(L)) \)
2. \( h(\Lambda, \sigma_e^2) \to \frac{1}{2} \log(2\pi e \sigma_e^2) \)

Given \( \sigma_b^2 > \sigma_e^2 \), all strong secrecy rates \( R \) satisfying

\[
R < \frac{1}{2} \log \left( \frac{\sigma_e^2}{\sigma_b^2} \right)
\]

are achievable as \( N \to \infty \), using the polar lattice \( L \) on the mod-\( \Lambda_e \) Gaussian wiretap channel.

### 4.2.2 Reliability Analysis

In the original polar coding scheme for the binary wiretap channel [10], how to assign set \( D \) is a problem. Assigning frozen bits to \( D \) guarantees reliability but only achieves weak secrecy, whereas assigning random bits to \( D \) guarantees strong secrecy but may violate the reliability requirement because \( D \) may be nonempty. In order to ensure strong secrecy, \( D \) is assigned with random bits \( (D \leftarrow R) \), which makes this scheme failed to accomplish the theoretical reliability. For any \( l \)-th level channel \( V(\Lambda_{l-1}/\Lambda_b, \sigma_b^2) \) at Bob’s end, the probability of error is upper bounded by the sum of the Bhattacharyya parameters \( Z(V(\Lambda_{l-1}/\Lambda_b, \sigma_b^2)) \) of subchannels that are not frozen to zero. For each bit-channel index \( j \) and \( \beta < 0.5 \), we have

\[
j \in \mathcal{G}(V(\Lambda_{l-1}/\Lambda_b, \sigma_b^2)) \cup \mathcal{D}_l
\]

By the definition of \( \mathcal{G} \), the sum of \( Z(V(\Lambda_{l-1}/\Lambda_b, \sigma_b^2)) \) over the set \( \mathcal{G}(V(\Lambda_{l-1}/\Lambda_b, \sigma_b^2)) \) is bounded by \( 2^{-N\beta} \), therefore the error probability of the \( l \)-th level channel under the SC decoding, denoted by \( P_{e}^{SC}(\Lambda_{l-1}/\Lambda_b, \sigma_b^2) \), can be upper bounded by

\[
P_{e}^{SC}(\Lambda_{l-1}/\Lambda_b, \sigma_b^2) \leq N2^{-N\beta} + \sum_{j \in \mathcal{D}_l} Z(V^{(j)}_n(\Lambda_{l-1}/\Lambda_b, \sigma_b^2))
\]

Since multistage decoding is utilized, by the union bound, the final decoding error probability for Bob is bounded as

\[
Pr(M \neq \tilde{M}) \leq \sum_{l=1}^{r} P_{e}^{SC}(\Lambda_{l-1}/\Lambda_b, \sigma_b^2)
\]
Unfortunately, a proof that this scheme satisfies the reliability condition cannot be attained here because the bound of the sum \(\sum_{j\in\mathcal{D}_l}Z(V_n^{(j)}(A_{l-1}/A_l,\sigma^2_0))\) is not known. Note that significantly low probabilities of error can still be achieved in practice since the size of \(\mathcal{D}_l\) is very small for sufficiently large \(N\).

The reliability problem was recently solved in [11], where a new scheme dividing the information message into several blocks was proposed. For a specific block, \(\mathcal{D}_l\) is still assigned with random bits and transmitted in advance in the set \(\mathcal{A}_l\) of the previous block. This scheme involves negligible rate loss and finally realizes reliability and strong security simultaneously. In this case, if the reliability of each partition channel can be achieved, i.e., for any \(l\)-th level partition \(\Lambda_{l-1}/\Lambda_l\), \(P_e^{\text{SC}}(A_{l-1}/\Lambda_l,\sigma^2_0)\) vanishes as \(N\to\infty\), then the total decoding error probability for Bob can be made arbitrarily small. Consequently, based on this new scheme of assigning the problematic set, the error probability on level \(l\) can be upper bounded by

\[
P_e^{\text{SC}}(A_{l-1}/\Lambda_l,\sigma^2_0) \leq \epsilon^{N'} + k_l \cdot O(2^{-N'^\beta}),
\]

where \(k_l\) is the number of information blocks on the \(l\)-th level, \(N'\) is the length of each block which satisfies \(N' \times k_l = N\) and \(\epsilon^{N'}\) is caused by the first separate block on the \(l\)-th level consisting of the initial bits in \(\mathcal{D}_l\). Since \(|\mathcal{D}_l|\) is extremely small comparing to the block length \(N\), the decoding failure probability for the first block can be made arbitrarily small when \(N\) is sufficiently large. Meanwhile, when \(h(\Lambda_0,\sigma^2_0) \to \log(\text{Vol}(L))\), \(h(\Lambda_0,\sigma^2_0) \to \frac{1}{2}\log(2\pi e\sigma^2_0)\), and \(R_c \to C(\Lambda)/\Lambda_0,\sigma^2_0)\), we have \(\gamma_{\Lambda_0}(\sigma_0) = 2\pi e\). Therefore, \(\Lambda_0\) is an AWGN-good lattice.

Note that the rate loss incurred by repeatedly transmitted bits in \(\mathcal{D}_l\) is negligible because of its small size. Specifically, the actual secrecy rate in the \(l\)-th level is given by \(\frac{k_l}{k_l+1}[C(\Lambda_{l-1}/\Lambda_l,\sigma^2_0) - C(\Lambda_{l-1}/\Lambda_l,\sigma^2_0)]\). Clearly, this rate can be made close to the secrecy capacity by choosing sufficiently large \(k_l\) as well.

### 4.3 Secrecy-good polar lattices with discrete Gaussian shaping

In this section, we apply Gaussian shaping on the AWGN-good and secrecy-good polar lattices. The idea of lattice Gaussian shaping was proposed in [17] and then implemented in [19] to construct capacity-achieving polar lattices. For wiretap coding, the discrete Gaussian distribution can also be utilized to satisfy the power constraint. In simple terms, after obtaining the AWGN-good lattice \(\Lambda_0\) and the secrecy-good lattice \(\Lambda_c\), Alice still maps each message \(m\) to a coset \(\tilde{\lambda}_m \in \Lambda_0/\Lambda_c\) as mentioned in the previous section. However, instead of the mod \(\Lambda_c\) operation, Alice samples the encoded signal from \(D_{\Lambda_0+\tilde{\lambda}_m,\sigma^2_0}\), where \(\tilde{\lambda}_m\) is the coset representative of \(\tilde{\lambda}_m\) and \(\sigma^2_0\) is arbitrarily close to the signal power \(P_c\) (see [14] for more details). Based on the lattice Gaussian shaping, we will propose a new partition for the genuine GWC. We will also show that this shaping operation does not hurt the secrecy rate and that the proposed scheme is semantically secure.

#### 4.3.1 Gaussian shaping over polar lattices

As shown in [19], the shaping scheme is based on the technique of polar codes for asymmetric channels. We give a brief review in this section.

Similarly to the polar coding on symmetric channels, the Bhattacharyya parameter for a binary memoryless asymmetric (BMA) channel is defined as follows.

Let \(W\) be a BMA channel with input \(X \in \mathcal{X} = \{0,1\}\) and output \(Y \in \mathcal{Y}\). The input distribution and channel transition probability is denoted by \(P_X\) and \(P_{X|Y}\) respectively. The Bhattacharyya parameter \(Z\) for \(W\) is defined as

\[
Z(X|Y) = 2 \sum_{y} \sqrt{P_{X|Y}(0,y) \cdot P_{X|Y}(1,y)}
\]

It is not difficult to find that adding observable at the output of \(W\), \(Z\) will not decrease, i.e., \(Z(X|Y,Y') \leq Z(X|Y)\).

When \(X\) is uniformly distributed, the Bhattacharyya parameter of BMA channels coincides with that of BMS channels defined in Definition. Moreover, the calculation of \(Z\) can be converted to the calculation of the Bhattacharyya parameter for a related BMS channel.
Let \( \tilde{W} \) be a binary input channel corresponding to the asymmetric channel \( W \) with input \( \tilde{x} \in X = \{0,1\} \) and output \( \tilde{y} \in (Y,X) \). The input of \( \tilde{W} \) is uniformly distributed, i.e., \( P_\tilde{x}(\tilde{x}=0) = P_\tilde{x}(\tilde{x}=1) = \frac{1}{2} \). The relationship between \( \tilde{W} \) and \( W \) is shown in Fig. 7. Then \( \tilde{W} \) is a binary symmetric channel in the sense that

\[
P_{\tilde{y}|\tilde{x}}(y|x) = P_{y|x}(y|x)
\]

For a BMA channel \( W \) with input \( X \sim P_x \), let \( \tilde{W} \) be its symmetrized channel constructed according to Fig. 8. Suppose \( X^{[N]} \) and \( Y^{[N]} \) be the input and output vectors of \( W^N \), and let \( \tilde{X}^{[N]} \) and \( \tilde{Y}^{[N]} = (\tilde{X}^{[N]} \oplus \tilde{X}^{[N]}), Y^{[N]} \) be the input and output vectors of \( \tilde{W}^N \). Consider polarized random variables \( U^{[N]} = X^{[N]G_N} \) and \( \bar{U}^{[N]} = X^{[N]}G_N \), and denote by \( U_N \) and \( \bar{U}_N \) the combining channel of \( N \) uses of \( W \) and \( \tilde{W} \), respectively. The Bhattacharyya parameter for each subchannel of \( W_N \) is equal to that of each subchannel of \( \tilde{W}_N \), i.e.,

\[
Z(U_i|U_i^{1:i-1},Y^{[N]}) = Z(\bar{U}_i|\bar{U}_i^{1:i-1},X^{[N]}G_N,Y^{[N]}).
\]

To obtain the desired input distribution of \( P_x \) for \( W \), the indices with very small \( Z(U_i|U_i^{1:i-1}) \) should be removed from the information set of the symmetric channel. Following [19], the resultant subset is referred to as the information set \( I \) for the asymmetric channel \( W \). For the remaining part \( I^c \), we further find out that there are some bits which can be made independent of the information bits and uniformly distributed. The purpose of extracting such bits is for the interest of our lattice construction. We name the set that includes those independent frozen bits as the independent frozen set \( F \), and the remaining frozen bits are determined by the bits in \( I \). We name the set of all those deterministic bits as the shaping set \( S \). The three sets are formally defined as follows:

\[
\begin{align*}
\text{the independent frozen set: } F &\quad= \{i \in [N]; Z(U_i|U_i^{1:i-1},Y^{[N]}) \geq 1 - 2^{-N^P}\} \\
\text{the information set: } I &\quad= \{i \in [N]; Z(U_i|U_i^{1:i-1},Y^{[N]}) \leq 2^{-N^P} \text{ and } Z(U_i|U_i^{1:i-1}) \geq 1 - 2^{-N^P}\} \\
\text{the shaping set: } S &\quad= (F \cup I)^c
\end{align*}
\]

To find these three sets, one can use channel symmetrization to calculate \( Z(U_i|U_i^{1:i-1},Y^{[N]}) \) using the known constructing techniques for symmetric polar codes [20] [21]. We note that \( Z(U_i|U_i^{1:i-1}) \) can be computed in a similar way, by constructing a symmetric channel between \( X \) and \( X \oplus \tilde{X} \). Besides the construction, the decoding process for the asymmetric polar codes can also be converted to the decoding for the symmetric polar codes.

This polar coding scheme can be viewed as an extension of the scheme proposed in [22], has been proved to be capacity-achieving in [19]. Moreover, it can be extended to the construction of multilevel asymmetric polar codes.

Consider a polar code with the following encoding and decoding strategies for the channel of the \( l \)-th (\( 1 \leq l \leq r \)) level \( \tilde{W} \) with the channel transition probability \( P_{\tilde{y}|\tilde{x}_{l:t-1}}(y|x) \).

- **Encoding**: Before sending the codewords \( x_i^{[N]} = u_i^{[N]}G_N \), the index set \( N \) are divided into three parts: the independent frozen set \( F_i \), information set \( I_i \), and shaping set \( S_i \), which are defined as follows:

\[
\begin{align*}
F_i &\quad= \{i \in [N]; Z(U_i^{[N]}|U_i^{1:i-1},X_i^{[N]},Y^{[N]}) \geq 1 - 2^{-N^P}\} \\
I_i &\quad= \{i \in [N]; Z(U_i^{[N]}|U_i^{1:i-1},X_i^{[N]},Y^{[N]}) \leq 2^{-N^P} \text{ and } Z(U_i^{[N]}|U_i^{1:i-1},X_i^{[N]}) \geq 1 - 2^{-N^P}\} \\
S_i &\quad= (F_i \cup I_i)^c
\end{align*}
\]

The encoder first places uniformly distributed information bits in \( I_i \). Then the frozen set \( F_i \) is filled with a uniform random sequence which are shared between the encoder and the decoder. The bits in \( S_i \) are generated by a mapping \( \Phi_{S_i} \) from a family of randomized mappings \( \Phi_{S_i} \), which yields the following distribution:

---

**Figure 8**: The relationship between \( \tilde{W} \) and \( W \)
Decoding: The decoder receives $y^{[N]}$ and estimates $\hat{u}_i^{[N]}$ based on the previously recovered $x_i^{[N]}$ according to the rule

$$ u_i^t = \begin{cases} 
0 & \text{with probability } P_{u_i^t|u_i^t=x_i^{[N]}\cap x_i^{[N]-1}}(0|u_i^t=x_i^{[N]}\cap x_i^{[N]-1}) \\
1 & \text{with probability } P_{u_i^t|u_i^t=x_i^{[N]}\cap x_i^{[N]-1}}(1|u_i^t=x_i^{[N]}\cap x_i^{[N]-1}) 
\end{cases} $$

- Decoding: The decoder receives $y^{[N]}$ and estimates $\hat{u}_i^{[N]}$ based on the previously recovered $x_i^{[N]}$ according to the rule

$$ \hat{u}_i^{[N]} = \begin{cases} 
\hat{u}_i, & \text{if } i \in F_t \\
\phi_i(\hat{u}_i^{[N]}, x_i^{[N]-1}), & \text{if } i \in S_t \\
\text{argmax}_u P_{u_i^t|u_i^t=x_i^{[N]}\cap x_i^{[N]-1}, y^{[N]}}(u|u_i^t=x_i^{[N]}\cap x_i^{[N]-1}, y^{[N]}), & \text{if } i \in I_t 
\end{cases} $$

Note that probability $P_{u_i^t|u_i^t=x_i^{[N]}\cap x_i^{[N]-1}, y^{[N]}}(u|u_i^t=x_i^{[N]}\cap x_i^{[N]-1}, y^{[N]})$ can be calculated by the SC decoding algorithm efficiently, treating $\hat{u}_i$ and $x_i^{[N]}$ (already decoded by the SC decoder at previous levels) as the outputs of the asymmetric channel. With the above encoding and decoding, the message rate can be arbitrarily close to $I(X_i; Y|X_{1:i-1})$ and the expectation of the decoding error probability over the randomized mappings satisfies $E_{\phi_S}[P_e(\phi_S)] = O(2^{-N\beta})$ for any $\beta^r \beta < \beta < 0.5$. Consequently, there exists a deterministic mapping $\phi_S$ such that $P_e(\phi_S) = O(2^{-N\beta^r})$.

In practice, to share the mapping $\phi_S$ between the encoder and the decoder, we can let them have access to the same source of randomness, which can be achieved by initializing the pseudorandom number generators on both sides the same state.

Now let us pick a suitable input distribution $P_{X_i^r}$ to implement the shaping. As shown in [17], the mutual information between the discrete Gaussian lattice distribution $D_{\Lambda, \sigma}$ and the output of the AWGN channel approaches $P_{X_i^r}$ as the flatness factor is negligibly small. Therefore, we use the Gaussian distribution $P_{X_i^r} \sim D_{\Lambda, \sigma}$ as the constellation, which gives us $\lim_{N \to \infty} P_{X_i^r} = P_{X_i^r}^{\Lambda, \sigma}$. When $N \to \infty$, the mutual information $I(X_i; Y|X_{1:i-1})$ at the bottom level goes to 0 if $r = O(|\log \log N|)$, and using the first $r$ levels would involve a capacity loss $\sum_{i=1}^{r} I(X_i; Y|X_{1:i-1}) \leq O(\frac{1}{N})$.

From the chain rule of mutual information,

$$ I(X_{1:i}; Y) = \sum_{i=1}^{r} I(X_i; Y|X_{1:i-1}) $$

we have $r$ binary-input channels and the $i$-th channel according to $I(X_i; Y|X_{1:i-1})$ is generally asymmetric with the input distribution $P_{X_i^r}(1 \leq i \leq r)$. Then we can construct the polar code for the asymmetric channel at each level according to Fig. 8. As a result, the $i$-th symmetrized channel is equivalent to the MMSE-scaled $\Lambda_{1-1}/\Lambda_i$ channel in the sense of channel polarization. (See [19] for more details.)

Therefore, when power constrain is taken into consideration, the multilevel polar codes before shaping are constructed according to the symmetric channel $V(\Lambda_{1-1}/\Lambda_i, \tilde{\sigma}_b^2)$ and $W(\Lambda_{1-1}/\Lambda_i, \tilde{\sigma}_b^2)$, where $\tilde{\sigma}_b^2 = \left(\frac{\sigma_b}{\sqrt{\sigma_b^2 + \sigma_b^2}}\right)^2$ and $\tilde{\sigma}_b^2 = \left(\frac{\sigma_b}{\sqrt{\sigma_b^2 + \sigma_b^2}}\right)^2$ are the MMSE-scaled noise variance of the main channel and of the wiretapper’s channel, respectively. This is similar to the mod-$\Lambda_b$ GWC scenario mentioned in the previous section. The difference is that $\sigma_b^2$ and $\sigma_b^2$ are replaced by $\tilde{\sigma}_b^2$ and $\tilde{\sigma}_b^2$ accordingly. As a result, we can still obtain an AWGN-good lattice $\Lambda_b$ and a secrecy-good lattice $\Lambda_b$ by treating $V(\Lambda_{1-1}/\Lambda_i, \tilde{\sigma}_b^2)$ and $W(\Lambda_{1-1}/\Lambda_i, \tilde{\sigma}_b^2)$ as the main channel and wiretapper’s channel at each level.

### 4.3.2 Three-dimensional partition

Now we consider the partition of the index set $[N]$ with shaping involved. According to the analysis of asymmetric polar codes, we have to eliminate those indices with small $Z(U_i|U_i^{[N]}\cap X_i^{[N]-1})$ from the information set of the symmetric channels. Therefore, Alice cannot send message on those subchannels with $Z(U_i|U_i^{[N]}\cap X_i^{[N]-1}) \leq 1 - 2^{-N^\beta}$. Note that this part is the same for $\tilde{V}_i$ and $\tilde{W}_i$, because it only depends on the shaping distribution. At each level, the index set which is used for shaping is given as

$$ S_i \triangleq \left\{i \in [N] : Z(U_i|U_i^{[N]}\cap X_i^{[N]-1}) \leq 1 - 2^{-N^\beta} \right\} $$
and the index set which is not for shaping is denoted by $S^c_l$. Recall that for the index set $[N]$, we already have two partition criteria, i.e., reliability-good and information-bad. We rewrite the reliability-good index set $\mathcal{G}_l$ and information-bad index set $\mathcal{N}_l$ at level $l$ as

$$\mathcal{G}_l \triangleq \{ i \in [N] : Z \left( U_i^{1:l-1}, X_i^{1:l-1}, Y_i^{[N]} \right) \leq 2^{-N^B} \}$$

$$\mathcal{N}_l \triangleq \{ i \in [N] : Z \left( U_i^{1:l-1}, X_i^{1:l-1}, Y_i^{[N]} \right) \geq 1 - 2^{-N^B} \}$$

Note that $\mathcal{G}_l$ and $\mathcal{N}_l$ are defined by the asymmetric Bhattacharyya parameters. Nevertheless, by channel symmetrization and the channel equivalence, we have $\mathcal{G}_l = \mathcal{G}(\tilde{V}_l)$ and $\mathcal{N}_l = \mathcal{N}(\tilde{W}_l)$ as defined in section 3.2, where $\tilde{V}_l$ and $\tilde{W}_l$ are the respective symmetric channels or the MMSE-scaled $\Lambda_l^{-1/2}$ channels for Bob and Eve at level $l$. The four sets $A_l, B_l, C_l$ and $D_l$ are defined in the same fashion as before, with $\mathcal{G}_l$ and $\mathcal{N}_l$ replacing $\mathcal{G}(\tilde{V}_l)$ and $\mathcal{N}(\tilde{W}_l)$, respectively. Now the whole index set $[N]$ is divided like a cube in three directions, which is shown in Fig. 9.

Clearly, we have eight blocks:

$$A^c_l = A_l \cap S_l A^c_s = A_l \cap S^c_l$$

$$B^c_l = B_l \cap S_l B^c_s = B_l \cap S^c_l$$

$$C^c_l = C_l \cap S_l C^c_s = C_l \cap S^c_l$$

$$D^c_l = D_l \cap S_l D^c_s = D_l \cap S^c_l$$

We can observe that $A^c_l = C^c_l = \emptyset$, $A^c_s = A^c_l$, and $C^c_s = C^c_l$. The shaping set $S_l$ is divided into two sets $B^c_l$ and $D^c_l$. The bits in $S^c_l$ are determined by the bits in $S^c_l$ according to the mapping. Similarly, $S^c_l$ is divided into the four sets $A^c_s$, $B^c_s$, $C^c_s$, and $D^c_s$. Note that for wiretap coding, the frozen set becomes $C^c_s$, which is slightly different from the frozen set for channel coding. To satisfy the reliability condition, the frozen set $C^c_s$ and the problematic set $D^c_s$ cannot be set uniformly random any more. Recall that only the independent frozen set $F_i$ at each level, which is defined as $\{ i \in [N] : Z(U_i^{1:l-1}, X_i^{1:l-1}, Y_i^{[N]} \geq 1 - 2^{-N^B}) \}$, can be set uniformly random (which are already shared between Alice and Bob), and the bits in the unpolarized frozen set $F_i$, defined as $\{ i \in [N] : 2^{-N^B} <$
$Z(U_t^i|U_t^{1:i-1}, X^{[N]}_{1:i-1}, Y^{[N]}_{1:i}) < 1 - 2^{-N^R}$, should be determined according to the mapping. Moreover, we can observe that $F_t \subseteq C_t^c$ and $D_t^c \subseteq D_t \subseteq F_t$. Here we make the bits in $F_t$, uniformly random and the bits in $C_t^c \setminus F_t$ and $D_t^c$ determined by the mapping. Therefore, from now on, we adjust the definition of the shaping bits as:

$$S_t \triangleq \{i \in [N]: Z(U_t^i|U_t^{1:i-1}, X^{[N]}_{1:i-1}) < 1 - 2^{-N^R} \text{ or } 2^{-N^R} < Z(U_t^i|U_t^{1:i-1}, Y^{[N]}_{1:i}) < 1 - 2^{-N^R} \}$$

To sum up, at level $t$, we assign the sets $A_t^c$, $B_t^c$, and $F_t$ with message bits $M_t$, uniformly random bits $R_t$, and uniform frozen bits, respectively. The rest bits $S_t$ will be fed with random bits according to $P_{U_t^i|U_t^{1:i-1}, X^{[N]}_{1:i-1}}$. Clearly, this shaping operation will make the input distribution arbitrarily close to $P_{X_t|X_{1:t-1}}$. In this case, we can obtain the equality between the Bhattacharyya parameter of asymmetric setting and symmetric setting. This provides us a convenient way to prove the strong secrecy of the wiretap coding scheme with shaping because we have already proved the strong secrecy of a symmetric wiretap coding scheme using the Bhattacharyya parameter of the symmetric setting. A detailed proof will be presented in the following section. Before this, we show that the shaping will not change the message rate.

For the symmetrized main channel $\bar{V}_t$ and wiretapper’s channel $\bar{W}_t$, consider the reliability-good indices set $\mathcal{G}_t$ and information-bad indices set $\mathcal{N}_t$. By eliminating the shaping set $S_t$ from the original message set, we get the new message set $A_t^c = \mathcal{G}_t \cap \mathcal{N}_t \cap S_t$. The proportion of $|A_t^c|$ equals to that of $|A_t|$, and the message rate after shaping can still be arbitrarily close to $\frac{1}{2} \log (\frac{\alpha_2^2}{\beta_2^2})$.

### 4.3.3 Strong secrecy

In [10], an induced channel is defined in order to prove strong secrecy. Here we call this induced channel randomness-induced channel because it is induced by feeding the subchannels in the sets $B_t$ and $D_t$ with uniformly random bits. However, when shaping is involved, the set $B_t$ and $D_t$ are no longer fed with uniformly random bits. In fact, some subchannels (covered by the shaping mapping) should be fed with bits according to the distribution $P_{U_t^i|U_t^{1:i-1}, X^{[N]}_{1:i-1}}$. We define the channel induced by the shaping bits as the shaping-induced channel.

The shaping-induced channel $Q_N(W, S)$ is defined in terms of $N$ uses of an asymmetric channel $W$, and a shaping subset $S$ of $[N]$ of size $|S|$. The input alphabet of $(Q_N(W, S))$ is $\{0, 1\}^{N-|S|}$ and the bits in $S$ are determined by the input bits according to a specific shaping mapping $\phi$. A block diagram of the shaping induced channel is shown in Fig. 10.

![Block diagram of the shaping-induced channel $Q_N(W, S)$](image)

Based on the shaping-induced channel, we define a new induced channel, which is caused by feeding a part of the input bits of the shaping-induced channel with uniformly random bits.
The new induced channel $Q_{\psi}(W, S, R)$ is specified in terms of a randomness subset $R$ of size $|R|$. The randomness is introduced into the input set of the shaping-induced channel. The input alphabet of $Q_{\psi}(W, S, R)$ is $\{0,1\}^{N-|S|-|R|}$ and the bits in $R$ are uniformly and independently random. A block diagram of the new induced channel is shown in Fig. 11.

The new induced channel is a combination of the shaping-induced channel and randomness-induced channel. This is different from the definition given in [10] because the bits in $S$ are neither independent to the message bits nor uniformly distributed. As long as the input bits of the new induced channel are uniform and the shaping bits are chosen according to the shaping mappings, the new induced channel can still generate $2^N$ possible realizations $x_i^{[N]}$ of $X_i^{[N]}$ as $N$ goes to infinity, and those $x_i^{[N]}$ can be viewed as the output of $N$ i.i.d binary sources with input distribution $P_X(x_{1:N})$. Specifically, we have $Z(U[|U^{1:i-1}, Y^{[N]}] = Z(U[|U^{1:i-1}, X^{[N]}]) \oplus X^{[N]}, Y^{[N]})$. In simple words, this equation holds when $x_i^{[N]}$ and $x^{[N]} \oplus X^{[N]}$ are all selected from $\{0,1\}^N$ according to their respective distributions. Then we can exploit the relation between the asymmetric channel and the corresponding symmetric channel to bound the mutual information of the asymmetric channel. Therefore, we have to stick to the input distribution (uniform) of our new induced channel and also the distribution of the mappings. This is similar to the setting of the randomness induced channel in [10], where the input distribution and the randomness distribution are both set to be uniform. In [10], the randomness-induced channel is further proved to be symmetric; then any other input distribution can also achieve strong secrecy and the symmetry finally results in semantic security. In this work, however, we do not have a proof of the symmetry of the new induced channel. For this reason, we assume for now that the message bits are uniform distributed. To prove semantic security, we will show that the information leakage of the symmetrized version of the new induced channel is in section 4.2.

Let $M_i$ be the uniformly distributed message bits and $F_i$ be the independent frozen bits at the input of the channel at the $l$-th level after shaping, we have

$$I(M_i; F_i; Z^{[N]}, X_{1:l-1}^{[N]}) \leq 2N2^{-N^\beta}$$

Finally, the strong secrecy (for uniform message bits) can be proved in the same fashion as shown previously:

$$I(M_i; Z^{[N]}) \leq \sum_{i=1}^{r} I(M_i; Z^{[N]}, X_{1:l-1}^{[N]}) \leq \sum_{i=1}^{r} I(M_i; F_i; Z^{[N]}, X_{1:l-1}^{[N]}) \leq 2rN2^{-N^\beta}$$

An unavoidable question is that whether the shaping bits $S_l$ make the message $M_i$ insecure when Eve knows the frozen bits and the mapping beforehand. It is possible for Eve to decode some bits in $S_l$ because the mutual information $I(M_i; F_i; S_l)$ is not vanishing. In this case, it seems that Eve may use the decoded bits to decide other shaping bits and then get some information about the message bits. In fact, this is not the case, because bits in $S_l$ which are decodable for Eve turn out to be irrelevant to the message $M_i$. Let us assume a specific shaping bit $s$ which can be decoded by Eve from the knowledge of frozen bits and mapping. Then the index of this $s$ should precede that of any message bits. We can easily obtain that $I(U_i; U_i^{1:i-1}, X_{1:l-1}^{[N]}) \to 0$ for any $i \in A_i$. This means that $s$ is almost independent of the message and is determined by the other bits. Knowing this $s$ will not help Eve to get extra information about the message bits. Therefore we conclude that the whole shaping scheme is secure in the sense that the mutual information leakage between $M$ and $Z^{[N]}$ vanishes with the block length $N$.

4.3.4 Achieving the Secrecy Capacity
The reliability analysis in section 4.2.2 holds for the wiretap coding without shaping. When shaping is involved, the problematic set $D_i$ at each level is included in the shaping set $S_i$. The bits in $D_i$ can be recovered by Bob simply by the shared mapping but not requiring the blocking technique [11]. Consequently, the reliability at each level can be guaranteed by uniformly distributed independent frozen bits and random mapping with distribution $P_{u_i|u_i=x_i^{[N]}}$. By the multilevel decoding and union bound, the expectation of the block error probability of our wiretap coding scheme is vanishing as $N \to \infty$.

It remains to illustrate that the secrecy rate approaches the secrecy capacity. For some $\epsilon' \to 0$, we have

$$
\lim_{N \to \infty} R = \sum_{i=1}^{r} \lim_{N \to \infty} \frac{|A_i^{s_i}|}{N} = \sum_{i=1}^{r} I(X_i; Y|X_1, \ldots, X_{i-1}) - I(X_i; Z|X_1, \ldots, X_{i-1})
$$

$$
= \frac{1}{2} \log \left( \frac{\sigma^2_c}{\sigma^2_b} \right) - \epsilon'
$$

$$
\geq 1/2 \log \left( \frac{1 + SNR_b}{1 + SNR_c} \right) - \epsilon'
$$

Consider a multilevel lattice code constructed from polar codes based on asymmetric channels and lattice Gaussian shaping $D_{A,\sigma^2}$. Given $\sigma^2_c \geq \sigma^2_b$, let $\epsilon_A(\tilde{\epsilon}_c)$ be negligible and set the number of levels $r = O(\log \log N)$ for $N \to \infty$. Then all strong secrecy rates $R$ satisfying $R < \frac{1}{2} \log \left( \frac{1 + SNR_b}{1 + SNR_c} \right)$ are achievable for the Gaussian wiretap channel, where $SNR_b$ and $SNR_c$ denote the $SNR$ of the main channel and wiretapper's channel, respectively.
5 Wiretap Coding for MIMO and Fading Channels

In this section, we consider MIMO (Multiple Input Multiple Output) wiretap channels, where a legitimate transmitter Alice is communicating with a legitimate receiver Bob in the presence of an eavesdropper Eve, and communication is done via MIMO channels. We suppose that Alice’s strategy is to use a codebook which has a lattice structure, which then allows her to perform coset encoding. We analyse Eve’s probability of correctly decoding the message Alice meant to Bob, and from minimizing this probability, we derive a code design criterion for MIMO lattice wiretap codes. The case of block fading channels is treated similarly, and fast fading channels are derived as a particular case. The Alamouti code is carefully studied as an illustration of the analysis provided.

Similarly to the Gaussian wiretap coding case mentioned in section 4, we consider the case where Alice transmits lattice codes using coset encoding, which requires two nested lattices \( \Lambda_s \subset \Lambda_b \). Alice encodes her data in the coset representatives of \( \Lambda_b / \Lambda_s \). Both Bob and Eve try to decode using coset decoding. It was shown in [23] for Gaussian channels that a wiretap coding strategy is to design \( \Lambda_b \) for Bob (since Alice knows Bob’s channel, she can ensure he will decode with high probability), while \( \Lambda_s \) is chosen to maximize Eve’s confusion, characterized by a lattice invariant called secrecy gain, under the assumption that Eve’s noise is worse than the one experienced by Bob. The contribution of this work is to generalize this approach to MIMO channels (and in fact block and fast fading channels as particular cases). We compute Eve’s probability of making a correct decoding decision, and deduce how the lattice \( \Lambda_e \) should be designed to minimize this probability. A MIMO wiretap channel will then consist of two nested lattices \( \Lambda_s \subset \Lambda_b \) where \( \Lambda_b \) is designed to ensure Bob’s reliability, while \( \Lambda_e \) is a subset of \( \Lambda_b \) chosen to increase Eve’s confusion.

5.1 The MIMO Case

We now consider the case when the channel between Alice and Bob, resp. Eve, is a quasi-static MIMO channel with \( n_t \) transmitting antennas at Alice’s end, \( n_b \) resp. \( n_e \) receiving antennas at Bob’s, resp. Eve’s end, and a coherence time \( T \), that is:

\[
Y = H_b X + V_b,
\]

\[
Z = H_e X + V_e,
\]

where the transmitted signal \( X \) is a \( n \times T \) matrix, the two channel matrices are of dimension \( n_b \times n_t \) for \( H_b \), and \( n_e \times n_t \) for \( H_e \), and \( V_b \), \( V_e \) are \( n_b \times T \), resp. \( n_e \times T \) matrices denoting the Gaussian noise at Bob, respectively Eve’s side, both with coefficients zero mean, and respective variance \( \sigma_b^2 \) and \( \sigma_e^2 \). The fading coefficients are complex Gaussian i.i.d. random variables, and in particular \( H_e \) has covariance matrix \( \sum_{\ell} \sigma_{\ell}^2 I_{n_b} \). As for the Gaussian case (as described in Section IV), we assume that Alice transmits a lattice code, via coset encoding, and that the two receivers are performing coset decoding of the lattice, thus \( n_b \geq n_t \). Indeed, if the number of antennas at the receiver is smaller than that of the transmitter, the lattice structure is lost at the receiver. This case will not be treated. That \( n_e \geq n_t \) might be assumed without loss of generality, since in this case Eve is in a more advantageous situation than if she had less antennas. Finally, we denote by \( \gamma_e = \sigma_e^2 / \sigma_b^2 \) Eve’s SNR. We do not make assumption on knowing Eve’s channel or on Eve’s SNR, since we will compute bounds which are general, though their tightness will depend on Eve’s SNR.

In order to focus on the lattice structure of the transmitted signal, we vectorize the received signal and obtain
We now interpret the $nT$ codeword $X$ as coming from a lattice. This is typically the case if $X$ is a space-time code coming from a division algebra [24], or more generally if $X$ is a linear dispersion code as introduced in [25] where $Tn$ symbols QAM are linearly encoded via a family of $Tn$ dispersion matrices. We write

$$\text{vec}(H_x X) = M_b u,$$

where $u \in \mathbb{Z}[i]^{Tn}$ and $M_b$ denotes the $Tn \times Tn$ generator matrix of the $\mathbb{Z}[i]$-lattice $\Lambda_b$ intended to Bob. Thus, in what follows, by a lattice point $x \in \Lambda_b$, we mean that

$$x = \text{vec}(X) = M_b u,$$

and similarly for a lattice point $x \in \Lambda_e$, we have

$$x = \text{vec}(X) = M_e u.$$

By setting

$$M_{b,h_i} = \text{diag}(H_b, \ldots, H_b)M_b,$$
$$M_{b,h_i} = \text{diag}(H_e, \ldots, H_e)M_e,$$

we can rewrite $\text{vec}(Y)$ and $\text{vec}(Z)$ as

$$\text{vec}(Y) = M_{b,h_i} u + \text{vec}(V_b),$$
$$\text{vec}(Z) = M_{b,h_i} u + \text{vec}(V_e),$$

where $M_{b,h_i}$, resp. $M_{b,h_i}$ can be interpreted as the lattice generators of the lattices $\Lambda_{b,h_i}$, resp. $\Lambda_{b,h_i}$, representing the transmitted lattice seen through the respective receivers’ channel, with by definition volume

$$\text{vol}(\Lambda_{b,h_i}) = |\text{det}(M_{b,h_i}M_{b,h_i}^-)| = |\text{det}(H_b H_b^*)|^T \text{vol}(\Lambda_b),$$
$$\text{vol}(\Lambda_{b,h_i}) = |\text{det}(M_{b,h_i}M_{b,h_i}^-)| = |\text{det}(H_e H_e^*)|^T \text{vol}(\Lambda_b).$$

Similarly, the lattices $\Lambda_{b,h_i}$, resp. $\Lambda_{b,h_i}$ describe the lattices intended to Eve, seen through Bob’s, resp. Eve’s channel, with respective generator matrix $M_{b,h_i} = \text{diag}(H_b, \ldots, H_b)M_b$ and $M_{b,h_i} = \text{diag}(H_e, \ldots, H_e)M_e$. Note that for $r \in \Lambda_{e,h_i}$, we have

$$||r||^2 = ||\text{diag}(H_b, \ldots, H_b) M_b u||^2 = ||\text{diag}(H_b, \ldots, H_b) x||^2 = ||H_b X||_F^2$$

where $||H_b X||_F^2 = \text{Tr}(H_b X X^H)$ is the Frobenius norm, and $x = \text{vec}(X) \in \Lambda_e$.

For a given realization of the channel matrices $H_b$ and $H_e$, the channel can be seen as the Gaussian wiretap channel

$$y = M_{b,h_i} u + v_b,$$
$$y = M_{b,h_i} u + v_e,$$

where $y = \text{vec}(Y), z = \text{vec}(Z), v_b = \text{vec}(v_b), v_e = \text{vec}(v_e)$. We now focus on Eve’s channel, since we know from [26] how to design a good linear dispersion space-time code, and the lattice $\Lambda_b$ is chosen so as to correspond to this space-time code. We also know that Eve’s probability of correctly decoding is

$$P_{c,e,h_i} \leq \frac{\text{vol}(\Lambda_{b,h_i})}{(2\pi\sigma_e^2)^{nT}} \sum_{r \in \Lambda_{b,h_i}} e^{- ||r||^2 / 2\sigma_e^2} = \frac{\text{vol}(\Lambda_b)}{(2\pi\sigma_e^2)^{nT}} \sum_{r \in \Lambda_b} e^{- ||r||^2 / 2\sigma_e^2}.$$
with \( x = \text{vec}(X) \in \Lambda_e \). Note that as mentioned at the end of Section 2, the exponent of \( 2 \pi^2 \sigma_e^2 \) depends on the dimension of the transmitted lattice, which is here \( Tn_T \). Using Equation of error probability, we derive Eve’s average probability of correct decision:

\[
\overline{P}_{e,c} = \mathbb{E}_{H}\left[ P_{e,c,H} \right] \leq \frac{\text{vol}(\Lambda_e)}{(2 \pi^2 \sigma_e^2)^{2n_T}} \sum_{x \in \Lambda_e} \int_{e^*} \det(H_x H_x^\dagger) e^{-Tn_T \left( \frac{1}{2 \sigma_e^2} \frac{1}{2 \sigma_e^2} \right) XX^\dagger} \text{d}H_x.
\]

By setting \( W = H_x H_x^\dagger \), we note that the above integral can be rewritten as

\[
\int_{W \in D_w} \int_{H \in W} \det(H_x H_x^\dagger)^T e^{-Tn_T \left( \frac{1}{2 \sigma_e^2} \frac{1}{2 \sigma_e^2} \right) XX^\dagger} \text{d}H_x \text{d}W,
\]

where \( D_w \) is the set of all \( n_i \times n_i \) positive definite Hermitian matrices. Now, \( \overline{P}_{e,c} \) becomes

\[
\overline{P}_{e,c} = \frac{\text{vol}(\Lambda_e) \Gamma_e(n_e + T)}{\Gamma_e(n_e)(2 \pi^2 \sigma_e^2)^{n_T/2} \Gamma_e^2(2 \pi \sigma_e^2)^{n_T/2}} \sum_{x \in \Lambda_e} \det \left( \frac{1}{2 \sigma_e^2} I_{n_T} + \frac{1}{2 \sigma_e^2} XX^\dagger \right)^{n_T - T},
\]

where the last equality comes from [27]

\[
\int \det(W)^k \exp\left(-\text{Tr}(\Sigma^{-1}W)\right) \text{d}W = \pi^{2(p-1)} \Gamma(p+k) \cdots \Gamma(1+k)(\det \Sigma)^{k},
\]

where \( D_w \) is here the set of all \( p \times p \) positive definite Hermitian matrices.

We finally obtain that an upper bound on the average probability of correct decoding for Eve is

\[
\overline{P}_{e,c} \leq C_{\text{MIMO}} \gamma_e^{Tn_T} \sum_{x \in \Lambda_e} \det \left( I_{n_T} + \gamma_e XX^\dagger \right)^{n_T - T},
\]

where we set \( \gamma_e = \frac{\sigma_e^2}{\sigma_e^2} \) for Eve’s SNR, and \( C_{\text{MIMO}} = \frac{\text{vol}(\Lambda_e) \Gamma_e(n_e + T)}{\pi^{n_T/2} \Gamma_e^2(2 \pi \sigma_e^2)^{n_T/2}} \).

In order to design a good lattice code for the MIMO wiretap channel, we try to derive a code design criterion from the above equation:

\[
\overline{P}_{e,c} \leq C_{\text{MIMO}} \gamma_e^{Tn_T} \left[ 1 + \sum_{x \in \Lambda_e(0)} \det \left( I_{n_T} + \gamma_e XX^\dagger \right)^{n_T - T} \right].
\]

We can suppose that the space-time code used to transmit data to Bob is designed according to the so-called “rank criterion” of [26]. This means that, if \( X \neq 0 \) and \( T \geq n_T \), then, \( \text{rank}(X) = n_T \). If we assume now that Eve’s SNR \( \gamma_e \) is high compared to the minimum distance of \( \Lambda_e \), or actually design \( \Lambda_e \) that way assuming Alice knows Eve’s channel, we get

\[
\overline{P}_{e,c} \leq C_{\text{MIMO}} \gamma_e^{Tn_T} \left[ 1 + \gamma_e^{n_T} \sum_{x \in \Lambda_e(0)} \det \left( XX^\dagger \right)^{n_T - T} \right].
\]

We thus conclude that to minimize Eve’s average probability of correct decoding, the design criterion is now

\[
\min_{\Lambda_e} \frac{1}{\sum_{x \in \Lambda_e(0)} \det \left( XX^\dagger \right)^{n_T - T}}.
\]

**Remark 1** We discuss the meaning of the bound of \( \overline{P}_{e,c} \). The higher \( \gamma_e \), the higher should be Eve’s probability of correct decoding. The expression of this bound is decreasing as a function of \( \gamma_e \) around the origin, a regime which we do not consider (as we just derived the expression assuming \( \gamma_e \) big enough), and is then indeed increasing elsewhere as expected. The minimum value of this upper bound (computed by taking its derivative) is achieved for

\[
\gamma_{e,\text{min}} = \left( \frac{n_T}{T} \sum_{x \in \Lambda_e(0)} \det \left( XX^\dagger \right)^{n_T - T} \right)^{1/(n_T - T)}.
\]
Remark 2 It is important to notice that the upper bound was computed using an infinite lattice \( \Lambda_e \). In some rare cases, as for an example in the case of the Alamouti code discussed later on, the bound happens to be finite even though the lattice is not. In general, it is not, in which case the bound refers not to the infinite lattice \( \Lambda_e \), but instead a finite subset carved from \( \Lambda_e \) via a shaping region. The same holds for the bounds derived below for block and fast fading channels.

5.2 Block and Fast Fading Channels

As a corollary of the analysis done for the MIMO case, we consider the particular fading channels where \( H_b, H_e \) are diagonal matrices. In this case, setting \( n_t = n_r = n \), the channel can be rewritten as

\[
Y = \text{diag}(h_b)X + V_b,
\]

\[
Y = \text{diag}(h_e)X + V_e,
\]

which corresponds to a block fading channel with \( n \) transmit antennas emitting one after the other, coherence time \( T \) and

\[
\text{diag}(h_b) = \begin{bmatrix} h_{b,1} & \cdots & h_{b,n} \\
\end{bmatrix},
\]

\[
\text{diag}(h_e) = \begin{bmatrix} h_{e,1} & \cdots & h_{e,n} \\
\end{bmatrix}.
\]

However, we cannot use the final result for MIMO channels immediately, since the integral over all positive definite Hermitian matrices does not hold anymore. Moreover, the general expression does not hold either since it assumes that \( H_e \) (here \( \text{diag}(h_e) \)) is i.i.d distributed. We thus start from the generic equation, which gives, using a polar coordinates change, and the change of variables \( u_{e,j} = \rho_{e,j} \)

\[
\overline{P}_{e,x} \leq \frac{\Gamma(1+T)^n \text{vol}(\Lambda_e)}{(2\pi\sigma_e^2)^n(2\pi\sigma_b^2)^n} \sum_{x_b, x_e} \left[ \frac{1}{2\sigma_b^2} + \frac{1}{2\sigma_e^2} \|x_e\|^2 \right]^{1-T}.
\]

We finally obtain an upper bound of the average probability of correct decision for Eve for the wiretap block fading channel, given by

\[
\overline{P}_{e,x} \leq C_{BF} \gamma_e^{\gamma_e^n} \sum_{x_b, x_e} \left[ 1 + \gamma_e \|x_e\|^2 \right]^{1-T},
\]

where \( C_{BF} = \frac{(T)^n \text{vol}(\Lambda_e)}{\pi^{nT}} \) and similarly to the MIMO case, \( \gamma_e = \frac{\sigma_b^2}{\sigma_e^2} \).

In order to design a good lattice code for the block fading wiretap channel, we now try to derive a code design criterion:

\[
\overline{P}_{e,x} \leq C_{BF} \gamma_e^{\gamma_e^n} \left[ 1 + \sum_{x_b, x_e \in \Lambda_e \setminus \{0\}} \left[ 1 + \gamma_e \|x_e\|^2 \right]^{1-T} \right].
\]

We can suppose that the code used to transmit data to Bob is designed according to the minimum product distance criterion. This means that, if \( x \neq 0 \), then \( x_i \neq 0 \) for any \( i \). If we assume this time that Eve’s SNR \( \gamma_e \) is high compared to the minimum distance of \( \Lambda_e \), or actually design \( \Lambda_e \) that way assuming Alice knows Eve’s channel, we get

\[
\overline{P}_{e,x} \leq C_{BF} \gamma_e^{\gamma_e^n} \left[ \frac{1}{\gamma_e^n} \sum_{x_b, x_e \in \Lambda_e \setminus \{0\}} \left[ \|x_e\|^2 \right]^{1-T} \right].
\]
This expression is decreasing as a function of $\gamma_e$ around the origin, a regime which we do not consider (as we again just derived the expression assuming $\gamma_e$ big enough), and is then indeed increasing as expected. The minimum value of this upper bound is achieved for

$$\gamma_{e,\text{min}} = \frac{\sum_{x \in \Lambda_s \setminus \{0\}} \prod_{i=1}^{n} (\|x_i\|^2)^{-1-T}}{T}.$$

We thus conclude that to minimize Eve’s average probability of correct decoding, the design criterion is now

$$\min_{\Lambda_s} \sum_{x \in \Lambda_s \setminus \{0\}} \left( \prod_{i=1}^{n} (\|x_i\|^2)^{-1-T} \right).$$

When furthermore $T=1$ (and $X$ is thus a $n \times 1$ vector $x$), we get a fast fading channel:

$$y = \text{diag}(h_e)x + v,$$

where all vectors are $n$-dimensional complex vectors corresponding to $n$ usages of the channel. We can thus immediately apply the previous result to deduce that

$$\overline{P}_{e|\theta} \leq C_{FF}^n \gamma^n \sum_{x \in \Lambda_s \setminus \{0\}} \left( 1 + \gamma_e \|x_i\|^2 \right)^{-2},$$

where $C_{FF} = \frac{\text{vol}(\Lambda_s)}{\pi^n}$ and still again, $\gamma_e = \frac{\sigma_h^2}{\sigma_e^2}$. We thus recover the expressions presented in [28], though here in the complex case, which explains the difference in the exponent.

### 5.3 A MIMO example: The Alamouti Code

In this section, we illustrate the code design criterion derived above using the Alamouti code [29] with QAM constellation, $n_1 = 2$, $n_2 \geq 2$ and $T = 2$. Note that the Alamouti code does not form a $\mathbb{Z}[i]$-lattice, but a $\mathbb{Z}$-lattice. We choose the Alamouti code nevertheless since this is the best understood and the simplest MIMO code available in the literature. It is not difficult to check that our analysis, and thus the resulting code design, holds for real lattices as well. An Alamouti codeword is then of the form

$$X = \begin{bmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{bmatrix}, \quad x_1, x_2 \in \mathbb{Z}[i],$$

so that $\det(XX^*) = (\|x_1\|^2 + \|x_2\|^2)^2 = \|x\|^4$, where

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{Z}[i] = \Lambda_s.$$

The design criterion requires to study

$$\sum_{x \in \Lambda_s \setminus \{0\}} \det(XX^*)^{-n - T} = \sum_{x \in \Lambda_s \setminus \{0\}} \|x\|^{2(n+2T)} = \zeta_{\Lambda_s}(2(n+2T)),$$

where we recognize the Epstein zeta function of a scaled lattice $\mu \Lambda$ ($\mu > 0$), defined by

$$\zeta_{\mu \Lambda}(s) = \sum_{x \in \Lambda \setminus \{0\}} \frac{1}{\mu^s \|x\|^s} = \frac{1}{\mu^s} \zeta_{\Lambda}(s).$$

Since $x \in \mathbb{Z}[i]^2 = \mathbb{Z}^4$, we will consider as possible lattices $\Lambda_s$, either $\mathbb{Z}^4$ itself, with Epstein zeta function

$$\zeta_{\mathbb{Z}^4}(s) = B(1 - 4^{1-s}) \zeta(s) \zeta(s-1)$$

or $D_4$, in which case the vector $x$ above is coded, and belongs to $D_4$ instead of $\mathbb{Z}^4$, which in turn involved the Epstein zeta function of $D_4$

$$\zeta_{D_4}(s) = 3 \cdot 4^{2-s} \zeta(s) \zeta(s-1).$$

In both cases, $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ is the Riemann zeta function.
In order to compare the Epstein zeta function of the two lattices $\mathbb{Z}^4$ and $D_4$, we rescale $D_4$ so that its fundamental volume is equal to the fundamental volume of $\mathbb{Z}^4$, that is 1. Since $\text{vol}(D_4) = 2$, the scaling factor is $\mu = 1/\sqrt{2}$. We then obtain

$$
\zeta_{D_4}(s) = \left(\sqrt{2}\right)^{-2s} \cdot 3 \cdot 4^{s/2} \left(2^{s-1} - 1\right) \zeta(s) \zeta(s-1) = 3 \cdot 2^{-4s} \cdot \left(2^{s-1} - 1\right) \zeta(s) \zeta(s-1),
$$

where $s = 2n_e + 4$, which we have to compare with $\zeta_{\mathbb{Z}^4}(s) = 8 \left(1 - 4^{s-1}\right) \zeta(s) \zeta(s-1)$. We eventually define the gain $\zeta_{D_4}$ obtained by using $D_4$ instead of $\mathbb{Z}^4$ (the uncoded case) as

$$
\zeta_{D_4} = \frac{\zeta_{D_4}(s)}{\zeta_{\mathbb{Z}^4}(s)} = \frac{-2^{4s-1} \left(4^{s-1} - 1\right)}{3 \cdot 2^{-4s} \cdot \left(2^{s-1} - 1\right)} = \frac{1}{3} \cdot 2^{2n_e+1} = \frac{4}{3} 2^{2n_e}.
$$

We illustrate the obtained results on Fig. 12 by plotting the upper bound on Eve's probability $P_{\text{ce}}$ of correct decision, divided by the constant $c_{\text{MIMO}}$, when the Alamouti code is used with as coarse lattice $\Lambda_e$ either $\mathbb{Z}^4$ or $D_4$.

![Alamouti Code - 2 receive antennas](image)

Figure 12: An upper bound on $P_{\text{ce}}/C_{\text{MIMO}}$: the Alamouti code with $n_e = 2$.

Notice that when $\gamma_e$ is small, this upper bound becomes of course very loose as it is a decreasing function in $\gamma_e$, while we expect on the contrary $P_{\text{ce}}(\gamma_e)$ to be an increasing function. This motivates the following discussion the tightness of the upper bound. We go back to the tighter upper bound on $P_{\text{ce}}$:

$$
\overline{P}_{\text{ce}} \leq C_{\text{MIMO}} T_n \sum_{x \in \Lambda_e} \text{det} \left( I_n + \gamma_e X X^T \right)^{-n_e} = C_{\text{MIMO}} T_n \sum_{x \in \Lambda_e} \text{det} \left( \frac{1}{\gamma_e} I_n + X X^T \right)^{-n_e}.
$$

When $X$ is a codeword from the Alamouti code, with $T = n_e = 2$, then

$$
\text{det} \left( \frac{1}{\gamma_e} I_n + X X^T \right) = \left( \frac{1}{\gamma_e} + \|X\|^2 \right)^2,
$$

so that

$$
\varphi_{\Lambda_e}(\gamma_e) \overset{def}{=} \gamma_e^{-n_e} \sum_{x \in \Lambda_e} \text{det} \left( \frac{1}{\gamma_e} I_n + X X^T \right)^{-n_e} = \gamma_e^{-2n_e} \sum_{x \in \Lambda_e} \left( \frac{1}{\gamma_e} + \|X\|^2 \right)^{-2(n_e+2)} = \gamma_e + \gamma_e^{-2n_e} \sum_{x \in \Lambda_e \setminus \{0\}} \left( \frac{1}{\gamma_e} + \|X\|^2 \right)^{-2(n_e+2)}.
$$

We are thus interested in the calculation of

$$
\zeta_{\Lambda_e}(s,a) \overset{def}{=} \sum_{x \in \Lambda_e} \left( a + \|X\|^2 \right)^{-s} = \frac{1}{a^s} + \sum_{x \in \Lambda_e \setminus \{0\}} \left( a + \|X\|^2 \right)^{-s}, a = \frac{1}{\gamma_e}, s = 2(n_e + 2).
$$
which will be done via the Mellin transform \( \mathcal{M}(f)(s) = \int_0^\infty f(t) t^{s-1} dt \). Therefore, if \( a < \| x \|^2 \), we have that

\[
(a + \| x \|^2)^{-\nu} = \sum_{k=0}^{\infty} (-1)^k \frac{a^k}{k!} \Gamma(s+k) \left( \| x \|^2 \right)^{-\nu-k}.
\]

We are now ready to prove the following result for \( Z^4 \). The equivalent result for \( D_4 \) follows, and the consequences of both computations for the bound on the error probability. Suppose \( 0 < a < q \). For the lattice \( Z^4 \), we have

\[
\zeta_{r^2}(s,a) = \sum_{x \in Z^4} \left( a + \| x \|^2 \right)^{-\nu} = \sum_{j=0}^{\infty} \frac{r(j)}{(a+j)^2} + \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\Gamma(s+k)}{\Gamma(s)} \left( -8 \cdot 4^{s+k} \sum_{j=0}^{\infty} \frac{r(j)}{j^k} \right),
\]

where \( r(j) \) denotes the number of vectors of norm \( j \) in \( Z^4 \). In particular, if \( 0 < a < 2 \), we have

\[
\zeta_{r^2}(s,a) = \sum_{x \in Z^4} \left( a + \| x \|^2 \right)^{-\nu} = \frac{1}{a^\nu} + \frac{8}{(1+a)^\nu} - 8 \cdot 4^{s+\nu} \left( 1 - \frac{a}{4} \right)^\nu,
\]

and

\[
\zeta_{r^2}(s,a) = \sum_{x \in Z^4} \left( a + \| x \|^2 \right)^{-\nu} = \frac{3}{a^\nu} + \frac{8}{(1+a)^\nu} - 3 \cdot 4^{s+\nu} \left( 1 - \frac{a}{4} \right)^\nu.
\]

Suppose \( 0 < a < q \). For the lattice \( D_4 \), we have

\[
\zeta_{r^2}(s,a) = \sum_{x \in D^4} \left( a + \| x \|^2 \right)^{-\nu} = \sum_{j=0}^{\infty} \frac{r(j)}{(a+j)^2} + \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\Gamma(s+k)}{\Gamma(s)} \left( -3 \cdot 4^{s+k} \sum_{j=0}^{\infty} \frac{r(j)}{j^k} \right),
\]

where \( r(j) \) denotes the number of vectors of length \( j \) in \( D_4 \). In particular, if \( 0 < a < 2 \), we have

\[
\zeta_{r^2}(s,a) = \sum_{x \in D^4} \left( a + \| x \|^2 \right)^{-\nu} = \frac{1}{a^\nu} + \frac{8}{(1+a)^\nu} - 3 \cdot 4^{s+\nu} \left( 1 - \frac{a}{4} \right)^\nu,
\]

and

\[
\zeta_{r^2}(s,a) = \sum_{x \in D^4} \left( a + \| x \|^2 \right)^{-\nu} = \frac{3}{a^\nu} + \frac{8}{(1+a)^\nu} - \frac{1}{4} \sum_{j=0}^{\infty} r(j) \left( 1 - \frac{a}{4} \right)^\nu.
\]

for \( 0 < a < 4 \).

The implications of the above computations for the error probability are summarized below for \( \gamma_e > 1/2 \). Similar expressions can be obtained for \( \gamma_e > 1/4 \) (or smaller values of \( \gamma_e \)). See Fig.13 for an illustration of the new bounds.

![Figure 13: A tighter upper bound on \( P_{\text{error}}^{\text{CMMO}} \): the Alamouti code with \( n_e = 2 \).](image)

To conclude, we compare the loose upper-bounds with the tight ones in Fig.14, and our bounds on the probability of correct decision for the eavesdropper with simulations in Fig.15. The coarse lattice \( \Lambda_e \) is \( Z^4 \) (resp. \( D_4 \)) while the fine...
lattice $\Lambda_b$ is $1/2Z^4$ (resp. $1/2D^4$) giving rise to a secret spectral efficiency equal to 1 bit per real dimension. For simulations, we used the linear ML decoder of the original Alamouti paper [29]. Decoding of $D_4$ has been done using the Wagner decoder of the binary parity check (4,3) code.

Figure 14: Loose upper bounds versus tight upper bounds: the Alamouti code with $n_e = 2$.

Figure 15: Upper bounds versus simulations: the Alamouti code with $n_e = 2$. 
6 Secret key generation from Gaussian sources using lattice hashing

In this section, we propose a simple yet complete lattice-based scheme for secret key generation from Gaussian sources in the presence of an eavesdropper, and show that it achieves strong secret key rates up to $1/2$ nat from the optimal in the case of “degraded” source models. The novel ingredient of our scheme is a lattice-hashing technique, based on the notions of flatness factor and channel intrinsic randomness. The proposed scheme does not require dithering.

6.1 Lattice Hashing for Gaussian Sources

Consider now a source model for secret key generation with public discussion, in the presence of an eavesdropper. For simplicity, we first assume that Alice and Bob observe the same i.i.d. Gaussian random variable $X^n = Y^n$ of variance $\sigma^2$ per dimension. Eve observes a correlated i.i.d. random variable $Z^n$. We assume that $X^n$ and $Z^n$ are jointly Gaussian, according to the following model

$$X^n = Z^n + W^n,$$

where $W^n$ is an i.i.d. zero-mean Gaussian random vector of variance $\sigma^2$ per dimension. We suppose that $W^n$ and $Z^n$ are independent.

Our aim is to extract from $X^n$ a random number that is almost uniform on $\mathbb{R}^n$ and almost independent of $Z^n$. To do this, we apply the mod $R(\Lambda)$ operation. Then, the conditional density of $\widetilde{X} = X^n \mod R(\Lambda)$ given $Z^n$ is

$$p_{\widetilde{X}|Z}(\widetilde{x}|z) = \frac{1}{V(\Lambda)} \sum_{x \in \mathbb{R}^n \mod R(\Lambda)} p_{X^n}(x|z) 1_{R(\Lambda)}(\widetilde{x}).$$

From the definition of the flatness factor, it then follows that

$$\forall z \in \mathbb{R}^n, \forall x \in R(\Lambda), \left| p_{\widetilde{X}|Z}(\widetilde{x}|z) - \frac{1}{V(\Lambda)} \right| \leq \frac{\varepsilon(\sigma)}{V(\Lambda)}.$$

With a similar reasoning, we also find

$$p_{\widetilde{X}}(\widetilde{x}) = f_{\varepsilon(\sigma)}(\widetilde{x}) 1_{R(\Lambda)}(\widetilde{x})$$

and again by definition of the flatness factor, we find

$$\forall x \in R(\Lambda), \left| p_{X}(x) - \frac{1}{V(\Lambda)} \right| \leq \frac{\varepsilon(\sigma)}{V(\Lambda)}.$$

So, if the flatness factor is small, $\widetilde{X}^n$ is almost uniformly distributed over $R(\Lambda)$, and also almost independent of $Z^n$. It is worth mentioning that unlike other works which use dithering or the high-resolution assumption [30], we obtain uniformity and independence from the flatness factor. One can now bound the mutual information

$$I(X^n;Z^n) = \int_{\mathbb{R}^n} \int_{R(\Lambda)} p_{X^n}(x,z) \log \frac{p_{X^n}(x,z)}{p_{X^n}(x)} \, dx \, dz \leq \varepsilon(\sigma) + 2\varepsilon(\sigma)$$

if $\varepsilon(\sigma) \leq 1/2$. Since $\sigma_\Lambda \geq \sigma$, and recalling Remark 1, we have $I(\widetilde{X}^n;Z^n) \leq 3\varepsilon(\sigma)$. The sufficient condition for the existence of secrecy-good lattices ensures that $I(\widetilde{X}^n;Z^n)$ vanishes exponentially if $\gamma(\sigma) < 2\pi$. Observe that depending on the choice of $\Lambda$, the rate of extracted randomness can be arbitrarily large.

The asymptotic differential entropy rate of $\widetilde{X}^n$ is

$$r = \liminf_{n \to \infty} \frac{1}{n} h(\widetilde{X}^n) \geq \liminf_{n \to \infty} \frac{1}{n} \left[ \log V(\Lambda) - \log(1 + \varepsilon(\sigma)) \right].$$
Taking a sequence of secrecy-good lattices such that \( \gamma_\Lambda (\sigma) \to 2\pi \) as \( n \to \infty \), which is compatible with the condition required, we can obtain the asymptotic rate \( r = \log(\sqrt{2\pi \sigma}) \), which is only 1/2nat from the asymptotic differential entropy rate of the Gaussian noise \( W^n \) (i.e., \( \log(\sqrt{2\pi e}) \)).

Note that neither nearest-neighbour quantization nor dither is used in our lattice-hashing scheme, and we only need to implement the mod \( R(\Lambda) \) operation, which can be performed in polynomial time for many fundamental regions \( R(\Lambda) \). In particular, we can choose the fundamental parallelepiped. Moreover, Remark 1 implies that if the lattice \( \Lambda \) is chosen randomly in a mod-\( p \)-ensemble, it is secrecy-good with high probability, so that one can obtain explicit schemes.

### 6.2 Secret Key Agreement

From the above discussion, it seems that one can get an arbitrarily high rate of the secret key, since \( X^n \) is continuous. However, this is fictitious, because

\[
X^n \neq Y^n.
\]

This requires Alice and Bob to agree on the key over a public channel, which will lead to a finite key rate.

We consider an i.i.d. memoryless Gaussian source \( p_{xyz} \) whose components are jointly Gaussian with zero mean. The distribution is fully described by the variances \( \sigma_x^2, \sigma_y^2, \sigma_z^2 \) and the correlation coefficients \( \rho_{xy}, \rho_{xz}, \rho_{yz} \). We can write:

\[
X^n = \frac{\sigma_y}{\sigma_x} Y^n + W_1^n,
\]

\[
X^n = \frac{\sigma_z}{\sigma_x} Z^n + W_2^n
\]

respectively. Further, \( W_1^n \) is independent of \( Y^n \), and \( W_2^n \) is independent of \( Z^n \).

![Figure 16: Secret key generation in the presence of an eavesdropper with communication over a public channel [46], where \( W_1^n \) and \( W_2^n \) are i.i.d. zero-mean Gaussian noise vectors of variance.](image)

The results of the previous section allow to extract from \( X^n \) a random variable \( \tilde{X}^n \) that is almost statistically independent of \( Z^n \), with \( \sigma_x^2 \) replaced by \( \sigma_z^2 \). Note that the coefficient \( \rho_{xz} \sigma_x / \sigma_z \) does not affect the argument based on the flatness factor. However, not all the extracted randomness can be exploited to generate the key, because Bob has to reconstruct \( X^n \) with side information \( Y^n \), which requires Wyner-Ziv coding. Also, in secret key generation, we are not concerned with the standard rate-distortion function, but with the error probability of the key.
We assume that only one round of one-way public communication (from Alice to Bob) takes place. More precisely, Alice computes a public message $S$ and a secret key $K$ from her observation $X^n$; she then transmits $S$ over the public channel. From this message and his own observation $Y^n$, Bob reconstructs a key $K$.

Let $K_n$ and $S_n$ be the sets of secret keys and public messages respectively. A secret key rate - public rate pair $(R_K, R_S)$ is achievable if there exists a sequence of protocols with

$$\lim_{n \to \infty} \frac{1}{n} \log |K_n| \geq R_K, \quad \lim_{n \to \infty} \frac{1}{n} \log |S_n| \geq R_S,$$

such that the following properties hold:

$$\lim_{n \to \infty} \log |K_n| - H(K) = 0 \quad \text{(uniformity)};$$

$$\lim_{n \to \infty} P\left( K \neq \hat{K} \right) = 0 \quad \text{(reliability)};$$

$$\lim_{n \to \infty} (K, S, Z^n) = 0 \quad \text{(strong secrecy)}.$$

To define our key generation scheme, we use the lattice partition chain $\Lambda_1 / \Lambda_2 / \Lambda_3$, where

- $\Lambda_1$ is quantization-good, which serves as the “source-code” component of Wyner-Ziv coding;
- $\Lambda_2$ is AWGN-good, which serves as the “channel-code” component in Wyner-Ziv coding;
- $\Lambda_3$ is secrecy-good with respect to $\sigma_2$, which serves as the extractor of randomness.

The existence of such a chain of lattices will be shown in the next section. We suppose that the lattices are scaled so that their volumes $V_1, V_2, V_3$ satisfy

$$|\Lambda_2 / \Lambda_3| = \frac{V_3}{V_2} = e^{\rho R_S}, \quad |\Lambda_1 / \Lambda_2| = \frac{V_2}{V_1} = e^{\rho R_K}.$$

The procedure of secret key generation is described as follows:

- Alice quantizes $X^n$ to $X^n_{\Lambda_1} = Q_{\Lambda_1}(X^n) \in \Lambda_1$. She then computes $S = X^n_{\Lambda_1} \mod V(\Lambda_2)$, which belongs to a set of coset leaders of $\Lambda_1 / \Lambda_2$ in $V(\Lambda_2)$, and transmits its index to Bob. Moreover, Alice computes the key $K = Q_{\Lambda_2}(X^n_{\Lambda_1}) \mod R(\Lambda_3)$, which belongs to a set of coset leaders of $\Lambda_2 / \Lambda_3$ in $R(\Lambda_3)$. Note that $X^n = E^n_q + S + K + \lambda_3$ for some $\lambda_3 \in \Lambda_3$, where $E^n_q = X^n - X^n_{\Lambda_1} \in V(\Lambda_3)$ is the quantization error.

- Bob receives $S$ and reconstructs $X^n_{\Lambda_1} = S + Q_{\Lambda_2} \left( \frac{\sigma_x}{\sigma_y} Y^n - S \right)$.

He then computes his version of the key $\hat{K} = Q_{\Lambda_1}(X^n_{\Lambda_1}) \mod R(\Lambda_3)$.

Note that $K$ and $S$ are functions of $\overline{X} = X^n \mod R(\Lambda_3)$ generalizing [31], we have

$$K = (Q_{\Lambda_2}(Q_{\Lambda_1}(X^n_{\Lambda_1}) \mod R(\Lambda_3))) \mod R(\Lambda_3) = f(\overline{X}).$$

Moreover, it is not hard to see that

$$X^n \mod V(\Lambda_2) = \overline{X} \mod V(\Lambda_2)$$

and thus

$$S = Q_{\Lambda_1}(\overline{X}) \mod V(\Lambda_2) = g(\overline{X}).$$

### 6.3 Uniformity

Using the results of the previous section, we can show that $K$ is almost uniformly distributed on $\Lambda_2 / \Lambda_3$: from Eq. (10) we have
\[ \overline{X}^n = X^n \mod R(A_\gamma) = (E_0^n + S + K) \mod R(A_\gamma), \] 
where \( Y(k+s) = (Y(A_\gamma) + k+s) \mod R(A_\gamma) \). We then find that
\[
\forall k \in A_2 \cap R(A_\gamma),
\left| p_e(k) - \frac{V_z}{V_0^2} \right| \leq \frac{\varepsilon_{A_\gamma}(\sigma_x)}{e^{\sigma_y}}.
\]
Consequently, the entropy of the key is lower bounded by
\[
H(K) \geq \sum_{k \in A_2 \cap R(A_\gamma)} p_e(k) \log \left( \frac{e^{\sigma_y}}{1 + \varepsilon_{A_\gamma}(\sigma_x)} \right) = n R_K - \log(1 + \varepsilon_{A_\gamma}(\sigma_x)) \geq n R_K - \varepsilon_{A_\gamma}(\sigma_x).
\]

### 6.4 Strong secrecy and Reliability

We recall the following bound from [32]:
\[
I(K; S, Z^n) \leq d_{av} \log |K^n|, \]
where \( d_{av} = \sum_{k \in \mathcal{K}_n} p_e(k) V(p_{S^n|K=k, Z^n}, p_{S^n}) \), and \( V \) denotes the variational distance. Observe also that \((S, K) \to \overline{X}^n \to Z^n\) is a Markov chain. Therefore we have
\[
p_{S^n|K=k}(s, z | k) = \frac{p_{S^n|K=k}(s, z, k)}{p_e(k)} = \frac{1}{p_e(k)} \int_{Y(k+s)} p_{Z^n}(z|x) \overline{x} d\overline{x}.
\]
Similarly,
\[
p_{S^n}(s, z) = \sum_{k \in \mathcal{K}_n} p_{S^n|K=k}(s, z | k) = \sum_{k \in \mathcal{K}_n} \int_{Y(k+s)} p_{Z^n}(z|x) \overline{x} d\overline{x}.
\]
From the bounds on \( p_e(k) \) and flatness factor, and noticing that \( \sigma_2 \leq \sigma_x \), we have,
\[
\left| \sum_{k \in \mathcal{K}_n} \int_{Y(k+s)} p_{S^n}(z|x) \overline{x} d\overline{x} - \frac{p_e(z)}{e^{\sigma_y}} \right| \leq \frac{\varepsilon_{A_\gamma}(\sigma_2)}{e^{\sigma_y}} - p_e(z),
\]
provided that \( \varepsilon_{A_\gamma}(\sigma_x) \leq 1/2 \). Consequently,
\[
V(p_{S^n|K=k}, p_{S^n}) \leq \sum_{z} \int_{Z^n} \frac{5\varepsilon_{A_\gamma}(\sigma_x)}{e^{\sigma_y}} p_{S^n}(z) d\overline{z} = 5\varepsilon_{A_\gamma}(\sigma_x).
\]
Therefore \( d_{av} \leq 5\varepsilon_{A_\gamma}(\sigma_x) \). If \( A_\gamma \) is secrecy-good, we find
\[
I(K; S, Z^n) \leq 5\varepsilon_{A_\gamma}(\sigma_x) (n R_K - \log 5\varepsilon_{A_\gamma}(\sigma_x)) \to 0.
\]
Let us analyze the error probability \( P[K \neq \hat{K}] \). Note that \( K = \hat{K} \) if \( \hat{X}_0^n = X_0^n \). Since \( X_0^n = S + Q_{A_\gamma}(X_0^n) \), we have
\[
\hat{X}_0^n = X_0^n \iff Q_{A_\gamma} \left( \rho_{\sigma_y} \frac{\sigma_{X^n}}{\sigma_y} - S \right) = Q_{A_\gamma} (X_0^n).
\]
Since \( Q_{A_\gamma} \left( \rho_{\sigma_y} \frac{\sigma_{X^n}}{\sigma_y} - S \right) = Q_{A_\gamma} (\rho_{\sigma_y} \frac{\sigma_{X^n}}{\sigma_y} - X_0^n + Q_{A_\gamma}(X_0^n)) = Q_{A_\gamma} (\rho_{\sigma_y} \frac{\sigma_{X^n}}{\sigma_y} - X_0^n) + Q_{A_\gamma}(X_0^n) \), we derive
\[
\hat{X}_0^n = X_0^n \iff Q_{A_\gamma} \left( \rho_{\sigma_y} \frac{\sigma_{X^n}}{\sigma_y} - X_0^n \right) = 0
\]
\[
\iff Q_{A_\gamma} (E_0^n - W^n_1) = 0.
\]
When $\epsilon_{\Lambda_1}(\sigma_1)$ and $\epsilon_{\Lambda_1}(\rho_1\sigma_1)$ are small, $E_q^n = X^n \text{Mod } \mathcal{V}(\Lambda_1)$ is almost uniformly distributed on $\mathcal{V}(\Lambda_1)$ and almost independent of $W_1^n$. The variance per dimension of $E_q^n - W_1^n$ is asymptotically $G(\Lambda_1)V_1^{2/n} + \sigma_1^2$, where $G(\Lambda_1)$ is the normalized second moment. According to [30], if $\Lambda_1$ is good for quantization, then the effect of $E_q^n$ on the decoding error probability is sub-exponential in $n$ relative to the AWGN of the same power.

By the AWGN-goodness of $\Lambda_2$, the error probability $P(K = 1) \leq P(Q_{\Lambda_1}(E_q^n - W_1^n) \neq 0)$ will vanish exponentially as long as

$$\frac{V_2^{2/n}}{2\pi e} > 2\pi e.$$ 

On the other hand, the secrecy-goodness of $\Lambda_2$ requires $V_2^{2/n} / \sigma_2^2 < 2\pi$. Therefore, the rate of the secret key is bounded by

$$R_K \leq \frac{1}{n} \log \frac{\sigma_2^2}{\sigma_1^2} - \frac{1}{2}.$$ 

This corresponds to the rate of public communication

$$R_p = \frac{1}{n} \log \frac{V_2}{V_1} > \frac{1}{2} \log \left(1 + \frac{2\pi e \sigma_1^2}{V_2^{2/n}}\right).$$ 

If we make $\Lambda_1$ sufficiently fine such that $G(\Lambda_1)V_1^{2/n} \ll \sigma_1^2$, then the key rate approaches

$$R_K \leq \frac{1}{2} \log \frac{\sigma_2^2}{\sigma_1^2} - \frac{1}{2}.$$ 

For degraded sources, i.e. $\sigma_1 < \sigma_2$ or equivalently, the secret key rate is upper-bounded by $I(X;Y) - I(X;Z)$. The key-rate $R_K$ is only $1/2$ nat away from this bound. Achieving this rate requires a high rate $R_p$ of public communication, but $\Lambda_1$ need to be very fine in practice. To see this, we substitute $V_2^{2/n} = 0.2\pi e \sigma_1^2$ so that $R_K$ is almost the same as above while $R_p \approx 1.2$ nats/dimension. Note, however, that our scheme does not achieve the optimal trade-off between $R_K$ and $R_p$ identified yet. Achieving the optimal trade-off will be our future work.

### 6.5 Existence of a Sequence of Nested Lattices for Secret Key Generation

We begin by showing the existence of a suitable coarse lattice $\Lambda_3^n$. For the notions of AWGN-good, Rogers-good and quantization-good lattices we refer the reader to [33, 31]. Following the same reasoning as in [14], it can be shown that there exists a sequence $\delta_n \to 0$ and an ensemble of sequences of lattices $\Lambda_3^n$ with second moment $\sigma^2(\Lambda_3^n) = \sigma_2^2 / e$ which are AWGN-good, Rogers-good and quantization-good and such that

$$E \left[ \Theta_{\Lambda_3^n} \left( \frac{1}{2\pi \sigma_2^2} \right) \right] \geq 1 + \delta_n + \frac{(2\pi \epsilon_2^2)^2}{V_3^{2/n}}.$$ 

Quantization-goodness then implies that $G(\Lambda_3^n) = \sigma^2(\Lambda_3^n) / (V_3^{2/n})^2 \to 1 / 2\pi e$, and consequently $\Lambda_3^n$ tends to $2\pi \sqrt{2}$ from below, as required to achieve optimal rate while satisfying the required condition. From the average bound above, and recalling the relation between theta series and flatness factor, we can deduce the existence of a sequence $\Lambda_3^n$ which is also secrecy-good.

By applying twice the technique in [33], and its extension in [34], we can find two sequences of fine lattices $\Lambda_1^n$, $\Lambda_2^n$ with $\Lambda_1^n \supseteq \Lambda_2^n \supseteq \Lambda_3^n$ which are also Rogers, quantization and AWGN-good and such that the volume ratios are
arbitrarily close to the bounds. Note that since \( \epsilon_\Lambda^*(\sigma_2) \leq \epsilon_\Lambda^*(\sigma_2') \) whenever \( \Lambda \subset \Lambda' \), the lattices \( \Lambda_2^{(a)}, \Lambda_1^{(a)} \) are also secrecy-good with respect to \( \sigma_2' \).

## 7 Conclusion - future work and relations with other tasks

### 7.1 Summary

In summary, this deliverable has addressed two main aspects of the wiretap coding in the context of PHYLAWS objectives:

In the first part of the deliverable, we investigated the explicit wiretap coding design with good transition efficiency. First, we have introduced two standard models for discrete channels, based on LDPC codes and polar codes, respectively. Then we used polar codes to construct polar lattices and designed a lattice-based wiretap coding scheme for the continuous Gaussian channels. Moreover, an efficient and explicit lattice shaping method was also presented. In this work, we further introduced MIMO wiretap coding using a similar lattice structure. We analysed Eve’s probability of correctly decoding the message Alice meant to Bob, and from minimizing this probability, we derived a code design criterion for MIMO lattice wiretap codes. The case of block fading channels was treated similarly, and fast fading channels are derived as a particular case. The Alamouti code was carefully studied as an illustration of the analysis provided.

In the second part, as a compliment of the previous deliverable, we proposed a simple yet complete lattice-based scheme for secret key generation from Gaussian sources in the presence of an eavesdropper, and showed that it achieves strong secret key rates up to 1/2 nat from the optimal in the case of “degraded” source models. The novel ingredient of our scheme is a lattice-hashing technique, based on the notions of flatness factor and channel intrinsic randomness. The proposed scheme does not require dithering, indicating a convenient implementation in practice.

### 7.2 Future works and relations to other tasks

The theoretical aspects will be further developed to wiretap coding of the MIMO channel and fading channel. In fact, the current claim of security in these scenarios is based on the result that the decoding error probability at Eve’s end is large. However, from the perspective of information theoretical security, large decoding error probability does not necessarily ensure negligible information leakage. Therefore, more careful analysis is needed. We may still follow the prior work [14] and try to combine the flatness factor with the information leakage.

Experimental and simulation works will also intend to progress into the provision of databases of channel characteristics intended to evaluate the performance of the schemes investigated in PHYLAWS regarding PHYSEC. This will be described in the final version of this deliverable, as well as in the deliverables produced in the other related tasks. In particular, measurements/simulations/modeling of the radio channel are involved in WP3, WP4, WP5 and WP6, specifically in tasks

- T3.1 : Study and development and test of “propagation dependent” random generators
- T4.1 : New RATs based on SISO, MISO and MIMO technologies
- T5.1 : CIR measurement and modeling
- T6.1 : Modeling of LTE-based cellular system (Leader: VTT, Participants: TPT)

while the use of this data/models will mainly take place in:

- T3.2 : Study of TRANSEC upgrades in existing networks
- T4.3: New RATs based on IFF and double talk technologies, transposed to ground radio mobile
- T5.2 : Simulations of PHYSEC methods using measuredCIRs
- T6.2: Simulation of interception of waveform signals in LTE-based cellular system

These works will be in part based on the contents of the present deliverable.
8 References

8.1 Phylaws deliverables

Consultation of these documents is specifically relevant with respect to the present deliverable:

http://www.phylaws-ict.org/

[D2.3] Deliverable 2.3 - Fundamental aspects of physical layer security

8.2 Open literature


8.3 More relevant references


[72] A. Sibille, “Keyholes and MIMO Channel Modelling”, COST 273, 15-17 October 2001, Bologna (Italy), Document TD(01) 017


